Extreme-scale Vector Comparison Enables High-resolution Climate Analysis

Background

- Quantifying high-resolution climatic zones and longitudinal environmental change provides a blueprint to assess relationships between micro- and meso-climates and the processes they drive, e.g., greenhouse gas (GHG) emission, zoonotic transmission, ice-melt/sea-level rise, deforestation, agronomic efficiency, and many other applications. However, previous approaches have relied on heuristic decision rules instead of data-driven methods.

Approach

- Computational advancements leveraging nearly every aspect of the Summit and JUWELS Booster architectures are used to compute exhaustive pairwise and trinary comparisons of high-dimensional vectors comprising 60 years of historical climate data. The results are captured as massive networks, which are used in downstream scientific analysis, e.g., high-performance unsupervised graph clustering and longitudinal network analysis.

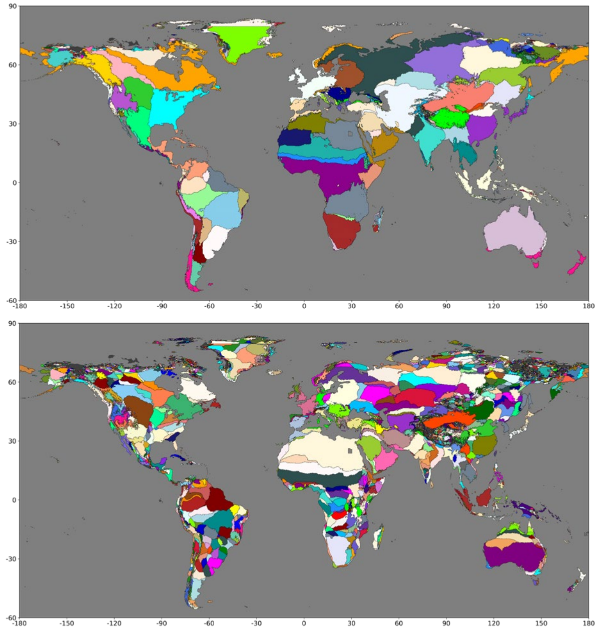
Results

- The methods performed here resulted in the fastest (9.37×10^{18} operations per second) scientific computation ever performed, with more than 100 quadrillion edges considered for a single climatic network.
- We applied Markov clustering and our novel *Correlation of Correlations (Cor-Cor)* method to the resulting networks, yielding unprecedented agglomerative and longitudinal views of relativistic changes in global climatic relationships.
- These analyses revealed areas of the world experiencing rapid environmental changes and highlight potential “danger zones” for future pandemics – where we may expect new diseases to emerge or propagate – and a host of applications, including bioenergy systems optimization, improved understanding of weather systems and hydrodynamics, and the evolution of geopolitics.

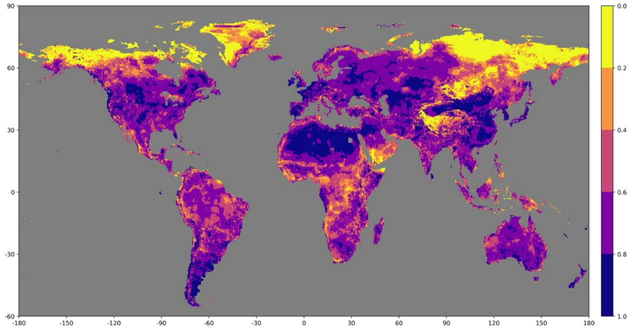
Significance

- Exhaustive vector comparison across data sets of this scale was previously unachievable. This work provides a methodology for applying graph algorithms to datasets of virtually unlimited complexity, and lays the foundation for applying AI algorithms, particularly deep graph networks, to pressing data-driven scientific problems.

Lagergren, J., et al. 2021 Climatic clustering and longitudinal network analysis with impacts on food, bioenergy, and pandemics. *Phytobiomes* (2021).



Pairwise (top) and trinary (bottom) clustering results, indicating environmentally similar climatic zones.



Cor-Cor results tracking longitudinal relativistic climatic change.