

Using data mining to make sense of climate change

January 18 2018, by Jason Maderer



Credit: Georgia Institute of Technology

Big data and data mining have provided several breakthroughs in fields such as health informatics, smart cities and marketing. The same techniques, however, have not delivered consistent key findings for

climate change.

There are a few reasons why. The main one is that previous [data mining](#) work in [climate science](#), and in particular in the analysis of climate teleconnections, has relied on methods that offer rather simplistic "yes or no" answers.

"It's not that simple in climate," said Annalisa Bracco, a professor in Georgia Tech's School of Earth and Atmospheric Sciences. "Even weak connections between very different regions on the globe may result from an underlying physical phenomenon. Imposing thresholds and throwing out weak connections would halt everything. Instead, a [climate scientist's](#) expertise is the key step to finding commonalities across very different data sets or fields to explore how robust they are."

And with millions of data points spread out around the globe, Bracco said current models rely too much on human expertise to make sense of the output. She and her colleagues wanted to develop a [methodology](#) that depends more on actual data rather than a researcher's interpretation.

That's why the Georgia Tech team has developed a new way of mining data from [climate data](#) sets that is more self-contained than traditional tools. The methodology brings out commonalities of data sets without as much expertise from the user, allowing scientists to trust the data and get more robust—and transparent—results.

The methodology is open source and currently available to scientists around the world. The Georgia Tech researchers are already using it to explore sea surface temperature and cloud field data, two aspects that profoundly affect the planet's climate.

"There are so many factors—cloud data, aerosols and wind fields, for example—that interact to generate climate and drive [climate change](#),"

said Athanasios Nenes, another College of Sciences climate professor on the project. "Depending on the model aspect you focus on, they can reproduce climate features effectively—or not at all. Sometimes it is very hard to tell if one model is really better than another or if it predicts climate for the right reasons."

Nenes says the Georgia Tech methodology looks at everything in a more robust way, breaking the bottleneck that is typical of other model evaluation and analysis algorithms. The methodology, he says, can be used for observations, and scientists don't need to know anything about computer code and models.

"The methodology reduces the complexity of millions of data points to the bare essentials —sometimes as few as 10 regions that interact with each other," said Nenes. "We need to have tools that reduce the complexity of model output to understand them better and evaluate if they are providing the correct results for the right reasons."

To develop the methodology, the climate scientists partnered with Constantine Dovrolis and other data scientists in Georgia Tech's College of Computing. Dovrolis said it's exciting to apply algorithmic and computational thinking in problems that affect everyone in major ways, such as global warming."

"Climate science is a 'data-heavy' discipline with many intellectually interesting questions that can benefit from computational modeling and prediction," said Dovrolis, a professor in the School of Computer Science, "Cross-disciplinary collaborations are challenging at first—every discipline has its own language, preferred approach and research culture—but they can be quite rewarding at the end."

The paper, "Advancing [climate](#) science with knowledge-discovery through [data](#) mining," is published in *Climate and Atmospheric Science*.

More information: Annalisa Bracco et al. Advancing climate science with knowledge-discovery through data mining, *npj Climate and Atmospheric Science* (2017). [DOI: 10.1038/s41612-017-0006-4](https://doi.org/10.1038/s41612-017-0006-4)

Provided by Georgia Institute of Technology

Citation: Using data mining to make sense of climate change (2018, January 18) retrieved 10 April 2024 from <https://phys.org/news/2018-01-climate.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.