TOOLKIT FOR EXTREME CLIMATE ANALYSIS (TECA)

Risks of extreme weather events pose some of the greatest hazards to society and the environment as the earth system changes. Extreme weather has recently focused public attention on the dramatic consequences that follow from these events. In 2011, unusually high precipitation, combined with high snowpack, caused extensive flooding throughout the central United States. Heat waves across the same region in 2012 and 2015 produced the country's hottest and second-hottest years in recorded history. Most recently, the year 2017 brought intense record precipitation in California and enormous impacts to the U.S. Gulf Coast and Caribbean Islands from Hurricanes Harvey, Irma, and Maria.

As the severity of extreme weather events continues to increase, this will constitute one of the most stressing forms of change for society and the environment. Therefore, it is crucial to predict with greater reliability how extreme events might change in the future and, in order to advance this objective, to determine with as much certainty as possible whether and why extreme events have already changed.

SCIENTIFIC FOCUS

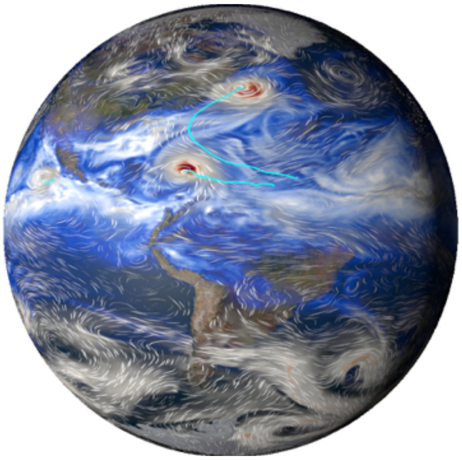
The intersection of climatic extremes with critical water and energy resources for the United States is a key focal area for earth system research in the U.S. Department of Energy (DOE). This priority is reflected in the DOE *Climate and Environmental Sciences Division Strategic Plan*, the 2012 *DOE Workshop on Community Modeling and Long-Term Predictions of the Integrated Water Cycle*, and a 2013 DOE report on *U.S. Energy Sector Vulnerabilities to Climate Change and Extreme Weather*. Sponsored by DOE's Regional and Global Model Analysis, the Calibrated and Systematic Characterization, Attribution, and Detection of Extremes (CASCADE) project addresses the critical knowledge gaps on weather extremes needed to advance DOE's mission.

CASCADE is developing the following capabilities to accelerate DOE's research portfolio in earth system analysis and to advance scientific capabilities in climate analysis:

1. Understanding drivers of observed changes in extremes
2. Characterization of the dominant sources of uncertainty in extremes
3. Understanding and simulating the physical behavior of extreme events
4. High-performance software toward exascale analysis of extreme events.

These capabilities will be used to answer several key science questions:

1. How has the nature of extreme events changed in recent history (e.g., the frequency, duration, intensity, and spatial extent)?
2. What has contributed to this change?
3. How can the nature of extreme events change in the future?

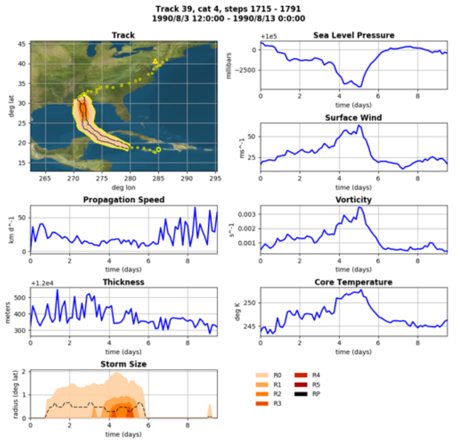


CASCADE is investigating what causes some years to have more hurricanes or stronger hurricanes than other years, and how hurricanes might change in the future. The CASCADE team created TECA for such investigations (two automatically detected hurricane tracks are shown in cyan). TECA allows researchers to examine many distinct weather events—hurricanes, in this case—in an automated way that is efficient for earth system model data sets terabytes to petabytes in scale.

CASCADE COMMUNITY RESEARCH TOOLS

High-performance software toward exascale analysis of extreme events

CASCADE is a science-focused research project, and many of the team's research problems require new software tools and data sets.



TECA uses advanced parallel computing techniques to automatically and efficiently detect hurricanes in large climate data sets. The image above shows one such detection (track shown in upper left) and various hurricane properties during the lifetime of the storm (panels). TECA outputs data in a format that simplifies aggregating results from multi-decadal simulations by a variety of criteria, including: magnitude, basin, season, etc.

As one of four key focus areas within CASCADE, the team creates high-performance, open-source computational and statistical tools that can be shared, reused, and further developed for research beyond the project's central research challenges. The effort is designed to produce significant new capabilities for earth system science: (1) creation of high-fidelity statistical tools for quantifying extremes; (2) development of a high-throughput tool for identifying and tracking weather features in terabytes to exabytes of earth system model data; and (3) extension of uncertainty quantification frameworks to treat a wide of variety of extreme phenomena.

Toolkit for Extreme Climate Analysis: TECA

TECA is a general-purpose, high-performance tool for detecting discrete events in climate model output. It leverages a map-reduce framework for efficient parallelization at large scales (order 10K+ cores). Currently, TECA contains detection algorithms for tropical cyclones (TC), atmospheric rivers (AR), and extratropical cyclones (ETC); plans are underway to implement algorithms for mesoscale convective complexes, African Easterly waves, atmospheric blocks, and fronts.

A Programmable Interface for Detecting Weather Events

TECA has two main interfaces aimed at two distinct use-cases: use of 'canned' algorithms (TC, AR, ETC, etc.), and development of user-defined algorithms. The TECA code base is written in modern C++ (i.e., using C++ 2011 standards), and contains a Python API.

The first is a command-line interface, with command-line programs for each implemented detection algorithm. The command-line interfaces take file paths as input and contains several arguments for customization (e.g., setting of detection thresholds, setting output files). A typical parallel run on NERSC systems looks like the following:

```
srun -n 29200 teca_tc_detect   \
        --input_regex ${data_dir} ${files_regex} \
        --candidate_file candidates_1990s.bin \
        --track_file tracks_2000s.bin
```

The second interface leverages the Python API to allow non-C++-proficient developers to easily prototype parallel performance algorithms using Python—a widely adopted language in the scientific community. An example TC detector is implemented in the Python API, and efforts are currently underway to implement several different atmospheric river detection algorithms using the Python interface, both for scientific use and to highlight the flexibility of the TECA pipeline when integrated into Python. The Python API offers the option to design custom detection workflows with minimal complexity and software engineering on the scientist's part.

COLLABORATIONS

The CASCADE project is a multidivisional, collaborative work at Lawrence Berkeley National Laboratory (LBNL), drawing upon expertise of scientists in the lab's Computational Research Division and Climate and Ecosystem Sciences Division as well as the University of California, Berkeley and University of California, Davis campuses. CASCADE scientists collaborate with related projects at LBNL and across DOE's Office of Biological and Environmental Research modeling efforts. These projects include earth system modeling efforts; land, ocean, and atmosphere diagnostics projects; and stakeholder-driven science projects. The resulting connections and related projects ensure tight integration of observations, experiments, and modeling of extreme weather events. CASCADE is also active in national and international scientific activities, including CMIP, SAMSI, ARTMIP, etc.

CONTACTS

Renu Joseph, Ph.D.
DOE Program Manager
Regional and Global Model Analysis
renu.joseph@science.doe.gov

William D. Collins, Ph.D.
Principal Investigator
Lawrence Berkeley National Laboratory
wdcollins@lbl.gov

Travis A. O'Brien, Ph.D.
Technical Co-Manager
Lawrence Berkeley National Laboratory
taobrien@lbl.gov

Project Website

Users can download TECA and learn more at *https://github.com/LBL-EESA/TECA*