Managing Large-scale Atmospheric and Oceanic Climate Data for Efficient Analysis and On-the-fly Interactive Visualization

Aashish Panta, Valerio Pascucci, Giorgio Scorzelli, Amy A. Gooch, Nina McCurdy*, David Ellsworth*, Kyo Lee**
University of Utah, NASA Ames Research Center*, NASA JPL**

Introduction

Traditional tools often cannot handle extensive datasets, creating bottlenecks in data analysis and visualization. We present a groundbreaking visualization dashboard designed to address these issues by enabling progressive streaming and visualization.

1. Scalable Visualization Framework: An innovative, scalable dashboard framework that enables progressive visualization with advanced analytical tools

2. Efficient Data Reduction and Optimization: We enable efficient storage and the transfer of large-scale data with analysis-ready cloud-optimized format

3. Data Democratization: Successfully converted and migrated more than a Petabyte of raw data from Pleiades, a NASA Supercomputer, to the cloud, enhancing public access and collaborative opportunities.

Interactive Visualization and Dashboards

- We worked on two large-scale climate simulation datasets: the DYAMOND global atmospheric model and global ocean model dataset and the LLC4320 Ocean Dataset, including several fields like **ocean velocity, salinity, and sea surface temperature**.
- We developed an innovative data conversion pipeline to convert more than **a Petabyte** of this data to OpenVisus IDX format, known for its fast and progressive streaming capabilities. We also tested various compression algorithms, both lossy and lossless, on these datasets
- To make the data publicly available, we uploaded over **600TB of compressed data** to SealStorage, a decentralized cloud storage service. We deployed several **dashboards and Jupyter notebooks** to fetch the data from cloud storage.

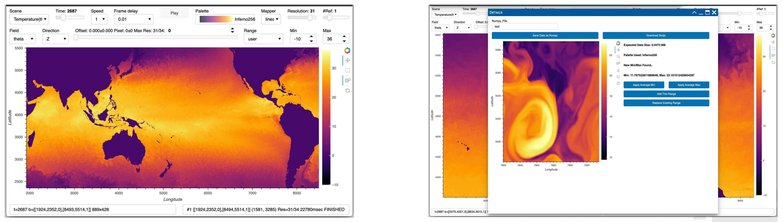

Fig. 1: Interactive visualization of sea surface temperature, along with inset for selecting the regions of interest (left). The dashboard provides the ability to directly download the data locally or to download a Python script that fetches the region from the cloud (right).

Workflow

1. Our dashboards are dynamically adjusted to fetch the data from the appropriate source.
2. For NASA scientists with access to Pleiades, a NASA supercomputer, the dashboard will fetch the data from their file system.
3. Other users will be fetching the data from the cloud, and this data will be cached as needed to reduce data transmission over networks.

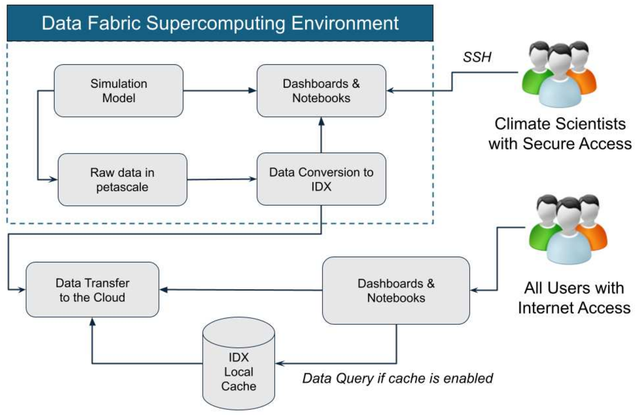

Fig 2: An example workflow showing data conversion and retrieval pipeline for Jupyter notebooks and dashboards from a supercomputing environment. Simulation data is typically restrictive due to being in raw format and requiring credentials, which impacts real-time capabilities. Our framework makes the data readily available for on-the-fly analysis and provides users with access to the data as needed through interactive dashboards and Jupyter notebooks.

Examples

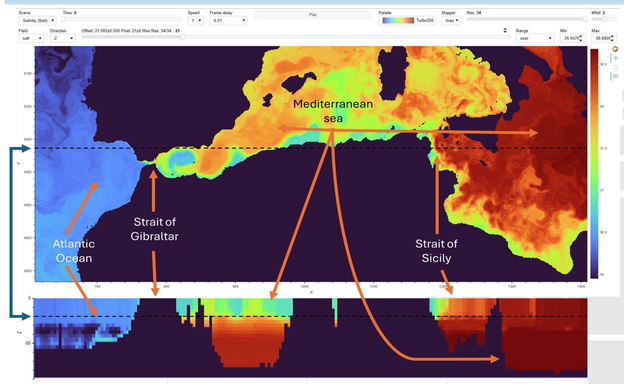

Fig 3: Zoomed-in view of the general water circulation through the Strait of Gibraltar connecting the Mediterranean with the Atlantic Ocean. The low-salinity water enters the Mediterranean Sea from the Atlantic through the Strait of Gibraltar. The salinity increases, and the water sinks as the current moves east. This type of on-the-fly selection of interesting regions from a massive dataset and playing through the time facilitates a deeper understanding of complex climatic phenomena, which was not practically accessible before the implementation of our framework.

Fig4: Interactive features available from the dashboards

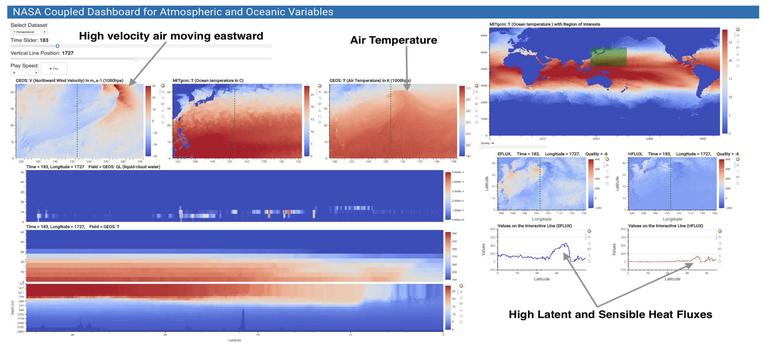

Fig. 5: Increasing heat fluxes (two plots and images at the bottom right) and air temperature (image at top middle) creating a high-velocity wind (image at top left) in the atmosphere moving eastward for the Kuroshio region

2026 IEEE SciVis Contest

- Challenge: Visualize NASA's petascale climate data, including DYAMOND LLC2160, ECCO LLC4320, and NEX GDDP CMIP6, for climate prediction and environmental insights.
- Tools: Leverage open-source platforms like OpenVisus, ParaView, Python, VTK, or anything you know.
- Prize: Compete for a **$1000 cash prize** and showcase cutting-edge visualization techniques!
- All data are publicly available! Link to the contest here: https://sciviscontest2026.github.io/

or Scan this:

Try our Dashboards

Scan these from your phone. to access the interactive dashboards with petabytes of climate data