Contribution ID: **14** Type: **Long Talk**

Exploring Climate Data Analysis with MATLAB in the ENES Data Space

Tuesday, 1 October 2024 15:15 (15 minutes)

The escalating volume and complexity of Earth and environmental data necessitate an effective, interdisciplinary partnership among scientists and data providers. Achieving this requires the utilization of research infrastructures that offer sophisticated e-services. These services enhance data integration and interoperability, enable seamless machine-to-machine data exchanges, and leverage High-Performance Computing (HPC) along with cloud capabilities.

In this presentation, we will demonstrate a case study focused on the import, analysis, and visualization of geodata within the ENES Data Space (https://enesdataspace.vm.fedcloud.eu), a cutting-edge cloud-enabled data science environment designed for climate data analysis. This platform is ingeniously constructed atop the European Open Science Cloud (EOSC) Compute Platform. By integrating with either an institutional or social media account, users gain entry to the ENES Data Space. Here, they can initiate JupyterLab, accessing a personal workspace equipped with computational resources, analytical tools, and pre-prepared climate datasets. These datasets, which include historical data recording and future projections, are primarily sourced from the CMIP (Coupled Model Intercomparison Project).

Our case study will utilize global precipitation data derived from the Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC) experiments, analyzed within the ENES workspace through two distinct approaches:

1. Direct MATLAB Online Integration: Users can launch MATLAB Online directly from the ENES Data Space JupyterLab. Utilizing a Live Script (.mlx), the process involves importing, filtering, and manipulating data, creating visual maps, comparing results, and conducting hypothesis testing to ascertain the statistical significance of the project findings. Live Scripts serve as interactive notebooks that facilitate the clear articulation of research methodologies and goals by integrating data, hyperlinks, figures, text, and code. These scripts also incorporate UI tools for intuitive, point-and-click data analysis and visualization, eliminating the need for extensive programming expertise.

2. MATLAB Kernel within Jupyter Notebook: This method demonstrates the analysis process using a MATLAB kernel executed from a Jupyter notebook (.ipynb) within the same JupyterLab environment.

In both scenarios, the results can be exported in multiple formats (e.g., PDF, markdown, LaTeX, etc.), allowing for easy downloading and sharing with other researchers, educators, and students. This entire workflow is seamlessly executed in MATLAB within the ENES Data Space, without the need for software installation or data downloads on local (non-cloud) devices. This case study exemplifies the power of cloud-based platforms in enhancing the accessibility, efficiency, and collaborative potential of climate data analysis.

Topic

Needs and solutions in scientific computing: Platforms and gateway

Primary authors: LEPTOKAROPOULOS, Kostas (MathWorks); Dr CHAKRABARTI, Shubo (MathWorks); ANTONIO, Fabrizio (CMCC)

Presenter: LEPTOKAROPOULOS, Kostas (MathWorks)

Session Classification: Unlocking the Potential of Environmental Data