Growing a community analysis platform with JASMIN and the Community Inter-comparison Suite

Stephen Pascoe¹, Phil Stier², Nick Schutgens², Alan Iwi¹

¹Centre of Environmental Data Archival, Rutherford Appleton Laboratory
²Atmospheric, Oceanographic and Planetary Physics, University of Oxford

The JASMIN Analysis Platform

CIS Extensible Design

CIS is designed to be extensible through plug-ins.

- Co-location algorithms
- Co-location plug-ins can be written by non-experts to meet the needs of particular science domains

CIS Tool

- Co-location
- Reduction
- Plotting

Data Products

User-Provided Extension Points

Future Extension Points in Phase II & III

Co-location Plug-in

Colocation is the spatio-temporal remapping of one dataset onto another. CIS uses a single customizable command (cis col) to accomplish this. Afterwards, another command (cis plot) can be used for comparison of the now colocated datasets.

Installation

Getting started with the platform is as simple as downloading a Virtual Machine image and running it on your laptop using the VM host software VirtualBox. For scaling out to a server of cluster you can configure any RPM-based Linux system to install the platform RPMs from our public repository. Our RPMs are built for 64-bit Red Hat Enterprise Linux 6 and should be compatible with RHEL derivatives including CentOS 6.4.

Moving to JASMIN

Using a consistent and portable set of tools allows researchers to move their analysis routines to the JASMIN infrastructure where they can take advantage of JASMIN’s peta-scale parallel data processing capabilities. The JASMIN Analysis Platform is installed across all JASMIN’s shared analysis VMs and the LOTUS cluster. It is available on request on any project Vms hosted on JASMIN.

More details at: http://goo.gl/BQKOd

JASMIN Community Inter-comparison Suite (CIS)

A central component of the JASMIN Analysis platform is the Community Inter-comparison suite (CIS). A high-level analysis tool developed for CEDA and Oxford University (AOPP), CIS enables inter-comparison of diverse atmospheric and EO datasets through a command-line and Python interface.

CIS is...

- Generic tool for visualising and co-locating datasets used in atmospheric sciences
- Handling of complex gridded and ungridded data in many formats (netCDF, hdf5, text,...)
- Simple syntax with many options
- Flexible approach through plug-ins
- Python source code readily available

CEDA is developing a software platform which enables researchers to use a consistent set of tools whether running their analyses at local research institutions or on JASMIN hardware.

The JASMIN Analysis Platform brings together community tools such as NCO, CDO, CF-Python and IRIS with open source tools SciPy, IPython, R and Octave.