Call: H2020-SC5-2014-two-stage

Topic: SC5-01-2014

PRIMAVERA

Grant Agreement 641727

PRocess-based climate sIMulation: AdVances in high resolution modelling and European climate Risk Assessment

Deliverable D9.2

Initial training for JASMIN users

Deliverable Title	*D9.2 Initial training for JASMIN users*	
Brief Description	*Deliver initial training for users of the JASMIN platform, to include on-site training course and web-based documentation, the latter to also support Open Data Access.*	
WP number	9	
Lead Beneficiary	*Matt Mizielinski Met Office*	
Contributors	*Matt Mizielinski (Met Office)* *Ag Stephens (STFC CEDA)* *Paul van der Linden (Met Office)*	
Creation Date Version Number Version Date	18-10-2016 1 18-10-2016	
Deliverable Due Date Actual Delivery Date	31-10-2016 31-10-2016	
Nature of the Deliverable	R	*R - Report*
		P - Prototype
		D - Demonstrator
	O	*O – Other – includes an online element*
Dissemination Level/ Audience	PU	*PU - Public*
		PP - Restricted to other programme participants, including the Commission services
		RE - Restricted to a group specified by the consortium, including the Commission services
		CO - Confidential, only for members of the consortium, including the Commission services

Version	**Date**	**Modified by**	**Comments**
0.5	10-10-16	PVDL	First version for WP9 internal comment
1.0	31-10-16	MM	Submission to EC

Table of Contents

1. Summary

This report describes the initial training made available to PRIMAVERA scientists on how to use the JASMIN system for project work. The training is available online as a number of linked, but self-contained lessons, which users can participate in at their own pace. The training was developed around the project's data flows and work plan (D9.1) with the aim of being clear and relevant to users' needs and it also included feedback in the development stage which led to an improved focus on the desired outcomes of the training.

2. Project Objectives

With this deliverable, the project has contributed to the achievement of the following objectives (DOA, Part B Section 1.1) WP numbers are in brackets:

No.	Objective	Yes	No
A	To develop a new generation of global high-resolution climate models. *(3, 4, 6)*		No
B	To develop new strategies and tools for evaluating global high-resolution climate models at a process level, and for quantifying the uncertainties in the predictions of regional climate. *(1, 2, 5, 9, 10)*	Yes	
C	To provide new high-resolution protocols and flagship simulations for the World Climate Research Programme (WCRP)'s Coupled Model Intercomparison Project (CMIP6) project, to inform the Intergovernmental Panel on Climate Change (IPCC) assessments and in support of emerging Climate Services. *(4, 6, 9)*		No
D	To explore the scientific and technological frontiers of capability in global climate modelling to provide guidance for the development of future generations of prediction systems, global climate and Earth System models (informing post-CMIP6 and beyond). *(3, 4)*	Yes	
E	To advance understanding of past and future, natural and anthropogenic, drivers of variability and changes in European climate, including high impact events, by exploiting new capabilities in high-resolution global climate modelling. *(1, 2, 5)*		No
F	To produce new, more robust and trustworthy projections of European climate for the next few decades based on improved global models and advances in process understanding. *(2, 3, 5, 6, 10)*		No
G	To engage with targeted end-user groups in key European economic sectors to strengthen their competitiveness, growth, resilience and ability by exploiting new scientific progress. *(10, 11)*		No
H	To establish cooperation between science and policy actions at European and international level, to support the development of effective climate change policies, optimize public decision making and increase capability to manage climate risks. *(5, 8, 10)*		No

3. Detailed Report

Introduction and scope

The purpose of this report is to document the scope and content of the training made available to project partners to help them use the JASMIN platform in PRIMAVERA. The JASMIN platform is the computing facility hosted by project partner STFC which allows high volume transfer and processing of data, and is a key component of the project. JASMIN provides both interactive and batch computing environments for scientists to develop and test workflows interactively before running those workflows efficiently at scale. The training should allow users (who are conversant with handling climate model data on computers) to:

- Learn how to set up and administer a user account
- Become familiar with the JASMIN system, its features and capabilities
- Understand how to make the best use of JASMIN storage
- Understand how to make the best use of JASMIN analysis tools
- Make appropriate use of the LOTUS batch and parallel processing tool
- Know how to make data transfers to and from JASMIN

All within the context of the climate data needs of the PRIMAVERA project

The PRIMAVERA Description of Action describes this deliverable as having a workshop element. However following discussions within WP9 and the science leads of the project it was decided to provide first the training online, where at the point of use it would be more relevant, and later follow this up with a workshop (possibly virtual) if project researchers deem this necessary.

Getting started

To access the online training users are directed to the following public web page: https://light.ceda.ac.uk/primavera/jasmin_docs/ where they are presented with the page in Figure 1. From this page they can access the six training modules as named in the left hand side of the web page. Each of these modules comprises a video of about 10 minutes duration along with written support available from the “Transcript” link below the video window. In subsequent lessons the text in the ‘Transcript’ box contains useful commands and scripts which can be cut and pasted by the user in to their work as they progress through the lesson. This allows them to observe and interact with JASMIN in real time with the lesson.

Also available from the home page are links to CEDA help pages on JASMIN, aspects of its use and related systems hosted by STFC.

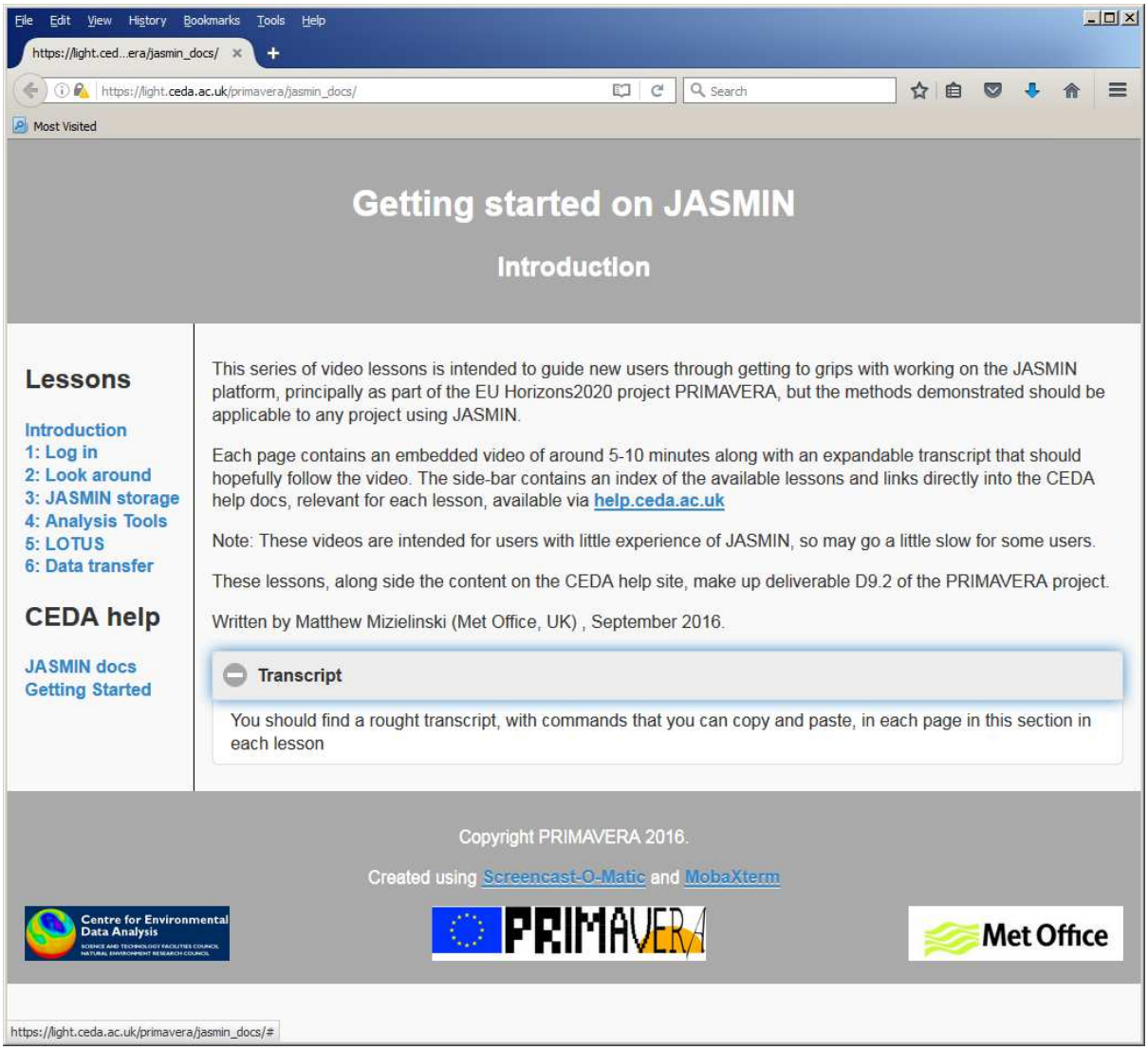

Figure 1: Homepage for the JASMIN training in PRIMAVERA.

Lesson 1: connecting to JASMIN

This lesson can be accessed as the first of the online tutorial via the home page, or directly at: https://light.ceda.ac.uk/primavera/jasmin_docs/lesson-1.html It covers logging in to the system using a line command secure remote access with relevant protocols enabled. A screenshot is shown in Figure 2.

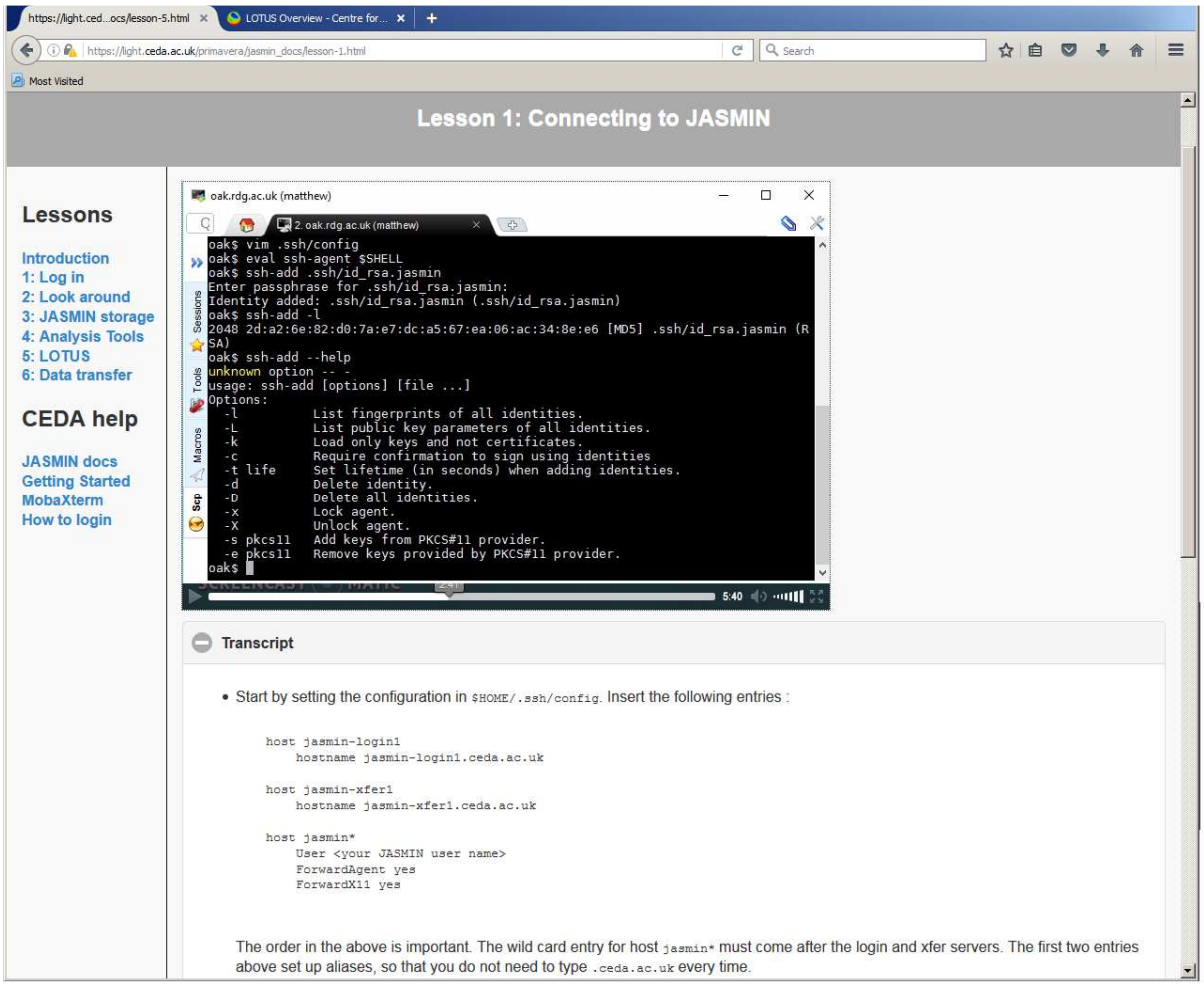

Figure 2: A screenshot of Lesson 1 “connecting to JASMIN”.

Lesson 2: familiarisation

Information about key commands for navigating around the system, and understanding the layout and structure for use are given in this tutorial. See Figure 3 for a screenshot of this webs page.

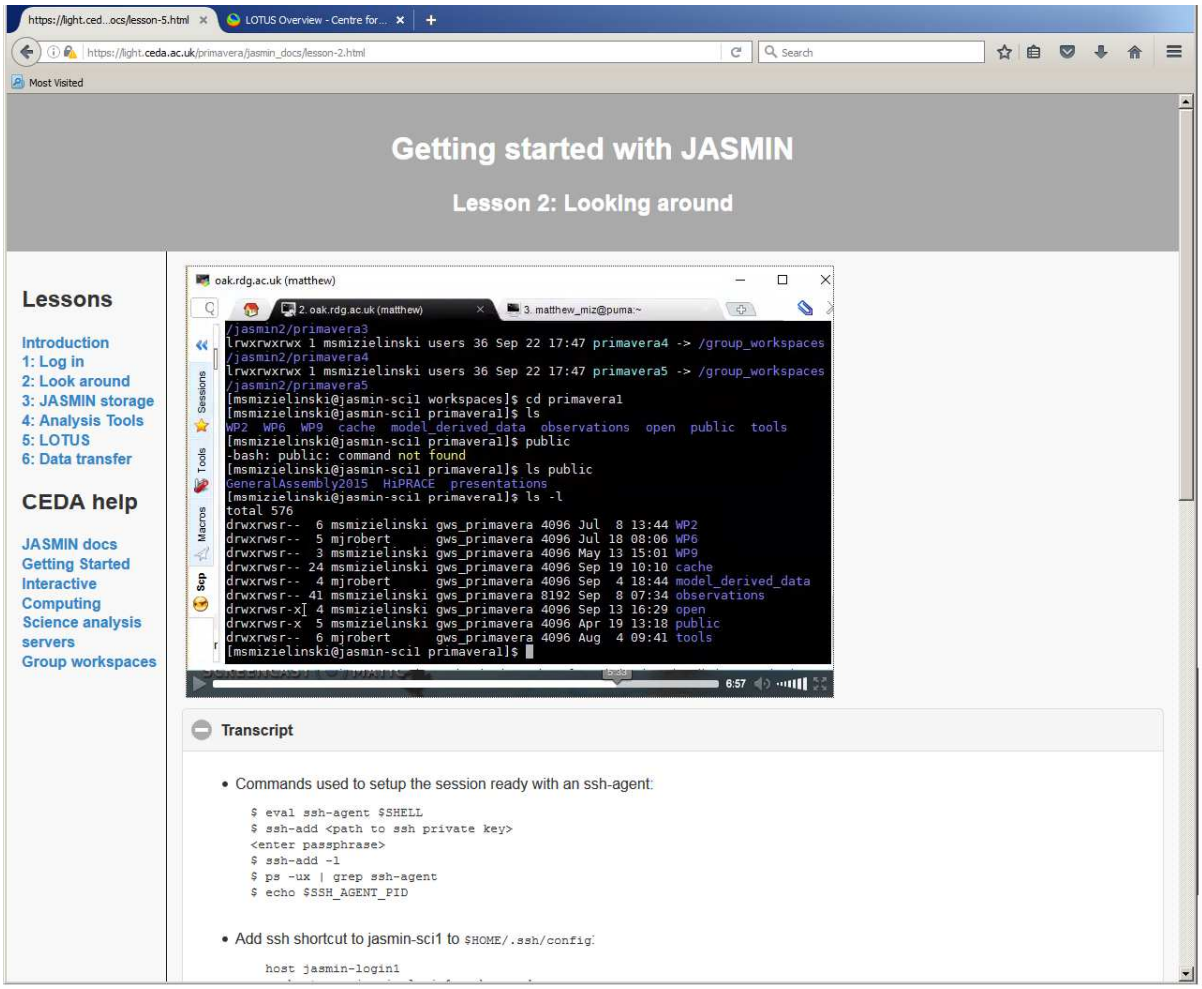

Figure 3: A screenshot of Lesson 2 “looking around” in the JASMIN training.

Lesson 3: storage

This lesson provides the user with information about the main commands used for accessing, understanding and using storage on JASMIN. This includes a description of the different storage types and how they can be optimised at different scales. A screenshot (Figure 4) of this lesson shows part of this demonstration. The lesson can be accessed directly at: https://light.ceda.ac.uk/primavera/jasmin_docs/lesson-3.html

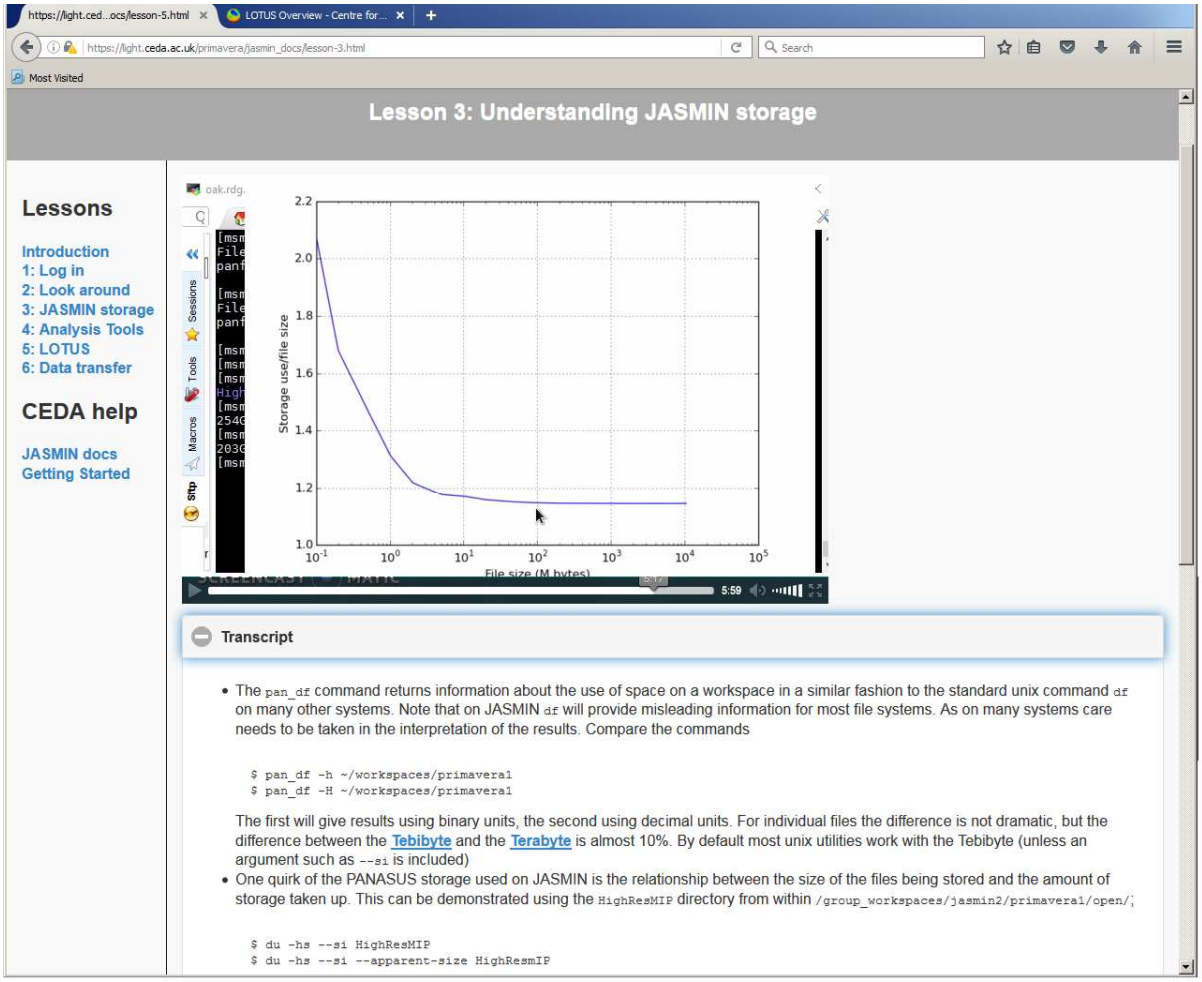

Figure 4: A screenshot of Lesson 3 “storage” in the JASMIN training.

Lesson 4: analysis tools

In Lesson 4 the user is familiarised with the suite of data analysis tools available on JASMIN including the standard set of installed software called JAP (JASMIN analysis platform). Finding information about which versions of software tools and which libraries are available is also given. The main tools of interest to PRIMAVERA researchers available on JAP are Python and R, and the video tutorial explains how to select the appropriate version, choose from available libraries (e.g. numpy, scipy, matplotlib, Sci Tools, idl) and run jobs remotely including the environment to run them in. Figure 5 shows a screenshot from Lesson 4.

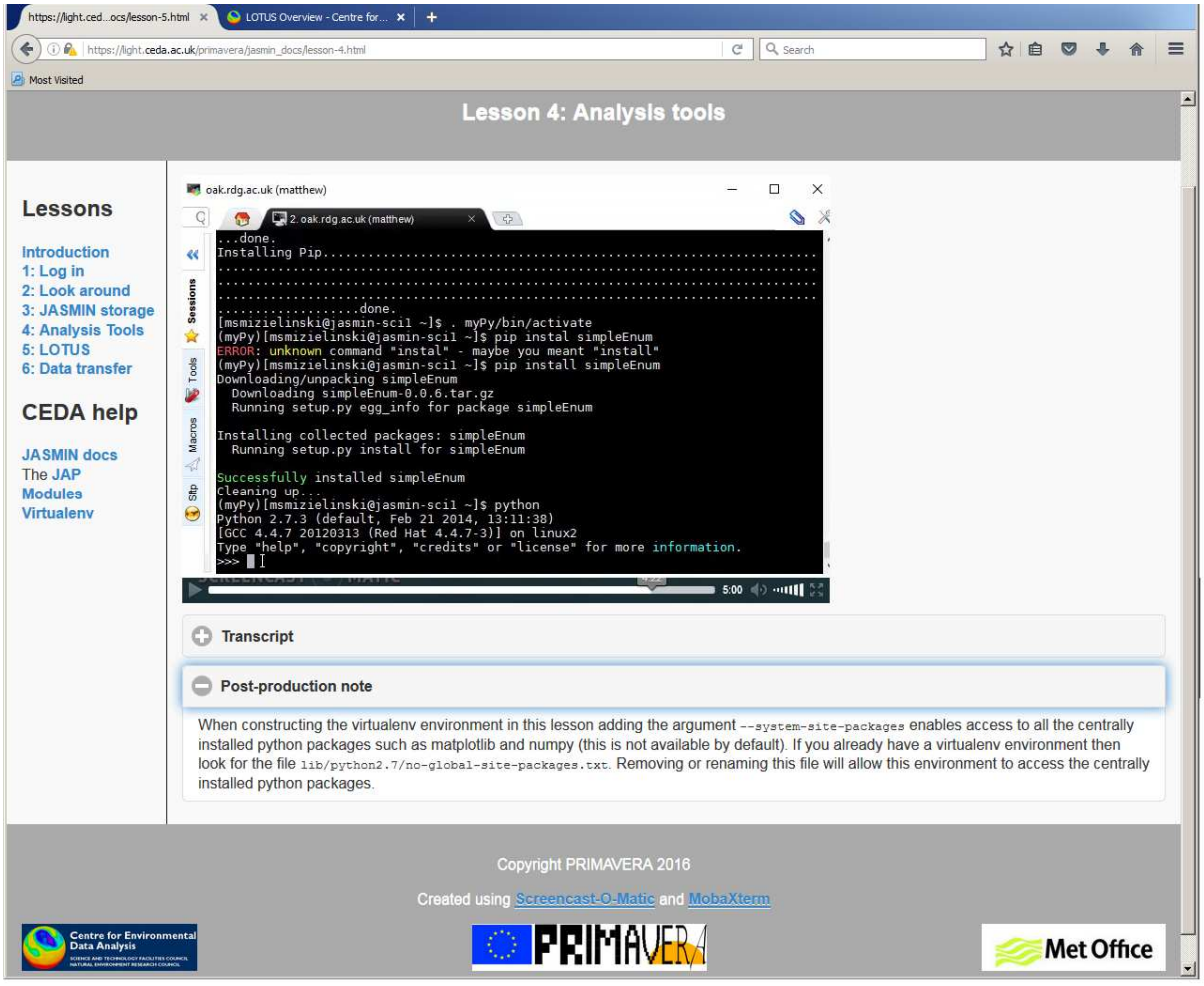

Figure 5: A screenshot of Lesson 4 “analysis tools” in the JASMIN training.

Lesson 5: LOTUS

LOTUS provides the batch and parallel processing component of the JASMIN data-intensive scientific analysis environment. It is a group of machines that enable efficient scheduling of larger data analysis tasks including parallel processing. Lesson 5 allows users to learn how to use LOTUS, and shows them it's main features. A screenshot of the lesson is shown in Figure 6.

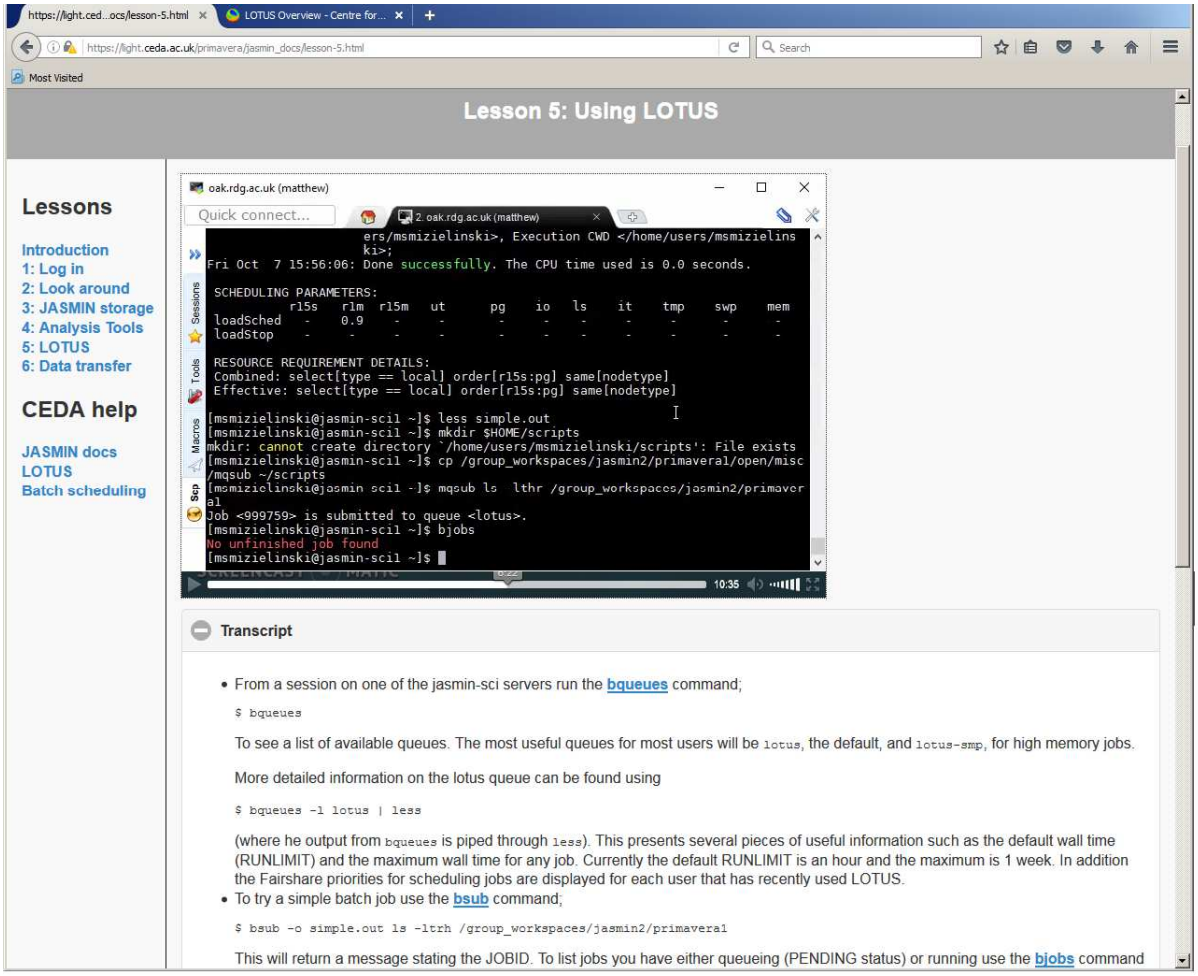

Figure 6: A screenshot of Lesson 5 "using LOTUS" in the JASMIN training.

Lesson 6: data transfer

This lesson demonstrates how to transfer data to or from a remote node (the example is from a UK based research institute) and covers the topics of synchronising files and directory structures; single-, parallel- and batch copying; and transfers. It gives examples of different situations and the commands best suited to use in them. A screenshot of the lesson is shown in Figure 7.

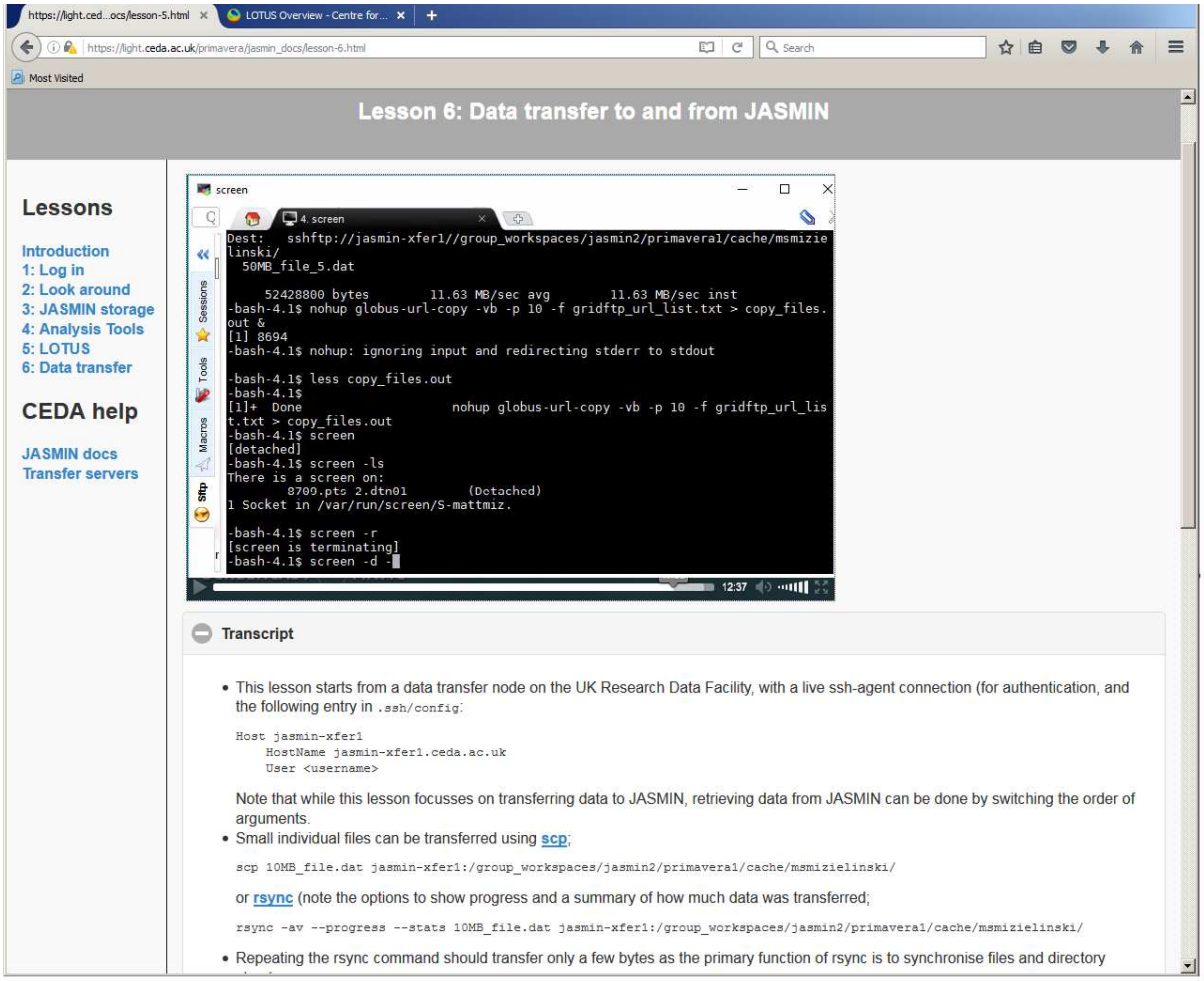

Figure 7: A screenshot of Lesson 6 “data transfer” in the JASMIN training.

Help for CEDA systems

There are links in the left hand panel of all of the web page lessons to relevant help provided by the CEDA team (the JASMIN providers and support group). These CEDA help files are narrative based web pages with more specific information about using JASMIN ranging from getting started, through detailed technical information, to general tips and advice.

User feedback

The onset and amount of use by PRIMAVERA researchers will be logged by the CEDA team allowing WP9 to judge if the system is being used by the project as envisaged. At an appropriate point an informal survey will be circulated to the PRIMAVERA research teams asking the for feedback on JASMIN and the training. These two things will be used to infor the future possible direction of this deliverable, such as including new lessons, revising the existing lessons, or holding a user workshop to give hands on training.

4. Lessons learnt

At the time of writing the modelling work has not started in PRIMAVERA, so there is no data to transfer/process on JASMIN. When this has happened there will be an opportunity to examine the activity and learn lessons.

5. Links Built

This report involved interactions between the modelling centres producing the data and CEDA who are hosting all of the data from the core simulations, plus other non-contractual data that are being produced. There are commonalities between the data centres that need to be understood (to avoid duplication of effort).

It was discussed at CEDA that these lessons prepared for RIMAVERA could form part of the formal CEDA user tutorials.