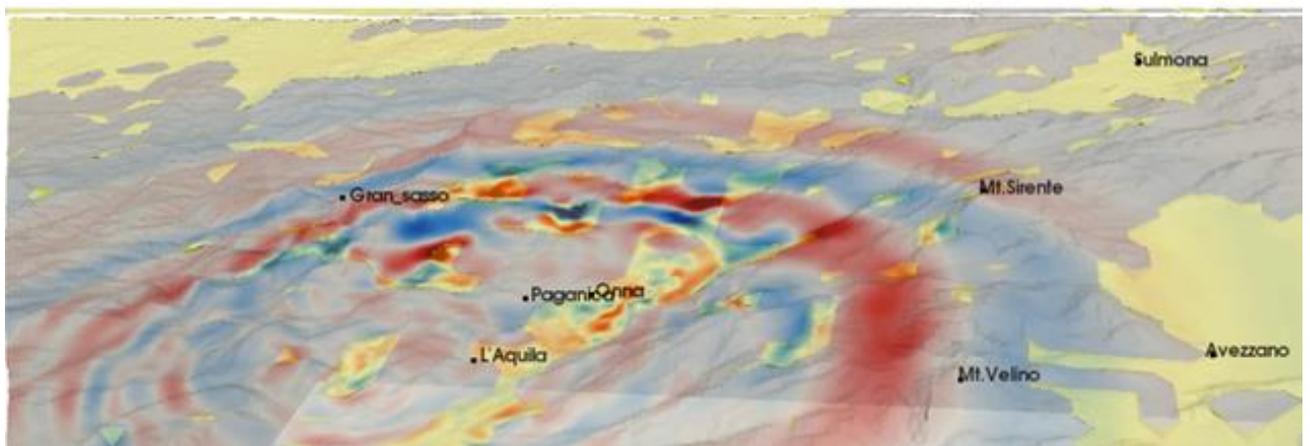




The VERCE portal – a user's guide



Version 1.1

September 2015

Contributors

The VERCE project was funded by the EU as an EU infrastructure project, involving a wide range of institutions from across the EU.



The following individuals have contributed material to this guide

T. Garth, A. Spinuso, F. Magnoni, E. Casarotti, A. Gemund, S. Hoon
Leong, J. Holt, L. Krishner, A. Kraus, R. Filgueira, M. Aktinson, H. Igel,
A. Rietbrock, A.

The cover image shows seismic waves propagating from an earthquake in central Italy.
Contributed by Federica Magnoni, INGV.

The VERCE portal – a user’s guide

Contents

Contributors	2
1. Introduction to the VERCE platform	6
2. Introduction to full waveform modelling.....	7
2.1 Why full waveform?.....	7
2.2 Calculating the wave field.....	7
2.3 High performance computing in seismic waveform modelling.....	7
2.4 What do I need to run a simulation?	8
3. Registering for the platform and certification.....	10
3.1 Registering for the platform	10
3.2 Getting a certificate	10
3.3 Installing your certificate in your browser.....	11
3.4 Registering for super computing and data resources.....	11
3.4.1 Registering for superMUC (and LRZ hosted services).....	11
3.4.2 Registering for SCAI.....	12
3.4.3 Registering for iRODS.....	12
3.5 Creating and uploading proxy certificates	12
3.5.1 Creating a proxy certificate with ‘My Proxy tool’	13
3.5.2 Creating a proxy certificate externally.....	13
3.5.3 Uploading a proxy certificate to the VERCE platform.....	15
3.5.4 Certificate Association	15
4. A tour of the VERCE platform.....	17
4.1 ‘Security’ tab.	17
4.2 ‘Forward Modelling’ tab	17
4.2.1 Group 0 - Basic inputs	18
4.2.2 Group 1 – Inverse problem	19
4.2.3 Group 2 – UTM projection	20
4.2.4 Group 3 – Attenuation	20
4.2.5 Group 4 – CPML absorbing conditions.....	21
4.2.6 Group 5 – Movie	22
4.2.7 Group 6 – Point source	23

4.2.7 Group 7 – Advanced.....	23
4.3 Statistics tab.....	26
4.4 The Meshes and models already uploaded.....	26
5. Running a simple simulation: a step by step guide.....	29
5.1 Preparing the portal.....	29
5.2 Selecting a solver, mesh and velocity model.....	29
5.3 Specifying the input parameters for SPECFEM.....	30
5.4. Selecting earthquakes.....	30
5.5. Selecting stations.....	31
5.6. Submitting the job.....	32
5.7. Monitoring the job.....	33
6. Accessing the results.....	34
6.1 Waveform outputs.....	36
6.2 Animation outputs.....	36
6.3 Other output.....	37
7. Running simulation using your own data.....	38
7.1 Creating your own velocity model.....	38
7.2 Creating a bespoke mesh for your area.....	39
7.2.1 Meshing parameters.....	39
7.2.2 Meshing software.....	40
7.3 Submitting a mesh and a velocity model.....	41
7.4 Submitting a new earthquake catalogue.....	42
7.5 Submitting a new station catalogue.....	43
VERCE glossary.....	44
Appendix one – using ObsPy.....	45
A1.1 Installing Python and ObsPy.....	45
A1.2 Other dependencies you need to run ObsPy are listed below;.....	46
A1.3 Using ObsPy.....	46
Appendix two – using dispel4py.....	47
A2.1 Installing Dispel4Py.....	47
A2.2 Using Dispel4Py.....	48
Appendix three – other VERCE supported resources.....	49
A3.1 ObsPy.....	49
A3.2 Pyflex.....	49

A3.3 LASIF.....	49
A3.4 SEIS-PROV	49
A3.5 dispel4py.....	50
A3.5 Whisper.....	50

1. Introduction to the VERCE platform

The VERCE portal is an online resource that allows large scale full waveform simulations to be easily run on a variety of high performance computers, and the associated large data sets to be managed easily. This resource is designed to make the ever increasing potential of full waveform seismological techniques available to a much wider spectrum of seismology and earthquake science communities.

This guide provides a brief description of the science and technologies behind the VERCE portal, and a step by step guide to using the portal. The written description of each step is complemented by online tutorial videos and presentations, and provides a stepping off point for users who want to use other tools supported by VERCE such as, 'Obspy', a python toolbox for dealing with seismic data, and 'Dispel4Py', a python toolbox designed to manage large data sets in seismology (and other disciplines). The resources supported by VERCE should then enable a seismologist of any specialism run large scale waveform simulations for an area of interest, and deal with large seismological data sets.

The VERCE project is an EU infrastructure project, and the development of the VERCE portal has been a collaborative project involving a wide range of project partners from across the EU.

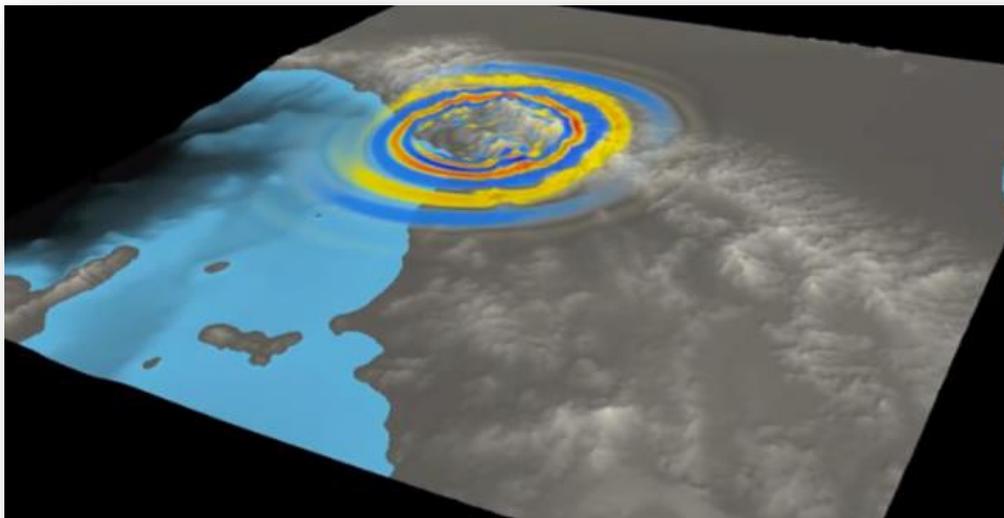


Figure 2.1: A snapshot of a simulation of the 5.2 MI Lunigiana Earthquake which occurred on the 21st June 2013 in Northern Italy. Image produced by the INGV.

2. Introduction to full waveform modelling

2.1 Why full waveform?

Much of the modern seismological methods used today rely on very little of the information contained within the seismogram. Earthquakes can be located using just the arrival time of the P- and S- waves, and this same information from many earthquakes in a given area can be used to invert for a travel time tomography. These methods however neglect the vast majority of the data within the waveform.

The seismic waveform recorded from an earthquake principally depends upon the seismic structure of the medium it passes through (i.e. the earth) , and the characteristics of the earthquake source itself. If this waveform data is utilised, it could therefore lead to a much greater resolution of seismic properties (e.g. velocity and attenuation) of the Earth, as well as better models of earthquake sources and ground motion from potentially hazardous earthquakes.

2.2 Calculating the wave field

In order to use the full waveform to constrain a seismic velocity model of the Earth, we need a way of calculating the response of the 3D velocity models we already have. The physics of how a seismic wave propagates through the earth is well understood, and is described by the wave equation.

A number of different methods exist to solve the wave equation (or its constituent equations) in order to model the propagation of a waveform from its source through a velocity model to a receiver.-In order to do this, the modelled area is broken up into a grid of points with given seismic properties, and the motion of each point in that grid is calculated at each time step.

The spacing of these points is controlled by the velocity model used and the frequency of the seismic wave that is modelled, with tighter grid spacing and a smaller time step being required to model higher frequency seismic waves.

The time step and grid spacing are also dependent on the method used to solve the wave propagation. The codes supported within the VERCE platform use the spectral element method to solve the constituent equations of the wave equation, hence calculating the response of the seismic wave-field to the velocity and attenuation parameters of the input model. Details on how to ensure that the model is stable to the frequencies required are given in section 6 of this guide.

2.3 High performance computing in seismic waveform modelling

The main problem/issue with this sort of simulation is that calculating the motion of a tight grid of points at thousands of time steps is extremely computationally expensive. For this reason both 2D and 3D waveform models are run on high performance computers. These simulations are run in parallel, meaning that the simulation runs on many different processing cores. Simulations run through the VERCE platform can be run on a range of different super computers from across Europe. The supercomputer pictured in figure 2.1 is superMUC, hosted by the LRZ in Munich, Germany.



Figure 2.1: Inside superMUC, one of the high performance computers which is available through the VERCE platform. SuperMUC is hosted at the LRZ, Munich.

2.4 What do I need to run a simulation?

In order to run a full waveform model you need to know the following details.

i. Velocity model

In order to calculate the wave field you need to know the velocity structure of the area of interest. The velocity models already available through the portal are all based on published travel time tomography models, but it is also possible to load your own velocity models into the portal.

ii. Mesh

The mesh is a grid of points within the volume that is to be modelled. The mesh takes into account the changes in required resolution with depth, but also the topography and bathymetry of the area being modelled. Several meshes are already loaded into the portal for specific regions. However, if you wish to run models in a new area you will need to create and submit a new mesh and velocity model. Details of how to do this are given in section 6 of this guide.

iii. Focal mechanisms

For any earthquake you want to model you must know the hypo-central location as well as the full focal mechanism. The VERCE portal allows you to download focal mechanisms from the gCMT catalogue and other local earthquake mechanism catalogues, but it is also possible to load in your own focal mechanisms as described in section 6.

iv. Station locations

Again there are several seismic networks already loaded into the VERCE platform. But if there are any networks or station locations that are not pre-loaded to the portal, you can also add your own station locations.

All of these are provided for certain study areas, such as the area we consider in in the simple example in part 5 of this guide. You can however use your own velocity model, your own focal mechanisms or add additional station locations that are not already included in the portal.

If you are working on a study area that is not currently supported in the VERCE portal you will also need to create a new mesh, velocity model, station locations and possible earthquake focal mechanisms (although there are global catalogues supported in the platform).

3. Registering for the platform and certification

The VERCE platform can be accessed through a normal web browser. The portal has been well tested with browsers such as *Mozilla Firefox* and *Safari*, and should also work with *Google Chrome*. At this stage the portal doesn't support *Internet Explorer* or *Microsoft Edge*.

3.1 Registering for the platform

Registering for the portal is exceptionally easy. Simply go to the VERCE portal website on the link below and click the **SIGN UP** icon. Fill in your details ensuring that you use your university email address (i.e. your academic email address) if at all possible. This can take a one or two working days to be processed so please leave yourself time for this.

<https://portal.verce.eu/home>

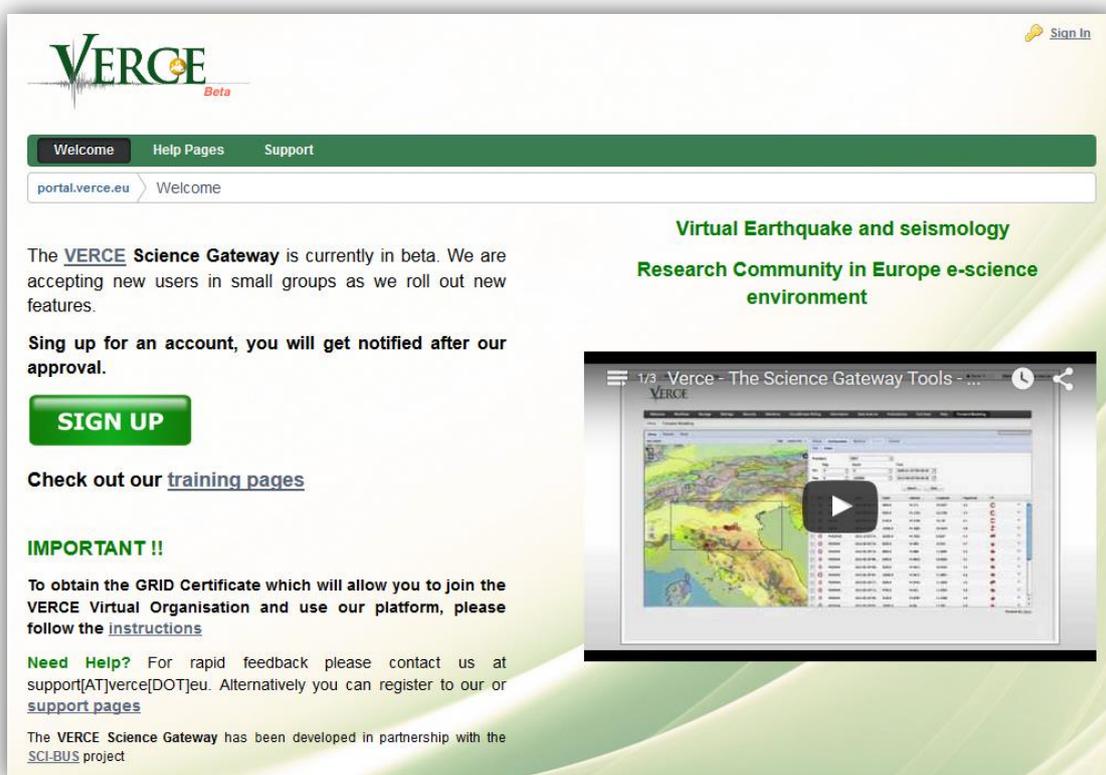


Figure 3.1: The VERCE sign up page

3.2 Getting a certificate

In addition to registering for the platform you must also get an e-science certificate from the relevant authority. This is an internationally recognised certification scheme that will allow you to submit simulations to a wide range of supercomputers. Without an e-science certificate you will be able to log into the portal website, but you will not be able to submit jobs to any of the supercomputers, or access any data through the portal.

The procedure for getting an e-science certificate depends on the where you are based as the e-science certificates are distributed by national certification authorities. First you must request a e-science certificate from your national provider, following the instructions on the relevant website,

- For the UK: <http://www.ngs.ac.uk/ukca/certificates>
- For Germany: https://pki.pca.dfn.de/grid-root-ca/cgi-bin/pub/pki?RA_ID=101
- For the Netherlands: <http://ca.dutchgrid.nl/>
- For France: <https://igc.services.cnrs.fr/GRID2-FR>
- For Italy: <http://security.fi.infn.it/CA>

You then have to go to an administrator (usually in your university) who confirms who you are, and (hopefully) approves the certificate. This can also take a day or two, so please leave time for this.

It is important to back up this certificate in a different location to the host computer (i.e. the computer that you used to request and download the certificate). Your local certifying authority will provide full details of how to do this.

3.3 Installing your certificate in your browser

To make the next stages of registration easier it is recommended that the user installs their new certificate to the browser. Ideally this should be done on the computer you are most likely to be using the VERCE portal from. The certificate must first be exported from the certificate manager (see instructions from your national certifying authority). Both the 'Private Key' and the Certificate should be exported in 'PKCS#12' format. The certificate is exported to the local machine, and is protected with a new password.

To install the certificate on Mozilla you must first select '*Options/Preferences*' from the menu button (☰) in the top right of the browser. Select the '*Advanced*' tab from the left hand panel and then the '*Certificates*' tab under the '*Advanced*' menu. Finally select '*View Certificates*' and click '*Import*' to upload your certificate from its location on your local machine using the new password set when exporting the certificate.

Other browsers can be used (see start of section for browser options), but the procedure for uploading the certificate will vary. You must then ensure that you use this browser for the validation steps outlined below, and ideally when you log in to the portal.

3.4 Registering for super computing and data resources

The VERCE portal is currently hosted by super computers on two sites. The first is the LRZ in Germany who run '*superMUC*', '*coolMUC*', and '*Linux cluster*' which are all available through the platform. The second site is CINECA in Italy, who run '*SCAI*'. All of these machines are part of the PRACE initiative, and this is the most likely source of extra resources should they be needed.

3.4.1 Registering for superMUC (and LRZ hosted services)

Send the following information Cerlane Leong (siew-hoon.leong@lrz.de):

- Title
- First Name
- Last Name
- Nationality
- Address
- Telephone Number
- Email address
- Certificate DN (distinguished name, also called subject) of your certificate

The LRZ will then email you to confirm the details of your certificate. Use the link they provide to confirm your registration and follow the instructions provided in the email. If the certificate has been installed correctly Firefox will ask you to confirm the certificate and you will be able to continue. Otherwise, the user certificate is either not valid or has been installed incorrectly.

3.4.2 Registering for SCAI

To register for SCAI go to <https://userdb.hpc.cineca.it/user/register> and fill out your details (similar to above). The SCAI administrators will then email you to confirm your identity, by following a link in their email. Again the certificate must be correctly installed in your browser.

3.4.3 Registering for iRODS

Lastly please send the same information listed in section 3.4.1 to André Gemünd (andre.gemuend@scai.fraunhofer.de), and request to be registered for iRODS. iRODS is a suite of data managements software that is embedded within the VERCE platform, and allows you to easily access your data regardless of where you submitted your simulation. Andre will be able to register you for this service with the data provided to him from the step above.

3.5 Creating and uploading proxy certificates

Once you have got your national e-science certificate, you will need to load a proxy certificate up to the portal to allow you to access data and submit jobs to the available supercomputers. A proxy certificate is essentially a copy of the full certificate that will expire after a short time period (usually 24 hours). This allows you to upload and use your certificate, while the limited life span of the proxy minimises the risk of the certificate falling into the wrong hands.

There are a number of ways of creating a proxy certificate and loading into the portal, and the easiest method depends on the national certifying authority that provided your certificate.

1. **Using My Proxy tool:** You can directly create a proxy through 'My Proxy tool' which is embedded within the portal. This is the easiest way to upload a proxy certificate if you have a German, Dutch or EGI hosted certificate. The procedure for this is described below in section 3.4.1.
2. **Creating a proxy externally:** Alternatively, you can create the proxy certificate through the certificate managing software of your national provider, and upload the certificate to the VERCE portal manually. The procedure for this is described below in section 3.4.2.

3.5.1 Creating a proxy certificate with 'My Proxy tool'

This is by far the simplest and easiest overall method to create a proxy certificate to use within the VERCE portal, and should be the same regardless of the CA (certificating authority) that provided your certificate.

Before starting, make sure that you have your certificate downloaded on the machine you are using. You also need to have Java 1.7 or later installed on your machine. Under the security tab, select the 'MyProxy Tool' option. If your Java is up to date it should bring up an applet as shown below. If not, please update your Java library.

Please select the CA where you wish to upload your proxy certificate in the server, set the port number to 7512, and enter the user name and password of your certificate. The applet will then look for your certificate in the local .globus directory in your machine to create a proxy certificate from it. If the proxy certificate does not successfully upload, please change the port number to 21 and retry.

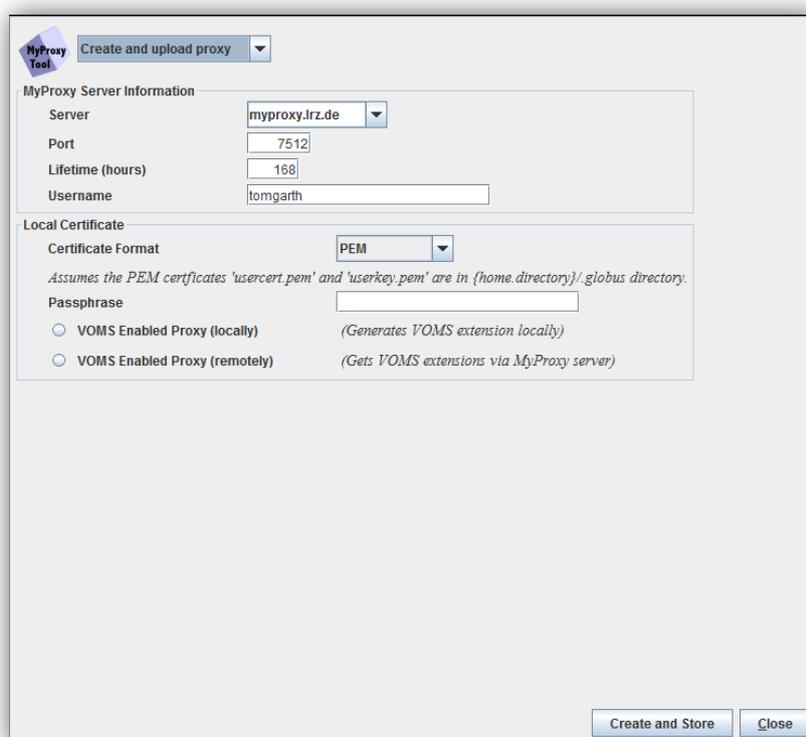


Figure 3.2: The 'MyProxy Tool' Java applet.

3.5.2 Creating a proxy certificate externally

First you have to create a proxy certificate using the certificate managing software of your certificate host. This guide demonstrates how to do this with the software used by the NGS (the UK certificating authority), which is called 'CertWizard', but other national certificate managers should be very similar. When downloading the proxy certificate the procedure will be the same, but clearly a host name must be used.

Once you have successfully requested your e-science certificate you must install it into the certificate wizard, by clicking 'install' once you have selected the appropriate certificate under the 'Apply For/Manage Your Certificate' tab of the Certificate wizard.

You can then navigate to the 'Use Your Certificate' tab at the top of the certificate manager and select the 'MyProxy' tab on the bottom panel.

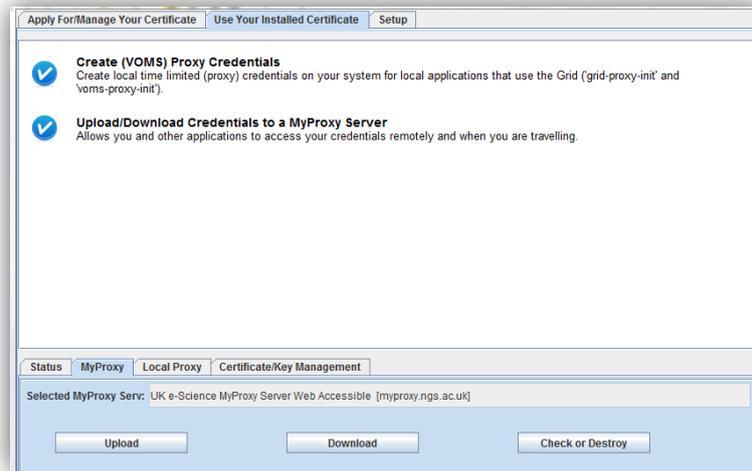


Figure 3.3: The 'Use Your Installed Certificate' tab of the UK Certificate wizard app.

This will bring up the pop up panel shown below. You must then input the password for your certificate under '**Private Key Password**', and create a new user name and password for the proxy certificate which can be entered under '**MyProxy Username**' and '**MyProxy Password**'. Click 'OK' and the proxy certificate will be uploaded to the proxy server, in this case at the UK NGS (national grid service).

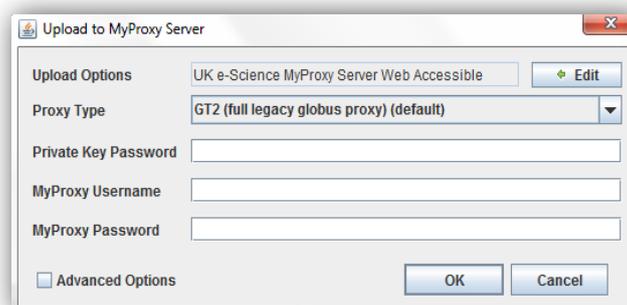


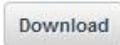
Figure 3.4: Proxy server input form.

If you wish to use another proxy server for any reason, you can manually change this using the '**edit**' tab. Alternatively, if you consistently want to use a different server you can change the default proxy server under the '**setup**' tab at the top right of the Certificate wizard.

3.5.3 Uploading a proxy certificate to the VERCE platform

Once you have created your proxy certificate using one of the methods above, you need to load the certificate into the VERCE platform in order to be able to submit jobs and access the data.

To do this you need to navigate to the certificate page under the security tab. On this page click the

 button which will bring you to the proxy certificate upload page shown below.

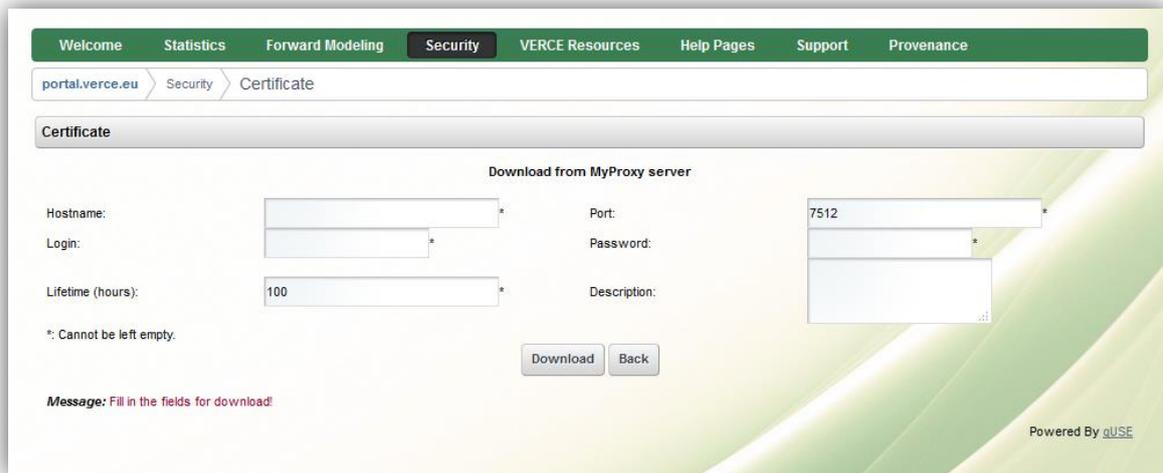


Figure 3.5: The proxy certificate upload page.

Here you need to enter the address of the institution hosting your proxy certificate in the **'hostname'** box (e.g. myproxy.lrz.de). The username and password you set for your proxy certificate must then be entered in the **'Login'** and **'Password'** boxes respectively. Clicking the  button then adds the proxy certificate to the portal, allowing you to access the high performance computing (HPC) resources and data.

The proxy certificate will only be valid for up to 24 hours, so you will need to repeat this process of creating and uploading the proxy certificate every time you wish to use the VERCE portal to run simulations or access data.

3.5.4 Certificate Association

Once you have successfully uploaded your certificate, you must associate the certificate with the platform (verce.eu) and any resources you intend to use in this session (e.g. super computing resources). This is done by clicking, security > certificate > associate VO. The proxy certificate will then authenticate you as a user, and allow you to access the HPC and memory resources.

First navigate to the Certificate page under the security tab. Here you will see details of the proxy certificate you have just uploaded as shown below. Click the 'Associate to VO' button to bring up the page shown in figure 3.6. You can now select the resource you wish to associate your proxy certificate to from the drop down menu located below your certificate details. You can see in both figure 3.5 and figure 3.6 that the proxy certificate here is associated to verce.eu, SuperMUC and SCAI_Cluster2.

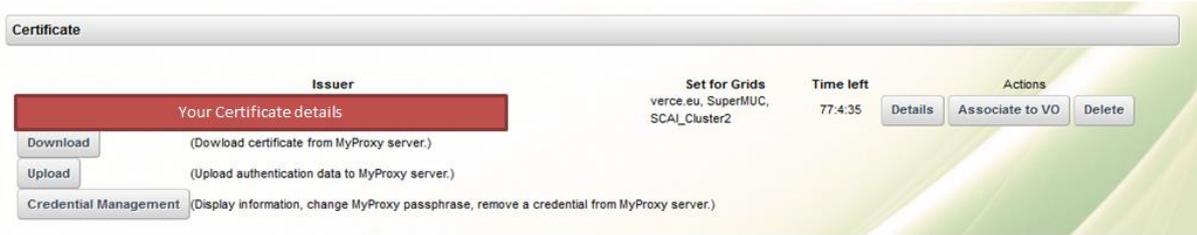


Figure 3.5: The proxy certificate loaded into the certificates page of the portal. This proxy certificate has been associated with the VERCE portal (verce.eu) as well as the SuperMUC and SCAI computing resources.

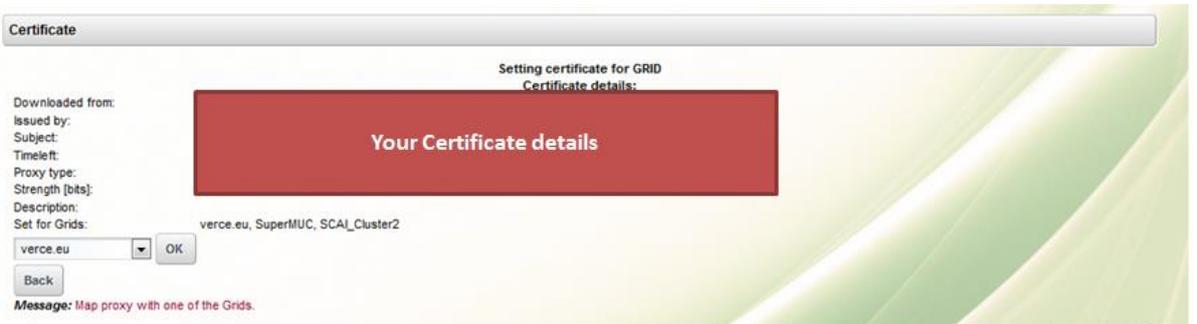


Figure 3.6: The certificate association page. The resource to which the certificate may be associated are listed in the drop down box shown in the bottom right of the figure.

4. A tour of the VERCE platform



In this section we will introduce the different parts of the VERCE platform, and show the models and meshes that are already loaded into the VERCE portal.

4.1 'Security' tab.



We have already used many of the features available in this tab in order to upload a proxy certificate. The main security page though gives an overview of how to register for and get certificated for the platform, as covered in section 3 of this guide. The main tools you need to be aware of are the **'MyProxy tool'** (section 3.4.1) and the **'Certificate'** upload tab (section 3.4.3).

4.2 'Forward Modelling' tab



The forward modelling tab is the main feature of the portal, and it is from here that you can setup and run full waveform simulations. Jobs can be run from the setup tab shown in figure 4.1 below. On the right hand side of the forward modelling panel, existing pre-loaded mesh and associated velocity model can be selected from the drop down menu. Earthquake sources and seismic stations can be selected from the catalogues that are pre-installed into the portal under the 'Earthquakes' and 'Stations' tabs respectively. This process is described in more detail later in section 5 of this guide.

Alternatively, you can add your own velocity and mesh using the blue link below the drop down boxes in figure 4.1. You can then add in your own earthquake focal mechanisms and station locations. Details on how to create and submit a more advanced bespoke job like this are given in section 7 of this guide.

The left hand side of the panel shows a summary map of the area you are running your model for, currently showing the default view of Europe. The map also shows details of existing geological maps, hazard maps and fault traces. The relative weight of these can be adjusted using the drop down menu from the layer info button in the top right of the map area.

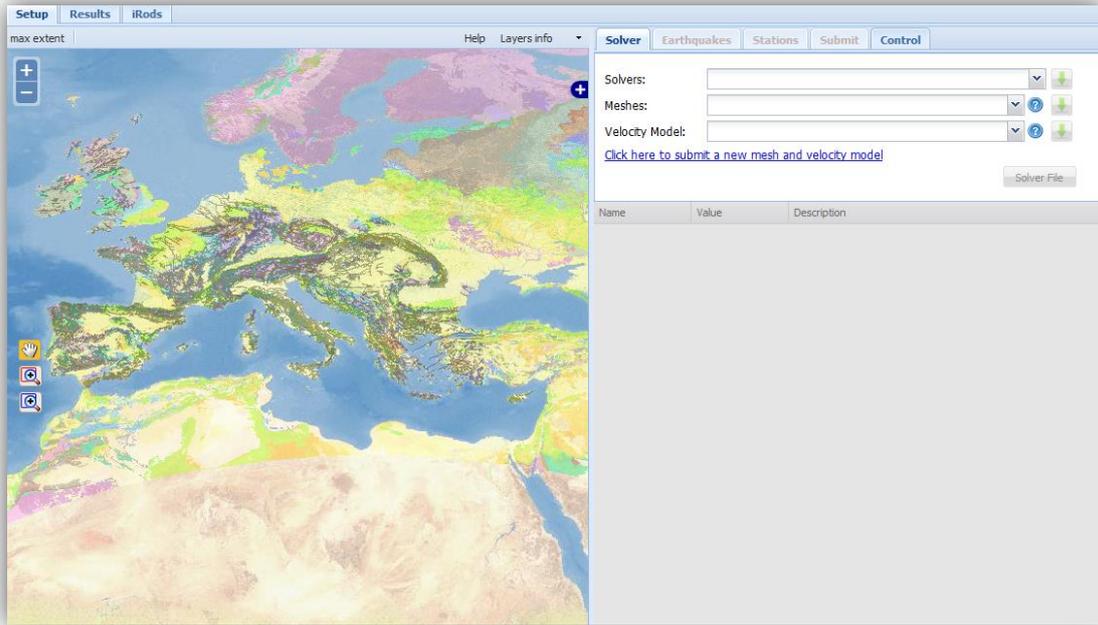


Figure 4.1: The forward modelling interface 'setup' page.

4.2.1 Group 0 - Basic inputs

Name	Value	Description
group: 0 - Basic		
NPROC	256	number of MPI processors
NSTEP	60000	The number of time steps
DT	0.001	The time increment in seconds
MODEL	default	setup the geological models, options are: default (model para...
STACEY_ABSOR...	<input checked="" type="checkbox"/>	Stacey absorbing boundary conditions for a regional simulation
GPU_MODE	<input type="checkbox"/>	set .true. for GPU support

Figure 4.2: Parameter form for 'Group 0 - Basic inputs'.

NPROC is the number of processors that the simulation is run on. This is essentially dependent upon the high performance computer, and work flow you intend to submit your job to.

NSTEP is the number of time steps that you want to run your model for. This should be set so that $(NSTEP * DT)$ is equal to the time in seconds you want to simulate. So the model setup shown above will run the simulation for 60 seconds, and provide synthetic seismograms for 60 seconds after the origin time of the simulated earthquake.

DT is the time step used in the solver. This must be small enough to ensure that the simulated waveform is properly sampled. The equations this is based on are given in section 6 of this guide. For the meshes and models that are already available in the portal though, the recommended DT is given in figure 4.4.

MODEL allows you to select the velocity model that is used in the simulation. Leaving this to default will select the 3D velocity model that is specified in the drop down menu next to ‘Velocity Model’, at the top of the input parameters panel. It is however also possible to select from a range of 1D models that are pre-loaded into the solver, SPECIFEM.

STACEY_ABSORBING_CONDITIONS is selected to activate absorbing boundary conditions on the sides and bottom of the simulated areas. This is designed to prevent artificial reflections from the model edges from affecting the simulated waveforms. An alternative method for suppressing these side reflections is using CPML absorbing conditions which can be set in group 4 of the input parameters as described in section 4.2.5. For a full discussion of the relative merits of the two methods, please see the SPECIFEM3D Cartesian manual.

GPU_MODE allows SPECIFEM to be run on high performance computers that use graphical processing units (GPUs) rather than the more conventional CPU (central processing unit). All the workflows currently available on the VERCE platform use CPUs, so you should always leave this box unchecked.

4.2.2 Group 1 – Inverse problem

Name	Value	Description
group: 1 - Inverse problem		
SIMULATION_TYPE	forward	forward or adjoint simulation, 1 = forward, 2 = adjoint, 3 = both simultaneously
NOISE_TOMOGRAPHY	earthquake	noise tomography simulation, 0 = earthquake simulation, 1/2/3 = three steps in noise simulation
SAVE_FORWARD	<input type="checkbox"/>	save forward wavefield

Figure 4.3: Parameter form for ‘Group 1 – Inverse problem’.

In addition to calculating the wavefield from an earthquake source (referred to as a ‘forward model’), SPECIFEM can also be used to calculate the adjoint wavefield, as well as being able to simulate noise sources for ambient noise tomography applications. These options are controlled by this group of parameters.

SIMULATION_TYPE is set to ‘forward’ to model the wavefield from an earthquake.

NOISE_TOMOGRAPHY is set to ‘earthquake simulation’ as the noise tomography applications of SPECIFEM are not currently supported within the VERCE platform.

SAVE_FORWARD is selected if the last step of the wavefield is to be saved. This enables a simulation to be started at this point, but requires a large amount of storage space, and so should not be checked in this example.

4.2.3 Group 2 – UTM projection

As SPECIFEM3D Cartesian uses, unsurprisingly, Cartesian coordinates, you must specify the UTM zone that your model falls in. this is described in more detail in section 6 when we consider uploading new meshes and models. For the pre-loaded meshes and models though the correct Cartesian coordinates are given by the table shown in figure 4.5, and Be set correctly by default when the mesh is selected.

Name	Value	Description
group: 2 - UTM Projection		
UTM_PROJECTION_ZONE	32	set up the utm zone, if SUPPRESS_UTM_PROJECTION is false
SUPPRESS_UTM_PROJECTION	<input type="checkbox"/>	suppress the utm projection

Figure 4.4: Parameter form for ‘Group 2 – UTM projection’.

UTM_PROJECTION_ZONE is where the UTM zone is specified.

SUPPRESS_UTM_PROJECTION is not enabled in the VERCE platform, meaning the model ranger must always be specified in geographical coordinates (not Cartesian coordinates).

4.2.4 Group 3 – Attenuation

In the Earth seismic waves are attenuated by the visco-elastic deformation as the wave propagates. If we are to gain simulated seismic waves with a similar amplitude to the recorded waves, we must include this attenuation in our waveform simulation. This

Name	Value	Description
group: 3 - Attenuation		
ATTENUATION	<input type="checkbox"/>	see SPECIFEM3D_CARTESIAN manual
FULL_ATTENUATION_SOLID	<input type="checkbox"/>	see SPECIFEM3D_CARTESIAN manual
USE_OLSEN_ATTENUATION	<input checked="" type="checkbox"/>	use the Olsen attenuation, $Q_{\mu} = \text{constant} * v_s$ attenuation rule
OLSEN_ATTENUATION_RATIO	0.05	see SPECIFEM3D_CARTESIAN manual

Figure 4.5: Parameter form for ‘Group 3 – Attenuation’.

ATTENUATION controls whether attenuation is simulated or not. Turning attenuation on means that extra variables are generated, and therefore will increase the time taken for the simulation to run.

FULL_ATTENUATION_SOLID alters how attenuation of the wavefield is calculated, but has an insignificant effect unless the wavefield is being calculated in a fluid medium. This option therefore should not be selected for any earth models and all models currently loaded into the VERCE platform.

USE_OLSEN_ATTENUATION_RATIO can be used to define the attenuation model from the S-wave velocity using the empirical relationship proposed by Olsen et al. (2003).

OLSEN_ATTENUATION_RATIO is defined here and should be in the range or 0.02-0.1.

4.2.5 Group 4 – CPML absorbing conditions

CPMLs, or ‘convolutional perfectly matched layers’ are an absorbing boundary condition that can be used instead of the Stacey absorbing conditions set in parameter group 0. CPMLs are the most effective and therefore computationally efficient boundary condition, and should be considered for all new meshes that are uploaded. It is especially important that they are used in models where you are particularly worried about side reflections (e.g. models where receivers or particularly sources are very close to the model edge).

Name	Value	Description
group: 4 - CPML absorbing condition		
PML_CONDITIONS	<input type="checkbox"/>	C-PML boundary conditions for a regional simulation
PML_INSTEAD_OF_FREE_SURFACE	<input type="checkbox"/>	see SPECSEM3D_CARTESIAN manual
f0_FOR_PML	12.7	C-PML dominant frequency, see manual
PML_WIDTH_MIN	10000	C-PML thickness min
PML_WIDTH_MAX	25000	C-PML thickness max
ROTATE_PML_ACTIVATE	<input type="checkbox"/>	
ROTATE_PML_ANGLE	0	

Figure 4.6: Parameter form for ‘Group 4 – cPML absorbing conditions’.

PML_CONDITIONS select whether PMLs are implemented. Please ensure that Stacey absorbing conditions are unchecked (in group 0) if you do this.

PML_INSTEAD_OF_FREE_SURFACE replaces the free surface at the top of the model with a PML absorbing layer. This can be useful if you are simulating a deep model, rather than a model that includes the Earth’s surface.

f0_FOR_PML is the dominant frequency, or the frequency at which the PML will be the most effective. It should therefore be set to the dominant frequency of the waveforms being simulated. This largely depends on the resolution of your model but for the pre-loaded meshes should generally be set to 0.25 Hz

4.2.6 Group 5 – Movie

One of the SPECfEM which can be requested through the VERCE platform is a movie of the waveform simulation. This is usually output on a 2D surface such as a slice through the model or the surface topography of the model, but can be output for the whole 3D volume. This last option is extremely demanding on memory though, and not recommended for normal simulations.

Name	Value	Description
group: 5 - Movie		
MOVIE_SURFACE	<input type="checkbox"/>	save velocity snapshot files only for surfaces
MOVIE_TYPE	Show top s	Set this flag to 1 to show the top surface only, to 2 to show all external faces of the mesh in s...
MOVIE_VOLUME	<input type="checkbox"/>	save the entire volumetric velocity snapshot files
USE_HIGHRES_FOR_MOVIES	<input type="checkbox"/>	save high resolution snapshot files (all GLL points)
NTSTEP_BETWEEN_FRAMES	200	number of timesteps between 2 consecutive snapshots
HDUR_MOVIE	1	half duration for snapshot files

Figure 4.7: Parameter form for ‘Group 5 – Movie’.

MOVIE_SURFACE sets the output movie for just the surfaces you define in MOVIE_TYPE.

MOVIE_TYPE selects whether the movie is generated for the top surface of the model (e.g. how the seismic waves interact with the surface topography of the model), or the side surface of the model (e.g. a vertical slice through the Earth)

MOVIE_VOLUME allows 3D snapshots of the entire model area to be output. This would allow the entire wavefield to be imaged, which could be useful. But this would also be hugely demanding on memory and so should be left unchecked by default.

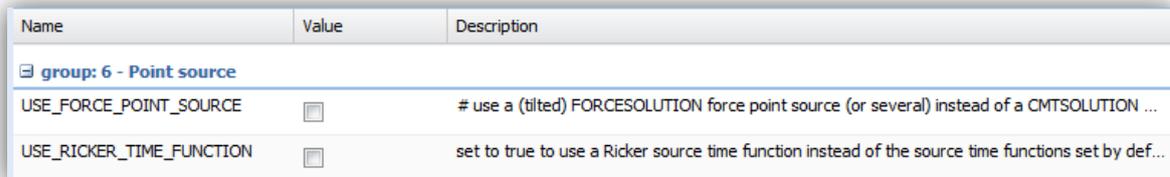
USE_HIGHRES_FOR_MOVIES saves all of the grid points so that the resolution of the movie is the same as the resolution of the model. Selecting this option requires a large amount of memory, so should not be selected by default.

NTSTEPS_BETWEEN_FRAMES sets the number of time steps between snapshots of the wavefield. The spacing of the frames in seconds is given by (NTSTEPS_BETWEEN_FRAMES*DT).

HDUR_MOVIE is the half duration of the source time function for the movie simulation.

4.2.7 Group 6 – Point source

The VERCE platform is very much configured to simulate earthquake sources. However there are other types of seismic source such as active sources, explosions or impacts that you may want to simulate. This can be done using the options described below.



Name	Value	Description
group: 6 - Point source		
USE_FORCE_POINT_SOURCE	<input type="checkbox"/>	# use a (tilted) FORCESOLUTION force point source (or several) instead of a CMTSOLUTION ...
USE_RICKER_TIME_FUNCTION	<input type="checkbox"/>	set to true to use a Ricker source time function instead of the source time functions set by def...

Figure 4.8: Parameter form for ‘Group 6 – Point Source’.

USE_FORCE_POINT_SOURCE simulates a force source (e.g. impact source) rather than an earthquake source. If you are using this option the source must be defined in a FORCESOLUTION file, rather than in the CMT solution convention used for earthquakes sources. See the SPECFEM manual for full details.

USER_RICKER_TIME_FUNCTION this inputs the source as a Ricker wavelet, rather than the default delta function that is designed to represent the slip on a fault during an earthquake. Again this option is useful for simulating non earthquake seismic sources.

4.2.7 Group 7 – Advanced

There are a large amount of other functions within SPECFEM that can be altered using the VERCE platform. A brief description of the functions that can be changed in the platform is given below, but in most cases if you intend to use these advanced option you should also refer to the SPECFEM manual.

Name	Value	Description
group: 7 - Advanced		
NGNOD	8	number of nodes for 2D and 3D shape functions for hexahedral, we use either 8-node mesh el...
APPROXIMATE_OCEAN_LOAD	<input type="checkbox"/>	see SPECSEM3D_CARTESIAN manual
TOPOGRAPHY	<input type="checkbox"/>	see SPECSEM3D_CARTESIAN manual
ANISOTROPY	<input type="checkbox"/>	see SPECSEM3D_CARTESIAN manual
GRAVITY	<input type="checkbox"/>	see SPECSEM3D_CARTESIAN manual
TOMOGRAPHY_PATH	./velocity/	path for external tomographic models files
STACEY_INSTEAD_OF_FREE_SU...	<input type="checkbox"/>	Stacey absorbing top surface (defined in mesh as 'free_surface_file')
CREATE_SHAKEMAP	<input type="checkbox"/>	save shakemap files
SAVE_DISPLACEMENT	<input type="checkbox"/>	save displacement instead velocity in the snapshot files
SAVE_MESH_FILES	<input type="checkbox"/>	Set this flag to true to save ParaView mesh files for subsequent viewing.
LOCAL_PATH	./OUTPUT_FIL	path to store the local database file on each node
NTSTEP_BETWEEN_OUTPUT_INFO	500	interval at which we output time step info and max of norm of displacement
NTSTEP_BETWEEN_OUTPUT_SEI...	10000	interval in time steps for writing of seismograms
LOCAL_PATH	./OUTPUT_FIL	path to store the local database file on each node
NTSTEP_BETWEEN_OUTPUT_INFO	500	interval at which we output time step info and max of norm of displacement
NTSTEP_BETWEEN_OUTPUT_SEI...	10000	interval in time steps for writing of seismograms
NTSTEP_BETWEEN_READ_ADJSRC	0	interval in time steps for reading adjoint traces, 0 = read the whole adjoint sources at the sam...
PRINT_SOURCE_TIME_FUNCTION	<input type="checkbox"/>	
COUPLE_WITH_EXTERNAL_CODE	<input type="checkbox"/>	
EXTERNAL_CODE_TYPE	1	1 = DSM, 2 = AxISEM, 3 = FK
TRACTION_PATH	./DATA/DSM_1	
ADIOS_ENABLED	<input type="checkbox"/>	
ADIOS_FOR_DATABASES	<input type="checkbox"/>	
ADIOS_FOR_MESH	<input checked="" type="checkbox"/>	
ADIOS_FOR_FORWARD_ARRAYS	<input checked="" type="checkbox"/>	
ADIOS_FOR_KERNELS	<input checked="" type="checkbox"/>	
MESH_A_CHUNK_OF_THE_EARTH	<input checked="" type="checkbox"/>	

Figure 4.9: Parameter form for 'Group 7 – Advanced'.

NGNOD controls the number of nodes for each part of the hexagonal mesh. For all meshes loaded into the VERCE platform, and all meshes created using CUBIT this should be left at the default value of 8.

APROXIMATE_OCEAN_LOAD is a relatively computationally cheap method of modelling the effect of oceans on the wavefield. It is however only effective at relatively low frequencies (20 second and below). For higher frequencies if the effects of the water column are to modelled, the ocean must be included in the model itself.

TOPOGRAPHY is only needed if the 'APPROXIMATE_OCEAN_LOAD' option above is selected, and reads in the topography/bathymetry files needed to define that surface.

ANISOTROPY is selected if you want to include seismic anisotropy. You will also need to provide an anisotropy model to include this, and this has not been done for any of the pre-loaded meshes and models, and so cannot be selected for these cases.

STACY_INSTEAD_OF_FREE_SURFACE is largely the same as the 'PML_INSTEAD_OF_FREE_SURFACE' option, apart from the free surface is replaced with the less effective Clayton-Enquist style absorbing boundary conditions.

CREATE_SHAKEMAP creates a map of peak ground velocity for the area modelled.

SAVE_DISPLACEMENT saves displacement in the movie snapshots, rather than the default which is to save velocity for the movie.

SAVE_MESH_FILES saves mesh files in a 'paraview' format for later use.

NSTEPS_BETWEEN_OUTPUT_INFO controls the frequency that information about a model that is running is output to the log file.

NSTEPS_BETWEEN_OUTPUT_SEISMOS controls the frequency (in number of time steps) that the seismograms are written to disk. Fewer disk writes will allow the model to run quicker, but will also increase the amount of data that is lost if the code does crash.

NSTEPS_BETWEEN_OUTPUT_ADJSRC this is as above, but controls the number of time steps between information on the adjoint sources being written.

PRINT_SOURCE_TIME_FUNCTION outputs the source time function input to the model as a text file.

COUPLE_WITH_EXTERNAL_CODE allows information to be passed to other solvers (e.g. axiSEM), and simulations to be run with this information. This function is not supported within the VERCE platform.

EXTERNAL_CODE_TYPE selects which of the external codes the model should interact with.

4.3 Statistics tab

This tab allows you to assess how the proportion of jobs that are successfully run through the portal (figure 4.10) as well the reliability and successful rate of various ,HPC resources (figure 4.11) and workflows. Developers can also see the relative success rate of a given user.



Figure 4.10: The success rate statistics for the whole portal, displayed through the statistics tab.

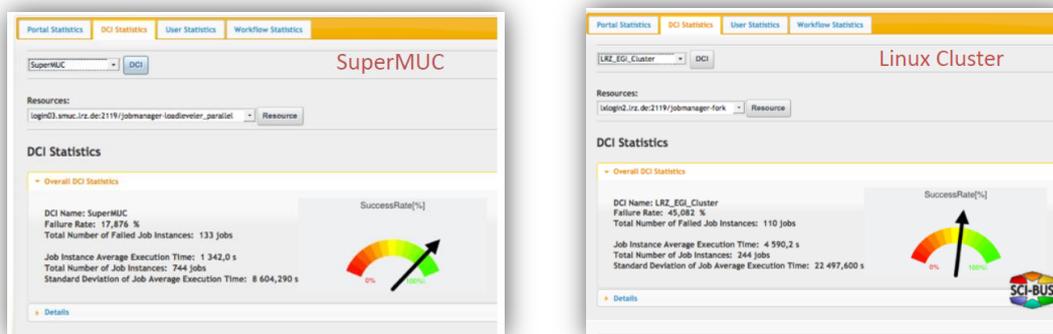


Figure 4.11: The comparative success rate statistics for the two HPC resources SuperMUC and Linux cluster, displayed through the statistics tab.

4.4 The Meshes and models already uploaded

Currently there are several meshes and velocity models loaded for Italy, and a mesh loaded for the Maule area of Central Chile. The frequency to which the seismic wavefield can be simulated is partly controlled by the time step of the model, but is also controlled by the spacing of grid points within the model mesh. For this reason, there is a maximum frequency (or minimum period) of waveform that a given mesh and velocity model can support. This minimum period (maximum frequency) resolvable is shown below (figure 4.12) for each of the meshes currently available in the VERCE portal.

Other details of these meshes such as the suggested time step (suggest DT) to make each model stable, the number of points in the mesh (Num. of HEX), and the approximate time that a 1 minute simulation would take if it was run on 100 cores (CPU time) are shown in figure 4.13. The UTM zone for each the meshes is also shown as this can be useful when using the output data.

Finally figure 4.14 gives details of the velocity models that are uploaded, along with the meshes, to the VERCE portal. The minimum and maximum P-wave and S-wave velocities are given as these are required to calculate the grid spacing and time step needed to resolve a given frequency of seismic wave.

Model \ Mesh	ITALY (5.8 MHex)	ITALY 2 (0.8 MHex)	C. ITALY (0.2 MHex)	N. ITALY (1.6 MHex)	S. ITALY (1.8 MHex)	CHILE (1.7 MHex)
FULL ITALY (Di Stefano 2014)	3s			4s		
FULL ITALY 2 (Di Stefano 2014)		4s				
CENTRAL ITALY (Chiababba 2010)			3s			
NORTH ITALY (Di Stefano 2011)				4s		
SOUTH ITALY (Di Stefano 2011)					3s	
CHILE (UJUV)						3s

Figure 4.12: The mesh and velocity model combinations currently available through the VERCE portal, and the period to which the wave-field can be resolved in each of these mesh model combinations.

	ITALY	ITALY 2	C. ITALY	N. ITALY	S. ITALY	CHILE
Num of Hex (Million)	5.8	0.8	0.2	1.6	1.8	1.7
UTM zone	33	33	33	32	33	-18
Suggested DT	1E-02	1E-02	5E-03	4E-03	5E-03	1E-03
CPU time (100 cores - 60 s)	~3 h	~1 h	25 min	~2 h	~1.5 h	~20 h

Figure 4.13: Details of the meshes currently available through the VERCE platform.

	FULL ITALY (for ITALY mesh)	FULL ITALY (for N. ITALY mesh)	FULL ITALY 2 (for ITALY 2 mesh)	CENTRAL ITALY (for C. ITALY mesh)	NORTH ITALY (for N. ITALY mesh)	SOUTH ITALY (for S. ITALY mesh)	CHILE (for CHILE mesh)
UTM zone	33	32	33	33	32	33	-18
Min Resolved Period (s)	3	4	4	3	4	3	3
vp range (m/s)	768 - 8806	804 - 8774	3600 - 8807	1782 - 8000	1039 - 8609	1732 - 8245	4360 - 6703
vs range (m/s)	428 - 4725	434 - 4493	2000 - 4715	1000 - 4444	600 - 4971	1000 - 4761	2450 - 3818

Figure 4.14: Details of the velocity models that are currently available through the VERCE platform.

5. Running a simple simulation: a step by step guide

5.1 Preparing the portal

Before you are able to run a simulation you must log in to the VERCE portal, and upload a proxy certificate as described in section 3.4 of this guide. Once you have done this you should be able to use all parts of the portal for the next 24 hours (or the lifetime of your proxy certificate).

5.2 Selecting a solver, mesh and velocity model

The waveform simulations are run from the Forward Modelling tab of the VERCE portal, shown in figure 4.1. First you must select the solver tab from the top of the forward modelling panel. In the first drop down menu you must select the solver. This is the code that will perform the full waveform simulation. Currently the VERCE platform support SPECSEM3D Cartesian, which is designed to simulate waveforms on the regional scale. Specify this model by selecting SPECSEM3D_CARTESIAN in the drop down menu labelled 'Solvers'. Once you have done this you will see that the right hand side of the panel is populated with the input parameters for SPECSEM, which are categorised into eight groups.

You can now select the area that you wish to run a simulation for from the drop down menu labelled 'Meshes'. Once you have selected the relevant mesh, the map on the left of the forward modelling panel will zoom to the area concerned, and the area the mesh covers will be outlined with a black box as shown below, in figure 5.1.

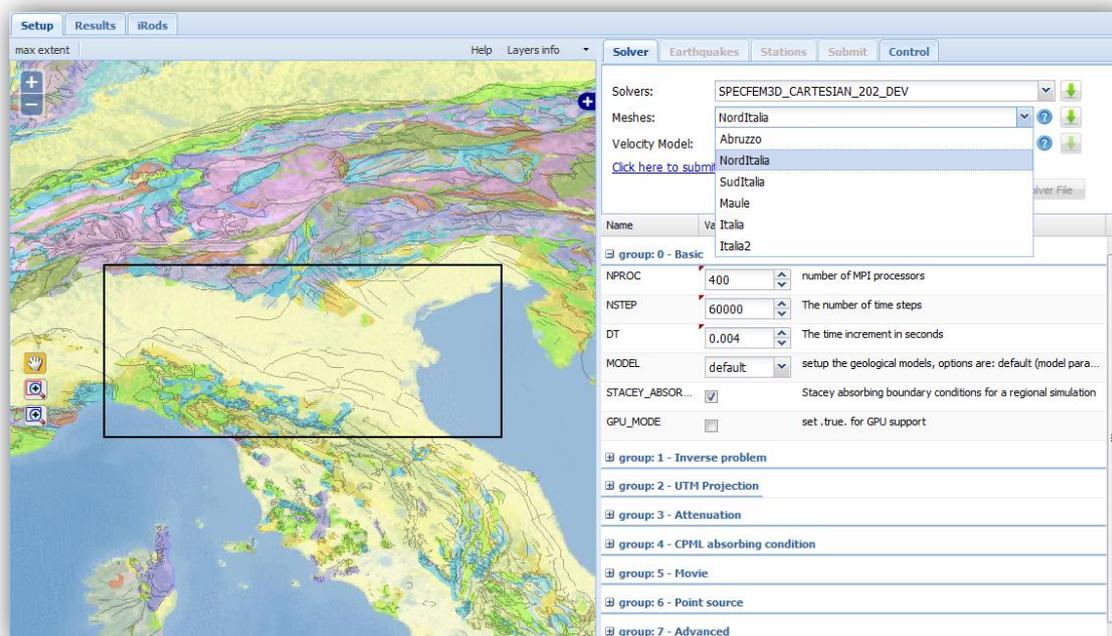


Figure 5.1: Selecting a solver and mesh for Northern Italy. On the left panel the colours show the local geology, and known faults are plotted in black. On the right panel the drop down menu showing the meshes that are currently loaded in are is shown. The input parameters for the solver can also be seen at the bottom of the right panel.

Once you have selected a solver and a mesh you can then select a velocity model for the given area. Most of the meshes currently only have one velocity model associated with them. But in theory it is possible to have more than one velocity model for each mesh, as long as the range of velocities within each velocity model is not different. This could for instance allow different tomographic models to be compared.

5.3 Specifying the input parameters for SPECfEM

The input parameters for SPECfEM are broken up into 8 categories, which are briefly described below. For a basic simulation, many of the parameters can be left at the default setting within the portal, but the meaning of these inputs is described below. It is however important that you check that the parameters in group 0, especially insuring that the number of processors (NPROC), the time step (DT) used in the simulation, and length of the simulation are all set correctly. For this example you also need to set the attenuation parameters as specified in section 5.3.4. It is worth checking the other parameters, but they should be set correctly by default. The final set up of the workflow is shown in figure 5.2.

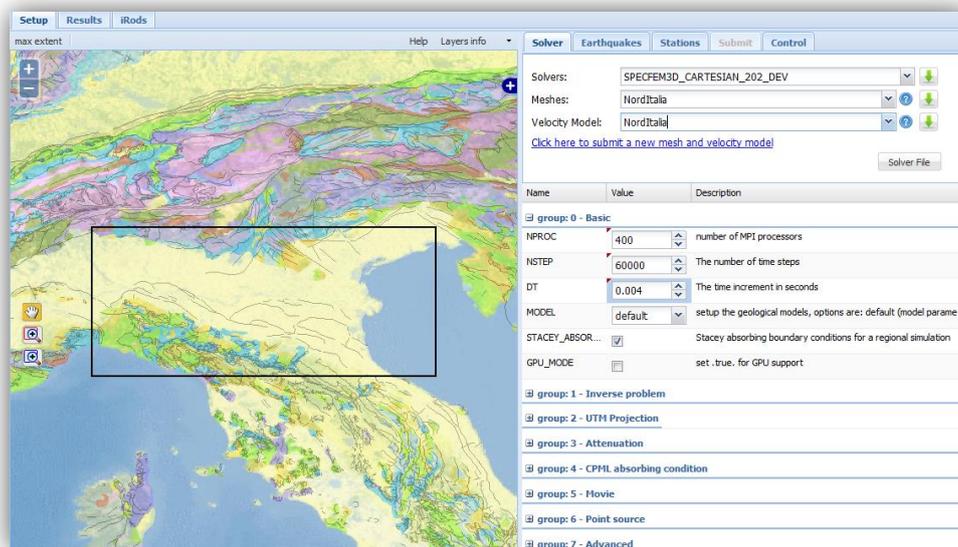


Figure 5.2: Final setup of the workflow for the simple example. The area of Northern Italy which is to be simulated is outlined by the black box.

5.4. Selecting earthquakes

Next the earthquakes to be simulated are defined in the ‘earthquakes’ tab at the top right of the forward modelling page. There are a number of earthquake catalogues pre-loaded into the VERCE platform, including the global CMT catalogue, which provides a starting data set for any area that we have a model and mesh for. It is also possible to load in your own bespoke catalogue of earthquake moment tensors, as described in greater detail in section 7.4. Given that the model and mesh selected here is in Northern Italy we can however use the INGV focal mechanism data set, which is likely to have a larger range of events down to smaller magnitudes.

The earthquake catalogue you wish to search is selected using the drop down menu at the top of the earthquakes tab. You can then search for earthquakes in a range of magnitudes, depths and time. In

this example we have searched for all earthquakes from magnitude Mw 4.0 – 9.0, at up to 1000 m depth that occurred within the model and mesh area in the year 2012. The earthquake of interest can then be selected either by ticking the box next to the earthquake in the list in the bottom right of the panel, or by selecting the location of the earthquake on the map in the left of the panel. It is also possible to select multiple events from this page. This will then submit the same number of jobs as events that you have selected, and produce waveforms for each of them. This would then allow multiple events to be used in an inversion for instance.

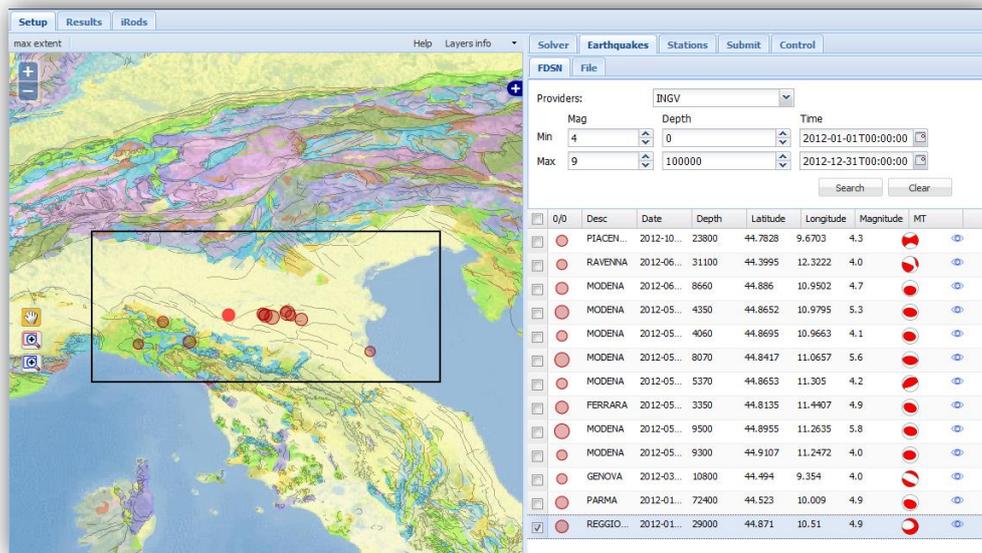


Figure 5.3: The earthquake selection page of the forward modelling tool. Events are shown from the INGV catalogue. The location of the events are shown in the summary map on the left, and details of the events are shown in the bottom right hand part of the panel. The event or events to be modelled can be selected from either of these panels.

For now though, just submit one event. In figure 5.3 an earthquake in the centre of the mesh has been selected so that we see a nice clear waveform on all stations. You can select any event you are interested in, but be aware that events close to the limits of the mesh may be more greatly affected by the absorbing boundary conditions at the edges of the model.

5.5. Selecting stations

The seismometers where you want to simulate the recorded seismogram can then be selected under the 'Stations' tab on the right hand side of the forward modelling panel. The portal is configured to output the synthetic waveforms at points where real seismometers exist so that the synthetic waveforms can be directly compared to the observed waveforms recorded at these stations.

There are many seismic networks loaded into the VERCE portal that can be used. To see all of the stations that are available within the mesh and model area you can simply select 'Any Network' in the drop down box at the top of the 'Stations' panel. Alternatively you can select a given network you are interested in (or have the data for), for instance the INGV network (network code IV). You

can then manually select the stations you are interested in by clicking on the stations in the map view, or selecting the tick box next to the station information in the right hand panel. All of the stations in this list can be selected by selecting the tick box at the top of the list.

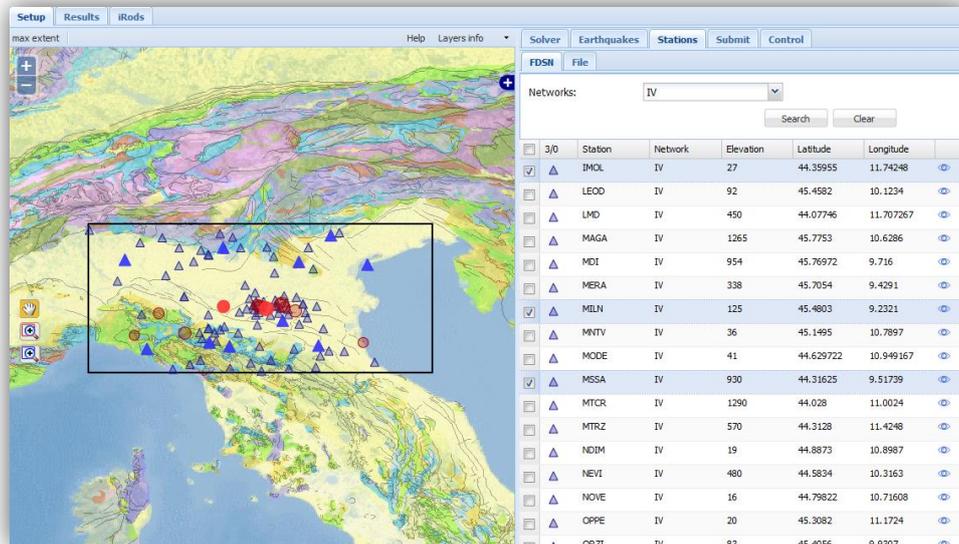


Figure 5.4: The station selection panel. Stations are shown inside the area of the model and mesh (shown by the black box) for the INGV network. The stations (shown in blue) can be selected individually from the map or from the station list. All the available stations can be selected by ticking the tick box in the title bar (labelled 0/0 when no stations are selected).

While selecting a large number of stations will not greatly affect the overall time taken or computational cost of the simulation, the more stations you select the longer it will take to move the model data to memory where you can then access it. For large simulations it is most efficient therefore to output seismograms for all the stations that you may be interested in.

5.6. Submitting the job

You can then select the work flow you wish to use and submit your job. Each workflow is configured to a different machine, and some machines have more than one workflow available on them. In the drop down menu shown below, you can see workflows for the three main supercomputers available through the portal.

- CoolMUC / Linux cluster : LRZ, <https://www.lrz.de/services/compute/linux-cluster/>
- SMU (SuperMUC):LRZ, <https://www.lrz.de/services/compute/supermuc/systemdescription/>
- SCAI – Cineca, <http://www.hpc.cineca.it/>

You can then select the relevant workflow (here we have selected SMU_test_512_30m_SPECFEM) for the HPC resource you wish to submit to, and enter a name and description of the model. Please note that the name of the model must be 30 characters or less, and can only consist of letters, numbers and decimals. Other characters are not accepted.

If you have selected more than one event to simulate you may select the 'Process the events in parallel' box. This submits each of the events to the queue of the resources you have selected as separate jobs, rather than running the jobs in serial (on after another). This can speed up the process of simulating several earthquake events for the same velocity model.

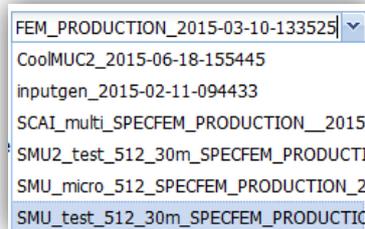


Figure 5.5: The list of workflows available from the top drop down menu of the 'submit' tab of the forward modelling page.

5.7. Monitoring the job

Once the job has been submitted, the status and progress of the job can be monitored from the 'Control' tab. This brings up a list of all of the jobs that have recently been submitted as shown below.

Name	Desc	Workflow	Grid	Status	Date	
LUCCA_25th_Jan_201301442833077367	-	n/a	n/a	INIT	21 - 09 - 2015	
LUCCA_25th_Jan_201301442832806990	Strike slip EQ in the LUCCA region	SMU_micro_512_SPECFEM_PRODUCTI...	SuperMUC	ERROR	21 - 09 - 2015	
12.02_LUCCA01442833343719	strike slip EQ in the LUCCA region	SCAI_multi_SPECFEM_PRODUCTION_...	SCAI_Cluster2	ERROR	21 - 09 - 2015	
_LUCCA_25th_Jan_201301442833233...	strike slip EQ in the LUCCA region	SCAI_multi_SPECFEM_PRODUCTION_...	SCAI_Cluster2	ERROR	21 - 09 - 2015	
certificate_test01442502075634	-	n/a	n/a	INIT	17 - 09 - 2015	
Maule23092012_test01417435387337	-	n/a	n/a	INIT	01 - 12 - 2014	

Figure 5.6: Jobs listed in the 'control' tab.

Clicking the following symbols at the end of the row allows information on the model to be accessed among other things. The symbols are described below,

- Gives access to the log files from a given job.
- Takes you to the results tab to view the results from this model run.
- Reloads the setup of the model so that it can be reused or modified and resubmitted.
- Deletes the record of this model run.

6. Accessing the results

Once your experiment has run, or even during the computation, you will be able to explore the output and execution information of a simulation, and validate its results by accessing the 'Results' tab. This can be found at the top of the Forward Modelling Panel. This interface provides full interactive access to the information about the processes and the data produced by your experiment, also known as Provenance. You can directly open the results of a run from the Control tab, by clicking on the blue 'eye' icon associated to the run of interest. The 'Run Activity Monitor' will show you the list of terminating processes. Runs, whose status is still on RUNNING, may show more processes as the computation proceeds. Use the Refresh button at run-time to update the 'Run Activity Monitor' view.

Selecting a process from the 'Run Activity Monitor' will populate the 'Data Products' view with the list of all the data generated by the selected process. For each item, a variety of information is available, such as data properties and values (metadata), parameters and errors associated to the generating process. It is also possible to download or visualise the files containing the data here, if these have been produced. To help identifying the availability of data files, the 'Run Activity Monitor' shows names of the processes in *Italic* and **Bold Italic** text. They mean respectively, that the specific activity produced intermediate data products that will be made available at the end of the computation, and that these are instead already available for download also while the submitted RUNNING status of the submitted experiment. Simple text is used for processes that did not produce any concrete files, but nonetheless, we provide the metadata associated to its output, since they can be of great value to help understanding the final outcome of the experiment.

In the 'Data Products' view each element gives you the possibility to 'Navigate the Data Derivation Graph' which means accessing the data via an interactive graph shown in the 'Data Derivation View'. This can be done backwards, viewing all the data and processes that contributed to the generation of the specific data. Alternatively the graph can show the forward progression of a process that has used the data as an input. The yellow circle means that the navigation can proceed further to the next objects in the derivation, while blue circles represent fully expanded nodes. Right click on a node of the graph to display the information about the data associated to that node. This is analogous to the 'Data Products' view. An example of a combined visualisation of data derivation and visualisation is shown in Figure 6.1.

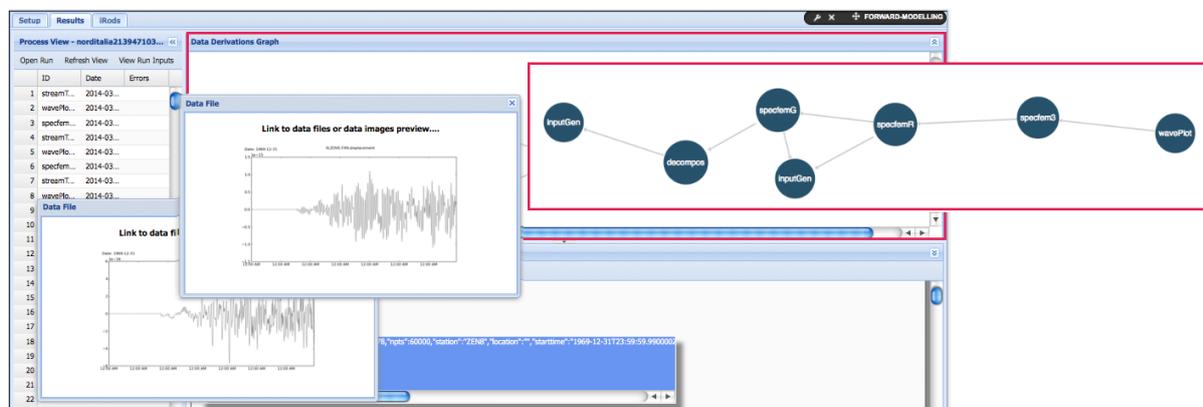


Figure 6.1: Example of data derivation graph and output visualisation

On the menu bar of the 'Data Products View', you can choose to 'Search' for data and to 'Filter' the element presented in the current list of data products. A combination of terms, min and max values and mime-type can be used to 'Search' across all the data of the current run. This is shown in the *tooltip* that appears when hovering the input fields. You can download all the files associated to the visualised list of data products by selecting the 'Produce Download Script' Function. This will generate a list of GridFTP commands that you can use from a computational cluster where you want to download the data to perform further analysis and processing.

Alternatively, simulations and experiments can be searched and opened using the 'Open Run' button, which can be found on the top left of the 'Results' tab. As for the 'Search' function explained before, this enables you to search for whole runs. For instance you can search for events within a range of magnitudes, depths, latitudes, longitudes, or simulated for a given stations or network, or using a given meshes and models etc. For a list of suggested parameters for which it is possible to search for, please see the drop down menu. The terms in the drop down contain only a subset of all the terms that could be used. You can use also new ones, as long as they exist among the metadata describing your results. The list of runs that are shown allow also for the re-editing of the 'Description'.

From the same window you can also open other users' runs. You will need to know the ID of the run and the username. In this way you can share results metadata with others. To allow access also to the data please use the iDROP application, which is available by clicking the 'iDROP' button on the top left menu bar. It will provide access to a tool that will allow advanced administration options associated to your own data-space, including access control settings, as shown in figure 6.2

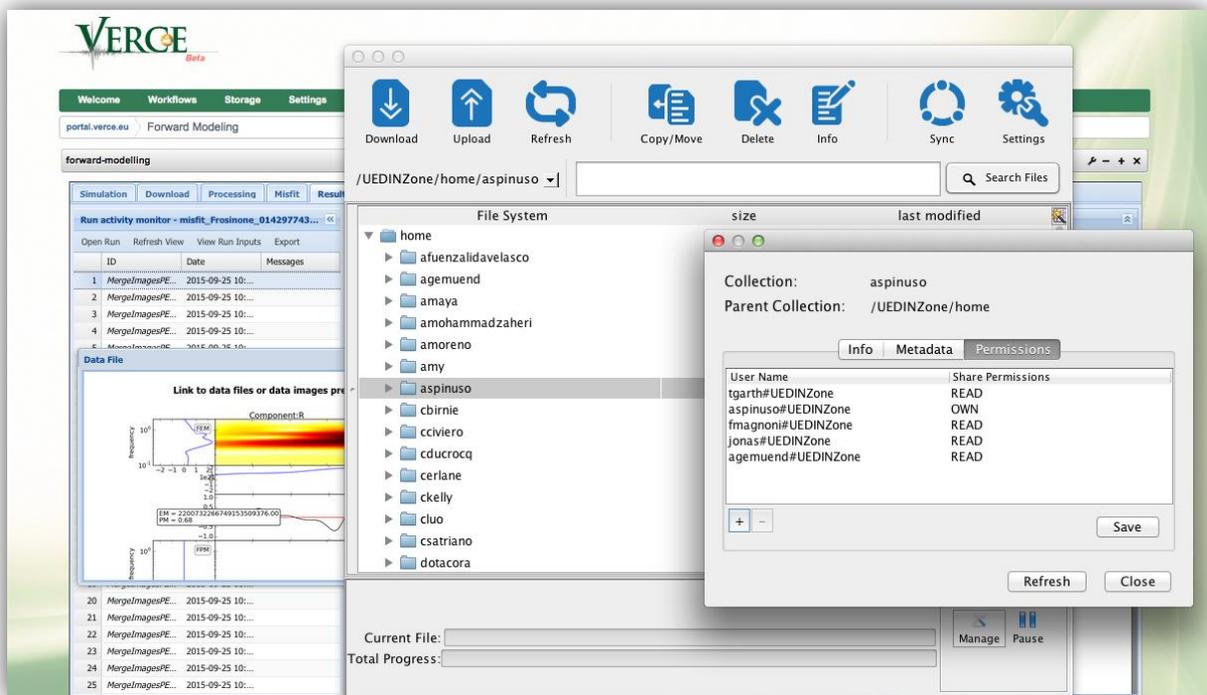


Figure 6.2: Data Sharing: Users can share and visualise metadata of their runs with others and data access permissions can be easily managed with the generic iDROP tool

You can also visualise the input files for the current run or directly open the workflows that contributed to its results, by using the 'View Run Inputs' function. Here you can access to files in various formats, like *QuakeML*, which contains information about the source or sources that are input into the model.

6.1 Waveform outputs

The primary output of any waveform model is the recorded waveforms. These can be viewed most simply as a .jpg file as shown in top panel of figure 6.3. The three components of different seismometers that are output can also be viewed in a more interactive form, by downloading the KMZ file that is automatically output from the model run and viewing it in Google Earth as shown in bottom panel of figure 6.3.

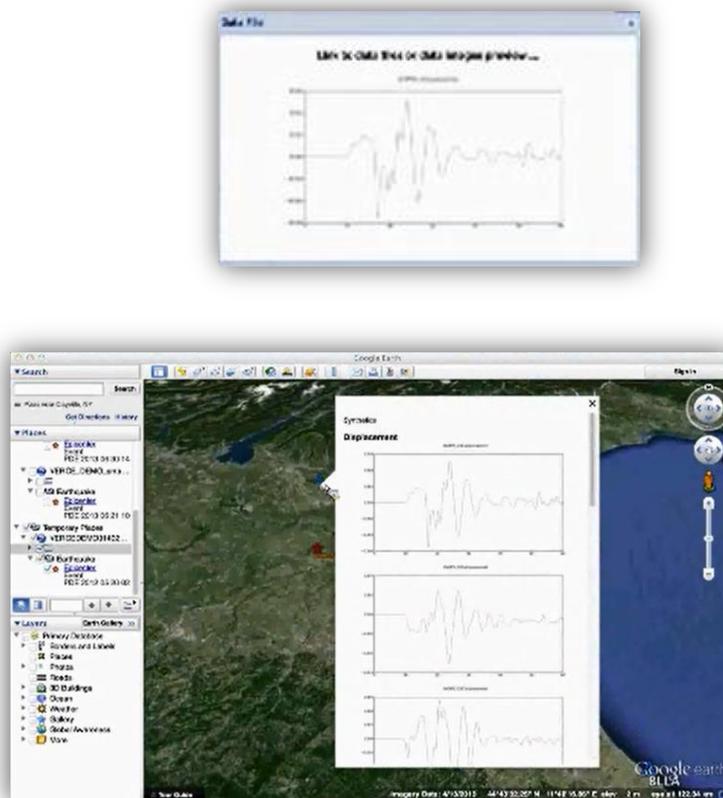


Figure 6.3: *Top:* Example waveform output. *Bottom:* Three components of a synthetic seismogram produced for an earthquake in Northern Italy, observed on the interactive google earth tool.

6.2 Animation outputs

The VERCE platform can also be used to produce animations or movies of the waveforms propagating out from the simulated earthquake event. These animations can be projected onto the Earth's surface as shown in the snapshot example on figure 6.3, or can be projected onto a vertical surface that cuts through the model area as described in section 4.2.6.

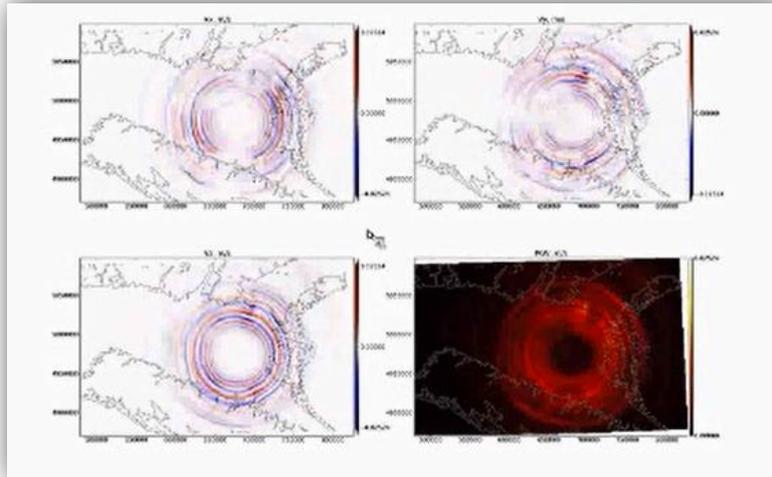


Figure 6.3: Snapshot movie of an Earthquake in Central Italy produced from the VERCE platform.
Other outputs

6.3 Other output

The portal output can also be processed externally to produce products such as ground motion maps as shown for an event in Northern Italy in figure 6.4.

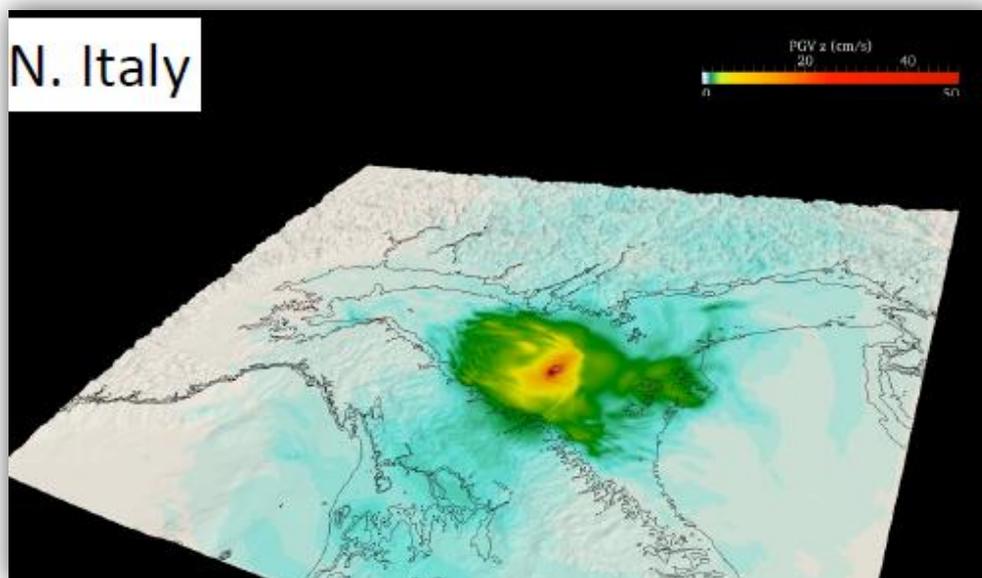


Figure 6.4: A ground shaking map produced with SPECfEM for an event occurring in Northern Italy.

7. Running simulation using your own data

So far we have covered how to run a relatively simple model using data that is already loaded into the portal. While we hope to keep increasing the areas where meshes and models are currently available, it is also essential that users can upload their own velocity models, meshes, earthquake catalogues and station locations. In this section we will give a brief overview of what is needed for each of these inputs, and how they are submitted to the portal.

7.1 Creating your own velocity model.

All of the velocity models currently loaded into the VERCE portal are based on regional travel time tomography models of the area concerned. It would be possible to construct an input velocity model based tomography on a surface wave tomography or noise tomography, or even from a velocity model produced from active seismic techniques form refraction of reflection seismic surveys. If you interested in running a full waveform simulation in a new area, you will therefore acquire a published velocity model, or use a preliminary model from your own work or collaborations.

The velocity model you submit should be based on your tomography, and should deviate back to the regional or global 1D starting model at the edges of the 3D volume. The tomography file should be formatted as shown in the below (Figure 7.1), and saved as a text file. The grid spacing defined on the second row of the file depends upon the frequency of seismic wave you intend to simulate, as described in section 7.2. Once this file has been created the velocity model can be uploaded along with the corresponding mesh as described in section 7.3.

The variables input into the velocity model text file above are defined below.

ORIG_X	ORIG_Y	ORIG_Z	END_X	END_Y	END_Z						
255066.763887	4551257.50051	-32000.0	455066.763887	4751257.50051	2717.0						
1000.0	1000.0	1000.0									
201	201	34									
1782.	7930.	1000.	4361.11	2000.	2785.46						
						SPACING_X	SPACING_Y	SPACING_Z			
						NX	NY	NZ			
						VP_MIN	VP_MAX	VS_MIN	VS_MAX	RHO_MIN	RHO_MAX
255066.763887	4551257.50051	-32000.0	7800.0	4333.33333333	2070.90976436						
256066.763887	4551257.50051	-32000.0	7800.0	4333.33333333	2070.90976436						
257066.763887	4551257.50051	-32000.0	7800.0	4333.33333333	2070.90976436						
258066.763887	4551257.50051	-32000.0	7800.0	4333.33333333	2070.90976436						
259066.763887	4551257.50051	-32000.0	7800.0	4333.33333333	2070.90976436						
260066.763887	4551257.50051	-32000.0	7800.0	4333.33333333	2070.90976436						
261066.763887	4551257.50051	-32000.0	7800.0	4333.33333333	2070.90976436						
262066.763887	4551257.50051	-32000.0	7800.0	4333.33333333	2070.90976436						
263066.763887	4551257.50051	-32000.0	7800.0	4333.33333333	2070.90976436						
264066.763887	4551257.50051	-32000.0	7800.0	4333.33333333	2070.90976436						
265066.763887	4551257.50051	-32000.0	7800.0	4333.33333333	2070.90976436						
⋮	⋮	⋮	⋮	⋮	⋮						
x	y	z	vp m/s	vs m/s	rho kg/m ³						

Figure 7.1: Format for an input velocity model. Image re-produced from the SPECfEM manual.

ORIG_X, ORIG_Y, & ORIG_Z: are the coordinates of the initial grid points in the tomographic model in the x, y and z directions respectively.

END_X, END_Y, & END_Z: are the coordinates of the final grid points in the tomographic model in the x, y and z directions respectively.

SPACING_X, SPACING_Y, & SPACING_Z: describe the spacing between points of the tomography file in the x, y and z directions respectively.

NX, NY, & NZ: describe the number of grid points in the x, y and z directions respectively.

VP_MIN & VP_MAX: describe respectively the minimum and maximum p-wave velocity of the input file in ms^{-1} .

VS_MIN & VS_MAX: describe respectively the minimum and maximum s-wave velocity of the input file in ms^{-1} .

RHO_MIN & RHO_MAX: describe respectively the minimum and maximum density of the input file in kg/m^3 .

7.2 Creating a bespoke mesh for your area

Creating a mesh is the most complicated step in setting up a model in a new area. The mesh must be created so that it can account for the frequency of seismic wave at the seismic velocities that are found in the velocity model defined as above.

7.2.1 Meshing parameters

The spacing of the grid (Δh) depends upon the minimum seismic velocity in the velocity model (v_{min}) and the frequency (f) to which you wish to resolve the wavefield in your simulations as shown in the equation below.

$$\Delta h = v_{min} T_o \frac{n + 1}{f(n)}$$

Where T_o is the shortest period that can be resolved, and n is related to the number of points that per wavelength of frequency f .

$$\Delta t < C_{max} \frac{v_{min}}{v_{max}} T_o \frac{n + 1}{f(n)}$$

Where C_{max} is the highest possible value of the Courant number, given by the equation below.

$$C_{max} = \Delta t (v/\Delta h)_{max}$$

Where $(v/\Delta h)_{max}$ denotes the maximum ratio between the p-wave velocity and the grid spacing.

The resolution of the mesh therefore can reduce with depth, as the average seismic velocity decreases. The mesh must also account for the topography of the Earth's surface, or bathymetry if the modelled area includes oceanic areas.

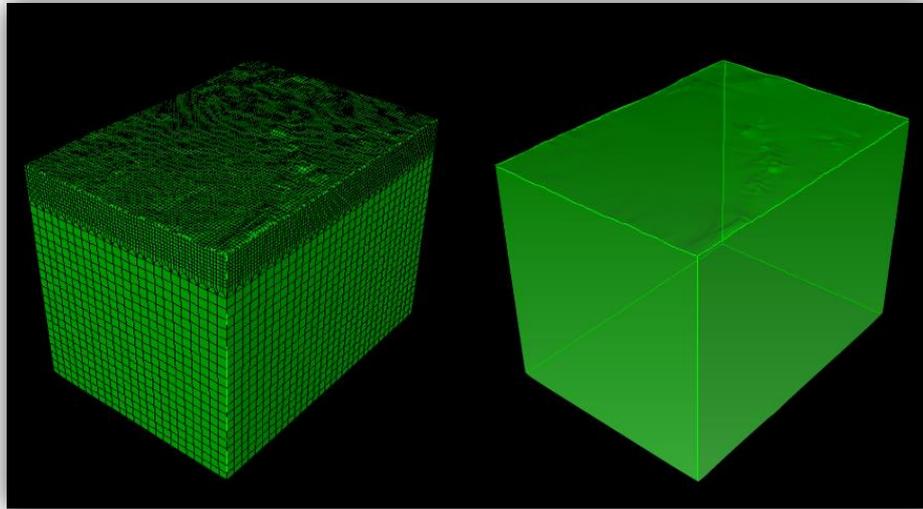


Figure 7.2: An example hexagonal for the produced in GeoCubit. On the left of the figure we can see that the resolution of the mesh decreases with depth, therefore the maximum resolved frequency also increases. The top surface of the mesh represents the topography of the area to be modelled.

7.2.2 Meshing software

Meshes that can be used with SPECSEM and within the portal can be produced using CUBIT/ TRELIS software. Unfortunately the CUBIT/TRELIS software is not free, although a 30 day trial licence can be downloaded. Full details of the commercial software can be found at the following link.

<http://www.csimsoft.com/>

This software is then used in conjunction with the free python based GeoCubit software developed at the INGV.

https://github.com/geodynamics/specfem3d/tree/devel/CUBIT_GEOCUBIT

A full description of how to create a mesh for SPECSEM or the VERCE platform can be found here.

<http://verce.eu/Training/UseVERCE/2015-7-VERCE-hexmeshing101.pdf>

7.3 Submitting a mesh and a velocity model

Once you have produced your mesh you can take the ten mesh files listed in figure 7.3 and put them into a single zip file named 'mesh_MyMeshArea', where 'MyMeshArea' is the name of the area you are studying. Do the same with the tomography file formatted as outlined in section 7.1. the tomography file should be formatted as shown in figure 7.3, with the UTM zone of the tomography specified in the file name.

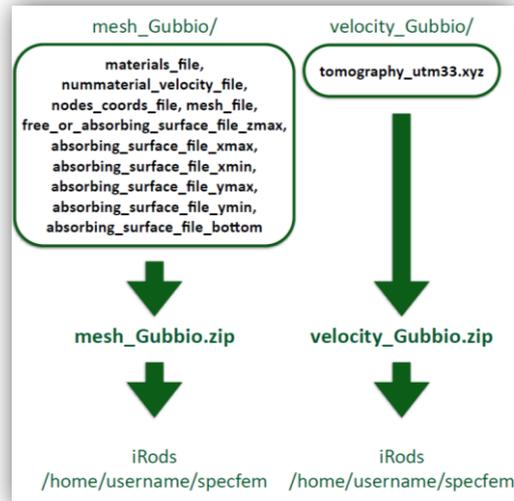


Figure 7.3: Creating the zip files needed to upload a new model to the VERCE platform. The left hand side shows the 10 mesh files that need to be included and uploaded. The right hand side shows the single velocity file that needs to be uploaded, along with the naming conventions for these files.

Once these zip files have been created, they can be uploaded by clicking the link labelled '**Click here to submit a new mesh and velocity model**' in the solver tab of the forward modelling page. This will bring up the parameter form shown in figure 7.4.

The zipped mesh and velocity model files can then be uploaded to the portal from the local machine. The limits of the model area in latitude and longitude can then be input in the 'mesh bounds' section of the file. Finally click 'submit' for the mesh and velocity model to be uploaded to the portal. The meshes and models are manually validated before they are made available to you, so it can take several days for the mesh and model to be uploaded and ready to use.

It is of course possible to upload meshes and models that only you can use, or to upload meshes and models to be used by the community. If you wish to keep the mesh and model available for only you or a group of users please specify this in the note box at the bottom of the parameter form.

Submit a new mesh and velocity model

You can submit a new mesh and velocity model here. Currently we will check your submission by hand before adding them to the list of available meshes and models.

Mesh

Upload a file...: Browse...

...or paste a link:

Mesh Bounds

Minimum latitude: Maximum latitude:

Minimum longitude: Maximum longitude:

Velocity Model

Upload a file...: Browse...

...or paste a link:

note:

Submit Cancel

Figure 7.4: The parameter form for inputting a new mesh and model into the VERCE platform.

7.4 Submitting a new earthquake catalogue

If you wish to submit your own earthquake catalogue you can do this on the file tab of the earthquakes page shown in figure 7.5. The catalogue must be uploaded in quakeML format. The easiest way to convert other earthquake catalogue formats to quakeML is using ObsPy. The earthquakes must have full details of the location and focal mechanism of all the events you are interested in modelling.

The ObsPy command `'readEvents'` can be used to read in events that are in a range of text based formats (e.g. NDK & ZMAP). The event data can then be written to a quakeML file using `'writeQuakeML'`. Full details on how to install ObsPy, and access tutorials on ObsPy are given in appendix one.

Solver Earthquakes Stations Submit Control

FDSN File

Open

Name: DefaultName

File: Select a file Browse...

Upload

<input type="checkbox"/>	0/0	Desc	Date	Depth	Latitude	Longitude	Magnitude	MT

Figure 7.5: The input form for bespoke user input earthquake catalogues.

7.5 Submitting a new station catalogue

New seismic stations and networks can be input in a similar way using the upload form shown in figure 7.7. The format of the input data is simple a list of the station name, group (i.e. network name) longitude, latitude, depth and meters buried. Figure 7.6 shows an example of the station locations for the temporary Maule network in Central Chile, which is already loaded into the platform.

NAME	GROUP	LON	LAN	DEPTH	BURIED
U01B	MAU	-72.490400	-37.287700	0	0
U02B	MAU	-72.980600	-37.205900	0	0
U03B	MAU	-72.330200	-37.701200	0	0
U04B	MAU	-72.569800	-37.986700	0	0

Figure 7.6: The input format for stations and networks to be added to the VERCE platform.

The screenshot shows a web-based interface for submitting seismic station data. The main navigation tabs are 'Solver', 'Earthquakes', 'Stations', 'Submit', and 'Control'. The 'Stations' tab is active, and within it, the 'File' sub-tab is selected. Below the sub-tab, there is an 'Open' dropdown menu. The form includes a 'File name' input field containing 'DefaultName', a 'File format' dropdown menu, and a 'File' input field with the text 'Select a file' and a 'Browse...' button. An 'Upload' button is positioned below the 'File' field. At the bottom of the interface, a table header is visible with columns for 'Station', 'Network', 'Elevation', 'Latitude', and 'Longitude', and a '0/0' indicator on the left.

Figure 7.7: The input form for new seismic station networks.

VERCE glossary

Workflow – refers to a sequence of jobs that can be submitted. In the VERCE project we have a number of workflows prepared that can run your job on a specific high performance computer.

HPC – high performance computing. This usually refers to parallel computers, where a given computational task is spread over many separate processing cores.

iRODS – a suite of data management software that is embedded within the VERCE platform, and allows you to easily access your data regardless of where you submitted your simulation.

superMUC – a super computer hosted by the LRZ in Munich, Germany.

SCAI – a super computer hosted by CINECA in Italy.

DT – the time step of the waveform model.

Solver – shorthand for the code that does the forward calculation. The solver currently hosted in the VERCE platform is SPECFEM.

Mesh – the grid over which the wavefield is calculated. The modelled space is broken up into a grid of points, each with specific seismic properties (e.g. p-wave velocity, s-wave velocity, seismic attenuation). The spacing of this grid is able to change especially with depth (due to the increasing seismic velocity). The structure of these grid points is referred to as the mesh.

Velocity model – this is the seismic velocity model that is input for an area, and includes the p-wave and s-wave velocities. Most models in the VERCE portal are 3D, though there are some 2D models (for subduction zones) and 1D models (global 1D velocity models) available.

CPML – ‘convolutional perfectly matched layers’ are a type of absorbing boundary condition

Absorbing Boundary Conditions – are at the edges of the model that shouldn’t reflect the seismic energy. These boundary conditions are designed to absorb an incoming wave simulating an infinite medium.

CUBIT – an external program used for making meshes for a variety of scientific and engineering modelling disciplines.

TRELIS – the commercial name for the CUBIT package.

GeoCubit – a python based program that uses CUBIT command to create meshes for geographical bodies. Particularly, GeoCubit can produce meshes that include topography and/or bathymetry.

Appendix one – using ObsPy

ObsPy is a python based seismology toolbox, which can be used to deal with seismic waveform data and earthquake catalogue information. The toolbox can read a large range of seismic data formats and perform a large range of processing applications such as filtering the data, and misfit calculation.

Many of the pre- and post-processing applications within the portal use python and ObsPy. The toolbox can also be used to format input data, for instance producing an earthquake catalogue in quakeML format.



Figure A1.1: The ObsPy logo

A1.1 Installing Python and ObsPy

1. **Installing Python:** Depending on your operating system this can be installed through the software manager. Alternatively the latest version of python can be installed here, <https://www.python.org/downloads>.

You will also require to specific python libraries that can be installed from GitHub as described below.

2. **Installing anaconda:** if you do not have super user privalages on your machine then both ObsPy and dispel4py (see appendix two) can be installed using Anaconda, a package to manage and deploy Python packages. Anaconda can be installed on mac and Linux operating systems as described here,

<http://docs.continuum.io/anaconda/install.html>

Installing ObsPy: Make sure you have a C and fortran compile installer. You can then install ObsPy with the command,

```
conda install -c obspy obspy
```

or, if you have not installed anaconda,

```
sudo pip install obspy
```

Full instructions can be found at

<https://github.com/obspy/obsipy/wiki#installation>

A1.2 Other dependencies you need to run ObsPy are listed below;

numPy – a toolbox for doing numerical applications in python.

```
sudo apt-get install python-numpy
```

SciPy – a generic scientific programming toolbox for applications in python.

```
sudo apt-get install python-scipy
```

Additionally, in order to do the plotting parts of the tutorial below you will need to install the following;

Matplotlib – a toolbox for creating plots and figures for python applications. Figures are customised very easily and intuitively, and can be exported in a number of different. Files can also be exported easily to a .mat file for Matlab users.

```
sudo apt-get install python-matplotlib
```

A1.3 Using ObsPy

A full online python tutorial, that covers everything from a basic introduction to ObsPy, up to more advanced applications such as developing an automated processing workflow, can be found at the link below;

<http://docs.obspy.org/tutorial/>

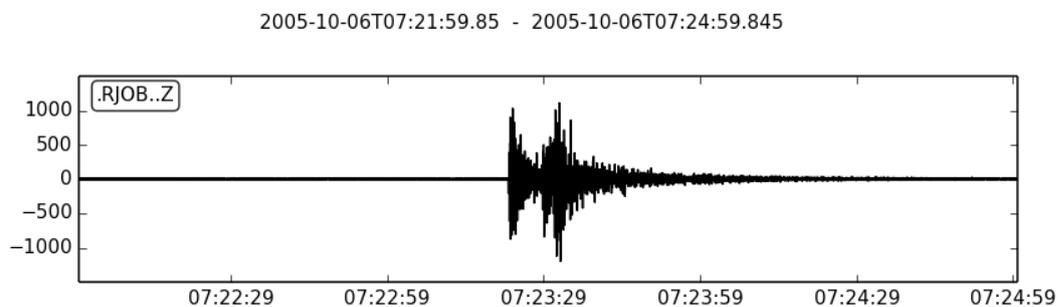


Figure A1.2: Z-component data plotted using ObsPy. Image re-produced from http://docs.obspy.org/tutorial/code_snippets/reading_seismograms.html

Appendix two – using dispel4py

Dispel4Py is a python library that allows workflows to be written that can easily scale to different sizes of computational resource, ranging from your own laptop, to a large parallel supercomputer. This means that you can devolve and test your workflow locally on your own machine before transferring the workflow to a much larger computer to process a large amount of data.

In seismology this could be particularly useful for performance calculations with very large data sets, such as noise correlation. Dispel4Py is used widely in the VERCE portal in the pre- and post-processing workflows that are implemented there. This toolbox is especially useful for any seismologist who is looking to speed up a data application.



Figure A2.1: The Dispel4Py logo.

A2.1 Installing Dispel4Py

Firstly please ensure that you have an up to date version of Python installed as described in appendix one. You can then install Dispel4Py and its dependencies (such as MPI) as described below.

1. **Installing dispel4py:** Dispel4py can be installed using the command,

```
pip install dispel4py
```

Full instructions for installing dispel4py can be found at,

<https://github.com/dispel4py/dispel4py>

2. **Installing MPI and mpi4py (optional):** If you wish to explore the parallel mapping of dispel4py to MPI you may want to install these on your machine. Different implementations of MPI are available, for example Open MPI or MPICH2. Depending on your operating system MPI can be installed through the software manager. You can then install mpi4py with the command:

```
pip install mpi4py
```

Full instructions can be found at:

<http://mpi4py.scipy.org/docs/usrman/install.html>

A2.2 Using Disple4Py

A full tutorial describing how to create workflows and perform sequential and parallel processing using Disple4Py can be found at the link below.

<http://www2.epcc.ed.ac.uk/~amrey/VERCE/Disple4Py/>

For more information on the basic principles of Disple4Py please see the following presentation.

<http://www.verce.eu/Training/UseVERCE/2015-7-VERCE-disple4py-basic.pdf>

An advanced introduction for those who wish use Disple4Py to create workflows for their own applications can be found at the link below.

<http://www.verce.eu/Training/UseVERCE/2015-7-VERCE-disple4py-advanced.pdf>

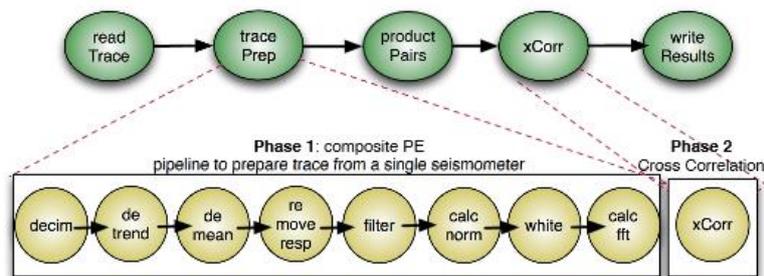


Figure A2.2: An example of a workflow to perform a cross correlation of seismic noise data

Appendix three – other VERCE supported resources

As well as the VERCE portal there are many programs that have been developed to support computationally intensive and data intensive tasks in seismology. Many of these programs have in fact been used in the VERCE portal to manage and process the data produced. More advanced users may want to use these resources directly for more bespoke applications than are supported by the VERCE platform. A summary of the online resources available is given below.

A3.1 ObsPy

ObsPy is a python library for managing seismic data. See appendix one for more details

<https://github.com/obspy/obspy/wiki>

A3.2 Pyflex



Pyflex is a python port for the program FLEXWIN (Maggi et al 2009), a tool for comparing synthetic and recorded seismograms. Full documentation can be found at,

<http://krischer.github.io/pyflex/>

A3.3 LASIF



LASIF (Large-scale Seismic Inversion Framework) is a tool for managing the large data sets associated with seismic waveform inversion.

<https://github.com/krischer/LASIF>

A3.4 SEIS-PROV



SEIS-PROV is a proposed format of documenting and describing the provenance of seismic data, in order to increase the reproducibility of seismological results.

<http://seismicdata.github.io/SEIS-PROV/>

A3.5 dispel4py

disple4py is a tool for managing large data sets, and scaling large tasks to the computational resources available. Full details are given in appendix two, and further documentation can be found at the link below.

<http://www2.epcc.ed.ac.uk/~amrey/VERCE/Dispel4Py/>

A3.5 Whisper

Whisper is a set of tools for processing seismic ambient noise, and full details can be found below.

Whisper website: <https://whisper.obs.ujf-grenoble.fr/>

Whisper documentation: <http://code-whisper.isterre.fr/>