Site Effects Assessment Using Ambient Excitations

SESAME

European Commission – Research General Directorate
Project No. EVG1-CT-2000-00026 SESAME

Final report

WP07

Array measurements:
Inversion of velocity profile

Deliverable D21.07

May 2004

List of Contents

List of Contributors (in alphabetical order after the last name)

Marc Wathelet   GEOMAC, Liège, Belgium
Denis Jongmans   LIRIGM, Grenoble, France

Project coordinator: Pierre-Yves Bard   LGIT, Grenoble, France
Task B Leader: Pierre-Yves Bard   LGIT, Grenoble, France
WP07 Leader: Denis Jongmans   LIRIGM, Grenoble, France

Summary

In the following we present the inversion program developed in the framework of the SESAME project (Site Effects Assessment Using Ambient Excitations, EC-RGD, Project No. EVG1-CT-2000-00026 SESAME), Task B (Array measurements), Work Package 07 (WP07 – Inversion of velocity profile). It has been implemented in platform independent software (based of Qt libraries): Linux OS (free libraries), Windows (proprietary libraries), Mac OS (free libraries), Solaris (not tested).

Chapter 1: Introduction

The aim of the work package is to develop a flexible software which allows to retrieve the compressional (Vp) and the shear (Vs) velocities as function of depth from dispersion curves. They can be either obtained from array data processing (WP06) or from active source experiments.

The principles of the implemented method were presented in a preceding report (deliverable 14-07, June 2003). This short report mainly focus on the developed tool themselves and how to use them.

The available binary executable files are for the main ones:

`os_forward:` computation of theoretical dispersion curves and the Rayleigh ellipticity from a layered ground model.

`os_na:` inversion of the velocity profile from a given dispersion curve and according to the parameterization provided by the user through the `param.in` file (see `os_na` chapter).

`na_viewer:` graphical user interface that ease the preparation of input files for `os_na` and that allows to browse the inversion results. Only one type of parameterization can be loaded at a time (see `report_viewer` for a more general visualization).

There are also some secondary programs developed for special tasks:

`report_viewer:` as for `na_viewer`, this tool is designed for browsing the results, but dedicated to the merging of inversion results of the same object (dispersion curve) having various parameterizations.

`spac_forward:` computation of theoretical autocorrelation curves from a layered model.

`spac_na:` same as `os_na` but for autocorelation curves.

`os_forward_grid:` this tool is especially developed for debug purposes: as the dispersion curve is not a stable computation in all cases, it is sometimes interesting to investigate the dispersion curve computation. It samples the function of frequency and slowness whose roots are the dispersion curves at all available modes.

<u>Note:</u> the syntax of the options of the preceding commands is described in the next chapters using the following conventions:

[op1 op2 ...] means optional parameters, default values will be used for these parameters.

op1 | op2 means that either option op1 or op2 must be specified, but not both (operators can be chained).

An option always starts with "-", followed by a letter (capital or not), and may end with a value. For instance,

`-F city1.model -s 15.23 -g`

means that option "F" is set to value "city1.model", option "s" to value 15.23, and option "g" is set but has no value.

Chapter 2: os_forward

This command needs arguments and it is usually launched from a console. It is an interface to the library functions used by os_na to compute the dispersion curves during inversions. It allows the user to calculated the theoretical dispersions curves (velocity versus frequency) and, for the Rayleigh case, the ellipticity curves (horizontal over vertical ratio of eigen functions) as well.

2.1. Calling syntax

os_forward [-L nModes -R nModes -p|-f|-l -k -d -t -N -s min -e max -n nsamp -g -q -o outputReport -O -r rmsReport -h] -F modelfile | -P parameterFile [-b functionName@scriptFile.qs]

2.2. Arguments

Description

-R   Specify the number of Rayleigh modes to calculate (default=1).

-L   Specify the number of Love modes to calculate (default=0).

-p   If set, the samples of the curves will be regular in period (default=not set).

-f   If set, the samples of the curves will be regular in frequency (default=not set).

-l   If set, the samples of the curves will be regular in log(frequency). If none of -p, -f or -l is set -l is the default. If there are various occurrences of one of these options, only the last one will have an effect.

-k   When performing Monte-Carlo analysis, if set, it keeps the Poisson's ratios constant (default=not set).

-d   When performing Monte-Carlo analysis, if set, it lets underlying depths of layers changing when modifying the thickness of more superficial layers (default=not set).

-t   If set, it allows retrograding velocity profile (default=not set). It is especially useful for Monte-Carlo analysis (default=not set).

-N   When performing Monte-Carlo analysis, defines the number of random models to generate (default=5000).

-s   Defines the minimum frequency or period (depends upon the type of sampling) for the frequency range (default=0.2 Hz).

-e   Defines the maximum frequency or period (depends upon the type of sampling) for the frequency range (default=20 Hz).

-n   Defines the number of samples over the frequency range(default=100).

-g   Plot reference curve using the group velocity (default=not set).

-q   Do not show a window with the calculated curves, only provide the results through the standard output (default=not set). It is usually set in conjunction with option -b.

Description

-o   Specify the output file name where the results will be stored (default=modelfile.report). The output format is a binary file called "inversion report" that can be read by all other programs of this package.

-O   When option -o is not specified, and when the output file already exists, if set, it overwrite the existing file (default=not set).

-r   Specify the report file used as a reference to calculate the misfits (default=use the curves calculated for the first generated model to calculate the misfits).

-h   Print a short help message through the standard output (default=not set). It overrides all other options and the program stops immediately after showing the help message.

-F   Specify the file containing the models (several models are accepted). If any of the model parameter has a standard deviation greater than 0, a Monte-Carlo analysis is performed. The format for the model file is (usually with extension ".model"):

```
Line 1:
  n
Line 2:
  h(0) dh(0 vp(0) dvp(0) vs(0) dvs(0) rho(0) drho(0)
Line 3:
  h(1) dh(1) vp(1) dvp(1) vs(1) dvs(1) rho(1) drho(1)
...
Line n:
  h(n-2) dh(n-2) vp(n-2) dvp(n-2) vs(n-2) dvs(n-2) rho(n-2) drho(n-2)
Line n+1:
  0 0 vp(n-1) dvp(n-1) vs(n-1) dvs(n-1) rho(n-1) drho(n-1)
```

Where, n stands for the number of layers, h for the thickness (m), dh for the standard deviation of the thickness, vp for the P-wave velocity (m/s), dvp for the standard deviation of the P-wave velocity, vs for the S-wave velocity (m/s), dvs for the standard deviation of the S-wave velocity, rho for the density (T/m^3), and drho for the standard deviation of the density.

The last layer is an half-space and its thickness is ignored but must be specified. This file format is the general ".model" format admitted in this package. Many models can be concatenated in the same file, os_forward will calculate the dispersion curves for all of them.

For this option no default is accepted, either -F or -P must be specified.

Description

^{-P} It is possible to specify a parameterization file as input to generate the models. The format is exactly the same as for the parameters of the inversion, see `os_na` (section `# parameterization description`). A first section called `# options` must contain the values for parameters under the keyword `parameters`. For instance,

```
# options
is_depth n
parameters 5 30
# parameterization description
3 layers
#H_min  H_max  Vp_op  Vp_min  Vp_max  Vs_op  Vs_min  Vs_max  Vsp_min  Vsp_max
Rho_min Rho_max DV_min DV_max nSubLayers
2 20 0 500 500 0 250 250 0.5 0.5 2 2 0 0 1
25 50 0 800 800 0 400 400 0.5 0.5 2.1 2.1 1000 1000 100
30 30 0 3500 3500 0 1750 1750 0.5 0.5 2.5 2.5 0 0 1
```

Great care must be paid to the second section as no format checking is performed. It is better to generate it with `na_viewer` (see chapter 4). This option is useful to construct models with velocity gradients inside layers.

^{-b} This option is used to launch a QSA script once the dispersion curves are computed (e.g. automatic capture of computed graphs). This option is available if you install Qt together with qsa (www.trolltech.com). The sesarray package must be compiled with qsa already installed.

2.3. Graphical user interface

Once the dispersion curves are calculated, a ".report" file is available and a window appears on the screen with the Vp and Vs profiles, the dispersion curves and the ellipticity curves. The grayed dispersions curves are Love modes.

Graphs have 4 distinct parts: x-axis, y-axis, content, and the graph it-self (blank corner, usually the lower left corner). You can double-click on whatever part to change its properties. A right-click will display a popup menu different for each 4 parts. For instance, to print an individual graph, right-click in its lower-left corner and select "Print". To get a bitmap (jpeg, png, tif, ...), do the same job and select "Export image as ...". To zoom-in on a particular rectangle, right-click in the content, select "Zoom", left-click and drag the mouse to draw the zoom-in rectangle. You can travel across the axis' ranges by scrolling the vertical and horizontal scroll bars. To go back to the original view, select "Unzoom".

If you left-click on the lower left corner, and maintain the mouse button down, you can move the graphs on the sheet to re-organize the presentation of the results.

A content of a graph is a stack of layers. Each layer contains a different type of plot (grids, curves, points, circles, error bars,...). You can hide or show them in the content properties ("layers" tab). It is also possible to save visible layers to an ASCII file and to add new layers defined in an existing file (".layer" files). This way you can exchange information from one plot to another and you can easily compare distinct results.

The "layout" is a file format that can contain any properties of any object. It is an easy way to save the current properties of a graph and to apply them to another.

2.4. Output

The dispersion curves for all modes are produced through the standard output together with the ellipticity curves. There are some additional information like the integral averaged shear velocity (Vs) over the layer stack (except the half-space) and over the first 30 m, the frequency of the fundamental ellipticity peak, the dispersion and the ellipticity misfit.

Chapter 3: os_na

This command needs arguments and it is usually launched from `na_viewer` graphical user interface or from a console. This program uses the Neighbourhood Algorithm to inverse the dispersion curves provided either as simple average values or as values with their standard deviations. The ellipticity frequency peak can be also inversed together with the dispersion curve.

This program needs 2 or 3 arguments (see next section) and 2 files, `param.in` and `na.in`, which must exist in the current directory. The first one describes the parameterization of the model and some processing options. The second file specify the options for the inversion algorithm itself. We strongly advise to generate the input files through the graphical user interface (`na_viewer`, menu "Inversion/Run").

3.1. Calling syntax

os_na curve_to_fit.report output.report [--bayes]

The order of the argument is important.

3.2. Arguments

The curve to fit must be provided in a binary format ("`.report`"). `na_viewer` has tools to create those report files from an ASCII source or from results of noise array processing. The output report will contain all the randomly generated models with their dispersion curves, their misfit values, and eventually their ellipticity curves. It will be created without warning even if the file already exists.

If the optional keyword `--bayes` is added at the end of the command line, the resampling algorithm will be used. It generates new models according to the misfit function considered here as probability density function. The multi-dimensional misfit function (whose input ensemble is the parameter space used for inversion) is defined as constant over the Voronoi cells computed from the ensemble of models produced by any inversion. The ensemble of models may be any `na.nad` file renamed to `ensemble.nad`. The results are the 1D-marginals, 2D-marginals, the resolution and the the covariance matrices. This former option is implemented without any warranty. The parameters for the Neighbourhood Algorithm must be gathered in a `nad.in` file (see Sambridge's home page for details, http://rses.anu.edu.au/~malcolm).

3.3. Output

At least 5 files are created:

`output.report` as specified in the arguments

`na.nad` produced by the Neighbourhood Algorithm, it is the complete list of models.

`na.log` comments produced by the Neighbourhood Algorithm during the inversion.

`forward_sampler.log` comments produced by the dispersion computation algorithm.

`forward_bayes.log` comments produced by the dispersion computation algorithm (`--bayes`).

`na.sum` comments produced by the Neighbourhood Algorithm at the end of the inversion.

Chapter 4: na_viewer

This command does not need any arguments and can be launched from the Desktop. It is a graphical user interface to prepare the inversion and to analyze its results.

4.1. Calling syntax

na_viewer [-b functionName@scriptFile.qs]

4.2. Arguments

An optional QSA script can be launched on startup. This option is available if you install Qt together with qsa (www.trolltech.com). The sesarray package must be compiled with qsa already installed.

4.3. Starting an inversion

The first thing to define is the curve or the ensemble of curves to fit. In a second stage, you need to prepare the input files for `os_na`. It is then ready for inversion.

Select "Set dispersion to fit" from menu "Inversion". "Set autocorr to fit" is a very similar tool dedicated to the autocorrelation curves. The following dialog box will appear:
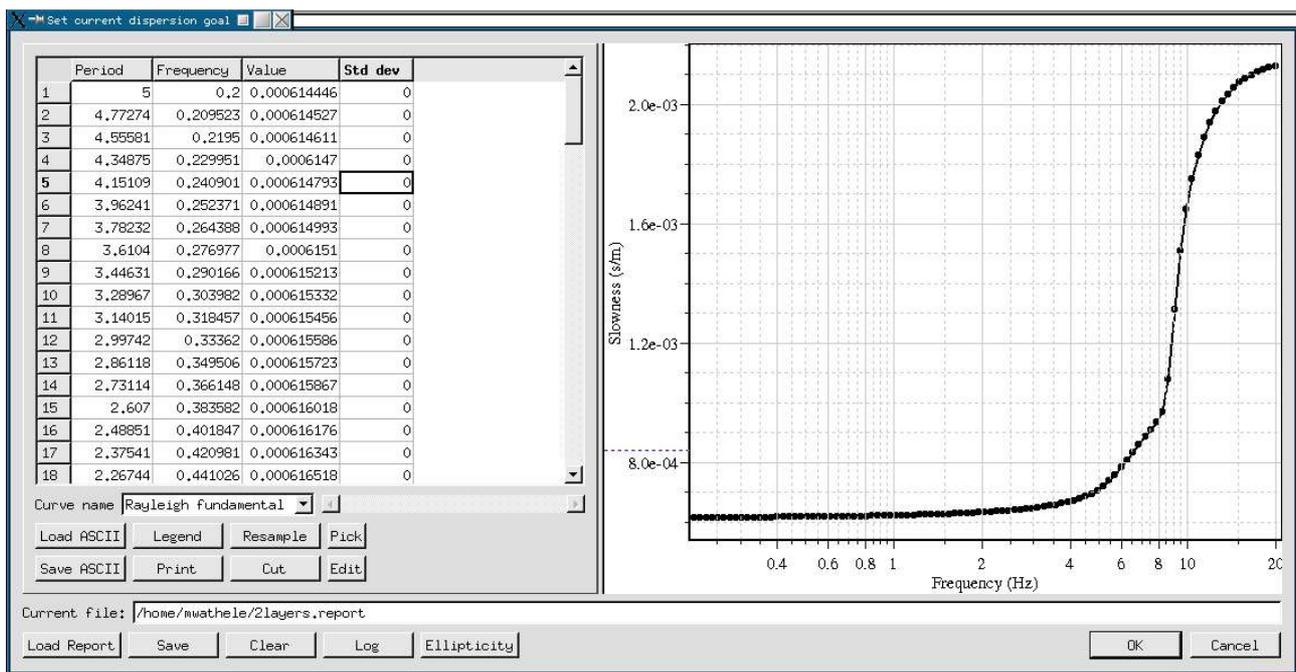
You can load ASCII files (2 or 3 columns, period or frequency, velocity or slowness, m or km) as well as existing ".report" files. In the above case, it is a ".report" generated by `os_forward` for 1 sediment layer model. The horizontal scroll bar lets you browse the currently opened curves. The graph on the right always shows all opened curves (except the curves with name "### Remove ###", see combo box's content), while the table on the left only contains the current active curve. You can edit whatever cell of the table and watch at the graphical results of the left. See chapter 2.3 for details on how to use the graph. "Resample" helps you resampling the active curve, "Cut" throws away some parts of it.

To start an inversion you must select one current ".report" file whose name appears at the bottom of the above dialog box. If you changed an existing report or if you imported an ASCII file, you must save your changes in a new ".report" file. The curve name is important when saving: only curves with standard names will be included in the saved ".report". The standard names are given in the combo box. For higher modes you can enter any number greater than 1. If more than 1 curve is detected for a particular mode at a certain frequency, an average is calculated. The final sample set is defined by the union of the sample sets of each curve. To avoid over-sampling, try to set a common sampling set for all curves. The resulting curves are added to the ensemble of opened curves when the saving operation succeed (they appear in thin red).

Click on "Ok", if you are satisfied with the current file. It is possible to add an HV curve to the current report by clicking on "Ellipticity". You are now ready for the second stage: preparing the input files and running the inversion.

Select "Run" in menu "Inversion". A wizard will help you to define the inversion options step by step, as reproduced here below:
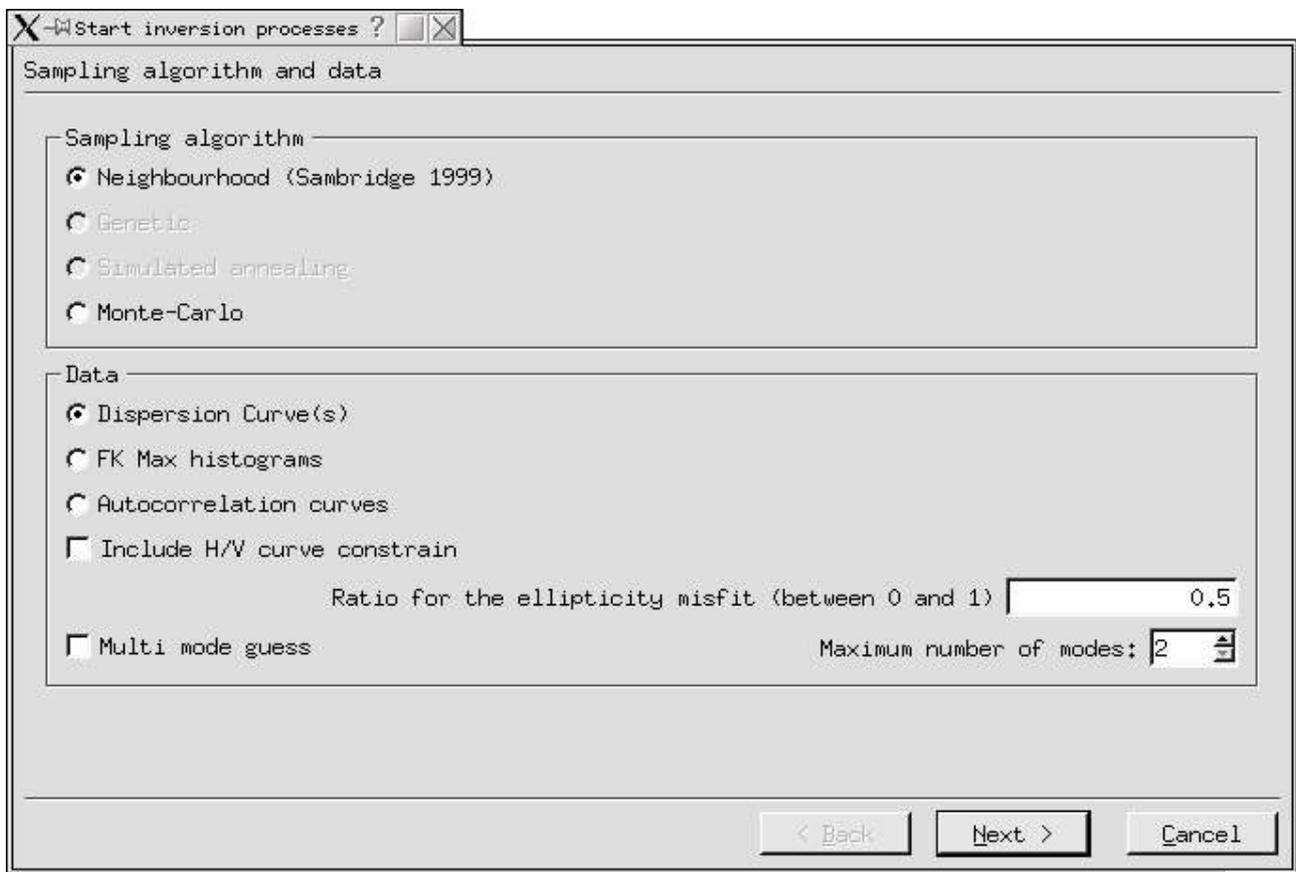
Select the type of data you want to invert (dispersion or autocorrelation, with or without H/V constrain). The program used for inversion will be different (`os_na` or `spac_na`).
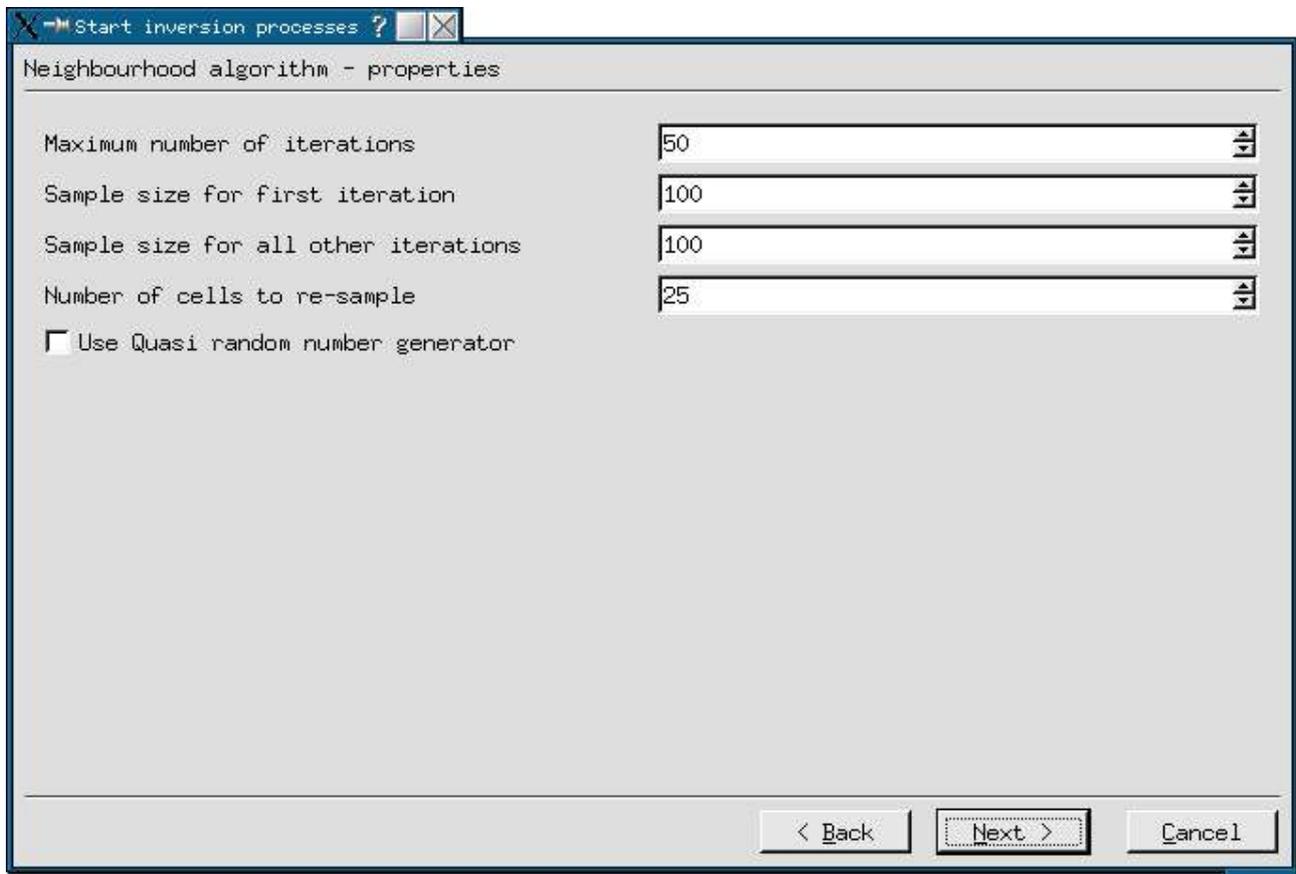
The meaning of these parameters is given in Deliverable 14-07 (June 2003).

The next wizard page is the tool that helps you creating parametrized models which can be used by `os_na` and `os_forward`. Use the "Load" and "Save" commands to restore an old parameterization (".param" or "param.in" directly) and to save the current parameterization. For the inversion, this is the definition of the a-priori probability density function. The probability is set as uniform inside the specified ranges and null outside.

The first table defines the ranges of the parameter of one particular layer. When a parameter is fixed, min and max are equal, it is usually the case for the density. The last column is the variation of Vp across the layer. In a usual case, the velocity is constant over the thickness and the column is inactive. If you choose a "velocity gradient across layer", this last column lets you set the range for the variation of Vp from the top to the bottom of the layer. This variation is equivalent to the power law exponent as explained in Deliverable_14.07 (June 2003). The power law is internally represented by a stack of sub-layers, do not forget to set its number to a reasonable value (between 5 and 10).

The second table summarizes the properties for all layers of the model. To edit them first select the layer, change the values in the first table and click on "Modify layer".

The "incremental velocity" option can be used for all layers except for the first one. The parameter Vp is changed into dVp which is the velocity gap between the bottom of the upper layer and the top of this layer.

Using option "Depth" requires that the ranges of depth for all layers are consistent. Actually, overlapping ranges may lead to strange and undetermined results (negative thicknesses).

The Neighbourhood Algorithm is a random process which must be initialized with a random seed. Due to its random nature, it is mandatory to execute several inversions over the same data with the same parameterization but with distinct seeds. The misfit threshold for storing the models in the report can be by-passed by setting it to 1e99 (all models will be stored).
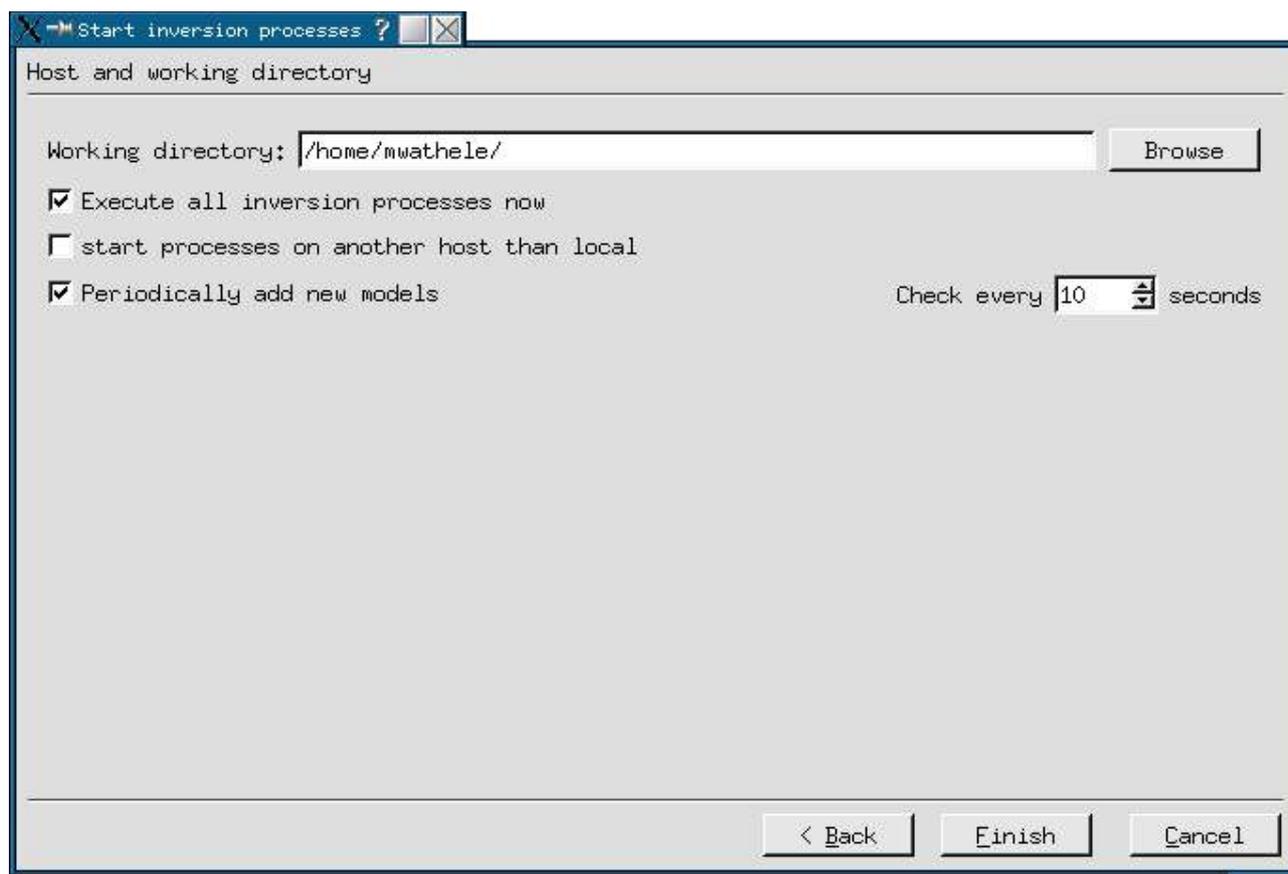
The last page lets you choose the working directory where to save all input files and results. It is also recommended to start all "processes now" to view them directly in the `na_viewer` interface. Random models are simultaneously stored in the reports corresponding to each launched inversion process. The `na_viewer` interface will periodically check if new models are available for plotting until the last process completes.

### 4.4. Viewing parameter space results

Once the first models are detected in the reports, the 1D and the 2D marginals will appear in the main sheet. As it is only projections of a multi-dimensional parameter space, it is necessary to sort the models by their misfit values to have a look at the best fitting part of the space ("Sort" from menu "View"). "Set limits" from the same menu automatically sets the limits of axis in order to show all model points.

Menus "Models" and "Filter" works on the same basis: different ways of selecting the models. The first one selects them to plot in a "profile" way (see next section), the former just hides non selected models. It is possible to calculate the Sh response for the visible models ("Compute Sh response" from menu "Filter").

When running an inversion, the corresponding reports are automatically opened. It is also possible to manually open existing reports to view their content ("Add report" from menu "File"). The "Add models" just synchronize all opened reports, useful when looking at a running inversion that has been launched manually inside a console.

4.5. Viewing profiles and dispersion results

The P-wave and S-wave profiles of models are plotted on separate graphs. The dispersion and eventually the ellipticity curves are plotted on the right. See example below:
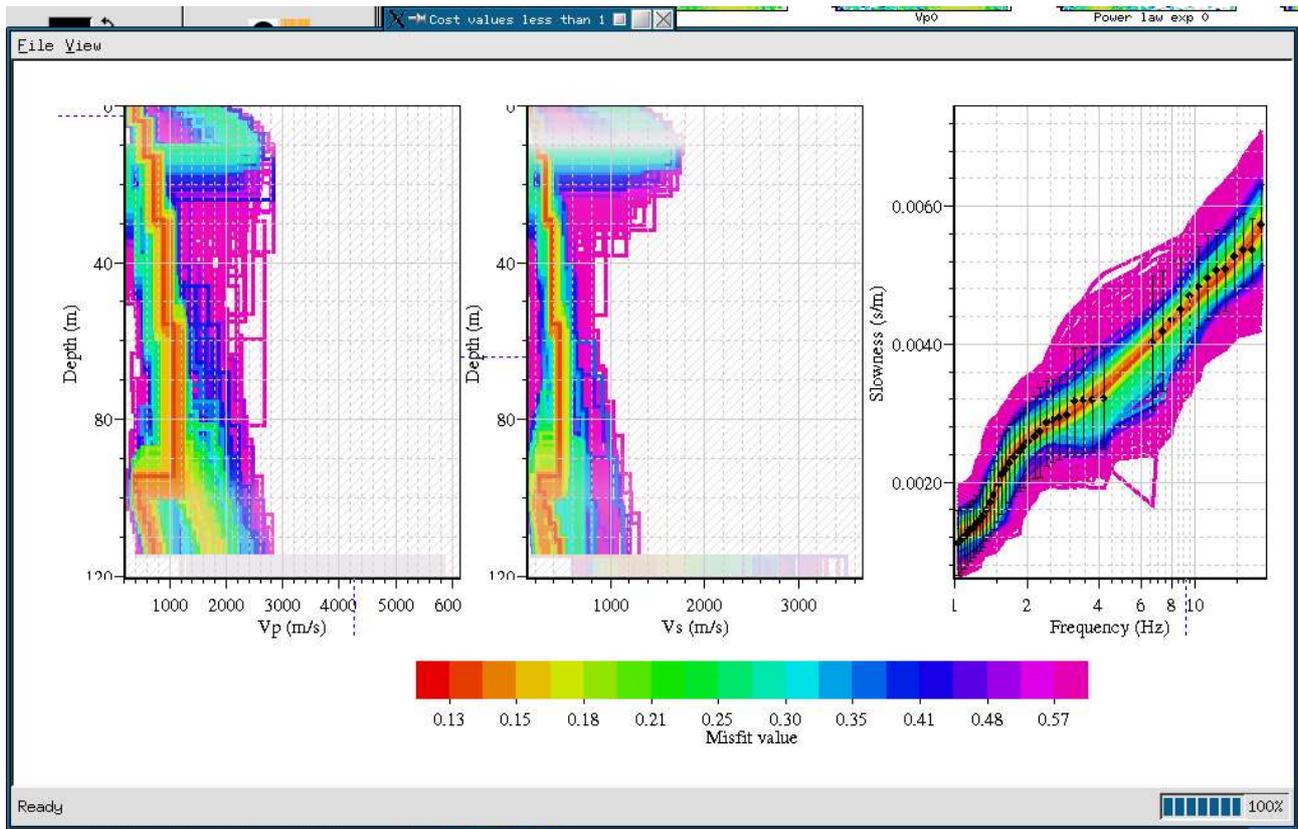
The gray shading on the left plots just indicates where the variance of the slowness is the highest. It is just a relative and not an absolute criterion to determine the valid depth range. The hashed part is quarter wave length criterion to estimate the valid depth range as well (minimum and maximum wave length deduced from the frequency range and the velocity range of the best model).

You can select a particular group of models by right-clicking in the graph's contents (menu "Edit", then draw the selection rectangle). In menu "File", select "New window" to select a particular mode (if more than one has been used for inversion) or to switch to the Likelihood plot (next section).

4.6. Viewing likelihood results

This is a more quantitative way of plotting the resulting models, it produces the average, the median, the 16% and 84% quartiles of the Vp and Vs distributions as a function of depth. The likelihood (L) is defined by the following equation:

$$L = k \exp\left(\frac{-\text{misfit}^2}{2}\right)$$

where the misfit is calculated relative to the standard deviation of the dispersion curve. The slowness and the depth axis are divided into a certain number of classes. The maximum likelihood is computed for each cell. The likelihood function is normalized versus the slowness for each depth class, and a marginal probability function is hence constructed (the colored grids in figure below). The statistical results, plotted in gray(median) and black (average), are deduced from these grids.
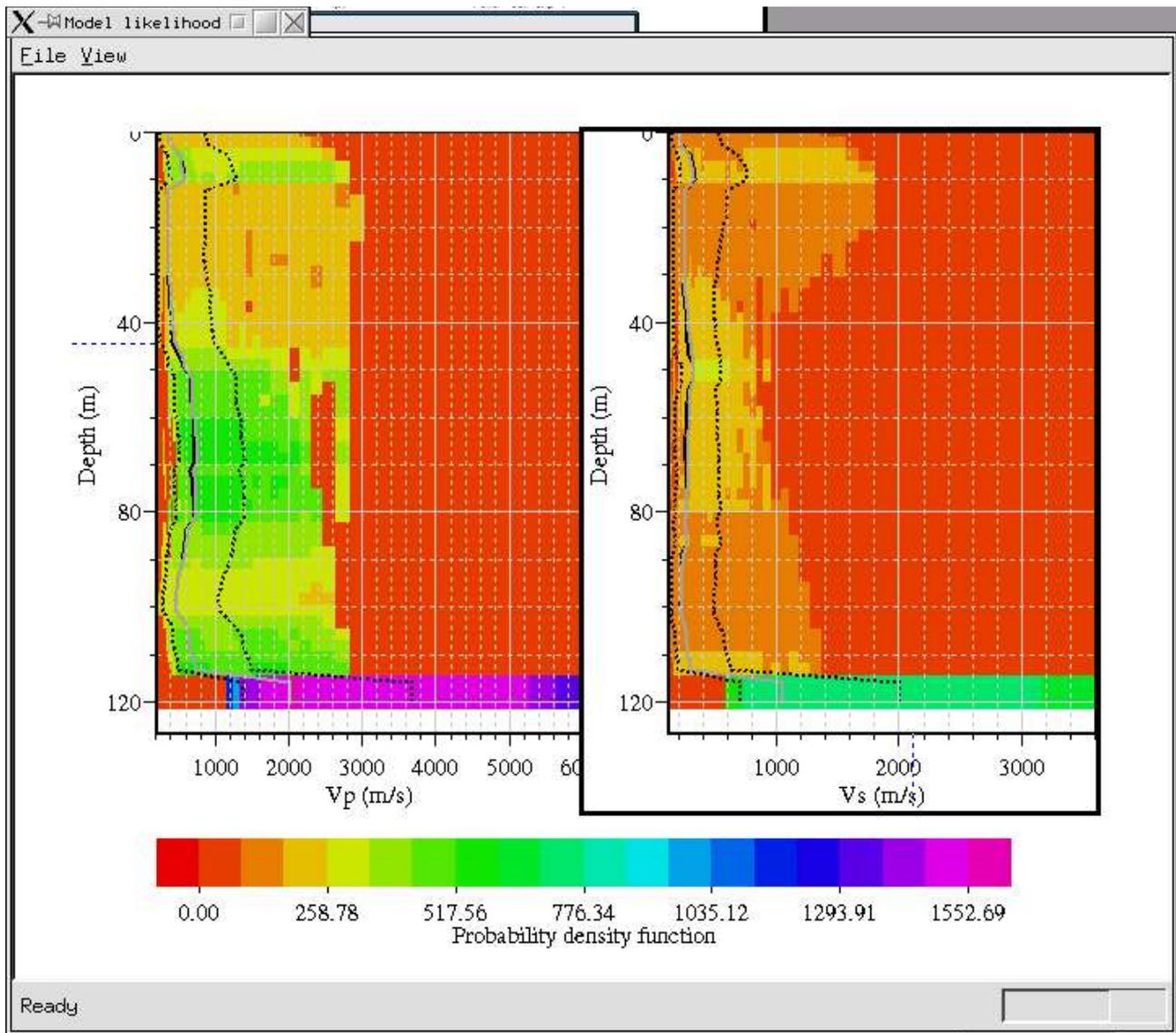
Chapter 5: report_viewer

This command needs arguments and it is usually launched from a console. It is equivalent to `na_viewer` sections 4.5 and 4.6, except that the general parameter space 1D and 2D marginals are not produced. Contrary to `na_viewer`, inversions with distinct parameterizations can be plotted together. It is interesting when trying to obtain the a-posteriori probability density function in a robust way, this is to say, independently of the parameter selection.

5.1. Calling syntax

report_viewer `[-D -m maxMisfit -i max -n ncolors -r misfitReport -l -h -b functionName@scriptFile.qs] reportName...`

5.2. Arguments

Description

−D   Switch to autocorrelation versus distance rather than frequency, valid only for autocorrelation reports produced by `spac_na` or `spac_forward` (default=not set).

−m   Select only models with a misfit less than the specified value (default=1e99).

−i   Maximum number of models selected per report (default=0, which means no limit).

−n   Number of colors for the misfit color palette (default=15).

−r   Specify the report files used as a reference. The misfits are not recalculated, but the reference curves are just pasted on the ensemble of models. Various references are allowed (default=use the reference of the first suitable report).

−l   Switch directly to Likelihood plot (section 4.6), it is quicker and sometimes mandatory for great number of models (memory space).

−h   Print a short help message through the standard output (default=not set). It overrides all other options and the program stops immediately after showing the help message.

−b   This option is used to launch a QSA script once the dispersion curves are computed (e.g. automatic capture of computed graphs). This option is available if you install Qt together with qsa (www.trolltech.com). The sesarray package must be compiled with qsa already installed.

5.3. Graphical user interface

See sections 2.3, 4.5 and 4.6 for details, exactly the same graphical object are used by this program.

Chapter 6: spac_forward

This program is organized in the same way as `os_forward`, except that it computes the autocorrelation curves once the dispersion is obtained. The calling syntax is the same and only two options differ. For a complete description, please refer to section 2.2.

6.1. Calling syntax

```
spac_forward [-L nModes -R nModes -p|-f|-l -k -d -t -N -s min -e max -n nsamp
-g -q -o outputReport -O -r rmsReport -a -h] -F modelfile | -P parameterFile
[-b functionName@scriptFile.qs]
```

6.2. Arguments

Description

−a   The list of rings with their minimum and maximum radii in a 2 column ASCII file (this option is mandatory except when option -r is present).

−r   Specify the report file used as a reference to calculate the misfits and specify the list of rings if not already set by option -a (default=use the curves calculated for first generated model to calculate the misfits).

6.3. Graphical user interface

See section 2.3 for details, exactly the same graphical object are used by this program. The plots of the autocorrelation curves are just added at the bottom of the sheet. They are working in the same way as all other plots.

Chapter 7: spac_na

This command needs arguments and it is usually launched from `na_viewer` graphical user interface or from a console. This program uses the Neighbourhood Algorithm to inverse the autocorrelation curves provided either as simple average values or as values with their standard deviations. The ellipticity frequency peak can be also inversed together with the autocorre curve.

This program needs 2 arguments (see next section) and 2 files, `param.in` and `na.in`, which must exist in the current directory. The first one describes the parameterization of the model and some processing options. The second file specify the options for the inversion algorithm itself. We strongly advise to generate the input files through the graphical user interface (`na_viewer`, menu "Inversion/Run").

7.1. Calling syntax

spac_na curve_to_fit.report output.report

The order of the argument is important.

7.2. Arguments

The curve to fit must be provided in a binary format ("`.report`"). The report file must be of type "autocorr", this means created by the tool "Set autocorr to fit" in `na_viewer`. The output report will contain all the randomly generated models with their dispersion curves, their misfit values, their autocorrelation curves and eventually their ellipticity curves. It will be created without warning if the file already exists.

7.3. Output

5 files are created:

`output.report` as specified in the arguments

`na.nad` produced by the Neighbourhood Algorithm, it is the complete list of models.

`na.log` comments produced by the Neighbourhood Algorithm during the inversion.

`forward_sampler.log` comments produced by the dispersion computation algorithm.

`na.sum` comments produced by the Neighbourhood Algorithm at the end of the inversion.

Chapter 8: os_forward_grid

This program is intended for debugging the computation of dispersion curves, which may diverge in certain circumstances. It samples the function of frequency and slowness whose roots are the dispersion curves at all available modes.

8.1. Calling syntax

os_forward_grid [-L|R -p|f|l -h -b functionName@scriptFile.qs] -F modelfile -s min -e max -n nsamp -v vmin -m vmax -o nvsamp

8.2. Arguments

Description

-R Specify that Rayleigh modes are to be calculated (default=set).

-L Specify that Love modes are to be calculated (default=not set).

-p If set, the samples of the curves will be regular in period (default=not set).

-f If set, the samples of the curves will be regular in frequency (default=not set).

-l If set, the samples of the curves will be regular in log(frequency). If none of -p, -f or -l is set -l is the default. If there are various occurrences of one of these options, only the last one will have an effect.

-s Defines the minimum frequency or period (depends upon the type of sampling) for the frequency range.

-e Defines the maximum frequency or period (depends upon the type of sampling) for the frequency range.

-n Defines the number of samples over the frequency range.

-v Defines the minimum slowness.

-m Defines the maximum slowness.

-o Defines the number of samples over the slowness range.

-h Print a short help message through the standard output (default=not set). It overrides all other options and the program stops immediately after showing the help message.

-F Same as for os_forward.

-b This option is used to launch a QSA script once the dispersion curves are computed (e.g. automatic capture of computed graphs). This option is available if you install Qt together with qsa (www.trolltech.com). The sesarray package must be compiled with qsa already installed.

8.3. Graphical user interface

It is just a plot of the frequency-slowness grid obtained. Comparison with the computed dispersion curves is very easy with the save and add layer mechanism.

Acknowledgments