The Toolbox besa_tfc_analysis

1 Requirements and Usage

The toolbox *besa_tfc_analysis* performs grand averages and performs a statistical comparison of two groups of time-frequency data (TFC data) exported from BESA Research to MATLAB using the BESA-MATLAB interface.

1.1 Software requirements

- *BESA Research 5.3* for the MATLAB export feature
- *MATLAB*, including the Image Processing Toolbox (the toolbox is required for the statistics options, but not for grand averaging)
- *Fieldtrip*, the MATLAB toolbox of the university of Nijmegen (download for free at http://fieldtrip.fcdonders.nl/).

1.2 Data requirements

The TFC data to be analyzed need to coincide in type (e.g. TSE or Coherence), frequency (range and spacing), latency (range and spacing) and channel labels.

1.3 Usage

Use this toolbox in MATLAB as:

besa_tfc_analysis(besa_tfc_all):

Here, *besa_tfc_all* is a struct containing TFC data transferred from BESA Research. *besa_tfc_analysis(besa_tfc_all)* calls the dialog window for the selection of the two groups and statistics parameters. User settings for a later run without dialog window can also be saved.

or

besa_tfc_analysis (besa_tfc_all, n, 'UserCfg_TFC'):

Does not open the dialog window. Rather, the statistics parameters are specified by the second and third parameter: *n* is an integer specifying the number of Monte Carlo draws (see below). *UserCfg_TFC* is the name of a previously saved user configuration. This syntax is useful if *besa_tfc_analysis* is to be run in automatic mode and a manual toolbox start by dialog window would halt the calculation (see below).

2 First steps

To use the toolbox, the following steps are recommended:

2.1 Preparation

1. Download the latest *Fieldtrip* version and unzip it.
2. Add the *Fieldtrip* folder to the MATLAB search path (e.g. using the MATLAB menu entry *File/Set Path*). The "Add with subfolders" option is not necessary.

3. Optionally, modify the input and output path in file *besa_tfc_analysis.m*.

2.2 Data analysis

1. **Create TFC data** in BESA Research and transfer them to MATLAB. The data is then collected in MATLAB in variable *besa_tfc_all*.
2. In MATLAB, call *besa_tfc_analysis(besa_tfc_all)*
3. The **dialog window** of the toolbox opens. If there is more than one TFC data type in the loaded data, please make a selection. This updates the datafile, condition and channel lists. Edit group names if desired.
4. Select your **data and analysis settings:**
    - Channel selection will influence statistics output but not appear in all plots.
    - Channel labels: By default, channel labels are displayed and evaluated up to the last underscore in the original name.
    - Check settings. Panels highlighted in blue are more likely to require modification. If desired, enter the name of the MATLAB variable that will contain the result of the analysis (default: stat_TFC)
    - If there is a hypothesis as to the time and frequency ranges of interest, edit the fields in the "**Range"** panel. For an activation-against-baseline hypothesis test, time ranges for both activation and baseline have to be edited. They must be non-overlapping and of equal duration.
    - Decide whether to use **channel neighbourhood** information (usually 'Yes' if the montage consists of surface sensors, 'No' if it is a source montage).
    - Hitting the "**Advanced**" pushbutton opens an additional menu. If the "Plot all channels" option is checked, plots displaying all channels of all group members will be calculated.
5. Press the *Start* button to start the analysis and display the analysis results. This may take some time as plots are calculated in the background.
6. A user interface for plot selection appears. Select which plots to display.

"Close Figures" in the Dialog Window closes all plots and the TFC_selectplots figure.

3 Parameter specification – The dialog window

In the dialog window two time-frequency data groups can be selected for comparison. Information on the available image data (datafile, condition, channels, type) is displayed. In each list (with the exception of type), multiple entries can be selected. If no entry in a list is highlighted, all entries are interpreted as selected. Names for both groups can be edited and will appear in figure titles.

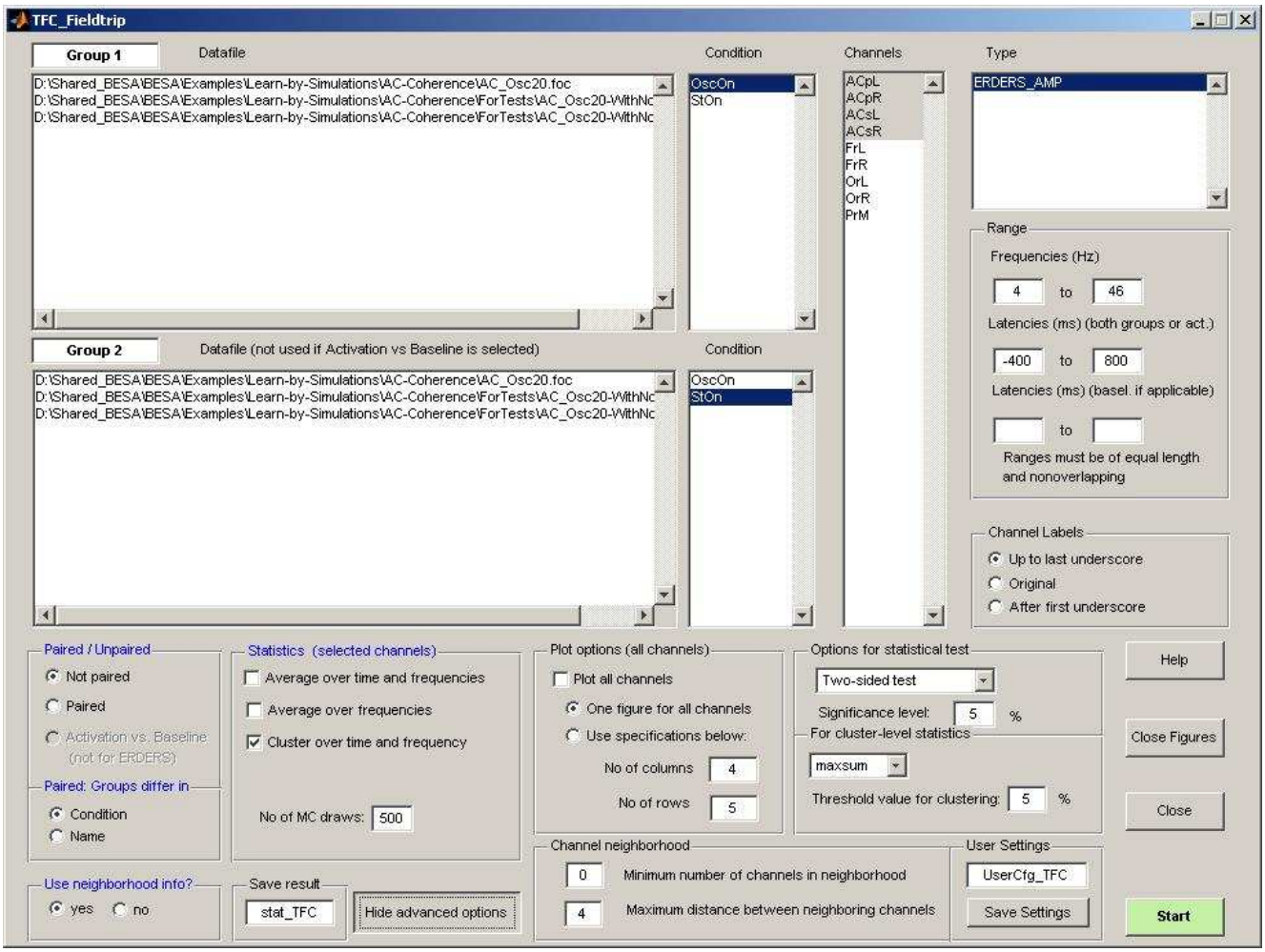

Fig 1: The dialog window

3.1 Data selection

3.1.1 Group selection - Listboxes

The dialog window contains the following lists for TFC data selection:

- **Type**: All available time-frequency data types, e.g. “ERDERS” or “COHERENCE_SQUARED”.

If the loaded structure contains more than one data type, it is necessary to first select a type. Selection then restricts the displayed datafiles, conditions and channels to those available for this image type and grid size.

- **Datafile**: Names of the datafiles for which the TF data have been generated.
- **Condition**: Available conditions.

- **Channels**: All channels in the loaded data structure. To allow for patient-specific montages, channel labels are displayed and compared up to the last underscore in the original label by default. After type selection: All channels available for the selected type. Please note that the presence of different montages for the same type will lead to an error message. To allow for patient-specific montages, channel labels are displayed and compared up to the last underscore in the original label by default.

3.1.2 Group selection – The Paired/Unpaired panel

Select whether images in the two groups are paired. If so, also specify in which property (name or condition) the two members of the pair differ from each other. A paired comparison requires that each image of the first group have a matching partner in the other group differing only in the specified property. ***Activation vs. Baseline***: The criteria datafile, condition and type are used for both groups. In the edit fields within the "Range" panel, time ranges for activation and baseline have to be specified, which must be of equal size and non-overlapping. Group 1 is then filled with the data in the range given for the baseline, group 2 contains the activation range. The option "Activation vs. Baseline" is not available for TFC data of the ERDERS type.

3.1.3 Group selection - The Range panel

The **edit fields** in the **Range** panel display the time and frequency ranges present in the TFC structure. They can be edited to define the time and frequency of interest for the analysis. If the option "Activation vs. Baseline" is chosen, both latency ranges need to be specified, and they have to be of equal size and non-overlapping.

3.2 Statistics panel

3.2.1 What is computed?

- **Grand average:** A grand average of the selected group members over the selected channels and the specified time-frequency range is always computed.
- **Statistics**: Carries out a nonparametric statistical test. For all samples (e.g. (channel, time) pairs or (channel, time, frequency) triplets the experimental conditions are compared using a t-statistic, e.g., possibly selected for clustering and clustered based on temporal, spatial and spectral adjacency. Significance probability is calculated by means of a Monte Carlo method on the cluster level.

3.2.2 Options in the Statistics panel

- **Checkboxes**: Depending on which of the three mutually exclusive **checkboxes** is checked, clustering takes place over the channels, time and frequency unless they are averaged over.
- **Nr of MC draws** (Fieldtrip: *cfg.numrandomization*): Number of draws for the Monte Carlo estimate of the permutation p-value. Default: 500

3.2.3 Further options

- **Use neighborhood info**: Decides whether clustering takes place over adjacent channels ('yes', e.g. for surface sensors) or whether channels are treated separately ('no', e.g. for source montages).

Store results of the statistical test:

- **Edit field in panel "Save result"**: The output of the statistical test is saved in a MATLAB variable 'stat' that is stored to disc to a file whose name can be specified here. The default name is 'stat_TFC', the directory is the editable output path (default: userpath).
- **Show advanced options**: Toggle button for additional options.

3.3 Control buttons:

- **Start:** Start the analysis. Grand averages are calculated and the statistical group comparison is performed. A second window for the selection of figures appears.
- **Close Figures**: Closes the figures generated by the toolbox in the current or earlier sessions..
- **Close**: Closes the graphical user interface (both windows).
- **Advanced**: Opens a dialog window for modification of standard settings (see below).

4 Output

4.1 Output - Plot selection

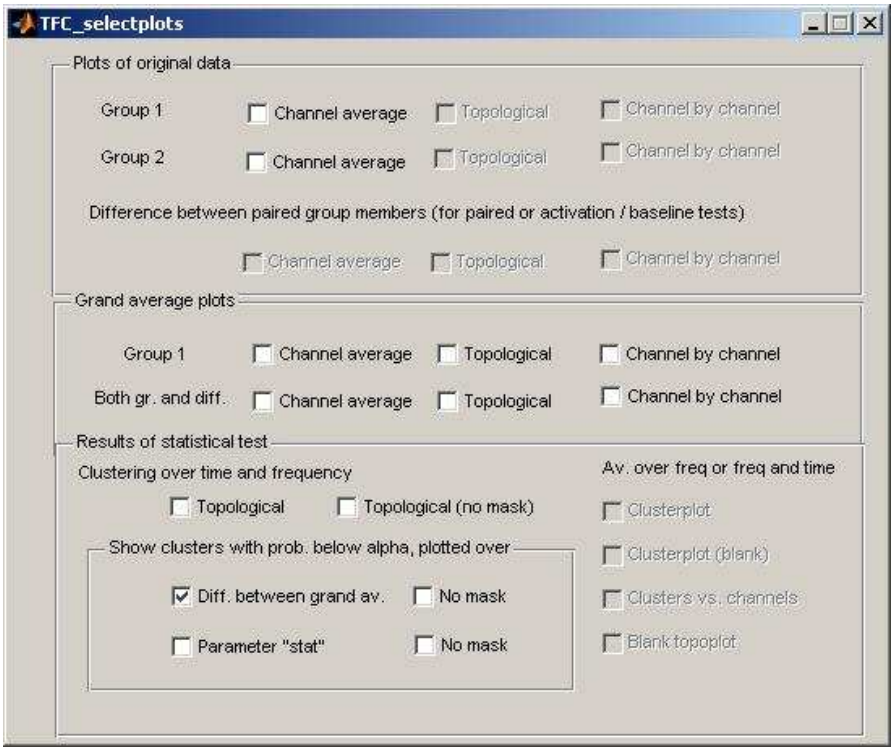

Fig 2: TFC_selectplots: Dialog for plot selection

After calculations are finished, the figure **TFC_selectplots** appears and shows which pre-calculated figures are available. They are set visible if the corresponding checkbox is checked. On unchecking, figures are set invisible again. (Please note that closing the figures by hitting the Close button on the figure window itself closes the figure. Checking the checkbox again will thus result in an error message.)

4.1.1 Panel “Plots of original data”

Displays plots for all members in group 1 and group 2, respectively.

- **Channel average**: One subplot for each group member (data averaged over channels).
- **Topological / Channel by channel**: One figure for each group member; subplots show channels (only enabled if the checkbox “Plot all channels” was checked before the calculation.
    - Topological: Channels plotted in topological arrangement, if possible.

- Channel by channel: Channels plotted in side-by-side subplots. (This option uses the selections in the Plot options panel, viz. “One figure for all channels” or “User settings”).

- ***Difference between paired group members*** Basically the same, with the original data replaced by the difference between both groups. This is possible only if each group member has a matching partner, i.e. for settings “Paired” or “Activation vs. Baseline”.

4.1.2 Panel “Grand average”

- Grand average of group 1
- Grand average of both groups and difference between both groups (group 2 – group 1)

As above: options Channel average, topological plot of all channels and side-by-side plots of all channels.

4.1.3 Panel “Results of statistical test”:

Available plots depend on the test choices.

Average over frequency or frequency and time:

- **Clusterplot** - significant channels highlighted, background: difference between grand averages of both groups
- **Clusterplot (blank)** - significant channels highlighted, background empty for better visibility
- **Clusters vs. channels** - Shows cluster numbers against time or time and channels. Numbers for negative clusters are negative. Clusters with lower absolute numbers have a lower probability, hence are more significant. Significant clusters are shown in red, others in blue.
- **Blank topoplot** - Only for average over time and frequency. Similar to blank clusterplot, but shows all channels.

Clustering over time and frequency:

- **Topological:** Channels in topological order. Background: Difference between grand averages of both groups. White: Time-frequency regions that are not significant.
- **Topological (no mask):** Same without the white regions for better comparison.
- **Diff. between both groups:** Side-by-side plots for all channels. Background: Difference between grand averages of both groups. Only significant clusters shown, remaining regions are white (“masked”)
- **Without mask:** Same without masking for comparison.
- **Parameter “stat”:** Side-by-side plots for all channels. Background: Parameter “stat” (t-statistic value for each sample – decides whether a sample is a candidate for clustering). Only significant clusters shown, remaining regions are white.
- **Without mask:** Same without masking for comparison.

4.2 Output - Text output in the MATLAB command window

During execution of the toolbox, information about the process is displayed in the command window. Some of this is feedback from *Fieldtrip* applications, e.g. information about clustering.

Also provided is information

- on the groups prior to sorting
- on the group members (including name, condition and type)
- on the test settings: paired or not, the factor in which groups differ - important only for paired comparisons (and batch mode) - , and number of Monte Carlo draws
- on the number of clusters found by the test, and whether any are significant.

Remark: As stated in the quoted Fieldtrip tutorial http://fieldtrip.fcdonders.nl/tutorial/cluster_permutation_freq, one concludes that the data in the experimental conditions are different if the p-value is smaller than the critical alpha level. It is not correct – although common usage – to say that a cluster is significant if its p-value is less than the critical alpha-level.

4.3 Output - File output

The result of the statistical test is also stored in a MATLAB variable 'stat' (generated by *Fieldtrip*) that is saved on disc in a file named as specified in the dialog window under 'Save result'. This variable 'stat' contains intermediate and final results of the test. For example, the fields '*posclusters'* and '*negclusters*' of this variable contain the number of detected clusters together with their p-values. (For a detailed explanation, please refer to the *Fieldtrip* tutorials).

5 Advanced options

5.1 The Advanced menu

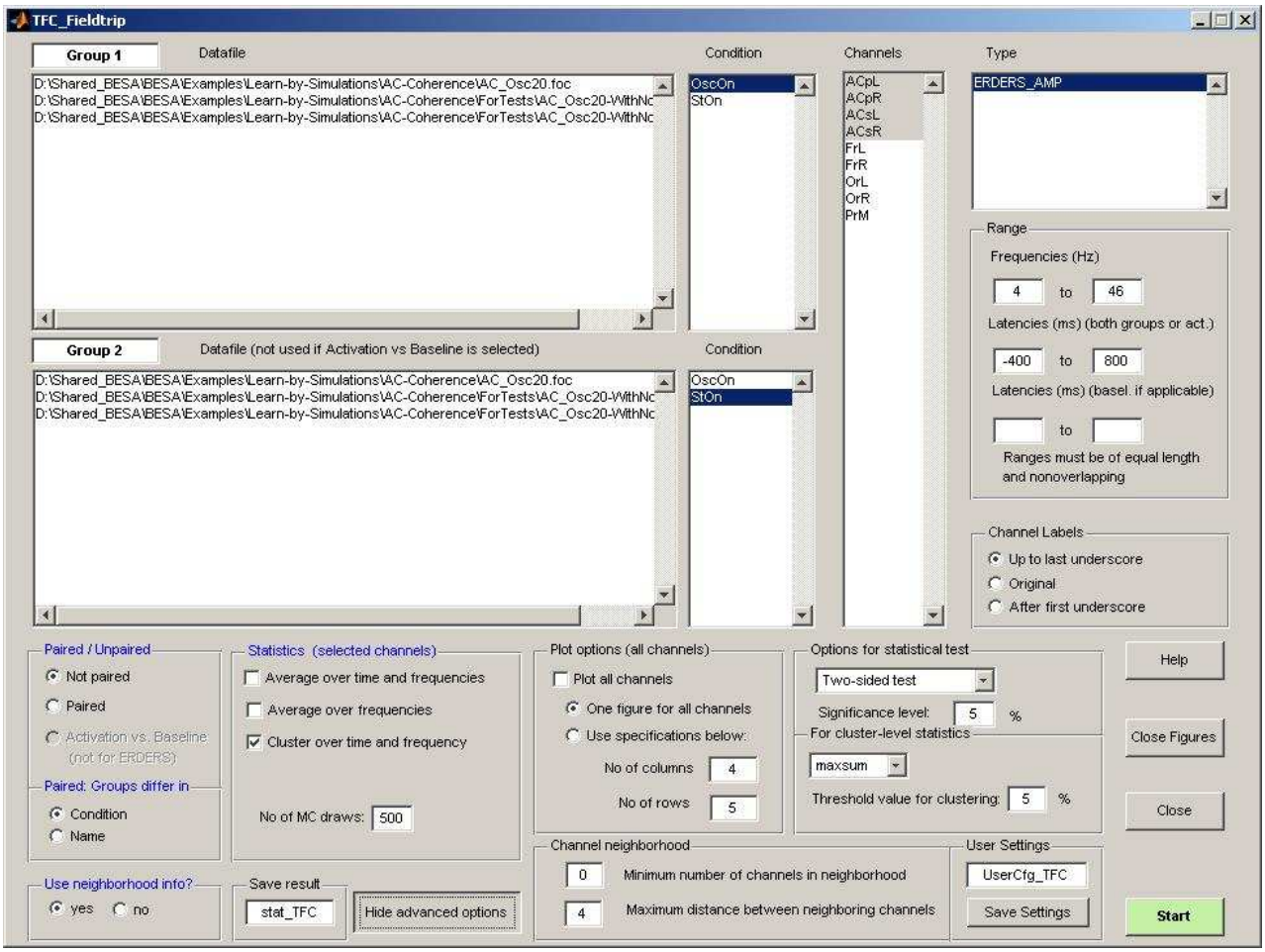

Fig 3: Dialog window with advanced settings

5.1.1 Plot options:

Most plots available for the user settings are pre-calculated while the program runs but stay invisible. After the computation has finished, a second graphical user interface appears. This allows the user to make a choice of all available plots.

Checkbox "Plot all channels": Plots displaying all channels for all group members are not pre-calculated by default. If the checkbox is checked, they are calculated (but invisible). Channels are displayed in a topological arrangement and also in side-by-side plots.

User settings: In some plots, subplots for each channel are displayed in one or more figures. By default, the subplot number is adjusted so that all channels fit into one figure. Since this may result in small subplots – depending on channel number – the *User settings* option allows for the specification of desired numbers and rows in a figure.

5.1.2 Options for the statistical test:

The statistical analysis is performed using the advanced statistical tests implemented in the *Fieldtrip* toolbox. They address the multiple comparisons problem by first clustering the data and performing the hypothesis test using a cluster-based test statistic. It is strongly advisable to read the tutorials on the *Fieldtrip* homepage (e.g., http://fieldtrip.fcdonders.nl/tutorial/cluster_permutation_timelock) and the introductory articles describing the background and the technique on http://fieldtrip.fcdonders.nl/references_to_implemented_methods#statistical_inference_by_means_of_permutation for an understanding of the parameters in the Test Options panel.

- **Drop-down menu** (Fieldtrip: *cfg.tail*): Specify whether to perform a two- or one-sided test.
- **Significance level** (Fieldtrip: *cfg.alpha*): Critical alpha-level for controlling false alarm rate. (Note: In Fieldtrip, the alpha parameter has to be explicitly halved for a two-sided test. Here, this is done within the application). Default: 5%.

Parameters for cluster-level statistics:

- **Threshold value for clustering:** (Fieldtrip: *cfg.clusteralpha*): Threshold for determining whether sample is candidate for clustering (based on t-values quantifying the effect of an experimental condition at each sample)
- **Drop-down menu** (Fieldtrip: *cfg.clusterstatistic*): Select the method to determine test statistic after selected samples are clustered and the sum of the t-values in each cluster is taken (default: maxsum).
- **Channel neighbourhood panel**:
    - **Minimum number of channels in neighborhood** (Fieldtrip: *cfg.minnbchan*): Tuning parameter for forming clusters; minimum number of channels with a positive t-test required for a selected sample. Details in http://fieldtrip.fcdonders.nl/tutorial/cluster_permutation_timelock
    - **Maximum distance between neighboring channels** (Fieldtrip: *cfg.neighbordist*): minimum neighborhood distance, default 4 cm

5.1.3 Channel labels:

To allow for patient-specific montages, channel labels are displayed and compared up to the last underscore in the original label by default. Original channel labels are optionally available through the Channel labels panel.

5.2 Defining the default output folder (optional):

Optionally, the output path of the toolbox can be specified by editing the first section of the MATLAB file *besa_tfc_analysis.m*

- ***OutputPath***: The directory in which the output of the statistical test as well as the user configuration are saved. Default: MATLAB userpath.

6 Automation mode without dialog window

The toolbox can also be run from the command line or from a batch, bypassing the dialog window for parameter selection. This may be useful e.g. if the statistics calculations are to run directly after data preparation. In this case, settings for the calculation must have been stored earlier from the dialog window, using the ***Save Settings*** button in the User Settings panel.

Preparation:

1. Create TFC data with the same condition, time and frequency ranges, type and montage as the data which will be analyzed later on. The filenames do not have to be the same, so the groups can be much smaller than the actual analysis data.
2. In *besa_tfc_analysis.m*, edit the output path if desired.
3. Call *besa_tfc_analysis(besa_tfc_all)* to open the dialog window.
4. Specify the two groups and statistics parameters. (Choices in the datafile lists will not be saved.)
5. Edit a name in the User Settings panel (default: UserCfg_TFC) and hit the “Save Settings” button. The GUI may now be closed.
6. Create the TFC data for testing. Call *besa_tfc_analysis(besa_tfc_all, n; ‘UserCfg_TFC’)* where *n* is an integer specifying the number of Monte Carlo draws and *UserCfg_TFC* is the name under which the settings have been saved.