

TRENTOOL workshop

TRansfer **EN**tropy **TOOL**box

Dr. Michael Lindner

TRENTOOL

TRENTOOL

- TRENTOOL
 - is a Matlab based toolbox
 - is designed to use transfer entropy (TE) on any kind of time series data.
 - is designed to make the non trivial handling of TE easy for all users.
 - offers solutions for optimizing parameters of TE calculation.
 - Infos and download: www.trentool.de

The main functions of TRENTOOL

- TEprepare
- TEsurrogatestats
- TEconditionstatssingle
- TEgroup_prepare
- TEgroup_calculate
- TEgroup_stats
- TEplot2D

The main functions of TRENTOOL

- TEprepare
- TEsurrogatestats
- TEconditionstatssingle
- TEgroup_prepare
- TEgroup_calculate
- TEgroup_stats
- TEplot2D

The main functions of TRENTOOL

- TEprepare

single subject
analyses

- TEsurrogatestats
- TEconditionstatssingle

- TEgroup_prepare
- TEgroup_calculate
- TEgroup_stats
- TEplot2D

The main functions of TRENTOOL

- TEprepare
- TESurrogatestats
- TEconditionstatssingle

group
analyses

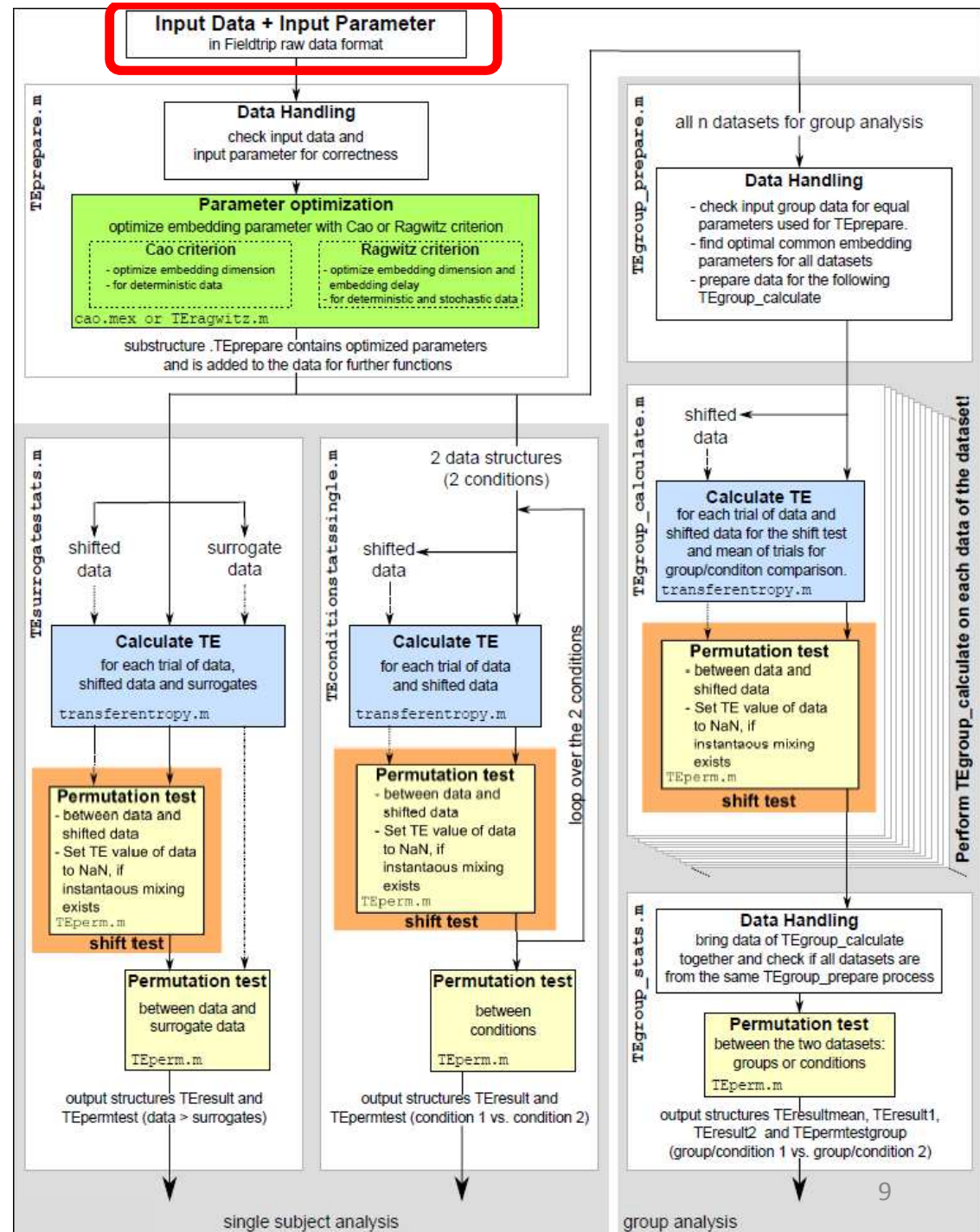
- TEgroup_prepare
 - TEgroup_calculate
 - TEgroup_stats
- TEplot2D

The main functions of TRENTOOL

- TEprepare
- TEsurrogatestats
- TEconditionstatssingle
- TEgroup_prepare
- TEgroup_calculate
- TEgroup_stats
- TEplot2D

Plotting results

Overview TRENTOOL



Input - Data

- In Fieldtrip raw data structure:

Matlab structure containing:

- {trials} ← (channel x time)
- {time} ← (time indices)
- {labels} ← (labels of the channels)
- fsample

{ } = Cells

Input - Data

- In Fieldtrip raw data structure:

Matlab structure containing:

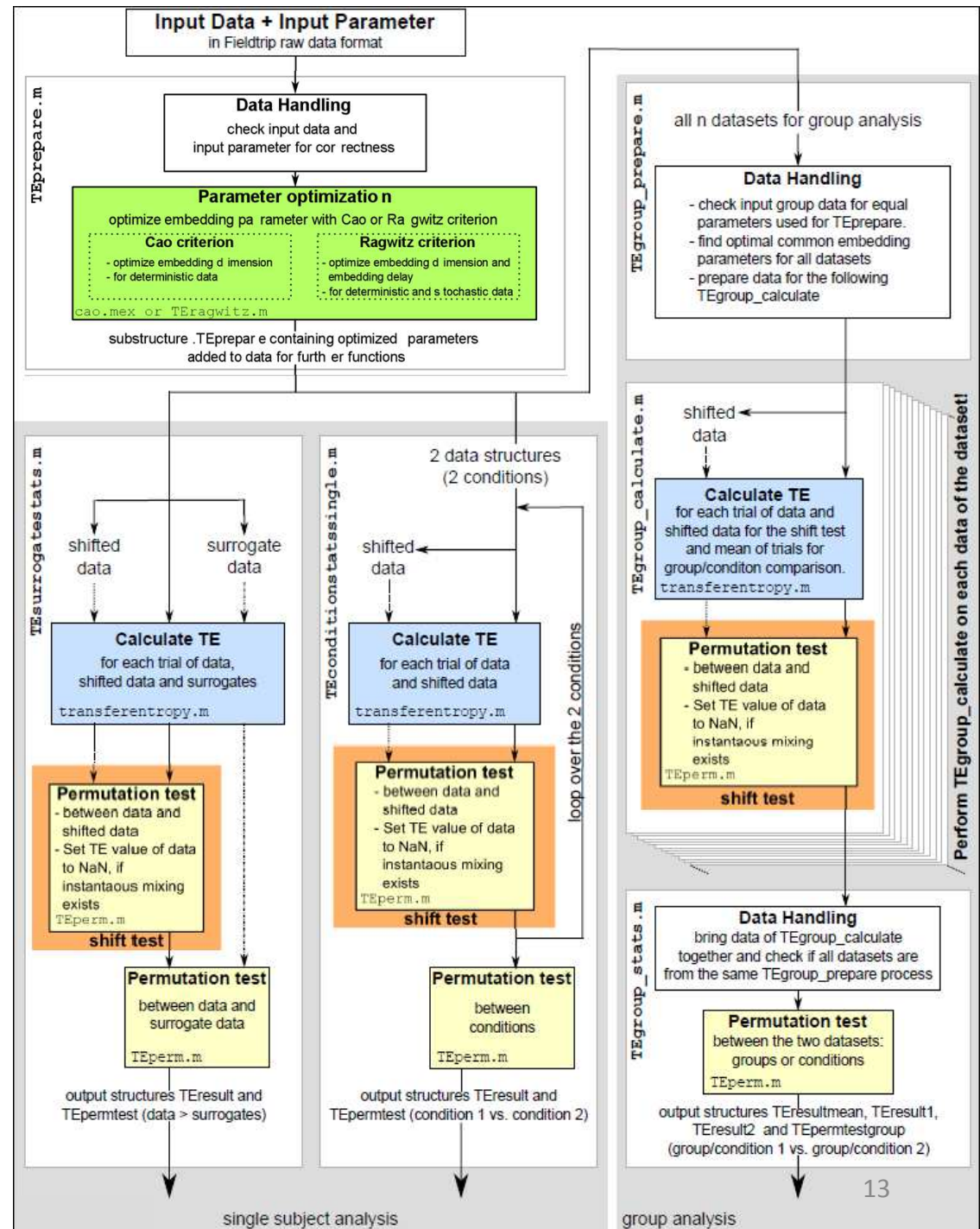
- {trial} (1xnumber of trials) {} = Cells
- {time} (1xnumber of trials)
- {labels} (1xnumber of channels)
- fsample

help TEprepare

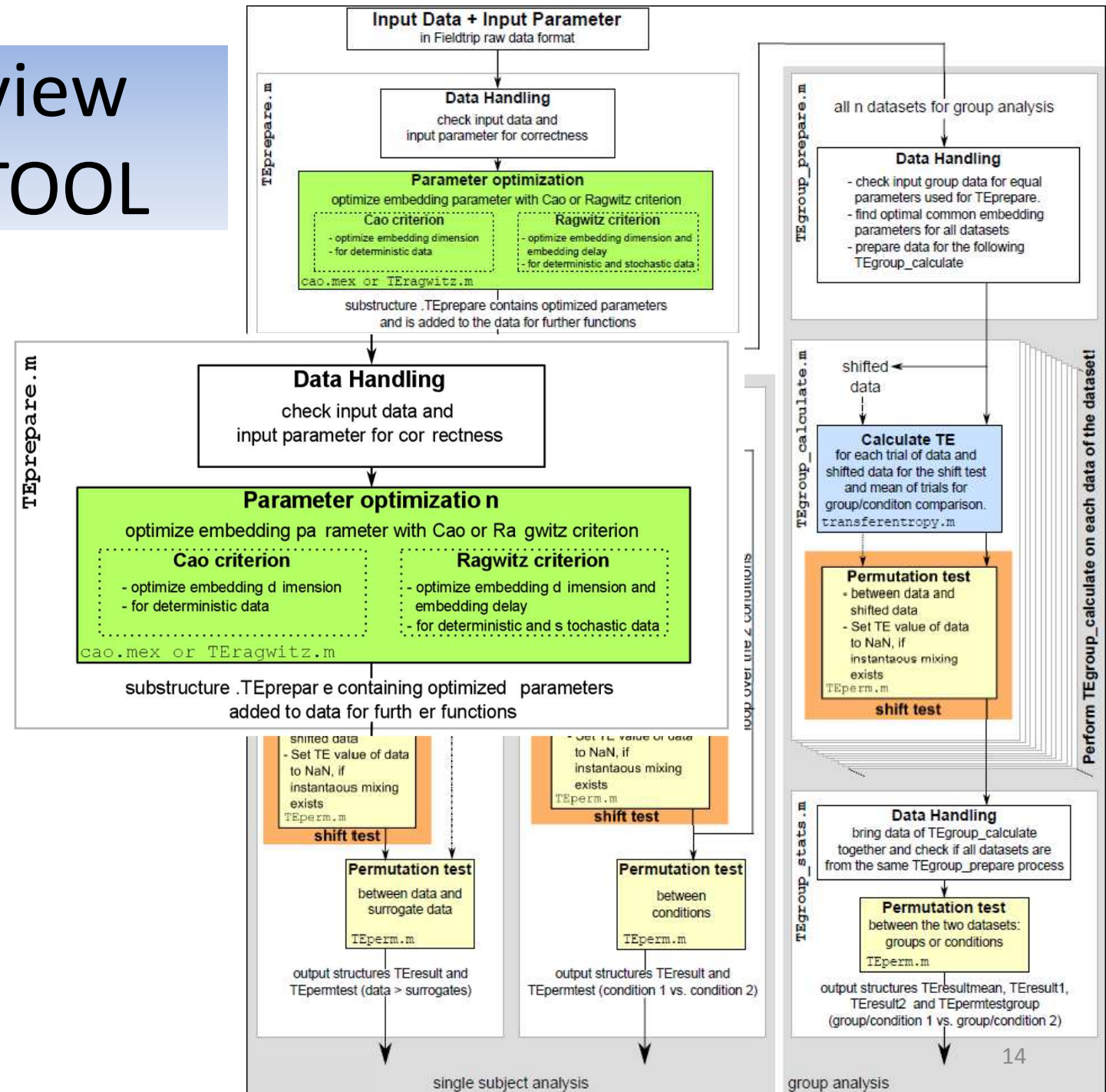
Input - Parameter

- Parameters are given using a MATLAB structure (typically called „cfg“ - configuration)
- Type: `help <functionname>` to see which parameter is needed.

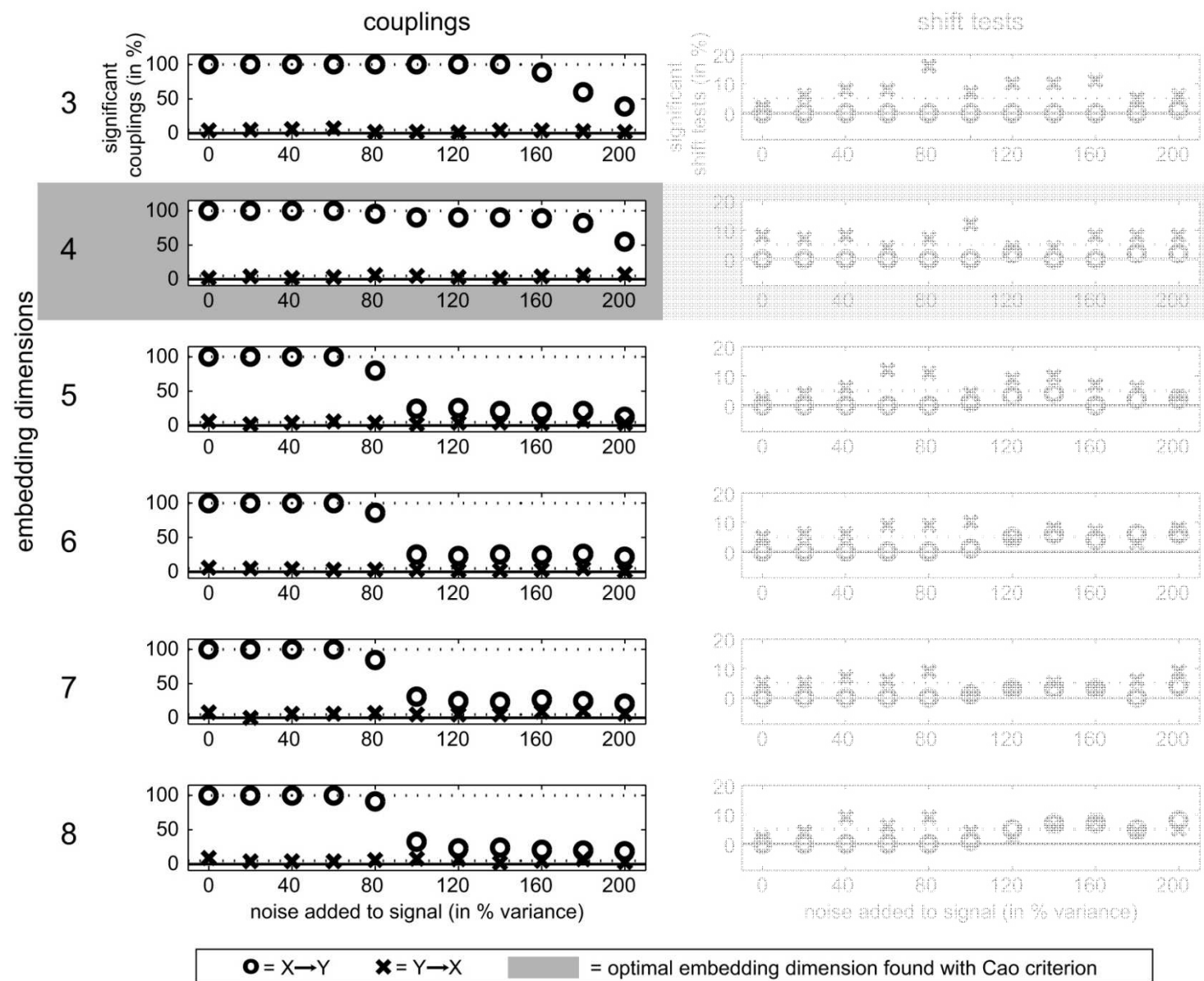
Overview TRENTOOL



Overview TRENTOOL



Parameter optimization



All parameters for TPrepare

cfg.sgncmb = list of channelpairs
cell array (Nx(source, target))

or

cfg.channel = list of channels - testing will be done all-by-all

and

cfg.Path2TSTOOL = Path to the folder including the TSTOOL package

cfg.toi = the time range of interest (vector 1 x 2) in seconds
e.g. (time_from, time_to) (units: seconds)

cfg.predicttime_u = time ahead for the advance prediction (scalar, in ms)

cfg.optimizemethod = Method to optimize parameters: 'ragwitz' or 'cao'
if you choose 'ragwitz':

cfg.ragdim = range of embedding dimensions to scan vector
(1xnumdim)

cfg.ragtaurange = vector (1x2) of min and max embedding delays (in multiples of the autocorrelation decay time)

cfg.ragtausteps = number of equidistant steps in ragtaurange
(min 5) (default = 10)

cfg.flagNei = 'Range' or 'Mass' type of neighbor search

cfg.sizeNei = Radius or mass for the neighbor search according to flagNeighborhood

cfg.repPred = repPred represents the number of points for which the prediction is performed (it has to be smaller than length(timeSeries)-(dimEmb-1)*tauEmb-u)

if you choose 'cao' (recommended for fMRI Data):

cfg.caodim = range of embedding dimension to scan with stepwidth 1 (vector 1 x numdim)
(default = [1,2,3,4,5,6,7,8,9,10])

cfg.caokth_neighbors = number of neighbors for fixed mass search for cao (controls balance of bias/statistical errors) (default = 4)

cfg.caotau = embedding delay in units of ACT (x*ACT)
(default = 1.5)

cfg.kth_neighbors = number of neighbors for fixed mass search
(controls balance of bias/statistical errors)
(default = 4)

cfg.TheilerT = number of temporal neighbors excluded to avoid serial correlations (Theiler correction) (default = ACT)

cfg.trialselect = ACT thresholding of trials - 'ACT', 'range' or 'no'
(default = 'ACT')

if you chose 'ACT' (or nothing):

cfg.actthvalue = max threshold for ACT; min threshold

cfg.minnrtrials = minimum Nr of trials with ACT < actthrest used to calculate transfer entropy

if you chose 'range':

cfg.trial_from = Inferior limit for the trials to be considered

cfg.trial_to = Superior limit for the trials to be considered

cfg.maxlag = the range of lags for computing the auto correlation time: from -MAXLAG to MAXLAG (default = 1000)

cfg.TEcalctype = 'V' : self-prediction of the target signal and cross-prediction are both made from states in source and target that precede the target state to be predicted by cfg.predicttime_u.
'VW' : the self-prediction time for the target is tau and cross-predictions are made from source states that precede the target state to be predicted by cfg.predicttime_u.
(to solve the problem of decreasing self-prediction accuracy for large prediction times)
(default = 'VW')

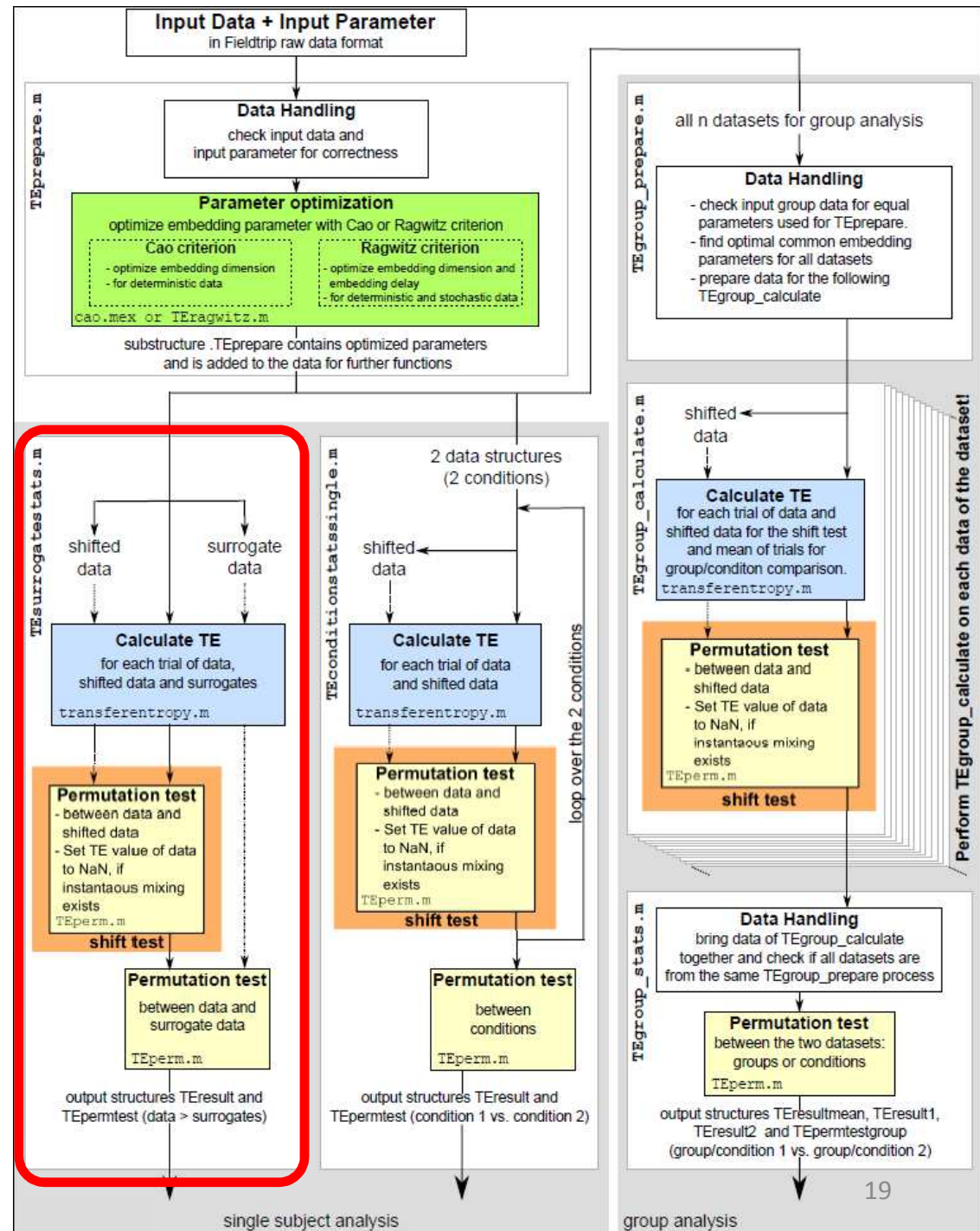
Possible parameter set for TEprepare

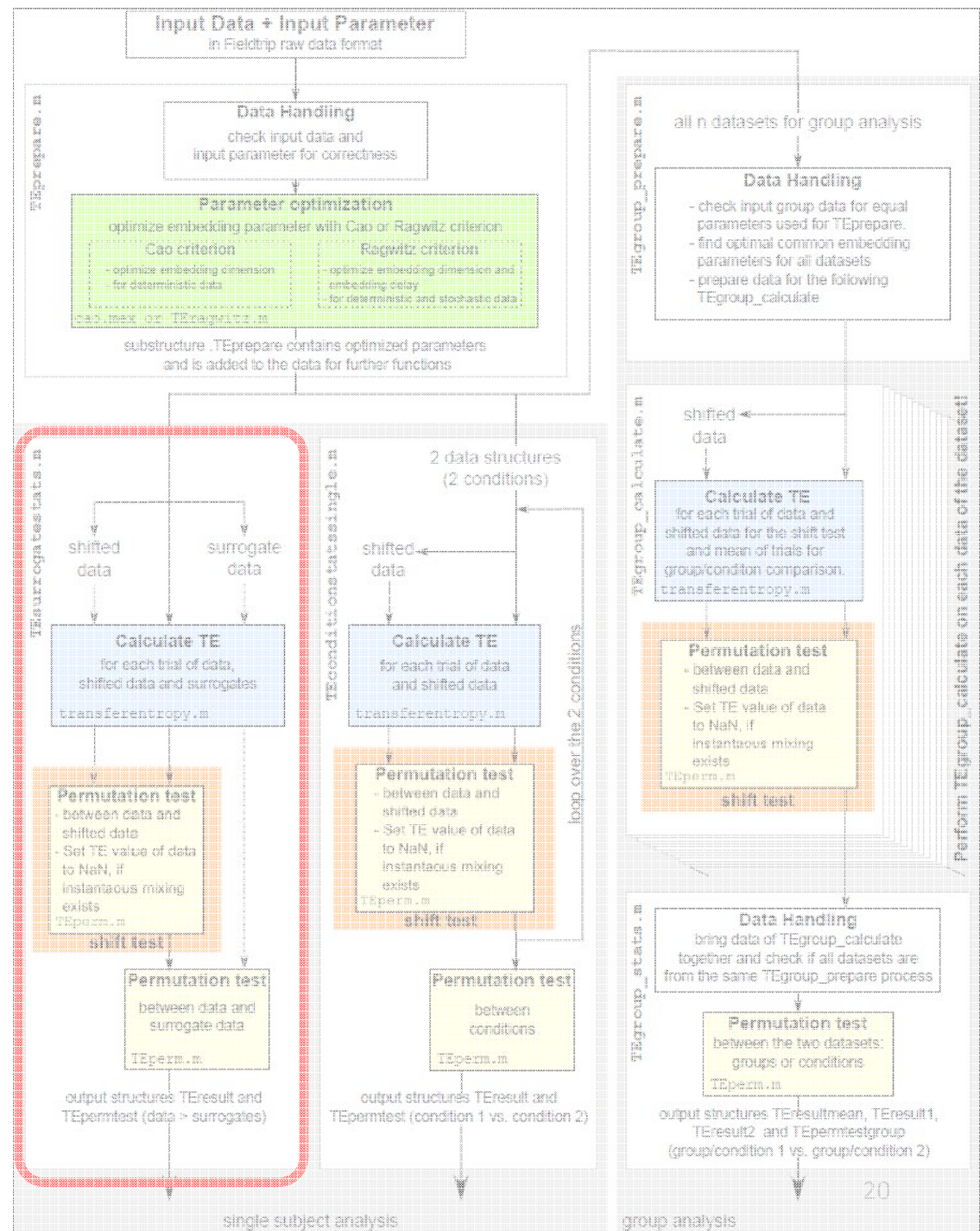
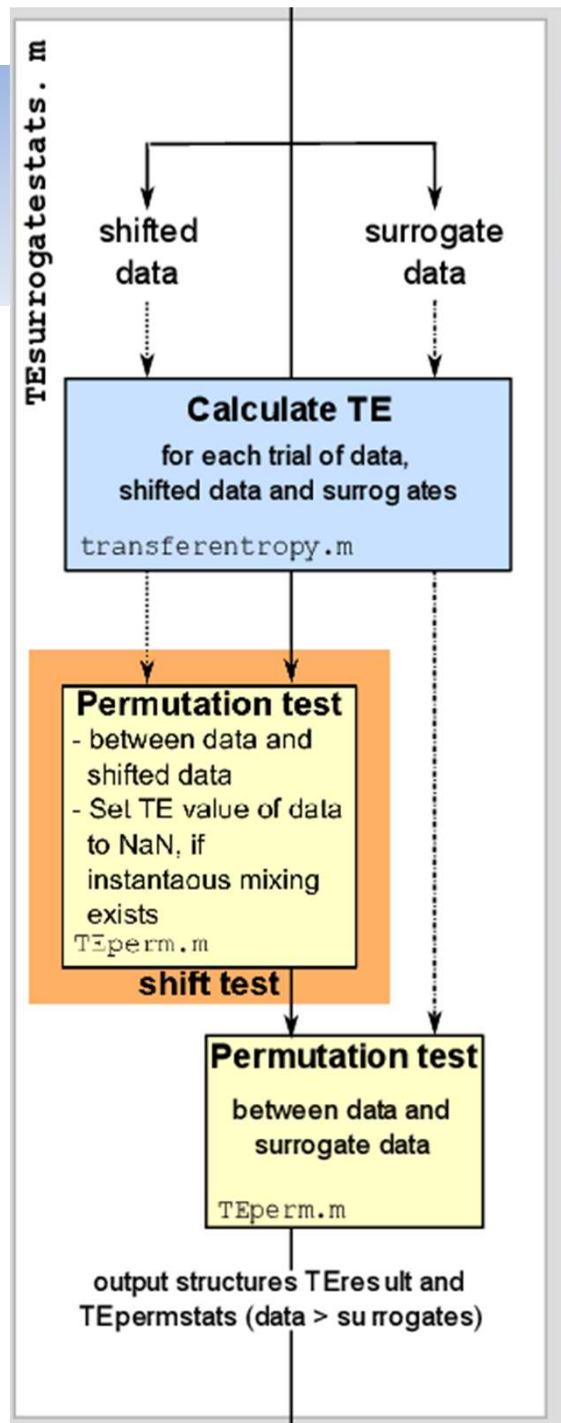
```
cfg.toi = [0 2.999];  
cfg.sgncmb = {'X', 'Y';  
              'Y', 'X'};  
cfg.optimizemethod = 'cao';  
cfg.caodim = 1:6;  
cfg.trialselect = 'ACT';  
cfg.actthrvalue = 120;  
cfg.minnrtrials = 30;  
cfg.Path2TSTOOL = '/data/common/OpenTSTOOL';  
cfg.predicttime_u = 21;
```

Run TEprepare

```
Data_prepared = TEprepare(cfg,Data)
```

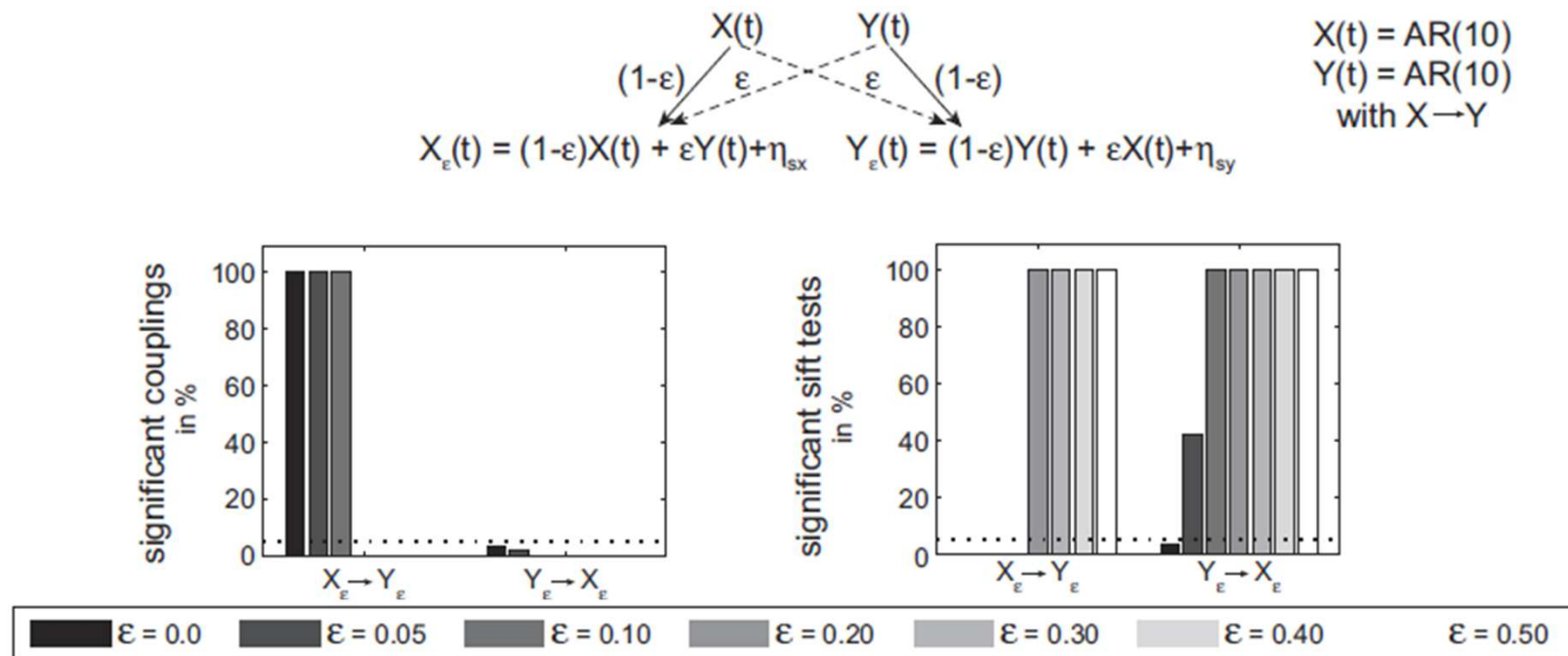
Overview TRENTOOL



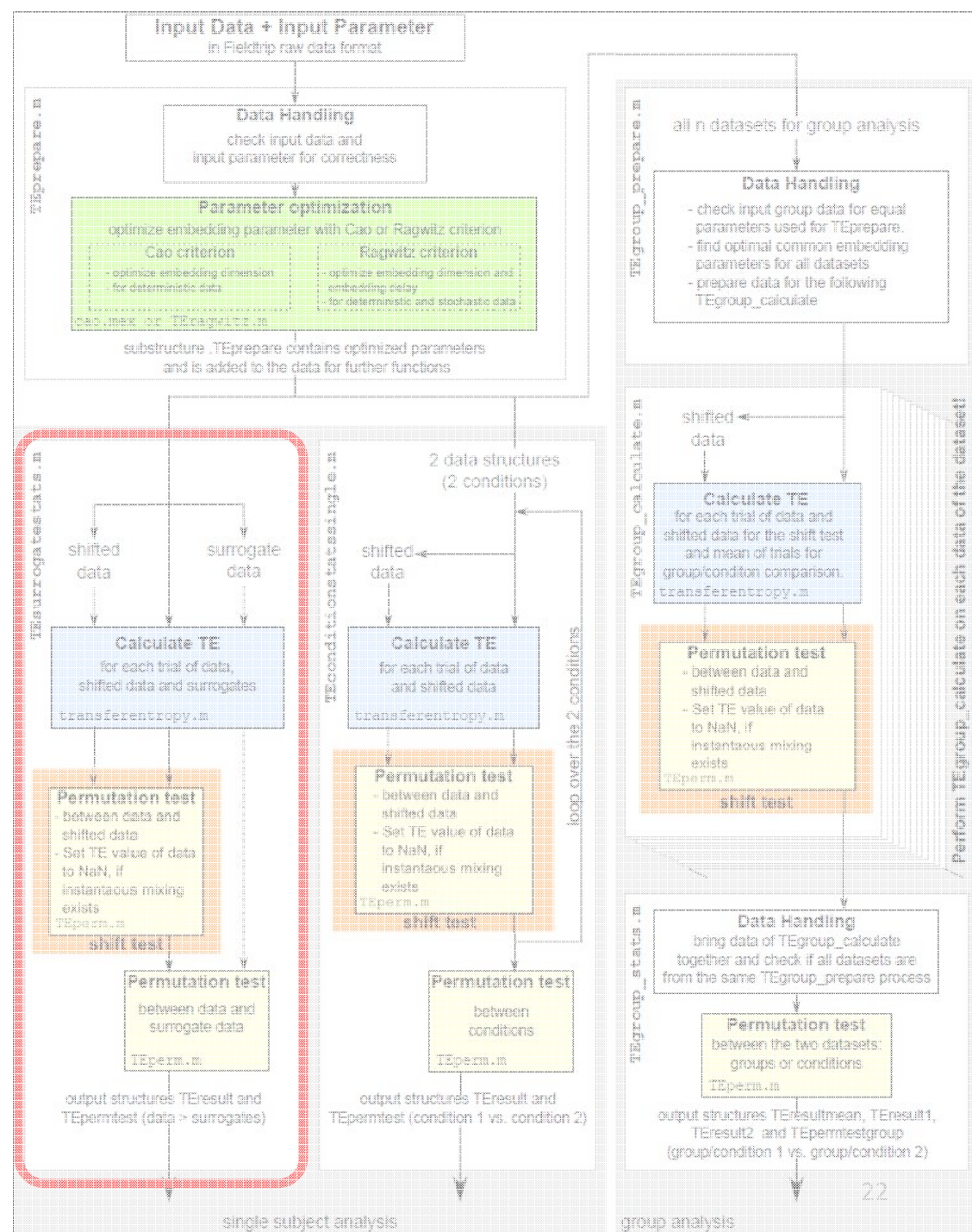
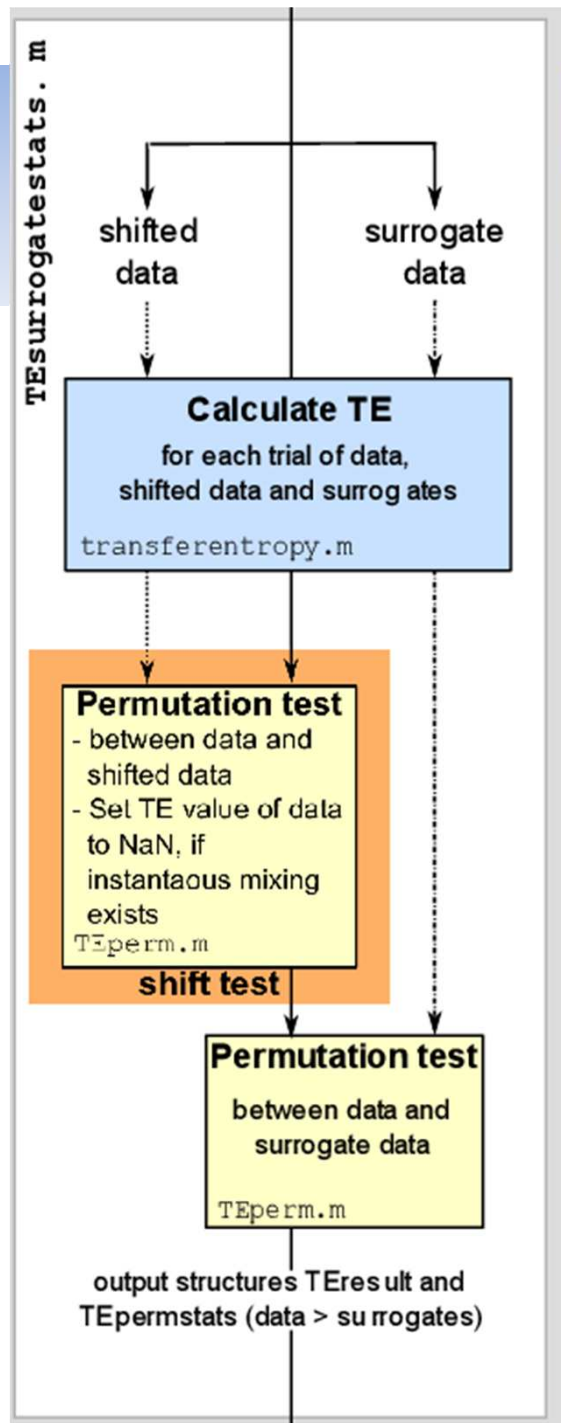


Shift test

- detecting volume conduction in mixed signals (e.g. MEG)
- Permutation test between TE of the data and of the shifted data



Wibral, Rahm, Rieder, Lindner, Vicente, Kaiser (2011)



TEresult

- in file with the suffix ,_TE_output'
- Matlab structure
- contains
 - .TEmat : Transfer entropy values
 - .MImat : Mutual informaton values
 - .act : autocorrelation decay time values
 - .trials : trial numbers used for TE calculation
 - .TEprepare
 - ...

TEresult

- in file with the suffix ,_TE_output'
- Matlab structure
- contains
 - .TEmat : Transfer entropy values
 - .MIimat : Mutual informaton values
 - .act : autocorrelation decay time values
 - .trials : trial numbers used for TE calulation
 - .TEprepare
 - ...

TEpermtest

- in file with suffix ,_TEpermtest_output'
- MATLAB structure
- contains
 - .TEpermvalues (number of channelpairs x 5) :
 - .TEprepare
 - (.TEgrouppprepare)
 - ...

TEpermtest

- in file with suffix ,_TEpermtest_output'
- MATLAB structure
- contains
 - .TEpermvalues (number of channelpairs x 5) :
 - .TEprepare
 - (.TEgrouppprepare)
 - ...

TEpermtest

- in file with suffix ,_TEpermtest_output'
- MATLAB structure
- contains
 - .TEpermvalues (number of channelpairs x 5) :
 1. p-value
 2. significance (uncorrected)
 3. significance (corrected for multiple comparisons)
 4. statistic value (mean or t-value)
 5. Instantaneous mixing exists or not
 - .TEprepare
 - (.TEgrouppprepare)
 - ...



All parameters for TEsurrogatestats

cfg.optdimusage = 'maxdim' to use maximum of optimal dimensions over all channels for all channels, or 'indivdim' to use the individual optimal dimension for each channel. In case of using ragwitz criterion also the optimal embedding delay tau per channelcombi is used.

cfg.dim = Value(s) for embedding dimension. In case of using cfg.optdimusage = 'maxdim' this has to be a scalar value. In case of cfg.optdimusage = 'indivdim' this has to be a vector of the size (channelcombi x 1). If not specified, the optimal dimension(s) found in TEprepare will be used, which is the recommended option!

cfg.tau = embedding delay in units of act (x*act). If not specified (recommended option), the tau is used as followed:
Depending optimizemethod in TEprepare:
 'ragwitz' = optimal tau found via ragwitz criterion
 'cao' = cfg.tau given by user in TEprepare
If not specified, the optimal embedding delay found in TEprepare will be used, which is the recommended option!

cfg.alpha = significance level for statistical permutation test and correction for multiple comparison (default = 0.05)

cfg.surrogatetype = 'trialshuffling', 'trialreverse', 'blockresampling', 'blockreverse1', 'blockreverse2', or 'blockreverse3', surrogate data for trial(n) will be created as following:
trialshuffling: trial(n+1)
trialreverse: reverse of trial(n)
blockresampling: cuts trial(n) at random point and resamples the trial
blockreverse1: reverse after blockresampling
blockreverse2: reverse first block after blockresampling
blockreverse3: reverse second block after

blockresampling
swapneighbors: pair odd trials with the higher neighbor and 3even with the lower neighbor

cfg.shifttest = perform shift test to identify instantaneous mixing between the signal pairs. Values: 'yes' or 'no' (default = 'yes')
This shift test is important for EEG and MEG data, because linear mixing is always present in the data. In case of instantaneous mixing transfer entropy should not be calculated for the affected channelpairs with the corresponding parameter sets, because it could result in false positive results. Hence the TE values for these cases will be set to NaN and the corresponding p-values of the permutation test to 1.

cfg.shifttesttype = The shift test can be calculated for the direction TE value of original data > TE values of shifted data (value = 'TE>TEshift') or for the other direction (value = 'TEshift>TE'). In this case the alpha is set to 0.1 . (default = 'TE>TEshift')

cfg.shifttype = Shifting the data 'onesample' or the length of the 'predicttime' (default = 'predicttime')

cfg.numpermutation = nr of permutations in permutation test (default = 190100)

cfg.permstatstype = 'mean' to use the distribution of the mean differences and 'depsamplesT' or 'indepsamplesT' for distribution of the t-values. (default = 'indepsamplesT')

cfg.tail = 1 tail or 2 tailed test of significance (for the permutation tests) (default in TEsurrogatestats= 1)

cfg.correctm = correction method used for correction of the multiple comparison problem - False discovery rate 'FDR' or Bonferroni correction 'BONF' (default = 'FDR')

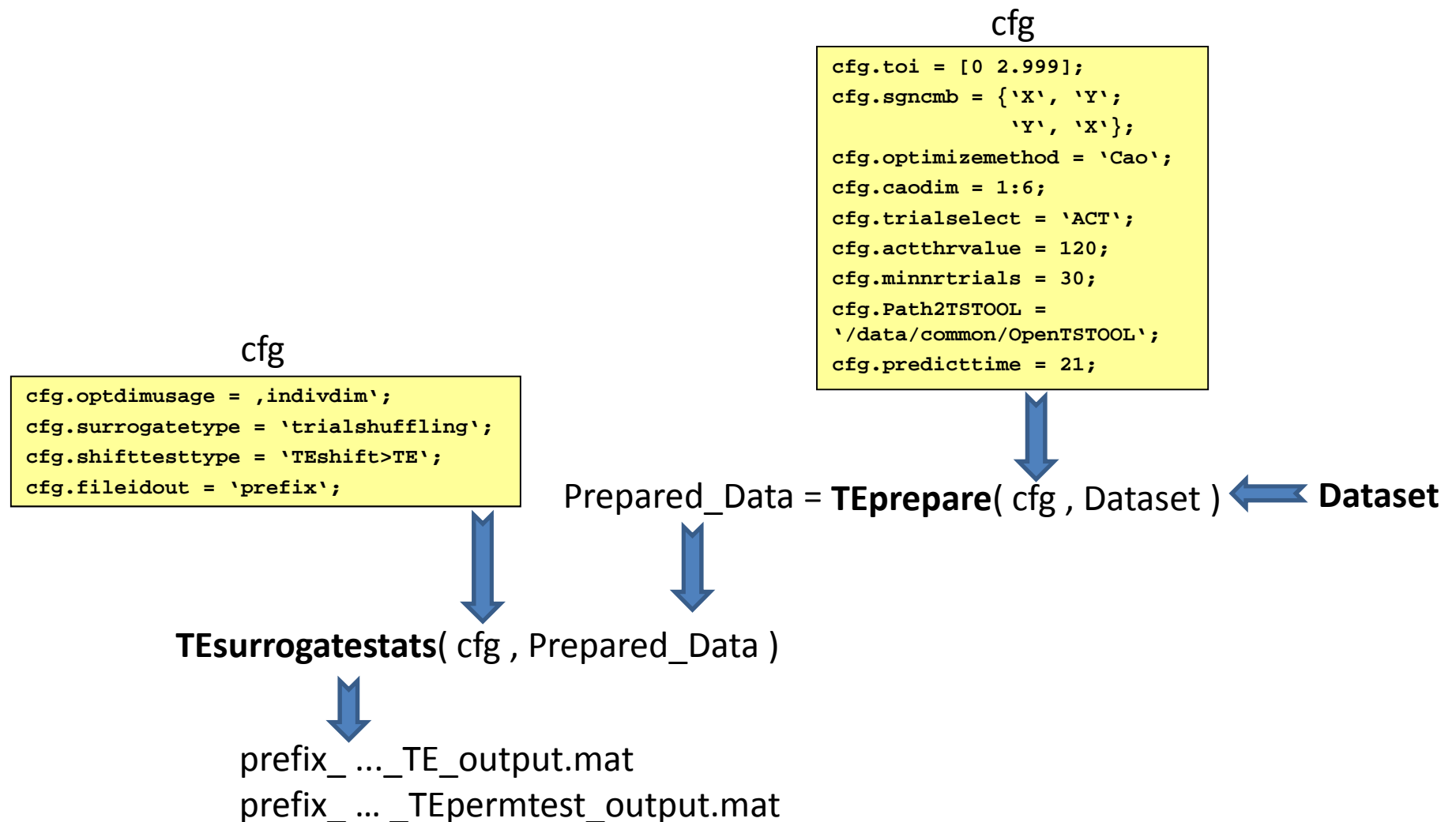
cfg.fileidout = string for the first part of the output filename.

Example parameter set for TEsurrogatestats

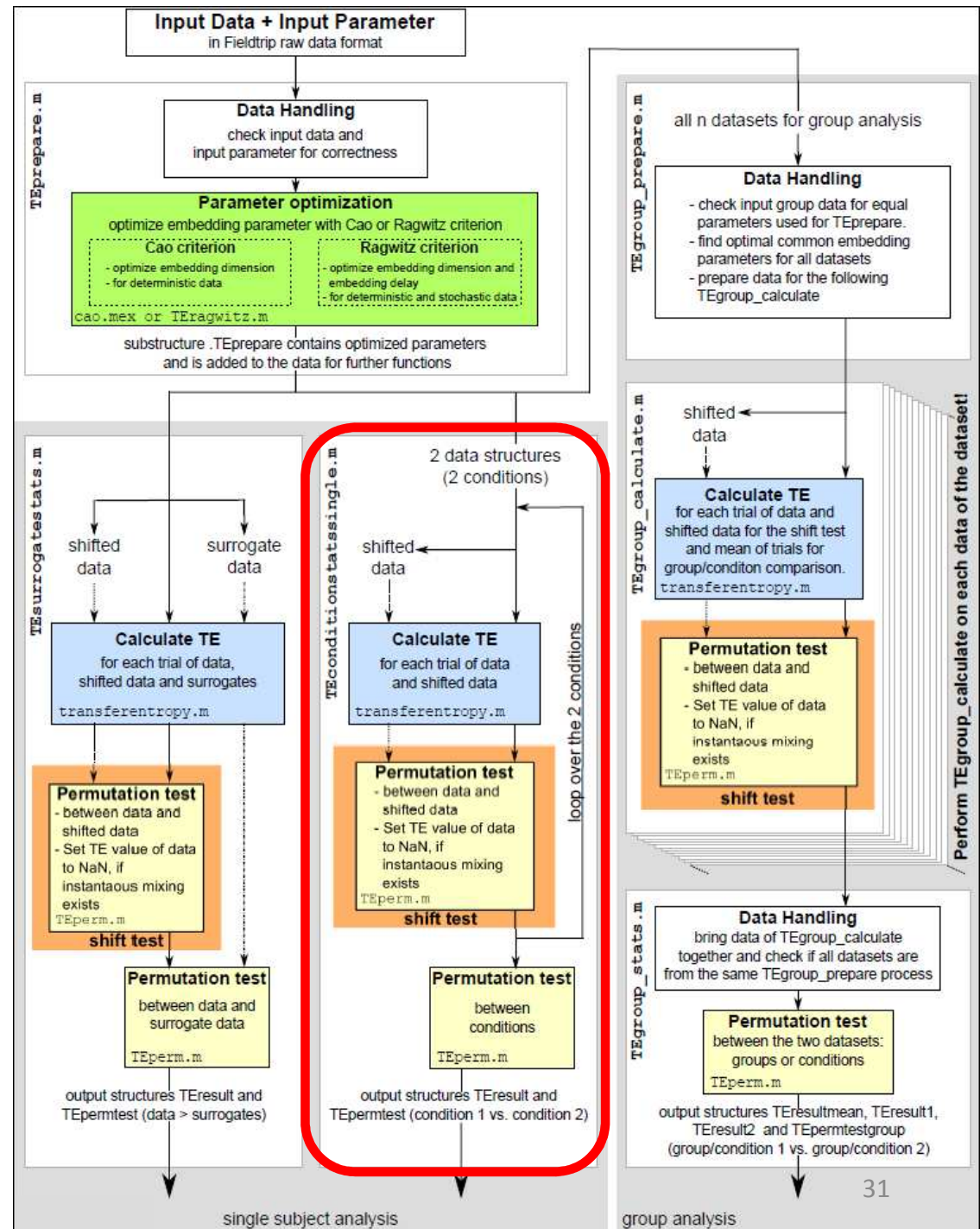
```
cfg.optdimusage = 'indivdim'  
cfg.surrogatetype = 'trialshuffling';  
cfg.shifttesttype = 'TEshift>TE';  
cfg.fileidout = 'your_choice';
```

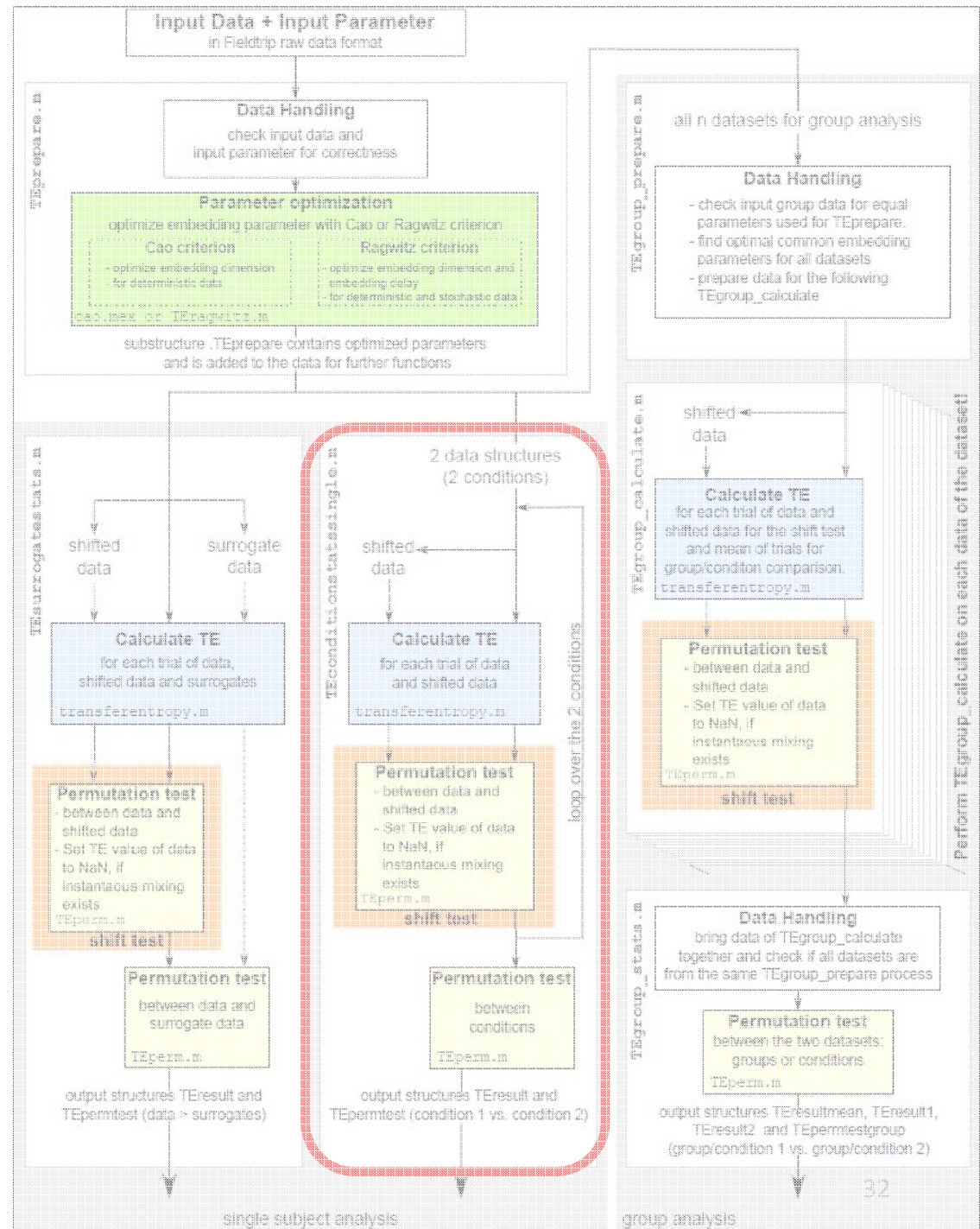
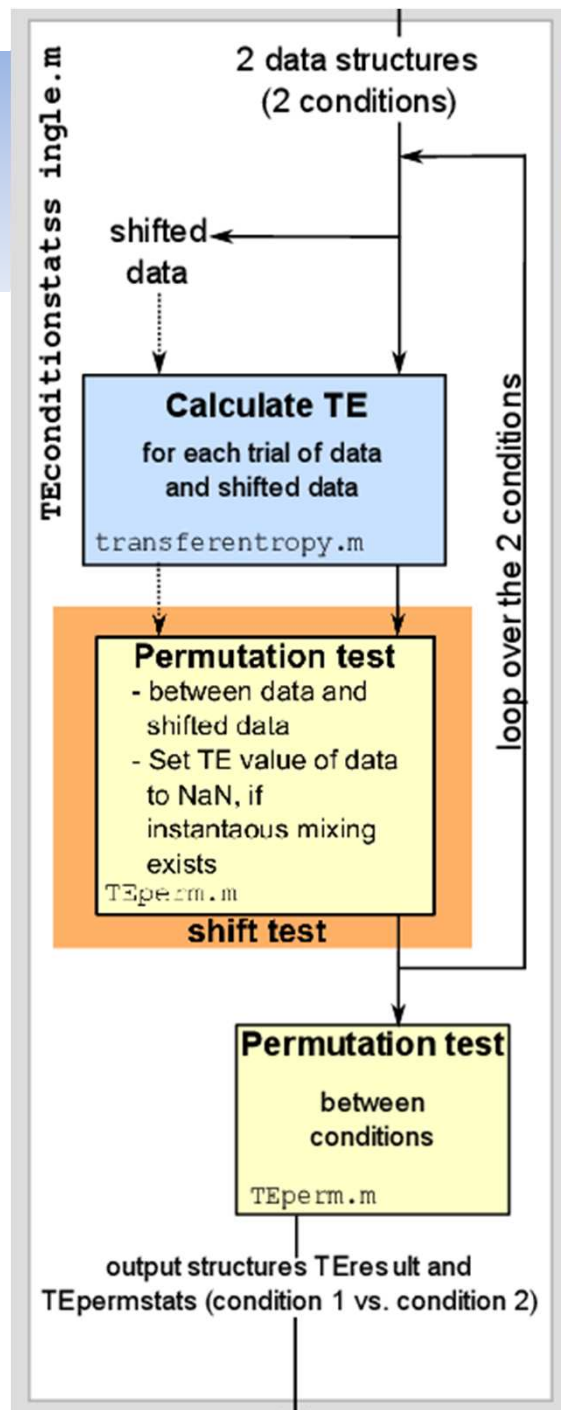
```
TEsurrogatestats(cfg, Data_prepared)
```

Example 1 - TEsurrogatestats



Overview TRENTOOL





All parameters for TEconditionstatssingle

cfg.dim = Scalar value for embedding dimension.
If not specified, the optimal dimension(s) found in TEprepare will be used, which is the recommended option!

cfg.tau = embedding delay in units of act ($x \cdot \text{act}$). If not specified (recommended option), the tau is used as followed:
In case of optimizemethod in TEprepare:
 'ragwitz' = optimal tau found via ragwitz criterion
 'cao' = cfg.tau given by user in TEprepare

cfg.alpha = significance level for statistical shift test, permutation test and correction for multiple comparison (default = 0.05)

cfg.shifttest = perform shift test to identify instantaneous mixing between the signal pairs. Values: 'yes' or 'no' (default = 'yes')
This shift test is important for EEG and MEG data, because linear mixing is always present in the data. In case of instantaneous mixing transfer entropy should not be calculated for the affected channelpairs with the corresponding parameter sets, because it could result in false positive results. Hence the TE values for these cases will be set to NaN and the corresponding p-values of the permutation test to 1.

cfg.shifttesttype = The shift test can be calculated for the direction TE value of original data > TE values of shifted data (value = 'TE>TEshift') or for the other direction (value = 'TEshift>TE'). In this case the alpha is set to 0.1. (default = 'TE>TEshift')

cfg.shifttype = Shifting the data 'onesample' or the length of the 'predicttime' (default = 'predicttime')

cfg.numpermutation = nr of permutations in permutation test (default = 190100)

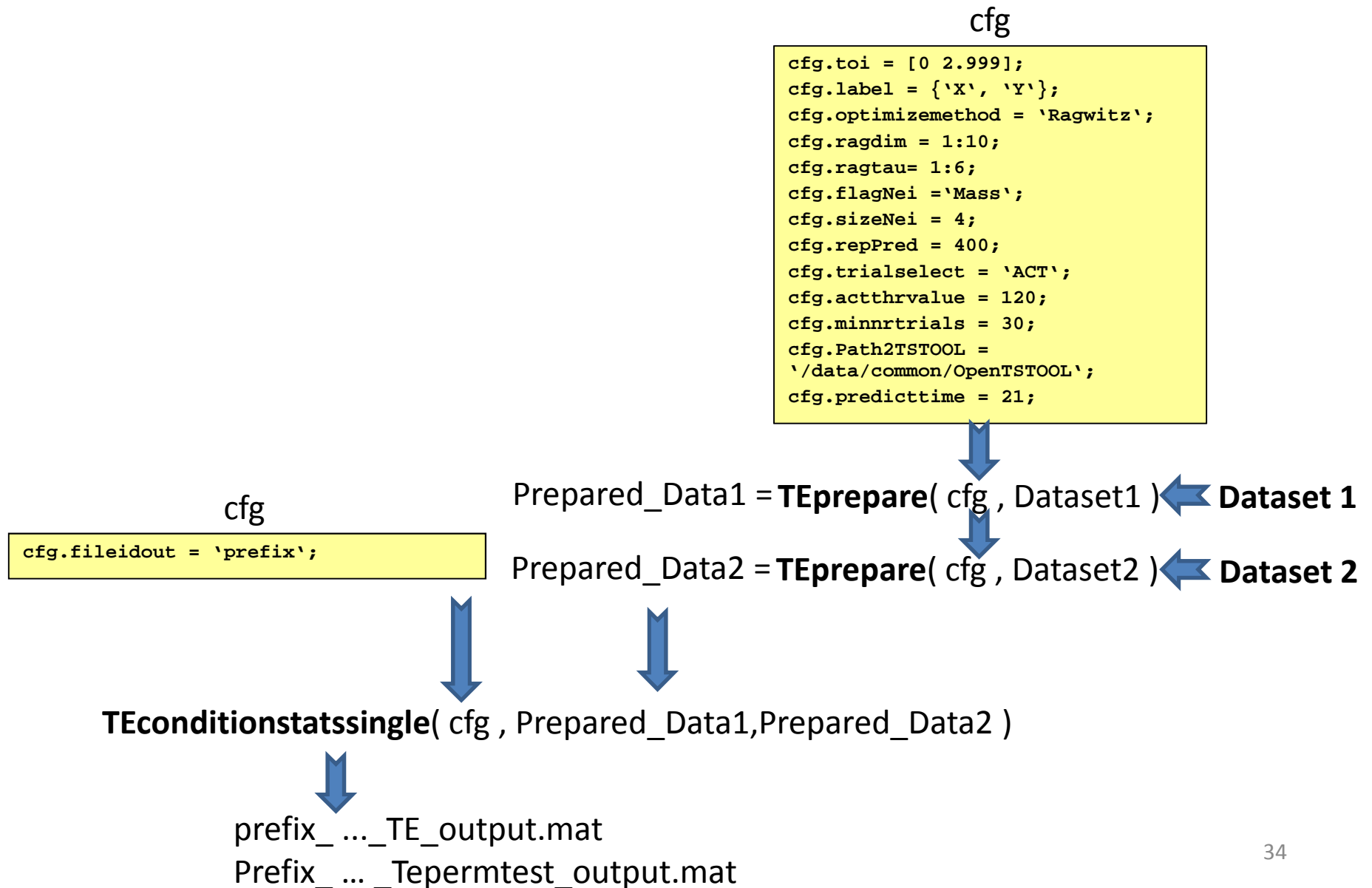
cfg.permstatstype = 'mean' to use the distribution of the mean differences and 'depsamplesT' or 'indepsamplesT' for distribution of the t-values. (default = 'mean')

cfg.tail = 1 tail or 2 tailed test of significance (for the permutation tests) (default = 2)

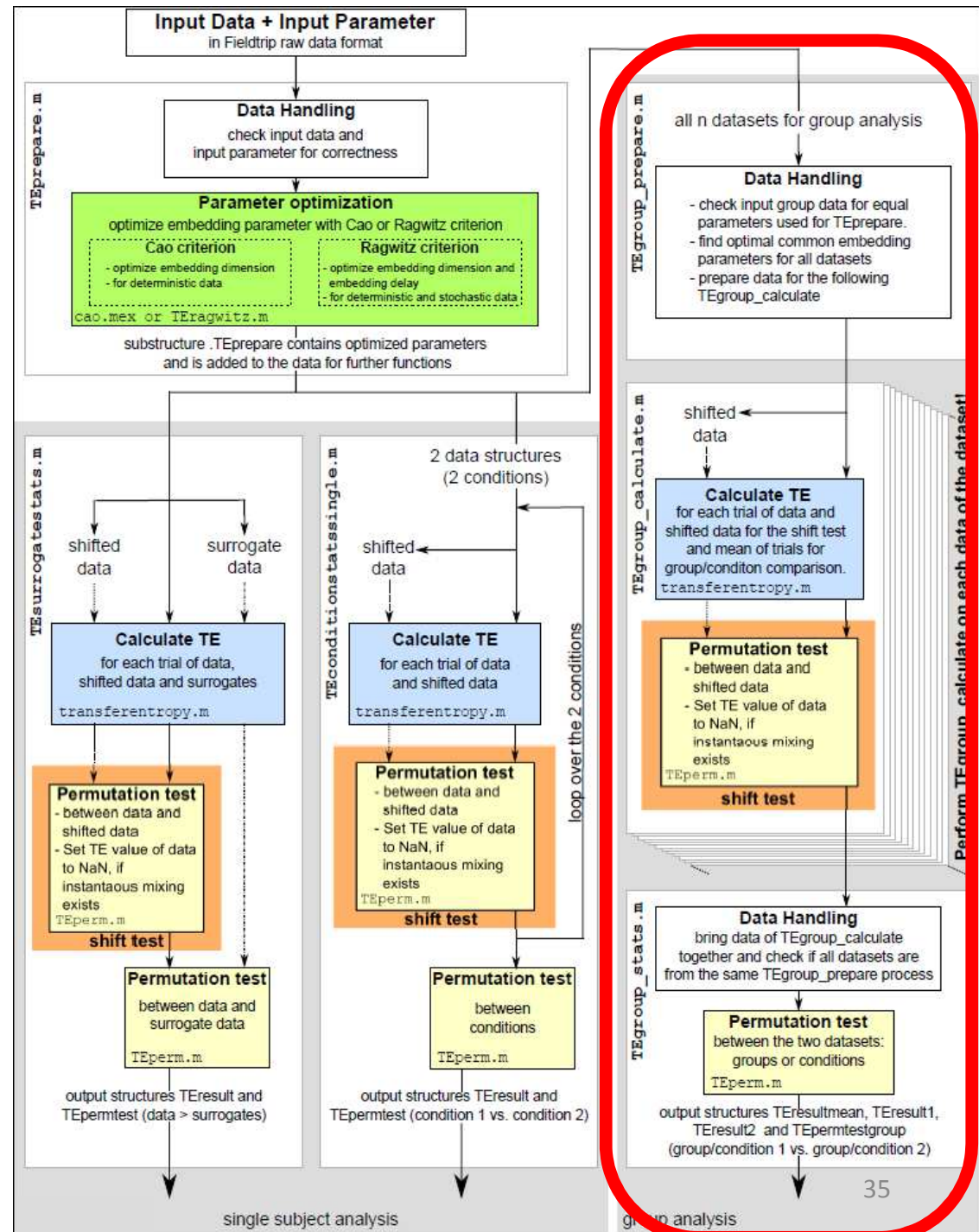
cfg.correctm = correction method used for correction of the multiple comparison problem - False discovery rate 'FDR' or Bonferroni correction 'BONF' (default = 'FDR')

cfg.fileidout = string for the first part of the output filename.

Example 2 - TEconditionstatssingle



Overview TRENTOOL



Overview TRENTOOL

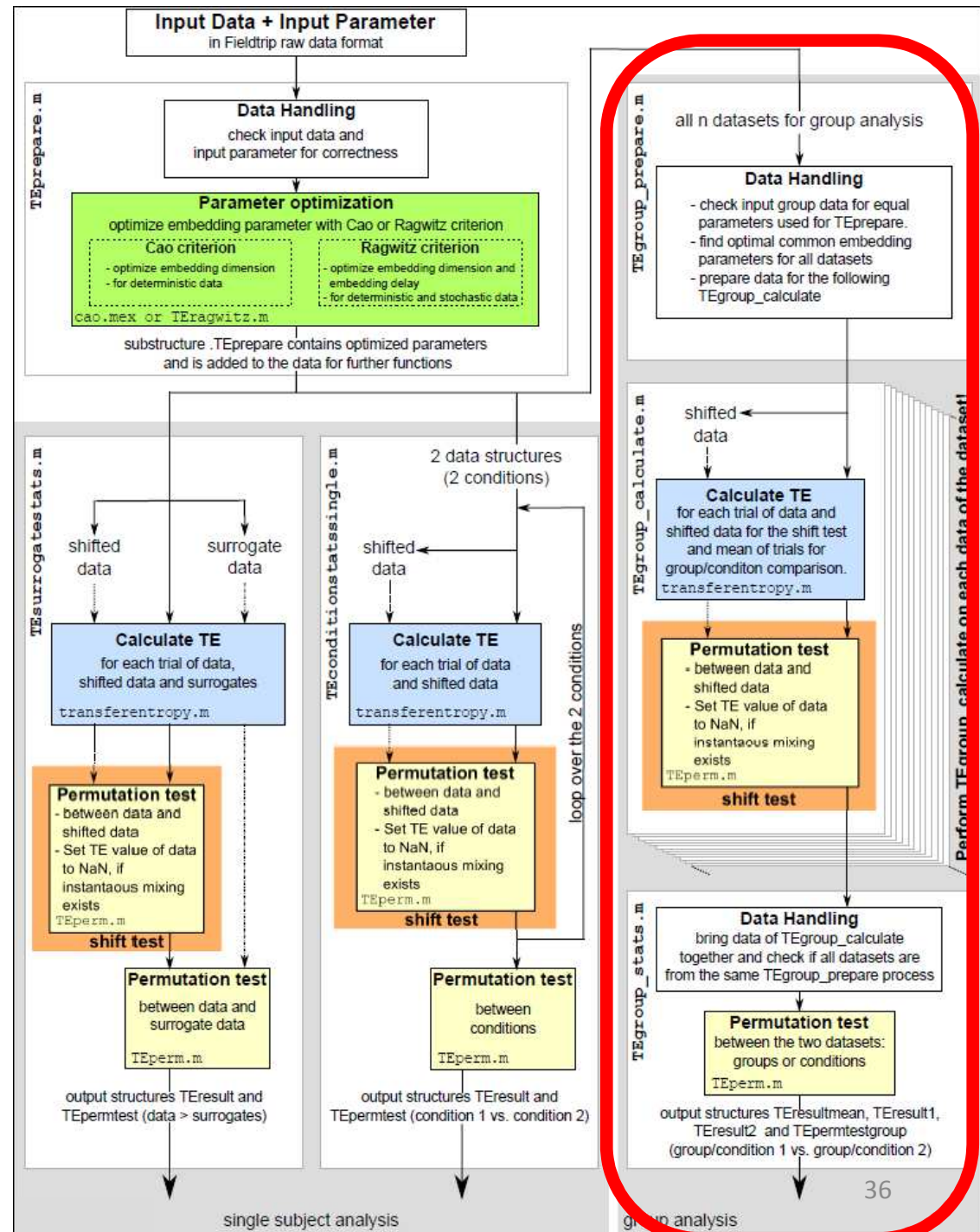
Important:

**Data_prepared =
TEprepare(cfg,Data)**

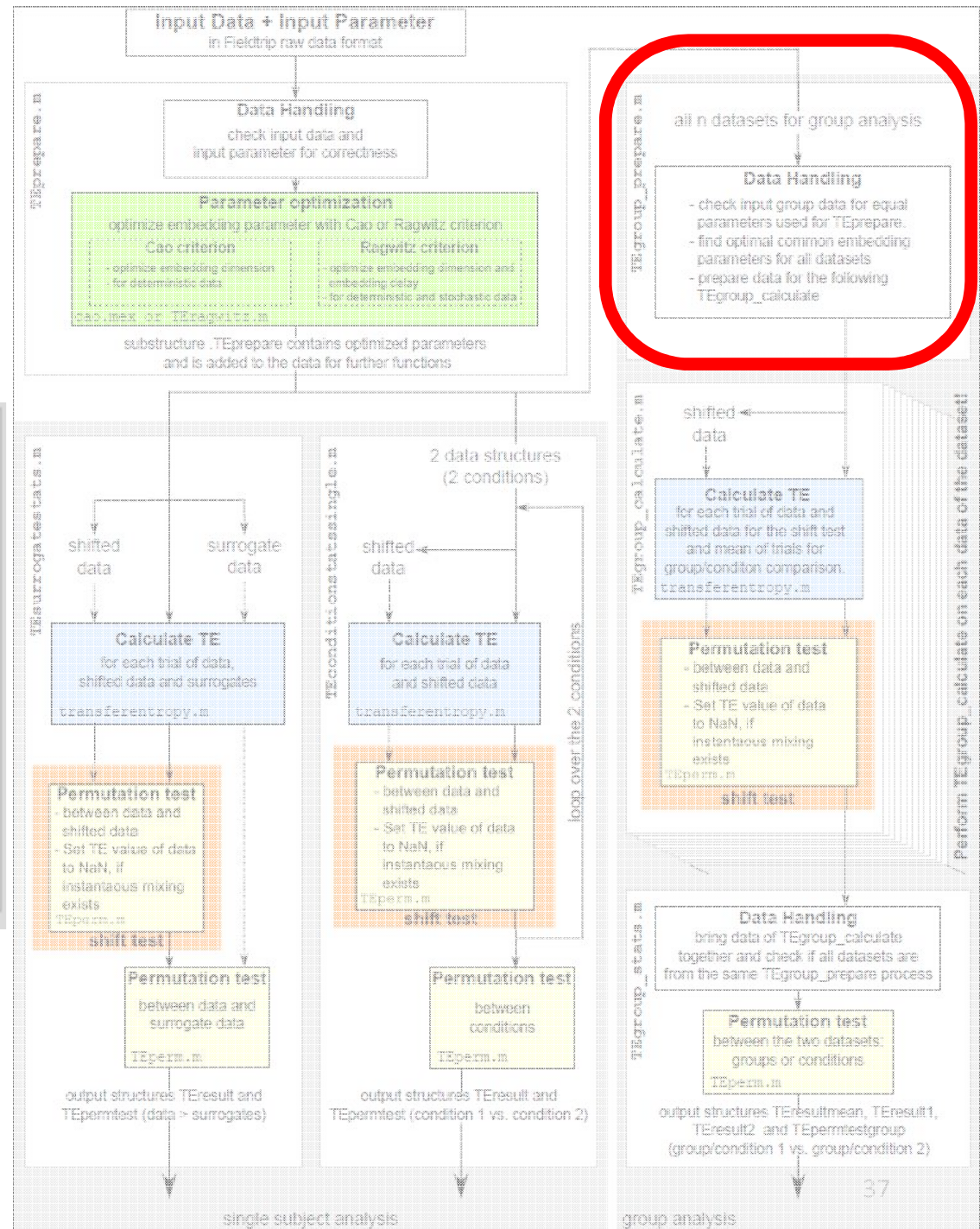
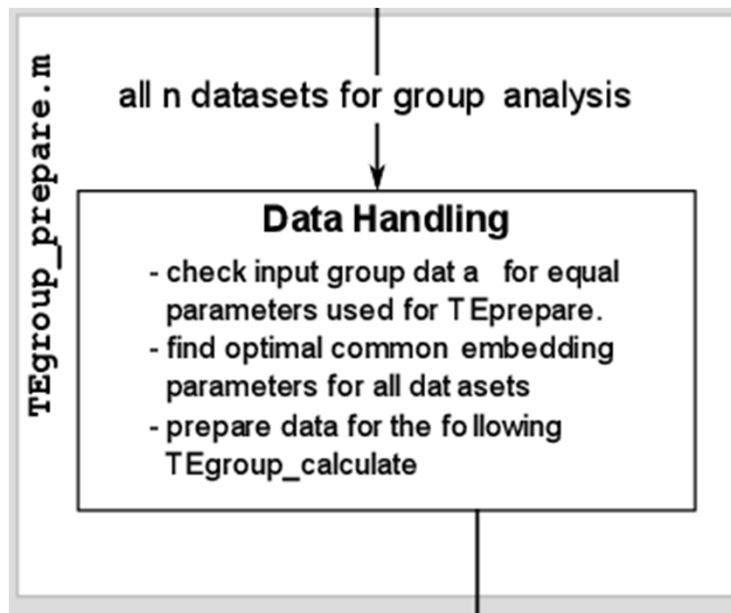
for EACH dataset with identical
parameters!!!

&

**SAVE EACH Prepared_Data
MANUALLY !!!**



Overview TRENTOOL



All parameters for TEgroup_prepare

cfg.dim = Number of embedding dimensions. If not specified, the maximum of the optimal dimensions found in TEprepare will be used, which is the recommended option!

cfg.tau = embedding delay in units of act ($x \cdot \text{act}$). If not specified (recommended option), the tau is used as followed:
In case of optimizemethod in TEprepare:
 'ragwitz' = optimal tau found via ragwitz
 critrion
 'cao' = cfg.tau given by user in TEprepare

cfg.shiftest = perform shift test to identify instantaneous mixing between the signal pairs. Values: 'yes' or 'no' (default = 'yes')
This shift test is important for EEG and MEG data, because linear mixing is always present in the data. In case of instantaneous mixing transfer entropy should not be calculated for the affected channelpairs with the corresponding parameter sets, because it could result in false positive results. Hence the TE values for these cases will be set to NaN and the corresponding p-values of the permutation test to 1. Alpha is set to 0.05.

cfg.shiftesttype = The shift test can be calculated for the direction TE value of original data > TE values of shifted data (value = 'TE>TEshift') or for the other direction (value = 'TEshift>TE'). In this case the alpha is set to 0.1. (default = 'TE>TEshift')

cfg.shifftype = Shifting the data 'onesample' or the length of the 'predicttime' (default = 'predicttime')

Run TEgroup_prepare over all datasets

minimum parameter structure:

```
cfg = [ ];
```

```
TEgroup_prepare(cfg,FilesCell)
```

FilesCell is a MATLAB Cell containing all filenames of the already prepared Data sets for the group analyses

TEgroup_prepare

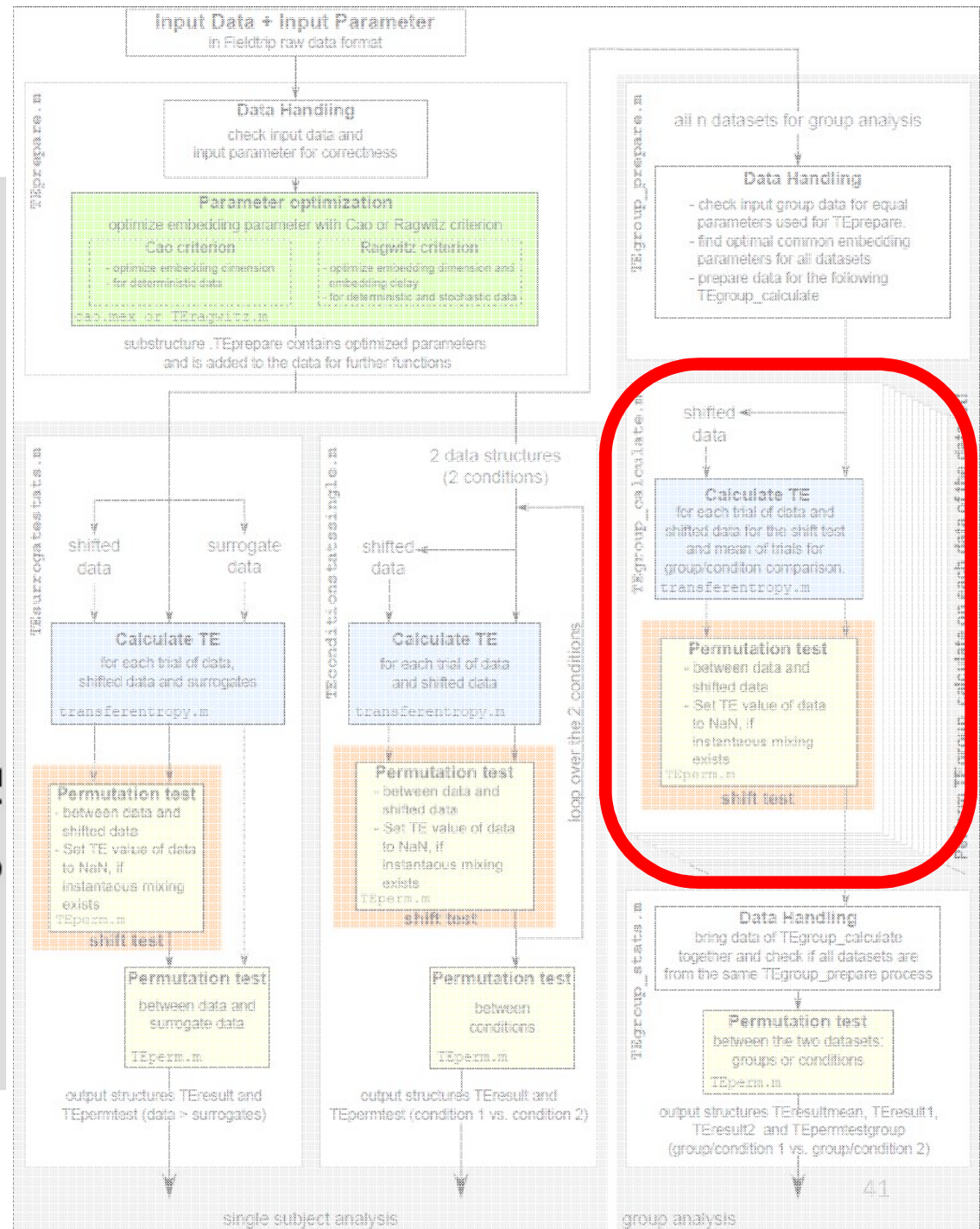
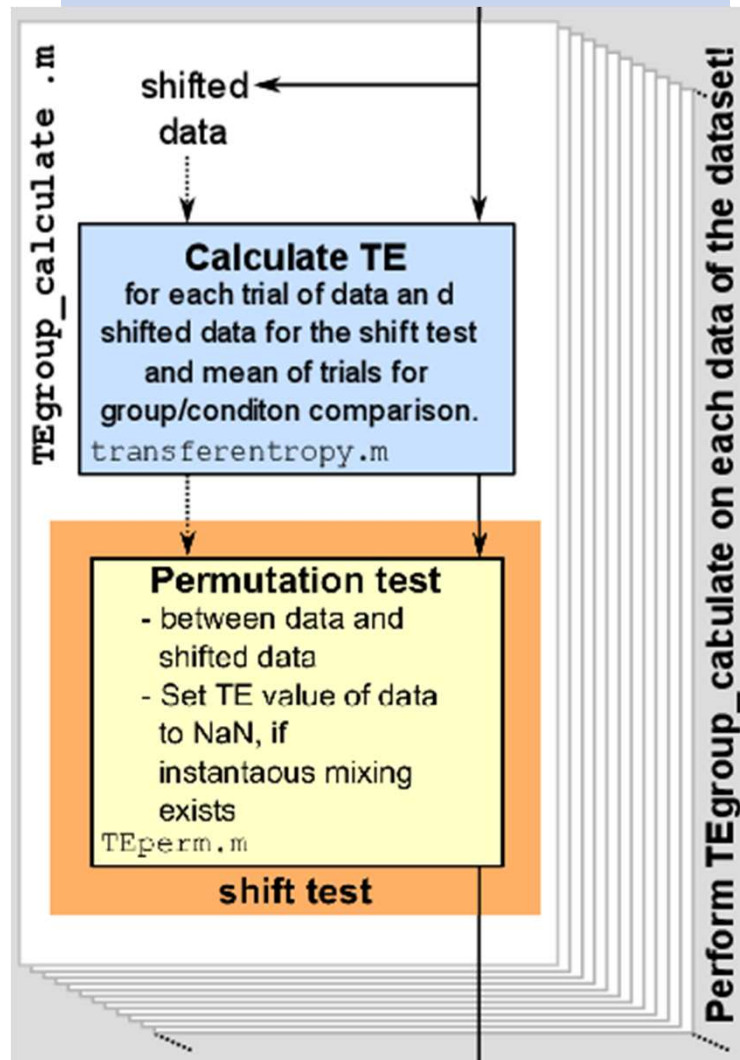
Output :

New version of each dataset:

filename_for_TEgroup_calculate.mat

containing an additional substructure .TEgroupprepare
that is needed for the further functions.

Overview



Run TEgroup_calculate

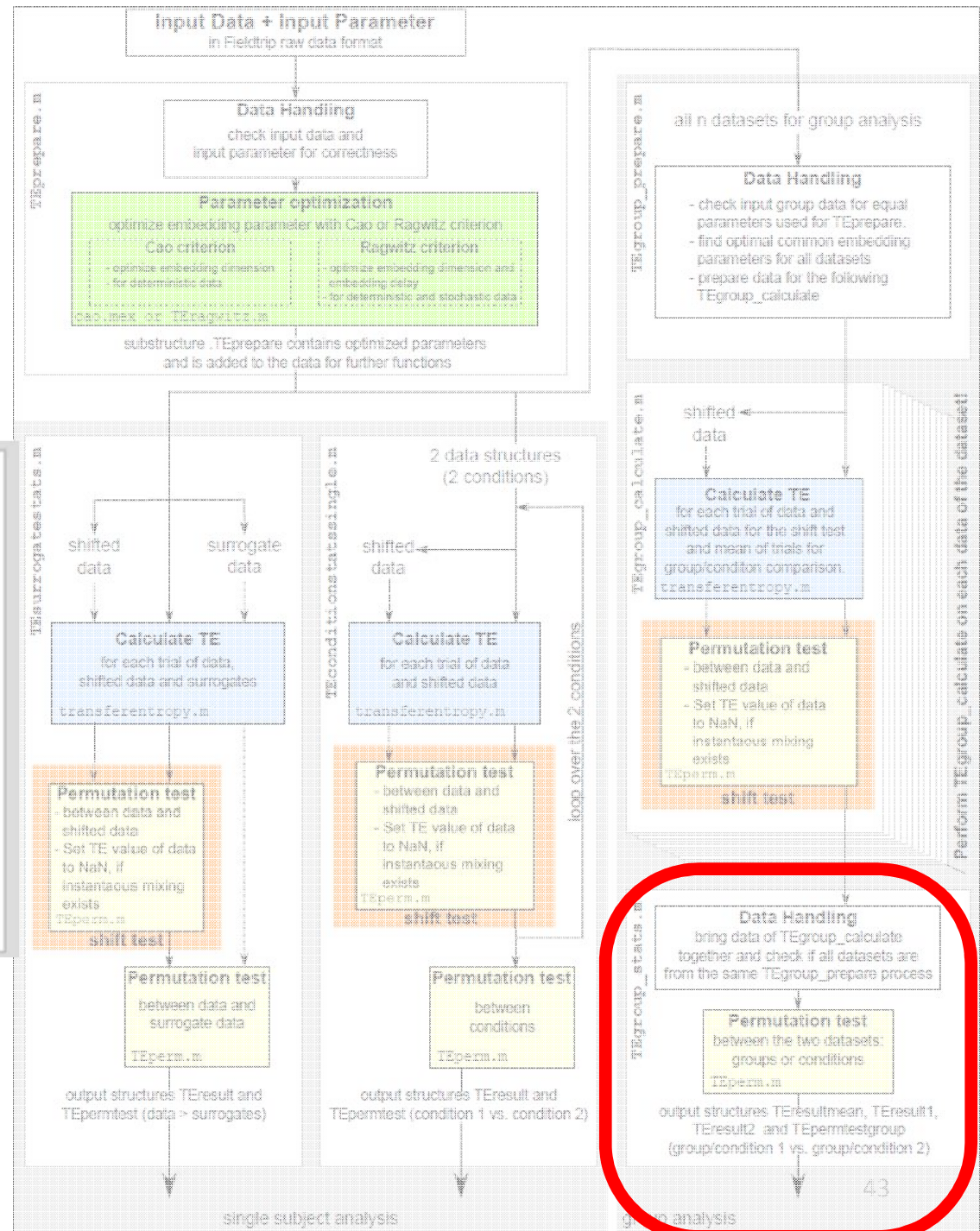
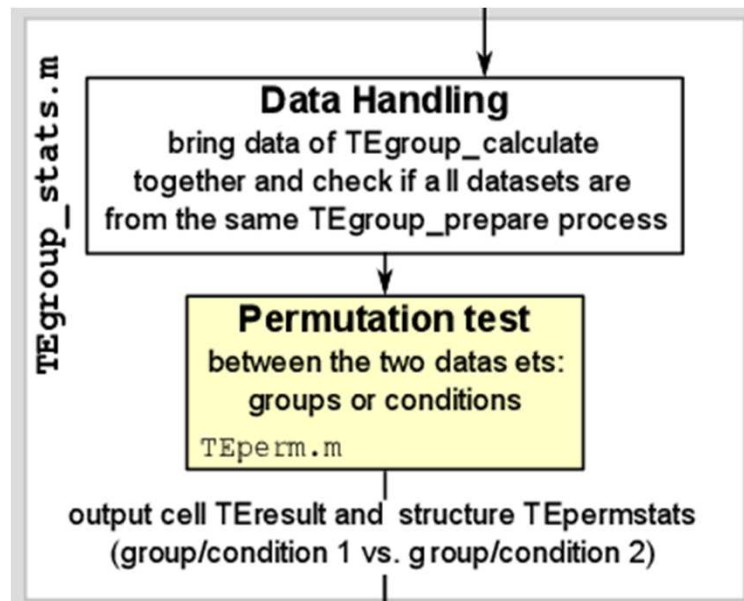
for all datasets separately
(possibility to pseudo-parallelize):

TEgroup_calculate(filename)

Output:

filename_TE_output.mat

Overview TRENTOOL



All parameters for TEgroup_stats

cfg.design = matrix containing a row with subject number and a row with independent variable representing the order of the data input.
example:
datasets: 1 2 3 4 5 1 2 3 4 5
conditions: 1 1 1 1 1 2 2 2 2 2

cfg.uval = row in cfg.design which contains the dataset number (in the example: 1)

cfg.ival = row in cfg.design which contains the independent variable (in the example: 2)

cfg.alpha = significance level for statistical shift test, permutation test and correction for multiple comparison (default = 0.05)

cfg.numpermutation = nr of permutations in permutation test (default = 190100)

cfg.permstatstype = 'mean' to use the distribution of the mean differences and 'depsamplesT' or 'indepsamplesT' for distribution of the t-values. (default = 'mean')

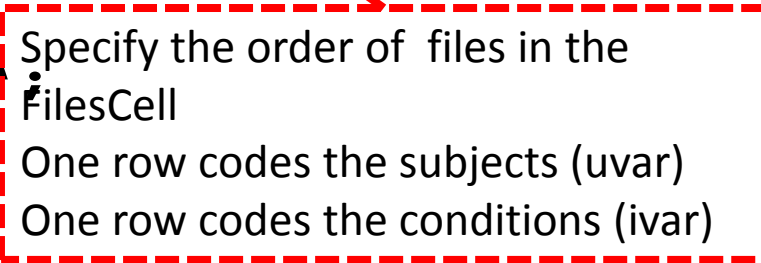
cfg.tail = '1' tail or '2' tailed test of significance (for the permutation tests) (default = 2)

cfg.correctm = correction method used for correction of the multiple comparison problem - False discovery rate 'FDR' or Bonferroni correction 'BONF' (default = 'FDR')

cfg.fileidout = string for the first part of the output filename.

Example parameter set for TEgroup_stats

```
cfg.design =  
[1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15;  
 1,1,1,1,1,1,1,1,1,1,1 ,1 ,1 ,1 ,1 ,1 ,2,2,2,2,2,2,2,2,2 ,2 ,2 ,2 ,2 ,2];  
cfg.uvar = 1;  
cfg.ivar = 2;  
cfg.permstatstype = 'depsamplest';  
cfg.tail = 2;  
cfg.fileidout = 'your_choice';
```



Specify the order of files in the
FilesCell
One row codes the subjects (uvar)
One row codes the conditions (ivar)

Run TEgroup_stats over all datasets

TEgroup_stats(cfg,FilesCell)

Output :

prefix_ ... _
TEpermtestgroup_output.mat
prefix_ ... _ TE_output.mat

Example 3 - Group analyses

Prepared_Data1 = **TEprepare**(cfg , Dataset1) ← **Dataset 1**

Prepared_Data2 = **TEprepare**(cfg , Dataset2) ← **Dataset 2**

...

...

...

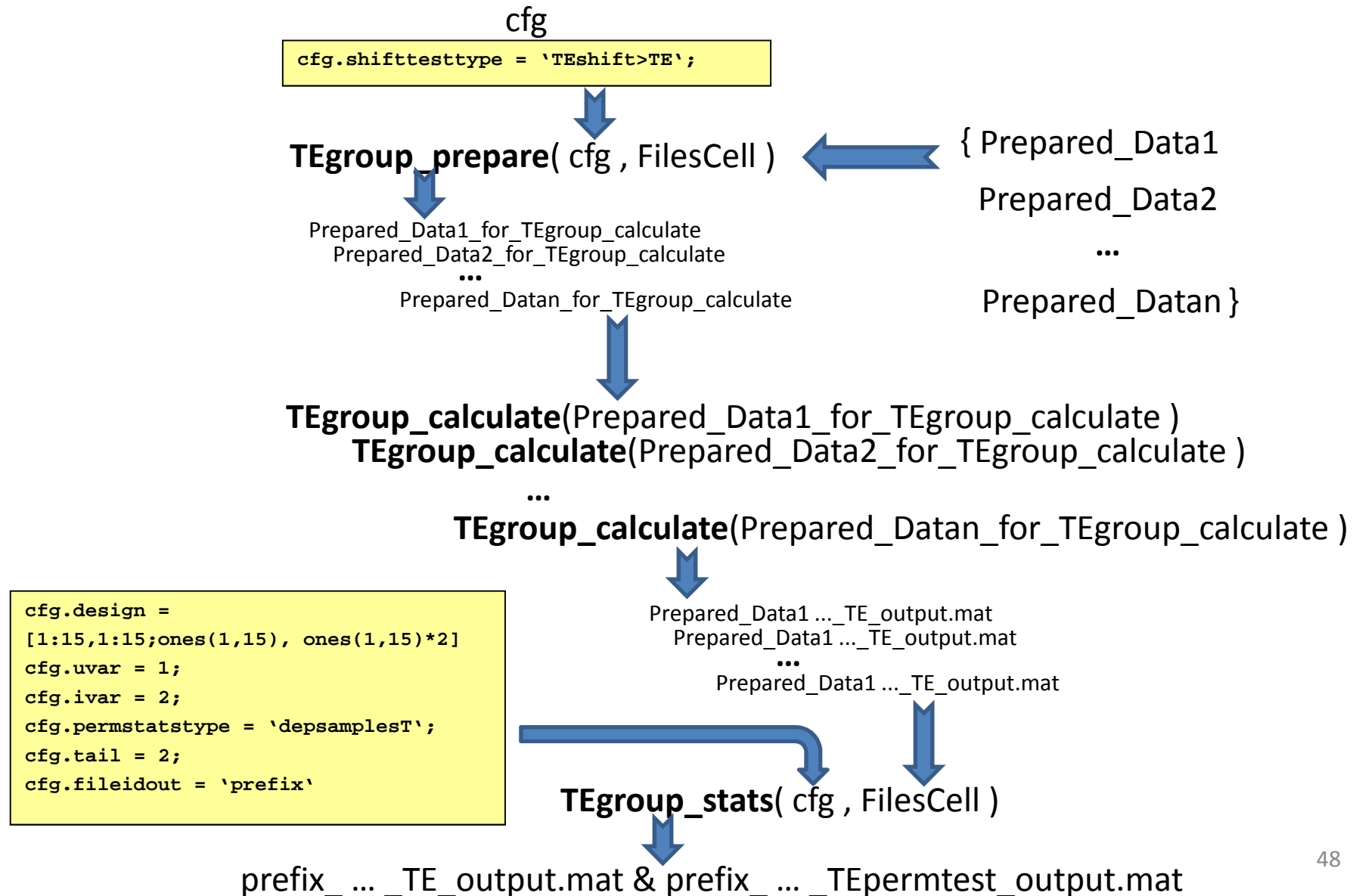
Prepared_Datan = **TEprepare**(cfg , Datasetn) ← **Dataset n**



cfg

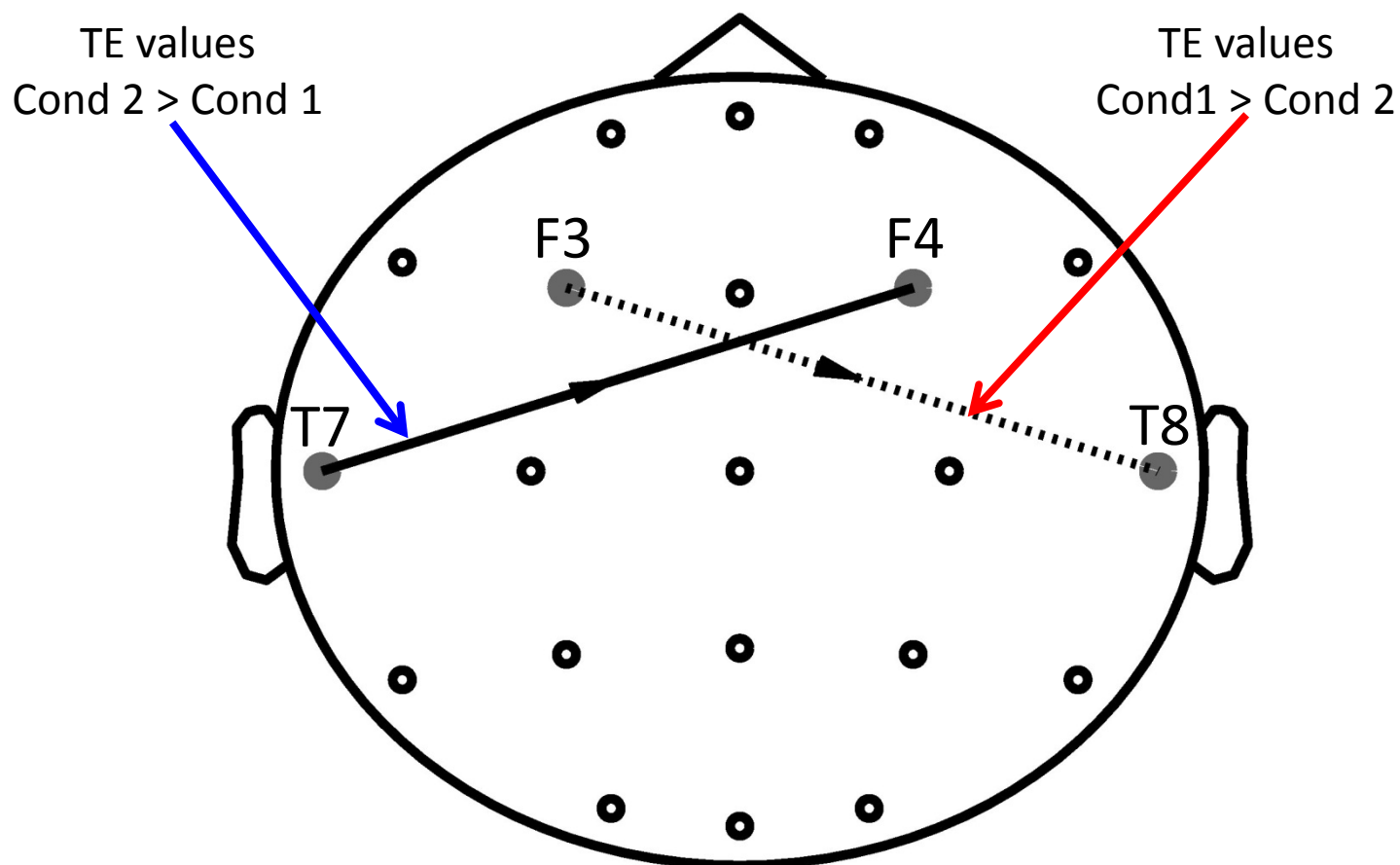
```
cfg.toi = [0 2.999];  
cfg.labels = {'X', 'Y'};  
cfg.optimizemethod = 'Cao';  
cfg.caodim = 1:6;  
cfg.caokth_neighbors = 3;  
cfg.trialselect = 'ACT';  
cfg.actthrvalue = 120;  
cfg.minnrtrials = 30;  
cfg.Path2TSTOOL =  
'/data/common/OpenTSTOOL';  
cfg.predicttime = 21;
```

Example 3 - Group analyses



TEplot2D

Example for the results of TEconditionstatssingle:



Additional functions

- time reconstruction (detecting optimal prediction time u)

TEInteractionDelayReconstruction_calculate

TEInteractionDelayReconstruction_analyze

TEInteractionDelayReconstruction_plotting

- Graph theoretical detection of cascade effects

Some additional features we are working on:

- TRENTOOL for fMRI
 - Nifti to TRENTOOL data converter
 - Spatial embedding optimization
 - Usage for normal block or event related designs
- 3D plotting function
- ...

Conclusion

- TRENTOOL
 - robustly detects the directed interactions.
 - is quite easy to use.
 - for neuroscientific applications: it works well on MEG data, single cell recording, intracranial recording, etc.
 - at the moment we are testing the usage of TRENTOOL on fMRI data

Requirements

- MATLAB (<http://www.mathworks.de>)
 - statistic toolbox
 - (Parallel processing toolbox – only to make it much faster)
- Fieldtrip (<http://fieldtrip.fcdonders.nl>)
- OpenTSTOOL (<http://www.physik3.gwdg.de/tstool/>)
- Aplenty of calculating capacity

Course program

- Day 1
 - TEprepare
 - Create cfg and script, run the function, inspect output
 - TESurrogatestats
 - Create cfg and script , run the function, inspect output
 - Run prepared script for time reconstruction and let it run over night
- Day 2
 - Go through the results
 - Open questions

Username = tuser
password = 3Nt00l_2

Course folders:

/data/home1/tuser/user01_vnc28
.../user02_vnc29
.../user03_vnc30
.../user04_vnc31
.../user05_vnc32
.../user06_vnc33
.../user07_vnc34
.../user08_vnc35
.../user09_vnc36
.../user10_vnc37
.../user11_vnc38

Each user folder contains:

- Toolboxes**
 - TRENTOOL2**
 - OpenTSTOOL**
 - Fieldtrip**
- Data**
 - data_save**
- Prepared scripts (for time reconstruction)**

Example script:

```
addpath( '~/user01_vnc28/toolboxes/TRENTOOL2' )
addpath( '~/user01_vnc28/toolboxes/fieldtrip' )
ft_defaults

load( '~/user01_vnc28/data/Lorenz_bidirectional_data.mat' )

% TEprepare
cfg = [ ];
cfg.Path2TSTOOL =
    '~/user01_vnc28/toolboxes/TRENTOOL2';
cfg.toi = [0.001 3];
cfg.channel = data.label;
cfg.predicttime_u = 46;
cfg.actthrvalue = 50;
cfg.minnrtrials = 12;
cfg.optimizemethod = 'ragwitz';
cfg.ragdim = 2:8;
cfg.ragtaurange = [0.5 1];
cfg.ragtausteps = 15;
cfg.repPred = 1000;
cfg.flagNei = 'Mass';
cfg.sizeNei = 4;

Prepared_Data = TEprepare(cfg,Data)
save( '~/user01_vnc28/data/data_save/Prepared_Data.mat', 'Prepared_Data' )

% TESurrogatestats
cfg = [ ];
cfg.optdimusage = 'indivdim';
cfg.tail = 1;
cfg.numpermutation = 50000;
cfg.shifttesttype = 'TEshift>TE';
cfg.surrogatetype = 'trialshuffling';
cfg.fileidout =
    '~/user01_vnc28/data/data_save/TEsur_output';

TESurrogatestats(cfg,Prepared_Data)
```


Thank you for your attention and your upcoming
directed interactions with the speaker!

Have fun with TRENTOOL!