

Manual

Computation Anatomy Toolbox - CAT12



QUICK START GUIDE	3
INTRODUCTION AND OVERVIEW	3
GETTING STARTED	5
Download and Installation	5
Starting the Toolbox	5
Basic VBM analysis (overview)	6
BASIC VBM ANALYSIS (DETAILED DESCRIPTION)	8
Second Module: Display one slice for all images	8
Third Module: Check sample homogeneity	8
Fourth Module: Smooth	10
Estimate Total Intracranial Volume (TIV)	10
Building the statistical model	10
Two-sample T-Test	12
Using the Full Factorial Model (for a 2x2 Anova)	13
Multiple Regression (Correlation)	14
Using the Full Factorial Model (for an Interaction)	15
Estimating the statistical model	16
Defining Contrasts	16
SPECIAL CASES.....	18
CAT12 for longitudinal data	18
Change Settings for Preprocessing	18
Preprocessing of longitudinal data	19
Statistical Analysis of longitudinal data in one group	20

Statistical Analysis of longitudinal data in two groups	21
Adapting the workflows	25
Customized Tissue Probability Maps	25
Customized DARTEL-template	26
OTHER VARIANTS OF COMPUTATIONAL MORPHOMETRY	28
Deformation-based morphometry (DBM)	28
Surface-based morphometry (SBM)	28
Region of interest (ROI) analysis	30
ADDITIONAL INFORMATION ON NATIVE, NORMALIZED AND MODULATED VOLUMES.....	30
NAMING CONVENTION OF OUTPUT FILES.....	32
TECHNICAL INFORMATION	33
REFERENCES	34

Quick start guide

VBM data

- Segment data using defaults (for longitudinal data use longitudinal pipeline)
- Check data quality using sample homogeneity for VBM data
- Smooth data (suggested starting value 8mm)
- Estimate total intracranial volume (TIV) in order to correct for different head size and volume
- Build 2nd-level model: Use "Full factorial" for cross-sectional data and "Flexible factorial" for longitudinal data and use TIV as covariate and select threshold masking with an absolute value of 0.1.
- Estimate model
- Optionally transform and threshold SPM-maps to (log-scaled) p-maps or correlation maps
- Optionally overlay selected slices

Additional surface data

- Segment data and additionally select "Surface and thickness estimation" in "Writing options"
- Check data quality using sample homogeneity for surface data
- Optionally extract additional surface parameters (e.g. sulcus depth, gyration index, cortical complexity)
- Resample and smooth surface data (suggested starting value 15mm)
- Build 2nd-level model model: Use "Full factorial" for cross-sectional data and "Flexible factorial" for longitudinal data
- Estimate surface model

Introduction and Overview

This manual is intended to help any user to perform a computational anatomy analysis using the CAT12 Toolbox. Although it will mainly focus on voxel-based morphometry (VBM) other variants of computational analysis such as deformation-based morphometry (DBM), surface-based morphometry (SBM), and region of interest (ROI) morphometric analysis will be also introduced and can be applied with a few changes.

Basically the manual may be divided into four main sections:

- Naturally, a quick guide of how to *get started* is given at the beginning. This section provides information how to *download and install* the software and *start* the Toolbox. Furthermore, a short *overview* on the steps of a VBM analysis is given.
- A *detailed description of a basic VBM analysis* is subsequently given, which will guide the user step by step through the whole process – from preprocessing to the selection of contrasts. This description should provide all necessary information to analyze most studies successfully.
- There are a few *special cases* of VBM analyses, for which the basic analysis workflow has to be adapted. These cases are *longitudinal studies* and studies in *children or special patient populations*. Relevant changes to a basic VBM analysis are described here and a description of how to apply these changes is provided. Importantly, only the changes are described – steps like for example quality control or smoothing are the same as in the basic analysis and not described a second time.
- The manual closes with *information on native, normalized and modulated volumes*, which determines how the results may be interpreted. Furthermore an overview of the naming conventions used as well as technical information is given.

Getting Started

DOWNLOAD AND INSTALLATION

- The CAT12 Toolbox runs within SPM12. That is, SPM12 needs to be installed and added to your Matlab search path before the CAT12 Toolbox can be installed (see <http://www.fil.ion.ucl.ac.uk/spm/> and <http://en.wikibooks.org/wiki/SPM>).
- Download (<http://dbm.neuro.uni-jena.de/cat12/>) and unzip the CAT12 Toolbox. You will get a folder named “cat12”, which contains various matlab files and compiled scripts. Copy the folder “cat12” into the SPM12 “toolbox” folder.

STARTING THE TOOLBOX

- Start Matlab
- Start SPM12 (i.e., type “spm fmri”)
- Select “cat12” from the SPM menu (see Figure 1). You will find the drop-down menu between the “Display” and the “Help” button (you can also call the Toolbox directly by typing “cat12” on the Matlab command line). This will open the CAT12 Toolbox as additional window (Fig. 2).

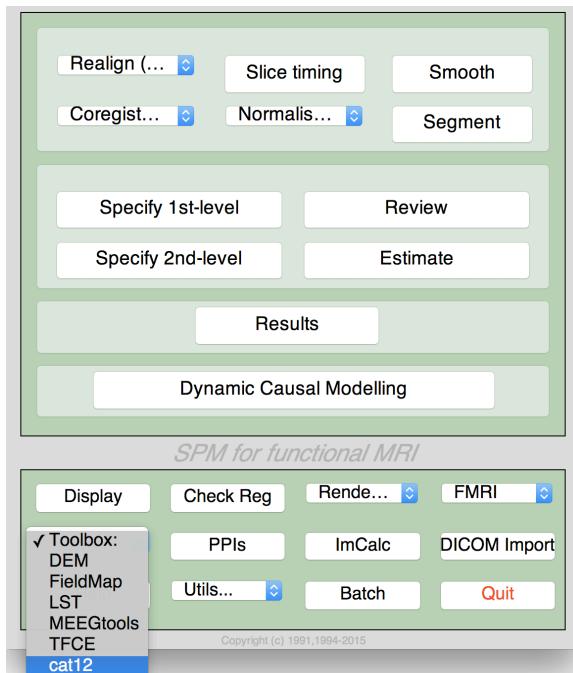


Figure 1: SPM menu

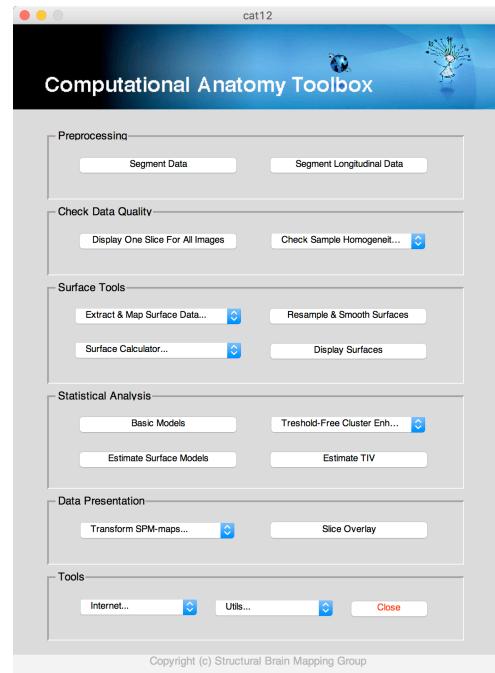


Figure 2: CAT12 Window

BASIC VBM ANALYSIS (OVERVIEW)

The CAT12 Toolbox comes with different modules, which may be used for an analysis. Usually, a VBM analysis comprises the following steps

(a) Preprocessing:

1. T1 images are **normalized** to a template space and **segmented** into gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF). The preprocessing parameters can be adjusted via the module “Estimate and write”.
2. After the preprocessing is finished, a **quality check** is highly recommended. This can be achieved via the modules “Display one slice for all images” and “Check sample homogeneity”. Both options are located in the CAT12 window under “Check Data Quality”.
3. Before entering the GM images into a statistical model, image data need to be **smoothed**. Of note, this step is not implemented into the CAT12 Toolbox but achieved via the standard SPM module “Smooth”.

(b) Statistical analysis:

4. The smoothed GM images are entered into a statistical analysis. This requires building a statistical model (e.g., T-Tests, ANOVAs, multiple regressions). This is done by the standard SPM modules “Specify 2nd Level” or “Basic Models” in the CAT12 window covering the same function.
5. The statistical model is estimated. This is done by the standard SPM module “Estimate”.
6. After estimating the statistical model, contrasts will be defined to get the results of the analysis. This is done by the standard SPM module “Results”.

The sequence of “preprocessing → quality check → smoothing → statistical analysis” remains the same for every VBM analysis, even when different steps are adapted (see “special cases”).

A few words about the Batch Editor...

- As soon as you select a module from the CAT12 Toolbox menu, a new window (the Batch Editor) will open. The Batch Editor is the environment where you will set up your analysis (see **Figure 3**). For example, an “<-X” indicates where you need to select files (e.g., your image files, the template, etc.). Other parameters have either default settings (which can be modified) or require input (e.g., choosing between different options, providing text or numeric values, etc.).
- Once all missing parameters are set, a green arrow will appear on the top of the window (the current snapshots in **Figure 3** show the arrow still in gray). Click this arrow to run the module or select “File → Run Batch”. It is very useful to save the settings before you run the batch (click on the disk symbol or select “File → Save Batch”).
- Of note, you can always find helpful information and parameter-specific explanations at the bottom of the Batch Editor window.¹
- All settings can be saved either as .mat file or as .m script file and reloaded for later use. The .m script file has the advantage to be editable with a text editor.

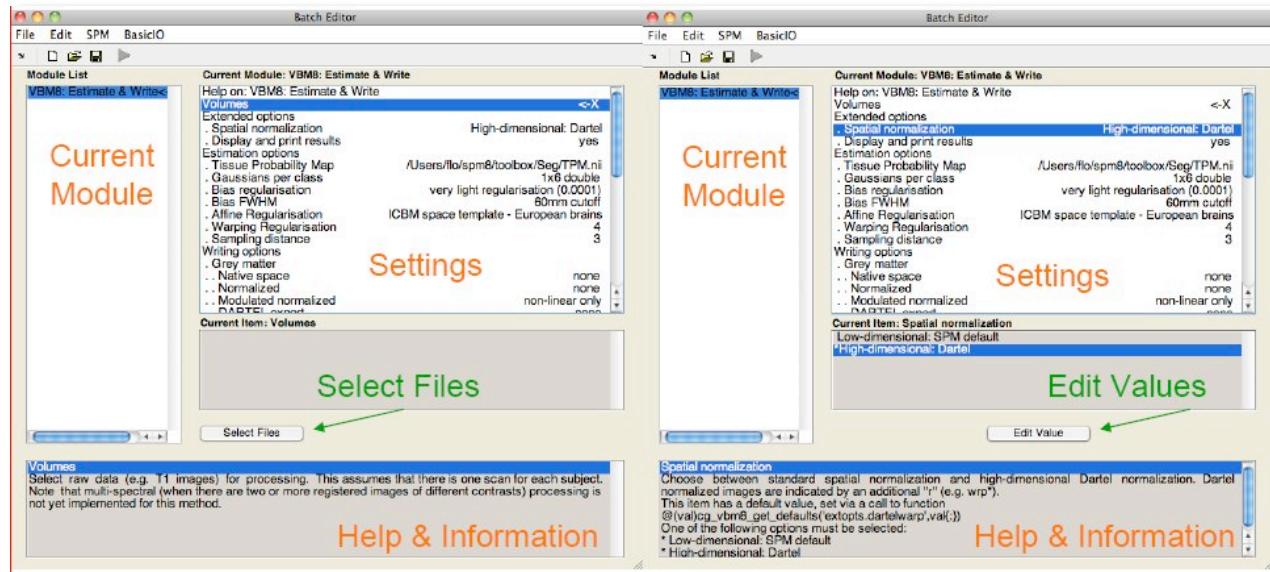


Figure 3: The Batch Editor is the environment where the analysis is set up. *Left:* For all settings marked with “<-X”, files have to be selected (“Select Files”). *Right:* Parameters can be edited and adapted (“Edit Value”).

¹ Additional CAT12-related information can be found by selecting “VBM Website” in the CAT12 window (Tools → Internet → VBM Website”). This will open a website. Here, look for “VBM subpages” on the right.

Basic VBM analysis (detailed description)

Text in preparation...

SECOND MODULE: DISPLAY ONE SLICE FOR ALL IMAGES

CAT12 → Check data quality → Display one slice for all images

Parameters:

- Sample data <-X → Select Files → *[select the new files]* → Done
 - Select the newly written data [e.g. the “wm*” files, which are the normalized bias corrected volumes]. This tool will display one horizontal slice for each subject, thus giving a good overview if the segmentation and normalization procedures yielded reasonable results. For example, if the native volume had artifacts or if the native volumes had a wrong orientation, the results may look odd. Solutions: Use “Check Reg” from the SPM main menu to make sure that the native images have the same orientation like the MNI Template (“SPM → templates → T1”). Adjust if necessary using “Display” from the SPM main menu.
- Proportional scaling → *[use defaults or modify]*
 - Check “yes”, if you display T1 volumes.
- Spatial orientation

Show slice in mm → *[use defaults or modify]*

- This module displays horizontal slices. This default setting provides a good overview.
 - File → Save Batch
 - File → Run Batch [the outcomes will be displayed in SPM’s graphic window]
-

THIRD MODULE: CHECK SAMPLE HOMOGENEITY

CAT12 → Check data quality → Check sample homogeneity → VBM data

Parameters:

- Data → New: Sample data <-X → Select Files → *[select gray matter volumes]* → Done
 - Select the newly written data [e.g. the “mwp1*” files, which are the modulated (m) normalized (w) GM segments (p1)]. This tool visualizes the correlation between

the volumes using a boxplot (or violin plot if the Statistics Toolbox was found) and correlation matrices. Thus, it will help identifying outliers. Any outlier should be carefully inspected for artifacts or pre-processing errors using “Check Reg” from the SPM main menu. If you specify different samples the mean correlation is displayed in separate plots for each sample.

- Load quality measures → New: XML files → *[optionally select xml-files with quality measures]*
 - Optionally select the xml-files that are saved for each data set. These files contain useful information about some estimated quality measures that can be also used for checking sample homogeneity. Please note, that the order of the xml-files must be the same as the other data files.
 - Separation in mm → *[use defaults or modify]*
 - To speed up calculations you can define that correlation is estimated only every x voxel. Smaller values give slightly more accurate correlation, but will be much slower.
 - Nuisance → *[enter nuisance variables if applicable]*
 - For each nuisance variable which you want to remove from the data prior to calculating the correlation, select “New: Nuisance” and enter a vector with the respective variable for each subject (e.g. age in years). All variables have to be entered in the same order as the respective volumes. You can also type “spm_load” to upload a *txt file with the covariates in the same order as the volumes. A potential nuisance parameter can be TIV if you check segmented data with the default modulation.
- File → Save Batch
- File → Run Batch
 - A window with a correlation matrix will open, which depict the correlation between the volumes. The correlation matrix shows the correlation between all volumes. High correlation values mean that your data are more similar to each other. If you click in the correlation matrix the corresponding data pairs will be displayed at the right bottom corner and allow a more careful inspection. The slider below the image changes the displayed slice. The popup menus at the right top corner provide more options. Here you can select other measures that are displayed in the boxplot (e.g. mean squared error or optionally quality measures if loaded), can change the order of the correlation matrix (by filename or mean correlation). Finally, the worst data can be shown in the SPM graphics window to check the data more carefully.
 - The boxplot (or violin plot) in the SPM graphics window averages all correlation values for each subject and shows the homogeneity of your sample. A small overall correlation in the boxplot not always means that this volume is an outlier or contains an artifact. If there are no artifacts in the image and if the image quality is reasonable you don't have to exclude this volume from the sample. This tool is intended to utilize the process of quality checking and there is no clear criteria defined to exclude a volume only based on the overall correlation value. However,

volumes with a noticeable lower overall correlation (e.g. below two standard deviations) are indicated and should be checked more carefully.

FOURTH MODULE: SMOOTH

SPM menu → Smooth

Parameters:

- Images to Smooth <-X → Select Files → [select grey matter volumes] → Done
 - Select the newly written data [e.g. the “m0wrp1” files, which are the normalized (wr) grey matter segments (p1) modulated for the non-linear components (m0)].
 - FWHM → [use defaults or modify]
 - 8-12mm kernels are widely used for VBM. To use this setting select “edit value” and type “8 8 8” (or “12 12 12”, respectively) for a kernel with 8mm (with 12mm) FWHM.
 - Data Type → [use defaults or modify]
 - Filename Prefix → [use defaults or modify]
-
- File → Save Batch [this will save the parameters as *.m script file]
 - File → Run Batch [the outcomes will be written to the same directory like the original data]

Estimate Total Intracranial Volume (TIV)

Text in preparation...

Building the statistical model

Although there are many potential designs offered in the 2nd-level analysis I recommend to use the “Full factorial” design because it covers most statistical designs. For cross-sectional VBM data you have usually 1..n samples and optionally covariates and nuisance parameters:

Number of factor levels	Number of covariates	Statistical Model
-------------------------	----------------------	-------------------

1	0	one-sample t-test
1	1	single regression
1	>1	multiple regression
2	0	two-sample t-test
>2	0	Anova
>1	>0	Ancova (for nuisance parameters) or Interaction (for covariates)

TWO-SAMPLE T-TEST

SPM menu → Specify 2nd-level

Parameters:

- Directory <-X → Select Files → [select the working directory for your analysis] → Done
- Design → “Two-sample t-test”
 - Group 1 scans → Select Files → [select the smoothed grey matter volumes for group 1; following this script these will be the “sm0wp1” files] → Done
 - Group 2 scans → Select Files → [select the smoothed grey matter volumes for group 2] → Done
 - Independence → Yes
 - Variance → Equal or Unequal
 - Grand mean scaling → No
 - ANCOVA → No
- **Covariates ***
 - Masking
 - Threshold Masking → Absolute → [specify value (e.g., “0.1”)]
 - Implicit Mask → Yes
 - Explicit Mask → <None>
 - Global Calculation → Omit
 - Global Normalization
 - Overall grand mean scaling → No
 - Normalization → None
- File → Save Batch [this will save the parameters as *.m script file]
- File → Run Batch [this will create an “SPM.mat” file in your working directory]

*You could specify one or many covariates (i.e., partial out the variance of specific factors when looking at group differences).

It is strongly recommended to always use total intracranial volume (TIV) as covariate if you use modulated data in VBM in order to correct for different brain sizes.

- Covariates → New Covariate
- Vector <-X → enter the values of the covariates (e.g., TIV and optionally age in years) in the same order as the respective file names or type “spm_load” to upload a *.txt file with the covariates in the same order as the volumes
- Name <-X → Specify Text (e.g., “age”)
- Interactions → None
- Centering → No centering

USING THE FULL FACTORIAL MODEL (FOR A 2X2 ANOVA)

SPM menu → Specify 2nd-level

Parameters:

- Directory <-X → Select Files → [select the working directory for your analysis] → Done
- Design → “Full Factorial”
 - Factors → “New: Factor; New: Factor”
 - Factor
 - Name → [specify text (e.g., “sex”)]
 - Levels → 2
 - Independence → Yes
 - Variance → Equal or Unequal
 - Grand mean scaling → No
 - ANCOVA → No
 - Factor
 - Name → [specify text (e.g., “handedness”)]
 - Levels → 2
 - Independence → Yes
 - Variance → Equal or Unequal
 - Grand mean scaling → No
 - ANCOVA → No
 - Specify Cells → “New: Cell; New: Cell; New: Cell; New: Cell”
 - Cell
 - Levels → [specify text (e.g., “1 1”)]
 - Scans → [select files (e.g., the smoothed GM volumes of the left-handed males)]
 - Cell
 - Levels → [specify text (e.g., “1 2”)]
 - Scans → [select files (e.g., the smoothed GM volumes of the right-handed males)]
 - Cell
 - Levels → [specify text (e.g., “2 1”)]
 - Scans → [select files (e.g., the smoothed GM volumes of the left-handed females)]
 - Cell
 - Levels → [specify text (e.g., “2 2”)]
 - Scans → [select files e.g., the smoothed GM volumes of the right-handed females)]
- Covariates*
- Masking
 - Threshold Masking → Absolute → [specify value (e.g., “0.1”)]

- Implicit Mask → Yes
 - Explicit Mask → <None>
 - Global Calculation → Omit
 - Global Normalization
 - Overall grand mean scaling → No
 - Normalization → None
- File → Save Batch [this will save the parameters as *.m script file]
- File → Run Batch [this will create an “SPM.mat” file in your working directory]
-

MULTIPLE REGRESSION (CORRELATION)

SPM menu → Specify 2nd-level

Parameters:

- Directory <-X → Select Files → [*select the directory for your analysis*] → Done
 - Design → “Multiple Regression”
 - Scans → [*select files (e.g., the smoothed GM volumes of all subjects)*] → Done
 - Covariates → “New: Covariate”
 - Covariate
 - Vector → [*enter the values in the same order as the respective file names of the smoothed GM images*]
 - Name → [*specify test (e.g., “age”)*]
 - Centering → No centering
 - Intercept → Include Intercept
 - Covariates*
 - Masking
 - Threshold Masking → Absolute → [*specify value (e.g., “0.1”)*]
 - Implicit Mask → Yes
 - Explicit Mask → <None>
 - Global Calculation → Omit
 - Global Normalization
 - Overall grand mean scaling → No
 - Normalization → None
- File → Save Batch [this will save the parameters as *.m script file]
- File → Run Batch [This will create an “SPM.mat” file in your selected folder]

USING THE FULL FACTORIAL MODEL (FOR AN INTERACTION)

SPM menu → Specify 2nd-level

Parameters:

- Directory <-X → Select Files → [select the working directory for your analysis] → Done
- Design → “Full Factorial”
 - Factors → “New: Factor”
 - Factor
 - Name → [specify text (e.g., “sex”)]
 - Levels → 2
 - Independence → Yes
 - Variance → Equal or Unequal
 - Grand mean scaling → No
 - ANCOVA → No
 - Specify Cells → “New: Cell; New: Cell”
 - Cell
 - Levels → [specify text (e.g., “1”)]
 - Scans → [select files (e.g., the smoothed GM volumes of the males)]
 - Cell
 - Levels → [specify text (e.g., “2”)]
 - Scans → [select files (e.g., the smoothed GM volumes of the females)]
 - Covariates → “New: Covariate”
 - Covariate
 - Vector → [enter the values in the same order as the respective file names of the smoothed GM images]
 - Name → [specify test (e.g., “age”)]
 - Interactions → With Factor 1
 - Centering → No centering
 - Masking
 - Threshold Masking → Absolute → [specify value (e.g., “0.1”)]
 - Implicit Mask → Yes
 - Explicit Mask → <None>
 - Global Calculation → Omit
 - Global Normalization
 - Overall grand mean scaling → No
 - Normalization → None

- File → Save Batch [this will save the parameters as *.m script file]

- File → Run Batch [this will create an “SPM.mat” file in your working directory]

ESTIMATING THE STATISTICAL MODEL

SPM menu → Estimate

Parameters:

- Select SPM.mat <-X → Select Files → [select the SPM.mat which you just built] → Done
 - Method → “Classical”
- File → Save Batch
- File → Run Batch [This will create an “SPM.mat” file in your selected folder]

DEFINING CONTRASTS

SPM menu → Results → [select the SPM.mat file] → Done (this opens the Contrast Manager) → Define new contrast (i.e., choose “t-contrast” or “F-contrast”; type the contrast name and specify the contrast by typing the respective numbers, as shown below):

T-contrasts:

a. Simple group difference

⇒ Use SPM.mat from model **“2 sample T-test”**

- For Group A > Group B: specify “1 -1”
- For Group A < Group B: specify “-1 1”

b. 2x2 ANOVA

⇒ Use SPM.mat from model **“2x2 ANOVA”**

- For left-handed males > right-handed males: specify “1 -1 0 0”
- For left-handed females > right-handed females: specify “0 0 1 -1”
- For left-handed males > left-handed females: specify “1 0 -1 0”
- For right-handed males > right-handed females: specify “0 1 0 -1”
etc.
- For males > females: specify “1 1 -1 -1”
- For left-handers > right-handers: specify “1 -1 1 -1”

c. Multiple Regression (Correlation)

⇒ Use SPM.mat from model “**CORRELATION**”

- For positive correlation: specify “1”
- For negative correlation: specify “-1”

In case that the first column in the design matrix is a constant (sample effect) you have to prepend a “0” to all contrasts (e.g. “0 1”).

d. Interaction

⇒ Use SPM.mat from model “**INTERACTION**”

- For regression slope Group A > Group B: specify “0 0 1 -1”
- For regression slope Group A < Group B: specify “0 0 -1 1”

→ Done

F-contrasts:

If you would like to use the old SPM2 F-contrast “Effects of interest” the respective contrast vector is:

$$\text{eye}(n)-1/n$$

where n is the number of columns of interest. This F-contrast is often helpful for plotting parameter estimates of effects of interest.

Getting Results:

SPM menu → Results → [select a contrast from Contrast Manager] → Done

- Mask with other contrasts → No
- Title for comparison: [use the pre-defined name from the Contrast Manager or change it]
- P value adjustment to:
 - None (uncorrected for multiple comparisons), set threshold to 0.001
 - FDR (false discovery rate), set threshold to 0.05, etc.
 - FWE (family-wise error), set threshold to 0.05, etc.
- Extent threshold: [either use “none” or specify the number of voxels²)

² In order to *empirically* determine the extent threshold (rather than saying 100 voxels or 500 voxels, which is completely arbitrary), simply run this first without specifying an extent threshold. This will give you an output (i.e., the standard SPM glass brain with significant effects). When you click “Table” (SPM main menu) you will get a table with all relevant values (MNI coordinates, p-values, cluster size etc). Below the table you will find additional information, such as “Expected Number of Voxels per Cluster”. Remember this number (this is your empirically determined extent threshold). Re-run SPM → Results etc. and specify this number when asked for the “Extent Threshold”. There is also a hidden option in “CAT12 → Data presentation → Threshold and transform spmT-maps” to define the extent threshold in terms of a p-value or to use the “Expected Number of Voxels per Cluster”.

Special Cases

CAT12 for longitudinal data

BACKGROUND

The majority of VBM studies are based on cross-sectional data, where one image is acquired for each subject. However, in order to track e.g. learning effects over time longitudinal designs are necessary, where additional time-points are acquired for each subject. The analysis of these longitudinal data requires a customized processing, that considers the characteristics of intra-subject analysis. While for cross-sectional data images can be processed independently for each subject longitudinal data has to be registered to the baseline image (or mean image) for each subject. Furthermore, spatial normalization is estimated for the baseline image only and applied to all images (Figure 4). Additional attention is then needed for the setup of the statistical model. The following section will therefore describe data preprocessing and model setup for longitudinal data.

Fig 4.: Flow diagram for processing longitudinal data with CAT12. This figure demonstrates the steps for processing longitudinal data. After an initial realignment, the mean of the realigned images is calculated (mean) and used as reference image in a subsequent realignment. The realigned images (r_{1x}) are then corrected for signal inhomogeneities with regard to the reference mean image. Spatial normalization parameters are estimated in the next step using the segmentations of the mean image. These normalization parameters are applied to the segmentations of the bias-corrected images ($p1mri_x$). The resulting normalized segmentations ($wp1mri_x$) are finally again realigned.

Preprocessing of longitudinal data - overview

The CAT12 Toolbox supplies a batch for longitudinal study design. Here, for each subject the respective images need to be selected. Intra-subject realignment, bias correction, segmentation, and normalization are calculated automatically. Preprocessed images are written as $wp1mr^*$ and $wp2mr^*$ for grey and white matter respectively. To define the segmentation and normalization parameters, the defaults in `cg_defaults.m` are used.

CHANGE SETTINGS FOR PREPROCESSING

Change your working directory to “/toolbox/CAT12” in your SPM directory:

→ select “Utilities → cd” in the SPM menu and change to the respective folder.

Then type “open `cg_CAT12_defaults.m`” in your matlab command window. The file will open in the editor. If you are unsure how to change the values, open the module “estimate and write” in the batch editor for reference.

The most interesting parameters to change here may be:

- DARTEL or SPM12 default normalization:
Line 65 → set to “0” for SPM12 default normalization
- The Tissue Probability Maps:
Line 16 → replace “[fullfile(spm('dir'), 'toolbox', 'Seg', 'TPM.nii')]” with the path to your Maps.

It is always a great idea to memorize or record the changes and change all values back to default after the Analysis.

PREPROCESSING OF LONGITUDINAL DATA

CAT12 → Process longitudinal data

Parameters:

- Data <-X → New: Subject → Subject → Longitudinal data for one subject → Select Files → [select raw data] → Done
 - Select all volumes for each subject. As the Toolbox does not support multispectral data yet (i.e., different imaging methods for the same brain, such as T1-, T2-, diffusion-weighted or CT images), it is recommended to choose a T1-weighted image.
 - Select “New: Subject” to add data for a new subject
The data for each subject should be listed as one “subject” in the Batch Editor, i.e. there are as many subjects listed as included in the analysis.
- File → Save Batch [this will save the parameters as *.m script file]
- File → Run Batch [this will create an “SPM.mat” file in your working directory]

For the naming conventions of all written files see p.29 “Naming convention of output files”. The final GM segments are wp1mr*, the final WM segments are named wp2mr*.

Statistical analysis of longitudinal data - overview

The main interest in longitudinal studies is the common change of tissue volume over time in a group of subjects or the difference in these changes between two or more groups. The setup of the statistical model needed to assess these questions will be described on two examples. First, the case of only one group of 4 subjects with 2 time points each (e.g. normal aging) is presented. Subsequently, the case of two groups of subjects with 4 time points per subject will be described. These examples should cover most analyses – the number of time points / groups just have to be adapted.

STATISTICAL ANALYSIS OF LONGITUDINAL DATA IN ONE GROUP

SPM menu → Specify 2nd-level

Parameters:

- Directory <-X → Select Files → [select the working directory for your analysis] → Done
- Design → “Flexible Factorial”
 - Factors → “New: Factor; New: Factor”
 - Factor
 - Name → [specify text (e.g., “subject”)]
 - Independence → Yes
 - Variance → Equal or Unequal
 - Grand mean scaling → No
 - ANCOVA → No
 - Factor
 - Name → [specify text (e.g., “time”)]
 - Independence → No
 - Variance → Equal or Unequal
 - Grand mean scaling → No
 - ANCOVA → No
 - Specify Subjects or all Scans & Factors → “Subjects” → “New: Subject; New: Subject; New: Subject; New: Subject;”
 - Subject
 - Scans → [select files (the smoothed GM volumes of the first Subject)]
 - Conditions → “1 2” [for two time points]
 - Subject
 - Scans → [select files (the smoothed GM volumes of the second Subject)]
 - Conditions → “1 2” [for two time points]
 - Subject
 - Scans → [select files (the smoothed GM volumes of the third Subject)]
 - Conditions → “1 2” [for two time points]
 - Subject
 - Scans → [select files (the smoothed GM volumes of the fourth Subject)]
 - Conditions → “1 2” [for two time points]
 - Main effects & Interaction → “New: Main effect”
 - Main effect
 - Factor number → 2
 - Main effect
 - Factor number → 1
- Covariates (see “basic VBM analysis (detailed description)”)
- Masking
 - Threshold Masking → Absolute → [specify value (e.g., “0.1”)]

- Implicit Mask → Yes
 - Explicit Mask → <None>
 - Global Calculation → Omit
 - Global Normalization
 - Overall grand mean scaling → No
 - Normalization → None
 - File → Save Batch [this will save the parameters as *.m script file]
 - File → Run Batch [this will create an “SPM.mat” file in your working directory]
-

STATISTICAL ANALYSIS OF LONGITUDINAL DATA IN TWO GROUPS

SPM menu → Specify 2nd-level

Parameters:

- Directory <-X → Select Files → [*select the working directory for your analysis*] → Done
- Design → “Flexible Factorial”
 - Factors → “*New: Factor; New: Factor; New: Factor*”
 - Factor
 - Name → [*specify text (e.g., “subject”)*]
 - Independence → Yes
 - Variance → Equal
 - Grand mean scaling → No
 - ANCOVA → No
 - Factor
 - Name → [*specify text (e.g., “group”)*]
 - Independence → Yes
 - Variance → Unequal
 - Grand mean scaling → No
 - ANCOVA → No
 - Factor
 - Name → [*specify text (e.g., “time”)*]
 - Independence → No
 - Variance → Equal
 - Grand mean scaling → No
 - ANCOVA → No
- Specify Subjects or all Scans & Factors → “*Subjects*” → “*New: Subject; New: Subject; New: Subject; New: Subject; New: Subject; New: Subject*”

- Scans → [select files (the smoothed GM volumes of the 1st Subject of first group)]
- Conditions → “[1 1 1 1; 1 2 3 4]” [first group with four time points]
Do not forget the additional single quote!

Subject

- Scans → [select files (the smoothed GM volumes of the 2nd Subject of first group)]
- Conditions → “[1 1 1 1; 1 2 3 4]” [first group with four time points]

Subject

- Scans → [select files (the smoothed GM volumes of the 3rd Subject of first group)]
- Conditions → “[1 1 1 1; 1 2 3 4]” [first group with four time points]

Subject

- Scans → [select files (the smoothed GM volumes of the 4th Subject of second group)]
- Conditions → “[2 2 2 2; 1 2 3 4]” [second group with four time points]

Subject

- Scans → [select files (the smoothed GM volumes of the 1st Subject of second group)]
- Conditions → “[2 2 2 2; 1 2 3 4]” [second group with four time points]

Subject

- Scans → [select files (the smoothed GM volumes of the 2nd Subject of second group)]
- Conditions → “[2 2 2 2; 1 2 3 4]” [second group with four time points]

Subject

- Scans → [select files (the smoothed GM volumes of the 3rd Subject of second group)]
- Conditions → “[2 2 2 2; 1 2 3 4]” [second group with four time points]

Subject

- Scans → [select files (the smoothed GM volumes of the 4th Subject of second group)]
- Conditions → “[2 2 2 2; 1 2 3 4]” [second group with four time points]

Subject

- Scans → [select files (the smoothed GM volumes of the 1st Subject of second group)]
- Conditions → “[2 2 2 2; 1 2 3 4]” [second group with four time points]

Subject

- Scans → [select files (the smoothed GM volumes of the 2nd Subject of second group)]
- Conditions → “[2 2 2 2; 1 2 3 4]” [second group with four time points]

Subject

- Factor numbers → 2 3 [Interaction between group and time]

Main effect

- Factor number → 1

- Covariates (see “basic VBM analysis (detailed description)”)

- Masking

- Threshold Masking → Absolute → [specify value (e.g., “0.1”)]
- Implicit Mask → Yes
- Explicit Mask → <None>

- Global Calculation → Omit

- Global Normalization
 - Overall grand mean scaling → No
 - Normalization → None
- File → Save Batch [this will save the parameters as *.m script file]
 - File → Run Batch [this will create an “SPM.mat” file in your working directory]

Altered Workflows for VBM-analyses

Background

For most analyses the VBM Toolbox will supply all tools needed. That is, as the new segmentation algorithm is not dependent on Tissue Probability Maps (TPMs) anymore and as predefined DARTEL-templates for healthy adults exist, most questions can be assessed using the toolbox settings. However, for some special cases such as analyses in children or special patient populations, the toolbox settings might not be optimal. For these cases the CAT12 Toolbox provides an integration into the SPM12 environment that can be used to optimize the preprocessing. In the following, we will present strategies how to deal with these special cases.

Standard VBM preprocessing: Input, Output and where to modify

The first module of the CAT12 Toolbox (“Estimate and Write”) processes all preprocessing steps except for the smoothing. Basically, it takes structural volumes and TPMs as input. It will then segment the data, apply a registration to MNI Space (either rigid or affine) and subsequently a non-linear deformation. The non-linear deformation parameters can be calculated via the low dimensional SPM default approach or the high dimensional DARTEL algorithm and the predefined templates. Figure 5 depicts this preprocessing workflow and highlights possibilities where to modify.

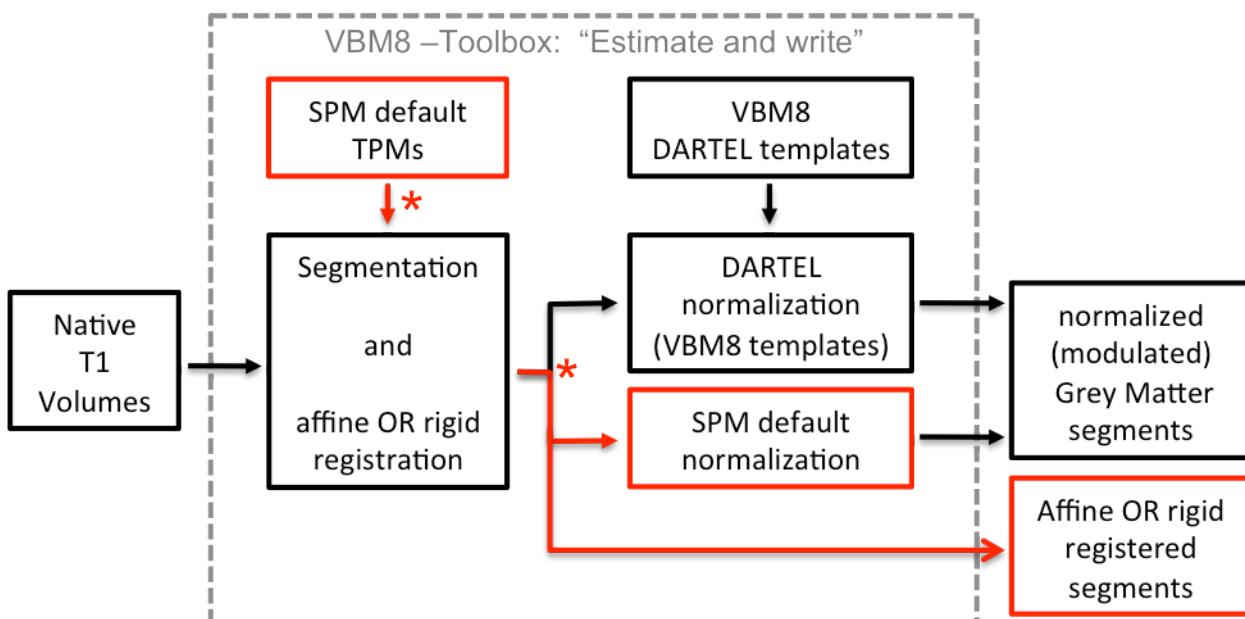


Fig. 5: Flow-chart of the preprocessing steps within the module “Estimate and Write”. Marked in red are those steps, where the preprocessing can be customized. Per default, the built-in DARTEL normalization works with the CAT12 DARTEL templates of 550 healthy adult control subjects. Affine registered tissue segments can be used to create customized DARTEL-templates, which can then be used to replace the default DARTEL template.

Adapting the workflows

Customized tissue probability maps - overview

For data on children it will be a good idea to create customized TPMs, which reflect age and gender of the population. The TOM8 Toolbox (available via: <https://irc.cchmc.org/software/tom.php>) provides the means to customize these TPMs. To learn more about the TOM toolbox, see also <http://dbm.neuro.uni-jena.de/software/tom/>.

CUSTOMIZED TISSUE PROBABILITY MAPS

About the TOM8 Toolbox:

→ select Module “create new template”

 → select “TOM.mat” (you will have to download this file together with the toolbox)

 → write priors/template as single file

 → for all others use default settings or modify. For “Age” either a vector or a mean age (when using the average approach) must be specified.

Implementation into CAT12:

Module: “Estimate and write”

 → “Estimation Options”.

 → Tissue Probability Maps (→ Select your customized TPMs here)

Customized DARTEL-template - overview

For all cases that include at least 50-100 subjects a customized DARTEL template can be created. That is, grey matter and white matter tissue segments of all subjects are used to create a mean template of the study sample. As the CAT12 toolbox writes all files needed to create these templates (“DARTEL export”), this requires only two additional steps. In order to use these newly created DARTEL-Templates with the CAT12 Toolbox, an affine registration of the DARTEL export should be used. From these affine registered segments customized DARTEL templates can then be created and used with the CAT12 module “Write already estimated segmentations”.

CUSTOMIZED DARTEL-TEMPLATE

Three steps are needed to create normalized tissue segments with customized DARTEL Templates. These steps can be batched using dependencies in the Batch Editor. The last step can be rerun using the customized templates, if additional output files are needed. In the first step the T1 images are segmented, and the tissue segments normalized to the Tissue Probability Maps using an affine transformation. Start with selecting the module “Estimate and Write”.

Module: “Estimate and Write”:

- for all options except “writing options” use settings like for a “standard” VBM analysis.
- “Writing Options”
 - “Grey Matter” → “DARTEL export” → “affine”
 - “White Matter” → “DARTEL export” → “affine”

These settings will produce the volumes “rp1*-affine.nii” and “rp2*-affine.nii”, which are the grey (rp1) and white (rp2) matter segments after affine registration. The following modules can be chosen directly in the batch editor (SPM → Tools → DARTEL Tools → Run DARTEL (create Templates), and SPM → Tools → CAT12 → CAT12: Write already estimated segmentations). It makes sense to add and specify these modules together with the “Estimate and Write” module within the Batch Editor and to set dependencies.

Module “Run DARTEL (create Templates)”

- Images → select two times “new: Images”
 - Images: → select the “rp1*-affine.nii” files or create a dependency.
 - Images: → select the “rp2*-affine.nii” files or create a dependency.
- all other options: use defaults or modify

The thus created customized DARTEL Templates are now used in the CAT12 Toolbox:

Module “Write already estimated segmentations”

Parameters:

- Volumes <-X → Select the original T1 images like in the first module “Estimate and Write”.

- Extended Options → “High-dimensional:DARTEL” (default) → Select DARTEL Template
 - if you use dependencies in the Batch Editor create a dependency on the first Template in the batch editor. Templates 2-6 will be used automatically.
 - If you have already created your DARTEL Templates, chose the fist Template (named “Template_1” per default). Attention: If you want to rename your template, you may either add a prefix to the default name, or make sure that the template number is embedded in the name between two underscores (e.g. “abc_1_xyz”).
 - For all other options use the same settings as in the first module, or modify.
- Writing Options → select the output files just like in any standard VBM analysis:
 - For **GM, WM, and CSF image volumes** see page 14: “Additional Information on native, normalized and modulated normalized volumes”.
 - A **bias corrected image volume**, in which MRI inhomogeneities and noise are removed, can be written in normalized or native space.
 - A **partial volume effect (PVE) label image volume** can also be written in normalized or native space or as a DARTEL export file.
 - The **Jacobian determinant** for each voxel can be written in normalized space.
 - Finally, **deformation fields** can be written. This option is useful to re-apply normalization parameters to other images or particular regions of interest.

- File → Save Batch
- File → Run Batch

How to proceed:

All steps described above are just an adaption of the CAT12 Toolbox module “Estimate and Write”. A subsequent smoothing is necessary before statistics can be calculated. As always it is a good idea to save the applied modules and to perform quality control. Here, the modules “Display one slice for all images” and “Check sample homogeneity” from the CAT12 Toolbox will be helpful.

Other variants of computational morphometry

Deformation-based morphometry (DBM)

DBM is based on the application of non-linear registration procedures to spatially normalise one brain to another one. The simplest case of spatial normalisation is to correct the orientation and size of the brains. In addition to these global changes, a non-linear normalisation is necessary to minimise the remaining regional differences by means of local deformations. If this local adaptation is possible, the deformations now reveal information about the type and localization of the structural differences between the brains and can undergo subsequent analysis.

Differences between both images are minimized and are now coded in the deformations. Finally, a map of local volume changes can be quantified by a mathematical property of these deformations – the Jacobian determinant. This parameter is well known from continuum mechanics and is usually used for the analysis of volume changes in flowing liquids or gases. The Jacobian determinant allows a direct estimation of the percentage change in volume in each voxel and can be statistically analyzed (Gaser et al. 2001). This approach is also known as tensor-based morphometry because the Jacobian determinant represents such a tensor.

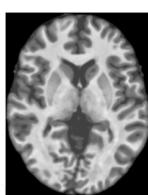
A deformation-based analysis can be carried out not only on the local changes in volume but also on the entire information of the deformations, which also includes the direction and strength of the local deformations (Gaser et al. 1999). Since each voxel contains three-dimensional information, a multivariate statistical test is necessary for analysis. A multivariate general linear model or Hotelling's T2 test is commonly used for this type of analysis (Gaser et al. 1999; Thompson et al. 1997).

Text in preparation...

Surface-based morphometry (SBM)

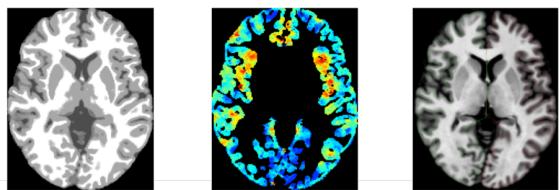
Surface-based morphometry has several advantages over using volumetric data alone. For instance, brain surface meshes have been shown to increase the accuracy of brain registration compared with Talairach registration (Desai et al. 2005). Brain surface meshes also permit new forms of analyses, such as gyration indices that measure surface complexity in 3D (Yotter et al. 2011b) or cortical thickness. Furthermore, inflation or spherical mapping of the cortical surface mesh raises the buried sulci to the surface so that mapped functional activity in these regions can be easily visualized.

Workflow for SBM:



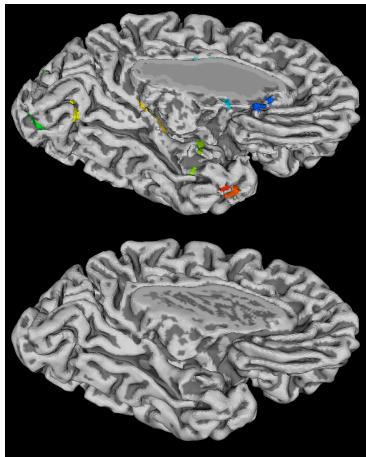
Local adaptive segmentation

Gray matter regions with high iron concentration, like the motor cortex and the occipital regions, often have increased intensities that lead to misclassifications. In addition to our adaptive MAP approach for partial volume segmentation we use an approach that allows adaptation of local intensity changes in order to deal with varying tissue contrast (Dahnke et al. 2012a).



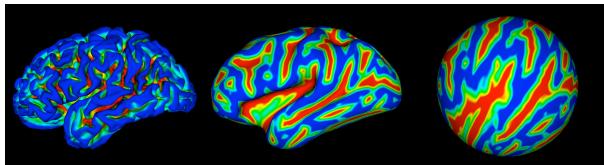
Cortical thickness and central surface estimation

We use a fully automated method that allows for measurement of cortical thickness and reconstructions of the central surface in one step. It uses a tissue segmentation to estimate the white matter (WM) distance, then projects the local maxima (which is equal to the cortical thickness) to other gray matter voxels by using a neighbor relationship described by the WM distance. This projection-based thickness allows the handling of partial volume information, sulcal blurring, and sulcal asymmetries without explicit sulcus reconstruction (Dahnke et al. 2012b).



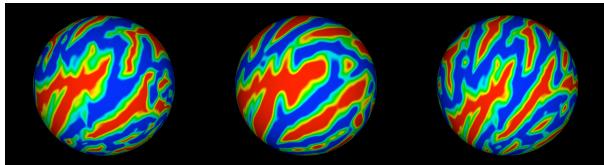
Topological correction

In order to repair topological defects we use a novel method that relies on spherical harmonics (Yotter et al. 2011a). First, the original MRI intensity values are used as a basis to select either a “fill” or “cut” operation for each topological defect. We modify the spherical map of the uncorrected brain surface mesh, such that certain triangles are favored while searching for the bounding triangle during reparameterization. Then, a low-pass filtered alternative reconstruction based on spherical harmonics is patched into the reconstructed surface in areas that previously contained defects.



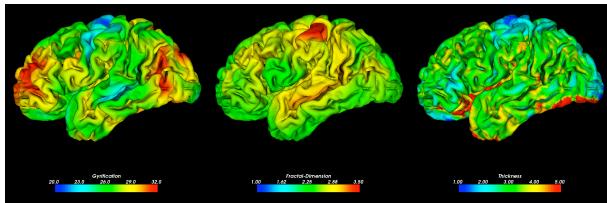
Spherical mapping

A spherical map of a cortical surface is usually necessary to reparameterize the surface mesh into a common coordinate system to allow intersubject analysis. We use a fast algorithm to reduce area distortion resulting in an improved reparameterization of the cortical surface mesh (Yotter et al. 2011c).



Spherical registration

We have adapted the volume-based diffeomorphic Dartel algorithm to the surface (Ashburner, 2007) to work with spherical maps (Yotter et al. 2011d). We apply a multi-grid approach that uses reparameterized values of sulcal depth and shape index defined on the sphere to estimate a flow field that allows deforming a spherical grid.



Statistical analysis

We use the core statistical functions of SPM12 to provide the same functionality for surface-based analysis. This allows analysis of parameters such as cortical thickness, gyrification index, fractal dimension (Yotter et al., 2011b) or even fMRI measurements that are projected onto the cortical surface.

Text in preparation...

Region of interest (ROI) analysis

Text in preparation...

Additional information on native, normalized and modulated volumes

When preprocessing the images (see “First Module: Estimate and write”, on pages 5-6), the decision about the normalization parameters will determine the interpretation of the analysis outcomes:

“**Native space**” produces tissue class images in spatial correspondence to the original data. Although this could be useful for estimating global tissue volumes (e.g., GM+WM+CSF=TBV) it is not suitable to conduct VBM analyses due to the missing voxel-wise correspondence across brains). Of note, if one is interested in these global tissue volumes in native space (“raw values”), it won’t be necessary to actually output the tissue class images in native space. The “Estimate and write”-function automatically generates a text file for each subject (*_seg8.txt), which contains the raw values for GM, WM, and CSF. The subject-specific values can be combined (i.e., integrated into a single text file) by using the function “Calculate raw volumes for GM/WM/CSF”: CAT12 → Tools → Calculate raw volumes for GM/WM/CSF

“**Normalized**” produces tissue class images in spatial correspondence to the template. This is useful for VBM analyses (e.g., “concentration” of gray matter; Good et al. 2001; Neuroimage).

“**Modulated normalized**” gives two different options:

- **Affine+non-linear** produces tissue class images in alignment with the template, but multiplies (“modulates”) the voxel values by the Jacobian determinant (i.e., linear and non-linear components) derived from the spatial normalization. This is useful for VBM analyses and allows comparing the *absolute amount of tissue* (e.g., “volume” of gray matter; Good et al. 2001; Neuroimage).
- **Non-linear only** produces tissue class images in alignment with the template, but multiplies the voxel values by the non-linear components only. This is useful for VBM analyses and allows comparing the *absolute amount of tissue corrected for individual brain sizes*. Of note, this option is similar to using “Affine+non-linear” (see above) in combination with “global normalization” (when later building the statistical model using the traditional PET designs). That is, when building the statistics, one would specify “Global normalization” → “Overall grand mean scaling

– No” → “Normalization – Proportional”. It is also similar to using “Affine+non-linear” (see above) and including the numeric brain volume (as given in the *_seg8.txt file for each subject) as covariate (when later building the statistical model). Although all 3 approaches allow comparing tissue volumes while correcting for individual brain size, it is recommended to use the option “non-linear only” as it applies the correction directly to the data, rather than to the statistical model.

For further explanation see also:

<http://www.neuro.uni-jena.de/vbm/segmentation/modulation>

Naming convention of output files

<i>segmented images:</i>	<i>bias corrected images:</i>
m[0]wp[0123]* m - modulated m0 - modulated non-linear only w - (dartel) warped p - segmented 0 - PVE label 1 - GM 2 - WM 3 - CSF _affine - affine registered only	wm[*] m - bias corrected w - (dartel) warped

For longitudinal data (default settings):

<i>segmented images:</i>	<i>bias corrected images:</i>	<i>mean images (useful for overlaying results):</i>
wp[12]mr* w - warped p - segmented 1 - GM 2 - WM m - bias corrected r - realigned	mr* m - bias corrected r - realigned <i>estimated raw volumes:</i> pmrxxx_seg.txt	wmrmean* w - warped m - bias corrected r - realigned mean - mean <i>estimated raw volumes (mean):</i> pmeanxxx_seg.txt

Technical information

This toolbox is an extension of the “New Segment Toolbox” in SPM12, but uses a completely different segmentation approach.³

1. The segmentation approach is based on an adaptive **Maximum A Posterior (MAP) technique** without the need for *a priori* information about tissue probabilities. That is, the Tissue Probability Maps are not used constantly in the sense of the classical unified segmentation approach, but just for spatial normalization³. The following MAP estimation is adaptive in the sense that local variations of the parameters (i.e., means and variance) are modelled as slowly varying spatial functions (Rajapakse et al. 1997). This not only accounts for intensity inhomogeneities but also for other local variations of intensity.
2. Additionally, the segmentation approach uses a **Partial Volume Estimation (PVE)** with a simplified mixed model of at most two tissue types (Tohka et al. 2004). We start with an initial segmentation into three pure classes: gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) based on the above described MAP estimation. The initial segmentation is followed by a PVE of two additional mixed classes: GM-WM and GM-CSF. This results in an estimation of the amount (or fraction) of each pure tissue type present in every voxel (as single voxels – given by their size – probably contain more than one tissue type) and thus provides a more accurate segmentation.
3. Furthermore, we apply **two denoising methods**. The first method is a spatially adaptive non-local means (SANLM) denoising filter (Manjon et al. 2010). This filter will remove noise while preserving edges and is implemented as preprocessing step. The second method is a classical Markov Random Field (MRF) approach, which incorporates spatial prior information of adjacent voxels into the segmentation estimation (Rajapakse et al. 1997).
4. Another important extension to the SPM12 segmentation is the **integration of the Dartel normalisation** (Ashburner 2007) into the toolbox. If high-dimensional spatial normalisation is chosen, an already existing Dartel template in MNI space will be used. This template was derived from 550 healthy control subjects of the IXI-database (<http://www.brain-development.org>) and is provided in MNI space⁴ for six different iteration steps of Dartel normalisation. Thus, for the majority of studies the creation of sample-specific Dartel templates is not necessary anymore⁵.

³ The classic SPM12 segmentation is still used in addition, but only to initially remove non-brain tissue from the image and to get a starting estimate for the segmentation.

⁴ Thus, no additional MNI normalization is necessary.

⁵ For studies investigating data of children I still recommend creating a customized Dartel template. Of note, for this option a representative sample with a sufficient number of subjects is required (n>50-100). Alternatively, if a sufficient sample size cannot be achieved, the low-dimensional SPM12 normalization approach combined with customized Tissue Probability Maps (e.g. from the TOM8 toolbox) can be selected.

References

- Good, C.D. et al. (2001): A voxel-based morphometric study of ageing in 465 normal adult human brains. *Neuroimage*. **14(1 Pt 1)**:21-36.
- Rajapakse, J.C. et al. (1997): Statistical Approach to Segmentation of Single-Channel Cerebral MR Images. *IEEE Trans. Med. Imag.* **16(2)**:176-186.
- Tohka, J. et al. (2004): Fast and robust parameter estimation for statistical partial volume models in brain MRI. *Neuroimage* **23(1)**:84-97.
- Ashburner, J. (2007): A fast diffeomorphic image registration algorithm. *Neuroimage* **38(1)**:95-113.
- Manjon, J.V. et al. (2010). Adaptive Non-Local Means Denoising of MR Images With Spatially Varying Noise Levels *Journal of Magnetic Resonance Imaging*, **31**: 192-203.