Manual
Computational Anatomy Toolbox - CAT12

Quick Start Guide

VBM data

- **Segment** data using defaults (use **Segment Longitudinal Data** for longitudinal data).
- **Get Total Intracranial Volume** (TIV) to correct for different brain sizes and volumes.
- Check the data quality with **Check Sample** for VBM data (optionally consider TIV and age as nuisance variable).
- **Smooth** data (recommended start value 6-8mm[1]).
- **Specify the 2nd-level** model and check for design orthogonality and sample homogeneity:
    - Use "Full factorial" for cross-sectional data.
    - Use "Flexible factorial" for longitudinal data.
    - Use TIV as covariate (confound) to correct different brain sizes and select centering with overall mean.
    - Select threshold masking with an absolute value of 0.1. This threshold can ultimately be increased to 0.2 or even 0.25 if you still notice non-brain areas in your analysis.
    - If you find a considerable correlation between TIV and any other parameter of interest it is advisable to use global scaling with TIV. For more information, refer to the section "Building the statistical model".
- **Estimate** the model and finally call **Results**.
- Optionally, **Transform SPM-maps** to (log-scaled) p-maps or correlation maps and apply thresholds.
- Optionally, you can try **Threshold-Free Cluster Enhancement** (TFCE) with the SPM.mat file of a previously estimated statistical design.
- Optionally, **Overlay Selected Slices**. If you are using log-p scaled maps from **Transform SPM-maps** without thresholds or the TFCE_log maps, use the following values as the lower range for the colormap for the thresholding: 1.3 (P<0.05); 2 (P<0.01); 3 (P<0.001).
- Optionally, estimate the results for ROI analysis using **Analyze ROIs**. Here, the SPM.mat file of a previously estimated statistical design is used. For more information, see the online help "Atlas creation and ROI based analysis".

Additional surface data

- **Segment** data and also select "Surface and thickness estimation" under "Writing options" (for longitudinal data use **Segment Longitudinal Data**).
- Optionally, **Extract Additional Surface Parameters** (e.g. sulcal depth, gyrification index, cortical complexity).
- **Resample & Smooth Surfaces** (recommended start value 12mm for cortical thickness and 20-25mm for folding measures[1], use the default merging of hemispheres).
- Check data quality using **Check Sample** for surface data.
- **Specify 2nd-level** model for (merged) hemispheres and check for design orthogonality and sample homogeneity:
    - Use "Full factorial" for cross-sectional data.
    - Use "Flexible factorial" for longitudinal data.

- ○ It is not necessary to use TIV as a covariate (confound) because cortical thickness or other surface values are usually not dependent on TIV.
  - ○ It is not necessary to use threshold masking.
  - ○ If you find a considerable correlation between a nuisance parameter and any other parameter of interest it is advisable to use global scaling with that parameter. For more information, refer to the section "Building the statistical model".
- **Estimate the Surface Model** for (merged) hemispheres and finally call **Results**.
- Optionally, **Transform SPM-maps** to (log-scaled) p-maps or correlation maps and apply thresholds.
- Optionally, you can try **Threshold-Free Cluster Enhancement** (TFCE) with the SPM.mat file of a previously estimated statistical design.
- Optionally, **Surface Overlay** for both hemispheres. Select the results (preferably saved as log-p maps with "Transform and threshold SPM-surfaces" or the TFCE_log maps with the different methods for multiple comparison correction) to display rendering views of your results.
- Optionally **Extract ROI-based Surface Values** such as gyrification or fractal dimension to provide ROI analysis. Since version 12.7 extraction of ROI-based thickness is not necessary anymore because this is now included in the segmentation pipeline.
- Optionally, estimate the results for ROI analysis using **Analyze ROIs**. Here, the SPM.mat file of a previously estimated statistical design is used. For more information, see the online help "Atlas creation and ROI based analysis".

Errors during preprocessing

Please use the **Report Error** function if any errors during preprocessing occurred. You first have to select the "err" directory, which is located in the folder of the failed record, and finally, the specified zip-file should be attached manually in the mail.

Additional options

Additional parameters and options are displayed in the CAT12 expert mode. Please note that this mode is for experienced users only.

[1]**Note to filter sizes for Gaussian smoothing**

Due to the high accuracy of the spatial registration approaches used in CAT12, you can also try to use smaller filter sizes. However, for very small filter sizes or even no filtering, you have to apply a non-parametric permutation test such as the TFCE-statistics.

Please also note that for the analysis of cortical folding measures such as gyrification or cortical complexity the filter sizes have to be larger (i.e. in the range of 15-25mm). This is due to the underlying nature of this measure that reflects contributions from both sulci as well as gyri. Therefore, the filter size should exceed the distance between a gyral crown and a sulcal fundus.

Version Information

Preprocessing should remain unaffected until the next minor version number. New processing of your data is not necessary if the minor version number of CAT12 remains unchanged.

Changes in version CAT12.8 (1800)

- Changes in preprocessing pipeline (which affects your results compared to CAT12.7)
  - Volumetric templates, atlases, and TPMs are now transformed to MNI152NLin2009cAsym space to better match existing standards. The templates_volume folder is now renamed to ''templates_MNI152NLin2009cAsym'' to indicate the template space used. The Dartel and Geodesic Shooting templates are renamed or relocated:
    - templates_volumes/Template_0_IXI555_MNI152_GS.nii -> templates_MNI152NLin2009cAsym/Template_0_GS.nii
    - templates_volumes/Template_1_IXI555_MNI152.nii -> templates_MNI152NLin2009cAsym/Template_1_Dartel.nii
    - templates_volumes/TPM_Age11.5.nii -> templates_MNI152NLin2009cAsym/TPM_Age11.5.nii
    - templates_volumes/Template_T1_IXI555_MNI152_GS.nii -> templates_MNI152NLin2009cAsym/Template_T1.nii
    - spm12/toolbox/FieldMap/T1.nii -> templates_MNI152NLin2009cAsym/T1.nii
    - spm12/toolbox/FieldMap/brainmask.nii -> templates_MNI152NLin2009cAsym/brainmask.nii
  - The volumetric atlases have been revised and are now defined with a spatial resolution of 1mm, except for the Cobra atlas, which is defined with 0.6mm resolution. The labels of the original atlases were either transformed from the original data or recreated using a maximum likelihood approach when manual labels were available for all subjects (Cobra, LPBA40, IBSR, Hammers, Neuromorphometrics). In addition, the original labels are now used for all atlases if possible. Some atlases were updated to include new regions (Julichbrain, Hammers). Please note that this will also result in slight differences in ROI estimates compared to previous versions.
  - The bounding box of the Dartel and Geodesic Shooting templates has been changed, resulting in a slightly different image size of the spatially registered images (i.e. modulated normalized segmentations). Therefore, older preprocessed data should not (and cannot) be mixed with the new processed data (which is intended).
  - Transformed T1 Dartel/GS surface templates to the new MNI152NLin2009cAsym space:
    - templates_surfaces/lh.central.Template_T1_IXI555_MNI152_GS.gii -> templates_surfaces/lh.central.Template_T1.gii
    - templates_surfaces/rh.central.Template_T1_IXI555_MNI152_GS.gii -> templates_surfaces/rh.central.Template_T1.gii
    - templates_surfaces_32k/lh.central.Template_T1_IXI555_MNI152_GS.gii -> templates_surfaces_32k/lh.central.Template_T1.gii
    - templates_surfaces_32k/rh.central.Template_T1_IXI555_MNI152_GS.gii -> templates_surfaces_32k/rh.central.Template_T1.gii
  - The surface pipeline has been optimized to better handle data at different spatial resolutions.
  - Older preprocessing pipelines (12.1, 12.3, 12.6) were removed because their support became too difficult.
- Important new features
  - The Mahalanobis distance in the quality check is now replaced by the normalized ratio between overall weighted image quality (IQR) and mean correlation. A low ratio indicates good quality before and after preprocessing and means that IQR is highly

rated (resulting in a low nominal number/grade) and/or mean correlation is high. This is hopefully a more intuitive measure to combine image quality measurement before and after preprocessing.

- CAT12 now allows the use of the BIDS directory structure for storing data (not possible for the longitudinal pipeline). A BIDS path can be defined relative to the participant level directory. The segmentation module now supports the input of nii.gz files (not possible for the longitudinal pipeline).

- The "Basic models" function has been completely restructured and simplified. There are now only two models available for: (1) cross-sectional data and (2) longitudinal data. Options that are not relevant for VBM or SBM have been removed. In addition, a new experimental option has been added that allows a voxel-wise covariate to be defined. This can be used (depending on the contrast defined) to (1) remove the confounding effect of structural data (e.g., GM) on functional data or (2) examine the relationship (regression) between functional and structural data. Additionally, an interaction can be modeled to investigate whether the regression between functional and structural data differs between two groups. Please note that the saved vSPM.mat file can only be evaluated with the TFCE toolbox.

- Added a new function cat_io_data2mat.m to save spatially registered volume or resampled surface data as Matlab data matrix for further use with machine learning tools. Volume data can be resampled to lower spatial resolutions and can optionally be masked to remove non-brain areas.

- Added a new function cat_vol_ROI_summarize.m to summarise co-registered volume data within a region of interest (ROI). This tool can be used in order to estimate ROI information for other (co-registered) modalities (i.e. DTI, (rs)fMRI) which can be also defined as 4D data. Several predefined summary functions are available, as well as the possibility to define your own function.

Changes in version CAT12.7 (1700)

- Changes in preprocessing pipeline (which affects your results compared to CAT12.6)
  - Geodesic shooting registration and surface estimation are now used by default.
  - The surface pipeline is largely updated: **(1)** Parameters for surface reconstruction were optimized. **(2)** Freesurfer distance method Tfs is now implemented, which is computed as the average of the closest distances from the pial to the white matter surface and from the white matter to the pial surface. This reduces the occurrence of larger thickness values and results in more reliable thickness measures. For mapping of 3D data the old thickness metric from PBT is more appropriate and is still used.
  - An additional longitudinal model is implemented that also takes into account deformations between time points. The use of deformations between the time points makes it possible to estimate and detect larger changes, while subtle effects over shorter periods of time in the range of weeks or a few months can be better detected with the model for small changes.
  - Minor changes were made to the segmentation approach to improve accuracy and reliability.
  - Internal resampling is now using a new "optimal" resolution setting to better support high-resolution data.
  - Changed recommendations and defaults for smoothing size to smaller values.
  - Renamed template folder for volumes to templates_volumes.

- - Atlases installed in spm12/atlas are now called cat12_atlasname.
    - The old AAL atlas has been replaced by the AAL3 atlas.
    - Cobra atlas is updated because of some previous inconsistencies.
- Important new features
    - New GUI
    - Added older preprocessing pipelines (12.1, 12.3, 12.6) to provide compatible versions to previous preprocessing. These older preprocessing pipelines are available via the SPM batch editor (SPM → Tools → CAT12) or through expert mode.
    - Added simple batch for cross-sectional and longitudinal data that combines several processing steps from preprocessing to smoothing. These batches are available via the SPM batch editor (SPM → Tools → CAT12) or through expert mode.
    - The function "Display Surface Results" can now also visualize results from VBM analysis and map the 3D volume information onto the surface using an absmax mapping function inside the cortical band. The function is now renamed to "Surface Overlay".
    - Added shell script cat_standalone.sh to call CAT12 batch files (in expert mode) for SPM12 standalone installation (e.g. for HPCs).
    - CAT12 now uses Piwik for anonymized CAT12 user statistics (i.e. version information, potential errors). See "CAT12 user statistics" in the online help for more information. This can be optionally disabled in cat_defaults.m.
    - The extraction of ROI-based thickness is not necessary anymore because this is now included in the segmentation pipeline in cat_main.m.
    - Changed gifti-format after resampling to use external dat-files. This increases processing speed and prevents the SPM.mat file for surfaces from becoming too large. This can be optionally disabled in cat_defaults.m.
    - The use of own atlases for ROI processing is now supported.
    - Updated and extended examples for statistical designs and respective contrasts in the CAT12 manual.

Changes in version CAT12.6 (1445)

- Changes in preprocessing pipeline (which affects your results compared to CAT12.5)
    - Two main parts of the preprocessing of CAT12 were largely updated: **(1)** Incorrect estimates of the initial affine registration were found to be critical for all subsequent preprocessing steps and mainly concerned skull-stripping and tissue segmentation. This was a particular problem in the brains of older people or children, where the thickness of the skull differs from that of the template. The new estimate of the initial affine registration should now be more robust. In the CAT report, the registered contour of the skull and the brain is now overlayed onto the image to allow for easier quality control. **(2)** Skull-stripping now uses a new adaptive probability region-growing (APRG) approach, which should also be more robust. APRG refines the probability maps of the SPM approach by region-growing techniques of the gcut approach with a final surface-based optimization strategy. This is currently the method with the most accurate and reliable results.
    - The longitudinal pipeline should now also be more sensitive for the detection of effects over longer time periods with VBM (ROI and SBM approaches are not affected by the length of the period). In earlier versions, the average image was used to estimate the spatial registration parameters for all time points. Sometimes this average image was not as accurate if the images of a subject were too different (e.g. due to large

ventricular changes). Now, we rather use the average spatial registration parameters (i.e. deformations) of all time points, which makes the approach more robust for longer periods of time. However, the SPM12 Longitudinal Toolbox can be a good alternative for longer periods of time if you want to analyze your data voxel by voxel. Surface-based preprocessing and also the ROI estimates in CAT12 are not affected by the potentially lower sensitivity to larger changes, as the realigned images are used independently to create cortical surfaces, thickness, or ROI estimates.

- Important new features
  - CAT report now additionally plots the contour of the registered skull and brain and the central surface onto the image and visualizes the result of skull-stripping. "Display Surface Results" is largely updated.
  - Parallelization options in CAT12 now enable subsequent batch jobs and are also supported for longitudinal preprocessing.

Changes in version CAT12.5 (1355)

- Changes in preprocessing pipeline (which affects your results compared to CAT12.3)
  - Detection of white matter hyperintensities (WMHs) is updated and again enabled by default.
  - The default internal interpolation setting is now "Fixed 1 mm" and offers a good trade-off between optimal quality and preprocessing time and memory demands. Standard structural data with a voxel resolution around 1 mm or even data with high in-plane resolution and large slice thickness (e.g. 0.5x0.5x1.5 mm) will benefit from this setting. If you have higher native resolutions the high-resolution option "Fixed 0.8 mm" will sometimes offer slightly better preprocessing quality with an increase of preprocessing time and memory demands.
- Important new features
  - CAT12 can now deal with lesions that have to be set to "0" in your image using the Stroke Lesion Correction (SLC) in expert mode. These lesion areas are not used for segmentation or spatial registration, thus these preprocessing steps should be almost unaffected.

Changes in version CAT12.4 (1342)

- This version had some severe errors in spatial registration which affected all spatially registered data and should not be used anymore.

Changes in version CAT12.3 (1310)

- Changes in preprocessing pipeline (which affects your results compared to CAT12.2)
  - Skull-stripping is again slightly changed and the SPM approach is now used as default. The SPM approach works quite well for the majority of data. However, in some rare cases, parts of GM (i.e. in the frontal lobe) might be cut. If this happens the GCUT approach is a good alternative.
  - Spatial adaptive non-local mean (SANLM) filter is again called as a very first step because noise estimation and de-noising works best for original (non-interpolated) data.
  - Detection of white matter hyperintensities (WMHs) is currently disabled by default, because of unreliable results for some data.
- Important new features

- ○ Cobra atlas has been largely extended and updated.

Changes in version CAT12.2 (1290)

- Changes in preprocessing pipeline (which affects your results compared to CAT12.1)
  - ○ Skull-stripping now additionally uses SPM12 segmentations by default: The default gcut approach in CAT12.1 removed too much of the surrounding (extracranial) CSF, which led to a slight underestimation of TIV for atrophied brains. The skull-stripping approach based on the SPM12 segmentations prevents this through a more conservative approach. However, sometimes parts of the meninges (i.e. dura mater) or other non-brain parts remain in the GM segmentation. By combining both approaches a more reliable skull-stripping is achieved.
  - ○ More reliable estimation of TIV: The changed skull-stripping also affects the estimation of TIV, which is now more reliable, especially for atrophied brains.
- Important new features
  - ○ Automatic check for design orthogonality and sample homogeneity using SPM.mat in **Basic Models**
  - ○ Added equi-volume model by Bok and a multi-save option for mapping native volumes to individual surfaces.
  - ○ Added Local-Global Intrinsic Functional Connectivity parcellation by Schaefer et al. for resting-state fMRI data.

Introduction and Overview

This manual is intended to help any user to perform a computational anatomy analysis using the CAT12 Toolbox. Although it mainly focuses 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 are also presented and can be applied with few changes.

The manual can be divided into four main sections:

- Naturally, a quick guide of how to *get started* is given at the beginning. This section provides information about *downloading and installing* the software and *starting* the Toolbox. Also, a brief *overview* of the steps of a VBM analysis is given.

- This is followed by a *detailed description of a basic VBM analysis* that guides the user step-by-step through the entire process – from preprocessing to contrast selection. This description should provide all the information necessary to successfully analyze most studies.

- There are some *specific 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 how these changes can be applied. Only the changes are described here – steps such as quality control or smoothing are the same as those described in the basic analysis and are not repeated a second time.

- The guide concludes with *information about native, normalized, and modulated volumes* that determine how the results can be interpreted. Also, an overview of the naming conventions used and technical information is given.

Getting Started

DOWNLOAD AND INSTALLATION

- The CAT12 Toolbox runs within SPM12. That is, SPM12 must 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.

- 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 an additional window (Fig. 2).
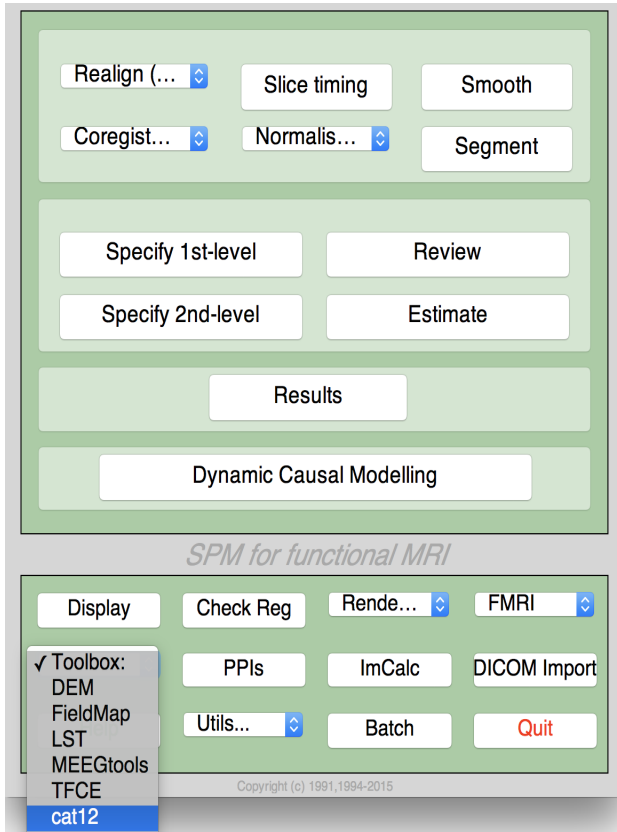
Figure 1: SPM12 menu

Figure 2: CAT12 Window

BASIC VBM ANALYSIS (OVERVIEW)

The CAT12 Toolbox comes with different modules, which may be used for 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 "Segment Data".

2. After the preprocessing is finished, a **quality check** is highly recommended. This can be achieved via the modules "Display slices" and "Check sample". Both options are located in the CAT12 window under "Check Data Quality". Furthermore, quality parameters are estimated and saved in xml-files for each data set during preprocessing. These quality

parameters are also printed on the report PDF-page and can be additionally used in the module "Check sample".

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 with the standard SPM module "Estimate" (except for surface-based data where the function "Estimate Surface Models" should be used instead).

6. If you have used total intracranial volume (TIV) as a confound in your model to correct for different brain sizes it is necessary to check whether TIV reveals a considerable correlation with any other parameter of interest and rather uses global scaling as an alternative approach.

7. After estimating the statistical model, contrasts are defined to get the results of the analysis. This is done with the standard SPM module "Results".

The sequence of "preprocessing → quality check → smoothing → statistical analysis" remains the same for every VBM or SBM 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.

---

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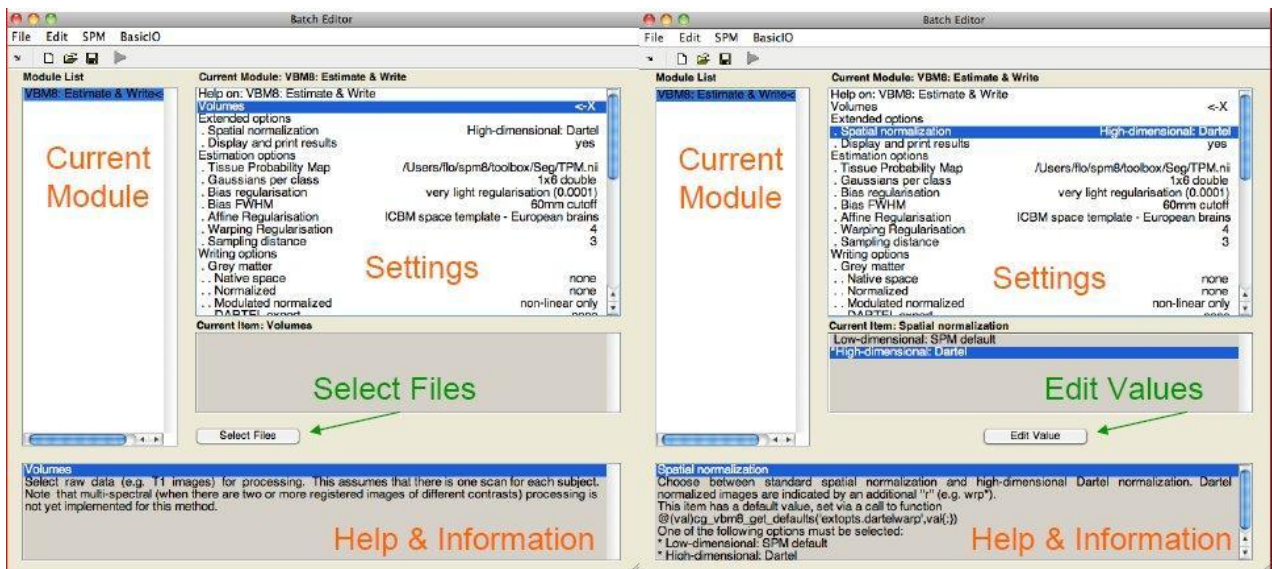
Figure 3: The Batch Editor is the environment where the analysis is set up. *Left*: For all settings marked with "<-X", files have to be selected ("Select Files"). *Right*: Parameters can be edited and adapted ("Edit Value").

Overview of CAT12 Processing

CAT12 MAJOR PROCESSING STEPS

As shown in **Figure 4**, CAT's processing workflow comprises of two main steps: voxel-based processing and surface-based processing. The former is a prerequisite for the latter, but not the other way round. That is, while voxel-based processing is always required for surface-based analyses, users not interested in surface-based analyses can simply omit this second step to save processing time. The '*Voxel-based processing*' step (**Figure 4**, left) can be thought of as one module for tissue segmentation and another one for spatial registration. An optional third module allows for the generation of ROIs and the calculation of ROI-based measures. The '*Surface-based processing*' step (**Figure 4**, right) can be thought of as one module for surface creation and another one for surface registration. An optional third module allows for the generation of surfaced-based ROIs and the calculation of ROI-based measures. As shown in **Figure 4**, the different modules utilize different priors, templates, and atlases. Those are briefly explained in the next paragraph.

Voxel-based processing: While the final tissue segmentation in CAT is independent of tissue priors, the segmentation procedure is initialized using Tissue Probability Maps (TPMs). The standard TPMs (as provided in SPM) suffice for the vast majority of applications, and customized TPMs are only recommended for data obtained in young children. Please note that these TPMs should contain 6 classes: GM/WM/CSF and 3 background classes. For spatial registration, CAT uses DARTEL (Ashburner, 2007) or Geodesic Shooting (Ashburner and Friston, 2011) with predefined templates. Those templates are an appropriate choice for most studies and, again, sample-specific DARTEL or Geodesic Shooting templates may only be advantageous for young children. For the voxel-based ROI analyses, CAT offers a selection of volume-based atlases in the predefined template space. Thus, any atlas-based ROI analysis requires a normalization of the individual scans to CAT's default DARTEL or Geodesic Shooting template. On that note, the aforementioned creation and selection of a customized (rather than the predefined) DARTEL or Geodesic Shooting template will disable the third module for the voxel-based ROI analysis.

Surface-based processing: During the surface registration, the cortical surfaces of the two hemispheres are registered to the Freesurfer 'FsAverage' template[2] (as provided with CAT). For surface-based ROI analyses, CAT provides a selection of surface-based atlases. As the surface registration module uses the Freesurfer 'FsAverage' template, surface-based ROI analyses are not impacted by any template modification during voxel-based processing. In contrast to voxel-based ROI analyses, surface-based ROI analyses can, therefore, be applied regardless of whether the predefined or customized versions of the DARTEL or Geodesic Shooting template have been used.

---

[2] https://surfer.nmr.mgh.harvard.edu/fswiki/FsAverage

CAT major processing steps
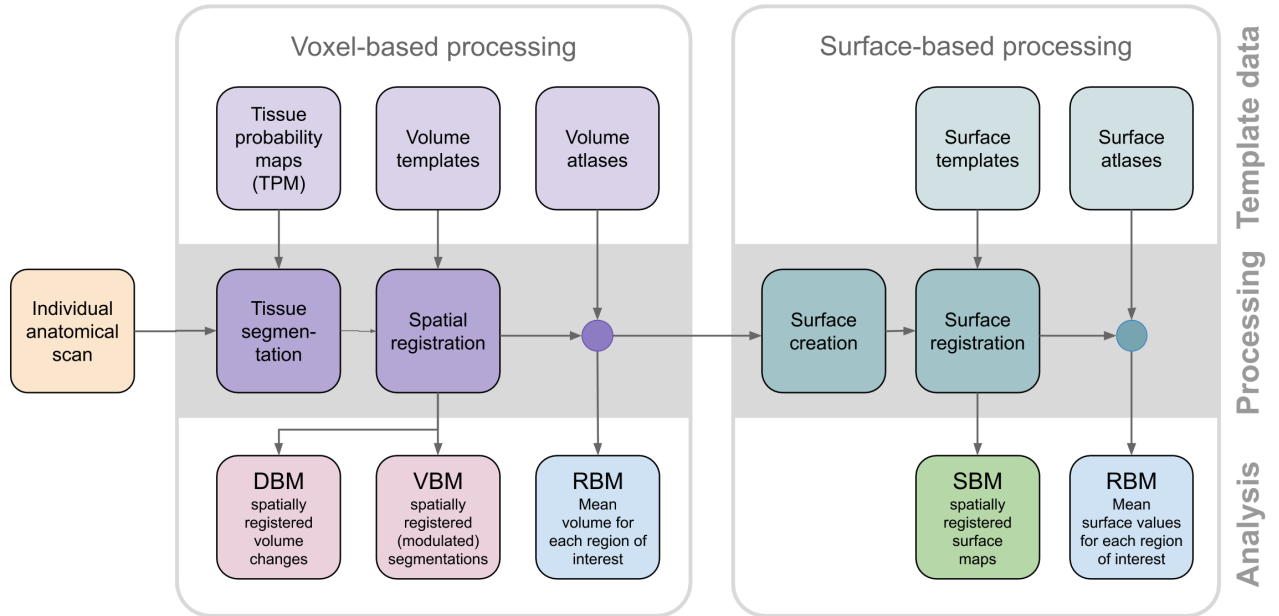

Figure 4: Overview of CAT's major processing steps.

CAT12 Processing Steps in Detail

The modules described in the previous section help understand the CAT's overall processing workflow, including its priors, templates, and atlases. Also, data processing in CAT can be separated into three main processes: (1) the initial voxel-based processing, (2) the main voxel-based processing, and (3) the surface-based processing (optional), as further detailed below and visualized in **Figure 5**.

The '*initial voxel-based processing*' begins with a spatial adaptive non-local means (SANLM) denoising filter (Manjón et al., 2010), which is followed by internal resampling to properly accommodate low-resolution images and anisotropic spatial resolutions. The data are then bias-corrected and affine-registered (to further improve the outcomes of the following steps) followed by the standard SPM "unified segmentation" (Ashburner and Friston, 2005). The outcomes of the latter step will provide the starting estimates for the subsequent refined voxel-based processing.

The *'refined voxel-based processing'* uses the output from the unified segmentation and proceeds with skull-stripping of the brain. The brain is then parcellated into the left and right hemisphere, subcortical areas, and the cerebellum. Furthermore, local white matter hyperintensities are detected (to be later accounted for during the spatial normalization and cortical thickness estimation). Subsequently, a local intensity transformation of all tissue classes is performed, which is particularly helpful to reduce the effects of higher gray matter intensities in the motor cortex, basal ganglia, or occipital lobe before the final adaptive maximum a posteriori (AMAP) segmentation. This final AMAP segmentation step (Rajapakse et al., 1997), which does not rely on *a priori* information of the tissue probabilities, is then refined by applying a partial volume estimation (Tohka et al., 2004), which effectively estimates the fractional content for each tissue type per voxel. As a last default step, the tissue segments are spatially normalized to a common reference space using DARTEL (Ashburner, 2007) or Geodesic Shooting (Ashburner and Friston, 2011) registrations.

Optionally, the *'surface-based processing'* will be run following the completion of the voxel-based processing steps. Here, the cortical thickness estimation and reconstruction of the central surface occur in one step using a projection-based thickness (PBT) method (Dahnke et al., 2013). Importantly, this PBT allows the appropriate handling of partial volume information, sulcal blurring, and sulcal asymmetries without explicit sulcus reconstruction (Dahnke et al., 2013). After the initial surface reconstruction, topological defects[3] are repaired using spherical harmonics (Yotter et al., 2011). The topological correction is followed by a surface refinement, which results in the final central surface mesh. This mesh provides the basis to extract folding patterns (i.e., based on the position of mesh nodes relative to each other), where the resulting local values (e.g., absolute mean curvature) are projected onto each node. Subsequently, the individual central surfaces are spatially registered to the Freesurfer 'FsAverage' template using a spherical mapping with minimal distortions (Yotter et al., 2011b). In the last step, the local thickness values are transferred onto the Freesurfer 'FsAverage' template. While this last step is performed by default during the surface-based processing, it can be repeated to also transfer measurements of cortical folding (e.g., gyrification) as well as other measurements or data (e.g., functional or quantitative MRI) to the Freesurfer 'FsAverage' template. Note that the spatial registration of cortical measurements is combined with a spatial smoothing step in CAT to prepare the data for statistical analysis.

CAT processing pipeline
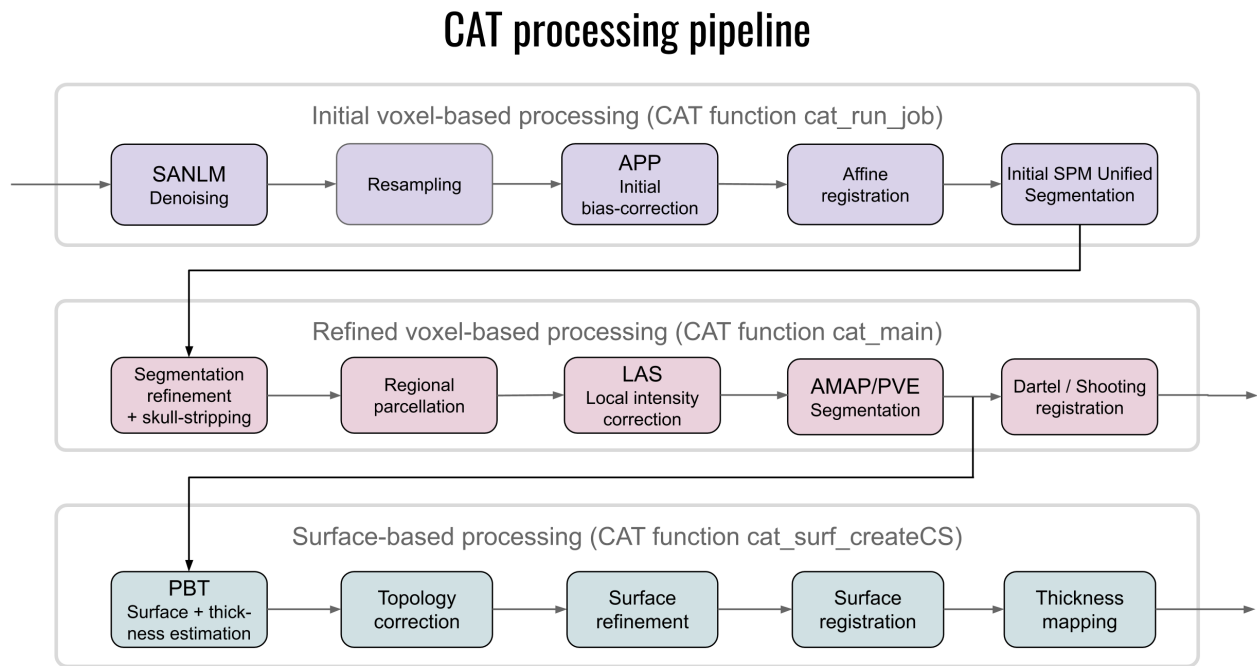
Figure 5: Schematic flowchart of CAT's processing pipeline.

---

[3] Topology defects are holes and handles that prevent the surface from being homeomorphic with a sphere.

Basic VBM Analysis (detailed description)

Preprocessing Data

Please note that additional parameters for expert users are displayed in the GUI if you set the option cat.extopts.expertgui to "1" in cat_defaults.m or call cat12 by:

cat12('expert')

CAT12 → Preprocessing → Segment Data

Parameters:

- o Volumes <-X → Select Files → *[select the new files]* → Done

    - Select one volume for each subject. As the Toolbox does not support multispectral data (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.

    - Importantly, the images need to be in the same orientation as the priors; you can double-check and correct via using "Display" in the SPM menu. The priors are located in your SPM folder "SPM12 → tpm → TPM.nii")

- o Split job into separate processes →*[use defaults or modify]*

    - To use multi-threading the CAT12 segmentation job with multiple subjects can be split into separate processes that run in the background. If you don't want to run processes in the background then set this value to 0.

    - Keep in mind that each process needs about 1.5..2GB of RAM, which should be considered to choose the appropriate number of processes.

- o Options for initial SPM12 affine registration → *[use defaults or modify]*

    - The defaults provide a solid starting point. The SPM12 tissue probability maps (TPMs) are used for the initial spatial registration and segmentation. Alternatively, customized TPMs can be chosen (e.g. for children data) that were created with the Template-O-Matic (TOM) Toolbox. Please note that these TPMs should contain 6 classes: GM/WM/CSF and 3 background classes.

- o Extended options for CAT12 segmentation → [*use defaults or modify*]

    - Again, the defaults provide a solid starting point. Using the extended options you can adjust specific parameters or the strength of different corrections ("0" means no correction and "0.5" is the default value that works best for a large variety of data).

    - CAT12 provides a template for the high-dimensional DARTEL registration that should work for most data. However, a customized DARTEL template can be selected (e.g. for children data) that was created using the DARTEL toolbox. For more information on the necessary steps, see the section "Customized DARTEL Template".

- Writing options → [*use defaults or modify*]

  - For **GM, and WM image volumes** see at the end of the document: "Additional information on native, normalized and modulated normalized volumes". *Note: The default option "Modulated normalized" results in an analysis of relative differences in regional GM volume, that have to be corrected for individual brain size in the statistical analysis using total intracranial volume (TIV).*

  - A **Bias, noise, and globally intensity corrected T1 image**, in which MRI inhomogeneities and noise are removed and intensities are globally normalized, can be written in normalized space. This is useful for quality control and also to create an average image of all normalized T1 images to display/overlay the results. *Note: For a basic VBM analysis use the defaults.*

  - A **partial volume effect (PVE) label image volume** can also be written in normalized or native space or as a DARTEL export file. This is useful for quality control. *Note: For a basic VBM analysis use the defaults.*

  - The **Jacobian determinant** for each voxel can be written in normalized space. This information can be used to do a Deformation-Based Morphometry (DBM) analysis. *Note: For a basic VBM analysis this is not needed.*

  - Finally, **deformation fields** can be written. This option is useful to re-apply normalization parameters to other co-registered images (e.g. fMRI or DTI data). *Note: For a basic VBM analysis this is not needed.*

Note: If the segmentation fails, this is often due to an unsuccessful initial spatial registration. In this case, you can try to set the origin (anterior commissure) in the Display tool. Place the cursor roughly on the anterior commissure and press "Set Origin" The now displayed correction in the coordinates can be applied to the image with the button "Reorient". This procedure must be repeated for each data set individually.

SECOND MODULE: DISPLAY SLICES (OPTIONALLY)

CAT12 → Check data quality → Display slices

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 of whether 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 as 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.

- o Spatial orientation
- o Show slice in mm → [*use defaults or modify*]
  - - This module displays horizontal slices. This default setting provides a good overview.

CAT12 → Statistical Analysis → Estimate TIV

Parameters:

- o XML files <-X → Select Files → [*select xml-files]* → Done
  - - Select the xml-files in the report-folder [e.g. the "cat_*.xml"].
- o Save values → *TIV only*
  - - This option will save the TIV values for each data set in the same order as the selected xml-files. Optionally you can also save the global values for each tissue class, which might be interesting for further analysis, but is not recommended if you are interested in only using TIV as a covariate.
- o *Output file → [use defaults or modify]*

Please note that TIV is strongly recommended as a covariate for all VBM analyses to correct different brain sizes. This step is not necessary for deformation- or surface-based data. Please also make sure that TIV does not correlate too much with your parameters of interest (please make sure that you use "Centering" with "Overall mean", otherwise the check for orthogonality in SPM sometimes does not work correctly). In this case, you should use global scaling with TIV.

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

Parameters:

- o 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)]. It is recommended to use the unsmoothed segmentations that provide more anatomical details. This tool visualizes the correlation between the volumes using a boxplot and correlation matrices. Thus, it will help to identify outliers. Any outlier should be carefully inspected for artifacts or pre-processing errors using "Check worst data" in the

GUI. If you specify different samples the mean correlation is displayed in separate boxplots for each sample.

- o Load quality measures (optional) → *[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.
- o 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 a slightly more accurate correlation but are much slower.
- o Nuisance → *[enter nuisance variables if applicable]*
    - For each nuisance variable which you want to remove from the data before 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.

A window opens with a correlation matrix in which the correlation between the volumes is displayed. 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 are displayed in the lower right corner and allow a closer look. The slider below the image changes the displayed slice. The pop-up menus in the top right-hand corner provide more options. Here you can select other measures that are displayed in the boxplot (e.g. optional quality measures such as noise, bias, weighted overall image quality if these values were loaded), change the order of the correlation matrix (according to the file name or mean correlation). Finally, most deviating data can be displayed in the SPM graphics window to check the data more closely. For surfaces, two additional measures are available (if you have loaded the xml-files) to utilize the quality check of the surface extraction. The Euler number gives you an idea about the number of topology defects, while the defect size indicates how many vertices of the surface are affected by topology defects. Because topology defects are mainly caused by noise and other image artifacts (e.g. motion) this gives you an impression about the potential quality of your extracted surface.

The boxplot 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 does not always mean 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 are 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. The same holds for all other measures. Data with deviating values have not necessarily to be excluded from the sample, but these data should be checked more carefully with the "Check most deviating data" option. The order of the most deviating data is changed for every measure separately. Thus, the check for the most deviating data will give you a different order for mean

correlation and other measures which allows you to judge your data quality using different features.

If you have loaded quality measures, you can also display the ratio between weighted overall image quality (IQR) and mean correlation. These two are the most important measures for assessing image quality. Mean correlation measures the homogeneity of your data used for statistical analysis and is therefore a measure of image quality **after** pre-processing. Data that deviate from your sample increase variance and therefore minimize effect size and statistical power. The weighted overall image quality, on the other hand, combines measurements of noise and spatial resolution of the images **before** pre-processing. Although CAT12 uses effective noise-reduction approaches (e.g. spatial adaptive non-local means filter) pre-processed images are also affected and should be checked.

The ratio between IQR and mean correlation makes it possible to combine these two measures of image quality before and after pre-processing. A low ratio indicates good quality before and after preprocessing and means that IQR is highly rated (resulting in a low nominal number/grade) and/or mean correlation is high. Both measures contribute to the ratio and are normalized before by their standard deviation. In the respective plot, the ratio is colour-coded and each point can be selected to get the filename and display the selected slice to check data more closely.

FIFTH MODULE: SMOOTH

SPM menu → Smooth

Parameters:

- o Images to Smooth <-X → Select Files → *[select grey matter volumes]* → Done
    - Select the newly written data [e.g. the "mwp1" files, which are the modulated (m) normalized (w) grey matter segments (p1)].
- o 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.
- o Data Type → *[use defaults or modify]*
- o Filename Prefix → *[use defaults or modify]*

Building the Statistical Model

Although many potential designs are offered in the 2nd-level analysis I recommend using the "Full factorial" design as it covers most statistical designs. For cross-sectional VBM data you usually have 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)

CAT12 → Statistical Analysis → Basic Models

    Parameters:

- o Directory <-X → Select Files → *[select the working directory for your analysis]* → Done
- o Design → "Two-sample t-test"
    - ▪ Group 1 scans → *Select Files* → *[select the smoothed grey matter data for group 1; following this script these are the "smwp1" files]* → Done
    - ▪ Group 2 scans → *Select Files* → *[select the smoothed grey matter data for group 2]* → Done
    - ▪ Independence → Yes
    - ▪ Variance → Equal or Unequal
    - ▪ Grand mean scaling → No
    - ▪ ANCOVA → No
- o **Covariates*** (see the text box)
- o Masking
    - ▪ Threshold Masking → Absolute → *[specify a value (e.g. "0.1")]*
    - ▪ Implicit Mask → Yes
    - ▪ Explicit Mask → <None>
- o Global Calculation → Omit
- o Global Normalization
    - ▪ Overall grand mean scaling → No
- o Normalization → None

> *You could specify one or more 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 for modulated data in VBM to correct different brain sizes. This is not necessary for surface analysis and DBM.
>
> - ● 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 Overall mean

CAT12 → Statistical Analysis → Basic Models

Parameters:

- o Directory <-X → Select Files → *[select the working directory for your analysis]* → Done
- o 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. "age")]*
    - 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 data of the young males)]*
    Cell
    - Levels → *[specify text (e.g. "1 2")]*
    - Scans → *[select files (e.g. the smoothed GM data of the old males)]*
    Cell
    - Levels → *[specify text (e.g. "2 1")]*
    - Scans → *[select files (e.g. the smoothed GM data of the young females)]*
    Cell
    - Levels → *[specify text (e.g. "2 2")]*
    - Scans → *[select files e.g. the smoothed GM data of the old females)]*
- **o Covariates*** (see the text box in example for two-sample T-test)
- o Masking
  - Threshold Masking → Absolute → *[specify a value (e.g. "0.1")]*
  - Implicit Mask → Yes
  - Explicit Mask → <None>
- o Global Calculation → Omit
- o Global Normalization
  - Overall grand mean scaling → No
- o Normalization → None

CAT12 → Statistical Analysis → Basic Models

Parameters:

- o Directory <-X → Select Files → *[select the directory for your analysis]* → Done

- o Design → "Multiple Regression"
  - Scans → [select files (e.g. the smoothed GM data of all subjects)] → Done
  - Covariates → "New: Covariate"
    Covariate
    - Vector → *[enter the values in the same order as the respective file names of the smoothed images]*
    - Name → [specify (e.g. "*age*")]
    - Centering → Overall mean
  - Intercept → Include Intercept
- **o Covariates* (see the text box in example for two-sample T-test)**
- o Masking
  - Threshold Masking → Absolute → *[specify a value (e.g. "0.1")]*
  - Implicit Mask → Yes
  - Explicit Mask → <None>

- o Global Calculation → Omit

- o Global Normalization
  - Overall grand mean scaling → No

- o Normalization → Non

To use a polynomial model, you need to estimate the polynomial function of your parameter before analyzing it. To do this, use the function cat_stat_polynomial (included with CAT12 >r1140):

y = cat_stat_polynomial(x,order)

where "x" is your parameter and "order" is the polynomial order (e.g. 2 for quadratic).

Example for polynomial order 2 (quadratic):

CAT12 → Statistical Analysis → Basic Models

Parameters:

- o Directory <-X → Select Files → *[select the directory for your analysis]* → Done

- o Design → "Multiple Regression"
    - ▪ Scans → [select files (e.g. the smoothed GM data of all subjects)] → Done
    - ▪ Covariates → "New: Covariate"
      Covariate
        - ● Vector → *[specify linear term (e.g. "y(:,1)")]*
        - ● Name → [specify (e.g. "*age linear*")]
        - ● Centering → Overall mean
    - ▪ Covariates → "New: Covariate"
      Covariate
        - ● Vector → *[specify quadratic term (e.g. "y(:,2)")]*
        - ● Name → [specify (e.g. "*age quadratic*")]
        - ● Centering → Overall mean
    - ▪ Intercept → Include Intercept
- o **Covariates*** (see the text box in example for two-sample T-test)
- o Masking
    - ▪ Threshold Masking → Absolute → *[specify a value (e.g. "0.1")]*
    - ▪ Implicit Mask → Yes
    - ▪ Explicit Mask → <None>
- o Global Calculation → Omit
- o Global Normalization
    - ▪ Overall grand mean scaling → No
- o Normalization → None

CAT12 → Statistical Analysis → Basic Models

   Parameters:

- o   Directory <-X → Select Files → *[select the working directory for your analysis]* → Done
- o   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 data of the males)]*
       Cell
         - ●   Levels → *[specify text (e.g. "2")]*
         - ●   Scans → *[select files (e.g. the smoothed GM data of the females)]*
- o   Covariates → "New: Covariate"
     - ▪   Covariate
         - ●   Vector → *[enter the values in the same order as the respective file names of the smoothed images]*
         - ●   Name → [specify (e.g. "*age*")]
         - ●   Interactions → With Factor 1
         - ●   Centering → Overall mean
- **o   Covariates*** (see the text box in example for two-sample T-test)
- o   Masking
     - ▪   Threshold Masking → Absolute → *[specify a value (e.g. "0.1")]*
     - ▪   Implicit Mask → Yes
     - ▪   Explicit Mask → <None>
- o   Global Calculation → Omit
- o   Global Normalization
     - ▪   Overall grand mean scaling → No
- o   Normalization → None

To use a polynomial model you have to estimate the polynomial function of your parameter prior to the analysis. Use the function cat_stat_polynomial (provided with CAT12 >r1140) for that purpose:

     y = cat_stat_polynomial(x,order)

where "x" is your parameter and "order" is the polynomial order (e.g. 2 for quadratic).

Example for polynomial order 2 (quadratic):

CAT12 → Statistical Analysis → Basic Models

    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 data of the males)]*
    Cell
    - Levels → *[specify text (e.g. "2")]*
    - Scans → *[select files (e.g. the smoothed GM data of the females)]*
- Covariates → "New: Covariate"
  - Covariate
    - Vector → *[specify linear term (e.g. "y(:,1)")]*
    - Name → [specify (e.g. "*age linear*")]
    - Interactions → With Factor 1
    - Centering → Overall mean
- Covariates → "New: Covariate"
  - Covariate
    - Vector → *[specify quadratic term (e.g. "y(:,2)")]*
    - Name → [specify (e.g. "*age quadratic*")]
    - Interactions → With Factor 1
    - Centering → Overall mean
- **Covariates*** (see the text box in example for two-sample T-test)
- Masking
  - Threshold Masking → Absolute → *[specify a value (e.g. "0.1")]*
  - Implicit Mask → Yes
  - Explicit Mask → <None>
- Global Calculation → Omit
- Global Normalization

- ▪ Overall grand mean scaling → No
- o Normalization → None

ESTIMATING THE STATISTICAL MODEL

SPM menu → Estimate

Parameters:

- o Select SPM.mat <-X → Select Files → *[select the SPM.mat which you just built]* → Done
- o Method → "Classical"

CHECKING FOR DESIGN ORTHOGONALITY

If you have modeled TIV as confounding parameter, you must check whether TIV is orthogonal (in other words independent) to any other parameter of interest in your analysis (e.g. parameters you are testing for). This means that TIV should not correlate with any other parameter of interest, otherwise not only the variance explained by TIV is removed from your data, but also parts of the variance of your parameter of interest.

Please use "Overall mean" as "Centering" for the TIV covariate. Otherwise, the orthogonality check sometimes even indicates a meaningful co-linearity only because of scaling problems.

To check the design orthogonality, you can use the Review function in the SPM GUI:

SPM menu → Review

- o Select SPM.mat <-X → Select Files → *[select the SPM.mat which you just built]* → Done
- o Design → Design orthogonality
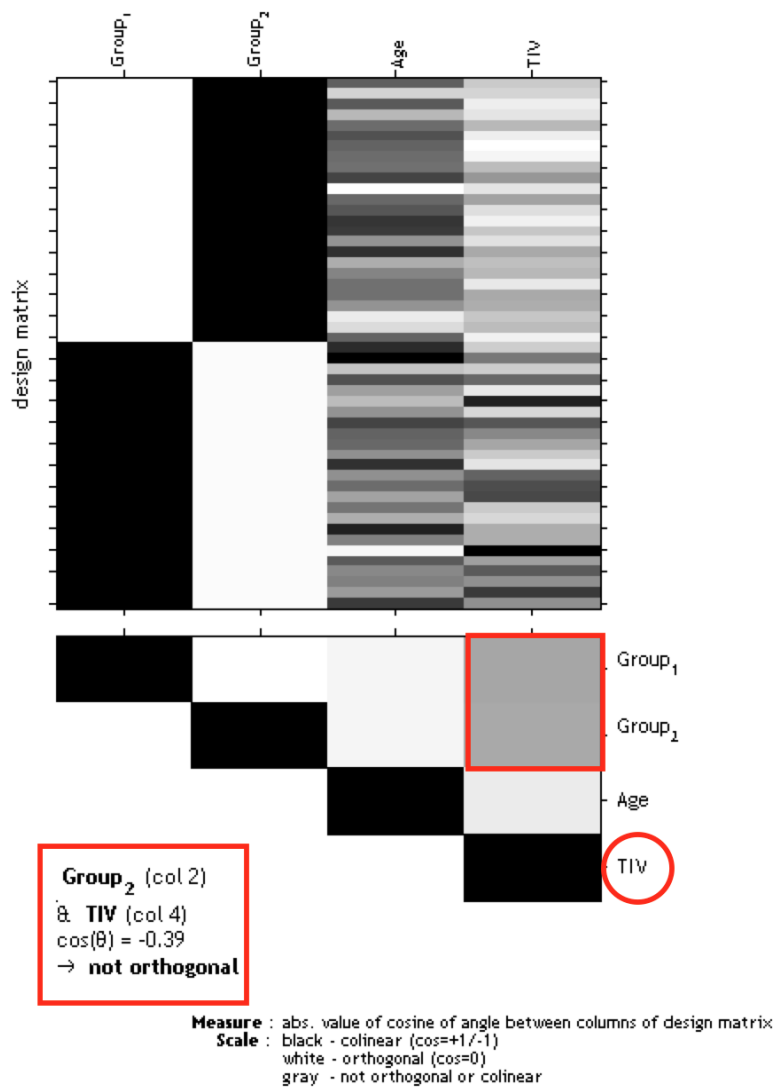
Figure 6: Gray boxes between the parameters point to a co-linearity (correlation): the darker the box the larger the correlation (which also holds for inverse correlations). If you click in the box the co-linearity between the parameters is displayed.

The larger the correlation between TIV and any parameter of interest the more the need to not use TIV as nuisance parameter. In this case an alternative approach is to use global scaling with TIV. Use the following settings for this approach:

- o Global Calculation → User → Global Values <-X → Define the TIV values here
- o Global Normalisation → Overall grand mean scaling → Yes → Grand mean scaled value → Define here the mean TIV of your sample (or as approximation a value of 1500 which might fit for the majority of data from adults)
- o Normalisation → Proportional

This approach is (proportionally) scaling the data according to the individual TIV values. Instead of removing variance explained by TIV all data are scaled by their TIV values. This is also similar to the approach that was used in VBM8 ("modulate for non-linear effects only"), where global scaling was internally used with an approximation of TIV (e.g. inverse of affine scaling factor).

Please note that global normalization also affects the absolute masking threshold as your images are now scaled to the "Grand mean scaled value" of the Global Normalization option. If you have defined the mean TIV of your sample (or an approximate value of 1500) here, no change of the absolute threshold is required. Otherwise, you will have to correct the absolute threshold because your values are now globally scaled to the "Grand mean scaled value" of the Global Normalization option.

An alternative approach for all models with more than one group should not be unmentioned. This is to use TIV as nuisance parameter with interaction and mean centering with factor "group". That approach will result in separate columns of TIV for each group. The values for the nuisance parameters are here mean-centered for each group. As the mean values for TIV are corrected separately for each group, it is prevented that group differences in TIV have a negative impact on the group comparison of the image data in the GLM. Variance explained by TIV is here removed separately for each group without affecting the differences between the groups.

Unfortunately, no clear advice for either of these two approaches can be given and the use of these approaches will also affect the interpretation of your findings.

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).*

You can also use the SPM Batch Editor to automate that procedure:

SPM Batch Editor → SPM → Stats → Contrast Manager

Please note that all zeros at the end of the contrast don't have to be defined, but are sometimes remained for didactic reasons.

Two-sample T-test

T-test

- For Group A > Group B   **1 -1**
- For Group A < Group B   **-1 1**

F-test

- For any differences between Group A and B   **eye(2)-1/2**
- Effects of interest (for plotting)   **0 1**
   **1 0**

2x2 ANOVA

T-test

- For young males > old males   **1 -1 0 0**
- For young females > old females   **0 0 1 -1**
- For young males > young females   **1 0 -1 0**
- For old males > old females   **0 1 0 -1**
- For males > females (main effect sex)   **1 1 -1 -1**
- For young > old (main effect age)   **1 -1 1 -1**
- For young > old & males > females   **1 -1 -1 1**
- For young > old & males < females   **-1 1 1 -1**

F-test

- For main effect age in males   **1 -1 0 0**
- For main effect age in females   **0 0 1 -1**
- For main effect sex in young subjects   **1 0 -1 0**
- For main effect sex in old subjects   **0 1 0 -1**
- For main effect sex   **1 1 -1 -1**
- For main effect age   **1 -1 1 -1**
- For interaction sex by age   **1 -1 -1 1**
- Any difference between the 4 groups   **eye(4)-1/4**
- Effects of interest (for plotting)   **eye(4)**

Multiple Regression (Linear)
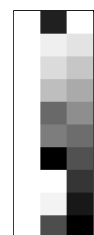
T-test

- For positive regression slope (correlation)   **0 0 1**
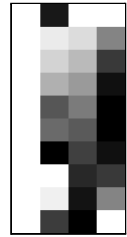- For negative regression slope (correlation)   **0 0 -1**

F-test

- Any regression slope (effects of interest)   **0 0 1**

The two leading zeros in contrast indicate the constant (sample effect, 1st column in the design matrix) and TIV (2nd column in the design matrix). If no additional covariate such as TIV is defined you must skip one of the leading zeros (e.g. "0 1" and "0 -1").

Multiple Regression (Polynomial)

T-test

- For positive linear effect   **0 0 1 0**
- For positive quadratic effect   **0 0 0 1**
- For negative linear effect   **0 0 -1 0**
- For negative quadratic effect   **0 0 0 -1**

F-test

- For any linear effect   **0 0 1 0**
- For any quadratic effect   **0 0 0 1**
- For any linear or quadratic effect (effects of interest) **[zeros(2,2) eye(2)]**

The two leading zeros in the contrast indicate the constant (sample effect, 1st column in the design matrix) and TIV (2nd column in the design matrix). If no additional covariate such as TIV is defined you must skip one of the leading zeros (e.g. "0 1").

Interaction (Linear)
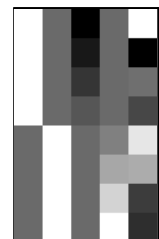
T-test

- For regression slope Group A > Group B **0 0 1 -1 0**
- For regression slope Group A < Group B **0 0 -1 1 0**

F-test

- For any difference in regression slope **0 0 1 -1 0**
- Effects of interest (for plotting)   **[zeros(2,2) eye(2)]**

The two leading zeros in the contrast indicate the main effect "group" and should be always set to "0" in the contrast, because we are rather interested in the interaction.

Interaction (Polynomial)

T-test

- For linear regression slope Group A > Group B   **0 0 1 -1 0 0 0**
- For linear regression slope Group A < Group B   **0 0 -1 1 0 0 0**
- For quadratic regression Group A > Group B   **0 0 0 0 1 -1 0**
- For quadratic regression Group A < Group B   **0 0 0 0 -1 1 0**

F-test

- For any difference in linear regression slope   **0 0 1 -1 0 0 0**
- For any difference in quadratic regression   **0 0 0 0 1 -1 0**
- For any difference in regression   **0 0 1 -1 0 0 0**
                                                     0 0 0 0 1 -1 0
- Effects of interest (for plotting)   **[zeros(4,2) eye(4)]**

The two leading zeros in the contrast indicate the main effect "group" and should be always set to "0" in the contrast, because we are rather interested in the interaction.

Defining F-contrasts

For Anovas (e.g. comparison between groups) the test for any difference between the groups can be defined with:

    eye(n)-1/n

where *n* is the number of columns of interest.

For interaction effects (e.g. any interaction between covariate and group) you need to add leading zeros to leave out the group effects:

    [zeros(n,m) eye(n)-1/n]

where m is the number of group effects.

For a regression design (e.g. any regression slope) you are not interested in differential effects as described above, but rather in any effect of a regression parameter:

    [zeros(n,m) eye(n)]

where *n* is the number of columns of interest (e.g. regression parameters) and m is the number of leading columns you need to exclude (i.e. effects of no interest such as the constant or any other covariates of no interest).

In order to plot any contrast estimate or fitted response, however, you have to define the effects of interest by using:

    [zeros(n,m) eye(n)]

where *n* is the number of columns of interest and m is the number of leading columns you need to exclude (i.e. effects of no interest such as the constant or any other covariates of no interest).

This F-contrast (effects of interest) is necessary to correctly plot contrast estimates or fitted responses. Thus, it's recommended to always define this contrast.

<u>**Getting Results**</u>
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 voxels[4])

---

[4] To determine the extent threshold *empirically* (instead of saying 100 voxels or 500 voxels, which is completely arbitrary), simply execute it first without specifying an extent threshold. This gives you an output (i.e., the standard SPM glass brain with significant effects). If you click on "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 enter this number when prompted for the "Extent Threshold". There is also a hidden option in "CAT12 → Data presentation → Threshold and transform spmT-maps" to define the extent threshold as p-value or to use the "Expected Number of Voxels per Cluster".

Special Cases

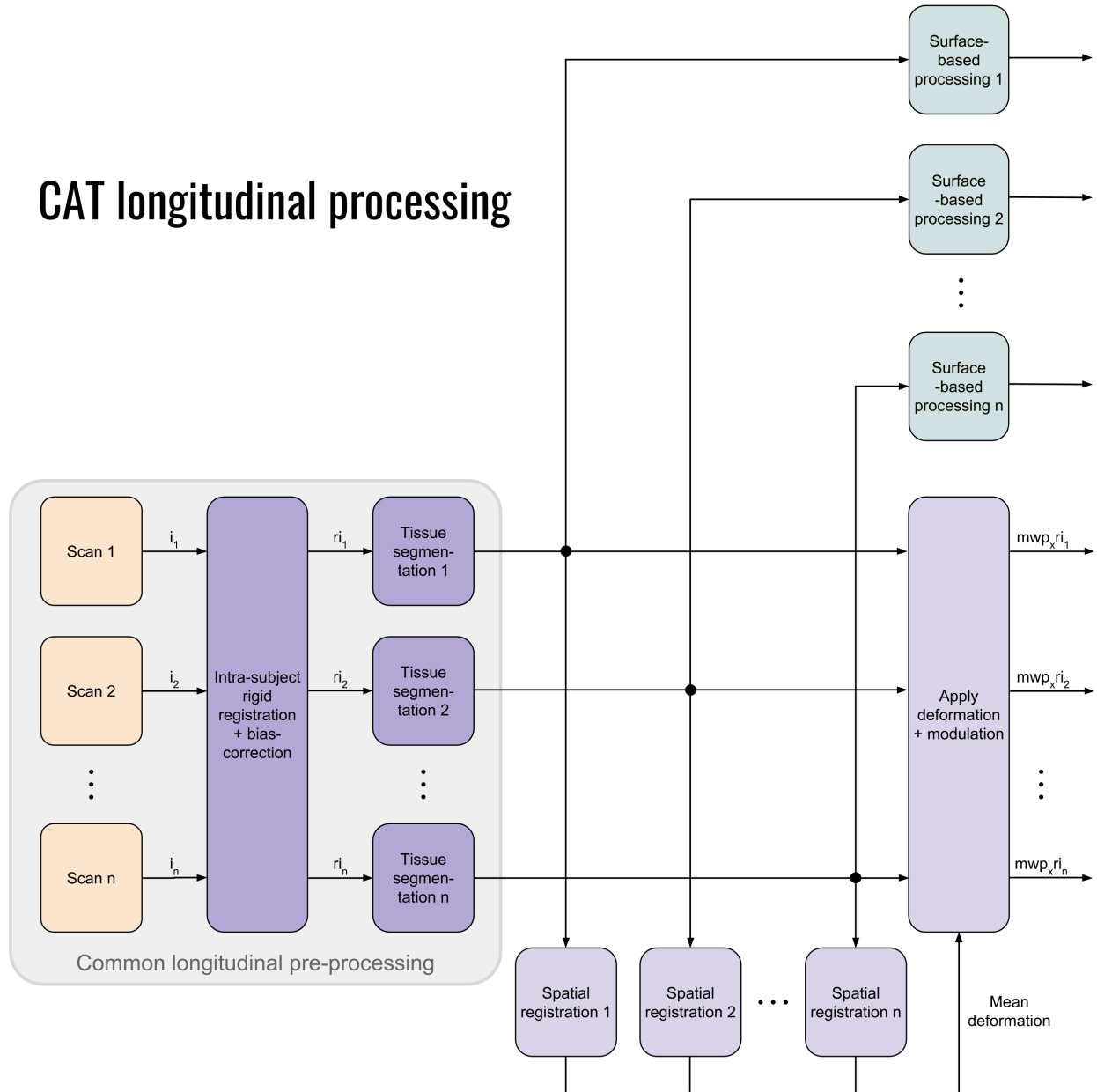## CAT12 for longitudinal data

BACKGROUND

The majority of morphometric studies are based on cross-sectional data in which one image is acquired for each subject. However, to map structural changes over time (e.g., in the framework of use-dependent plasticity, neurodevelopment, or aging), longitudinal designs that consider additional time points for each subject are necessary. The handling of these longitudinal data requires a customized processing, as a voxel- or point-wise comparability needs to be assured not only across subjects, but also across time points within subjects. This requires an inverse-consistent (or "symmetric") realignment as well as an intra-subject bias field correction (Ashburner and Ridgway, 2012; Reuter et al., 2010; Reuter and Fischl, 2011). Thus, CAT first realigns the images from all time points using inverse-consistent rigid-body registrations and applies intra-subject bias-field corrections. Subsequently, the resulting images are processed individually (see **Figure 4**, **5** and **7**) using the standard pipeline.

That is, images are tissue-classified and the necessary deformations for spatial registration to the standard template are estimated. However, instead of applying the individual spatial transformations (as done in a cross-sectional analysis), a mean transformation for all time points is calculated and applied to all individual images (Figure 7). The resulting normalized tissue segments can then be used for voxel-based analyses. Optionally, the processed images for all time points can be used to additionally extract the cortical surfaces and to calculate surface parameters, such as cortical thickness or diverse folding measures using the surface-based processing pipeline. Furthermore, the processing of the individual time points is initialized by the processed results from the (unbiased) average image. This reduces random variations in the processing procedure and improves the robustness and sensitivity of the overall longitudinal analysis. This latter processing stream of surface data is similar to the longitudinal approach employed in Freesurfer (Reuter et al., 2010; Reuter and Fischl, 2011).

Please note that two different longitudinal models are offered. The first longitudinal model is optimized for processing and capturing smaller changes over time in response to short-term plasticity effects (e.g. from learning and training). This model will probably not be as sensitive for larger longitudinal changes where large parts of the brain change over time (e.g. atrophy due to Alzheimer's disease or aging). This is due to the effect of estimating the parameters of spatial registration from the deformations of all time points and then applying them to all time points. If a large atrophy occurs between the time points, this can lead to a displacement of tissue boundaries and might result in areas with reduced volumes over time, which are surrounded by areas with increased volume due to these displacement problems. For data with larger volume changes over time, you should choose the longitudinal model, which is optimized to detect larger changes. This model also takes into account the deformations between time points and prevents the problems mentioned above.

Additional attention is then required for the setup of the statistical model. The following section therefore describes the data preprocessing and model setup for longitudinal data.

Figure 7: Flowchart for processing longitudinal data with CAT12 for detecting small changes (e.g. plasticity or learning effects). This figure shows the steps for processing longitudinal data. After an initial inverse-consistent rigid registration, which also includes bias-correction between time points, the realigned images of all time points (ri_x) are segmented into different tissue types (i.e. gray and white matter and CSF). In the next step, the registration parameters (i.e. deformation fields) are estimated using a non-linear spatial registration (Dartel or Geodesic Shooting) and then averaged. Finally, the resulting mean deformation is applied to the tissue segmentations of all time points and modulated with the Jacobian determinant of the deformation ($mwp_x ri_x$). Optionally, the tissue segmentations of all time points can be used to apply the surface-based processing for each time point separately to determine cortical thickness and other surface-related parameters.

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 spatial registration are calculated automatically. Preprocessed images are written as mwp1r* and mwp2r* for grey and white matter respectively. By selecting the optimized pipeline to detect larger changes an additional "mw" is prepended to the name due to the additional warping and modulation steps. To define the segmentation and registration parameters, the defaults in cat_defaults.m are used. Optionally surfaces can be extracted in the same way as in the cross-sectional pipeline and the realigned tissue segmentations are used for this step.

Choose of longitudinal model and comparison to SPM12 longitudinal registration

SPM12 also offers a batch for pairwise or serial longitudinal registration. In contrast to CAT12, this method relies on the deformations needed to non-linearly register the images of all time points to the mean image. Local volume changes can be calculated from these deformations, which are then multiplied (modulated) by the segmented mean image.

However, not only the underlying methods differ between SPM12 and the CAT12 longitudinal, but also the focus of possible applications. SPM12 additionally regularizes the deformations with regard to the time differences between the images and has its strengths more in finding larger effects over longer periods of time (e.g. ageing effects or atrophy). The use of deformations between the time points makes it possible to estimate and detect larger changes, while subtle effects over shorter periods of time in the range of weeks or a few months are more difficult to detect. In contrast, longitudinal preprocessing in CAT12 has been initially developed and optimized to detect more subtle effects over shorter periods of time (e.g. brain plasticity or training effects after a few weeks or even shorter times), but now also offers a longitudinal model that is optimized to detect larger changes. With these two different longitudinal models, CAT12 attempts to combine the advantages of both approaches, providing flexibility for a wide range of potential data.

Unfortunately, there is no clear recommendation for one method or the other method. As a rule of thumb, however, the following could apply: The shorter the time periods and the smaller the expected changes, the more likely the longitudinal model for small changes is the right choice. To find effects of ageing or atrophy due to a (neurodegenerative) disease after a longer period of time, I recommend the longitudinal model for large changes in CAT12.

Please note that surface-based preprocessing and ROI estimates are not affected by the selected longitudinal model, as the realigned images are used independently to generate cortical surfaces, thickness, or ROI estimates.

You can change the tissue probability map (TPM) via GUI or by changing the entry in the file cat_defaults.m (i.e. for children data). Any parameters that cannot be changed via the GUI must be set in the file cat_defaults.m:

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 enter "open cat_defaults.m" in your matlab command window. The file is opened in the editor. If you are not sure how to change the values, open the module "Segment Data" in the batch editor as a reference.

PREPROCESSING OF LONGITUDINAL DATA

CAT12 → Segment longitudinal data

Parameters:

- ○ Longitudinal Data
    - Select mode of longitudinal data selection for time points or subjects
    - In case of "timepoints" you can create multiple time points where each time point has to contain the same number and order of subjects.
    - If you have a varying number of time points for each subject you have to use the "subjects" mode where you have to define the files of each subject separately.
    - Select all data for each subject or time point. 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), choose a T1-weighted image.
    - Select "New: Subject" or "New: Timepoint" to add data for a new subject or time point
- ○ Longitudinal Model
    - Depending on your type of longitudinal data you can either select the pipeline that is optimized for detecting small changes (e.g. brain plasticity or training effects after a few weeks or even shorter periods of time) or large changes (e.g. ageing or developmental effects).

    → For all other options you can follow the instructions for preprocessing of cross-sectional data as described before. Please note that not all writing options are available for longitudinal data.

For the naming conventions of all written files see "Naming convention of output files". The GM segments are mwp1r* (or mwmwp1r* for the model with larger changes) , the WM segments are named mwp2r* (mwmwp2r*) if you have selected the option to modulate the data. Without modulation the leading "m" is omitted.

Statistical analysis of longitudinal data - overview

The main interest in longitudinal studies is in the 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 evaluate these questions is described in three examples. First, the case of only one group of 8 subjects with 2 time points each (e.g. normal aging) is shown. Then, the case of two groups of subjects with 4 time points is described. Here, we can test for a time by group interaction if we assume that the time points are acquired at the same times and are therefore comparable between the groups. If the time points are not acquired at the same times and are therefore not comparable between the groups, we can use the third design. This design models a time by group interaction using an additional covariate (e.g. age at scanning) and offers more flexibility.

These examples should cover most analyses – the number of time points / groups just needs to be adjusted. Unlike the analysis of cross-sectional data, as described above, we need to use the flexible factorial model, which takes into account that the time points for each subject are dependent data.

Please note that it's not strictly necessary to use TIV as a covariate (nuisance) because we are analyzing intra-individual effects that do not need to be corrected and TIV should not change over time.

LONGITUDINAL DATA IN ONE GROUP

This model supports the analysis between all time points for one group.

CAT12 → Statistical Analysis → Basic Models

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

    > SPM is internally handling some keyword factors such as "subject" or "repl". If you use "subject" as keyword for the first factor the conditions can be easier defined by only labeling the time points as input (see below).

    - Specify Subjects or all Scans & Factors → *"Subjects"* → *"New: Subject; New: Subject; New: Subject; New: Subject;"*

      Subject
        - Scans → *[select files (the smoothed GM data of the 1st Subject)]*

- Conditions → "1 2" [for two time points]

Subject

- Scans → *[select files (the smoothed GM data of the 2nd Subject)]*
- Conditions → "1 2" [for two time points]

Subject

- Scans → *[select files (the smoothed GM data of the 3rd Subject)]*
- Conditions → "1 2" [for two time points]

Subject

- Scans → *[select files (the smoothed GM data of the 4th Subject)]*
- Conditions → "1 2" [for two time points]

Subject

- Scans → *[select files (the smoothed GM data of the 5th Subject)]*
- Conditions → "1 2" [for two time points]

Subject

- Scans → *[select files (the smoothed GM data of the 6th Subject)]*
- Conditions → "1 2" [for two time points]

Subject

- Scans → *[select files (the smoothed GM data of the 7th Subject)]*
- Conditions → "1 2" [for two time points]

Subject

- Scans → *[select files (the smoothed GM data of the 8th Subject)]*
- Conditions → "1 2" [for two time points]

> It is also possible to deal with missing values for the data. For example, if you normally have three time points, but for some data only two time points for the baseline and the last scan are available, this can be defined by "1 3".

- Main effects & Interaction → *"New: Main effect"*

  Main effect
  - Factor number → 2

  Main effect
  - Factor number → 1

o Covariates

o Masking

- Threshold Masking → Absolute → *[specify a value (e.g. "0.1")]*
- Implicit Mask → Yes
- Explicit Mask → <None>

o Global Calculation → Omit

o Global Normalization

- Overall grand mean scaling → No

o Normalization → None

This model supports the analysis between all time points between two groups.

CAT12 → Statistical Analysis → Basic Models

Parameters:

- o Directory <-X → Select Files → *[select the working directory for your analysis]* → Done
- o Design → "Flexible Factorial"
    - ▪ Factors → *"New: Factor; New: Factor; New: Factor"*

      > SPM is internally handling some keyword factors such as "subject" or "repl". If you use "subject" as keyword for the first factor the conditions can be easier defined by only labeling the time points as input (see below).

      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;"*

      Subject
        - Scans → [select files (the smoothed GM data of the 1st 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 data of the 2nd Subject of first group)]

          > Do not forget the additional single quote! Otherwise you have to define the conditions as " [1 1; 1 2; 1 3; 1 4] " for the first group.

        - Conditions → " [1 1 1 1; 1 2 3 4]**'** " [*first group with four time points*]

      Subject
        - Scans → [select files (the smoothed GM data of the 3rd Subject of first group)]
        - Conditions → " [1 1 1 1; 1 2 3 4]**'** " [*first group with four time points*]

42

Subject

- Scans → [select files (the smoothed GM data of the 4th 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 data of the 1st Subject of second group)]
- Conditions → " [2 2 2 2; 1 2 3 4]' " [*second group with four time points*]

> It is also possible to deal with missing values for the data. For example, if you normally have three time points, but for some data only two time points for the baseline and the last scan are available, this can be defined by "1 3".

Subject

- Scans → [select files (the smoothed GM data 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 data 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 data of the 4th Subject of second group)]
- Conditions → " [2 2 2 2; 1 2 3 4]' " [*second group with four time points*]

- Main effects & Interaction → *"New: Interaction; New: Main effect"*

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

  Main effect
  - Factor number → 1

o Covariates

o Masking
  - Threshold Masking → Absolute → *[specify a value (e.g. "0.1")]*
  - Implicit Mask → Yes
  - Explicit Mask → <None>

o Global Calculation → Omit

o Global Normalization
  - Overall grand mean scaling → No

o Normalization → None

This model supports the analysis of linear increase or decreases over time that differ between two groups. In order to analyze non-linear time effects, a polynomial covariate can be used. Furthermore, this model can be also reduced to one group accordingly or you can use other parameters such as performance over time to analyze dependencies to that variable (e.g. linear increase w.r.t. to performance).

CAT12 → Statistical Analysis → Basic Models

Parameters:

- o Directory <-X → Select Files → *[select the working directory for your analysis]* → Done
- o Design → "Flexible Factorial"
  - ▪ Factors → *"New: Factor; New: Factor; New: Factor"*

    Factor

    SPM is internally handling some keyword factors such as "subject" or "repl". If you use "subject" as keyword for the first factor the conditions can be easier defined by only labeling the time points as input (see below).

    - 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;"*

    Subject

    - Scans → [select files (the smoothed GM data of the 1st Subject of first group)]

      Do not forget the additional single quote! Otherwise you have to define the conditions as " [1 1; 1 2; 1 3; 1 4] " for the first group.

    - Conditions → " [1 1 1 1; 1 1 1 1[5]]**'** " [*first group with four time points*]

    Subject

    - Scans → [select files (the smoothed GM data of the 2nd Subject of first group)]
    - Conditions → " [1 1 1 1; 1 1 1 1]**'** " [*first group with four time points*]

---

[5] Please note that, in contrast to the previous model, all time points are coded here with only one value. The differentiation between the time points is now made by the covariate.

Subject

- Scans → [select files (the smoothed GM data of the 3rd Subject of first group)]
- Conditions → " [1 1 1 1; 1 1 1 1]' " [*first group with four time points*]

Subject

- Scans → [select files (the smoothed GM data of the 4th Subject of first group)]
- Conditions → " [1 1 1 1; 1 1 1 1]' " [*first group with four time points*]

Subject

- Scans → [select files (the smoothed GM data of the 1st Subject of second group)]
- Conditions → " [2 2 2 2; 1 1 1 1]' " [*second group with four time points*]

Subject

- Scans → [select files (the smoothed GM data of the 2nd Subject of second group)]
- Conditions → " [2 2 2 2; 1 1 1 1]' " [*second group with four time points*]

Subject

- Scans → [select files (the smoothed GM data of the 3rd Subject of second group)
- Conditions → " [2 2 2 2; 1 1 1 1]' " [*second group with four time points*]

Subject

- Scans → [select files (the smoothed GM data of the 4th Subject of second group)]
- Conditions → " [2 2 2 2; 1 1 1 1]' " [*second group with four time points*]

> It is also possible to deal with missing values for the data. For example, if you normally have three time points, but for some data only two time points for the baseline and the last scan are available, this can be defined by "1 3".

- Main effects & Interaction → *"New: Interaction; New: Main effect"*

  Interaction

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

  Main effect

  - Factor number → 1

o Covariates → "New: Covariate"

- Covariate
  - Vector → *[enter the values for the time-points in the same order as the respective file names of the smoothed images]*
  - Name → [specify (e.g. "*time*")]
  - Interactions → With Factor 2
  - Centering → Overall mean

o Masking

- Threshold Masking → Absolute → *[specify a value (e.g. "0.1")]*
- Implicit Mask → Yes
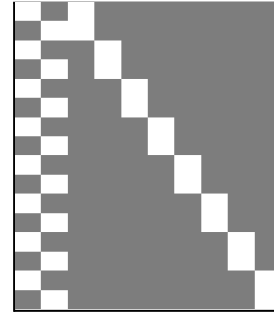- Explicit Mask → <None>

o Global Calculation → Omit

o Global Normalization

- Overall grand mean scaling → No

o Normalization → None

45

Contrasts:

Longitudinal data in one group (example for two time points)

T-test

- For Time 1 > Time 2   **1 -1**
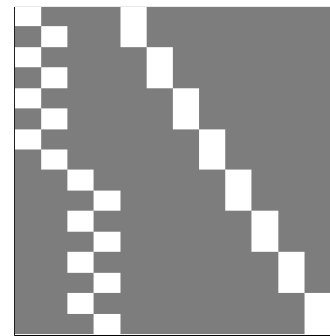- For Time 1 < Time 2   **-1 1**

F-test

- For any differences in Time (effects of interest)   **eye(2)-1/2**

Longitudinal data in two groups (example for two time points)

T-test

- For Time 1 > Time 2 in Group A   **1 -1 0 0**
- For Time 1 > Time 2 in Group B   **0 0 1 -1**
- For Time 1 > Time 2 in Groups A and B   **1 -1 1 -1**
- For Time 1 > Time 2 & Group A > Group B (interaction)   **1 -1 -1 1**

F-test

- For any differences in Time in Group A   **eye(k)-1/k**
- For any differences in Time in Group B   **[zeros(k) eye(k)-1/k]**
- For main effect Time   **[eye(k)-1/k eye(k)-1/k]**
- For interaction Time by Group   **[eye(k)-1/k 1/k-eye(k)]**
- For main effect Group[6]   **[ones(1,k)/k -ones(1,k)/k ones(1,n1)/n1 -ones(1,n2)/n2]**
- Effects of interest (for plotting)   **[eye(k)-1/k zeros(k); zeros(k) eye(k)-1/k]**

Here n1 and n2 are the number of subjects in Group A and B respectively and k is the number of time points.

---

[6] This contrast may artificially inflate your results. More information about that issue can be found here: https://doi.org/10.3389/fnins.2019.00352

Longitudinal data in two groups with interaction of covariate by group (example for two time points)
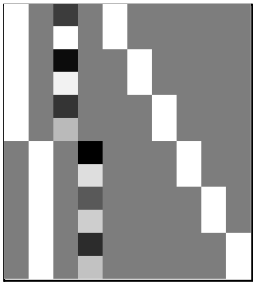
T-test

- For regression slope Group A > Group B   **0 0 1 -1**
- For regression slope Group A < Group B   **0 0 -1 1**

F-test

- For any differences in regression slope   **0 0 1 -1**
- Effects of interest (for plotting)   **0 0 1 0**

                                                 0 0 0 1

Adapting the CAT12 workflow for unusual populations

Background

For most analyses, CAT12 provides all necessary tools. Since the new segmentation algorithm is no longer dependent on Tissue Probability Maps (TPMs) and pre-defined DARTEL templates for healthy adults exist, most questions can be evaluated using the default Toolbox settings. However, the Toolbox settings may not be optimal for some special cases, such as analysis in children or certain patient groups. Below we present strategies for dealing with these specific cases.

Customized tissue probability maps - overview

For pediatric MRI data, it is a good idea to create customized TPMs, which reflect age and gender of the population. The TOM8 Toolbox (available via: http://dbm.neuro.uni-jena.de/software/tom) provides the necessary parameters to customize these TPMs. You can also try the CerebroMatic Toolbox by Marko Wilke that uses a more flexible template creation approach and additionally allows for the creation of customized DARTEL templates (https://www.medizin.uni-tuebingen.de/kinder/en/research/neuroimaging/software/?download=cerebromatic-toolbox). Please note that these TPMs should contain 6 classes: GM/WM/CSF and 3 background classes.

CUSTOMIZED TISSUE PROBABILITY MAPS

Within TOM8 Toolbox or Cerebromatic toolbox:

> → *create your own customized TPM based on the respective toolbox's instructions*

Implementation into CAT12:

CAT12 → Segment Data

- Parameters:
    - ○ Options for initial SPM12 affine registration
        - ▪ Tissue Probability Map (→ Select your customized TPMs here)

Overview

An individual (customized) DARTEL template can be created for all cases involving a representative number of subjects. This means that an average template of the study sample is created from the tissue segments of the grey matter and white matter tissue segments of all subjects.

Please note that the use of your own DARTEL template results in deviations and unreliable results for ROI-based estimations as the atlases are different. Therefore, all ROI outputs in CAT12 are disabled if you use your own customized DARTEL template.

Steps to create a customized DARTEL template based on your own subjects

Several steps are required to create normalized tissue segments with customized DARTEL Templates. These steps can be bundled using dependencies in the Batch Editor. The last step ("Normalise to MNI space") can be repeated with the customized templates if additional output files are required (e.g. you have only saved GM segmentations but are also interested in the WM segmentations or you would like to use a different smoothing size).

In the first step, the T1 images are segmented and the tissue segments are normalized to the Tissue Probability Maps by means of an affine transformation. Start by selecting the module "Segment Data".

CAT12 → Segment Data

Parameters:

→ for all options except "writing options" use the same settings as you would for a "standard" VBM analysis.

o Writing Options

- "Grey Matter"→ "Modulated normalized" → "No"
- "Grey Matter"→ "DARTEL export" → "affine"
- "White Matter"→ "Modulated normalized" → "No"
- "White Matter"→ "DARTEL export" → "affine"

These settings will generate the volumes "rp1*-affine.nii" and "rp2*-affine.nii", i.e. the grey (rp1) and white (rp2) matter segments after affine registration. The following modules can be selected directly in the batch editor (SPM → Tools → CAT12 → CAT12: Segment Data and SPM → Tools → DARTEL Tools → Run DARTEL (create Templates)). It makes sense to add and specify these modules together with the "Segment Data" module described above, within the Batch Editor and to set dependencies so that all processing takes place within one batch. To this effect, add the following two steps:

SPM → Tools → DARTEL Tools → Run DARTEL (create Templates)

Parameters:

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 as you wish*

SPM → Tools → DARTEL Tools → Normalise to MNI space

Parameters:

DARTEL Template → select the final created template with the ending "_6".

- Select according to → "Many Subjects"
  - Flow fields: → select the "u_*.nii" files or create a dependency.
    - Images: → New: Images → *select the* "rp1*-affine.nii" *files or create a dependency.*
    - Images: → New: Images → *select the* "rp2*-affine.nii" *files or create a dependency.*
- Preserve: → "Preserve Amount"
- Gaussian FWHM: → *use defaults or modify as you wish*

→ all other options: *use defaults or modify as you wish*

Please note, that a subsequent smoothing is not necessary before statistics if you have used the option "Normalise to MNI space" with a defined Gaussian FWHM.

Alternative approach

Instead of creating your own customized DARTEL template as described above you can also make use of the customized DARTEL templates that are created with the CerebroMatic Toolbox by Marko Wilke:

https://www.medizin.uni-tuebingen.de/kinder/en/research/neuroimaging/software/?download=cerebromatic-toolbox.

Initial or re-use of customized DARTEL-templates

Optionally you can also use either the externally-made or your newly created, customized DARTEL templates directly in the CAT12 Toolbox. This can be helpful if new data is age-wise close to the data used to create the DARTEL template (e.g. you have an ongoing study and some more new subjects are added to the sample). In this case, it is not absolutely necessary to always create a new template and you can simply use the previously created DARTEL template. Finally, an additional registration to MNI (ICBM) space will usually be applied:

SPM → Tools → DARTEL Tools → Run DARTEL (Population to ICBM Registration)

Parameters:

DARTEL Template → select the final created template with the ending "_6".

SPM → Util → Deformations

    Parameters:

    Composition → "New: Deformation Field"

- Deformation Field: → *select the* "y_*2mni.nii" *file from step above.*

      o Output → "New: Pushforward"

- Apply to → select all "Template" files with the ending "_0" to "_6".
- Output destination → Output directory → *select directory for saving files.*
- Field of View → Image Defined → *select final* "Template" *file with the ending* "_6".
- Preserve → Preserve Concentrations (no "modulation").

Finally, your customized DARTEL template can also be used as the default DARTEL template within

CAT12:

CAT12 → Segment Data

    Parameters:

      o Volumes <-X → Select the original T1 images (as in the first module "Segment Data" described above)

      o Extended Options for CAT12 segmentation → "Spatial normalization Template" → *Select normalized DARTEL Template "*wTemplate*_1.nii*"*

- For all other options use the same settings as in the first module, or modify as you wish.

      o Writing Options → *select the output files just like in any standard VBM analysis:*

Please note that now the output files have to be smoothed as usual with a Gaussian filter before statistical analyses. Also note that, while ROI analyses will not be disabled, their validity must be critically assessed depending on the DARTEL template used.

How to proceed

All the steps described above are only an adaption of the CAT12 Toolbox module "Segment Data". As always, it is a good idea to store the modules and perform a quality control. The modules "Display slices" and "Check sample" from the CAT12 Toolbox are helpful here.

Other variants of computational morphometry

Deformation-based morphometry (DBM)

Background

DBM is based on the application of non-linear registration procedures to spatially normalize one brain to another one. The simplest case of spatial normalization is to correct the orientation and size of the brain. In addition to these global changes, a non-linear normalization is necessary to minimize the remaining regional differences due to local deformations. If this local adaptation is possible, the deformations now provide information about the type and localization of the structural differences between the brains and can then be analyzed.

Differences between the two 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 analyzed statistically (Gaser et al., 2001). This approach is also known as tensor-based morphometry since the Jacobian determinant represents such a tensor.

A deformation-based analysis can be performed not only on the local volume changes, but also on the entire information of the deformations, which also includes the direction and strength of the local deformations (Gaser et al., 1999; 2001). Since each voxel contains three-dimensional information, a multivariate statistical test is required for analysis. For this type of analysis, a multivariate general linear model or Hotelling's T2 test is often used (Gaser et al., 1999; Thompson et al., 1997).

Additional Steps in CAT12

CAT12 → Segment Data

Parameters:

- o Writing options → Jacobian determinant → *Normalized* → Yes

  - To save the estimated volume changes change the writing option for the normalized Jacobian determinant to "yes".

Changes in statistical analysis

Follow the steps for the statistical analysis as described for VBM, select the smoothed Jacobian determinants (e.g. swj_*.nii) and change the following parameters:

CAT12 → Statistical Analysis → Basic Models

Parameters:

- o Covariates: **Don't use TIV as covariate**

- o Masking
    - Threshold Masking → **None**
    - Implicit Mask → Yes
    - Explicit Mask → **../spm12/tpm/mask_ICV.nii**

Surface-based morphometry (SBM)

Background

Surface-based morphometry has several advantages over the sole use of volumetric data. For example, it has been shown that using brain surface meshes for spatial registration increases the accuracy of brain registration compared to volume-based registration (Desai et al., 2005). Brain surface meshes also permit new forms of analyses, such as gyrification indices which measure surface complexity in 3D (Yotter et al., 2011b) or cortical thickness (Dahnke et al., 2013; Gaser et al., 2020). In addition, 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 made easily visible.

Local adaptive segmentation

Gray matter regions with high iron concentration, such as the motor cortex and occipital regions, often have an increased intensity leading to misclassifications. In addition to our adaptive MAP approach for partial volume segmentation, we use an approach that allows to adapt local intensity changes to deal with varying tissue contrasts (Dahnke et al., 2012a).

Cortical thickness and central surface estimation

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

Topological correction

To repair topological defects, we use a novel method based on spherical harmonics (Yotter et al., 2011a). First of all, the original MRI intensity values are used to select either a "fill" or "cut" operation for each topological defect. We modify the spherical map of the uncorrected brain surface mesh in such a way that certain triangles are preferred when searching for the bounding triangle during reparameterization. Subsequently, a low-pass filtered alternative reconstruction based on spherical harmonics is patched into the reconstructed surface in areas where defects were previously present.

Spherical mapping

A spherical map of a cortical surface is usually necessary to reparameterize the surface mesh into a common coordinate system to enable inter-subject analysis. We use a fast algorithm to reduce area distortion, which leads to 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 the deformation of a spherical grid.

Additional Steps in CAT12

CAT12 → Segment Data

      Parameters:
- Writing options → Surface and thickness estimation → Yes

Use projection-based thickness to estimate cortical thickness and to create the central cortical surface for the left and right hemisphere.

CAT12 → Surface Tools → Resample Surfaces

      Parameters:
- Surface Data <-X → Select the surface data (e.g. [lr]h.thickness.*)

- Smoothing Filter Size in FWHM *[use defaults or modify]*
  - 12-15mm kernels are widely used for SBM and I recommend to start with a value of 15mm for thickness data and 20mm for folding data (e.g. gyrification, complexity).
- Split job into separate processes

  - To use multi-threading the CAT12 segmentation job with multiple subjects can be split into separate processes that run in the background. You can even close Matlab, which will not affect the processes that will run in the background without GUI. If you don't not want to run processes in the background then set this value to 0.
  - Keep in mind that each process needs about 1.5..2GB of RAM, which should be considered to choose the right number of processes.
  - Please further note that no additional modules in the batch can be run except CAT12 segmentation. Any dependencies are broken for subsequent modules.

Extract optional surface parameters

You can also extract additional surface parameters that need to be resampled and smoothed with the above tool.

CAT12 → Surface Tools → Surface Measures → Extract Additional Surface Parameters

   Parameters:

   o   Central Surfaces <-X → Select the central surface data (e.g. [lr]h.central.*)

   o   Gyrification index

   - Extract gyrification index (GI) based on absolute mean curvature. The method is described in Luders et al., NeuroImage, 29: 1224-1230, 2006.

   o   Cortical complexity (fractal dimension)

Extract Cortical complexity (fractal dimension) which is described in Yotter et al., Neuroimage, 56(3): 961-973, 2011.

Warning: Estimation of cortical complexity is very slow!

   o   Sulcus depth

Extract sulcus depth based on the Euclidean distance between the central surface and its convex hull. In addition, sulcal depth can be transformed with sqrt-function to render the data more normally distributed, which is the recommended option for further statistical analysis.

Changes in statistical analysis

For statistical analysis of surface measures you can use the common 2nd level models, which are also used for VBM. Follow the steps for the statistical analysis as described for VBM and change the following parameters:

CAT12 → Statistical Analysis → Basic Models

   Parameters:

   o   Covariates: **Don't use TIV as covariate**

   o   Masking

      ▪   Threshold Masking → **None**
      ▪   Implicit Mask → Yes
      ▪   Explicit Mask → <none>

As input you have to select the resampled and smoothed files (see above). A thickness file of the merged left and right hemispheres that was resampled and smoothed with the default FWHM of 15mm is named as:

   S15.mesh.thickness.resampled.name.gii

where "name" if the original filename without extension. Please note that these files are saved in the surf-folder as default.

Do not use the "Estimate" function in the SPM window, but the corresponding function in CAT12. This allows you to automatically overlay the results on the Freesurfer average surface:

CAT12 → Statistical Analysis → Estimate Surface Models

Region of interest (ROI) analysis

CAT12 enables the estimation of mean tissue volumes (and additional surface parameters such as cortical thickness) for different volume and surface-based atlas maps (Gaser et al., 2016). All of these results are estimated in the native space before any spatial normalization and the mean value inside the ROI is estimated. The results for each dataset are stored as XML files in the *label* directory. The XML file *catROI[s]_*.xml* contains information of all atlases as data structure for one dataset and the optional "s" indicates surface atlases. You can use the CAT function "cat_io_xml" to read the XML data as a structure.

Additional steps for surface data

While ROI-based values for VBM (volume) data are automatically stored in the *label* folder as XML file it is necessary to extract these values additionally for surface data. This must be done after preprocessing the data and creating cortical surfaces. You can extract ROI-based values for cortical thickness but also for any other surface parameter that extracted with the "Extract Additional Surface Parameters" function:

CAT12 → ROI Data → Extract ROI-based surface values

> Parameters:
> - o (Left) Surface Data Files <-X → Select surface data files such as lh.thickness.* for all subjects

Statistical analysis of ROI data

Finally, the XML files of several subjects can be analyzed using an existing SPM design with *CAT12 → Analyze ROIs*. Here, the SPM.mat file is used to get information about all corresponding label files, but also about your design (including all covariates/confounds you have modeled). In this way, the same statistical analysis saved in the SPM.mat file is applied to your ROI data. You can then select a contrast, a threshold value and a measurement type to be analyzed (e.g. Vgm, Vwm, thickness, gyrification...) and choose between different atlas maps. The results are printed and saved as thresholded log-p volume or surface map:

logP*Threshold_NameOfContrast_NameOfAtlas_Measure*.nii

[lr]h.logP*Threshold_NameOfContrast_NameOfAtlas_Measure*.gii

These maps can be optionally visualized using *CAT12 → Display Results → Slice Overlay* for volume maps or *CAT12 → Display Results → Surface Overlay* for surface maps.

To analyze different measures (e.g. Vgm/Vwm for volumes or thickness/gyrification for surfaces) you can use any existing volume-based analysis to extract different volume measures, or any existing surface-based analysis to extract different surface measures. To give you an example: An existing SPM.mat file with a VBM analysis of GM allows you to analyze ROI measures for both, GM as well as WM. Therefore, it is not necessary to have a SPM.mat file of a VBM analysis of WM. The same applies to surface-based analysis. If you already have an SPM.mat for the analysis of cortical thickness, you can also estimate the ROI analysis for gyrification or fractal dimension. For surfaces, however, it is necessary to extract ROI-based measures for each subject using *CAT12 → ROI Data → Extract ROI-based surface values* prior to this step.

For ROI analysis of surfaces you can select the SPM.mat file of the analysis, either for the left or right hemisphere, or for the merged hemispheres. The design should be the same and the ROI results are always estimated for both hemispheres.

Please note that if you have moved your data after estimating your original voxel- or surfaced-based statistics, the required ROI files cannot be found.

Optional extraction of ROI data

Finally, the XML files of several subjects can be combined and saved as CSV file using the "Estimate Mean Values inside ROI" function for further analysis:

CAT12 → ROI Data → Estimate mean values inside ROI

> Parameters:
> - o   XML files <-X → Select xml files for each subject that are saved in the label folder.
> - o   Output file → Define output name for csv file. This name is extended by the atlas name and the name of the measure (e.g. "Vgm" for gray matter volume)

For each measurement (e.g. "Vgm" for gray matter volume) and each atlas a separate CSV file is written. This works for both volume and surface data, but volume and surface data has to be processed separately with this function (surface-based ROI values are indicated by an additional "s", e.g. ''catROIs_''). You can use external software such as Excel or SPSS to import the resulting CSV files for further analysis. Also pay attention to the different interpretations between "." and "," depending on your region and language settings on your system.

Use of atlas functions in SPM12

You can also use the volume-based atlases delivered with CAT as atlas maps with *SPM atlas functions*. This is particularly useful if you have used the default VBM processing pipeline, since the CAT12 atlas maps are then in the same DARTEL space as your data. If you have used the default VBM processing pipeline, it is recommended to use CAT12 atlases in the DARTEL space instead of the SPM Neuromorphometrics atlas. To use CAT12 atlases, you must call the *cat_install_atlases* function once , which copies the atlases to SPM. By default, only Neuromorphometics, LPBA40, and Hammers atlas maps are used, which are identified by a leading ''dartel_'' in the name.

Atlas maps for surfaces can be used with the function *CAT12 → Display Results → Surface Overlay*. You can use the data cursor function to display atlas regions under the cursor. In addition, you can use the ''Atlas labeling'' function to print a list of atlas regions of the resulting clusters or overlay the atlas borders on your rendered surface.

Additional information on native, normalized and modulated volumes

When preprocessing the images (see "First Module: Segment Data", on pages 5-6), the decision on the normalization parameters determines the interpretation of the analysis results. Please note that some of the output options are only available in the expert mode.

"**Native space**" creates tissue class images in spatial correspondence to the original data. Although this may be useful for estimating global tissue volumes (e.g. GM+WM+CSF=TIV) it is not suitable for performing VBM analyses due to the lack of voxel-wise correspondence in the brain). If one is interested in these global tissue volumes in the native space ("raw values"), it is not necessary to actually output the tissue class images in native space. The "Segment Data"-function automatically

generates an xml file for each subject (cat_*.xml), which contains the raw values for GM, WM, and CSF. The subject-specific values can be calculated (i.e., integrated into a single text file) using the function: CAT12 → Statistical Analysis → Estimate TIV.

"**Normalized**" creates 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**" creates tissue class images in alignment with the template, but multiplies ("modulates") the voxel values with the Jacobian determinant (i.e., linear and non-linear components) derived from the spatial normalization. This is useful for VBM analyses and allows the comparison of the *absolute amount of tissue* (e.g. "volume" of gray matter; Good et al., 2001; Neuroimage). Please note that in this type of modulation you must use TIV as covariate to correct different brain sizes. An alternative approach is to use global scaling with TIV if TIV is too strongly correlated with your covariates of interest. For more information, see the section "Checking for Design Orthogonality".

If you use the expert mode, you can optionally choose whether you want to modulate your data for non-linear terms only, which was the default for previous VBM-toolbox versions. This produces tissue class images in alignment with the template, but the voxel values are only multiplied by the non-linear components. This is useful for VBM analyses and allows comparison of the *absolute amount of tissue corrected for individual brain sizes*. This option is similar to using "Affine+non-linear" (see above) in combination with "global normalization" (when building the statistical model with the traditional PET design late one). Although this type of modulation was used in as default previous VBM-toolbox versions it is no longer recommended as the use of standard modulation in combination with TIV as covariate provides more reliable results (Malone et al., 2015; Neuroimage).

Naming convention of output files

Please note that the resulting files of CAT12 are organized in separate subfolders (e.g. mri, report, surf, label). If you don't want to use subfolders you can change the option "cat12.extopts.subfolders" in cat_default.m to "0".

Images (saved in subfolder "mri")

Segmented Images:	[m[0]w]p[0123]*[_affine].nii
Bias, noise and intensity corrected T1 image:	[w]m*.nii
Jacobian determinant:	wj_*.nii
Deformation field (inverse field):	y_*.nii (iy_*.nii)

Image space prefix:

m	modulated
m0	modulated non-linear only (expert mode only)
w	warped (spatially normalized using Dartel/Shooting)
r	realigned (for longitudinal data only)

Image space extension:

_affine	affine registered only

Image data prefix:

p	partial volume (PV) segmentation
0	PV label
1	GM
2	WM
3	CSF

Surfaces in native space (saved in subfolder "surf")

SURF.TYPE.*.gii

SURF left, right hemisphere [lh | rh]

TYPE surface data file [central | sphere | thickness | gyrification | ...]

   central - coordinates and faces of the central surface

   sphere - coordinates and faces of the spherical projection of the central surface

   sphere.reg - coordinates and faces of the sphere after spherical registration

   thickness - thickness values of the surface

   pbt - distance measure that is internally used for mapping

   depth - sulcal depth based on the euclidean distance between the central surface and its convex hull

   gyrification - gyrification values based on absolute mean curvature

   fractaldimension - fractal dimension values (cortical complexity)

Surfaces in (normalized) template space (after resampling and smoothing; saved in subfolder "surf" by default)

FWHM.SURF.TYPE.RESAMP.*.gii/dat

FWHM filtersize in FWHM after smoothing (e.g. s12)

SURF left, right, or both hemispheres [lh | rh | mesh]

TYPE surface data file [thickness | gyrification | fractaldimension | ...]

   thickness - thickness values of the surface

   depth - sulcal depth based on the euclidean distance between the central surface and its convex hull

   gyrification - gyrification values based on absolute mean curvature

   fractaldimension - fractal dimension values (cortical complexity)

RESAMP resampling to 164k (Freesurfer) or 32k (HCP) mesh [resampled | resampled_32k]

Please note that the resampled surfaces are stored in normalized space as gifti file and an external dat-file to increase processing speed and minimize the size of the resulting SPM.mat for statistical analysis. Both files are required for each surface dataset.

Images and surfaces of longitudinal data

After processing longitudinal data, the filenames additionally contain a "r" between the original filename and the other prefixes to indicate the additional registration step.

Furthermore, if you select the longitudinal model for detecting larger changes (e.g. ageing or developmental effects) the segmented images are additionally indicated by a leading "mw" due to the additional warping and modulation steps.

Example:

mwp1r*　　　- gray matter

mwmwp1r*　　- gray matter (longitudinal model for detecting larger changes)

Reports (saved in subfolder "report")

Global morphometric and image quality measures are stored in the *cat_*.xml* file. This file also contains other useful information about software versions and options used to preprocess the data. You can use the cat_io_xml function to read data from xml-files. In addition, a report for each data set is saved as pdf-file *catreport_*.pdf*.

Regions of interest (ROI) data (saved in subfolder "label")

ROI data is optionally saved as xml-file *catROI [s]_*.xml*. The optional "s" indicates surface atlases.

Calling CAT from the UNIX command line

You can also call CAT as a shell script from the Unix command line. This makes it possible to run the process completely in the background without graphical output and distributes all given data to parallel jobs. This can be useful if you want to run CAT on a computer cluster.

You can use the shell scripts cat_batch_cat.sh or cat_batch_long.sh (for longitudinal data). Please call the scripts for more information on how to define optional defaults files, for example.

There is also a script distribute_to_server.sh that can be used to distribute jobs to different servers.

Standalone versions (precompiled)

CAT12 is also distributed as a precompiled version, which requires only the Matlab Compiler Runtime, but no Matlab license. Please check the standalone folder where the corresponding scripts can be found. This might be especially useful if you want to run CAT on a computer cluster.

Technical information

This toolbox is an extension of the segmentation in SPM12, but uses a completely different segmentation approach.[7] See also the section 'Overview about CAT12 Processing' for a detailed overview about the processing steps.

Denoising

We also use two noise reduction methods to make data processing, and the tissue segmentation in particular, more robust against noise. The first method is a spatial-adaptive Non-Local Means (SANLM) denoising filter and removes noise while maintaining edges (Manjón et al., 2010) and is implemented as pre-processing step. The second method is a classical Markov Random Field (MRF) approach, which includes spatial information from adjacent voxels in the segmentation estimation (Rajapakse et al., 1997) and is part of the AMAP segmentation. The strength of the filters is automatically determined by estimating the residual noise in the image or can be set manually.

Interpolation

CAT12 uses an internal interpolation to provide more reliable results even with low resolution images and anisotropic spatial resolutions. Although interpolation cannot add more details to the images, some of the functions used benefit from the higher number of voxels and the usual strip artefacts in modulated images are greatly reduced.

Affine Preprocessing (APP)

To improve the initial SPM segmentation, an initial affine registration is applied to a bias-corrected image and the intensity range is limited to avoid problems in special protocols. If the preprocessing fails a more aggressive version is available that applies a rough bias correction and removes non-brain parts the brain before the initial affine registration.

Local Adaptive Segmentation (LAS)

GM intensity can vary for different regions such as the motor cortex, the basal ganglia, or the occipital lobe. These changes have an anatomical background (e.g. iron content, myelenization), but are dependent on the MR-protocol and often lead to GM-underestimations at higher intensities and CSF-overestimations at lower intensities. Therefore, a local intensity transformation of all tissue classes is used to reduce these effects in the image before the final AMAP segmentation. The strength of the changes is controlled by the LASstr parameter, with 0 for no LAS, small values (0.01-0.5) for small adjustments, 0.5 for medium adjustments (default), and higher values (0.5-1) for strong adjustments.

AMAP Segmentation

The segmentation approach is based on an Adaptive Maximum A Posterior (AMAP) technique without the need for *a priori* information on the tissue probabilities. This means that the Tissue Probability Maps (TPM) are not constantly used in the sense of the classical Unified Segmentation approach (Ashburner and Friston 2005), but only for spatial normalization, initial skull-stripping, and as initial segmentation estimate. The subsequent AMAP estimation is adaptive in the sense that local variations of the parameters (i.e., means and variance) are modelled as slowly varying

---

[7] 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.

spatial functions (Rajapakse et al., 1997). This accounts not only for intensity inhomogeneities, but also for other local intensity variations.

Partial Volume Segmentation

In addition, the segmentation approach uses a Partial Volume Estimation (PVE) with a simplified mixed model of a maximum of two tissue types (Tohka et al., 2004). We begin with an initial segmentation into three pure classes: gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) based on the AMAP estimation described above. The initial segmentation is followed by a PVE consisting of two additional mixed classes: GM-WM and GM-CSF. This results in an estimate of the amount (or fraction) of each pure tissue type that is present in each voxel (since single voxels - given their size - probably contain more than one tissue type) and thus allows for more precise segmentation.

Skull-Stripping

CAT12 provides different approaches for skull-stripping. Currently, the method with the most accurate and reliable results is an adaptive probability region-growing (APRG) method. It refines the probability maps of the SPM approach by region-growing techniques of the GCUT approach with a final surface-based optimization strategy.

The SPM approach works also quite stable for the majority of data. However, in some rare cases parts of GM (i.e. in frontal lobe) might be cut. If this happens the GCUT approach is a good alternative. GCUT is a graph-cut/region-growing approach starting from the WM area.

If your segmentations still contain skull and other non-brain tissue (e.g. dura) using the GCUT approach you can try to increase the strength in expert mode. The strength is defined with 0 for a more liberal and wider brain masks and 1 for a more aggressive skull-stripping. The default setting is 0.5 and has been successfully tested on a variety of different images. Thus, if parts of the brain are missing in the segmentations, the strength can be decreased.

Cleanup

CAT12 includes a new cleanup routine that uses morphological, distance and smoothing operations to remove the remaining meninges after the final segmentation. The strength of the cleanup is controlled by the cleanupstr parameter, with 0 for no cleanup, low values <0.5 for light cleanup, 0.5 for medium cleanup (default), and 1 for strong cleanup.

Spatial Normalization

Another important extension of the SPM12 segmentation is the integration of the DARTEL (Ashburner, 2007) and Geodesic Shooting (Ashburner and Friston, 2011) normalization into the toolbox by already existing DARTEL and Geodesic Shooting templates in MNI space. These templates were derived from 555 healthy control subjects of the IXI-database (http://www.brain-development.org) and are available in the MNI space[8] for six different iteration steps of the DARTEL and Geodesic Shooting normalization. Therefore, the creation of sample-specific DARTEL and Geodesic Shooting templates is no longer necessary for most studies [9].

---

[8] Therefore, no additional MNI normalization is necessary.

[9] For studies investigating children data, I still recommend creating a customized DARTEL or Geodesic Shooting template. However, it should be noted that this option requires a representative sample with a sufficient number of subjects.

CAT12 Citation

If you use CAT12, please cite the following references in your work:

CAT12 (in General)
Gaser C, Dahnke R, Kurth K, Luders E, Alzheimer's Disease Neuroimaging Initiative. A Computational Anatomy Toolbox for the Analysis of Structural MRI Data. Neuroimage, in review.

Surface Estimation
Dahnke R, Yotter RA, Gaser C. Cortical thickness and central surface estimation. Neuroimage. 2013 Jan 15;65:336-48.
Yotter RA, Dahnke R, Thompson PM, Gaser C. Topological correction of brain surface meshes using spherical harmonics. Hum Brain Mapp. 2011 Jul;32(7):1109-24.

Cortical Gyrification
Luders E, Thompson PM, Narr KL, Toga AW, Jancke L, Gaser C. A curvature-based approach to estimate local gyrification on the cortical surface. Neuroimage. 2006 Feb 15;29(4):1224-30.

Cortical Complexity (Fractal Dimension)
Yotter RA, Nenadic I, Ziegler G, Thompson PM, Gaser C. Local cortical surface complexity maps from spherical harmonic reconstructions. Neuroimage. 2011 Jun 1;56(3):961-73.

Gray Matter Asymmetry
Kurth F, Gaser C, Luders E. A 12-step user guide for analyzing voxel-wise gray matter asymmetries in statistical parametric mapping (SPM). Nat Protoc. 2015 Feb;10(2):293-304.

Anatomical Automatic Labeling Atlas (AAL)
Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, et al. Automated anatomical labelling of activations in spm using a macroscopic anatomical parcellation of the MNI MRI single subject brain. Neuroimage 2002; 15: 273-289.
Rolls ET, Huang C, Lin C, Feng J, Joliot M. Automated anatomical labelling atlas 3. Neuroimage 2020; 206:116189.

Anatomy Atlas
Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. NeuroImage 2005, 25(4), 1325-1335.

CoBra Atlas: Hippocampus
Winterburn JL, Pruessner JC, Chavez S, Schira MM, Lobaugh NJ, Voineskos AN, Chakravarty MM. A novel in vivo atlas of human hippocampal subfields using high-resolution 3 T magnetic resonance imaging. Neuroimage. 2013 Jul 1;74:254-65.

CoBra Atlas: Amygdala
Entis JJ, Doerga P, Barrett LF, Dickerson BC. A reliable protocol for the manual segmentation of the human amygdala and its subregions using ultra-high resolution MRI. Neuroimage. 2012 Apr 2;60(2):1226-35.

CoBra Atlas: Subcortical Areas
Tullo S, Devenyi GA, Patel R, Park MTM, Collins DL, Chakravarty MM. Warping an atlas derived from serial histology to 5 high-resolution MRIs. Sci Data. 2018 Jun 19;5:180107.

CoBra Atlas: Cerebellum
Park MT, Pipitone J, Baer LH, Winterburn JL, Shah Y, Chavez S, Schira MM,Lobaugh NJ, Lerch JP, Voineskos AN, Chakravarty MM. Derivation of high-resolution MRI atlases of the human cerebellum at 3T and segmentation using multiple automatically generated templates. Neuroimage. 2014 Jul 15;95:217-31.

LONI Probabilistic Brain Atlas (LPBA40)
Shattuck DW, Mirza M, Adisetiyo V, Hojatkashani C, Salamon G, Narr KL, Poldrack RA, Bilder RM, Toga AW. Construction of a 3D probabilistic atlas of human cortical structures. Neuroimage. 2008 Feb 1;39(3):1064-80.

Mori atlas
Oishi K, Faria A, Jiang H, Li X, Akhter K, Zhang J, Hsu JT, Miller MI, van Zijl PC, Albert M, Lyketsos CG, Woods R, Toga AW, Pike GB, Rosa-Neto P, Evans A, Mazziotta J, Mori S. Atlas-based whole brain white matter analysis using large deformation diffeomorphic metric mapping: application to normal elderly and Alzheimer's disease participants. Neuroimage. 2009 Jun;46(2):486-99.

Desikan-Killiany Atlas (DK40)
Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, Albert MS, Killiany RJ. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage. 2006 Jul 1;31(3):968-80.

Destrieux Atlas
Destrieux C, Fischl B, Dale A, Halgren E. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. Neuroimage. 2010 Oct 15;53(1):1-15.

Human Connectome Project (HCP) Multi-Modal Parcellation
Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E, Ugurbil K, Andersson J, Beckmann CF, Jenkinson M, Smith SM, Van Essen DC. A multi-modal parcellation of human cerebral cortex. Nature. 2016 Aug 11;536(7615):171-178.

Local-Global Intrinsic Functional Connectivity Parcellation
Schaefer A, Kong R, Gordon EM, Laumann TO, Zuo XN, Holmes AJ, Eickhoff SB, Yeo BTT. Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. Cereb Cortex. 2018 Sep 1;28(9):3095-3114.

References

J. Ashburner, K.J. Friston (2005): Unified segmentation. Neuroimage 26(3):839-51.

J. Ashburner (2007): A fast diffeomorphic image registration algorithm. Neuroimage 38(1):95-113.

J. Ashburner, K.J. Friston (2011): Diffeomorphic registration using geodesic shooting and Gauss–Newton optimisation. Neuroimage 55(3):954-967.

J. Ashburner, G.R. Ridgway (2012). Symmetric diffeomorphic modeling of longitudinal structural MRI. Front. Neurosci. 6: 197.

E. Luders, P.M. Thompson, K.L. Narr, A.W. Toga, L. Jancke, C. Gaser (2006): A curvature-based approach to estimate local gyrification on the cortical surface. Neuroimage, 29(4): 1224-30.

C. Gaser, H.-P. Volz, S. Kiebel, S. Riehemann, H. Sauer (1999): Detecting structural changes in whole brain based on nonlinear deformations – application to schizophrenia research. Neuroimage, 10:107-113.

C. Gaser, I. Nenadic, B. Buchsbaum, E. Hazlett, M. S. Buchsbaum (2001): Deformation-based morphometry and its relation to conventional volumetry of brain lateral ventricles in MRI. Neuroimage, 13:1140-1145.

C. Gaser, R. Dahnke (2016). CAT - A Computational Anatomy Toolbox for the Analysis of Structural MRI Data. HBM 2016. http://www.neuro.uni-jena.de/hbm2016/GaserHBM2016.pdf

C. Gaser, R. Dahnke, F. Kurth, E. Luders, Alzheimer's Disease Neuroimaging Initiative. A Computational Anatomy Toolbox for the Analysis of Structural MRI Data. Neuroimage, in review.

R. Dahnke, R. Yotter, C. Gaser (2012a). Cortical thickness and central surface estimation. Neuroimage, 65: 336-348.

R. Dahnke, G. Ziegler, C. Gaser (2012b). Local Adaptive Segmentation. HBM 2012. http://www.neuro.uni-jena.de/hbm2012/HBM2012-Dahnke02.pdf

R. Desai, E. Liebenthal, E.T. Possing, E. Waldron, J.R. Binder (2005): Volumetric vs. surface-based alignment for localization of auditory cortex activation. Neuroimage 26(4):1019-29.

C.D. Good, I.S. Johnsrude, J. Ashburner, R.N. Henson, K.J. Friston, R.S. Frackowiak (2001): A voxel-based morphometric study of ageing in 465 normal adult human brains. Neuroimage. 14(1 Pt 1):21-36.

I.B. Malone, K.K. Leung, S. Clegg, J. Barnes, J.L. Whitwell, J. Ashburner, N.C. Fox, G.R. Ridgway (2015): Accurate automatic estimation of total intracranial volume: a nuisance variable with less nuisance. Neuroimage 104:366-72.

J. Manjon, P Coupe, L. Marti-Bonmati, D.L. Collins, M. Robles (2010). Adaptive Non-Local Means Denoising of MR Images With Spatially Varying Noise Levels Journal of Magnetic Resonance Imaging, 31: 192-203.

J.C. Rajapakse, J.N. Giedd, J.L. Rapoport (1997): Statistical Approach to Segmentation of Single-Channel Cerebral MR Images. IEEE Trans. Med. Imag. 16(2):176-186.

M. Reuter, B. Fischl (2011). Avoiding asymmetry-induced bias in longitudinal image processing. NeuroImage 57: 19–21.

M. Reuter, H.D. Rosas, B. Fischl (2010). Highly accurate inverse consistent registration: a robust approach. NeuroImage 53: 1181–1196.

J. Tohka, A. Zijdenbos, A. Evans (2004): Fast and robust parameter estimation for statistical partial volume models in brain MRI. Neuroimage 23(1):84-97.

R.A. Yotter, R. Dahnke, P.M. Thompson, C. Gaser (2011a): Topological Correction of Brain Surface Meshes Using Spherical Harmonics. Human Brain Mapping, 32(7): 1109-24.

R.A. Yotter, G. Ziegler, I. Nenadic, P.M. Thompson, C. Gaser (2011b): Local cortical surface complexity maps from spherical harmonic reconstructions. Neuroimage, 56(3): 961-973.

R.A. Yotter, P.M. Thompson, C. Gaser (2011c): Algorithms to Improve the Re-Parameterization of Spherical Mappings of Brain. Journal of Neuroimaging, 21(2):e134-47.

R.A. Yotter, G. Ziegler, P.M. Thompson, C. Gaser (2011d): Diffeometric Anatomical Registration on the Surface. HBM 2011.

Christian Gaser christian.gaser@uni-jena.de

Florian Kurth   f.kurth@auckland.ac.nz