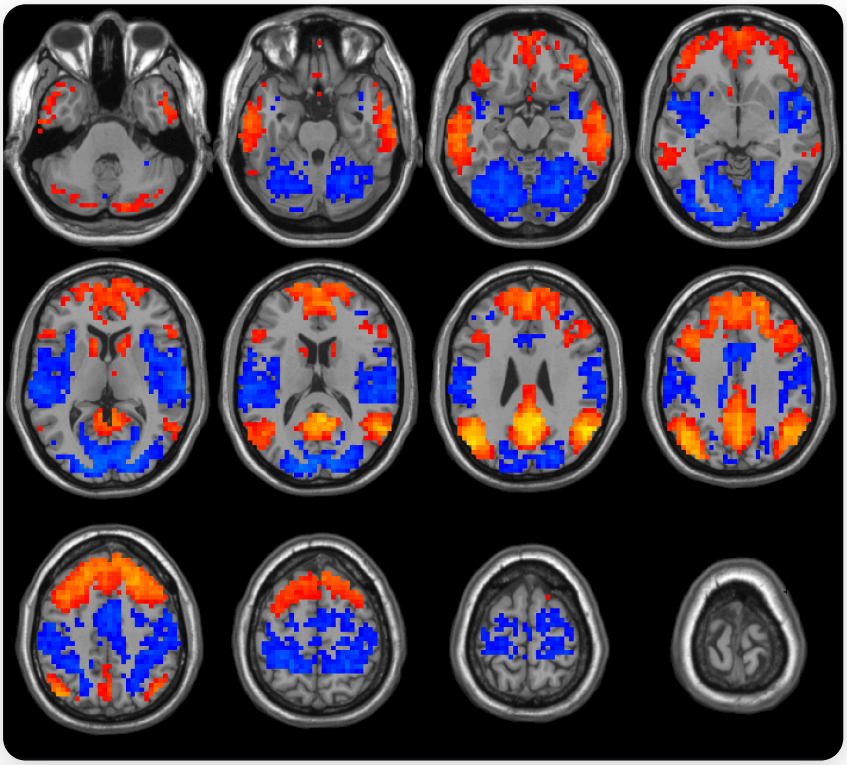
SeeCAT
Brain Functional Connectivity Analysis Toolbox

User Manual



Version 2.0 · November 2026

Resting-state fMRI from DICOM or BIDS to group statistics: preprocessing, quality control, seed connectivity, degree, gradients, multi-site harmonisation and statistics.

Mingrui Xia
National Key Laboratory of Cognitive Neuroscience and Learning
Beijing Key Laboratory of Brain Imaging and Connectomics
IDG/McGovern Institute for Brain Research
Beijing Normal University, Beijing, China

GPL-3.0-or-later · www.nitrc.org/projects/seecat

Contents

SeeCAT is a MATLAB toolbox for brain functional connectivity analysis of resting-state functional MRI. It takes a study from DICOM files or a BIDS dataset through preprocessing, quality control, connectivity measures and multi-site harmonisation to group statistics and figures, in one window or from scripts. This manual describes version 2.0: what each part does, how it does it, every option, the files it reads and writes, and worked examples on real data.

1 Introduction 4
- 1.1 What SeeCAT does 4
- 1.2 What is new in 2.0 4
- 1.3 Licence and citation 5

2 Installation and start 6
- 2.1 Requirements 6
- 2.2 Install 6
- 2.3 What is in the release 7
- 2.4 Two ways to work 7

3 Concepts 8
- 3.1 BIDS and BIDS-derivatives 8
- 3.2 Spaces and templates 8
- 3.3 Units, resuming and skipping 9
- 3.4 Logs, configurations and summaries 9
- 3.5 Parallel processing and memory 9
- 3.6 Quality-control exclusions 9
- 3.7 participants.tsv 10

4 The main window 11
- 4.1 Layout 11
- 4.2 Running 12
- 4.3 Kinds of fields 13

5 Workflow 14

6 The modules 15
- 6.1 DICOM to BIDS 15
- 6.2 Data check 17
- 6.3 Preprocessing 17
- 6.4 fMRIPrep denoising 22
- 6.5 Quality control 24
- 6.6 Seed FC 29
- 6.7 Degree 31
- 6.8 Gradients 34
- 6.9 Harmonize 38
- 6.10 Statistics 42
- 6.11 Viewer 47
- 6.12 Shared: denoising, censoring and smoothing 49
- 6.13 Shared: quality control 51

6.14 Shared: selection of runs 52
6.15 Shared: output and run 53

7 Worked examples 54
7.1 The example dataset 54
7.2 Example A: seed connectivity of the default-mode network 54
7.3 Example B: grey-matter volume and age (VBM) 57
7.4 Example C: degree centrality and gradients 58
7.5 Time for the whole study 60

8 Batch scripting 61
8.1 Functions 61
8.2 A complete study 61
8.3 Configuration files 63
8.4 Running on a server 63

9 Output reference 64
9.1 The derivatives folder 64
9.2 File names 64
9.3 Confounds table 65
9.4 JSON sidecars 65

10 Performance and common questions 66
10.1 Time and memory 66
10.2 Common questions 66

11 References 68

12 Appendix 70
12.1 Index of options 70
12.2 Coming from SeeCAT 1.x 73

1 Introduction

1.1 What SeeCAT does

SeeCAT covers the whole path of a resting-state study:

- **DICOM to BIDS**: converts DICOM folders in any common layout into a BIDS dataset with dcm2niix (bundled), finding subjects and sessions by itself.
- **Data check**: lists every BOLD run with its volumes, TR and head motion.
- **Preprocessing**: dummy-scan removal, slice timing, realignment, coregistration, unified segmentation, normalisation to MNI space, nuisance regression with band-pass filtering, handling of high-motion frames, smoothing; MRIQC image quality metrics, an HTML quality report per run and per study, and optional VBM maps.
- **fMRIPrep denoising**: the same denoising, censoring and smoothing applied to fMRIPrep outputs.
- **Quality control**: rules suggest which runs to exclude (motion, coverage, registration), image quality outliers are flagged per site, and a table lets you confirm each decision. Later steps leave excluded runs out.
- **Seed FC**: seed-based functional connectivity from spheres or masks.
- **Degree**: voxel-wise degree centrality and strength.
- **Gradients**: connectivity gradients by diffusion map embedding, aligned across runs.
- **Harmonize**: removes site and scanner effects from per-subject maps with ComBat and related models.
- **Statistics**: group GLM (one- and two-sample t tests, paired tests, one-way and repeated-measures ANOVA, regression, custom designs) with covariates, random field theory and FDR correction as in SPM, or TFCE with permutations.
- **Viewer**: slices, thresholds, cluster tables with atlas labels, mosaics, time series and surfaces.

Everything the window does is a configuration (a MATLAB struct, saved as JSON) passed to `seecat.run`, so every analysis can be repeated from a script.

1.2 What is new in 2.0

- A rewrite: SeeCAT needs only SPM, and works on BIDS data instead of the `FunImgARW…` folder layout of 1.x.
- Inputs are BIDS datasets and outputs follow BIDS-derivatives naming, so results can be read by other tools and other tools' results (fMRIPrep) can be read by SeeCAT.
- New: DICOM conversion, fMRIPrep input, MRIQC-style quality metrics, quality-control exclusions, connectivity gradients, multi-site harmonisation, TFCE, VBM maps, parallel processing that sizes itself to the computer's memory, and a new viewer.
- Steps resume after an interruption, and finished work with unchanged settings is skipped.

1.3 Licence and citation

SeeCAT is free software under the GNU General Public License, version 3 or later. It is developed by Mingrui Xia and colleagues at Beijing Normal University and distributed on GitHub and NITRC. When you publish results, please cite SeeCAT and the methods behind the steps you used (SPM, and the papers listed in References).

2 Installation and start

2.1 Requirements

Item	Needed	Notes
MATLAB	developed and tested on R2025a	Windows, macOS or Linux. Older releases are not tested; SeeCAT runs on them and notes this once in the Command Window. No other MathWorks toolbox is required.
SPM	SPM25, or SPM12 r7771 or later	On the MATLAB path. Used for slice timing, realignment, coregistration, segmentation, normalisation, smoothing and random field theory.
Parallel Computing Toolbox	optional	Processes runs at the same time; without it SeeCAT runs them one after another.
BrainNet Viewer 2.0	optional	Surface figures from the viewer.
dcm2niix	bundled	Windows, macOS and Linux builds are in the release.
Memory	8 GB or more	Degree and Gradients at fine resolution need more; see Performance.

2.2 Install

1. Download `SeeCAT-2.0.0.zip` from GitHub or NITRC and unzip it.
2. In MATLAB:

```
addpath('path/to/spm');
addpath('path/to/SeeCAT-2.0.0/SeeCAT');
savepath   % optional: keep the paths for later sessions
```

3. Open the main window:

```
SeeCAT
```

> **macOS.** If dcm2niix is blocked the first time, allow it in System Settings → Privacy & Security, or run `xattr -dr com.apple.quarantine SeeCAT-2.0.0/external`.

> **Linux.** The bundled dcm2niix needs glibc 2.29 or later (Ubuntu 20.04+, Debian 11+, RHEL/Rocky 9+). On older systems install dcm2niix (e.g. `pip install dcm2niix`) and put it on the PATH or set the environment variable `SEECAT_DCM2NIIX` to its path.

2.3 What is in the release

```
SeeCAT-2.0.0/
  SeeCAT/              the toolbox: add this folder to the path
    SeeCAT.m           opens the main window
    +seecat/           all functions (seecat.config, seecat.run, seecat.viewer, ...)
    templates/         brain, GM, WM, CSF masks and underlays (IXI549Space, MNI 2009c)
    resources/         icons and the window's web parts
  external/dcm2niix/   dcm2niix for win, mac, lnx
  Doc/manual/          this manual (HTML and PDF)
  README.md, LICENSE, CHANGELOG.md
```

2.4 Two ways to work

The window (`SeeCAT`) has a page per step. Each page builds a configuration; Run starts it in a separate MATLAB process so the window stays responsive, and Save writes the configuration as a JSON file.

Scripts use the same configurations:

```
cfg = seecat.config('prep');                % every option at its default
cfg.bids_dir  = 'D:/study/bids';
cfg.subjects  = {'01', '02'};                % empty = all
cfg.n_workers = 'auto';
seecat.run(cfg);                            % writes D:/study/bids/derivatives/seecat

seecat.run('my_prep.json');                 % a configuration saved from the window
```

Chapter 8 lists the batch functions and complete scripts.

3 Concepts

3.1 BIDS and BIDS-derivatives

SeeCAT reads raw data organised in the Brain Imaging Data Structure and writes its results as BIDS-derivatives. A raw dataset looks like this:

```
bids/
  dataset_description.json
  participants.tsv                 participant_id, and columns such as group, age, sex, site
  sub-01/
    anat/sub-01_T1w.nii.gz (+ .json)
    func/sub-01_task-rest_bold.nii.gz
         sub-01_task-rest_bold.json  RepetitionTime, SliceTiming, ...
```

File names are made of **entities**, key-value pairs separated by underscores, and end with a **suffix** (`bold`, `T1w`, `mask`, `connectivity`, ...). SeeCAT's outputs keep the entities of their source and add their own:

Entity	Meaning	Example
`sub`, `ses`, `task`, `acq`, `run`	subject, session, task, acquisition, run of the source	`sub-01_ses-1_task-rest_run-2`
`space`, `res`	template space and resolution label	`space-IXI549Space_res-3`
`desc`	which variant of the data	`desc-preproc`, `desc-denoised`, `desc-denoisedSm6`, `desc-PCC`
`stat`	which statistic a map holds	`stat-r`, `stat-z`, `stat-t`, `stat-p`
`label`	tissue of a probability map	`label-GM_probseg`

Every image comes with a JSON sidecar that records how it was made (sources, settings, frames used).

3.2 Spaces and templates

SeeCAT's own preprocessing normalises to SPM's template space, `IXI549Space` (the space of SPM's tissue probability maps, close to MNI152). Outputs are written at the chosen voxel size (3 mm by default) over SPM's bounding box `[-90 -126 -72; 90 90 108]`. fMRIPrep outputs stay in their space, usually `MNI152NLin2009cAsym`. The `templates` folder holds brain, GM, WM and CSF masks at 2 and 3 mm, made from SPM's TPMs, and underlays for the viewer (ch2, ch2better, ICBM152 2009c).

3.3 Units, resuming and skipping

Each pipeline splits its work into **units**: one BOLD run for Preprocessing, fMRIPrep denoising, Seed FC and Degree; one unit for a group step (Gradients, Harmonize, Statistics). When a unit finishes, SeeCAT records its settings. Running again:

- skips a unit that finished with the same settings (Overwrite runs it again);
- runs a unit again when its settings changed, after deleting the files it wrote before;
- resumes a unit that failed at the step where it stopped, from its work folder;
- never leaves a failed unit's partial outputs among the derivatives.

A failed unit does not stop the others; the summary table and the logs say what failed and why.

3.4 Logs, configurations and summaries

Every run writes, under `<output>/logs`: the configuration as JSON (`<pipeline>_<time>_config.json`), a main log, one log per unit, and a list of every file each unit wrote. `<output>/reports/<pipeline>_<time>_summary.tsv` has one row per unit: status, time, message and the unit's key results (frames kept, mean FD, ...).

3.5 Parallel processing and memory

The **Parallel workers** option of each page sets how many units run at the same time:

- `auto` (default): as many as fit in memory. SeeCAT estimates the memory one unit needs (from the size of the largest run, the brain mask, the node count of Gradients, the maps of Statistics), adds about 1 GB per worker, keeps what the main MATLAB needs, and uses at most 85 % of the free memory, one worker per core and one per unit. The log states the estimate and the choice. With only one unit (Harmonize, Statistics without TFCE), with too little memory for two workers, or without the Parallel Computing Toolbox, it runs one after another without a pool;
- `0`: one after another, e.g. to leave the computer free for other work;
- a number: at most that many workers.

Each worker is given its share of the computer's cores as threads (e.g. 5 workers on 32 cores get 6 threads each), so matrix work keeps its speed. For Gradients and TFCE, which are one unit, the workers share the runs or the permutations of that unit. If a pool fails to start, SeeCAT tries once more with half the workers and only then runs one after another, and the log says why.

3.6 Quality-control exclusions

After Preprocessing or fMRIPrep denoising, the group report writes `reports/qc_exclusions.tsv`, a list of every run (and T1w image) with a decision: keep or exclude. Rules suggest exclusions; on the **QC reports** page you confirm or change them. Seed FC, Degree, Gradients, Harmonize and Statistics then leave excluded runs out (their option **Leave out QC exclusions**, on by default) and list them in the log. Chapter 6.5 describes the rules.

3.7 participants.tsv

Group information comes from `participants.tsv` of the BIDS dataset: the group of each subject, age, sex, site, scores. Statistics and Harmonize read it next to the derivatives, two folders up (`bids/derivatives/seecat` → `bids`), or from the dataset named in the derivatives' `dataset_description.json`; a different file can be chosen. Columns with numbers are used as numbers, text columns as categories. Tissue volumes from Preprocessing (`gm_ml`, `wm_ml`, `csf_ml`, `tiv_ml`, in `reports/anat_volumes.tsv`) can be used as if they were columns of participants.tsv.

4 The main window

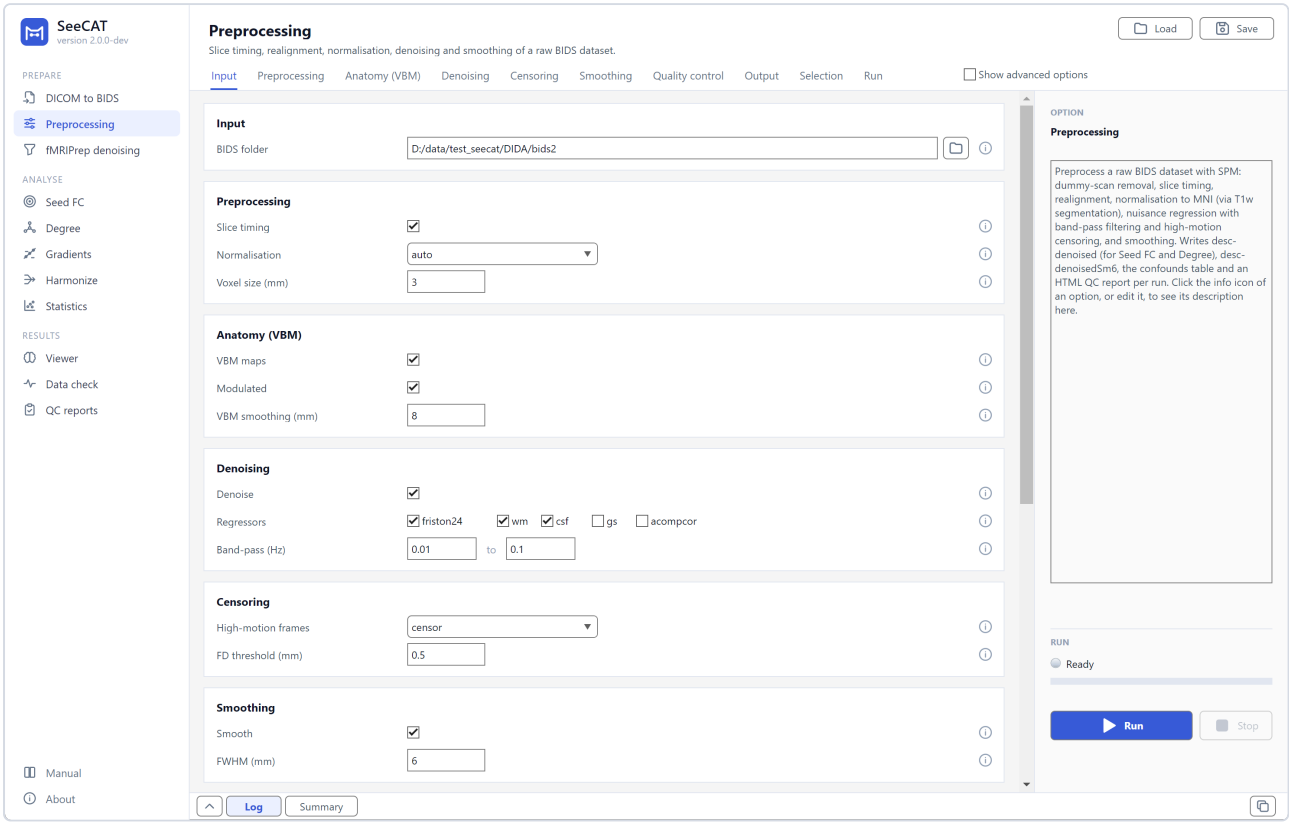


Figure 1. The main window on the Preprocessing page. Left: the pages. Centre: the options in cards, with tabs that jump to each card. Right: the description of the option in use, and the Run panel. Bottom: the log and summary drawer.

4.1 Layout

- **Pages** (left) in three groups: PREPARE (DICOM to BIDS, Preprocessing, fMRIPrep denoising), ANALYSE (Seed FC, Degree, Gradients, Harmonize, Statistics) and RESULTS (Viewer, Data check, QC reports). Manual opens this document, About shows the version.
- **Options** (centre), in cards such as Input, Preprocessing, Denoising. The tabs above jump to a card. **Show advanced options** adds the options most studies keep at their defaults.
- **Description** (right): click the **i** icon of an option, or edit it, to read what it does, with examples, its default and its batch field name. Hovering over an option shows a short tooltip.
- **Run panel** (right, bottom): Run, Stop, the progress of the units, and the state (Ready, Running, Finished, Failed).
- **Drawer** (bottom): the Log and the Summary table of the last run; the arrow opens and closes it.
- **Load** and **Save** (top right): read or write the page's configuration as JSON. A configuration saved here runs unchanged with `seecat.run`.

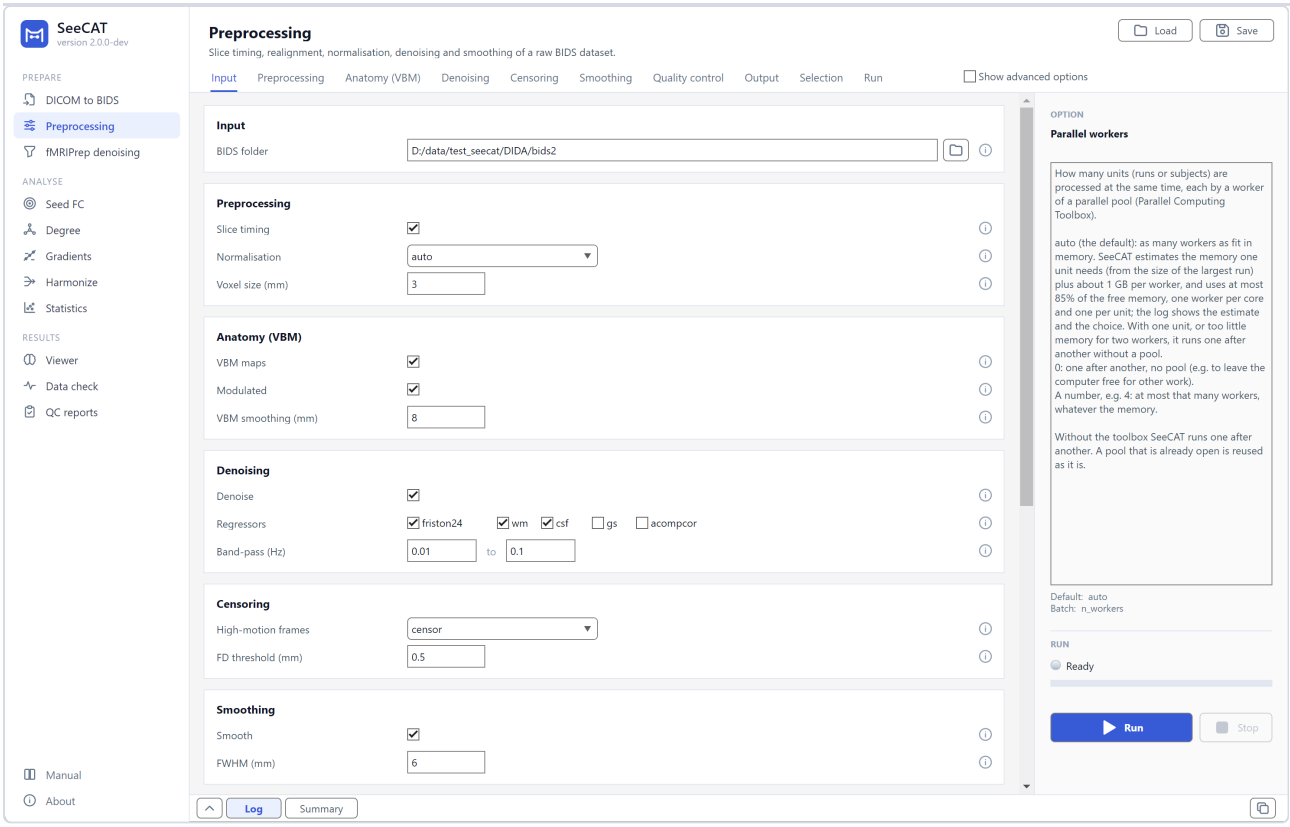


Figure 2. The description column: what the option does, examples, the default, and the batch field (here Parallel workers).

4.2 Running

Run checks every option first and lists all problems at once (an empty required folder, a value out of range). The pipeline then runs in a separate MATLAB process: the window stays usable, the log streams into the drawer, and Stop ends the process (units that finished keep their outputs; the unit that was interrupted resumes next time). When the run ends, the Summary tab of the drawer shows the run's table (one row per unit, as in `reports/<pipeline>_<time>_summary.tsv`); each unit's own log is in `logs/<pipeline>_<time>/`.

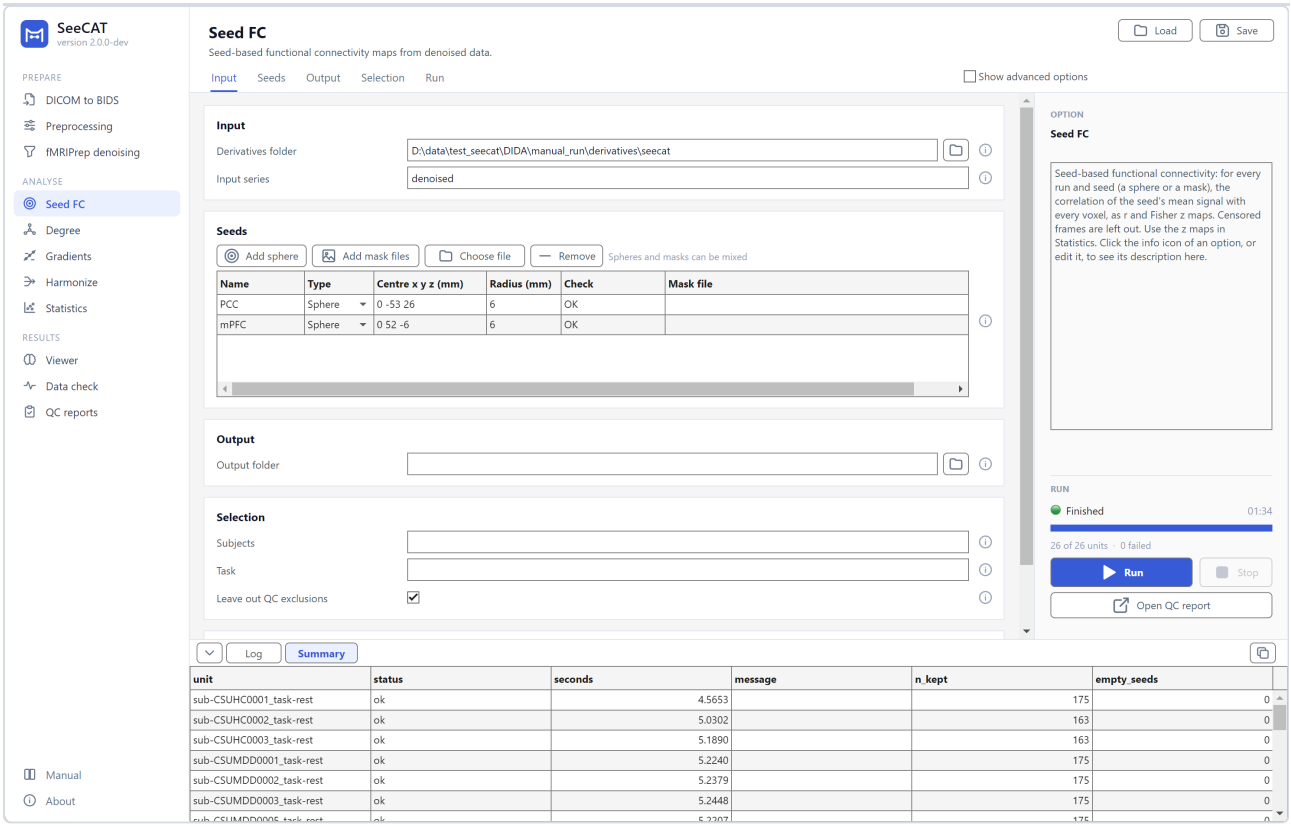


Figure 3. After a run: the summary table with one row per unit and the Run panel reporting the result.

4.3 Kinds of fields

Field	How to fill it
Folder, file	Type a path or use the folder button.
Number	One number, or several separated by spaces (e.g. `6 6 8`). Some accept a word: `auto`.
Range	Two numbers (e.g. band-pass `0.01` to `0.1`); both empty switch the step off.
Labels	A comma-separated list, e.g. subjects `01, 02, 15`.
Ticks and menus	On/off options and choices.
Mask	Grey matter, Whole brain, or Mask file with a file button.
Number with On	A value that can be switched off, e.g. smoothing of degree maps.
Maps	Fixed fields (Map, Statistic, Suffix, Space, Task, Other) that choose per-subject maps by their file names; the lists offer the values found in the folder and a line counts the matching maps.
Tables	Seeds (Seed FC) and contrasts (Statistics), with buttons to add and remove rows.

5 Workflow

DICOM → DICOM to BIDS → BIDS dataset → Data check → Preprocessing → QC reports → Seed FC / Degree / Gradients → Harmonize (multi-site) → Statistics → Viewer

fMRIPrep users start at **fMRIPrep denoising** instead of Preprocessing; everything after it is the same. The table shows what each step reads and writes:

Step	Reads	Writes
DICOM to BIDS	DICOM folders	BIDS dataset: `sub-*/anat`, `sub-*/func`, participants.tsv
Data check	BIDS dataset or derivatives	nothing (a table on screen)
Preprocessing	BIDS dataset	`derivatives/seecat`: preprocessed, denoised and smoothed BOLD, masks, confounds, T1w tissue maps, VBM maps, QC reports, exclusion list
fMRIPrep denoising	`derivatives/fmriprep`	`derivatives/seecat`, as above without the anatomical outputs
QC reports	`reports/qc_exclusions.tsv`	the same list with your decisions
Seed FC, Degree, Gradients	denoised BOLD in `derivatives/seecat`	per-run maps next to the runs; Gradients also `group/gradients…`
Harmonize	per-subject maps, participants.tsv	harmonised maps next to them (`desc-…ComBat`)
Statistics	per-subject maps, participants.tsv	`group/<name>/` : statistic maps, thresholded maps, cluster tables
Viewer	any NIfTI	figures, thresholded maps

By default every step writes into the same `derivatives/seecat` folder, so each next step finds its input there: set the derivatives folder once and leave Output folder empty.

6 The modules

Each section below describes one page: what it is for, how it works, every option (the tables are generated from the software, so they show exactly what the window offers), what it reads and writes, and how to run it from a script. Options marked advanced appear after ticking **Show advanced options**. Options shared by several pages (denoising, censoring, quality control, output, selection, run) are described once, in sections 6.12 to 6.15.

6.1 DICOM to BIDS

Converts DICOM files into a BIDS dataset with dcm2niix. It reads the DICOM headers to find subjects, visits and series, names each series by rules, and writes the images with their JSON sidecars into the BIDS layout.

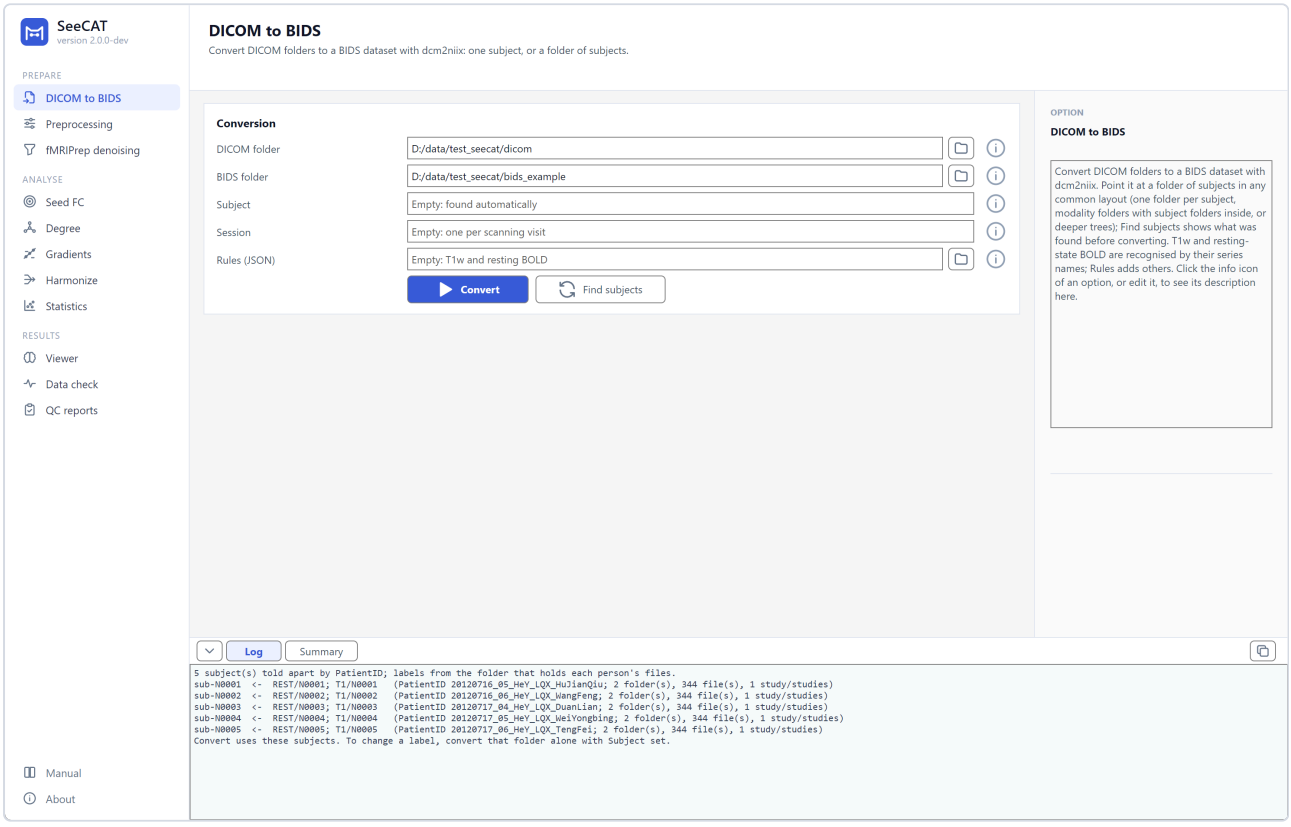


Figure 4. DICOM to BIDS after Find subjects: the layout it recognised and one line per subject.

How subjects, sessions and series are found

1. **Reading headers.** SeeCAT reads the first and last file of every folder that holds DICOM files (a fast reader of the header up to the series fields; other files, such as .txt or .nii, are ignored). A folder whose first and last files belong to different people is reported as an error.

2. **Subjects.** Folders with the same PatientID belong to one subject. The subject is named after the folder that holds all of that person's folders and no one else's, keeping letters and digits (`S_001` → `sub-S001`). When IDs are missing or anonymised (e.g. `anonymous` , `xxxx` , `0000`), the folder tree decides. Layouts that work include one folder per subject (`S001/…`), modality folders with subject folders inside (`T1/S001` , `REST/S001`) and deeper trees (`S001/visit1/T1`). **Find subjects** shows the layout it recognised and the subjects before anything is converted.
3. **Sessions.** A subject scanned in one visit (one DICOM study) gets no session folder. Several visits become `ses-1` , `ses-2` , … by scan date; the assignment is kept in `code/seecat/from_dicom_sessions.tsv` so a later visit gets the next number.
4. **Series.** dcm2niix converts each subject; every series is named by the first rule that matches its sidecar. The default rules take a 3D T1-weighted image (MPRAGE, SPGR, BRAVO, …) as `anat/…_T1w` and resting-state BOLD (series names containing rest, rsfMRI, or RS as a word) as `func/…_task-rest_bold` . A resting run must also be a 2D EPI with at least 20 volumes. Series that match no rule are listed, not copied.
5. **Writing.** Images go to `sub-<label>/[ses-<label>/]anat|func` as .nii.gz with their sidecars; `dataset_description.json` is written if missing and participants.tsv gets the new subjects. Images already present are kept, so converting again only adds what is missing. A table of every series and what happened to it is written to `code/seecat/` .

Options

Option	Batch	Default	Meaning
DICOM folder	`from_dicom(dicom_dir, bids_dir)`	(required)	The folder with the DICOM files, in any of the layouts above: one subject, or a folder of subjects.
BIDS folder	`bids_dir`	(required)	The BIDS dataset to write into; created if missing.
Subject	`'sub', '01'`	empty	Given: everything in the DICOM folder is this subject. Empty: subjects are found as described above.
Session	`'ses', 'pre'`	empty	Given: every image goes into this session. Empty: no session for one visit, `ses-1` , `ses-2` , … for several.
Rules (JSON)	`'rules', 'my_rules.json'`	empty	A JSON list of naming rules; each matches sidecar fields with regular expressions (case-insensitive) and gives datatype, suffix and entities. Empty: T1w and resting BOLD.

A rules file that also keeps a task run and a field map:

```
[{"match": {"SeriesDescription": "rest"}, "datatype": "func", "suffix": "bold", "entities":
{"task": "rest"}},
 {"match": {"SeriesDescription": "nback"}, "datatype": "func", "suffix": "bold", "entities":
{"task": "nback"}},
 {"match": {"SeriesDescription": "mprage|t1"}, "datatype": "anat", "suffix": "T1w"}]
```

Script

```
seecat.bids.scan_dicom('D:/study/dicom')                    % what would be converted
S = seecat.bids.from_dicom('D:/study/dicom', 'D:/study/bids');   % convert every subject
S = seecat.bids.from_dicom('D:/study/dicom/S_001', 'D:/study/bids', 'sub', '01', 'ses', 'pre');
```

6.2 Data check

Before preprocessing, look at what the dataset holds. Data check lists every BOLD run of a BIDS dataset (or of a derivatives folder) with its number of volumes and TR; where a confounds table exists (after preprocessing or fMRIPrep), also the mean and largest framewise displacement and the frames flagged for motion. Runs with mean FD above 0.25 mm or more than 20 % of frames flagged are highlighted. **Open in viewer** steps through the volumes of the selected run with its motion trace; **QC report** opens its report.

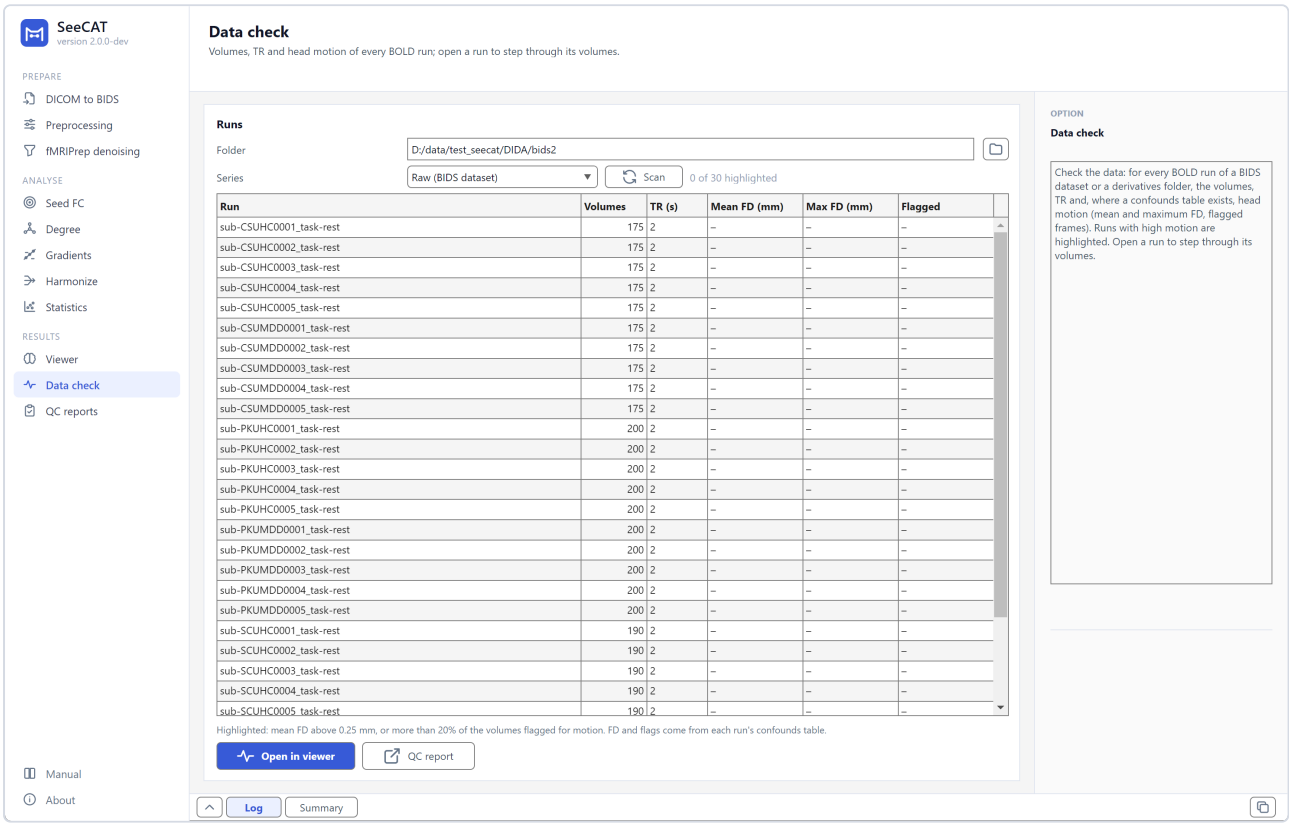


Figure 5. Data check on the example dataset: 30 runs from three sites with their volumes and TR.

The **Series** menu chooses which images are listed: preprocessed (`desc-preproc`), denoised, raw (the BIDS dataset) or all BOLD images. Script: `R = seecat.qc.runs('D:/study/bids', 'raw')`.

6.3 Preprocessing

Preprocesses every BOLD run of a BIDS dataset with SPM and writes analysis-ready data to `derivatives/seecat`. Each run is a unit; its T1w (from the same session if there is one) is used for registration and segmentation.

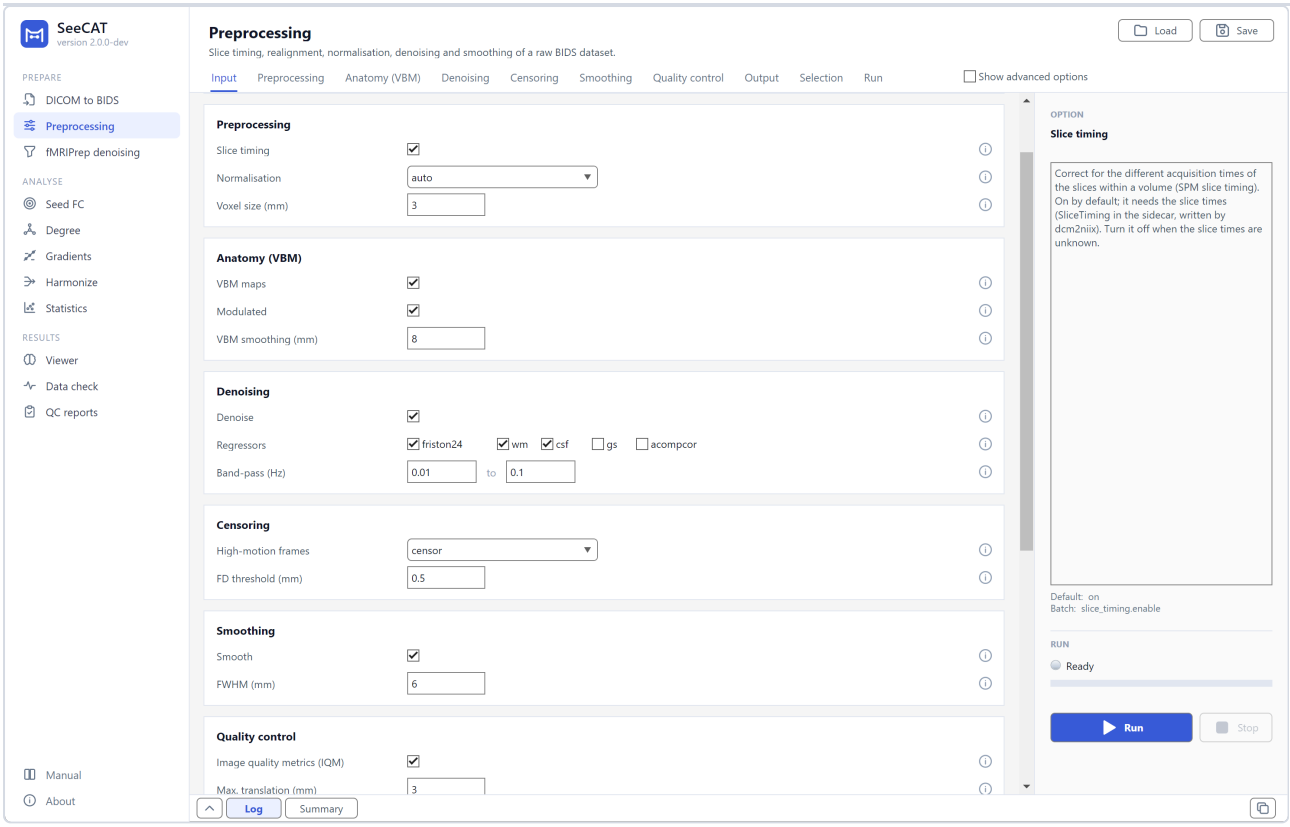


Figure 6. The Preprocessing page of the example: input, preprocessing, VBM, denoising and quality-control cards.

Steps

1. **Input.** The run and its T1w are copied into the work folder. Leading non-steady-state volumes are removed: with `dummy_scans = auto`, leading volumes whose mean brain intensity is a robust outlier ($z > 3.5$ against the median and MAD of the run, within the first 20) are dropped; volumes the scanner already discarded are not counted again.
2. **Slice timing** (SPM). The slice acquisition times come from `SliceTiming` in the sidecar (seconds), so sequential, interleaved and multiband orders are all handled; the reference is the middle of the TR unless set. The times are read along `SliceEncodingDirection` (`k` when absent; `k-` is reversed to voxel order). Slices along the first or second voxel axis (`i`, `j`) are not supported and stop the run with an error; reorient such images, or turn slice timing off.
3. **Realignment** (SPM). Two passes, to the first volume and then to the mean. The six motion parameters go to the confounds table; the data are not resampled here.
4. **Coregistration.** The T1w is registered to the mean EPI by normalised mutual information (only its header changes).
5. **Segmentation.** SPM unified segmentation of the T1w with SPM's tissue probability maps gives native GM, WM and CSF maps, a bias-corrected T1w and the deformation to IXI549Space. Without a T1w (`normalise.method = epi`) the mean EPI is segmented instead.
6. **Nuisance masks.** WM ($p > 0.99$) and CSF ($p > 0.95$) masks in the subject's own space, eroded by a 2 mm ball, so they keep away from grey matter; they are resampled to the output grid.

7. **Normalisation.** The deformation is applied once to the BOLD data (4th-degree B-spline; the only resampling of the data) and to the tissue maps, at the chosen voxel size over SPM's bounding box.
8. **Masks** in output space: brain = GM + WM + CSF > 0.5, inside the field of view in every frame.
9. **Confounds.** A table with fMRIPrep's column names: 24 motion regressors (6 parameters, their derivatives and squares; Friston 1996), framewise displacement (Power 2012), the mean signals of brain (`global_signal`), WM and CSF, aCompCor components of WM and CSF (computed on good frames), one `motion_outlier_XX` column per flagged frame and a `censor` column.
10. **IQMs** (optional, on by default): MRIQC image quality metrics of the raw BOLD and T1w; see 6.5.
11. **Denoising.** One linear model per voxel: polynomial trend (order 1) + the chosen regressors, with data and regressors band-passed by the same filter, fitted on the good frames. High-motion frames are first filled in from the good frames (spectral interpolation) so that filtering does not spread them, then left out of the fit and marked for later steps. The voxel mean is added back. Output: `desc-denoised` .
12. **Smoothing** (optional): Gaussian kernel of the chosen FWHM; output `desc-denoisedSm6` . The unsmoothed series is kept.
13. **Outputs** and a **QC report** per run; when all runs have finished, the group report and the exclusion list.

> **Why denoise before smoothing, and why filter regressors too.** Regressing nuisance signals from smoothed data mixes signals of neighbouring tissue; band-passing the data but not the regressors reintroduces removed signal (Hallquist et al. 2013; Lindquist et al. 2019). SeeCAT filters both and fits them together. High-motion frames are interpolated before filtering because a spike would otherwise spread over its neighbours; the interpolated frames are never used by later statistics (Power et al. 2014).

Input options

Option (window)	Batch field	Default	Meaning and advice
BIDS folder	`bids_dir`	`(empty)`	The raw BIDS dataset, e.g. D:/study/bids, with sub-<label>/anat/*_T1w.nii.gz and sub-<label>/func/*_bold.nii.gz plus their JSON sidecars (DICOM to BIDS writes this). Results go to <BIDS folder>/derivatives/seecat unless Output folder is set. Use Data check first to see the runs, TR and head motion.
Dummy scans advanced	`dummy_scans`	`auto`	Volumes removed from the start of each run, before the signal is steady. auto (default): detected from the data (leading volumes whose mean brain intensity is an outlier). Scanners that already discard them give none. A number, e.g. 5: remove that many from every run.

Preprocessing options

Option (window)	Batch field	Default	Meaning and advice
Slice timing	`slice_timing.enable`	`true`	Correct for the different acquisition times of the slices within a volume (SPM slice timing). On by default; it needs the slice times (SliceTiming in the sidecar, written by dcm2niix). Turn it off when the slice times are unknown.
Slice times (s) advanced	`slice_timing.slice_times`	`auto`	auto (default): SliceTiming from each run's JSON sidecar. Otherwise the acquisition time of every slice in seconds, in slice order, e.g. for 4 interleaved slices and TR 2: 0 1 0.5 1.5.
Reference time advanced	`slice_timing.ref_time`	`middle`	The time within the TR that all slices are shifted to. middle (default): half the TR. A number in seconds, e.g. 0 for the first slice.
Realign advanced	`realign.enable`	`true`	Correct head motion by aligning every volume (SPM realign). The six motion parameters and framewise displacement go into the confounds table. Keep it on.
Register to mean advanced	`realign.register_to_mean`	`true`	Two passes: first to the first volume, then to the mean of the aligned volumes (SPM's "register to mean"). More accurate; on by default.
Normalise advanced	`normalise.enable`	`true`	Bring the data into MNI space (SPM's IXI549Space template) so subjects can be compared voxel by voxel. Needed by the group analyses.
Normalisation	`normalise.method`	`auto`	auto (default): segment when the subject has a T1w image, otherwise epi. segment: the T1w image is segmented (SPM unified segmentation) and its deformation applied to the functional data, after registering the mean EPI to the T1w. The most accurate. epi: the mean EPI is normalised directly to the tissue templates; for data without a T1w. `auto` | `segment` | `epi`
Voxel size (mm)	`normalise.voxel_size`	`3`	Voxel size of the normalised data, e.g. 3 (default) or 2. Smaller voxels need more disk and memory (a 2 mm run is 3.4 times as large) and make Degree much slower. Three numbers, e.g. 3 3 4, give different sizes per axis.
Bounding box advanced	`normalise.bounding_box`	`[-90 -126 -72;90 90 108]`	The extent of the normalised images in MNI mm, [xmin ymin zmin; xmax ymax zmax]. Default [-90 -126 -72; 90 90 108], SPM's box, which holds the whole brain and cerebellum.

VBM options

Option (window)	Batch field	Default	Meaning and advice
VBM maps	`vbm.enable`	`false`	Also write the grey- and white-matter (and CSF) maps of the T1w segmentation normalised to MNI (IXI549Space, 1.5 mm, as SPM writes them), for voxel-based morphometry with the Statistics page (Map: modulatedSm8, Suffix: probseg). Only for runs with a T1w; written once per T1w. Tissue volumes (GM, WM, CSF, TIV in mL) are always written (reports/anat_volumes.tsv); Statistics and Harmonize can use tiv_ml as a covariate without adding it to participants.tsv. This uses SPM's unified segmentation without DARTEL.
Modulated	`vbm.modulated`	`true`	Ticked (default): modulated maps (desc-modulated), scaled by the Jacobian of the warp so they keep the amount of tissue: the usual VBM input, analysed with TIV as a covariate. Unticked: unmodulated maps (desc-warped), tissue concentration.
VBM smoothing (mm)	`vbm.fwhm`	`8`	FWHM of the smoothed copy (desc-modulatedSm8 for 8 mm, the default); 0 writes none. 6 to 10 mm is usual for VBM.

Nuisance mask options

Option (window)	Batch field	Default	Meaning and advice
WM probability advanced	`masks.wm_threshold`	`0.99`	White-matter mask for the wm regressor (and aCompCor): voxels whose WM probability from segmentation is above this, 0.99 by default, then eroded. A strict mask keeps grey-matter signal out of the nuisance signal.
CSF probability advanced	`masks.csf_threshold`	`0.95`	CSF mask for the csf regressor: CSF probability above this, 0.95 by default (0.99 eroded can leave no voxels in small ventricles), then eroded.
Erosion (mm) advanced	`masks.erode_mm`	`2`	The WM and CSF masks are shrunk by a ball of this radius (2 mm by default) in the subject's own space, away from grey matter. 0: no erosion.

Denoising, censoring, smoothing and quality-control options are in sections 6.12 and 6.13; output, selection and run options in 6.15.

Outputs

```
derivatives/seecat/
  dataset_description.json
  sub-01/
    anat/  sub-01_desc-preproc_T1w.nii             bias-corrected T1w (coregistered to the EPI)
           sub-01_label-GM|WM|CSF_probseg.nii       native tissue maps (sidecar: volumes in mL,
TIV)
           sub-01_from-T1w_to-IXI549Space_mode-image_xfm.nii   deformation field
           sub-01_space-IXI549Space_res-1p5_label-GM_desc-modulated_probseg.nii      VBM
(vbm.enable)
           sub-01_space-IXI549Space_res-1p5_label-GM_desc-modulatedSm8_probseg.nii   VBM, smoothed
    func/  sub-01_task-rest_space-IXI549Space_res-3_desc-preproc_bold.nii      normalised, not
denoised
           sub-01_task-rest_space-IXI549Space_res-3_desc-denoised_bold.nii     input of Seed FC,
Degree, Gradients
           sub-01_task-rest_space-IXI549Space_res-3_desc-denoisedSm6_bold.nii  smoothed copy
           sub-01_task-rest_space-IXI549Space_res-3_desc-brain_mask.nii
           sub-01_task-rest_desc-confounds_timeseries.tsv (+ .json)
    sub-01_task-rest_desc-qc_report.html, sub-01_task-rest_desc-qc_metrics.json, figures/
  reports/  prep_qc.html, prep_qc.tsv, qc_exclusions.tsv, anat_volumes.tsv,
prep_<time>_summary.tsv
  logs/
```

Images are written compressed (.nii.gz) when the input is compressed, otherwise as .nii. Work files are in `seecat_work` next to the output folder and are deleted after a unit succeeds (unless Keep work files).

6.4 fMRIPrep denoising

For data already preprocessed by fMRIPrep. For each BOLD run in the chosen space it reads `*_desc-preproc_bold`, its brain mask and confounds table, removes fMRIPrep's non-steady-state volumes, replaces fMRIPrep's motion outlier columns by SeeCAT's own (the same FD rule for every input), and then denoises, censors and smooths exactly as Preprocessing. Outputs have the same names, so Seed FC, Degree, Gradients and Statistics work unchanged. IQMs are not computed (fMRIPrep's outputs hold no raw data); the motion, coverage and registration checks are (registration uses fMRIPrep's tissue maps in the output space when present).

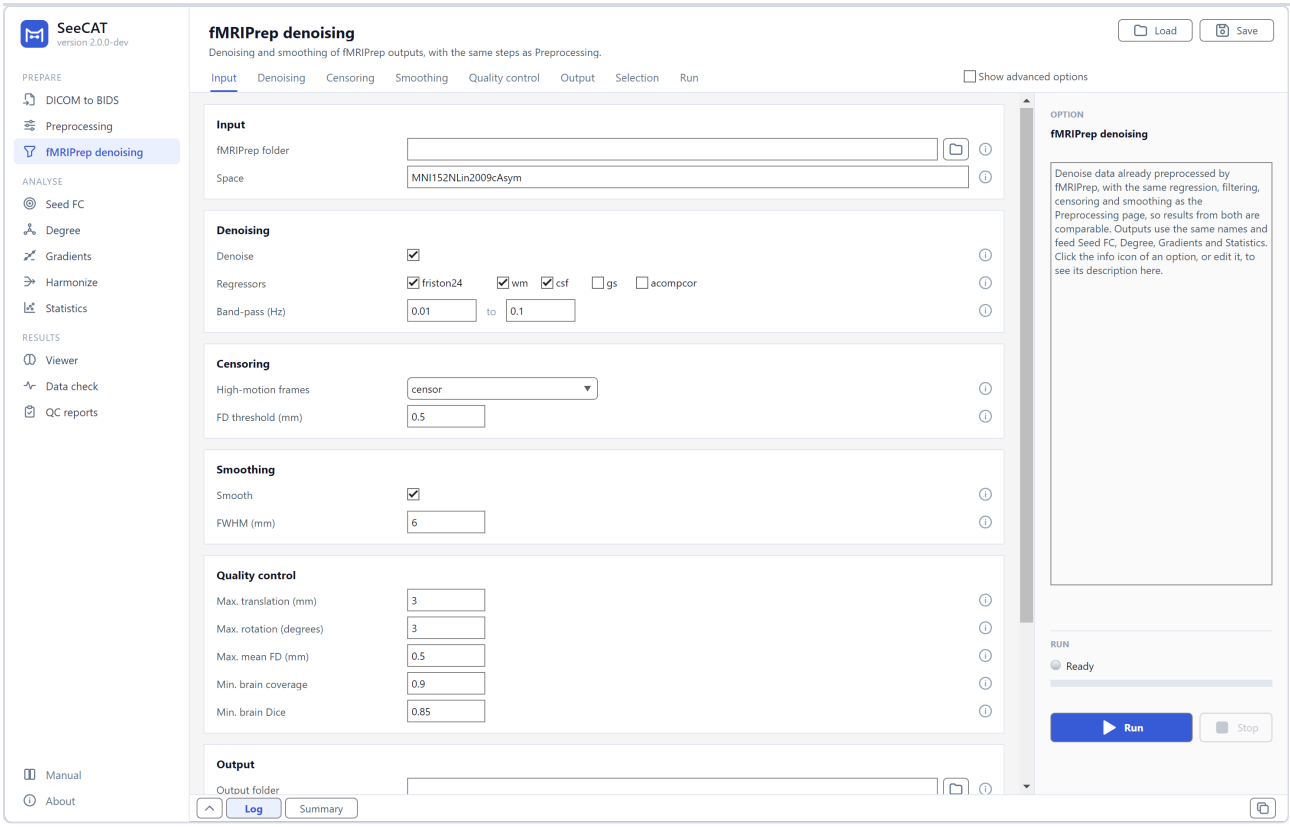


Figure 7. The fMRIPrep denoising page.

Option (window)	Batch field	Default	Meaning and advice
fMRIPrep folder	`fmriprep_dir`	`(empty)`	The output folder of fMRIPrep, e.g. D:/study/bids/derivatives/fmriprep. SeeCAT reads, for each BOLD run, *_space-<space>_desc-preproc_bold.nii.gz, its brain mask and the confounds table; fMRIPrep's own steps are not repeated. Results go to derivatives/seecat next to it unless Output folder is set, with the same names as the Preprocessing page, so Seed FC, Degree and Statistics work the same way.
Space	`space`	`MNI152NLin2009cAsym`	Which fMRIPrep output space to denoise, the space-<label> in the file names. Default MNI152NLin2009cAsym. Group analyses need a template space, not T1w or fsnative.
Resolution label advanced	`res`	`(empty)`	The res-<label> of the files, e.g. 2, needed only when fMRIPrep wrote the space at more than one resolution. Empty: the only one present.
Dummy scans advanced	`dummy_scans`	`auto`	Volumes removed from the start of each run. auto (default): those fMRIPrep marked as non-steady-state (non_steady_state_outlier columns). A number, e.g. 5: that many from every run.

`aCompCor` takes the first n `c_comp_cor` / `w_comp_cor` columns of fMRIPrep's table; older fMRIPrep versions write only `a_comp_cor_*`, which must then be listed by name in a config file (`denoise.regressors`).

6.5 Quality control

Quality control runs inside Preprocessing and fMRIPrep denoising; the **QC reports** page is where you review it and decide which runs are used.

The run report

Each run gets `sub-*/…_desc-qc_report.html` with:

- a summary: frames, dummy scans, mean and largest FD, flagged frames, minutes kept, tSNR, the correlation of FD with DVARS before and after denoising (a good denoising brings it close to 0), largest translation and rotation, coverage;
- head motion: FD over time with the threshold and flagged frames shaded, translations and rotations;
- DVARS before and after denoising;
- carpet plots (Power 2017) before and after denoising, voxels ordered GM, WM, CSF;
- registration and coverage: the mean normalised EPI with the template GM outline and the brain mask, and the registration measures;
- the WM and CSF masks used for the nuisance signals;
- the image quality metrics (MRIQC).



Figure 8. Part of a run report: summary, head motion and DVARS.

Image quality metrics (IQMs)

With `qc.iqm` (on by default) Preprocessing computes the metrics of MRIQC (Esteban et al. 2017) on the raw data, reusing its own segmentation, realignment and coregistration: the BOLD run after dummy-scan removal, head-motion corrected, without slice timing; the T1w with its bias-corrected

image and native tissue maps. The formula-based metrics give the same values as MRIQC's own functions on the same inputs; the head, air and brain masks are SeeCAT's own, so absolute values can differ from an MRIQC run. Compare runs within a dataset and within a site.

BOLD metric	Meaning	Better
`bold_tsnr`	median temporal SNR in the brain (mean / SD over time)	higher
`bold_snr`	SNR of the mean EPI in the brain	higher
`bold_fber`	foreground (brain) to background energy ratio	higher
`bold_efc`	entropy focus criterion: ghosting and blurring	lower
`bold_gsr_x` , `bold_gsr_y`	ghost to signal ratio along x and y (y is the usual phase-encoding direction)	lower
`bold_dvars_std` , `_nstd` , `_vstd`	DVARS standardised, raw and voxel-wise standardised (Nichols 2013); near 1 is good for the standardised one	lower
`bold_fd_mean` , `fd_num` , `fd_perc`	mean FD, frames and percentage with FD > 0.2 mm	lower
`bold_aor`	AFNI outlier ratio: mean share of outlier voxels per volume	lower
`bold_aqi`	AFNI quality index: mean of 1 − Spearman r of each volume with the median volume	lower
`bold_gcor`	global correlation: average correlation of all voxel pairs	lower
`bold_fwhm_*`	smoothness in voxels	–

T1w metric	Meaning	Better
`t1_cnr`	GM–WM contrast to noise	higher
`t1_cjv`	coefficient of joint variation of GM and WM (Ganzetti et al. 2016)	lower
`t1_snr_*`	SNR in CSF, GM, WM and their mean	higher
`t1_snrd_*`	Dietrich SNR, with the noise taken from the air around the head	higher
`t1_efc` , `t1_fber`	entropy focus criterion; foreground to background energy	lower; higher
`t1_inu_med` , `t1_inu_range`	median and spread (95th − 5th percentile) of the bias field in the brain	range lower
`t1_wm2max`	WM median over the 99.95th percentile of the image; about 0.6–0.8 is typical	–
`t1_icvs_*` , `t1_rpve_*`	tissue volume fractions; residual partial volume	–
`t1_tiv_ml` , `t1_gm_ml` , …	total intracranial volume and tissue volumes (mL)	–

> **Empty backgrounds.** Air-based T1w metrics (FBER, Dietrich SNR) are not comparable when the background holds only a few integer values (e.g. clipped when exported). The group report detects this (air MAD ≤ 1) and does not use those two metrics for warnings on such images.

Registration measures

Measure	What it checks	Typical (example dataset)
`reg_brain_dice`	Dice of the run's brain mask with the template brain mask: the overall fit of the normalisation	0.95–0.98
`reg_gm_corr`, `reg_wm_corr`, `reg_csf_corr`	spatial correlation of the normalised tissue maps with SPM's TPM	GM 0.64–0.74, WM 0.79–0.83
`reg_gm_dice`	Dice of GM > 0.5 with TPM GM > 0.5	0.70–0.75
`reg_epi_t1_dice`	Dice of a brain mask taken from the mean EPI's own intensity (above 20 % of its 98th percentile) with the T1w brain: how well the EPI sits on its T1w	0.85–0.92

Coverage

The share of the template brain inside the field of view in every frame, in total and separately for its top (z > 50 mm) and bottom (z < −30 mm: cerebellum and temporal poles). Low coverage means part of the brain was not scanned.

Exclusion rules and warnings

When every run has finished, the group report (`reports/prep_qc.html`) writes `reports/qc_exclusions.tsv`. A run is suggested for exclusion when it breaks a rule:

Rule	Option	Default
largest absolute translation	`qc.max_translation`	3 mm
largest absolute rotation	`qc.max_rotation`	3°
mean framewise displacement	`qc.max_mean_fd`	0.5 mm
too little data after censoring	`censor.min_minutes`	5 min
coverage of the template brain	`qc.min_coverage`	0.9
brain Dice with the template (failed normalisation)	`qc.min_brain_dice`	0.85

Warnings, which do not exclude: mean FD above 0.25 mm; bottom coverage below 80 %; an IQM or registration measure whose robust z (distance from the median in MADs) within the run's site is beyond 3 in the bad direction. Sites come from the `site` column of participants.tsv; without it, or with fewer than 5 runs in a site, the whole dataset is the reference. T1w images get a row of their own, with warnings only.

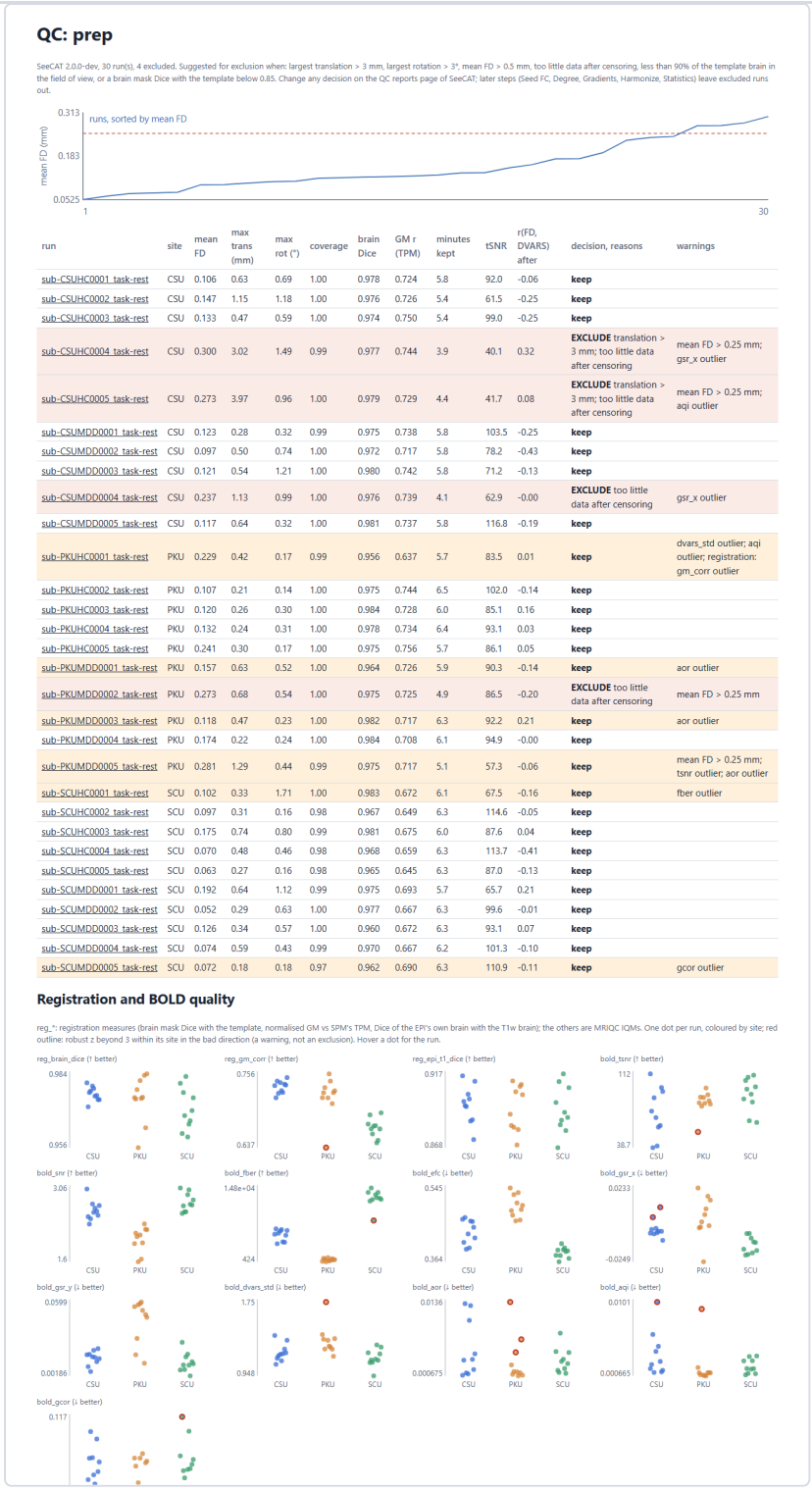


Figure 9. The group report: one row per run with its decision, reasons and warnings, then the distribution of each measure by site.

The QC reports page

Choose the derivatives folder: the table lists every run and T1w image with its site, mean FD, largest translation and rotation, coverage, brain Dice, GM correlation with the TPM, minutes kept, tSNR, T1w CNR, the reasons and the warnings. Tick or untick **Exclude**, add a **Note**, and **Save**

decisions. Rows you change are marked as your decision and keep it when Preprocessing is run again. **Group report** and **Run report** open the HTML pages.

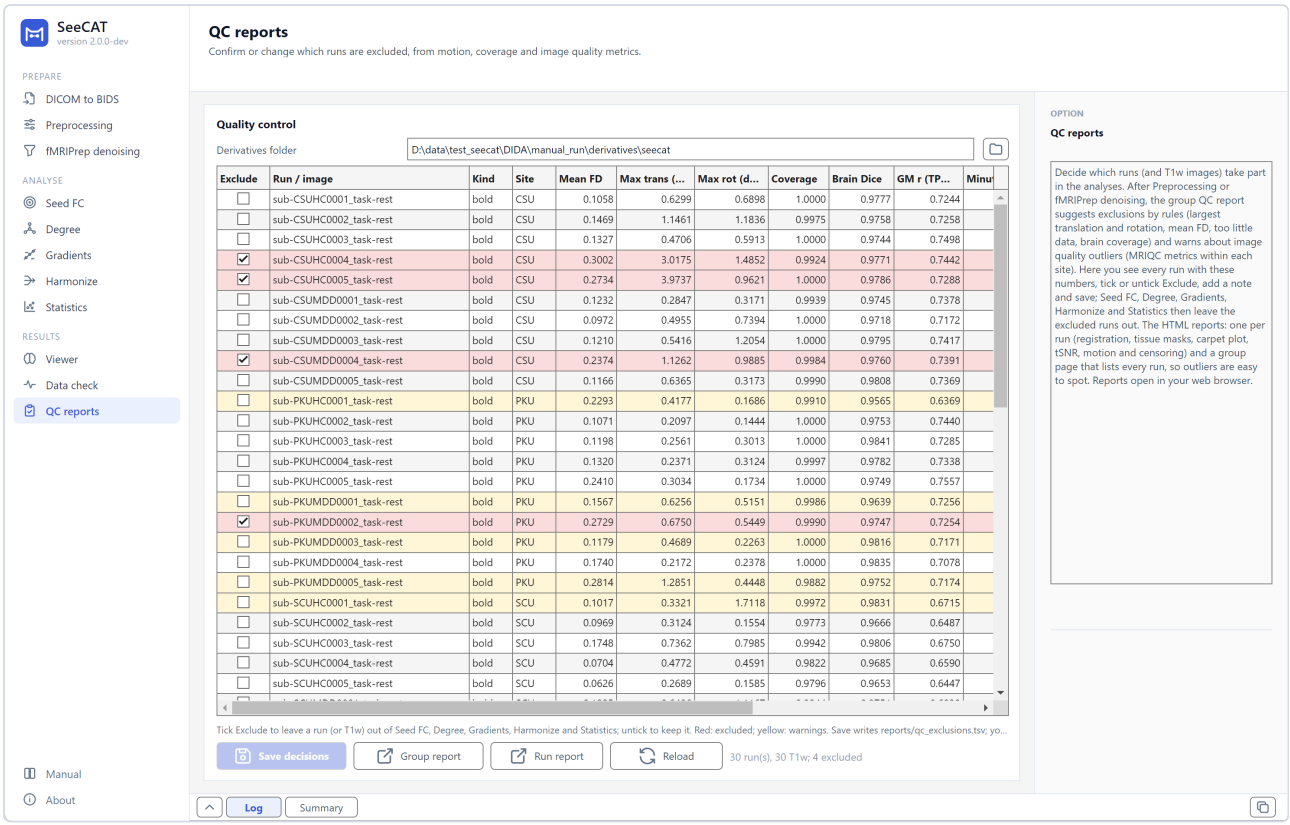


Figure 10. The QC reports page: rows in red are excluded, yellow ones have warnings.

The list is a plain table:

```
id                        kind  exclude  reasons            warnings              decision  note
sub-PKUHC0001_task-rest   bold  1        mean FD > 0.5 mm                         auto
sub-SCUMDD0003_task-rest  bold  0                           mean FD > 0.25 mm     user
checked carpet plot
sub-CSUHC0002_T1w         T1w   0                           cnr outlier           auto
```

Later steps match files to this list by the entities that identify an acquisition (sub, ses, task, acq, ce, rec, dir, run, echo), so every map derived from an excluded run, and the VBM maps of an excluded T1w, are left out. Editing the list does not make finished units run again.

6.6 Seed FC

Seed-based functional connectivity. For each run and seed, the mean signal of the seed's voxels (inside the brain mask, on the frames kept after censoring) is correlated with every voxel; the result is written as a Pearson r map and a Fisher z map (z = atanh r). Use the z maps in Harmonize and Statistics.

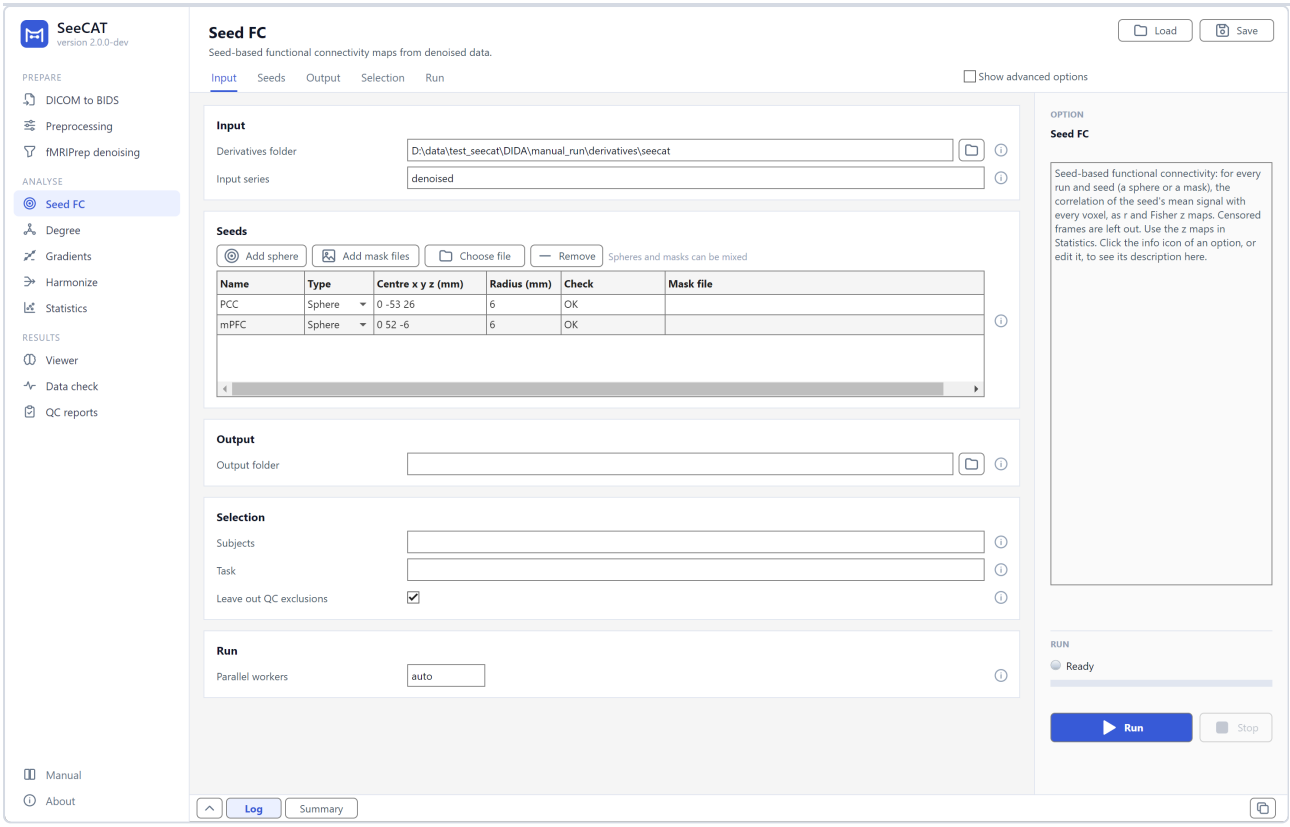


Figure 11. Seed FC with two spheres (PCC and mPFC). Spheres and masks can be mixed in the table; the Check column says what a row still needs.

Seeds

- **Sphere**: a centre in MNI millimetres (`x y z`) and a radius; voxels within the radius of the centre. Examples: PCC `0 -53 26` , mPFC `0 52 -6` , left motor cortex `-38 -22 56` .
- **Mask**: a NIfTI image in MNI space (any voxel size; voxels above 0 are the seed), resampled to the data by nearest neighbour. **Add mask files** adds one row per chosen file, named after it.

Names are letters and digits; they become the `desc` of the output maps.

Option (window)	Batch field	Default	Meaning and advice
Derivatives folder	`input_dir`	`(empty)`	The SeeCAT derivatives folder written by the Preprocessing or fMRIPrep page, e.g. D:/study/bids/derivatives/seecat. It holds, per run, the denoised BOLD series, the brain mask and the confounds table with the censored frames. The maps are written next to the runs (desc-<seed>_stat-r and stat-z), in the same folder unless Output folder is set.
Input series	`input`	`denoised`	Which series of each run is used: the desc label of the file name. denoised (default): denoised, band-passed, not smoothed. The seed signal then comes only from the seed's own voxels. denoisedSm6 (or another Sm label your preprocessing wrote): the smoothed series, as in many published FC studies; the maps are then smooth without a further step.

Option (window)	Batch field	Default	Meaning and advice
Space advanced	`space`	`(empty)`	Needed only when the derivatives hold the runs in more than one space. Empty: the only space present.
Resolution label advanced	`res`	`(empty)`	The res-<label> of the files, e.g. 2. Needed only when the derivatives hold the same runs at more than one resolution (e.g. fMRIPrep res-1 and res-2). Empty: the only one present.
Seeds	`seeds`	`(empty)`	One row per seed; spheres and masks can be mixed. For each run and seed, the mean signal of the seed's voxels (inside the brain mask, censored frames left out) is correlated with every voxel: a map of r and of Fisher z (desc-<name>_stat-z, the one to use in Statistics). Sphere: centre in MNI mm as "x y z" and a radius in mm, e.g. PCC: 0 -53 26, radius 6; mPFC: 0 52 -6; left motor cortex: -38 -22 56. Mask: a NIfTI image in MNI space (any voxel size; voxels above 0 are the seed), e.g. an atlas region saved as a binary mask. It is resampled to the data. Add sphere adds a row; Add mask files adds one row per chosen file, named after it; Choose file sets the file of the selected row; Remove deletes the selected rows. Names are letters and digits and name the output maps. The Check column says what a row still needs.

Outputs

```
sub-01/func/sub-01_task-rest_space-IXI549Space_res-3_desc-PCC_stat-r_connectivity.nii.gz (+ .json)
sub-01/func/sub-01_task-rest_space-IXI549Space_res-3_desc-PCC_stat-z_connectivity.nii.gz (+ .json)
```

The sidecar records the seed (centre and radius, or file), the input series and the frames used.

```
cfg = seecat.config('fc');
cfg.input_dir = 'D:/study/bids/derivatives/seecat';
cfg.seeds = struct('name', {'PCC', 'Amygdala'}, 'center', {[0 -53 26], []}, ...
                   'radius', {6, []}, 'file', {'', 'D:/masks/amygdala_L.nii'});
seecat.run(cfg);
```

6.7 Degree

Voxel-wise degree centrality. Every voxel of the mask is a node of the correlation network of a run (frames kept after censoring). The binary degree of a voxel counts its connections whose correlation reaches the threshold; the weighted degree (strength) sums their r values. As in SeeCAT 1.0: correlations from z-scored series, the sign rule applied (absolute: $|r|$; negative: $-r$), values below the threshold set to 0. Optionally connections shorter than a distance are ignored, or split into distance bins (short- and long-range degree). Maps are written raw and z-scored within the mask, and smoothed afterwards if asked (the network is built from unsmoothed data).

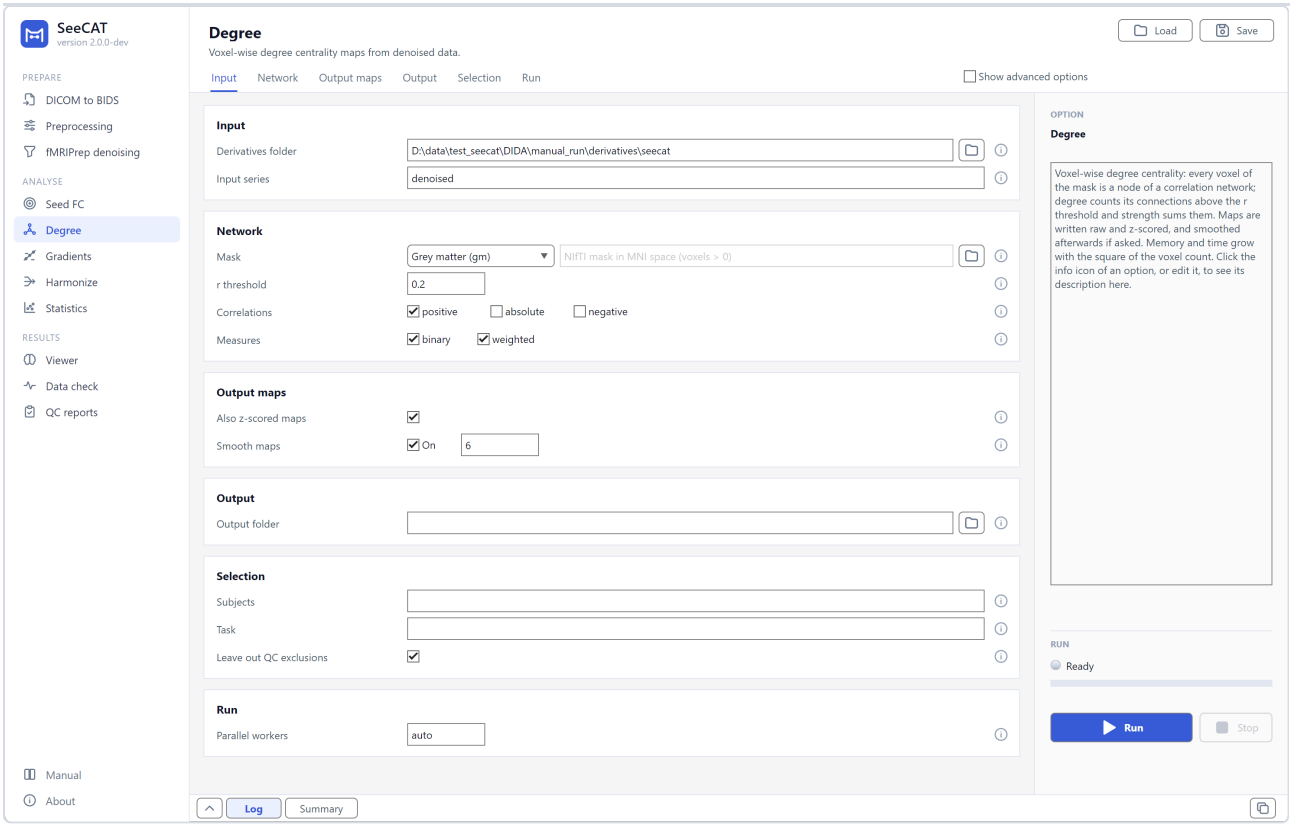


Figure 12. The Degree page.

Option (window)	Batch field	Default	Meaning and advice
Derivatives folder	`input_dir`	`(empty)`	The SeeCAT derivatives folder written by the Preprocessing or fMRIPrep page, e.g. D:/study/bids/derivatives/seecat. It holds, per run, the preprocessed BOLD series (*_desc-<input>_bold.nii.gz), the subject brain mask and the confounds table with the censored frames. Degree maps are written next to them, in the same folder unless Output folder is set.
Input series	`input`	`denoised`	Which series of each run the network is built from: the desc label of the file name. denoised (default): denoised and band-pass filtered, NOT smoothed. Use this one: smoothing before correlating inflates the connections between neighbouring voxels. Other labels, e.g. denoisedSm6, if your preprocessing wrote them.
Space advanced	`space`	`(empty)`	Needed only when the derivatives hold the runs in more than one space, e.g. MNI152NLin2009cAsym and T1w from fMRIPrep. Empty: the only space present.
Resolution label advanced	`res`	`(empty)`	The res-<label> of the files, e.g. 2. Needed only when the derivatives hold the same runs at more than one resolution (e.g. fMRIPrep res-1 and res-2). Empty: the only one present.

Option (window)	Batch field	Default	Meaning and advice
Mask	`mask`	`gm`	The voxels that make up the network: every voxel of this mask is a node and is correlated with every other. Grey matter (gm, default): SeeCAT's grey-matter mask (SPM TPM GM > 0.2) resampled to the data, as in SeeCAT 1.0. White matter and CSF voxels would add noise connections. Whole brain (brain): every voxel of the subject's brain mask. Mask file: your own NIfTI image in the space of the data (MNI); voxels > 0 are kept, e.g. a cortical mask or a mask shared by all subjects. Choose it with the folder button. The mask is always limited to the subject's brain mask. The time and memory grow with the square of the number of voxels: about 50 000 grey-matter voxels at 3 mm, 4 times more at 2 mm.
r threshold	`threshold`	`0.2`	A connection counts when its correlation r is at least this value (after the sign rule below), e.g. 0.2 (default, as in SeeCAT 1.0), 0.25 or 0.3. Higher thresholds keep fewer, stronger connections. The value appears in the output name, e.g. desc-degPosBinR0p2. Range 0 to below 1.
Correlations	`sign`	`positive`	Which correlations count; each ticked option gives its own maps. positive (default, desc-degPos...): r >= threshold. The usual choice, since negative correlations depend on the denoising (e.g. global signal regression). absolute (desc-degAbs...): |r| >= threshold, positive and negative together. negative (desc-degNeg...): -r >= threshold, the anti-correlations only. `positive` | `absolute` | `negative`
Measures	`weight`	`binary`, `weighted`	What is added up over a voxel's connections; each ticked option gives its own maps. binary = degree (desc-deg...Bin...): the number of connections above the threshold. weighted = strength (desc-deg...Wei...): the sum of their r values. Example: a voxel with 3 connections of r = 0.3, 0.5 and 0.6 has degree 3 and strength 1.4. `binary` | `weighted`
Exclude within (mm) advanced	`exclude_local_mm`	`0`	Connections between voxels closer than this distance are not counted. Neighbouring voxels correlate partly through the data's own smoothness, so short connections can dominate degree. 0 (default): count all. E.g. 20: leave out connections shorter than 20 mm (leaves out the immediate neighbourhood).

Option (window)	Batch field	Default	Meaning and advice
Distance bins (mm) advanced	`distance_edges`	`[0 Inf]`	Splits the connections by length, with one map per bin. Give the bin edges in mm, increasing, starting at 0 or more; Inf means no upper limit. [0 Inf] (default): one map with all connections. [0 75 Inf]: short-range (up to 75 mm) and long-range (beyond 75 mm) degree, e.g. desc-degPosBinR0p2D0to75 and …D75toInf. [0 25 75 Inf]: three bins. A connection of length d is in the bin lo < d <= hi.
Also z-scored maps	`zscore`	`true`	Also write each map z-scored within the mask (minus the mean, divided by the standard deviation over the mask voxels), as stat-z next to the stat-raw maps. Raw degree depends on the number of voxels and the data quality of each subject; z-scored maps are easier to compare between subjects and are the usual input to group statistics.
Smooth maps	`smooth_fwhm`	`6`	Smooth the finished maps with a Gaussian kernel of this FWHM in mm, e.g. 6 (default) or 4. Smoothed copies are written with Sm<fwhm> in the name (desc-degPosBinR0p2Sm6); the unsmoothed maps are always kept. The network itself is built from the unsmoothed series, and only the maps are smoothed, as group statistics usually need. Untick On to skip smoothing. Three values, e.g. 6 6 8, give a different FWHM per axis.

Outputs

```
…_desc-degPosBinR0p2_stat-raw_connectivity.nii.gz      binary degree, positive r ≥ 0.2
…_desc-degPosBinR0p2_stat-z_connectivity.nii.gz        the same, z-scored within the mask
…_desc-degPosWeiR0p2Sm6_stat-z_connectivity.nii.gz     strength, z-scored, smoothed 6 mm
…_desc-degPosBinR0p2D0to75_stat-raw_connectivity.nii.gz  with distance bins [0 75 Inf]
```

Time and memory grow with the square of the number of voxels: about 50 000 grey-matter voxels at 3 mm take seconds to a minute per run; at 2 mm there are 3.4 times as many voxels and over ten times the work.

6.8 Gradients

Connectivity gradients (Margulies et al. 2016): the main axes along which connectivity profiles change across the cortex, the first usually running from sensory and motor cortex to the default-mode network. For each run:

1. the node series: the mean of the run's brain-mask voxels in each node (voxels of a coarser MNI grid, or parcels of an atlas), frames kept after censoring;
2. the connectivity matrix: Pearson r between nodes;
3. each row keeps its strongest 10 % (sparsity 0.9) and the rest is set to 0;

4. the affinity between rows: normalised angle (1 − angle/π of their cosine similarity) or cosine;
5. diffusion map embedding with α = 0.5 (Coifman & Lafon 2006) and multiscale eigenvalues λ/(1−λ), as BrainSpace (Vos de Wael et al. 2020);

then the runs are aligned (generalized Procrustes: rotations and sign flips onto a reference, the reference updated to the mean, up to 10 rounds), starting from the gradients of the group-mean connectivity matrix (default) or from the first run. SeeCAT's embedding equals the exact diffusion map; it solves the symmetric form of the problem.

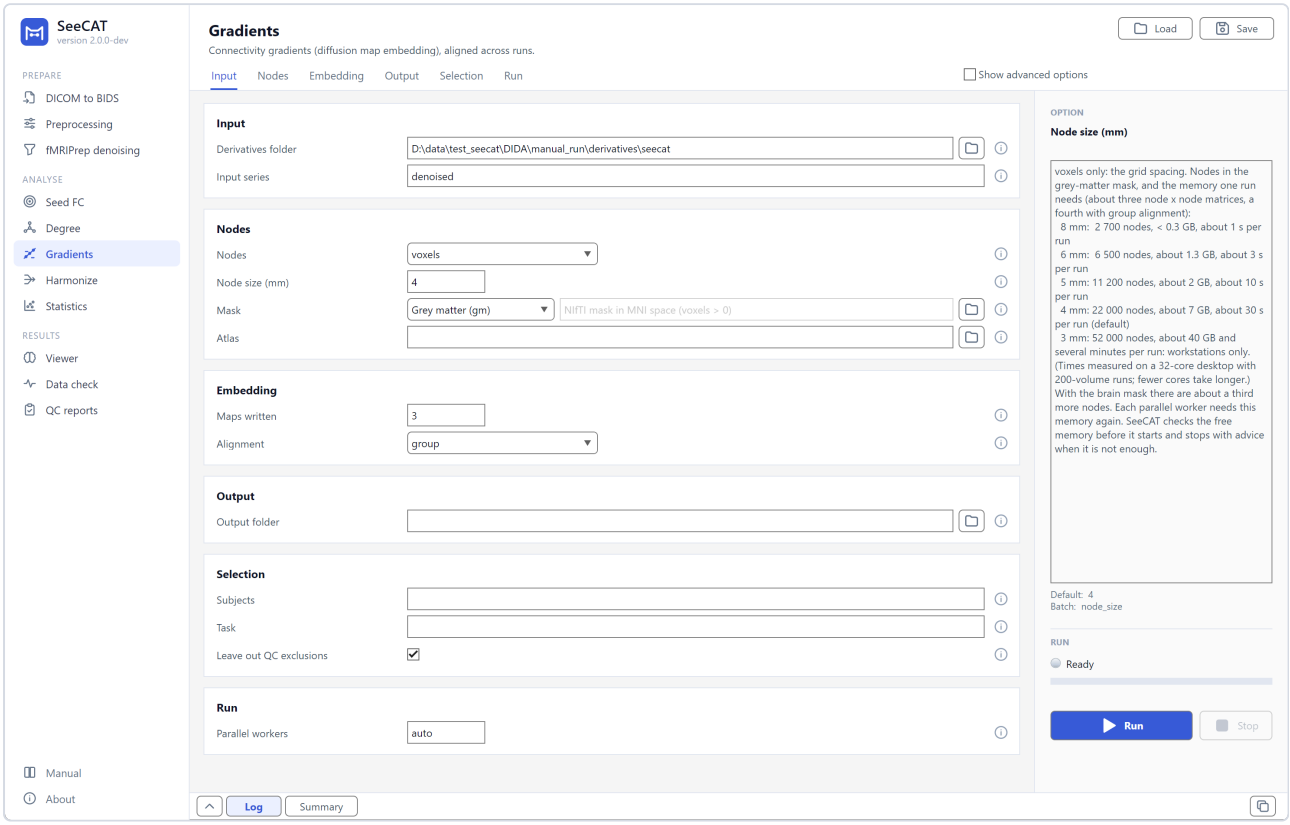


Figure 13. The Gradients page. The help for Node size lists the memory and time per run.

Option (window)	Batch field	Default	Meaning and advice
Derivatives folder	`input_dir`	`(empty)`	The SeeCAT derivatives folder written by the Preprocessing or fMRIPrep page, e.g. D:/study/bids/derivatives/seecat. Every run found there is used (or those chosen under Selection); they are analysed together because their gradients are aligned to each other.
Input series	`input`	`denoised`	Which series of each run is used: denoised (default; unsmoothed) or another desc label such as denoisedSm6. Voxel nodes average their voxels anyway, so smoothing adds little.
Space advanced	`space`	`(empty)`	Needed only when the derivatives hold the runs in more than one space. Empty: the only space present.
Resolution label advanced	`res`	`(empty)`	The res-<label> of the files, e.g. 2. Needed only when the derivatives hold the same runs at more than one resolution (e.g. fMRIPrep res-1 and res-2). Empty: the only one present.

Option (window)	Batch field	Default	Meaning and advice
Nodes	`nodes`	`voxels`	What the connectivity matrix is built between. voxels (default): the cells of a grid of Node size in MNI space inside the Mask; each node's signal is the mean of the data voxels in it. Gives voxel-wise gradient maps. atlas: the parcels of an atlas image (e.g. Schaefer 400 in MNI space); fast, and the values are also written as a table per run. Time and memory grow with the SQUARE of the number of nodes (see Node size). `voxels` | `atlas`
Node size (mm)	`node_size`	`4`	voxels only: the grid spacing. Nodes in the grey-matter mask, and the memory one run needs (about three node x node matrices, a fourth with group alignment): 8 mm: 2 700 nodes, < 0.3 GB, about 1 s per run 6 mm: 6 500 nodes, about 1.3 GB, about 3 s per run 5 mm: 11 200 nodes, about 2 GB, about 10 s per run 4 mm: 22 000 nodes, about 7 GB, about 30 s per run (default) 3 mm: 52 000 nodes, about 40 GB and several minutes per run: workstations only. (Times measured on a 32-core desktop with 200-volume runs; fewer cores take longer.) With the brain mask there are about a third more nodes. Each parallel worker needs this memory again. SeeCAT checks the free memory before it starts and stops with advice when it is not enough.
Mask	`mask`	`gm`	voxels only: where the nodes are. Grey matter (gm, default): cells at least half inside SeeCAT's grey-matter mask; the usual choice (Margulies et al. 2016 used cortex). Whole brain (brain): cells inside the brain mask, including white matter. Mask file: your own NIfTI mask in MNI space, e.g. cortex only.
Atlas	`atlas`	`(empty)`	atlas only: a NIfTI image in MNI space whose voxels hold integer parcel labels (0 = background), e.g. Schaefer2018_400Parcels_7Networks_order_FSLMNI152_2mm.nii.gz. Every label becomes a node; the gradient maps paint each parcel with its value.
Components advanced	`n_components`	`10`	How many gradients are computed and used in the alignment, 10 by default (BrainSpace). More components make the alignment more stable; they cost little.
Maps written	`n_maps`	`3`	How many gradients are written per run, 3 by default (desc-grad1, grad2, grad3 ...). Gradient 1 explains the most variance (usually sensory to default-mode cortex), gradient 2 the next (often visual to somatomotor), and so on. The eigenvalues of all components are in the JSON sidecars and the group table.
Sparsity advanced	`sparsity`	`0.9`	Share of each node's connections set to 0 before the affinity, keeping only its strongest ones: 0.9 (default) keeps the top 10%, as Margulies et al. 2016 and BrainSpace.

Option (window)	Batch field	Default	Meaning and advice
Affinity advanced	`kernel`	`normalized_angle`	How similar two nodes' connectivity profiles are. normalized_angle (default): 1 - angle/pi between the profiles; BrainSpace's recommendation. cosine: cosine similarity, negative values set to 0 (Margulies et al. 2016). none: the sparsified connectivity itself (symmetrised). `normalized_angle` | `cosine` | `none`
Diffusion alpha advanced	`alpha`	`0.5`	Anisotropy of the diffusion map: 0.5 (default) lowers the influence of how densely nodes are sampled (Coifman & Lafon 2006); 0 is a Laplacian eigenmap, 1 the Laplace-Beltrami operator.
Alignment	`align`	`group`	Gradients of separate runs can come out in a different order or sign; alignment makes gradient k comparable across runs, which group statistics need. group and procrustes both use generalized Procrustes, as BrainSpace: every run is rotated (rotation and sign flips only, no scaling) onto a reference, the mean of the rotated runs becomes the new reference, and this repeats up to 10 times or until the reference changes by less than 1e-5. They differ in the first reference: group (default): the gradients of the group-mean connectivity matrix, so gradient k keeps its group-level meaning. Needs one more node x node matrix of memory. These starting gradients are also written (desc-grad<k>_groupmatrix). procrustes: the first run's gradients; no group matrix. none: each run's own gradients; only for looking at single runs. `group` | `procrustes` | `none`

Choosing the nodes

Node size	Nodes (grey matter)	Memory per run	Time per run
8 mm	2 700	< 0.3 GB	about 1 s
6 mm	6 500	about 1.3 GB	about 3 s
5 mm	11 200	about 2 GB	about 10 s
4 mm (default)	22 000	about 7 GB	about 30 s
3 mm	52 000	about 40 GB	minutes

Times are for a 32-core desktop and 200-volume runs. SeeCAT checks the free memory first and stops with advice when it is not enough. An atlas (e.g. Schaefer 400) is fast and gives a value per parcel, also written as a table.

Outputs

```
sub-01/func/…_desc-grad1Vox4_connectivity.nii.gz   gradient 1 of the run, aligned (2, 3 … up to
Maps written)
group/gradientsVox4/gradientsVox4_desc-grad1_reference.nii    final reference (mean of the aligned
runs)
group/gradientsVox4/gradientsVox4_desc-grad1_groupmatrix.nii  gradients of the group-mean matrix
group/gradientsVox4/gradientsVox4_lambdas.tsv     eigenvalues of every run, variance explained, r
with the reference
```

6.9 Harmonize

Multi-site studies carry site and scanner differences in every map. Harmonize removes them from per-subject maps, between the measures (Seed FC, Degree, Gradients, VBM) and Statistics, while keeping the effects you will test. It writes the harmonised maps next to the originals with the desc label extended (`desc-PCC` → `desc-PCCComBat`).

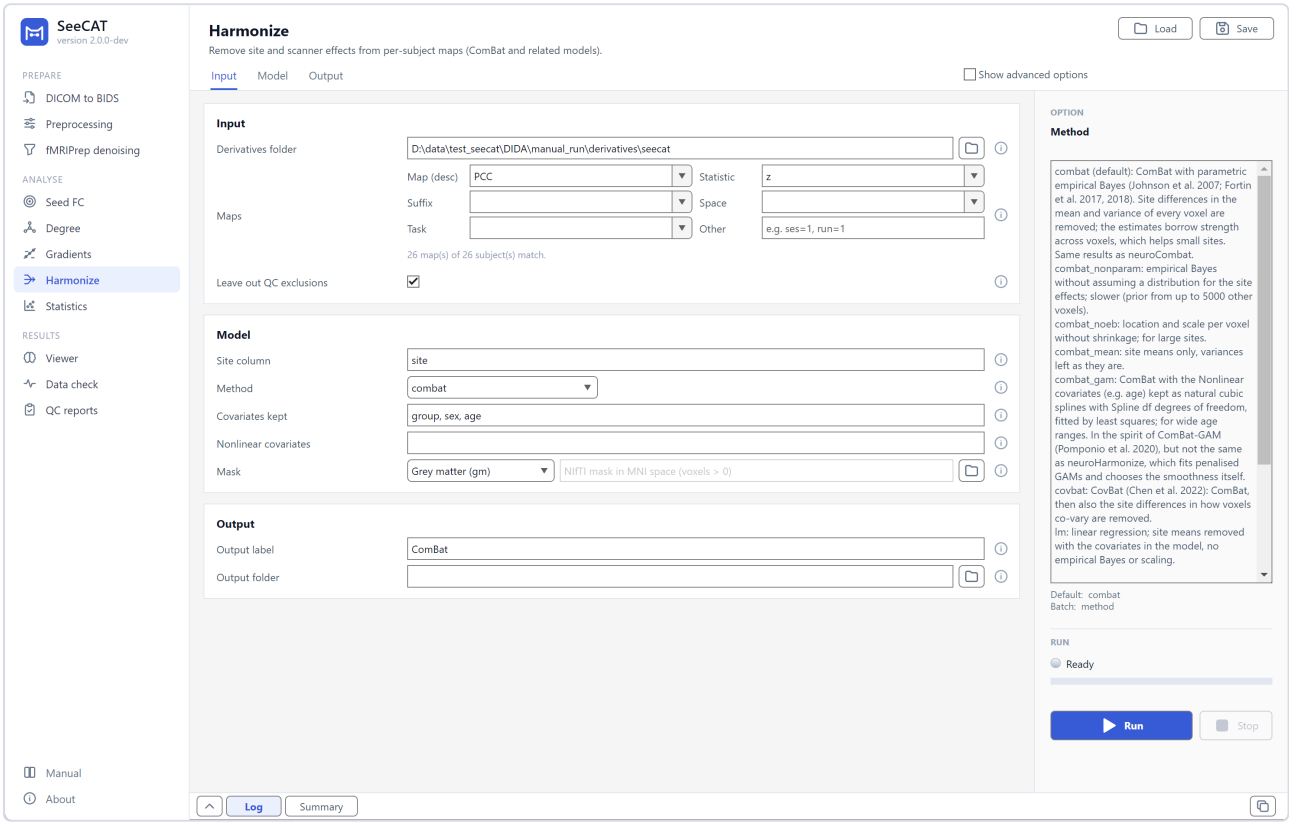


Figure 14. Harmonize for the PCC maps of the example: site column, ComBat, and the covariates whose effects are kept.

Methods

Method	What it does	When
`combat`	ComBat with parametric empirical Bayes (Johnson et al. 2007; Fortin et al. 2017, 2018): per voxel, data = site location + covariate effects + error scaled per site; the site location and scale are estimated per voxel and shrunk across voxels, then removed.	the default; small sites benefit from the shrinkage

Method	What it does	When
`combat_nonparam`	empirical Bayes without a distribution for the site effects (prior from up to 5000 other voxels)	when the site effects are far from normal
`combat_noeb`	location and scale per voxel without shrinkage	large sites
`combat_mean`	site means only	when variances should stay as they are
`combat_gam`	ComBat with the Nonlinear covariates as natural cubic splines of fixed df (Spline df), fitted by least squares. In the spirit of ComBat-GAM (Pomponio et al. 2020); neuroHarmonize fits penalised GAMs instead, so results differ from it	wide age ranges, non-linear effects
`covbat`	ComBat, then the site differences in the covariance between voxels are removed (Chen et al. 2022)	when network structure differs between sites
`lm`	linear regression: site means removed with the covariates in the model	a simple reference

The ComBat variants give the same results as neuroCombat. A **reference site** can be kept unchanged with the others mapped onto it. Every effect you will test later must be among the kept covariates; otherwise, where it differs between sites, it is partly removed with the site effect. The log and `group/harmonize<label>_<maps>_<hash>/…_report.json` (the hash tells map selections apart) give the share of voxels with a site effect (F test, $p < 0.05$, adjusted for the kept effects) before and after.

> **Small sites.** The scale (variance) step of `combat`, `combat_nonparam`, `combat_noeb`, `combat_gam` and `covbat` divides each site's residuals by its own standard deviation. With few subjects per site those estimates are noisy and the harmonised data look less variable than they are, so later tests of the kept covariates become too liberal (Nygaard et al. 2016). In Example B (three sites of ten, simulated ages) a GM–age regression on ComBat-harmonised VBM maps found three clusters that are not there: none with `combat_mean`, none with `lm`, none with the unharmonised maps and site as a covariate. With fewer than about 20 subjects per site, SeeCAT writes a warning in the log and in the report. Then prefer `combat_mean` or `lm`, or leave the maps unharmonised and add the site column as a covariate in Statistics; compare with the unharmonised result before reporting.

Option (window)	Batch field	Default	Meaning and advice
Derivatives folder	`input_dir`	`(empty)`	The SeeCAT derivatives folder with the per-subject maps to harmonise (output of Seed FC, Degree or Gradients), e.g. D:/study/bids/derivatives/seecat. The harmonised maps are written next to the originals unless Output folder is set.

Option (window)	Batch field	Default	Meaning and advice
Maps	`maps`	`(empty)`	Which maps to harmonise, by the parts of their file names, as in Statistics: e.g. Map PCC and Statistic z for seed FC, or Map grad1Vox4 for the first gradient. Open a list to pick from the values found; the line below counts the matching maps. Harmonise each kind of map separately (one run of this page per measure). All maps of all sites are needed together.
participants.tsv advanced	`participants`	`(empty)`	The table with participant_id, the site column and the covariates. Empty: participants.tsv of the BIDS dataset the derivatives come from.
Site column	`site`	`site`	The participants.tsv column that names the site or scanner of each subject, e.g. site, scanner or center. Every site needs at least 2 subjects; ComBat works best with 10 or more per site. A study scanned on two scanners at one centre has two sites.
Method	`method`	`combat`	combat (default): ComBat with parametric empirical Bayes (Johnson et al. 2007; Fortin et al. 2017, 2018). Site differences in the mean and variance of every voxel are removed; the estimates borrow strength across voxels, which helps small sites. Same results as neuroCombat. combat_nonparam: empirical Bayes without assuming a distribution for the site effects; slower (prior from up to 5000 other voxels). combat_noeb: location and scale per voxel without shrinkage; for large sites. combat_mean: site means only, variances left as they are. combat_gam: ComBat with the Nonlinear covariates (e.g. age) kept as natural cubic splines with Spline df degrees of freedom, fitted by least squares; for wide age ranges. In the spirit of ComBat-GAM (Pomponio et al. 2020), but not the same as neuroHarmonize, which fits penalised GAMs and chooses the smoothness itself. covbat: CovBat (Chen et al. 2022): ComBat, then also the site differences in how voxels co-vary are removed. lm: linear regression; site means removed with the covariates in the model, no empirical Bayes or scaling. Small sites (fewer than about 20 subjects each): the variance adjustment of combat, combat_nonparam, combat_noeb, combat_gam and covbat can make later tests of the kept covariates too liberal (Nygaard et al. 2016). In the manual's example (3 sites of 10, simulated age) a GM-age regression found 3 false clusters after combat and none after combat_mean or lm. Prefer combat_mean or lm for small samples; the log warns. `combat` | `combat_nonparam` | `combat_noeb` | `combat_mean` | `combat_gam` | `covbat` | `lm`
Covariates kept	`covariates`	`(empty)`	participants.tsv columns whose effects must survive harmonisation, comma separated, e.g. diagnosis, age, sex. Text columns become categories. Include every variable you will test later: an effect that is not listed and differs between sites is partly removed with the site effect.

Option (window)	Batch field	Default	Meaning and advice
Nonlinear covariates	`nonlinear`	`(empty)`	Numeric columns kept as smooth (natural spline) effects instead of straight lines, e.g. age over a lifespan sample. Needed by combat_gam; allowed with every method. Do not list a column in both boxes.
Spline df advanced	`spline_df`	`3`	Degrees of freedom of each nonlinear covariate's spline, 3 by default; more allow wigglier curves.
Reference site advanced	`ref_site`	`(empty)`	A site whose maps are left unchanged; the other sites are mapped onto it (e.g. the largest site, or the one a normative model was built on). Empty (default): all sites are mapped to their common average.
CovBat variance advanced	`covbat_var`	`0.95`	covbat only: share of the residual variance whose principal components are harmonised, 0.95 by default.
Mask	`mask`	`gm`	The voxels harmonised: grey matter (gm, default), whole brain, or a mask file; always only voxels that are non-zero in every map. Voxels outside are 0 in the outputs, so Statistics leaves them out.
Output label	`label`	`ComBat`	Added to the desc label of each map, e.g. desc-PCC -> desc-PCCComBat (default ComBat). Use another label to keep the results of different methods apart (e.g. CovBat). In Statistics, choose Map PCCComBat.
Output folder	`output_dir`	`(empty)`	Where the results go. Empty: the default of the pipeline, next to the input (for the analyses, the derivatives folder itself, so the maps sit beside the runs they come from). Set it to keep the results of different settings apart, e.g. D:/study/derivatives/degree_r03. Logs and the config of every run are written to its logs folder.
Work folder advanced	`work_dir`	`(empty)`	Intermediate files, one subfolder per unit. Empty: seecat_work next to the output folder. It is deleted after a unit succeeds (unless Keep work files) and kept after a failure, so a rerun does not redo finished steps. A fast local disk helps.
Leave out QC exclusions	`qc_exclusions`	`true`	Ticked (default): runs (and T1w images) marked "exclude" in the derivatives folder's reports/qc_exclusions.tsv are left out, and the log lists them. That list is made by the group QC report after Preprocessing or fMRIPrep denoising, from the rules there (head motion, coverage, too little data); confirm or change each decision on the QC reports page. Untick to use every run, e.g. to compare results with and without the excluded data.
Subjects advanced	`subjects`	`(empty)`	Only these subjects, labels without "sub-", comma separated. Empty: all with a matching map.

6.10 Statistics

Group statistics on per-subject maps: a general linear model at every voxel, contrasts, and correction for multiple comparisons. The model, smoothness estimate and random field theory follow SPM, so the numbers match SPM's results table for the same model.

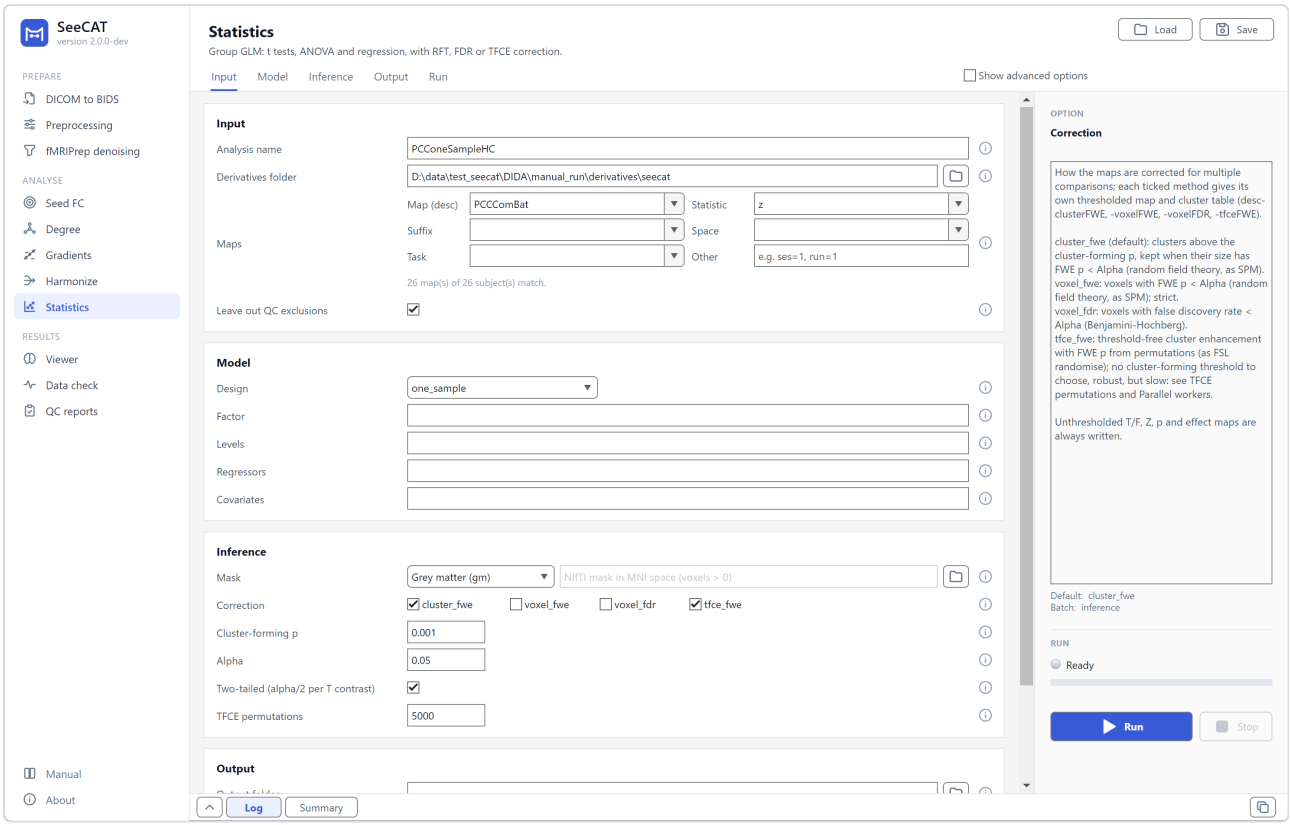


Figure 15. Statistics for the one-sample example: maps chosen by their name fields, the design and the corrections.

Designs

Design	Question	Set	Default contrasts
`one_sample`	is the mean map different from 0?	–	`pos`, `neg`
`two_sample`	do two groups differ?	Factor: participants column; Levels	`AgtB`, `BgtA`
`paired`	two maps per subject (e.g. ses-1, ses-2)	Factor: the entity that differs	both directions
`anova1`	three or more groups	Factor, Levels	F of equal means
`anova1_rm`	three or more maps per subject	Factor: entity	F of equal levels
`regression`	does the map vary with a score?	Regressors	`<name>pos`, `<name>neg`
`custom`	your own design matrix	Design TSV, Contrasts	–

Covariates are mean-centred; text covariates become centred dummy columns. With **Variance** unequal (default for anova1_rm), non-sphericity is estimated with ReML as in SPM.

Corrections

- **cluster_fwe** (default): clusters of voxels above the cluster-forming p (0.001), kept when their extent has a family-wise error p < alpha by random field theory.
- **voxel_fwe**: voxels with FWE p < alpha (random field theory).
- **voxel_fdr**: voxels with false discovery rate < alpha (Benjamini–Hochberg).
- **tfce_fwe**: threshold-free cluster enhancement (Smith & Nichols 2009; E = 0.5, H = 2, 26-connectivity) with voxel FWE p from permutations (Freedman–Lane: sign flips for one-sample, swaps within subjects for paired designs, row permutations otherwise), as FSL randomise. Parallel workers share the permutations.

With **Two-tailed** (default), each T contrast of a pair is tested at alpha/2. Maps excluded by QC are left out, and every map that is dropped (no participants.tsv row, a missing value) is listed in the log.

Option (window)	Batch field	Default	Meaning and advice
Analysis name	`name`	`(empty)`	A name for this analysis, letters and digits only, e.g. PCCzSCHZvsCTL. Results go to <output folder>/group/<name>/ and every file starts with it, so different analyses of the same data stay apart. Running again with the same name and settings skips the analysis (tick Overwrite to redo it).
Derivatives folder	`input_dir`	`(empty)`	The SeeCAT derivatives folder that holds the per-subject maps, e.g. D:/study/bids/derivatives/seecat (the output of Seed FC or Degree). Once it is set, the Maps fields below list the values found there.
Maps	`maps`	`(empty)`	Which per-subject map enters the analysis, by the parts of its file name. Every subject must have exactly one matching map (two for paired designs, one per level). Map (desc): the seed or measure, e.g. PCC (seed FC) or degPosBinR0p2Sm6 (degree). Statistic (stat): z (Fisher z of r, or z-scored degree; the usual choice) or raw. Suffix: connectivity for SeeCAT's maps. Space and Task: needed only when several are present. Other: further entities as key=value, comma separated, e.g. ses=1, run=1. Open a list to pick from the values in the derivatives folder, or type a value. Empty fields match anything. The line below the fields counts the matching maps. Example: sub-01_task-rest_space-MNI152NLin2009cAsym_desc-PCC_stat-z_connectivity.nii.gz is chosen by Map PCC, Statistic z, Suffix connectivity.

Option (window)	Batch field	Default	Meaning and advice
participants.tsv advanced	`participants`	`(empty)`	The table of subject information (group, age, sex, …) with a participant_id column (sub-01, …). Empty: participants.tsv of the BIDS dataset the derivatives come from. Factor, Regressors and Covariates name its columns. Subjects with an empty or n/a value in a used column are left out and listed in the log.
Design	`design`	`one_sample`	one_sample: is the mean map different from 0? (e.g. the PCC network of all subjects). Contrasts pos and neg. two_sample: do two groups differ? Factor = the participants column of the group; contrasts AgtB and BgtA. paired: two maps per subject (e.g. ses-1 and ses-2, before and after); Factor = the entity that differs (ses, desc, …). anova1: three or more groups; F contrast of equal means. anova1_rm: three or more maps per subject (repeated measures); F contrast. regression: does the map vary with a score (age, symptoms)? Regressors = the columns; contrasts <name>pos and <name>neg. custom: your own design matrix (Design TSV) and Contrasts. `one_sample` | `two_sample` | `paired` | `anova1` | `anova1_rm` | `regression` | `custom`
Factor	`factor`	`(empty)`	two_sample and anova1: the participants.tsv column that holds the group, e.g. diagnosis. paired and anova1_rm: the file-name entity whose value differs between a subject's maps, e.g. ses (sessions) or desc (two seeds or measures). Not used by the other designs.
Levels	`levels`	`(empty)`	The levels of Factor in the order the contrasts use, comma separated, e.g. CONTROL, SCHZ: AgtB is then CONTROL > SCHZ. Levels not listed are left out, so this also picks two groups out of several. Empty: all levels, in the order found.
Regressors	`regressors`	`(empty)`	regression only: participants.tsv columns of the scores tested, comma separated, e.g. age or PANSS_total, MoCA. They are mean-centred and fitted in one model, so each is tested adjusted for the others.
Covariates	`covariates`	`(empty)`	participants.tsv columns to adjust for (not tested), comma separated, e.g. age, sex, mean_fd. Numbers are mean-centred; text columns (e.g. sex, site) become centred dummy columns. The log warns when a column is strongly collinear with others (VIF > 5). Paired and repeated-measures designs cannot take between-subject covariates.

Option (window)	Batch field	Default	Meaning and advice
Design TSV advanced	`design_tsv`	`(empty)`	custom design only: a tab-separated table with a path column (the map of each row, absolute or relative to the table) and one numeric column per design column, e.g. path A B age sub-01_map.nii 1 0 23 ... Contrasts then have one weight per design column, in this order.
Contrasts advanced	`contrasts`	`(empty)`	One row per contrast: a name (letters and digits), type t or F, and the weights, one per design column (see <name>_design.tsv of a first run for the columns and their order). F weights have one row per line, separated by ";", e.g. [1 -1 0; 0 1 -1]. Empty: the design's defaults (both directions of each T test, or the F test of equal means).
Variance advanced	`variance`	`auto`	Whether the groups or levels may have different variances (and, for repeated measures, correlated levels), as SPM's non-sphericity. auto (default): equal, except unequal for anova1_rm. equal: one variance for all (ordinary least squares). unequal: estimated with ReML over the strongly responding voxels, as SPM does. `auto` | `equal` | `unequal`
Mask	`mask`	`gm`	The voxels tested. Grey matter (gm, default): SeeCAT's grey-matter mask (SPM TPM GM > 0.2). Whole brain (brain): no extra mask. Mask file: your own NIfTI image in MNI space (voxels > 0), e.g. a region of interest for a small-volume analysis; choose it with the folder button. In every case only voxels that are non-zero in every subject's map are kept. A smaller mask means a milder correction.

Option (window)	Batch field	Default	Meaning and advice
Correction	`inference`	`cluster_fwe`	How the maps are corrected for multiple comparisons; each ticked method gives its own thresholded map and cluster table (desc-clusterFWE, -voxelFWE, -voxelFDR, -tfceFWE). cluster_fwe (default): clusters above the cluster-forming p, kept when their size has FWE p < Alpha (random field theory, as SPM). voxel_fwe: voxels with FWE p < Alpha (random field theory, as SPM); strict. voxel_fdr: voxels with false discovery rate < Alpha (Benjamini-Hochberg). tfce_fwe: threshold-free cluster enhancement with FWE p from permutations (as FSL randomise); no cluster-forming threshold to choose, robust, but slow: see TFCE permutations and Parallel workers. Unthresholded T/F, Z, p and effect maps are always written. `cluster_fwe` | `voxel_fwe` | `voxel_fdr` | `tfce_fwe`
Cluster-forming p	`cdt_p`	`0.001`	cluster_fwe only: voxels below this uncorrected p form the clusters. 0.001 (default) is the recommended value; above 0.001 random field theory clusters are too liberal (Eklund et al. 2016), and the log warns.
Alpha	`alpha`	`0.05`	The corrected significance level of every method, e.g. 0.05 (default). With Two-tailed, each T contrast is tested at Alpha/2.
Two-tailed (alpha/2 per T contrast)	`two_tailed`	`true`	The default T contrasts come in pairs (AgtB and BgtA, pos and neg), one per direction. Ticked (default): each is tested at Alpha/2, so the pair together is a two-tailed test at Alpha. Untick only for a direction chosen before looking at the data.
TFCE permutations	`tfce.n_perm`	`5000`	tfce_fwe only: how many permutations (sign flips for one_sample, swaps within subjects for paired designs) build the null distribution. 5000 (default) as FSL randomise; the smallest possible p is 1/permutations. When fewer distinct permutations exist (e.g. 2^n sign flips of n maps), all are used. Time grows with the permutations, the contrasts and the mask: several minutes to an hour. Parallel workers (Run) share the permutations. The random seed is fixed, so results repeat exactly.
TFCE extent exponent E advanced	`tfce.E`	`0.5`	TFCE weight of cluster extent, 0.5 (default, Smith & Nichols 2009; as FSL for 3D data).
TFCE height exponent H advanced	`tfce.H`	`2`	TFCE weight of height, 2 (default, Smith & Nichols 2009; as FSL).

Option (window)	Batch field	Default	Meaning and advice
Leave out QC exclusions	`qc_exclusions`	`true`	Ticked (default): runs (and T1w images) marked "exclude" in the derivatives folder's reports/qc_exclusions.tsv are left out, and the log lists them. That list is made by the group QC report after Preprocessing or fMRIPrep denoising, from the rules there (head motion, coverage, too little data); confirm or change each decision on the QC reports page. Untick to use every run, e.g. to compare results with and without the excluded data.
Subjects advanced	`subjects`	`(empty)`	Only these subjects, labels without "sub-", comma separated. Empty: every subject with a matching map.

Outputs

```
group/<name>/
  <name>_design.tsv                                    maps and design columns
  <name>_desc-mask_mask.nii.gz                         voxels tested
  <name>_contrast-<c>_stat-t_statmap.nii.gz          T (or F) map, with df and smoothness in the
sidecar
  <name>_contrast-<c>_stat-z_statmap.nii.gz          Z of the same p
  <name>_contrast-<c>_stat-p_statmap.nii.gz          -log10 p
  <name>_contrast-<c>_stat-effect_statmap.nii.gz     contrast estimate
  <name>_contrast-<c>_desc-clusterFWE_stat-t_statmap.nii.gz   thresholded map
  <name>_contrast-<c>_desc-clusterFWE_dseg.nii.gz             surviving clusters, numbered
  <name>_contrast-<c>_desc-clusterFWE_clusters.tsv            cluster table (SPM style)
  <name>_contrast-<c>_desc-tfce_stat-tfce_statmap.nii.gz     TFCE values (tfce_fwe)
  <name>_contrast-<c>_desc-tfceFWE_stat-p_statmap.nii.gz     -log10 FWE p (tfce_fwe)
  <name>_report.json
```

The cluster table has, per cluster, its size, FWE and uncorrected cluster p, and up to three peaks more than 8 mm apart with statistic, Z, peak p and MNI coordinates.

6.11 Viewer

The viewer shows any NIfTI image on an MNI brain: SeeCAT's statistic maps, seed and degree maps, gradients, VBM maps, or a 4D BOLD series.

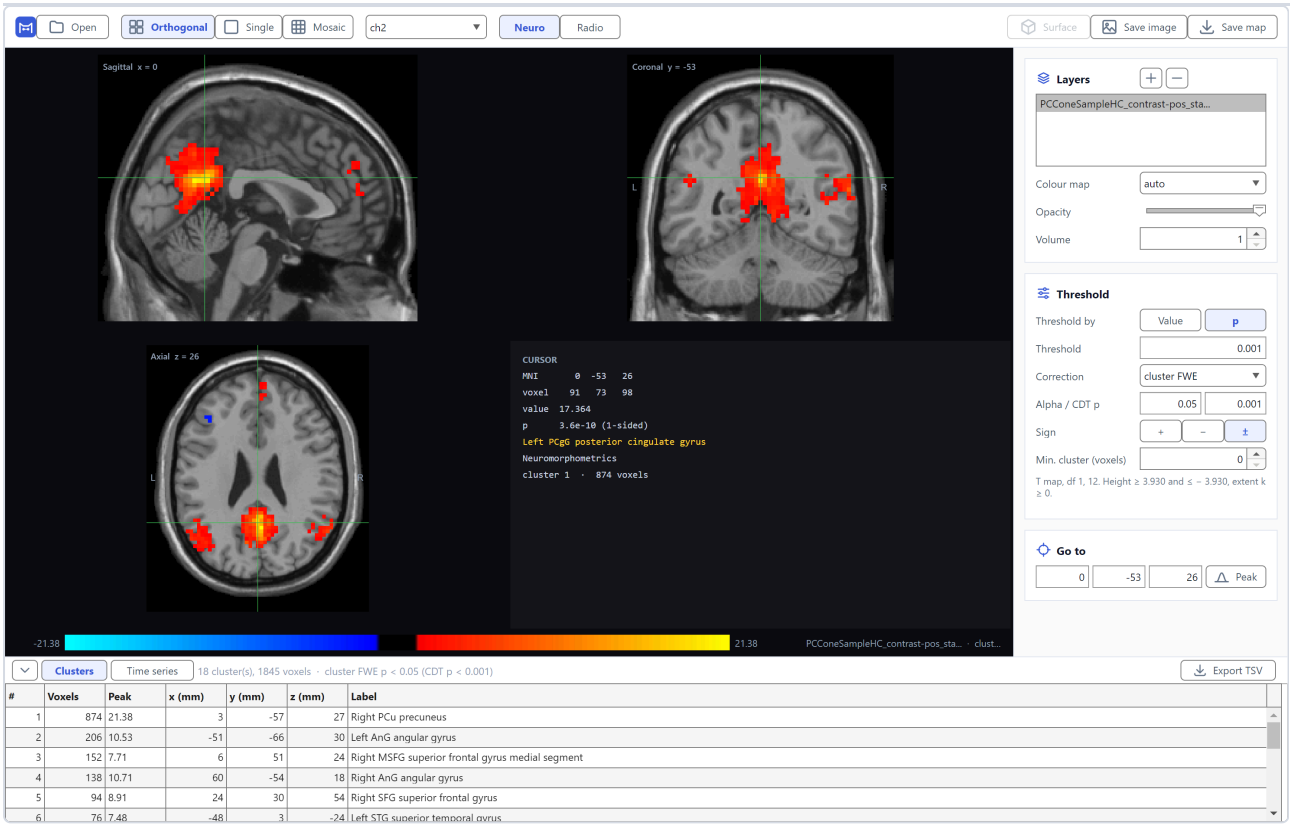


Figure 16. The one-sample PCC result in the viewer: three views, the cursor read-out with the atlas label, and the cluster table.

- **Layouts**: Orthogonal (three views and the cursor read-out), Single (one large view and three thumbnails), Mosaic (slices of one orientation).
- **Underlay**: chosen by the map's space (ch2 for IXI549Space, ICBM152 2009c for MNI152NLin2009cAsym), or picked from the menu.
- **Threshold** by value or by p. For SeeCAT statistic maps (and SPM-style maps with df in the header) p thresholds can be corrected (cluster FWE, voxel FWE, voxel FDR) with the same code as Statistics, so the viewer and the pipeline agree. Sign (+, −, ±), minimum cluster size, colour map and opacity.
- **Clusters**: the table of clusters with their peaks and atlas labels (Neuromorphometrics from SPM; AAL and Brodmann when BrainNet Viewer is installed); clicking a row moves the cursor. Export TSV saves it.
- **Time series**: for a 4D series, the cursor voxel's signal with FD and flagged frames; a slider and Play step through the volumes.
- **Save image** writes the current view as PNG; **Save map** writes the thresholded map as NIfTI; **Surface** shows the map on the cortex in BrainNet Viewer.
- **Neuro / Radio**: left on the left, or on the right.

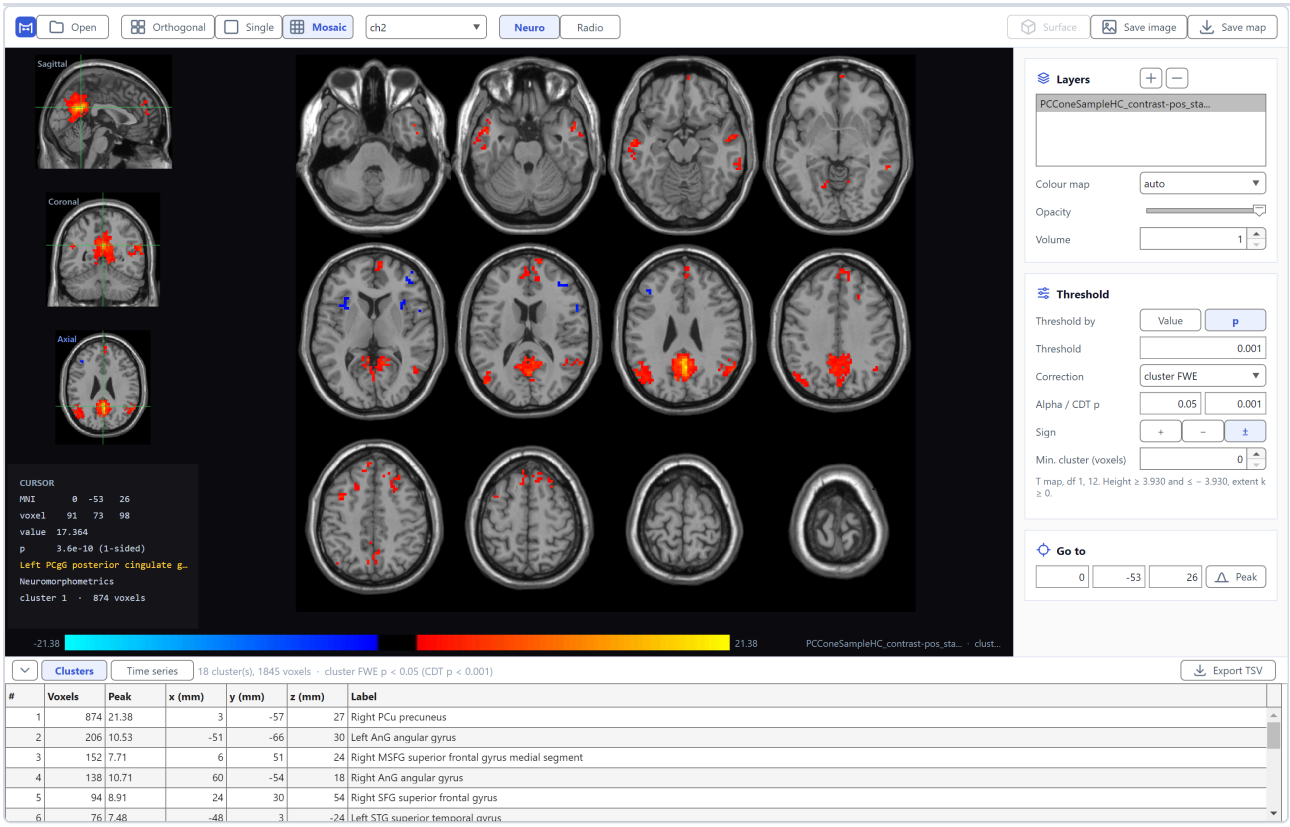


Figure 17. Mosaic view.

```
seecat.viewer('group/PCConeSampleHC/PCConeSampleHC_contrast-pos_stat-t_statmap.nii.gz')
```

6.12 Shared: denoising, censoring and smoothing

These options are the same on Preprocessing and fMRIPrep denoising.

Option (window)	Batch field	Default	Meaning and advice
Denoise	`denoise.enable`	`true`	Remove nuisance signal from every voxel: one regression of a polynomial trend and the chosen Regressors, with data and regressors band-passed by the same filter, so the filter does not bring removed signal back. The result is desc-denoised (unsmoothed), the input of Seed FC and Degree.

Option (window)	Batch field	Default	Meaning and advice
Regressors	`denoise.regressors`	`friston24, wm, csf`	friston24: the 6 head-motion parameters, their derivatives and squares (Friston 1996). wm, csf: mean signal of the eroded white-matter and CSF masks. gs: global signal (global signal regression). Off by default: it removes widespread neural signal too and creates negative correlations; tick it if your study uses it. acompcor: principal components of WM and CSF (aCompCor; see aCompCor components); use it instead of wm and csf. Default friston24 + wm + csf. Any other column of the confounds table can be added in a config file (denoise.regressors). `friston24` | `wm` | `csf` | `gs` | `acompcor`
aCompCor components advanced	`denoise.acompcor_n`	`5`	acompcor only: components taken per tissue (WM and CSF), 5 by default (Behzadi 2007).
Detrend order advanced	`denoise.detrend_order`	`1`	Order of the polynomial trend removed with the regressors: 0 = mean only, 1 = linear (default), 2 = quadratic, up to 3.
Band-pass (Hz)	`denoise.band`	`0.01 0.1`	Frequencies kept, low to high in Hz. 0.01 to 0.1 (default) is the usual resting-state band. Leave the second box 0 for a high-pass only (e.g. 0.01 to 0), and both empty for no filter (e.g. for ALFF-type measures computed later).
High-motion frames	`censor.mode`	`censor`	What happens to volumes with large head motion (framewise displacement above the FD threshold). censor (default): they are filled in from the good volumes before filtering (so the spike is not spread), left out of the regression, and marked in the confounds table; Seed FC and Degree then leave them out. Runs keep their length. interp_keep: the same, but the filled-in volumes are used by later analyses. none: no special handling. `censor` | `interp_keep` | `none`
FD threshold (mm)	`censor.fd_threshold`	`0.5`	Framewise displacement (Power et al. 2012) above which a volume counts as high-motion. 0.5 mm (default) is common; 0.2 to 0.3 is stricter (e.g. for children or patient groups, at the cost of more lost volumes).
Frames before advanced	`censor.before`	`1`	Also flag this many volumes before each high-motion volume, 1 by default.
Frames after advanced	`censor.after`	`2`	Also flag this many volumes after each high-motion volume, 2 by default (motion effects outlast the movement).

Option (window)	Batch field	Default	Meaning and advice
Interpolation advanced	`censor.interp`	`spectral`	How flagged volumes are filled in before filtering. spectral (default): a fit of sinusoids up to the low-pass edge to the good volumes; in tests it changed the good volumes and the seed maps less than spline. spline: cubic spline through the good volumes. `spectral` | `spline`
Minimum minutes advanced	`censor.min_minutes`	`5`	Runs with less good (unflagged) data than this many minutes, 5 by default, are flagged in the QC report so you can exclude them.
Smooth	`smooth.enable`	`true`	Also write a spatially smoothed copy of the denoised data (desc-denoisedSm<FWHM>). The unsmoothed desc-denoised is always kept: Seed FC and Degree use it by default.
FWHM (mm)	`smooth.fwhm`	`6`	Width of the Gaussian kernel, e.g. 6 (default; about twice the voxel size) or 4. Three numbers give a different FWHM per axis.

6.13 Shared: quality control

Option (window)	Batch field	Default	Meaning and advice
Image quality metrics (IQM)	`qc.iqm`	`true`	Compute MRIQC's image quality metrics of the raw T1w and BOLD data while preprocessing (the T1w segmentation, realignment and coregistration are shared, so it adds about half a minute per run). BOLD: tSNR, SNR, ghost-to-signal, EFC, FBER, standardised DVARS, AFNI outlier ratio and quality index, global correlation, FD, smoothness. T1w: CNR, CJV, SNR (tissue and Dietrich), EFC, FBER, bias-field spread, WM-to-max, volume fractions, partial-volume residuals, smoothness, and TIV. The values agree with MRIQC's formulas; the masks (head, air, brain) are SeeCAT's own, so absolute values can differ from an MRIQC run. The group report shows their distribution per site and warns about outliers.
Max. translation (mm)	`qc.max_translation`	`3`	A run whose largest absolute translation (any axis, any frame, from realignment) is above this is suggested for exclusion; 3 mm (default) is the common limit, about one voxel.
Max. rotation (degrees)	`qc.max_rotation`	`3`	As Max. translation, for the largest absolute rotation; 3 degrees by default.
Max. mean FD (mm)	`qc.max_mean_fd`	`0.5`	A run whose mean framewise displacement (Power) is above this is suggested for exclusion: 0.5 mm by default; 0.2 to 0.3 mm is stricter (Parkes et al. 2018 use 0.55 lenient, 0.25 stringent). Above 0.25 mm the report always warns.

Option (window)	Batch field	Default	Meaning and advice
Min. brain coverage	`qc.min_coverage`	`0.9`	Share of the template brain mask that must lie inside the scan's field of view after normalisation (in every frame); 0.9 by default. Lower means part of the brain (often the cerebellum or the top of the cortex) was not scanned; such runs are suggested for exclusion. The report also shows the coverage of the top and bottom separately.
Min. brain Dice	`qc.min_brain_dice`	`0.85`	Registration check: the Dice overlap of the run's brain mask with the template brain mask after normalisation (1 = identical). Runs below this, 0.85 by default, are suggested for exclusion as failed normalisations; good ones are usually above 0.9. The reports also give, for each run: the correlation of the normalised GM, WM and CSF maps with SPM's tissue templates (reg_gm_corr ...), and the Dice of a brain mask taken from the mean EPI's own intensity with the T1w brain (reg_epi_t1_dice, EPI-T1w alignment; about 0.85-0.92 when good). Outliers within a site are warnings.
IQM outlier z advanced	`qc.iqm_z`	`3`	An image quality metric is flagged (a warning, not an exclusion) when its robust z (distance from the median in MADs) within the run's site exceeds this in the bad direction, 3 by default. Sites come from the "site" column of participants.tsv; without it, or with fewer than 5 runs in a site, the whole dataset is the reference.

6.14 Shared: selection of runs

The pages that process runs choose them by subject, session, task and run. Seed FC, Degree, Gradients, Harmonize and Statistics also leave out what quality control excluded:

Option (window)	Batch field	Default	Meaning and advice
Subjects	`subjects`	`(empty)`	Which subjects to process: labels without "sub-", comma separated, e.g. 01, 02, 15. Empty: all subjects found.
Sessions advanced	`sessions`	`(empty)`	Labels without "ses-", comma separated, e.g. 1, 2. Empty: all sessions.
Task	`task`	`(empty)`	The task label of the runs, e.g. rest. Empty: the only task present; if the data hold several tasks, name one.
Runs advanced	`runs`	`(empty)`	Run labels, comma separated, e.g. 1, 2. Empty: all runs.

Option (window)	Batch field	Default	Meaning and advice
Leave out QC exclusions	`qc_exclusions`	`true`	Ticked (default): runs (and T1w images) marked "exclude" in the derivatives folder's reports/qc_exclusions.tsv are left out, and the log lists them. That list is made by the group QC report after Preprocessing or fMRIPrep denoising, from the rules there (head motion, coverage, too little data); confirm or change each decision on the QC reports page. Untick to use every run, e.g. to compare results with and without the excluded data.

6.15 Shared: output and run

Option (window)	Batch field	Default	Meaning and advice
Output folder	`output_dir`	`(empty)`	Where the results go. Empty: the default of the pipeline, next to the input (for the analyses, the derivatives folder itself, so the maps sit beside the runs they come from). Set it to keep the results of different settings apart, e.g. D:/study/derivatives/degree_r03. Logs and the config of every run are written to its logs folder.
Work folder advanced	`work_dir`	`(empty)`	Intermediate files, one subfolder per unit. Empty: seecat_work next to the output folder. It is deleted after a unit succeeds (unless Keep work files) and kept after a failure, so a rerun does not redo finished steps. A fast local disk helps.
Parallel workers	`n_workers`	`auto`	How many units (runs or subjects) are processed at the same time, each by a worker of a parallel pool (Parallel Computing Toolbox). auto (the default): as many workers as fit in memory. SeeCAT estimates the memory one unit needs (from the size of the largest run) plus about 1 GB per worker, and uses at most 85% of the free memory, one worker per core and one per unit; the log shows the estimate and the choice. With one unit, or too little memory for two workers, it runs one after another without a pool. 0: one after another, no pool (e.g. to leave the computer free for other work). A number, e.g. 4: at most that many workers, whatever the memory. Without the toolbox SeeCAT runs one after another. A pool that is already open is reused as it is.
Overwrite finished units advanced	`overwrite`	`false`	A unit that already finished with the same settings is skipped. Tick to compute it again anyway. A unit whose settings changed is always redone.
Keep work files advanced	`keep_work`	`false`	Keep the work folder of units that succeeded, e.g. to inspect intermediate images.

7 Worked examples

This chapter runs a complete study and reports what SeeCAT produced, with the real numbers. Every figure in this manual comes from these runs. The script that ran them is in section 8.2; the configurations can also be loaded into the window with **Load**.

7.1 The example dataset

A public test subset of a multi-site depression study: 30 subjects from three sites (CSU, PKU, SCU), five healthy controls (HC) and five patients with major depression (MDD) at each site. Every subject has one T1-weighted image and one resting-state run (TR 2 s). The three per-site folders were merged into one BIDS dataset with a `participants.tsv`:

```
participant_id   site   group   age   sex
sub-CSUHC0001    CSU    HC      28    M
sub-CSUHC0002    CSU    HC      34    M
…
```

> **Simulated covariates.** Age and sex in this table were generated at random for the demonstration. They are fine as covariates, but no real effect of age or sex is in the data: a significant age effect would be a false positive.

All runs used `n_workers = 'auto'` on a 32-core workstation with about 200 GB of free memory. The outputs go to `manual_run/derivatives/seecat`.

7.2 Example A: seed connectivity of the default-mode network

From BIDS to a group map of posterior cingulate (PCC) connectivity, and a comparison of the two groups.

A1. Check the data

On **Data check**, choose the BIDS folder and press **Scan**. The table lists 30 runs, one per subject, with their volumes and TR (section 6.2). Nothing is missing, so preprocessing can start.

A2. Preprocess

On **Preprocessing**: BIDS folder `bids2`, Output folder `manual_run/derivatives/seecat`, tick **VBM maps** (for Example B), Parallel workers `auto`; all other options at their defaults. Press **Run**.

```
Parallel workers (auto): about 5.5 GB per worker (4.5 GB task + 1 GB worker), 0.5 GB kept by the
main MATLAB,
    206.4 GB free: 30 worker(s) (32 cores, 30 task(s))
Parallel pool: A pool of 30 workers did not start (…). Started 15 workers instead.
QC: …/derivatives/seecat/reports/prep_qc.html
Tissue volumes: …/derivatives/seecat/reports/anat_volumes.tsv
Done: 30 ok, 0 skipped, 0 failed.
```

The memory estimate allowed 30 workers. Windows could not start that many at once, so SeeCAT retried with half and noted it in the log. The 30 runs took about 19 minutes.

A3. Review quality

The group report `reports/prep_qc.html` (figure in section 6.5) suggests excluding four runs:

Run	Reason	Mean FD (mm)	Largest translation (mm)	Minutes kept
sub-CSUHC0004	translation > 3 mm; too little data after censoring	0.300	3.02	3.9
sub-CSUHC0005	translation > 3 mm; too little data after censoring	0.273	3.97	4.4
sub-CSUMDD0004	too little data after censoring	0.237	1.13	4.1
sub-PKUMDD0002	too little data after censoring	0.273	0.68	4.9

Other runs have warnings but are kept. Examples:

- sub-PKUHC0001: DVARS and AQI outliers, and a low grey-matter correlation with the TPM (0.637, the lowest of the study). Open its run report and look at the carpet plot and registration before deciding.
- sub-PKUMDD0005: mean FD 0.281 mm, above the 0.25 mm warning level, but enough clean data remain.
- sub-SCUHC0001: an FBER outlier. Its run report (figure in section 6.5) shows one movement about 2.5 minutes into the run; the six censored frames around it are shaded.

The brain mask Dice with the template is 0.956 to 0.984 for all runs, so no registration failed. We keep the suggestions as they are. The QC reports page shows the same list and can change any decision before the next step.

Every later step reads the exclusion list and logs what it leaves out:

```
Left out by QC (reports/qc_exclusions.tsv): sub-CSUHC0004_task-rest, sub-CSUHC0005_task-rest,
    sub-CSUMDD0004_task-rest, sub-PKUMDD0002_task-rest
```

This leaves 26 runs: 13 HC (CSU 3, PKU 5, SCU 5) and 13 MDD (CSU 4, PKU 4, SCU 5).

A4. Seed FC

On **Seed FC**, add two spheres of 6 mm radius: PCC at (0, −53, 26) and mPFC at (0, 52, −6). The 26 runs took 7 seconds. Each run gets r and Fisher z maps:

```
sub-SCUHC0001_task-rest_space-IXI549Space_res-3_desc-PCC_stat-z_connectivity.nii.gz
sub-SCUHC0001_task-rest_space-IXI549Space_res-3_desc-mPFC_stat-z_connectivity.nii.gz
```

A5. Harmonize

On **Harmonize**: Map `PCC`, Statistic `z`, Site column `site`, Method `combat`, Covariates kept `group, sex, age`.

```
[harmonizeComBat_PCCz_3821b739] 26 maps, 45150 voxels, 3 sites (CSU: 7, PKU: 9, SCU: 10)
[harmonizeComBat_PCCz_3821b739] kept effects: group_MDD, sex_M, age
[harmonizeComBat_PCCz_3821b739] voxels with a site effect (p < 0.05): 9.9% before, 0.1% after
```

Before harmonisation, 9.9% of grey-matter voxels had a site effect, about twice the 5% expected by chance. After ComBat, 0.1% did. The new maps are named `desc-PCCComBat`. With sites this small, SeeCAT also logs the small-site warning of section 6.9. Here we keep ComBat for the FC maps; Example B shows when this matters.

A6. One-sample test in the healthy controls

Is PCC connectivity positive in the 13 controls? On **Statistics**: Name `PCConeSampleHC`, Map `PCCComBat`, Statistic `z`, Design `one_sample`, Subjects: the 13 HC labels, Correction `cluster_fwe` and `tfce_fwe`.

```
[PCConeSampleHC] 13 maps, 45150 voxels in the analysis mask
[PCConeSampleHC] df 12.00; FWHM [6.31 6.48 7.38] mm; resels 2812.9
[PCConeSampleHC] TFCE: 5000 permutations (flip), 2 contrast(s), 16 worker(s), 93 s
[PCConeSampleHC] pos, cluster_fwe: height 3.930, 12 cluster(s) / 1705 voxel(s) survive
[PCConeSampleHC] pos, tfce_fwe: height 2.780, 33 cluster(s) / 3867 voxel(s) survive
[PCConeSampleHC] neg, cluster_fwe: height 3.930, 6 cluster(s) / 140 voxel(s) survive
[PCConeSampleHC] neg, tfce_fwe: height 6.340, 4 cluster(s) / 7 voxel(s) survive
```

The cluster-forming threshold $p < 0.001$ is $t \geq 3.93$ at 12 degrees of freedom. The largest positive clusters are those of the default-mode network:

Voxels	Peak t	Peak Z	x, y, z (mm)	Region
874	21.38	6.53	3, −57, 27	precuneus / posterior cingulate (the seed)
206	10.53	5.20	−51, −66, 30	left angular gyrus
152	7.71	4.55	6, 51, 24	medial superior frontal gyrus
138	10.71	5.23	60, −54, 18	right angular gyrus
94	8.91	4.85	24, 30, 54	right superior frontal gyrus

TFCE finds the same network and extends it (3867 voxels). Its height is the smallest t that survives (2.78), so lower than the cluster-forming threshold. In the cluster file, a row with `n/a` voxels is a further peak of the cluster above it (e.g. −9, −69, 27 and 3, −63, 39 in the PCC cluster). Open the result in the viewer (figures in section 6.11): at p 0.001 with cluster FWE and both signs, the viewer shows the same 18 clusters (12 positive, 6 negative; 1845 voxels).

A7. Two-sample test, HC versus MDD

Name `PCCtwoSample`, Design `two_sample`, Factor `group`, Levels `HC, MDD`, Covariates `age, sex`.

```
[PCCtwoSample] 26 maps, 45150 voxels in the analysis mask
[PCCtwoSample] df 22.00; FWHM [6.09 6.23 7.18] mm; resels 3115.0
[PCCtwoSample] HCgtMDD, cluster_fwe: height 3.505, 0 cluster(s) / 0 voxel(s) survive
[PCCtwoSample] MDDgtHC, cluster_fwe: height 3.505, 0 cluster(s) / 0 voxel(s) survive
```

No cluster survives in either direction. With 13 subjects per group this is expected and is not evidence that the groups have the same connectivity. The unthresholded t map (`…_contrast-HCgtMDD_stat-t_statmap`) can still be shown in the viewer, or used to plan a larger study.

7.3 Example B: grey-matter volume and age (VBM)

Because **VBM maps** was ticked in A2, every T1w already has modulated grey-matter maps in MNI space (1.5 mm), smoothed with 8 mm:

```
sub-SCUHC0001_space-IXI549Space_res-1p5_label-GM_desc-modulatedSm8_probseg.nii
```

Total intracranial volume per subject is in `reports/anat_volumes.tsv`. Statistics and Harmonize read it automatically as the column `tiv_ml`. The QC exclusions are for fMRI runs, so all 30 T1w images are used here.

B1. Regression on age

On **Statistics**: Name `VBMage`, Map `modulatedSm8`, Label `GM`, Suffix `probseg`, Design `regression`, Regressors `age`, Covariates `sex, tiv_ml, site`, Mask `gm`.

```
[VBMage] 30 maps, 414177 voxels in the analysis mask
[VBMage] df 24.00; FWHM [11.43 11.81 11.25] mm; resels 797.0
[VBMage] agepos, cluster_fwe: height 3.467, 0 cluster(s) / 0 voxel(s) survive
[VBMage] ageneg, cluster_fwe: height 3.467, 0 cluster(s) / 0 voxel(s) survive
```

The site column is text, so it enters the model as two indicator columns. With intercept, age, sex, TIV and site, that leaves 30 − 6 = 24 degrees of freedom. Since the ages are random, no effect is the correct result.

B2. Why not ComBat here

Site differences in VBM are large: in 49.8% of grey-matter voxels before harmonisation (0.2% after ComBat with group, sex, age and TIV kept). It seems natural to harmonise first. We ran the same regression after each choice:

Maps	Site in the model	Clusters for age (cluster FWE)
unharmonised	as a covariate	0
ComBat (`combat`)	no	**3**
ComBat (`combat`)	as a covariate	**2**
location only (`combat_mean`)	no	0
linear regression (`lm`)	no	0

The age effect is not real, so the ComBat clusters are false positives. ComBat rescales each site's residual variance using only ten subjects per site. That makes the harmonised data look less variable than they are, and tests of the kept covariates become too liberal (Nygaard et al. 2016). Methods that only remove site means do not have this problem. With fewer than about 20 subjects per site:

- use `combat_mean` or `lm`, or keep the maps unharmonised with site as a covariate;
- compare any ComBat result with the unharmonised one before reporting it.

SeeCAT writes this warning in the Harmonize log and report.

7.4 Example C: degree centrality and gradients

C1. Degree

Degree with its defaults: grey-matter mask, positive correlations above r = 0.2, binary and weighted degree, also as z-scored maps, smoothed with 6 mm.

```
Parallel workers (auto): about 3.9 GB per worker (2.9 GB task + 1 GB worker), …: 26 worker(s)
Done: 26 ok, 0 skipped, 0 failed.
```

The 26 runs took about 2 minutes. Each run gets four maps, each as raw values and as z:

```
…_desc-degPosBinR0p2_stat-z_connectivity.nii.gz        binary, unsmoothed
…_desc-degPosWeiR0p2Sm6_stat-z_connectivity.nii.gz     weighted, smoothed: the usual choice for
statistics
```

Harmonising `degPosWeiR0p2Sm6` z (with group, sex and age kept) reduced voxels with a site effect from 17.2% to 0.0%. Degree is more site-dependent than seed FC here, probably because it summarises the whole connectivity distribution, which depends on noise and smoothness.

C2. Gradients

Gradients with the defaults: Nodes `voxels` with Node size 4 mm (the unit is named `gradientsVox4`), 10 components with 3 maps written, Alignment `group`.

```
Parallel workers (auto): about 8.2 GB per worker (7.2 GB task + 1 GB worker), 7.7 GB kept by the
main MATLAB,
    187.4 GB free: 18 worker(s) (32 cores, 26 task(s))
[gradientsVox4] 26 run(s), 21972 nodes (Vox4); about 7.2 GB of memory per run being computed
[gradientsVox4] node series done; 19564 of 21972 nodes have data in every run
[gradientsVox4] gradients of 26 run(s) in 413 s (15 worker(s))
[gradientsVox4] finished (472.0 s)
```

Each run needs a 21972 × 21972 connectivity matrix, so memory, not cores, set the number of workers (18). The reference is the gradient of the group-mean matrix. Each run is aligned to it, a new reference is formed from the mean of the aligned runs, and this repeats until the reference no longer changes.

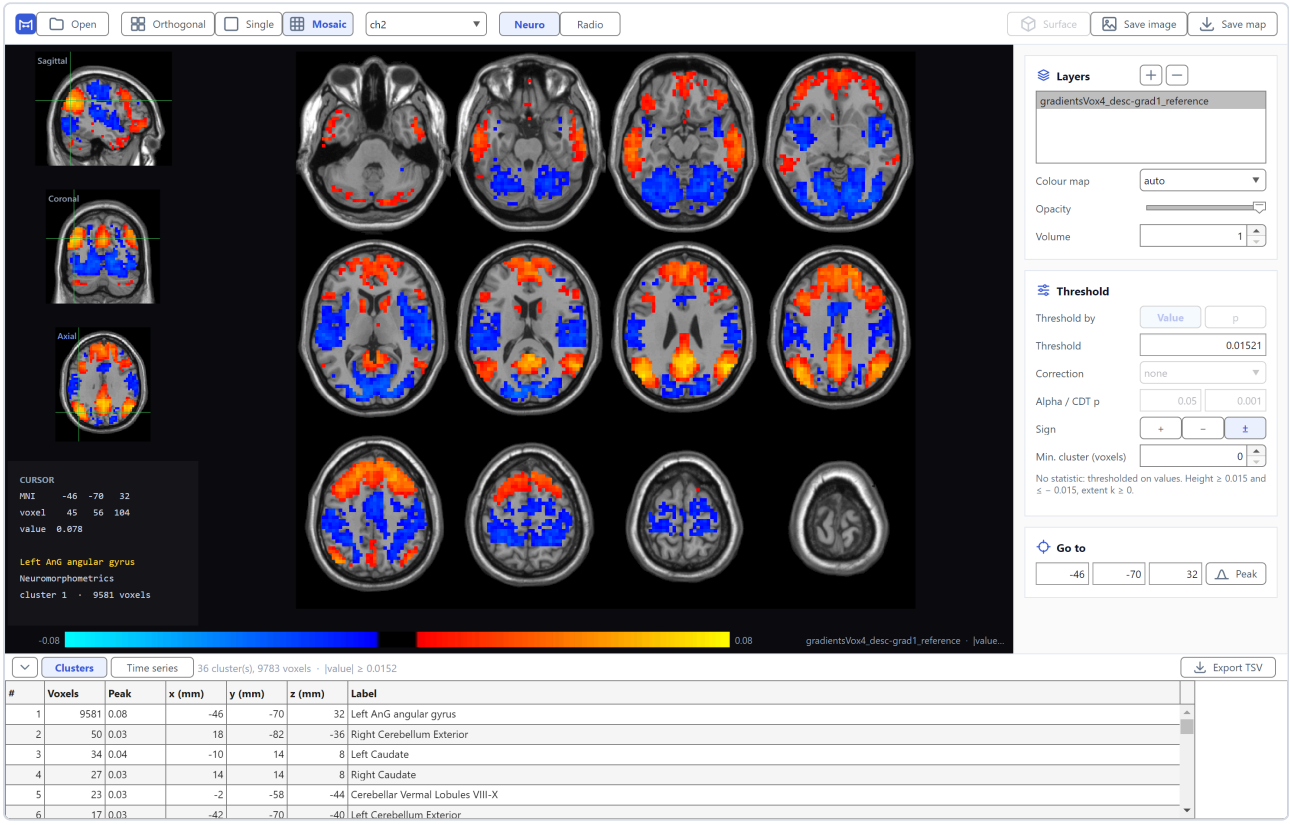


Figure 18. The group reference of gradient 1 (group/gradientsVox4/gradientsVox4_desc-grad1_reference). It runs from visual and sensorimotor cortex (blue) to the default-mode network (red), the principal gradient of Margulies et al. (2016).

`group/gradientsVox4/gradientsVox4_lambdas.tsv` gives, per run:

- the eigenvalues;
- the variance explained by gradient 1: 0.18 to 0.31, median 0.24;
- the correlation of the run's aligned gradient 1 with the reference: 0.43 to 0.69, median 0.58.

A run with a much lower correlation than the others deserves a look at its QC report. The per-run maps (`desc-grad1Vox4` … `grad3Vox4`) can be harmonised and tested like any other map.

7.5 Time for the whole study

Step	Units	Workers	Time
Preprocessing (with IQMs and VBM)	30 runs	15	19 min
Seed FC (2 seeds)	26 runs	26	7 s
Degree	26 runs	26	2 min
Gradients (Vox4)	26 runs	15	8 min
Harmonize (3 measures)	26–30 maps each	1	3–10 s each
Statistics, one-sample with TFCE (5000 permutations)	13 maps	16	96 s
Statistics, two-sample / VBM regression	26 / 30 maps	1	2 s / 5 s

8 Batch scripting

Everything in the window is available from MATLAB code. A pipeline is configured by a struct and run by `seecat.run`; the window's Save button writes the same struct as JSON.

8.1 Functions

Function	Purpose
`SeeCAT`	open the main window
`cfg = seecat.config('prep')`	every option of a pipeline at its default (`prep`, `fmriprep`, `fc`, `degree`, `gradient`, `harmonize`, `stat`)
`cfg = seecat.config('file.json')`	load a saved configuration; missing options take their defaults
`seecat.check_config(cfg)`	report every invalid option at once
`seecat.save_config(cfg, 'file.json')`	save a configuration
`S = seecat.run(cfg)`	run a pipeline; `S` is the summary table (one row per unit)
`seecat.run('file.json')`	run a saved configuration
`seecat.bids.scan_dicom(dir)`, `seecat.bids.from_dicom(dir, bids)`	DICOM to BIDS
`R = seecat.qc.runs(folder, 'raw')`	Data check table
`seecat.qc.exclusion_list(folder)`	read the QC exclusion list with metrics; `(folder, 'set', ids, exclude, notes)` changes it
`seecat.viewer(file, ...)`	open the viewer with maps

8.2 A complete study

The scripts of chapter 7, in one file. Each `seecat.run` call can be run on its own; finished units are skipped when run again.

```
bids = 'D:/data/test_seecat/DIDA/bids2';
out  = 'D:/data/test_seecat/DIDA/manual_run/derivatives/seecat';

% 1. Preprocessing with image quality metrics and VBM maps
cfg = seecat.config('prep');
cfg.bids_dir   = bids;
cfg.output_dir = out;
cfg.vbm.enable = true;
cfg.n_workers  = 'auto';
seecat.run(cfg);
% ... review reports/prep_qc.html and the QC reports page ...

% 2. Seed FC
cfg = seecat.config('fc');
cfg.input_dir = out;
cfg.seeds = struct('name', {'PCC', 'mPFC'}, 'center', {[0 -53 26], [0 52 -6]}, ...
                   'radius', {6, 6}, 'file', {'', ''});
cfg.n_workers = 'auto';
seecat.run(cfg);

% 3. Degree and gradients
cfg = seecat.config('degree');   cfg.input_dir = out;   cfg.n_workers = 'auto';   seecat.run(cfg);
cfg = seecat.config('gradient'); cfg.input_dir = out;   cfg.n_workers = 'auto';   seecat.run(cfg);

% 4. Harmonize the PCC maps across the three sites
cfg = seecat.config('harmonize');
cfg.input_dir  = out;
cfg.maps       = struct('desc', 'PCC', 'stat', 'z');
cfg.site       = 'site';
cfg.covariates = {'group', 'sex', 'age'};
seecat.run(cfg);                                % writes desc-PCCComBat maps

% 5. One-sample t test in the HC group, with TFCE
P  = seecat.io.read_tsv(fullfile(bids, 'participants.tsv'));
hc = erase(P.participant_id(strcmp(P.group, 'HC')), 'sub-')';
cfg = seecat.config('stat');
cfg.name      = 'PCConeSampleHC';
cfg.input_dir = out;
cfg.maps      = struct('desc', 'PCCComBat', 'stat', 'z');
cfg.design    = 'one_sample';
cfg.subjects  = hc;
cfg.inference = {'cluster_fwe', 'tfce_fwe'};
cfg.n_workers = 'auto';
seecat.run(cfg);

% 6. VBM: grey matter and age, unharmonised maps with site as a covariate
cfg = seecat.config('stat');
cfg.name       = 'VBMage';
cfg.input_dir  = out;
cfg.maps       = struct('desc', 'modulatedSm8', 'label', 'GM', 'suffix', 'probseg');
cfg.design     = 'regression';
cfg.regressors = {'age'};
cfg.covariates = {'sex', 'tiv_ml', 'site'};
cfg.mask       = 'gm';
seecat.run(cfg);

seecat.viewer(fullfile(out, 'group', 'PCConeSampleHC', 'PCConeSampleHC_contrast-pos_stat-
t_statmap.nii.gz'));
```

8.3 Configuration files

A configuration file is the struct as JSON. Only the options you change need to be in it; the others take their defaults. Numbers with Inf are written as text (`"[0 75 Inf]"`) and read back as numbers.

```
{
  "pipeline": "stat",
  "name": "PCCtwoSample",
  "input_dir": "D:/data/test_seecat/DIDA/manual_run/derivatives/seecat",
  "maps": {"desc": "PCCComBat", "stat": "z"},
  "design": "two_sample",
  "factor": "group",
  "levels": ["HC", "MDD"],
  "covariates": ["age", "sex"]
}
```

Run several configurations in sequence:

```
for f = {'prep.json', 'fc.json', 'harmonize.json', 'stat.json'}
    seecat.run(fullfile('D:/study/configs', f{1}));
end
```

8.4 Running on a server

Without a display, start MATLAB with `-batch`:

```
matlab -batch "addpath('/opt/spm'); addpath('/opt/SeeCAT/SeeCAT');
seecat.run('/data/configs/prep.json')"
```

Logs are written as the run goes (`<output>/logs`), so progress can be followed with `tail -f`.

9 Output reference

9.1 The derivatives folder

```
derivatives/seecat/
  dataset_description.json                names the source dataset (SourceDatasets)
  sub-<label>/[ses-<label>/]
    anat/                                 T1w outputs (Preprocessing)
    func/                                 run outputs and per-run maps (all pipelines)
    sub-<label>_…_desc-qc_report.html     run reports, qc_metrics.json, figures/
  group/
    <name>/                               Statistics
    harmonize<label>_<maps>/                     Harmonize: maps.tsv, report.json
    gradients<nodes>/                     Gradients: reference maps, lambdas.tsv
  reports/
    prep_qc.html, prep_qc.tsv             group QC (or fmriprep_qc.*)
    qc_exclusions.tsv                     the exclusion list
    anat_volumes.tsv                      GM, WM, CSF, TIV per T1w
    <pipeline>_<time>_summary.tsv         one per run of a pipeline
  logs/
    <pipeline>_<time>.log, _config.json   main log and the configuration used
    <pipeline>_<time>/<unit>.log          one log per unit
    done/                                 finished units and their settings (for skipping)
    outputs/                              files written by each unit
seecat_work/                              work files (next to the output folder)
```

9.2 File names

File	Written by	Content
`*_desc-preproc_bold`	Preprocessing	normalised BOLD, not denoised
`*_desc-denoised_bold`	Preprocessing, fMRIPrep denoising	denoised, band-passed, unsmoothed
`*_desc-denoisedSm<fwhm>_bold`	same	denoised and smoothed
`*_desc-brain_mask`	same	brain mask in output space
`*_desc-confounds_timeseries.tsv`	same	confounds; the `censor` column marks frames left out of later analyses
`*_label-GM_probseg`	Preprocessing	native tissue map; sidecar holds the tissue volumes and TIV

File	Written by	Content
`*_space-IXI549Space_res-1p5_label-GM_desc-modulated[Sm8]_probseg`	Preprocessing (VBM)	normalised, modulated GM (and WM, CSF)
`*_desc-<seed>_stat-r|z_connectivity`	Seed FC	seed correlation, Fisher z
`*_desc-deg…_stat-raw|z_connectivity`	Degree	degree or strength
`*_desc-grad<k><nodes>_connectivity`	Gradients	aligned gradient k of the run
`*_desc-<desc>ComBat_*`	Harmonize	the harmonised map
`group/<name>/…_statmap`	Statistics	statistic maps (t, F, z, p, effect, TFCE)

9.3 Confounds table

Columns	Content
`trans_x … rot_z`, `*_derivative1`, `*_power2`, `*_derivative1_power2`	Friston 24 motion regressors (rotations in radians)
`framewise_displacement`	Power FD (mm; head radius 50 mm)
`global_signal`, `white_matter`, `csf`	mean signals of the brain, WM and CSF masks
`c_comp_cor_XX`, `w_comp_cor_XX`	aCompCor components of CSF and WM
`motion_outlier_XX`	one column per flagged frame (1 at that frame)
`censor`	1 = frame left out by Seed FC, Degree and Gradients

9.4 JSON sidecars

Each image's sidecar describes how it was made. Common fields: `Sources` (the files it came from: BIDS URIs such as `bids:raw:sub-01/func/…` for the per-run outputs; Harmonize and Statistics list the paths of their input maps), `Software`, and step-specific fields, for example `FramesUsed`, `FramesTotal` and `Seed` (Seed FC); `Threshold`, `Sign`, `Measure` (Degree); `Lambdas`, `VarianceExplained`, `Alignment` (Gradients); `Harmonization`, `Site`, `KeptEffects` (Harmonize); `DegreesOfFreedom`, `FWHMmm`, `Resels`, `Contrast`, `Inference` (Statistics).

10 Performance and common questions

10.1 Time and memory

Step	Per run (typical)	Memory
Preprocessing, 3 mm, 200 volumes	4–6 minutes (segmentation about 2)	2–5 GB; with IQMs of a large T1w (e.g. 512×512×156) more
Seed FC	seconds	< 1 GB
Degree, 3 mm grey matter	seconds to a minute	about 3 GB
Gradients	see the table in 6.8	grows with the square of the node count
Harmonize	seconds for the whole study	maps × voxels
Statistics, TFCE	about 0.2 s per permutation and contrast (52 000 voxels)	maps × voxels per worker

With `n_workers = auto`, SeeCAT chooses the number of workers from these estimates and the free memory; the log shows the choice.

10.2 Common questions

A subject failed. Do I have to run everything again?

No. Run the same configuration again: finished subjects are skipped and the failed one resumes at the step where it stopped. The unit's log (`logs/<pipeline>_<time>/<unit>.log`) gives the reason.

I changed an option. What is recomputed?

Every unit whose result depends on it. Changing only Parallel workers or Keep work files recomputes nothing; ticking Overwrite recomputes every unit and clears its work folder. A finished unit is also run again when its inputs changed, when SeeCAT itself changed, or when one of its outputs was deleted; the log says why. Inputs are every file the unit reads, including the files named in the options (seed and ROI images, masks, atlases, image covariates, the participants table), the set of maps selected, the participant values and the runs excluded or restored by QC. A file up to 64 MB is compared by its contents, a larger one (a BOLD series) by its size and date. SeeCAT itself is the code that computes the results and the bundled templates: after an update, finished units are computed again with the new code. Files outside these, for example an SPM installed elsewhere, are not followed; tick Overwrite after changing them.

Can I write the results into the input folder, or run SeeCAT twice at once?

The output folder may not be the BIDS folder or the fMRIPrep folder; SeeCAT refuses it before writing anything. In every other case a unit that would write over one of its own input files stops with an error and leaves the file as it is. Two SeeCAT sessions may work in the same output folder at the same time on different units (for example different subjects); a unit that another session is computing fails with a message naming that session, and is computed when you run the configuration again. If MATLAB was closed while a unit ran, its lock (`logs/locks/`) is removed by the next run on the same computer.

Seed FC says "no desc-denoised bold" or finds no runs.

Check that the derivatives folder is the output of Preprocessing or fMRIPrep denoising and that Denoise was on. If the folder holds runs in more than one space, set Space; with several tasks, set Task. If every run is excluded by QC, check the QC reports page.

Statistics finds several maps per subject.

The map fields must pick exactly one map per subject (two for paired designs). Add Space, Task or Other (e.g. `run=1`) until the line below the fields counts one map per subject.

Where are age, sex and group taken from?

From participants.tsv of the BIDS dataset (see 3.7). Values that are empty or `n/a` drop that subject from the analysis that uses the column, and the log lists it.

Should I smooth before Seed FC?

The default input is the unsmoothed `desc-denoised`; use `denoisedSm6` to follow studies that correlate smoothed data. Degree and Gradients should use unsmoothed data; Degree smooths its maps afterwards.

Why harmonize, and when not?

Harmonize when maps from several sites or scanners enter one analysis and the sites have enough subjects (ComBat works best with 10 or more per site). Keep every effect you will test among the covariates. With a single site, skip it.

How do I report the quality control?

The group report gives the rules, the runs excluded and why; `reports/qc_exclusions.tsv` and `prep_qc.tsv` hold the numbers. Report the exclusion rules, how many runs were excluded per group and site, and mean FD per group.

The window does not respond during a run.

It should: runs happen in a separate MATLAB process. If MATLAB itself is busy (e.g. a long command in the command window), the window waits for it.

11 References

- Ashburner J, Friston KJ (2005). Unified segmentation. NeuroImage 26:839–851.
- Behzadi Y, Restom K, Liau J, Liu TT (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. NeuroImage 37:90–101.
- Buckner RL et al. (2009). Cortical hubs revealed by intrinsic functional connectivity. J Neurosci 29:1860–1873.
- Chen AA et al. (2022). Mitigating site effects in covariance for machine learning in neuroimaging data (CovBat). Hum Brain Mapp 43:1179–1195.
- Coifman RR, Lafon S (2006). Diffusion maps. Appl Comput Harmon Anal 21:5–30.
- Eklund A, Nichols TE, Knutsson H (2016). Cluster failure. PNAS 113:7900–7905.
- Esteban O et al. (2017). MRIQC: advancing the automatic prediction of image quality in MRI from unseen sites. PLoS ONE 12:e0184661.
- Esteban O et al. (2019). fMRIPrep: a robust preprocessing pipeline for functional MRI. Nat Methods 16:111–116.
- Fortin JP et al. (2017). Harmonization of multi-site diffusion tensor imaging data. NeuroImage 161:149–170.
- Fortin JP et al. (2018). Harmonization of cortical thickness measurements across scanners and sites. NeuroImage 167:104–120.
- Freedman D, Lane D (1983). A nonstochastic interpretation of reported significance levels. J Bus Econ Stat 1:292–298.
- Friston KJ et al. (1996). Movement-related effects in fMRI time-series. Magn Reson Med 35:346–355.
- Gorgolewski KJ et al. (2016). The brain imaging data structure. Sci Data 3:160044.
- Hallquist MN, Hwang K, Luna B (2013). The nuisance of nuisance regression. NeuroImage 82:208–225.
- Johnson WE, Li C, Rabinovic A (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 8:118–127.
- Li X, Morgan PS, Ashburner J, Smith J, Rorden C (2016). The first step for neuroimaging data analysis: DICOM to NIfTI conversion (dcm2niix). J Neurosci Methods 264:47–56.
- Lindquist MA, Geuter S, Wager TD, Caffo BS (2019). Modular preprocessing pipelines can reintroduce artifacts into fMRI data. Hum Brain Mapp 40:2358–2376.
- Margulies DS et al. (2016). Situating the default-mode network along a principal gradient of macroscale cortical organization. PNAS 113:12574–12579.
- Nichols TE (2013). Notes on creating a standardized version of DVARS. arXiv:1704.01469.
- Nygaard V, Rødland EA, Hovig E (2016). Methods that remove batch effects while retaining group differences may lead to exaggerated confidence in downstream analyses. Biostatistics 17:29–39.

- Parkes L, Fulcher B, Yücel M, Fornito A (2018). An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI. NeuroImage 171:415–436.
- Pomponio R et al. (2020). Harmonization of large MRI datasets for the analysis of brain imaging patterns throughout the lifespan (ComBat-GAM). NeuroImage 208:116450.
- Power JD et al. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. NeuroImage 59:2142–2154.
- Power JD et al. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. NeuroImage 84:320–341.
- Power JD (2017). A simple but useful way to assess fMRI scan qualities. NeuroImage 154:150–158.
- Smith SM, Nichols TE (2009). Threshold-free cluster enhancement. NeuroImage 44:83–98.
- Vos de Wael R et al. (2020). BrainSpace: a toolbox for the analysis of macroscale gradients in neuroimaging and connectomics datasets. Commun Biol 3:103.
- Worsley KJ et al. (1996). A unified statistical approach for determining significant signals in images of cerebral activation. Hum Brain Mapp 4:58–73.
- Xia M, Wang J, He Y (2013). BrainNet Viewer: a network visualization tool for human brain connectomics. PLoS ONE 8:e68910.
- Zuo XN et al. (2012). Network centrality in the human functional connectome. Cereb Cortex 22:1862–1875.

12 Appendix

12.1 Index of options

Every option of every pipeline, by batch field. The tables in chapter 6 explain each one.

Batch field	Option	Pages
`align`	Alignment	Gradients
`alpha`	Diffusion alpha	Gradients, Statistics
`atlas`	Atlas	Gradients
`bids_dir`	BIDS folder	Preprocessing
`cdt_p`	Cluster-forming p	Statistics
`censor.after`	Frames after	Preprocessing, fMRIPrep
`censor.before`	Frames before	Preprocessing, fMRIPrep
`censor.fd_threshold`	FD threshold (mm)	Preprocessing, fMRIPrep
`censor.interp`	Interpolation	Preprocessing, fMRIPrep
`censor.min_minutes`	Minimum minutes	Preprocessing, fMRIPrep
`censor.mode`	High-motion frames	Preprocessing, fMRIPrep
`contrasts`	Contrasts	Statistics
`covariates`	Covariates kept	Harmonize, Statistics
`covbat_var`	CovBat variance	Harmonize
`denoise.acompcor_n`	aCompCor components	Preprocessing, fMRIPrep
`denoise.band`	Band-pass (Hz)	Preprocessing, fMRIPrep
`denoise.detrend_order`	Detrend order	Preprocessing, fMRIPrep
`denoise.enable`	Denoise	Preprocessing, fMRIPrep
`denoise.regressors`	Regressors	Preprocessing, fMRIPrep
`design`	Design	Statistics
`design_tsv`	Design TSV	Statistics
`distance_edges`	Distance bins (mm)	Degree
`dummy_scans`	Dummy scans	Preprocessing, fMRIPrep
`exclude_local_mm`	Exclude within (mm)	Degree
`factor`	Factor	Statistics

Batch field	Option	Pages
`fmriprep_dir`	fMRIPrep folder	fMRIPrep
`inference`	Correction	Statistics
`input`	Input series	Seed FC, Degree, Gradients
`input_dir`	Derivatives folder	Seed FC, Degree, Gradients, Harmonize, Statistics
`keep_work`	Keep work files	Preprocessing, fMRIPrep, Seed FC, Degree, Gradients, Harmonize, Statistics
`kernel`	Affinity	Gradients
`label`	Output label	Harmonize
`levels`	Levels	Statistics
`maps`	Maps	Harmonize, Statistics
`mask`	Mask	Degree, Gradients, Harmonize, Statistics
`masks.csf_threshold`	CSF probability	Preprocessing
`masks.erode_mm`	Erosion (mm)	Preprocessing
`masks.wm_threshold`	WM probability	Preprocessing
`method`	Method	Harmonize
`n_components`	Components	Gradients
`n_maps`	Maps written	Gradients
`n_workers`	Parallel workers	Preprocessing, fMRIPrep, Seed FC, Degree, Gradients, Statistics
`name`	Analysis name	Statistics
`node_size`	Node size (mm)	Gradients
`nodes`	Nodes	Gradients
`nonlinear`	Nonlinear covariates	Harmonize
`normalise.bounding_box`	Bounding box	Preprocessing
`normalise.enable`	Normalise	Preprocessing
`normalise.method`	Normalisation	Preprocessing
`normalise.voxel_size`	Voxel size (mm)	Preprocessing
`output_dir`	Output folder	Preprocessing, fMRIPrep, Seed FC, Degree, Gradients, Harmonize, Statistics
`overwrite`	Overwrite finished units	Preprocessing, fMRIPrep, Seed FC, Degree, Gradients, Harmonize, Statistics
`participants`	participants.tsv	Harmonize, Statistics

Batch field	Option	Pages
`qc.iqm`	Image quality metrics (IQM)	Preprocessing
`qc.iqm_z`	IQM outlier z	Preprocessing, fMRIPrep
`qc.max_mean_fd`	Max. mean FD (mm)	Preprocessing, fMRIPrep
`qc.max_rotation`	Max. rotation (degrees)	Preprocessing, fMRIPrep
`qc.max_translation`	Max. translation (mm)	Preprocessing, fMRIPrep
`qc.min_brain_dice`	Min. brain Dice	Preprocessing, fMRIPrep
`qc.min_coverage`	Min. brain coverage	Preprocessing, fMRIPrep
`qc_exclusions`	Leave out QC exclusions	Seed FC, Degree, Gradients, Harmonize, Statistics
`realign.enable`	Realign	Preprocessing
`realign.register_to_mean`	Register to mean	Preprocessing
`ref_site`	Reference site	Harmonize
`regressors`	Regressors	Statistics
`res`	Resolution label	fMRIPrep, Seed FC, Degree, Gradients
`runs`	Runs	Preprocessing, fMRIPrep, Seed FC, Degree, Gradients
`seeds`	Seeds	Seed FC
`sessions`	Sessions	Preprocessing, fMRIPrep, Seed FC, Degree, Gradients
`sign`	Correlations	Degree
`site`	Site column	Harmonize
`slice_timing.enable`	Slice timing	Preprocessing
`slice_timing.ref_time`	Reference time	Preprocessing
`slice_timing.slice_times`	Slice times (s)	Preprocessing
`smooth.enable`	Smooth	Preprocessing, fMRIPrep
`smooth.fwhm`	FWHM (mm)	Preprocessing, fMRIPrep
`smooth_fwhm`	Smooth maps	Degree
`space`	Space	fMRIPrep, Seed FC, Degree, Gradients
`sparsity`	Sparsity	Gradients
`spline_df`	Spline df	Harmonize
`subjects`	Subjects	Preprocessing, fMRIPrep, Seed FC, Degree, Gradients, Harmonize, Statistics
`task`	Task	Preprocessing, fMRIPrep, Seed FC, Degree, Gradients
`tfce.E`	TFCE extent exponent E	Statistics

Batch field	Option	Pages
`tfce.H`	TFCE height exponent H	Statistics
`tfce.n_perm`	TFCE permutations	Statistics
`threshold`	r threshold	Degree
`two_tailed`	Two-tailed (alpha/2 per T contrast)	Statistics
`variance`	Variance	Statistics
`vbm.enable`	VBM maps	Preprocessing
`vbm.fwhm`	VBM smoothing (mm)	Preprocessing
`vbm.modulated`	Modulated	Preprocessing
`weight`	Measures	Degree
`work_dir`	Work folder	Preprocessing, fMRIPrep, Seed FC, Degree, Gradients, Harmonize, Statistics
`zscore`	Also z-scored maps	Degree

12.2 Coming from SeeCAT 1.x

1.x	2.0
folders by processing stage (`FunRaw` , `FunImgARW…`)	a BIDS dataset; DICOM to BIDS converts DICOM folders
preprocessing through external toolbox functions	Preprocessing with SPM only; outputs in `derivatives/seecat`
VoxDeg	Degree (same definition: thresholded correlations, binary and weighted)
FuncConnect	Seed FC (spheres and masks, r and Fisher z)
Stat with GRF/AlphaSim correction	Statistics with SPM's random field theory, FDR or TFCE
an external viewer	SeeCAT's own viewer, BrainNet Viewer for surfaces
settings in the GUI only	every setting is a JSON configuration that runs from scripts