**E5912**

**A PRIMER FOR DATA ANALYSIS: PREPROCESSING PIPELINE FOR RS-FMRI**

**Behroze Vachha, MD PhD**

**Memorial Sloan Kettering Cancer Center**

[vachhab@mskcc.org](mailto:vachhab@mskcc.org)

**STEP 1. ORGANIZING YOUR FOLDERS AND SETTING THE PATH FOR SPM**

* Download SPM12 using the following link: <https://www.fil.ion.ucl.ac.uk/spm/software/download/>
* Put SPM12 in your downloads folder. Unzip it and leave it in the DOWNLOADS folder
* Use the links below to download anatomical and functional data from a single subject from the NKI Rockland sample (<http://fcon_1000.projects.nitrc.org/indi/enhanced/>) using the following links (you may need to copy paste the URL if clicking on it doesn’t work):

[http://fcp-indi.s3.amazonaws.com/data/Projects/RocklandSample/RawDataBIDSLatest/sub-A00029126/ses-BAS1/anat/sub-A00029126\_ses-BAS1\_T1w.nii.gz](http://fcp-indi.s3.amazonaws.com/data/Projects/RocklandSample/RawDataBIDSLatest/sub-A00029126/ses-BAS1/anat/sub-A00029126_ses-BAS1_T1w.nii.gz" \t "_blank)

[http://fcp-indi.s3.amazonaws.com/data/Projects/RocklandSample/RawDataBIDSLatest/sub-A00029126/ses-BAS1/func/sub-A00029126\_ses-BAS1\_task-rest\_acq-CAP\_bold.nii.gz](https://nam02.safelinks.protection.outlook.com/?url=http%3A%2F%2Ffcp-indi.s3.amazonaws.com%2Fdata%2FProjects%2FRocklandSample%2FRawDataBIDSLatest%2Fsub-A00029126%2Fses-BAS1%2Ffunc%2Fsub-A00029126_ses-BAS1_task-rest_acq-CAP_bold.nii.gz&data=04%7C01%7Cgohelsu%40shp.rutgers.edu%7C95078db59c0546a8d19208d8ffbf8ca8%7Cb92d2b234d35447093ff69aca6632ffe%7C1%7C0%7C637540545497162143%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C1000&sdata=hSuA%2BjqnTnerIeQqv0JCvBu6wkpzePm%2BSiLxHmAnJdk%3D&reserved=0" \t "_blank" \o "Original URL: http://fcp-indi.s3.amazonaws.com/data/Projects/RocklandSample/RawDataBIDSLatest/sub-A00029126/ses-BAS1/func/sub-A00029126_ses-BAS1_task-rest_acq-CAP_bold.nii.gz. Click or tap if you trust this link.)

* Once you download these you should get two “.nii.gz” files.
  + Unzip the files (in MacOS you can just double click and it will unzip to .nii files).
  + Now, go to your **DOWNLOADS** folder and create a new folder called **SUB01.**
  + Next put both the .nii files into the SUB01 folder and rename them as follows:

Name the file ending **with T1W.nii to “ANAT.nii**”

Name the file ending with **bold.nii to “REST.nii**”

* + So now you have your anatomical and resting state data in the SUB01 folder
* Now open MATLAB
  + Click the third icon from the left (“Browse for folder”) in the top bar
  + ” 🡪navigate to the Downloads folder 🡪Select SUB01🡪 Click open. Now you should see both ANAT.nii and REST.nii in the pane under “Current Folder” in Matlab
* Now we will add SPM12 and SETPATH (skip this step if you already have SPM12 on your path)
  + Click on SETPATH🡪Click on “Add with Subfolders” 🡪Downloads 🡪SPM12🡪Open🡪Save🡪Close
  + Now in matlab command window type

>> rehash (so now you added the path and this command reloads the path)

>> ls (this tells you that the SUB01 folder has ANAT.nii and REST.nii files)

* Launch SPM in Matlab by typing:
  + >> “spm fmri”
* SPM will open up

***Things to do:***

1. *Inspect your Anatomical Images:*

*In SPM, click on “Display”. When the window opens, select ANAT.nii, click done. A window with the anatomical images in three planes will open. Click within each pane and see if there is any obvious motion or distortion.*

*Look for bright and dark spots.*

*b. Inspect your Functional Images:*

*In SPM, click on “Display”. When the window opens, select REST.nii, click done. A window with the resting state functional images in three planes will open. Again look for obvious distortion and bright/dark spots.*

*Now check for motion by right clicking in one of the image panes🡪click on the “Browse” button. A window will open up. Now type 1:120 in the Frames field. All the REST images will open up. Select all the REST images (select the first REST file, press the shift key and select the last REST file) and click “Done”. Now the same window with three panes will open up with a horizontal scrolling bar at the bottom of the window. Use the scrolling bar to quickly move through your images to see if there is a lot of motion or not.*

*Note: Checking for motion as described above may not be possible with older versions of Matlab.*

**STEP 2. SETTING THE ORIGIN/AC ALIGNMENT**

* In SPM, click “Display”
* Select the image you want to display
  + Select “REST.nii 1”
  + Click “Done”
* A new window will pop up with the REST scan images
  + Click the “Origin” button to move the crosshair to the 0 0 0 position on the images
  + **Reset the 0 0 0 position to the anterior commissure of the brain** 
    - Use the mouse to move the crosshair to the anterior commissure
    - The crosshair position will have changed now -> make note of the new coordinates. For example, if it is “-8.3 6.9 -25.4”
    - Enter these coordinates as follows:
      * In right (mm) 🡪 type “8.3”
      * In forward (mm) 🡪 type “-6.9”
      * In up (mm) 🡪 type “25.4”
    - Press “Enter”
    - The crosshair will move
    - Click the “Origin” button again to move the crosshair to the AC and now the crosshair position coordinates will be 0 0 0
  + Click the “Reorient” button
    - In the frame field that has “1”, “1:120” and hit enter
    - Select all the REST.nii files (select the first, press shift key and select the last - 120 files in total)
    - Click “Done”
    - A window will pop up that says, “Save reorientation matrix for future reference?” 🡪 click “yes”
    - A new window will open asking you to save the file as “REST\_reorient”🡪 click the “Save” button
  + Program will run in the background
  + When it is done running, click the “Origin” button again to double check that the crosshair 0 0 0 position has been reset to the AC
* Now repeat the above steps with ANAT.nii (i.e. click Display, select “ANAT.nii”, click Done, etc.).

**STEP 3. SLICE-TIMING CORRECTION**

**For Slice-timing correction you will need to know the number of slices, the TR and the Slice Order for the dataset. For the data provided, the values are:**

1. Number of Slices: 38
2. TR: 2.5s
3. Slice Order: interleaved bottom🡪up formula

* Click the “Slice Timing” menu box🡪 new window will open
* In the new window, double click “Data”
* Double click “Session”
* In the frame field that has “1”, type “1:120” hit enter
* Now select all “Rest” files by clicking the first Rest file, holding “shift” and click the last Rest file
* Click “done”
* “.Session” under “Data” in the Batch Editor should now have “120 files” listed next to it
* Now double click on “Number of Slices”🡪A box will open up. Enter “38” (no. of slices for this dataset)
* Next double click “TR” ”🡪A box will open up. Enter “2.5” (TR for this dataset in seconds)
* Next double click “TA” ”🡪A box will open up. Enter the formula “2.5 – (2.5/38)” .
* Next double click “Slice order” ”🡪A box will open up. Enter the formula “1:2:38 2:2:38”
* Finally double click “Reference Slice” and enter 1
* Once everything is entered click on the green arrow (“run”) in the top bar

**STEP 4. REALIGNMENT**

* In the left-hand box, click the dropdown arrow in the “Realign” menu box
* Select “Realign (Est & Res)” 🡪 new window will open
* In the new window, double click “Data”
* Double click “Session”
* In the frame field that has “1”, type “1:120” hit enter.
* Now select all “aREST.nii” files by clicking the first “aREST.nii” file, holding “shift” and clicking the last “aREST.nii” file
* Click “done”
* “.Session” under “Data” in the Batch Editor should now have “120 files” listed next to it
* Click the green arrow (“run”) in the top bar
* Wait for the program to run in the background
* When done running, a results window will pop up
  + Check the “translation” motion plot
  + The maximum value on the y-axis of the graph should not be >1 otherwise this indicates too much movement and the data will be unusable
* Leave the results window open and move to the next step

***Things to do:***

*For our realignment step we set the Num Passes to “Register to mean” which means the images are registered to the mean of the images after the first realignment. Now re-run the realignment step and click on Num Passes and pick “Register to first”. In the SPM explanation bar read what the differences are. Also did you notice that now it takes a little less time to run than when you pick the “Register to mean”?*

**STEP 5. COREGISTRATION**

* In the left-hand box, click the dropdown arrow in the “Coregister” menu box
* Select “Coregister (Estimate)” 🡪 new window will open
* In the new window, double click “Reference Image”
* Single click “ANAT.nii” and click “Done”
* Reference Image” should list a pathway for “ANAT.nii”
* Double click “Source Image”
* Click “meanaREST.nii,1” and click “Done”
* “Source Image” should list a pathway for “meanaREST.nii,1”
* Double click “Other Images”
  + In the frame field that has “1”, type “:120” after it (1:120) and hit enter
* In the new box, select all “raRest” files by clicking the first raRest file, holding “Shift” and click the last raRest file
* Click “done”
* “Other Images” should now have “120 files” listed next to it
* Click the green arrow (“run”) in the top bar
* Wait for the program to run in the background
* When done running, a results window will pop up, titled “Normalised Mutual Information Coregistration”

***Things to do:***

*In the window that opens click around the edges of the ventricles in several scans to ensure that the anatomical and functional map well to one another.*

**STEP 6. SEGMENTATION**

* In the SPM window click “Segment”
* In the new window that pops up, double click “Volumes” (listed under “Data”)
* In the new window that pops up, click “ANAT.nii” and then click “Done”
* Under “Tissues” in the first right-hand box, find the first (of 6) tissue probability maps – a pathway should be listed to the right ending in “1”
  + Click “Native Tissue” underneath it
    - In the new window that pops up, select “Native + Dartel Imported”
  + Click “Warped Tissue” underneath it
    - In the new window that pops up, select “Modulated + Unmodulated”
  + Repeat these “Native Tissue” and “Warped Tissue” steps with all 6 tissue probability maps.
* Scroll down and click “Deformation Fields”
  + In the new window that pops up, select “Inverse + Forward”
* Click the green arrow (“run”) in the top bar
* Wait for the program to run in the background
* When done running, command window will say “done segment”

***Things to do:***

*If you want to look at your data, click on “Check Reg”. In the window you can select c1ANAT.nii to c6ANAT.nii. These are segmentations in subject space.*

*Now look at these images and note GM, WM and CSF (C1 to C3) images.*

**STEP 7. NORMALIZATION**

* Click the dropdown arrow in the “Normalize” menu box
* Select “Normalize: (Write)” 🡪 new window will open
* In the new window, double click “Data”
* Double click “Deformation Field”
  + - In the new window, select “y\_ANAT.nii” and then click “Done”
* Below “Deformation Field”, double click “Images to Write”
* In the frame field that has “1”, type “:120” after it (1:120) and click enter
* In the new box, select all “raREST” files by clicking the first raRest file, holding “Shift” and click the last raREST file
* Click “done”
* “Images to Write” should now have “120 files” listed next to it
* Double click “Voxel Sizes” under “Writing Options”
  + Change [2 2 2] to [3 3 3]
* Click the green arrow (“run”) in the top bar
* Wait for the program to run in the background
* When done running, the main results window will update automatically, and the command window will say “done normalize”

***Things to do:***

*Click “Display”, scroll down and click “wraREST.nii” and click “done”. Now a window with images in three panes will open up. Scroll through these images and see that they line up.*

**STEPS 8,9,10. DISCARD “n” TIME POINTS, REGRESSION AND BAND PASS FILTERING**

**The next three steps are not covered in the presentation so I have attached screenshots to illustrate the various steps.**

1. **Create** **masks for white matter and CSF in SPM**

CREATE A MASK FOR WM:

* In SPM click on ImCalc

**Graphical user interface

Description automatically generated**

* Click on Input images

**Graphical user interface, text, application

Description automatically generated**

* Now select wraREST.nii (you only need the first timepoint)
* Next select mwc2ANAT.nii (this is the WM)
* Click “done”

**Graphical user interface, text, application

Description automatically generated**

* Now double click “Output Filename” and in the box write MASK\_WM

**Graphical user interface, text, application

Description automatically generated**

* Now click “Output Directory”

**Graphical user interface, text, application

Description automatically generated**

* Click on single dot in the right pane
* Then click “done”

**Graphical user interface, application, Word

Description automatically generated**

* This sets your output directory path to SUB01

**Graphical user interface, text, application, email

Description automatically generated**

* Next click on expression and in the box type “i2>0.9” and hit OK

**Graphical user interface, text, application

Description automatically generated**

* Now click the green arrow (run)
* Now if you click display you will see MASK\_WM.nii

CREATE A MASK FOR CSF:

* In SPM click on ImCalc
* Click on Input images
* Now select wraREST.nii (you only need first timepoint)
* Next select mwc3ANAT.nii (this is the CSF)
* Click “done”
* Now double click “Output Filename” and in the box write MASK\_CSF
* Now click “Output Directory”
* Click on single dot in the right pane
* Then click “Done”
* This sets your output directory path to SUB01
* Next click on expression and in the box type “i2>0.9” and hit OK
* Now click the green arrow (run)
* Now if you click display you will see MASK\_CSF.nii

**For the Regression, Removal of n timepoints and Bandpass filtering, we will be using scripts provided courtesy of Dr. Suril Gohel, Assistant Professor, Rutgers University.**

**These have been created for MAC OS. There may be changes required to adapt this for Windows.**

1. **Extract PCA for WM and CSF:**

* Open matlab
* Open a new matlab editor window and copy paste the script below into the editor window.
* Hit Save and save the script as a matlab (.m) file in your SUB01 folder
* You will need to change the DataPath in the script to be specific to where your SUB01 is stored.
* Then click the green arrow “run” button.

DataPath='/Users/bvbg/Downloads/'

subname = 'SUB01'

DataPath1=([DataPath '/' subname '/']);

cd(DataPath1);

v\_WM = spm\_vol([ DataPath1 '/MASK\_WM.nii']);

y\_WM = spm\_read\_vols(v\_WM);

v\_CSF = spm\_vol([ DataPath1 '/MASK\_CSF.nii']);

y\_CSF = spm\_read\_vols(v\_CSF);

image\_dim = v\_WM.dim;

a\_WM = reshape(y\_WM ,image\_dim(1)\*image\_dim(2)\*image\_dim(3),1);

a\_CSF = reshape(y\_CSF,image\_dim(1)\*image\_dim(2)\*image\_dim(3),1);

clear a\_BOLD\_WM a\_BOLD\_CSF

for imagei = 1:115

v\_BOLD = spm\_vol([DataPath1 '/wraREST.nii,' num2str(imagei+5) '''']);

y\_BOLD = spm\_read\_vols(v\_BOLD);

a\_BOLD = reshape(y\_BOLD,image\_dim(1)\*image\_dim(2)\*image\_dim(3),1);

a\_BOLD\_WM(:,imagei) = a\_BOLD(find(a\_WM > 0.9));

a\_BOLD\_CSF(:,imagei) = a\_BOLD(find(a\_CSF > 0.9));

end

[c,s\_WM] = pca(a\_BOLD\_WM');

[c,s\_CSF] = pca(a\_BOLD\_CSF');

pca\_WM = s\_WM (:,1:5);

pca\_CSF = s\_CSF(:,1:5);

save([ DataPath1 '/covariance\_wm\_csf.mat'], 'pca\_WM', 'pca\_CSF');

fprintf(' done! \n')

1. **Removal of n timepoints (5), Regression, and Band-pass filter**

* Copy paste the script below into a new matlab editor window.
* Hit Save and save the script as a matlab (.m) file in your SUB01 folder
* You will need to change the DataPath in the script to be specific to where your SUB01 is stored.
* Then click the green “run” button.

DataPath='/Users/bvbg/Downloads/'

subname = 'SUB01'

cd(DataPath)

tp\_remove=5

num\_tp = 120

TR = 2.5

sf=1/TR

cH = 0.1

cL = 0.01

DataPath1=([DataPath '/' subname '/']);

cd(DataPath1)

v\_pair = spm\_vol([ DataPath1 '/wraREST.nii,4']);

y\_pair = spm\_read\_vols(v\_pair);

image\_dim = v\_pair.dim;

for imagei = 1:num\_tp-tp\_remove

v\_image = spm\_vol([ DataPath1 '/wraREST.nii,' num2str(imagei+tp\_remove)]);

y\_image(:,:,:,imagei) = spm\_read\_vols(v\_image);

end

fprintf(' done! \n')

fprintf('loading covariates...')

rp\_REST=load([DataPath1 '/rp\_aREST.txt']);

rp\_REST\_temp=rp\_REST(tp\_remove+1:num\_tp,:);

rp\_REST = zscore(rp\_REST\_temp);

rp\_previous = [0 0 0 0 0 0; rp\_REST(1:end-1,:)];

rp\_auto = [rp\_REST rp\_REST.^2 rp\_previous rp\_previous.^2];

load([DataPath1 '/covariance\_wm\_csf.mat']);

fprintf(' done! \n')

fprintf('regressing out covariates, & filtering...')

mean\_image = mean(y\_image,4); % calculate mask

y\_mean\_image = mean\_image(:);

imp\_mask = mean(y\_mean\_image(find(y\_mean\_image>(mean(y\_mean\_image)/8))))\*.3;

for vi = 1:image\_dim(1)

fprintf('%d ',vi)

for vj = 1:image\_dim(2)

for vk = 1:image\_dim(3)

if y\_image(vi,vj,vk)<imp\_mask

y\_image\_regressed(vi,vj,vk,:) = zeros(1,1,1,num\_tp-tp\_remove);

else

b = zscore([pca\_WM(:,1:5) pca\_CSF(:,1:5) rp\_auto]);

[b,bint,y\_image\_regressed(vi,vj,vk,:)] = regress(shiftdim(y\_image(vi,vj,vk,:),3),[b ones(num\_tp-tp\_remove,1)]);

y\_image\_regressed\_filtered(vi,vj,vk,:) = fmri\_bandpass\_filter(y\_image\_regressed(vi,vj,vk,:),sf,cH,cL);

end

end

end

end

fprintf(' done! \n')

%varname=([subjfolder(subji).name '\_WS'])

V2.dim = v\_pair.dim;

V2.dt = v\_pair.dt;

V2.mat = v\_pair.mat;

V2.pinfo = v\_pair.pinfo;

for imagei = 1:num\_tp-tp\_remove

V2.n=[imagei 1]

V2.fname = [DataPath1 '/filt\_wraREST.nii'];

V2 = spm\_write\_vol(V2,y\_image\_regressed(:,:,:,imagei)+mean\_image);

end

function [fmat]=fmri\_bandpass\_filter(mat,sf,cH,cL)

% Band-pass filter

% [fmat]=fmri\_banpass\_filter(mat,sf,cH,cL);

% mat : data [data point, regions]

% sf : sampling frequency [Hz]

% cH : high cut-off frequency [Hz]

% cL : low cut-off frequency [Hz]

%dim=size(mat);

% change the cut-off frequency into normalized frequency

% cH, cL should be the value within [0,1].

% 1 is corresponding to the (sampling time)/2

cL2 = cL/(sf/2);

cH2 = cH/(sf/2);

% parameter setting

N=2; % order of butterworth filter 1\*n

Wn=[cL2,cH2]; % cut-off frequency

% construct a butter worth filter

[B,A] = butter(N,Wn);

% figure; freqz(B,A); % plot frequency response

fmat=filtfilt(B,A,mat); % filter for both direction

end

**11. Smoothing**

* In the SPM window click “Smooth”
* In the new window that pops up, double click “Images to smooth”
* In the frame field that has “1”, type “:120” after it (1:120) and click enter
* In the new box, select all “filt\_wraREST” files by clicking the first filt\_wraRest file, holding “Shift” and click the last filt\_wraREST file
* Click “done”
* “Images to smooth” should now have “115 files” listed next to it
* Double click “FWHM”
  + Change [8 8 8] to [6 6 6]
* Click the green arrow (“run”) in the top bar
* Wait for the program to run in the background
* Smoothing is complete

***Things to do:***

1. *To see the images, you can click on “Check Reg” and select the filt\_wraREST and the sfilt\_wraREST. Click done and the image will open up showing the unsmoothed and smoothed images. Compare them and see the blur that has been introduced in the smoothed images.*
2. *Now repeat the Smoothing step by using different FWHM ie [8 8 8] and*

*[4 4 4].  Remember to give each a distinct Filename Prefix. Then click “Check Reg”. Select these files and see how different they look from each other.*