Comparison of Normalisation Techniques for fMRI Preprocessing: SPM, DARTEL, and CAT12

Paul Fisher

11628653

Department of Psychology, Ludwig-Maximilians-Universität in Munich (LMU)

Author Notes:

Paul Fisher, Department of Neuro-Cognitive Psychology, in conjunction with the Graduate School of Systemic Neurosciences (GSN).

Correspondence concerning this article should be addressed to Paul Fisher

Ludwig-Maximilians-Universität Muenchen

Leopoldstasse 13

D-80802 Muenchen

Email: paulozfisher@gmail.com

Phone: +44 7835575116

Abstract

Neurological research, particularly fMRI research has struggled to develop a concrete preprocessing pipeline in regards to normalisation techniques and their application. This raises questions over the efficiency of techniques that are used by different departments and organisations. The focus of this paper relates to SPM, DARTEL, and CAT12 normalisation techniques and their usage within an fMRI preprocessing pipeline.

The study incorporated three identical sets of data (20 participants) acquired from a navigation study conducted in 2014. The objective was to perform stages of preprocessing on each version of the data and compare the results in terms of: efficiency, optimisation, usability, and quality of fMRI normalisation. Results demonstrated that CAT12 produced the highest quality of normalisation out of the three techniques. However, this was limited by methodological constraints, such as the usage of cropping. One major success of the paper was the usage of a tool for setting the origin point automatically across the data. The key take-away message was that the most effective normalisation technique CAT12, was also the least intuitive and most error prone technique, with SPM being significantly easier, but with lower quality, and DARTEL being the moderate approach.

Keywords: *{normalisation; fMRI; CAT12; SPM; DARTEL}*

Comparison of Normalisation Techniques for fMRI Preprocessing: SPM, DARTEL, and CAT12

Functional Magnetic Resonance Imaging (fMRI) currently sits at the forefront of medical and neurological research. Across different schools of academia it has become synonymous with disciplined experimentation and a reliable tool for examining, at a high resolution, connectivity and brain activity (Soares et al., 2016). First developed in the nineties (Kwong, 2012), fMRI has grown to be one of the most commonly used methods of data collection and boasts a repertoire of strong spatial resolution, consistent reproducibility, and - fundamental to the practise - a non-invasive methodology (Soares et al., 2016). A key aspect of any respectable fMRI study is the diligence of researchers to apply reliable and relevant preprocessing steps. This is due to the low contrast-to-noise ratios (CNR) that fMRI data notoriously produces (Churchill et al., 2012) which are most notably seen in noise patterns and spatiotemporal signals. Consequently, a selection of reliable and well defined steps should be followed to prepare raw fMRI data for evaluation and to avoid these confounds. This insistence of standardised preprocessing caused a divide amongst researchers, with a plethora of results demonstrating that parameters chosen for a preprocessing pipeline can have drastically different effects on data (Poline et al., 2006;Tanabe et al., 2002; Zhang et al., 2009). It is therefore pertinent to explore the differences between opposing preprocessing pipelines and evaluate their efficacy in a comprehensive manner.

One important preprocessing step is the spatial normalisation of data (Brett, Johnsrude, & Owen, 2002). At this stage, the objective is to transform the data to a standardised anatomical reference space template. Thus matching each subject to an established stereotaxic space and allowing for comparisons between them. This is accomplished by pairing white and grey matter to their respective reference space in the

given template. Most commonly, the template is either the Monteal Neurological Institute (MNI) template, or the Talairach atlas (Talairach & Tournoux, 1988). This enables the researcher to look for general trends and differences between subjects and to average signals across individuals, it therefore increases the statistical power of the data (Saraiva & Hoffland, 2009). Conversely, poorly conducted or error prone normalisation has been demonstrated to increase false positives due to artifacts and inaccurate alignments (Soares et al., 2016). One major application of spatially normalised data is the ability to perform functions such as voxel-based morphometry (VBM) which executes area based calculations of volumes amongst populations (Ashburner & Friston, 2005).

The Talairach space computes the principal axis to the anterior commissure - posterior commissure (AC-PC) with the coordinates (x = 136mm, y = 172mm, z = 118mm) and the origin typically resides at the AC (Laird et al., 2010). The MNI template coordinates are registered differently depending on the version attributed to the data set. One example of the template is the BIC MNI-305 template (x = 142mm, y = 181mm, z = 132mm) which is a high resolution image generated during the ICBM project (Mazziotta et al., 2001b; as cited in Lancaster at al., 2007). Both methods are present in literature, however, more recently the trend has moved towards the MNI template due to its presence across fMRI analysis software such as Statistical Parametric Mapping (SPM 12, Wellcome Department of Imaging Neuroscience, http://www.fil.ion.ucl.ac.uk/spm/ ; Penny et al., 2011).

The current study evaluated three commonly used normalisation techniques: the built-in SPM toolbox normalisation, Diffeomorphic Anatomical Registration using Exponentiated Lie algebra (DARTEL), and the Computational Anatomy Toolbox (CAT12). The SPM normalisation calculates data-sets using EPI templates in MNI space, this technique has come under scrutiny due to the poorer anatomical resolution of functional images

creating higher contrast features on the outer regions of the brain (Poldrak & Mumford & Nichols, 2011). One method that has garnered interest involves the process of computational anatomy. In this approach, a transformation style known as diffeomorphism is applied to the images, this uses a complex mathematical formula to help account for anatomical constraints in images (see Holden, 2008).

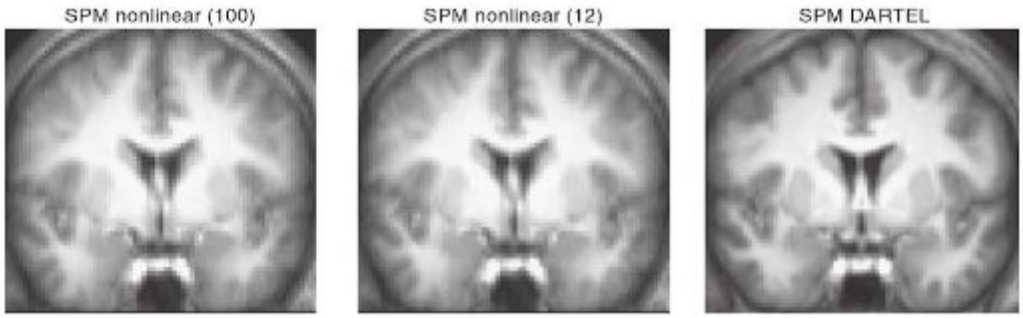

Figure 1: Adapted from Poldrack (2011), demonstrating examples of mean anatomical images for a group of 8 participants . The number in brackets refers to nonlinear frequency cutoff (higher number = less warping).

The diffeomorphic method is employed by DARTEL and as shown in Figure 1, the DARTEL approach appears to have a clearer and better aligned normalisation than SPM. The clarity is easier to identify around the sulci and ventricles. Poldrack et al.,(2011) explain that this is likely due to a poor registration of finer structures in the SPM approach. DARTEL incorporates a large deformation framework in order to retain clarity. Furthermore, the template is dynamic and iterative, meaning that an updated template is generated upon the completion of one iteration (Yassa & Stark, 2009). It is speculated that one of the most useful applications of DARTEL normalisation relates to data that struggles to cross-reference the deformation with regional information. DARTEL has also been recommended as a high ranking application in a meta-study of different nonlinear deformation algorithms (Klein et

al., 2009). However they did not examine the techniques observed for the current paper (SPM and CAT12).

CAT12 is primarily a toolbox that maintains its own segmentation process in order to perform normalisation. The literature pertaining to CAT12 efficacy is quite scarce, however, it has been used as a technique in recent fMRI studies (Abulafia et al., 2017; Franke et al., 2016; Igata et al., 2017) but with little to no evaluation. Due to its recent popularity, one can assume that researchers have had success in using the technique for preprocessing pipelines, it is therefore reasonable to assume that it offers improvements over SPM and DARTEL.

The current study is aiming to examine the three techniques from three perspectives: the ease of use, the clarity and optimisation of the normalisation, and the data quality of the functional images. These will be evaluated in terms of the researcher's experiences and using comparative analysis techniques. Ultimately the research question is "What normalisation methods within SPM produce the best overall results in terms of consistency, workload/logistics, and timeframe?". Based on contemporary literature in the field of fMRI preprocessing, we would hypothesise that DARTEL would out-perform the default SPM and that CAT12 would out-perform DARTEL. However, it is quite probable that due to its recency, that the CAT12 technique may encounter unexpected hurdles or difficulties.

Methods

Materials

This experiment was conducted using a personal computer, running on a 4.4.0-72-generic LINUX system at the Forchungshaus neurology department in conjunction with LMU and the Graduate School of Systemic Neurosciences (GSN).

Image Acquisition

Raw TAR file data collected by a researcher were used for the purposes of this experiment. This included twenty healthy participants (9 male, age $M = 25.4$, $SD = 2.72$) who had undergone an fMRI study examining time and distance estimation in a virtual environment (for a more detailed description of participant information, procedure or any further details, see; Ries, 2014). The TAR files were accompanied by scanner protocol sheets detailing: name, age, education, handedness, sex, weight, etc. This information was used to generate a patient-ID for both privacy and anonymity assurance. Each participant scan series included: four function runs (fMRI Time A&B and fMRI Distance A&B), one anatomical run (FSPGR), a localiser and an asset run. For a total of seven sequences on average (in some circumstances runs had to be re-ordered or were incomplete due to errors). The number of image volumes collected ranged between 200-500 for each function run. The resulting TAR-files (including the .nii files) were used as the testing data for this experiment.

Three identical versions of the data were created and separated according to their testing methodology. These datasets were hosted on the Forchunshaus server network and were only accessible by authorised experimenters. One participant (MR10) had incomplete data and was not included in the preprocessing analysis.

Toolboxes

For analysis, different software had to be implemented into MATLAB (version R2015B). Firstly, SPM was used for the application of all preprocessing stages of this experiment. Secondly, in order to evaluate alternative methods of normalisation, two toolboxes were installed onto MATLAB which have different methods of preprocessing, namely, the DARTEL toolbox (Ashburner, 2007) and the CAT12 toolbox (Computational Anatomy Toolbox, 2016). Furthermore, an fMRI imaging quality assurance toolbox was installed (http://www.mccauslandcenter.sc.edu/crnl/tools/qa) which included tools that were

helpful for the analysis of the transformed data, such as the nii_setOrigin12x.m, nii_mean_std.m and nii_qa_moco.m scripts.

fMRI

A 3T Signa HDx (GE,USA) system was used, incorporating an eight-channel head volume coil. Functional data were acquired using T2-weighted MRI EPI volumes and BOLD contrast were used (TR = 2500ms, TE = 32ms, flip angle = 80°, matrix 64 x 64, FOV = 22cm, axial slice (total 38) thickness = 3.5mm) and a further high-resolution, T1-weighted anatomical scan, created using a fast-spoiled gradient-echo sequence, FSPGR (TR = 8.008ms, TE = 3.15ms, flip angle = 15°, matrix 256 x 256 x 248, and 0.8 x 0.8 x 0.7 mm voxel size) (adapted from Ries, 2014).

Procedure

This experiment focused around conducting three different normalisation techniques on the same set of data in order to examine the strengths and weaknesses of each method. Before beginning this task, it was pertinent to evaluate the images by performing artifact detection. As can be seen from Figure 2, the images were very clean when examining at a threshold of 2 standard deviations. This confirmed that there had been no concerning movement artifacts or

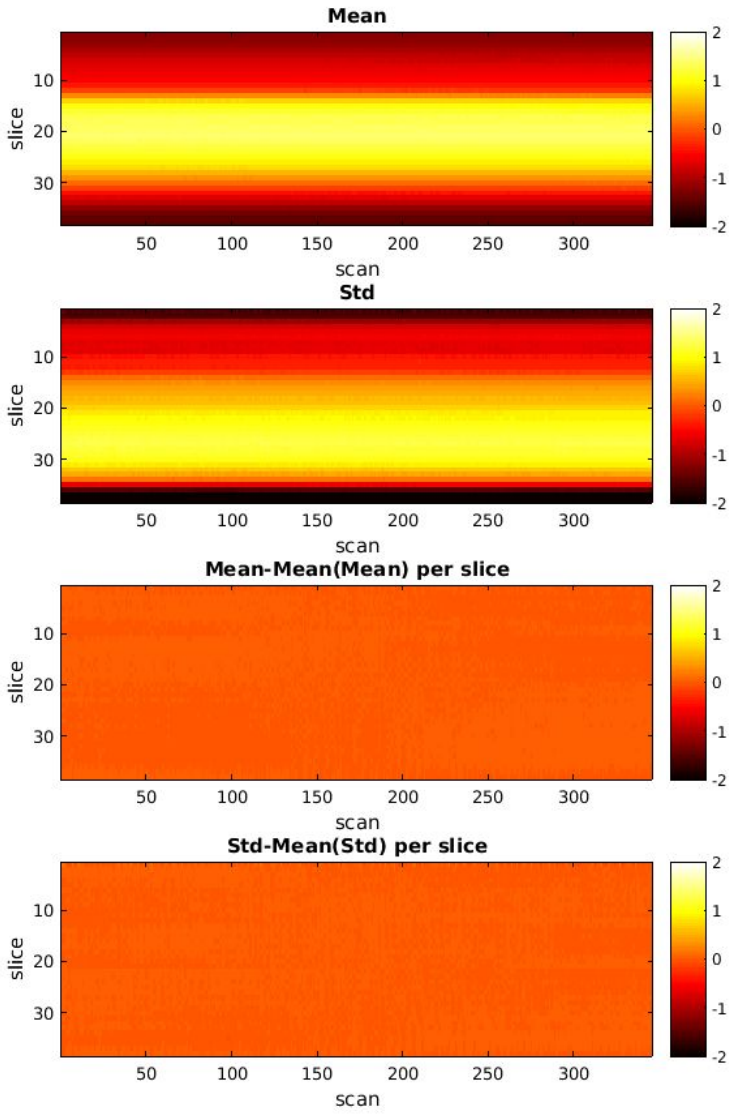

Figure 2: Example of artifact detection for raw data before preprocessing. Calculated including mean, standard deviation, mean-mean(mean) per slice, and std-mean(Std) per slice. This scan indicated that there were few artifacts or errors across the data.

problems with the data-set. This was done as a formality, as this data were procured and used for another study. The next step was to perform the origin set for each participant data-set. Due to the large number of trials per participant and based on the logistics of manual origin setting. It was decided to use a MATLAB script for setting the origin in order to speed up the process. The script (a part of the quality assurance toolbox, named "nii_setOrigin12x") was able to set the origin to the Anterior Commissure to a high precision and was capable of running through an entire data-set without interruption. For the purposes of this experiment, the cropping function (included in the script) was integrated into the procedure (something that, in hindsight, was regrettable - see Discussion). Following this, the three different preprocessing pipelines were created.

SPM Preprocessing

In order to run the necessary scripts on the data-set, a preprocessing skeleton had to be created in SPM batch editor. The SPM skeleton consisted of standardised preprocessing modules. This included, Slice Timing (using settings defined by the functional run data), e.g. Number of slices = 38, TR = 2.5, TA = 22.8, the slice order was bottom-up interleaved at 1x38 double, with the reference slice set to 1. Realign: Estimate & Reslice generated mean images only with the Slice Timing images used as dependency. Coregister: Estimate used the anatomical scan and the mean image dependencies with otherwise default settings. Segmentation used the default anatomical image and required the input of SPM12 template images for transforming to MNI standard template space. These templates corresponded to grey matter, white matter, and cerebrospinal fluid (CSF) as well as templated probability maps in MNI space. Following this the normalisation was performed for the first pipeline. Normalise: Write took the segmented images as dependencies for the deformation field and selected the slice time corrected images to write. The images where then smoothed based on

the normalisation that was performed using the warped images. Finally a further Normalise: Write module was conducted taking bias corrected images as dependencies in order to generate the final normalised functional and anatomical data.

DARTEL Preprocessing

For the DARTEL preprocessing, two batch skeletons had to be created due to functionality of DARTEL and MATLAB and how they deal with 4D images. Therefore, the Skeleton A consisted of: Slice Timing (following identical guidelines to SPM preprocessing), Realign & Unwarp using dependencies from the Slice Timing and otherwise default settings. Coregister: Estimate using the anatomical file as the reference image, the source image as the unwarped mean image dependency, and the other images set as dependencies of the unwarped function runs. For the Segmentation, DARTEL uses the SPM12 template images. The same corresponding image types were entered into the batch editor, however, the tissue class images that are produced need to be specified to Native + DARTEL imported in order for them to be a format that the DARTEL toolbox can recognise and utilise later. The final stage of skeleton A is Run Dartel (existing Templates), this runs the DARTEL nonlinear image registration procedure which matches pre-existing templates to the individual images. This requires the selection of templates to different tissue probability maps, corresponding to the correct tissue type (white matter, etc.). After the creation of Skeleton A, the relevant scripting is performed in MATLAB and the batch is executed. This generates the images suitable for the normalisation in Skeleton B. For Skeleton B, the first module was Normalise to MNI Space. In order to have this run successfully, the Dartel Template section had to be cleared as MNI space had already been specified in Skeleton A. Next the functional EPI images were added to the flow-field as a dependency and the realigned and slice time corrected functional images are selected that are needed to normalise. The anatomical image

needs to be normalised to DARTEL template space so a final Normalise: Write is performed. This section of the skeleton is then scripted and executed in order to generate the DARTEL normalised functional and anatomical data.

CAT12 Preprocessing

In a similar fashion to DARTEL, CAT12 required two sets of preprocessing batch skeletons to be created (referred to here as CATSkeleton A and B). For CATSkeleton A, there are similarities to both the SPM and DARTEL pipeline. Slice Timing involved identical settings. Realign & Unwarp followed the same structure as DARTEL, as did Coregister: Estimate. Then CAT12: Segmentation was performed on the data. This involved setting the tissue probability map to the SPM template, and allowing for Affine transformation of grey matter and white matter images. This was then scripted and executed through MATLAB in order to generate the images for CATSkeleton B. CATSkeleton B included a Run Dartel (existing templates) and required adding the CAT12 toolbox versions of the templates for each image type. This was then followed by Normalise to MNI space. This was then scripted and executed to create the CAT12 normalised functional and anatomical data.

Both the DARTEL and CAT12 methods have automatic smoothing functions built into them. So at this stage the images are complete are ready for analysis. The images will be compared both individually and across methods using both SPM functions and quality assurance toolbox scripts. At this stage it is important to evaluate the success of the preprocessing steps using the SPM tool Check Registration, this way you can examine both the functional and anatomical runs in comparison to a predefined template, search for inconsistencies as well as identify any artifacts.

Checking Image Registries

Before analysing the data, it was important to manually examine the accuracy of the techniques by overlaying functional and anatomical runs from each trial onto templates used for normalisation. This is demonstrated below, in Figure 3, showing anatomical scans of all three techniques. As can be seen, skull-stripping was performed for the CAT12 data-set making any true comparison of the FSPGR images difficult. However, the researcher checked every run and found no major artifacts or errors across the trials. Some known effect phenomenon were demonstrated, this will be considered in the discussion.

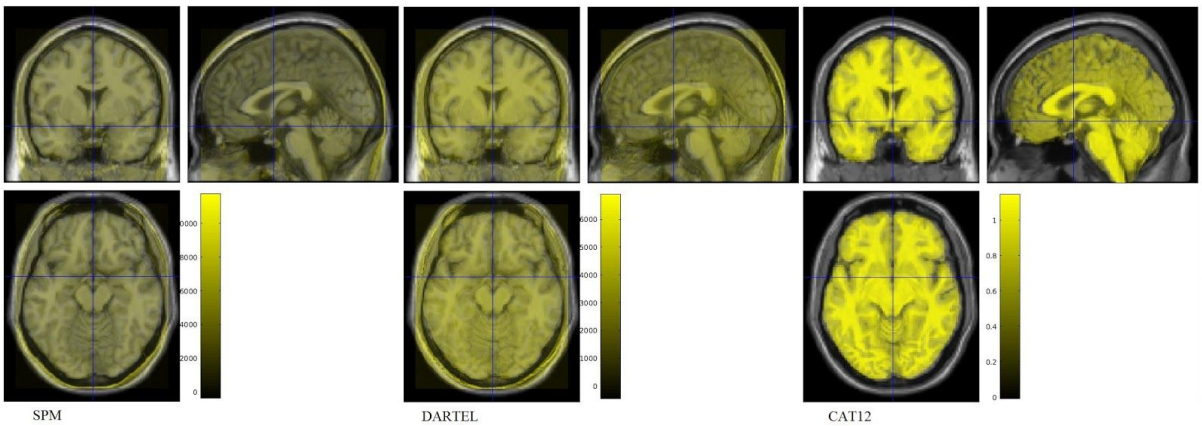

Figure 3: The normalised FSPGR (anatomical) scans taken from a random participant for SPM (left), DARTEL (centre), and CAT12 (right). One key difference to note is the skull stripped style of the CAT12 anatomical scans.

Results

For analysis, the current paper will be divided into three sections in order to evaluate the data from an empirical perspective. Firstly, the researchers notes will be examined from a descriptive perspective. Following this, the quality of the normalisation will be assessed from the data collected, and finally the data quality of the subsequent fMRI images will be examined as objectively as possible.

Descriptives

It was important to demonstrate the effectiveness of each normalisation technique using a scale that was at least moderately reliable. Therefore, each method was ranked according to metrics measuring three qualities: efficiency (number of redos due to software problems), optimisation (length of time to generate and execute MATLAB scripts), and useability (number of errors or issues encountered). Note that this section does not evaluate the actual tools, rather their ease of use. Each metric is rated out of 5.

SPM efficiency (5) was a perfect score with no issues causing a redo of specific parts or participants. SPM optimisation (4) was of high quality. The scripting and execution of the script took ~12 hours in total and completed without errors. SPM useability (5) was also a perfect score, the script worked effectively and no problems were encountered from beginning to completion. This makes the SPM total score 14.

DARTEL efficiency (4) was also good. One section had to be recomputed due to an error with the way MATLAB registers 4D images and therefore caused the files to need replacing. DARTEL optimisation (2) was fairly poor. In total, after re-configuring the files, running the setOrigin script for a second time, and writing out the second preprocessing skeleton and corresponding script, the total time of work and execution was close to ~24 hours. DARTEL useability (3) was adequate. The impression of the researcher was that DARTEL becomes more intuitive after a longer stint of familiarity, a total of ~5 major issues occurred during the preprocessing pipeline. This brings the total score for DARTEL to 9.

CAT12 efficiency (4) had a similar problem to DARTEL relating to the generation of a secondary script and subsequent redoing of setOrigin scripting. CAT12 optimisation (2) was also poor due to these issues, the total runtime including scripting and execution was ~30 hours. CAT12 useability (1) was problematic. The predominant issue related to the software package requiring updating due to it being new and constant patches being applied. There

was a real problem with the continuation of the script after Segmentation that caused several major issues (~8 in total). CAT12 scored a total of 7.

Normalisation

The anatomical scans of the three techniques were examined for their standard deviation and outliers were reported, consistently across all methods participant 15 was reported with the highest values (SPM, Z = 1.18; DARTEL, Z = 1.2; CAT12, Z = 0.89). Participant 14 was, conversely, reported as the lowest standard deviation across all participants (SPM, Z = 0.64, DARTEL, Z = 0.64, CAT12, Z = 0.69). Across the datasets, reliability analysis determined that the figures were acceptable ($\alpha = 0.87$) which would have increased to $\alpha = 0.91$ with outliers removed. SPM anatomical scans were examined in relation to the mean intensity values ($M = 0.77$, $SD = 0.13$, *variance* = .017), as were DARTEL scans ($M = 0.76$, $SD = 0.12$, *variance* = .017) and CAT12 ($M = 0.75$, $SD = 0.04$, *variance* = .002). CAT12 was demonstrated to have the most consistent dataset in terms of relation to the mean intensity.

Another way to examine the success of normalisation is by examining the difference between the template grey-matter images and the normalised grey-matter images created. However, due to the way that SPM generates the segmented versions of the images prior to any normalisation, these do not properly align. Therefore, only DARTEL and CAT12 were examined.

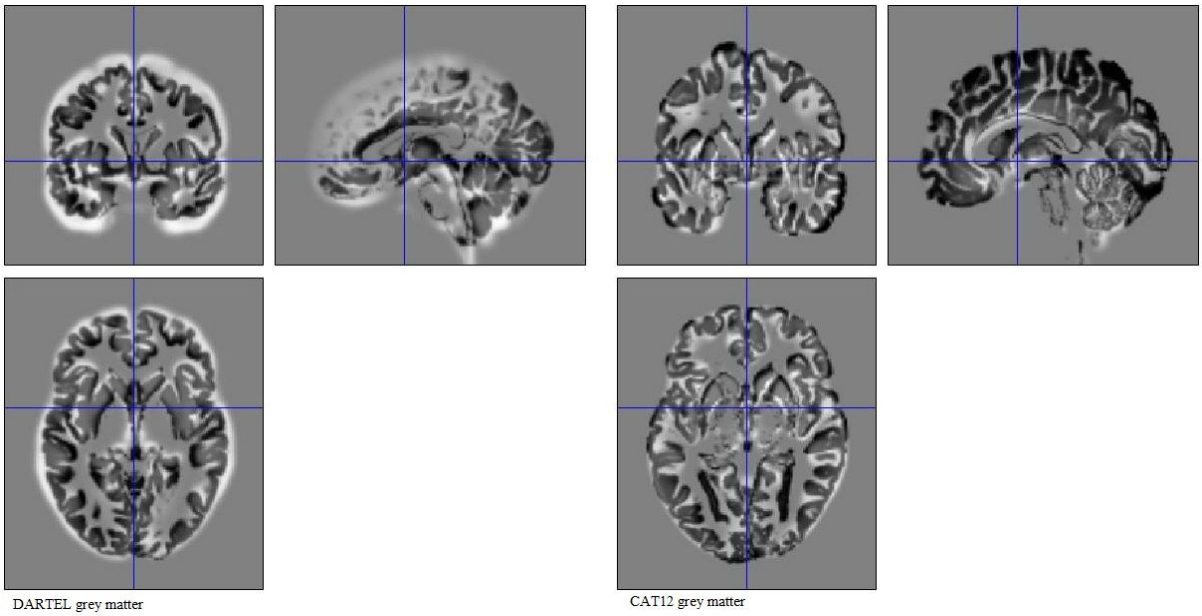

Figure 4: DARTEL (left) and CAT12 (right) mean grey matter anatomical scans that have been subtracted from the toolbox template versions in preprocessing. As viewed from the AC origin, it is clear that CAT12 has retained a more consistent overall structural value and it can be seen that DARTEL has brighter intensity on the extremities.

As can be seen from Figure 4, there are clear differences between the consistency of the DARTEL and CAT12 anatomical normalisation, this, in conjunction with the more consistent Z scores firmly places CAT12 as the frontrunner for the anatomical scans.

fMRI Images

For the functional runs, a different quality assurance script was applied to the data (nii_qa_moco.m). This script enabled the researcher to examine the Z-scores and standard deviations across functional runs. For the preprocessing stages, there was no use of the unwarp (during realignment) function which may have caused differences across trials or techniques. Calculations were made for SPM fMRI images ($M = 0.8$, $SD = 0.16$, *variance* = .027), DARTEL fMRI images ($M = 0.78$, $SD = 0.15$, *variance* = .023), and CAT12 images ($M = 0.78$, $SD = 0.13$, *variance* = .018) which once again demonstrated that CAT12 data was the more consistent technique.

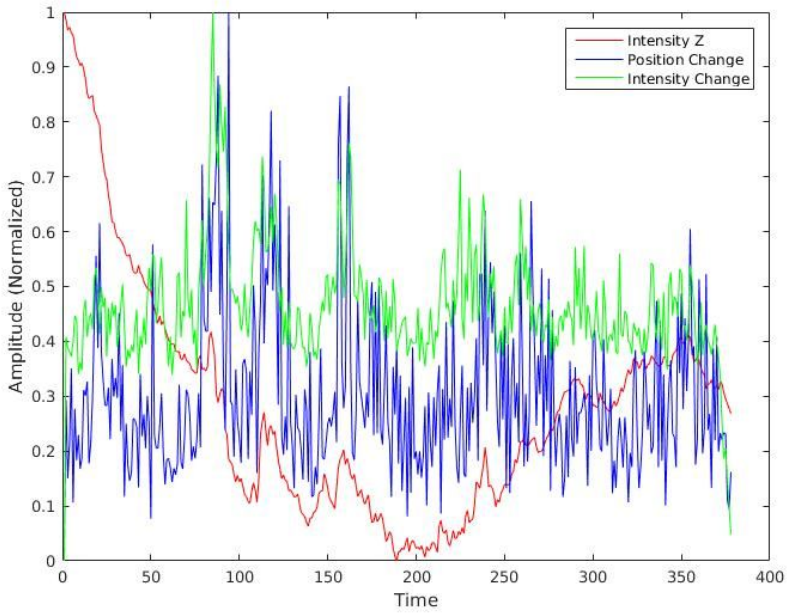

Figure 5: The amplitude of intensity z-scores (red), position change (blue) and intensity change (green) of the SPM normalised function run across time.

Figure 5 demonstrates the differences in intensity and position across the different images in an average function run for the SPM method. One important to note phenomenon is the dropping intensity Z score across the runs. An ANOVA was calculated for the three functional run scores to determine whether they differed in any distinguishable way beyond simple consistency. There was no significant effect found of the mean intensity values across the three functional runs [$F(2.18) = 0.08$, $p = 0.91$]. Which is to be expected when examining these intensity values.

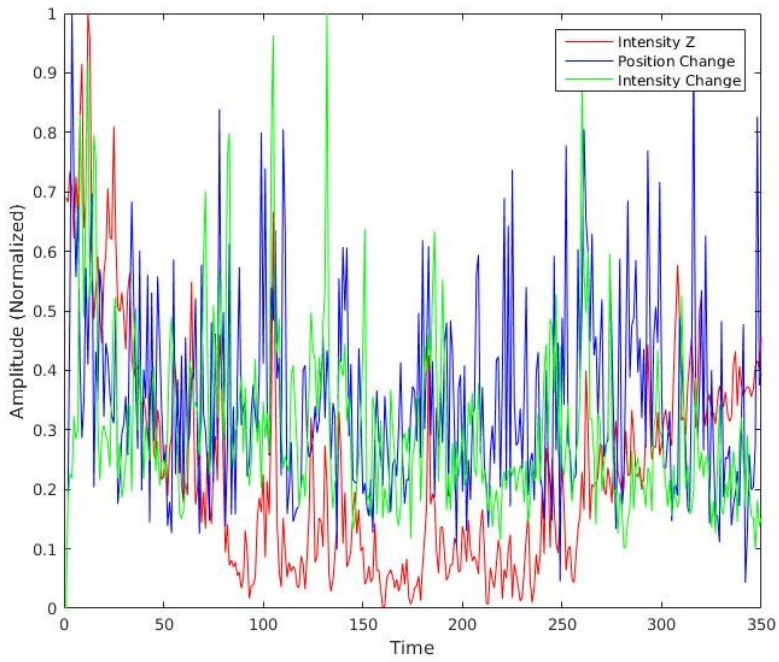

Figure 6: The amplitude of intensity z-scores (red), position change (blue) and intensity change (green) of the DARTEL normalised function run across time.

As can be seen from both Figure 6 and 7, there are numerous differences between the fMRI amplitudes across the three methods. The DARTEL has a much stronger overlap of position and intensity value changes whereas the CAT12 is characterised with a much sharper reduction in amplitude but with numerous strong spikes across different time values. This is likely due to the different style of segmentation that is performed by CAT12 using its own toolbox templates to normalise the images. The distribution across CAT12 is clearly more uniform and organised, which helps corroborate with the lower standard deviation and variance scores that it demonstrated across both anatomical and functional scans.

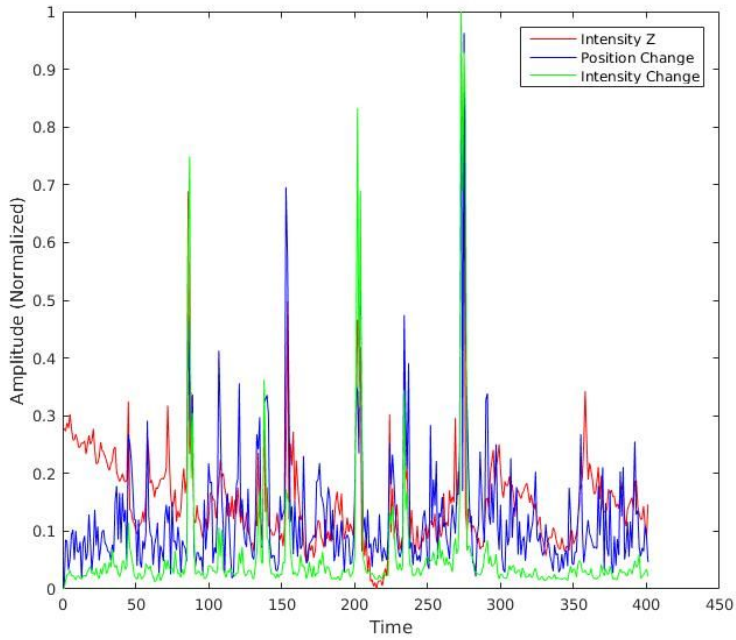

Figure 7: The amplitude of intensity z-scores (red), position change (blue) and intensity change (green) of the CAT12 normalised function run across time.

Discussion

This experiment set out to identify the best method for normalisation preprocessing by examining a dataset and comparing the results of each method. From the results, it is clear that in terms of efficiency, optimisation, and usability, SPM was by far the prefered method for the researcher. Fundamentally, the SPM technique was less of a logistical problem and far more streamlined. This result is unsurprising, as the SPM toolset is already built into the fMRI analysis software that is commonly used and has a user-friendly interface. Furthermore, the SPM technique required less preliminary setup as the majority of functions and options are optimised by default for fMRI analysis. This paper hypothesised that CAT12 would out-perform both DARTEL and SPM in overall normalisation quality and it would appear that this was an accurate presumption. Moreover, the caveat suggesting that CAT12 may have been more susceptible to errors and software issues held true, as the researcher ranked

CAT12 poorly in terms of the efficiency, optimisation, and usability scores. DARTEL appears to fit somewhat in the middle of the other two techniques, performing respectably in terms of analysis and researchers scores. It may be fair to say that DARTEL strikes a better balance between usability and strength of normalisation (FSPGR and fMRI).

The assertion by Poldrack et al.,(2011) that one of the main issues plaguing SPM is the poorer anatomical resolution of finer structures holding back the clarity of normalised images appears to hold true. It was further explained that one of the benefits of DARTEL normalisation is the much larger deformation framework applied. CAT12, also following a computational anatomy approach, appears to have refined this methodology even further, producing less variance and a much smoother intensity across iterations. However, one issue that was found to perhaps confound results for CAT12 was an anatomical phenomenon known as fountaining. This has been described by researchers as a signal mixture of CSF and venules (small veins) impacting the upper frontal cortex. Often one positions the EPI slice package to a degree that causes an image corruption due to cartilage and bone which creates a "fountain of signal" from the top of the skull. This effect was found during the image registration checking phase when applying global colour blobs to anatomical templates and was not demonstrated in SPM (although, a less pronounced variation was seen in DARTEL).

One major triumph of the current experiment was the success of the origin script (nii_setOrigin12x.m). Logistically speaking, the presence of this apparatus lead to a far smoother sequence of events. Setting the origin, usually a manual practise, is seen by many as a tedious and unrewarding task. With the demonstrably positive influence of the set origin script to this experiment, it is clear that it should be examined more closely in the future. Overall, it receives a glowing recommendation from the researcher.

Limitations

The most obvious limitation of this research relates to the poorly optimised usage of the crop function of the origin script. Due to the fact that cropping was used for some of the data, but not all of the data, and due to time-constraints this caused a lot of the datasets to be incomparable. This meant that analysis was lacking in some areas as a true reflection of the normalisation could not be achieved with differently sized image dimensions across techniques. Another limitation is that both DARTEL and CAT12 required significantly more execution time due to the incompatibility of the scripting. This caused a further delay in generating the correct preprocessing skeletons, this fault is predominantly the researchers for being unaware of the CAT12 executable running in the background without user knowledge.

With more time, it would have been preferable to look at signal to noise ratio of temporal signals, as a higher result would have been more likely to indicate significant results between normalisation techniques, this is a key area of interest for future research.

Future Directions

As previously stated, time constraints were one of the main burdens of this research. Therefore, one future consideration is to reexamine the data but be more consistent with the usage of cropping and brain stripping tools. This way the data can be analysed from a more empirical standpoint, with matching dimensions between SPM, DARTEL, and CAT12. Furthermore, it became increasingly apparent throughout the experiment that more attention should be paid to scripting and functionality of the different methods in order to make the process more streamlined.

More research should be attributed to the set origin script. The main takeaway message from this experiment is certainly the success of this technique for minimising time issues and retaining a level of consistency (avoiding human error). However, particular attention should be focused on whether or not to crop the images using this technique, failure

to plan this in advance, as evidenced by this experiment, can cause real limitations in data analysis.

Normalisation Technique Recommendation

The experiment has demonstrated that each technique has certain merits and negatives. Firstly, if when performing an experiment, a research would like a quick, easy to use technique with little to no major expertise, or for a beginner of fMRI experimentation, the SPM is the clear winner. If a researcher is very experienced and savvy with the problems they may encounter and is not highly restricted by time, the CAT12 provides the best empirical normalisation. If a researcher finds themselves somewhere in-between these two descriptions, the DARTEL is the more moderate approach to normalisation from the three tested techniques. The researcher would personally recommend DARTEL as the safer middle-ground, it is more closely explored from an academic point of view (Klein et al., 2009; Poldrack et al., 2011) and more commonly understood.

To conclude, this paper set out to examine the efficiency of three normalisation techniques, the SPM, DARTEL, and CAT12. It became clear during research that each method has its own strengths and weaknesses and the most successful normalisation technique (CAT12) was, incidentally, the least intuitive software/toolbox and least user-friendly. The most positive outcome of the paper was by some margin, the success of the origin setting tool, nii_setOrigin12x.m. Which served as a powerful apparatus for ensuring a reliable and error free AC origin configuration for anatomical and fMRI datasets. Future directions should focus on optimising the origin script and exploring a manner of implementing this feature into fMRI analysis software.

References

Abulafia, C., Duarte-Abritta, B., Villarreal, M. F., Ladrón-de-Guevara, M. S., García, C., Sequeyra, G., Sevlever, G., Fiorentini, L., Bar, K. J., Gustafson, D. R., Vigo, D. E., & Guinjoan, S. M. (2017). Relationship between Cognitive and Sleep–wake Variables in Asymptomatic Offspring of Patients with Late-onset Alzheimer's Disease. *Frontiers in Aging Neuroscience*, *9*(93), 1-8.

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

Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *Neuroimage*, *26*(3), 839-851.

Brett, M., Johnsrude, I. S., & Owen, A. M. (2002). The problem of functional localization in the human brain. *Nature reviews neuroscience*, *3*(3), 243-249.

Churchill, N. W., Oder, A., Abdi, H., Tam, F., Lee, W., Thomas, C., ... & Strother, S. C. (2012). Optimizing preprocessing and analysis pipelines for single‐subject fMRI. I. Standard temporal motion and physiological noise correction methods. *Human brain mapping*, *33*(3), 609-627.

Computational Anatomy Toolbox. (2016). Retrieved May 04, 2017, from http://www.neuro.uni-jena.de/cat

Franke, K., Dahnke, R., Clarke, G., Kuo, A., Li, C., Nathanielsz, P., Schwab, M., & Gaser, C. (2016). MRI based biomarker for brain aging in rodents and non-human primates. In *Pattern Recognition in Neuroimaging (PRNI), 2016 International Workshop on* (pp. 1-4). IEEE.

Holden, M. (2008). A review of geometric transformations for nonrigid body registration. *IEEE transactions on medical imaging*, *27*(1), 111-128.

Igata, N., Kakeda, S., Watanabe, K., Nozaki, A., Rettmann, D., Narimatsu, H., ... & Korogi, Y. (2017). Utility of real-time prospective motion correction (PROMO) for segmentation of cerebral cortex on 3D T1-weighted imaging: Voxel-based morphometry analysis for uncooperative patients. *European Radiology*, 1-9.

Klein, A., Andersson, J., Ardekani, B. A., Ashburner, J., Avants, B., Chiang, M. C., ... & Song, J. H. (2009). Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage*, *46*(3), 786-802.

Kwong, K. K. (2012). Record of a single fMRI experiment in May of 1991. *Neuroimage, 62*(2), 610-612.

Laird, A. R., Robinson, J. L., McMillan, K. M., Tordesillas-Gutiérrez, D., Moran, S. T., Gonzales, S. M., ... & Lancaster, J. L. (2010). Comparison of the disparity between Talairach and MNI coordinates in functional neuroimaging data: validation of the Lancaster transform. *Neuroimage*, *51*(2), 677-683.

Lancaster, J. L., Tordesillas‐Gutiérrez, D., Martinez, M., Salinas, F., Evans, A., Zilles, K., ... & Fox, P. T. (2007). Bias between MNI and Talairach coordinates analyzed using the ICBM‐152 brain template. *Human brain mapping*, *28*(11), 1194-1205.

Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., & Nichols, T. E. (Eds.). (2011). *Statistical parametric mapping: the analysis of functional brain images*. Academic press.

Poldrack, R. A., Mumford, J. A., & Nichols, T. E. (2011). *Handbook of functional MRI data analysis*. Cambridge University Press.

Poline, J. B., Strother, S. C., Dehaene‐Lambertz, G., Egan, G. F., & Lancaster, J. L. (2006). Motivation and synthesis of the FIAC experiment: reproducibility of fMRI results across expert analyses. *Human brain mapping*, *27*(5), 351-359.

Ries, A. K. (2014). *Distinct and common brain networks recruited during human magnitude estimation.* (Master's thesis). Retrieved from: https://edoc.ub.uni-muenchen.de/

Saraiva, A. & Hoffland, B. (2009). Coregistration and spatial normalisation [PowerPoint slides]. Retrieved from http://www.fil.ion.ucl.ac.uk/spm/course/slides08-zurich/

Soares, J. M., Magalhães, R., Moreira, P. S., Sousa, A., Ganz, E., Sampaio, A., Alves, V., Marques, P., & Sousa, N. (2016). A Hitchhiker's Guide to Functional Magnetic Resonance Imaging. *Frontiers in Neuroscience*, *10*.

Talairach, J., & Tournoux, P. (1988). Co-planar stereotaxic atlas of the human brain. 3-Dimensional proportional system: an approach to cerebral imaging. *Stuttgart, Germany: Theime.*

Tanabe, J., Miller, D., Tregellas, J., Freedman, R., & Meyer, F. G. (2002). Comparison of detrending methods for optimal fMRI preprocessing. *NeuroImage*, *15*(4), 902-907.

Yassa, M. A., & Stark, C. E. (2009). A quantitative evaluation of cross-participant registration techniques for MRI studies of the medial temporal lobe. *Neuroimage*, *44*(2), 319-327.

Zhang, J., Anderson, J. R., Liang, L., Pulapura, S. K., Gatewood, L., Rottenberg, D. A., & Strother, S. C. (2009). Evaluation and optimization of fMRI single-subject processing pipelines with NPAIRS and second-level CVA. *Magnetic resonance imaging*, *27*(2), 264-278.