



Evaluation of Six Phase Encoding Based Susceptibility Distortion Correction Methods for Diffusion MRI

Xuan Gu^{1,2*} and Anders Eklund^{1,2,3}

¹ Division of Medical Informatics, Department of Biomedical Engineering, Linköping University, Linköping, Sweden, ² Center for Medical Image Science and Visualization, Linköping University, Linköping, Sweden, ³ Division of Statistics and Machine Learning, Department of Computer and Information Science, Linköping University, Linköping, Sweden

Purpose: Susceptibility distortions impact diffusion MRI data analysis and is typically corrected during preprocessing. Correction strategies involve three classes of methods: registration to a structural image, the use of a fieldmap, or the use of images acquired with opposing phase encoding directions. It has been demonstrated that phase encoding based methods outperform the other two classes, but unfortunately, the choice of which phase encoding based method to use is still an open question due to the absence of any systematic comparisons.

Methods: In this paper we quantitatively evaluated six popular phase encoding based methods for correcting susceptibility distortions in diffusion MRI data. We employed a framework that allows for the simulation of realistic diffusion MRI data with susceptibility distortions. We evaluated the ability for methods to correct distortions by comparing the corrected data with the ground truth. Four diffusion tensor metrics (FA, MD, eigenvalues and eigenvectors) were calculated from the corrected data and compared with the ground truth. We also validated two popular indirect metrics using both simulated data and real data. The two indirect metrics are the difference between the corrected LR and AP data, and the FA standard deviation over the corrected LR, RL, AP, and PA data.

Results: We found that *DR-BUDDI* and *TOPUP* offered the most accurate and robust correction compared to the other four methods using both direct and indirect evaluation metrics. *EPIC* and *HySCO* performed well in correcting b_0 images but produced poor corrections for diffusion weighted volumes, and also they produced large errors for the four diffusion tensor metrics. We also demonstrate that the indirect metric (the difference between corrected LR and AP data) gives a different ordering of correction quality than the direct metric.

Conclusion: We suggest researchers to use *DR-BUDDI* or *TOPUP* for susceptibility distortion correction. The two indirect metrics (the difference between corrected LR and AP data, and the FA standard deviation) should be interpreted together as a measure of distortion correction quality. The performance ranking of the various tools inferred from direct and indirect metrics differs slightly. However, across all tools, the results of direct and indirect metrics are highly correlated indicating that the analysis of indirect metrics may provide a good proxy of the performance of a correction tool if assessment using direct metrics is not feasible.

Keywords: susceptibility distortion, diffusion MRI, opposing phase encoding, diffusion tensor, diffusion MRI simulation

OPEN ACCESS

Edited by:

Ludovico Minati,
Tokyo Institute of Technology, Japan

Reviewed by:

Carlo Pierpaoli,
Independent Researcher, Bethesda,
United States

David Atkinson,
University College London,
United Kingdom

***Correspondence:**

Xuan Gu
xuan.gu@liu.se

Received: 11 September 2019

Accepted: 21 November 2019

Published: 05 December 2019

Citation:

Gu X and Eklund A (2019) Evaluation of Six Phase Encoding Based Susceptibility Distortion Correction Methods for Diffusion MRI. *Front. Neuroinform.* 13:76. doi: 10.3389/fninf.2019.00076

1. INTRODUCTION

Analysis of diffusion MRI data is confounded by the presence of susceptibility distortions, caused by an off-resonance field induced by differences in magnetic susceptibility at the air-tissue interface. There are a number of techniques available for correcting susceptibility distortions. Broadly, these techniques can be divided into three types: registration based (RB) methods, fieldmap based (FB) methods and phase encoding based (PB) methods. The first approach involves registration of the distorted image to a structural image without distortions. The second approach involves estimating a map of the B_0 inhomogeneity, and using this along with information about the diffusion acquisition protocol to correct for the distortions. The third approach is based on estimating the underlying distortions using additional data acquired with a different phase encoding direction. For example, it is common to collect LR (left right) and RL (right left) images, or AP (anterior posterior) and PA (posterior anterior) images. Phase encoding based techniques have been demonstrated to outperform the other two approaches (Esteban et al., 2014; Graham et al., 2017), at the cost of a longer scan time.

There are many software packages providing phase encoding based tools for correcting susceptibility distortions, e.g., *animaDistortionCorrection* (*aDC*) (Voss et al., 2006), *animaBMDistortionCorrection* (*aBMDC*) (Hedouin et al., 2017), *DR-BUDDI* (Irfanoglu et al., 2015), *EPIC* (Holland et al., 2010), *HySCO* (Ruthotto et al., 2013), and *TOPUP* (Andersson et al., 2003), summarized in **Table 1**. To date, there is no systematic comparison of existing phase encoding based methods for susceptibility distortion correction. See **Table 2** for an overview of previous comparisons of different distortion correction tools.

The lack of ground truth means that evaluations are typically indirect or qualitative (Jezzard and Balaban, 1995; Wu et al., 2008; Bhushan et al., 2012; Ruthotto et al., 2013; Fritz et al., 2014; Irfanoglu et al., 2015, 2018; Taylor et al., 2016; Hedouin et al., 2017; Wang et al., 2017). Only a few investigations have been carried out with the presence of a ground truth for evaluation of susceptibility distortion correction (Andersson et al., 2003; Esteban et al., 2014; Graham et al., 2017). Hedouin et al. (2017) compared *aDC*, *aBMDC*, and *TOPUP* using phantom data and human data. For the phantom data, both the *aBMDC* and *TOPUP* corrected images appear visually similar for correcting

the distortion, while *aDC* gives visually poorer results. *aBMDC* outperformed *aDC* and *TOPUP* by obtaining smaller landmark position errors. For human data, images corrected using *aDC* contain a mismatch around the lateral ventricles compared with respect to a structural (T1-weighted) image. *aBMDC* and *TOPUP* both obtain a corrected image very close to the structural T1-weighted image. *aBMDC* and *TOPUP* show a very high similarity between the two corrected images C_{AP+PA} and C_{LR+RL} , outperforming *aDC*. Irfanoglu et al. (2015) compared *EPIC*, *DR-BUDDI*, and *TOPUP* using human data. *DR-BUDDI* produced sharper images than *EPIC* and *TOPUP*, showing clearly visible tissue interfaces. Areas such as the inferior temporal lobes and the olfactory bulbs were more accurately reconstructed by *DR-BUDDI* than *EPIC* and *TOPUP*. *DR-BUDDI* resulted in the lowest variability between the two corrected images C_{AP+PA} and C_{LR+RL} , followed by *TOPUP* and then *EPIC*. Overall, *DR-BUDDI* corrected images showed a higher correlation with the undistorted T2-weighted image than did *EPIC* and *TOPUP*.

In this work, we undertake a comparison of six phase encoding based methods for susceptibility distortion correction using both simulated diffusion data and real diffusion data, see **Table 2** for differences between our study and previous comparisons. A brief summary of the six methods is as follows.

- *aDC*: The *aDC* method estimates the displacement field based on the image cumulative distribution function along each line in the PE direction. This method is rather sensitive to noise and it realigns every line independently, so in order to reduce the effect of noise and increase the coherence between the corrected lines, a 3D Gaussian smoothing is applied on the estimated displacement field, which leads to a trade-off between regularity and precision.
- *aBMDC*: The *aBMDC* method adopts the symmetric diffeomorphic image registration idea from Avants et al. (2008) to make the transformed LR (or AP) and RL (or PA) images as similar as possible. The transformation field is obtained using a block-matching algorithm (Ourselin et al., 2000). The squared correlation coefficient is used as the similarity measure between blocks.
- *DR-BUDDI*: The *DR-BUDDI* method also adopts the symmetric diffeomorphic image registration idea from Avants et al. (2008). The first part of the cost function is designed to maximize the similarity between the structural image and

TABLE 1 | Phase encoding based susceptibility distortion correction tools evaluated in this paper.

Tool	Software package	Webpage	References
<i>animaDistortionCorrection</i>	ANIMA	https://github.com/Inria-Visages/Anima-Public/wiki/Registration-tools#epi-distortion-correction	Voss et al., 2006
<i>animaBMDistortionCorrection</i>	ANIMA	https://github.com/Inria-Visages/Anima-Public/wiki/Registration-tools#epi-distortion-correction	Hedouin et al., 2017
<i>DR-BUDDI</i>	TORTOISE	https://tortoise.nibib.nih.gov/tortoise/v313/10-step-31-after-diffprepdr-buddi	Irfanoglu et al., 2015
<i>EPIC</i>		https://github.com/dominicholland/EPIC	Holland et al., 2010
<i>HySCO</i>	SPM	https://bitbucket.org/siawoosh/acid-artifact-correction-in-diffusion-mri/wiki/ACIDSC_wiki_hysco	Ruthotto et al., 2013
<i>TOPUP</i>	FSL	https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/topup/TopupUsersGuide	Andersson et al., 2003

TABLE 2 | A list of susceptibility distortion correction evaluation papers.

Method class	Ground truth availability	Software	Dataset	Conclusion	References
RB, FB	No	RB: In-house C++ code using ITK FB: FSL function <i>PRELUDE</i> and <i>FUGUE</i>	5 human subjects	RB showed an overall better performance than FB RB and FB showed different performance in different brain regions	Wu et al., 2008
FB, PB	No	FB: FSL function <i>PRELUDE</i> and <i>FUGUE</i> PB: <i>EPIC</i>	Human subjects	PB provided superior accuracy than FB	Holland et al., 2010
FB, PB	No	FB: SPM fieldmap toolbox PB: SPM function <i>HySCO</i>	1 human subject	FB gave a better performance than RB	Ruthotto et al., 2013
RB, FB, PB	Yes	RB: ANTs FB: In-house code PB: FSL function <i>TOPUP</i>	A simulated phantom dataset	PB is the most accurate method	Esteban et al., 2014
FB, PB	No	FB: SPM fieldmap toolbox PB: FSL function <i>TOPUP</i> and SPM function <i>HySCO</i>	4 human subjects	PB outperformed FB <i>TOPUP</i> outperformed <i>HySCO</i>	Fritz et al., 2014
PB	No	<i>TORTOISE</i> function <i>DR-BUDDI</i> <i>EPIC</i> FSL function <i>TOPUP</i>	12 human subjects 1 mouse dataset	<i>DR-BUDDI</i> performed the best	Irfanoglu et al., 2015
RB, FB, PB	Yes	RB: <i>NiftyReg</i> function <i>reg_f3d</i> FB: FSL function <i>PRELUDE</i> and <i>FUGUE</i> PB: FSL function <i>TOPUP</i>	A simulated dataset 10 human subjects	FB and PB outperformed RB FB was sensitive to partial volume with air	Graham et al., 2017
PB	No	<i>ANIMA</i> function <i>animaDistortionCorrection</i> <i>ANIMA</i> function <i>animaBMDistortionCorrection</i> FSL function <i>TOPUP</i>	A phantom dataset 5 human subjects	<i>animaBMDistortionCorrection</i> performed the best	Hedouin et al., 2017
RB, FB	No	RB: ANTs function <i>SyN</i> FB: FSL function <i>FUGUE</i>	71 human subjects	RB resulted in higher reliability	Wang et al., 2017
PB	Yes	<i>ANIMA</i> function <i>animaDistortionCorrection</i> <i>ANIMA</i> function <i>animaBMDistortionCorrection</i> <i>TORTOISE</i> function <i>DR-BUDDI</i> <i>EPIC</i> SPM function <i>HySCO</i> FSL function <i>TOPUP</i>	5 simulated datasets 40 human subjects	<i>DR-BUDDI</i> and <i>TOPUP</i> performed better than the other four methods in several cases.	This paper

RB, registration based methods; FB, fieldmap based methods; PB, phase encoding based methods.

the transformed LR/RL (or AP/PA) images. The second part of the cost function is designed to maximize the similarity between the structural image and the geometric average of the transformed LR and RL images. The transformation is constrained along the PE direction. Two co-dependent transformations [one for mapping LR to RL (or AP to PA), the other for mapping RL to LR (or PA to AP)] are used to account for differences in the inhomogeneity between LR (or AP) and RL (or PA) acquisitions. Unlike other methods, *DR-BUDDI* estimates the transformation not only for b_0 images but based on a weighted sum over different diffusion weighted images, in order to preserve tract structures.

- *EPIC*: The *EPIC method* estimates the displacement field by minimizing the sum of squared differences between the transformed LR (or AP) and RL (or PA) images. Two regularization terms are used to regularize the smoothness and amplitude of the displacement. The undistorted image can then be restored using the estimated displacement field.
- *HySCO*: The *HySCO method* adopts the physical distortion model from Chang and Fitzpatrick (1992) and minimizes a distance function to estimate the inhomogeneity, such that the transformed LR (or AP) and RL (or PA) images become as similar as possible. Two regularization terms are used to ensure a differentiable inhomogeneity and positive intensity modulations.
- *TOPUP*: The *TOPUP method* models the displacement field as a linear combination of basis warps consisting of a truncated 3D cosine transform. The weights of the basis warps are estimated using an iterative procedure by minimizing the sum of squared differences between the transformed LR (or AP) and RL (or PA) images. Two options are available for obtaining the distortion free image, they are based on least-squares and Jacobian modulation. The resolution of the displacement field is limited by the highest frequency component of the cosine transform.

We used the POSSUM (Drobnjak et al., 2006, 2010) based diffusion MRI simulator (Graham et al., 2016, 2017), in order to produce realistic diffusion data with susceptibility distortions typically seen in real data. Simulated data can provide ground truth that enables direct and quantitative evaluation. Our analysis directly measures the ability to correctly recover distortion-free data by comparing the corrected b_0 image with its ground truth. We also investigate the suitability of two commonly used indirect metrics, i.e., the difference of the corrected data from the LRRL and APPA pairs (Ruthotto et al., 2013; Graham et al., 2017), and the FA standard deviation over the corrected LR, RL, AP, and PA data (Wu et al., 2008; Irfanoglu et al., 2015). We hope that this work will enable researchers to make more carefully informed choices when designing their processing pipelines.

2. DATA

2.1. Simulated Data

The diffusion data was simulated with 11 volumes of $b = 700$ s/mm², 12 volumes of $b = 2,000$ s/mm² and 1 volume

TABLE 3 | A list of MR parameters (relaxation times T1, T2*, spin density, and chemical shift value) used in our simulations within POSSUM.

	T1 (ms)	T2* (ms)	Spin density	Chemical shift	T2 (ms)
GM	1,331	51	0.86	0	75
WM	832	44	0.77	0	70
CSF	3,700	500	1	0	500

of $b = 0$. The input to the POSSUM diffusion MRI simulator is a collection of three 3D anatomical volumes: gray matter, white matter and cerebro-spinal fluid (CSF). The voxel values in these segmentations reflect the proportion of tissue present in each voxel, in the range [0,1]. The input was generated from the T1-weighted structural image from HCP using the *FSL* function *FAST* (Zhang et al., 2001). The representation of diffusion weighting was achieved by a spherical harmonic fit of order $n = 8$ to the $b = 1,000$ s/mm² shell of the diffusion data, using the *Dipy* (Garyfallidis et al., 2014) module *reconst.shm*. The MR parameters used are listed in **Table 3**. We used a matrix size of 72×86 , 55 slices and a voxel size of 2.5 mm isotropic. The TE was 109 ms, the TR was 9.15 s and the flip-angle was 90°. The PE bandwidth per pixel was 17.1 Hz for LR and RL data, and 14.3 Hz for AP and PA data. The fieldmap was generated from one phase difference volume and two magnitude volumes (one for each echo time) from HCP using the *FSL* function *fsl_prepare_fieldmap* (Jenkinson, 2003). To generate a tight brain extraction for *fsl_prepare_fieldmap*, the brain mask created by the *FSL* function *BET* (Smith, 2002) was further eroded using a 5 mm box kernel. The generated fieldmap was linearly registered to the T1-weighted structural image using the *FSL* function *FLIRT* (Jenkinson and Smith, 2001; Jenkinson et al., 2002). Diffusion data was simulated with four PE directions, i.e., left-right (LR), right-left (RL), anterior-posterior (AP) and posterior-anterior (PA). No other distortions (e.g., eddy-currents and head motion) were included in the simulations. We also simulated a ground truth set, acquired with the same acquisition parameters but no input susceptibility fieldmap. We simulated diffusion data for five subjects (100206, 100307, 100408, 100610, 101006) from the Human Connectome Project (HCP)¹ (Glasser et al., 2013; Van Essen et al., 2013).

2.2. Real Data

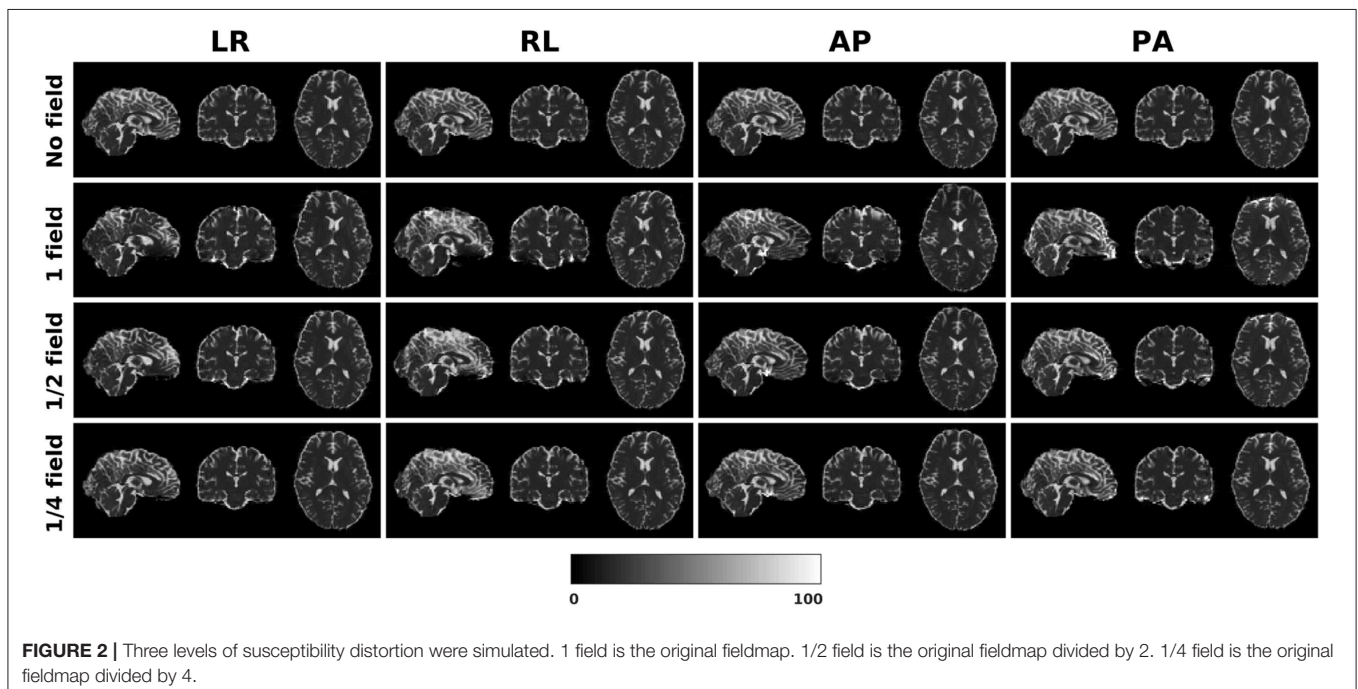
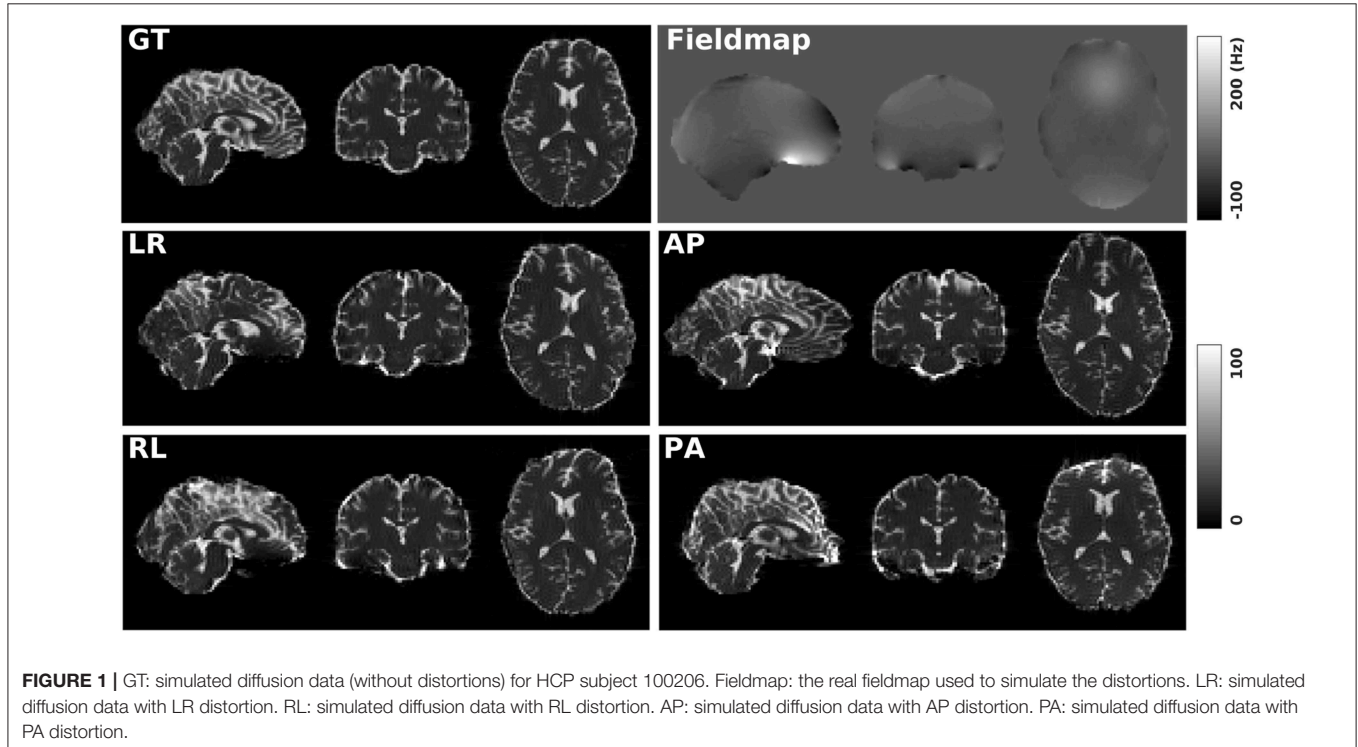
We used 40 subjects from the developing HCP project (Hughes et al., 2017; Bastiani et al., 2019). It provides diffusion data acquired with four PE directions: AP, PA, LR and RL, enabling evaluation using the indirect metric (i.e., comparing APPA corrected to LRRL corrected). The data was acquired on a 3T

¹Data collection and sharing for this project was provided by the Human Connectome Project (U01-MH93765) (HCP; Principal Investigators: Bruce Rosen, M.D., Ph.D., Arthur W. Toga, Ph.D., Van J. Weeden, MD). HCP funding was provided by the National Institute of Dental and Craniofacial Research (NIDCR), the National Institute of Mental Health (NIMH), and the National Institute of Neurological Disorders and Stroke (NINDS). HCP data are disseminated by the Laboratory of Neuro Imaging at the University of Southern California.

Philips Achieva scanner and consists of 4 shells: 20 volumes of $b = 0$, 64 volumes of $b = 400 \text{ s/mm}^2$, 88 volumes of $b = 1,000 \text{ s/mm}^2$ and 128 volumes of $b = 2,600 \text{ s/mm}^2$. The data was acquired using $\text{TR} = 3,800 \text{ ms}$ and $\text{TE} = 90 \text{ ms}$. The matrix size is 128×128 , the number of slices is 64 and the acquired voxel size is $1.17 \times 1.17 \times 1.5 \text{ mm}$.

3. METHODS

For the simulated data, we used the *FSL* function *BET* (Smith, 2002) to create the brain mask from the distortion-free b_0 image. Diffusion tensor fitting and FA calculation were performed using the *FSL* function *dtfit*. For simulated data, we evaluated the



ability of each method to recover the correct intensity at each voxel, by computing error maps between the distortion corrected data and ground truth. We also investigated two indirect metrics. One is the difference of the corrected data from the LRRL and APPA pairs (Ruthotto et al., 2013; Graham et al., 2017), and the other is the FA standard deviation over the corrected LR, RL, AP and PA data (Wu et al., 2008; Irfanoglu et al., 2015). Comparing the uncertainty of data from different preprocessing pipelines is a way to determine if one pipeline is better than the other (Sjölund et al., 2018; Gu et al., 2019). For the real data, we used the brain mask provided with the dataset. We corrected for head motion between LR, RL, AP and PA scans by registering all volumes to the first volume using the *FSL* function *FLIRT* (Jenkinson and Smith, 2001; Jenkinson et al., 2002). For real data, we used indirect evaluation. We compared the corrected data from the LRRL pair with the result from the APPA pair, since ideally the corrected results from the two pairs should be the same.

For each susceptibility distortion correction tool we used default settings and steps provided by the software's basic help

documentation unless otherwise stated. For *DR-BUDDI*, we used the command *DR_BUDDI_withoutGUI* because it is faster than *DR_BUDDI*. By default, the first step in *DIFFPREP* performs Gibbs ringing correction, denoising, head motion correction and eddy-current distortion correction, prior to the second step susceptibility distortion correction by *DR-BUDDI*. However, the simulated data in this paper have only susceptibility distortions, corrections by *DIFFPREP* are therefore not necessary and can even be counterproductive. Additionally, the use of a structural image in *DR-BUDDI* will introduce deviation from the ground truth due to the registration between diffusion and structural images. Therefore, we ran *DR-BUDDI* for the simulated data without the corrections from *DIFFPREP* and without a structural image. For the real data, we ran *DR-BUDDI* with the default setting together with *DIFFPREP*. There is no method referred to in any of the *EPIC* documentation for choosing an alternate phase direction than the *y*-direction (or the LRRL direction with regards to this study). We manually rotated the LR and RL data 90 degrees in the $x - y$ plane before feeding them to *EPIC*.

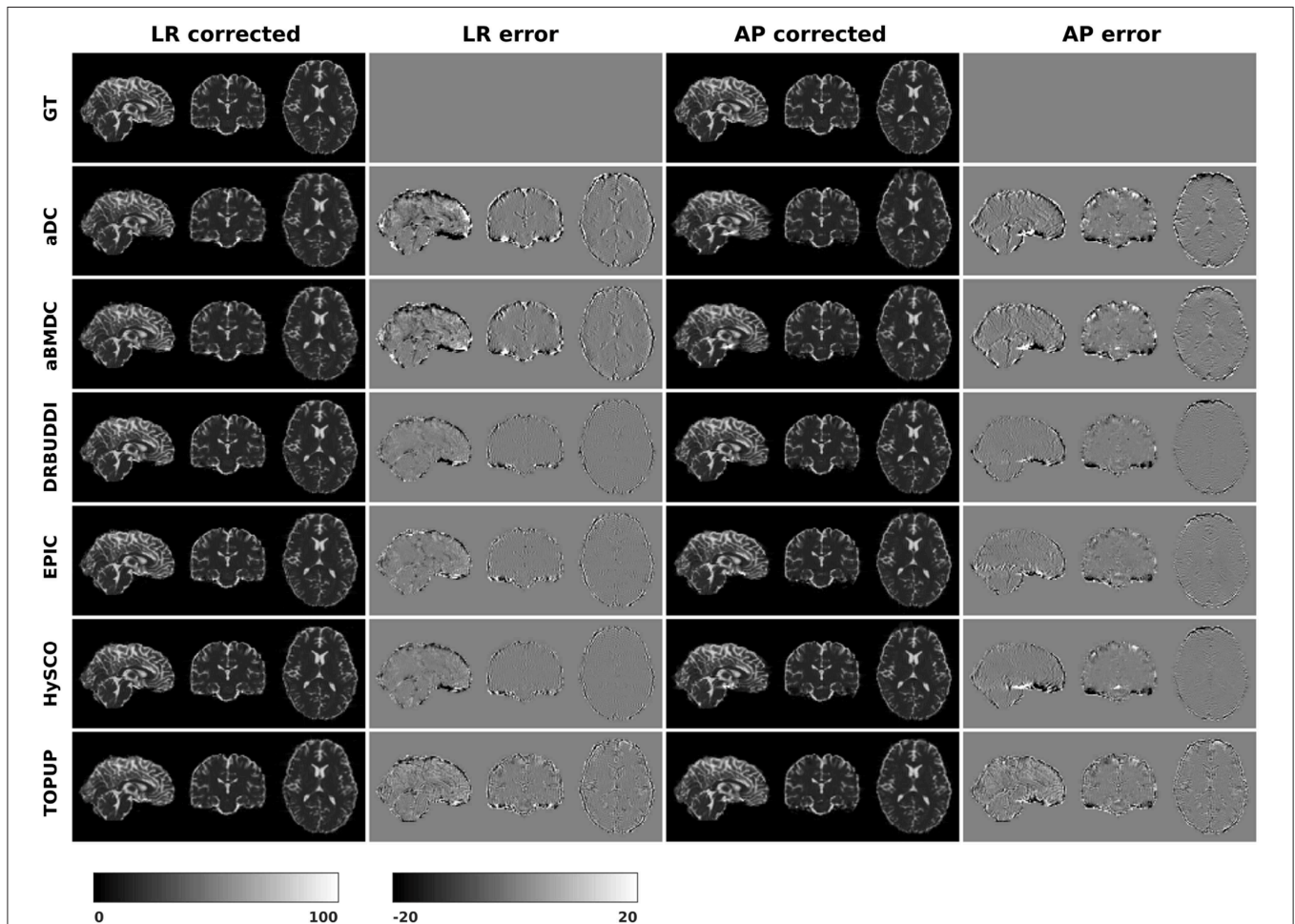


FIGURE 3 | Corrected b_0 volumes for simulated HCP Subject 100206 using the six methods. Error maps were obtained by calculating the difference compared to the ground truth. Correction was carried out for LRRL and APPA pairs, respectively.

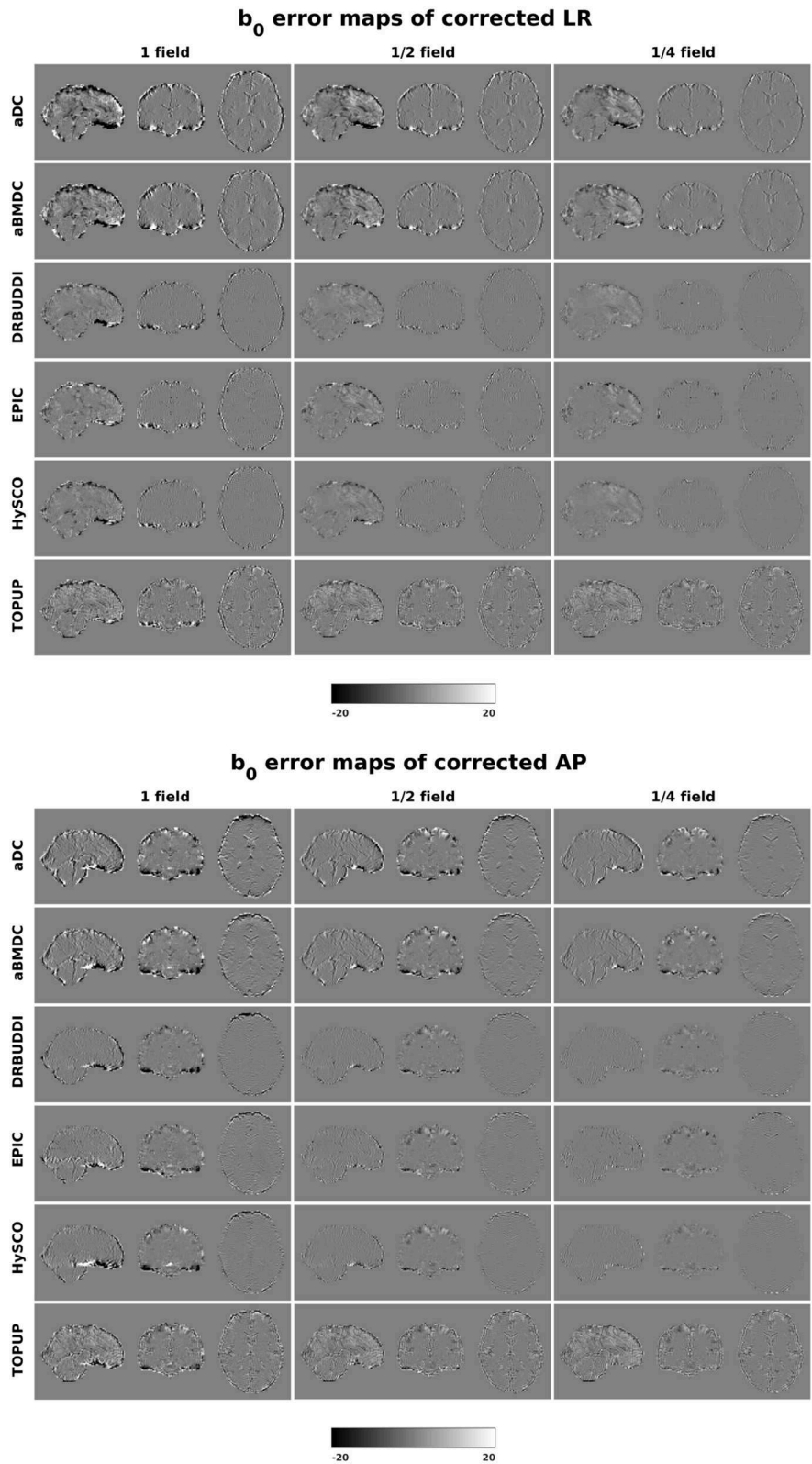
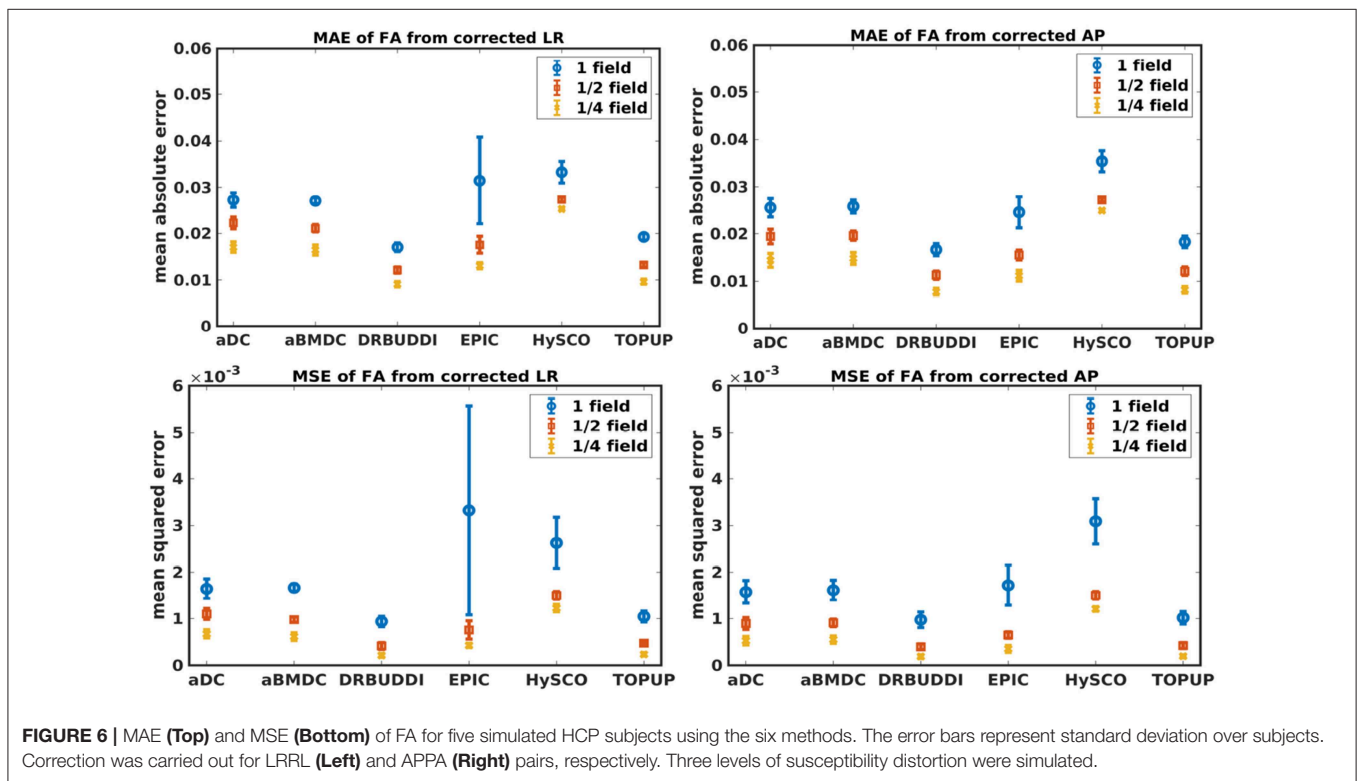
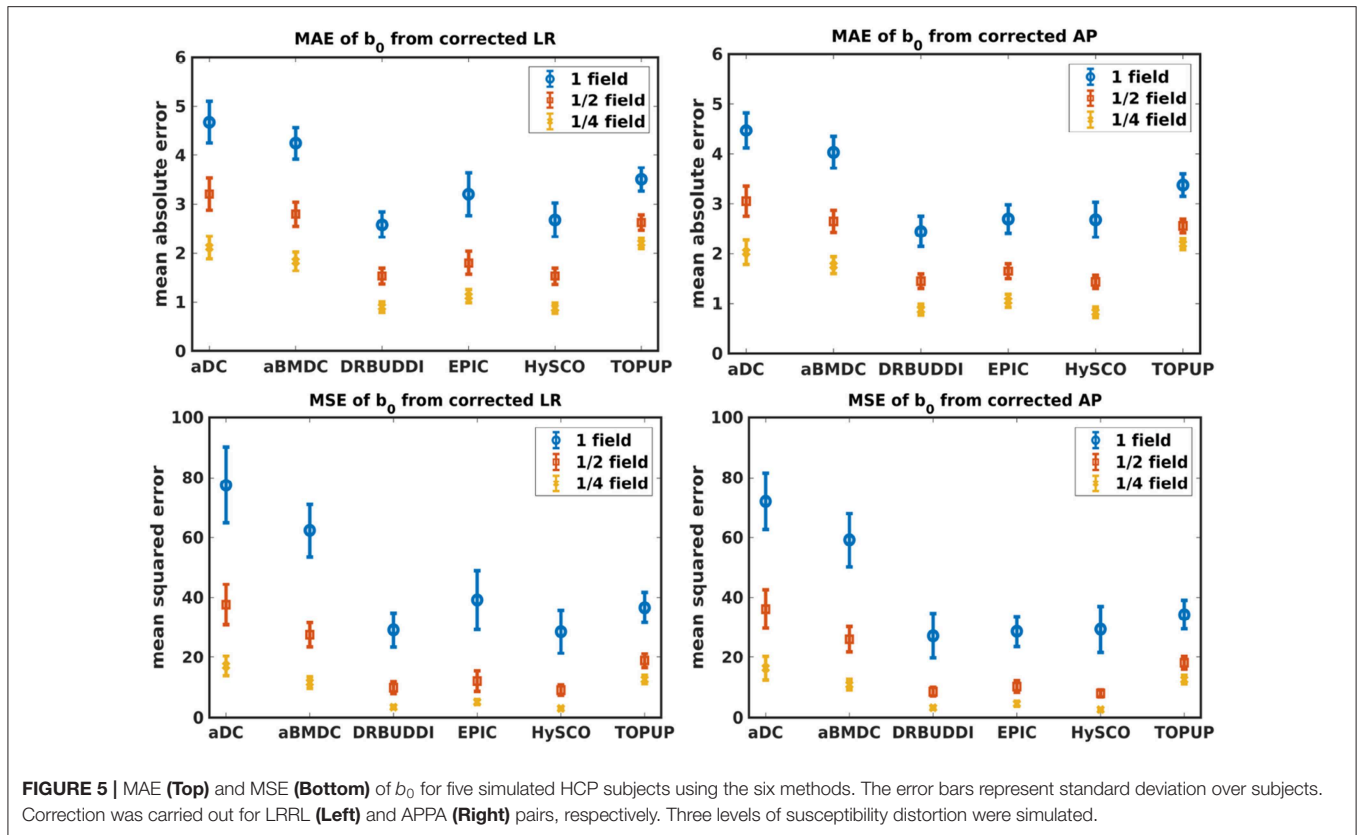
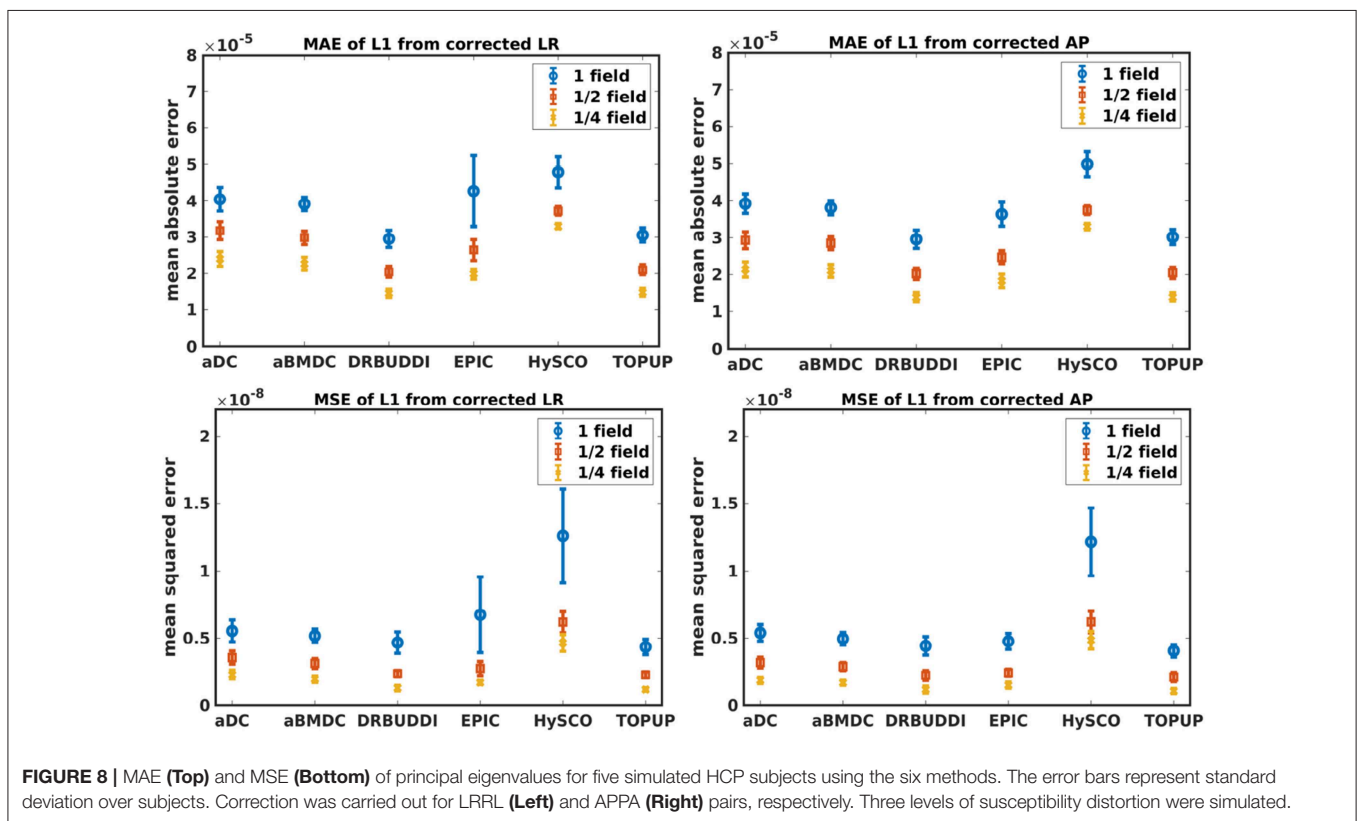
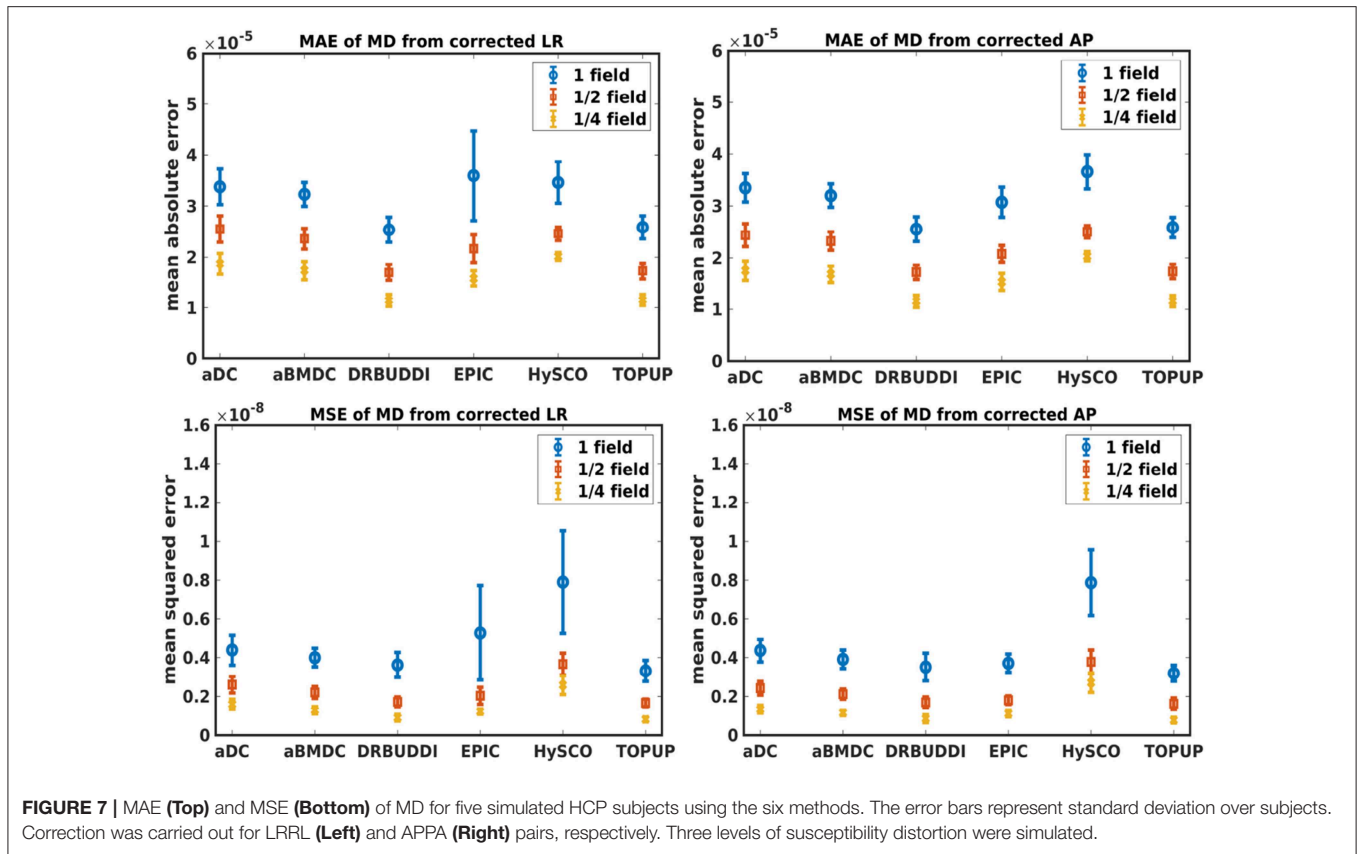


FIGURE 4 | b_0 error maps for simulated HCP Subject 100206 using the six methods. Three levels of susceptibility distortion were simulated. Correction was carried out for LRRL (left) and APPA (right) pairs, respectively. Three levels of susceptibility distortion were simulated.





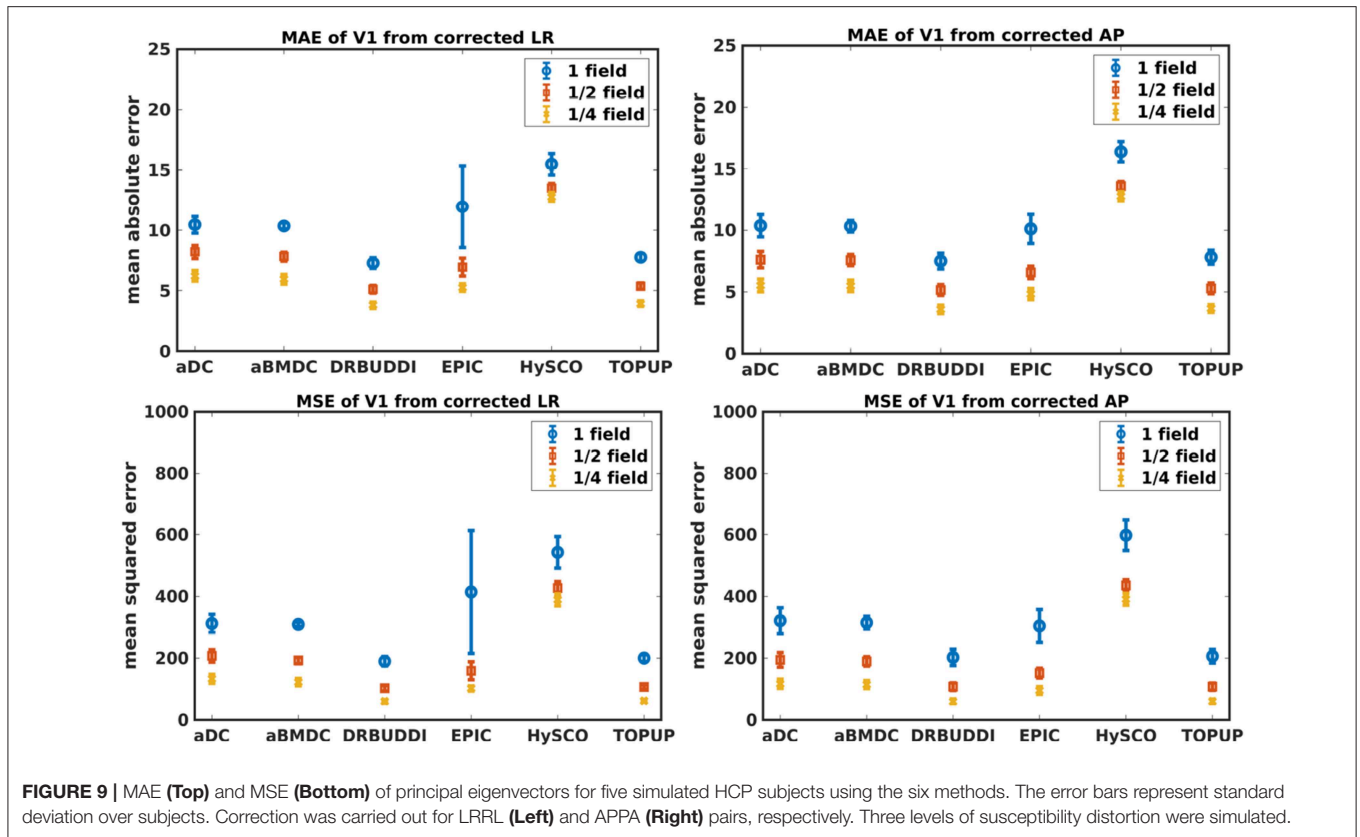


FIGURE 9 | MAE (Top) and MSE (Bottom) of principal eigenvectors for five simulated HCP subjects using the six methods. The error bars represent standard deviation over subjects. Correction was carried out for LRRL (Left) and APPA (Right) pairs, respectively. Three levels of susceptibility distortion were simulated.

TABLE 4 | MAE and MSE of b_0 , FA, MD, principal eigenvalues and principal eigenvectors using the six methods.

Tool / Metric	aADC	aBMDC	DR-BUDDI	EPIC	HySCO	TOPUP
b_0 MAE, LR	4.675	4.242	2.582	3.198	2.680	3.499
b_0 MAE, AP	4.470	4.029	2.449	2.698	2.683	3.369
FA MAE, LR ($\times 10^{-2}$)	2.719	2.698	1.700	3.143	3.326	1.923
FA MAE, AP ($\times 10^{-2}$)	2.550	2.577	1.664	2.454	3.540	1.829
MD MAE, LR ($\times 10^{-5}$)	3.371	3.224	2.535	3.588	3.454	2.584
MD MAE, AP ($\times 10^{-5}$)	3.346	3.197	2.555	3.070	3.653	2.587
λ_1 MAE, LR ($\times 10^{-5}$)	4.043	3.914	2.954	4.272	4.785	3.057
λ_1 MAE, AP ($\times 10^{-5}$)	3.926	3.815	2.959	3.639	4.985	3.019
v_1 MAE, LR ($\times 10$)	1.045	1.034	0.728	1.196	1.550	0.776
v_1 MAE, AP ($\times 10$)	1.037	1.032	0.750	1.011	1.640	0.781

MAE and MSE are averaged across five simulated HCP subjects. For each row, the method with the best performance are in bold.

The data was finally rotated 90° in the other direction after correction. We share our processing scripts on Github², such that other researchers can reproduce and extend our findings (Eklund et al., 2017).

4. RESULTS

4.1. Simulated Data

Figure 1 shows the simulated diffusion data and the fieldmap for HCP Subject 100206, the simulated distortions look very

realistic. We investigated how different levels of distortion affect the correction performance by simulating three levels of susceptibility distortion, as shown in Figure 2. The distortion level was controlled by dividing the fieldmap by a factor 1, 2, or 4.

4.1.1. Direct Metric

In this section, we present the results for the direct metric, including the error maps of b_0 , FA, MD, principal eigenvalues and principal eigenvectors. Figure 3 shows the corrected b_0 images using six different methods, along with error maps obtained by calculating the difference compared to the ground truth images. Correction was carried out for LRRL and APPA pairs,

²<https://github.com/xuagu37/SusceptibilityDistortionCorrection>

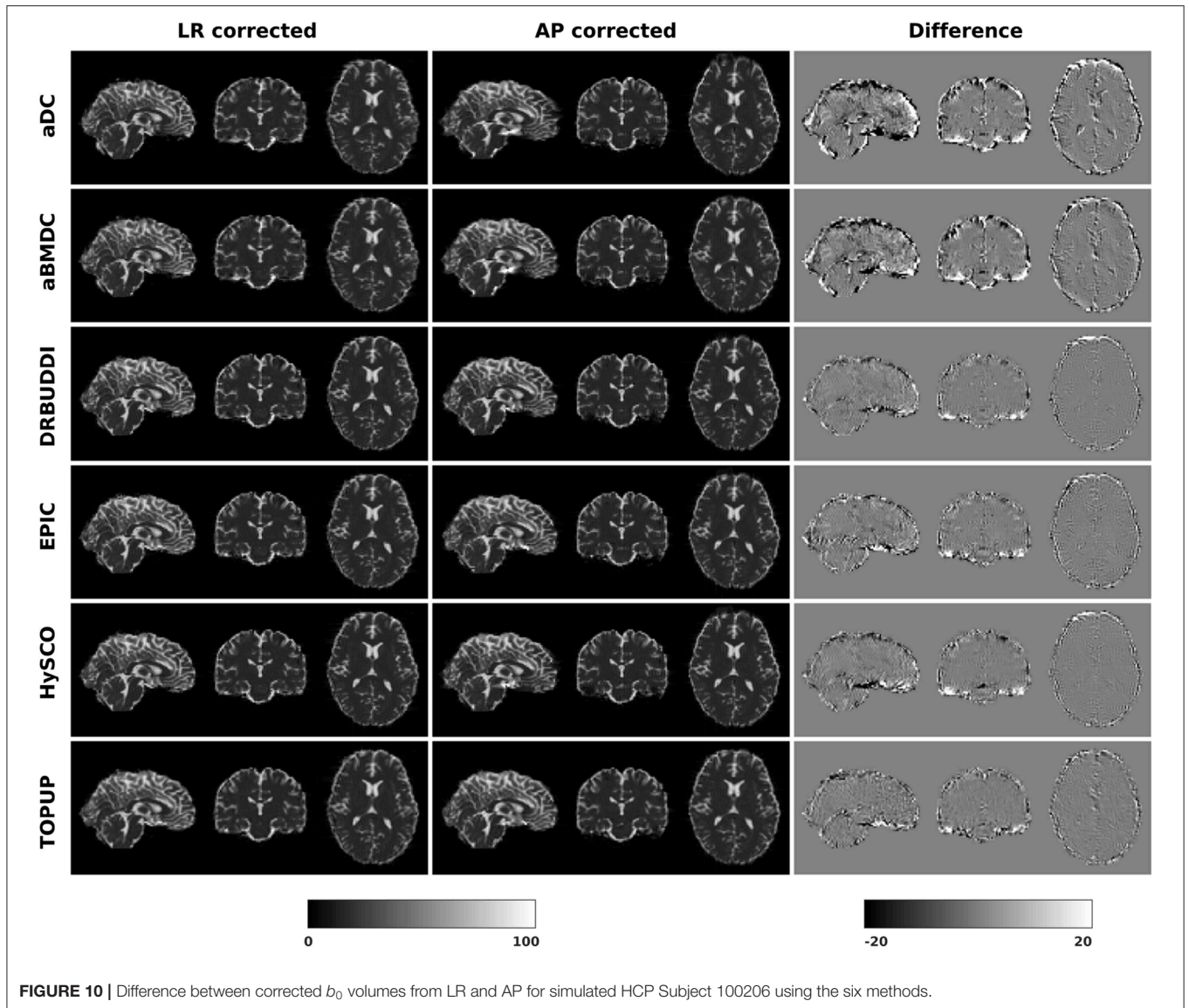


FIGURE 10 | Difference between corrected b_0 volumes from LR and AP for simulated HCP Subject 100206 using the six methods.

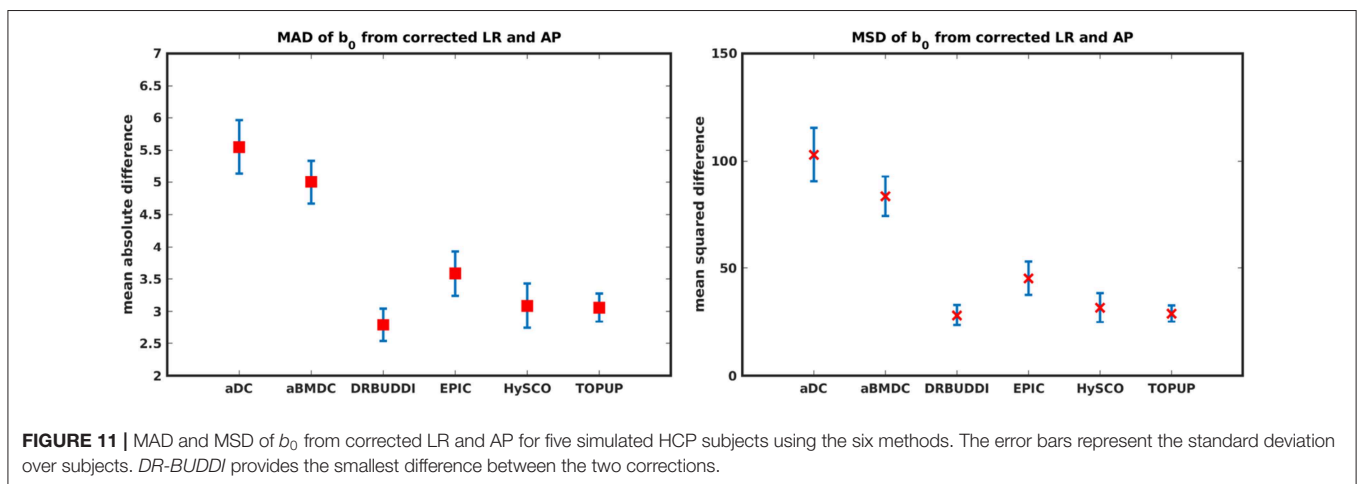


FIGURE 11 | MAD and MSD of b_0 from corrected LR and AP for five simulated HCP subjects using the six methods. The error bars represent the standard deviation over subjects. *DR-BUDDI* provides the smallest difference between the two corrections.

respectively, and we used the corrected LR and AP images as the results for the two pairs. *aDC* and *aBMDC* show large errors for edge voxels. *DR-BUDDI*, *EPIC*, *HySCO*, and *TOPUP* produced very small errors for both LR and AP cases, mainly along edges.

Figure 4 shows the error maps using the six methods for the three levels of distortion. Correction was carried out for LRRL and APPA pairs, respectively. In addition to visual inspection, we computed the mean absolute error (MAE) and the mean squared error (MSE) within the brain and their standard deviations (error bars) for five simulated HCP subjects using the six methods, as shown in **Figure 5**. The results confirmed what we observed in **Figure 4** and quantitatively demonstrates the accuracy and robustness of the six methods. The MAE and MSE decreased with decreasing inhomogeneity field strength, aligned with our predictions. Similarly, **Figures 6–9** show the error maps of FA, MD, principal eigenvalues and principal eigenvectors after diffusion tensor fitting of the corrected data. The error map of principal eigenvectors was obtained by calculating the angle (in degree) between the principal eigenvector and its ground truth. All error maps are quantitatively summarized in **Table 4**.

4.1.2. Indirect Metric

In this section, we present the results for the indirect metric, including b_0 difference maps and the FA standard deviation over the corrected LR, RL, AP, and PA data. To investigate the suitability of LRRL-APPA differences as an indirect metric we plot the corrected LR and AP data, and their differences, as shown in **Figure 10**. Ideally, the corrected LR and the corrected AP

would be identical. The whole-brain mean absolute difference (MAD) and mean squared difference (MSD) were computed for every correction method, as shown in **Figure 11**. The results show larger MAD and MSD for *aDC* and *aBMDC* compared to the other four methods *DR-BUDDI*, *EPIC*, *HySCO*, and *TOPUP*. The results are not completely consistent with the previous results in **Figure 5** for the direct metric. These results demonstrate that the indirect metric (difference maps) shows a

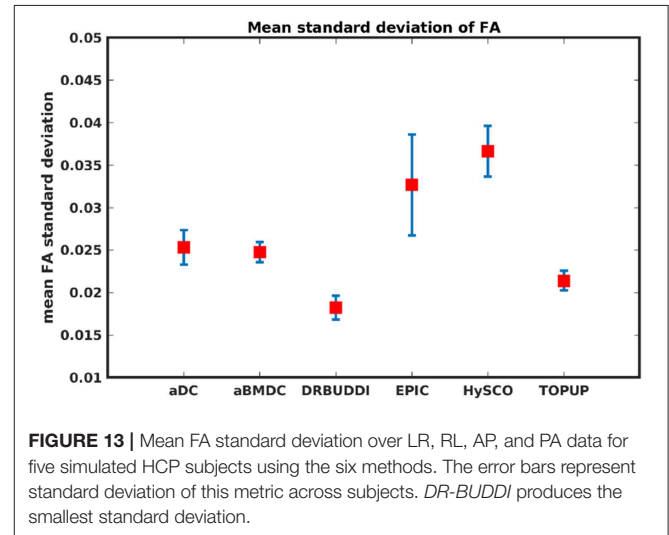


FIGURE 13 | Mean FA standard deviation over LR, RL, AP, and PA data for five simulated HCP subjects using the six methods. The error bars represent standard deviation of this metric across subjects. *DR-BUDDI* produces the smallest standard deviation.

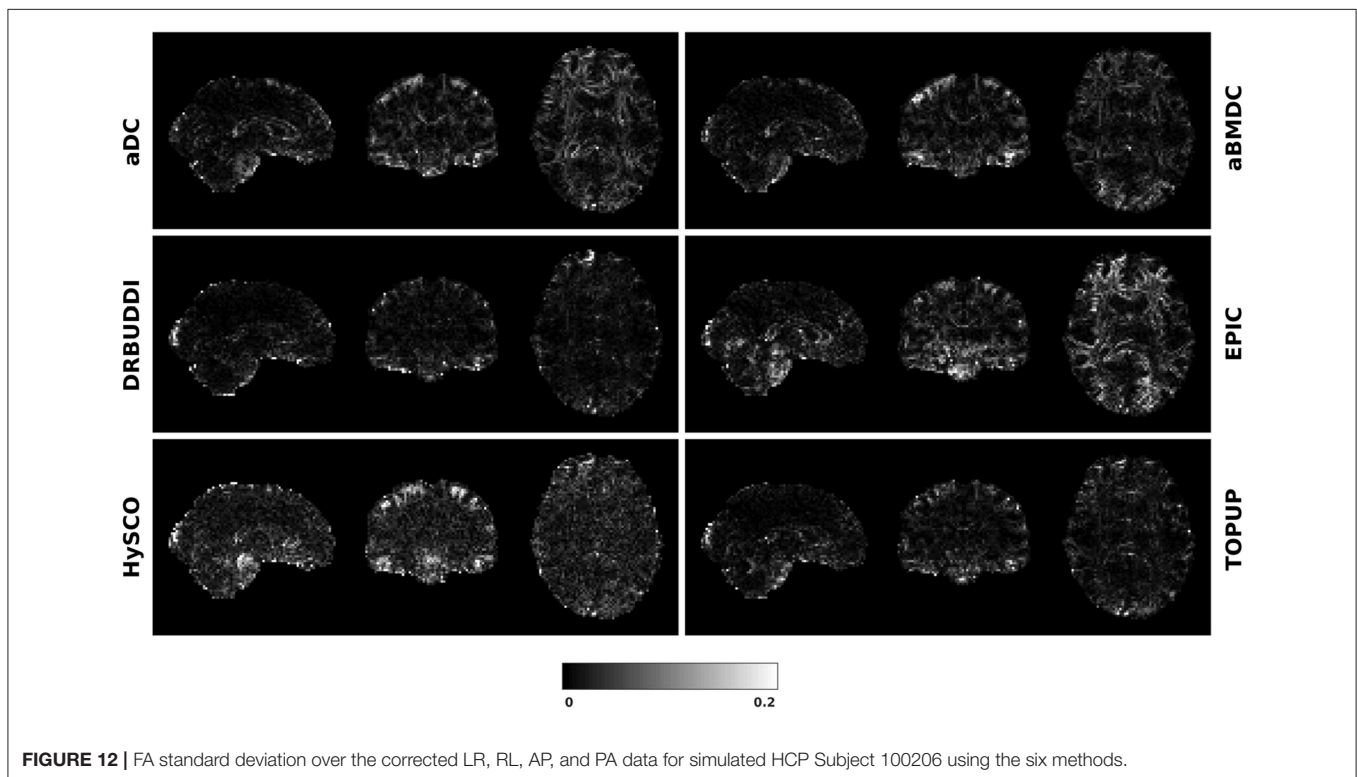
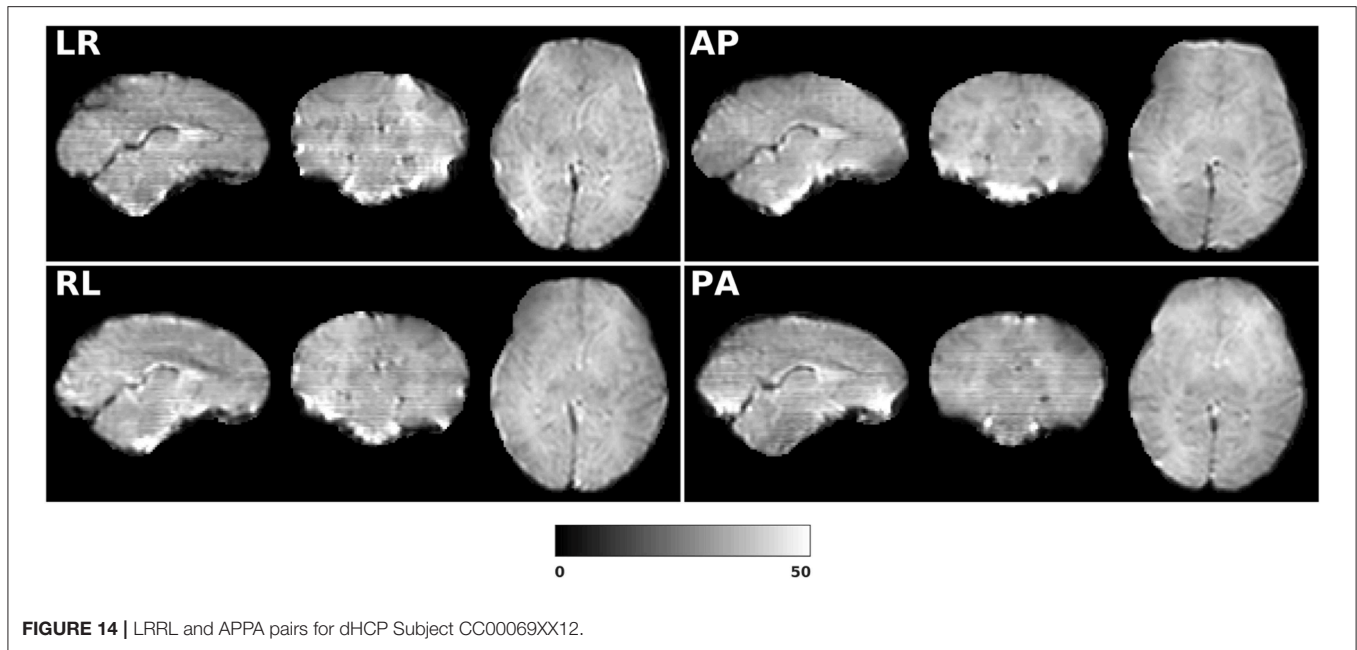


FIGURE 12 | FA standard deviation over the corrected LR, RL, AP, and PA data for simulated HCP Subject 100206 using the six methods.



different ordering of correction performance compared to the direct metric (b_0 error maps).

To investigate the suitability of FA standard deviation as an indirect metric we plot the FA standard deviation over the corrected LR, RL, AP, and PA data, as shown in **Figure 12**. Ideally, the corrected LR, RL, AP, and PA would be identical, which would result in zero FA standard deviation. The whole-brain mean of the standard deviation was computed for every correction method, as shown in **Figure 13**. The results show the largest mean FA standard deviation for *EPIC*. *DR-BUDDI* and *TOPUP* produced very small FA standard deviation over LR, RL, AP, and PA data, and the standard deviation of this metric is also low across subjects. The indirect metric (FA standard deviation) served as a good indication for correction performance compared to the direct metric (FA error maps) as shown in **Figure 6**.

4.2. Real Data

Figure 14 shows the distorted diffusion data from a dHCP subject scanned with four different PE directions. The distortions are more pronounced in regions close to tissue-air interfaces, such as the frontal poles and the temporal lobes near the petrous bone. To evaluate the six methods, we computed the difference between the corrected LR and AP for the dHCP data, as shown in **Figure 15**. Higher difference values are visible in regions most affected by magnetic susceptibility variations such as the boundary regions of the brain. **Figure 16** reports the whole-brain mean of the difference, for the 40 processed subjects. *DR-BUDDI*, *EPIC*, *HySCO*, and *TOPUP* performed almost equally well, which resembled the results for simulated data given in **Figure 11**. We plot the FA standard deviation over the corrected LR, RL, AP, and PA data for the 40 processed subjects in **Figure 17**. It confirmed what we observed in **Figure 13** for results of simulated data.

4.3. Processing Time

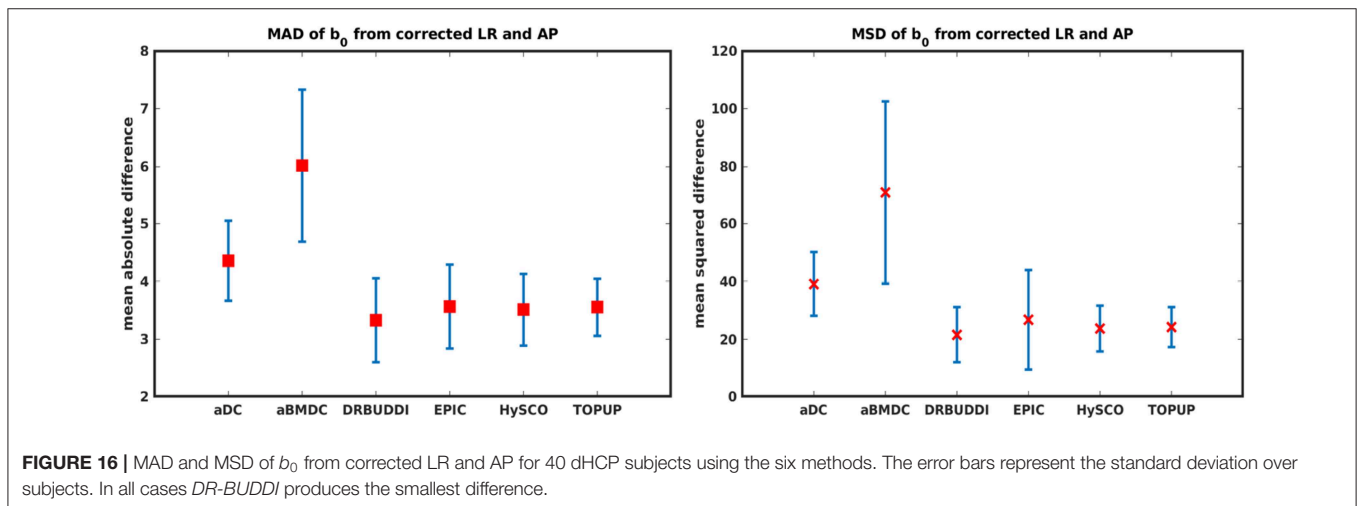
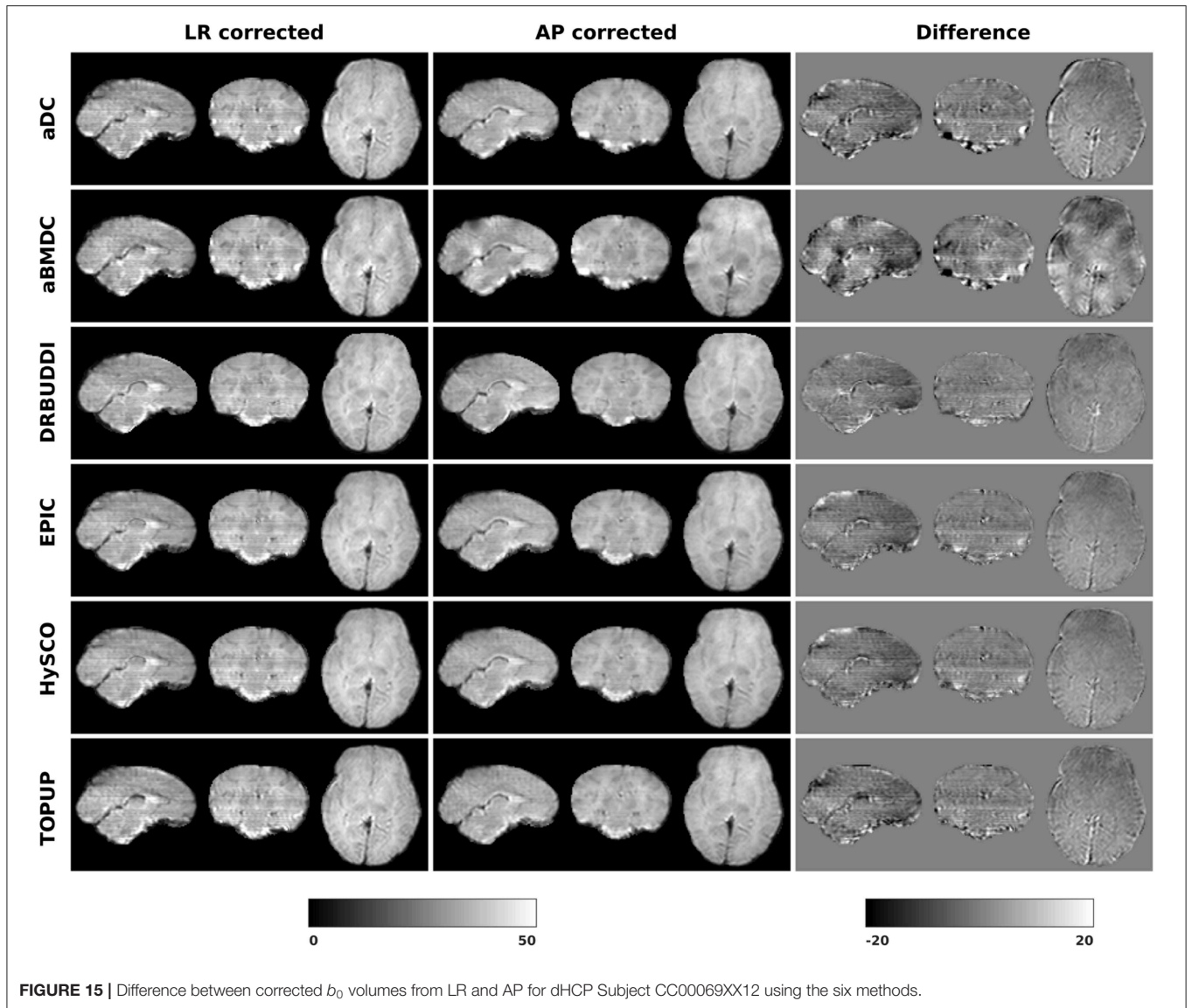
The processing time of one simulated HCP dataset for the six softwares are 2.8 (*aDC*), 35.5 (*aBMDC*), 153.3 (*DR-BUDDI* without its pre-step *DIFFPREP*), 38.0 (*EPIC*), 8.1 (*HySCO*), and 195.7 (*TOPUP*) seconds, respectively. Please note that *aBMDC*, *DR-BUDDI*, *EPIC*, and *HySCO* use several CPU threads to speedup the processing. We used a computer with 32 GB RAM and an Intel(R) Xeon(R) Silver 4114 2.20 GHz CPU (containing 10 cores, which can run 20 threads in parallel).

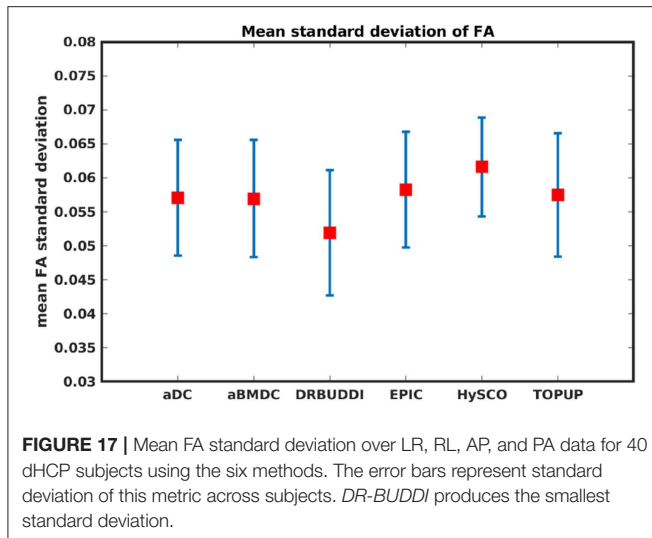
5. DISCUSSION

5.1. Discussion

In this paper, we used both simulated and real data to evaluate six phase encoding based methods for correcting susceptibility distortions. This work is important given that phase encoding based methods have been demonstrated to outperform the other two classes of approaches, and are very frequently used in diffusion data analysis pipelines. It is thus essential to carefully evaluate phase encoding based correction techniques and their limitations. By this work we aim to answer the following two questions. Which method of the six provides the best distortion correction? Are the indirect metrics suitable for measuring the distortion correction performance?

The error map directly measures the ability to correctly recover distortion-free data for different methods. Based on our experiments, we found that *DR-BUDDI*, *EPIC*, *HySCO*, and *TOPUP* were generally superior to the other two methods in achieving better performance for the error map of b_0 , see **Figure 5**. We then applied the transformation field obtained from b_0 images to diffusion weighted volumes and calculated FA, MD, eigenvalues and eigenvectors after diffusion tensor fitting. The





error maps of these four metrics demonstrated that *DR-BUDDI* and *TOPUP* provided the most accurate and robust distortion corrections among the six methods. *EPIC* and *HySCO* performed well in correcting b_0 images but produced poor corrections for diffusion weighted volumes, and these two methods produced rather large errors in terms of the four diffusion tensor metrics as shown in **Figures 6–9**. Therefore, the error map of b_0 should be interpreted together with the error maps of diffusion metrics for a better evaluation of the correction quality. It is notable that *EPIC* showed large errors and very large variance of performance over subjects for the LRRL case when the inhomogeneity field strength was 1. The reason might be that *EPIC* was originally designed for the APPA case and there is no method referred to in any of the *EPIC* documentation for choosing an alternate phase direction than the y -direction (or the LRRL direction with regards to this study). With the phase encode direction chosen in the APPA dimension, susceptibility distortion is manifested as stretching or compression of the image in the APPA direction, which can be more desirable than asymmetric distortion in the LRRL direction (Glover et al., 2012). This might explain that *EPIC* was designed only for the APPA case. It is indeed often the case that the phase encoding direction is chosen to be the y -direction, however there are many images where this is not the case (Kemper et al., 2015; Hughes et al., 2017). *TOPUP* was found to work only for data volumes with even size on x , y , and z directions since the default configuration file (*b02b0.cnf*) performs a factor of 2 subsampling. There are two solutions suggested to eliminate this problem; either cropping or adding dummy data to the 3D volume.

Indirect metrics are often used to evaluate distortion correction for real data due to the absence of ground truth. We investigated the ability of two indirect metrics to measure the correction quality. We validated two of the most promising indirect metrics for correction quality, i.e., the difference between corrected LR and AP data, and the FA standard deviation over the corrected LR, RL, AP, and PA data. The first indirect metric, as shown in **Figures 11, 16**, roughly confirmed what we observed

for the error map of b_0 in **Figure 5**, although the ordering of correction quality is slightly different. Similarly, the FA standard deviation over the corrected LR, RL, AP, and PA data, as shown in **Figure 13**, confirmed what we observed for the error map of FA in **Figure 6**.

5.2. Limitations

Before comparing the quality of the corrections provided by each of these tools, it is important to note that each tool has its customizable parameters. The default parameter settings were used in this work, as it was reasoned that this would be representative of the way the tools were most often used. It was also reasoned that changing default parameters or influencing the inputs prior to correction would lead to a less fair comparison.

It should be noted that we used different *DR-BUDDI* settings for simulated data and real data, because the default pipeline is designed for real data with different types of distortions, but our simulated data only contain susceptibility distortions. It is therefore possible that the corrections in this work may not represent the best corrections attainable by use of these tools.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://db.humanconnectome.org/data/projects/HCP_1200, <http://www.developingconnectome.org/project/data-release-user-guide/>.

AUTHOR CONTRIBUTIONS

XG and AE contributed to the conception and design of the study. XG performed the diffusion MRI data analysis, statistical analysis, and wrote the draft of the manuscript. AE provided comments to the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This research was supported by the Swedish Research Council (grant 2017-04889), Linköping University Center for Industrial Information Technology (CENIIT) and the ITEA3/VINNOVA funded project Intelligence based iMprovement of Personalized treatment And Clinical workflow support (IMPACT).

ACKNOWLEDGMENTS

The authors would like to thank Mark Graham for providing help with the POSSUM diffusion MRI simulator. The authors would also like to thank Renaud Hedouin, Olivier Commowick, Mustafa Okan Irfanoglu, Ivar Thokle Hovden, Lars Ruthotto, and Jesper Andersson for providing comments on the usage of their softwares. We used the data provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research;

and by the McDonnell Center for Systems Neuroscience at Washington University. Part of these results were obtained using data made available from the Developing Human

Connectome Project funded by the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement no. [319456].

REFERENCES

- Andersson, J. L., Skare, S., and Ashburner, J. (2003). How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *NeuroImage* 20, 870–888. doi: 10.1016/S1053-8119(03)00336-7
- Avants, B. B., Epstein, C. L., Grossman, M., and Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med. Image Anal.* 12, 26–41. doi: 10.1016/j.media.2007.06.004
- Bastiani, M., Andersson, J. L. R., Cordero-Grande, L., Murgasova, M., Hutter, J., Price, A. N., et al. (2019). Automated processing pipeline for neonatal diffusion MRI in the developing Human Connectome Project. *NeuroImage* 185, 750–763. doi: 10.1016/j.neuroimage.2018.05.064
- Bhushan, C., Haldar, J. P., Joshi, A. A., and Leahy, R. M. (2012). “Correcting susceptibility-induced distortion in diffusion-weighted MRI using constrained nonrigid registration,” in *Proceedings of the Asia Pacific Signal and Information Processing Association Annual Summit and Conference* (Hollywood, CA: IEEE), 1–9.
- Chang, H., and Fitzpatrick, J. M. (1992). A technique for accurate magnetic resonance imaging in the presence of field inhomogeneities. *IEEE Trans. Med. Imaging* 11, 319–329. doi: 10.1109/42.158935
- Drobnjak, I., Gavaghan, D., Süli, E., Pitt-Francis, J., and Jenkinson, M. (2006). Development of a functional magnetic resonance imaging simulator for modeling realistic rigid-body motion artifacts. *Magn. Resonan. Med.* 56, 364–380. doi: 10.1002/mrm.20939
- Drobnjak, I., Pell, G. S., and Jenkinson, M. (2010). Simulating the effects of time-varying magnetic fields with a realistic simulated scanner. *Magn. Resonan. Imaging* 28, 1014–1021. doi: 10.1016/j.mri.2010.03.029
- Eklund, A., Nichols, T. E., and Knutsson, H. (2017). Reply to brown and behrmann, cox, et al., and kessler et al.: data and code sharing is the way forward for fMRI. *Proc. Natl. Acad. Sci. U.S.A.* 114, E3374–E3375. doi: 10.1073/pnas.1620285114
- Esteban, O., Daducci, A., Caruyer, E., O'Brien, K., Ledesma-Carbayo, M. J., Bach-Cuadra, M., et al. (2014). “Simulation-based evaluation of susceptibility distortion correction methods in diffusion MRI for connectivity analysis,” in *International Symposium on Biomedical Imaging* (Beijing: IEEE), 738–741. doi: 10.1109/ISBI.2014.6867976
- Fritz, L., Mulders, J., Breman, H., Peters, J., Bastiani, M., Roebroek, A., et al. (2014). “Comparison of EPI distortion correction methods at 3T and 7T,” in *Poster Presented at the Annual Meeting of the Organization for Human Brain Mapping* (Hamburg).
- Garyfallidis, E., Brett, M., Amirbekian, B., Rokem, A., Van Der Walt, S., Descoteaux, M., et al. (2014). Dipy, a library for the analysis of diffusion MRI data. *Front. Neuroinformatics* 8:8. doi: 10.3389/fninf.2014.00008
- Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., et al. (2013). The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage* 80, 105–124. doi: 10.1016/j.neuroimage.2013.04.127
- Glover, G. H., Mueller, B. A., Turner, J. A., Van Erp, T. G., Liu, T. T., Greve, D. N., et al. (2012). Function biomedical informatics research network recommendations for prospective multicenter functional MRI studies. *J. Magn. Reson. Imaging* 36, 39–54. doi: 10.1002/jmri.23572
- Graham, M. S., Drobnjak, I., Jenkinson, M., and Zhang, H. (2017). Quantitative assessment of the susceptibility artefact and its interaction with motion in diffusion MRI. *PLoS ONE* 12:e0185647. doi: 10.1371/journal.pone.0185647
- Graham, M. S., Drobnjak, I., and Zhang, H. (2016). Realistic simulation of artefacts in diffusion MRI for validating post-processing correction techniques. *NeuroImage* 125, 1079–1094. doi: 10.1016/j.neuroimage.2015.11.006
- Gu, X., Eklund, A., Özarslan, E., and Knutsson, H. (2019). Using the wild bootstrap to quantify uncertainty in mean apparent propagator MRI. *Front. Neuroinformatics* 13:43. doi: 10.3389/fninf.2019.00043
- Hedouin, R., Commowick, O., Bannier, E., Scherrer, B., Taquet, M., Warfield, S. K., et al. (2017). Block-matching distortion correction of echo-planar images with opposite phase encoding directions. *IEEE Trans. Med. Imaging* 36, 1106–1115. doi: 10.1109/TMI.2016.2646920
- Holland, D., Kuperman, J. M., and Dale, A. M. (2010). Efficient correction of inhomogeneous static magnetic field-induced distortion in echo planar imaging. *NeuroImage* 50, 175–183. doi: 10.1016/j.neuroimage.2009.11.044
- Hughes, E., Cordero-Grande, L., Murgasova, M., Hutter, J., Price, A., Gomes, A. D. S., et al. (2017). “The Developing Human Connectome: announcing the first release of open access neonatal brain imaging,” in *Poster Presented at the Annual Meeting of the Organization for Human Brain Mapping* (Vancouver, BC).
- Irfanoglu, M. O., Modi, P., Nayak, A., Hutchinson, E. B., Sarlls, J., and Pierpaoli, C. (2015). DR-BUDDI (diffeomorphic registration for blip-up blip-down diffusion imaging) method for correcting echo planar imaging distortions. *NeuroImage* 106, 284–299. doi: 10.1016/j.neuroimage.2014.11.042
- Irfanoglu, M. O., Sarlls, J., Nayak, A., and Pierpaoli, C. (2018). Evaluating corrections for eddy-currents and other EPI distortions in diffusion MRI: methodology and a dataset for benchmarking. *Magn. Resonan. Med.* 81, 2774–2787. doi: 10.1002/mrm.27577
- Jenkinson, M. (2003). Fast, automated, n-dimensional phase-unwrapping algorithm. *Magn. Reson. Med.* 49, 193–197. doi: 10.1002/mrm.10354
- Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage* 17, 825–841. doi: 10.1006/nimg.2002.1132
- Jenkinson, M., and Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Med. Image Anal.* 5, 143–156. doi: 10.1016/S1361-8415(01)00036-6
- Jezzard, P., and Balaban, R. S. (1995). Correction for geometric distortion in echo planar images from B_0 field variations. *Magn. Reson. Med.* 34, 65–73. doi: 10.1002/mrm.1910340111
- Kemper, V. G., De Martino, F., Vu, A. T., Poser, B. A., Feinberg, D. A., Goebel, R., et al. (2015). Sub-millimeter T_2 weighted fMRI at 7 T: comparison of 3D-GRASE and 2D SE-EPI. *Front. Neurosci.* 9:163. doi: 10.3389/fnins.2015.00163
- Ourselin, S., Roche, A., Prima, S., and Ayache, N. (2000). “Block matching: a general framework to improve robustness of rigid registration of medical images,” in *International Conference on Medical Image Computing And Computer-Assisted Intervention* (Springer), 557–566.
- Ruthotto, L., Mohammadi, S., Heck, C., Modersitzki, J., and Weiskopf, N. (2013). “Hyperelastic susceptibility artifact correction of DTI in SPM,” in *Bildverarbeitung für die Medizin* (Springer), 344–349. doi: 10.1007/978-3-642-36480-8_60
- Sjölund, J., Eklund, A., Özarslan, E., Herberthson, M., Bänkestad, M., and Knutsson, H. (2018). Bayesian uncertainty quantification in linear models for diffusion MRI. *NeuroImage* 175, 272–285. doi: 10.1016/j.neuroimage.2018.03.059
- Smith, S. M. (2002). Fast robust automated brain extraction. *Hum. Brain Mapp.* 17, 143–155. doi: 10.1002/hbm.10062
- Taylor, P. A., Alhamud, A., van der Kouwe, A., Saleh, M. G., Laughton, B., and Meintjes, E. (2016). Assessing the performance of different DTI motion correction strategies in the presence of EPI distortion correction. *Hum. Brain Mapp.* 37, 4405–4424. doi: 10.1002/hbm.23318
- Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K., et al. (2013). The WU-minn human connectome project: an overview. *NeuroImage* 80, 62–79. doi: 10.1016/j.neuroimage.2013.05.041
- Voss, H. U., Watts, R., Ulug, A. M., and Ballon, D. (2006). Fiber tracking in the cervical spine and inferior brain regions with reversed gradient diffusion tensor imaging. *Magn. Reson. Imaging* 24, 231–239. doi: 10.1016/j.mri.2005.12.007
- Wang, S., Peterson, D. J., Gatenby, J. C., Li, W., Grabowski, T. J., and Madhyastha, T. M. (2017). Evaluation of field map and nonlinear registration methods for

- correction of susceptibility artifacts in diffusion MRI. *Front. Neuroinformatics* 11:17. doi: 10.3389/fninf.2017.00017
- Wu, M., Chang, L.-C., Walker, L., Lemaitre, H., Barnett, A. S., Marengo, S., et al. (2008). "Comparison of EPI distortion correction methods in diffusion tensor MRI using a novel framework," in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (New York, NY: Springer), 321–329.
- Zhang, Y., Brady, M., and Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans. Med. Imaging* 20, 45–57. doi: 10.1109/42.906424

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Gu and Eklund. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.