Comparing the Accuracy and Precision of Multi-echo Combination Methods for Quantitative Susceptibility Mapping Using Laplacian-Based Methods at 3 T

Emma Biondetti[1,2], Anita Karsa[2], Francesco Grussu[3,4], Marco Battiston[3], Marios C. Yiannakas[3], David L. Thomas[5,6], and Karin Shmueli[2]

[1]Institute for Advanced Biomedical Technologies, Department of Neuroscience, Imaging and Clinical Sciences, "D'Annunzio University" of Chieti-Pescara, 66100 Chieti, Italy
[2]Department of Medical Physics and Biomedical Engineering, University College London, London WC1E 6BT, United Kingdom
[3]NMR Research Unit, Queen Square MS Centre, Department of Neuroinflammation, UCL Queen Square Institute of Neurology, University College London, London WC1N 3BG, United Kingdom
[4]Radiomics Group, Vall d'Hebron Institute of Oncology, Vall d'Hebron Barcelona Hospital Campus, 08035 Barcelona, Spain
[5]Dementia Research Centre, UCL Queen Square Institute of Neurology, University College London, London WC1E 6BT, United Kingdom
[6]Wellcome Centre for Human Neuroimaging, UCL Queen Square Institute of Neurology, University College London, London WC1E 6BT, United Kingdom

Corresponding author
Emma Biondetti, PhD
Institute for Advanced Biomedical Technologies, Department of Neuroscience, Imaging and Clinical Sciences, "G. D'Annunzio University" of Chieti-Pescara, via Luigi Polacchi 11, 66100 Chieti, Italy
E-mail: emma.biondetti@unich.it

Funding information
E. Biondetti was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) (award number: 1489882). A. Karsa was supported by the EPSRC-funded UCL Centre for Doctoral Training in Medical Imaging (EP/L016478/1) and the Department of Health's National Institute for Health Research funded Biomedical Research Centre at University College London Hospitals. D. L. Thomas was supported by the UCL Leonard Wolfson Experimental Neurology Centre (PR/ylr/18575). The Queen Square MS Centre, where part of the MRI scans for this work were performed, is supported by grants from the

UK MS Society and by the National Institute for Health Research University College London Hospitals Biomedical Research Centre. F. Grussu was supported by PREdICT, a study at the Vall d'Hebron Institute of Oncology in Barcelona funded by AstraZeneca, and currently receives funding from the postdoctoral fellowships program Beatriu de Pinós (2020 BP 00117), funded by the Secretary of Universities and Research (Government of Catalonia).

Disclosures

The authors declare no conflicts: AstraZeneca was not involved in any aspect concerning this work and has not influenced its content or the decision to submit it for publication.

Abstract

Purpose: To compare different multi-echo combination methods for MRI quantitative susceptibility mapping (QSM). Given the current lack of consensus, we aimed to elucidate how to optimally combine multi-echo gradient-recalled echo (GRE) signal phase information, either before or after applying Laplacian-base methods (LBMs) for phase unwrapping or background field removal.

Methods: Multi-echo GRE data were simulated in a numerical head phantom, and multi-echo GRE images were acquired at 3 T in ten healthy volunteers. To enable image-based estimation of GRE signal noise, five volunteers were scanned twice in the same session without repositioning. Five QSM processing pipelines were designed: one applied nonlinear phase fitting over echo times (TEs) before LBMs; two applied LBMs to the TE-dependent phase and then combined multiple TEs via either TE-weighted or signal-to-noise ratio (SNR)-weighted averaging; two calculated TE-dependent susceptibility maps via either multi-step or single-step QSM and then combined multiple TEs via magnitude-weighted averaging. Results from different pipelines were compared using visual inspection; summary statistics of susceptibility in deep gray matter, white matter, and venous regions; phase noise maps (error propagation theory); and, in the healthy volunteers, regional fixed bias analysis (Bland-Altman) and regional differences between the means (nonparametric tests).

Results: Nonlinearly fitting the multi-echo phase over TEs before applying LBMs provided the highest regional accuracy of χ and the lowest phase noise propagation compared to averaging the LBM-processed TE-dependent phase. This result was especially pertinent in high-susceptibility venous regions.

Conclusion: For multi-echo QSM, we recommend combining the signal phase by nonlinear fitting before applying LBMs.

Abstract word count: 250/250 words

Keywords: magnetic resonance imaging, quantitative susceptibility mapping, multi-echo QSM

Word count (body of the text, excluding the title page, abstract, figure captions and references): 4992/5000

Introduction

MRI quantitative susceptibility mapping (QSM) aims to determine the underlying spatial distribution of tissue magnetic susceptibility (χ) from the phase (ϕ) of a single-echo or a multi-echo gradient-recalled echo (GRE) sequence. It is well-established that multi-echo QSM provides more accurate susceptibility estimates than TE-dependent QSM (1-3). However, for multi-echo QSM, there is still a lack of consensus on how to combine the information acquired at different echo times (TEs) and when to perform this combination in the processing pipeline.

Generally, ϕ measured using a multi-echo GRE acquisition evolves linearly over time:

$$\phi(\boldsymbol{r}, TE) = \gamma \Delta B_{Tot}(\boldsymbol{r}, TE) TE + \phi_0(\boldsymbol{r}) \#(1)$$

where $\boldsymbol{r}$ is a vector of image space coordinates, γ the proton gyromagnetic ratio, ΔB_{Tot} the χ-induced total field perturbation along the scanner's z-axis, and ϕ_0 the TE-independent phase offset at a nominal TE=0 ms.

Based on Equation 1, previous studies using multi-echo GRE protocols and multi-step reconstruction pipelines for brain QSM at 3 T have combined the signal from multiple echoes by either averaging (4-7) or fitting (8-11) before or after the application of Laplacian-based methods (LBMs). For QSM, LBMs are commonly used to unwrap the phase, i.e., remove 2π phase aliasing over TEs or in space (12,13). LBMs are also commonly used to remove background fields (ΔB_{Bg}), i.e., mostly slowly varying field variations inside the region of interest (ROI) generated by sources outside the ROI (14). Based on the properties of the Laplacian of the wrapped phase or ΔB_{Bg}, LBMs perform nonlinear operations, thus they must be applied carefully to avoid introducing inaccuracies in the GRE phase (1).

For multi-echo combination in QSM, this study focuses on approaches based on weighted averaging (4-7) or complex nonlinear fitting (NLFit) (8,9), which outperform simpler approaches based on unweighted averaging (15) or linear fitting (10).

Multi-echo combination of the phase over time via weighted averaging has been performed using either TE-based weighting factors (4,5) or phase SNR-based weighting factors (6). Notably, TE-weighted averaging (TE-wAvg) requires temporal unwrapping of the input multi-echo phase and results in a combined total field map (ΔB_{Tot}), which still contains ΔB_{Bg} contributions. Instead, SNR-weighted averaging (SNR-wAvg) requires performing temporal and spatial unwrapping as well as background field removal on the input multi-echo phase, thus resulting in a combined local field map (ΔB_{Loc}) free from ΔB_{Bg}. Alternatively, a distinct χ map has been calculated at each TE and then combined by weighted averaging with magnitude-based weighting factors (7). To determine the most accurate multi-echo combination method for QSM, this study aimed to compare the effect of these different

approaches on the bias of the estimated tissue χ. To achieve this, in addition to NLFit, TE-wAvg and SNR-wAvg, we implemented two different pipelines based on weighted averaging of TE-dependent χ images: in one, TE-dependent χ images were calculated using a multi-step procedure similar to NLFit, TE-wAvg, and SNR-wAvg (Susc-wAvg); in the other, TE-dependent χ images were calculated in a single-step using total generalized variation (Susc-TGV-wAvg) (16). We hypothesized that combining multiple TEs by fitting the temporally unwrapped multi-echo phase before applying LBMs for spatial phase unwrapping or ΔB_{Bg} removal (NLFit) gives more accurate χ estimates than averaging after applying LBMs (TE-wAvg, SNR-wAvg, Susc-wAvg, or Susc-TGV-wAvg), because LBMs alter the temporal evolution of the phase.

To determine the most precise multi-echo combination method for QSM, this study also aimed to compare the effect of fitting or averaging-based approaches on the propagation of noise from the multi-echo GRE phase into the ΔB_{Tot}, ΔB_{Loc}, and χ maps. Previous studies using fitting (9) or SNR-wAvg (6) have calculated expressions for the noise in the total field map ($\sigma(\Delta B_{Tot})$). For TE-wAvg or SNR-wAvg, expressions for $\sigma(\Delta B_{Tot})$ have not been calculated and were, therefore, derived here. For each multi-echo combination method, expressions for noise propagation from ΔB_{Tot} to the corresponding ΔB_{Loc} and χ images are also presented here.

Bias and noise propagation characteristics of different multi-echo combination methods were investigated using numerically simulated data and images acquired *in vivo*, incorporating LBMs for both phase unwrapping and ΔB_{Bg} removal into the QSM processing pipelines.

Theory

This section presents the expressions for multi-echo combination and noise propagation from the multi-echo signal phase into the ΔB_{Tot}, ΔB_{Loc}, and χ maps.

Multi-echo combination

Multi-echo combination (with n echoes) of the temporally unwrapped phase over time via TE-weighted averaging (TE-wAvg) has been performed as:

$$\Delta B_{Bg}(\boldsymbol{r}) = \frac{1}{\gamma}\sum_{i=1}^{n} w(\boldsymbol{r}, TE_i)\frac{\phi(\boldsymbol{r}, TE_i)}{TE_i} \#(2)$$

with TE-based weighting factors equal to (4,5):

$$w(\boldsymbol{r}, TE_i) = \frac{TE_i}{\sum_{j=1}^{n} TE_j} \#(3)$$

Multi-echo combination of the local field over time via SNR-weighted averaging (SNR-wAvg) has been performed as:

$$\Delta B_{Loc}(\boldsymbol{r}) = \frac{1}{\gamma}\sum_{i=1}^{n} w(\boldsymbol{r}, TE_i)\frac{\Delta B_{Loc}(\boldsymbol{r}, TE_i)}{TE_i} \#(4)$$

With SNR-based weighting factors equal to (6):

$$w(\boldsymbol{r}, TE_i) = \frac{TE_i \exp\left(-TE_i R_2^*(\boldsymbol{r})\right)}{\sum_{j=1}^{n} TE_j \exp\left(-TE_j R_2^*(\boldsymbol{r})\right)} \#(5)$$

An alternative multi-echo combination approach has calculated a distinct χ map at each TE and then applied weighted averaging as (7):

$$\chi(\boldsymbol{r}) = \frac{1}{n}\sum_{i=1}^{n} w(\boldsymbol{r}, TE_i)\chi(\boldsymbol{r}, TE_i) \#(6)$$

with weighting factors (Susc-wAvg and Susc-TGV-wAvg):

$$w(\boldsymbol{r}, TE_i) = \frac{M^2(\boldsymbol{r}, TE_i)TE_i^2}{\sum_{j=1}^{n}\left(M^2\left(\boldsymbol{r}, TE_j\right)TE_j^2\right)} \#(7)$$

where M denotes the signal magnitude.

Multi-echo combination by fitting the (complex) signal phase (Equation 1) enables estimation of both ΔB_{Tot} and ϕ_0 contributions from the coil-combined phase, whereas combining the phase by averaging according to Equations 2-7 does not enable ϕ_0 estimation. Moreover, multi-echo combination using NLFit does not require spatial phase unwrapping, thus avoiding the application of LBMs or other phase unwrapping algorithms before multi-echo combination.

Noise in the total field map

Fitting

Regressing the phase on TE using a weighted linear least square fit enables calculation of a field map ($\omega^{Fit} = \gamma\Delta B_{Tot}^{Fit}$ in rad/s) and a phase offset (ϕ_0^{Fit}) by minimizing the following weighted residual sum of squares (WRSS):

$$WRSS(\boldsymbol{r}) = \sum_{i=1}^{n} w_i\left(\phi(\boldsymbol{r}, TE_i) - \left(\omega^{Fit}(\boldsymbol{r})TE_i + \phi_0^{Fit}(\boldsymbol{r})\right)\right)^2 \#(8)$$

Where w_i is a weighting factor equal to the inverse variance of the phase, i.e. $w_i = 1/\sigma^2\left(\phi(TE_i)\right)$ and n is the total number of TEs (17).

The values of ω^{Fit} and ϕ_0^{Fit} that jointly minimize Equation 8 are calculated by jointly setting to zero the partial derivatives of WRSS relative to ω^{Fit} and ϕ_0^{Fit}:

$$\begin{cases} \dfrac{\partial WRSS(\boldsymbol{r})}{\partial \omega^{Fit}} = 0 \\ \dfrac{\partial WRSS(\boldsymbol{r})}{\partial \phi_0^{Fit}} = 0 \end{cases} \#(9)$$

Omitting the $\boldsymbol{r}$ dependency (because fitting is performed voxel by voxel) solving the system in Equation 9 yields (9):

$$\omega^{Fit} = \frac{\sum_{i=1}^{n} w_i \sum_{i=1}^{n} w_i TE_i \phi(TE_i) - \sum_{i=1}^{n} w_i \phi(TE_i) \sum_{i=1}^{n} w_i TE_i}{\sum_{i=1}^{n} w_i \sum_{i=1}^{n} w_i TE_i^2 - \left(\sum_{i=1}^{n} w_i TE_i\right)^2} \#(10)$$

Based on error propagation, the noise in ΔB_{Tot} calculated using a linear least squares fitting approach (9) is:

$$\sigma\left(\Delta B_{Tot}^{Fit}\right) = \frac{\sqrt{\sum_{i=1}^{n} \frac{1}{\sigma^2\left(\phi(TE_i)\right)}}}{\gamma \sqrt{\sum_{i=1}^{n} \frac{1}{\sigma^2\left(\phi(TE_i)\right)} \sum_{i=1}^{n} \frac{TE_i^2}{\sigma^2\left(\phi(TE_i)\right)} - \left(\sum_{i=1}^{n} \frac{TE_i}{\sigma^2\left(\phi(TE_i)\right)}\right)^2}} \#(11)$$

Equation 11 also corresponds to the a priori noise estimate found in nonlinear least squares fitting (8), thus $\sigma\left(\Delta B_{Tot}^{NLFit}\right) = \sigma\left(\Delta B_{Tot}^{Fit}\right)$.

Weighted averaging

Based on Equations 2-5, ΔB_{Tot} calculated using TE-weighted averaging and ΔB_{Loc} calculated using SNR-weighted averaging respectively have variances equal to:

$$\sigma^2\left(\Delta B_{Tot}^{TE-wAvg}\right) = \sigma^2\left(\frac{\sum_{i=1}^{n} \phi(TE_i)}{\gamma \sum_{i=1}^{n} TE_i}\right) \#(12)$$

$$\sigma^2\left(\Delta B_{Loc}^{SNR-wAvg}\right) = \sigma^2\left(\frac{1}{n\gamma} \sum_{i=1}^{n} \frac{TE_i \exp(-TE_i R_2^*)}{\sum_{j=1}^{n} TE_j \exp\left(-TE_j R_2^*\right)} \frac{\phi(TE_i)}{TE_i}\right) \#(13)$$

Assuming that noise in the single-echo phase is temporally uncorrelated and based on error propagation, the noise in ΔB_{Tot}/ΔB_{Loc} calculated using TE-weighted/SNR-weighted averaging is respectively equal to:

$$\sigma\left(\Delta B_{Tot}^{TE-wAvg}\right) = \frac{\sum_{i=1}^{n} \sigma\left(\phi(TE_i)\right)}{\gamma \sum_{i=1}^{n} TE_i} \#(14)$$

$$\sigma\left(\Delta B_{Loc}^{SNR-wAvg}\right) = \frac{1}{\gamma} \sum_{i=1}^{n} \frac{\exp(-TE_i R_2^*)}{\sum_{j=1}^{n} TE_j \exp\left(-TE_j R_2^*\right)} \sigma\left(\phi(TE_i)\right) \#(15)$$

Equations 14 and 15 omit the $\boldsymbol{r}$ dependency, because they combine multi-echo phase voxel by voxel.

Noise in the local field map

For the NLFit and TE-wAvg pipelines, based on error propagation and the orthogonality (18) of ΔB_{Loc} and the ΔB_{Bg} in the ROI, e.g., the brain:

$$\sigma(\Delta B_{Loc}) = \sqrt{\sigma^2(\Delta B_{Tot}) - \sigma^2\left(\Delta B_{Bg}\right)} \Leftrightarrow \sigma(\Delta B_{Loc}) \leq \sigma(\Delta B_{Tot}) \#(16)$$

where ⇔ denotes an “if and only if” relationship, and $\sigma(\Delta B_{Loc}) = \sigma(\Delta B_{Tot})$ is the worst-case scenario.

Noise in the susceptibility map

The noise in the susceptibility map $\sigma(\chi)$ is calculated based on error propagation and the following relationship (19):

$$\Delta B_{Loc}(\boldsymbol{r}) = d(\boldsymbol{r}) \star \chi(\boldsymbol{r}) \#(17)$$

where d is the magnetic dipole and $\star$ denotes a spatially-dependent convolution. Owing to the circular convolution theorem, the deconvolution operation needed to calculate χ from ΔB_{Loc} (Equation 17) can be performed by point-wise division in the Fourier domain. Thus, if the regularized inverse dipole kernel in k space $\widetilde{D}^{-1}$ can be analytically derived independent of the Fourier transforms of χ or ΔB_{Loc}, $\sigma(\chi)$ can be calculated as (see Supporting Information):

$$\sigma\big(\chi(\boldsymbol{r})\big) = \frac{1}{B_0}\sqrt{FT^{-1}\left(\widetilde{D}^{-2}(\boldsymbol{k})FT\left(\sigma^2\big(\Delta B_{Loc}(\boldsymbol{r})\big)\right)\right)} \#(18)$$

where FT and FT^{-1} respectively denote the direct and inverse Fourier transforms, and $\boldsymbol{k}$ denotes k-space coordinates. Deriving an analytical expression for $\widetilde{D}^{-1}$ is possible, for example, when considering thresholded k-space division (TKD), or the Tikhonov-regularized minimal norm solution (9,19,20).

Weighted averaging

For weighted averaging of TE-dependent χ (Equations 6 and 7), based on phase error propagation over time and Equation 16, $\sigma\big(\Delta B_{Loc}(TE_i)\big)$ equals:

$$\sigma\big(\Delta B_{Loc}(TE_i)\big) = \frac{\sigma\big(\phi(TE_i)\big)}{\gamma TE_i} \#(19)$$

Based on Equations 6, 7, 18, and 19, $\sigma(\chi^{Susc-wAvg})$ is the square root of:

$$\begin{aligned}
\sigma^2(\chi^{Susc-wAvg}) = \\
&= \frac{1}{(B_0\gamma n)^2}\sum_{i=1}^{n}\chi^2(TE_i)\left(\frac{4M^2(TE_i)TE_i^4\sigma^2\big(M(TE_i)\big)}{\sum_{j=1}^{n}\big(M^2(TE_j)TE_j^2\big)^2}\right.\\
&+\frac{M^4(TE_i)TE_i^4}{\sum_{j=1}^{n}\big(M^2(TE_j)TE_j^2\big)^4}\sum_{j=1}^{n}4M^2(TE_j)TE_j^4\sigma^2\left(M(TE_j)\right)\\
&\left.-\frac{2M^2(TE_i)TE_i^2}{\big(\sum_{j=1}^{n}M^2(TE_j)TE_j^2\big)^3}4M^2(TE_i)TE_i^4\sigma^2\big(M(TE_i)\big)\right)\\
&+\left(\frac{M^2(TE_i)TE_i^2}{\sum_{j=1}^{n}M^2(TE_j)TE_j^2}\right)^2\sigma^2\big(\chi(TE_i)\big)
\end{aligned} \tag{20}$$

Notably, Equation 20 only applies to the Susc-wAvg pipeline, as it requires analytically describing $\sigma(\chi)$ (Equation 18). In the Susc-TGV-wAvg pipeline, such an equation could not be calculated, because the total generalized variation (TGV) method estimates χ iteratively (16).

Methods

When not otherwise stated, image analysis was performed using Matlab (R2017b, The MathWorks, Natick, MA) and statistical analysis using Stata (R15, StataCorp LP, College Station, TX). A preliminary version of this study was presented at the 2016 annual meeting of the International Society for Magnetic Resonance in Medicine (21).

In vivo data acquisition

Multi-echo 3D GRE imaging of ten healthy volunteers (average age/age range: 26/22-30 years, five females) was performed in two centers (University College London Hospital and Queen Square Multiple Sclerosis Centre, University College London) equipped with the same 3 T MRI system (Philips Achieva, Philips Healthcare, Best, NL; 32-channel head coil). Five subjects were acquired in each center. All the volunteers provided written informed consent, and the local research ethics committees approved the experimental sessions. Images were acquired using a transverse orientation, field of view=240x240x144 mm^3, voxel size=1 mm^3 isotropic, flip angle=20 degrees, repetition time=29 ms, five evenly spaced echoes (TE_1/TE spacing=3/5.4 ms), bandwidth=270 Hz/pixel, SENSE (22) along the first/second phase encoding directions=2/1.5, flyback gradients along the frequency encoding (FE) direction, and without flow compensating gradients (23). Five subjects were scanned twice within the same session without repositioning to enable image-based calculation of SNRs in the magnitude and phase images.

Data simulation from a numerical head phantom

To ensure the availability of ground-truth χ values against which to test the accuracy of QSM pipelines, a piece-wise constant Zubal numerical head phantom (24) was used to generate simulated noisy multi-echo complex data (1). The resulting phantom had matrix size=256x256x128 $voxels^3$, resolution=1.5 mm isotropic, and the following regions of interest (ROIs): the caudate nucleus (CN), globus pallidus (GP), putamen (PU), thalamus (TH), superior sagittal sinus (SSS), grey and white matter (GM and WM), and cerebrospinal fluid (CSF). Compared to our previous study (1), the numerical phantom was updated by adding noise to achieve realistic regional standard deviations (SDs) for χ, the proton density (M_0), and the transverse relaxation rate (T_2^*) (see Supporting Information).

Data preprocessing

A brain mask was calculated for each subject by applying FSL BET (25,26) with robust brain center estimation (threshold=0.3) to the magnitude image at the longest TE. This choice of TE accounted for the greater amount of signal dropout near regions of high-χ gradients compared to shorter TEs.

A whole-brain mask for the Zubal phantom was calculated by applying FSL BET with robust brain center estimation (threshold=0.5) to the T_2^* map of the numerical phantom (1).

Processing pipelines for QSM

Five distinct processing pipelines (Figure 1) were applied to both the numerically simulated and the healthy volunteer data. Three of these pipelines (NLFit, TE-wAvg, and SNR-wAvg) combined the phase across TEs at different stages before performing the ΔB_{Loc}-to-χ inversion. The two other pipelines (Susc-wAvg and Susc-TGV-wAvg) first calculated a distinct χ map at each TE and then combined the χ maps. The following paragraphs describe each processing pipeline in detail.

The NLFit pipeline (1) first combined the complex GRE signal by nonlinear fitting over TEs (8) using the Cornell QSM software package's Fit_ppm_complex function (27). It then applied simultaneous phase unwrapping and ΔB_{Bg} removal using Sophisticated Harmonic Artifact Reduction for Phase data (SHARP) (28), a direct solver of the discretized Poisson equation (28).

SHARP was chosen because it has been widely used in the literature on QSM and is both robust and numerically efficient (14). Moreover, in our recent study comparing multi-echo and TE-dependent QSM, a multi-echo pipeline incorporating SHARP gave highly accurate multi-echo QSM values (1). As in this previous study (1), SHARP was implemented using the minimum-size 3-voxel isotropic 3D Laplacian kernel (28), and it was applied using a threshold for truncated singular value decomposition (tSVD) equal to 0.05 and a brain mask that was eroded by two (numerical phantom) or four voxels (healthy volunteers).

The TE-wAvg processing pipeline first applied Laplacian unwrapping to the phase at each TE using a threshold for tSVD equal to 10^{-10} (i.e., the default value in (27)). Second, it calculated ΔB_{Tot} by averaging the unwrapped phase according to Equations 2 and 3 (4,5). Then, it calculated ΔB_{Loc} by applying SHARP with the same parameters as in NLFit. As substantial residual background fields remained, SHARP was applied again with variable kernel size (V-SHARP) with an initial kernel radius of approximately 40 mm, and a step size and final kernel radius of 1 voxel (29,30).

The SNR-wAvg processing pipeline first applied simultaneous phase unwrapping and ΔB_{Bg} removal to the phase at each TE using SHARP with the same parameters as in NLFit. It

then calculated ΔB_{Loc} by averaging the unwrapped and background-field-free phase according to Equations 4 and 5 (6) and removing residual background fields by applying V-SHARP as in TE-wAvg.

In these three pipelines, ΔB_{Loc}-to-χ inversion was performed using Tikhonov regularization with correction for susceptibility underestimation and using the L-curve method to determine the optimal value for the regularization parameter (9,28,31). This inversion method was chosen because it is computationally efficient and substantially reduces streaking artifacts relative to TKD (32).

The Susc-wAvg processing pipeline calculated a distinct χ map at each TE by applying simultaneous phase unwrapping and ΔB_{Bg} removal using SHARP as in NLFit, removing residual background fields by applying V-SHARP as in TE-wAvg and SNR-wAvg, and performing the ΔB_{Loc}-to-χ inversion using Tikhonov regularization as in NLFit, TE-wAvg and SNR-wAvg. This pipeline then calculated a combined χ map by applying Equation 6 with the magnitude-based weights in Equation 7 (7).

The Susc-TGV-wAvg processing pipeline applied one-step TGV (16) to the phase at each TE, and then calculated a combined χ map as in Susc-wAvg. The TGV method was chosen because it avoids stair-casing artifacts in the resulting χ map while correctly preserving structural borders (16). Moreover, in our recent study comparing multi-echo and TE-dependent QSM, TGV provided highly accurate TE-dependent QSM images (1). TGV was implemented using Singularity (Sylabs, Albany, CA, (33)) with the default parameter values $(\alpha_1, \alpha_0) = (0.0015, 0.005)$, which are optimal for medical imaging applications (16).

ROI segmentation in the healthy volunteer images

Regional χ values were compared within the simulated and *in vivo* data, similarly to our previous study comparing multi-echo and TE-dependent QSM (1). ROIs similar to those in the numerical phantom were segmented *in vivo*: the CN, GP, PU, TH, posterior corona radiata (PCR) as a WM ROI, and the straight sinus (StrS) as a venous ROI. Briefly, for each subject, the CN, GP, PU, TH and PCR were segmented based on the Eve χ atlas (34), whose GRE magnitude image was aligned to each subject's fifth-echo magnitude image using NiftyReg (35,36) (TE_{Eve}/TE_5=24/24.6 ms). The ITK-SNAP active contour segmentation tool (37) was used to segment the StrS in each subject based on the fifth-echo magnitude image, which presented the best contrast between the StrS and the surrounding brain tissue.

Quantitative evaluation of the measured χ

In the numerical phantom simulations, each QSM pipeline's performance relative to the ground truth was visually assessed by calculating a difference image between the corresponding χ map and χ^{True}. In the volunteers, owing to the lack of a ground truth, similar difference images were calculated relative to χ^{NLFit}, because the NLFit pipeline provided the lowest regional susceptibility bias in the numerical phantom simulations (see the Results section).

In the numerical phantom simulations, the means and SDs of χ were calculated for each pipeline in each ROI with $\chi^{True} \neq 0$, i.e., the CN, GP, PU, TH, SSS, and WM. The root mean squared errors (RMSEs) of χ relative to χ^{True} were also calculated throughout the brain volume as previously described (1).

In the volunteers, regional means and SDs of χ were calculated for each processing pipeline and compared against χ values in subjects of a similar age from the QSM literature. In the volunteers, the RMSEs could not be calculated because of the lack of a ground truth.

For visualization purposes in the healthy volunteers, if it was possible to pool the regional means and SDs of χ across subjects (i.e., all *intra*subject SDs of χ were larger than the *inter*subject SD of χ), the pooled averages and SDs were calculated as previously described (1).

Noise propagation maps

Only the healthy volunteers scanned twice were considered for this analysis. To enable image SNR calculation, in one healthy volunteer, five 20x20-voxel ROIs were drawn on a sagittal slice of the first-echo magnitude image (38), including both the GM and WM and excluding regions with artifacts induced by SENSE, motion, or flow. All five ROIs were applied across the other four volunteers by using rigid alignment transforms (NiftyReg (36)) between the first-echo magnitude images.

In each subject, ROI-based magnitude (M_{ROI}), magnitude noise ($\sigma_{ROI}(M)$) and phase noise values ($\sigma_{ROI}(\phi)$) were calculated based on the SNR difference method (39):

$$M_{ROI}(TE,\boldsymbol{r}) = \frac{1}{2} mean\big(M_1(TE,\boldsymbol{r}) + M(TE,\boldsymbol{r})\big), \qquad \boldsymbol{r} \in ROI \#(22)$$

$$\sigma_{ROI}\big(M(TE,\boldsymbol{r})\big) = \sqrt{\frac{R}{2}} SD\big(M_1(TE,\boldsymbol{r}) - M_2(TE,\boldsymbol{r})\big), \qquad \boldsymbol{r} \in ROI\ \#(23)$$

$$\sigma_{ROI}\big(\phi(TE,\boldsymbol{r})\big) = \sqrt{\frac{R}{2}} SD\big(\phi_1(TE,\boldsymbol{r}) - \phi_2(TE,\boldsymbol{r})\big), \qquad \boldsymbol{r} \in ROI \#(24)$$

Here, M_1/ϕ_1 and M_2/ϕ_2 respectively denote the magnitude/phase images from the first and second scan, $R = 3$ is the 2D SENSE factor calculated by multiplying the SENSE factors applied along the two phase encoding directions (40). The values calculated based on Equations 22-24 were averaged across the five ROIs, to calculate summary values of magnitude, magnitude noise, and phase noise at each TE.

For both the numerical phantom simulations and the healthy volunteers, a phase noise $\sigma(\phi)$ map was calculated at each TE as (17,41):

$$\sigma\big(\phi(TE,\boldsymbol{r})\big) = c(TE)\frac{1}{SNR(TE,\boldsymbol{r})} = c(TE)\frac{\sigma_{ROI}\big(M(TE,\boldsymbol{r})\big)}{M(TE,\boldsymbol{r})} \#(24)$$

where $c(TE)$ was a constant equal to 1 (by definition) in the numerical phantom simulations and equal to:

$$c(TE) = \frac{\sigma_{ROI}(TE)}{\frac{1}{SNR_{ROI}(M)}} = \frac{\sigma_{ROI}\big(\phi(TE)\big)\sigma_{ROI}\big(M(TE)\big)}{M(TE)} \#(25)$$

in the healthy volunteers. Notably, calculating phase noise analytically as in Equation 24 enabled the direct comparison of noise propagation between numerical simulations and data acquired in vivo.

For the NLFit, TE-wAvg and SNR-wAvg pipelines, the $\sigma(\Delta B_{Loc})$ map was calculated based on the multi-echo $\sigma(\phi)$ maps according to Equations 11 and 14-16. Then, the $\sigma(\chi)$ map was calculated according to Equation 18 with $B_0 = 3\ \mathrm{T}$ and the Tikhonov-regularized inverse magnetic dipole kernel with regularization parameter α (9):

$$\widetilde{D}^{-2}(\boldsymbol{k};\alpha) = \frac{D(\boldsymbol{k})}{D(\boldsymbol{k})^2 + \alpha} \#(26)$$

and correction for χ underestimation (28). For the Susc-wAvg pipeline, the $\sigma(\chi)$ map was calculated based on Equation 20. Note that, as explained in the Theory section, $\sigma(\chi)$ maps were not calculated for the Susc-TGV-wAvg pipeline, because an analytical expression for $\sigma(\chi)$ could not be derived using TGV.

To compare $\sigma(\chi)$ maps across pipelines, a line profile was traced in the same location of all $\sigma(\chi)$ images and the $\sigma(\chi)$ value of each voxel along this profile was plotted. To compare the noise intensity and its variability between processing pipelines, the mean and SD of this representative line profile were also calculated.

Statistical analysis

Statistical analyses were performed based on the healthy subject data. For each ROI and each pair of pipelines, potential symmetric biases between distinct processing pipelines were investigated using Bland-Altman analysis of the average χ. For each ROI, statistically significant differences between pipelines were tested by considering the corresponding

distributions of average χ values across subjects. To assess whether to apply parametric paired t-tests or nonparametric sign tests, the normal distribution of the differences between paired χ values was assessed using the Shapiro-Wilk test. All statistical tests were two-tailed and an uncorrected *P*-value of 0.05 or less was considered significant.

Results

Pooling of χ measurements

For each ROI and each processing pipeline, all *intra*subject SDs of χ were larger than the *inter*subject SD. Thus, pooled means and SDs were calculated (1).

Performance of pipelines for multi-echo QSM

In the numerical phantom, Figure 2 shows the QSM images calculated by each processing pipeline for multi-echo QSM (A-L), their difference relative to the ground truth (M-V), and the RMSEs of χ throughout the brain volume (bottom row). Figure 3A shows the means and SDs of χ calculated using each processing pipeline in each ROI.

For the numerical phantom simulations, the error between the calculated and ground-truth χ appeared fairly similar for all processing pipelines, although slightly larger SDs were always observed for the NLFit pipeline (Figures 2M-V and Figure 3A). The SSS, which was the highest χ structure (χ^{True}=0.30 ± 0.06 ppm), showed the largest susceptibility errors for all the processing pipelines (arrowheads in Figures 2G-L, R-V and Figure 3A).

For one representative volunteer, Figures 4A-J show the susceptibility images calculated using each processing pipeline for multi-echo QSM. Additionally, Figures 4K-R show the differences between each QSM image relative to the χ^{NLFit} image. For each ROI in the healthy volunteers, Figure 3B shows the pooled means and SDs of χ calculated using each processing pipeline. In the volunteers, susceptibility differences between processing pipelines were most prominent in the StrS (arrowheads in Figures 4F-J, O-R).

In general, the average χ measured in the deep-GM ROIs and in the PCR (Figure 3B) had values within the ranges reported by previous studies (34,42-46): 0.01–0.13 ppm for the CN, 0.06–0.29 ppm for the GP, 0.02–0.14 ppm for the PU, -0.02–0.08 ppm for the TH, and -0.06–0.03 ppm for the PCR. In the StrS, only χ^{NLFit} had an average value close to the previously reported range for venous blood, namely, 0.17–0.58 ppm (32,45,47,48) (Figure 3B).

Similar to the numerical phantom simulations, in the healthy volunteers, χ calculated using the NLFit pipeline (χ^{NLFit}) had slightly larger SDs in the GP, PU and StrS (Figure 3B). In contrast with the numerical phantom simulations, in the healthy volunteers, χ calculated

using the SNR-wAvg pipeline ($\chi^{SNR-wAvg}$) had slightly larger SDs in the CN, TH and PCR compared to χ^{NLFit}.

Finally, in the numerical phantom simulations, the χ map calculated using the NLFit pipeline had the largest RMSE value (108.8%) followed, in decreasing order, by the χ maps calculated using the SNR-wAvg (94.7%), TE-wAvg (94.0%), Susc-wAvg (91.7%) and Susc-TGV-wAvg pipelines (79.6%) (Figure 2, bottom row).

Phase noise propagation into the χ maps

Figures 5 and 6 show the estimated $\sigma(\chi)$ profiles in the numerical phantom simulations and a representative healthy subject, respectively. In the phantom, NLFit and Susc-wAvg had the $\sigma(\chi)$ profiles with the lowest means and SDs (Figures 5A, D, E, H). In both the phantom and the healthy volunteers, the NLFit and SNR-wAvg pipelines resulted in $\sigma(\chi)$ images with the lowest number of streaking artifacts (Figures 5A, E, C, G and 6A, E, C, G). Although all four pipelines had a certain degree of streaking artifacts, especially in the healthy volunteer, these were considerably worse in the TE-wAvg and Susc-wAvg pipelines around high-χ venous structures.

Statistical analysis

In the healthy volunteers, for all processing pipelines and ROIs, the Shapiro-Wilk test always rejected the hypothesis of normally distributed paired differences of χ. Therefore, pairwise comparisons between pipelines were always evaluated using the nonparametric sign test. Significant differences between pipelines are highlighted in Figure 3B, whereas the between-pipeline biases are shown in Figure 7. Biases smaller than |0.01| ppm (|·| denotes the absolute value) were considered negligible.

A bias greater than |0.01| ppm was observed for the NLFit pipeline relative to all other pipelines in the CN, GP and StrS, for the Susc-TGV-wAvg pipeline relative to the TE-wAvg, SNR-wAvg and Susc-wAvg pipelines in the GP, and between the SNR-wAvg and Susc-wAvg pipelines in the GP (Figure 7).

In general, the bias was greater than |0.01| ppm for the same pairs of pipelines that differed according to the sign test (Figures 3B and 7).

Discussion

This study compared multi-echo combination methods for QSM, with the goal of elucidating the optimal combination method. The combination methods were applied at different stages of the QSM processing pipeline before or after LBMs for phase unwrapping or background field removal. Methods for multi-echo combination were compared by applying each pipeline to numerically simulated data and images from healthy volunteers.

The relative accuracy of χ from different processing pipelines *in vivo* was similar to that in phantom simulations in the CN, GP and the venous ROIs, where the NLFit pipeline provided the highest and most accurate ROI means of χ (Figure 3). Slightly different relative accuracies were observed in the PU, TH and PCR. Here, however, no significant difference or systematic bias was detected between processing pipelines (Figures 3 and 7). Two exceptions were observed for the Susc-TGV-wAvg pipeline in the GP and the StrS, as this pipeline had the second highest ROI mean of χ in the phantom simulations but the lowest ROI mean of χ in the healthy volunteers (Figure 3). Both the TGV reconstruction pipeline and shortcomings of the numerically simulated data could explain these discrepancies in the performance of the Susc-TGV-wAvg pipeline. The numerically simulated data were generated based on a digital phantom containing a piece-wise constant χ distribution. As previously observed (49), piece-wise constant geometries allow good recovery of the underlying χ distribution using TGV-based algorithms, because the piece-wise constant constraints exactly match the underlying χ distribution. However, in regions with flow, anisotropic χ distributions or microstructure, these numerical models are likely to depart from a realistic representation of the tissue χ. Notably, the GP, unlike the PU and CN, is traversed by multiple myelinated axons of the striato-pallido-nigral tract. In the GP in vivo, additional anisotropic χ contrast dependent on axonal orientation could explain the lower accuracy of the Susc-TGV-wAvg pipeline compared to the phantom simulations, in which the GP had isotropic χ (50). As for the StrS, the lower accuracy of the Susc-TGV-wAvg pipeline compared to the phantom simulations could be due to the presence of flow *in vivo* and the absence of flow compensation in the acquisition sequence, jointly violating the assumption of a piece-wise constant venous χ. To overcome the limitations of this assumption, future studies could exploit a newly developed realistic head phantom for QSM, which does not have a piecewise constant χ distribution and also incorporates microstructural effects (51). Generally, the higher accuracy of the NLFit pipeline suggests that, for QSM, combining the temporally unwrapped multi-echo phase before applying LBMs for spatial phase unwrapping or ΔB_{Bg} removal is preferable to averaging the TE-dependent LBM-processed phase or χ. This suggestion appears to conflict with the higher RMSE associated with the NLFit pipeline compared to other pipelines in the phantom simulations (Figure 2). However, the RMSE jointly reflects systematic and random errors, as it measures the bias between the estimated χ and χ^{True} but also reflects the variability of the estimated χ relative to its average value (49). Thus, the RMSE must always be interpreted in combination with complementary measurements of bias and precision.

There are several potential explanations as to why LBMs applied before multi-echo combination reduce the overall accuracy of QSM. Firstly, in contrast with path-based or region-growing-based phase unwrapping, Laplacian phase unwrapping usually removes some ΔB_{Bg} components from the input phase image (28). Thus, the consistently lower accuracy of χ calculated using the TE-wAvg processing pipeline was probably driven by the incorrect assumption that the TE-dependent LBM-unwrapped phase corresponded to the true unwrapped phase. LBMs for ΔB_{Bg} removal also applied tSVD (with a larger truncation threshold) but here the high pass filtering effect was expected, as background fields are slowly varying. Finally, the similar accuracy and values of χ calculated using the TE-wAvg, SNR-wAvg and Susc-wAvg pipelines suggest a negligible difference between averaging the Laplacian unwrapped phase over TEs before (TE-wAvg pipeline) or after TE-dependent Laplacian ΔB_{Bg} removal (SNR-wAvg and Susc-wAvg pipelines).

In the numerical phantom simulations, all processing pipelines resulted in higher SDs of χ compared to the ground-truth, suggesting that they all intrinsically amplified the noise in the multi-echo signal phase. This result is in line with the known noise amplification of ill-posed inverse problems. However, the estimated SDs of χ varied across pipelines, with the Susc-TGV-wAvg pipeline generally having the lowest SDs and $\sigma(\chi)$ in both the numerical simulations and healthy volunteers (Figure 3 and Figures 5D, H and 6D, H). This result was in agreement with a previous study showing that, for QSM, L_1-regularization (as in TGV) provides better noise suppression than L_2-regularization (e.g., Tikhonov regularization) (9). In both numerical simulations and healthy volunteers, based on the line profiles traced on the noise maps and streaking artifact reduction, the NLFit pipeline had better noise mitigation than the TE-wAvg, SNR-wAvg and Susc-wAvg pipelines (Figures 5 and 6). This result suggests that combining the temporally unwrapped multi-echo phase by nonlinear complex fitting, designed to account for noise in the complex signal (8), results in better noise management than combining the multi-echo phase by averaging. As previously shown (8), errors in the combined field map mainly result from both noise in the signal phase and phase unwrapping errors near high-χ regions (e.g., the veins). Both sources of error were successfully managed by nonlinear complex fitting, as shown by the dramatic reduction of streaking artifacts in Figure 6A. In line with previous observations (30), this result also suggests that the regularization strategy employed by the ΔB_{Loc}-to χ step can mitigate artifactual streaking errors only when major sources of error in the field map have been tightly constrained. In both the numerical simulations and healthy volunteers, the SNR-wAvg pipeline had the best noise mitigation of all averaging-based pipelines (Figures 5 and 6).

Thus, if applying nonlinear complex fitting is not possible, averaging using the SNR-weighting method could offer the best alternative for noise reduction.

All these indications do not necessarily apply to χ estimation in WM tissue. Indeed, due to WM's ordered microstructure, a comprehensive estimation of χ in WM requires acquiring GRE images at multiple head orientations and modelling χ as a tensor (52,53). Thus, further studies are needed to evaluate the applicability of these results to WM tissue.

In the present study, all experiments were limited to one field strength (i.e., 3 T). As tissue relaxation times (e.g., T_2^*) shorten with increasing field strength but the signal phase at a given TE increases, further work is needed to assess the relevance of these results at ultrahigh fields. Finally, it must be noted that in the numerical phantom simulations, all processing pipelines underestimated χ^{True} (Figure 3). However, in QSM, some degree of underestimation is always expected, because of the ill-posed nature of the ΔB_{Loc}-to-χ inverse problem (19).

Conclusions

The higher accuracy of regional χ values and better noise management of the NLFit pipeline suggest that, for QSM, combining the multi-echo phase by nonlinearly fitting over TEs before applying LBMs is preferable to combining the TE-dependent LBM-processed phase or χ by averaging.

Acknowledgments

We thank all the healthy volunteers who participated in this study. We also thank Dr. Rosa Cortese and Dr. Floriana De Angelis (Queen Square MS Centre, UCL Queen Square Institute of Neurology, UCL, London, UK) for their help with the MRI scans. We are indebted to Prof. Claudia Gandini Wheeler-Kingshott (NMR Unit, Queen Square MS Centre, UCL Queen Square Institute of Neurology, UCL, London, UK) for her support with the MRI scans performed at the Queen Square MS Centre, UCL.

References

1. Biondetti E, Karsa A, Thomas DL, Shmueli K. Investigating the accuracy and precision of TE-dependent versus multi-echo QSM using Laplacian-based methods at 3 T. Magn Reson Med 2020.
2. Liu T, Surapaneni K, Lou M, Cheng L, Spincemaille P, Wang Y. Cerebral microbleeds: burden assessment by using quantitative susceptibility mapping. Radiology 2012;262(1):269-278.
3. Haacke EM, Liu S, Buch S, Zheng W, Wu D, Ye Y. Quantitative susceptibility mapping: current status and future directions. Magn Reson Imaging 2015;33(1):1-25.

4. Wei H, Dibb R, Zhou Y, Sun Y, Xu J, Wang N, Liu C. Streaking artifact reduction for quantitative susceptibility mapping of sources with large dynamic range. NMR Biomed 2015;28(10):1294-1303.
5. Li W, Wang N, Yu F, Han H, Cao W, Romero R, Tantiwongkosi B, Duong TQ, Liu C. A method for estimating and removing streaking artifacts in quantitative susceptibility mapping. Neuroimage 2015;108:111-122.
6. Wu B, Li W, Avram AV, Gho SM, Liu C. Fast and tissue-optimized mapping of magnetic susceptibility and T2* with multi-echo and multi-shot spirals. Neuroimage 2012;59(1):297-305.
7. Chen Y, Liu S, Wang Y, Kang Y, Haacke EM. STrategically Acquired Gradient Echo (STAGE) imaging, part I: Creating enhanced T1 contrast and standardized susceptibility weighted imaging and quantitative susceptibility mapping. Magn Reson Imaging 2018;46:130-139.
8. Liu T, Wisnieff C, Lou M, Chen W, Spincemaille P, Wang Y. Nonlinear formulation of the magnetic field to source relationship for robust quantitative susceptibility mapping. Magn Reson Med 2013;69(2):467-476.
9. Kressler B, de Rochefort L, Liu T, Spincemaille P, Jiang Q, Wang Y. Nonlinear regularization for per voxel estimation of magnetic susceptibility distributions from MRI field maps. IEEE Trans Med Imaging 2010;29(2):273-281.
10. de Rochefort L, Brown R, Prince MR, Wang Y. Quantitative MR susceptibility mapping using piece-wise constant regularized inversion of the magnetic field. Magn Reson Med 2008;60(4):1003-1009.
11. Karsa A, Punwani S, Shmueli K. An optimized and highly repeatable MRI acquisition and processing pipeline for quantitative susceptibility mapping in the head-and-neck region. Magn Reson Med 2020;84(6):3206-3222.
12. Robinson SD, Bredies K, Khabipova D, Dymerska B, Marques JP, Schweser F. An illustrated comparison of processing methods for MR phase imaging and QSM: combining array coil signals and phase unwrapping. NMR Biomed 2017;30(4).
13. Schofield MA, Zhu Y. Fast phase unwrapping algorithm for interferometric applications. Opt Lett 2003;28(14):1194-1196.
14. Schweser F, Robinson SD, de Rochefort L, Li W, Bredies K. An illustrated comparison of processing methods for phase MRI and QSM: removal of background field contributions from sources outside the region of interest. NMR Biomed 2017;30(4).
15. Qiu D, Brown RC, Sun B, Palasis S, Burns TG, Jones RA. Abnormal Iron Levels in the Brain of Pediatric Sickle Cell Disease Patients: a Study using Quantitative

Susceptibility Mapping (QSM). Proceedings of the International Society for Magnetic Resonance in Medicine 2014;22:897.

16. Langkammer C, Bredies K, Poser BA, Barth M, Reishofer G, Fan AP, Bilgic B, Fazekas F, Mainero C, Ropele S. Fast quantitative susceptibility mapping using 3D EPI and total generalized variation. Neuroimage 2015;111:622-630.
17. Conturo TE, Smith GD. Signal-to-noise in phase angle reconstruction: Dynamic range extension using phase reference offsets. Magnetic Resonance in Medicine 1990;15(3):420-437.
18. Liu T, Khalidov I, de Rochefort L, Spincemaille P, Liu J, Tsiouris AJ, Wang Y. A novel background field removal method for MRI using projection onto dipole fields (PDF). NMR Biomed 2011;24(9):1129-1136.
19. Wang Y, Liu T. Quantitative susceptibility mapping (QSM): Decoding MRI data for a tissue magnetic biomarker. Magn Reson Med 2015;73(1):82-101.
20. Shmueli K, de Zwart JA, van Gelderen P, Li TQ, Dodd SJ, Duyn JH. Magnetic susceptibility mapping of brain tissue in vivo using MRI phase data. Magn Reson Med 2009;62(6):1510-1522.
21. Biondetti E, Thomas DL, Shmueli K. Application of Laplacian-based Methods to Multi-echo Phase Data for Accurate Susceptibility Mapping. Proceedings of the International Society for Magnetic Resonance in Medicine 2016;24:1547.
22. Pruessmann KP, Weiger M, Scheidegger MB, Boesiger P. SENSE: Sensitivity encoding for fast MRI. Magnetic Resonance in Medicine 1999;42(5):952-962.
23. Berg RC, Preibisch C, Thomas DL, Shmueli K, Biondetti E. Investigating the effect of flow compensation and quantitative susceptibility mapping method on the accuracy of venous susceptibility measurement. Neuroimage 2021;240:118399.
24. Zubal IG, Harrell CR, Smith EO, Smith AL, Krischlunas P. Two dedicated software, voxel-based, anthropomorphic (torso and head) phantoms. Proceeding of the International Conference at the National Radiological Protection Board1995. p 105-111.
25. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. Fsl. Neuroimage 2012;62(2):782-790.
26. Smith SM. Fast robust automated brain extraction. Hum Brain Mapp 2002;17(3):143-155.
27. Cornell MRI Research Group. http://pre.weill.cornell.edu/mri/pages/qsm.html. Updated January 15, 2020. Accessed January 21, 2022.

28. Schweser F, Deistung A, Sommer K, Reichenbach JR. Toward online reconstruction of quantitative susceptibility maps: superfast dipole inversion. Magn Reson Med 2013;69(6):1582-1594.
29. Wu B, Li W, Guidon A, Liu C. Whole brain susceptibility mapping using compressed sensing. Magn Reson Med 2012;67(1):137-147.
30. Acosta-Cabronero J, Milovic C, Mattern H, Tejos C, Speck O, Callaghan MF. A robust multi-scale approach to quantitative susceptibility mapping. Neuroimage 2018;183:7-24.
31. Hansen PC, O'Leary DP. The Use of the L-Curve in the Regularization of Discrete Ill-Posed Problems. SIAM Journal on Scientific Computing 1993;14(6):1487-1503.
32. Biondetti E, Rojas-Villabona A, Sokolska M, Pizzini FB, Jager HR, Thomas DL, Shmueli K. Investigating the oxygenation of brain arteriovenous malformations using quantitative susceptibility mapping. Neuroimage 2019;199:440-453.
33. Centre for Advanced Imaging software repository. https://github.com/CAIsr/qsm. Updated March 17, 2021. Accessed January 25, 2022.
34. Lim IA, Faria AV, Li X, Hsu JT, Airan RD, Mori S, van Zijl PC. Human brain atlas for automated region of interest selection in quantitative susceptibility mapping: application to determine iron content in deep gray matter structures. Neuroimage 2013;82:449-469.
35. Modat M, Ridgway GR, Taylor ZA, Lehmann M, Barnes J, Hawkes DJ, Fox NC, Ourselin S. Fast free-form deformation using graphics processing units. Comput Methods Programs Biomed 2010;98(3):278-284.
36. Ourselin S, Roche A, Subsol G, Pennec X, Ayache N. Reconstructing a 3D structure from serial histological sections. Image and Vision Computing 2001;19(1-2):25-31.
37. Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, Gerig G. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage 2006;31(3):1116-1128.
38. Lerski RA, de Wilde J, Boyce D, Ridgway J. Quality control in magnetic resonance imaging. York, UK: The Institute of Physics and Engineering in Medicine; 1998.
39. Dietrich O, Raya JG, Reeder SB, Reiser MF, Schoenberg SO. Measurement of signal-to-noise ratios in MR images: influence of multichannel coils, parallel imaging, and reconstruction filters. J Magn Reson Imaging 2007;26(2):375-385.
40. Weiger M, Pruessmann KP, Boesiger P. 2D SENSE for faster 3D MRI. MAGMA 2002;14(1):10-19.
41. Gudbjartsson H, Patz S. The Rician distribution of noisy MRI data. Magn Reson Med 1995;34(6):910-914.

42. Li W, Wu B, Liu C. Quantitative susceptibility mapping of human brain reflects spatial variation in tissue composition. Neuroimage 2011;55(4):1645-1656.
43. Wharton S, Bowtell R. Whole-brain susceptibility mapping at high field: a comparison of multiple- and single-orientation methods. Neuroimage 2010;53(2):515-525.
44. Bilgic B, Pfefferbaum A, Rohlfing T, Sullivan EV, Adalsteinsson E. MRI estimates of brain iron concentration in normal aging using quantitative susceptibility mapping. Neuroimage 2012;59(3):2625-2635.
45. Deistung A, Schäfer A, Schweser F, Biedermann U, Turner R, Reichenbach JR. Toward in vivo histology: A comparison of quantitative susceptibility mapping (QSM) with magnitude-, phase-, and R2*-imaging at ultra-high magnetic field strength. NeuroImage 2013;65:299-314.
46. Schweser F, Deistung A, Lehr BW, Reichenbach JR. Quantitative imaging of intrinsic magnetic tissue properties using MRI signal phase: an approach to in vivo brain iron metabolism? Neuroimage 2011;54(4):2789-2807.
47. Fan AP, Bilgic B, Gagnon L, Witzel T, Bhat H, Rosen BR, Adalsteinsson E. Quantitative oxygenation venography from MRI phase. Magn Reson Med 2014;72(1):149-159.
48. Cetin S, Bilgic B, Fan A, Holdsworth S, Unal G. Vessel Orientation Constrained Quantitative Susceptibility Mapping (QSM) Reconstruction. Ourselin S, Joskowicz L, Sabuncu MR, Unal G, Wells W, editors. Cham: Springer International Publishing; 2016.
49. Langkammer C, Schweser F, Shmueli K, Kames C, Li X, Guo L, Milovic C, Kim J, Wei H, Bredies K, Buch S, Guo Y, Liu Z, Meineke J, Rauscher A, Marques JP, Bilgic B. Quantitative susceptibility mapping: Report from the 2016 reconstruction challenge. Magn Reson Med 2018;79(3):1661-1673.
50. Duyn JH, Schenck J. Contributions to magnetic susceptibility of brain tissue. NMR Biomed 2017;30(4).
51. Marques JP, Meineke J, Milovic C, Bilgic B, Chan KS, Hedouin R, van der Zwaag W, Langkammer C, Schweser F. QSM reconstruction challenge 2.0: A realistic in silico head phantom for MRI data simulation and evaluation of susceptibility mapping procedures. Magn Reson Med 2021;86(1):526-542.
52. Li W, Liu C, Duong TQ, van Zijl PC, Li X. Susceptibility tensor imaging (STI) of the brain. NMR Biomed 2017;30(4).
53. Chen WC, Foxley S, Miller KL. Detecting microstructural properties of white matter based on compartmentalization of magnetic susceptibility. Neuroimage 2013;70:1-9.

Figure captions

Figure 1. Processing pipelines for multi-echo QSM. For each multi-echo combination method (NLFit [nonlinear phase fitting], TE-wAvg [TE-weighted phase averaging], SNR-wAvg [SNR-weighted phase averaging], Susc-wAvg and Susc-TGV-wAvg [magnitude-weighted susceptibility averaging]) the processing steps are described as separate processing streams in the grey box.

Figure 2. χ maps calculated using distinct multi-echo combination methods in the numerical phantom simulations. The same transverse and sagittal slices are shown for the ground-truth susceptibility map (A, G), and for the susceptibility maps calculated using NLFit (B, H), TE-wAvg (C, I), SNR-wAvg (D, J), Susc-wAvg (E, K) and Susc-TGV-wAvg (F, L). The figure also shows the difference between each susceptibility map and the ground truth (M-V). The bottom row shows the root mean squared errors (RMSEs) of χ for each pipeline. In all the sagittal images (G-L, R-V), the yellow and blue arrowhead respectively point at the same posterior and anterior locations in the superior sagittal sinus.

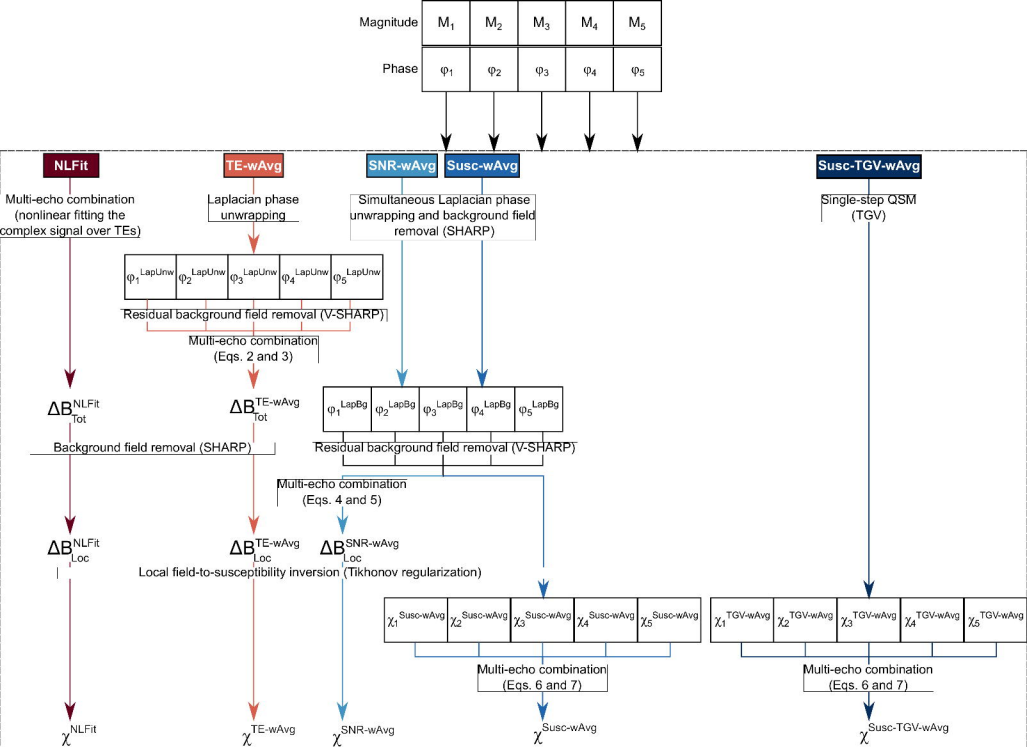
Figure 3. Means and SDs of χ in the phantom and healthy volunteer ROIs. The means and SDs (error bars) of χ are shown in each ROI of the numerically simulated (A) and pooled healthy volunteer data (B) for each processing pipeline. In the numerical phantom the ground-truth χ is also shown. In the healthy volunteers, significant differences between pairs of pipelines are denoted using the symbol ** (P-value < 0.05). Abbreviations: CN, caudate nucleus; GP, globus pallidus; PCR, posterior corona radiata; PU, putamen; StrS, straight sinus; TH, thalamus; WM, white matter; SSS, superior sagittal sinus.

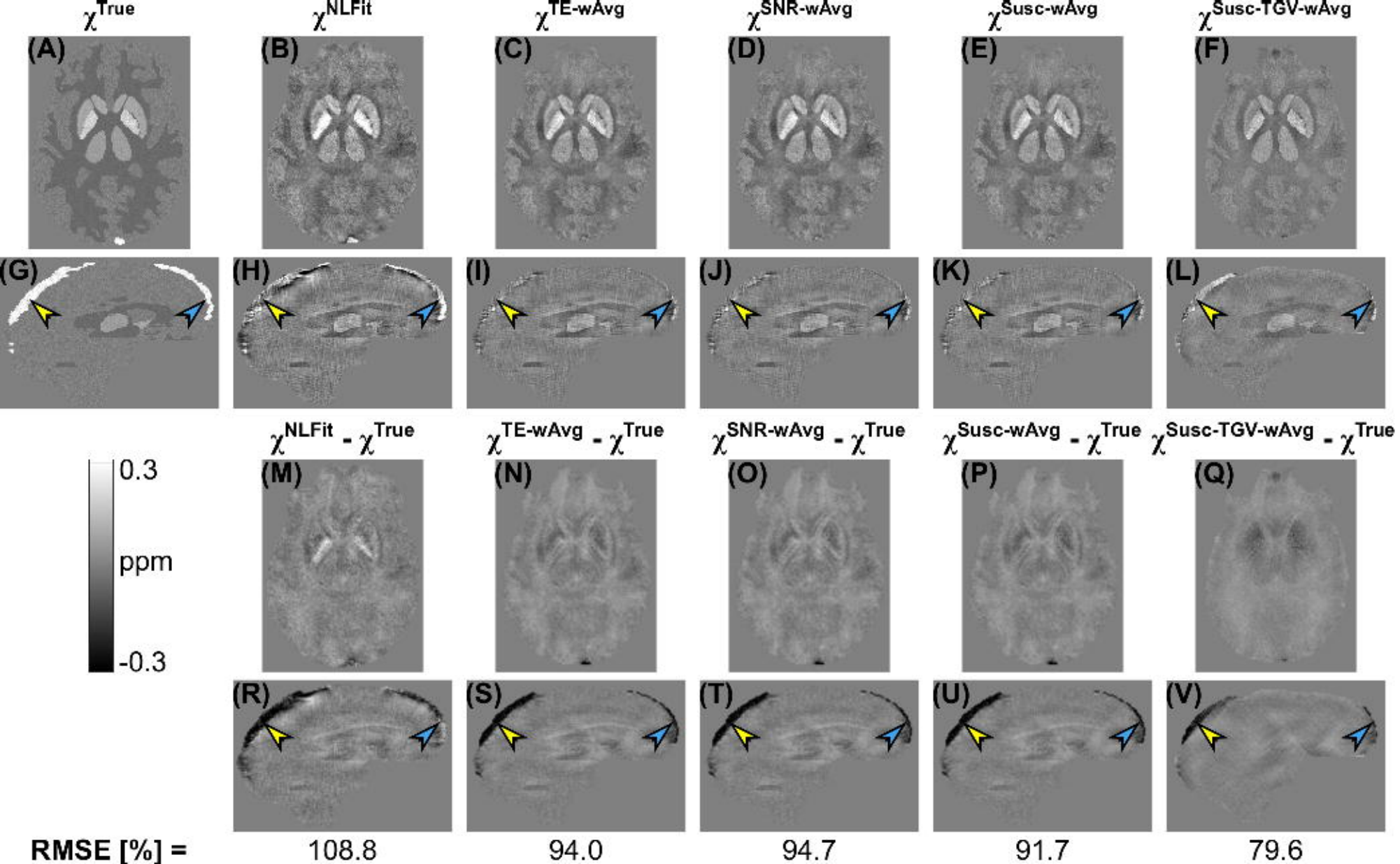
Figure 4. χ maps calculated using distinct multi-echo combination methods in a representative healthy volunteer. The same transverse and sagittal slices are shown for the susceptibility maps calculated using NLFit (A, F) [nonlinear phase fitting], TE-wAvg (B, G) [TE-weighted phase averaging], SNR-wAvg (C, H) [SNR-weighted phase averaging], Susc-wAvg (D, I) and Susc-TGV-wAvg (E, J) [magnitude-weighted susceptibility averaging]. The figure also shows the differences between the TE-wAvg, SNR-wAvg, Susc-wAvg and Susc-TGV-wAvg maps and the NLFit map (K-R). In all the sagittal images (F-J, O-R), the yellow arrowheads point at the same location in the StrS.

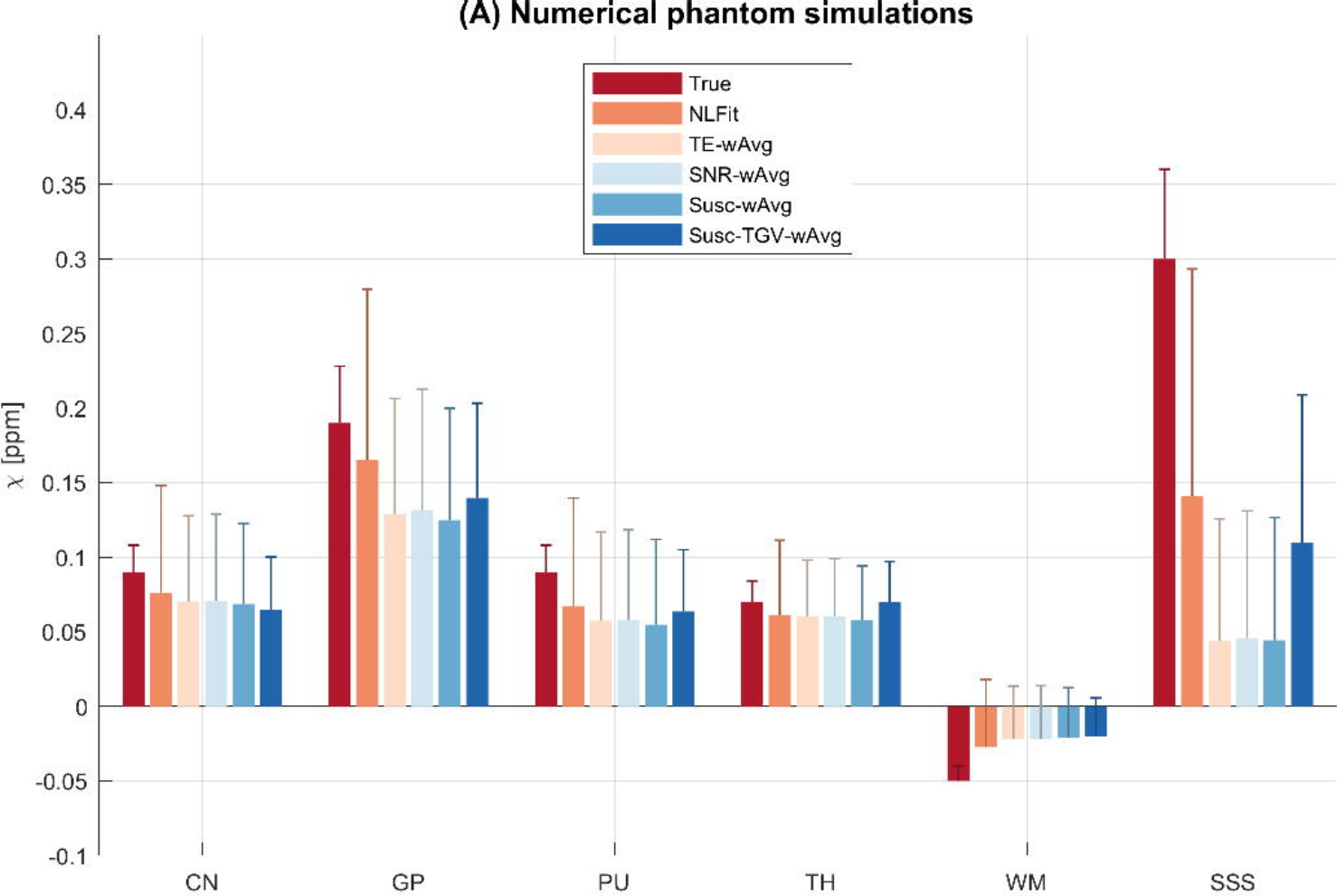
Figure 5. Susceptibility noise profiles in the numerical phantom simulations. The same sagittal slice is shown for susceptibility noise ($\sigma(\chi)$) images calculated using the NLFit, TE-wAvg, SNR-wAvg and Susc-wAvg pipelines (A-D). The susceptibility noise is plotted for a line profile traced on the $\sigma(\chi)$ images (E-H) including the high-χ SSS. The mean and the standard deviation of $\sigma(\chi)$ are shown for each line profile.

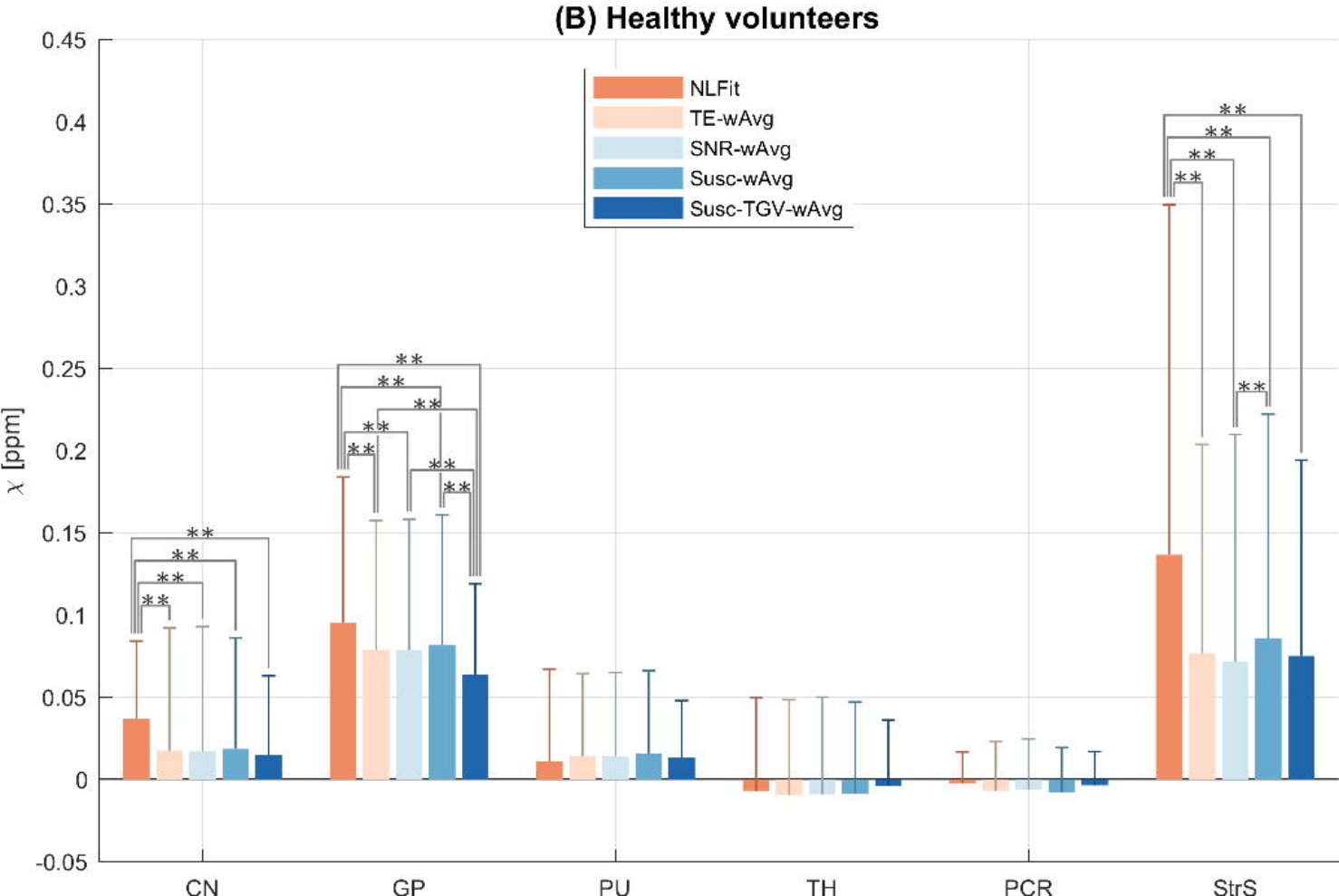
Figure 6. Susceptibility noise profiles in a representative healthy volunteer. The same sagittal slice is shown for susceptibility noise ($\sigma(\chi)$) images calculated using the NLFit, TE-wAvg, SNR-wAvg and Susc-wAvg pipelines (A-D). The susceptibility noise is plotted for a line profile traced on the $\sigma(\chi)$ images (E-H) including the high-χ StrS. The mean and the standard deviation of $\sigma(\chi)$ are shown for each line profile.

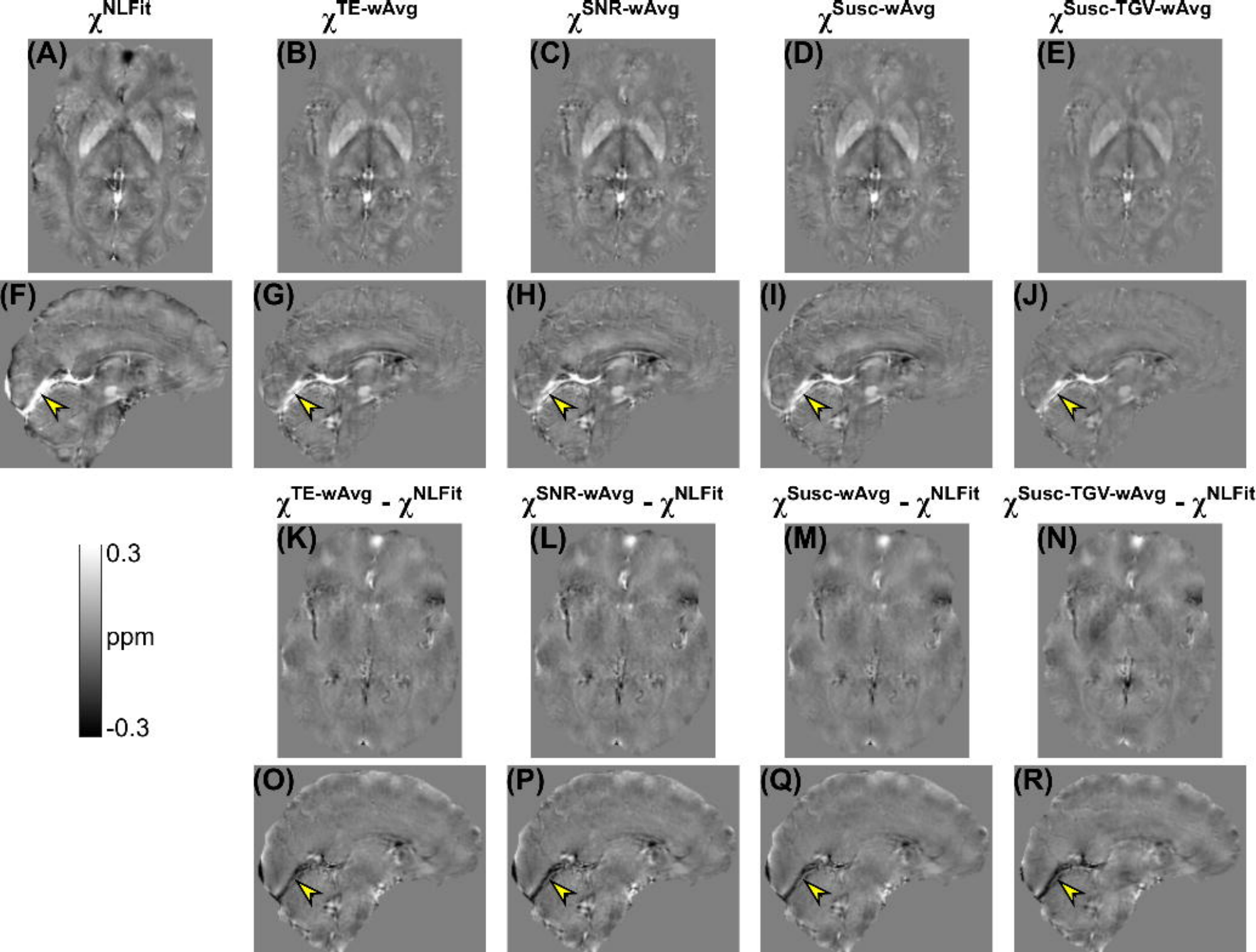
Figure 7. Bias between multi-echo pipelines for QSM in each ROI. The mean and SDs (error bars) of the bias are shown in each healthy volunteer ROI for all pairs of multi-echo processing pipelines. The grey band denotes the [-0.01 – 0.01] ppm interval. If the mean of the bias was within this interval, the difference between the corresponding pair of QSM pipelines was considered negligible.

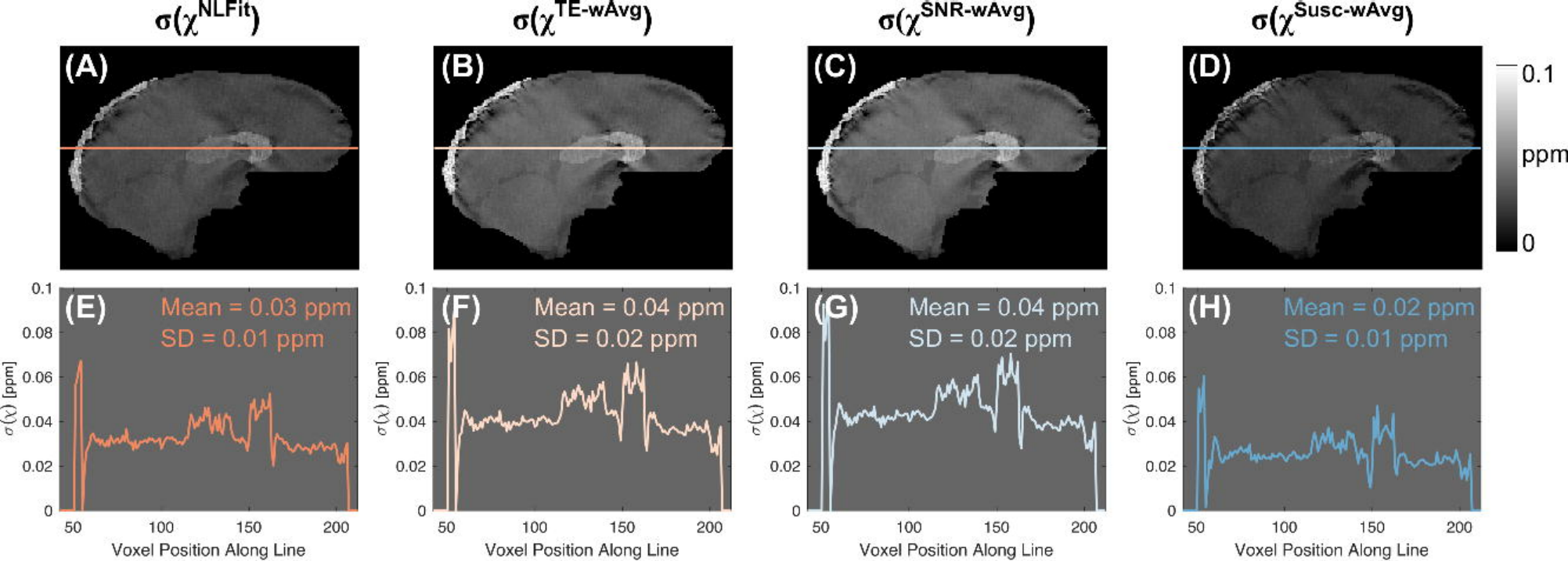

Magnitude
M_1 M_2 M_3 M_4 M_5
Phase
φ_1 φ_2 φ_3 φ_4 φ_5
NLFit
TE-wAvg
SNR-wAvg
Susc-wAvg
Susc-TGV-wAvg
Multi-echo combination (nonlinear fitting the complex signal over TEs)
Laplacian phase unwrapping
Simultaneous Laplacian phase unwrapping and background field removal (SHARP)
Single-step QSM (TGV)
φ_1^{LapUnw} φ_2^{LapUnw} φ_3^{LapUnw} φ_4^{LapUnw} φ_5^{LapUnw}
Residual background field removal (V-SHARP)
Multi-echo combination (Eqs. 2 and 3)
ΔB_{Tot}^{NLFit}
$\Delta B_{Tot}^{TE\text{-}wAvg}$
φ_1^{LapBg} φ_2^{LapBg} φ_3^{LapBg} φ_4^{LapBg} φ_5^{LapBg}
Background field removal (SHARP)
Residual background field removal (V-SHARP)
Multi-echo combination (Eqs. 4 and 5)
ΔB_{Loc}^{NLFit}
$\Delta B_{Loc}^{TE\text{-}wAvg}$
$\Delta B_{Loc}^{SNR\text{-}wAvg}$
Local field-to-susceptibility inversion (Tikhonov regularization)
$\chi_1^{Susc\text{-}wAvg}$ $\chi_2^{Susc\text{-}wAvg}$ $\chi_3^{Susc\text{-}wAvg}$ $\chi_4^{Susc\text{-}wAvg}$ $\chi_5^{Susc\text{-}wAvg}$
$\chi_1^{TGV\text{-}wAvg}$ $\chi_2^{TGV\text{-}wAvg}$ $\chi_3^{TGV\text{-}wAvg}$ $\chi_4^{TGV\text{-}wAvg}$ $\chi_5^{TGV\text{-}wAvg}$
Multi-echo combination (Eqs. 6 and 7)
Multi-echo combination (Eqs. 6 and 7)
χ^{NLFit}
$\chi^{TE\text{-}wAvg}$
$\chi^{SNR\text{-}wAvg}$
$\chi^{Susc\text{-}wAvg}$
$\chi^{Susc\text{-}TGV\text{-}wAvg}$

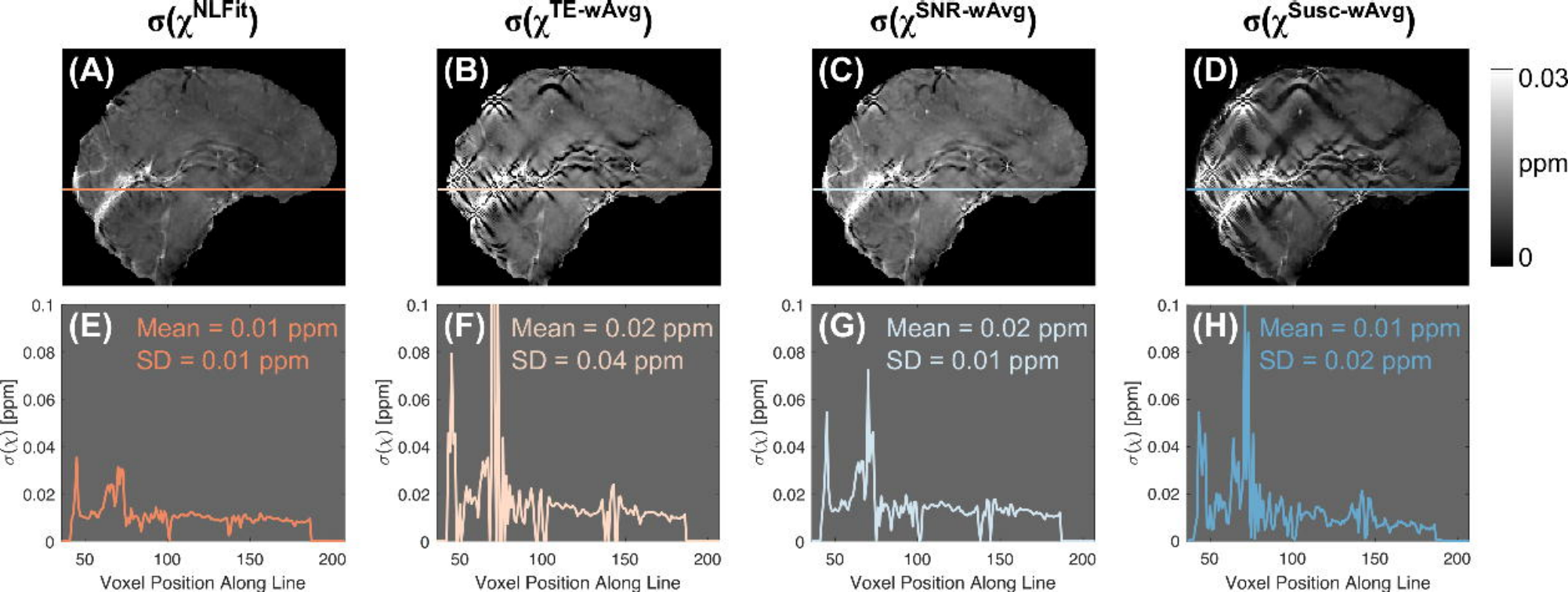

χTrue
χNLFit
χTE-wAvg
χSNR-wAvg
χSusc-wAvg
χSusc-TGV-wAvg
(A)
(B)
(C)
(D)
(E)
(F)
(G)
(H)
(I)
(J)
(K)
(L)
χNLFit - χTrue
χTE-wAvg - χTrue
χSNR-wAvg - χTrue
χSusc-wAvg - χTrue
χSusc-TGV-wAvg - χTrue
0.3
ppm
-0.3
(M)
(N)
(O)
(P)
(Q)
(R)
(S)
(T)
(U)
(V)
RMSE [%] =
108.8
94.0
94.7
91.7
79.6

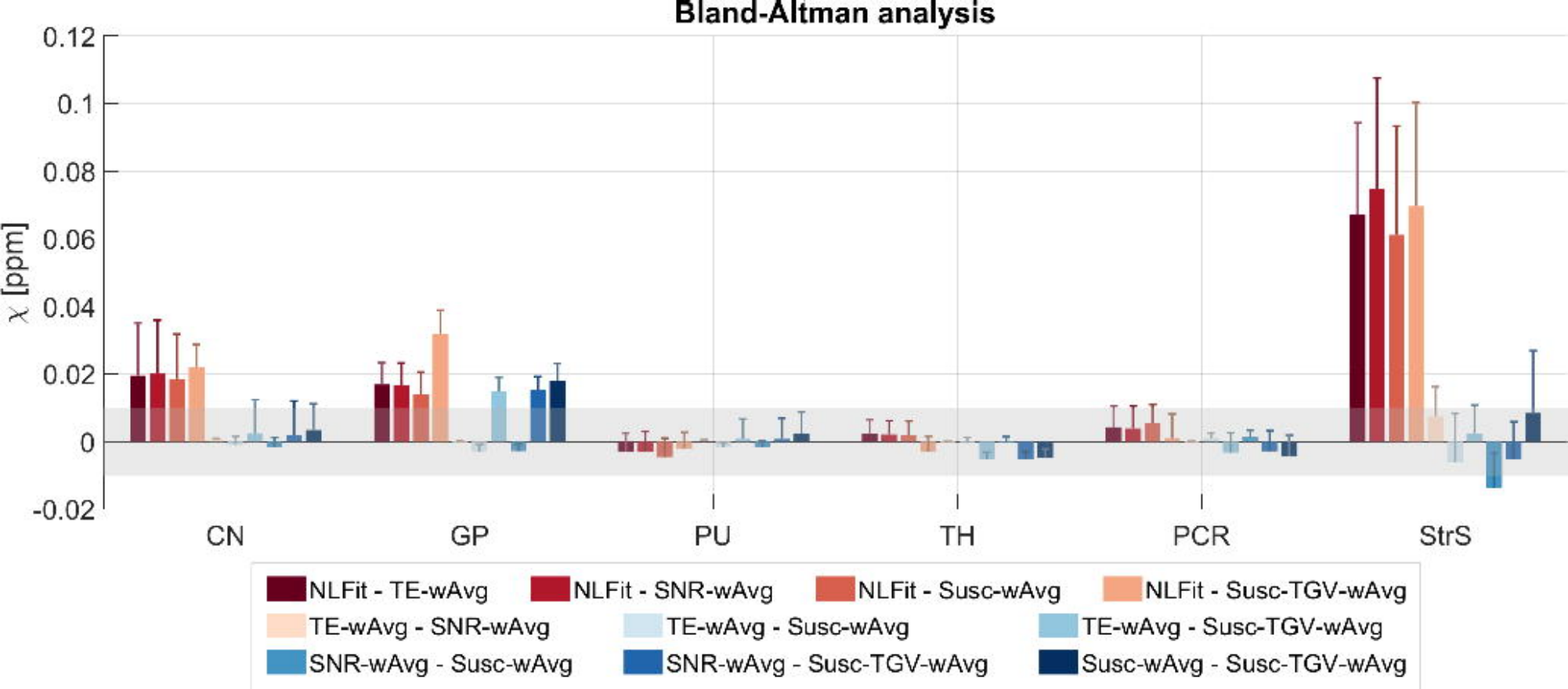

(A) Numerical phantom simulations
True
NLFit
TE-wAvg
SNR-wAvg
Susc-wAvg
Susc-TGV-wAvg
χ [ppm]
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-0.05
-0.1
CN
GP
PU
TH
WM
SSS


(B) Healthy volunteers
NLFit
TE-wAvg
SNR-wAvg
Susc-wAvg
Susc-TGV-wAvg
χ [ppm]
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-0.05
**
CN
GP
PU
TH
PCR
StrS


χ^{NLFit}
$\chi^{TE\text{-}wAvg}$
$\chi^{SNR\text{-}wAvg}$
$\chi^{Susc\text{-}wAvg}$
$\chi^{Susc\text{-}TGV\text{-}wAvg}$
(A)
(B)
(C)
(D)
(E)
(F)
(G)
(H)
(I)
(J)
$\chi^{TE\text{-}wAvg} - \chi^{NLFit}$
$\chi^{SNR\text{-}wAvg} - \chi^{NLFit}$
$\chi^{Susc\text{-}wAvg} - \chi^{NLFit}$
$\chi^{Susc\text{-}TGV\text{-}wAvg} - \chi^{NLFit}$
0.3
ppm
-0.3
(K)
(L)
(M)
(N)
(O)
(P)
(Q)
(R)


σ(χ^NLFit)
σ(χ^TE-wAvg)
σ(χ^SNR-wAvg)
σ(χ^Susc-wAvg)
(A)
(B)
(C)
(D)
0.1
ppm
0
(E)
Mean = 0.03 ppm
SD = 0.01 ppm
(F)
Mean = 0.04 ppm
SD = 0.02 ppm
(G)
Mean = 0.04 ppm
SD = 0.02 ppm
(H)
Mean = 0.02 ppm
SD = 0.01 ppm
σ(χ) [ppm]
0.1
0.08
0.06
0.04
0.02
0
50
100
150
200
Voxel Position Along Line


σ(χ^NLFit)
σ(χ^TE-wAvg)
σ(χ^SNR-wAvg)
σ(χ^Susc-wAvg)
(A)
(B)
(C)
(D)
0.03
ppm
0
(E)
Mean = 0.01 ppm
SD = 0.01 ppm
(F)
Mean = 0.02 ppm
SD = 0.04 ppm
(G)
Mean = 0.02 ppm
SD = 0.01 ppm
(H)
Mean = 0.01 ppm
SD = 0.02 ppm
σ(χ) [ppm]
0.1
0.08
0.06
0.04
0.02
0
50
100
150
200
Voxel Position Along Line


Bland-Altman analysis
0.12
0.1
0.08
0.06
0.04
0.02
0
-0.02
χ [ppm]
CN
GP
PU
TH
PCR
StrS
NLFit - TE-wAvg
NLFit - SNR-wAvg
NLFit - Susc-wAvg
NLFit - Susc-TGV-wAvg
TE-wAvg - SNR-wAvg
TE-wAvg - Susc-wAvg
TE-wAvg - Susc-TGV-wAvg
SNR-wAvg - Susc-wAvg
SNR-wAvg - Susc-TGV-wAvg
Susc-wAvg - Susc-TGV-wAvg