Appendix

“All models are wrong, some are useful” by George E. Box

Q1.1.1 Figures

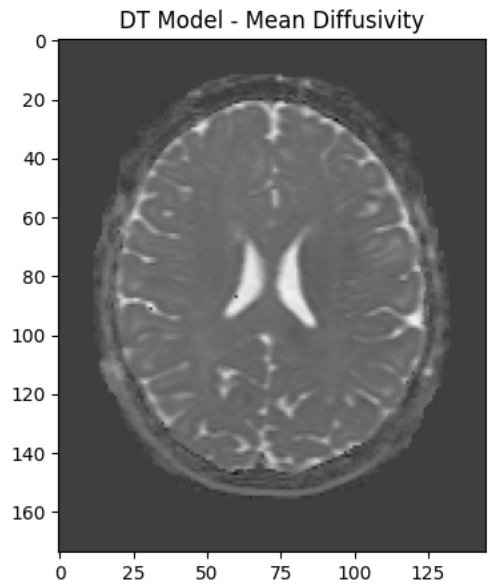

Figure 1: Mean diffusivity (MD) map computed using the DT model for slice 72. MD yields insights into water dispersion across brain white matter, where low MD could indicate water obstruction owed to structural damage owed to a brain tumor. With one patient’s scan, not much can be said. But if compared with other patients’ MDs, we can identify which brain cell barriers have been destroyed and thus water can’t move.

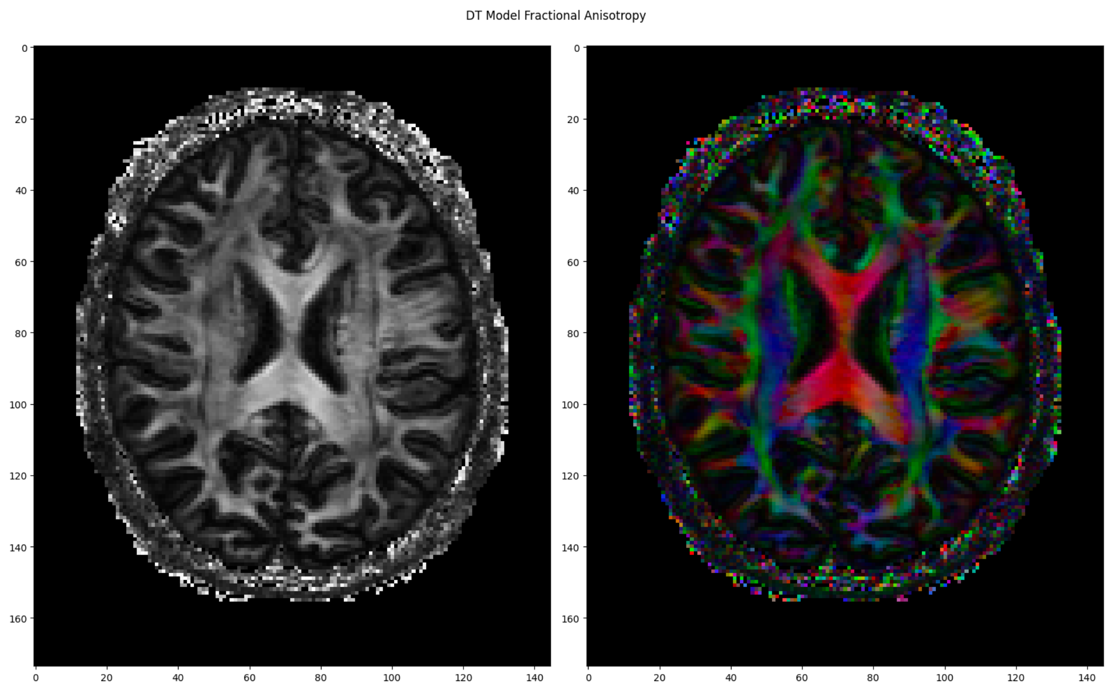

Figure 2: On the left we plot the FA map on the same brain slice 72 using the DT model. FA is like the variance of the MD, and it measures how elongated each diffusion tensor. FA highlights structures like the corpus callosum where there is oblate/prolate diffusion. The right corresponds to the FA map color-coded with the orientation of the eigenvalues decomposed from the diffusion tensor. Red corresponds to the primary eigenvector indicating horizontal diffusion (x-axis), Green to diffusion along the vertical y axis, and Blue corresponds to diffusion along the z axis.

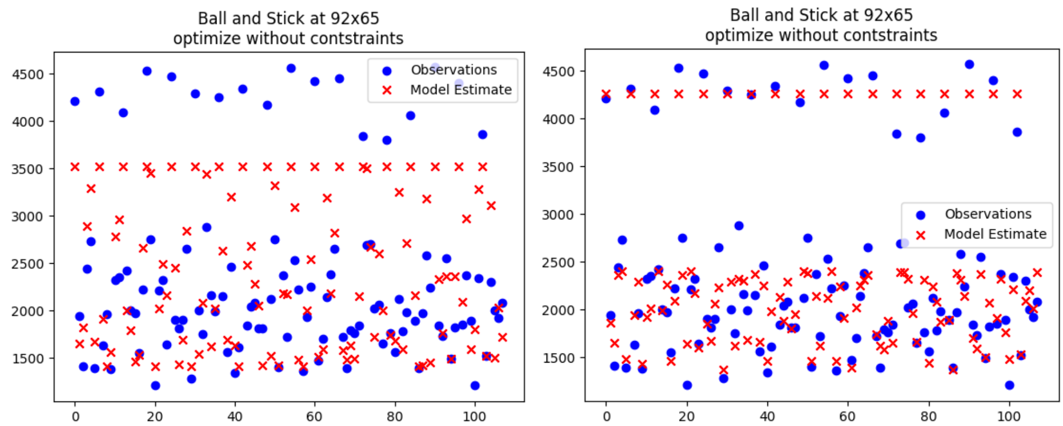

Figure 3: Ball-stick model fitted on the same voxel without any constraints during optimization using 2 different starting points: the left panel corresponds to the one given in the specifications (`x = [3.3e3,0.001,0.45,1,1]`), which results in a poor fit. The right panel corresponds to x2=`[4200, 4e-4, 0.25, 0, 0]` given in the lectures, which fits the data better at a RESNORM of `5.9e6`. In both cases, however, we get biophysically implausible parameter estimates.

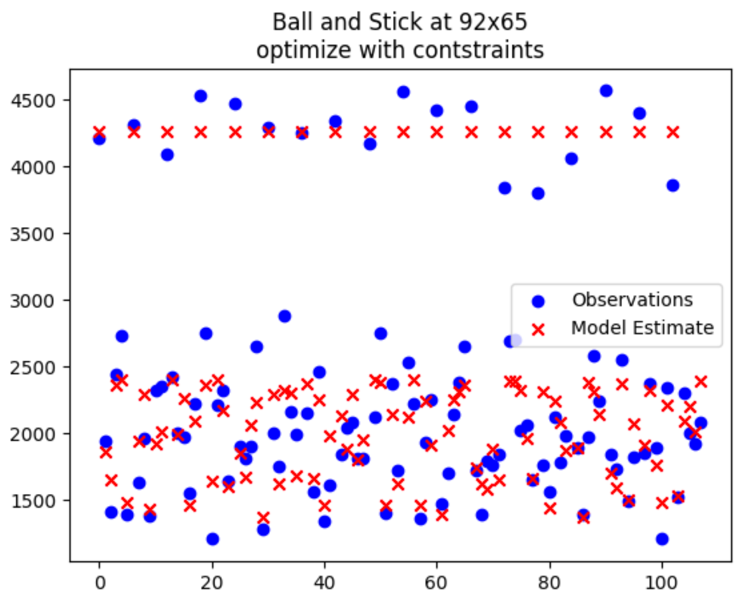

Figure 4: Ball-stick model fitted on the same voxel with constraints during optimization

using `x2=[4200, 4e-4, 0.25, 1, 1]` as a starting point. The only difference to the ones shared on the lectures is that θ, ϕ are set to 1 instead of 0. Experiments on our side show this prevents errors during optimization in further questions. We get a RESNORM of `5.9e6` whilst obtaining biophysically plausible parameters `x = [4257,0.00114,0.36,2.2,0.58]`. We likely hit a global minimum because we converged to the RESNORM as before.

Voxel location	Ratio of runs finding the global minimum	Number of runs required
92, 65	0.79	2
71, 70	0.98	1
41, 81	0.71	3
81, 61	0.86	2
86, 56	0.3	9
86, 73	0.33	8

Table 1: Results for calculating the likelihood of reaching the global minimum computed across 6 different voxels on the ball-stick model.

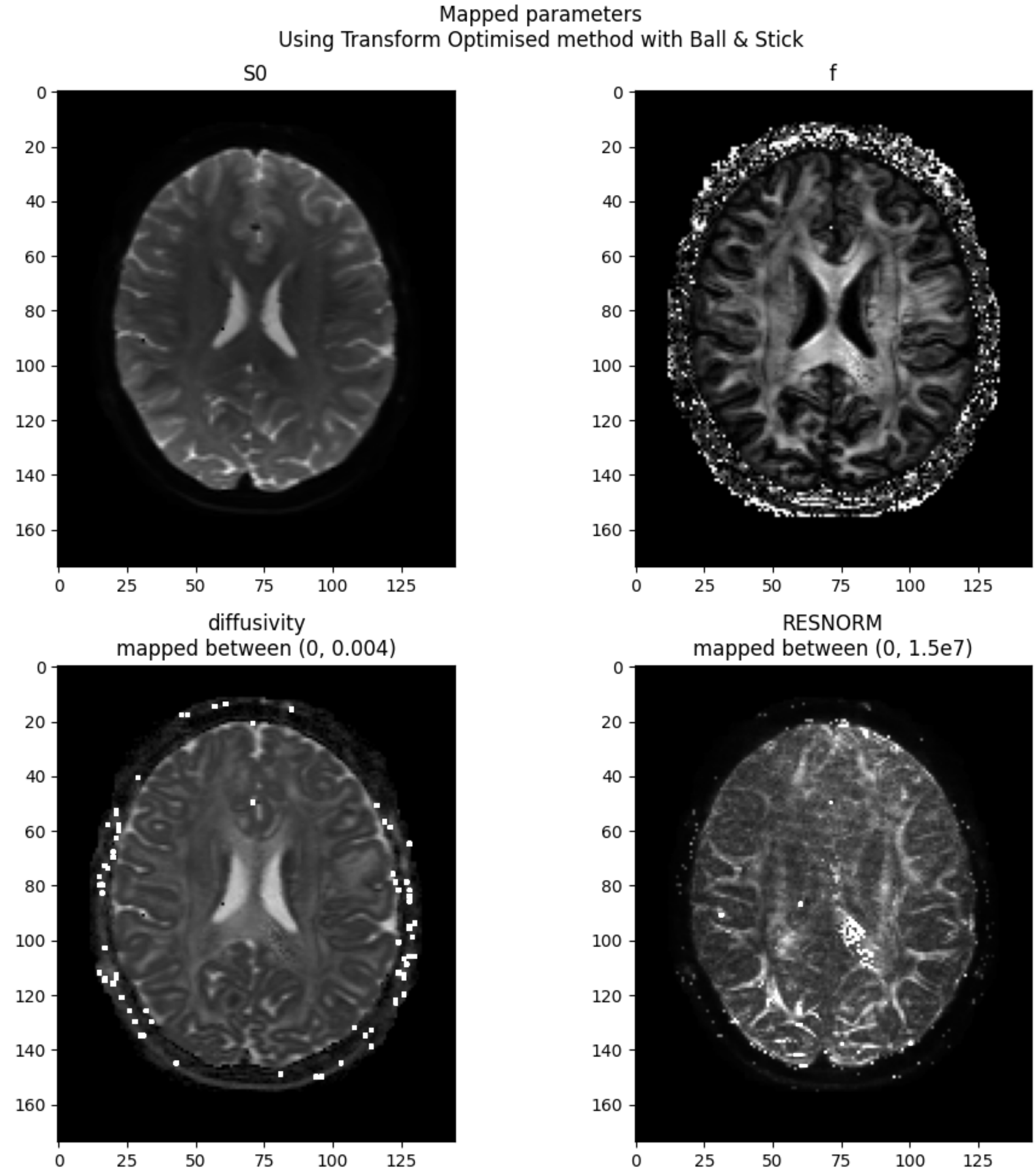

Figure 5: Parameter maps of S0, diff, and f, along with a map of RESNORM.

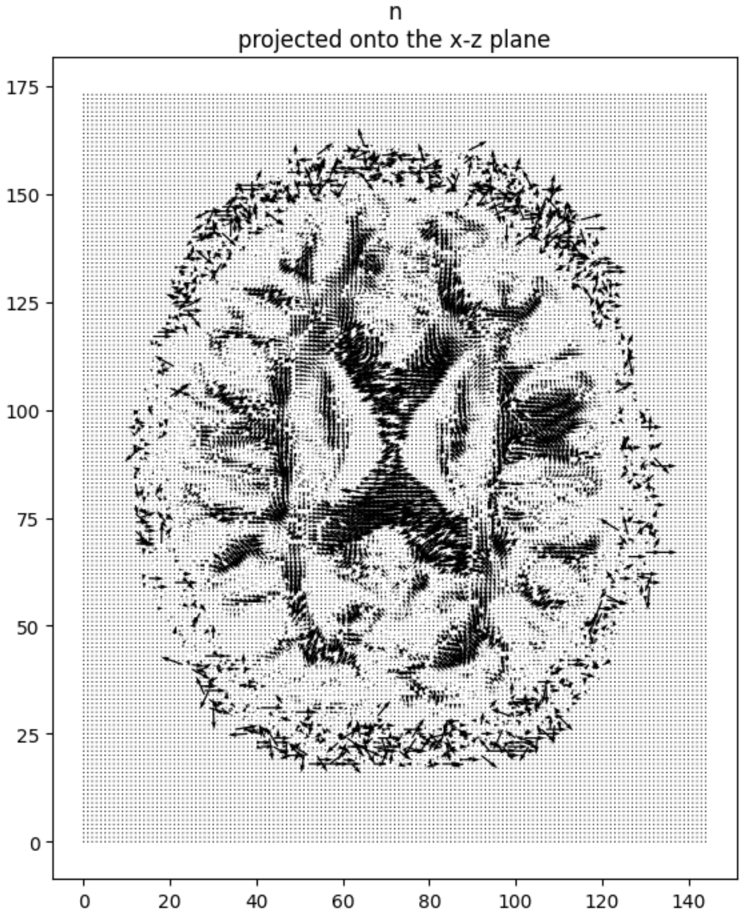

Figure 6: Map of the fiber directions of slice 72.

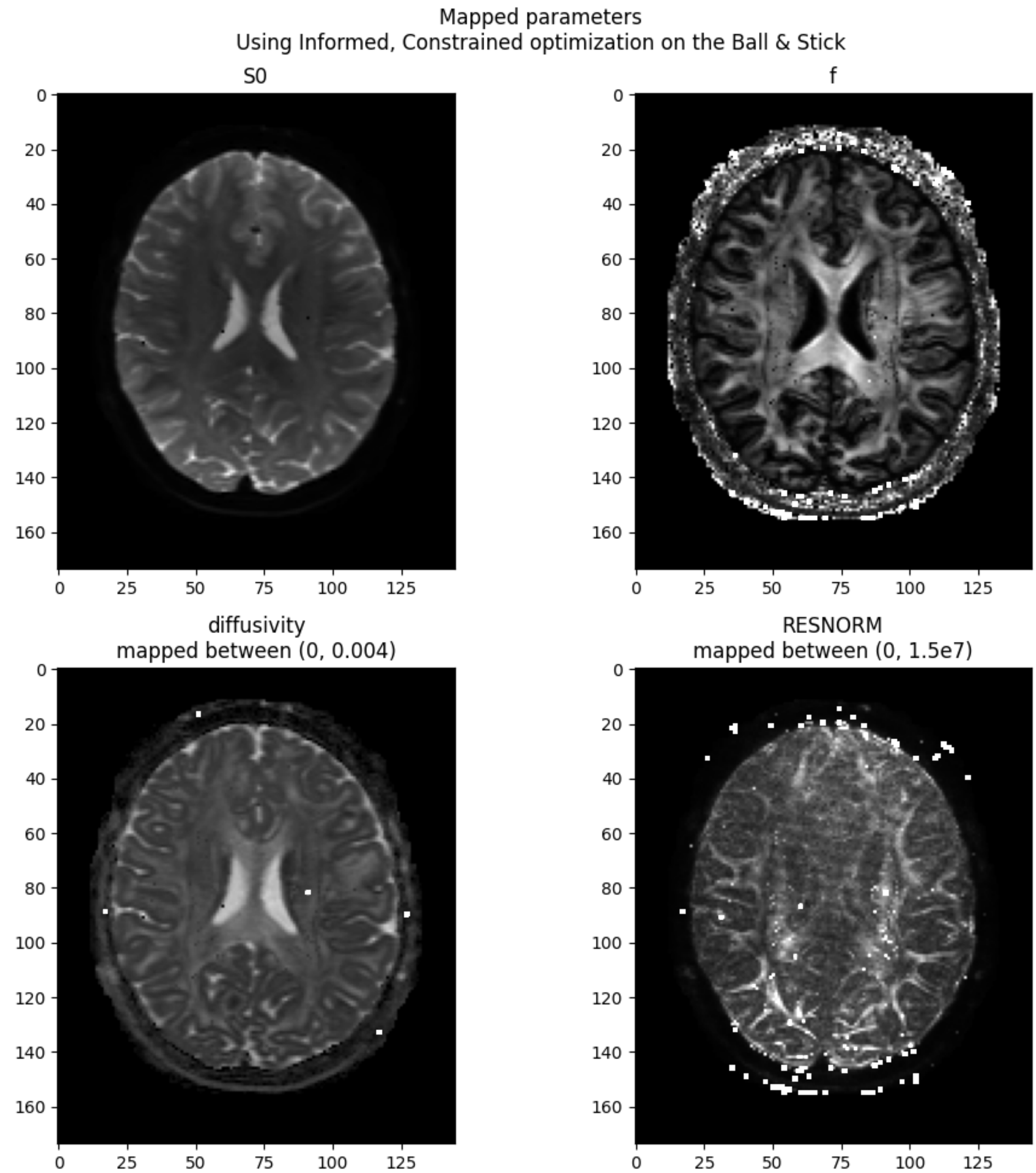

Figure 7: Parameter maps of S0, diff, and f, along with a map of RESNORM obtained via constrained optimization with informed starting point with the ball-stick model.

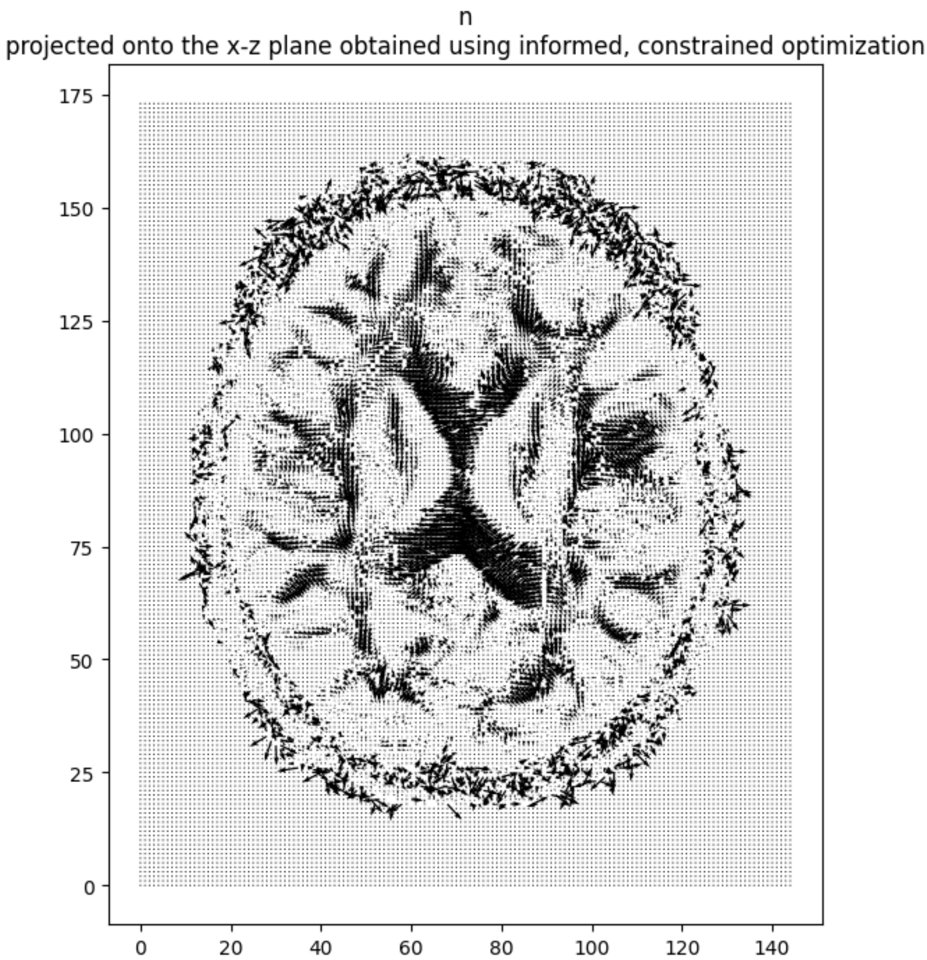

Figure 8: Fiber directions obtained via constrained optimization with informed starting point with the ball-stick model.

Experiment	**Max Number of Iterations**	**Prob_global_min**	**Time to compute parameter map**
Uninformed startx + Unconstrained optimization	14	Often 80%	N/A
Uninformed startx + Constrained optimization	6	Often 90%	~30 mins
Informed startx + Unconstrained optimization	14	Between 2%-60%	N/A
Informed startx + constrained optimization	4	Often 90%	~25 mins
uniformed startx + constrained optimization + analytical derivatives	6	Often 100%	N/A

informed startx + constrained optimization + analytical derivatives	4	Often 100%	~5 mins

Table 2: All uninformed starting points are `x3=[4200, 4e-4, 0.25, 0.9, 1]` based on the one given in the lectures, where we only modify the orientational angles to be close to 1, finding they yield stable estimations after experimenting with several other values. N/A indicates the experiment was not run owed to time constraints. The first column indicates the experimental setting, e.g., informed startx + constrained optimization + analytical derivatives indicate we use the DT to initialize the parameters that are constrained for the ball-stick, where we further compute the derivates analytically and passed them into `scipy's minimize's jac` parameter. The second column indicates the expected, maximum number of iterations to run to reach a global minimum obtained with experiments on 6 random voxels of slice 72 `[92,64], [71, 70], [41, 81], [81,61], [86,56], [86, 73]`. The third column indicates the probability of reaching the global minimum computed with the 6 voxels. The fourth column indicates the time it took to compute the parameter maps for each of the $145 * 174$ voxels, with the quickest to be the setting with informed startx + constrained optimization + analytical derivatives with only 5 minutes.

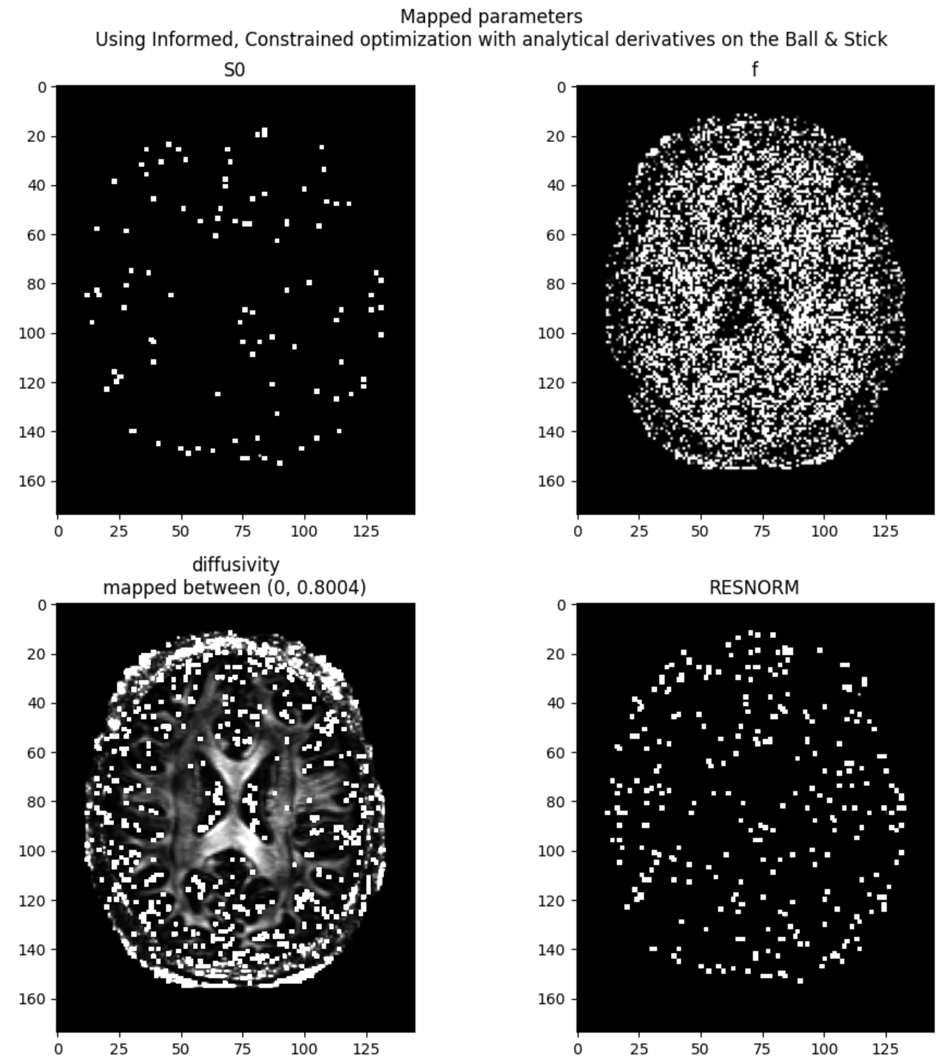

Figure 9: Parameter maps of S0, diff, and f, along with a map of RESNORM obtained via constrained optimization with informed starting point with the ball-stick model with analytical derivatives.

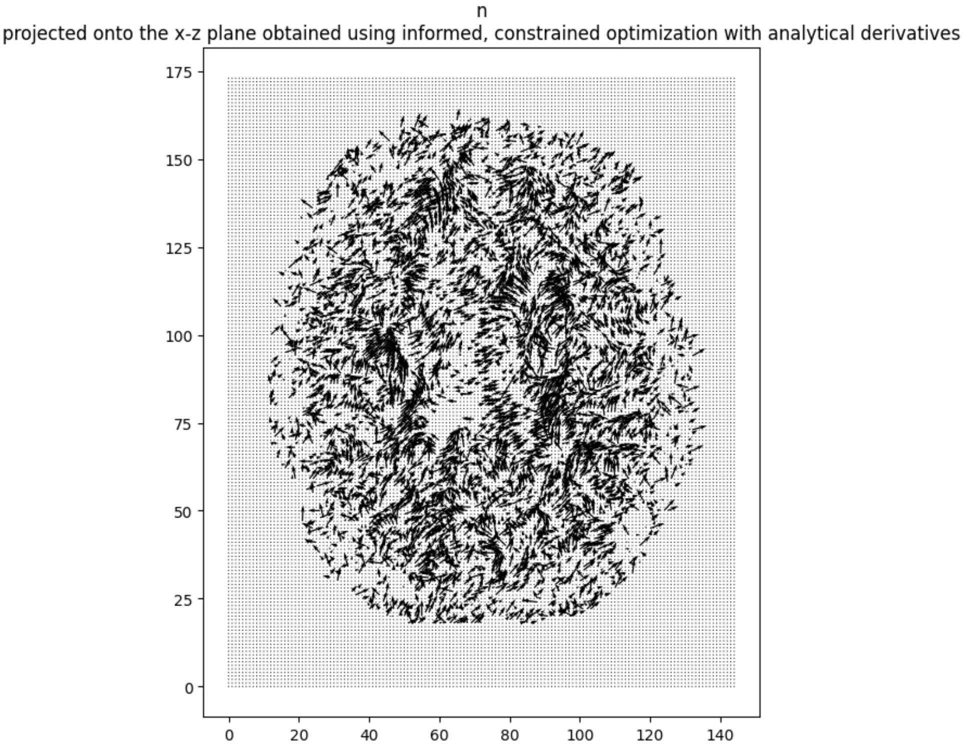

Figure 10: Fiber directions obtained via constrained optimization with informed starting point with the ball-stick model with analytical derivatives.

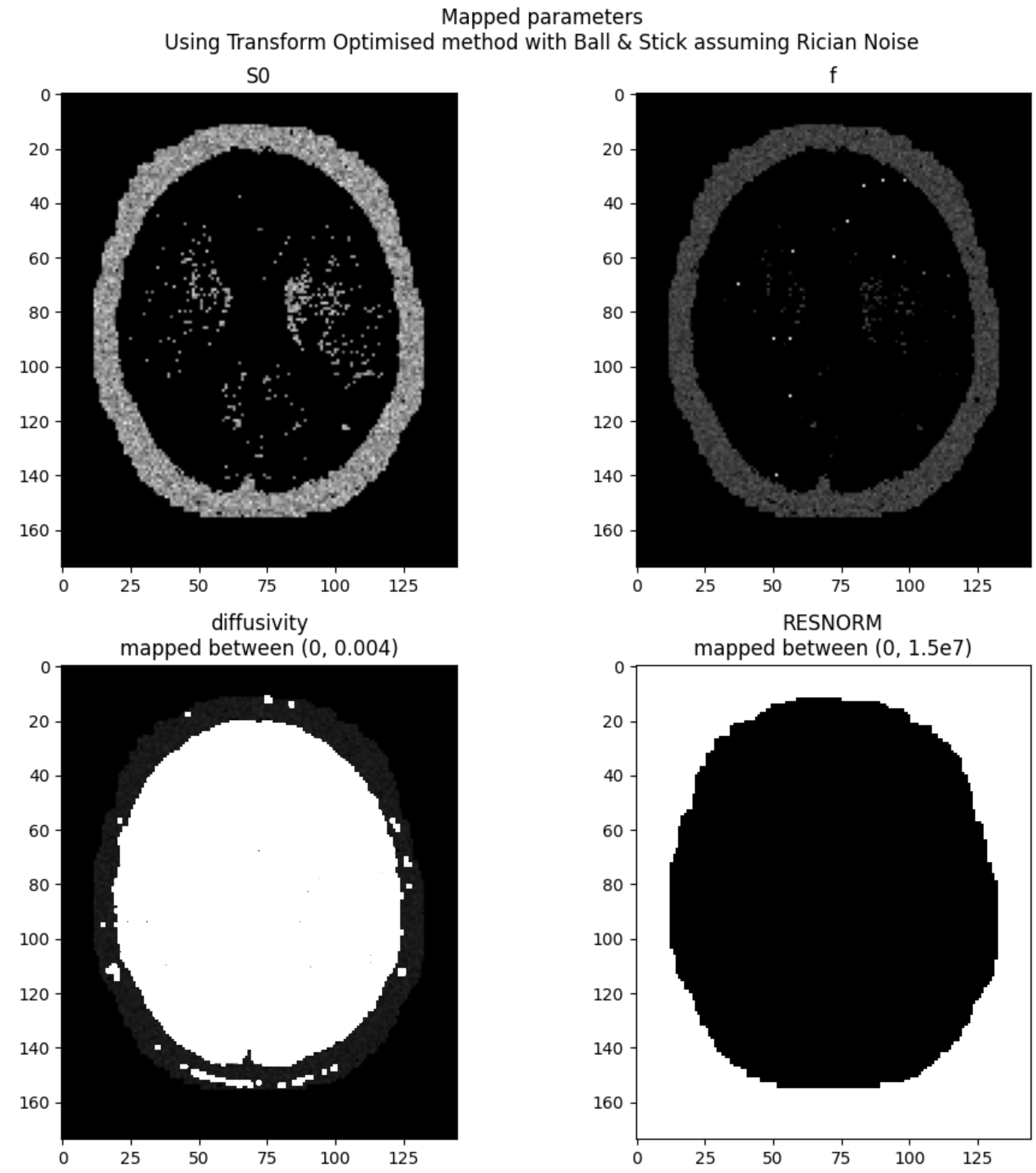

Figure 11: Parameter maps obtained from constrained optimization with the Rician Likelihood

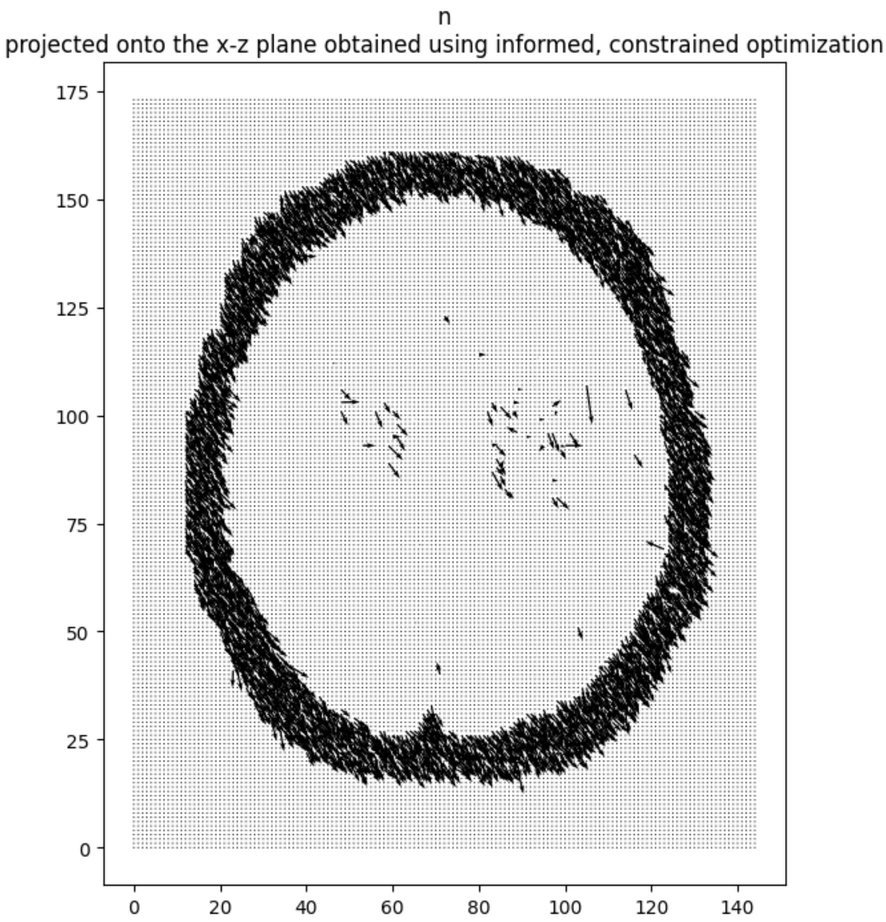

Figure 12: Fiber directions obtained from constrained optimization with Rician likelihood.

Parameter	Bootstrap mean	95% range	2 sigma range
S0	2507.49	[2268.1, 2830.7]	[2226.5, 2788.5]
Diff	0.000218	[2.8e-5, 4.1e-4]	[2.9e-5,4.1e-4]
f	0.934	[0.5, 1]	[0.6, 1.2]

Table 3: We bootstrap all parameters and report results for the required `S0, Diff, and f.` We show results for voxel $[92, 65]$,

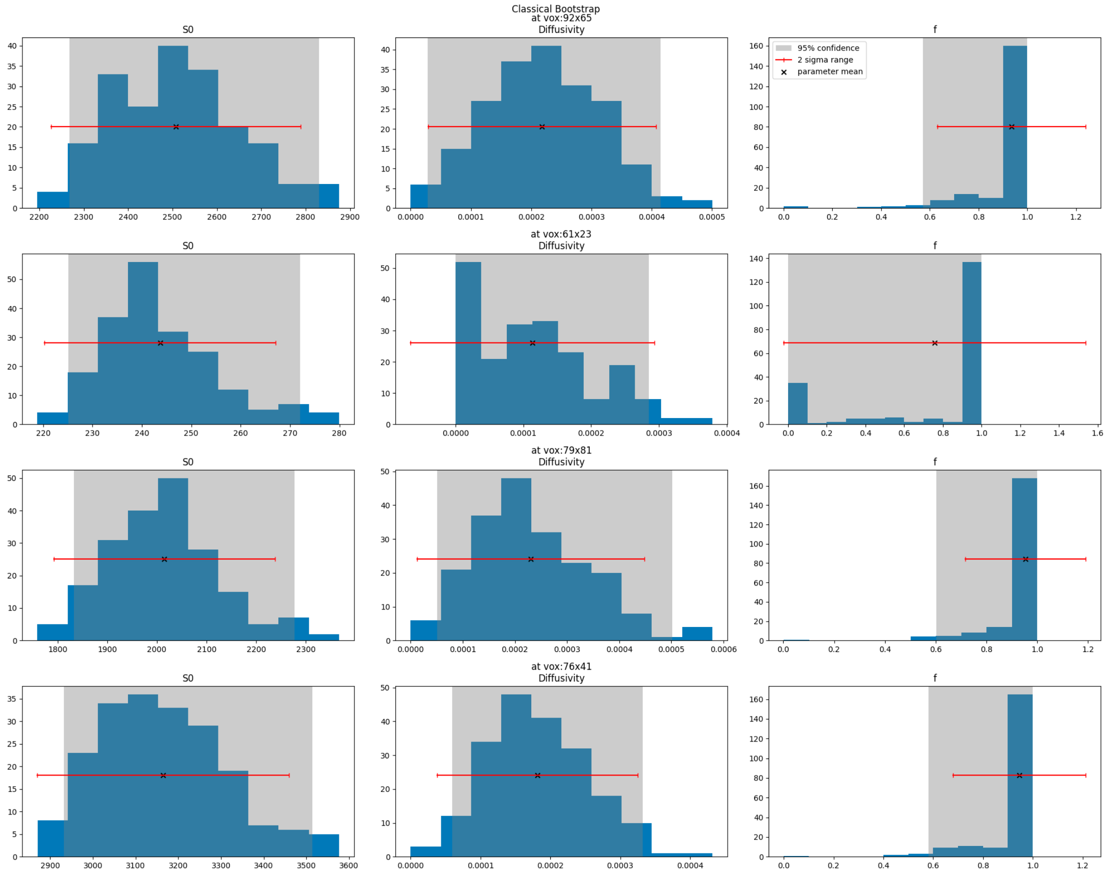

Figure 13: Classical bootstrap distributions for the 3 required parameters.

Parameter	Posterior mean	95% range	2 sigma range
`S0`	`4257.2`	`[4197.2, 4311.4]`	`[4199.8, 4314.7]`
`Diff`	`0.00114`	`[0.00110, 0.00118]`	`[0.00110,0.00118]`
`f`	`0.358`	`[0.334, 0.380]`	`[0.333, 0.380]`

Table 4: Ball-stick parameters estimated using MCMC, where we fit all 5 parameters and report results for the required `S0, Diff, and f.` We show results for voxel $[92, 65]$,

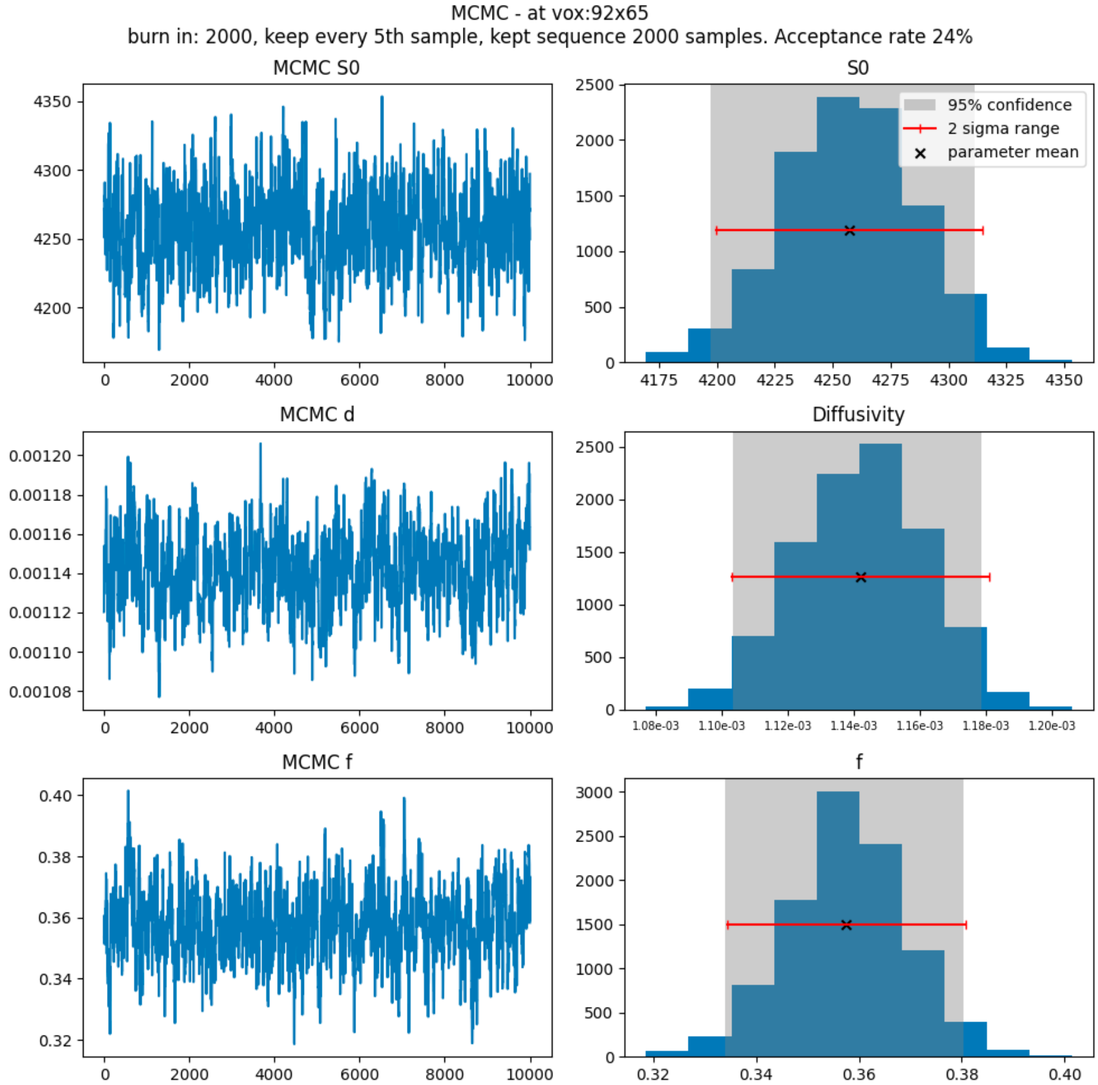

Figure 14: MCMC plots for voxel [92, 65]. The hyperparameters we use are a burn-in of `2000` to get rid initial samples of the random walk which might be of poor quality as well as making sure we sample closer to the true $P(x|A)$. The interval is set to `5` for thinning, so we only keep every 5^{th} sample to reduce autocorrelation of subsequent samples. A high sample size of 2000 is set to achieve a stable precision in parameter estimates. We manipulate the perturbation noise to be at 120 to get an acceptance rate of 23%, though we note that this depends on the voxel from which to estimate parameters.

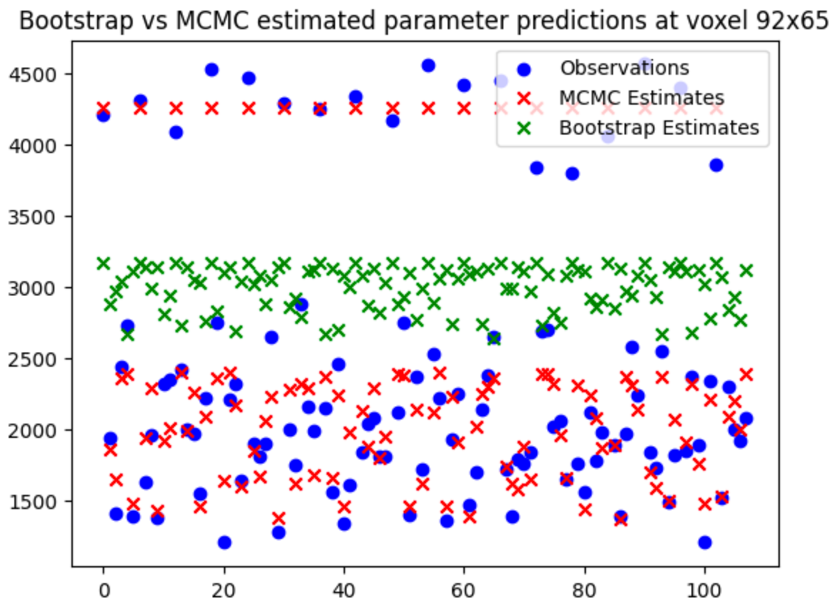

Figure 15: Comparison of predictions on signal observations (blue) made by parameters estimated from bootstrap (green) vs. from MCMC estimates (red)

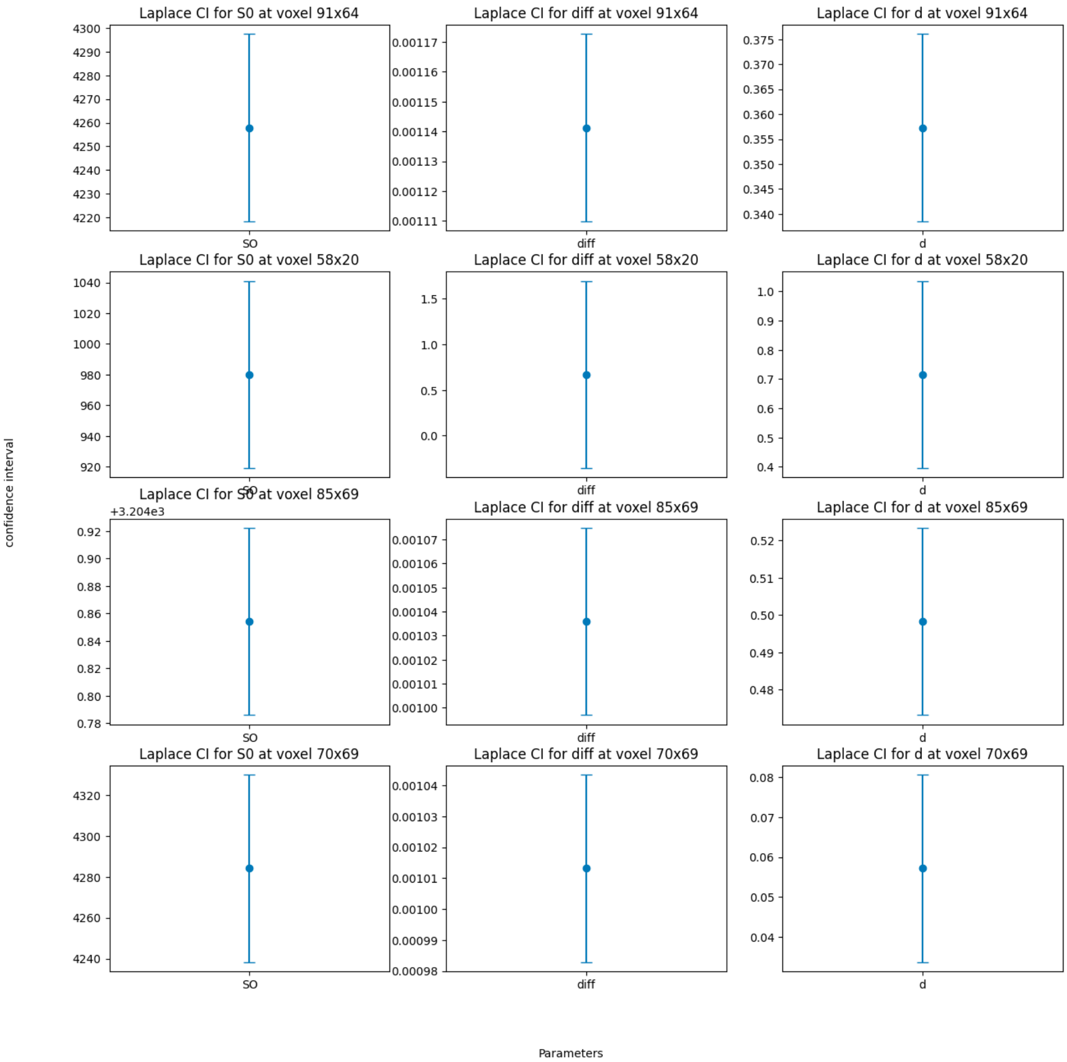

Figure 16: For Laplace's method, we rescale the diagonal entries of the Hessian inverse by $2 * \sigma^2$ where the noise is around 200 for this dataset. These diagonal entries serve as confidence intervals for the parameters fit.

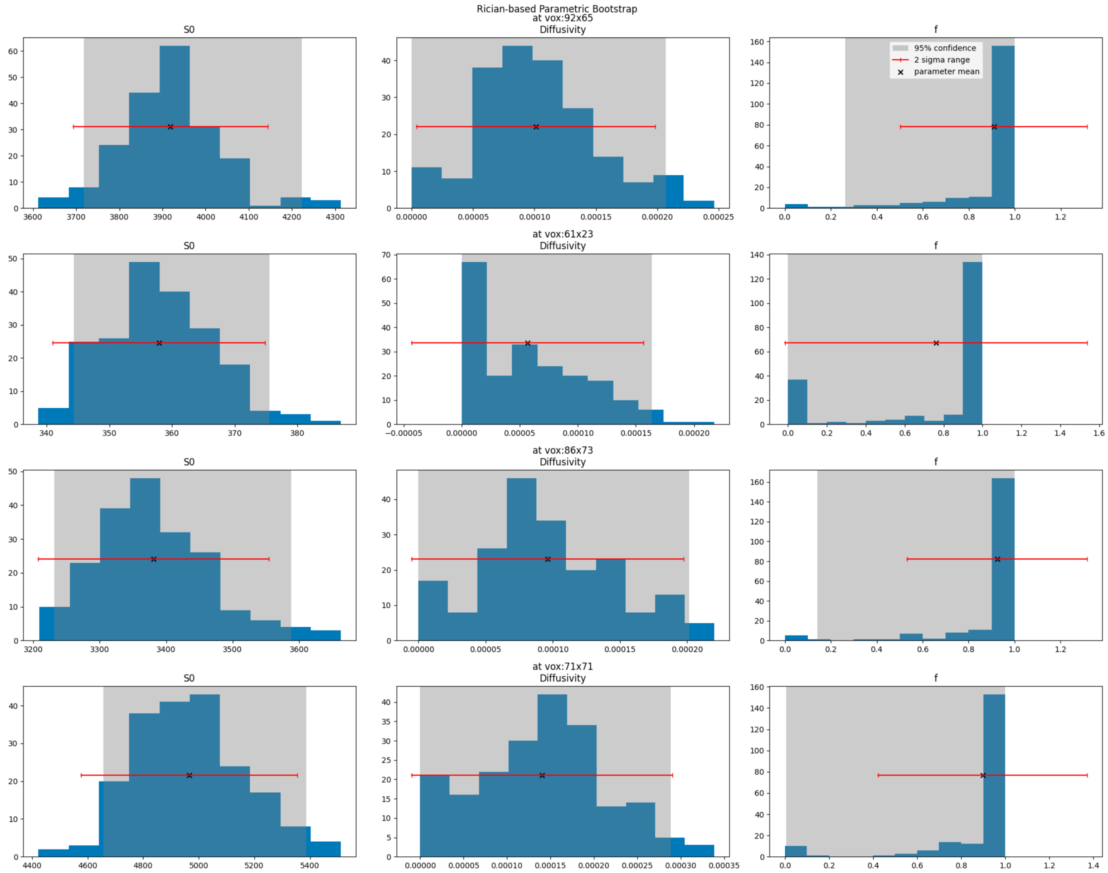

Figure 17: Parametric, Rician distribution based bootstrap estimates of the 3 parameters S0, diff, and f from the ball-stick model.

Parameter	Rician Bootstrap mean	95% range	2 sigma range
S0	3918.07	[3717.9, 4222.8]	[3692.5,4143.6]
Diff	0.0001012	[4.1e-18, 2.1e-4]	[4.1e-6 1.9e-4]
f	0.909	[0.26, 1.]	[0.5, 1.3]

Table 5: Rician-based parametric bootstrap estimates

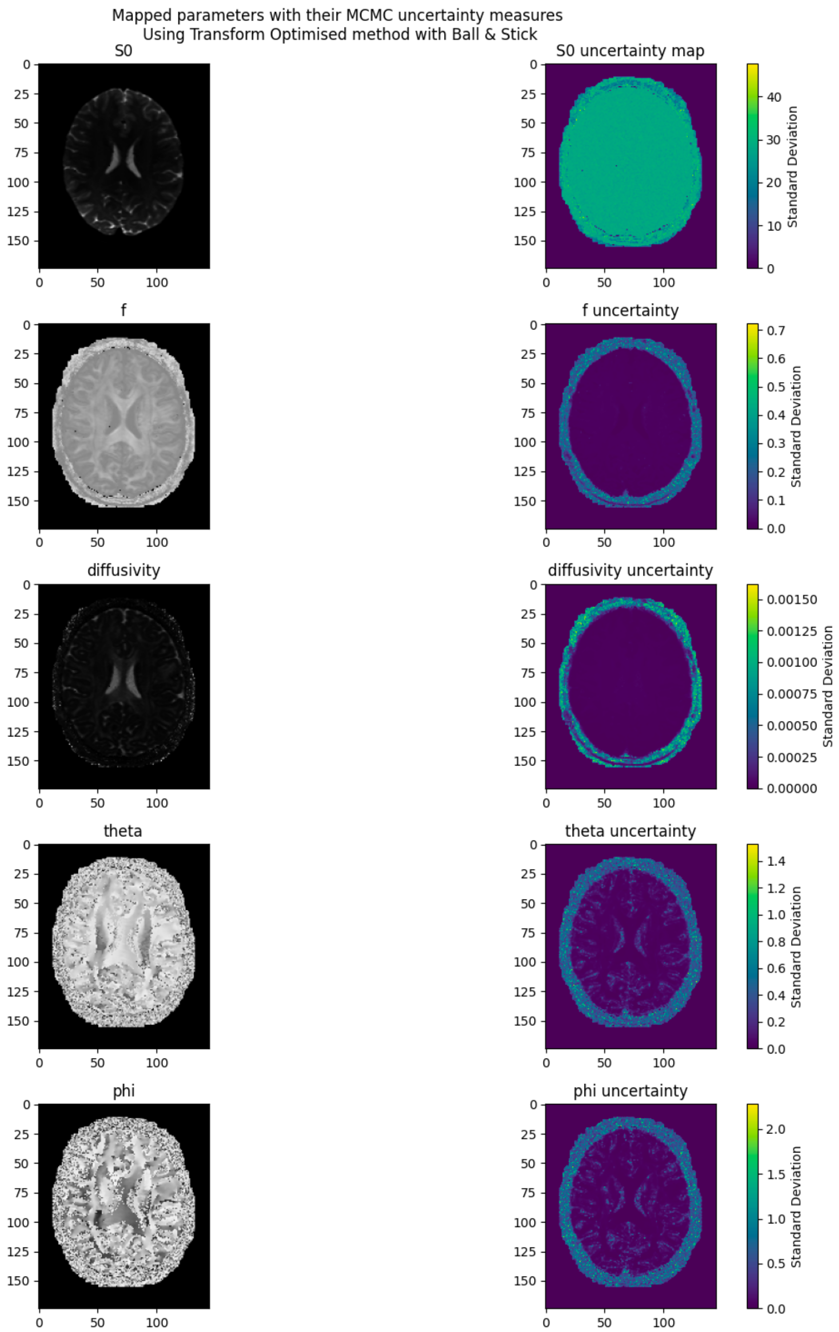

Figure 18: We plot uncertainty maps of the ball-stick parameters `S0, diff, f, theta, phi` by running the MCMC over each voxel of slice 72 and plotting both the MAP estimate of the parameters on the left, and on the right the standard deviation of each parameter color-coded as follows: the more blue a region (ignore background) is, the less uncertainty there is in the corresponding parameter, and yellow otherwise.

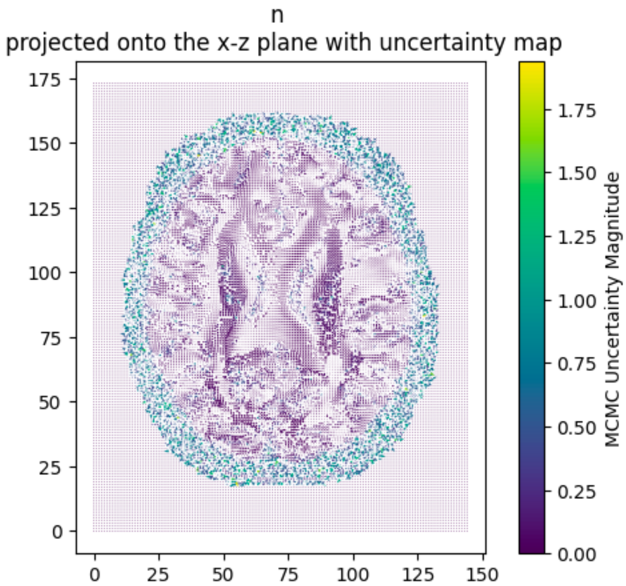

Figure 19: We plot uncertainty map of the fiber directions computed using MCMC parameter estimates. The uncertainty is computed via a propagation of errors that use the second derivatives and standard deviations computed using MCMC. The color code is the same as before.

Voxel	**Prob_global_min**	**RESNORM**
1	0.06	15
2	0.11	17
3	0.03	20
4	0.06	22
5	0.12	19
6	0.14	17

Table 6: For each voxel, the ball-stick model is fit for $N = 100$ iterations. We observe a very low probability of finding a global minimum, and RESNORMs are often 3-4 times higher than what's expected 5.8. This indicates the ball-stick model is perhaps not appropriate for this particular data due to mismatched assumptions.

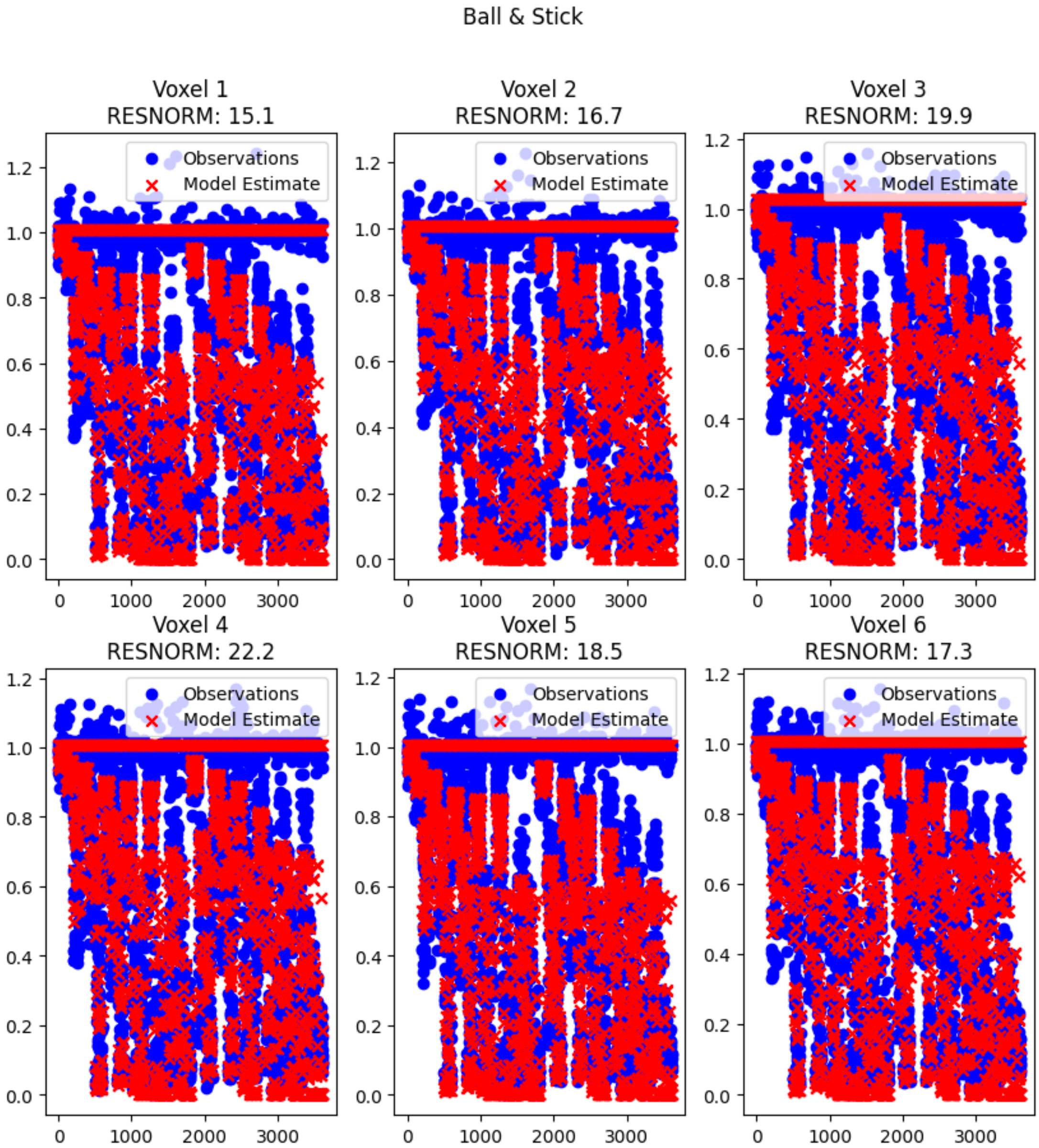

Figure 20: Predictions made by B-S model on 6 voxels with their RESNORM

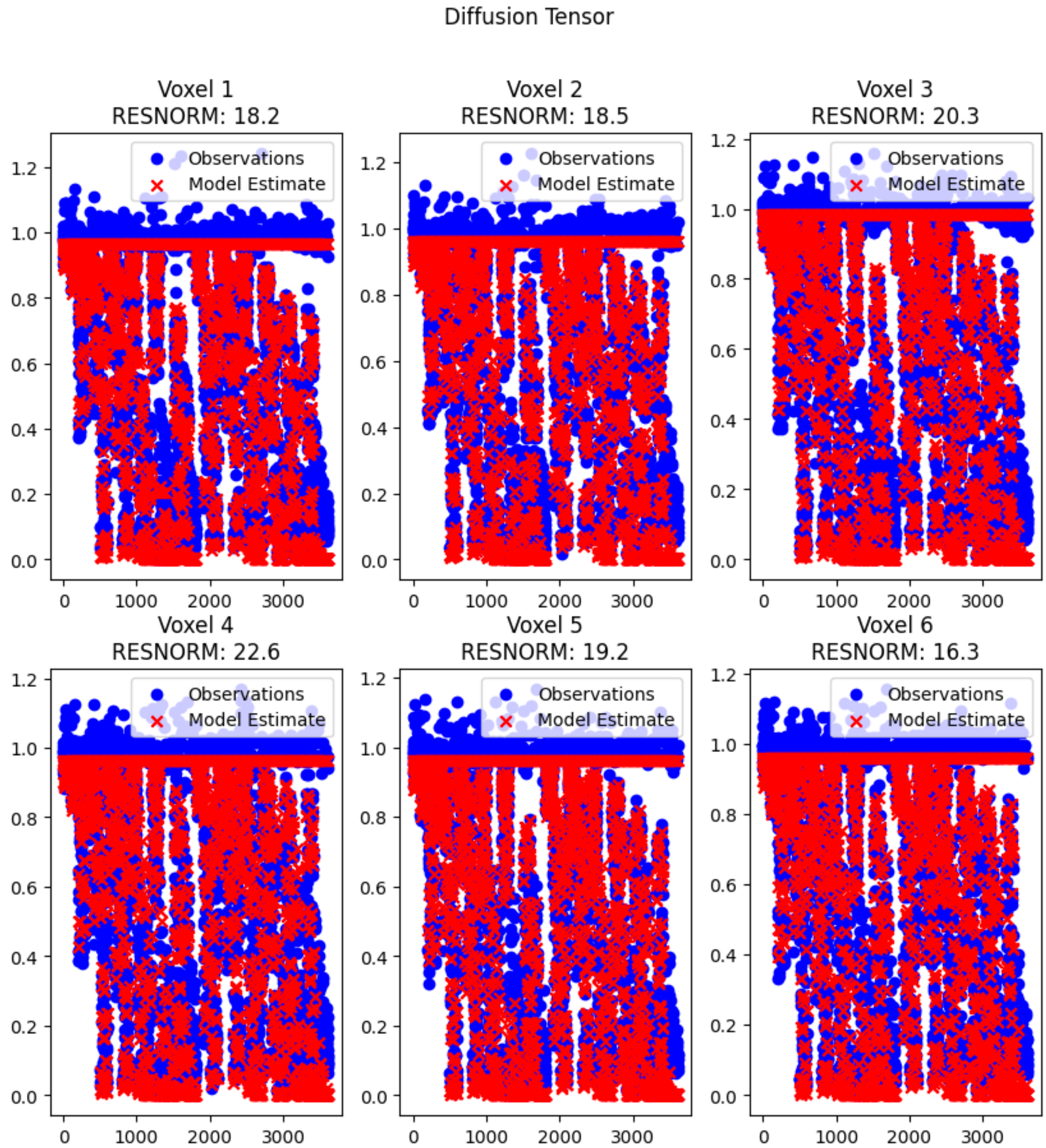

Figure 21: Predictions made by DT model on 6 voxels with their RESNORM

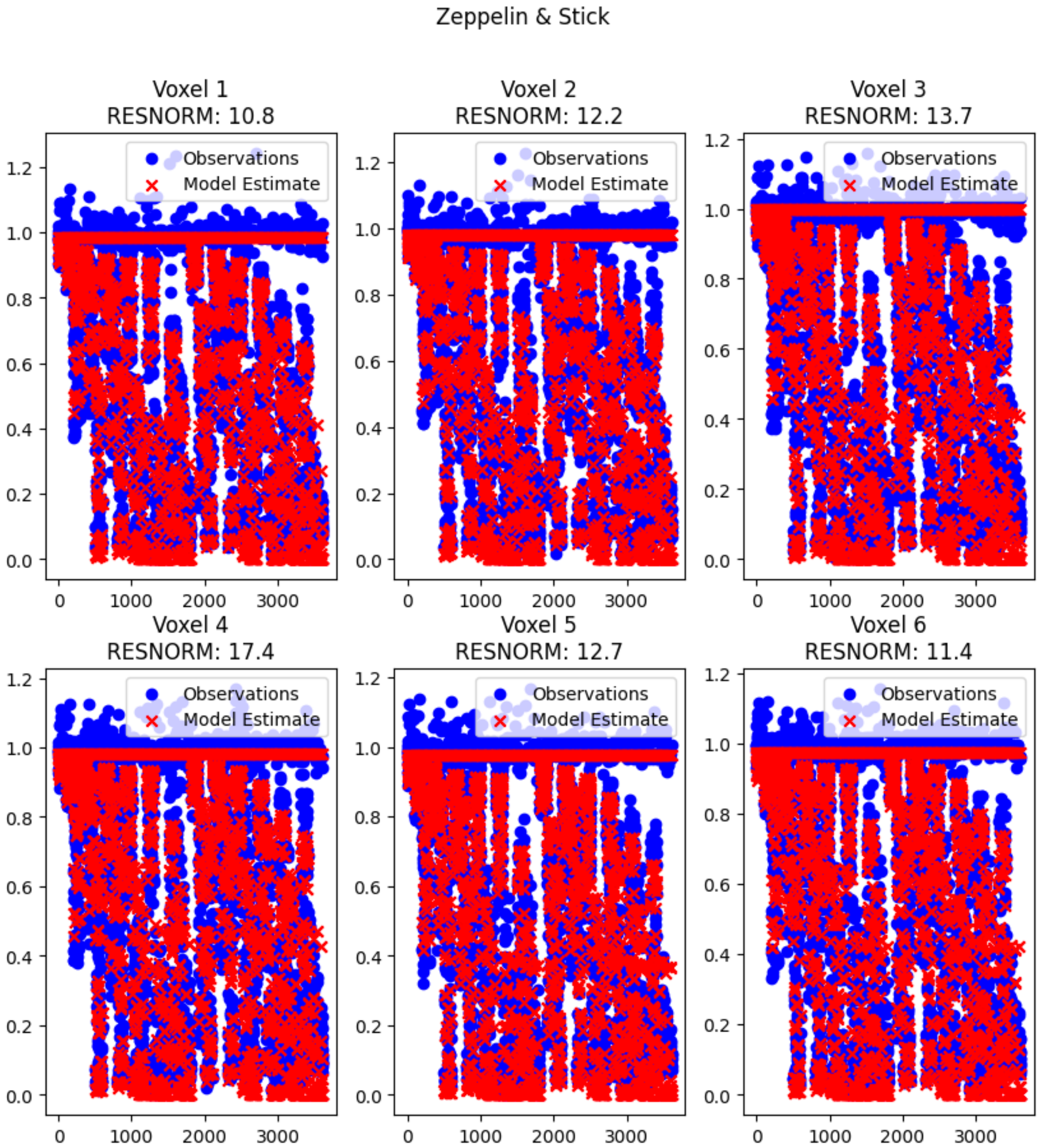

Figure 22: Predictions made by ZS model on 6 voxels with their RESNORM

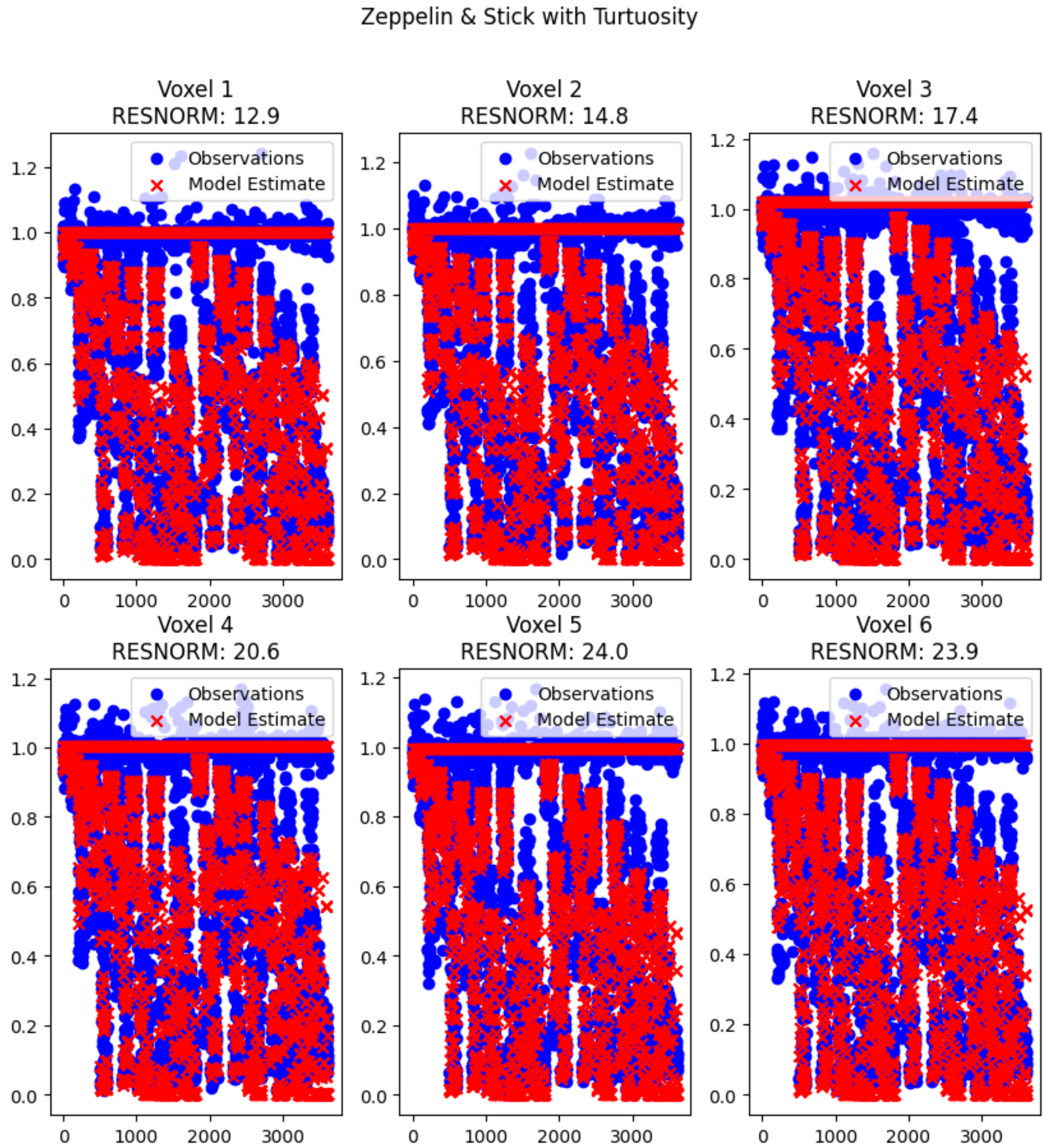

Figure 23: Predictions made by ZS with Tortuosity model on 6 voxels with their RESNORM

Model\Voxel	**1**	**2**	**3**	**4**	**5**	**6**
Ball & Stick	2	3	2	3	2	3
Diffusion Tensor	4	4	3	4	3	2
Zeppelin & Stick	1	1	1	1	1	1
Zeppelin & Stick with Tortuosity	3	2	4	2	2	4

Table 7: AIC and BIC rankings. Since they're the same, we report them in the same table. The AIC and BIC are computed using a noise std. deviation of 0.04 set for this data. The degrees of freedom for the Ball-Stick, DT, Zeppelin, and Zeppelin with Tortuosity models are 5, 7, 6 and 5 respectively.

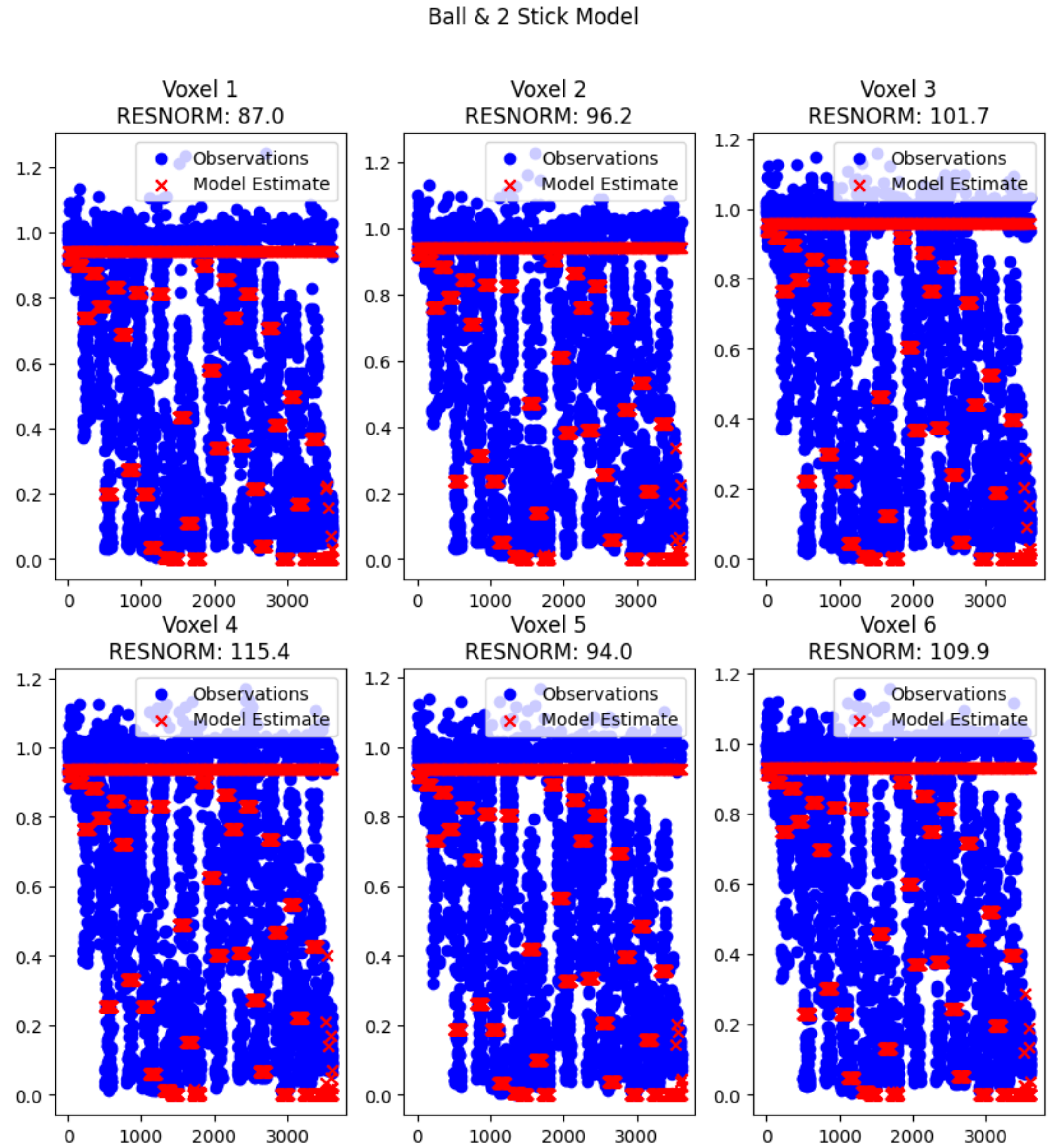

Figure 24: Predictions made by Ball-2 Stick model on 6 voxels with their RESNORM

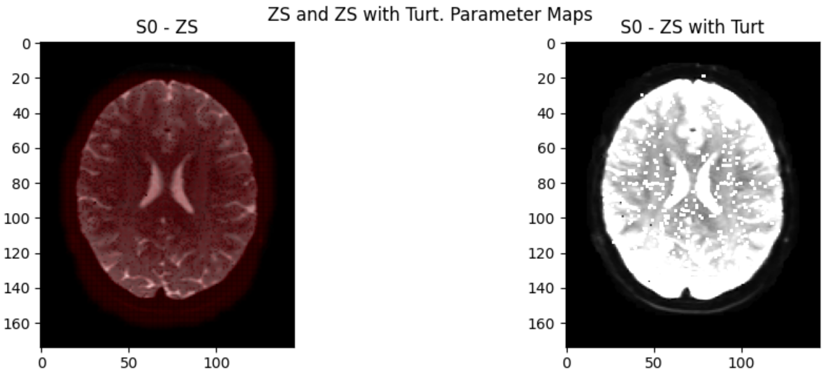

Figure 25: Parameter map for the S0 parameter estimated voxel-wise by the Zeppelin-Stick model on the left, and ZS with tortuosity in the right. In red we highlight regions where the AIC is greater for the ZS model than the ZS with tortuosity counterpart.