Evaluating the accuracy of diffusion models at multiple b-values with cross-validation

Ariel Rokem[1], Kimberly L Chan[1], Jason D Yeatman[1], Franco Pestilli[1], Aviv Mezer[2], and Brian A Wandell[2]
[1]*Stanford University, Stanford, CA, United States,* [2]*Stanford University, Stanford, California, United States*

PURPOSE: Models of diffusion-weighted MRI (DWI) are useful for making inferences about the properties of neural tissue and inferring fiber orientation distribution functions (fODFs) used by tractography algorithms. There is a substantial literature examining the reliability of the parameter estimates of the most common diffusion models[1], but there are no extensive assessments of model accuracy. We used cross-validation (CV) to assess the accuracy of two widely used classes of DWI models in the white matter. There is a great deal of interest in making measurements at high diffusion weighting, because increased b-values provide higher contrast between directions. But the potential improvement in angular resolution comes at the cost of reduced signal-to-noise ratio. One hypothesis is that a combination of the information from high-SNR low b-value data and high angular resolution high b-value would lead to increased accuracy in modeling the fODFs[5]. To test this hypothesis, we evaluated DWI models using multi-b-value measurements collected at standard resolution, as well as measurements with high spatial and angular resolution provided by the Human Connectome Project (HCP).

METHODS: Two datasets were evaluated. In data1, we measured DWI at standard spatial (2 mm isotropic) and angular (150 directions) using a GE 750 scanner at 3T, applying multiple different b-values (1/2/4 ms/mm^2) each in separate scans. Test-retest measurements were conducted at each b-value/direction. Data2 are from the Human Connectome Project (HCP), and contain DWI measured at high spatial (1.5 mm isotropic) and angular (270 directions) resolution, with multiple b-values (1/2/3 ms/mm^2; 90 distinct directions in each b-value)[5]. For data1, models were fit to one measurement. To assess model accuracy, the model predictions were tested on the repeat measurement. We calculated the relative root mean squared error: $rRMSE = \frac{\sqrt{<(model-data_1)^2>}}{\sqrt{<(data_2-data_1)^2>}}$, where <> denotes averaging across directions in each voxel. This index normalizes the model error with respect to test-retest reliability. When *rRMSE* is smaller than 1, the model predicts the data better than test-retest reliability. In data2, we used k-fold CV: in each iteration, 10% of the data was left out and the model was fit to the remaining 90%. The model was used to make a prediction for the left-out data set, and this was repeated until model predictions were made for all of the data. We evaluated the accuracy of two widely used models. One is the classic diffusion tensor model (DTM)[6]. More recent extensions of this model assume that there may be multiple fiber populations (or fascicles) within the voxel, each contributing to the diffusion signal. To prevent over-fitting, these models impose sparseness constraints on the number of distinct fascicles in each voxel; hence, we refer to these models as Sparse Fascicle Models (SFM). We fit the SFM using a regularized regression procedure[7], deconvolving the signal in each b-value with a single-fascicle impulse-response function. In data1, deconvolution of the fODF is done separately in each b-value. In data2, a single fODF was also fit to the data in all b-values, using a combination of the individual b-value single-fascicle response functions. To account for non-Gaussian diffusion, a multi-compartment model of diffusivity was fit to the mean signal across b-values.

RESULTS: In data1 the DTM model prediction is better than test-retest reliability (*rRMSE* < 1) in more than 98% of the white matter voxels in all b-values measured. However, at high b values (>1 ms/mm^2) there are clusters of relatively low accuracy (*rRMSE*>1). These clusters are located in regions known to contain fiber crossings, such as the centrum semiovale (depicted in Figure 1), and in the optic radiations. In these regions, the SFM more accurately predicts the data. Despite the high spatial resolution of the measurements in data2, these measurements have comparable SNR to the measurements conducted at standard resolution. In data2, the SFM is also more accurate than the DTM.

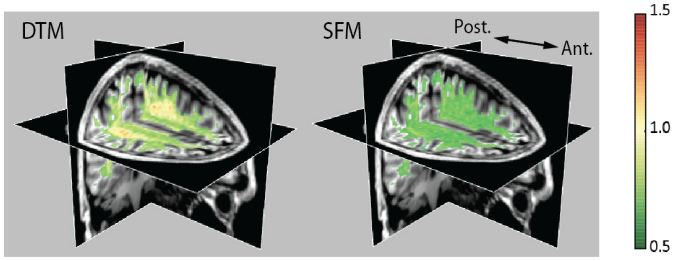

Figure 1: Relative RMSE at b=4 ms/mm² in data1. At the crossing of the corticospinal tract and other major fiber bundles SFM fits the data better than DTM.

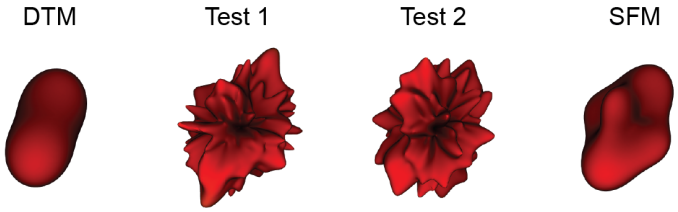

Figure 2: DWI signal profile in a single voxel in data1 (b-value=4 ms/mm²). DTM does not predicit the signal in this voxel, because it cannot capture the multiple 'dimples' in the signal, due to crossing fascicles. SFM predictions are more accurate than test-retest reliability, because the model is regularized to ignore the noise in different samples.

DISCUSSION and CONCLUSION: We conducted the first extensive study of model accuracy of DWI models in the white matter. We find that the two most commonly used models are more accurate than test-retest reliability in most of the white matter. The SFM is more accurate than the DTM, particularly for measurements with (a) a b-value above 1 ms/mm^2 (b) in locations containing fiber crossings. SFM also has the important theoretical advantage that it reliably estimates the (fODF) in each voxel, which is useful for fiber tracking. To conclude, model accuracy is an essential metric to assess and compare different models. The high accuracy of the models, relative to the repeatability of the data, underscores the utility of using models to analyze DWI data, compared to "model-free" reconstruction techniques, such as diffusion spectrum imaging (DSI).

REFERENCES: 1. Jones, D. K. Determining and visualizing uncertainty in estimates of fiber orientation from diffusion tensor MRI. *MRM* **49,** 7–12 (2003). 2. Frank, L. R. Characterization of anisotropy in high angular resolution diffusion-weighted MRI. *MRM* **47,** 1083–99 (2002). 3. Behrens, T. E. J., Berg, H. J., Jbabdi, S., Rushworth, M. F. S. & Woolrich, M. W. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage* **34,** 144–55 (2007). 4. Tournier, J.-D., Calamante, F. & Connelly, A. Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. *Neuroimage* **35,** 1459–72 (2007). 5. Sotiropoulos, S. N. *et al.* Advances in diffusion MRI acquisition and processing in the Human Connectome Project. *Neuroimage* **80,** 125–43 (2013). 6. Basser, P., Mattiello, J. & Le Bihan, D. Estimation of the effective self-diffusion tensor from the NMR spin echo. *J. Magn. Reson.* **103,** 247–254 (1994). 7. Zou, H. & Hastie, T. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. B* 301–320 (2005).