Machine Learning Based Compartment Models with Permeability for White Matter Microstructure Imaging

Gemma L. Nedjati-Gilani[1], Torben Schneider[2] Matt G. Hall[1], Claudia A.M. Wheeler-Kingshott[2], and Daniel C. Alexander[1]

[1] Centre for Medical Image Computing and Dept. of Computer Science, University College London, Gower St, London, UK
g.nedjati-gilani@ucl.ac.uk
[2] Dept. of Neuroinflammation, Institute of Neurology, University College London, Queen Square, London, UK

Abstract. The residence time τ_i of water inside axons is an important biomarker for white matter pathologies of the human central nervous system, as myelin damage is hypothesised to increase axonal permeability, and thus reduce τ_i. Diffusion-weighted (DW) MRI is potentially able to measure τ_i as it is sensitive to the average displacement of water molecules in tissue. However, previous work addressing this has been hampered by a lack of both sensitive data and accurate mathematical models. We address the latter problem by constructing a computational model using Monte Carlo simulations and machine learning in order to learn a mapping between features derived from DW MR signals and ground truth microstructure parameters. We test our method using simulated and in vivo human brain data. Simulation results show that our approach provides a marked improvement over the most widely used mathematical model. The trained model also predicts sensible microstructure parameters from in vivo human brain data, matching values of τ_i found in the literature.

1 Introduction

Numerous white matter (WM) pathologies of the human central nervous system (CNS), such as multiple sclerosis, spinal cord injury and leukodystrophies, are characterised by damage to the insulating myelin sheaths around the axonal fibres. As the breakdown of myelin is hypothesised to lead to an increase in axonal permeability, there is widespread interest in developing imaging biomarkers based on permeability or intra-axonal water residence time τ_i in order to improve diagnosis of and treatment for these conditions. Diffusion-weighted (DW) MRI is potentially amenable to estimating τ_i as it is sensitive to the dispersion of water molecules within tissue. However, due to a lack of both sensitive data and sufficiently accurate mathematical models, progress has been limited. Whereas improvements in modern hardware and the development of specialised imaging sequences are beginning to address the former issue, there is still a need

P. Golland et al. (Eds.): MICCAI 2014, Part III, LNCS 8675, pp. 257–264, 2014.

to develop models that accurately characterise water exchange within the brain. Many commonly used compartment models for microstructure imaging, such as AxCaliber [4] and ActiveAx [3], ignore the effects of permeability completely and overestimate axon diameter. Including the effects of permeability in these techniques not only gives us another clinically useful parameter τ_i, but may also improve the estimation of other microstructure indices.

Mathematical models that explicitly incorporate τ_i include the Kärger model (KM) and apparent exchange rate (AXR) imaging. The KM is commonly used [8,12] as it is compatible with data acquired using widely available pulsed gradient spin echo (PGSE) and stimulated echo (STE) imaging sequences. It accounts for intercompartmental water exchange by coupling the DW MR signals due to the separate compartments via τ_i. However, it relies on the assumption that the two pools of water are well mixed and it does not model restriction. These conditions are not valid in WM tissue where the intra- and extra-axonal compartments are spatially separate and the axonal membranes restrict the motion of water molecules. Even though these limitations have been known for over 20 years, there have been no improvements to the model due to the mathematical intractability of the problem. AXR imaging [9] has recently been introduced as an alternative to the KM; however it requires specialised double PGSE imaging sequences, and it again relies on strong assumptions about the compartmentation of water into a 'fast' and 'slow' pool. The estimated AXR parameter also conflates τ_i with intra-axonal volume fraction f, making it difficult to disentangle the origin of any measured change.

Given the inherent difficulties involved in deriving accurate analytical models of permeability, we approach this problem by constructing a computational model to learn the mapping between microstructural features of interest and the data, bypassing the need for a mathematical model altogether. We use Monte Carlo (MC) simulations to generate synthetic signals from a library of histologically relevant microstructure indices. A random forest (RF) regressor then learns the mapping between features derived from DW MR signals and ground truth microstructure parameters using the synthetic data, providing an efficient and accurate method for predicting microstructure parameters, including τ_i.

Previous related work [8] generated libraries of microstructure parameters and their corresponding DW MR signals from MC simulations, and used them to find the nearest-neighbour microstructure parameters that matched unseen signals; however nearest-neighbour matching typically has poor generalisation to unseen input data and the method has only been demonstrated on synthetic data. We extend this approach using RF regression which has better generalisation to unseen data [5], i.e. combinations of tissue parameter values not explored in the training set. We compare our approach to the KM using simulated data and demonstrate that the trained RF can be used to predict sensible estimates of microstructure indices from in vivo human brain WM.

2 Methods

Protocol Optimisation. We use an orientationally invariant (OI), DW-STE protocol, previously optimised [2] for a two-compartment KM assuming a maximum imaging time of 30 minutes. The protocol was optimised for the following biophysically plausible tissue parameters: f=0.7, parallel diffusivity $d_{\parallel}$=2×10^{-9} m^2s^{-1}, perpendicular diffusivity $d_{\perp}$=0.7×10^{-9} m^2s^{-1}, axon radius R=1 μm, $\tau_i\in\{0.05,\ 0.1,\ 0.2,\ 0.4,\ 1\}$ s, $T1$=832 ms. The resulting protocol contains 89 measurements divided into 4 distinct shells, with Δ ranging from 95 to 398 ms. The final protocol, accounting for the effects of the additional STE imaging gradients, is shown in table 1.

Monte Carlo Simulations. We use MC simulations [6], in combination with the OI protocol in table 1, to generate DW MR signals from 12,500 WM tissue substrates. We model WM as a collection of 100,000 non-abutting, parallel cylinders with radii drawn from a gamma distribution (with mean μ_R, standard deviation σ_R). Each substrate is described by a unique combination of μ_R, σ_R, f, τ_i and $d_{\parallel}$ which are randomly selected in the ranges: $\mu_R \in [0.2, 5]$ μm, $\sigma_R \in[\min(0.1,\frac{\mu_R}{5}),\ \frac{\mu_R}{2}]$ μm (to ensure that the distributions have a non-zero mode, matching the distributions observed in histology [1]), $f \in [0.4, 0.7]$, $\tau_i \in [20, 950]$ ms, $d_{\parallel} \in [0.8, 2.2]\times10^{-9}$ m^2s^{-1}. All simulations are performed using 100,000 spins and 2,000 time steps. We generate two sets of signals: noise-free and noisy. As spins undergo $T1$ relaxation during the mixing time TM between the two diffusion gradients, measurements made using longer Δ (and so longer TM) experience more relaxation leading to lower signal intensities and signal to noise ratios (SNR). For the noisy data set we scale the signals by $\exp\left(-\frac{TM}{T1}\right)$ where $T1$=832 ms for WM at 3T. We then add Rician noise, choosing the standard deviation of the noise σ so that the SNR of the $b = 0$ images with Δ=95 ms is 20.

Data Acquisition and Pre-processing. We acquire DW images, using the protocol in table 1, from one healthy subject (male, age 31) on a 3T Phillips Achieva scanner with the following imaging parameters: TE=55 ms, TR=12000 ms, FOV=256mm×256mm, matrix size=128×128, slice thickness=4 mm, no. of slices=30. The total imaging time is approximately 30 minutes. As the model

Table 1. OI STE protocol parameters, optimised for Kärger model parameter estimation

Shell	# $b = 0$ ($s\ mm^{-2}$)	# gradient directions	b ($s\ mm^{-2}$)	$\lvert G\rvert$ ($mT\ m^{-1}$)	Δ (ms)	δ (ms)	TM (ms)
1	1	32	1592	70	95	7	69
2	1	6	1746	62	398	4	370
3	1	32	3538	71	394	5	370
4	1	15	3950	84	162	7	135

that we learn here is specific to axons that resemble parallel cylinders, it is not applicable in regions containing CSF, grey matter or highly dispersed or crossing WM fibres. We mask out these voxels by computing maps of linearity $C_L = \frac{\lambda_1 - \lambda_2}{\lambda_1}$ and planarity $C_P = \frac{\lambda_2 - \lambda_3}{\lambda_1}$ [13] from diffusion tensor (DT) fits, and select only those voxels with $C_L > 0.5$ and $C_P < 0.3$. The SNR of the selected WM region in the $b = 0$ image with $\Delta = 95$ ms is approximately 19.

Random Forest Regression. RF regression works by averaging the predictions from an ensemble of randomly trained decision trees [5]. We use it here to learn a mapping between rotationally invariant features derived from DT (all shells) and 4th order spherical harmonic (SH) fits (shells 1,3,4) to the simulated DW MR signals and the ground truth microstructure parameters. From the DT fits we calculate the eigenvalues $\lambda_1, \lambda_2, \lambda_3$, mean diffusivity and fractional anisotropy. From the SH fits we calculate the mean, peak, dispersion (i.e. the eigenvalues of the hessian matrix at the peak), anisotropy, skewness and kurtosis of the apparent diffusion coefficient profile, as well as simple combinations of the SH coefficients given by $I_k = \sum_{i=-k}^{k} |a_{k,i}|^2$ for $k = 0, 2, 4$ where $a_{k,i}$ is the coefficient of SH order k and index i. This gives a vector with 50 features for each measurement. The RF regressor, containing 100 trees of maximum depth 20, is trained [10] on 10,000 of the 12,500 feature vectors from noisy and noise-free data separately. The remaining 2,500 previously unseen feature vectors are used for testing. When predicting the microstructure indices from the noise-free test sets, we use the RF trained on noise-free data; when predicting from the noisy test and in vivo feature vectors (which have similar noise characteristics), we use the RF trained on noisy data. The microstructure parameters we estimate during the RF regression are f, $d_\parallel$, τ and α, a single axon radius index [3] which reflects both the mean and spread of the radius distribution.

Kärger Model Fitting. We fit a two-compartment KM to the 2,500 noise-free and noisy test data sets, and to the masked WM voxels from the in vivo human data set. The intra-axonal compartment, with volume fraction f, is modelled using randomly packed, parallel cylinders with radius R and an intra-axonal water residence time τ_i. The extracellular space is modelled as a cylindrically symmetric DT with diffusivities, $d_\parallel$ and $d_\perp$. $d_\parallel$ is assumed to be the same in both compartments. The model is simplified using the tortuosity model as in [3]. Prior to model fitting, each measurement is normalised by the $b = 0$ measurement with the same TM to eliminate $T1$ effects. The model is fit using Markov Chain Monte Carlo (MCMC) with an offset Gaussian noise model (assuming different noise standard deviations σ for each shell of data, which we estimate a priori) to sample from the posterior distribution over the model parameters. The burn-in phase for the MCMC consists of 10,000 steps, after which we collect 1,000 samples at an interval of 100 steps.

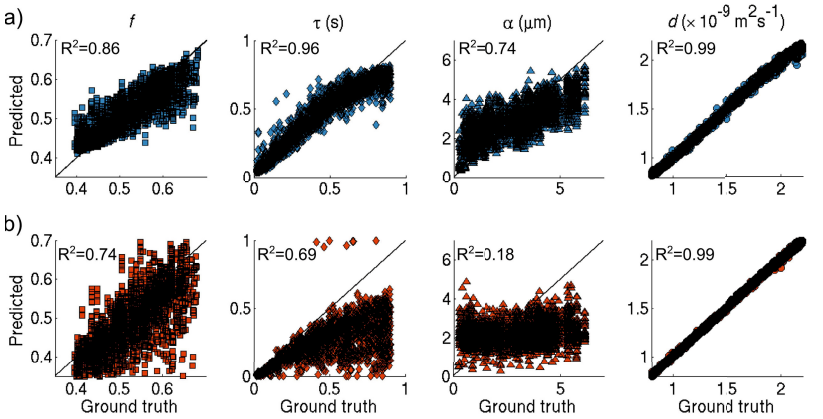

Fig. 1. Scatter plots of ground truth values of f, τ, α, $d_{\parallel}$ against a) predictions from the RF and b) the KM for noise-free data with correlation coefficients shown

3 Results

Figure 1 shows scatter plots of f, α, τ_i and $d_{\parallel}$ against predictions from a) the RF regressor and b) the KM in the case of noise-free simulated data. The correlations between ground truth and RF predictions are strong for all parameters. Even so, we do not get a perfect recovery, in part because of the statistical nature of the model, but also because parameters such as τ_i and α have a very weak influence on the DW MR signals and thus the features we derive from them. However, this provides an indication of the best predictions we can make given the measurements we have. For the KM, we observe good correlations for f and $d_{\parallel}$, but virtually no sensitivity to α. The protocol was optimised for sensitivity to τ_i rather than α, so this is in line with expectations. The KM can estimate τ_i for residence times less than 200 ms as, due to the long Δ used in the protocol, the intra- and extra-axonal compartments appear well mixed at this timescale for smaller τ_i; beyond this, the assumption breaks down and the KM severely underestimates large τ_i.

Figure 2 shows similar scatter plots, but for noisy data. Predictions from the RF are noisier, resulting in lower correlations with the ground truth values. Estimates of $d_{\parallel}$ are still strong, but large f in particular are consistently underestimated. There is still a positive correlation between ground truth and predicted values of τ_i until ≈400-500 ms, after which the estimates level off. This is because the small differences in the features due to τ_i are now overshadowed by the differences in the features due to noise. However, the RF is still able to distinguish short and long τ_i, unlike the KM, which is the key requirement for a useful imaging biomarker. The correlation for α is again much lower, although there still a slight positive trend indicating that the RF may be able to distinguish between large and small α. The KM generally provides good estimates of f and $d_{\parallel}$ but has almost no sensitivity to α and τ_i. Given that the KM fails to

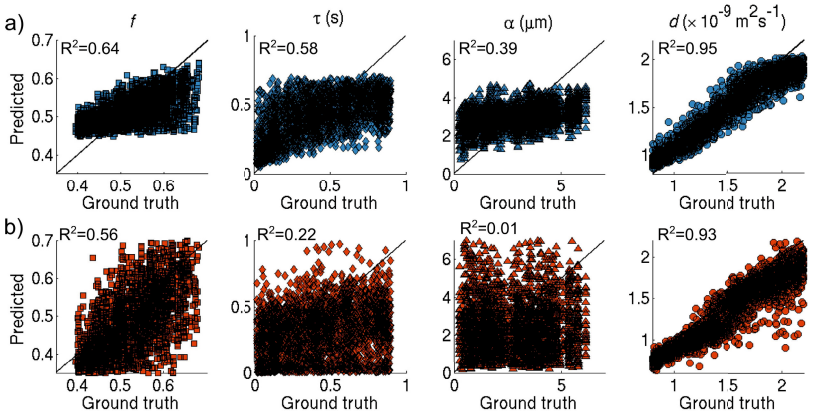

Fig. 2. Scatter plots of ground truth values of f, τ, α, $d_{\|}$ against a) predictions from the RF and b) the KM for noisy data with SNR=20, with correlation coefficients shown

estimate these parameters from data generated using a simple geometric model and realistic SNR, it is unlikely to be sensitive to these parameters in real brain tissue, which is much more complex.

We note that there is bias in this comparison, as the test data for the RF is generated in the same manner as the training data, which is not the case for the KM. However, it clearly illustrates the limitations of the KM's assumptions in situations where they are known to be violated, as in WM.

Figure 3 shows in vivo estimates of human WM microstructure parameters using a) the RF and b) the KM across the same sagittal and axial slices. The values of $d_{\|}$ estimated by both models are consistent across the WM. Predictions of f using the RF are lower than expected for WM, ranging from 0.45-0.63, reflecting the results from the noisy simulated data. Across the mid-sagittal corpus callosum (CC), we see slightly higher f in the genu and splenium compared to the midbody [1], but the trend is not as clear as that predicted by the KM. Estimates of α and τ_i from the KM are very noisy, as expected, and show no obvious patterns across the WM. In contrast, the values of α and τ_i predicted by the RF are much less noisy and we can identify trends across WM tracts. For example, across the mid-sagittal CC, we see the characteristic low-high-low trend in α [1]. Estimates of τ_i are consistently in the range 400-600 ms across the genu and splenium of the CC, and slightly lower in the corticospinal tract (CST), where we predict τ_i=300-500 ms. It is inherently difficult to validate these values as accurate estimates of τ_i are not obtainable via histology. However, a study of intra- to extra-axonal water exchange across the whole in vivo rat brain using relaxometry and contrast agents suggests a mean τ_i of approximately 550 ms [11]. This is similar to the values predicted here, although how well these numbers correspond to human tissue is unknown.

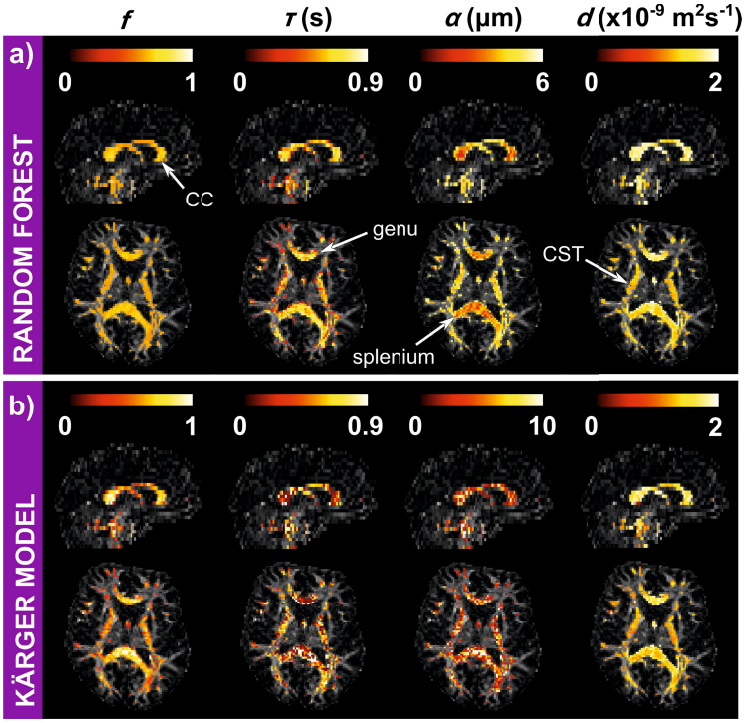

Fig. 3. Predicted values of f, τ, α, $d_{\|}$ from in vivo human brain WM across slice x=66, $z = 15$ using a) the RF regressor and b) the KM

4 Discussion

This study demonstrates that we can learn a mapping between microstructure parameters and simulated DW MR signals using RF regression, even when the data is noisy. The parameter correlations, particularly for τ_i and α, are higher for the RF than the KM, even though acquisition protocol we use was optimised for the latter approach. Furthermore, the trained RF predicts sensible microstructure parameters from in vivo human data, even for parameters such as α and τ_i which only weakly influence the DW MR signals available from human scanners. The model gives a way of obtaining plausible, if noisy, estimates of τ_i in vivo for the first time. Given the mathematical difficulty of deriving accurate analytical models of permeability, this approach is very promising; however further work and validation is needed.

Although the mapping we learn here is specifically for randomly packed, parallel, non-abutting cylinders and a STE imaging sequence, the approach can be easily extended to other tissue configurations and pulse sequences. In future we plan to incorporate models of fibre dispersion into our MC simulations, e.g. by using undulating cylinders, allowing us to extend our technique to dispersed white matter fibre regions as well as grey matter. We also plan to investigate

more specialised pulse sequences, such as the AXR sequence [9], which may improve sensitivity as well as allowing us to compare other analytic models.

The protocol used here was optimised specifically for the KM. It is therefore unlikely that this protocol is optimal for our approach, especially as some of the DW shells do not have high enough angular resolution to support feature calculation. Improving the angular resolution of the data, adding a longer diffusion time (subject to SNR constraints) and increasing $|G|$ should improve the RF predictions, particularly of τ_i and α.

Acknowledgments. This work is supported by EPSRC grant EP/I027084/1.

References

1. Aboitiz, F., Scheibel, A.B., Fisher, R.S., Zaidel, E.: Fiber composition of the human corpus callosum. Brain Res. 598, 143–153 (1992)
2. Alexander, D.C.: A general framework for experiment design in diffusion MRI and its application in measuring direct tissue-microstructure features. Mag. Res. Med. 60, 439–448 (2008)
3. Alexander, D.C., Hubbard, P.L., Hall, M.G., Moore, E.A., Ptito, M., Parker, G.J., Dyrby, T.B.: Orientationally invariant indices of axon diameter and density from diffusion MRI. Neuroimage 52, 1374–1389 (2010)
4. Assaf, Y., Blumenfeld-Katzir, T., Yovel, Y., Basser, P.J.: AxCaliber: a method for measuring axon diameter distribution from diffusion MRI. Mag. Res. Med. 59, 1347–1354 (2008)
5. Criminisi, A., Shotton, J., Konukoglu, E.: Decision Forests for Classification, Regression, Density Estimation, Manifold Learning and Semi-Supervised Learning. Microsoft Research technical report (2011)
6. Hall, M.G., Alexander, D.C.: Convergence and parameter choice for Monte-Carlo simulations of diusion MRI. IEEE Trans. Med. Im. 28, 1354–1364 (2009)
7. Kärger, J., Pfeifer, H., Wilfried, H.: Principles and application of self-diffusion measurements by nuclear magnetic resonance. Adv. Mag. Res. 12, 1–89 (1988)
8. Nilsson, M., Alerstam, E., Wirestam, R., Ståhlberg, F., Brockstedt, S., Lätt, J.: Evaluating the accuracy and precision of a two-compartment Kärger model using Monte Carlo simulations. J. Mag. Res. 206, 59–67 (2010)
9. Nilsson, M., Lätt, J., van Westen, D., Brockstedt, S., Lasič, S., Ståhlberg, F., Topgaard, D.: Noninvasive mapping of water diffusional exchange in the human brain using filter-exchange imaging. Mag. Res. Med. 69, 1573–1581 (2013)
10. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., Pedregosa, F., et al.: Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. 12, 2825–2830 (2011)
11. Quirk, J.D., Bretthorst, G.L., Duong, T.Q., Snyder, A.Z., Springer, C.S., Ackerman, J.J.H., Neil, J.J.: Equilibrium water exchange between the intra- and extracellular spaces of mammalian brain. Mag. Res. Med. 50, 493–499 (2003)
12. Stanisz, G.J., Wright, G.A., Henkelman, R.M., Szafer, A.: An analytical model of restricted diffusion in bovine optic nerve. Mag. Res. Med. 37, 103–111 (1997)
13. Westin, C.-F., Maier, S.E., Mamata, H., Nabavi, A., Jolesz, F.A., Kikinis, R.: Processing and visualization for diffusion tensor MRI. Med. Im. Analysis 6, 93–108 (2002)