

# Microstructural Parameter Estimation of Skeletal Muscle Using Random Forest Model of dMRI

Noel M. Naughton<sup>1</sup>, Nicolas R. Gallo<sup>2</sup>, Aaron T. Anderson<sup>3,4</sup>, and John J. Georgiadis<sup>1,2</sup>

<sup>1</sup>Mechanical Science and Engineering, University of Illinois at Urbana Champaign, Urbana, IL, United States, <sup>2</sup>Biomedical Engineering, Illinois Institute of Technology, Chicago, IL, United States, <sup>3</sup>Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana Champaign, Urbana, IL, United States, <sup>4</sup>Radiology, Carle Foundation Hospital, Urbana, IL, United States

## Synopsis

Results are presented for a random forest model to estimate skeletal muscle microstructure parameters from dMRI data. The model exhibits the ability to estimate microstructural parameters from both numerically simulated and experimentally acquired dMRI data suggesting that random forests, and machine learning more generally, may be a useful tool in dMRI microstructure estimation of skeletal muscle.

## Introduction

Employing diffusion MRI (dMRI) to estimate microstructure of skeletal muscle in vivo requires an accurate model of dMRI's relationship to the muscle microstructure. Current models often reduce the number of microstructural parameters to allow for more stable fitting, for example by discarding the extracellular matrix. However, extracellular properties play an important role in how force is transmitted from the individual myocyte to the tendon during activation.<sup>1,2</sup> Machine learning (ML) techniques allow estimation of structural parameters by training models with numerically simulated dMRI data and have shown success in estimating myelin permeability in white matter.<sup>3</sup> Combining our previous work on dMRI simulations<sup>4</sup> with ML techniques may enable estimations of microstructural properties with accuracy not available from current models. Here preliminary results are presented for the development of a random forest model for skeletal muscle microstructure estimation from dMRI measurements.

## Methods

Using a numerical model based on a Lattice Boltzmann Method solution of the Bloch-Torrey equation,<sup>4</sup> dMRI signals were computed on the San Diego Supercomputing Center's Comet cluster<sup>5</sup> over a domain of periodically packed hexagonal cells surrounded by a permeable membrane and embedded in an extracellular matrix. The domain was parameterized by intra- and extracellular diffusion, cell diameter, membrane permeability, and volume fraction which were sampled with a Sobol sequence<sup>6</sup> to define 1600 parameter combinations (Table 1). PGSE sequences were simulated for 5 b-values (500, 750, 1000, 1500, and 2000 s/mm<sup>2</sup>). At each b-value 5 diffusion times ( $\Delta$ ) were simulated (10, 25, 50, 75, and 100 ms) with a gradient timing of 7.5 ms and the gradient strength adjusted to maintain a constant b-value. Six gradient directions were used to estimate a diffusion tensor from which fractional anisotropy (FA), apparent diffusion coefficient (ADC) and the tensor's eigenvalues ( $\lambda_1, \lambda_2, \lambda_3$ ) were extracted, leading to 125 features for the model. This dataset was used to train a random forest (RF) model using the open source scikit-learn package in python.<sup>7</sup> For testing data, simulations were performed with a single microstructural parameter varied and all other parameters held constant. Additionally, a data set using a Sobol sequence to generate 320 parameter combinations was constructed. Both testing sets were compared with the RF model for noise-free signal and with added Gaussian noise (SNR=40). Finally, the RF model was used to estimate microstructural parameters from an ex-vivo porcine myocardium, used as a substitute for skeletal muscle due to its similar microstructure, and imaged with a Siemens 3T TRIO scanner using a monopolar diffusion-weighted sequence with b-values of 800, 1500 and 2500 s/mm<sup>2</sup>, TR/TE/ $\Delta/\delta$  = 2930/94/44.35/22.93 ms, 2x2x2 mm resolution and 64 gradient directions. Because only three b-values at a single diffusion time were acquired, the model was retrained from the training data for only these three sequences (requiring additional simulations for b=2500 s/mm<sup>2</sup>).

## Results and Discussion

The RF model is compared with the two testing data sets in Figures 1 and 2. Figure 1 compares how well the RF model tracks changes in particular microstructural parameters. Figure 2 shows how the RF model compares with the Sobol sequence test data. Figure 3 shows the RF model's microstructural estimations for the ex-vivo porcine phantom. The RF model was trained with 100 trees as it was found that more than 100 trees did not increase its ability to explain parameter variation (Table 2). The RF model is most accurate for estimating cell diameter and intracellular diffusion, which are also the parameters current models are best able to estimate. The model struggles to estimate volume fraction, likely due to low contrast between the intra- and extracellular diffusion. This RF model predicts parameters from 25 different dMRI sequences, however, using only three b-values at one diffusion time, the model can estimate tissue parameters from an experimental phantom demonstrating that the model can generalize beyond numerical data, an important first step. Future work is required to validate these predictions with histology and to reduce the number of sequences used to generate features.

## Conclusion

We have developed a random forest model to predict microstructural parameters of skeletal muscle from dMRI signal. The model has reasonable accuracy compared with numerically simulated data and provides physiologically reasonable values when applied to ex-vivo porcine heart. These results show that RF models are capable of predicting microstructural variations and that the model may be generalizable beyond the numerical data used to train it. More work is needed to accurately extract extracellular structural parameters, but random forest models appear to be a promising method of microstructure estimation in skeletal muscle.

## Acknowledgements

Funding for this work was provided by NSF Grant CMMI-1437113 and NSF Graduate Research Fellowship for NMN. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562 and provided access to the SDSC Comet Cluster under allocation #TG-MCB180044.

## References

[1] Gillies, A. R., & Lieber, R. L. (2011). Structure and function of the skeletal muscle extracellular matrix. *Muscle & nerve*, 44(3), 318-331.

[2] Purslow, P. P. (2002). The structure and functional significance of variations in the connective tissue within muscle. *Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology*, 133(4), 947–966.

[3] Nedjati-Gilani, G. L., Schneider, T., Hall, M. G., et al. (2017). Machine learning based compartment models with permeability for white matter microstructure imaging. *NeuroImage*, 150, 119-135.

[4] Naughton, N. M., Tennyson, C. G., & Georgiadis, J. G., Lattice Boltzmann method for simulation of diffusion magnetic resonance imaging physics in heterogeneous tissue models. *Journal of Computational Physics*, (submitted 2018)

[5] Towns, J., Cockerill, T., Dahan, M., Foster, I., et al. (2014). XSEDE: accelerating scientific discovery. *Computing in Science & Engineering*, 16(5), 62-74.

[6] Herman, J. and Usher, W. (2017) SALib: An open-source Python library for sensitivity analysis. *Journal of Open Source Software*, 2(9).

[7] Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011). Scikit-learn: Machine learning in Python. *Journal of machine learning research*, 12(Oct), 2825-2830.

### Figures

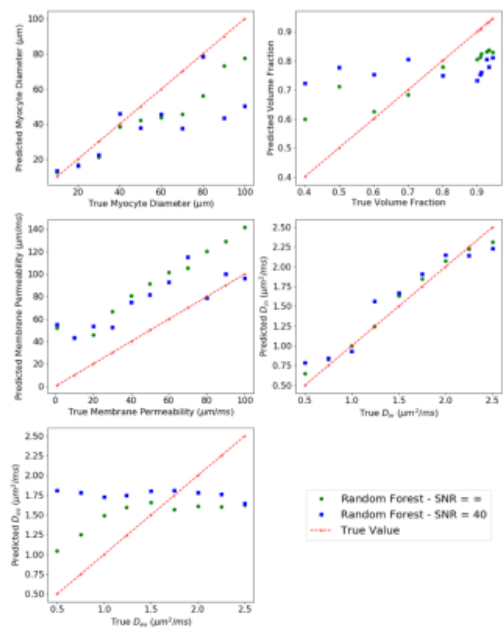


Figure 1. Comparison of random forest model's microstructural parameter prediction for noise-free simulations and with added SNR of 40. The random forest model is less sensitive to parameter changes related to extracellular organization such as volume fraction or extracellular diffusion. Diameter is underestimated for larger myocyte sizes and permeability is estimated with an offset. When noise is added to the data all parameters fits are affected with packing fraction and extracellular diffusion the most affected.

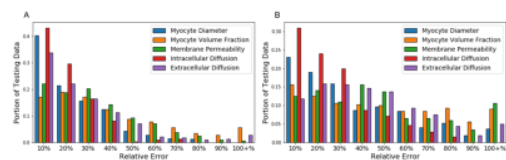


Figure 2: Histogram of random forest model error compared with testing data set generated from a Sobol sequence for A) noise-free data and B) added SNR of 40. The random forest model's fit is worse when noise is added. Overall accuracy of the RF model for each parameter is similar to the results for Figure 1 with diameter and intracellular diffusion most accurate and volume fraction, permeability and extracellular diffusion the least accurate.

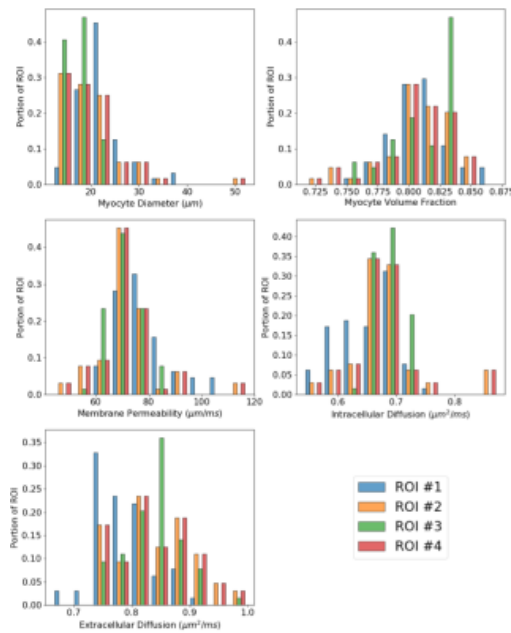


Figure 3. Histogram of results for random forest model's microstructural estimates for porcine myocardium for four ROI's (4x4x4 voxels) manually selected from the center of 2 different porcine hearts. Ranges of the fitted parameters are substantially smaller than the parameter ranges used for the training data and are of values expected for ex-vivo porcine myocardium (volume fraction ~ 0.80, larger permeability from membrane degradation).

Parameter	Parameter Range	Units
Intracellular Diffusion ( $D_{in}$ )	0.50 – 3.0	$\mu\text{m}^2/\text{ms}$
Intracellular Diffusion ( $D_{ex}$ )	0.50 – 3.0	$\mu\text{m}^2/\text{ms}$
Myocyte Diameter	10 – 100	$\mu\text{m}$
Membrane Permeability	1.0 – 150	$\mu\text{m}/\text{s}$
Myocyte Volume Fraction	0.40 – 0.95	

Table 1: Parameter ranges of the Sobol sequence used to generate training and testing data sets. Ranges were extended to include physiologically typical parameter values as well as values beyond these ranges in order not to constrain the random forest model.

$R^2$	Myocyte Diameter	Myocyte Volume Fraction	Membrane Permeability	Intracellular Diffusion	Extracellular Diffusion
10 trees	0.7951	0.3866	0.6419	0.8799	0.6645
100 trees	0.8323	0.4362	0.6673	0.8977	0.6840
1000 trees	0.8263	0.4358	0.6699	0.8979	0.6898

Table 2:  $R^2$  coefficients of random forest model for increasing number of trees showing a diminished increase in the predictive power of the random forest model for more than 100 trees.  $R^2$  coefficients also show that the random forest model is most accurate for cell diameter and intracellular diffusion, a result also illustrated in Figure 3.