Fully Bayesian Multi-Model Inference for Parameter Estimation in DCE-MRI

Tammo Rukat[1] and Stefan A. Reinsberg[2]

[1]University of Oxford, [2]University of British Columbia
Contact: tammo.rukat@kellogg.ox.ac.uk

http://tammor.github.io/

Motivation

- The **reliability of physiological parameters** that are estimated from DCE-MRI experiments strongly depends on the choice of the analysis model.
- We have shown before that the simple Akaike information criterion facilitates the improvement of parameter estimates, in particular between the simple Tofts model and the extended Tofts model [**Rukat2014a**]. Others reported the lack of a systematic benefit [Luypaert et al. 2012].
- Here we present promising results for weighted across-model averages, using the Watanabe-Akaike information criterion (WAIC) based on the **full posterior likelihood of the model**.

Procedure

Data simulation with a simple analysis model: 2CXM

- The ground truth tissue response is given by the two-compartment exchange model (2CXM, 4 free parameters).
- This model is also used fo fitting in the subsequent analysis. This allows to isolate the effect of overfitting the data despite the use of a fully accurate model.
- We simulate the following parameter combinations:
  K^{trans} [1/min] = 0.25, 0.45
  plasma flow [ml/g/min] = 0.5, 1.5, 2.5
  fractional plasma volume = 0.02, 0.08
  fractional interstitial volume = 0.45

Data simulation, approximating tissue physiology: MMID4

- The Multiple Path, Multiple Tracer, Indicator Dilution, 4 region model (MMID4) [Kroll and Wilke 1996] provides the ground truth tissue response, accounting for tissue heterogeneity and flow dispersion.
- It resembles tissue physiology more closely than any of the analysis models and as such mimics an experimental scenario.
- We simulate the following parameter combinations:
  K^{trans} [1/min] = 0.16, 0.18, 0.35, 0.47
  plasma flow [ml/g/min] = 0.5, 1.2
  fractional plasma volume = 0.001, 0.05
  fractional interstitial volume = 0.4.

Mimicking experimental constraints

- White noise (SNR=20, if not mentioned otherwise) is added to the signal with due regard to the FLASH signal equation.
- Sampling intervals of 4s (2CXM) and 2.5s (MMID4) are chosen.
- If not mentioned otherwise the measurement time is set to 100s.
- The tissue response is simulated based on typical arterial input function in the cephalic vein [Parker et al. 2006] and in a mouse tail [Moroz et al. 2012].

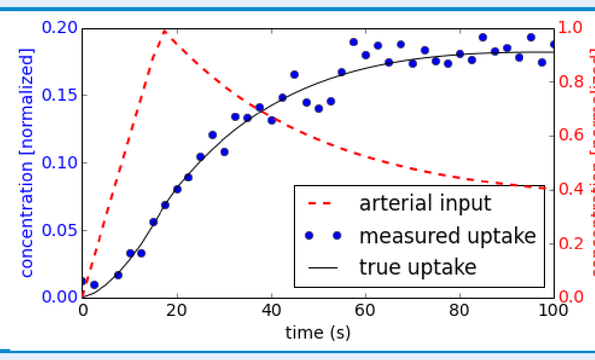

Figure 1: AIF and typical tissue response (here for the 2CXM) with white noise.

Iterate (10x)

MCMC to approximate the parameter posterior

- We fit the two-compartment exchange model (**2CXM**, 4 free parameters) and the extended Tofts model (**ETM**, 3).
- Posterior distributions of the model likelihood and the fit parameters are approximated by Markov chain Monte Carlo (MCMC) sampling within the Stan Modeling Language [Stan Dev. Team 2014].
- Convergence is assured by sampling until the scale reduction statistic ($\hat{R}$) is smaller 1.05 ($\hat{R} = 1.00$ in most cases).

Calculation of the weigthed across-model average

- The Watanabe-Akaike information criterion (**WAIC**) is based on the posterior predictive density [Watanabe 2012] and approximates the out-of-sample prediction error [**Gelman2014b**].
- Maximum likelihood estimates of the parameters are averaged across different models, weighted by their WAIC score [Burnham and Anderson 2004].

Results

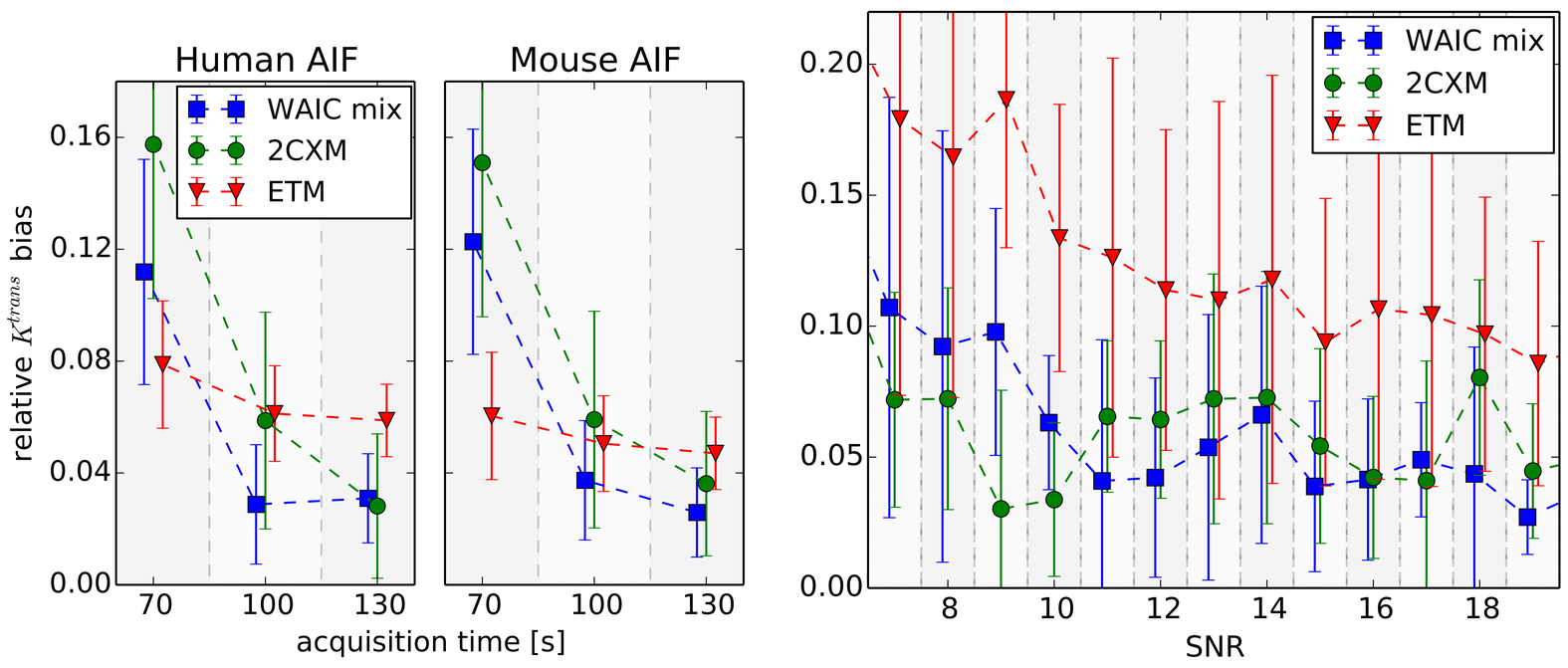

Figure 2: Mean relative K^{trans} deviation from the true value. Error bars show the mean SD of the parameter posterior distributions. **Left**: averaged over eight MMID4 parameter sets under variation of the acquistion time for a human and a mouse AIF (SNR=20). **Right**: averaged over twelve 2CXM parameter sets under variation of the SNR for a human AIF (acq. time = 100s). x-jitter for visibility.

- In the regime of intermediate and high SNR/acqusition time WAIC-weighted K^{trans} averages almost always increase the parameter accuracy.
- WAIC-weighted K^{trans} averages perform even better than the model that was used to simulate the data and are thus shown to be successfull in weighing models based on their sole complexity.
- Only for data of low SNR/acquision time do the simpler model or the known true model outperform WAIC-weighted K^{trans} estimates.

References

Burnham, Kenneth P. and David R. Anderson (2004). In: *Sociological Methods & Research* 33.2, pp. 261–304.

Kroll, K. and N. Wilke (1996). In: *American Journal of Physiology* 271, H1634–H1655.

Luypaert, Robert et al. (2012). In: *Physics in Medicine and Biology* 57.11, pp. 3609–28.

Moroz, Jen et al. (2012). In: *Proc. Intl. Soc. Mag. Reson. Med.* Vol. 20. University of British Columbia, p. 239.

Parker, Geoff J. M. et al. (2006). In: *Magnetic Resonance in Medicine* 56.5, pp. 993–1000.

Stan Dev. Team (2014). Stan: A C++ Library for Probability and Sampling, Version 2.5.0.

Watanabe, Sumio (2012). In: *arXiv:1208.6338 [cs, stat]*.

Conclusion

WAIC-weighted estimates improve K^{trans} accuracy for the two-compartment exchange model and the extended Tofts model in the regime of a reasonably high SNR and acquisition time. This holds true, even in the unrealistic case of a known true tissue response model.

In most realistic scenarios, we expect the demonstrated procedure to perform better than any fixed model choice.