Open Access MAGNETIC RESONANCE Discussions

Magn. Reson. Discuss., referee comment RC1
https://doi.org/10.5194/mr-2021-28-RC1, 2021

Comment on mr-2021-28

Anonymous Referee #1

Referee comment on "Bootstrap aggregation for model selection in the model-free formalism" by Timothy Crawley and Arthur G. Palmer III, Magn. Reson. Discuss., https://doi.org/10.5194/mr-2021-28-RC1, 2021

This is a very timely and impactful paper describing an insightful approach to model-free fitting of relaxation data. I predict that the proposed approach will be hugely influential and likely replace the seminal model-selection protocol published by the Palmer group in 1995, which has some 1000+ citations.

Here Crawley and Palmer implement bootstrap aggregation as a novel way to perform model-free fitting, taking into consideration the specific demands imposed by NMR relaxation data sets, where individual data points (i.e. J(w) sampled at specific w values) have very different impact on the fitted model.

The paper is comprehensive and easy to follow. Well-chosen examples are presented highlighting the advantages of the bootstrap aggregation approach. Nonetheless, I have a few suggestions for slight improvements in presentation and two questions.

- Could you please comment on the number of field strengths required to successfully implement bootstrap aggregation. Assuming I get the math right, using 3 field strengths, the total number of possible samples is 7^3= 343, which seems to be at the low end of the number of samples normally employed, but perhaps this would suffice?
- How do you determine the correlation time for overall rotational diffusion, tau_m (in the case of a globular protein)? The suggested protocol fits tau_m individually for each residue. Do you forego the concept of fitting a global tau_m as part of the MF fits? And subsequently fit a rotational diffusion model (isotropic or anisotropic) to these individual values (while taking into account the orientation of the HN bond vectors in the molecular frame in the case of anistropic models)?

Minor points:

line 19, Suggestion: spell out Akaike and Bayesian Information Criterion when introducing AIC and BIC.

l. 145-148, the mean J(0) is presumably only used in the conventional MF protocol with MC error analysis(?) This should be stated here to avoid confusion.

The bootstrap aggregation protocol is well described on p. 6, but I still feel that it might be beneficial to include flow-chart type figure outlining the construction of the bootstrap sample datasets.

The tables are not easily interpreted without referring back to the text. Please add footnotes to define p_ij and Y_ij in words (Table 1). Please add text to indicate that "Smooth" refers to the percentage of selected models in Tables 2-4.

Typos: "paramaters" (l. 86); "interogating conformional" (l. 250)