Open Access MAGNETIC RESONANCE Discussions

Magn. Reson. Discuss., referee comment RC1
https://doi.org/10.5194/mr-2022-21-RC1, 2022

Comment on mr-2022-21

Anonymous Referee #1

Referee comment on "DEEP Picker1D and Voigt Fitter1D: a versatile tool set for the automated quantitative spectral deconvolution of complex 1D-NMR spectra" by Da-Wei Li et al., Magn. Reson. Discuss., https://doi.org/10.5194/mr-2022-21-RC1, 2022

This paper introduces a tool for the analysis of 1D NMR spectra of complex mixtures. It is based on a concept which has been recently published by the same authors in the case of 2D NMR. The method is based on machine learning and deep neural networks. It includes peak picking, fitting and deconvolution, and offers an impressive performance in the case of highly overlapped peaks. In particular, it seems very promising for metabolomics applications, and the code availability should allow a large community of users to test it on a variety of matrices (please do not make it a commercial tool!)

Compared to the previous 2D NMR paper, the present article gives more details, and it also provides a critical demonstration, analysis and discussion of the method's performance. It is suitable for publication and I have only minor comments below.

-page 6: in addition to the manufacturer-dependent pulse sequence names, please add a comprehensive name (such as "water suppression using excitation sculpting with a perfect echo")

-page 9: "Peak parameters, such as peak position, peak height and peak volume can then be directly used for downstream analysis, such as compound identification and quantitative NMR applications" --> this sentence gives the feeling that with the authors' tool, anyone can retrieve quantitative data from signals...this is only true if data have been acquired in quantitative conditions. This may sound obvious for most readers, but I would still suggest modifying the sentence as follows (or something similar): "Peak parameters, such as peak position, peak height and peak volume can then be directly used for downstream analysis, such as compound identification and quantitative NMR applications -when incorporated into a quantitative NMR workflow"

-page 13: "It can be hard to assess the deconvolution accuracy of experimental spectra,

since the ground truth, i.e. the individual isolated peaks and their parameters, are often unknown." --> in fact this is not so difficult, relying on synthetic mixtures of metabolites (for instance synthetic urine or plasma) where concentrations are known. Such samples can be prepared in the lab with accurate gravimetric measurements, or even commercial. Since this is a methodological paper, I am not requesting the authors to perform such additional experiments, but that will be important for future papers relying on this tool (in particular when it comes to analytical chemistry considerations). Here I would just suggest to moderate the beginning of the paragraph by removing the first sentence: "To assess the deconvolution accuracy of our tool, we constructed experimental spectra with overlaps..."

-in the discussion, could the authors add a comment on how the method is expected to perform on a series of samples from a metabolomics cohort, and how it would account from peak shifting (due to pH variations, etc) across a metabolomics cohort?