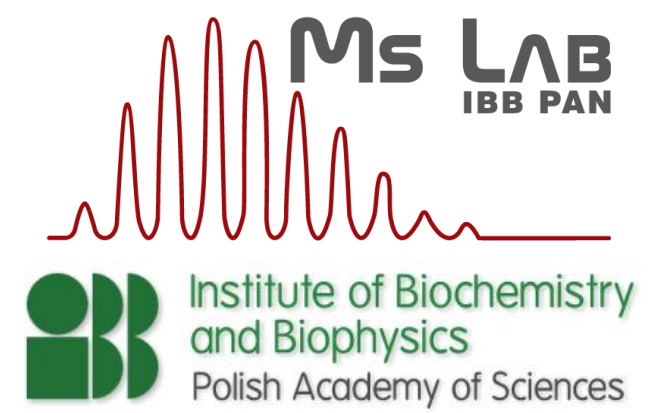




Just give me a p-value: statistical analysis of deuterium uptake curves in HDX-MS

Michał Burdukiewicz^{1,*}, Mateusz Staniak², Weronik Puchała³, Dominik Cysewski³, Michał Dadlez³
*michal.burdukiewicz@gmail.com

²Faculty of Mathematics and Information Science, Warsaw University of Technology, Poland, ³Institute of Mathematics, University of Wrocław, Poland, ¹Institute of Biochemistry and Biophysics Polish Academy of Sciences, Poland.



Introduction

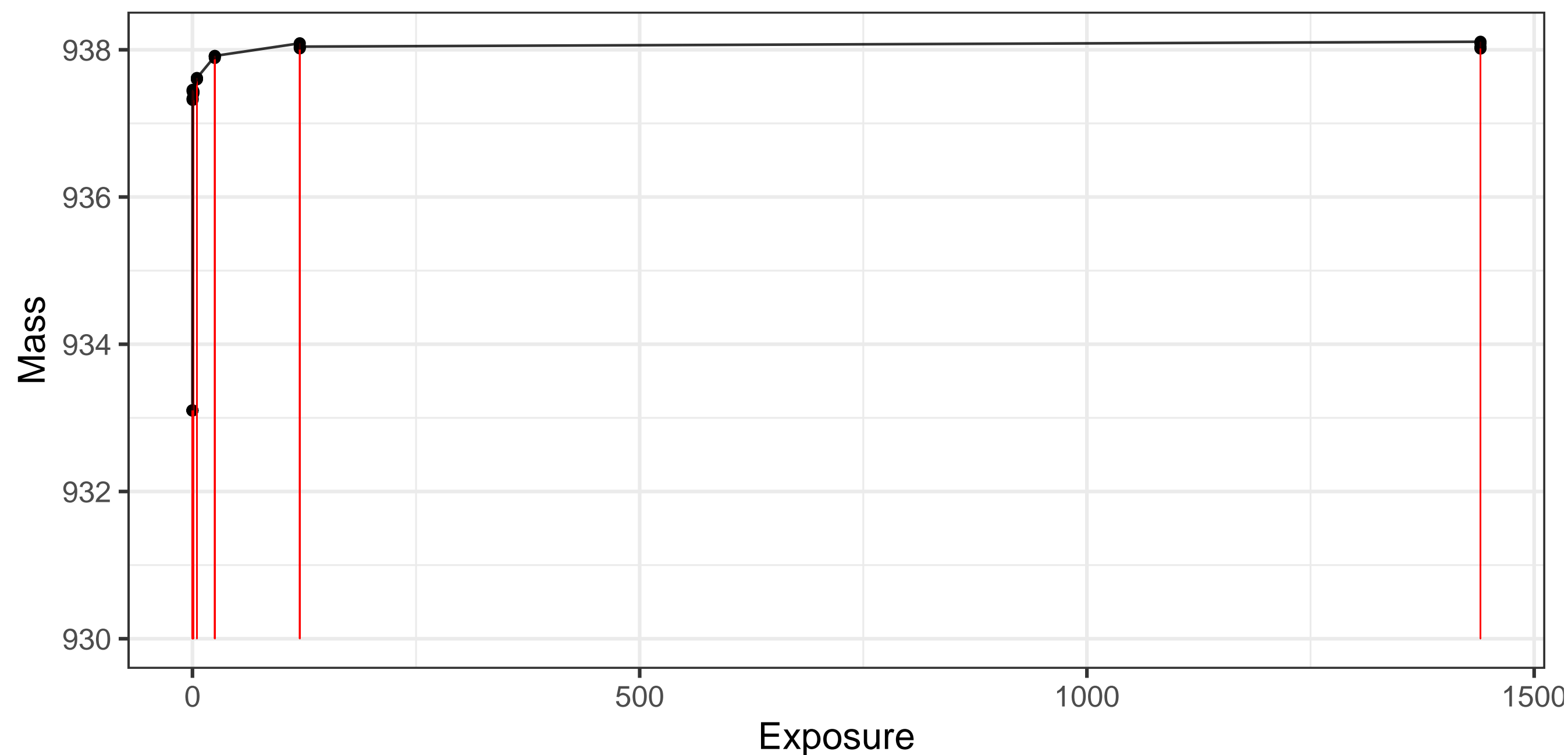
The misuse of statistical testing is one of the pitfalls of contemporary science (Krueger and Heck, 2019). It is especially prevalent in rapidly growing fields as the HDX-MS, where the development of experimental techniques precedes the formalization of good practices in data analysis.

Methods

We have proposed and investigated 30 statistical models on several HDX-MS datasets, partially based on models proposed elsewhere (Hourdel et al., 2011; Hourdel et al., 2016). As the theoretical model of deuterium uptake distributions is still unclear, we compare our models to empirical measures.

We use the following measures:

- Mean Absolute Error (mean length of red lines in the Figure)
- Squared Error (mean square length of red lines in the Figure)
- Area Under the Curve (sum of areas of trapezoids between the red lines in the Figure)



Houde's test

Houde's test is based on the following confidence interval:

$$0 + / - \left(\sqrt{\sum_{t=1}^n \frac{\sigma_t^2}{N_t}} \right) \times \alpha$$

Linear Models

In general, a test based on a linear model is derived from a model

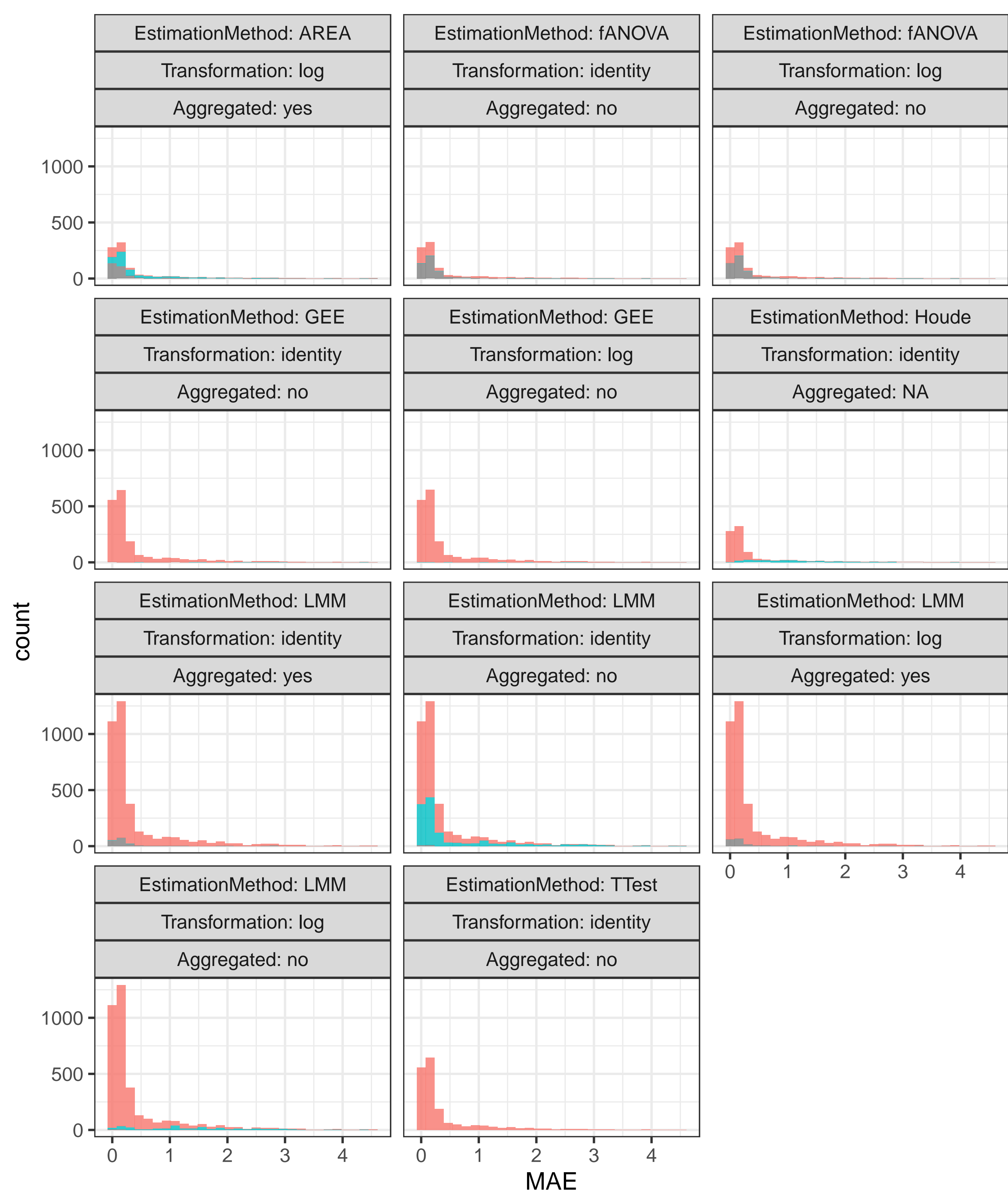
$$Y = \beta_{Time} X_{Time} + \beta_{State} X_{State} + \beta_{Time, State} X_{Time} X_{State} + \varepsilon$$

Where ε denotes a random error. The time variable X_{Time} can be transformed. We will compare models with untransformed and log-transformed time. Additionally, the model can include a term that corresponds to replications within a state. In this case, the following model formula is used.

$$Y = \beta_{Time} X_{Time} + \beta_{State} X_{State} + \beta_{Time, State} X_{Time} X_{State} + \beta_{Rep} X_{rep} + \varepsilon$$

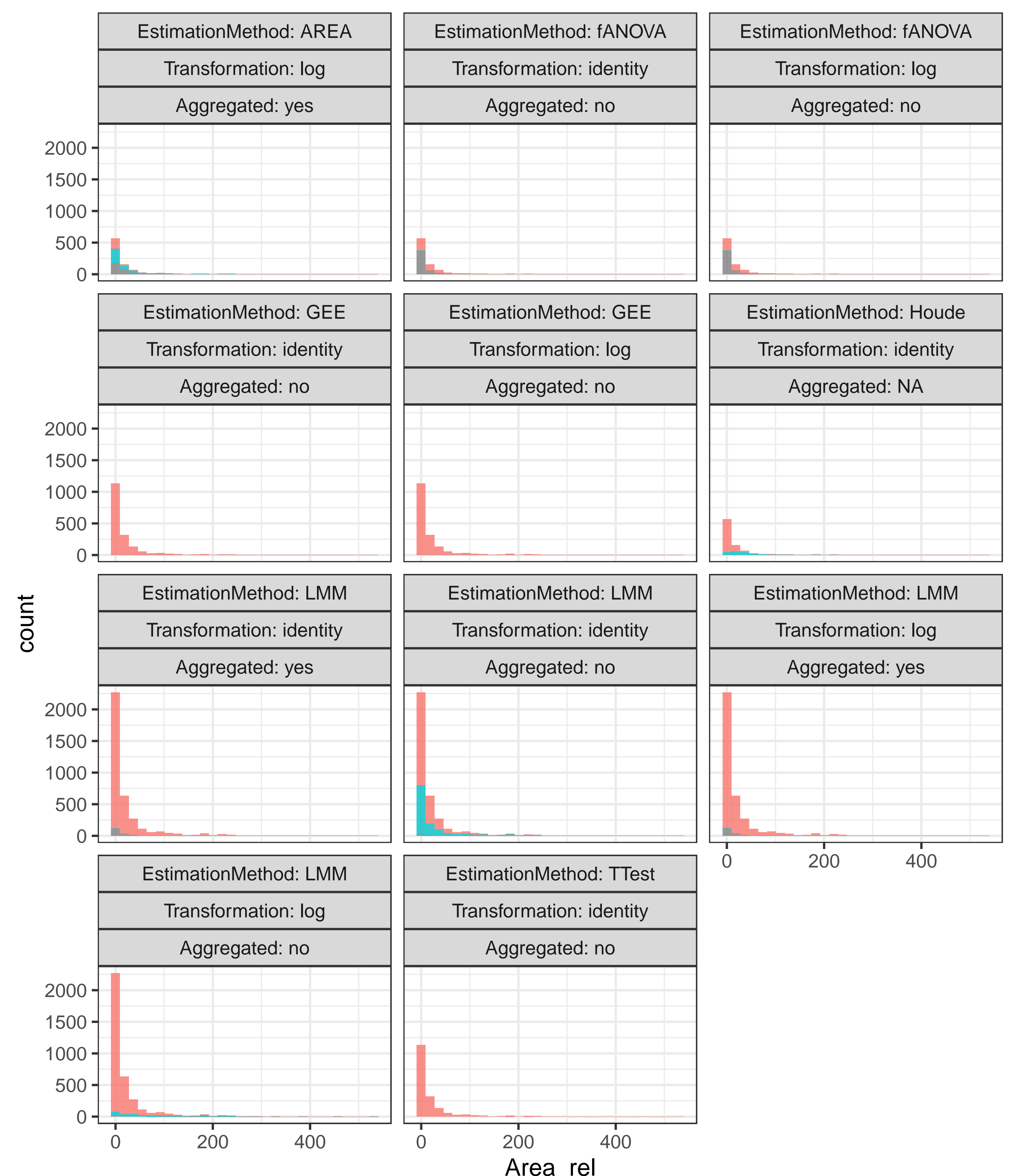
The coefficient β_{Rep} can be treated as either a fixed or a random effect. We test both of these options.

Mean Absolute Error



Decision ■ FALSE ■ TRUE ■ NA

Area Under the Curve

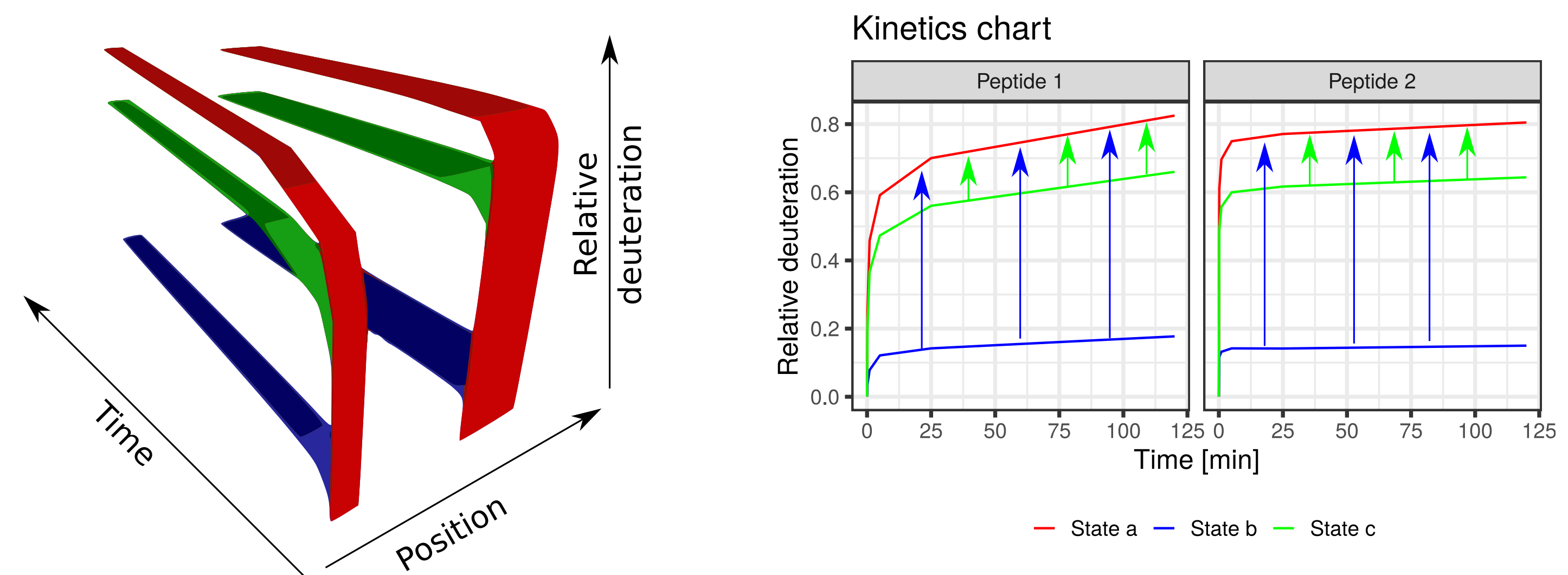


Decision ■ FALSE ■ TRUE ■ NA

We point out that tests based on non-aggregated deuterium uptake are more sensitive, especially to the differences of shapes of deuterium uptake curves. In consequence, such tests can find not only significant differences in absolute deuterium uptake but also in the speed of hydrogen-deuterium exchange.

Multi-state variability test (MSVT)

A multi-state variability test offers a statistical assessment of regional variability. A user can compare relative deuteration between multiple states in specific regions of the interest (ROI).



MSVT relies on information coming from all peptides which at least partially cover the specified ROI. It means that instead of per-peptide results, we take into account a whole regional variability, which leads to the increase of sensitivity.

Aside from these benefits, MSVT is calibrated to the area between kinetic curves. Thanks to that it offers a straightforward interpretation for experimentalists, while preserving full statistical validity.

As MSVT is not yet a part of the HaDeX software suite and is available only upon request.

Acknowledgements

The project is carried out within the TEAM TECH programme of the Foundation for Polish Science co-financed by the European Union under the European Regional Development Fund.(TEAM TECH CORE FACILITY/2016-2/2) Mass Spectrometry of Biopharmaceuticals-improved methodologies for qualitative, quantitative and structural characterization of drugs, proteinaceous drug targets and diagnostic molecules"



Bibliography

- MEMHDX: an interactive tool to expedite the statistical validation and visualization of large HDX-MS datasets. Hourdel V, Volant S, O'Brien DP, Chenal A, Chamot-Rooke J, Dillies MA, Brier S. Bioinformatics. 2016 Nov 15;32(22):3413-3419. Epub 2016 Jul 13
- The Utility of Hydrogen/Deuterium Exchange Mass Spectrometry in Biopharmaceutical Comparability Studies. Houde, D.et al.(2011). Journal of pharmaceutical sciences,100(6), 2071–2086