

Interactive comment on “Technical Note:
TimeFRAME - A Bayesian Mixing Model to
Unravel Isotopic Data and Quantify Trace Gas
Production and Consumption Pathways for
Timeseries Data” by Eliza Harris et al.

This manuscript by Harris et al. documents and aims to validate a useful software package called “TimeFRAME”. TimeFRAME can be used to analyse production, mixing and loss of trace gas compounds using isotope composition time series. It builds on the existing “FRAME” with 4 Bayesian hierarchical models. TimeFRAME provides a relevant and useful package to the scientific community, within the scope of biogeosciences. Although I believe the package can be useful, I see a major issue with the validation strategy of TimeFRAME as a discuss below. This needs to be addressed before considering this for publication. Additionally, I currently do not possess the expertise to fully evaluate the correctness of section 2.3. The manuscripts, written in clear language, is a bit long and at times somewhat unstructured, for which I have some remarks below as well.

A major issue I had after reading the manuscript is the applicability of TimeFRAME due to the assumption made to arrive at equation 3 on line 45. The authors acknowledge already that in this equation it is assumed that mixing and fractionation are separable. In addition, they acknowledge that this is an unlikely scenario, and it would very much depend on the use case if it is applicable. It is well known, that for most applications, when one or more loss terms is involved, accurate source apportionment becomes complicated (see for example Kaiser et al., 2006; Röckmann et al., 2011) and usually requires time resolved 2D or 3D modelling. Yet, the authors conclude that consumption pathways can be estimated using TimeFRAME. For this final generalized conclusion no proof is presented in this manuscript. In section 2.4.1 the data simulation experiments are explained. The data is generated by using equation 3. I.e. the forward model for data already has mixing and the loss process separated. It is thus expected that, given small uncertainty you retrieve back the inputs. So consumption pathways can be estimated only for the case where mixing and loss processes can be separated. For most applications, which the authors acknowledge (and there I fully agree), this is unlikely. So the experiment does

not support the general conclusion that TimeFRAME can quantify mixing and loss in general. I'm sceptical that it can, as soon as the leading assumption is violated. Moreover, this significantly reduces the impact of this package for the cases where it is validated. Then I wonder, with the current validation: to what extent TimeFRAME is different from the other packages mentioned by the authors (MixSIAR, simmr)?

Why did the authors not use a model that does consider mixing and loss at the same time to use for data generation? With a limited number of parameters, a system of differential equations can be designed to trace the relevant parameters. Then, you can see to what extent the linearised model used in TimeFRAME is capable of retrieving the relevant parameters in such a case. This would fully validate and allow the authors to quantify the circumstances under which the model is no longer applicable. Making the manuscript far more valuable. Moreover, it would present users with a methodology to come up with their own validation strategies when applying TimeFRAME to their own data.

General comments

Equations 1 and 2 (and therefore also 3 but equation 3 is discussed separately above) are approximations. See for example Chapter 4 Mook, 2000. Mook, 2000 noted that for equation 1 as presented in the manuscript by Harris et al., the induced error is indeed small. Nonetheless, as this paper is the fundamental reference for "TimeFRAME" this should at least be mentioned. Additionally, use " \approx " for equations where approximations are made to signal the user that the relation is not exact.

As mentioned I'm not fully equipped to judge the validity and correctness of section 2.3. Though I certainly see the relevance of documenting the techniques used in section 2.3, I don't understand it's purpose in the current text. I would be more interested to see a paragraph describing the strength and weaknesses of each of method. This can then be reflected on in the discussion. The technical description can be put in an appendix.

Results and Discussion section is long. The results are sometimes difficult to find between discussion paragraphs and also some parts that belong in the methodology section. Would it not be better to separate results and discussion? Some more suggestions:

1. Section 3.1 is not a result, it is a software design choice. A single paragraph in Section 2 could convey the same information: "We have settled on Stan over JAGS because of performance." Also, another column in table 3 could be used to specify runtime per experiment to give the author an idea of how long each computation takes. Then Table 2 can be removed altogether. (note I also have a question in the minor comments below on this topic)
2. Section 3.2 starts off with the need to test the prior distribution for the fraction remaining r . Shouldn't that be part of the methodology?

3. Table 3 is part of the methods. I suggest introducing the application of the different models in a separate section after 2.4.1. This section could alternatively benefit from a clear presentation of the different settings for the parameters in the model. (in other words what arguments are passed on to the function in the software package).
4. A paragraph introducing the different subsections in the results section would be very helpful to guide the reader, i.e. prior to section 3.1.

I miss a paragraph discussing the application of TimeFRAME to other trace gas isotope timeseries. They are mentioned in the introduction and abstract (CH_4 , CO_2), but not discussed, this also clearly relates to my major issue. What can the authors see about applications to other isotopes, for example atmospheric CH_4 , which has several loss terms in the Stratosphere, with temperature dependent isotope effects? The abstract seems to suggest that this software package is capable of dealing with that.

Although the technical description is certainly relevant, I miss a bit of an explanation how the user would determine the prior. For example, in Figure 2, if the user were to know a priori that at $f_1(t = 0) = 1$ how is that achieved? When should I use a Gaussian process prior and when a Dirichlet-Gaussian prior? It would be helpful to potential users, what kind of information needs to be defined per model. Throughout the text (For example L 457) recommendations on the usage of the different models can be found. It would be helpful to have a single paragraph (table?) that summarize the recommendations for different applications.

Some minor details caught my attention:

L10 here it states “..production, mixing and consumption..”, however the rest of the text considers only 2 endmember mixing with a single consumption pathway. What if there are three? or what if there are two loss processes? Please update the abstract, title, and manuscript to be in line with what is presented.

L18 Consider updating this after addressing my comment on L468.- I think an additional isotope can be very useful in cases where endmembers are significantly different.

L 39 “... substrate being consumed before consumption...” reads a bit odd, simply put “..the substrate before its consumption..”

L 44 I could not find and access Fischer 2023. So the derivation of this is missing.

L 62 MixSIAR and simmr are mentioned here, but are not reflected on in the discussion. Would it not be informative to include a comparison without a sink? Or at least reflect on those packages? For mixing-only application is there a substantial difference? Does TimeFRAME offer a unique functionality?

L 79 The remainder of the text suggest that only one fractionation factor can be used?

L 219 Please refer to the original text, as I'm not sure if Fischer 2023 is appropriate here. (potentially other places as well)

L 221 Is it necessary for this text to fully explain the DGP process. Would it not be sufficient to simply state that a Dirichlet-Gaussian process is modelled?

L 272 Why not test with three sources. How well would your model distinguish the b1 and b2 sources for oxygen-18? Also, fractionation factors are often variable, for example with a temperature dependence. There is no discussion on the applicability of this model to that specific case.

Figure 3. You have only two sources whereas Table 1 lists three. Why is there no test case with all three sources? What can you conclude about more sources?

L 304 "... is a useful metric for to evaluate ..." is bad sentence, please rephrase

L 316 and elsewhere, it is recommended that mathematical operators should be written upright Roman. See (Cohen et al., 2007) chapter 4. So the differential d in the integral should be d .

L 329 The conclusion to use Stan is based on timing tests as presented in Table 2. Yet there seems to be a fundamental difference in the sampling strategies. I assume that those don't matter for your application?

L 347 Doesn't this simply reflect the fact that, for large fractions remaining the effect on isotope measurements is simply so small that it is difficult to see? Especially in the light of the noted uncertainties in Table 1. In other words, if the fractionation constants are well know, and much larger, the change in isotope measurements is much larger and you would be able to see that. Therefore, I think it is strange to attribute this to the nature of the logarithm. In other words, you simply can't resolve r very well given the values in Table 1.

L 370 what about uncertainty in the fractionation factors?

L 420 To what extent does the described experiment here relate to the experiment described in section 3.2

L 427 This is what I meant with my comment on L 347... Those two descriptions are in conflict. This seems to be the valid analysis.

L 468 "Very little improvement is seen for the estimation of f ." I think this is to be expected. Looking at table 1, $\delta(^{18}\text{O})$ source signatures are very similar for

nitrification and denitrification with in there respective uncertainties. I think you would see improvements when adding the third source, fungal denitrification. Why didn't the authors consider that for this experiment?

L 470 Is this in any case realistic? Natural variability can put a limit on end-member uncertainty reduction through measurements.

References

- Cohen, E., T. Cvitas, J. Frey, B. Holmstrom, K. Kuchitsu, R. Marquardt, I. Mills, F. Pavese, M. Quack, J. Stohner, H. Strauss, M. Takami, and A. Thor (2007). *Quantities, Units, and Symbols in Physical Chemistry*. RSC Publishing. DOI: 10.1039/9781847557889.
- Kaiser, J., A. Engel, R. Borchers, and T. Röckmann (2006). "Probing stratospheric transport and chemistry with new balloon and aircraft observations of the meridional and vertical N₂O isotope distribution". In: *Atmospheric Chemistry and Physics* 6.11, pp. 3535–3556. DOI: 10.5194/acp-6-3535-2006.
- Mook, W. G. (2000). *Environmental isotopes in the hydrological cycle*. Vol. 1. IHP-V, Technical Documents in Hydrology. UNESCO.
- Röckmann, T., M. Brass, R. Borchers, and A. Engel (2011). "The isotopic composition of methane in the stratosphere: high-altitude balloon sample measurements". In: *Atmospheric Chemistry and Physics* 11.24, pp. 13287–13304. DOI: 10.5194/acp-11-13287-2011.