In this manuscript submitted to as special issue in paleoclimate time series analysis, the authors consider different analysis methods to estimate the H scaling index of paleoclimate series. For this they consider fractional Brownian motion simulations, with adequate modifications in order to introduce an irregular time step, and they systematically check all proposed methods. This work is well explained and the systematic work is convincing. I have some minor comments and one suggestion of significant change.

Major comment:
To understand and appreciate the results I need to see the estimated \hat{H} versus H, for H \in (-1,1). This plot is the most important. If relevant the authors may plot in the same graph the standard deviation. The bias-std plots (Figs 3,4, 5 and 7) are only providing information on the fluctuations of the estimates. The mean value is the more important information for the reader. As I see from Figure 1, it seems that the spectral method may be less accurate when H<0. This plot should help to visualize this property and any other…

We agree with the reviewer that the bias of an estimator is important, which can be depicted as the plot of estimated Ĥ versus true H. We might not have made the connection clear enough. The bias is currently defined in Sect. **2.2.2**:
B =< Ĥ > −H
Where B is the bias, Ĥ is the estimator and H is the input "true" H. Therefore, B corresponds to both the deviations from the 1-to-1 line on a Ĥ vs H plot, and to the x-axis on the bias-std plots we provided, although on the latter, one needs to carefully look at the legend to see the dependence with H.

To provide an easier to understand characterization of the mean behaviour of the bias as a function of H, we thus propose, as suggested, to include the Ĥ vs H figure in the main text for the irregularity experiment (Figure 3):

,
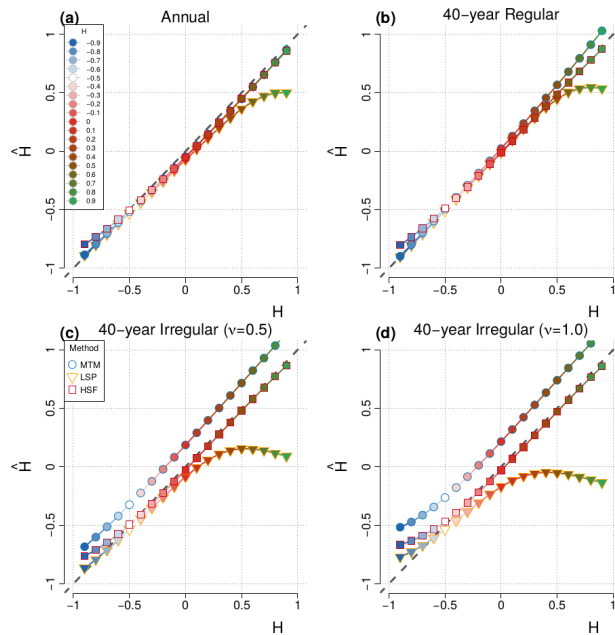
and also a similar figure for the length experiment (Figure 5) would be added in the supplement, along with the analogous figures for the block average method, which correspond to the results shown on Figures S1 and S3.

The confusion might also have stemmed from a slight mix up on our part regarding Figure 1 which incorrectly displayed results using the Block Average method in order to degrade the annual timeseries to an irregular resolution, rather than the Fitlering+Sub-Sampling. As a result, the displayed result on Figure 1 shows a bias for the regular case which is not reflected on Figures 3,4,5 and 7 as they were produced with the Filtering+Sub-Sampling method.

The reason for this bias is the aliasing of power from frequencies below the Nyquist frequency which is large in the case of blue noise with β<0 (H<-0.5) since the power keeps increasing below the Nyquist frequency. This is in fact one of the main motivations (lines 250-254) for us to show the Filtering+Sub-Sampling method as the main result, instead of the Block Average, since this allows us to identify other (more subtle) effects more easily rather than just the aliasing bias. The negative bias of the spectral method for β<0 when using the Block Average method can be seen in the supplement on Fig. S1, S2, S3.

We thus propose to update Figure 1 using the Filtering+Sub-Sampling method and introduce an analogous corrected figure in the supplement using the Block Average method in order to provide the same intuitive visualization of the mean behaviour for the Block Average results shown in the supplement.

 Minor comments:
- In line 27, replace log-linear by log-log
"Therefore, processes that fulfill this property show a log-linear relationship in the power spectrum (Schuster, 1898; Percival and Walden, 1993) over a given range of timescales."
By this we meant that once the log is taken we see a linear relationship between the log. We agree that the wording log-linear is indeed confusing.
We thus propose to rewrite as:
"Therefore, the power spectrum of such processes will appear linear on a log-log plot over a given range of timescales (Schuster, 1898; Percival and Walden, 1993). "

- In the review lines 46-52, the authors should cite the literature coming from other fields where irregular sampling had importance, such as astrophysics (where the Lomb-Scargle spectrum was introduced) and also fluid mechanics where Laser Doppler Velocimetry produces irregular sampled velocity data; in this domain an important literature is devoted to the question of adequate procedures for estimating Fourier spectra from LDV measurements.
We propose to add the following citations on lines 46-52:
"Secondly, the estimator can be adjusted for arbitrary sampling times. The so-called Lomb-Scargle Periodogram (LSP; Lomb, 1976; Scargle, 1982; Horne and Baliunas, 1986) was developed in the field of astronomy to identify periodic components in noisy astronomical timeseries with sampling hiatus and was often used to analyze Laser Doppler Velocimetry experiments which produce irregularly sampled velocity data (Benedict et al., 2000; Munteanu et al., 2016; Damaschke et al., 2018) as well as for the detection of biomedical rhythms (Schimmel, 2001). The LSP has sometimes been used in paleoclimatological context (Schulz and Stattegger, 1997; Trauth, 2020), although it may introduce additional bias and variance (Schulz and Stattegger, 1997; Schulz and Mudelsee, 2002; Rehfeld et al., 2011)."

- reference Corral and Gonzalez seems to have a problem of LateX writing - in the reference list I think it is not necessary to systematicaly provide the web address of each paper: the doi was introduced for this. Doi itself is sufficient (without https://doi.org). Example for the first reference: "doi: 10.1038/nclimate1456"

The mistake in Corral and Gonzalez was fixed.
We agree with the comment regarding web addresses. This was done automatically by the copernicus template. We should verify with the typesetting team if this should be removed.

- in the Discussion section it is necessary to mention the limitation of the present work which addresses only the H exponent for scaling processes, and not the intermittency. They should cite works that showed that the climate proxy data have multifractal statistics and discuss the fact that time series with intermittent fluctuations may react differently to the different methods proposed here, and also that an adequate method must be used to extract intermittency parameters (i.e. all the moment function and not only one moment order). The spectral methods cannot do this.
In the method section we already have a statement acknowledging this (line 89-91):
"On the other hand, highly intermittent archives which clearly display multifractality, such as volcanic series \citep{lovejoy_scaling_2016} or dust flux \citep{lovejoy_spiky_2019}, would require the ``intermittency correction'' from the moment scaling function $K(q)$."

We propose to further add in the discussion the following statements (in green following the original text in red copied here for context):

Around line 401:
"Our conclusions however rely on the real data having similar properties to the surrogate data used to validate the methods, and the conclusions may be different if the real data have properties different than the assumed power-law scaling. It is difficult to say for example which method would perform best if the data analyzed would contain two different scaling regimes, one with low H and the other with high H (Lovejoy and Schertzer, 2013). Another limitation of our surrogate data is that it is Gaussian-distributed, whereas real paleoclimate data can exhibit multifractality (Schmitt et al., 1995; Shao and Ditlevsen, 2016)."

Around line 425:
"On the other hand, the longer timescales of the MTM are relatively well estimated and it is appropriate to estimate the absolute variance over a timescale band well above the mean resolution (Rehfeld et al., 2018; Hébert et al, 2021). Further, the effect of proxy biases are well studied for the power spectrum (Kunz et al, 2020; Dolman et al., 2020) mainly due to the known properties of the Fourier transform. However, the MTM, and the LSP, do not allow for the characterization of intermittency through the study of all statistical moments, contrary to the HSF which is thus better suited for the analysis of timeseries displaying multifractality, such as paleoclimate timeseries at glacial-interglacial timescales comprising Dansgaard–Oeschger events (Schmitt et al., 1995; Shao and Ditlevsen, 2016)"