Author response to SC1 by Daniel Gavin

Dear Daniel Gavin,

we appreciate your helpful comment about the robust CHAR approach used in our manuscript and your constructive thoughts on improvements! Please find below your original remarks in black, and our responses in green:

I wish to submit a short comment on one part of the paper regarding the robust charcoal method originally described in Dietz et al. 2019 in PlosOne.

Incorporating uncertainties into proxy records, including both the age uncertainty and the uncertainty of the proxy itself, is important especially when comparing periods within a core and when comparing sites. It is overdue to include uncertainty in the analysis of individual sediment charcoal records. So, it is great to see this extension of the methods from the 'ensemble' approach from Blarquez et al.

My comment addresses the resampling methods used for estimating the uncertainty of the sediment accumulation rates. The robust method uses the age estimate of each sample (described as a mean and sd, but it could also be a PDF from an age-depth model), and selects ages from that PDF. Ages are generated independently for all samples, and only ages in adjacent samples that are in chronological order are retained. This results in some very slow sedimentation rates. This is acknowledged in the 2019 paper: "A comparison showed that robust fluxes were smoothed, but underestimated absolute mean fluxes due to strongly overlapping pdfage of adjacent samples at 1 cm sample resolution. Hence, we averaged the raw proxy and age values of three adjacent samples before robust flux calculation."

I am not sure how the averaging as described makes the CHAR influx values more comparable to the original influx values. did you average three samples in nonoverlapping segments, thus increasing the age difference of adjacent composite samples? The presented robust CHAR values are small compared to the raw data. I do not see this averaging step in the supplied code.

Thank you for drawing our attention to this important part of the method! Indeed, as you assumed, we did not apply an averaging of multiple adjacent samples in the current version of the manuscript. We have now run additional tests of robust CHAR parameters with different resampling intervals on our charcoal record, and concluded that the influence on the general trends is limited, whereas overall CHAR values are increased the more samples are aggregated. However, it rather seems that the different magnitudes of age uncertainties of the ^{14}C compared to $^{210}Pb/^{137}Cs$ ages of the record have a much higher impact on the resulting trend than the resampling. The current approach of scaling both to more comparable dimensions of 1σ and 2σ ranges, respectively, seems reasonable to us, which is why we suggest sticking to it. However, including an averaging of adjacent samples makes sense based on the high temporal resolution of the record and leads to CHAR magnitudes more comparable to the classic approach. For this reason, we will exchange the current version of robust CHAR in Fig. 3 with a slightly different one in the revised manuscript, which includes averaging over three adjacent samples following Dietze et al. (2019) (see Fig. 1 below). Some peaks of classic CHAR, for example in the latter half of phase 2, are still not mirrored in robust CHAR. However, we would expect to only see a conservative estimate in robust CHAR, meaning that only phases of increased CHAR that stand out even with the relatively large added uncertainties remain visible. Apart from that, the revised diagrams do provide a better fit to classic CHAR, especially in phase 4. Furthermore, we included this resampling step in the revised R script. It is now possible to easily choose at the beginning of the script whether a sample aggregation should be included, and across how many adjacent samples the averaging should take place. (script is available at https://github.com/rglueckler/CharcoalFireReconstructionR/tree/revised).

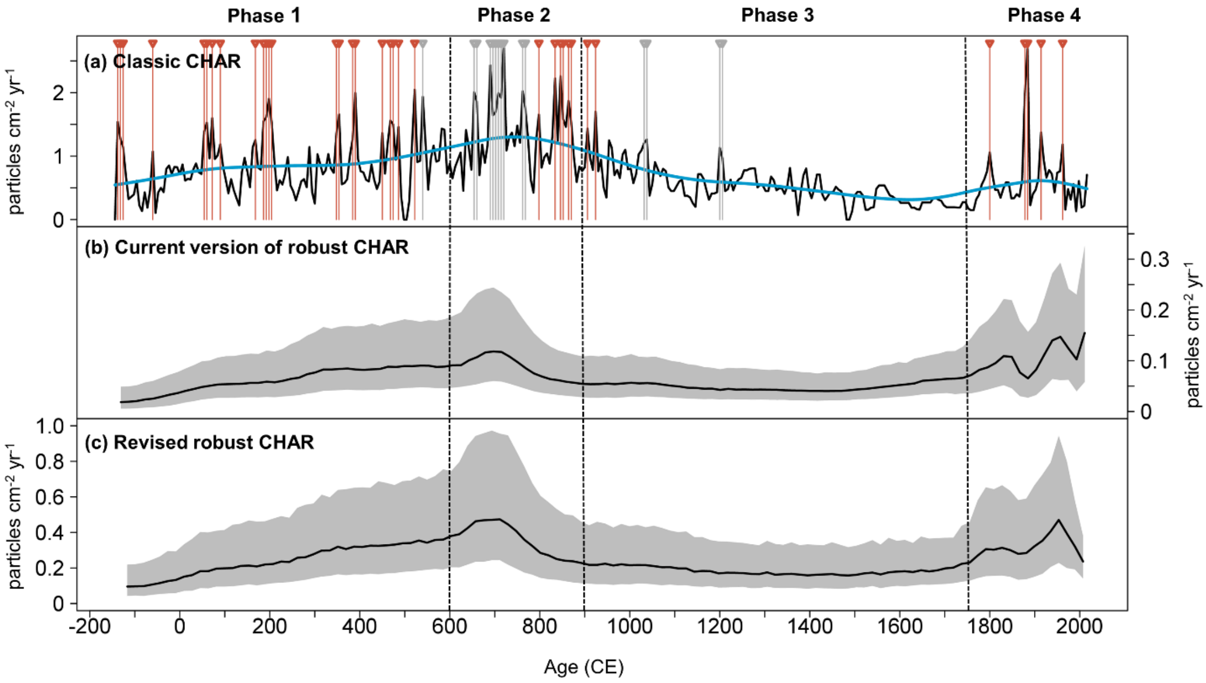
When the PDFs of adjacent samples are overlapping (<2 sd), the median age difference of the simulated ages is greater than the difference in the mean ages of the best-fit age-depth relationship. This is demonstrated in the attached figure. I think such small age differences occur in the majority of Holocene sediment records. The net effect is that as a core varies in sedimentation rate, the simulated sedimentation rate will have an increasing effect from the overlapping PDFs as the sedimentation rate decreases. This results in different effects of the analysis occurring in different parts of the same core. Variability in simulated sedimentation rates will not vary directly with the variation in the sedimentation of the best-fit age-depth model.

You are right. The difference is related to using only positive values for unit deposition times (the inverse of the sedimentation rate) and for the flux density distributions. We have now preserved the whole PDFs of unit deposition times in part 1 and of the proxy-flux calculations in part 2 of the robust CHAR script, as at these stages it is not necessary to remove the negative part of the PDFs (i.e. two dispensable lines of code in the robust CHAR function). Hence, only when we calculate the empiric flux density function we keep the positive values of the flux PDF. This results in slightly higher overall influxes but only in a small order of magnitude. The more pronounced effect is related to the relation of age uncertainties and sample resolution, as large age uncertainties will also lead to a wide spread of sedimentation rates for adjacent samples. We can account for this to a certain limit with the added sample aggregation.

An alternative approach to simulating sedimentation rates: use the output from bacon or clam, which saves many simulated runs of age-depth relationships. These can be used directly in the robust char calculations. The advantage here is that the simulated age-depth relationships preserves the monotonic age-depth pattern. The necessary ages for using this approach are in objects saved by the bacon and clam programs. (objects called info, dat, or chron). Clam and bacon can apply age uncertainties to proxy records directly. However, you have more flexibility by using the set of simulated age-depth relationships.

We appreciate this suggested alternative approach! It seems advantageous to use the age-depth model output directly to include the full range of multi-modal age distributions per depth. However, after some testing we think this is not a trivial task and requires more fine-tuning of our current method in order to be fully realized, which is beyond the scope of our present study. We will therefore work on implementing this approach in upcoming studies to further improve the robust CHAR methodology!

Figure 1: Comparison of **revised robust CHAR (including aggregation of three consecutive samples)** with its current version. Vertical dashed lines mark the different phases of the fire regime. (a) Classic CHAR peak component (dark grey bars = signal, light grey bars = noise, dashed horizontal line = threshold). (b) Current version of robust CHAR. (c) Revised robust CHAR. For (b) and (c): black line = median, grey area = interquartile range.



References mentioned in this response:

Dietze, E., Brykała, D., Schreuder, L. T., Jażdżewski, K., Blarquez, O., Brauer, A., Dietze, M., Obremska, M., Ott, F., Pieńczewska, A., Schouten, S., Hopmans, E. C. and Słowiński, M.: Human-induced fire regime shifts during 19th century industrialization: A robust fire regime reconstruction using northern Polish lake sediments, PLOS ONE, 14(9), 1–20, https://doi.org/10.1371/journal.pone.0222011, 2019.