Earth Syst. Sci. Data Discuss.,
https://doi.org/10.5194/essd-2020-184-RC1, 2020

Earth System Science Data Discussions

Interactive comment on "δ^{13}C values in stalagmites from tropical South America for the last two millennia" *by* Valdir Felipe Novello et al.

Anonymous Referee #1

Received and published: 27 August 2020

Review of manuscript: ESSD-2020-184 by Novello et al., about a data set publication entitled 'd13C values in stalagmites from tropical South America for the last two millennia'

The authors present a data set of d13C values for several speleothems from South America. Some records in the data set were already earlier published, some are new. But in all cases the according d18O values were published elsewhere. The time interval of provided data in the new data set are strictly limited to the last 2ka, even when the stalagmites started their growth earlier. Furthermore, the authors provide some interpretation on the d13 C data sets, by applying correlation methods (with T and prcp) and principal component analysis. They conclude that the d13C values in speleothems from South America reflect changes in hydrology, which is closely related to changes

in monsoon variability.

I like the approach about focusing on stable C isotopes in speleothem, as this is an underrepresented field, as this isotope system is usually much more difficult to interpret. Unfortunately, I think the paper is not ready to be accepted in ESSD. In fact, I suggest to make major edits to the data base and the manuscript text. Please find my suggestions below. I hope they are of some help to improve this interesting work.

First, I want to evaluate the data set. To my knowledge, this data set indeed seems to include all speleothems, which are published so far from this region and grew during the last 2ka. From this point of view, the data set seems to be complete. I, however, do not understand, why the authors decided to publish only the d13C values for those parts of the speleothems covering the last 2 ka - even when some speleothems showed some growth in earlier periods. I understand that the interpretation is only focused on the last 2ka, but this shouldn't be a reason to not publish the older portions of the speleothem data (>2ka). In fact, my understanding of this journals scope is its interest in the data sets and their description and not so much on interpretation of data. For me it makes no sense to publish only parts of the available d13C data for some speleothems. I agree, that the data base must not contain data from additional South American speleothems, which grew exclusively earlier than 2 ka BP, but I would expect to find at least the full data sets of those speleothems, which are already included.

Another reason, why the d13C values of the older parts of the speleothems should be included, is the usefulness of the data set. I think, you intend to publish this data set with wanting the data to be used. But why should anyone use incomplete data and instead has to search for the full data set elsewhere?

In addition, I would also favor, if you could add the according d18O values to this data set. I know that they are already published in individual papers (but this is also true for some of the d13C records) and even in SISAL (see Comas-Bru et al., 2020, ESSD), but it would make things much easier, when trying to use the data, as most users are

also interested in d18O. And it would be most inconvenient, to extract the d18O data from SISAL and d13C data from this data set (and maybe even to search for the older parts of the d13C data sets elsewhere) to obtain both time series.

Otherwise the data are well stored and easy to access in Pangea. I like this. However, as this journal is more about data sets instead of interpretation, it should at least contain a section in the manuscript, which describes how to access the data (even, when the section will be short, as it is very easy to access them). You might write about the file format, which program is suited to access the data. What other parameters are provided. What are the metadata? How were the age models constructed? (Describe in more detail how you constructed the composites if you decide to include them in your data base (see below).)

Furthermore, in the data set itself, you should:

* check if all the d13C values are really reported against the PDB standard as stated in the column header. (I think, they might be measured against the VPDB standard, aren't they?)

* remove the 'age (ka BP)' column. First this column is redundant and second it is not helpful at all, as you provide only two digits. This is very unfortunate for your sub-decadal, high resolution records.

* provide more digits for the 'depth' column. Otherwise this column is also only of limited help.

* add information where to find the according d18O values (if you decide against including them here)

*add information, where to find the original U-Th data sets. (Those are important for researchers wanting to establish alternative age-depth models.)

* add the stacked records as well, which you constructed here. You are free to follow this advice or not, but to my opinion this would save a lot time for others, which want

to have such composite records and would make your data base much more attractive and unique. And you already have the composites at hand. So this shouldn't be much of an effort.

Second, I want to focus this review on the interpretation part of the data set. Line 60: Please be more precise and rephrase. Using a sentence like 'The initial source of C for speleothems is the soil CO2 and tree roots' is very confusing. More correct would be that root respiration produces soil gas CO2. But it is also produced by microbial decomposition of SOM. So, please rephrase.

Line 63: '. . . models were proposed to explain the dissolution . . .' To my opinion, more correct would be '. . . models are the extreme cases of the dissolution . . .', as there are a lot of intermediate conditions, under which carbonate can be dissolved and those are much more common.

Line 64: Please replace 'Initially' with 'In an open system'.

Line 66: You should finish the sentence by something like: '. . . fingerprint from the d13C composition of soil gas CO2.'

Line 69 to 71: You could be more precise, if you would add something like that the mixing ratio is expected to be approximately 1:1.

Line 72: You might should replace 'partially open system' by 'intermediate dissolution system'.

Line 74: I would not agree with using the <10% value anymore. In 2001, it was indeed the case that investigated speleothems showed 'usually' values like that. But, I wouldn't use this finding anymore, as since then, there were a lot of additional studies, showing that quite often the host rock contributes a lot more carbon to the speleothem than 10% (e.g. Jackson and McDermott, 2008; Hoffmann et al., 2010; Griffiths et al., 2012; Lechleitner et al., 2016a;2016b; Spötl et al., 2016; Bajo et al., 2017; some examples in Markowska et al., 2019; Therre et al., 2020).

Line 93 and 94: I have to admit, that I haven't read the paper by Mickler et al. (2019), but I wonder, if they really say that PCP removes preferentially 12C from the solution. Usually and according to published fractionation factors for C, during carbonate precipitation 13C is preferentially removed. Nevertheless, the d13C values will increase in the solution during PCP. That part of the sentence is correct. But this effect would be due to the simultaneous degassing of CO2. This process is responsible for an increase in 13C in DIC as the light C isotope is preferentially removed during degassing.

Line 95-96: '... exposure of seepage solution to air pockets along the epikarst flow routs ...' Maybe that is also a reason, but usually, I thought, it is assumed that PCP is increased, due to larger drip intervals – leading to longer time intervals, where water is in contact with the cave atmosphere leading to more CO2 degassing and CaCO3 precipitation before the water reaches the speleothem top.

Line 98: Maybe it is more appropriate to talk about 'variations in PCP are climate related'. To my opinion, the occurrence of PCP itself is independent of climate, but its variations depend very well on climate conditions.

Line 174-175: I think it is appropriate to add 'seasonal' in this sentence as you refer to T-variations 'throughout the year' in the lines before.

Line 179-181: Here, you are describing how you calculate the average d13C data, which you use later (Fig. 2) to analyse them with respect to T and precipitation. I think you need to be a bit more precise here. To my understanding you calculate the mean d13C of the full record (for the last 2ka) and compare this with recent T. I doubt that this should be done. You might want to use the d13C values of the last 50 or 100 years or something like that. But using the mean d13C values of the last 2ka, shouldn't be compared with present-day T and precipitation - especially, after you mention a global T increase of 1.44 °C after 1850 CE (see line 145 of your manuscript).

Line 190: You describe here, how you constructed composite records. I understand from those lines, that you normalize the data before the stacking procedure starts. But

the stacking procedure itself is not described. Please elaborate more on this. Maybe it would also be a good idea to also present the stacked records in your data base (see above). This would be really a step forward.

Line 191: What do you mean by 'the inverse operation'?

Line 193: 'do not cover the entire period of the last 2000 years'. But you are calculating the PCA only from the period between 650-1950 CE (see line 197)? So, it shouldn't be a problem that those speleothems do not cover the full period of the last 2 ka. Or are they even shorter than this more limited period?

Line 197 to 199: I am not very familiar with the PCA, but my understanding of a PCA was, that it compares only relative changes in data, not absolute values. Thus, I do not understand your argument to exclude high-altitude sites due to large offsets. In addition, I do not understand, why you need a linear regression before you perform a PCA. Please explain.

Fig. 2: Indeed that are interesting relationships, however, I still wonder if you compare the mean d13C values calculated from the speleothems grown over the last 2ka with recent T and precipitation? Please clarify and recalculate the average of d13C over a more appropriate period (if not already done so). Furthermore, I do not fully understand why there are 9 dots in a) (d13C vs T) and 11 dots in b) (d13C vs prcp). Why both plots do not have the same number of data? In addition, according to Tab. 1 there exist data from 18 different caves. Please explain, why there are much less points in Fig. 2.

Line 282-284: I am not sure, if it is correct to argue that both regions experienced a larger isotopic contribution from host rocks to the d13C values. I think, there is much more what influences d13C than the host rock contribution. Such a claim could only be proved with additional 14C measurements. But usually other processes are more important for a change in d13C, as e.g., soil gas CO2 production rate, fractionation processes in the cave. Both of those processes lead to elevated d13C values during drier conditions and should have a higher impact on d13C as changes in carbonate

dissolution processes.

Line 310-311: Maybe it is better to be more specific and add ‘derived soil gas CO2’ after ‘soil organic matter’, as this is more interesting for speleothem d13C than the amount of SOM alone.

Line 336-337: You argue about T and atmospheric CO2 to be responsible for the missing coupling between d13C and d18O after $\sim$1750 CE. Around 1200 CE, there is a similar decrease in the PC1 of d18O and d13C. While d18O goes back to values similar to before this events, d13C does not. So there appears to be a similar decoupling of both isotope system as well. However, there is nothing comparable like T or CO2 increase. Thus, I wonder, how the decoupling at this period can be explained and if T and CO2 are indeed the most likely explanation for the decoupling after 1750CE. I do not have another explanation myself, but this observation leave me back in some doubt about the T and CO2 argument. That is really puzzling.

Line 340-341: You argue, that d13C is influenced by temperature here and refer to Fig. 2. But to my opinion, this plot shows only that the level of d13C values appear to depend on T (but still, you have to answer for which period you have calculated the average d13C values). I think, this is not a proof that the time series for d13C will also react on temperature in the same way. Therefore, please rephrase this sentence or leave it away, as you do not really need this.

Technical corrections:

Line 90: Please pay attentions to the spelling of ‘Schubert and Jahren’.

Line 311: Replace ‘low’ by ‘decreasing’

Line 330: You are using ‘fraction’ of explained variance, which calls for numbers between 0 and 1, but in Fig. 4 you provide numbers larger than 1. Please fix this.

Line 362: ‘much larger’

Interactive comment on Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2020-184, 2020.