Reply to the comments of referee #1:

This review starts out with the general comment that the paper is "somewhat speculative" and encounters "serious analytical difficulties on the way". However, in our opinion the following specific comments do in no way support or justify the above evaluation. While in our view some of the comments are factually incorrect, as explained below, other comments are interesting but not substantially relevant for our study. In the following we will address each specific comment and suggest potential modifications to our manuscript.

Referee #1 Comment 1: *Why is there such a big difference between _18O of the three stalagmites? Mukalla Cave has lower precipitation which should make its _18O higher but the opposite is true. And even the neighboring samples on Socotra I. differ by about 2 ‰*

The difference in delta-18-O between all stalagmites is to be expected and can be explained by several factors. Firstly, the Mukalla Cave stalagmite Y99 was deposited during the Eemian (130-120 kyr ago) when precipitation was considerably higher than during the entire Holocene (Fleitmann et al., 2011). Secondly, Mukalla Cave is further inland and received moisture not only from the Indian Ocean but also from Northern Africa (Herold and Lohmann, 2009), which was more depleted in delta-18-O due to the continental effect. Thirdly, stalagmites from Socotra Island exhibit more positive delta-1-8O values due to their close proximity to the moisture source (Indian Ocean). The isotopic offset between stalagmites D1 (Dimarshim Cave) and P3 (Pit Cave) is fairly small and most likely related to cave specific effects. Pit Cave is a fairly well-ventilated cave and therefore delta-18-O values are more affected by kinetic fractionation.

Referee #1 Comment 2: *The differences in water yield by the two methods are extraordinary. It seems unlikely that all of the water released at 480 _C is coming from the microscopically visible inclusions. There should be mention made of comparisons of yield to other studies, which generally give results comparable to what is obtained by heating to 320_C. It seems as if there are multiple storage sites for water in speleothem, something which has been suggested before and needs to be further investigated.*

The above mentioned differences in water yields are to be expected as not only the extraction temperatures of the two methods are different, but also the grain diameters of the respective calcite separates (350 microns for the 320 °C extractions; ≤ 100 microns for the 480°C extractions). Note that for example Yonge (1982) found that already different extraction temperatures alone cause substantial differences in extraction efficiency of water from stalagmite samples.

The statement that literature data on water yields are comparable to our 320°C extractions is incorrect. However, the results from our 480°C extractions, which we conclude to be close to quantitative, are indeed very similar to literature water yields from quantitative extractions (Yonge (1982): (24 – 100) x 10^{-4} g water / g of stalagmite rock, Demény et al. (2012): (1 – 10) x 10^{-4} g/g. Our 480°C results for comparison: (7 – 47) x 10^{-4} g/g. This similarity actually further corroborates the robustness of our water yield data, and thus this literature comparison would be included in a revised version of the manuscript.

We would like to emphasize generally, that our study focuses on the RELATIVE changes of water yields from one sample to another WITHIN one stalagmite, which is clearly stated in the manuscript. For this comparison it is crucial that extraction of water from each sample within this stalagmite is performed under the exact same experimental conditions, while the absolute water yields are of secondary importance to our conclusions. The same is true for the statement that not all water is stored in microscopically visible inclusions. This might well be the case but is not relevant for our study (see also reply to comment 6).

Referee #1 Comment 3: *The pixel counts gives about 10x more volume than the 320_C yields. This suggests that these are not very accurate estimates of fluid volume. There should be a table of these data, rather than only referring to them in text and figure captions.*

The 320°C water yields are by far not quantitative, as stated clearly in our manuscript. Therefore, results from pixel counts must only be compared to the 480°C extraction data. These are 10x higher than the 320°C extractions and therefore in excellent agreement with the results from pixel counts. The agreement of results obtained from these two fully independent methods further illustrates the robustness and resilience of our data, which should be mentioned in a revised version of the manuscript. However, we again emphasise that also here the important result is that the pixel count results support the RELATIVE changes of the stalagmites' water content as reflected by the water yields.

Referee #1 Comment 4: *The main conclusion of the paper is drawn from Fig. 6….*

This statement is incorrect. Data interpretation and conclusion are predominantly based on results visualised in Fig. 7 where regime shift analyses are applied to the delta-18-O data, and Students t-tests are applied to the average delta-18-O and water yield values on both sides of the respective major regime shift points. While Fig. 6 is discussed in 10 lines in chapter 3, Fig. 7 and the associated method are discussed on 32 lines in the same chapter, which also expresses the different relative importance of both Figures and associated methods.

…My impression is that there is too much scatter in these data to make a strong argument for a control by precipitation rate. For example, the data for P3 show essentially no correlation except that the last point is significantly higher. Likewise, for D1, there are really two regimes: low _18O and higher _18O but the uncertainty in each point is so large that it seems that only linear regression saves the day! Why is no error shown in the X-axis? Given the discussion in the text, this would seem to have a significant error as well. One would have to say that statistics suggests some correlation but it doesn't look like a method one could rely on very strongly. I think that the authors should be more forthright about this problem and discuss the sources of error and the reason for the scatter in the data. This is also important.

First we would like to clarify that Fig. 6 shows a plot of delta-18-O vs. water yield, not precipitation rate. After having analysed the direct correlation between water yield and delta-18-O (visualized in Fig. 6) we also concluded that there might be more appropriate (and complex) ways to investigate the potential correlations, such as regime shifts (see Fig. 7 and associated method). However, following Occam's Razor, the analysis in terms of a direct correlation between water yield and delta-18-O must not be ignored, and is already indicated from visual inspection of the plots in Fig. 6.
However, in a revised version we would statistically treat the data in a more rigorous way (see below, text in smaller font). Having done the respective calculations we conclude that the new results strengthen the direct correlations between average delta-18-O and water yield for D1 and Y99, but do not support a significant direct correlation for P3.
The errors on the water yields are in the range of 1-2%, in cases of very small water yields they can increase to up to several % (see table 2). These translate into error bars in Fig. 6 that can be seen in some cases, however more often than not error bars are smaller than the

symbol sizes. This could be stated in the figure captions of Fig. 6 in a revised version of the manuscript.

Adapted statistical treatment of the potential correlations of the parameters delta-18-O and water yield:

1. when averaging the delta-18-O values in order to "artificially" harmonize the different resolutions of the delta-18-O and water yield records we now use the more appropriate standard errors of the means with the averages instead of the standard deviations (by the way, this averaging procedure is, as stated in the figure captions of Fig. 6, the source of the rather large uncertainties of the delta-18-O values)
2. now we fit straight lines to the data by error-weighted least-squares regression, which results in slopes and errors (1σ) of 2.6 ± 0.2 (D1), 0.8 ± 1.0 (P3), and 1.2 ± 0.2 (Y99). The values of D1 and Y99 are clearly not consistent with a slope of zero, but P3 is.
3. Furthermore, we statistically compare by F-tests the two models "linear trend" (using the error-weighted regressions) and "no trend" (using an error weighted mean value of the data) for each data set. For D1 and Y99 the results clearly show that the "no trend" model should be rejected in favour of the "linear trend" model for D1 ($p = 0.00001$) and for Y99 ($p = 0.00068$), however, this is not the case for P3 ($p = 0.67$).

Referee #1 Comment 5: *One further detail in this regard: the paper repeatedly states that there is correlation between _18O and precipitation and cites Fleitmann et al., 2007 as the source. It seems that in that paper the "evidence" for this is the well known negative correlation between _18Oppt and rain volume ("amount effect"). There is not, as far as I could perceive, any independent evidence for this correlation. So this paper should really state from the outset that the correlation is attributable to the amount effect (identified by Dansgaard in 1964). That effect is known to be valid in tropical enviroinments typified by high rainfall, and where average annual temperature is > _25 _C. I am unaware if it has been demonstrated for desertic environments; has it?*

It has been demonstrated in several publications that the isotopic composition of precipitation in deserts and semi-deserts is strongly influenced by the amount effect. For instance, isotopic measurements of precipitation in Bahrain and Oman show that "…the amount effect is best observed in arid regions." and "…the low slope for rainfall data from Bahrain in the Arabian Gulf is clearly affected by secondary evaporation during rainfall. This effect on s (slope) is greatest for light rains"(Clark and Fritz, 1997).

Referee #1 Comment 5: *In trying to account for the correlation between _18O and water % the authors suggest that these effects are also relatable to differences in growth rate. However, for two of the stalagmites the U-Th dates provide direct evidence for growth rate. Does there seem to be any relationship between growth rate and _18O of calcite.*

We would like to clarify that nowhere in the manuscript we suggest a correlation between water yield or delta-18-O with the growth rate. We do propose, however, a correlation between water yield and drip rate (which, under suitable conditions may be correlated with the rate of precipitation).
While investigating the relationship between drip rate and growth rate is – again – not the scope of our manuscript, it might certainly be interesting to look into such a data set (water yield vs. growth rate). However, a prerequisite would be that age determinations and water yields should be analyzed from samples taken at the same spatial resolution, which is not at all the case for our data sets.
However, we speculate that a systematic correlation between the two parameters 'water yield' and 'growth rate' within a stalagmite is not necessarily to be expected. The primary question would always be if a slow and continuous or a fast but only periodic growth leads to higher overall growth rates. Without further knowledge, both scenarios seem plausible.

Referee #1 Comment 6: *On p. 2901 the authors state "We thereby imply that a sample with a high volume fraction of fluid inclusions (i.e., comprising both water- and gas-filled inclusions), is also characterized by a high volume fraction of water-filled inclusions alone and vice versa." why? What information do you have about gas-filled inclusions. Also: "imply" should perhaps be "infer".*

Considering the similarity of the results from our 480°C water extractions and the pixel counts for D1 and P3, the assumption that a large volume fraction of inclusion corresponds to higher water content and vice versa doesn't seem to be farfetched. Therefore the "imply" mainly refers to our third stalagmite Y99, for which no pixel counts exist. In a revised version, we would state this clearer in the text.

References:

Clark, I.D., Fritz, P.E., 1997. Environmental isotopes in hydrogeology. CRC press LLC, New York.

Demény, A., Czuppon, G., Siklosy, Z., Leél-Össy, S., Lin, K., Shen, C.-C., Gulyas, K., 2012. Mid-Holocene climate conditions and moisture source variations based on stable H, C, and O isotope compositions of speleothems in Hungary. Quarternary Int. in press,

Fleitmann, D., Burns, S.J., Pekala, M., Mangini, A., Al-Subbary, A.A., Al-Aowah, M., Kramers, J., Matter, A., 2011. Holocene and Pleistocene pluvial periods in Yemen, southern Arabia. Quart. Sci. Rev. 30, 783-787.

Herold, M., Lohmann, G., 2009. Eemian tropical and subtropical African moisture transport: an isotope modelling study. Climate Dynamics 33, 1075-1088.

Yonge, C.J., 1982. Stable isotope studies from the water extracted from speleothems. Thesis. Type, PhD Thesis, McMaster University.