

Response to Reviewers' comments

We are grateful to Yuri Dublyansky and John Mavrogenes for careful reading our manuscript, for their positive, insightful and constructive comments and suggestions and their supportive remarks that helped us to improve the clearness and the quality of our revised manuscript. Below we provide a point-by-point response to their comments:

Comments of Yuri Dublyansky:

Assumption 1: The modern-day temperature in the cave, obtained through 1 year-long monitoring, is assumed to correspond to the mean annual temperature (MAT) on the surface. This assumption is supported by references to general publications on cave climatology (McDermott, 2004; Fairchild et al., 2006). Although in many cases the cave T, indeed, corresponds to MAT, it is not ALWAYS the case, and significant discrepancies between the two values have been reported. This assumption must be supported by site-specific data (compare measured cave T's with independently derived MAT's from the area).

We totally agree with this comment. To show that the Milandre Cave temperature actually corresponds to the mean annual temperature on the surface we will add additional temperature measurements from Fahy station. The mean annual temperature at Fahy station (~ 550 m above sea level) in the vicinity (approx. 10 km) of Milandre Cave is 8.7 °C (1961-2011). Because Milandre cave is about 150 m lower than Fahy, the lapse rate (0.6 °C per 100 m) corrected mean annual temperature at our cave site should be around 9.6 °C, which is identical to measured cave air temperatures of 9.6 ± 0.2 °C. We therefore are convinced to state that the cave air temperature corresponds to the mean annual temperature above Milandre Cave.

Firstly, what is assumed by the authors to be the MAT from “the vicinity of Milandre cave” (p. 3699, l. 18; p. 3700, l. 14 and l. 24 for example) has little to do with the local area of the study. The temperature reconstructions extracted by the authors from Luterbacher et al. (2004) are averages for an area of ca. 15 million square km, stretching from Iceland to Syria and from northern Sweden to southern Spain. Portraying mean temperatures obtained from such a vast territory as representing mean annual temperature “in the vicinity of the Milandre Cave” is clearly inappropriate, if not misleading. As a minimum, the authors must present a convincing arguments as to why they believe the MAT averaged over the Europe can be attributed to one specific location in the Europe (with very small assumed uncertainty of tenths of a degree).

The temperature reconstructions published by Luterbacher and co-authors are based on gridded data and only the grid box corresponding to the location of Milandre Cave was used for the temperature reconstruction shown in figure 4. We will clarify this in the revised version of the manuscript.

Secondly, the original paper of Luterbacher et al. (2004) reported the temperature anomalies (i.e., relative values). The latter were converted by the authors into absolute temperatures. Methodology of the conversion is not presented in the paper, so there is no way of assessing the reliability of the derived temperatures.

Yuri you are absolutely right, but to calculate anomalies you need absolute temperature values. These absolute temperature data were provided by Prof. Jürg Luterbacher. The dataset was published in a peer-reviewed journal (Science, 2004: Vol. 303 no. 5663 pp. 1499-1503) to give a reader the possibility to test the reliability of the reconstruction. We will add this information into the revised manuscript.

Out of the two studied stalagmites, one (M1) has no associated geochronological data. Its growth rate is simply assumed to be similar to the second stalagmite, M2, on the basis of similar growth conditions (drip rate, drip height, T, etc.). In my opinion, this similarity does not represent a sufficiently robust basis for the assumption. For this second stalagmite, M2, the growth model was purportedly established (Schassmann, 2010; this is a Master Thesis, which means it is difficult to access). The growth model is not presented in the paper. All we are told is that the M2 had an average growth rate of approximately 0.02 mm per year and that the growth model “relies on U-Th dating in the lower part of the stalagmite and assumes a constant growth rate in the upper part” (p. 3695, l. 5-7). A number of questions arise here, the most relevant ones being: can the growth rates established in one part of the stalagmite be simply propagated throughout the stalagmite, and how reliable is the age estimates derived through extrapolation of the growth rates? (One must recall that at the assumed growth rate the stalagmite M2 must have been growing for 13.5 thousand years, and M1 – for 18.5 thousand years). Summary: the growth model purportedly exists for one stalagmite, but it is not available to a reader. The model is based on the U-Th dates from the lower (older) part of the stalagmite. Arguments why the determined growth rates must be constant (which means the age of the outer layers of the stalagmite can be determined by extrapolation) are not presented. The same growth rates are assumed to be valid for the second stalagmite, but basis of this assumption is poor (the presumed similarity of growth conditions). Under such circumstances the opinion of the authors about the growth period of the studied part of the stalagmites (assumed to be 0 to 350 years, as far as I understand; cf. p. 3700, l. 23-24) seems to be highly uncertain.

We agree that not knowing the exact growth rate of both stalagmites is a weak point of our study. Considering however the mean growth rates of all the studied stalagmites in Milandre Cave we found that it ranges from 0.02 to 0.25 millimeter per year (Schmassmann et al., 2010; Hasenfratz et al., in prep.; Häuselmann et al., in prep). These growth rates are very similar to those of stalagmites in caves with similar cave air temperatures. Furthermore, stalagmites M1 and M2 were both actively growing when collected in 2007 and there are no signs (e.g., dust layers) of discontinuities in the studied thin sections (see figure 6 in the manuscript). The age of the top of both stalagmites is therefore very well defined and it seems reasonable to use the observed growth rates of 0.02 to 0.25 millimeter to calculate a rough age model (as stated in the text). If we take the lowest estimate of 0.02 millimeter per year, the top of stalagmite M2 should comprise the last 350 years. If an extremely high growth rate of 0.25 millimeter per year is used, the analysed section should comprise the last 28 years. However, we consider this estimate to be highly unlikely as the calcite of stalagmite M2 is very dense and rather typical for a slow growing stalagmite. It is therefore likely that the analysed thin section of stalagmite M2 comprise the last 100 to 350 years. To clarify this we will add a short paragraph in the revised manuscript. Furthermore, we would like to emphasize that our intention was not to develop a precisely-dated temperature reconstruction for Milandre Cave, but our focus was clearly on the development of a

measurement routine to evaluate bubble radii. That is why our manuscript is submitted as a Technical Note.

Comment of John Mavrogenes:

I won't go into the technical aspects of this manuscript because Yuri Dublyansky has already done an excellent job, pointing out some significant issues that require the authors' attention. Thus, I am assuming that the measurements are basically correct, which makes this a very impressive piece of work. That the temperatures they obtain match historical temperatures in the cave is astounding. There are, however, some implications of this work that I feel are worth pointing out. Let me backtrack first. A few years ago I was asked to review a paper on temperatures acquired from halite in evaporate sequences. As with the current paper, the fluid inclusions contained no vapor bubble since they were metastable. In the case of the halite inclusions, they froze the samples for days to weeks at which time a bubble appeared. In this paper they use a femtosecond laser to nucleate a bubble, but both situations were similar; inclusions formed at temperatures so low that no bubbles nucleated. As I recall, it had been established that when measuring T_h of NaCl-hosted fluid inclusions only the highest temperatures were taken as the temperature of formation. Thus one measured away until one felt certain that no additional measurements would be higher and used the highest T as T_f . I questioned at the time how one can be sure that measuring 20 more might not have changed the T_f estimate and I ask the same question of the current work. In this case only the lowest temperature measured is used and all other temperatures are discarded. So, my question is: when have you measured enough fluid inclusions? And how could you ever be certain that enough have been measured. Interestingly, in this case measuring more might lower the T , while in the case of the halite measurements, more stood to raise the T . Given the contentious nature of global warming it is scary to think that one could unwittingly modify their results depending on their preferred model. Those measuring halite who would like to see higher T_s would diligently measure more inclusions while those measuring stalactites might stop sooner. I don't mean to disparage anyone here, as in both cases the integrity of the workers is above reproach, but given the strange politics associated with this issue, it is troubling. However, leaving aside the moral questions, I really would like to know how researchers using this technique would ever know if they have measured enough inclusions.

Dear John, I assume that the paper you had to review a few years ago as mentioned in your review is the paper by Tim K. Lowenstein et al "Paleotemperatures from fluid inclusions in halite: method verification and a 100'000 year paleotemperature record, Death Valley, CA" published in *Chemical Geology* 150 (1998) pp 223-245. I agree that at the first glance both studies seem to handle similar situations since both determine homogenization temperatures in fluid inclusions present in a monophasic state that contain only a liquid phase. However there is a difference between both studies. In case of the halite inclusions, the low temperature tail is explained by "either collapse of fluid inclusion walls or leaking of brine into fluid inclusions due to the pressure difference between the inside and the outside of inclusions". This strong pressure difference is caused by the extreme cooling of the samples to temperatures of -20°C . The highest homogenization temperatures measured therefore correspond to samples that kept their original volume. Any undesirable increase of the volume of the inclusions caused by stretching due to overheating far above the

homogenization temperature, or leaking of the fluid inclusions due to sample manipulation would lead to even higher homogenization temperatures.

In our study, we cool down the sample only to a temperature of 5°C which minimizes the mechanical load on the fluid inclusion and therefore the chance of a volume change of the inclusion. Besides these kinds of pressure induced volume changes we do not know any physical mechanisms that could lead to a decrease in the inclusion volume which would result in a decrease of the homogenization temperature. We therefore can assume that the lowest values measured correspond to the formation temperature of the stalagmite. Furthermore, since for each specific $T_{h\infty}$ there exist an inclusion volume below which no bubble can be induced (the surface tension is strong enough to prohibit bubble nucleation), the measured homogenization temperatures approach a lower limit. The lowest stalagmite formation temperature our technique can be applied is about 9°C. Of course, the determination of the stalagmite formation temperature used for palaeoclimatic interpretation assumes that neither natural processes nor sample handling have systematically altered the original fluid densities of the inclusions. If this is the case then we can assume that the closest approximation of the stalagmite formation temperature is derived from inclusions that display the lowest T_{hobs} values within a given growth band and are not influenced by any political or **socio-political** factors.

We sincerely thank Yuri Dublyansky and John Mavrogenes for their detailed and highly constructive review that will significantly improve the quality of the final manuscript or which we are indebted.