

## ***Interactive comment on “Estimating epikarst water storage by time-lapse surface to depth gravity measurements” by Cédric Champollion et al.***

**Anonymous Referee #1**

Received and published: 3 August 2017

### General comments

This paper aims at measuring gravity differences between the surface and caves, in order to measure the ground water mass changes within the unsaturated zone (and in particular, the epikarst) in three different karstic areas in the south of France. This is achieved by using CG5 Scintrex instruments, and, I'm afraid, the reason why, in my opinion, the results cannot be used, or at least should be discussed in much more details, for quantitative interpretation of differential gravity in terms of water storage, given the low signal to noise ratio.

Estimating the epikarst water storage is a very important objective, but this study suffers from a weak signal to noise ratio. I'm convinced that the authors made every effort to achieve very careful, state-of-the-art measurements, I'm just afraid that the

C1

limits of the CG5 possibilities (changing calibration, changing drift behavior, changing relaxation behavior, but also aliasing [see “specific comments”]) are such that relevant hydrogeological information cannot be extracted, or only at a barely significant level. In my opinion, the best way to support the case of this manuscript, as well as to confirm the results of Jacob et al. 2009, would be to repeat measurements on more seasonal cycles. Results could have been dramatically strengthened by performing measurements during more seasonal cycles at the 3 sites, from 2011 to summer 2017, and by discussing the orders of magnitude as a function of the numerous repeated absolute gravity measurements and the continuous series of the superconducting gravimeter installed in the same area.

Presently, I think that the best part worth publishing, e.g., as a short note in Journal of Geodesy, is the comparison between the two protocols of measurements. And, in that case, it will be valuable to discuss why Jacob et al. (2009) could achieve a much better result by using a similar “short time” strategy” (comparing Figure 4 with the Table 1 of Jacob et al. (2009), it seems that Jacob's process is as efficient as the “long strategy” of this manuscript). I also wonder to which extend the short method is more appropriate to mitigate for different relaxations effects: on Figure 5 of Jacob (or based on my experience with CG5s, or worse, on Figure 3b in Flury et al, 2006 (<http://bit.ly/2f7INXU>), there are clear jumps between different groups of points, which may not have been evidenced by the “long strategy”. Similarly, looking at Figure 4  $a_{t1}$  and  $a_{t2}$  in this manuscript, there are small groups of outliers centered on -0.008 ( $a_{t1}$ ) and -0.005 ( $a_{t2}$ ): do they belong to the same occupation? In that case, should this happen at  $t3$ - $t4$ - $t5$ , this would appear as a systematic, hidden error. Note also that one “bad” measurement, at  $t2$  or  $t5$ , can bias two of the three values on Table1. Same remark for Jacob with the  $t3$  measurements performed using the less precise CG5#323, which induces the largest error. The results briefly mentioned on L315-319, but not shown, must be discussed in details as well.

In its present stage, apart from the measurement protocol, the paper does not provide

C2

convincing results: the methodology was published by Jacob et al., 2009, at a location where the signal to noise was probably large enough (possibly thanks to the dolomite, but more cycles would better support the case) to infer relevant insight on the epikarst zone. This is not (yet) the case here. A way to publish the present results would be to discuss in details the significance of the results as a function of the signal to noise ratio. You may create synthetic data based on hydrogeological models and look at the influence of systematic errors caused by nonlinear relaxation, jumps, calibration, and aliasing issues (this latest point could be supported by using water balance models and meteorological time series). This will allow estimating the threshold in water content above which statistically significant information can be retrieved. Hence, you may provide at least an upper limit on the ground water changes.

#### Specific comments

-Figures 1-2: Add a regional map showing the 4 sites (i.e., include Jacob (2009) and Fores (2017) ones) altogether. Then discuss to which extent the results of Fores (2017) and present measurements of the flux tower and of the superconducting and absolute gravimeters can be used to improve the knowledge of the crop coefficient. For example, is the crop cover similar on the 3 sites? Given that all sites are located in a same Mediterranean area, one may expect that using the results from the superconducting gravimeter and the flux tower, one can provide an estimate of evapotranspiration parameters that is certainly better than the poor estimates used here and more generally, in hydrogeology. Then, the paper may be modified as a function of improved estimates of ET. Incidentally, in Table 1, I do not understand why at BESS and SEOU, the cumulative evapotranspiration values reach such low levels, actually comparable to winter ones and much lower than at the BEAU site, which, if I'm right, experiences similar climatic conditions (I'd even expect warmer and possibly drier conditions at SEOU, located at much lower altitude than BESS and BEAU; anyway, elaborate, please).

-L280: "stations are measured many times": do you mean: occupations? Is it exactly the same protocol as applied by Jacob?

C3

-Table 1: I do not understand why at BESS site, the differences for all depth have been added, as below 12 m results are not significant –according to the authors themselves– and hence, random signals and associated errors are summed rather than actual geophysical values. Why not taking the difference 0-12 m and then arguing that no significant gravity changes could be measured at greater depths?

-Table 1: why are the gravity differences divided by a factor 2, which is not applied to Figure 6?

-Why are measurements at t3 never discussed in the paper (however, it supports the discussion on the protocol)?

-How do you compute the final error? Is it the RMS of the STDs at each  $t_i$ ? Hence, according to Annex 1, I would expect, for example at SEOU,  $Dgt1-Dgt2 = -17 \pm 3.9 \mu\text{Gal}$  ( $3.9 = \sqrt{1.4^2 + 3.6^2}$ ). I assume that the errors provided in Annex 1 take into account the error due the calibration factor ( $1 \mu\text{Gal}$  in Jacob), is it? Elaborate. Same comment for BEAU: considering the  $\sigma_{STD}$  of Jacob (Table 1) I obtain 2.1 rather than 1.6.

-Figure 6: does "recharge" stand for  $t2-t4$  and discharge for  $t4-t5$ ? Mention it, and, in that case, one should read (a) +9, +3, -2, 0  $\mu\text{Gal}$  according to Annex 1.

-The reasoning leading to equation (7) repeats what was written by Jacob. Hence, section 3 can be significantly shortened by just referring to Jacob et al., 2009.

-L526: the mentioned 30 m are not significant; they belong to the error bars given in Table 1.

-L602: not only absolute measurements: this should be possible by performing continuous, relative gravity measurements both at the surface and in caves.

-What about possible aliasing effects? What would be the influence of a strong rainfall just before a gravity survey? Can you rule out this artifact, e.g., based on meteorological series and on the way the superconducting gravimeter behaves in the Larzac (close

C4

by BEAU site)? The numerous outstanding AG measurements may help as well.

-What is the role of the saturated zone? At SEOU, underground water is pretty close to the gravimeter when it is measuring in the cave, hence, it may be much more sensitive to that water changes than the surface site, hence biasing the results. This depends on the spatial extension of the water table; this should be modeled and discussed.

-Annex 1: why is the calibration factor around 0.9995, while being at 1.0005 in Jacob for CG5#1687? Slow evolution as evidenced by e.g., Meurers, 2017 <https://doi.org/10.1016/j.geog.2017.02.009>? Or artifact? Does gravity at experimental sites belong to the gravity ranged along the calibration line? If not, what would be the consequence? An error of 0.001 on the calibration factor would result in a 4  $\mu\text{Gal}$  error at SEOU, and nearly 6 at BEAU. Incidentally, I do not understand Table 1/Figure 4 of Jacob: on Figure 4 the calibration factor changes continuously by 800 ppm (from 400 to 1200) between  $t_0$  and  $t_5$  (first and last grey lines), whilst the calibration corrections mentioned on Table 1 randomly fluctuate from 1.0003 to 1.00065, i.e. 350 ppm. But, 800 ppm on a gravity difference of 6000  $\mu\text{Gal}$  as in BEAU would mean an error of about 5  $\mu\text{Gal}$ .

Technical corrections

-Missing spaces between several numerical values and unit symbol (e.g., <https://physics.nist.gov/cuu/pdf/sp811.pdf>)

-Inappropriate use of the present perfect for events that happened in the past.

-Table 1: according to Jacob, Sep07-Feb08 should be 12.9 (\*2?) instead of 13.0 (\*2?)  $\mu\text{Gal}$ .

-Table 1: for a better legibility make use of colors or gray background to evidence recharge and discharge periods; provide also an additional column providing the ratio  $\text{EqW}/\text{NW1}$ , which is discussed in the manuscript.

-Figure 4: provide the dates corresponding to each  $t_i$ ; invert (a) and (b) such that SEOU

C5

and BESS appear in a same order in the whole manuscript. How do those histograms compare with Jacob? As all authors of Jacob and this paper have belonged to the same lab, you may compute the histograms from Jacob's data directly, and include it in the discussion.

-Avoid useless repetitions, especially about the vulnerability of karst aquifers (introduction, L521-545).

-Deville et al., 2012, in press: published in 2013.

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., <https://doi.org/10.5194/hess-2016-355>, 2017.