

Interactive comment on “Accelerated gravity testing of aquitard core permeability and implications at formation and regional scale” by W. A. Timms et al.

X. Sanchez-Vila (Referee)

xavier.sanchezvila@upc.edu

Received and published: 13 April 2015

This paper presents how tests under accelerated gravity can be used to obtain fast values of the local conductivity of low permeable media. Then some effort is devoted to assess the relevance of the results as a method to evaluate the recharge through an aquitard in a field site located in NSW.

We start from the title; it involves the words “aquifer testing” and “implications”. Thus, the emphasis is placed on the aquifer testing. The way I see this paper, it focuses on the testing of a low permeability site with emphasis in the technique used. Then it

C995

turns to try to convince the reader that the data is correct as it agrees with other existing one obtained from different methods at different scales (local values from standard permeability test, as well as the calibration of a regional model), and to write a number of implications regarding the relevance of assessing properly the overall (effective) vertical conductivity of aquitards.

I am very much convinced about the need to properly characterize the hydraulic conductivity of aquitards, particularly at the regional scale, but I do not think this paper contributes toward this objective. I will try to explain myself next.

There is little to say regarding the part corresponding to the testing part. The centrifugal test was already presented in the work of Nimmo and Mello (WRR, 1991) and it eventually became the basis for the ASTM D6527, as correctly pointed out by the authors. In these papers it was already reported the benefits of using such method to evaluate K_v in low permeability media based on the enhanced acceleration. Centrifugal apparatus are present in a number of geotechnical labs (at least the big ones). Thus, I even question the need to include section 4.3 at all; maybe a simple reference to ASTM method, with a reference to equation (7) would be enough. As a summary of this part, the paper cannot be centered in the testing part, but rather on the implications regarding recharge through thick continuous aquitards.

The geological setup and the sampling process are also relevant and should be included in any paper, but again it is not the center of the paper.

Connectivity is therefore the issue. From local values you cannot address connectivity, obviously. Connectivity is better obtained if there is a large number of local data and the value obtained from calibration of a model is too large to fit any reasonable average of the local quantities. But in order to address this problem you need a lot of data, or else, little data and a clear model of connectivity or presence of high conductive features (e.g., fractures) that can account for it. This is missing here. The number of K data is too low to provide a real comparison in terms of upscaling techniques, and thus it is

C996

mostly done in terms of order of magnitude approximation.

There are 14 data points. Actually it is $3 + 5 + 6$. Notice the great variability (except in CL). The variance is quite large in the NR and BF sites. So, if the variance is so large, it is difficult to assess how representative the values are. Yet, with these few data, you cannot assess that some site (northern sites) are more heterogeneous than another one. The authors talk about values of 10^{-7} , but actually there is only one point per site with such value. Are these representative? Notice the sentence in pg 2817 lines 16 to 19; this is quite a strong assertion, so how many data were actually used to produce this statement? Is the statement statistically correct? Did they pass any statistical test to disprove all data belong to the same population?

All the upscaling effort at the CL site is reduced to the paragraph in pg 2820, lines 1-6.

The finding reported in pg 2821 regarding a very low K value derived from the regional model (and contradicting the observations reported in many papers such as the ones listed in the last paragraph of pg 2821), makes me suspect that the number of samples in this work is too low (so the mean obtained from samples is not representative). Of course it could be the opposite, and then the result is of very high significance and I would love to know why, but with this amount of data it becomes speculative (blaming the model). The speculative part I am saying can be seen in the sentence "[...] the clayey sediments in this region may lack preferential flow paths at some sites, and in other areas preferential flow may occur through features such as fractures and heterogeneity at a range of scales"

The last discussion part is very interesting. In any case I do not understand why in the discussion there is a large part devoted to stress the application of the centrifuge technology. In my opinion this is mostly known, and so it would be part of the introduction.

Page 2024 last paragraph. It reads "The availability of core scale facies measurements enables the up-scaling of bore log and geophysical data to determine upper and lower hydraulic conductivity bounds for regionally up-scaled aquitard units. Any differences

C997

between K values at various scales are important for indicating the possibility of preferential flow through heterogeneous strata or aquitard defects. The availability of these bounded estimates helps to constrain the uncertainty analyses conducted on regional groundwater flow models to yield more confident predictions. [...] mapping a regional aquitard- aquifer system by combining stochastic fluvial process modelling and a geostatistical simulation technique." This is precisely my point. I would expect the use of Kv data from an aquitard to indicate the possibility of preferential flow by mapping it into a geostatistical simulation model of some sort. It cannot be a semiquantitative indication in my opinion, in particular when the actual results (lower K in the regional model) contradict most of the examples found in the literature, indicating the need to find a good alternative explanation.

Some additional comments, some minor, some moderate

Pg 2800 line 5 "...rapid and reliable..." In my opinion you cannot assess reliability from this work. You get a number of data and you compare them with other values but only in a semiquantitative way.

Pg 2805 line 4 I find very strange to reference a figure in an introduction section. This should be moved to section 4.2.

Equation (8) is mathematically incorrect. You cannot use the same variable in the limit of the integral and also as the integrating variable. Thus, both r and dr within the integral should be r' and dr' (or any other letter). Same in eq (9).

Pg 2816 line 19. It is Figure 3, not 5. This is interesting, as there is a transient behavior observed in the flow rate (or apparent K) vs time. The transient behavior is maybe something to be explored in the future. If it is caused by sealing, what is the real structure at the site itself; can you guarantee that all cracks were sealed?

Pg 2816 lines 23-25. Somewhat confusing. If some data is not representative does it mean it was not in steady-state conditions?

C998

Pg 2818 I am sure the authors checked that the samples did not show at the end of the test any impact from overconsolidation; maybe a sentence here would be nice.

Table 3. Different g-levels used. Why? How were they selected?

I do not see the relevance of figure 4. Everything is linear, no need for a plot.

Figure 6, it is questionable whether it should be included in the paper. The scale of geophysics tends to oversmooth any particular variable or parameter, so unless it is stated how they built the map with reference to volume of data support and interpolation method used for drawing, it cannot be used as an external confirmation of continuity.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 12, 2799, 2015.