Reply to Reviewer #3

We thank the reviewer for the useful feedback which improves the paper. Unlike the reviewer, we are convinced that the method is sound. We believe that this assessment is the result of our inability to convey the method completely and clearly in the paper. We apologize for that and have tried our utmost to describe the methods more clearly.

Hereafter, all comments from the reviewer are repeated in italic face, our response is thereafter in regular face.

General comments

This manuscript discusses a method of estimating vertical hydraulic conductivity and thickness of groundwater aquitards. This topic should be of interest to HESS readers, because resistance of aquitard has not been well studied. The authors seem to have a large database (REGIS) that should be useful for them to pursue their research topic. The manuscript is also well organized. However, I am afraid that the manuscript does not meet the requirements for publishing in HESS, and below are my review comments.

R3.1 One misinterpretation of our paper seems to be the aim of the study. We do not want to calibrate or improve the vertical resistance of an aquitard of a groundwater flow model, but we want to improve the (hydro)geological model, making use of the calibration results of a groundwater flow model.

Major Comments:

The methodology presented in this manuscript does not appear to be sound. For example, the authors presented prior and posterior distribution. I assumed that they used Bayesian approaches, but I did not find anything Bayesian in the text.

R3.2 We failed to mention this explicitly, but the method is essentially a Bayesian method. Using a prior distribution of the conductivities (assumed spatially uniform) and a prior distribution of layer thicknesses (which is interpolated) we build a (prior) joint distribution of litho-layer thicknesses and conductivities. If we have n litho-layers, this is a $2n$-dimensional joint distribution $f(K_l, D_l, l = 1, \ldots, n)$. The posterior distribution for a given value of the observed total resistivity c_m, i.e. $f(K_l, D_l, l = 1, \ldots, n|c_m)$ can be found by sampling from $f(K_l, D_l, l = 1, \ldots, n)$ combinations of $2n$ values that result in the observed c_m-value. These combinations with their associated densities (normalized) form the conditional = posterior distribution $f(K_l, D_l, l = 1, \ldots, n|c_m)$. This is thus essential Bayesian, but without using Bayes' theorem explicitly. Next, we take the combination of $2n$-values with the maximum value of $f(K_l, D_l, l = 1, \ldots, n|c_m)$ as a maximum posterior estimate. This is essentially the same as a Maximum Likelihood estimate.

However, finding the conditional distribution, or even its maximum probability density value, in a $2n$-dimensional space (in our case 14-dimensional) is cumbersome. MCMC-methods could bring a solution here, but we have chosen for a method that exploits the properties of algebraic combinations of independent random variables and a piece-wise approximation of the probability density functions (PDFs; see our appendix) to arrive at the ML-estimate. The flow chart as given in response to reviewer #2 and Figure 1 of the paper provide insight into our method:

- We build for each location and for each litho-layer l that makes of the total resistivity of the aquitard, separately a joint a priori probability density function (Figure 1a) of K_l,D_l from the a priori marginal PDFs of D_l (obtained from interpolation) and K_l (assumed spatially homogeneous for a given litho-class).
- From the a priori PDF we obtain a maximum likelihood function (MLF) providing all combinations of K_l,D_l that provide a value of resistivity C_l with maximum probability density (please note that this is not the likelihood function as in Bayes' Theorem, i.e the prob[observations

given the parameter values]. The MLF is not a PDF in fact).

- The MLFs for all C_l are combined to obtain the MLF for the entire aquitard (see Flow chart). From this the MLF-value for the observed value c_m is obtained. Using this MLF-value, the MLF values and the associated K_l, D_l of each litho-layer can be back-calculated.
- The resulting $2n$-dimensional combination of K_l,D_l is the ML-estimate. This method is much faster than using MCMC-methods. However, the price we pay is that we do not obtain the full posterior (conditional) PDF $f(K_l, D_l, l = 1, \ldots, n | c_m)$, but only its maximum value.

We realize that our explanation of the method has not been clear enough. We hope that the explanation in this rebuttal is, and we will also change the method description accordingly.

The authors also talked about maximum likelihood methods, but I did not find anything about the likelihood function as well as observations and simulations used to build the likelihood function. Without such information, it is impossible to evaluate whether this research is technically sound.
R3.3 See also the answer to R3.2.
The likelihood function is defined as a joint density function (e.g. Mood et al., 1974, p.278). This is what we define in Sec. 2.1 (page 4196, line 21-23). Since we have all marginal distributions available ($D_l(u)$ and K_l) this distribution is fully defined. From this joint distribution, the relations between the marginal distributions described in Eqs. 1 and 2, and the calibrated vertical resistance $c_m(u)$, the maximum likelihood of all marginal distributions is uniquely defined.

Figure 1 seems the basis of the methodology used by the authors, but the figure is really hard to be understood. First of all, what is relative density? It is said in the text that all depicted densities are proportional to the maximum density of the joint PDF, but I do not see any joint PDF. Figure 1a looks like joint PDF but the legend is for relative density not for joint PDF.
R3.4 Figure 1a is indeed a bivariate joint probability density function with as marginal distributions the thickness (x-axis) and vertical conductivity (y-axis). The PDFs of these respective axis are not shown. The colors indicate the value of the joint density. It is not unusual to depict a joint density function in this way (e.g. Mood et al., 1974, p. 140,163,187)
The PDF of the vertical conductivity has very low values, hence the densities are very high. Depicting the real densities would yield a harder to read picture, therefore we divided all densities by the maximum joint density in the joint distribution, which yields the relative density. This affects only the legend, not the picture.

In addition, Figure 1(a) showed vertical conductivity but Figure 1(b) showed vertical resistance. To convert one variable to the other requires knowing aquitard (aquifer) thickness, but the thickness is not discussed anywhere in the text or the figure. The text in page 4197 is also unclear how to obtain the curves of Figures 1(b) and 1(c) based on Figure 1(a). Figure 1(a) is also questionable. How is Figure 1(a) obtained? How we can know whether Figure 1(a) is justifiable/accurate/reliable?
R3.5 In Eq. 1, the formula is given for calculating the vertical resistance out of the thickness ($D_l(u)$) and the conductivity (K_l). So, in Fig. 1a, the vertical resistance is the quotient of the layer thickness (x-axis) and the vertical conductivity (y-axis). The gray lines in Fig. 1a show this quotient for constant values of $C_l(u)$. By definition, integration of this joint distribution of Fig. 1a, with respect to $C_l(u)$, yields the PDF in Fig. 1b.
It is not completely clear to us what you mean with justifiability of Fig. 1a. We assume that the marginal distributions are accurate. If that is the case, the joint distribution in Fig 1a is correct too.
We will use your comment to rephrase parts of section 2.1.

For Figure 1(c), the authors said that it is based on the maximum likelihood. If so, what is the likelihood function? What are the related data and model simulations?
R3.6 Figure 1c is the maximum likelihood function for values of $C_l(u)$ in Fig. 1a. The relation

between Fig. 1c and Fig. 1a is direct. The black dashed line in Fig. 1a crosses the gray lines at the highest density available at this gray line. This position is denoted by the black dots. Herewith, for every value of C the maximum likelihood, or density, is known in the graph of Fig. 1c. Herewith, Fig. 1c shows the likelihood function of C, Fig. 1c is not a PDF.
We didn't do any model simulations.
See also the reply to R3.3.

As to the research itself, some assumptions used in this study do not seem to be reasonable. In L1-3 of P4196, the authors assumed that the calibrated resistance is the best estimate of the vertical resistance. This sentence is really confusing. If the assumption is correct, then why did the authors need to estimate vertical hydraulic conductivity, as in the L7-8 of P4208? If I am not mistaken, the authors used the calibrated resistance, but estimated layer thickness. If this is the case, the authors may make it clear and consistent about the goal of their study.
R3.7 If a groundwater flow model is calibrated, usually the assumption is that the calibrated version is better than the uncalibrated version. This is assumed because extra data is used for the calibration. Without this assumption, the calibrated model should be discarded. Therefore, we think it correct to make this assumption.
In the introduction, page 4193, line 19 to 21, we describe that we aim to find the most likely hydrogeological parameters of the subsoil. Every aquitard is an aggregation of one or more litho-layers. Each litho-layer has a thickness and a conductivity. This is the description of the hydrogeological model. On the other hand, the groundwater flow model has only one parameter per grid block, the vertical resistance. The proposed method uses one calibrated value of the vertical resistance to find the most likely parameters, layer thickness and conductance, of the constituting litho-layers of the aquitard. Herewith, the knowledge of the hydrogeological model can be improved.
For a revised introduction, with a more clear description of the aim of our study, please find a draft version attached to the reply to reviewer #1.

However, this may lead to another confusion. If the layer thickness is important, varying it may change the calibrated resistance (or conductance). In this case, the assumption made by the authors is not reasonable any more.
R3.8 The litho-layer thickness and the vertical conductance are described by PDFs. The calibrated vertical resistance is, in the current study, a deterministic value. Varying the layer thickness will vary the vertical conductance too, but their quotient (the vertical resistance) keeps constant. This is depicted in Fig. 1a: the gray line denotes the calibrated value, the marginal values are found at the respective axis. The location with the black dot has the highest density along the gray line and has therefore the most likely marginal values.

In addition, the authors probably made a number of assumptions, but the assumptions are not discussed. For example, in L10-15 of P4199, it is unclear how the PDF is obtained using kriging results. The authors probably assumed Gaussian distributions, but it is not mentioned anywhere in the text.
R3.9 We have no information available about the type of distribution of the uncertain observations of the layer thickness. Therefore, we tested different distributions as described in Sec. 4.3. We choose the lognormal distribution for the observations (page 4205, line 26).
You are right that we omitted to give the type of distribution of the interpolation error, we used a lognormal distribution. We will add this to the text.
On page 4198, line 23, we refer to previous work in preparation for publication. That paper describes in more detail the kriging method with interpolation of PDFs. It is yet accepted for publication and available on line (Lourens and van Geer, 2015).

The research goal itself is questionable. The authors here separate vertical hydraulic conductivity and layer thickness, and treated the former as a randomly homogeneous field but the latter as a randomly heterogeneous field. But the separation is not needed, if hydraulic head in the aquitard is not of

interest. In this case, one can build a pseudo- 3D model and only need to calibrate one variable, i.e., the vertical resistance (or con- ductance).
R3.10 We agree that for the calibration process it is useful to calibrate only one parameter of the aquitard. But we do not aim to improve the calibration result, but we want to improve the quality of the description of the subsoil, the hydrogeological model. In this model, an aquitard may be described by multiple litho-layers. We want to improve this description, making use of the results of a calibrated groundwater flow model.

At last, the manuscript is lack of an evaluation of the results. The authors presented a large number of figures and tables, but did not discuss how reliable the results are. For example, the authors did not present any variograms for readers to evaluate the variogram models shown in Table 5. For estimating the PDF of layer thickness using the kriging method, it requires estimating variogram. However, the estimation requires a relatively large number of data, and such data is always unknown. Based on the description in Section 2.4, it seems that the REGIS database has the data. If so, the authors need to present the sample variograms estimated from the data (not the fitted model in Table 3). More importantly, the authors need to process the data, at least, to investigate whether the layer thickness is Gaussian, because the PDF estimation based on kriging assumes Gaussian distribution.
R3.11 Reviewer #2 (R2.8) also asked for these experimental variograms. We proposed to add them as supplementary material to this paper.
Please find them attached to the reply to reviewer #2.

Minor Comments:

L1-5 of P4201. If there are 475,000 litho-layer thickness at 16,000 borehole locations, it is confusing that the authors said that no quantitative information is available about the uncertainty of the litho-layer thickness. What kind of uncertainty is the authors discussing here?
R3.12 We do have much borehole information. Every borehole description describes the individual litho-layers. The thickness of each individual layer is given as a single value, but it is not known how accurate this thickness value is. That is the uncertainty we consider.

It is unclear what subsoil is. What is the difference between soil and subsoil? The authors also used litho-layers later on. Is it the same as the subsoil?
R3.13 We think we should avoid 'soil' and only use 'subsoil'. Because of remarks of Reviewers #1 and #2 we wrote a new introduction. There, definitions of these terms are given.
Please see the reply to Reviewer #1 for a draft version of the new introduction.

Spell out PDF when it is used for the first time.
R3.14 This indeed should have been done in the introduction. We will take care of it.

L25 of P4195. Define the variables u and cm, and the function u(cm(u))
R3.15 There is no function u() at all, but we understand the misinterpretation of the text. We put the variable $c_m(u)$ between braces because it is "the vertical resistance of an aquitard at grid block u". So this sentence is meant as the definition of c_m and u.
We will rephrase the sentence to avoid this confusion.

Dependence of layer thickness (D) and hydraulic conductivity (K) may not be reasonable
R3.16 In the current study, we agree that it is unlikely that this dependency exists. But, for instance, when layers have a wide range of thickness, differences in compaction of the sediments can influence the conductivity. At this moment, this subject is not of great importance to us.

Concerning the remarks, we conclude that especially the introduction and Sec. 2.1 are not clear to the readers. We have written a draft version of a new introduction (see reply to Reviewer #1) and an introduction to Sec. 2 (see reply to Reviewer #2).

Furthermore, we propose to rewrite the other sections too, considering the remarks of all reviewers.

References

Lourens, A. and van Geer, F.: Uncertainty Propagation of Arbitrary Probability Density Functions Applied to Upscaling of Transmissivities, Stoch Environ Res Risk Assess, doi:10.1007/s00477-015-1075-8, 2015.

Mood, A. M., Graybill, F. A., and Boes, D. C.: Introduction to the Theory of Statistics, McGraw-Hill, Singapore, 3rd edition edn., 1974.