**Review of:**

Zech, A, P Dietrich, S Atttinger, and G Teutsch, 2020, A Field Evidence Model: How to Predict Transport in a Heterogeneous Aquifers at Low Investigation Level? Hydrology and Earth Sciences (in review)

By: Dr Joost C. Herweijer – joost.herweijer@gmail.com
Date: 2 September 2020.

---

## Summary

Below are the three main issues that I identified with respect to this paper:

1- The conceptualization is poor, it does not take into account available geological information. The modelling process does not adhere to well-published standard practices for this type of hierarchical modelling. Some key features of the data which underpin the conceptualization are not addressed, most critically, inconsistencies regarding the basic data (K values).
2- The analysis is conducted in a single 2D-vertical cross-section assuming symmetry in one of the horizontal directions. This is not in accordance to observed piezometric level and tracer test data showing flow and transport in both horizontal directions. The results of the 2D model cannot be used for quantitative analysis.
3- The paper incompletely references geological issues very relevant to conceptualization of heterogeneous aquifers. Especially when it comes to options for geological analysis and modelling, some statements in the paper are not well documented and potentially misleading

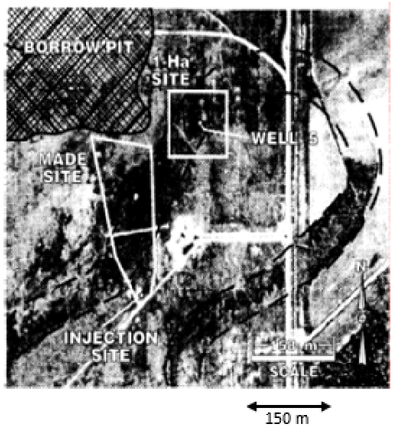
Based on point 1 and point 2, I reject this manuscript.

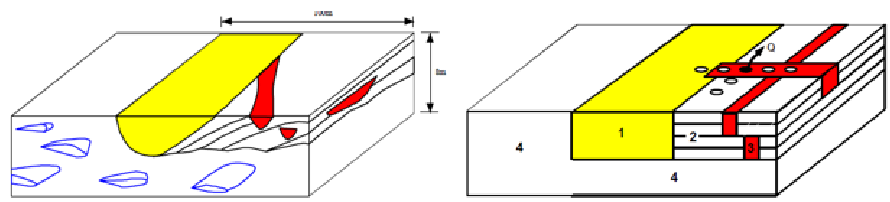
The MADE data set is the result of significant efforts to collect a 3D dataset of hydraulic conductivity and tracer transport observations. Many papers have been published addressing various aspects of the data and the geological concepts, showing significant variability in 3D of hydraulic conductivity and tracer transport. Creating models in 3D that take into account these data is well within the realm of what is technically possible. No scientific rationale is presented to explain why the model has been restricted to only 2D.

## 1 – Conceptualization is incomplete, inadequate and not novel

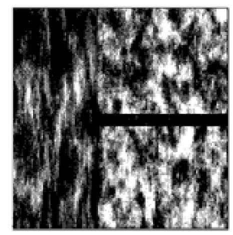
The paper promises a 'novel' conceptualization. However, the paper falls short of an adequate conceptualization. As discussed in 1a below, very adequate conceptualizations have been provided for the MADE aquifer (eg. Herweijer, 1997) and the MADE site in particular (eg. Julian et al., 2001), and these should be referenced and compared with the proposed approach. Also, as discussed in 1d below, the 'nested' scale' approach chosen by the authors is not novel.

1a) Creating K zones based on the change in piezometric surface was earlier discussed for the MADE site by Rehfeldt (1992) and Herweijer (1997). These K zones are related to the main geological feature at the site that has been reported about in several papers (eg. Herweijer and Young, 1991; Young 1995; Herweijer, 1996, 1997; Julian 2001; Bowling et al., 2005). The single vertical K zone model presented in the paper is incorrect. Herweijer (1997) established that the MADE aquifer consists of a two layer system where a high K channel deposit incised in somewhat older lower K deposits (see figure below).




| facies or architectural element | log-K measured min ... med ... max | model |
|---|---|---|
| 1 main meandering channel deposits | 1 -5.0 ... -2.9 ... -0.5 | -3.0 |
| 2 pointbar deposits | 2 -5.6 ... -3.4 ... -1.3 | -4.0 |
| 3 chute (cross-pointbar) channel deposits | 3 no field observations | -2.0 |
| 4 braided stream deposits | 4 -6.7 ... -3.6 ... -1.8 | -4.0 |



Aerial photograph of Columbus Air Force Base and the MADE and the MADE-1HA test site.

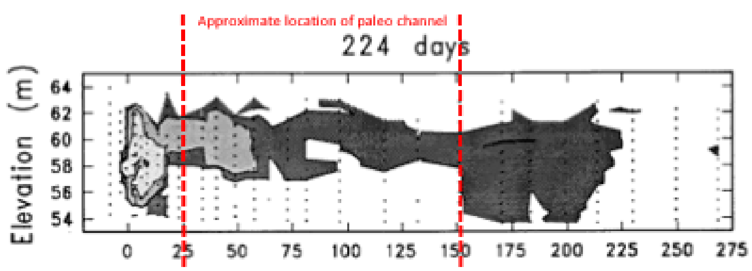
Also visible is the paleo-channel that intersects both test sites.

Reproduced from Herweijer, 1996 (fig 1&2) and 1997 (fig 6.25) showing a '*hierarchical approach combining large-scale deterministic structures and simple stochastic methods*'. This model was developed for the MADE-1HA site 100 m away from the MADE site

**Figure 6.25:** Example of conductivity (K) distribution for Layer-1 of the nested-facies model. Logarithmic gray-scale: black for $^{10}$log K = -1, white for $^{10}$log K = -7 (K in m/s, also see Table 6.5 for distribution parameters of $^{10}$log K.).

Bowling et al. (2005) shows via GPR data a similar two-layer system with a unit of coarser sand and high energy depositional bed-forms overlaying a more stratified lower sand unit. The same two-layer system can also be seen in figure 2 of the paper under review, as the borehole flowmeter K data of wells F20 and F40 show a sharp increase of K above the 56-67 m depth level.

The two-layer system is also clearly reflected in the Tritium plume of the MADE-2 experiments (see figure below from Boggs et al., 1993).




***In order to obtain a much more realistic conceptual model, the authors should use the above references on the geological background of the high K contrast. The conceptual model should also include the vertical K contrast at ~ the 57-58 m depth level as indicated by figure 2 in their paper and shown by various references cited above. The paper should also present a cross-section of the modelled vertical tracer distribution to compare with the observed vertical tracer distribution.***

1b) At the scale below the main zonation scale referred to in 1a above, the authors then insert binary K contrast 'inclusions' representing a medium scale of heterogeneity. This binary (Boolean) technique has been commonly used (eg. Haldorsen and Lake, 1984, Desbarats 1987) to create a heterogeneous architecture at various scales. The dimensions of these 'inclusions' assumed by the authors, seem not to be based on any field evidence or analogue systems. Rehfeldt (1992) shows a section of a nearby quarry with potential dimensions of these inclusions. Herweijer and Young (1991) show how pumping test data reveal some insights regarding the hydraulic continuity of these inclusions. Herweijer (1997) specifically mentions some scenarios for dimensions based on sedimentary analogues for the same aquifer. Bowling (2005) shows detailed data for the same site, where GPR sections show some of the sedimentary structures controlling heterogeneity on this scale

***In order to improve the conceptual model, the authors should elaborate on the nature of these inclusions and use the dimensional information for the inclusions contained in the above references.***

1c) At the next lower scale level, the authors use randomized K values from the borehole flow meter data. Figure 4 shows several datasets for K distribution, and there is a significant discrepancy between the mean and the range of borehole flowmeter data and the K datasets. For the borehole flowmeter the log-normal mean is a factor 5 lower than the pumping test value that represents the bulk of these BHF data (the pumping test in the high K zone – the channel deposit). The mean of the K derived from grain-size is a factor 2 higher than the pumping test value, which could indicate that the K values derived from grainsize would be more reliable to represent the high end of K values. This type of differences between hydraulic conductivity data is not uncommon. Apart from data acquisition issues, the differences between differently measured K values can often be traced back to scale effects, which are of utmost relevance to conceptualization, and should be addressed in the paper. The authors should also review for example Rehfeldt ea. (1989) and Young (1998) regarding some issues with the borehole flowmeter data specific to the MADE aquifer. Young (1995) and Herweijer (1997) show at the neighbouring MADE-1HA test site borehole flowmeter K-values for the same sediments. They publish values with a higher log-normal mean and maximum, which corroborate the grain-size K values for the MADE site as shown in figure 4 of the paper under review. This extreme end of the K distribution has at the MADE-1HA site a significant effect forming high-K pathways (Young, 1995; Herweijer, 1997)

***In order to support the conceptual model, the authors should review the meaning of the different hydraulic conductivity values as measured for the MADE site and how that impacts the conceptual model. They also should explain specifically why the borehole flowmeter data were selected, why the deviation between the mean of the borehole flowmeter data and the relevant pumping test is acceptable and, why the borehole flowmeter data are preferred to the grainsize K data (which seem to better represent this pumping test and the high end K values).***

1d) The hierarchical/nested scale approach using deterministic zonation with various levels of binary and continuous stochastic infill is not 'novel'. It has been widely used before, and the authors should reference some earlier work applying this approach (see eg: Damsleth et al, 1990; Herweijer, 1997 section 6.6, specific to the MADE aquifer – see also first figure of this review; Smith et al., 2001; Yupeng & Shenhe, 2013)

***The authors should quote above references as examples of the hierarchical method they employ, and not refer to their approach as novel, neither in general nor specific to the MADE aquifer.***

## 2 – Use of 2D model for a 3D plume

The paper presents a 2D cross-sectional model along the main axis of flow.

The field data show a major component of flow transversal to the main flow. The tritium plume picture below from Boggs et al., 1993) shows that initially the tracer released at the 5 injection wells converges and subsequently diverges. An elongated finger of the plume shows further downstream, but probably already developed closer to the source area. This finger is probably related to some very high K pathways related to sedimentary structures that are highly anisotropic and directionally variable (Young, 1995; Herweijer, 1997).

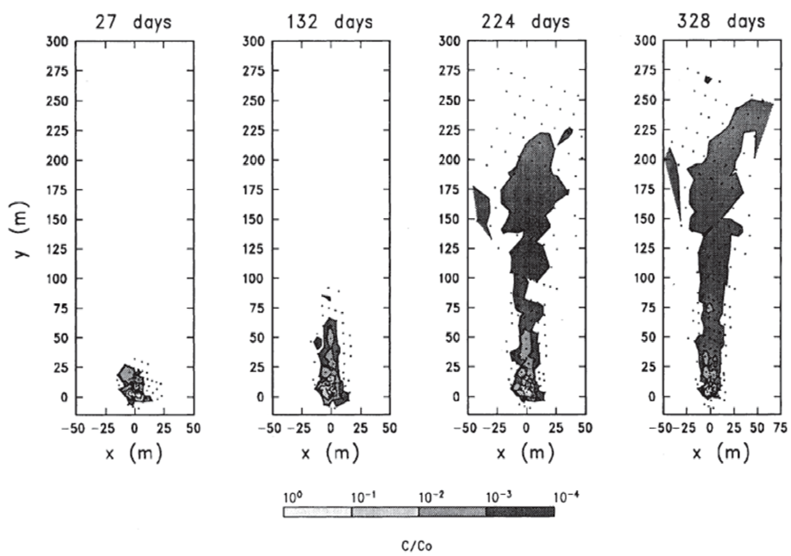



Figure 4-2 Horizontal Sections Through Tritium Plume at Elevation 59.5 m at 27, 132, 224, and 328 Days

x

The Bromide plume shows similar transversal movement (including a sharp sideways movement close to the source) and downstream patchiness (see eg. Julian et al., 2001, fig 5 &13)

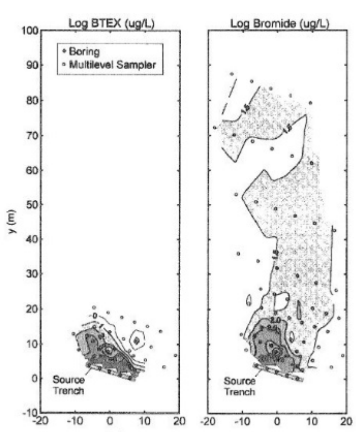



Figure 5. Depth-averaged total BTEX (left) and bromide (right) concentration distributions 278 days after source emplacement.

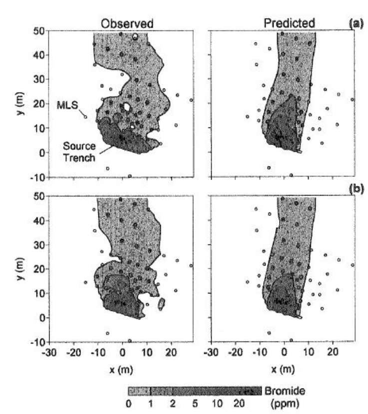
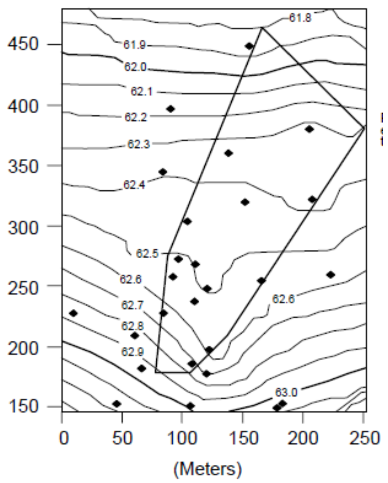



Figure 13. Comparison of observed and calculated bromide peak concentrations for (a) snapshot 3, 152 days after source release, and (b) snapshot 4, 278 days after source release.

The piezometric surface map around the injection zone (Rehfeldt et al., 1992, see figure below) shows a complex convergent and then divergent pattern. This is probably due to refraction of the flow lines at the boundary of two zones with very different K values and which occurs at an angle to the regional flow direction (Freeze and Cheery, 1979).




***Given a clear horizontally anisotropic flow and transport pattern, a 2D analysis/model is very limited and unrepresentative. A 3D model should be used. The results of the 2D model cannot be used for quantitative analysis.***

## 3 – Potentially misleading remarks re use of geological data and geological models

The paper makes (line 53-57, 64) statements re. the use of geological data (training images) and geological models referring papers that are 20+ years old (Koltermann and Gorelick, 1996; Herweijer, 1997).

In line 54, the authors dismiss training images as limited available and unrepresentative. This statement is incorrect: training images are quite widely available sourced from satellite images (Google Earth) and extensive literature on geology, sedimentology and paleogeography. This holds especially for relative recent shallow deposits which form the MADE aquifer, and which are often the subject of groundwater modelling efforts.

Herweijer (1997) provides a number of references to the sedimentology and paleogeography of the MADE site, which would provide a very good start with respect to representative training images. Ronayne et al. (2008) give a good example of a model based on training images to model a hydrogeological test site.

The paper also states that geological models as used in the petroleum industry have not found their way into applied hydrogeology (line 64), a statement which is quite strong, and in my view incorrect. Alloisio, (2011), Dowling et al. (2013) and Peereboom (2018) are examples of applications for a variety of shallow to deep aquifers. Even if it is the case that this type of modelling has not 'found its way' into widespread use in hydrogeology, this is not a reason to simply to set it aside from a research point of view. The authors should explain why their approach is 'better' than the standard geological modelling methods used in the petroleum industry and the hydrogeological applications of these methods referenced earlier in this paragraph.

***The authors should re-assess the literature on the above matter and correct their statements about the use of geological data and models.***

# Some further issues in need of clarification

## Line 335 and last section of supplement

As earlier discussed, the MADE plumes are variable in 3D and should really be modelled in 3D

The sections explaining the 3D modelling effort to confirm the validity of the 2D model are confusing. It is unclear if the 2D model was simply copied in the 3rd dimension, ie. is completely symmetric in the 3rd dimension, or if some sort of heterogeneity in the 3rd dimension was included. The supplement also states that the extension in the 3rd dimensions has no impact because of the binary nature of the ‘inclusions’ vs. K values that change gradually. If transport being restricted to a 2D cross-section of a 3D model is a significant impact of the binary conceptualization, than the binary conceptualization should not be used, as gradual changes of K are the norm for non-fractured/fissured aquifers such as the MADE aquifer

It would also be informative if the authors would present a visualization (map or x-section slices) of the 3D model and the modelled plume. It should also be looked at if a small level of transversal dispersion (representing very small scale heterogeneity) would have a significant impact on the 3D version of the 2D model.

## Line 330- 334 and section ‘details on flow and transport’ section in supplement

The injection rate modelled is Qin = 1.166e-5 m3/sec. The injection rate quoted by Boggs et. al. is 10.07 m3/48.5 hr which converts to 5.57e-5 m3/sec. If any adjustment has been made to the injection rate (perhaps to adapt for the 2D model setting?) this has to be clarified.

The paper states: *We use a flux related injection representing natural conditions. For technical details, the reader is referred to the Supporting Information.* This is unclear and there seems to be no further specific discussion on ‘*flux related injection’* in the supporting documentation.

## Section 3.4 - calibration and predictive capacity of model

The model matches data in a line downstream of the injection, but as discussed earlier in the review (point 2) does not represent any 3D movement of the plume close to the injection point and further downstream. The paper also does not show a vertical cross-section of the observed vs the modelled tracer distribution. Hence any calibration is quite limited and potentially has limited predictive value

As discussed in the paper (line 270) there is significant uncertainty due to the mass balance issues with the bromide tracer (~50% in snapshots past 300 days). The calibration could be (should be?) tested using the MADE-2 tritium tracer test (Boggs et al., 1993), which seems to have resulted in a better mass balance (77% in final snapshot at 328 days).

## Title

There is a grammar error in the title (the part which reads ‘in a Heterogeneous Aquifers’). It should be either ‘in a Heterogeneous Aquifer’ or ‘in Heterogeneous Aquifers’

## References


Alloiso, S, 2011, Integrated use of Petrel© and Modflow in the modelling of water injection and effects on a Quaternary aquifer - https://www.researchgate.net/publication/263611730_Integrated_use_of_PetrelC_and_Modflow_in_the_modelling_of_water_injection_and_effects_on_a_Quaternary_aquifer_Usage_integre_de_PetrelC_et_Modflow_dans_la_modelisation_de_l'injection_d'eau_et_effets_dan

Boggs, J.M, Beard, L.M., Waldrop, W.R., 1993 Transport of tritium and four organic compounds during a natural-gradient experiment (MADE-2), report TR-101998, EPRI, Palo Alto, CA, USA.

Bohling, G. C., Liu, G., Knobbe, S. J., Reboulet, E. C., Hyndman, D. W., Dietrich, P., & Butler, J. J. (2012). Geostatistical analysis of centimeter-scale hydraulic conductivity variations at the MADE site. Water Resources Research, 48(2). doi:10.1029/2011wr010791

Bowling, J. C., Rodriguez, A. B., Harry, D. L., & Zheng, C. (2005). Delineating Alluvial Aquifer Heterogeneity Using Resistivity and GPR Data. Ground Water, doi:10.1111/j.1745-6584.2005.00103.x

Damsleth, E.; Tjolsen, C. B.; and Omre, K. H., 1990, A two stage stochastic reservoir model applied to a North sea reservoir: SPE paper 20605 presented at the 1990 SPE Annual Technological Conference and Exhibition, New Orleans, La., September 23-26.

Desbarats, A.J., 1987, Numerical estimation of effective permeability in sand-shale formations: Water Resources Research, v. 23, no. 2, p. 273-286.

Desbarats, A.J., 1990, Macro-dispersion in sand-shale sequences: Water Resources Research, v. 26, no. 1, p. 153-163.

Dowling, J. , B. Zimmerlund, J. Keneally, E. van den Burg, 2013. An integrated model for mine dewatering at the Bagdad mine. In Wolkensdorfer, Brown and Figuora (editors), INWA, https://www.imwa.info/docs/imwa_2013/IMWA2013_Dowling_545.pdf

Freeze R.A. and Cherry J.A., 1979, Groundwater, Prentice Hall

Haldorsen, H. H., & Lake, L. W. (1984). A New Approach to Shale Management in Field-Scale Models. Society of Petroleum Engineers Journal, 24(04), 447–457. doi:10.2118/10976-pa

Herweijer, J.C., and Young, S.C., 1991, Use of detailed sedimentological information for the assessment of aquifer tests and tracer tests in a shallow fluvial aquifer: Proceedings of the 5th Annual Canadian American Conference on Hydrogeology, Calgary, September, 18-20, 1990. NWWA, Dublin (Ohio)

Herweijer, J.C, 1996. Use of sedimentology and geostatistical modelling to estimate uncertainty of groundwater models. Proc. International Conference on Calibration and Reliability in Groundwater modelling (ModelCARE 96), Golden (CO, USA), September, 1996. https://pdfs.semanticscholar.org/a5a5/25d8da8091bb59a59795262d932f0b4a6333.pdf

Herweijer, J.C., 1997. Sedimentary heterogeneity and flow towards a well. Ph.D. dissertation (in English), Free University, Amsterdam, NL, https://www.hydrology.nl/images/docs/dutch/1997.01.07_Herweijer.pdf

Julian, H. E., Boggs, J. M., Zheng, C., & Feehley, C. E. (2001). Numerical Simulation of a Natural Gradient Tracer Experiment for the Natural Attenuation Study: Flow and Physical Transport. Ground Water, 39(4), 534–545. doi:10.1111/j.1745-6584.2001.tb02342.x

Koltermann, C. E. and Gorelick, S. M.: Heterogeneity in Sedimentary Deposits: A Review of Structure-Imitating, Process-Imitating, and Descriptive Approaches, Water Resour. Res., 32, 2617–2658, https://doi.org/10.1029/96WR00025, 1996.

Peereboom E, 2018, Exploring electrofacies for property mapping of the Triassic: an improved approach for geothermal development in Brabant, MSc Thesis (in English), University of Utrecht, NL

Rehfeldt KR, Hufschmied P, Gelhar LW, Schaefer ME, 1989, Measuring hydraulic conductivity with the borehole flowmeter, Report EN-6511, EPRI, Palo Alto, CA, USA.

Rehfeldt, K.R.; Boggs, J.M.; and Gelhar, L.W., 1992, Field study of dispersion in a heterogeneous aquifer – 3 - Geostatistical analysis of hydraulic conductivity: Water Resources Research, v. 28, no. 12, p 3309-3324.

Ronayne, M.J., S.M. Gorelick, and J. Caers. 2008. Identifying discrete geologic structures that produce anomalous hydraulic response: an inverse modeling approach. Water Resources Research 44. DOI:10.1029/2007WR006635

Smith, R., W.A. Bard, J. Corredor, J.C. Herweijer, S.R. McGuire, A. Antunez, T. Block, N. Lazarde, 2001. Geostatistical Modeling and Simulation of a Compartmentalized Deltaic Sequence, Ceuta Tomoporo Field, Lake Maracaibo, Venezuela. SPE paper 69572, presented at the 2001 SPE Latin American and Caribbean Petroleum Engineering Conference, Buenos Aires, Argentina, 25–28 March 2001

Young, S.C., 1995, Characterization of high-K pathways by borehole flowmeter and tracer tests: GroundWater, v. 33, no. 2, p. 311-318
https://ngwa.onlinelibrary.wiley.com/doi/abs/10.1111/j.1745-6584.1995.tb00286.x

Young, S.C., 1998, Impacts of Positive Skin Effects on Borehole Flowmeter Tests in a Heterogeneous Granular Aquifer GroundWater, v. 36, no. 1, p. 67- 75
https://ngwa.onlinelibrary.wiley.com/doi/epdf/10.1111/j.1745-6584.1998.tb01066.x