

This work is of great interest and provides new insights in coupling hydrological and geochemical processes at watershed scale. The authors explore the development of a dimensionally-reduced model coupled with a reactive transport model. The modelling approach is validated by field data collected in springs and piezometers at wet and dry periods, which allows for the investigation of the spatial and temporal variability of transit times and reaction rates. The authors conclude on a hydrological control on the chemostatic behaviour of the watershed, due to fairly constant transit times despite highly varying flow dynamics. The conclusions of the paper thus strongly rely on the hydrological model simulating a low variability of water transit times, from which the geochemical model logically simulate a low variability of geochemical concentrations. To my point of view, a more detailed description of the hydrological model is needed to strengthen the confidence in the results (see specific comments). I recommend moderate revisions before publication of this interesting work, mainly because information is missing in the method section and because the study would gain from a more detailed analysis of the results.

Specific comments:

Figure 1. Please add a x scale.

I. 175 refer to Figure 1

METHODS

I.179 A lot of geochemical data seem available, both for springs and piezometers, but only some of them are used to validate the models. Why are you only using some of the available data? The model would be improved with a validation on all available data that span a long period of time. And if not, the authors should justify why they use only a limited part of their dataset and how they chose the data used.

I.200 “The exchange of water between the surface and subsurface flows are addressed via the hydraulic head differences between the compartments.” A hydraulic conductivity value of the interface between the two compartments must be also considered. Which value was used? Was it calibrated or fixed? Please clarify.

The model parametrization paragraph needs clarification and additional precisions.

- “several zones of heterogeneity” please refer to the Figure
- Please specify all parameters instead of “other parameters” I.206 for clarity
- I guess the aquifer thicknesses given Fig. 2 correspond to the values obtained after the calibration. If so, please be specific on the figure or in the legend. Also add in the legend that the grey lines show the mesh grid of the hydrological model (if it is the case).
- The hydraulic conductivity is calculated using with the Van Genutchen model (which involves parameters such as the saturated hydraulic conductivity, n and α), and therefore is a result of the simulations at the grid size, is it correct? I don’t think it is clearly said in the paper. As the temporal variability of the hydraulic conductivity is an important point of this study, I think a very clear explanation of how the depth-integrated hydraulic conductivity is calculated would make it easier for the reader to follow the logic. Fig.6, do the red colour (highest hydraulic conductivity) correspond to the value of the calibrated saturated hydraulic conductivity? If yes, it might be worth saying it, as it shows that the subsurface compartment is fully saturated at high flows.
- How is the calibration realized? Which algorithm was used? Could the authors add some uncertainty estimates on the calibrated parameters?

- The hydrological model over the whole catchment is calibrated only on the discharge time series located at the outlet of the catchment. Do the authors have other hydraulic data they could use to strengthen their calibration, such as piezometric heads or spring discharge rates? If not, I am worried that equifinality might not be negligible, which also points for a serious estimation of parameter uncertainties and/or sensitivity analysis (see previous comment).

The hydrogeochemical modelling strategy presented Figure 5 deserves more details. For each flow line, how much is “several”? What is the value of the “constant distance along the flow line” that is used? Do the boxes drawn Figure 5 correspond to the grid of the NIHM model or not? KIRMAT simulations were performed for different flow lines independently and then mixed at the outlet of the flow lines, if I got it correctly (Figure 5). The mixing could also occur all along the flow line, each percolated soil water mixing with the water coming from the upstream “box”. Would the results be different? What is the justification for no mixing between flow lines within the subsurface compartment?

RESULTS

I. 320 Where are the results of the water velocities?

I. 328 Please specify which characteristics of the flow lines you consider similar (geometric characteristic? length, position...), as the velocity along the flow lines for instance differ from given dates (4.2). Maybe not talking about water velocity along flow lines in this paragraph might be less confusing.

I. 341 “or parameters” Parameters should not change under transient conditions as they have been previously calibrated, right?

I. 344 Please describe the driving factors of the spatial and temporal change in the simulated hydraulic conductivity. How much is this result related to the geometry of the watershed (small thickness, steep topography...)?

I. 385 refer to Figure 7

Figure 7. Where do the uncertainty bars come from in the KIRMAT simulated concentrations? Please, clarify what you mean by “induced” in the legend. A priori it could not be induced by the hydrological model as uncertainties are not taken into account. Is it only coming from the propagation of soil solution uncertainties?

Figure 8. The CS1 geochemistry was simulated at 6 different dates, but we cannot tell which of the observation blue point corresponds to each simulated orange point on the Figure. This information is needed to assess the reliability of the modelling. Maybe using different colours, please link each simulated point to the corresponding observation point.

Figure 9. and I. 403-405. The sentence is quite vague. The authors chose to present the differences between elements in a figure, which is very interesting. But then their explanation for the differences remain very vague and not specific to any element. Could the authors expand on the geochemical mechanisms yielding to the different C-MTT relations?

PZ3 and PZ5 piezometers. You already specify paragraph 4.2 that, such as for CS1 and CS3 springs, single flow lines can be used, so you can shorten the first part of the section. Maybe this section could be merged with the previous one as the approach and the results are similar.

CS2 and CS4 springs. Same comment as above, you can probably shorten the justification of the scattered flow line distributions as it is a repetition of paragraph 4.1

DISCUSSION

Have the standard kinetic constants used here been determined on minerals all coming from the Strengbach catchment? Then would you recommend to use site-specific kinetic constants to account for local aging effects?

I am not sure if Figure 13 brings anything new. It might be more interesting to show the distribution of dissolution rates over the catchment.

The discussion on the chemostatic behaviour would gain from a more concise and straightforward argumentation. Some repetitions with the result section could be avoided. For instance, the whole paragraph describing changes in the simulated hydraulic conductivities, in water velocities and in mean transit times should not be detailed as much or details should be moved to the result section 4.2. Lines 693-699 aim to justify the modelling approach and results (same holds true for the discussion on the clay solid solution), which could be moved elsewhere for clarity. I would advise to refocus the whole paragraph only on the discussion on the potential origins for the chemostatic behaviour.