Author's Response on review of our manuscript HESS-2014-207 "Iron oxidation kinetics and phosphate immobilization along the flow-path from groundwater into surface water" by Søren Jessen

First of all, we would like to thank Søren Jessen for his effort for thoroughly reviewing our manuscript and his valuable comments which we will take in full consideration in improving our paper. Our replies to his comments and the changes in our manuscript are listed below.

General comment

1: The model for Fe oxidation rate is applied with average pH (and temperature and pO_2*) while the kinetics are in fact non-linearly dependent on pH (Eq. 1). For example, using Eq. 1, the amount of oxidized Fe over, say, two days at pH 7 would be much less than the amount of Fe oxidized during one day at pH 6 plus one day a pH 8, all other parameters fixed; the vast majority of the Fe in this example would be oxidized during the one day at pH 8. This makes direct comparison of models and observations (Fig. 6) somewhat difficult. To this one may add the non-linear isotherm for phosphate adsorption to iron oxides. The error thus introduced in the model prediction must be addressed thoroughly by the authors, and the manuscripts conclusions reevaluated accordingly. This is my only major concern regarding the manuscript; everything below is minor*

It is true that we modeled the Fe oxidation rate with a non-transient pH value and that in the real world the pH might change during the Fe oxidation process. The pH of the reservoir water varies roughly between 6 and 6.9 and change during oxidation will be small because it is buffered by the alkalinity of the exfiltrated groundwater. As the kinetics are non-linearly dependent on pH this influences the Fe oxidation rate during the oxidation process. In this, we agree with the reviewer. For reasons described below, the objective of our model calculations was not to reconstruct the individual Fe measurements in the reservoirs but to compare the range and annually average trend of measured Fe concentrations with increasing transit time of the water within the reservoirs with model scenarios. Therefore, we performed three model scenarios representing a typical summer situation, a typical winter situation and an annually average situation with their accompanying pH values (resp. 6.2, 6.6 and 6.4) and other input parameters. The results are plotted together with observed concentrations against the transit time. We improved the description of this objective in the manuscript (p.6 l. 21-26)

We agree with the reviewer that modeling of the individual Fe measurements followed by a comparison of measured data points and modeled data points would be the ultimate model goal. Ideally, physically-based distributed models are applied and all parameters are independently obtained in field application of transport models. In a previous study (Van der Grift & Griffioen, 2008) we argue that this can seldom be reached in field studies considering water quality. In fact, this can only be reached for systems that are rather homogeneous in their spatial characteristics, have a temporally constant input, and where a limited amount of processes is operational. If this is not the case, calibration of process parameters may be applied to improve the model performance. A problem with calibration is that non-unique calibration solutions may exist even for these relatively simple field systems or that there are no objective calibration criteria. Another kind of approach is based upon use of best available information using independently estimated or measured parameters whether or not combined with a sensitivity analysis. We use this approach to model the Fe oxidation (and P immobilization) process as function of the reservoir transit time and compared the 'statistics' of the dataset with measured values with typical model scenarios. To our opinion the modeling exercise at our experimental field can best be carried out according to this approach for the following reasons: 1) the Fe_i and P_{O2} spatial and temporal variable and not exactly known; 2) the pH value is known at time of sampling but is probably transient during Fe oxidation; 3) there is some uncertainty in the calculated transit time due to uncertainty in reservoir discharge and reservoir volume measurement. We are sure that it is possible to reconstruct the individual Fe measurement points by modeling with a good match between model predictions and observations by adjusting the Fe_i and P_{O2} in a reasonable range (the Fe concentrations in the groundwater ranging between 0.2 and 45.5 $mg.L^{-1}$). However, we don't see added value of this kind of model calculations above our approach Using ''best available information'', that is even based on fundamental kinetic parameters

Because pH and pOH are log values, it would be more correct to use in Eq. 1 the average of [OH-]^2 (average of 'OH- activity squared') instead of a value arrived to via an average pH of pOH.

Regarding the use of the average pH vs. the average of the $[OH^-]^2$: The reviewer is right that we used the average of measured pH values as input instead of the average of $[OH^-]^2$ as input for the model. PHREEQC is used to calculate the $[OH^-]^2$ from a given pH to deal with the temperature dependency of the dissociation constant of water and the ion strength. For representing a typical summer, winter and year average situation we prefer to use a typical pH value instead of an average $[OH^-]^2$. We as well calculated the average pH based on average $[H^+]$ from measured pH values. These values (6.17 for winter; 6.53 for summer and 6.31 for the year) slightly deviate from the values that we used in the manuscript (6.19 for winter, 6.60 for summer and 6.36 for the year). We followed the reviewer's suggestion and plot the lines for average pH based on the average $[H^+]$. This is a brief sensitivity analysis for the model results for the pH. The results are shown in the graph below. The black and blue lines are de Fe concentrations according the model from our original HESSD paper. The red lines give the model results adjusted pH values. The differences in Fe oxidation rates between the original and adjusted model are minor and to our opinion do not impact our conclusions. We use the red lines of the revised version of our manuscript.

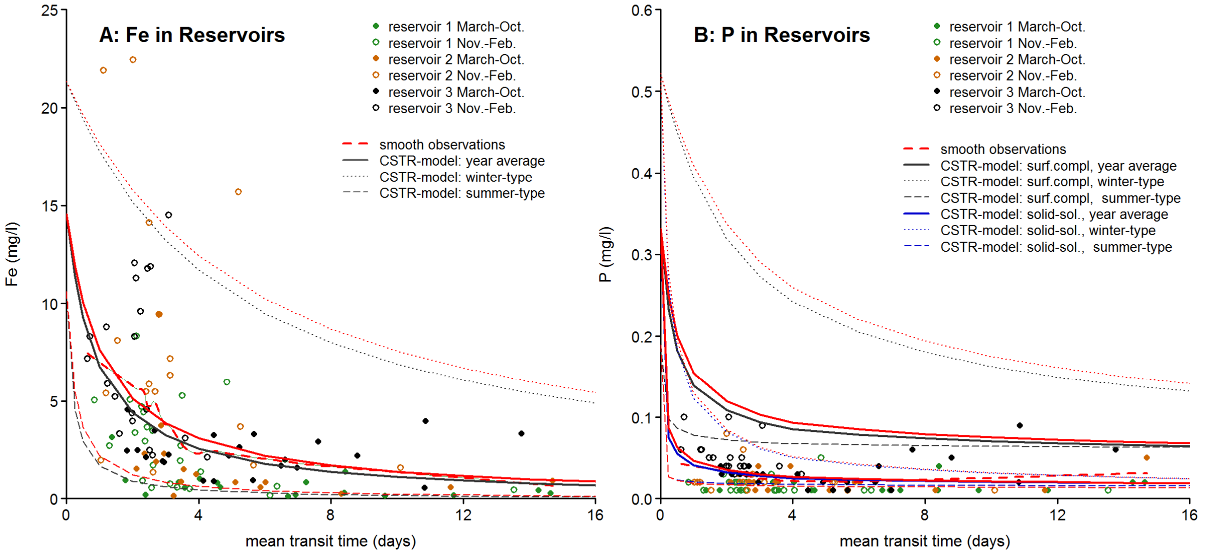

Figure 1. Sensitivity analysis of model results on pH, black and blue lines are de Fe concentrations according the model from our original HESSD paper, with pH values of 6.19 for winter, 6.60 for summer and 6.36 for the yearly average conditions. The red lines give the model results adjusted pH values of 6.17 for winter; 6.53 for summer and 6.31 for the yearly average conditions.

2) *The authors need to adapt a clearly defined terminology for “surface water”, “groundwater”, “reservoir water” and “ditch water”. For instance, in p. 6650 l. 12 the term “surface water” apparently means “reservoir water”; in other places (e.g., Fig. 2) “surface water” refers specifically to water sampled at the “catchment” and “subcatchment” sampling points. In my opinion, the moment groundwater has entered the reservoir it is to be considered as “surface water”; albeit I do acknowledge the complication that tube drain water is not included in the reservoir.*

Strictly, the “reservoir water” is, as the reviewer suggests, surface water. To avoid confusion between the reservoirs, and the sampling locations at the sub-catchment and catchment outlet we prefer to use terms “reservoir water” for the reservoirs and “surface water” specifically for the sub-catchment outlet and the catchment outlet sampling points. Then it is indeed better to avoid the term surface water for other situations like p. 6650 l. 12. In these cases it is better to use the term ditch water. We rechecked the manuscript and revised when needed.

3) *Related to the above: The authors should be more clear about the scope. I see some difference in studying the P retention in surface water as compared to during exfiltration to surface water or along the flow path from groundwater to surface water (the latter in my view is also solely before the groundwater enters the ditch). An example is the first sentence of the abstract (p. 6638), but other examples are p. 6640 l. 25 and p. 6641 l. 3-4. With the sampling points used, the authors are both capable of making conclusions about Fe and P retention along the flow path from groundwater to surface water, and about Fe and P retention at increasing flow distances (or “reaction times”) downstream the in the catchment. The differences should be made clear and could be utilized more actively in the interpretation. The title could be changed to comply with the above changes, e.g. to*

something along the lines of " Iron oxidation kinetics and phosphate immobilization during exfiltration of anaerobic groundwater and in surface water".

We agree with the reviewer that this might be confusing. The scope of the manuscript is (bio)geochemical processes affecting Fe and P concentrations, and therefore P retention, during flow from groundwater into surface water and not something like P retention downstreams a catchment due to sedimentation and resuspension of particulate P. We deliberately chose the words '*into* surface water' in the title instead of '*to* surface water' to point out that we consider the processes in the surface water as well. We checked if this is consequently done throughout the manuscript and revised when needed.

4). *Could the authors please provide in the manuscript more detail of the models used, including the solid solution model. It is for instance unclear if the solid solution model used in combination with the surface model, as might be perceived from p. 6654 l. 26-29.*

We provided more detail of the models used p.7 l.18-24.

5) *The sections 4.5 and 4.6 contain mainly literature review on subjects where the authors have presented little if any data from the present study to compare with. I suggest the two sections to be either omitted (to make room for other improvements) or rewritten and combined (together with parts of section 4.4) into a shorter "implications" section.*

These sections describe as the reviewer points out 'implications' of our study that are not supported by data. To our opinion these implications are a valuable part of the discussions. We do support the suggestion of the reviewer and rewritten this into a 'implications' section of the discussion.

6) *When it comes to flow direction, quite a lot of deviation from the general flow direction can happen over the 20 m distance from the groundwater wells and to the ditch/reservoirs. Therefore, the hydrogeology and hydrology, including groundwater head distribution, needs to be better described in the text and Fig. 1 needs some improvement. Especially since some effort is done discussing which well represents groundwater arriving to which reservoir (cf. p. 6646 l. 4-5). For example, Van der Velde et al. (2010)'s similar figure is much better. On the other hand, if some discussion on the relation between groundwater well observations and individual reservoir concentrations is omitted, then a thorough description of hydrology/hydrogeology may not be needed.*

We agree with the reviewer that the groundwater flow can be quite complex in this part of the field. Although it might be interesting, we do not aim to do a sophisticated study of the groundwater flow over the 20 m from the wells to the ditch. The groundwater flow in the field is approximately from northwest to southeast (see fig 5 in Rozemeijer et al., 2010). Therefore, we just stated that it is conceivable that the difference in measured Fe concentrations between the reservoirs can be attributed to differences in Fe concentrations within the field and that the variation of Fe concentrations between the different wells and the general flow direction within the field is consistent with the high Fe concentration in reservoirs 2 and 3.

We made a reference to Rozemeijer et al., 2010 and we improved Fig. 1

7) *Quite a lot of speculation of what is going on in tube drain 3 is presented, without substantial supporting data (e.g., p. 6650 l. 6-9). I suggest this speculation be reduced to increase focus on the very many other interesting subjects of the study.*

The objective of discussing the data from our tube drain (especially drain 3 where the redox change occurs) is that the tube drains are an important route of groundwater flow into surface water and this topic lies within the scope of our study. We want to point out that the domain where a major part of the Fe oxidation takes place can shift in time from the soil to the surface water (after discharge to the ditch). For our field site this has, surprisingly, minor implications for P in drain water. Like the reservoir water, P concentrations of the tube drain water were always low, also during periods of high Fe concentrations. The dynamics of P in the drain water is consistent with that in the reservoir water.

Specific comments

9) *p. 6638 l. 24-27: Please clarify whether the term particulate phosphate as used here covers still mobile or immobilized phosphate. As I recall, no original data for mobile particulate P is presented in the manuscript, so perhaps conclusions on this regard should be left out from the abstract.*

Although our model calculation point to a rapid turnover of dissolved phosphate to a precipitated Fe(III)-phosphate type of solid during oxidation of Fe, we only collected a single Fe-oxide sample from a tube drain (p. 6653 l.20) This sample had the P/Fe ratio comparable with the groundwater. So, we agree with the reviewer that this conclusion is not thoroughly supported by measured data. Therefore we used the words '*seems an important geochemical mechanism'.* Moreover, we agree with the reviewer to distinguish in terminology between "particulate" and "precipitated" (specific comment 35.) and see this as a valuable comment. We use *structural phosphate* instead of *particulated phosphate* in the revised manuscript.

10) p. 6639 l. 12: Please clarify if they infer a difference the term particulate bound phosphate and the term particulate phosphate used in the abstract. If the same thing is meant, I suggest using the same term for clarity

We use *particulate phosphate*

11. p. 6640 l. 1: Could the authors please consider not introducing the term iron oxidation process. In my opinion it will not help improve clarity in the text. The term oxidative precipitation of Fe(II) (used shortly below in l. 14) describe the same thing and is better in my view.

We removed the introduction of them iron oxidation process.

12. p. 6640 l. 7-9: Could the authors please provide a reference for the statement to be valid specifically for lowland catchments?

We changed this statement and rearranged this part of the introduction : 'To our knowledge there are no field studies on the mechanisms and rates of iron oxidation with associated binding of phosphate in lowland catchments that drain anaerobic groundwater while these processes may have an important control on mobility of P in surface water (Baken et al., 2013;Griffioen, 2006;Fox, 1989). Understanding these processes is important to quantify P retention in such areas.'

13. p. 6640 l. 13-19. The authors closely cite Spiteri et al (2006) to argue that Fe oxidation rate is controlled by pH more than by pO2 and that a pH increase in mixed freshwater (pH 5.7) seawater (pH 7.9) resulted in a seven-fold Fe(II) oxidation rate increase. The citation is correct a little bit confusing: the seven fold increase reported by Spiteri et al. is the result of increase in pH of just 0.4-0.5 pH unit which take place at initial part of the pH gradient in Spiteri al.'s modelled cross section. An increase from pH 5.7 to 7.9 would increase the rate by about 50000-fold. I suggest the authors save space by just citing Spiteri et al. for the fact that pH is a more controlling factor than pO2, and then just make reference to Eq. 1 to explain why (OH- activity is squared, pO2 is not).

Good suggestion, we have done this

14. p. 6641 l. 4-5: I suggest the authors delete the second experimental, because only the field set-up is experimental, not the catchment.

We have done this

15. p. 6642 l. 10 (and elsewhere): in-stream could be replaced by in-ditch for improved clarity.

We prefer to use in-stream to be consistent with the papers from Van der Velde and Rozemeijer on the same experimental field site.

16. p. 6643 l. 9: Could the authors please elaborate on how the pH and temperature (not Temperature) were carried out. For instance, if a flow cell was used...?

We have done this

17. p. 6643 l. 21: I assume that concentrations should be replaced by, e.g. oxidation kinetics.

We have done this

18. p. 6643 l. 22-23: The authors can omit homogeneous (cf. Vollrath et al., 2012) and in laboratory systems (because laboratory systems are very different from each other).

We have done this

19. p. 6644 l. 8: If k in Eq. 2 is the same as K in Eq. 1, please consider using the same case (either k for both or K for both).

It is not exactly the same we now use *k* and *k'*

20. p. 6644 l. 20: ...an ideal solid solution...

done

21. p. 6647 l. 8: ditch water or reservoir water?

This was not clear in the manuscript. The water temperature was measured continuously direct downstream the reservoirs. We adjusted this in the study area section

22. p. 6647 l. 9-14: The author's discussion of the extent of oxygen saturation should be moved to the Discussion (since no data are available to back up these considerations) and merged with the part on *p. 6652 l. 11-17. Also, the authors should provide references to support their statements. For instance, the present formulations does not seems to acknowledge that O2 concentrations must be controlled by a balance between the rate of O2 consuming processes such as Fe(II) oxidation, respiration, and organic matter degradation, and O2 supply by diffusion and photosynthesis.*

We agree with the reviewer, merged p. 6647 l. 9-14 in the discussion and provide references to support the statement: ‘Biological, chemical and physical processes control the dissolved oxygen concentration of drainage ditches (Kersting and Kouwenhoven, 1989). Therefore the oxygen concentration of our reservoirs will be variable over the year and even over the day. However, low oxygen saturation levels or even hypoxia are not likely for the reservoir water. Hypoxia is not an uncommon phenomenon in ditches with free-floating plant mats (Verdonschot and Verdonschot, 2014;Kersting and Kouwenhoven, 1989) but this was not the situation for our reservoirs. Moreover, because of a high degree of atmosphere–water contact (the reservoirs have typically a small wetted volume as opposed to its wetted perimeter) equilibrium with atmospheric oxygen might be possible’.

23. p. 6647 l. 12: …93% of the time period.

We removed this sentence

24. p. 6647 l. 21: Please rewrite; reservoir volumes appear twice.

done

25. p. 6648 l. 4: replace felt dry by went dry.

done

26. p. 6648 l. 23: I guess the authors mean the OH- activity rather than concentration.

done

27. p.6648 l. 13: I suggest the authors use the phrase consistent with instead of indicating, because the observation mentioned does not indicate where the removal takes place. This might be partly alleviated by presenting the supporting data, namely Fig. 7, in the Results section.

We removed this part

28. p. 6649 l. 6: Please consider replacing changes by increases to improve clarity of the text. Also replace of surface water by from surface water.

done

29. p. 6649 l. 8-10: Please move to Results section.

done

30. p. 6649 l. 17: …degassing of groundwater with a pH of 6.16…

done

31. p. 6649 l. 17-21: Please comment on how this compares with Fig. 4?

These pH values are somewhat lower than a majority of pH measurements of the reservoir water during the summer months (Fig. 4). So degassing of the surface water is not the only process that induces the seasonal pH increase.

32. p. 6649 l. 24 (and elsewhere, including Fig. 6a+6b+caption): Please make sure that to distinguish correctly between mean transit time and transit time.

done

33. p. 6649 l. 26: The authors may consider to replace at moments on which with when.

done

34. p. 6652 l. 18: The authors might consider moving the statements about microbial Fe(II) oxidation down to p. 6653 l. 4.

done

35. p. 6654 l. 5-6: The authors did not measure particulate P, so they cannot make conclusions regarding whether lost dissolved P is particulate (and hence quite mobile in my terminology) or adsorbed to immobile surfaces or precipitated.

done

36. p. 6654 l. 7: depleted means totally "totally gone" in my word book (but alright, English is not my native language; I don't think neither P nor Fe became depleted.

We changed this into: 'This resulted in nearly complete P depletion in the water before Fe(II) was depleted.'

37. p. 6654 l. 9: ...black and blue...

done

38. p. 6654 l. 12: ...of a solid solution...

done

39. p. 6654 l. 17, first word: replace de by the.

done

40. p. 6654 l. 20: ...precipitation...

done

41. p. 6654 l. 23: Delete the after The.

done

42. p. 6655 l. 17-18: Please clarify whether the reduction of the Fe(II) oxidation rate and the increased inflow took place in winter or summer for ease of reading.

done

43. p. 6655 l. 21: I believe the figure reference should be to Fig. 2c and not Fig. 3, right?,

right

44. p. 6656 l. 1: Please clarify the scale under consideration: e.g., do the authors mean the sediment-interface of a grain or mineral.

We changed this into: ‘at the groundwater-surface water interface of a small river’

45. p. 6656 l. 3-4: Please rephrase sentence.

We changed this into: ‘We show that dynamics in redox processes as well may impact the Fe(II) oxidation process and phosphorus immobilization during flow from groundwater to surface water’.

46. p. 6656 l. 12-14: To me it is surprising that only three times as much P (units of e.g. mg/kg sediment?) was found in the “Fe oxide ring”, compared to unaffected surrounding soil, if this zone is thought to be continuously P retarding. Could the authors please elaborate on how this makes sense?

We added: ‘Although the Fe content of the soil samples and the P concentrations of the groundwater were not measured this indicates at least a partial retardation of dissolved P in this soil domain”.

47. p. 6656 l. 16-17: No data for the amount of Fe-oxides flocs created or their contribution to the suspended sediments is presented. If I am wrong, the authors are encouraged to make the data much more visible.

We indeed do not have quantitative data on the amount of Fe-oxide floces. However, it is clear that these flocs will contribute to the suspended sediment concentration. Therefore, we changed this sentence into: ‘The formation of Fe-oxide flocs in the surface water will contribute to the suspended sediment concentration of this surface water’.

48. Fig. 4, caption: Please replace groundwater with reservoir water.

done

49. Fig. 5, caption: …inside… (not insight).

done

50. Fig. 6: Isn’t it the mean transit time on the x-axes?

done

Author's response on review of our manuscript HESS-2014-207 "Iron oxidation kinetics and phosphate immobilization along the flow-path from groundwater into surface water" by Steeve Bonneville

First of all, we would like to thank Steeve Bonneville for his effort for reviewing our manuscript and his valuable comments which we will take in full consideration in improving our paper. Our replies to his addressed issues and the changes in our manuscript are described below.

1) *The preservation of the samples from oxidation*. We indeed did this in the field directly after filtration by acidification with $HNO_{3.}$ We agree with the reviewer that Fe colloids will go through the 0.45 µm filters and dissolve in the acid media. So, our measured Fe concentration may partly (for the drain water and reservoir water samples) or largely (for the surface water samples) be attributed to dissolved Fe(III) colloids. For the surface water samples taken at the sub-catchment outlet and catchment outlet we discussed this on p. 6651 l. 3-11. The Fe concentration of the surface water samples ranges around 1 mg/l. This value range that is commonly found in (oxygenated) surface water in the Netherlands. For example, Fe concentration in filtered samples in the FOREGS database for the Netherlands ranges between 0.014 and 0.787 mg/l (Salminen et al., 2005). The Fe concentrations of the anaerobic drain water and reservoir water during winter time exceed this value range by an order of magnitude or more. Combined with the observation of the decrease of 'dissolved' Fe concentrations in the continuum groundwater to sub-catchment/catchment outlet this leads us premise that a large majority of the measured Fe concentration in the drain water and reservoir water consist of Fe(II).

   We agree with the reviewer that this in not addressed properly in the manuscript and we updated the field procedure description and discussed the likeliness of dissolved Fe(III) colloids attributing to a small extent to the measured Fe concentrations in the drain water and reservoir water in the revised manuscript (p 11.l5-15).

   Secion 4.1: 'Not all the dissolved Fe in the tube drain water, reservoir water and surface water samples can be assumed as Fe(II). The presence of dissolved Fe after 0.45 µm filtration in oxygenated water has been variously attributed to Fe(III) in colloidal phases (Lyvén et al., 2003) or complexed Fe(II) (Lofts et al., 2008). The particulate colloidal Fe(III) can exist as both organic complex and small hydroxide particle (Lyvén et al., 2003;Allard et al., 2004;Benedetti et al., 2003). Colloidal Fe(III) is stabilized against aggregation by binding of dissolved organic carbon (DOC) on its surface. Neal et al. (2008) found positive correlations between Fe and DOC concentrations of river water. However, dissolved Fe concentration in streams and rivers with near neutral pH values do seldom exceed 0.5 mg/l (Salminen, 2005;Neal et al., 2008). The Fe concentration in our drain water and reservoir water samples are commonly much higher and can, therefore, largely be attributed to Fe(II).'

2) *Assume all the dissolved iron to be Fe(II)*: This is in line with the previous issue and we revised this. The measured Fe concentration data in Fig. 6 may be (partial) dissolved Fe(III) colloid, especially the low concentration range below 1 mg/l. Therefore, our reported oxidation rate might indeed be overestimated slightly (p15 l.13-17)

Following nitrate patterns from groundwater to surface water doesn't have much added value for our situation. The nitrate concentration in the (Fe-rich) groundwater that exfiltrate to the reservoirs, is very low and the groundwater can be assessed as anaerobic. We now have described this in the results section. As a consequence the nitrate concentrations in the reservoirs are low as well. However, the nitrate concentration in the shallow groundwater can be fairly high (up to 100 mg/l in uppermost groundwater in some parts of the field). This groundwater is predominantly discharged to the ditch via the tube drains and does not end up in the reservoirs. At the sampling locations of the catchment outlet the nitrate concentrations ranges between 1 and 80 mg/l. This is mainly due to discharge of shallow groundwater via the tube drains (Rozemeijer, et al., 2010)

3) *pCO_2 pressures of the groundwater:* these kind of values are (Griffioen et al., 2013). Water with a pH = 6.4 and HCO_3 = 140 mg/l (which is very common as well) has a P_{CO2} of 0.059 atm. We described this in the manuscript (p.12 l5-6).

4) *Autocatalytic oxidation of Fe(II):* improved the manuscript regarding this topic (p.15 l.24 – p16.l3):

References

Rozemeijer, J.C., Van Der Velde, Y., Van Geer, F.C., Bierkens, M.F.P., Broers, H.P. Direct measurements of the tile drain and groundwater flow route contributions to surface water contamination: From field-scale concentration patterns in groundwater to catchment-scale surface water quality (2010) Environmental Pollution, 158

Salminen, R, Batista, MJ, Bidovec et al (2005) Geochemical Atlas of Europe. Part 1: background information, methodology and maps. Geological Survey of Finland, Espoo

Griffioen, J., Vermooten, S., and Janssen, G.: Geochemical and palaeohydrological controls on the composition of shallow groundwater in the Netherlands, Applied Geochemistry, 39, 129-149, http://dx.doi.org/10.1016/j.apgeochem.2013.10.005, 2013.

Rozemeijer et al, 2010. Integrated modeling of groundwater-surface water interactions in a tile-drained agricultural field: The importance of directly measured flow route contributions. WRR 46.

Van der Grift, .B and J. Griffioen (2008) Modelling assessment of regional groundwater contamination due to historic smelter emissions of heavy metals. Journal of Contaminant Hydrology, (96) 48-68.

Van der Velde, Y., J.C. Rozemeijer, G.H. de Rooij, F.C. van Geer, H.P. Broers (2010) Field-scale measurements for separation of catchment discharge into flow route contributions Vadose Zone Journal 9 (1), 25-35.