

Review HESS-2015-69

Phosphorus dynamics in lowlands streams as a response to climatic hydrological and agricultural land use gradients

General response to Editor and Reviewers:

We would like to thank both reviewers and editor for their thorough comments to our manuscript. Several concerns have been raised. In light of these comments, we agree to conduct major changes to the organization and logic of the article.

We are aware about our dataset is limited and restricted to only four watersheds varying in a series of characteristics. This is a weak point that is hard to overcome in any field study of this nature and conducted at the watershed level. However, each of them can be considered as a representative of its type regarding contrasting climate-hydrologic scenarios. We also consider that our dataset has other strengths given by the high frequency monitoring strategy. In this line, we decided to center the new version of the ms in the comparison of monitoring strategies for stream phosphorus considering the contrasting climate driven flow regimes as main factor. Land use intensity (please see below) will become considered as contrasting conditions for each climatic/hydrologic scenario. We aim at focusing on the overall behavior of these watersheds in the dynamics of nutrients, not necessarily on the specific mechanisms within each of them. Therefore we will now focus on monitoring strategies for nutrients, in the line of the special issue of HESS, describing the observed patterns and highlighting the potential of our results to generate hypothesis for future works, down toning the extent of our conclusions. These changes will cause modifications in the overall structure of the ms.

Henceforth: R1: reviewer 1; A: authors.

R1: Four small agricultural streams are compared for hydrological characteristics and phosphorus concentrations in four different forms. Additionally, the effect from two different water sampling strategies is compared and different methods are used to estimate P export. The assessment and conclusion are very general, based on only two years study and need to be more specific and deeper.

E.g. What is meant by agricultural intensification? The title suggests agricultural land use and in the discussion the number of livestock units is mentioned. However, there is very poor information about land used and nothing about manure and fertilization use, number

cattle and number of inhabitants in the M+M section. It is also very hard to directly compare catchments with different soil types and under two totally contrasting climates. The Danish climate might include snow accumulation and the hydrology is probably much affected by the existence of tile drains. Uruguay may have both summer and winter crops and high production of grass for feeding all year around, while in Denmark summer crops harvested in summer/autumn probably is dominating. Finally environmental regulations are hard to discuss only based on rough modelling of the relative contribution from point sources and diffuse sources. The discussion of sampling technique is clear but overall are too many conclusions based on too little data from just 4 catchments.

A: Please see General response to Editor and Reviewers. Given the recognized limitations, we will moderate the degree of certainty used in our previous conclusions. As the reviewer states and we agree, we will expand the rationale of the choice taken and operational definitions, adding more information about aspects as cropping systems, livestock type and density, kind and use of fertilizers and tillage operations.

The use of “agricultural intensification” concept in the ms will be changed by “land use intensity” (LUI). We are aware that LUI is a complex, multidimensional term, and measure the intensity of land use is not a trivial issue. We have selected really contrasting LUI conditions, and we are able to classify them in a robust way using LUI categorical descriptors such as fertilization, mowing, and livestock, following the concepts of a big number of previous works (e.g. Blüthgen et al., 2012; Brown and Vivas, 2005). In this context, we will use the term LUI as synonymous with farming intensity, integrating all the productive set and management activities in each climatic scenario (country).

At the same time, we intend to characterize real and typical productive systems representative from the each selected country/climatic condition, and will contribute to generate basic knowledge about consequences of LUI over water quality. Even though intensive land use is a worldwide phenomenon (Alexandratos and Bruinsma, 2012; Foley et al., 2005), the knowledge about the relationships between LUI and environmental consequences, has not been homogeneously generated. Most of the studies have been conducted in Europe and USA, and within these, in developed countries and temperate climatic conditions. Otherwise, a low number of studies were generated in developing countries, with warmer and wetter conditions as Uruguay. In this context, what are the environmental consequences of the farming intensity in subtropical South America constitute still an open question.

Specific comments

R1: **Title** suggest: Phosphorus concentrations in four lowland streams representing two contrasting climates and agricultural conditions.

A: We agree with the suggestion that the title has to be changed. Our proposal is this:

“Comparing monitoring strategies for stream phosphorus under contrasting climate driven flow regimes”

Abstract

R1: Line 16 why freshwater? The streams may also reach seas - Gudenå river reach Kattegat

A: We agree with the comment and changed the phrase referencing “downstream aquatic systems” instead freshwaters.

R1: Line 17 the word ‘intensification’ may imply that you have monitored the streams for a long time under which the agriculture has changed.

A: We agree with the comment and changed “intensification” by “intensity” in this line and throughout all of the manuscript. This change was accompanied with more general changes as described above.

R1: Line 17 dynamics and dynamics of what?

A: We referred to “temporal dynamics” or “temporal variability” and will change the text as a consequence.

R1: Line 27 Intensification (of agriculture) had a significant impact on subtropical climate. Should be reformulated. The agriculture is rather a result of the climatic conditions

A: In this case the authors refer to the impact of the specific high LUI productive systems representatives from a specific country/climatic condition. We will change the text following the general changes as narrated above.

R1: Line 30 increased contribution .. to P export? or higher concentration of ...P? Please reformulate

A: This point is consistent for all the studied variables. Applies to concentration, flow-weighted concentration and export. We will change the text.

R1: Line 32. Sanitary risks are not in focus in this manuscript

A: We will delete it.

R1: Line 33hand, temperate ... One or two streams in temperate climate?

A: The sentence refers to the comparison between both temperate streams, and specifically to the located in high LUI catchment. We will change the text following the general changes as narrated above.

R1: Line 34 farmed stream change to farmed catchment

A: We will change the manuscript text in the suggested way.

R1: Line 34 lack of environmental regulations Can this conclusion really be drawn from your study?

A: We agree that we might have extended our discussion and conclusions too far given the low number of watersheds included in this study. Our knowledge on the management practices in both countries surely influenced our conclusions. Considering that both reviewers highlighted this weakness (and that we agree upon re-reading our ms), we have decided to remove the section that discusses the “goodness” of different farming systems and their implications for management in the next version.

R1: Introduction **Line 43** Reference altering components of the hydrological cycle is about irrigation – What about the draining systems in Denmark?

A: Drainage systems are in place in most catchments in Denmark either as ditches or as tile drains. They may alter the hydrological response of catchments in different ways – as they decrease the maximum events by preventing surface runoff (tile drains) and lowering evaporation due to the lower groundwater table. On the other hand, tile drains is a quick conduit (shortcut) for water that via macropores from fields can reach surface water with pollutants very quickly (see Grant et al., 1996; Laubel et al., 1998). So, in case of hydrology there is no trivial answer to the question raised about importance of tile drainage on the hydrology of catchments. There need to be a catchment specific answer we are afraid as local conditions such as geology, topograh, soil types and groundwater aquifers are part of the answer as well as type of tile drain systems used.

Grant, R., Laubel, A., Kronvang, B., Andersen, H.E., Svendsen, L.M. and Fuglsang, A., 1996: Loss of dissolved and particulate phosphorus forms in drainage water from four arable catchments on structured soils in Denmark. *Water Research*, 30(11), 2633-2642..

Laubel, A., Jacobsen, O.H., Kronvang, B. and Grant, R., 1998. Subsurface drainage loss of

particles and phosphorus from field plot experiments and a tile-drained catchment. *Journal of Environmental Quality*, 28, 576-584.

R1: Line 44 several controlling systems Information about those in the current catchments should have been given in the M+M section and discussed

A: We will include more information about cropping system, livestock type and density, use of fertilizers, manure tillage operations, and other watershed characteristics as topography.

R1: Line 52 Explain how changed land use may alter flashiness

A: We will include information about this point in the revised manuscript.

R1: Line 72 Agricultural land use intensity This is not clear seems that a main difference between the Danish catchments was the proportion of arable land not necessary the intensity (?)

A: We wrote arable not as a potential, but as a realized condition. We agree with the reviewer that the chosen expression is not clear and we are going to modify accompanying the more general changes narrated above.

R1: Line 76 Suggest that you delete 'evaluate the direction and magnitude' and change to compare....

A: We will make this change.

R1: Line 83 Question iii ; a synergistic interaction..... Very vague clear aims/hypothesis are needed

A: A new set of hypothesis will be formulated and inserted:

Hypothesis 1. Low frequency sampling programs will have lower performance for estimation of stream P exportation in systems with higher hydrological variability (flashiness) regardless of land use intensity (LUI).

Hypothesis 2. Hydrologic behavior will promote stronger differences in P export than LUI.

Hypothesis 2.1. Hydrologically variable streams (those with higher flashiness), will export higher proportion of particulate P and with a higher temporal variability.

Hypothesis 2.2. The highest P exports from diffuse sources will be related to streams with higher flashiness.

R1: Line 86 Suggest cathment characterization

A: We agree and will change the text in this way.

R1: Lines 103-108 Suggest that as complete information as possible about point sources are given in a Table.

A: We will include the requested information.

M+M

R1: Suggests a special paragraph with statistical methods – It is not clear for me the time-span used for estimating significance between median values. Also explain why median and not mean values were used.

A: In section 2.4.2, is explained the rational of the statistical analysis election [“As the parametric test assumptions generally failed to be accomplished even with transformed data, the statistical comparisons were made using Kruskal-Wallis tests (Zar, 2010). When differences in the median values among treatment groups were greater than would be expected by chance, a post hoc pairwise multiple comparison procedure was applied to isolate the group or groups differing from the others (Dunn, 1964).”]. The time span used for estimating significance between median values was the entire studied lapse (2 years). We will work for clarify the point.

R1: 176 There are several info about flashiness in a recent paper: Deelstra et al., 2014 Agriculture, Ecosystem and Environment 195, 211-219.

A: Thank you by the suggestion. We will include information from this paper.

R1: Line 195 One method for total P export and two methods for source separation.

A: Sorry, we do not understand the comment. Three different methods were used for stream P load calculations. We used: 1) mean P values from composite samples and integrated fortnightly discharge, 2) P values from grab fortnightly samples, daily linear interpolation of concentrations between sampling dates and integrated discharge and 3) instantaneous discharge (each 10 minutes) and source apportionment model to assign P concentrations. We will improve the text to be clearer.

A: About the following comments:

R1:

2.2 Suggests Water sampling and water flow measurement

Line 131 10 minutes .. might be expressed: discharge was estimated and stored each 10 minutes

Line 46 pore size

Line 158 please give any acid and oxidizing agent used

2.4.1 Climate and hydrology

Line 169 variation in water flow or characterization of flow regime rather than 'responsiveness'

2.4.3 suggest: **Estimates of diffuse and point sources and P export**

2.4.4. Suggest: **Relative contribution of**

242 One method for load and one method for concentration (delete in two ways)

A: In each case we will change the text in the suggested way.

Results

R1: Line 257... Is it possible to compare with long-term precipitation data of the regions?

A: Climate characteristics of the studied lapse could be considered as typical or "average" years both in Denmark and in Uruguay. So, we would prefer to clearly state this fact in the paper without include more analysis about it.

R1: Line 270 The result that intensively farmed catchments had more flashy water regimes should be further discussed in connection to the different P loads.

A: we will discuss this point in the revised manuscript.

R1: Line 281 Relationship could only be expected if the local precipitation was similar and if water sampling took place simultaneously

A: Probably the referred phrase is not clear and induces to misunderstand its content. Each set of P forms (TP, PP, TDP, SRP, NSRP) are analytically estimated from the same original sample. Correlations were analyzed only for each stream (one matrix by stream). Please see table 3. We will change the syntax.

R1: Line 291 Median concentrations were calculated per year? With just two years you cannot estimate any significance. Give the rational for using median and not mean values.

A: Median concentrations were calculated for all the period (biweekly samples by 2 years, 48 samples approximately). Please see Fig. 3. The rational was stated in 2.4.2 (M&M). "As the parametric test assumptions generally failed to be accomplished even with transformed data, the statistical comparisons were made using Kruskal-Wallis tests (Zar, 2010). When differences

in the median values among treatment groups were greater than would be expected by chance, a post hoc pairwise multiple comparison procedure was applied to isolate the group or groups differing from the others (Dunn, 1964).”

R1: Line 308, 314 Too many numerals (H=133.298 and 141.157)

A: We will reduce the number of figures.

R1: Line 331 Could not been explained in what respect? Incidental P losses following manure application on agriculture land is quite common

A: As others, the cause suggested by the reviewer may explain the observed outlier. But we have no grounds to assign with certainty that this is the case for the ‘outlier’ observation. This was the meaning of the included phrase.

R1: Line 334 Isn’t the dilution effect primarily a consequence of base-flow contribution and the size of the catchment rather than ‘intensive stream’? Do you rather mean intensive agriculture in the cathment?

A: The magnitude of the difference in the dilution effect is really larger than the difference of the catchment sizes or base-flow contribution, so the coauthors are very confident in the attribution to land use intensity as the main cause. We agree that we have to change the way to refer to the stream under the influence of intensive land use.

R1: Line 342 than the other 3 streams (?)

A: Yes. Please, see fig. 4.

R1: Line 363 What do you mean with field evidence? This is more a discussion point.

A: In this case, we used “field evidence” as synonymous of “results”. We will change the expression to not induce to confusion for the readers. We do not agree with the suggestion to move this point to discussion, because the phrase is included only as a description of the results about contribution of particulate to total phosphorus exported.

R1: Line 367 I do not understand the sentence: .. is repeats itself in the grab concentrations

A: The phrase refers to the pattern of major contribution of PP to TP is detected if we analyze grab samples data, composite samples data and flow-weighted concentrations. We will change the syntax to clarify this point.

R1: Line 368 higher SRP contribution should be discussed in connection to soil type and the soil P status

A: We agree with the comment and we will discuss this point in the revised manus.

R1: Line 374 of exported total P. A high proportion should have been in SRP form!

A: Yes. The correlation level between TP and SRP was high. But the correlation level with TDP and NSRP was even higher (see Table 3).

Discussion

R1: Line 380 Repeat the figures (0.3 versus 1.0).

A: The figures will be repeated.

R1: Line 386 Bee more precise- extreme in what sense? Warmer in which months? How does it change the crop production, snowmelt periods etc.....

A: We will include complementary information from the cited paper and other relevant papers.

R1: Line 398. Perhaps in Denmark but in many other European countries very poor improvements have taken place to reduce P losses from small point sources such as single household wastewater treatments. How many people, milk-rooms etc. with waste water treatments were included in the studied rural landscapes?

A: We will include more detailed info about it.

R1: Line 416. Do you mean that the diffuse load is underestimated but the point sources are overestimated for flashy streams with a low base-flow contribution?

A: Both methods based in grab samples (linear interpolation and concentration-discharge relationships) tends to underestimate P exports for TEMP streams, and overestimate it for SUBTR streams, when compare with the method selected as reference (composite samples). This is interpreted in relationship to the inadequacy of the method to depict stream dynamics, and is not discussed here in relation with the characteristic of the source.

R1: Line 425 Vague. Explain 'signs of interaction'. 'Such as lower deviation of LFS-LI estimates seems to be a weak proof. Do you suggest that Intense manure application needs to be considered when planning sampling strategy?

A: Intense manure application would be a practice with strong impact in the temporal

dynamics of stream P In Danish systems, but probably not in Uruguayan streams. We will discuss this aspect.

R1: Line 446. Discuss expected contribution of P from anaerobic groundwater. Is this a general overlooked problem?

A: We only suggested that this can be a possible explanation for the observed data.

R1: Line 448 Needs explanation. Were there less drainage, more anaerobic condition and less Fe-bound P in the extensive Danish catchments? The reference Leinweber et al 2002 is missing in the list.

A: I apologize especially by the missing references. This will be fixed in the next version. The sentence is stated in conditional way (“possibly be explained”). We cannot make an affirmation about your question.

R1: Line 462 Give a reference demonstrating the impact of cattle with direct access to the stream channel.

A: We are going to include a reference.

R1: Line 468 Did you notice such deterioration etc. in the catchment of Uruguay?

A: Yes, this kind of symptoms of cattle activity is evident. We will make more emphasis on these aspects in the new revised manuscript.

R1: Line 475 Rather the opposite – grab sampling may underestimate the diffuse P pollution during high discharge events.

A: The idea suggested by you is exactly the meaning of the included phrase. “..., the contribution of diffuse sources to exported P seems to be underestimated by the C-Q model. This is probably a consequence of the underrepresentation of the grab sampling programme during high discharge events...”

R1: Line 510 Explain further. Do you mean that an enrichment of P is taking place in the very topsoil and is reaching the stream via surface runoff? The reference Derpsch et al., 2010 is missing in the list.

A: Yes, this is the meaning, and this will be explicitly stated in the next version. I apologize again for the missing reference. This will be fixed in the revised manus.

A: As was stated above, we have decided to remove the section that discusses the “implications for management” in the next version. Therefore, we do not include here comments about lines 521, 524 and 555.

Conclusions

R1: How much did flow response regulated the TP load (80-90%?) and how much did agricultural management (10-20%?)

A: It is not possible to quantify it based on this study.

R1: Give examples of more precise environmental regulations

A: As was stated above, we have decided to remove the section that discusses the “implications for management” in the next version.

R1: How precise do need the models to be? What about other strategies P balances, P surplus P-index etc. mentioned at the end of the discussion?

A: As was stated above, we have decided to remove the section that discusses the “implications for management” in the next version.

R1: Figure 2. ‘Flashiness is clear’ support with FI values. The temperate stream seems quite flashiness too.

A: We will include the values (they are currently included in Table 2).

R1: Figure 3 What is grab samples? Better to combine flashiness with P concentrations directly, now the figure is not necessary.

R1: Figure 5-6. Not necessary.

A: Sorry, this would be “grab samples”. The necessity of figures 3, 5-6 will be change in function of the modifications (title, hypothesis and emphasis) that we will make in the ms.

References

Alexandratos, N., and Bruinsma, J.: World agriculture towards 2030/2050: the 2012 revision, FAO, Rome, 2012.

Blüthgen, N., Dormann, C. F., Prati, D., Klaus, V. H., Kleinebecker, T., Hölzel, N., Alt, F., Boch, S., Gockel, S., Hemp, A., Müller, J., Nieschulze, J., Renner, S. C., Schöning, I., Schumacher, U., Socher, S. A., Wells, K., Birkhofer, K., Buscot, F., Oelmann, Y., Rothenwöhrer, C., Scherber, C., Tscharntke, T., Weiner, C. N., Fischer, M., Kalko, E. K. V., Linsenmair, K. E., Schulze, E.-D., and Weisser, W. W.: A quantitative index of land-use intensity in grasslands: Integrating mowing, grazing and fertilization. *Basic and Applied Ecology*, 13, 207-220, <http://dx.doi.org/10.1016/j.baae.2012.04.001>, 2012.

Brown, M. T., and Vivas, M. B.: Landscape development intensity index. *Environmental Monitoring and Assessment*, 101, 289-309, 2005.

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., Chapin, F. S., Coe, M. T., Daily, G. C., Gibbs, H. K., Helkowski, J. H., Holloway, T., Howard, E. A., Kucharik, C. J., Monfreda, C., Patz, J. A., Prentice, I. C., Ramankutty, N., and Snyder, P. K.: Global consequences of land use. *Science*, 309, 570-574, 10.1126/science.1111772, 2005.