

Overall this study is very nice, novel, and useful and should be published after minor revisions (see comments).

General comments

This paper focusses on improving water quality by using aquatic plants systems in combination with different sediment types. In the study, the authors use several aquatic plant with contrasting growth forms and three different sediment types. They assessed how these systems perform in the removal of nutrients (focus on N&P) and where these nutrients end up in the system, with a focus on the plants and sediment. They furthermore studied how many of the nutrients coming into the system (loading) can be removed by harvesting the plant biomass. The fact that they combine different nutrient loading, different species with contrasting growth forms, and different sediment type makes this study very interesting, useful and novel and will be of interest to scientists (ecological and biotechnological), ecosystem managers and wastewater treatment specialists.

The authors conclude that the selected species can be used to remove free N and P from the water up to a certain nutrient loading, above which the plants only sequester a very small portion of the nutrients introduced to the water. The amount of nutrients removed from the water not only depends on species, but also interacts with sediment characteristics. This stresses the importance to consider the whole system, but in particular the plant and sediment characteristics in successfully designing a sustainable CWS.

Overall the quality of the research and presentation is very good and should be published after mainly textual revisions. Some parts could be presented in a bit more detail or more concisely for clarity, as I was confused about some sentences and terms used. (See specific comments.)

For example, the title does fit the content very well, providing the reader know the difference between water polishing and water treatment. To me this was not clear, so perhaps the authors could changes this for something like: plants can be used to remove nutrients from surface water up to certain nutrient loading levels.

The paper also cites many relevant references, but could be improved by adding a few more recent ones and adding a few more references of other research where plants have been used to remove nutrients. This will then allow for more discussion on the general applicability of this study and the influence of plant growth form in nutrient removal. (Also see specific comments.)

I've added quite some specific comments and suggestions to aid the authors in revising the manuscript. Most points concern sentences / terms that were not completely clear to me, while a few are about questions I had about the methods, results and discussion. I hope this will help in a swift revision and publication.

Specific comments

Title: I'm not sure that the difference between polishing and purification are clear to all readers. Perhaps change for low and high loading (See general comments).

Line 25-26; here you make it very clear what you mean with polishing VS purifying, perhaps also include this in the introduction or just use the loading terms instead of polishing / purifying. Additionally, in the title you mention polishing vs treatment. I think it would be good to choose the same terms throughout the text or leave them out.

Line 27: you only mention the importance of soil, but I think that the effect of plant species / plant growth form is also an important conclusion which could be added here.

Line 40: I'm confused about the definition of the terms 'free surface flow systems' and 'subsurface flow systems' you use. Is the only difference between the two systems that the soil can also take up nutrients in the subsurface systems? The name suggest that in the subsurface systems, the water doesn't flow over the sediment, but rather through it (helophyte CWS). If this is true, than the incorporation of sediment is not the only difference. Because you do not mention these terms in the rest of the text you could also remove them to avoid confusion.

Line 46-47: You mention that low maintenance leads to a saturated system. I think the reason why a low maintenance system is the same as an easily saturated system may not be evident to all readers, please elaborate. I think you mean that if you do not remove P from the sediment, the binding capacity will decrease and no additional P will be taken up, thus no water quality increase.

Line 47-48: I think seasonality is also an important limiting factor for these systems. You mention this in the discussion, but perhaps you can also include it here. If you want a system with plants to remove nutrients year round in the temperate regions of the world, you will need to add energy in the form of light and heating in the cold seasons, thereby increasing energy consumption and perhaps making it not such a 'low energy requiring system', as you mention before.

Line 50: You mention few studies have been performed, could you shortly give their main results?

Line 57: with adsorption, do you only mean adsorption to sediment, or more in general (also to waterborne particles?)

Line 60: How are the mechanisms affected? Is their speed affected, or efficiency? Please clarify.

Line 70: could you explain why you report the nutrient loading levels in unit per square metre? As the system is 3-dimensional, would it not be more logical to provide loading in units per litre or per cubic metre?

Line 74: If you want to use these terms (polishing / purifying), please specify the difference here as you did in the abstract.

Line 74-76: I miss the sediment in your aim / main research question.

Line 81: did you mix the sediment before adding to the mesocosm?

Line 82: could you summarize the basic characteristics for tap water, as you did for sediment?

Line 90-92: Did you add the nutrients to the surface water? I was also wondering why you included the natural deposition, generally it's very good to include it! But it seems that they are negligible compared to your real treatment. Also, you used relatively old references to determine the amount of background nutrient deposition, are they still relevant? I think you can also just report your loading treatment as is.

Line 96: What do you mean by environmentally relevant densities? Typical densities for this species in lakes, ditches? Please add a bit more detail or a reference. Could you also explain why you used different amounts per species? Would your results have been different if you've started with the same amount for all species? I feel that this section needs a bit more argumentation.

Line 97: Why do you mention *Chara hispida* here? It's not mentioned in line 68 of your introduction. If you want to include it, perhaps just mention that you also tested chara, but that it was outcompeted so is not a suitable species for water purification under your experimental conditions and will therefore be disregarded in the rest of this study.

Line 103: Were rooted species harvested including roots, or just the shoots?

Line 110: Because pH can vary over time, please provide information on the time of measurement.

Line 116: Please provide information why it was important to measure Al, Fe and Ca (They can bind P, but I'm not sure all readers will know). You can for example add this information near line 56/57.

Line 133: Do you mean total P or inorganic P (PO₄) here, please specify.

Line 135-136: could you provide a reference for these processes to explain why this assumption is valid.

I also wondered if you couldn't quantify (or at least estimate) the amount stored in the sediment, based on your sediment measurements. In that way, you don't have to make this assumption, but can provide proof.

Line 145: please clarify what you mean with: 'except for treatments also including time as main effect', which are they?

Line 148: Please mention which R packages/statistical tests you used for the regression models.

Line 148: you introduce the term 'sequestration' here. I'm not sure everyone will be familiar with this, could you explain or swap for a term you've mentioned before, such as nutrient uptake or absorption?

Line 152-153: I think you mean that there is a main effect; however, looking at the graphs not all treatments show an increase. Perhaps you could provide some more information on this. You performed a full-factorial experiment and especially the interactions are very interesting I think!

Line 158: Because you don't give the data, perhaps it's informative if you just provide the mean +/- SE values.

Paragraph 3.3 seems a bit long and sometimes unstructured; perhaps not all information is needed. e.g. is line 186-187 really relevant for your story? Also lines 183-184 may fit better at the end of the paragraph.

Results: I expected information on Fe, Al and Ca in the results, as you mention them in the M&M.

Results: I also expected fig. 6 to be mentioned here, for example in section 3.3.

Line 203: perhaps add that N removal also depends on species.

Line 206: are these averages calculated over all nutrient loading treatments?

Line 235-237: Could you comment a bit more on how the different growth forms impact nutrient removal and which species you would recommend under what loadings? Do you expect similar results for other floating / submerged plants? Perhaps you can use the 'few references' you mention are available on this topic.

Line 237-238: this sentence is hard to read, perhaps start with: "Low O₂ mobilizes PO₄ [...]"

Line 240: do they only take up nutrients by their roots or also via their shoots when nutrients are mobilized and leach into the water?

Line 241-242: please add reference to support this.

Line 245: How does this seasonality affect your maximum loading you can remove with plants? Should we divide the results of your study by 2 to get the year round

maximum nutrient loading to account for the winter influx of nutrients, assuming that they are then bound to the sediment? Please elaborate on this.

Line 247-248: how do you come to 6-24mg? In fig. 6 I don't see a maximum.

Line 248-249: You mention that your results are comparable to the results of Reddy & de Busk, but according to you, they found values 2times higher than yours, which sounds much higher to me. Could this be because you were not near the maximum potential of Azolla, as suggested by the linear relationship between loading and uptake (fig6)?

Line 264: Very nice that you mention N-fixation in Azolla! Could you add references of a study that researched the N-fixation capacity of Azolla to assess how much they can fix N and thus argue if this amount could have affected your results. N was still removed from the water, so I see no reason why this would be very detrimental to your story.

Line 263-265: This argument needs a bit more information, could you provide some mechanisms why you conclude that senescence is more important than soil leaching? I'm not sure the information provided is sufficient.

Line: 271- 274: Here you extrapolate your results to plant growth forms. Only your data does not allow for this, please provide literature on other species with the same growth form or on the mechanisms which will show why we can assume that other species with the same growth form with have similar nutrient removal rates.

Line 279: explain how the creation of anoxic conditions removes P from the sediment.

DISCUSSION: Do you have any idea why *C. demersum* performed so poorly?

Discussion: Could you compare your results to other studies about CWS with floating or submerged plants or with plant nutrient uptake studies? This will enable you to give more general recommendations and conclusions.

Perhaps also shortly compare your results with more traditional emergent CWS, as you've mentioned there's a lot of literature about those and readers will be more familiar with those.

Line 283-286: Perhaps remove the species specific information here, as the effects on N are different (Azolla is always better). Furthermore, looking at fig. 6 *M. spicatum* is only better in P uptake between your low and medium nutrient treatment, because you've fitted a log.-line, at 22mgP input azolla is already better (although not sign.) It's important to keep in mind how reliable your regression is when only using 3 points on the x-axis, especially when making these kinds of general statements and using the regression line to determine thresholds. I would tone this down a little bit, or at least acknowledge the uncertainty.

Technical comments

Line 54: change 'drained' to 'removed'

Line 55: add recent reference, several recent studies that have looked into this as well.

Line 63: you mention 'soil type'. Do you mean the same with this as with 'soil characteristics' in line 60? If so, it's better to use the same term throughout the whole text.

Line 68: because you specifically use macrophyte species with contrasting growth forms, I would stress this here.

Line 70: Please provide reference of 'environmentally relevant nutrient levels'.

Line 79: please check your calculation for volume. I'm not sure that 20cm in a 185cm diameter cylinder is 135 L.

Line 143-146: Long, complex sentence. Could you abbreviate this or split in 2 sentences?

Line 155 and further: you often write $p < 0.000$, which means that p is negative. I believe you mean $p < 0.001$.

Line 158-161: To remove some ambiguity, perhaps rewrite as: [...] even though total P and Olsen P concentrations were much higher in clay than in the other two soils [...].

Line 177: I think you mean *Azolla* OR *Spicatum*. Now it seems like they were both inside one quarter.

Line 179: 'on the other hand' seems to contradict the previous sentence where *M. spicatum* also didn't take up more than 20%

Line 188: Here you use 'absorbed', I was wondering if you mean the same with words like: take up nutrient, sequester nutrients and absorb nutrients. If so, it is more clear to choose one term and use that one throughout the whole text for clarity.

Line 195: 'As P loading [...]' can be removed as the first 2 lines of the paragraph provides the same information.

Line 251: It's hard to compare values in mols and grams, please choose one unit and use it throughout the text.

Line 257: add '[.] in our study.' to the end of the sentence to make sure that readers know your talking about your results.

Line 271: change beginning to: '[...] CWS, both submerged and floating plants are [...]'

Line 274-276: Too many dependent clauses, please reformulate. (Also check the rest of the text for these sentences, they are often hard to understand.)

Line 278: change with: [...] resulting in saturated soil and thus leading to an increase in water nutrient levels under continued nutrient input.

General: I'm not sure the term 'soil' is used often when referring to aquatic systems, sediment may be more appropriate in the context of your study.