Author Comment for BG-2015-331 "Nitrogen export from a boreal stream network following forest harvesting: seasonal nitrate removal and conservative export of organic forms" by Schelker, J., Sponseller, R., Ring, E., Högbom, L., Löfgren, S., Laudon, H.

This Author Comment will address the main issues raised by the three referees and provide additional information for the referees and the editor. Please note, that not all minor suggested changes are addressed in this document, as these will be explicitly discussed when the revised manuscript will be provided. This should happen shortly.

Yours Sincerely

J. Schelker

Reviewer #1

This study quantified nitrogen removal by the river system draining a third order watershed in Sweden. Nitrogen loading is elevated because of clearcutting in this primarily forested watershed. Because significant deforestation has occurred recently, loading into the headwater streams has increased. The amount of nitrogen entering the entire river network can be estimated based on the proportion of the watershed that has been clear cut (all in a similar time frame), and fluxes that are characteristic of forested and clear cut catchments. This modeled estimate of loading can then be compared to fluxes measured at the mouth of the watershed, and the difference is due to nitrogen retention by the watershed. The study found that DON is not retained, where a significant proportion of nitrate is net retained (from 30 to 100%). Highest retention appears to occur following spring snow melt, lowest during the winter, and intermediate during the summer growing season. Retention was not related to flow conditions. Results indicate that increased export from small catchments due to clear cutting can be retained by the river network, buffering the impact in larger rivers and downstream water bodies.This is an interesting study and well written manuscript. Overall, I believe that the analysis is sound. A few issues need to be addressed however to strengthen the paper.

Dear Dr. Wollheim, thank you very much for this overall positive evaluation. We will do our best respond to your comments and to address your concerns.

I was surprised that removal in this relatively small network is so high. I think it is important to report the surface area estimate of the river network. In addition, there are lakes in the watershed, which likely increase significantly the surface area of surface waters. The lake in the mainstem in particular could contribute to the high removal. What is the surface area of the lakes, and their residence time?

Yes, we have would like to add the information (see Table 1) of the estimated river surface area in an additional table that will be added to the site description.

Site Name	Short Name	Catchment Area	Proportion Clear-Cut*, 2004; 2011		Wetland Area	Total Stream Length	Lake Area*	Stream Surface Area	Total Aquatic Area
Unit		[ha]	[%]		[ha]	[m]	[m^2]	[m^2]	[m^2]
Balån River 1 Outlet	BA-1	2291	2%	11%	337	37521	87829	185738	273567
Balån River 2	BA-2	868	5%	18%	88	15754	6590	19249	25839
Southern Reference	RS-3	156	0%	3%	4	2195	0	2195	2195
Southern Clear Cut	CC-4	41	0%	56%	3	1650	0	660	660
Northern Catchment	NO-5	40	0%	33%	5	1386	0	554	554
Northern Reference	NR-7	24	0%	16%	4	835	0	334	334
* estimated from satellite data									

Table 1, catchment characteristics of the Balsjö catchments.

Unfortunately we have no detailed knowledge of the depth of the lake located between BA-1 and BA-2. Thus, we are not able to estimate the mean residence time. However, from other work we know that similar ponds in this landscape are commonly very shallow, i.e. don't exceed a depth of 2-3 m, but there is not really more here that we could add to the manuscript.

The high removal estimates hinge on the loading estimates from the two clear cut catchments. One of these (CC-4) had much higher loading estimates than the other one (NO-5) and this was attributed to riparian buffer in the latter removing the inputs. The mixing model uses the average of these two catchments and I believe assumes the average applies to all cleared land in the entire watershed. The issue here is that this amount is based only two catchments with very different loading estimates. If the catchment with smaller increases is more representative, then the estimate of removal by the river system would be an overestimate. Is there any additional data available to assess which of the catchments is more representative (or whether an average is)? If riparian removal is inferred as the reason why the second catchment does not have as high response to clearcutting, are there any data on what proportion of clear cuts maintain the riparian zone? Another way to address this uncertainty, is to look at the range in watershed removal by looking at two scenarios, one where all clear cuts have the low response, and the second where all clear cuts have the high loading response.

It is absolutely correct that the model results are strongly dependent on the clear-cut (CC) loading estimates. More specifically, the main question is how the clear-cuts in the landscape can be represented by the measured data from the two somewhat contrasting CCs. Here we would like to describe a few properties of the CCs in this study.

First, the difference between the two harvested catchments is not only that a forest buffer strip was kept, but also that CC-4 has a much smaller riparian zone than NO-5. The riparian zone within NO-5 has likely been a small peatland that was drained by 'shovel and spade' sometimes in the past with several meters of peat soil surrounding the stream (Schelker et al., 2013a). This peaty riparian zone is present almost along the full length of the stream; large parts of the stream bed also consist of peat. In contrast, the CC-4 stream drains mostly a more 'upland' like catchment with well-developed podzol soils. It does have some riparian zone, though with less peat. This streambed is more sandy, underlain by low permeable glacial till, that limits high vertical hyporheic exchange.

Second, the properties of the harvests outside of our two experimental harvests, but inside the stream networks drainage area have riparian areas of mixed character. Some of the harvests appear to be more of the upland type, some may have more extensive peat-rich riparian zones. Figure 1 below shows a map indicating the locations of wetlands combined with the satellite clear-cut data. To us, this map indicates the mixed character of both types of harvests, represented by our experimental harvest in the landscape.

Third, even though the foresters claimed that in most of the other clear-cuts, a riparian buffer zone was left intact, we have evidence from harvested areas within the network, but outside the experimental treatments, that harvests reached almost all the way to the stream and that severe soil damage was caused in locations very close to the stream (see Figure 2). Such damage of riparian soil from forestry machines originating from, for example crossing the stream, was also present within the CC-4 catchment (were stream crossings were done on purpose), but not within the NO-5 catchment and may be an additional factor for the difference in the response of the two treatments. Furthermore, narrow buffer strips were often found to be subject to wind through. We would also add this information to the methods section of the manuscript.

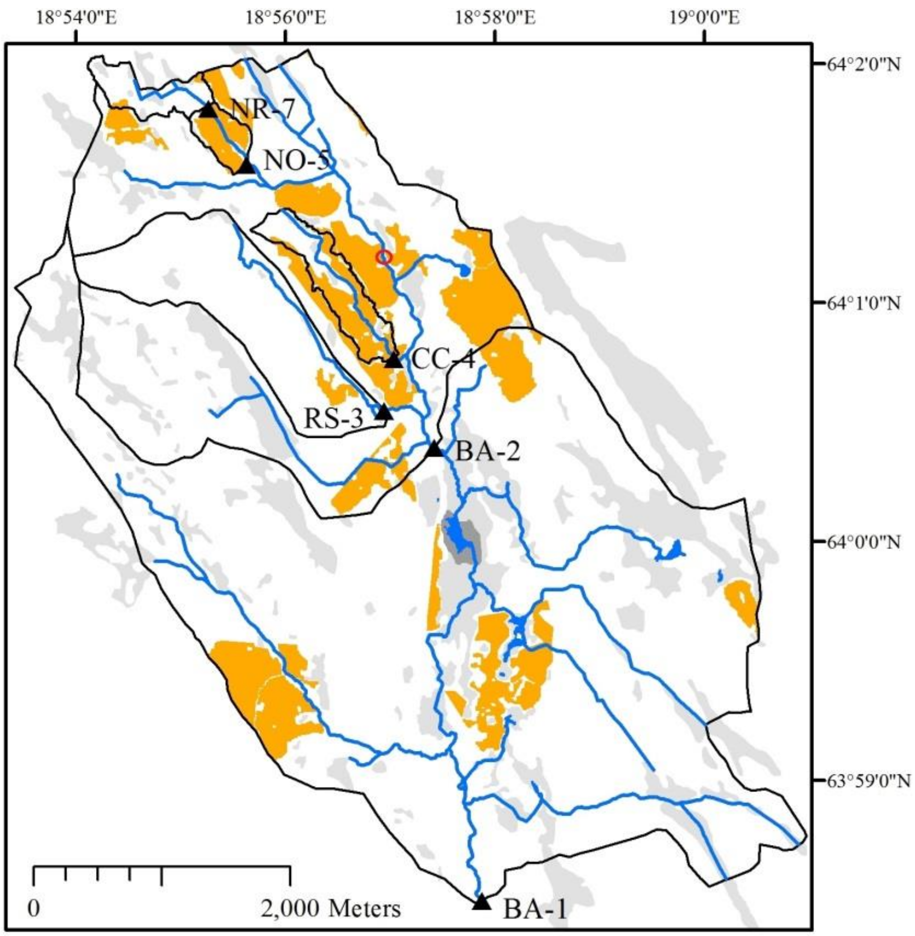

Figure 1, overlay of wetlands (grey shading), and forest harest (yellow) within the Balsjö stream network. The red circle denotes the location where Figure 2 was taken.

Figure 2, riparian soil damage in several locations within the Balsjö stream network, but outside the experimental harvests of NO-5 and CC-4 (approximate location of pictures given in Fig. 1).

Finally, the idea of using scenarios of each of the treatments to represent the CC end member in the mixing model was one we had developed previously, but decided against it. In short, the dataset we compare with contains at least one catchment that we know would not behave like this definition of the end member, i.e. if we would assume all CC areas to respond like NO-5, there would be at least the CC-4 catchment that we know does not follow this behavior. Considering that both harvest also account for an important fraction of the total drainage area of BA-1 and BA-2 that was harvested (CC4 for 9% and 15%; NO-5 for 5% and 9% of BA-1 and BA-2, respectively), applying the scenarios as suggested and comparing these to the measured values at the downstream sites would simply be wrong, as they cannot represent the physical system.

Another reason for not adopting the approach of different scenarios was given by the difficulty of presenting the resulting amount of model output data in a clear and concise manner. The effective amount of data to present would be threefold, which would make most of our plots difficult understand.

Overall we conclude that we are simply limited on more detailed data that could guide us on better choices of the input loading assumptions. Our current approach to estimate the loadings of clear-cuts as simple averages of the available data provides us with what we believe is the most robust estimate. We would therefore argue for keeping these assumptions. Furthermore, we argue that the use of model scenarios will be at the cost of clarity. We thus suggest to not introduce them into this study.

Greater confidence in the mixing model would be gained if there is also a conservative solute that responds to clearing, and then mixes throughout the network, where there is no removal. I believe that Shelker et al. 2014 may have this data. Use of a conservative tracer would also address the representativeness of the watersheds. I suppose the DON serves as a conservative tracer based on the result, but a priori this was not expected, whereas a solute like chloride would be conservative.

Yes, this is a good point. Figure 3, below provides this information for dissolved silica and chloride, both used as conservative tracers. We would also like to include this figure in the revised manuscript.

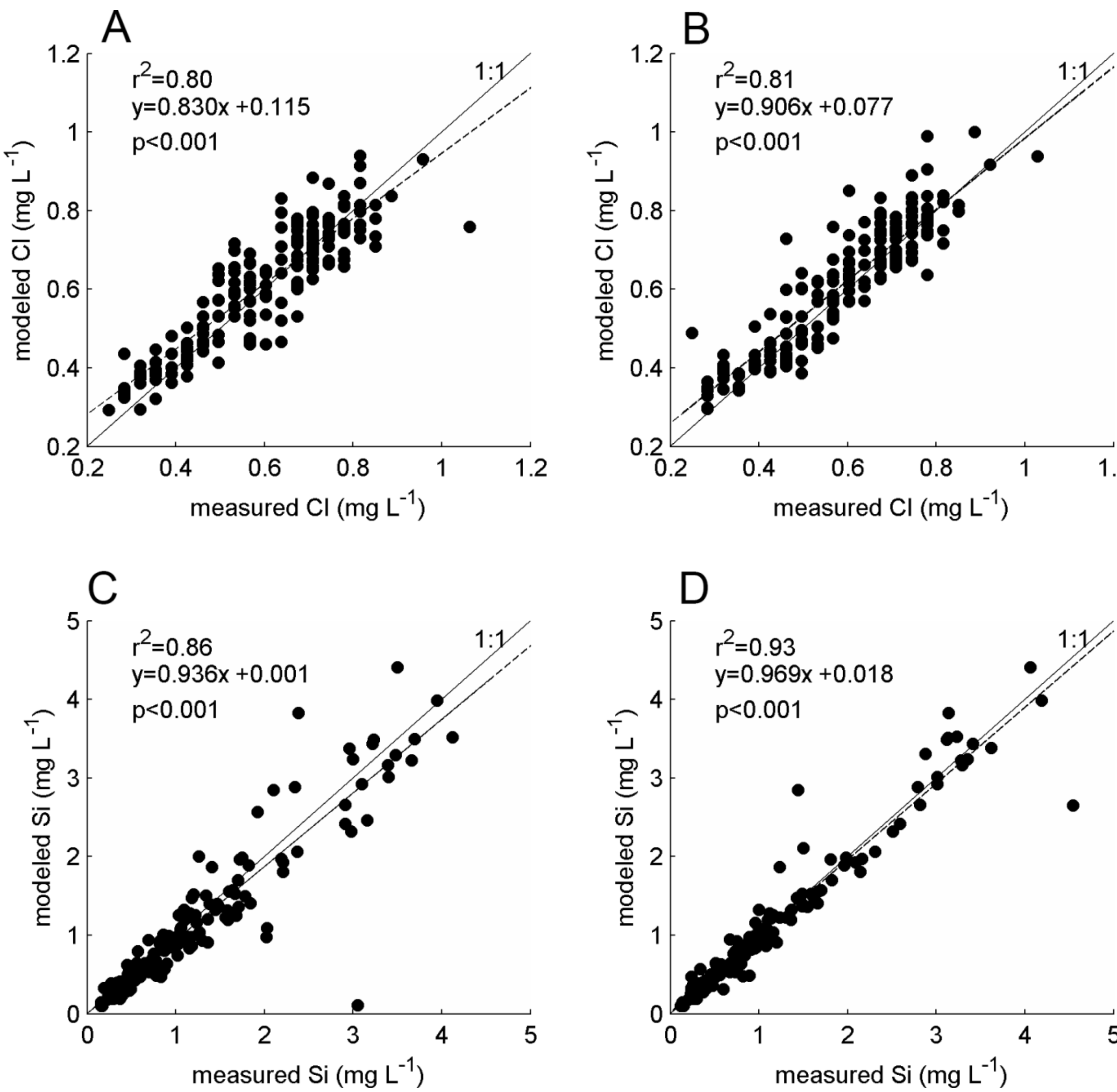

Figure 3, comparison of modelled and measured Cl and Si concentrations for BA-1 (panel A and C), and BA-2 (B and D).

These plots may be also seen in comparison to similar plots of the suggested model scenarios. The two figures (Figure 4 and 5) below present these. Panels A-C present the mixing model for the site BA-1, Panels D-F for the site BA-2. Thereby are the panels A and D the scenario with the highest load (CC-4, scaled to 100%), B and E, the 'average' scaled concentration load (as also shown above and used in the manuscript) and panels C and F the low loading scenario (response as NO-5, scaled to 100%).

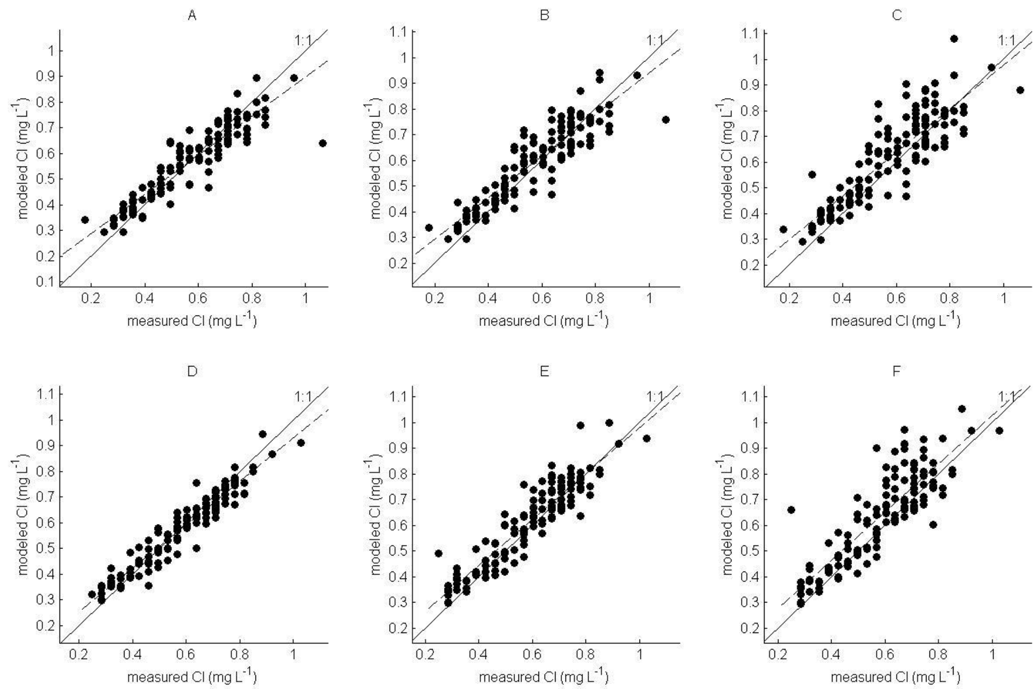

Figure 4, comparison of modelled and measured Cl concentrations for BA-1 (panel A to C), and BA-2 (D and F) for the three different loading scenarios.

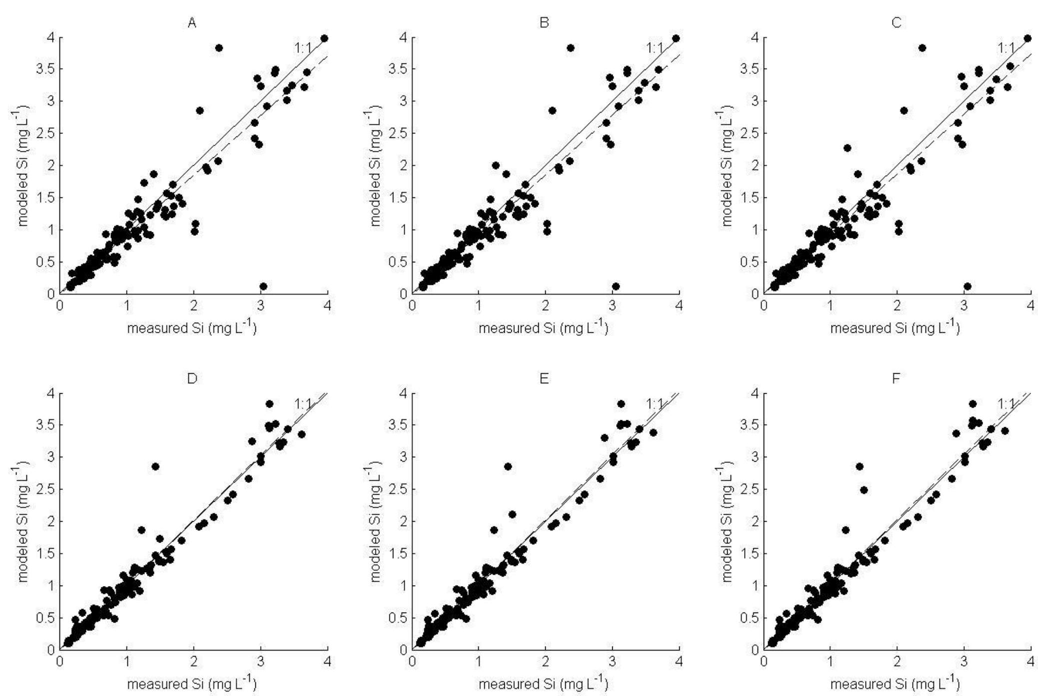

Figure 5, comparison of modelled and measured Si concentrations for BA-1 (panel A to C), and BA-2 (D and F) for the three different loading scenarios.

Overall we believe that the mixing models of Cl and Si using the different scenarios show the best overall visual fit (Figure 4, B and E, as well as Figure 5, B and E) and the smallest systematic deviation for the 'average scenario'. These observations lead us to conclude that our choice of the loading scenarios as the average concentrations, scaled to 100% is reasonable.

Although DON retention (or lack thereof) using the model is reported, no results on DON response to clear cutting are shown or presented in the results. I think this is important to include (perhaps adding a panel to Figure 2).

Yes, we fully agree to this point. We have added DON concentrations to the former Figure 2 (Figure 6, below) as a separate panel. Also, we are now presenting concentrations of NH4 and NO3 instead of only NO3, as requested by the other reviewers. These results will also be briefly presented in the revised results section.

Some more discussion of the mechanisms that contribute to the removal efficiency patterns (both over time and vs. flow) would also strengthen the paper. It is not clear what the mechanisms are so that the snow melt period would have the highest retention. Flows are high and temperatures are cold which should lead to low retention. Transfer to the hyporheic zone, riparian habitats, or groundwater is suggested very briefly. But could these explain such high losses, and why during spring only? The U term would incorporate net losses to these areas. If U is higher because there is more DOM or it is more labile why is there no DON retention then (or at least conversion of DON to DIN), especially when DIN supply is limited. What about light coming through the riparian canopy? Is it high and canopy cover low, so more primary producer uptake of nitrate during spring? If clear cut removes riparian this could be a mechanism contributing to temporary removal at least - but how common is riparian clearing. And why a more important factor in the spring? For Q to not be a factor means that as Q increases so does the uptake rate (or uptake velocity) in order for retention to remain high. A plot of uptake or uptake velocity over time or vs. flow would help to evaluate this.

When discussing the seasonal pattern of U, it should be remembered that others have shown very similar patterns before, such as high uptake in the spring (Roberts & Mulholland, 2007). Furthermore, we are somewhat limited in our data, so that not all mechanisms can be explicitly ruled out or confirmed. We will try to be more specific and less speculative in our revised version of the manuscript.

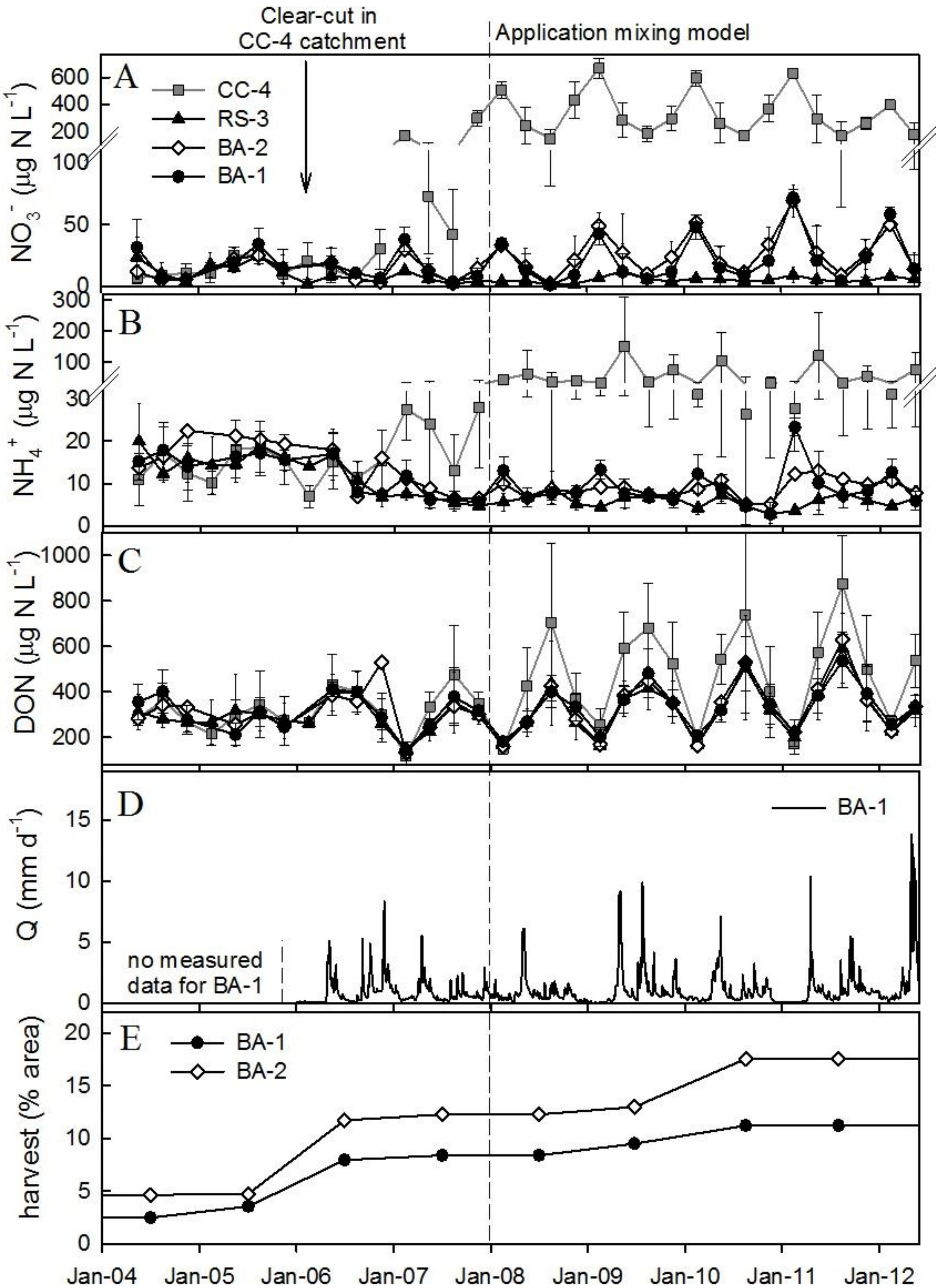

Figure 6, revised version of former Figure 2 of the manuscript.

As this is just an Author comment and not a Response to Reviewer Letter, we would like to state that we will revise the points listed as specific comments below in our revision of the manuscript.

Specific Comments

Equation 1. Units are confusing because of the use of mm/d (have mm, L, m2). Please use consistent units throughout (I suggest m), and make sure easy to see that all units cancel out correctly. Concentration units with equations given in mg, but data given in micrograms. Please be consistent.

What is size of the small catchments? What length of stream is above the sample site in these watersheds? Could some removal already have occurred at sample location?
The scaling to 100% harvested assumes (equation 2) assumes linear relationship between % harvest and concentration. Should state this explicitly.

12071.5-6. Unclear what the values in parentheses mean. Negative values are confusing. I understand the negative value is used because it is removing N from the water column, but areal uptake should be reported as a positive value.

12075.6. Denitrification is a dissimilatory process.

12075.11. Dissimilatory reduction to ammonium (DNRA) is also a dissimilatory process, but seems unlikely in this site. Mostly occurs where there is low OM, and very high N (much higher than here). This discussion seems too speculative. If keep, then add refs on this process from the literature.

12075.18. Should include more evidence of high DOC in this catchment if want to make this point. Seems too speculative.

Figure 1. Hard to see basin boundaries. Make darker lines.

Figure 2. Really hard to tell the lines apart. Especially important to see BA1 and BA2. Can't tell the two lines apart in bottom panel (symbols too small).

Figure 4. Points are very small so hard to tell them apart. So it is hard to make sense of what is happening. Not clear what points are (observed). Make points bigger. Add the seasonal demarcations so can tell evaluate result about high retention during spring, etc.

Figure 5. Uptake should be in positive units.

Reviewer #2

GENERAL COMMENTS:
This study evaluates nitrate and dissolved organic nitrogen removal along the river network of a boreal catchment in Sweden that has been altered by forest harvests. This is an important scientific question given current forestry practices in boreal regions and predictions of increased forest use in the near future.

The manuscript is generally well written and within the scope of the journal Biogeosciences. However, there are some issues that could be addressed to improve the manuscript.

Thank you for your review and for agreeing with us, that this manuscript is well in the focus of Biogeosciences. We will do our best to address the concerns and to improve the manuscript.

- The model used to estimate N removal would benefit from estimates of uncertainty. The general assumptions made are considerable. Some apparently smaller assumptions like using the average concentration of CC-4 and NO-5 to calculate Charvest seem dangerous without considering measurements of uncertainty or running different scenarios. Moreover, the model could be better explained to the reader. A figure may be helpful in this sense. For instance, it is unclear how dilution is accounted for. It seems that nitrate removal efficiency should be calculated with the flux rather than with the concentration. The area of stream network used to calculate U should be reported.

Whereas we generally agree with the statement that the model would benefit from uncertainty estimations, we would like to clarify that the model as such is not any type of model for which a 'fitting' is performed, but only a mass balance of the measured data. This means that typical tools to evaluate the uncertainty of model parameterizations (such as for example, GLUE, Beven (2008)), cannot be used, as there are no free model parameters to choose.

Instead the model uncertainty will be present in three forms. First, structural model uncertainty (i), that is, if the mechanisms that are assumed (such as conservative mixing, closed mass-balance etc.) are reasonable. Second, there is uncertainty relating to the spatio-temporal representativeness of the input datasets (ii). Third, there is the general data uncertainty (iii), such as uncertainties in Q estimates and chemical analysis.

We argue that the uncertainties of (i) are small, which appears to find also agreement by reviewer #2. This corresponds also to follow Occam's razor and is further our model hypothesis, as we could reject the model, if the model would have not performed well for conservative tracers (see comment to reviewer #1).
Similarly, we argue that the rather extensive dataset and the replicated character of the sampling (two end members were sampled per treatment, data from 2004-2012 which will allow for some time for space substitution) will minimize the uncertainties of (iii).

Thus, what likely remains as the largest driver of model uncertainty, and in our case also of the conclusions drawn, is the uncertainty related to (ii), also explicitly criticized by reviewer #1 and #3. As we stated in our extensive response to reviewer #1, we have developed the assumptions for the clear-cut loads on what we considered the most robust approach, given the available data. As we also pointed out, the use of different scenarios is difficult. Instead we would like to include the graphs for Si and Cl mixing to argue for the validity of our model assumptions. We hope this comment addressed these concerns regarding model uncertainty.

Also, to clarify here: All calculations are done as mass fluxes, that is, concentration times discharge, as explicitly stated in the methods section. This means, that simple effects such as dilution by higher runoff from CCs will be accounted for in the model.

- Given the availability of nitrate, nitrite and ammonium data and the fact that ammonium seems to be almost as important as the other forms, I suggest that the authors redo the their calculations to estimate the dynamics and removal of dissolved inorganic nitrogen (DIN) rather than nitrate. In this way, the manuscript would cover all dissolved nitrogen forms (inorganic=DIN and organic=DON), which according to the manuscript also represent most exported nitrogen because particulate nitrogen seems very low in these streams. Moreover, there will be no need of speculation on the processes that convert nitrate to ammonium or vice versa.

This is a very good and reasonable point. We have revised our calculations so that we now calculate the fluxes and removal of the DIN pools rather than just NO3. These new results will be incorporated into the revised manuscript.

- The study design seems not justified well enough for the objectives of the manuscript. For instance, it is unclear why those catchments were chosen and why the catchments were differently harvested. The authors should explain it more clearly.

There a few paired catchment studies in boreal regions and, in fact, to the best of our knowledge non that would provide and as detailed, multi-year record of water chemistry that is needed to close the mass balance so that DIN removal can be calculated in the way we do it here. The differences in the harvests, the choices of sampling locations etc. were mainly of practical matters (Schelker et al., 2012, Schelker et al., 2013a). Also, the study design was not only designed for the purpose of this very study and is by far not perfect. We will try to address this comment in an additional sentence or two.

Furthermore, the interested reader will find several references in the methods section (Löfgren et al., 2009, Schelker et al., 2013b) that describe the Balsjö-experiment in more detail. Thus we see little need for changes here.

- The discussion seems too speculative in some parts, especially when it refers to processes and mechanisms that have not been measured in this study to explain some of the observed patterns. The authors could tone down some sentences.

We assume the reviewer points towards some of the points within the discussion also raised by reviewer #1. If not, then we would be more than happy to get to know which additional sentences specifically need to be toned down.

- The use of some terms is confusing. The authors should clearly define the terms chosen and then use them consistently throughput the manuscript. For instance, the authors should clearly define what they consider the “stream network”, and then use terms like “in-stream”, “riparian”, “landscape” consistently. Another confusing use of terms occurs when using words like “uptake”, “removal”, “retention”, etc.

Thanks for pointing this out. We will do our best to be more consistent on these.

- The title could be improved to reflect more clearly the contents of the paper. It seems too long and confusing.

We did try to use the best title we could come up with. The current title is precise in i) describing the location, climatic region and ecosystem where the study is performed, ii) that the study will present the

results of a harvest experiment and iii) what the key results of the analysis are, that is, the removal of NO3, but downstream transport of DON.

We find this title to be a good choice; however, we would be more than grateful to receive further suggestions on how the title can be further improved. For the moment, and as no additional advice on how to improve it is given, we suggest to keep it as is. Also, the title has not been criticized by reviewer #1 or #3.

As in the response to Reviewer #1, we state that we will revise the points listed as specific comments below within our revised manuscript.

SPECIFIC COMMENTS:

P12062 L17: "Landscape" here means river network or the whole catchment (including terrestrial ecosystems). Please clarify.

L21: "Net removal" within the river network? Please specify.

L22-25: Unclear sentence. Especially the part that says "capacity and limitation of N-limited . . ." Please rephrase.

P12063 L17: Some studies have. You could cite here Bernhardt et al. 2003, Riscassi and Scanlon 2009, etc. L18: I suggest adding "stream" or "river" before "network.

P12065 L4: I think that the N limitation issue could be mentioned earlier in the introduction. Moreover, its consequences for this particular study should be explained. A hypothesis may emerge from here. L3-18: I miss some hypotheses and predictions here.

P12066 L1: Why was the riparian buffer left intact in this catchment and not in the other?
L7: It seems quite strange that the samples were analyzed unfiltered. Why?
Did you make some tests to see the influence of not filtering on your DIN and DON estimates?

P12068 L24: I understand that the efficiency can be set to zero but a negative value may also mean in-stream release of NO3 (i.e. negative U values).

P12069 L16-19: It would be nice to see these different seasons depicted on the figures. This would allow the reader to follow results more easily.

P12070 L9-17: The scale of the figure does not allow seeing most of the described patterns.

P12071 L3-6: Why are U values negative? Net uptake values are usually positive if there is net uptake and negative if there is net release. I suggest changing it.

P12072 L8-16: Confusing paragraph. The supplementary figure is quite unclear and there is no figure legend or number. Unclear what is meant by upstream and downstream here and what the purpose of this paragraph is.

P12073 L11: Change to “zero or near-zero”. L17: Did you try correlations with variables other than discharge?
L29: It would be interesting to see and integrated U for the whole year (in kgN) that could be compared to other variables in Table 1.

P12075 L10-15: The effect of DNRA seems quite irrelevant here. I suggest removing these lines.

P12076 L14-19: These conclusions are ok, but they do not refer to consequences on stream network (in-stream) N removal.

Fig. 2: In the first panel it is not possible to see the temporal trends of the sites other than CC-4. Maybe you could try to use a log scale or to add a new panel/figure. What does “estimated Q” in the second panel mean? Please explain.

Fig.5: Strange to see U values as negative values. Also, I do not see the pairs of letters mentioned in the figure legend.

Anonymous Referee #3

General Comments
This is an interesting paper focused on how forest disturbances impact on stream water chemistry in boreal regions. The topic is relevant and the study fits perfectly within the scope of Biogeosciences. In general, the paper reads well and the introduction is well framed. I miss some information in the Study Site section such as the areas of the experimental catchments which can be useful to the reader for doing some back of the envelope calculations. The Methods section needs some extra work. The results are supported by a quite large amount of field and satellite data, though there are some results that need to be worked further. Rather than reporting patterns exclusively, the authors have brought the paper to a higher level by adopting a quantitative approach, which I mostly like. The discussion includes some results than need to be moved earlier in the text. Overall, I have some concerns with the applied mixing model in its present form. There are few other major issues that the authors need to solve before the paper can be published.

Thank you for your review. We hope to be able to address all the concerns and revise the manuscript accordingly.

The authors focus their study on nitrate because they argue that forest harvesting increase the mobilization of inorganic nitrogen, primarily nitrate. However, they indicate (in the discussion section) that the contribution of ammonium to the total inorganic pool in stream water is pretty high (from 20 to >50%). Therefore, by modeling only nitrate concentrations, the authors may be missing an important piece of information. I recommend showing more clearly nitrate and ammonium concentrations for the two periods of study (2004-2006 and 2007-2012). If changes in ammonium concentrations are small between the two periods, this would support the approach considered by the authors. Yet, if ammonium concentrations change substantially between the pre-harvest and post-harvest period, the authors should consider the possibility of calculating the mixing model for DIN rather than for nitrate to get a more complete picture of how forest disturbances translate downstream.

Yes, we have followed this advice and would like to provide the mass balance model for DIN, instead of only for NO3. Also, Figure 2 was revised to explicitly show NO3, NH4 and DON concentrations.

One of the major issues the authors need to deal with is the uncertainty associated with the mixing model calculations because the response to clear-cut differed tremendously between the CC4 and NO5 catchments. This issue cannot be overlooked by the authors and requires careful consideration. For instance, the concentration of the clear-cut end member (C_harvest) is characterized by averaging nitrate concentration for CC4 and NO5. Yet, results and conclusions could differ markedly from the ones presented here if authors would have used nitrate concentrations either from CC4 or NO5 alone. According to the authors, the distinct response between these two catchments may relay on the fact that riparian strips were kept in NO5 but not in CC4. If “leaving small (5-10 m) buffer zones along headwater streams is common practice” (12066.4), then one would expect that, on average, the mean response of the whole harvested area would be closer to NO5 than to CC4, being the later a more extreme scenario (savage clear-cutting without protecting riparian areas). By using the average of the two clear-cut catchments, the authors may be magnifying the “forest derived nitrate" and consequently the nitrate removal efficiency (Er) that is potentially attributed to in-stream processing.

Please find our earlier reply to reviewer #1s comment on this issue.

Another issue that the authors need to address is the implicit assumption that chemistry for the clear-cut and control end members is representative of the water draining through the whole harvested and uncut area within BA2 and BA1. Or in other words, that the chemical signature of groundwater entering to the stream outside the experimental catchments is similar to the stream water chemistry of the end members. I understand that this assumption is needed for applying the proposed mixing model, but the authors need to include this assumption explicitly in the paper and discuss the advantages and limitations of their approach.

That is an interesting point. First of all, it should be noted that these catchments are underlain by highly compacted till layers that have generally low hydraulic conductivities. Runoff generation is thus primarily from shallow saturated soil water entering streams laterally (Bishop et al., 2004, Bishop et al., 2011). Thus, and in contrast to other stream systems, contributions of deeper GW are considered minor, at least at the given spatial scale of this third order stream network.

The typical characteristics of deep GW from the underlying granitic bedrock in the region is that it is essentially free of nitrogen (DON, NO3, NO2, NH4; all normally all below detection limit). This is likely the result of a very low population density combined with a low pressure land use of forestry (for example in comparison to agriculture). Thus it appears reasonable to ask, if not deep GW inputs could have diluted the stream water causing an effect that would then be (mis-)interpreted as high NO3 removal. Such a dilution would be most likely found between the sites BA-2 and BA-1, as the small lake may be a location of GW upwelling. However, whereas such a mechanism appears generally plausible, there is little evidence for this. For example, it is very likely that the concentrations of the two conservative tracers Cl and Si (presented in response to reviewer #1), would have different concentrations in GW as in the surface water. As a result, systematic derivations of the results of the performed mixing models from the measured concentrations would occur. However, as we did not observe such derivations, we concluded that deep GW plays a minor role in modifying the water chemistry in this small stream network

Assuming a minor, negligible role of GW in the Balsjö catchment is also in agreement with our previous work (Schelker et al., 2014), as well as other work from the region that has evaluated this question in the face of DOC concentrations and found little GW influence in the till dominated regions above the highest coast line (Tiwari et al., 2014), but a stronger influence further downstream.

The interpretation of the modelled results should be explained in the Methods section rather than in the captions of the Figures. For instance, explain how the differences between modeled vs measured concentrations were interpreted, or the reasoning of why Er and Q should be or should not be related to each other.

Yes, we have moved the explanation for the modelled vs. measured plots to the methods section and deleted the Er to Q explanation as it was not the right place.

Be consistent with the presentation of Figures and add letters to identify the different panels. The second panel in Figure 2 is not referred anywhere and it is not clear what the author mean by estimated and measured Q.

We have revised Figure 2. Also, we corrected the 'estimated and measured Q' issue.

Figure 3a and Figure 2b are redundant.

That is correct. The motivation to plot Q here again was to allow for quick comparisons of different Q conditions and Er.

The results in Table 1 are not included in the results section.

We are now referring to Table 2 (former table 1) in the results section of our revised manuscript.

Specific Comments

Introduction
12064.24 clarify what you mean by "these relationships"

12065.13-18. The "questions" proposed by the authors are somehow interrelated because questions (i) and (ii) are focused on patterns, while (iii) refers to the involved processes and mechanism which lead to those observed patterns. Thus, I suggest some rewording for improving the strength of this final introductory paragraph.

Methods
12066.20. Include some more quantitative information about the areas that were harvested within the different studied catchments.

12066.24 Include for which catchments C_modelled was calculated.

12065.24. Include drainage area for the 4 experimental catchments.

12066.6. Indicate that water samples were also analyzed for chloride and silica (hydrological tracers) and that results on that were reported in a previous study (see later comment).

12067.3-10. This info could be partially moved to the Study Site section; focus this section on the description of the mixing model.

12067.17 According to eq. 1 "percentage" should be "fraction" and units would be "over 1" rather than in "%". Otherwise the factor 100 should be included in eq. 1

12067.25. The response to clear-cut differed tremendously between the CC4 and NO5 catchments. Thus, there is a substantial uncertainty associated to these calculations. There are several possibilities to deal with this problem. For instance, the author could consider either an upper and lower limit for C_modelled or different harvest scenarios (with and without keeping riparian strips).

Please see our earlier reply to reviewer #1 on this specific assumption.

12067.26. "...each scaled to 100% harvest using a scaling equation" Why the authors expect that C_harvest will increase linearly with increasing the harvest area (eq 2)? And by how much the results obtained would change if another ecosystem response (e.g. asymptotic) would be considered? The reasons behind this assumption are not clear, especially when reading later in the text that Q_harvest may not change substantially between a catchment harvested 88% or 100% (12068.12).

The assumption of a linear increase of C_harvest is of course critical, but needs to be made to be able to apply the model. Also there are at least some 'good indications' that suggest that this assumption is reasonable.

First, the concentrations of DIN of CC-4 and NO-5 scale linearly (Figure 7), if one plots them after the first of January of 2009, that is, after the harvest effects have stabilized a bit. To us this is suggestive that an increase in the harvested area will also cause an increase in DIN concentrations.

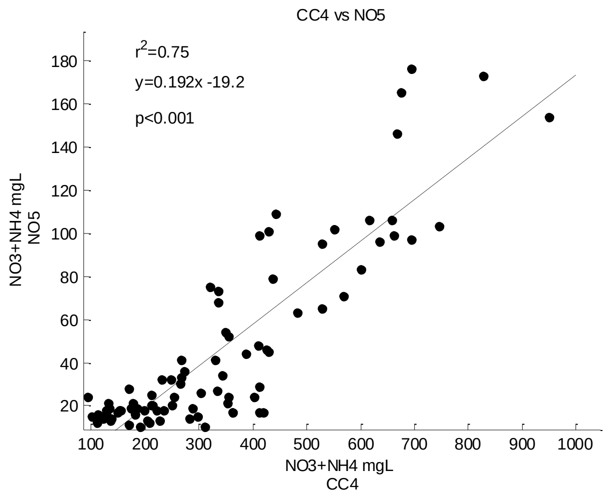

Figure 7, DIN concentrations of CC-4 vs. NO-5.

Second, the response of other solutes, such as DOC has been done using this very assumption in the past (Laudon et al., 2009, Schelker et al., 2012). Furthermore, other work on DOC concentrations along all the available Balsjö first-order streams also show a steady linear increase of the mean concentrations with increasing percentage harvesting (Schelker et al., 2014), at least for the range up to ~60% harvest.
Also, one should remember that the total percentage of the catchment area that is harvested within the catchments of the downstream sites BA-1 and BA-2 that are modelled within this study do not exceed 12% and 18% respectively (see new table 1 of the manuscript). We argue therefore that the assumption is reasonable.

12068.3 not clear what the authors mean by "reciprocal".

12068.4. Similar to C_harvested, the authors should consider some sort of confidence interval when characterizing the concentration of the control end member (C_control).

12068.19. According to the results presented Er was calculated the other way around:

(modeled – measured)/modeled.

12068.23. Values of Er < 0 could be indicating either in-stream nitrate release and/or groundwater inputs with higher nitrate concentrations than stream water. This information could be useful for discussing some of the obtained results. I recommend further considering this variable when working on the revised version of the mp.

As pointed out in one of the previous comments, there is very little nitrogen in GW in this region, commonly all below the detection limit.

12068.25. From here on, this info does not relate to the “Mixing Model”. Add a new subsection.

12069.14-16. By doing so, the authors are also assuming that stream water chemistry for the clear-cut and harvest end members is representative of the water draining through the whole harvested and uncut area within BA2 and BA1. Or in other words, that there may be no longitudinal changes in groundwater chemistry entering to the stream. Is this assumption reasonable? Do the authors have some additional data throughout the basin area to support this assumption? Could changes in groundwater inputs along the stream partially explain the observed patterns?

As pointed out before, we found no direct evidence so far, that there is an important role of GW for water chemistry in this rather small stream network. This assumption may quickly need to be revised, if one moves further downstream, where deeper lakes and larger streams are included in the stream network.

Results

12069.20 To improve the flow of this section, results could be divided in two subsections, one describing measured concentrations and fluxes; the other with the model results.

12070.15 This temporal pattern was also exhibited by CC4 but not for RS3 (as far as I can distinguish from the graph). You could reorganize these results in two paragraphs: the first focused on changes in concentration between 2004-2006 and 2007-2012 and the second focused on seasonal patterns.

12071.3-6. If U is the difference between modeled and measured fluxes, in-stream net areal uptake rates should be positive throughout the text. This would have more sense, since stream ecologists usually considered U> 0 when there is actually net nutrient uptake by stream biota.

Good point – we corrected this.

12071.3-6. Were the Us obtained for BA1 similar to those for BA2? And if not, why the bioreactive capacity of this stream may change along the longitudinal axis? The discussion of the paper would benefit if showing these results more clearly.

Discussion

12071.8-17. The authors are right in that the marked response in CC4 was not observed downstream. Yet, it will be interesting to highlight the differential response exhibited by the two harvested catchments, especially because if riparian areas are usually protected against clear-cut, the response observed for CC4 may not be widespread.

Please see our previous response to reviewer #1. One may also consider our updated Figure 2 of the manuscript that now provides a better possibility for readers to visually examine changes of DIN and DON concentrations at BA-1 and BA-2 as a response to harvests.

12071.13 The results contained in Table 1 should have been introduced in the earlier section.

12071.18-24. These are results and should be moved to the earlier section.

12071.23. Clarify to which season you refer when saying 54 and 46%.

1207120-24. The contribution of NH4 to the total inorganic N pool is quite substantial, and thus, the authors could be underestimating the potential of in-stream processes to retain and transform DIN in these catchments. I wonder how different the results would be if considering DIN rather than nitrate alone. Could the authors provide some insights on that?

12072.4. In this case, it may be clearer to refer to the years comprising the two periods than to “pre-treatment vs treatment”.

12072.8-10. These are results; moved them to the earlier section.

12072.11-16. What about NR7 and NO5? Did they show similar seasonal patterns than BA1, BA2? And if so, could one still say that “enhanced upstream inputs of nitrate in headwaters are translated downstream during the dormant season”?

12072.21-22. I recommend to briefly comment on that already in the methods section. The good match between measured and modeled concentrations for hydrological tracers would give consistency to the mixing model. Note, however, that the fact that the model works well for chloride but not for nitrate, does not necessarily imply in-stream nitrate retention because groundwater entering downstream the experimental catchments could have similar chloride concentrations but different nitrate/ammonia concentration than groundwater upstream.

Good point. It is absolutely right, that the model could even converge on one conservative tracer, but still be conceptually wrong. However, as we have now added a graph showing the model performance of two conservative tracers, the chances for the model conceptualization to be wrong are rather small.

12073.11-13. or that the concentration of nitrate was higher in downstream groundwater inputs.

See earlier comments on low GW NO3 concentrations.

12073.16-23. Avoid repeating results or adding new results in the discussion section.

Yes, thank you – we will revise accordingly.

12073.21-23. The authors should be cautious and take into consideration the uncertainties associated to these calculations before claiming that ca.70% of the nitrate inputs were removed. I suspect this figure is far too big, likely because the actual approach magnifies the effect of CC4.

Please see earlier comment on the representativeness of the CC end member. According to our understanding an 'over magnification' would have been given, if we would have only used CC-4 as the sole definition of C_harvest.

12073.24-29. According to Table1, the decrease in nitrate loads between BA2 and BA1 was <30%, which is a much lower number than the 70% proposed. Thus, and assuming that all nitrate retention was occurring within the stream channel and that there were no differences in groundwater inputs between BA2 and BA1, U values for this stream reach would be several times lower than 6 microg N/m2/min. How would the authors explain this shift in the in-stream bioreactive capacity along the stream?

We do not assume any change in bioreactive capacity along the stream, but simply close the mass-balance. Also, not all harvest are located upstream of BA-2, but some are also entering the network between BA-2 and BA-1 (see Figure 1 of the manuscript). May this be a reason for the inconsistency?

Overall we may add, that our additional table 1 will allow the interested reader to be able to make some back-of the envelope calculations with our data.

12074.11-26. These changes induced by forest harvest may be occurring only within the CC4 that (i) occupies a relatively small area of the BA2 and BA1 catchment and (ii) showed tremendous increases in nitrate concentrations. Thus I don't see how this explanation applies for patterns in BA2 and BA1.

Please find our previous comment on the representativeness of the CC end member.

12075.10-15. Too speculative. A more systematic analysis of the ammonium time data series will provide a clearer picture of whether seasonal changes in ammonium are terrestrially or stream derived. A table including ammonium and nitrate concentrations for the two periods could be useful.

We have added these concentrations to the revised version of Figure 2. We hope this allows the reader to evaluate this data.

12076.11. Not clear to which "two mentioned measures" the authors refer.

This needs clarification then also.

Figures and Tables

Figure 2. The second panel is not referred within the main text. Based on the Methods

section it is not clear what the authors mean by estimated vs measured Q.

We have revised the Figure, see above.

Figure 3. Differences between modelled and measured concentrations could indicate biogeochemical retention of the solute during transport downstream but also hydrological mixing with sources with different chemical signature. I recommend including the interpretation of the results in the Methods section where the authors can link the expected patterns to the assumptions underlying the model.

Yes, we have revised this.

Figure 4. Panel (A) is not introduced in the main text and it is redundant with Figure 2. If Er >0 means nitrate retention, then the differences is between modeled and measured concentrations nitrate concentration. Why did the authors explore the dependency of Er on Q? This should be explained in the Methods section.

Figure 5. Positivize U values. To avoid any confusion to the reader, highlight in the caption that this is a potential maximum value for in-stream uptake. The letters for statistical significance are not included in the figure. Show data for BA1; differences in U values between BA2 and BA1 can enrich the discussion by supporting (or not) the explanations given for in-stream nitrogen processing.

Table 1. This table and results therein should be included in the results section.

Figure SS1. Please include the caption of this figure. Include data for the snowmelt period to be consistent with the data analysis throughout the mp.

We will do our best to also address the comments below in our revised version.

Technical Corrections

12064.18. “photoautotrophic” rather than “autotrophic”

12067.7 or 2001-2011? (as in caption figure 1)

12068.26 Change “net uptake rates” by “net areal uptake rates” throughout the mp.
12069.7. change “treatment” by “clear-cut”.

12069.15. Change “loss” by “export”.

12069.24. Delete “buffer”

12070.1 and 4. Which treatment? Clarify.

12070.23. Change “nearly exclusively” by “usually”.

12072.4. Delete the “-“ sign.

References

*Schelker J, Öhman K, Löfgren S, Laudon H (2014) Scaling of increased dissolved organic carbon inputs by forest clear-cutting – What arrives downstream? Journal of Hydrology, **508**, 299-306.*

Tiwari T, Laudon H, Beven K, Ågren AM (2014) Downstream changes in DOC: Inferring contributions in the face of model uncertainties. Water Resources Research, n/a-n/a.

*Schelker J, Grabs T, Bishop K, Laudon H (2013a) Drivers of increased organic carbon concentrations in stream water following forest disturbance: Separating effects of changes in flow pathways and soil warming. Journal of Geophysical Research: Biogeosciences, **118**, 2013JG002309.*

*Schelker J, Kuglerová L, Eklöf K, Bishop K, Laudon H (2013b) Hydrological effects of clear-cutting in a boreal forest – Snowpack dynamics, snowmelt and streamflow responses. Journal of Hydrology, **484**, 105-114.*

*Schelker J, Eklöf K, Bishop K, Laudon H (2012) Effects of forestry operations on dissolved organic carbon concentrations and export in boreal first-order streams. Journal of Geophysical Research, **117**, G01011.*

*Bishop K, Seibert J, Nyberg L, Rodhe A (2011) Water storage in a till catchment. II: Implications of transmissivity feedback for flow paths and turnover times. Hydrological Processes, **25**, 3950-3959.*

*Laudon H, Hedtjärn J, Schelker J, Bishop K, Sørensen R, Ågren A (2009) Response of Dissolved Organic Carbon following Forest Harvesting in a Boreal Forest. Ambio, **38**, 381-386.*

*Löfgren S, Ring E, Von Brömssen C, Sørensen R, Högbom L (2009) Short-term Effects of Clear-cutting on the Water Chemistry of Two Boreal Streams in Northern Sweden: A Paired Catchment Study. Ambio, **38**, 347-356.*

Beven K (2008) Environmental Modelling: An Uncertain Future, Routledge: London.

*Roberts BJ, Mulholland PJ (2007) In-stream biotic control on nutrient biogeochemistry in a forested stream, West Fork of Walker Branch. Journal of Geophysical Research: Biogeosciences (2005–2012), **112**.*

*Bishop K, Seibert J, Köhler S, Laudon H (2004) Resolving the Double Paradox of rapidly mobilized old water with highly variable responses in runoff chemistry. Hydrological Processes, **18**, 185-189.*