

Dear referee

Thank you for taking your time to comment on this manuscript. First off, we are very sorry for the amount of technical errors in this manuscript! We have proofread the manuscript very carefully and hope that all the errors has been corrected. If the editor thinks that further proofreading could benefit the manuscript then we will have it proofread by a professional, when the final corrections has been added. We will address all the referee questions below.

Catchment tracers reveal discharge, recharge and sources of groundwater-borne pollutants in a novel lake modelling approach

General comments

I will give some more general comments in the following part while some further specific and technical aspects are listed below this section.

1. The introduction might need a bit more structure. There is a lack of background information especially on the functioning of the tracer methods and especially the model applied. It would be good to introduce the different tracers and their functioning in detection of groundwater-lake interactions in more detail. While the investigation of oxygen isotopes might be familiar to a lot of readers, the functions and applicability of the different DOC/DOC fractions as a tracer for groundwater interactions with lakes might be helpful to know. This is especially important for the development of the hypothesis/research question where the FDOM fraction plays an important role. Also, the reader might need more information about the meaning of WRTs in this context (elaborate this in more detail in Line 65).

Furthermore, the CATS model as the novelty in the study because it is applied for the first time in such a context, is not mentioned in the introduction at all. I feel that the quality of the manuscript would improve quite a bit if the authors try to step back from their own view and reflect the structure and the information given in the introduction from the point of the potential reader. Elaborate in more details the links between the different tracers, the WRT, the CATS model etc. Also, give a brief and general description of the study design and the methods applied in order to allow the reader to go for the details in the Material and Methods-section following.

We agree that the introduction needs to describe the tracer methods and the model used in more detail. This, is why we've changed 2/3 of the introduction and added further information regarding FDOM, CATS, CDOM, DOC and nutrients, which is specified below.

“FDOM:

“Some non-conservative tracers such as fluorescent dissolved organic matter (FDOM), which can be determined using parallel factor analysis (PARAFAC), has been used to trace dissolved organic matter (DOM) in aquatic environments (He et al., 2014; Massicotte and Frenette, 2011; Stedmon et al., 2003; Stedmon and

Markager, 2005b; Walker et al., 2009). PARAFAC analysis is a modelling tool which can separate multiple FDOM samples (emission and excitation matrices) into specific fluorescent components (Stedmon et al., 2003). The fluorescent components can be biological produced proteins derived from bacteria or molecules from the degradation of terrestrial organic material. These components has previously has been found visually using a single excitation emission matrix and the observed fluorescent peaks (Coble, 1996). The differentiation between the fluorescent components are both a strength and a weakness as we can isolate many different components, but all of them can differ in both degradation and production rate in the lake and groundwater. Furthermore, these FDOM components have not yet been investigated as tracers in groundwater fed lakes, as they, just as the rest of the non-conservative biological tracers, are volatile.”

CATS:

As the concentrations of both conservative and non-conservative tracers in a groundwater fed lake correspond to the mixed concentrations of discharging groundwater, while taking degradation and atmospheric deposition into account, it is possible to utilize the Community Assembly via Trait Selection approach (CATS). This model has been used to predict the relative abundances of a set of species from measures of community-aggregated trait values (average leaf area, root length etc.) for all plant species at a site (Shiple, 2010; Shipley et al., 2006, 2011). The CATS model has three main parameters: (1) it models the probabilities (2) that maximize the entropy (3) based on a set of constraints (Laliberté and Shipley, 2011; Shipley et al., 2011). In reality, the model (1) predicts the relative abundances of species at a location from their (3) average traits values by (2) minimizing the number of species that explain the mean traits values. The maximum entropy (2) is the maximizing of “new knowledge gained”, related to plant communities this means that we are moving from “all species has the same relative abundances” to “a few species has a high relative abundance”. When applying the model to the lake-groundwater interaction we use the measured tracer concentrations at groundwater well sites around the lake as the individual plant species and the estimated mixed lake concentration before degradation took place as the community-aggregated trait values.

CDOM, DOC and Nutrients:

This is observed as a change in tracer concentrations (often a decrease) after the groundwater is discharged to the lake. The speed of which the change in concentration happens are mostly related to seasonal variations (e.g. temperature, mixing of the water column and UV-radiation) and the WRT of the lake e.g. the amount of time the tracer is in the lake. The removal and degradation rates have been examined in many instances e.g. for phosphorus (Larsen and Mercier, 1976; Vollenweider, 1975), nitrate (Harrison et al., 2009; Jensen et al., 1995), CDOM and DOC (Madsen-Østerby et al., 2017). In a modelling approach these rates are important as they provides information about the change in tracer concentration, from the time when the tracer entered the lake. From this it is possible to back-calculate what the mixed inflowing concentration of specific tracers were when they were discharged to the lake. These estimations are crucial when working with non-conservative tracers as it enables a direct comparison between the tracer concentration found in the catchment and the estimated mixed lake concentration before degradation took place.”

2. Having said this, the authors start the method section by explaining groundwater sampling and sample/data analysis without introducing to the specific idea/concept/design of the study. This might be acceptable if the concept was already described in the introduction. However, as already mentioned above this is not the case which is why I strongly recommend adding the necessary information here.

We agree that the introduction needs to provide the specific idea, concept or design of the study. This is why we have added information regarding the CATS model (shown above) as well as information in the section, which describes the hypothesis of the study.

The section now reads:

“Determining groundwater movement using both conservative and non-conservative tracers found around the lake shore overcomes some fundamental shortcomings related traditional sampling. Firstly, we often measure tracers, which do not have a direct impact on the lake ecosystem and therefore do not provide meaningful information regarding the inflow of nutrients or CDOM. Furthermore, the sampling is only done in a few places throughout the catchment, which do not necessarily provide the all information regarding the groundwater flow patterns or to which degree water enters the lake and where. To overcome this, we measured conservative and relevant non-conservative tracers in and around a small lake with the aim of developing a novel approach to identify groundwater discharge and recharge areas on a high spatial scale. Thereby, pin-pointing areas which delivers pollutants to the lake, where groundwater recharge happens and where recharge occur with an increased flowrate of which the latter can spark further investigations into the lake WRT. Information regarding the WRT of the lake is especially useful when investigating how the concentrations of pollutants in the lake will develop after future restoration attempts. In the present study, we measured the eight following tracers: FDOM, CDOM, DOC, total dissolved phosphorus (TDP), total dissolved nitrogen (TDN), total phosphorus (TP), total nitrogen (TN) and $\delta^{18}\text{O}/\delta^{16}\text{O}$ isotope ratios and tested: (1) if groundwater discharge sites and pollutant sources can be estimated with the CATS model based on tracer concentrations, (2) if conservative and non-conservative tracers can be used to detect groundwater recharge areas as well provide insights into which areas have a high groundwater recharge rate and (3) if catchment-derived tracer concentrations can be used to estimate a range of WRTs which can be used with the CATS model. “

Since the application of the CATS model is a crucial part of the study it should be already described in the introduction.

We agree and have added it to the introduction (see above)

Also, since the reader until now is not aware of the concept of the study the headings “Groundwater recharge and areas of high recharge” (L132) is quite confusing since at first site such a differentiation does not make very much sense.

We agree and have changed the heading to:

“Groundwater recharge and areas with a high groundwater recharge rate “

Similar to that the heading “Groundwater discharge and lake WRT” (L135) makes the reader wonder why these two aspects are looked at together.

We agree and have changed the heading to:

Non-conservative tracer degradation and lake WRT”

Furthermore, we have included groundwater discharge in a separate section named **“The CATS model”**

3. About the CATS model:

- a. The model identifies those sites around the lake which contribute most significantly to the lake’s tracer concentrations. From this, you conclude that at these sites most groundwater discharge takes place. However, if concentrations are high discharge volumes do not necessarily have to be large as well to influence the lake concentrations. For these kinds of approaches there are usually end-member-mixing models applied which I guess also the CATS model is of such a kind.

Yes, the CATS model resembles an end-member-mixing model. In this regard, the number of tracers used and their differentiation in concentrations are very important. Thus, higher number of tracers and higher un-correlated differences between the tracers results in a more secure determination of groundwater discharge sites.

The following has been added to the materials and methods section:

“When choosing tracers, it is important that there is differentiation between the concentrations measured at the sites. This implies that a higher number of tracers and higher un-correlated concentration differences between the sites results in a more reliable determination of groundwater discharge sites.”

- b. L147: What are “entropy probability fractions of the groundwater...”

We have added this section to the introduction and changed the sentence to make it clearer:

“The CATS model has three main parameters: (1) it models the probabilities (2) that maximize the entropy (3) based on a set of constraints (Laliberté and Shipley, 2011; Shipley et al., 2011). In reality, the model (1) predicts the relative abundances of species at a location from their (3) average traits values by (2) minimizing the number of species that explain the mean traits values. The maximum entropy (2) is the maximizing of “new knowledge gained”, related to plant communities this means that we are moving from “all species has the same relative abundances” to “a few species has a high relative abundance”.”

c. L150f: “The model also predicts lambda values from the least squares regression explaining which tracers are most influential on the relative fractions of water originating from the groundwater well sites.”

→ do not fully understand what that means. To me this implies that depending on the tracer looked at the model output differs, i.e. if you look at P there are different sites relevant for lake concentrations than compared to CDOM. And to me that totally makes sense because the concentrations may spatially vary a lot within the same or among different parameters.

Regarding the lambda values. Lambda values are a measure of how much each tracer influence the model outcome. Meaning that, if all sites had the same tracer concentrations we would also have the same lambda values among tracers (which would be 0), but because we have differences in tracer concentration between sites we see a variation in lambda values. A higher lambda value means that the tracer varies a lot between sites and the sites varies from the estimated lake concentration – this means that a change in e.g. CDOM concentration would have a great impact on the model outcome.

We have added the following regarding lambda values to the introduction and the results section:

“The model also computes lambda values from the least squares regression measuring which tracers are most influential on the relative fractions of water originating from the groundwater well sites. Lambda values therefore quantifies how much the relative contribution from the sites change when one tracer is changed a unit while the rest of the tracers are kept constant. “

and

“Lambda values, explaining which tracers are the most important when predicting the fractions of water origination from groundwater well sites, showed that CDOM was the most important tracer when determining which sites delivered water to the lake with a mean lambda value for all WRTs of 24.2 versus 0.1-5.9 for the other tracers.”

Regarding the part related to 3a, the model does not see the tracers as individuals at a site it rather looks at them as a “package” with each site having a distinct fingerprint of tracers. The sites delivering water to the lake are therefore determined by the combination of tracers they have and the differences amongst them makes the model stronger.

However, as already mentioned in 3a. the lake concentrations do not only depend on the concentrations but also on the volumes of groundwater discharge. But the discharge volume of a site is (to my understanding) completely independent from the parameter concentration which means that it is not possible to deduce

discharge volumes/portions from parameter concentrations. By that the model does not give you any reliable output of relevant discharge sites.

We have described how each site are seen as a “package” and the differences between the tracers makes the model able to estimate discharge fractions in the answer to 3a.

- d. How is the water retention time implemented in the CATS model? I guess the CATS model initially didn't have such kind of parameter to be included in the algorithm.

The CATS model was developed for terrestrial plant communities and does therefore not have any parameter regarding WRT. Instead WRT is incorporated into the community aggregated trait mean. Meaning, that to every WRT there is a certain mixed groundwater discharge tracer concentration if we are to measure the concentration actually found in the lake.

We have added the following to the introduction and materials and method section to clarify this:

“This is observed as a change in tracer concentrations (often a decrease) after the groundwater is discharged to the lake. The speed of which the change in concentration occurs are typically related to seasonal variations (e.g. temperature, mixing of the water column and UV-radiation) and the WRT of the lake e.g. the amount of time the tracer has been in the lake. The removal and degradation rates have been examined in many instances e.g. for phosphorus (Larsen and Mercier, 1976; Vollenweider, 1975), nitrate (Harrison et al., 2009; Jensen et al., 1995), CDOM and DOC (Madsen-Østerbye et al., 2017). In a modelling approach these rates are important as they provide information about the change in tracer concentration, from the time when the tracer entered the lake. From this, it is possible to back-calculate the mixed inflow concentration of specific tracers when they were discharged to the lake. These estimations are crucial when working with non-conservative tracers, as it enables a direct comparison between the tracer concentration found in the catchment and the estimated mixed lake concentration before degradation took place.”

“WRT of the lake were found using traditional hydrological methods combined with non-conservative tracer concentrations which were related to their degradation rates to form a proxy for the maximum WRT. Previous hydrological models suggested that the lake had a WRT between 0.4 and 3.3 years. To further narrow this range, we estimated the WRT by relating the concentrations found in the lake to their respective degradation rates related to increasing WRT e.g. by adding the estimated removed tracer since the groundwater entered the lake to the measured concentration in the lake. This enabled us to narrow the span of the WRT based on the estimated mixed inflowing tracer concentration related to the actual catchment concentrations. E.g. if the estimated inflow concentration of a tracer is $100 \mu\text{mol l}^{-1}$, at a WRT of 2 years, and the highest catchment tracer concentrations is $50 \mu\text{mol l}^{-1}$ then the catchment do not support a WRT of 2 years.”

About the results section: Provide more evidence for what you describe in the text by figures, or tables.

We have added a flow diagram to ease the reading of the materials and method section as well as the results section. Furthermore, we can add the estimations of the mixed inflowing tracer concentrations at different WRTs if the editor wishes that?

4. About the discussion: You start the section by phrasing three questions which are not all corresponding to the research questions phrased in the introduction. Try to be consistent here.

The 1st question was referring to the main questions from the introduction while the 2nd and 3rd question were based on a more general discussion of the results.

To clarify this, we have changed the sentence so it now reads.

“Based on the model results and earlier hydrological studies we will discuss the main questions from the introduction (1) if groundwater discharge sites and pollutant sources can be estimated with the CATS model based on tracer concentrations, (2) if conservative and non-conservative tracers can be used to detect groundwater recharge areas as well provide insights into which areas have a high groundwater recharge rate and (3) if catchment-derived tracer concentrations can be used to estimate a range of the water retention times which can be used with the CATS model. Furthermore, we will discuss which of the tracers works and which could possibly work with refined methods? And how these findings could benefit lake restoration programs.”

Specific comments

L14 Within 5-45 m distance (?) to the shore

Yes. This has been corrected and now reads:

“installed every 50 m within a distance of 5-45 m to the shore”

L17 What is WRT?

This has been corrected and now reads:

“water retention times (WRTs)”

L30 The sentence implies that nutrients such as P, N, and C are contaminants which is not necessarily true. See also Line 10 in the abstract.

We agree that this is not always the case, at all, we hope it is clear as we write:

“Particularly in smaller lakes and ponds the groundwater contributes nutrients, dissolved organic carbon (DOC), colored dissolved organic matter (CDOM) or other contaminants, which can have a negative impact on the

biological quality of lakes”

Stating that it can have a negative impact.

L45 The water does not leave the lake bottom but the lake via the lake bottom.

This has been corrected

L52 about environmental tracers: Besides atmospheric tracers: What about lithospheric or pedosphere tracers?

The sentence has been changed to include lithospheric or pedospheric tracers from the catchment:

“These tracers are divided into three main categories: (1) environmental tracers (natural derived tracers from the atmosphere or catchment which are transported to the system)”

L58 I do not understand the last part of the sentence: What does “groundwater of different ages and origins” mean in the context of percolating groundwater?

“groundwater of different ages and origins” refers to that different catchments and flow rates influence the chemical levels in the groundwater.

The sentence has been changed to:

“Therefore, we propose a different approach utilizing both conservative and non-conservative tracers such as dissolved organic carbon and nutrients which are partly transferred to the percolating groundwater on its way to the lake (Kidmose et al., 2011). “

L60ff In this paragraph you are talking about the fate of inflowing components. However, (although for many readers probably quite obvious) you do not explain why this is important for your study.

The whole section has been rewritten and should now clearly state why the fate of the inflowing tracers is important for this study. The section now reads:

“This is observed as a change in tracer concentrations (often a decrease) after the groundwater is discharged to the lake. The speed of which the change in concentration occurs are typically related to seasonal variations (e.g. temperature, mixing of the water column and UV-radiation) and the WRT of the lake e.g. the amount of time the tracer is in the lake. The removal and degradation rates have been examined in many instances e.g. for phosphorus (Larsen and Mercier, 1976; Vollenweider, 1975), nitrate (Harrison et al., 2009; Jensen et al., 1995), CDOM and DOC (Madsen-Østerbye et al., 2017). In a modelling approach these rates are important as they provide information about the change in tracer concentration, from the time when the tracer entered the lake. From this, it is possible to back-calculate the mixed inflowing concentration of specific tracers when they were discharged to the lake. These estimations are crucial when working with non-conservative tracers, as it

enables a direct comparison between the tracer concentration found in the catchment and the estimated mixed lake concentration before degradation took place.”

L74 Measurements were done not only around the lake but also in the lake, i. e. of the lake water, right?

Yes, this has been corrected

L74f Why are the existing approaches/tracers not sufficient? Maybe add something like “... to overcome....” to the first sentence of that paragraph.

We have added the following to describe why existing approaches and tracers are now sufficient:

“Determining groundwater movement using both conservative and non-conservative tracers found around along the lake shore overcomes some fundamental shortcomings related traditional sampling. Firstly, we often measure tracers which do not have a direct impact on the lake ecosystem and therefore do not provide meaningful information regarding the inflow of nutrients or CDOM. Furthermore, the sampling is only done in a few places throughout the catchment which do not necessarily provide the all information regarding the groundwater flow patterns or to which degree water enters the lake and where. To overcome this, we measured conservative and relevant non-conservative tracers in and around a small lake with the aim of developing a novel approach to identify groundwater discharge and recharge areas on a high spatial scale. Thereby, pin-pointing areas which delivers pollutants to the lake, where groundwater recharge happens and where recharge occur with an increased flowrate of which the latter can spark further investigations into the lake WRT. Information regarding the WRT of the lake is especially useful when investigating how the concentrations of pollutants in the lake will develop after future restoration attempts.”

What methods/tracers do you actually combine and develop towards a new approach? Be more precise and at least list them here.

The section now reads:

“In the present study, we measured the eight following tracers: FDOM, CDOM, DOC, total dissolved phosphorus (TDP), total dissolved nitrogen (TDN), total phosphorus (TP), total nitrogen (TN) and $\delta^{18}\text{O}/\delta^{16}\text{O}$ isotope ratios”

L75ff Here you come up with a maximum entropy model which was not mentioned before. Is that the novel approach you are talking about before? So far, the reader might have considered the application of the tracers more or less described above to be the novel approach in this study.

Information on the maximum entropy and CATS model has now been added to the introduction and can be seen in the answer to question 1.

L77f What would be the benefit of a combined investigation of conservative and non-conservative tracers? And

again: Which tracers will you use for this investigation?

We have added this to the introduction which can be seen in the answer to L74f.

L78f Why are you interested in WRT? Where is the connection to the groundwater-lake interactions?

We have changed the sentence and added further information to the introduction to explain why WRT is of importance:

“(3) if catchment-derived tracer concentrations can be used to estimate a range of the water retention times which can be used with the CATS model. “

“Information regarding the WRT of the lake is especially useful when investigating how the concentrations of pollutants in the lake will develop after future restoration attempts.”

L86 “to bypass water”: Is that surface runoff that is collected in the drainage channel?

Before 1992 there were at stream going through the lake. We have specified that in the sentence:

“This led to a restoration attempt in 1992 where a channel was established to bypass the stream going through the lake, thus making the lake groundwater fed.”

L93 Does “5-45 m” relate to the distance to the shore?

Yes. We have adjusted the sentence:

“A total of 30 groundwater samples were taken every 50 meters around the lake, within a distance of 5-45 m to the shore, in temporary groundwater wells at 1.25 meters of depth in February 2016.”

L111 What exactly are the three steps of PARAFAC modelling?

The three-way modelling is done on the basis of (1) multiple (2) emission and (3) excitation FDOM signals. Information regarding PARAFAC has now been added to the introduction. Specifically, this sentence:

“PARAFAC analysis is a modelling tool which can separate multiple FDOM samples (emission and excitation matrices) into specific fluorescent components (Stedmon et al., 2003)”

L111 What information do the “specific fluorescent components” provide? Be more specific here (and maybe introduce this background information on FDOM already in the introduction)

Information regarding PARAFAC and FDOM has been added to the introduction including this sentence:

“The fluorescent components can be proteins derived from bacteria or molecules from the degradation of

terrestrial organic material. These components has previously been found visually using single excitation emission matrixes and their observed fluorescent peaks (Coble, 1996).”

L112 What is the “inner filter effect”?

An explanation of inner filter effect has been added to the materials and method section:

“The FDOM samples were initially diluted 2-12 times to account for self-quenching, also referred to as inner filter effect, which occurs with high CDOM absorbance in the sample (Kothawala et al., 2013).”

L113 “...which were measured spectrophotometrically ...”does this refer to CDOM or FDOM and if it refers to CDOM it implies that the samples from each site were measured twice for CDOM (see L99f). Why was that?

The sentence has been removed as this measurement of absorbance were only done to determine the dilution factor of the FDOM samples to correct for inner filter effect. This is done by measuring the total absorbance (the sum of absorbance at excitation and emission wavelengths) which should not be above 0.042.

L117 How can the data be “Raman normalized” when the “Raman scattering” was “removed” from the data before?

This sentence has been changed in the materials and methods section to explain how the normalization took place:

“The data were then Raman normalized by dividing the florescent intensities by the integral of the Raman peak of the blank sample (Lawaetz and Stedmon, 2009)”

L129f “Changes in ...” This sentence was probably supposed to explain the concept of using oxygen isotopes as tracers for groundwater-lake interaction but in fact it does not. The authors should rework and complement the description of the use of O2 isotopes (again, this general introduction of the approach should probably go to the introduction section of the manuscript).

The following section has been added to the introduction to account for isotope utilization as a tracer for groundwater movement and determination of WRT.

“Precipitation-derived environmental tracers, such as the isotope $\delta^{18}\text{O}$ (reported in the Vienna-standard mean ocean water (SMOW) where $\delta_{\text{sample}} \text{‰} = 1000((R_{\text{sample}}/R_{\text{smow}})-1)$ and R is the $\delta^{18}\text{O}/\delta^{16}\text{O}$ ratio (Turner et al., 1987)), have been used to trace the interaction between ground and surface-water. As evaporation occurs in the surface water it becomes enriched with $\delta^{18}\text{O}$ producing a unique lake $\delta^{18}\text{O}/\delta^{16}\text{O}$ ratio which can be traced in the areas with groundwater recharge (Krabbenhoft et al., 1990). The isotopic composition can be also be related to evaporation lines (from the $\delta^{18}\text{O}$ relationship $\delta^2\text{H}$) to estimate WRT (Gates et al., 2008; Gibson et al., 2002).“

L130 “However, deviations from lake ^{18}O concentration was not observed in areas with groundwater recharge due a sampling depth of 1.25 m close to the lake.” do not understand this sentence. ^{18}O values in groundwater similar to lake water values indicate lake water infiltration into the aquifer which is why the sentence to me seems to be a circular reference. Also, I am not sure if the last part of the sentence refers to the sampling depth or the distance to the lake shore.

We have removed this sentence which was previously added to explain why we chose $\delta^{18}\text{O}$ as a tracer for groundwater discharge. And added that to the introduction:

“As evaporation occurs in the surface water it becomes enriched with $\delta^{18}\text{O}$ producing a unique lake $\delta^{18}\text{O}/\delta^{16}\text{O}$ ratio which can be traced in the areas with groundwater recharge (Krabbenhoft et al., 1990).”

L131 “Groundwater well sites which clustered with the lake were considered as being groundwater recharge sites and were removed for the later estimations of groundwater discharge sites.” What does this sentence mean?

We have clarified this in the following sentence and have added a flow diagram which illustrates the work flow used in this study:

“Groundwater well sites which formed a cluster together with the lake samples were considered as being groundwater recharge sites, e.g. water originating from the lake, and was excluded for the later estimations of groundwater discharge sites. The groundwater recharge sites were further investigated using a range non-conservative tracers influenced by biological degradation.”

L132f If I understand the meaning of the sentence and its link to the o_2 isotope section above correctly I would recommend modifying it as follows:

“Hierarchical **Euclidean** clustering was also done for the fluorescence components from the PARAFAC modelling.”

To clarify the use of fluorescent components in a Hierarchical Euclidean cluster we have added the following to the material and methods section:

“The groundwater recharge sites were further investigated using a range of non-conservative tracers influenced by biological degradation. We found that some of the tracer concentrations changed when going from the lake to the groundwater. For example, CDOM which showed a decrease in concentration when entering the groundwater properly due to pH changes in the soil. An inspection of the results revealed that a protein based fluorescent component met our criteria of being (1) non-conservative, (2) not afflicted by the lake-groundwater interface and (3) not to easily degraded or produced in high amounts which could create false positive groundwater recharge sites. The PARAFAC component was related to the lake concentration with a Hierarchical Euclidean cluster dendrogram and the sites which clustered together with the lake samples indicated a high groundwater recharge rate.”

L133 This is not a complete sentence. It does not have a verb. I assume that the authors want to express that they applied two different tracers in the same way in order to validate their findings.

We have completely rewritten this part, please see the section above.

L136ff This applies only in exclusively groundwater-fed lakes which should be indicated in this sentence.

This part has been rewritten

L138ff “Probabilities” of what?

What are the “constraints” and the “linear features”? If possible try to explain these very abstract terms and relations a bit more. If this is not possible I would consider removing this sentence.

We have added a section regarding the CATS model in the introduction and made it much simpler:

“As the concentrations of both conservative and non-conservative tracers in a groundwater fed lake correspond to the mixed concentrations of discharging groundwater, while taking degradation and atmospheric deposition into account, it is possible to utilize the Community Assembly via Trait Selection approach (CATS). This model has been used to predict the relative abundances of a set of species from measures of community-aggregated trait values (average leaf area, root length etc.) for all plant species at a site (Shiple, 2010; Shiple et al., 2006, 2011). The CATS model has three main parameters: (1) it models the probabilities (2) that maximize the entropy (3) based on a set of constraints (Laliberté and Shiple, 2011; Shiple et al., 2011). In reality, the model (1) predicts the relative abundances of species at a location from their (3) average traits values by (2) minimizing the number of species that explain the mean traits values. The maximum entropy (2) is the maximizing of “new knowledge gained”, related to plant communities this means that we are moving from “all species has the same relative abundances” to “a few species has a high relative abundance”. When applying the model to the lake-groundwater interaction we use the measured tracer concentrations at groundwater well sites around the lake as the individual plant species and the estimated mixed lake concentration before degradation took place as the community-aggregated trait values.”

L140f I assume that “species” are “plant species”?

Yes, the sentence now says “plant species”

L142 At least give information on the meaning of the abbreviation of the R package mentioned.

This has been added: “FD (functional diversity)”

L143ff For which parameters/tracers have you run the CATS model? deltaO2, N, P, CDOM, FDOM, DOC? And

was the model run for each parameter separately or for all parameters simultaneously? In the first case: Would you not have to go for a more detailed discussion of the outcomes of the different model runs?

All parameters are incorporated in the model as a distinct fingerprint or package for each site. We have completely changed the section which now reads:

“In the present study, the concentrations of non-conservative tracers (DOC, CDOM, TDP and TDN) at groundwater well sites around the lake acted as the individual plant species at a site and the equilibrium tracer concentrations derived from Eq. 1 (DOC, CDOM, TP and TN) acted as the community-aggregated trait values. When choosing tracers, it is important that there is differentiation between the concentrations measured at the sites. Meaning that a higher number of tracers and higher un-correlated concentration differences between the sites results in a more secure determination of groundwater discharge sites. All tracers were investigated as a combined package, e.g. a single site are described by all the tracers mentioned above, and was run using the maxent function in the FD (functional diversity) package in R to compute the CATS model (Laliberté and Shipley, 2011).”

L145ff Isn't it more like the model identifies those sites along the lake which contribute most to the lake?

This has been clarified in the sentence:

“From this, the model predicts the minimum number of groundwater well sites along the lake shore that explains the measured concentrations in the recipient lake by maximizing the sites relative contribution.”

L153 Is “equilibrium” the tracer concentration in the lake which would be the same as MIC in Eq. 1?

Yes, this has been clarified in the material and methods section and in the result section:

“In the present study, the concentrations of non-conservative tracers (DOC, CDOM, TDP and TDN) at groundwater well sites around the lake acted as the individual plant species at a site and the equilibrium tracer concentrations derived from Eq. 1 (DOC, CDOM, TP and TN) acted as the community-aggregated trait values.”

And

“Equilibrium tracer concentrations of DOC, CDOM, TDP and TDN (found using Eq. 1-4) for water retention times between 0.25 and 3.5 years in 0.25 increments revealed that concentrations of TDN in the catchment are not high enough to support WRT-values over 2 years based on the nitrogen retention model used.”

L173ff “The combined summer UV-radiation and bacterial degradation rates of DOC and CDOM in groundwater from the dominating catchment vegetation type of the lake (Madsen Østerbye et al., 2017) were extrapolated to the rest of the year.”

- a. What does the vegetation type of the lake have to do with UV- and microbial degradation of C-fractions?

The vegetation type in the catchment of the lake matters as the degradation of CDOM and DOC are influenced by the molecular structure. The molecular structure change depending of the vegetation in the catchment (e.g. heathland or coniferous forest)

- b. Why was it necessary to extrapolate to the rest of the year? To get the specific removal rates for month of the sampling?

UV-radiation acts as a photo catalyst for the degradation of CDOM and DOC. Thus, degradation is correlated to the amount of UV-radiation which varies yearly (high levels during summer and low levels during winter). To estimate the degradation in the lake during the past months/years we therefore have to know how much UV-radiation there has been during the months of degradation.

L174f “This was done by relating the rates to the mean monthly UV index (DMI, 2015) while assuming a linear relationship between the UV-index and degradation rates.” is there are literature reference which proves that this a legitimate way of doing this?

Almost all UV-absorption in the lake is done by CDOM and DOC in the top 5 cm of the water column. CDOM concentrations are therefore highly correlated to the yearly UV-radiation. The degradation follows a seasonal pattern described in ““Photodegradation of DOC in a shallow prairie wetland: evidence from seasonal changes in DOC optical properties and chemical characteristics MARLEY J. WAISER* and RICHARD D. ROBARTS”.

L182f “Eq. 4 was solved in relation to τ_{inflow} and calculated using the same WRTs as the nitrate and phosphorus models.” do you mean “... was solved for τ_{inflow} ...”?

Yes, this has been corrected.

L191 The split-half analysis modelling has never been mentioned before. What is that and what does it do?

We have added a sentence regarding this in the materials and methods section which explains the role of the “split-half” analysis:

“A split-half analysis, were the dataset is split into two parts and compared multiple times, were used to test the results found in the PARAFAC model.”

L192 “Component C1 was similar to previously found humic-like material...” does this relate to material of this study site? Which one? Terrestrial, aquatic?

The component, a molecular structure dissolved in the sample water, here named component C1 has been found before. By comparing the specific excitation-emission signal shown in fig. 3 to other studies we are able classify it as a component that most likely originate from the degradation of terrestrial matter (leaves etc.). The component is washed into the groundwater and further into the lake from the soil.

L192ff “The component absorbs in the UV-C region which has low intensities at the ground surface (Diffey, 2002) and are is therefore expected to be photo-resistant (Ishii and Boyer, 2012).”

- a. Is “ground surface” = lake bottom?
- b. I am not sure why you can conclude that the material is photo-resistant from the fact that UV-C has low intensities at the lake bottom.

By ground surface we mean the surface of the earth. We have changed the sentence:

“The component absorbs in the UV-C region which is absorbed by the ozone layer and atmosphere (Diffey, 2002) and is therefore expected to be mainly photo-resistant (Ishii and Boyer, 2012).”

L199 “Component C3 may be an intermediate product since concentration changes even in open oceans”

- a. Intermediate product from what?
- b. What do changes in concentrations in the open ocean imply?

We have changed the sentence to address your question:

“Component C3 may be an intermediate product or produced biologically since changes in the concentration have been observed in the open oceans and in sea ice which has no apparent connection to the terrestrial environment (Ishii and Boyer, 2012).”

L200 “Component C4 was found to be similar to one in Stedmon et al. (2003) and are is believed to be a combination of peaks N and T produced biologically...”

- a. Please explain “peaks N and T produced biological”. Besides the fact that the sentence doesn’t seem to make sense grammatically it is also not clear what the peaks are and why they are able to biologically produce an FDOM component. Furthermore, T is not explained.

The term “peaks” are widely used when examining fluorescence. Before utilizing PARAFAC modelling researchers would often use “peak peaking” to describe fluorescent peaks. The total fluorescent signal of a sample is a combination of many peaks, but only with PARAFAC are we able to isolate the different peaks which we then name “components”. Coble (1996) isolated some of these peaks and named them, for example peak N and T.

We have changed the sentence to:

“Component C4 is similar to component 5 found in Stedmon et al. (2003) and is believed to be a combination of fluorescent labile materials named peak N and T which are produced biologically associated with DOM degradation (Coble, 1996; Stedmon and Markager, 2005b).”

L211f “Component C4 was chosen as a proxy for groundwater recharge as the concentration of the C4 component increase with biological activity and time in the groundwater.”

- a. How do you know that the concentration of C4 increases with biological activity and time in the groundwater? Since you have neither investigated biological activity nor groundwater residence times this information has to be from some literature. Please cite.

We have changed a sentence to clarify that component C4 is produced biologically. Furthermore, the degradation is related to the time spend in the soil/groundwater were biological activity are higher due to the increased surface area for bacteria.

“Component C4 is similar to component 5 found in Stedmon et al. (2003) and is believed to be a combination of fluorescent labile materials named peak N and T which are produced biologically associated with DOM degradation (Coble, 1996; Stedmon and Markager, 2005b).”

- b. I do not fully understand the conclusion, i.e. why you can conclude that C4 is suitable as a groundwater tracer because its concentrations increase with biological activity? Increased biological activities can be found in many environments and at very small scales (e.g., also in the lake sediments).

We found that some tracers changed very much when going from the lake to the groundwater. For example, CDOM which showed a decrease in concentration when entering the groundwater properly due to pH changes in the soil. An inspection of the results revealed that a protein based fluorescent component met our criteria of being (1) non-conservative, (2) not afflicted by the lake-groundwater interface and (3) not to easily degraded or produced in high amounts, which could create false positive groundwater recharge sites. The sediments in the areas are sandy, as stated in the introduction, and have been washed by the recharging groundwater for many years. This leads us to believe that the biological activity does not differ by a large margin between the sites. We have added the following to the materials and method section regarding the choice of tracer when determining groundwater recharge:

“The groundwater recharge sites were further investigated using a range of non-conservative tracers influenced by biological degradation. We found that some of the tracer concentrations changed when going from the lake to the groundwater. For example, CDOM which showed a decrease in concentration when entering the groundwater properly due to pH changes in the soil. An inspection of the results revealed that a protein based fluorescent component met our criteria of being (1) non-conservative, (2) not afflicted by the lake-groundwater interface and (3) not to easily degraded or produced in high amounts which could create false positive groundwater recharge sites. The PARAFAC component was related to the lake concentration with a Hierarchical Euclidean cluster dendrogram and the sites which clustered together with the lake samples indicated a high groundwater recharge rate.”

L216 “Tracer concentrations of the lake water narrowed down the possible WRT of the lake.”

This has been corrected

L216ff “Equilibrium tracer concentrations of DOC, CDOM, TDP and TDN for water retention times between 0.25 and 3.5 years in 0.25 increments revealed that concentrations of TDN in the catchment are not high enough to support WRT-values over 2 years.”

Does that mean that you calculated WRTs with each of the parameters individually and then picked the one most plausible? What is the range of results you got from those calculations? Show in table or graph and discuss why the same results differ (they shouldn't differ very much, right?).

We have investigated tracer concentrations in the lake and from these we calculated the mixed inflowing concentration if no degradation took place. This mixed groundwater discharge concentration increases as WRT increases to a point where no concentration in the catchment can match the estimated mixed concentration. The ranges of the estimated concentrations are described in the result section, but we can provide the estimated concentrations as a table if the editor feels that it will contribute to the manuscript?

L220f “Groundwater discharge areas were found using the CATS model combined with applied to (?) nutrient concentrations and dissolved organic matter fractions estimated in relation to WRTs between 0.25 and 2 years.”

We have changed the sentence:

“Groundwater discharge areas were found using the CATS model using nutrient concentrations and dissolved organic matter fractions estimated in with Eq. 1. related to WRTs between 0.25 and 2 years.”

L221f “The estimated phosphorus concentrations ranged from 46 to 80 $\mu\text{g P l}^{-1}$...”

- a. Does that refer to MIC in Eq. 1?

Yes and Eq. 2 which needs to be combined to estimate the mixed inflow concentration.

The following has been added to the sentences regarding P, N, CDOM and DOC:

“The estimated phosphorus concentrations ranged from 46 to 80 $\mu\text{g P l}^{-1}$ (Eq. 1 and Eq. 2) while nitrate concentrations ranged from 1113 to 2417 $\mu\text{g N l}^{-1}$ (Eq. 1 and Eq. 3).”

and

“Thus, estimated mixed inflow concentrations of CDOM ranged from $A_{\text{CDOM}(340)} = 0.43$ to 1.04 cm^{-1} and DOC ranged from 1205 $\mu\text{mol l}^{-1}$ to 3160 $\mu\text{mol l}^{-1}$ for WRT between 0.25 and 2 years (Eq. 1 and Eq. 4). “

- b. I assume the given concentration range refers to the time increments for which you run the model. Is that correct?

Yes, that is absolutely correct

L224 Explain $A_{CDOM}(340)$ and be consistent in using the same way/units for this parameter throughout the whole manuscript.

We have checked for consistency throughout the manuscript and corrected the mistakes. We have changed this sentence in the materials and method section to explain $A_{CDOM}(340) \text{ cm}^{-1}$:

“The CDOM absorbance was measured on a spectrophotometer (UV-1800, SHIMADZU, Japan) between 240 and 750 nm in 1 nm intervals in a 1 cm quartz glass cuvette and expressed as the absorbance at 340 nm ($A_{CDOM}(340) \text{ cm}^{-1}$).”

L228f “The model identified the sites 1, 9, 11, 13 and 14 as the possible groundwater discharge sites for all WRTs (Fig. 5).” Figure 5 does not show that. It shows modelling results for the sites mentioned but this does not show the results of the other sites.

The word “possible” has been removed and the sentence has been reworked to avoid any misunderstandings. We have also added information of the cut-off point for the sites we don’t include:

“The model identified the sites 1, 9, 11, 13 and 14 as the groundwater discharge sites delivering more than 0.1 % of the water for all WRTs (Fig. 5). Changes in site distribution and fractions of discharging water were observed between the different WRTs, but in general groundwater from 3-4 sites out of xx explain the estimated concentrations in the lake (Fig. 5).”

L236ff Add table or figure with lambda values

We have added the range of the lambda values for CDOM to the text:

“Lambda values, explaining which tracers are the most important when predicting the fractions of water origination from groundwater well sites, showed that CDOM was the most important tracer when determining which sites delivered water to the lake with a mean lambda value for all WRTs of 24.2 versus 0.1-5.9 for the other tracers.”

L251 “... thus facilitating.... “ does that refer to the drainage channels? Should be “which facilitate ...” or something?

Yes, it refers to the drainage channels. The sentence has been edited:

“The exact same areas are also the ones with adjacent drainage channels (Fig. 1a), which facilitate the areas as

recharge sites.”

L254ff “Sites resembling the fluorescence found in the lake will indicate flowing water, while a difference in components between lake and groundwater sites will indicate a lower flow rate where there is sufficient time for a significant modification of the components representing the DOM pool.”

- a. What is meant by flowing water? What is meant by differing between “flowing water” and “sites with lower flow rates”? Please be more specific and refer to the scientific correct terms.
- b. Which DOM pool do you refer to hear? Be more specific to facilitate easy understanding by the reader.

The sentence has been changed and flow rate is now termed groundwater recharge rate we have also added information to help the understanding regarding where the DOM pool originates from:

“Sites resembling the fluorescence found in the lake will indicate a high groundwater recharge rate, while a difference in components between lake and groundwater sites will indicate a lower groundwater recharge rate where there is sufficient time for a significant modification of the components representing the DOM pool of the lake.”

L257ff I still do not necessarily see the link between microbial activity and the use of this component as indicator for groundwater discharge since microbial activity is not at all a process restricted to groundwater or aquifers.

If the editor thinks that this will improve the manuscript we will add a section regarding higher microbial degradation lake sediments than in lake water?

L264f “CDOM generally showed much lower absorbance at groundwater recharge sites than in the lake making it less suitable for estimating recharge areas. ”Maybe I missed it but I haven’t found this information in the results section.

This information can be found in the supplementary and can be added to the manuscript if the editor finds it suited?

L267 “While component C1 was not particularly useful for estimating groundwater recharge, it could potentially be useful to estimate discharge sites.”

- a. This sentence sounds very vague (a lot of “could” “potentially” “estimate”). Consider rephrasing.
- b. I do not follow the argumentation that C1 is a discharge indicator. Isn’t photo-resistance irrelevant in groundwater environments?
- c. Also, when the component is photo-resistant and resistant to microbial degradation which other factors lead

to a degradation which can be related to increased groundwater discharge at sites 9 and 11?

The sentence has been changed to improve the readability:

“While component C1 was not particularly useful for estimating groundwater recharge, it could be useful to estimate discharge sites. To utilize the component for discharge estimates there is a need for an assessment of the degradation rate. While it has been shown that component C1 is largely photo-resistant, as it does not absorb in the UV-A radiation areas, and is largely resistant to microbial degradation processes (Ishii and Boyer, 2012) no reliable rates for the degradation has been found. In this study, we found that only sites number 9 and number 11 hold concentrations lower than the lake (Table S1) indicating that most groundwater discharge would originate from these sites if little to no degradation takes place.”

L282 “...hinting that the lake is influenced by other water sources.” What other sources could that be since the lake is solely groundwater fed? Besides atmospheric deposition there shouldn't be any other options.

We meant other sources than the southern bank. The sentence has been edited to ease the readers understanding.

“This does not correspond to tracer concentrations found in the southern area which show very high CDOM absorbance at 340 nm ($A_{\text{CDOM}(340)} = 1.3\text{-}3.1 \text{ cm}^{-1}$) and DOC concentrations ($3114\text{-}10467 \mu\text{mol l}^{-1}$) in relation to the lake ($A_{\text{CDOM}(340)} = 0.4 \text{ cm}^{-1}/\text{DOC } 1058 \mu\text{mol l}^{-1}$) hinting that the lake is influenced by groundwater discharge from other areas as well.”

L289f “Seepage meter measurements from this area showed both discharging and recharging of groundwater (Solvang, 2016).” Which area are you referring to? From the sentence before it could be any.

The sentence has been changed to illustrate that we're talking about the eastern shore:

“Seepage meter measurements from the eastern shore showed both discharging and recharging of groundwater (Solvang, 2016).”

L291f “...indicating an influence of newly precipitated water or discharge and recharge of groundwater.” Do the results indicate discharge or recharge? Indicating both at the same time is not possible.

We have changed the sentence so it now reads:

“The same was observed for $\delta^{18}\text{O}$ samples from the eastern part of the lake, which were lower than the southern groundwater, indicating an influence of newly precipitated water or influence from the lake.”

L298 Please elaborate how your approach is able to capture the dynamics you describe when you have only a

single sampling date.

When investigating groundwater discharge we often sample during short periods (with seepage meters for example) or during longer periods using hydraulic head (where we still have to measure the depth of the groundwater and the sampling points are often sparse and far apart). We're able to trace these events as we look into the lake tracer concentrations which is influenced by events taking place the last 0-2 years and by sampling at a representative time of the year we make sure that the groundwater tracer concentrations are representative.

We have changed the sentence to clarify:

"The model is therefore able to track uncommon events such as heavy precipitation events where large amount of water is discharged to the lake during a short period. These events are often difficult to track as seepage meters needs to be deployed in this period as well as in the right place."

L296 "because soils are generally wet at this time of year (Sand-Jensen and Lindegaard, 2004)."

a. I recommend saying "saturated" instead of wet.

This has been changed

b. What does that imply? That samples are more representative? But that doesn't go along with the findings that concentrations change during the year.

Yes, this means that the samples are representative. The changes in concentration of CDOM and DOC in the lake are related to the seasonal changes in UV-radiation.

L310 Did you show inter-annual DOC concentrations in the results? If yes refer to Figure, Table. If not provide evidence for the data you refer to.

Data is not shown – we have measured DOC and CDOM seasonally in the groundwater and the lake. Either can be added if the editor wants it?

L346f "Furthermore, huge reductions would occur for TP and TN, with a decrease of 82-96 % if diverting water from the eastern shore in contrast to the southern shore with a modelled decrease of 4-18 %." "What do the 4-18% relate to?"

We have added that the 4-18 decrease is TP and TN as well.

"Furthermore, huge reductions would occur for TP and TN, with a decrease of 82-96 % if diverting water from the eastern shore in contrast to the southern shore with a modelled decrease of 4-18 % in TP and TN."

Technical comments

All technical comments have been corrected

L17 ... was estimated to **be** 2 years.

L17 ...Isolation of groundwater recharge areas **was** based on...

L18 ... sites with a high degree... **were** isolated... (I would also consider saying “identified” instead of “isolated”)

L36 I recommend substituting “changes” by “differences”

L37 ... **particularly** in small water...

L19 I recommend exchanging “Although” by “However” L25 I recommend exchanging “in relation to lake water...” by “comparing to lake water...” L30 “... groundwater contributes nutrients....” skip “with”

L44 “... which **quantify**...”

L46 Exchange “although” by “However”

L50 “...to determine the groundwater input and influence.”

L59 “the fate **IS** well known”

L60 “...which **are** either remineralized...”

L98 “...hermetically **ly** closed...”

L119 “...corrected for **the** inner filter effect.”

L122 “This allows **for the** detection of components insufficiently represented....”

L123 Reference to software (probably R) is missing

L124ff “A contour map showing the measured FDOM concentrations in groundwater **was** plotted in ArcMap

(ArcMap 10.4.1, ESRI, U.S.A) using the inverse distance weighted (IDW) function”

L144f “... the average trait value **of** all species...”

L175ff “Thus, enabling estimations of the specific removal on a monthly basis related to the concentration measured in the lake at the sampling time following Eq. (4)” This is not a complete sentence. Please add a verb.

L192 “Component C1 was similar to previously found humic-like material...”

L192ff “The component absorbs in the UV-C region which has low intensities at the ground surface (Diffey, 2002) and are is therefore expected to be photo-resistant (Ishii and Boyer, 2012).”

L197f “The component absorbs in the UV-A region and are is susceptible to both microbial and photochemical degradation...”

L208 “...with a lowest value of only 0.1 R.U....”

L209 “Components C1, C2, and C3 had...”

L210 “Concentrations of C4 were generally higher....”

L253 “While 18O worked well as a general groundwater recharge estimator, it does not indicate which sites deliver more water.” about the last part of the sentence: Consider rephrasing including “quantitative information”

After this, I have given up to comment on each of the technical errors.