General comments:
The study by Morison et al. presents some interesting data on the variability of water chemistry in shallow Subarctic ponds. The Authors make a point that most studies in high latitude lakes are based rather on few water samples taken during the summer and do not account for temporal changes driven by hydrological events. This is a strong and valid statement, and the dataset presented in this manuscript is certainly unique by showing how variable water chemistry in thermokarst ponds can be during the ice-free season. Authors attempt to relate this variability to local hydrology by calculating index of hydrological dependence that is somehow arbitrary. Perhaps it would be possible to calculate a simplified mass balance for some ions of interests; knowing the catchment area for each pond (run-off) and temporal changes in concentrations through the water column?

The manuscript lacks context data that would allow better understanding of the driving processes. For example, many shallow thermokarst ponds show strong thermal and oxygen stratification; were the 6 studied ponds stratified or mixed? What was the approximate rate of water exchange both vertically and horizontally? Water samples were take from the shore or from the middle of each pond? How representative were studied ponds in terms of nutrient concentration and biological production for other ponds and lakes in the area. Authors discuss biological processed that may be responsible for rapid uptake of nutrients but do not include any indices of the trophic status or plankton biomass in studied ponds. Indeed part of the variability in chemistry between ponds can be probably explained by different patterns of uptake. All these informations are required to better appreciate presented data and would much improve the manuscript.

Authors attempt to compare temporal and spatial variability; Is 6 (5?) ponds enough to encompass spatial variability, particularly if temporal variability is assessed with 12 time points? This would probably bias the entire analysis toward higher temporal variability.

Finally, in the Discussion authors provide an outline on how to better plan field campaigns for sampling of high latitude lakes. Such guidelines are much needed and some unification of methods would help greatly to determine effects of recent environmental changes in Arctic freshwaters. However, I think that this can be done with instructive diagram that would be readily accessible for broad scientific community rather than by quoting and refuting methods used by others. Focus on the positives and how much your own research can add to future improvements.

Please find some detailed comments and suggestions attached.

Detailed Comments:

Abstract

This part can be much improved; mainly by streamlining and including some key point data. For example first sentence is confusing and introduces biology; biogeochemistry and climate change with mentioning feedbacks. This is probably not a right place for such a broad and somewhat conclusive statement. Focus on importance on variability in pond biogeochemistry and controlling processes. Please do not use very general descriptions such as “most ions; either mention % or which species.
Introduction

Again already in the first paragraph there is a need for editing. Two sentences start with similar phrase: “these ponds”.

Also first sentence from the abstract is repeated almost word-by-word

Page 2; Recently Przytulska et al (2016, Biogeosciences) described inter-annual variability in thermokarst pond chemistry

Page 3 First paragraph; this is all true and most of the studies in the Arctic are based on the few samples taken within short time frame. However please consider also recent research showing that some shallow high latitude ponds are remarkably stable in terms of oxygen and temperature stratification during the open water period (Deshpande et al 2015).

Page 4
Water for chemical analyses was taken in profiles? Why these profiles are not shown?

Statistical techniques
I question the design of the variability analysis. If I understand correctly for the spatial deviation score Authors used mean from all ponds sampled on the same date (n=6?) and for the temporal variability authors used mean from each single pond over the entire study period (n=12); this would probably introduce some bias. Anyway it should be clearly stated how many samples have been used for this calculations.

"and stage data to represent the geometry of conical bathymetry in the non-linear pond stage-volume relationship. In this case, the parameter $\beta_2$ is invariant under linear translation, such as spatial-normalization, and is expected to maintain a negative value with the magnitude depending on the bathymetry of the pond."

This is unclear to me. Data on bathymetry were available or modelled? Even if for one pond Authors do not record the average depth? How was this accomplished?

In this paragraph Authors introduced pond-stage; this should be clearly explained as pond stage is later used in analyses and discussion. Stage means volume, depth? and how was it classified based on the dataset from pressure transducers?

Definition of the PCA is not needed, citation would suffice.

PCA can be also used to elucidate relationships between control variables (hydrology) and controlled variables (chemistry). For this you would need to use PCA on your meteorology/hydrology dataset, extract components and relate these to concentrations of different ions using stepwise forward lineal regression. Perhaps this would be a clearer method for this analysis. Alternatively citations for the method used to differentiate between hydrologically driven and not-hydrologically driven ions is needed.
Results

Any data on changes in temperature and oxygen stratification?

Does this paragraph belong in the results?

"Seasonal patterns in pond hydrology were compared with water chemistry to determine if pond water chemistry parameters were hydrologically driven. The hydrologic control of different nutrient species was categorized by the performance of the fitting of a power curve to each set of normalized concentration data against pond stage for each of the five instrumented ponds. As noted earlier, a linear fit is not possible due to the somewhat conical bathymetry of ponds, which causes greater drops in water levels as ponds dry and sediments are exposed. In each case, species with a consistent pond-specific value of $\beta_2$ (the slope of the power curve) with low root mean square error (S) were categorized as hydrologically driven, where species 10 with either a $\beta_2$ value inconsistent with other hydrologically driven species or poor performance (greater root mean square error; Figure 5, Table 5) were categorized as non-hydrologically driven. DON, Cl-, Na+, K+ and Mg2+ were classified as hydrologically driven species and Ca2+, NH4+, NO3-, and SO42- were classified as non-hydrologically driven (Table 5). Slight differences in $\beta_2$ values between ponds for hydrologically driven species are driven by the differences in bathymetry driving the stage-volume relationship.

Discussion

In general well written; needs some trimming i.e., first and second paragraph are superfluous.

Some paragraphs rewind the results e.g, detailed PCA description

Tables:

I question utility of Table 1 and Table 3; I Authors need to provide an outline of similar studies
Table 1 can be included as Supplementary Material

Results from Table 2; showing coherence between difference ponds can be outlined in the text
Table 5 need to be revised, less spacing between numbers, less vertical lines

Figure 1 needs improvement, maybe provide a high quality map go the province
Figure 3 Description of axes not symmetrical, should be “temporal variation does not exceed spatial variation”

Figure 4 Caption unclear, what is pond stage? Ice-free rather than snow-free season.

Figure 6 Normalize fonts style and size with other figures, perhaps enlarge the symbols and remove thin horizontal lines

End of review