AE: Overall, reviewers were supportive of the manuscript, but had substantial concerns regarding the statistical approach, as well as overall ecological context. Please pay particular concern to the concerns expressed by Reviewer 2 regarding the completeness of the methods and data presentation.

Reviewers' comments:

Reviewer #1: This manuscript describes the lipid status of stocked and wild lake trout captured in 3 locations and from 3 seasons in Lake Champlain. The authors note that stocked fish are placed in the lake with a very high lipid content which dramatically drops over the coming seasons to asymptote in-line with their wild counterparts. Although the authors sought to use lipid content as a proxy for lake trout condition in the three locations to help explain why they see more naturally produced fish in the central basin no differences were seen among these locations leaving the question unanswered.

General Comments

1.) The flow of logic seems to avoid talking about many other reasons why fish may be more prevalent in the center of the lake vs in other areas considered to be spawning habitats.  Perhaps the lake trout are just migrating to the deeper colder parts or areas where they can be as far from shoreline areas as possible? Maybe they are avoiding predators or competitors in other locations? The alternative reasons were stated at the end of the introduction; we have expanded this section to make it clearer.

There seems to be a large focus in the introduction on the potential effects of spatial differences in diets due to spatial differences in the prey base but no actual discussion of what those differences are. Are more YOY alewife or smelt found in the central basin…are these even what YOY lake trout diets are composed of or do they mostly eat mysids and zoop? If lake trout are eating zoop, mysis, and fish which are more likely to lead to more lipid rich lake trout? Information on juvenile diets, and the lack of data on spatial differences in abundance of prey items, has been added to the second paragraph in the Introduction. YOY lake trout eat 99% Mysis, so the question of which prey has more lipid content is not relevant at this stage

Similarly, there is a lot of focus on diets of lake trout aged <4 as creating a bottleneck for the population yet no data or research is discussed that supports this claim. Where in the lake trout life cycle is the bottleneck for Lake Champlain? Are newly hatched swim-up fry abundant but nearly no age 1 (2 or 3) fish? Do we see large mortalities overwinter as is discussed repeatedly or do they experience more continual mortality? This has been added to the introduction.

Without the above two items the paper does not convince the reader of why the study was needed.

2.) I am unclear on if the analysis was done in a way that makes sense. Shouldn't the analysis be an ANCOVA with origin (stocked vs wild), Season, and Location as factors and length as the covariate? Maybe they did do this (lines 140 - 142) but the wording doesn't fully make sense of what happened and the results section does not include any results that indicate this was the case. Also with % data I have recently been seeing Beta Regression used but, if data is too sparse for the full model, perhaps some version of conditional inference trees or random forests can get at the idea of which of these factors best predicts lake trout lipid content. Yes, we used ANCOVA and have clarified this in the text. The logit transformation and the Beta Regression are both based off a Beta distribution; we have added a citation to Warton and Hui (2011) to support our use of the logit transformation for percent data.

Line Comments:

34-37 A more complete description of the lake trout population bottle neck is desired. It would then help describe why diets of juveniles are of interest in this study. What other factors may lead to differences in where lake trout are seen and why do we think it is not those that cause the differences? Added to introduction, see above.

40-41 "…risks…is"  grammar does not seem correct Changed

44-51 size and age information seems necessary here since lipid content is so size dependent. Are the stocked fish that are found of similar size to the wild ones? Size information has been added to the next paragraph

53-54 This topic sentence does not seem necessary Omitted

59-61 where are these studies? An overwinter study in florida is very different from one in Nova Scotia. Location of the studies in not relevant, lipids are still needed to cope with stress

62 Convention is to generally avoid the term "health" since it cant readily be defined versus condition, relative weight, or lipid content. The word ‘health’ has been removed on line 55 but we have left the word in line 62 because it is clarified later in the sentence; Adams (1999) uses ‘health’ in the title of his paper.

74 What is the prey base and how does it differ spatially? What do these fish eat at the size you are looking at? Added to introduction, see above

75-78 sentences seem to contradict each other a bit, maybe the sentences are just missing a better description of the timeline they are talking about. We disagree, we feel the sentences work well together, but have changed the word ‘recently’ (for stocked fish) to ‘newly’ to avoid timeline confusion.

87 what size trout are stocked into these areas? This information has been added to the third paragraph

109-110 Here is some vague size information, I would repeat this in the intro and discussion as lipid content is dependent on length This information has been moved to the introduction

114 what categories were used for these size classes? We have added a citation to a paper that gives these size classes

135-137 This doesn't seem fully described, I think its missing that they subtracted from 1 (or 100) at some point in the calculation. This has been corrected

139-141 This makes it sound like I should be seeing results of an interaction test in the results and results related to covariates of origin and length but neither appear in the results. The results read as if two separate anovas were done 1.)Lipids = Location + Origin +(Length covariate) and 2.)Lipids = Season + Origin +(Length covariate)    Although neither paragraph mentions the length covariate nor that origin and the other factor could have interactions. We have added a citation that supports use of the logit transformation, and have noted that we tested for interactions and added a statement to the Results that no interactions were found

152-165 This section talks about the results of testing for location differences as well as stocked vs not stocked. However it is difficult to visualize the results dicussed due to Fig2 depicting the interaction of these two items. Consider including bar charts of both main effects and the interaction in the figure or reporting the numbers herein or in a table. Figures should show significant differences in some manner as well.  Was there an interaction effect that was significant, if so it isn't discussed here. We have extensively revised this section to add clarity

208-213 If I am reading correctly this paragraph is only about stocked fish and this should be stated somewhere otherwise it seems contradicted by the next paragraph. This has been corrected by removing most of the cited paragraph and combining the remaining text with the following paragraph

244-246 Is the effect size shown a biologically important one?  It seems to be a lipid content difference of 2 or 3%, is this enough to cause differences in mortalities? We have restated the sentence to address the relationship between lipid content and survival rather than construe that the 2-3% difference will change survivorship

257 what metabolic burdon is being mentioned? This is not explained.

Text has been clarified with additional information in this paragraph to explain our two competing hypotheses

Figures 2 and 3 only depict the interaction effects (season by origin or location by origin) yet the results discuss the main effects only making it difficult to follow. Main effects could be shown alongside the current bar graphs. We have substantially changed the figures to clarify the interactions among factors and covariate

Reviewer #2. I like the ideas presented in this paper and believe that the results merit publishing.  Incomplete presentation of the methods and results, especially concerning the sample designs and the stats, needs to be corrected.  I think the authors should revisit their statistical tests and provide a more complete description of what they did and how the tests were structured.  The authors had 4 independent variables that require careful attention to interaction testing and the balance of the sampling design.  These issues have been changed, see comments above. The sample values presented on the figures seem too small in relation to the collections and give the appearance of an unbalanced sample structure that will often complicate comparisons by supporting independent variable interactions.   Lipid storage is age and size dependent in fish and careful explanation of their treatment of these 2 similar but separate effects should be included.  Even if the authors find no significant differences, I think for their results to be fully understandable as they relate to their hypotheses they need to present them for wild and hatchery fish separately.  I think they should only present grand means for all samples if they believe that the measured values are homogeneous across the population.  I also believe that their data will be valuable to future studies and would like to see some useable presentation of it within or attached to this paper. Presentation of the data has been substantially changed, with new figures that address these concerns.

Abstract

Line 13:  What is "relative abundance" relative to.  Suggest changing "The distribution of" to "The abundance of" and delete "in relative abundance". We made this change, except we retained the term ‘relative’, as absolute abundance was not measured and relative abundance is standard fisheries metric.

Lines 19-20:  Rather than defining spatial and temporal just delete those terms and the parentheses around "lake regions" and "seasonal". Changed

Line 22:  It would be helpful to give the month that pre-winter refers to. Done

Lines 22-23:  This sentence would be simple if it were written in chronological order with winter depletions followed by summer increase.  What is a cyclical summer increase - suggest deleting cyclical as the authors are already referring to times of year. These changes have been incorporated into the sentence

Line 24:  The high lipid content of hatchery fish?  The authors previously state that wild fish were higher.  Since "hatchery" isn't introduced as an origin grouping until later in the paper it's difficult to understand what the authors are presenting here especially since it seems to conflict with their previous statement about wild be higher in lipids than stocked. Sentence has been changed to add the necessary clarity

Line 25: Not sure what feeding alteration is being suggested here?  In other evaluations time of year when stocking took place was very important with fall and winter stockings being associated with greater post-stocking mortality.  Again a lack of the context here (that appears later in the paper) affects the clarity. We have removed this sentence from the abstract.

Introduction

Line 31: Change was to were. Species (lake trout) is singular, ‘was’ is correct

Line 35: Suggest changing "did not begin" to "was not observed". Changed

Line 40:  Over winter thermal stress is likely not a problem for lake trout; and for some prey species that move deeper in the winter their vulnerability to lake trout predation may increase. Changed sentence to state winter can be a period of high mortality – this is the conclusion reached by Hjort and generally understood in recruitment dynamics

Lines 44:  Need more information concerning what juvenile surveys? More text has been added

Lines 53-54:  Lipid storage would be a fish physiological response/activity, the metric would be lipid concentration. Done

Line 65:  Suggest deleting "how fish respond to winter depletion to energy reserves".  How does total lipid concentration tell you how fish respond, it's a measure of the presumed response. Done

Line 74:  The prey base also has to be attainable (not thermally separated) to lake trout. Summertime is the most productive season, it should be self-evident that we are referring to lake trout prey, i.e., coldwater species

Line 77-78:  "However, post-release stress could contribute" seems like discussion and I'd suggest moving the idea there (or earlier in the introduction) unless it bares directly on the statement of objectives. We disagree, this is part of the rationale/justification for the study

Methods

Line 85: Lake Champlain is along the borders of New York, Vermont and Quebec with waters in each jurisdiction rather than between them. This is incorrect, the border runs along the center of the lake

Lines 113-115:  Ageing by finclip and size class overlap was presented here and later in the discussion the authors indicate fish compared were of the same age class (line 191) but they don't indicate how size class overlap was measured, how ageing differed between wild and stocked fish and whether age entered the analyses. The authors do use TL as a scaling factor, but do not indicate how they did that and don't elaborate in the results or discussion on whether it mattered. These issues have been addressed; see response to comments from reviewer 1

Line 113: Suggest inserting "for" before total length and removing the parentheses. Done

Line 132:  How did the authors determine if samples were dry after the 3-day drying period?  Was the amount of water left inconsequential in comparison to the amount of lipid removed? These are standard methods; all samples were treated the same way and were all less than or equal to 1 g. We did preliminary tests to ensure drying time was adequate

Lines 135-136:  How did the authors come up with lipid percentage by dividing total sample weight by the weight of the extracted sample - that would give a value greater than 100%.  Percent lipids would be calculated by weight of the lipids extracted divided by initial sample weight.  So in order for the authors to get there they would need to subtract sample weight post lipid removal from sample weight prior to lipid removal and divide the difference by the pre-removal sample weight.  Not sure what the authors are referring to in their description and they should clarify what they mean by converting their quotient to a percent. This was a mistake and has been corrected

Line 142:  Is something missing from the parentheses? This is standard format for an R function.

Line 140:  The methods are a bit confusing and seem incomplete.  What were the interactive tests?  Were those tests for interactions among the independent variables - if so what interactions were tested? The author's description here indicates they may have run ANCOVAs with length as the covariate - true?  Why did they use a logit transformation?  It is hard for me to follow the authors test structure here and whether is was appropriate.  There is enough information published that indicate lipids are size and age related so they should present how specifically TL was used as a scaling factor.  Also the authors talk about aging the fish in the methods and comparing same aged fish in the discussion so how was age used in the analyses? I think I would have approached this analysis by running ANCOVAs first with origin, season and site as the independent factors and TL as the covariate and tested for all interactions using type III SS.  If interactions were present, I would have looked at the sample design to see if unbalanced samples were driving them or the factors themselves and then applied the appropriate tests for simple effects.  Also where possible I'd examine whether the slope of the TL vs MPTLC relation remained constant across factors. We have substantially changed the presentation of the results; the new text and figures address these issues (see replies to reviewer 1).

Results

What was the variance for extracted fat values among samples tested from each fish?  Was it generally low (e.g., <5%)?  It would be very helpful to include this info somewhere so the reader can judge how much measurement variation contributed to ecological variation. We have added coeeficient of variation in the methods section to address this issue.

Line 147:  A sample design table incorporating the distribution of samples across all factor levels (site, season, origin, TL (size range)) tested would help the reader understand how balanced the sample population was. We have added a table with this information.

Line 148:  What time periods, sites and fish sizes were the means representative of?  Was this just total sample means per fish origin?  Does the interaction results and size relationships indicate that these values are appropriate? Since size is an important factor in lipid accumulation, the size corrected lipid values would make more sense in these comparisons.  What were the size correction stats and how important was size to lipid values?  Also since the stocked fish have a unnaturally high level of lipids leaving the hatchery the authors should state here what samples (seasons in particular) were included to produce the means.  Why were the N values so low in the figures if 30 fish were used per collection? These questions are all addressed in the revised results section; the N of 30 fish was a target, not the final number sampled.

Lines 152-163:  The authors indicate in the methods that they did "interactive tests", were there any significant interactions among the independent variable effects?  Did the authors account for those in their comparisons (i.e., main vs. simple effects)?  Were comparisons corrected for fish size and/or age and how?  Were there significant size effects on lipid concentrations and were the slopes of those size effects consistent among comparisons?  Was age used as an independent variable? These questions are all addressed in the revised results section

Line 152:  Were origins tested for separately?  Without knowing if interactions were significant, whether the authors were using a main effects model incorporating all variable effects including length or some other model was used looking at simple effects, what test was used (the methods seem to point both at ANOVA and ANCOVA), and what sums of squares was looked at, it is difficult to follow the results. Initially we tested origins separately, but based on the suggestion from reviewer 1, we have now included origin as a factor in a three-way ANCOVA

Lines 154-155:  Where the authors state the wild fish showed more lipid content than stocked fish at the central and southern sites, were they referring to tests constrained within origins and seasons, or were those tests for values constrained within origins and grouped across seasons?  Figure 2 doesn't seem to agree with the results text, it looks like only the north (and maybe the central) had a significant difference between stocked and wild?  Also the caption on figure 2 does not give a complete description of what is being presented - are samples grouped across factor levels?  The sample n's seem too low for grouping 30 fish each across 2 or 3 collection dates. We have changed the analysis and the figure, and included sample sizes for each group in the new table to clarify group size.

Line 158:  Did the authors group wild with stocked fish for this result or did they group stocked and hatchery fish?  This paragraph doesn't provide useful information that bares directly on the hypotheses, did wild and stocked have different values within seasons and different trends across seasons?  Were comparisons size corrected?  Did independent variable effects interact?  What was the structure of the pairwise comparisons being referred to?  Stocked vs. Wild grouped across site and TL? As above, the new analysis addresses these issues; size is now a factor in the analyses, interactions have been tested and reported, etc. These comments have been very useful for helping us to improve the analyses and description of the methods, and we trust our revisions are clear.

General:  Why don't the authors look or present a closer look at the results for the wild and stocked fish separately? I suggest first looking within sites, seasons and origins to see if there is a significant length effect.  If so run ANCOVAs separately for stocked and wild origins incorporating length as a covariate and both class variables (site and season) and all interactions including with TL in the interaction term.  Examine seasonal and site trends, grouping where interaction tests indicate appropriate, still within origins.  If you can't get to a point where you can test between origins over levels of site and season due to interactions try to constrain origin tests within site and season.  Also examine the seasonal trend within origins especially if it is possible to group across sites.  Based on comments from reviewer 1 we have elected to run a single 3-way ANCOVA and report the results of the main effects and interactions.

If the fish are different ages the TL/lipid relation may be complicated as the stocked fish within their first year of life will lose lipids from their artificially high hatchery levels as they figure out how to find their own food (like the authors indicate).  After that first summer though the survivors should start making up lipid content and the overall seasonal cycle should be more natural and correspond more to environmental conditions.  Did the authors parse out those stocked fish in their first year at large, or group those with stocked fish that survived into their second year? This point has been clarified earlier in the manuscript by explaining the size-at-age difference between hatchery and wild fish

Discussion

Line 166:  Suggest changing "was" to "were" for smoother readability.  Did not change. Suggestion would result in incorrect grammar.

There are many instances where the authors point out differences in MPTLC between wild and stocked fish so to indicate that there were no differences among sites is confusing, this makes me think that this main effect discussion is inappropriate.  For example, In the results on lines 154-155 the authors state there were differences at two sites (although the data presented in figure 2 seems to conflict with that).  We have changed our text as no site differences were found. Additionally, the authors statement on lines 169-171 that wild lake trout had higher MPTLC refutes that there were no differences among sites, there had to be a difference somewhere if in general they were different. We did have a source effect, but that does not mean we need to have a site effect. This comment is confusing to us. However, our Results section has changed quite a bit so the text should be much more clear to the reader.

Lines 171-173:  Figure 3 suggests that MPTLC was rather stable from summer-fall for stocked fish?  In the original figure 3 (now replaced), lipids showed a steady decline from fall to the following fall. Fig 3 caption needs more explanation of the comparison structure for this graph, were samples grouped across sites? We have created new figures

Lines 175-181:  Figure 2 suggests that wild fish have significantly higher MPTLC at the north site? But there was no statistically significant difference, likely due to small sample size and high variability in the data.

Line 188:  Authors indicate hatchery fish lipids were higher than wild fish of the same size, but were those fish the same age? Age is not the relevant metric, as size varies substantially by age; this has been clarified in the additional text in the introduction.

Line 190:  The authors begin this discussion talking about size and finish talking about age, confusing.  Were the fish measured at the hatchery the same age as the fish they were compared to from wild samples? This has been clarified by the addition of text that indicates the ages of each of the fish being compared. This is the only comparison where age (not size) is particularly important.

Line 191:  The authors indicated in the methods that fish were aged from fin clips and non-overlapping size classes (they should clarify the size class determinations more), but did they explain somewhere that fish were of the same age class in comparisons prior to this instance?  In the methods they said TL was a scaling factor in analyses (but didn't explain how) which seems different to the statement on line 191.  How did they know this for wild unclipped fish when length varied from <150mm up to 300mm?  Please elaborate on whether same aged (in addition to size) fish were used in the analyses across each collection (e.g., within site and season)? This has been taken care of with the new analysis.

Lines 194-196:  Please clarify this statement on foraging efficiency to indicate that you are speaking about fish within the size and age ranges tested in your study. We have edited the paper as suggested.

Line 200:  I think that density dependent growth and condition should be removed from this statement because both current hatchery procedures and the authors own results suggest that this was not a factor for the fish in this study. Agreed, text has been removed

Line 208:  Again, this seems to conflict with the first sentence of the discussion. We hypothesized that total lipid content of wild juvenile lake trout would be highest in the summer, whereas in fact it was lower than other seasons; line 208 confirms that our hypothesis was not supported (first line of discussion)

Lines 208-209:  Clarify who you are talking about, hatchery, stocked or wild fish, or some combination of them.  Discussion of results from grouped samples of wild and stocked fish is confusing. Done

Lines 224-225:  It takes reading this sentence a couple of times to get the cohort delineation clear here.  It can seem from reading this fast that the authors sample the hatchery fish the following spring in the lake, but they actually got their fish after they took their in-lake samples, dates would help readability here.  Sentence has been rewritten

Line 233:  Suggest changing "necessary" to "advantageous". Done

Lines 237-241:  This reads like introduction, suggest connecting strongly to your results (e.g., "our results indicated that…..") or deleting theses 2 sentences. Text has been changed as suggested, and we deleted the first two sentence.

Lines 250-251:  There are examples in the literature where season of stocking was compared to recruitment.  See Elrod's Lake Ontario work on lake trout and the Cornell University's work on walleye for some local cites. We are not examining factors that affect recruitment of stocked fish – as stated in the paper, stocked fish survive well after stocking. We only discuss the issue of lipid content at stocking – possibly fish could be stocked with less lipid content and maintain survival.

Lines 251-254:  Sentence structure is clunky.  Maybe split this into 2 sentences at the semicolon.  Think about doing that also for the previous sentence.  You already include a subject in the phrase after the semicolon. Done, though we have kept the semicolon in the previous sentence.

Line 253:  Need a pronoun or a noun after "but" (e.g., "it" or "abundance"). Sentence has been changed

Lines 256-259:  The authors could also call these null and alternative hypotheses. We prefer to use ‘competing hypotheses’ as it is unclear which would be considered the null.