Responses to Anonymous Referee #1

Comments

R#1: The authors tested the effect of the wastewater treatment plant (WWTP) effluent inputs on Baltic Sea planktonic communities in 4 experiments. They did so sampling seawater during winter, spring, summer and autumn and observing the effect of different WWTP addition to natural communities. They observed that nitrogen-rich DOM inputs increased BP and decreased PP. This trend will drive to an increase of carbon consumption and shift the ecosystem toward heterotrophy. Although the experiment was well performed and that the authors analysed several variables, I had the feeling that the paper was written in a hurry, sometimes with lack of precision and details. However, the study and the results obtained here are strongly recommended to be available for scientific community. Then, I recommend this manuscript to be published with minor revisions.

Comment (C): We have made extensive changes following recommendations by 3 reviewers. We believe that the manuscript has improved considerably after incorporating all changes suggested by the reviewers. We hope that you find it suitable for publication.

Detailed comments:

Introduction

R#1: 65: For non-specialists in WWTP, I think that you should explain why you are talking about the Chesapeake Bay and its discharge limit. Is it a bay that receives important WWTP effluent at the US scale? Comparable with the study area?

C: We included Chesapeake Bay discharge limits because it's a bay with serious problems of lack of oxygen, like the Baltic Sea. It's an enclosed bay, and the Baltic Sea is an enclosed sea. We tried to highlight that in sensitive areas discharge limits could be stricter.

Action (A): We have included an explanation about the inclusion of discharge limits in Chesapeake Bay. The text reads (lines 67-70): "(…) Chesapeake Bay, the largest U.S. estuary that experiences severe hypoxic conditions, discharge limits (…) Both areas, the Baltic Sea and Chesapeake Bay, are enclosed water bodies with excessive anthropogenic nutrient inputs."

R#1: 58-73: You should explain a little bit more how is the WWTP effluent in the Baltic Sea. What is its average discharge? TN, DIM, DON, DOM concentrations? Do any treatment have been implemented in order to reduce the WWTP discharges into the Baltic Sea?

A: We have now further explained WWTP effluents characteristics. The text now reads (lines 70-78): "Wastewater treatment plants contribute 10-20% of total nutrient loading in the Baltic Sea (Hautakangas et al., 2014). Estimates of total nitrogen loads to the Baltic Sea due to WWTP effluents are about 110 000 tons of nitrogen per year, and for total phosphorus loads are around 11 000 tones of phosphorus per year (Hautakangas et al., 2014). Some Baltic countries have implemented nutrient reductions in their WWTP. Denmark and Germany have reduced both nitrogen and phosphorus loadings significantly. Sweden and Finland have reduced phosphorus loads but have failed so far in reducing nitrogen loads down to 70% as recommended by HELCOM (2009) (Hautakangas et al., 2014). "

R#1: Methods 87-90: Why did you choose to sample during the four seasons? Does the effluent discharge more during winter due to rainfall and enhance the WWTP inputs on Baltic Sea?

C: Environmental conditions differ between seasons (i.e. nutrient or DOC concentration). The amount of TN and DON differs between seasons, as it can be seen in table 2. Also, planktonic community differs between seasons too, and it can influence its responses to WWTP inputs. We also sampled during the four seasons to be able to acquire the full year variability.

A: We have explained in the text why we sampled during the 4 seasons. The text reads (lines 111-112): "(…) to be able to measure seasonal variation in both planktonic communities and effluent characteristics under different environmental conditions."

R#1: 93-96: Did you check your WWTP after being filtered and frozen if some organism remains into your sample (as some bacteria may be smaller than 0.2 μm) using for example flow cytometry?

C: We measured bacterial production in the WWTP water source for the experiment conducted in spring to test if some bacteria remained in the WWTP water. BP values were very low, lower than BP

in autoclaved milli-Q water for the same day (DMP 125.25 for WWTP and 200.53 for autoclaved milli-Q).

R#1: 102-103: In my opinion, the third treatment is unclear. 103-105: I am not sure that I understood how did you do CD. You said that CD consisting of seawater diluted with milli-Q to have the same portion of community that 1:10, 1:15 and IN. Does that mean that you made the same "dilution" than the WWTP treatment but instead of using WWTP, you used milliQ and you made a CD of 1:10 (100mL of milliq for 1000ml of seawater), another of 1:5 (200ml of milliq for 1000ml of seawater) and another as IN (?)? If this is the way you made, I don't understand how do you have just one CD. . .

C: We agree with the reviewer that the explanation of how the treatments were made was unclear in the previous version of the manuscript.

A: We have better explained the preparation of the different treatments. The text now reads (lines 116-119): "Those 3 treatments (1:10, 1:5 and IN) were performed to contain the same portion of community, so the 1:10 and the IN treatments were diluted with autoclaved milli-Q and salt solution to obtain the same community portion than the 1:5 treatment."

R#1: 122-137: In this paragraph, you didn't explain how did you do dark and light incubation, how many replicates did you have. You just said "bottles were incubated at the in situ temperature**. . .** during one week". Confinement methods prone to error as you might exclude zooplankton, enhance trophic interactions within the bottle and so on. Most of the incubations realized to estimate changes in DO in order to determine NCP and CR lasts between 12-48h. One week of incubation is a lot. Do you think that the community inside your bottles after 3-4 days of incubation is representative of in situ community that may receive once an amount of DON or DIN 3-4 days earlier? You did not talk in this paragraph about the confinement effect and the effect of 7-days incubation. I think that it is really important and that you should take time to write about it and give some criticism of your estimates.

C: Bottles were incubated under light/dark conditions in 2.3 L duplicate transparent bottles. We performed a pilot experiment lasting 20 days to decide how long the incubations should last. One week is a commonly used time for experimental incubations in micro/mescosms. In addition, water temperature in the Baltic is low, and a long period of time was needed to see changes in the metabolic rates and species composition due to relatively slow growth rates. Our microbial communities showed no growth of opportunistic OTUs, showing that the community present in our experiments was representative of in situ community.

A: We have included a better explanation in the number of replicates. The text now reads (lines 127-133): "Water from the respective treatments was siphoned carefully to avoid bubble formation into four 2.3 L glass bottles per treatment sealed with gas tight stoppers. Bottles were incubated at the in situ temperature (Tables 1 and S1) in a temperature-controlled chamber during one week. Oxygen was measured every minute in 2 of the 4 replicate bottles using optical oxygen sensors (optodes) and a 10-channel fiber optic oxygen transmitter (oxy-10, PreSens®). The remaining 2 bottles per treatment were used to sample nutrient and chlorophyll a concentrations." We have also added a paragraph to address the potential effects of incubation. The paragraph is as follows (Line 549-562): "The so-called "bottle-effect", in which confinement of water causes shifts in bacterioplankton community composition and physiological rates, is a factor to consider in interpreting results from experiments with natural microbial assemblages (Fuchs et al., 2000; Massana et al., 2001; Baltar et al., 2012). Such effects are typically detected by rapidly increasing proportions of fast-growing gammaproteobacterial populations and rate measurements across all treatments (including controls) (Pinhassi and Berman, 2003; Sjöstedt et al., 2012; Dinasquet et al., 2013). In our current experiments, microbial community composition remained relatively similar to in situ communities and we did not observe excessive increases in opportunistic bacterial populations in the controls. Rather, increases and decreases in relative abundance were observed among populations typical of Baltic Sea Proper, such as Rhodobacteraceae, Synechococcus and BAL58 (Lindh et al., 2015). Thus, although confinement per se surely had effects on microbial diversity and rates, our results indicate that such effects were minor relative to the actual treatment effects."

R#1: 130-133: You explained that incubations were illuminated by artificial light with a mean PAR of 1373.2 uW/cm2. Why did you choose this amount of light? Is it representative of the daily PAR that Baltic community receives at 2m depth? Does the illumination was constant during light exposition or does it increase until reaching a maximum natural irradiance and then decrease? The light hours range

for the summer experiment is about the double of light for winter experiment. Do you think that GPP is comparable in summer and in winter experiment? Would it not be better to express GPP per hour and not per day?.

C: Irradiation intensity was a constant 1373 uW/cm2 during the daylight hours of the experiment, as an artificial ramping up and down of light was not possible in our incubation room. Daylight hours were determined to match the daylight hours of the season in Sweden, since the goal was to understand the system during the seasonal cycle. Natural irradiation at high latitudes, such as in Sweden, has large variations in both intensity and duration (i.e. daylight hours). The irradiation rate here is equivalent to that received at 2.5m depth in the winter and 7m depth in the summer, at Kalmar, Sweden, the location at which the samples were collected.
As summer has much longer daylight hours, it has higher GPP in spring and summer than during winter and autumn. GPP is usually reported per day and not per hour. We believe that reporting it per day helps to see the seasonal differences of this metabolic rate.
Incident irradiation values in Kalmar, Sweden were obtained from the Strång model, from the Swedish Meterological and Hydrological Institute (SMHI).
http://strang.smhi.se/extraction/index.php?data=tmsrs&lev=2

A: We have added the depth information to the manuscript, and deleted the word "mean" describing the PAR dose, as it may have been misleading to the reader, since the dose was constant (lines 137-139):"This irradiation dose corresponds to the irradiation received at a depth of 2.5 m in the winter and 7 m in the summer, at Kalmar, Sweden."

R#1: 134-137: I understood that you estimated NCP this way: DOmin1-DOmin0, DOmin2-DOmin1, DOmin3-DOmin2, . . . DOmin1440-DOmin1439. Then, you had NCPmin1, NCPmin2, NCPmin3,. . ., NCPmin1440 and you sum it to have it per day. Is it right? Or did you make it directly DOmin1440-DOmin0=NCP24h? According to the calculation that you made, did you compare it with the other way? Is it similar?

C: We estimated NCP as dDO/dt. Just using DOmin1440-DOmin0 would lead to loss detailed information that we had, as we measure dissolved oxygen every minute.

A: We have added text to better explain how metabolic rates were calculated. The text now reads (lines 140-144): "NCP was estimated as the changes in dissolved oxygen content during 24 hours intervals (dDO/dt). CR was calculated from the rate of change in DO during the night from half an hour after lights went of to half an hour before light went on. CR was assumed to be the same during light and dark. NCP in darkness equals CR during night. GPP was estimated as the sum of NCP and CR (GPP = NCP + CR)."

R#1: 141: Can you explain what is the killed control please?

A: We now explain how the killed control was made. The text now reads (lines 160): "1 killed control with 5% trichloroacetic acid (TCA)"

R#1: 140-142: It is not specified during the 60min incubation the temperature and irradiance received by the samples.

A: We now explain that BP samples were incubated in the same conditions than the rest of the samples. The text reads (lines 161-163): "(…) in the temperature-controlled room, at the same incubation temperature and light irradiance than the rest of the samples."

Results
R#1: 211-245: In the Metabolic rates results section, I missed to read more about statistics. For example, does GPP was significantly different between each treatment for each season? The same for CR and NCP.

C: Our statistical approach consists of using mixed effects models to test which

physicochemical parameters influenced metabolic rates. We cannot use ANOVA tests to test for differences between treatments for each experiment as there was huge temporal variability and the number of replicates was only 2. Due to the limited replicates we cannot use repeated measures MANOVA for those metabolic rates. We could use repeated measures MANOVA for BP as there were 3 replicates for each treatment and day.

R#1: 223: I think you wanted to write "GPP decreased with Chl a content, it increased with DOC concentration" in reference with Table 3.

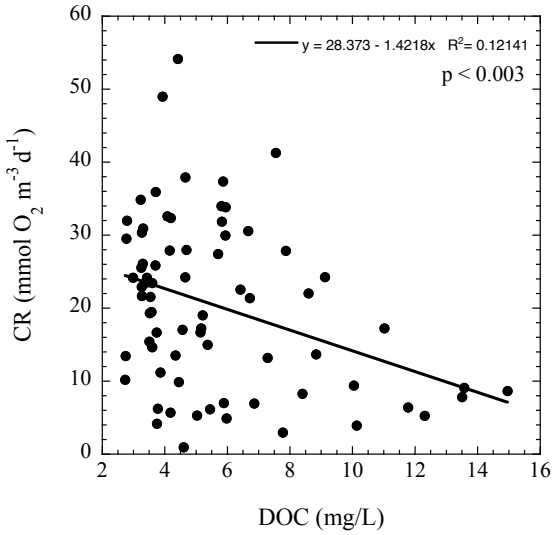
C: There was a mistake in the previous version of the manuscript, where Chl.a and DOC were changed.

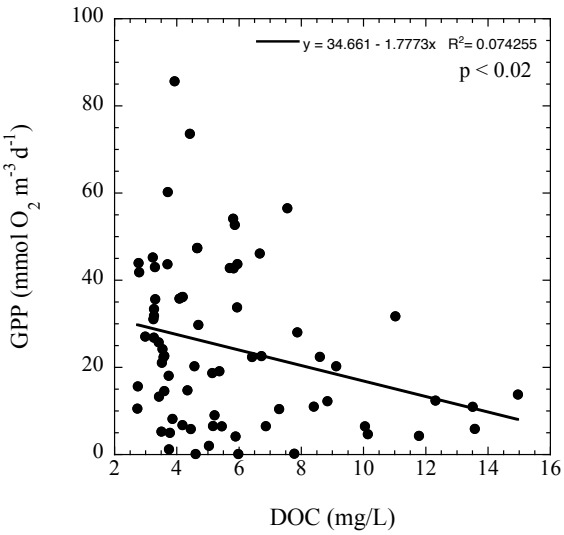
A: We have corrected the typo. We have done mixed effects again to include sampling day nested to season as random factor. Now, only DOC concentration is significant in the mixed effect model. We have removed chlorophyll a from the table.

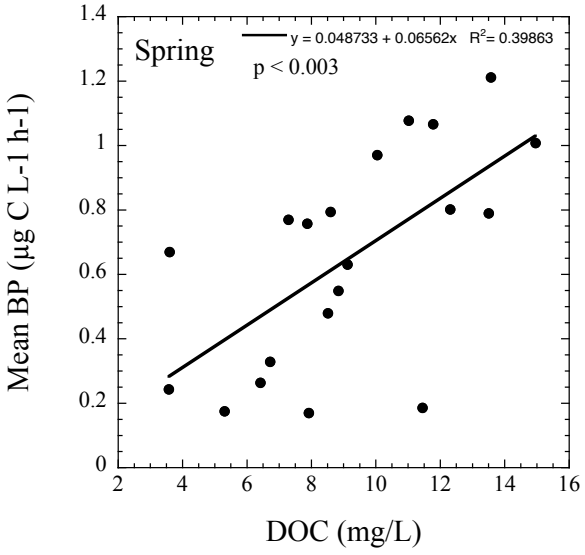
Discussion

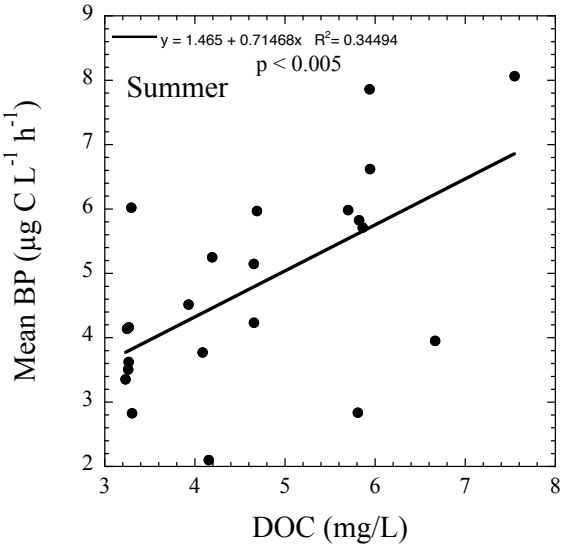
R#1: 345-348: I don't agree with this part of your discussion. I really don't see from which results you concluded that DOM significantly increased BP and decreased NCP, GPP and CR. Where are the statistical tests showing that GPP, NCP and CR were significantly different from control and CD? I understand that in BP results (247-260), you observed a tendency to increase with higher addition of effluent (nothing significant I guess), that you observed significant differences in BP for different sampling day, treatments, interaction between sampling day and treatment but there is no mention in this paragraph that BP was significantly higher at the end of the experiment with effluent addition. . . Furthermore you observed no significant differences between treatments for spring and winter experiments. I don't see in this results section any mention of "BP was significantly higher under effluent addition than control and CD at the end of the experiment" For GPP, NCP and CR, you even didn't show any statistical test in the results that can lead you to this conclusion. In Figs. 2, 4 and 5, we can see that the responses are different according to the season, that responses aren't linear along the experiment (response different at day 3 than at day 6) but you didn't talk about that in the result section and we missed that. Looking to these figures, I don't agree with the general conclusion that NCP, GPP and CR were suppressed by DOM addition, but if some statistical tests show me the inverse ok. . . but I need to see it! And, did you consider that the "good" response is at day 7? If it is so, why? Why not at day 3?

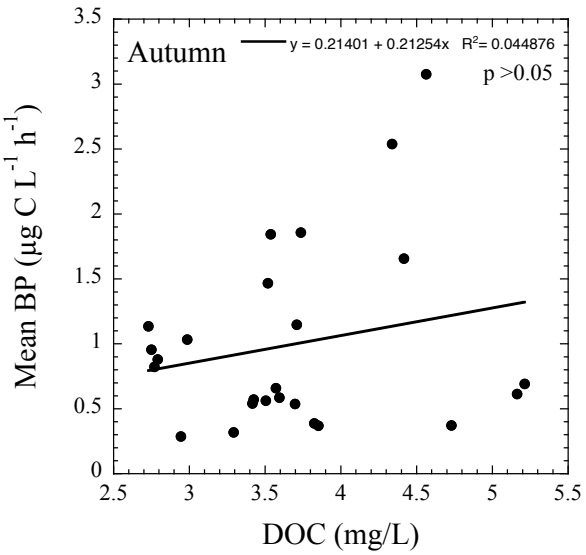
C: We conclude that DOM significantly increased BP and decreased NCP, GPP and CR from the results of the mixed effects models. We used DOC as a proxy for DOM. Mixed effects models showed that BP increased with the concentration of DOC (R^2= 0.91, $p < 0.0001$), and that NCP, CR and GPP decreased with DOC content (R^2= 0.79, $p < 0.0001$, R^2= 0.84, $p < 0.0001$ and R^2= 0.84, $p < 0.0001$, respectively). The estimate of the mixed effects models represents the slope associated to the given variable; p values were calculated comparing nested models with and without the inclusion of the response variable (i.e. ANOVA for the different models including or not the response variable). DOC concentration significantly decreased GPP and CR by itself without taking into consideration the other variables used in the models (i.e. random factor). DOC content significantly increased BP in summer, spring and winter, but not in autumn (see graphs bellow). NCP decreased with increasing DOC content, but this relationship was not significant when omitting the random factor included in the mixed effects model. Bellow you can find figures showing the relationships and the R^2 and p values for those relationships.

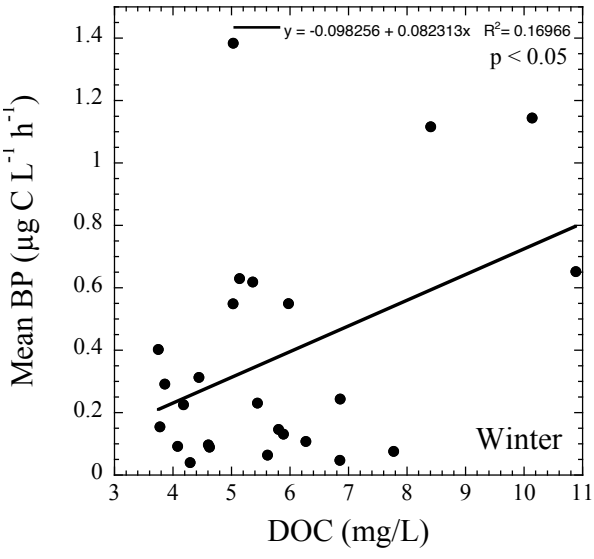

y = 28.373 - 1.4218x R²= 0.12141
p < 0.003
CR (mmol O2 m-3 d-1)
DOC (mg/L)

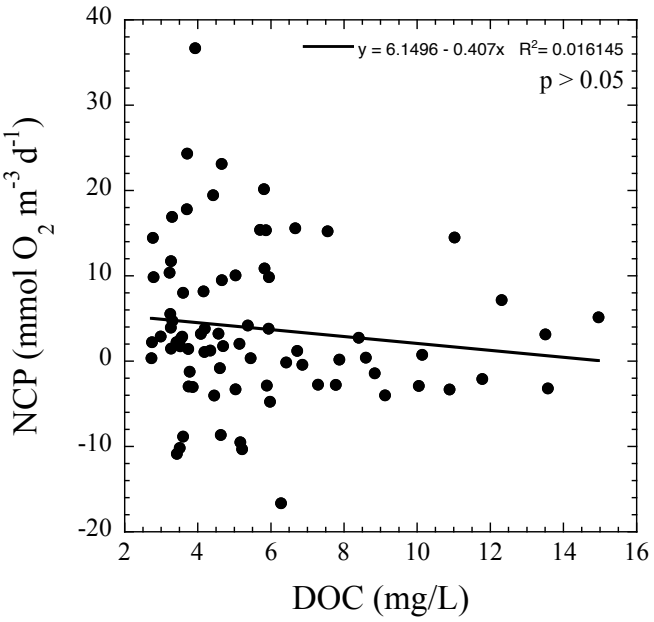

y = 34.661 - 1.7773x R²= 0.074255
p < 0.02
GPP (mmol O2 m-3 d-1)
DOC (mg/L)


Spring
y = 0.048733 + 0.06562x R²= 0.39863
p < 0.003
Mean BP (µg C L-1 h-1)
DOC (mg/L)


Summer
y = 1.465 + 0.71468x R²= 0.34494
p < 0.005
Mean BP (µg C L-1 h-1)
DOC (mg/L)


Autumn
y = 0.21401 + 0.21254x R²= 0.044876
p >0.05
Mean BP (µg C L-1 h-1)
DOC (mg/L)


y = -0.098256 + 0.082313x R²= 0.16966
p < 0.05
Winter
Mean BP (µg C L-1 h-1)
DOC (mg/L)



Figures showing the relationship between metabolic rates (CR, GPP, BP and NCP) and dissolved organic carbon (DOC) content. For BP individual graphs are plotted for each season.

R#1: 354: recent references?

A: We now also refer to work by Aranguren-Gassis et al. 2013, where they found that bacterial respiration contribution to total CR varies between 30 and 50% in mesotrophic conditions.

R#1: 359: Correct "deceased" by "decreased"

A: We have corrected the typo

R#1: 358-360: Did BGE increase and decrease significantly? R2, p? If not, you should specify it too.

C: On line 360 there were R^2 and p value of the linear mixed model. The text reads (current lines 432-434): "Estimated BGE increased with nitrate ($p < 0.003$) and DOC concentration ($p < 0.0009$) and decreased with phosphate content ($p < 0.02$, mixed effects model, $R^2 = 0.79$)."

R#1: 360-363: The example that you presented from the Bothnian Bay didn't seem really significant. . . Can you find other references showing significant increase of BGE with nutrient addition?

A: We have now added a reference showing an increase in BGE with DOM and nutrient additions in 3 Baltic Sea estuaries. The text reads (lines 437-438): "Other studies also report an increase in BGE with DOM and nutrient additions in three estuaries from the Baltic Sea (Asmala et al., 2013)"

R#1: 368-371, 381-393: Again, I disagree with the conclusion of a reduction of PP with effluent addition and that planktonic community in this region will shift toward heterotrophy. Either you have to improve your results showing statistical tests that can insure your conclusion that in general planktonic metabolism decreased under effluent addition or remove it.

C: Mixed effects models support a decrease of GPP and NCP with DOC content. The R^2 of the models are quite high (ranging from 0.79 to 0.91), explaining up to 91% of the variability of the given metabolic rate.

R#1: 432-434: There are more references about it and you should add few of them. . . not only yours.

C: We agree with the reviewer that we should include more references here.

A: We have included reference to Brown et al., 2004, Harris et al., 2006 and Yvon-Durocher et al. 2010.

R#1: 436: Use the same bibliography style than previously (Vaquer-Sunyer et al. 2015)

A: We have done so

Conclusion
R#1: 441: "DOM from WWTP effluent is nitrogen-rich." Remove this sentence.

A: We have removed the sentence.

R#1: 442-461: Same thing than previously, I disagree with your conclusion.

C: See above previous explanations on mixed effects models.

R#1: Table 1 Nutrients formulas should be written correctly with subscript and superscript.
Add C/N ratio to know if sampled communities were N limited or not.
Table 2 Idem (nutrient formulas)
Table 3 R2 not R2

A: We have done so

R#1: Figs, 1, 2, 4, 5, 6 For each figure, it could be better to have the four season plots with the same axis range in order to compare the seasonal variation of each variable.

C: As there are huge differences in the variables measured between seasons, plotting the figures with the same axis range will lead to loss of information, as it would be very difficult to visualize. As example, we have plotted 2 figures using the same axis scale, and as it can be seen in the figures enclosed bellow, most of the information is lost, especially for the figure 1 where chlorophyll a is plotted against experimental incubation day.

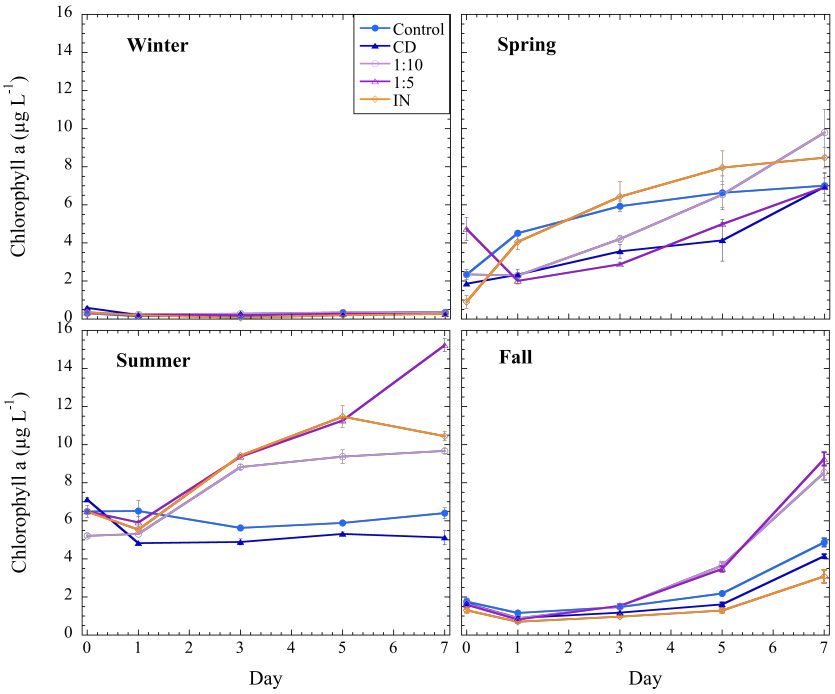


Figure 1 re-plotted using the same scale in all graphs.

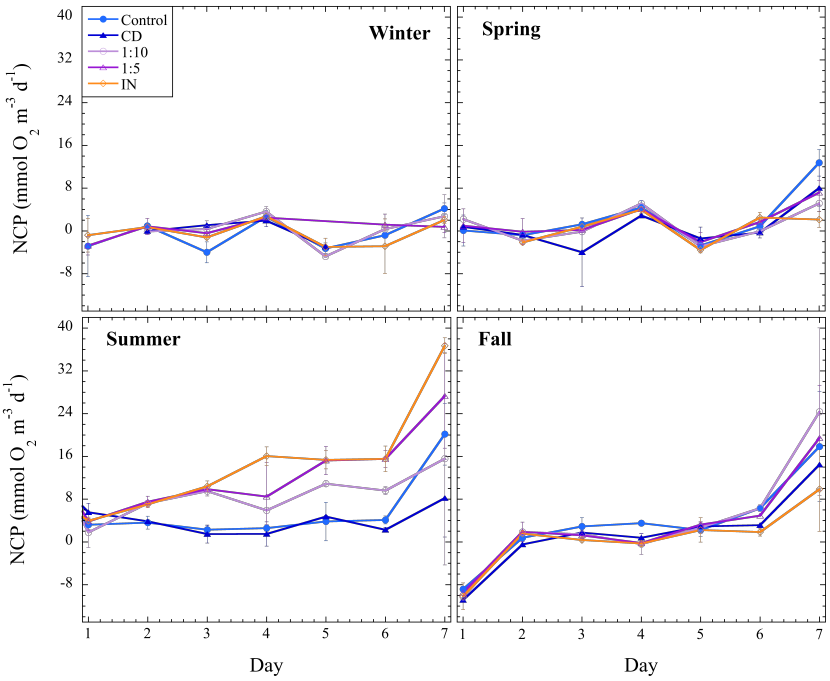


Figure 5 re-plotted using the same scale in all graphs.