Answer to review RC2 and RC3 for the manuscript "Sea Ice Deformation in a Coupled Ocean-Sea Ice Model and in Satellite Remote Sensing Data" by G. Spreen, R. Kwok, D. Menemenlis, and A.T. Nguyen

Dear Anonymous Referee #1,

Thank you for raising some concerns about our manuscript. It is our understanding that some of your concerns already could be reduced by the clarifications we made in SC1. These clarifications will be included in the revised version of the manuscript.

Find below your comments in blue and our answers to them in black. We will address as many of them as possible. We do not agree with all your comments and some of your them are in contradiction to recommendations by reviewer #2, which, if in doubt, we will follow. We will, however, try to incorporate as much of your criticism in a revised manuscript as possible.

1   General comments

This paper is essentially split into two parts. The first part discusses how modifications of the model parameter $P*$ affect modelled sea-ice volume, export, and   production and melt. The second part analyses the modelled sea-ice deformation in differently resolved model runs, comparing it to observations using the RGPS set of observations. The two parts are poorly linked, even though the authors do point out that such a link is possible. The splitting of the paper into two parts like this is cause for concern. An immediately obvious way to link the two would be to analyse the deformation patterns of the two low $P*$ runs used in part one in part two as well. With the current set-up I would recommend spliting the paper in two and expanding on each part. As it stands I will review the two parts of the paper independently, since this makes the most sense to me.

The main part of the manuscript is the model to satellite data comparison, which you and the other reviewers seem to agree have the strongest impact. We therefore will change the order of the manuscript and start with the model to data comparison. The influence of the ice strength parametrization on the model results will follow that.
The same 18 km simulation is used in the comparison and P* sensitivity study part. The simulations with reduced P* show higher deformation rates as expected. We do not want to focus on that but rather write about the maybe not so obvious effects on the Arctic ice export.

2   Specific comments

2.1   Part one

In part one the authors consider the effects of decreasing $P*$ on modelled sea-ice volume, export, and production and melt "to motivate the importance of sea ice deformation for the Arctic sea ice mass balance". I'm not sure the second part really needs this motivation. To me, seeing if we're modelling the deformation correctly is motivation enough. The question of to what extent modelled deformation affects the sea-ice mass balance is also interesting

The performance and realism of the 18 km solution has been assessed in detail in Nguyen et al. (2011). The mass balance is well reproduced in comparison to observations.

P* is commonly used as tuning knob in current VP/EVP models not only to get the deformation but also circulation in better agreement with observations. We do a sensitivity study by varying P* within the range of published values and look at the consequences. We do not claim that P* is the only knob to improve the modeled ice deformation. Actually the opposite. Fig. 8 shows that there is a stronger contrast in ice deformation between observations and model for seasonal, i.e., thinner ice. This contrast cannot be removed by changing P* alone.

The difference in the seasonal cycle of the deformation rate is similar for the second half 2000-2009 (see figure below) as for the complete simulation shown in Fig 2b. There is only a small change in the difference of D during the simulation period (see below). We will think about either adding the deformation rate time series below to Fig. 2 or exchange the seasonal cycle fig with it. Anyway, for the short sensitivity experiment section we actually like to focus on the effects on the sea ice export and ice production/melting as this in our opinion was not looked at yet. That D is increasing when P* is reduced is not really exciting or surprising.
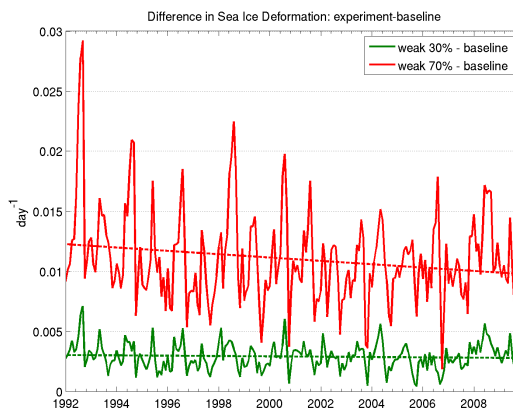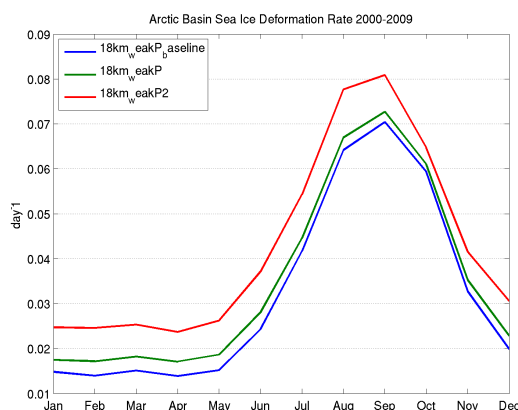
This leaves us with nothing much to judge the results of this experiment. It doesn't help that what we're looking at is essentially the model's transient response to a large change in its internal mechanics. Normally one would spin the model up to see the effect of a lower $P*$ on a model in equilibrium, but this is not done here.

We agree that it would be better to look at the solutions after some spin up phase. If we, however, would look only at the time series after a spin-up phase, e.g. after 2000, the three model solutions would already have a very different ice thickness distribution, which is part of the explanation for the found differences between the simulations. All three simulations start with a similar "initial shock". Also the baseline integration was not spun up before but all start from climatology.

The authors claim that the model has reached a new equilibrium after about 8 years, but the difference in "sea-ice production/melting" is still changing rapidly at the end of the model run (figure 3c).

We agree. We were talking about a new equilibrium in ice volume after about 8 years (see Fig. 3a. It is true that not all variables reached equilibrium after the complete simulation (as also the real world Arctic probably is not in equilibrium at the moment).

If we knew how the deformation rate changes from 1992 to 2009 and that the model does not capture that, and that tuning $P*$ correctly would give the right deformation rate, then we could say something about how simulating the wrong deformation rate gives the wrong mass balance, but the manuscript gives none of those building blocks.

We again only can refer to Nguyen et al. (2011) who show that the 18 km baseline simulation is capturing many aspects of the coupled Arctic ocean-sea ice system quite well. We therefore consider the strong deviations of the mass balance of the "weak" P experiments from the baseline degradation. Anyway, we are more interested in the sensitivity of the modeled mass balance on P* here not in finding the "best" P* value.

In terms of analysis of the low $P*$ runs the authors also miss what must be in my opinion the most obvious cause for increase in volume, and that is thickness increase due to excessive convergence. This is also pointed out by Steele et al . (1997), who performed a similar experiment. When the ice is artificially weakened (which is what we should consider is happening when using 30% of $P*$) it can be expected to ridge excessively and pile up at the north-Greenland and Canadian coasts. This effect is completely ignored by the authors, even though Steele et al. (1997) discuss it quite nicely and the authors cite that paper. In particular, the authors state that "[o]verall, the decrease in ice export E^- for both "weak ice" experiments explains most of the sea ice volume increase in the Arctic Basin shown in Section 3.1" —- a statement which seems to contradict the results of Steele et al. (1997) without giving due consideration to the piling up of ice. The pile-up of ice is, in my opinion clearly what causes the increased "sea-ice production" that the authors note in section 3.3. From the text it seems clear that the authors consider the sea-ice production(/melt) to be thermodynamic production, but there is no reason to assume that

We hope that this criticism was in large parts already resolved by our clarification in S1: "We do not consider the change in sea ice production/melt of the "weak" ice experiment in section 3.3 to be a thermodynamic process (and also do not write that in our opinion). We agree with the reviewer that this a combination of dynamic and thermo-dynamic effects. We agree with the reviewer and Steele et al. (1997) that dynamic ice thickening due to increased convergence for the weaker ice is causing the increased ice production, especially at the beginning of the experiment when the ice thicknesses is similar for all three experiments. As written in the introduction we want to add (not contradict) to the analysis of Steele et al. (1997) by also taking changes in ice export into account, which in our opinion was not done before. If the ice just would get dynamically thicker but the circulation, i.e., ice speed at the ice export gates would stay the same one would observe an increase in ice export. This is not the case. See also the discussion of possible different sea ice flow states in Hibler et al. (2006). The winter reduction in sea ice export as shown here is a positive feedback, which increases the sea ice volume for the weak experiment (in addition to initial dynamic sea ice thickening)."
We will add this discussion to the revised manuscript.

2.2   Part two

We will focus more on this part in agreement with also the other review and change the order of the sections and start with the model to data comparison.

Also this part hopefully should be clarified by our comment S1. We did not use monthly values for the analysis: "The ice deformation analysis in section 4 are not based on monthly

statistics. All analysis use the simulated RGPS dataset described in section 4.2, which has an about 3-daily time resolution. We then aggregate all deformations over one month (e.g. in Figs. 4-6) to not show a single day or show a noisy time series (e.g. Fig 8) (could be changed to other time ranges if important but would not change the results).”
We will make that more clear in a revised version by adding “monthly averages based on 3-daily deformation rates” were appropriate.

They should instead use the quantitative statistical tools and metrics previous authors have used. This would have made for much more solid conclusions and results that are quantitatively comparable to observations (e.g. Marsan et al., 2004 or Stern and Lindsey, 2009) and the model analysis mentioned above.

Our PDF analysis is similar to the one in Girard et al.. For the spatial power law scaling between different model solutions we are referring to a suggestion from Stern and Lindsay (2009). For the revised manuscript we will consider to add one power law scaling analysis as both you and the second review found this part particular important.

I want to stress, in particular that using monthly averages when studying deformation is inappropriate, since nearly all of the deformation happens at a much shorter time scale. This is a major problem with section 4.3.1. If the authors want to consider long-term differences in deformation then figure 7 is a more appropriate approach than figures 4, 5, and 6. I would even recommend taking a multi-month or seasonal average instead of only one month, in that case.

Clarified now. We do not use monthly means.

It is interesting how large the difference in deformation rate is between the seasonal and multi-year ice is.
I'm also left wondering if the deformation rates used in section 4.5 are monthly averages or not. Using monthly averages there would be inappropriate for the same reason as before, although it is not immediately clear how large an error we get using monthly averages in this case. Should the results in section 4.5 hold then they are a very interesting contradiction of the results of Girard et al. (2009).

Also in section 4.5 we did not use monthly averages but the original about 3-day RGPS time periods.

It does seem strange though, that the authors choose not to remove the noise of the RGPS data as prescribed by Bouillon and Rampal (2015a). They need to either remove the noise or justify not removing it.

Our analysis actually was done before the Bouillon and Rampal paper was published. Also currently the RGPS data is still available in its current form and still used in many studies. We prefer keep doing our analysis with the original RGPS data even if this means that they are a bit noisy. Qualitatively we would not expect different results by removing the noise.

It is also inappropriate to consider the percentage of area containing 80% of the deformation as a measure of localisation (section 4.3.3). It should be the largest 15% of the deformation, like Stern and Lindsay (2009) use. Using 80% of the deformation you essentially include all the deformation so this is no longer a measure of the localisation of deformation. The way it stands the metric is essentially meaningless.

We will follow this suggestion and also calculate the percentage of area containing the highest 15% of deformation rates.

The authors also do the power law scaling of deformation rate incorrectly (section 4.4). They use different model realisations (i.e. 4.5, 9, and 18 km resolutions) to determine the scaling, but the correct thing to do is to use a coarse graining method (like the authors named above) and calculate the scaling based on it. The authors of this manuscript argue that the high resolution model gives better results than the low resolution ones, but they then combine all three to calculate the scaling. This makes no sense.

The purpose of this exercise was to study how one can compare deformation rates originating from models with different grid resolutions, which we consider a common problem. Applying a power law one can bring the deformation rates closer together. Comparisons then might be possible if large scale (model domain) and long-term (yearly) averages are compared. We, however, also clearly state that this is by far not ideal due to the strong seasonal dependence and dependence on ice concentration and thickness and that statistical comparisons might be more appropriate. We will stress that even stronger in a revised version of the manuscript.

3   Conclusions

I am sorry to say that I will be recommending that this paper be rejected publication in The Cryosphere. The reasons for this decision are the poor structure of the paper, it being split into two unrelated parts, and the substantial shortcomings of both parts. This is quite disappointing since I believe that the comparison of the MITgcm results with RGPS data could be very interesting indeed. My recommendation to the authors is to thoroughly review Girard et al (2009) and the related literature, and then to revisit part two of the manuscript with the aim to refute or support the conclusions of Girard et al (2009) in the case of the 4.5 km resolution simulation, give an indication of the resolution dependence, and to provide contrast with the results of Bouillon and Ram- pal (2015b) and Rampal et al (2015). If this is properly done then that would make for an interesting paper and one that would be important for further evaluation and development of dynamical sea-ice models.

We hope that we could dispel and clarify some of the concerns the reviewer had. We believe that many of them were based on misunderstandings and we will work on making the manuscript clearer and easier to follow. We will restructure the manuscript as described above. As the second reviewer recommends to keep the P* sensitivity study we will not remove it as suggested here but rather move it more to the end of the manuscript. Together with the changes proposed in the answer to the second review we hope that a revised version will receive the reviewers approval.

4 References

All the papers I refer to here are already cited in the paper, with the exception of Rampal et al (2015), which is still under review at The Cryosphere Discussions: http://www. the-cryosphere-discuss.net/tc-2015-127/