Reply to Reviewer Comment 1 (RC1)

Comment:

This manuscript describes a coupled via C-Coupler2 Arctic ocean sea ice configuration of MITgcm and Polar WRF atmosphere. The model is intended for high quality Arctic sea ice seasonal predictions. There is large demand for high quality regional climate models of the Arctic basin and such activity must be strongly appreciated. In the manuscript the setup has been validated for year 2012 because of a strong storm formed off the coast of Alaska on 5 August 2012. The role of sea ice-ocean-atmosphere interaction has been addressed. Although the authors demonstrate good modelling skills, good knowledge of the Arctic Ocean system and impressive level of model validation the paper in it's resent form failed to convince me that the new climate model is ready to use and that is better than any existing global climate model. My main criticism is that (1) the performance of the model on different HPC systems, regarding scalability and the costs of all individual components is not addressed, (2) the presented configuration is not properly tuned. I encourage this paper for resubmission after these weaknesses have been fixed.

Reply:

The authors thank the reviewer for the insightful comments, and we completely agree with the questions and comments raised by the reviewer, which have helped us to improve the quality of the manuscript. We have carefully considered the reviewer's comments. Some paragraphs are rewritten and figures are redrawn. Regarding to the two main criticisms, we have added a paragraph of model scalability and the costs of all individual components in the revised manuscript. We also tuned sea ice albedo parameters in the MITgcm to get a better sea ice simulation and re-written Section 4 with new results. To further adequately address the two main criticisms, our replies are as follows:

1. Following the comments on the model scalability and the costs, we run several

experiments using different CPU numbers, each experiment runs for 7 model days, and we obtain the following results:

total_cpu_number	cpu_number_on_each_component_model	total_run_time (unit: s)	wrf_interface (unit: s)	mitgcm_interface (unit: s)	wrf_integration (unit: s)	wrf_time_alone (unit: s)	mitgcm_integration (unit: s)	mitgcm_time_alone (unit: s)
28	14	12840	4.8	12131	12835.2	/	709	/
56	28	12000	4.74	11196	11995.26	7140	804	317
112	56	10440	5.16	6477	10434.84	3960	3963	154
224	112	3780	5.26	3550	3774.74	2160	230	96
448	224	2460	5.21	2116	2454.79	1560	344	68
896	448	1380	358	48	1022	1320	1332	84

In our model configuration, the requested total CPUs are assigned equally to the component models, that is, if we request 448 CPUs, then 224 CPUs are assigned to the WRF and 224 CPUs are assigned to the MITgcm. The total_run_time of the coupled model decreases from 12840 s to 1380 s when total_cpu_number increases from 28 to 896. Limited by computational resource of our center, we can not perform experiment which uses more than 1000 CPUs.

In the above table, the "wrf_interface" expresses time of coupling process implemented by the WRF, then the time of integration process implemented by the WRF, i. e. "wrf_integration" can be calculated as : "total_run_time" minus "wrf_interface". The "mitgcm_interface" expresses time of coupling process implemented by the MITgcm, then the time of integration process implemented by the MITgcm, i. e. "mitgcm_integration" can be calculated as : "total_run_time" minus "mitgcm_interface". The "wrf_time_alone" expresses runtime of the standalone WRF which implements 7 model days integration. The "mitgcm_time_alone" expresses runtime of the standalone MITgcm which implements 7 model days integration.

We find that, when total_cpu_number is not larger than 448, the "mitgcm_integration" is substantially smaller than the "wrf_integration", meaning that the efficiency of the coupled model depends on the WRF component model. When total_cpu_number is larger than 448, the efficiency of the coupled model depends on the MITgcm component model. Additionally, both the integration efficiency of the component models in the coupled model are lower than those of the standalone model runs. The decrease in parallel efficiency results from the increase of communication time, load imbalance, and I/O (read and write) operation per CPU core (Christidis, 2015). By comparing the time cost of stand-alone WRF and MITgcm integration the parallel efficiency of the coupled model is higher than both ocean-alone or atmosphere-alone models with same numbers of grid points per CPU core.

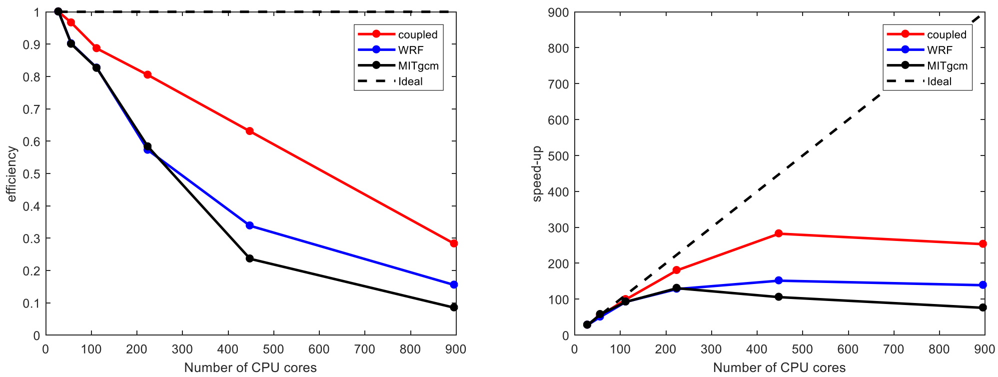


Fig.1 The parallel efficiency (left) and speed-up (right) test of the coupled model, employing up to 896 CPU cores. The simulation using 28 CPU cores is regarded as the baseline case when computing the speed-up.

2. In the MITgcm, sea ice albedo is a function of four kinds of ice/snow surface type: dry ice, wet ice, dry snow, wet snow. According the reference (Nguyen et al., 2011, Arctic ice ocean simulation with optimized model parameters: Approach and assessment, JGR-oceans), typical range of sea ice albedo in the AOMIP (Arctic Ocean Models Intercomparison Project) is: 0.6-0.75 for dry ice, 0.5-0.68 for wet ice, 0.8-0.84 for dry snow, 0.6-0.77 for wet snow. Typical sea ice albedo in the MITgcm under the JRA25 forcing is: (dry ice: 0.7, wet ice: 0.71, dry snow:0.87, wet snow: 0.81). They also found the optimized sea ice albedo parameters depend on the selected atmospheric forcing. In our previous studies (Liang et al., JGR-oceans, 2018, 2019), we use the sea ice albedo of (dry ice: 0.75, wet ice: 0.7, dry snow:0.86, wet snow: 0.8) when the JRA55 forcing is used. For this study using CFSR forcing, before manuscript submission, we tested the above albedo parameters for the coupled model and found that the model produced more sea ice than the observation. So we reduced the albedo parameters and tested the group of (dry ice: 0.65, wet ice: 0.55, dry snow:0.8, wet snow: 0.7) for

the coupled model, and we found this group of albedo parameters is appropriate for the coupled model when the CFSR forcing is used.

Besides, we add the standalone MITgcm simulation in 2012 for the comparison. To keep consistency between the coupled model and standalone MITgcm model, the standalone MITgcm simulation is forced by surface variables derived from the CFSR data and uses the same albedo group of (dry ice: 0.65, wet ice: 0.55, dry snow:0.8, wet snow: 0.7). Results of the modeled sea ice show that the two-way coupled model generates more rational sea ice distribution than the standalone MITgcm run. The modeled and observed monthly sea ice concentration are shown as follows:

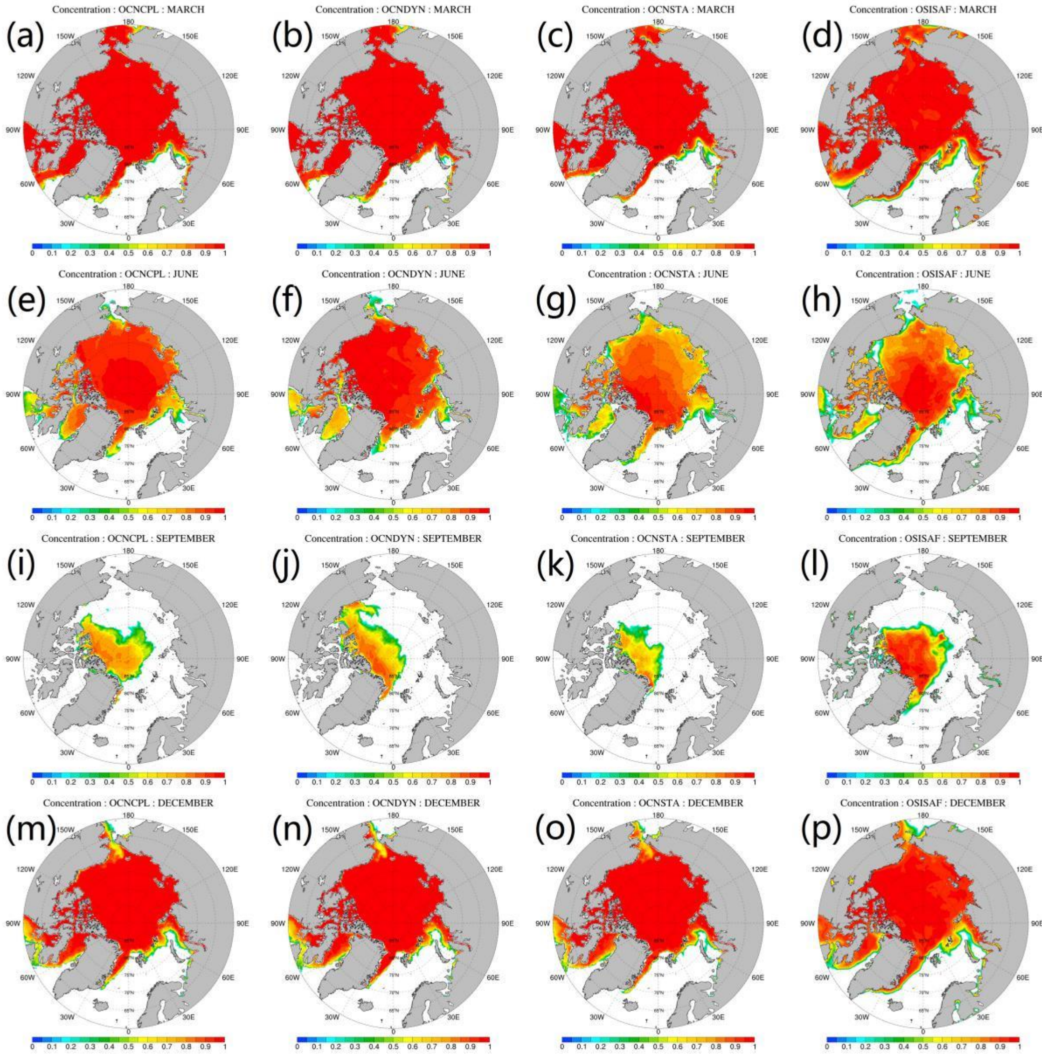


Fig.2 Monthly mean sea ice concentration in 2012. The 1st, 2nd, 3rd, 4th row denotes March, June, September, December. The 1st, 2nd, 3rd, 4th column denotes the two-way coupled run (OCNCPL), the one-way coupled run (OCNDYN), the standalone MITgcm

run (OCNSTA) and the observations (OSISAF).

Detailed replies to specific comments by the reviewer are presented below:

Comment :

In the model description, there need to be discussion on why fields and not fluxes are coupled. Do the authors guarantee the same bulk formulas are employed on the atmospheric and ocean sides? Which difference is expected if the fluxes are coupled? I guess that COSMO-CLM/NEMO group has some experience with it although not with the Arctic region. I find this aspect is more important than describing the computation of the corner geographic information for MITgcm. The latter piece I would even omit due to its simplicity. A following chapter after the model description, which gives more information about the model scalability and cost is required.

Reply:

In the original manuscript, we have cited the paper from COSMO-CLM/NEMO group (Van Pham et al. 2014). For COSMO-CLM/NEMO model, the exchanged fields from COSMO-CLM to NEMO are the flux densities of water, momentum, solar radiation, non-solar energy and sea level pressure; and from NEMO to COSMO-CLM are SST and the fraction of sea ice. Regarding to the MITgcm model configuration we used, instead of forcing the model with heat flux data, the model calculates these fluxes using the changing sea surface temperature and ice surface temperature. We need to read in some atmospheric data: 2 m air temperature, 2 m air humidity, downward shortwave radiation, downward longwave radiation, precipitation, 10 m wind speed. This combination of setups have been used in our ocean-seaice model for several years and showed reasonable results in aspects of sea ice forecasts (Liang et al., JGR-oceans, 2018, 2019). Therefore when we build the coupled model, the same setups from the standalone ocean-seaice model are kept. We have checked the manual and source code from both the WRF and the MITgcm.

(1) The heat fluxes calculated in the MITgcm are shown below

(https://mitgcm.readthedocs.io/en/latest/phys_pkgs/bulk_force.html):

Sensible heat flux (Qs):

$$Qs = \rho_{air} cp_{air} u^* T^*$$

Latent heat flux (Ql):

$$Q_l = \rho_{air} u^* q^*$$

Where

$$u^* = c_u u_s$$

$$T^* = c_T \Delta T$$

$$q^* = c_q \Delta q$$

$$c_u = c_T = c_q = \frac{\kappa}{\ln(z_{ref} / z_{rou})}$$

ρ_{air} : air density, cp_{air} : specific heat at constant pressure, u_s : wind speed, κ : Von Karman constant, z_{ref} : reference height and z_{rou} : roughness length scale which could be a function of type of surface.

(2) The heat fluxes calculated in the WRF are shown below (http://www2.mmm.ucar.edu/wrf/users/docs/user_guide_V3.8/contents.html) :

Sensible heat flux (H):

$$H = \rho\, c_p u_* \theta_*$$

Latent heat flux (E):

$$E = \rho\, u_* q_*$$

$$u_* = \frac{kV_r}{\ln(\frac{z_r}{z_0}) - \psi_m}$$

$$\theta_* = \frac{k\Delta\theta}{\ln(\frac{z_r}{z_{0h}}) - \psi_h}$$

$$q_* = \frac{k\Delta q}{\ln(\frac{z_r}{z_{0q}}) - \psi_h}$$

Where subscript r is reference level (the lowest model level, or 2 m or 10 m), Δ refers to difference between surface and reference level value, z_0 are the roughness lengths, k is the von Karman constant.

(3) The above calculations of bulk formula for sensible heat flux and latent heat flux are almost same. To further prove this, we also compare the sensible and latent flux from the WRF output and the MITgcm output within the coupled model based on the result on March 1, 2012. The results show very little discrepency, which mainly because that parameter setup is slightly different from atmosphere and ocean sides. In future work, we will use fluxes as exchange variables instead of fields, ensuring the energy balance.

We agree with the reviewer that the computation of the corner geographic information for the MITgcm is more like a technical issue and with simplicity. So the part of the computation of corner information and Fig. 3 are removed from the revised manuscript. We have already added a new chapter of model scalability and cost analysis in the revised manuscript.

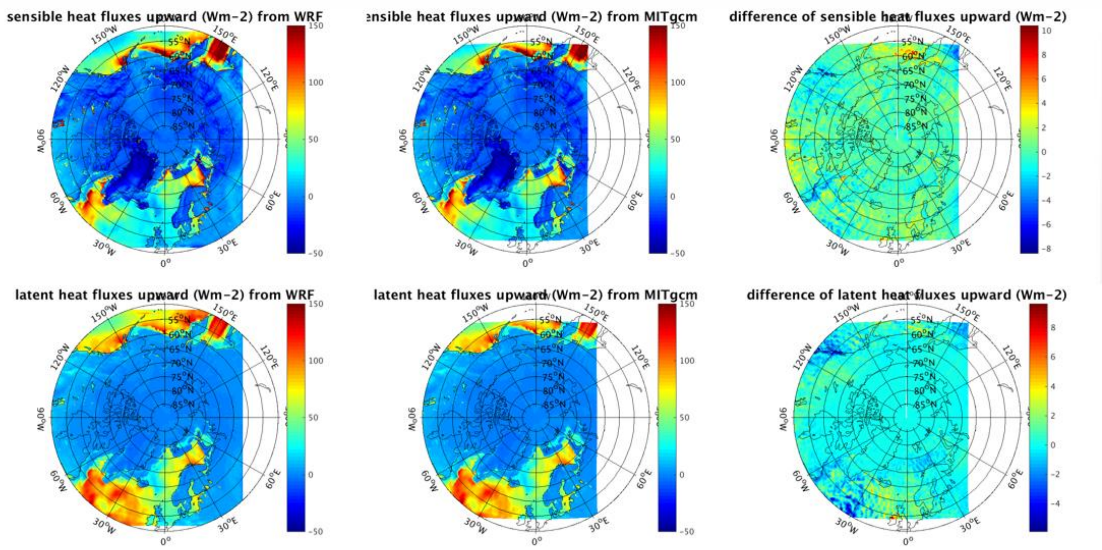


Fig.3 Sensible and latent heat fluxes derived from WRF and MITgcm output on March 1, 2012. Differences (WRF - MITgcm) of heat fluxes are calculate by interpolation to the same grid.

Comment:

I believe that C-Coupler is a good tool to use but the statement that a model produces bitwise identical results with a different coupler means that the coupler just works. Is it better in terms of performance? Which interpolation option do you use (question is more relevant for the wind stress)? As an illustration it would be good to see the curl of the wind stress on the atmospheric and oceanic meshes (instead of Fig.3).

Reply:

The innovations of C-Coupler are flexible and automatic coupling configuration and 3-D coupling capability, which is easier for users to build coupled models. For the interpolation (remapping) from a source horizontal grid to a target horizontal grid, users can use the remapping weights that are either automatically generated by C-Coupler2 in parallel, or read from an existing remapping weight file produced by an external software tool such as SCRIP, ESMF, YAC, CoR, etc. Remapping configuration files enable to flexibly and conveniently specify how to remap coupling fields between grids. For our model, we used the default remapping configuration, the bilinear remapping algorithm is used for remapping the “state” fields between the horizontal grids, the conservative remapping algorithm is used for remapping the “flux” fields between the horizontal grids, and the linear remapping algorithm is used for remapping in both the vertical dimension and the time dimension.

The momentum variable transferred in our model is 10m-wind instead of wind stress. We have calculated the wind stress in WRF and made a comparison with MITgcm output on March 1, 2012. The results show very little discrepency, which mainly because of parameter setup is slightly different from atmosphere and ocean sides.

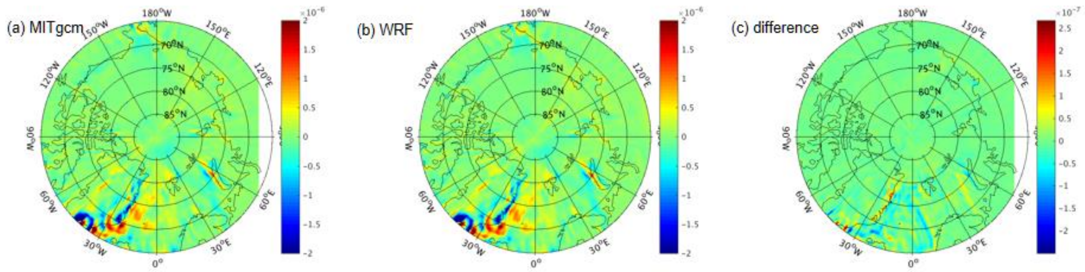


Fig.4 Wind stress curl (unit: Nm^{-2}) derived from (a) MITgcm and (b) WRF output on March 1, 2012. Difference (WRF - MITgcm) of wind stress curl is calculate by interpolation to the same grid.

Comment:

In the OCNDYN run is it just setting alpha to 1.0? I believe that one year of coupled simulation is too short to validate the model. Either ensemble of runs or a longer simulation is desired here. I don't want to force the authors to do much of additional work but the cheapest way would be to add the stand-alone MITgcm run for the comparison here. The spin up with JRA55 has been already computed.

Reply:

For the OCNDYN run , we switch off the export interface of coupling in the code to close the variable transfer from ocean to atmosphere. The alpha in section of coupling strategy is a relax coefficient weight to combine the boundary variable, to diminish the abrupt value changes from two sources. We have corrected the text and elaborate the description of the OCNDYN experiment to avoid the confusion. We agree that one year is short to validate the model. So we add the standalone MITgcm runcase for the comparison. The results are shown in the revised manuscript.

Comment:

Section 4.1, which introduces Fig. 5 illustrates that the model has not been tuned properly yet. Although the authors (line 254) claim that OCNCPL is closer to observations, which might be true, but I see that both runs failed. Here it is important to give (at least visually) the measure of the error. A stand-alone MITgcm run, hence, *would be a good choice. Provided the high skill of validation made in the following chapters I assume that the model has to be better tuned first.*

Reply:

We have already tuned the model. The appropriate albedo of (dry ice: 0.65, wet ice: 0.55, dry snow: 0.8, wet snow: 0.7) in the coupled model for the CFSR forcing is used. Compared with the one-way coupled and standalone run, the OCNCPL case shows better results. We re-write the section 4 of validation with new results. The measure of the sea ice concentration error between the model and observation are shown as follows:

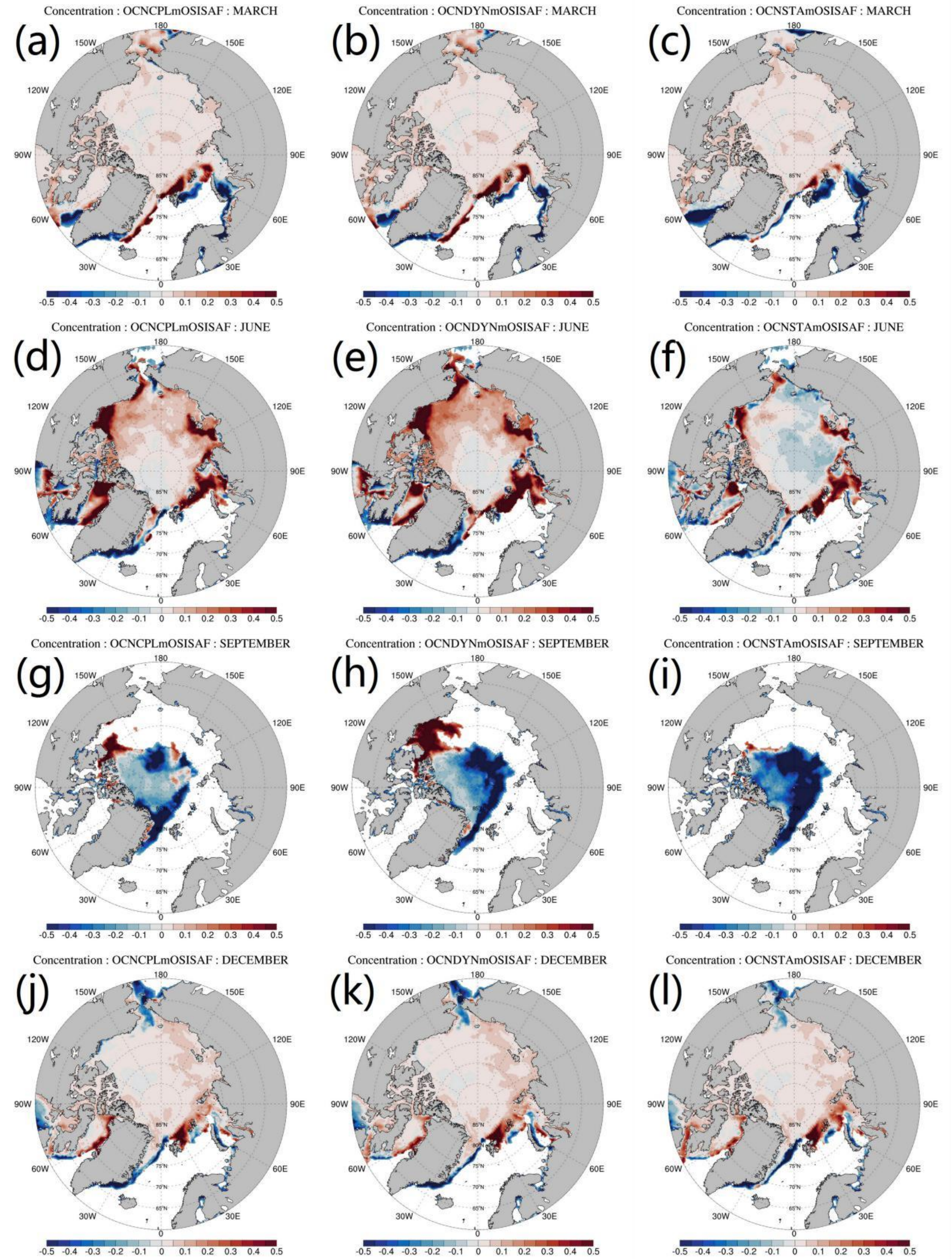


Fig.5 Sea ice concentration bias between the model and the OSISAF observation in 2012. The 1st, 2nd, 3rd, 4th row denotes March, June, September, December. The 1st, 2nd, 3rd column denotes the two-way coupled model, the one-way coupled model, the standalone MITgcm run.

Comment:

Minor things: line 181: "...without any data assimilation. . .". There was nothing said about new model is within an assimilation framework before or after. Why to mention this?

Reply:

In our institute, we have established an Arctic Ice Ocean Prediction System (ArcIOPS) based on the MITgcm and ensemble Kalman Filter data assimilation algorithm to carry out operational Arctic synoptic-scale sea ice forecast. Our future plan is to implement the coupled model with data assimilation algorithm to carry out Arctic seasonal sea ice prediction. We have re-written this paragraph to avoid confusion.

Comment:

line 182: "...the coupled model free simulations..." what do you mean?

Reply：

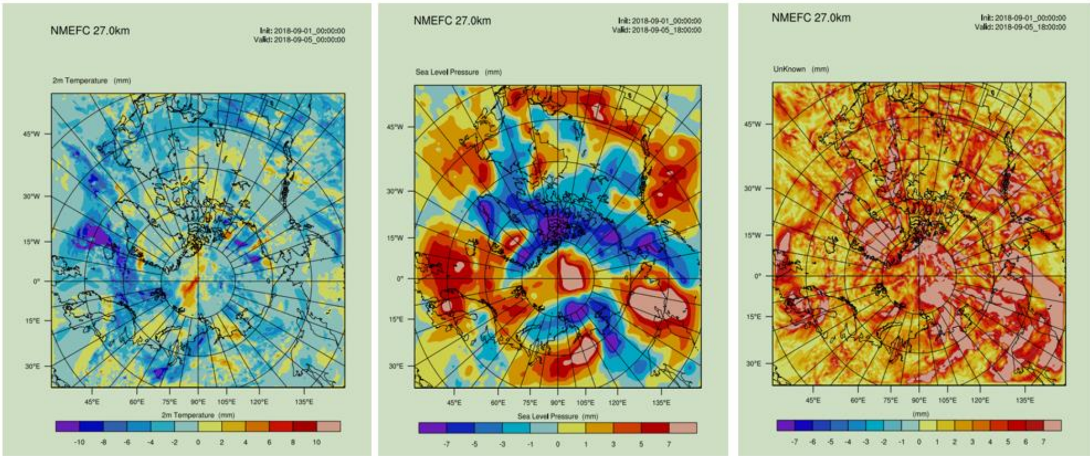
Free simulation is aiming at the experiment with data assimilation in planning. We have re-write this sentence with "… the coupled model simulations…"

Comment:

line 188: indeed, nothing about atmosphere. Patterns of SLP although would be of interest.

Reply:

We have compared the patterns of atmospheric variable (2m temperature, SLP and wind fields) in OCNCPL run and OCNDYN run.



2m temperature(°C) Sea Level Pressure(Pa) 10m wind speed (m/s)

Fig.6 Bias of 2m temperature, SLP and 10m wind speed from OCNCPL and OCNDYN on September 8, 2018.

Comment:

line 230: too speculative. did you do another run with different albedo?

Reply:

We have tried different sea ice albedo combinations and then chose a best albedo combination for the CFSR forcing. In our previous studies (Liang et al., JGR-oceans, 2018, 2019), we use the albedo combination of (dry ice: 0.75, wet ice: 0.7, dry snow:0.86, wet snow: 0.8) when JRA55 forcing is used. For this study using CFSR forcing, we tested the above albedo combination for the coupled model and found that the model produces more sea ice than the observation. So we reduced the albedo parameters and tested the combination of (dry ice: 0.65, wet ice: 0.55, dry snow:0.8, wet snow: 0.7) for the coupled model, and we found this group of albedo parameters is appropriate for the CFSR forcing.

Comment:

line 290: I would elaborate more on this if possible.

Reply:

Day et al., 2014 pointed out that sea ice behaves long-term memory of melting-freezing processes. Notz and Bitz (2017) indicated that summertime sea ice thickness has an important influence on sea ice state in the following spring through the ice thickness-ice growth feedback. A negative anomaly of sea-ice area in late summer induces larger heat losses in autumn and winter from the ocean to the atmosphere due to enhanced outgoing long-wave radiation and turbulent heat fluxes, this causes thinner snow and ice due to later freeze-up and hence larger heat-conduction fluxes through sea ice, eventually leading to larger ice-growth rates.

Comment:

line 366: not really or show that it is better than in other (global) models

Reply:

Indeed, we did not compare our results to other models. The motivation of our work is to enhance the operational sea ice seasonal prediction by coupling atmosphere, ocean and sea ice. Sea ice model intercomparison is a good index to weigh the performance of coupled models. In future, we hope that joining such international project to evaluate and improve our models.

Comment:

line 372: is it due to the albedo feedback? In OCECPL it is computed on the MITgcm side. What happens in OCEDYN? Again, figures 6c, 7 would have more value if the model is tuned. In the present form the model is not ready for this validation.

Reply:

We believe that two-way coupling between the WRF and the MITgcm provides a more rational representation of real air-ice-ocean physical processes, which includes the important ice-albedo feedback in early summer. In the MITgcm, sea ice albedo is calculated based on several variables, such as snow depth on ice, ice surface temperature. In the OCNCPL run, albedo is a coupling variable which affects both the

WRF and the MITgcm. In the OCNDYN run, albedo used in the WRF are directly read from the CFSR forcing data.

In our original manuscript, we make a mistake when calculate sea ice extent in Figure 4a and 4b, which leads to the negative sea ice extent bias in Figure 4a. We confuse sea ice area with sea ice extent. Actually the blue and red curves in Figure 4a and 4b in the original manuscript are sea ice area. We redraw the modeled sea ice extent and add the standalone MITgcm run for the comparison (see the following figure). Compared with the standalone MITgcm run, the modeled sea ice extent in the coupled runs are more closer to the observation. With respect to the one-way coupled run, the spatial distribution of summertime sea ice concentration in the two-way coupled run is more closer to the OSISAF observation.

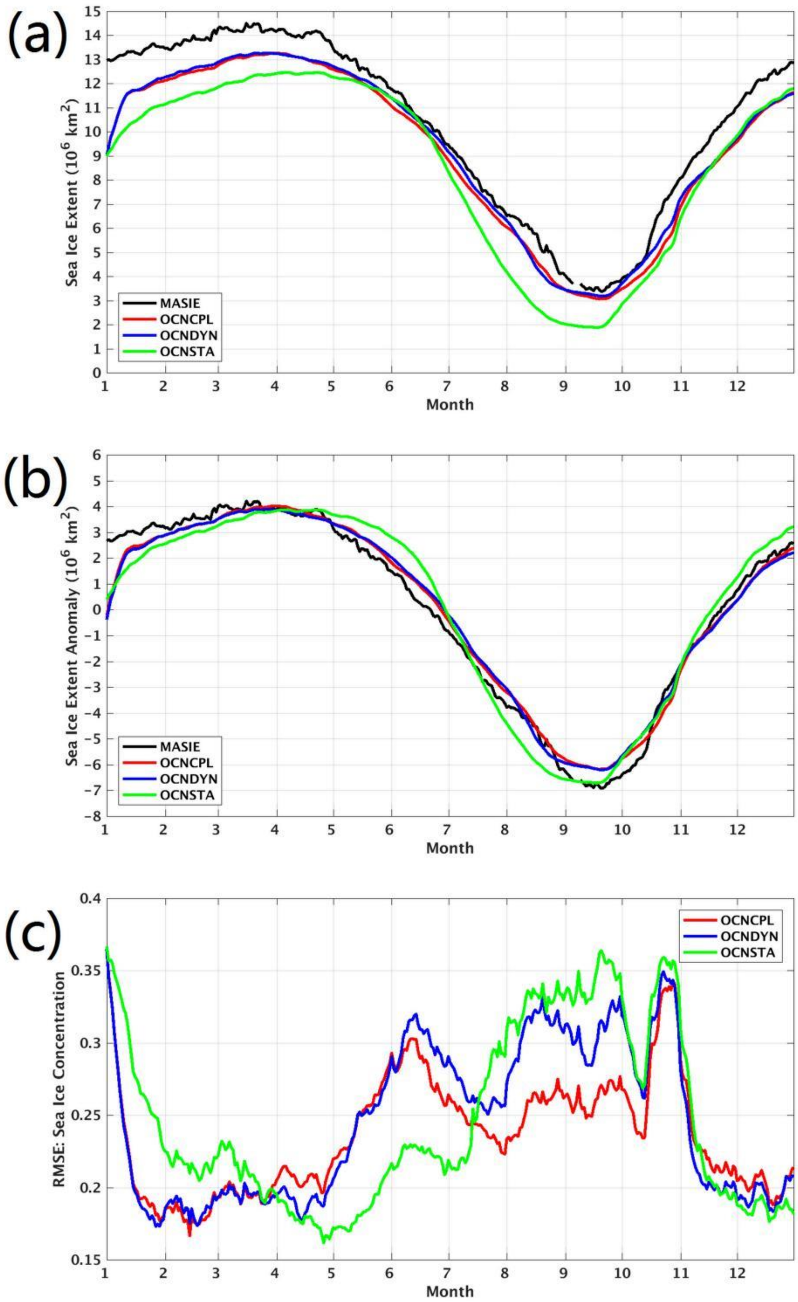


Fig.7 Time series of (a) sea ice extent, (b) sea ice extent anomaly, and (c) root mean square error (RMSE) of modeled sea ice concentration with respect to the OSISAF observation in 2012. The black, red, blue and green lines in (a) denote sea ice extent of the MASIE observation, the OCNCPL run, the OCNDYN run, and the OCNSTA run respectively. The black, red, blue and green lines in (b) denote sea ice extent anomaly of the MASIE observation, the OCNCPL run, the OCNDYN run, and the OCNSTA run respectively. The red, blue and green lines in (c) denote the sea ice concentration RMSE of the OCNCPL run, the OCNDYN run, and the OCNSTA run, respectively.

Reference:

Day, J. J., et al., Will Arctic sea ice thickness initialization improve seasonal forecast skill? Geophys Res Lett, 2014, 41: 7566-7575.

Liang, X. and M. Losch, On the Effects of Increased Vertical Mixing on the Arctic Ocean and Sea Ice. J. Geophys. Res., 2018. 123(12): p. 9266-9282.

Liang, X., et al., Using Sea Surface Temperature Observations to Constrain Upper Ocean Properties in an Arctic Sea Ice-Ocean Data Assimilation System. J. Geophys. Res., 2019. 124(7): p. 4727-4743.

Nguyen, A. T., et al., Arctic ice-ocean simulation with optimized model parameters: Approach and assessment. J. Geophys. Res., 2011, 116: C04025.

Notz, D. & Bitz, C. M. in *Sea Ice* (ed. Tomas, D. N.) (John Wiley & Sons, Chichester, 2017).

Van Pham, T., et al., New coupled atmosphere-ocean-ice system COSMO-CLM/NEMO: assessing air temperature sensitivity over the North and Baltic Seas. Oceanologia, 2014, 56(2): p. 167-189.

Christidis, Z., et al., Performance and Scaling of WRF on Three Different Parallel Supercomputers, in: International Conference on High Performance Computing, Springer, 514–528, 2015.