We would like to thank the reviewer for the evaluation of our study and the constructive comments that helped us to improve the manuscript. Please find below the reviewer's comments in black font and the author's response in blue font.

Responses to Referee Alek Petty

- In the revised manuscript, we added one new subsection 3.1 to show the seasonal difference of T2M, total precipitation and snowfall over the pan-Arctic sea ice and added one new figure (Figure 2). Accordingly, the original section 3.1 changed to 3.2, and the original subsections 3.1 and 3.2 changed to 3.2.1 and 3.2.2, respectively.
- We replotted the original Figure 2 and 3 for clarity, which are Figure 3 and 4 in the revised manuscript:
    - In the new Figure 3 we show the variation of T2M from ERA5, ERA-I and buoys and the differences of T2M between ERA5 and ERA-I for 5 buoys, plots for the other buoys provided in the Supplementary Information as Figure S1 and S2.
    - In the new Figure 4 we show the variation of T2M differences between ERA5/ERA-I and buoys for same 5 buoys as in the new Figure 2. Plots for the other buoys are provided in the Supplementary Information as Figure S3.
- We added a new Figure 6 to show the T2M difference between ERA5/ERA-I and buoys in four regions (Central Arctic, Atlantic sector, Pacific sector, and Laptev Sea) in the Arctic. Correspondingly, we added one paragraph at the end of the original section of 3.1 to describe the regional differences.
- We replotted the original Figure 5 and 6, the new figure named as Figure 7 and Figure S4-5. In these new figures, we show the accumulated total precipitation and snowfall from ERA5 and ERA-I and snow depth measured by buoys.
- The original Figure 7 (showing the FDD and sea ice growth) was moved to Supplementary Information as Figure S6 and the discussion on the FDD model was shortened.
- Forcing data of wind speed (V), relative humidity (Rh) and total cloud (CN) and ocean heat flux were plotted and provided in the Supplementary Information as Figure S7.

For our specific responses please see below.

General comments

I think it would help to show some more general comparisons between ERA5 and ERAI over the Arctic Ocean/sea ice, e.g. raw and difference maps/time series of air temperature, snowfall, precip, pressure. These could be just annual means but seasonal means might be good to see too. This could be included in the SI but I think it will be valuable to include in the main paper to help motivate the study (are there any big/obvious differences from the off?!). This doesn't need to be too detailed.

We now include a new figure showing seasonal mean differences of T2M, total precipitation and snowfall between ERA5 and ERAI over the Arctic (Figure 2). This figure is discussed in a new subsection 3.1 The new Figure 2 is shown below and the added subsection 3.1 reads as,

"3.1 Spatial distribution of seasonal difference of reanalysis near surface temperature and precipitation

Figure 2 shows the seasonal mean differences of T2M, total precipitation and snowfall between ERA5 and ERA-I over Arctic sea ice during 2010-2015. We classify spring as March, April and May, summer as June, July and August, autumn as September, October and November, and winter as December, January and February. The seasonal mean ice extent is obtained from the monthly sea ice concentration from NOAA/NSIDC during 2010-2015 (Meier et al., 2017).

The difference in T2M between ERA5 and ERA-I clearly varies with season (Fig. 2a-d). ERA5 is generally warner than ERA-I in spring and winter, and colder than ERA-I during summer and autumn over most regions of Arctic sea ice. These temperature differences are small during summer, but large during the other seasons. Near the North Pole, ERA5 is warmer than ERA-I in summer, but colder than ERA-I in winter. Whether warmer or colder, the differences between ERA5 and ERA-I are small ($<\pm 0.4$ °C) in this region.

ERA-I is known to be a relatively "dry" global reanalysis product in the Arctic compared with most other modern reanalyses (e.g. MERRA-2, CFSR, and JRA-55) (Lindsay et al., 2014; Merkouriadi et al., 2017; Boisvert et al., 2018). However, the total precipitation in ERA5 is lower than in EAR-I over Arctic sea ice in all seasons (Fig. 2e-h). The lower precipitation in ERA5 is most pronounced in summer, and in the eastern Arctic. Differences in the snowfall between ERA5 and ERA-I are smaller than for total precipitation (Fig. 2 i-j vs. Fig. 2e-h). The snowfall in ERA5 is lower than in ERA-I in spring, autumn and winter, but larger than ERA-I in summer."

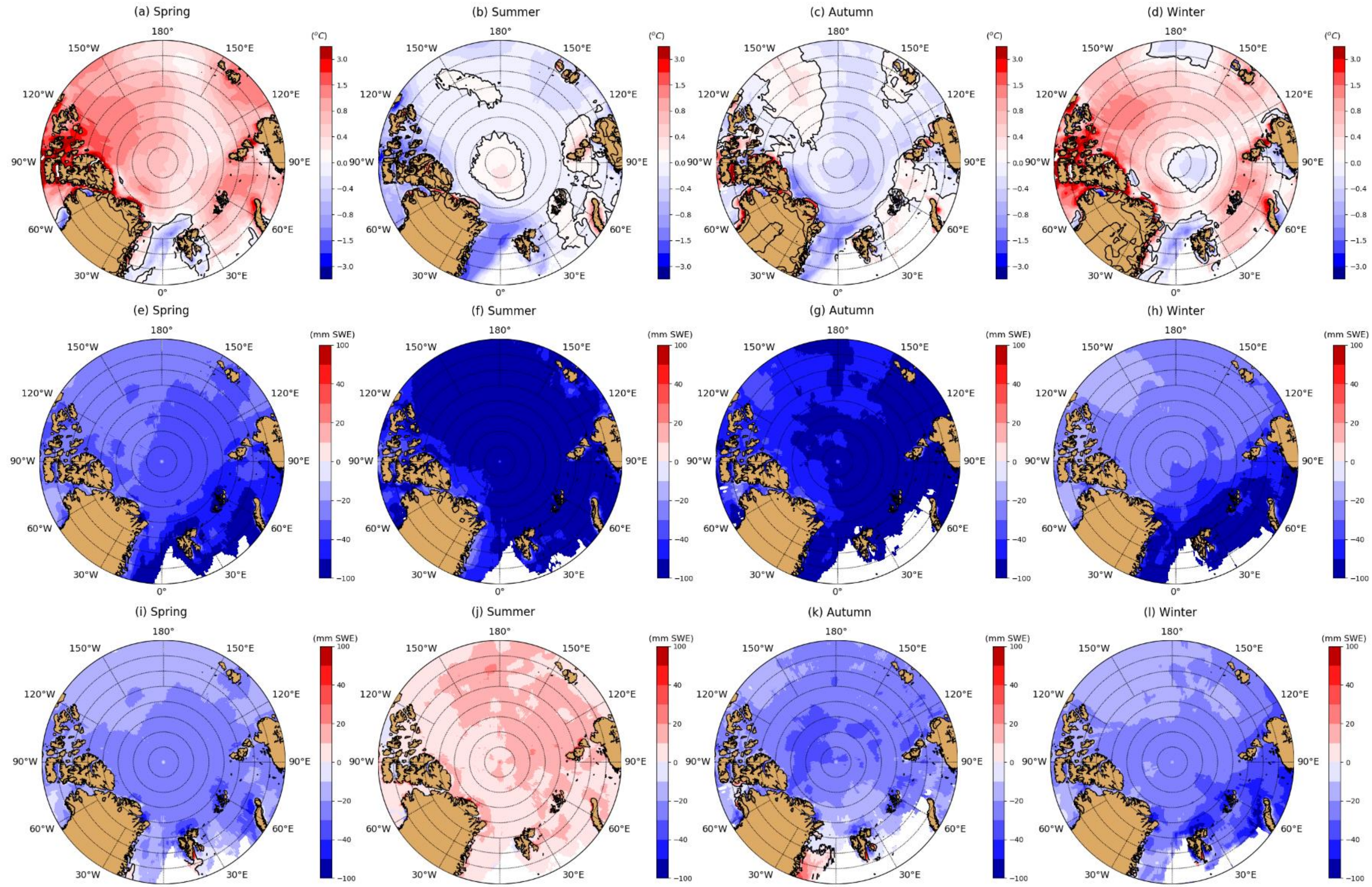

Figure 2. Seasonal mean difference (ERA5-ERA-I) of T2M (a-d), total precipitation (e-h), and snowfall (i-l) in spring (a, e, i), summer (b, f, j), autumn (c, g, k) and winter (d, h, l) over Arctic sea ice. The mean ice extent in the season is used for classification of sea ice and open ocean.

I was a bit disappointed in the FDD analysis and am unsure of its value. The main conclusion seems to be that the warm bias introduces a negative thickness bias, which is pretty obvious without the need for an FDD model. . .I would be tempted to drop this section entirely unless you can make it seem more value-added compared to the 1D modelling study that follows this (and does seem valuable despite my concerns).

We understand these concerns. We have rewritten this section to make it much more concise, and moved Figure 7 to Supplementary Information (now Figure S6). While it is intuitive that a warm bias introduces a negative thickness bias, we believe it is important to give an indication of how large we expect the magnitude of this bias to be based on a simple analytical model. Therefore we did not wish to entirely delete / move this section to supplementary material.

The revised text in this section reads as follows,
"The cumulative freezing degree days (FDD) model only needs air temperature as input and is often used to estimate sea ice growth (Δh) from zero (e.g., Huntemann et al., 2014; Lei et al., 2017). The sea ice growth is estimated based on Lebedev (Maykut, 1986), $\Delta h = 1.33 \sum (FDD)^{0.58}$, where $\sum FDD$ is daily average temperature below the freezing point of sea water (-1.8 °C), integrated over the time period from 1 October to 30 April.

The positive near surface air temperature bias in ERA5 and ERA-I results in a negative ice thickness bias at the end of the growth season. The cumulative FDD is smallest for ERA5 (Fig. S6, Table 2), corresponding to the largest warm bias in ERA5 during the freezing season. The differences in FDD between ERA5, ERA-I and buoys are large for buoys 2011M, 2012H and 2012L, but negligible for buoy 2012J. The ice growth is 0.08-0.12 m less, with a mean of -0.09 m for ERA-I T2M, and 0.13-0.20 m less, with a mean of -0.16 m for ERA5 T2M compared to when using buoy temperatures (Table 2).".

To me it's a shame you didn't show a complete regional Arctic sea ice model forced by both reanalyses as this could have been a useful way of showing regional biases in the reanalyses!

We agree that a complete regional Arctic sea ice model forced by both reanalysis is a useful way to show regional biases in the reanalysis. The work is our next step. However, in this study, the focus is on comparing reanalysis with buoy observations.

While I think the HIGHTSI section is useful and needed, I think it needs to be re-written and potentially expanded on to improve clarity. e.g.: - I know you cite those model papers but I think you to include at least a brief description of the model here and how those forcings were used. Why no downwelling for example? This is calculated form the cloud cover? I think you can provide some general information then cite the papers for more information.

We now provide further details on the HIGHTSI model and how it was forced.

"The snow and ice temperature regimes are solved by the partial differential heat conduction equations applied for snow and ice layers, respectively. The turbulent surface fluxes are parameterized taking the thermal stratification of the atmosphere surface layer into account. Downward short- and longwave radiative fluxes are parameterized, based on the total cloud cover."

We also now show the forcing data in Figure S7.

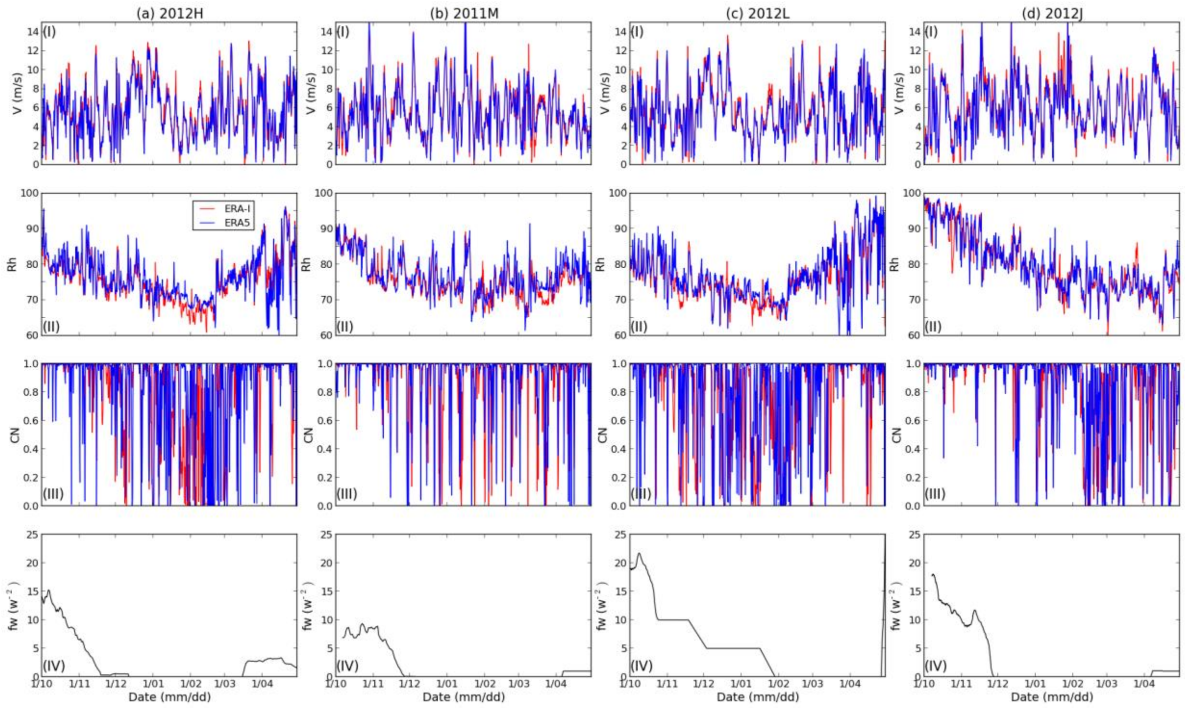

Figure S7. Forcing data of wind speed (V), relatively humidity (Rh), total cloud (CN) and ocean heat flux (fw) used in the model.

I got pretty confused by the use of Sp and Sf and think there could be some mistakes here. Why did you start by comparing the Sp runs and not the Sf runs? Sp is a bit of a confusing acronym so maybe you could try something like Tp2mt?

We apologize for this confusion. We have introduced a new naming system for the model runs. We now use TP to indicate total precipitation, and SF for snowfall. "I" indicates forcing is from ERAI, and "5" means it is from ERA5. Thus TPI_T2MI means total precipitation and T2M are both from ERAI. Further details are provided in Table 2, and the manuscript text.

Can you not also show some idealized Arctic mean simulations instead of just the buoy track simulations? This might give us a better sense of what the potential impact of these differences might be when we want to consider the Arctic as a whole.

Our simulations for the four buoys already cover FYI and MYI with a range of snow depths and ice thicknesses. We believe that these simulations already provide a realistic indication of the potential impacts of these differences over the wider Arctic, without the need for idealized simulations. For such idealized simulations we would neither have any direct observations to compare their validity, and such we do not believe they are of great value, as the focus in this study is to compare reanalysis to buoy observations.

‘a good representation of precip seems crucial’ seems like a pretty loose interpretation of the analysis you presented. I think this discussion needs to tie back better to what exactly your results demonstrated.

We have rewritten this sentence as

“Thus, not only the magnitude but also the frequency of the precipitation in the reanalysis data is crucial for the snow evolution in the simulation.”

I think you need to better justify early on why you only look at these two variables. Maybe mention earlier that you also looked at MSLP in ERAI and ERA5 (reanalyses tend to agree more in this regard as expected) and that you're limited by what the buoys can provide?

We focus on these two variables because these are the observations we have from the buoys. The final paragraph of the introduction has been rewritten as follows:

"In this study, we compare and evaluate the performance of ERA-I and ERA5 over Arctic sea ice. For this, we use data from Ice Mass Balance buoys (IMB) (Perovich et al., 2018) and Snow Buoys (Grosfeld et al., 2016; Nicolaus et al., 2017) deployed in 2010-2015. The buoys record position, the 2 m air temperature (T2M), mean sea level pressure (MSLP), and snow depth at regular intervals. Hence, these observations can be used to evaluate the variables of T2M, precipitation and MSLP in the reanalyses. The former two variables are critical parameters for sea ice simulation (Cheng et al., 2008; Wang et al. 2015), and form the focus of our study. We use the T2M and snow depth observations from these buoys to assess the performance of ERA5 and ERA-I over Arctic sea ice. We further use the reanalyses to force a 1-D thermodynamic sea ice model. The simulations are compared with snow and ice thickness observations from the buoys to evaluate how differences in the T2M and precipitation influence the evolution of sea ice in the model."

You later force the HIGHTSI model with other variables (e.g. cloud cover) so I think you should show these and their differences too, despite the lack of buoy obs to validate it against. These variables are shown in Figures S7 in the Supplementary Information.

Comparing 2 m and 10 m air temps might be illuminating.

We agree that comparing 2 m and 10 m air temperature might be illuminating. However, we not have observations from 10 m height, and so feel this goes beyond the scope of this study as the focus is to be able to compare buoy observations and reanalysis directly.

Any change in how the 2 m air temps are calculated in ERA5 (still not an explicit model level, right?).

Computation of temperature at 2 m level is based on interpolation between the lowest model level and the surface making use of the same profile functions as in the parametrization of the surface fluxes. Therefore there was no change in the computation of 2 m air temperature in ERA5 and ERAI. However, the lowest model levels are different in ERA5 and ERAI due to the higher vertical resolution in ERA5.

Confused why you need to interpolate the ERAI data to the ERA5 grid before interpolating to the buoy position. Guessing this won't be a big issue as ERA5 is of intermediate resolution but still seems odd to me. How do you deal with the temporal differences between ERAI and ERA5? ERA5 is hourly and ERA-I is 6-hourly?

The original text is indeed unclear, thank you for pointing this out. We do NOT interpolate ERAI to ERA5 for buoy comparison. To clarify, we have rewritten the text in the manuscript as follows: "For comparison and evaluation against buoy observations, ERA5 is bilinearly interpolated to the buoy positions, and ERA-I is first linearly interpolated to hourly data, and then bilinearly interpolated to the buoy positions. For the comparison between ERA-I and ERA5 over the Arctic sea ice, the ERA-I data are first bilinearly interpolated to the grid of ERA5, and then T2M is averaged in the season, and total precipitation and snowfall is integrated over the season to calculate the seasonal mean."

I'm confused why you don't show the actual ice thickness for the buoys (I think you just show the estimated ice thickness change from the FDD model?). Also confused as to whether you initialize the FDD model with zero ice thickness or not, as you show ice growth, not ice thickness. Any reason for this? Again, I see little value in this analysis so suggest dropping this and improving the rest of the analysis presented in the paper.

The new Figures 8 and 9 show the snow and ice thickness from the buoy observations and from the 1-D model runs (this is indicated in the legend), NOT for the FDD model. FDD is often used to estimate ice thermodynamic growth from zero, so ice thickness is from zero in the FDD model. For clarity, we have rewritten the FDD model part and made it much shorter as you suggest (Please refer to our response to your General Comments above).

Specific comments

P1 L3-4: 'The decline of Arctic sea ice has been attributed to various interrelated causes, including a general overall warming trend (Steel et al., 2008; Polyakov et al., 2010).' Seems pretty vague so would recommend you either improve this or drop it.

This sentence was deleted.

P1 L10: I would replace 'in-situ atmospheric observations' with something like 'direct observations of the atmosphere, sea ice and ocean conditions'?

Replaced with "direct observations of the atmosphere, sea ice and ocean conditions"

P1 L15: I suggest you combine this line 'Atmospheric reanalyses etc..' with the one about their use earlier on L12-13.

This sentence was moved to directly follow the earlier sentence about reanalyses other uses, and rewritten as follows: "In addition, reanalyses are also frequently used to force snow and sea ice models (Schweiger et al., 2011; Merkouriadi et al., 2017; Stroeve et al., 2018).".

P1 L22: I think you need to make the point here that the 1950 onwards data isn't yet available yet? Unless you're guessing it will be at the time of publication. . .

We now clarify the data availability as, "The entire ERA5 dataset, extending back to 1950, will be available for use in late 2019." on P4 L5-6 in the revised manuscript.

I also think you need to provide a better discussion of these supposed improvements and how they might increase ERA5's utility for Arctic studies, e.g.: - What do you mean by improved representation of troposphere and global balance? - More consistent? How?

We are limited in what we can say here, because there is still no official peer-reviewed publications documenting the ERA5 product, and neither many validation studies of ERA5 as far as we know, that is why this study is of value. This text has been revised as follows:

“There are several major improvements in ERA5 compared with ERA-I, including much higher spatial and temporal resolutions and more consistent sea surface temperature and sea ice concentration (Hersbach and Dee, 2016). Evaluations of the performance of ERA5 have been conducted over the land and revealed a higher performance of ERA5 than ERA-I (Albergel et al., 2018; Urraca et al., 2018), and other commonly used reanalysis, such as, MERRA-2 (the second version of the Modern-Era Retrospective Analysis for Research and Applications) (Olausen, 2018; Urraca et al., 2018). However, the performance of ERA5 over Arctic sea ice is yet to be fully investigated.”

P3 L20: Where were these buoys deployed?

This sentence has been revised to provide this information: “For additional coverage, we also use observations from 3 snow buoys deployed in 2015, two of which in the Laptev Sea and one in the Central Arctic (Table 1; Fig. 1).”. Please also refer to P3 L30-31 and P4 L1.

P4 L23: Boisvert not Biosvert P6 L3-4: Biosvert is corrected to Boisvert.

And any thoughts on how the cloud physics might have changed in ERA5 to cause this big change in snowfall/precip ratio?

We have added the following text in to the revised manuscript: “In ERA-I, the split between liquid and ice in clouds is determined diagnostically as a function of temperature from −23 to 0 °C, with ice-only below −23 °C and liquid-only above 0 °C. In contrast, the IFS Cy41r2 used in ERA5 includes a prognostic microphysics scheme, with separate cloud liquid, cloud ice, rain and snow prognostic variables (Sotiropoulou et al., 2015; see also ECMWF IFS documentation –Cy41r2; https://www.ecmwf.int/sites/default/files/elibrary/2016/16648-part-iv-physical-processes.pdf). Our findings indicate that ERA5 has significantly less Arctic rainfall than ERA-I, particularly in August-September (Fig. 7, Figs. S4-5).” (Please also refer to P7 L25-29 in the revised manuscript).

I think you should plot this ratio out as a separate figure as it seems like a crucial part of the story.

We feel that this moves beyond the scope of this study. The Snowfall/Total precipitation ratio can clearly be seen by comparing the cumulative snowfall /total precipitation plots in Figure 7 of the revised manuscript, and so does not necessitate an additional figure.

P6 L4-5: I’m not quite sure how Figure 5 shows it is less anomalous as these are just showing the reanalysis data not compared to anything. Think you need to plot the buoy results too despite the big issues of representation etc.

We now plot the buoy snow depth as snow water equivalent in Figure 7 (formerly figure 5). Further buoys are shown in Figure S4 and S5 in the supplementary information. The figures are shown below. We have removed the word “anomalous” from the manuscript.

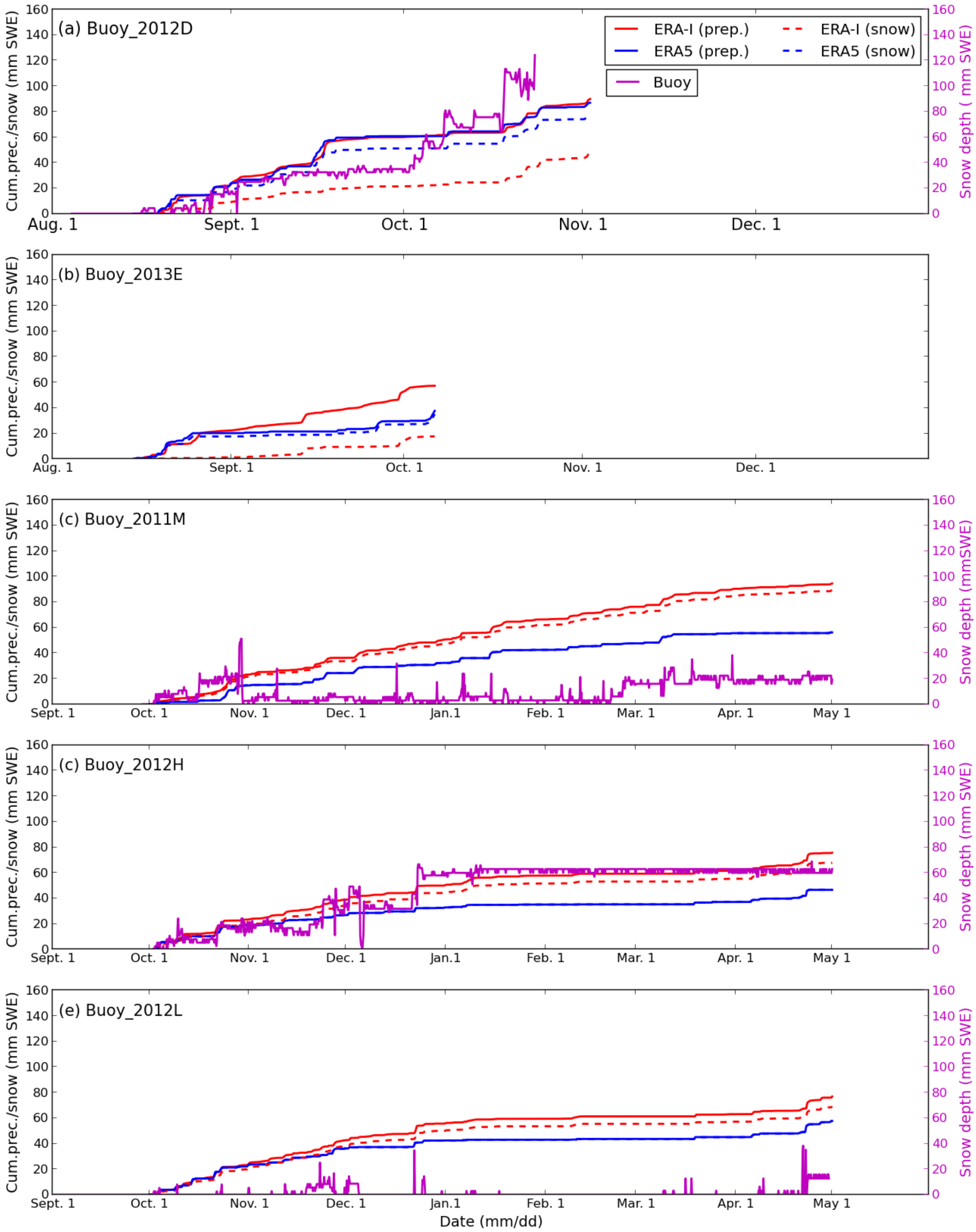

Figure 7. Cumulative total precipitation (prec.) and snowfall (snow) for ERA5 and ERA-I and snow depth for buoys (a) 2012D, (b) 2013E, (c) 2011M, (d) 2012H, and (e) 2012L. Accumulation starts from 15 August for panels (a) and (b), and from 1 October for panels (c)-(e). Note there was no snow depth data for buoy 2013E during the accumulation period.

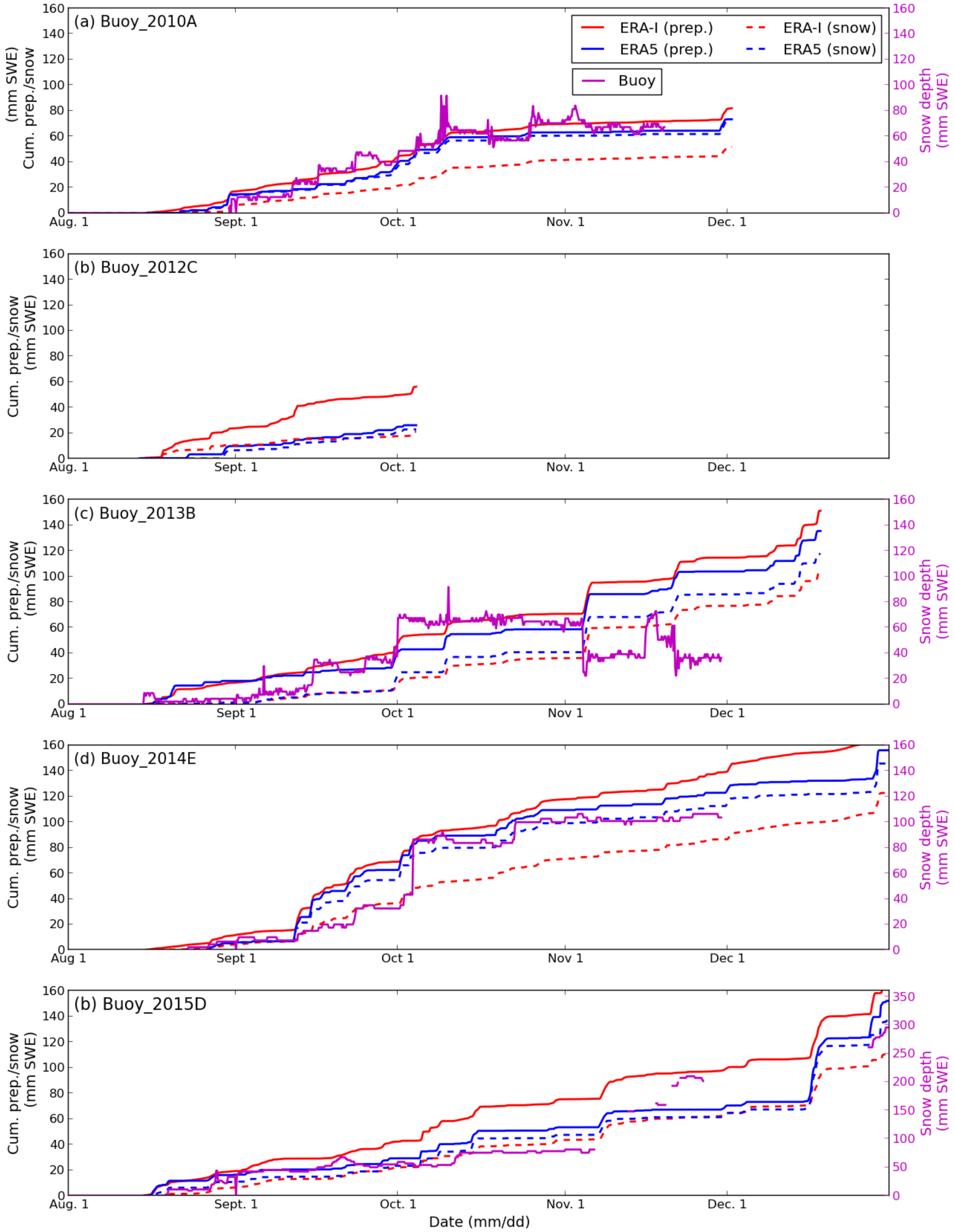

Figure S4. Cumulative total precipitation (prec.) and snowfall (snow) for ERA5 and ERA-I and snow depth for buoys (a) 2010A, (b) 2012C, (c) 2013B, (d) 2014E, and (e) 2015D. Accumulation starts from 15 August. Note that Buoy_2012C does not have snow depth.

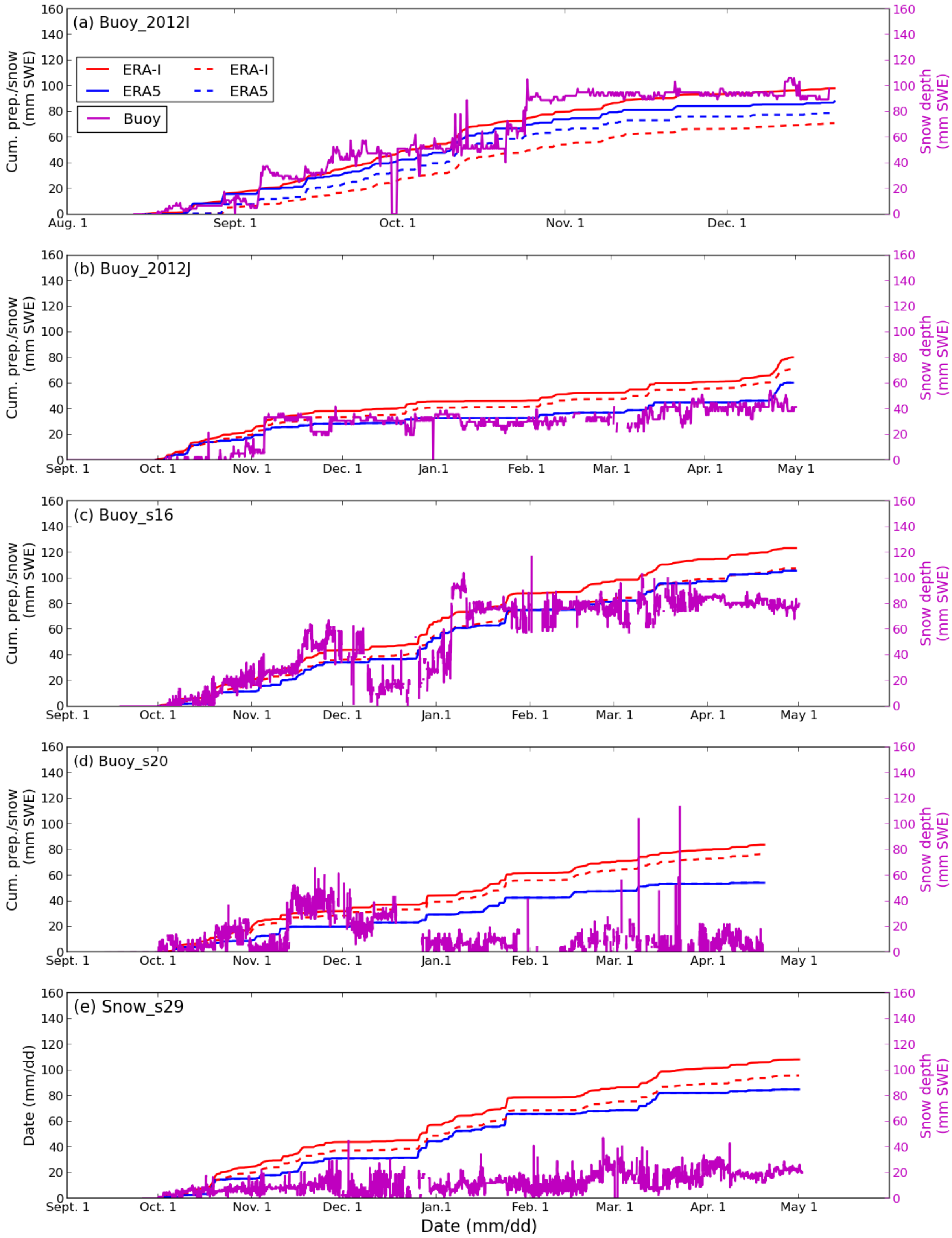

Figure S5. Same as Figure S4, but for buoys (a) 2012I, (b) 2012J, (c) s16, (d) s20, and (e) s29, and accumulation starts on 1 October.

P6 L10: why not use a daily climatology of density? You cite Warren1999 for the 350 value but this seems overly simplistic considering the results presented in Warren1999.

Thanks for the suggestion. The constant snow density was replaced with climatological monthly snow densities based on Warren et al. (1999). The results are shown in the new Figure 7, Figure S4 and S5, and Table 1.

P8, L4-5: can you briefly describe what this ocean heat flux is? E.g. 2 W/m2?

We have added description for the ocean heat flux in the manuscript on P10 L17-19 as "For all of the simulations we apply a seasonally variant ocean heat flux according to McPhee et al. (2003), which is large in October (10-20 Wm^{-2}), and decreases to nearly zero from mid-November." Accordingly, subplots for the used ocean heat flux were provided in the supplementary information Figure S7.

P9, L30: drop the warm summer bias comment here as you repeat it later.

This section now reads: "Overall, we find a warm bias in ERA-I and ERA5, when compared with the buoys. In both reanalyses, these biases are smallest in summer months, and larger during the autumn, winter and spring. The warm bias in ERA5 is smaller than ERA-I during the summer months. However, we find a larger warm bias in ERA5 than in ERA-I during the cold season, especially when the observed T2M was lower than -25 °C in the Atlantic sector and the Pacific sector."

P10, L1-3: think you should mention the caveat here that the buoy probably isn't giving the 2 m air temperatures.

We now include the sentence: "The near surface warm bias in ERA5 and ERA-I may also partly be attributed to the difference in height with observations."

Can we be sure resolving the boundary layer is the actual problem here?

We cannot be sure, as we do not analyze vertical profiles here. However, several studies have highlighted this problem in a range of reanalysis products, and many studies indicate this is the cause, for example (Beesley et al., 2000; Tjernstöm and Graversen, 2009; Graham et al., 2017b; Kayser et al., 2017)

P10, L5-14: I think this needs a bit of rewording for clarity. Really worth stressing that the total precip is lower but the snowfall is higher, right?! I would start with that difference then explain what it means in terms of the comparisons with the buoys. - Think you also need to make the point later regarding which precip was used to force the 1D model and that care needs to be taken regarding how precip is used in the products perhaps.

We have rewritten this section to clarify. The snowfall product in ERA5 is only larger than ERA-I during summer months. The text now reads: "The total precipitation over Arctic sea ice in ERA5 was lower than in ERA-I in all seasons. This is surprising, as ERA-I is known to be drier in the Arctic compared with some other recent reanalyses (Lindsay et al., 2014; Merkouriadi et al., 2017; Boisvert et al., 2018). However, the snowfall is higher in ERA5 than in ERA-I during the summer months. This indicates that ERA5 has a higher snowfall to precipitation ratio than ERA-I during summer. ERA-I is known to have an anomalously large fraction of liquid precipitation and low snowfall to precipitation ratio in the Arctic, especially during August-September (Dutra et al., 2011; Leeuw et al., 2015). The total precipitation accumulated along the buoys drift trajectories, during the cold season (from 15 August/1

October until buoy failed or 30 April), was lower in ERA5 than in ERA-I for every buoy. Similarly, the accumulated snowfall in ERA5 is lower than in ERA-I for all buoys with an accumulation date starting from 1 October. In contrast, the total accumulated snowfall in ERA5 is higher than in ERA-I for buoys with an accumulation date starting from 15 August, due to likely anomalous summer/autumn rainfall in ERA-I being classified as snow in ERA5. The accumulated total precipitation and/or snowfall in ERA5 are often closer to the SWE content of buoy measured snow pack, compared with ERA-I. Nonetheless, the lack of representative in-situ observations and difficulty in measuring snow accumulation on sea ice in the Arctic makes it a challenge to accurately evaluate precipitation products over sea ice (e.g. Rasmussen et al., 2012; Lindsay et al., 2014; Sato et al., 2017; Blanchard-Wrigglesworth et al., 2018; Boisvert et al., 2018). Given snow is such a critical factor in sea ice evolution, more representative observations are therefore needed (e.g. Merkouriadi et al., 2017; Webster et al., 2018)." (Please refer to P12 L22-31, P13 L1-8 in the revised manuscript)

Any particular recommendations here? I.e. do you think we should be using the snowfall product or deriving this from the total precip? There are other ways of doing this also (using higher level temps).

We do not make any specific recommendations here, because we do not have sufficient data to validate the reanalyses products. We simply highlight that the choice of snowfall / precipitation products will affect the simulation of snow and sea ice in a model. We conclude that further precipitation and snow pack measurements over Arctic sea ice are essential.

Figure 2: why does the green line seem dashed? Can you move the difference line lower so it's easier to see? - I also think you should show not the difference between ERA5 and ERA-I but two lines representing the differences between the reanalyses and the buoys. Maybe just pick a couple as good examples of the seasonal cycles you mention in the text and make these bigger/clearer, then put the rest in the supplementary information? As it is, it's hard to really get a sense of what these figures show quantitatively.

Thank you for the suggestions. We replotted Figures 2 and 3.

In the new Figure 3, we show the variation of 2 m air temperature in ERA5, ERA-I and the buoys. The differences of T2M between ERA5 and ERA-I are shown in a separate subplot below.

The differences between ERA5/ERA-I and the buoy measurements are snow shown in a separate new Figure (Figure 4).

We have reduced the number of buoys shown in the manuscript to five examples. All the other buoys are shown in supplementary Figure S1 and S2.

Figures 3-4 and Figure S1-S3 are shown below.

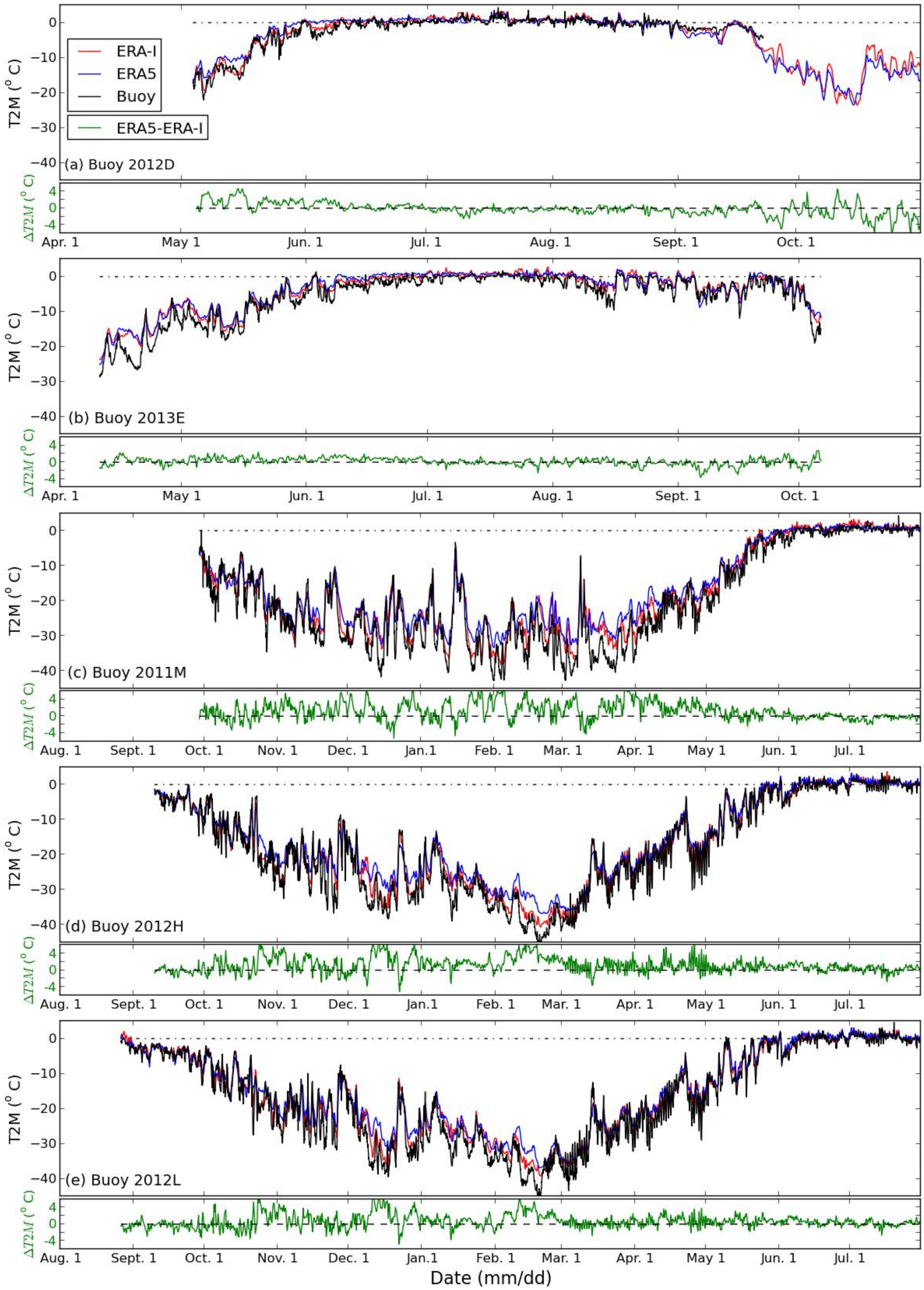

Figure 3. Variation of 2 m air temperature (T2M) in ERA5, ERA-I and the buoys and the differences of T2M between ERA5 and ERA-I for buoys (a) 2012D, (b) 2013E, (c) 2011M, (d) 2012H, and (e) 2012L. Note the different time-axis.

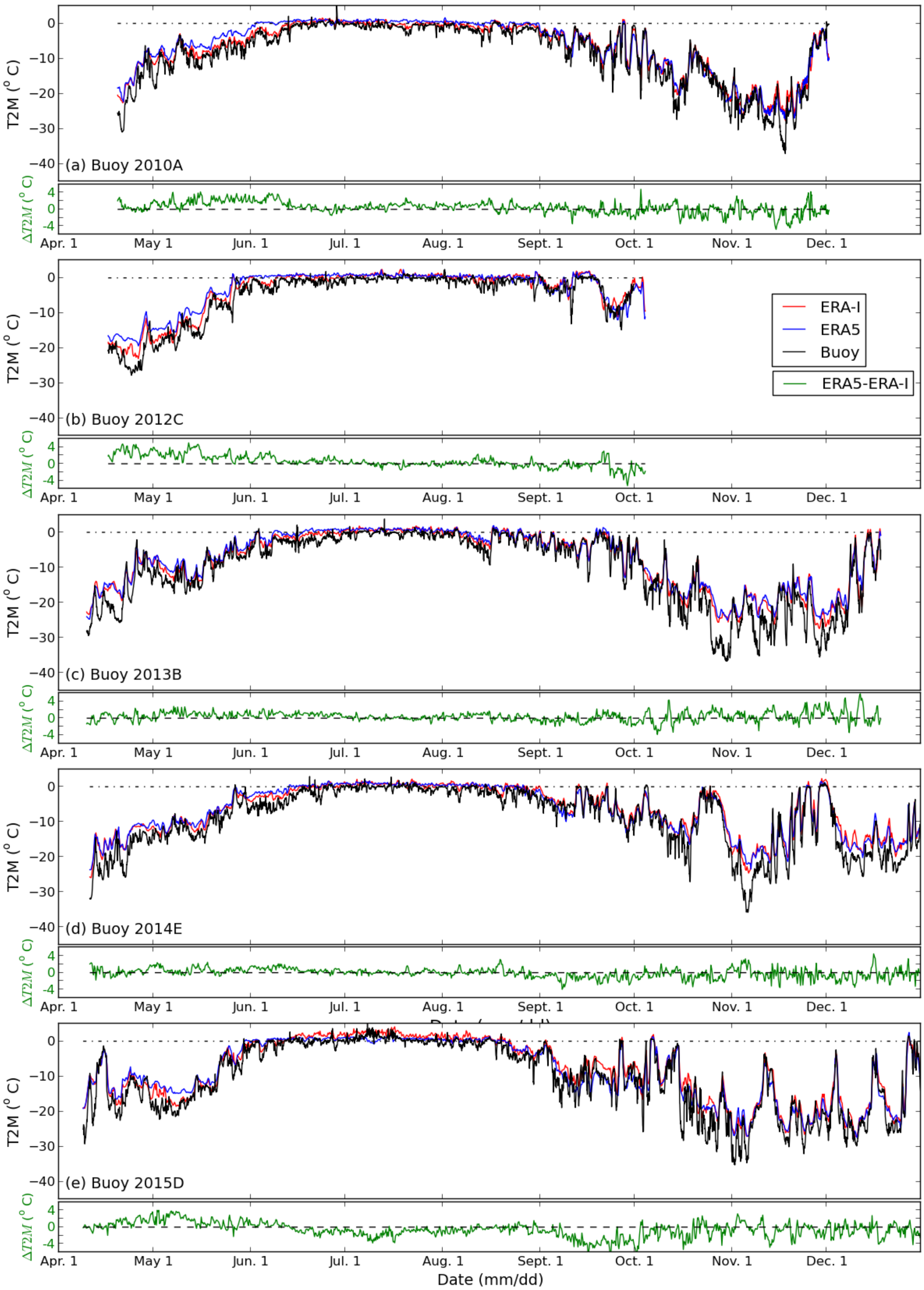

Figure S1. Variation of 2 m air temperature (T2M) in ERA5, ERA-I and the buoys and the difference of T2M between ERA5 and ERA-I for buoys (a) 2010A, (b) 2012C, (c) 2013B, (d) 2014E, and (e) 2015D.

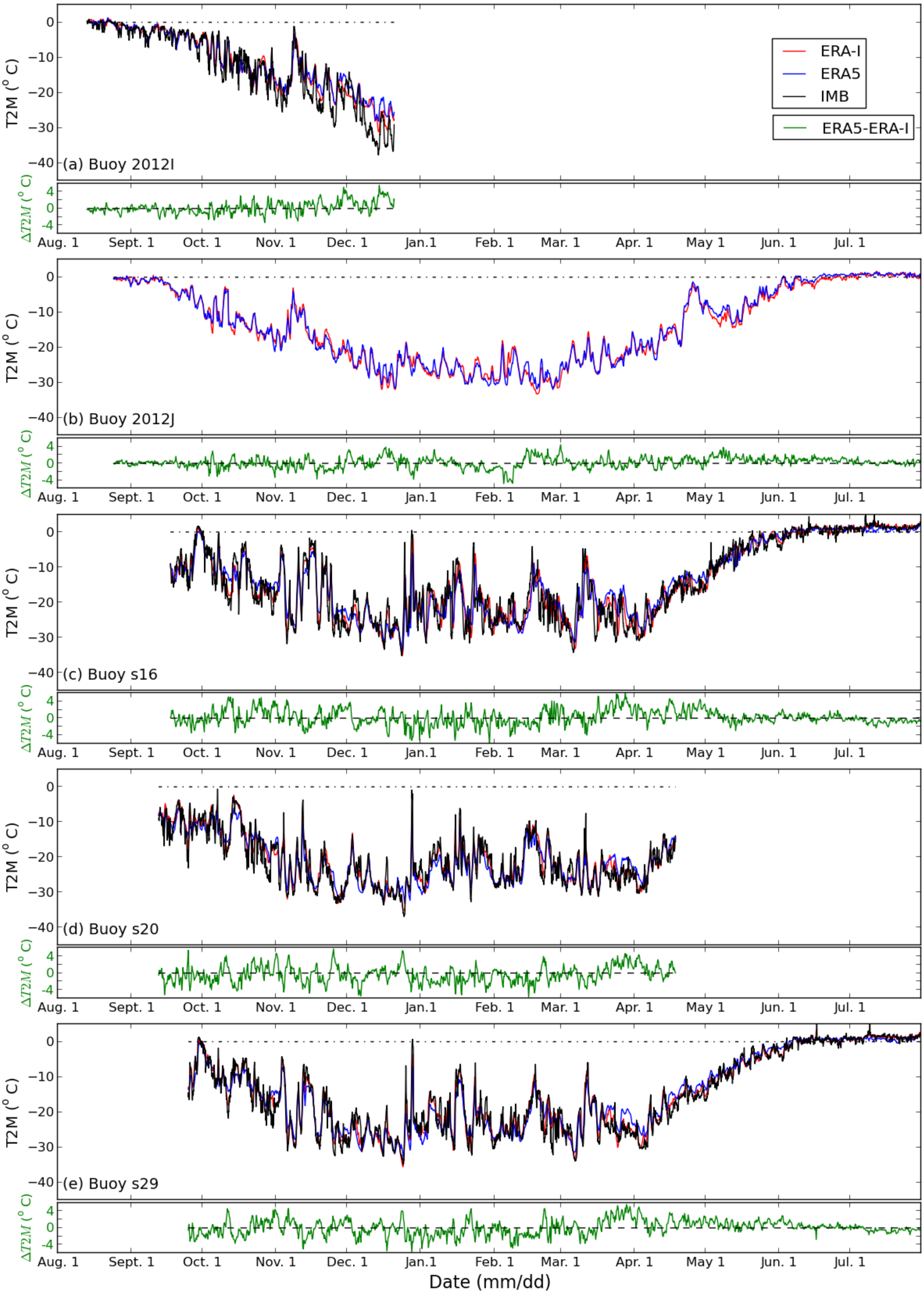

Figure S2. Same as Figure S1, but for buoys (a) 2012I, (b) 2012J, (c) s16, (d) s20, and (e) s29. Note there is no buoy data in Figure (b) for buoy 2012J.

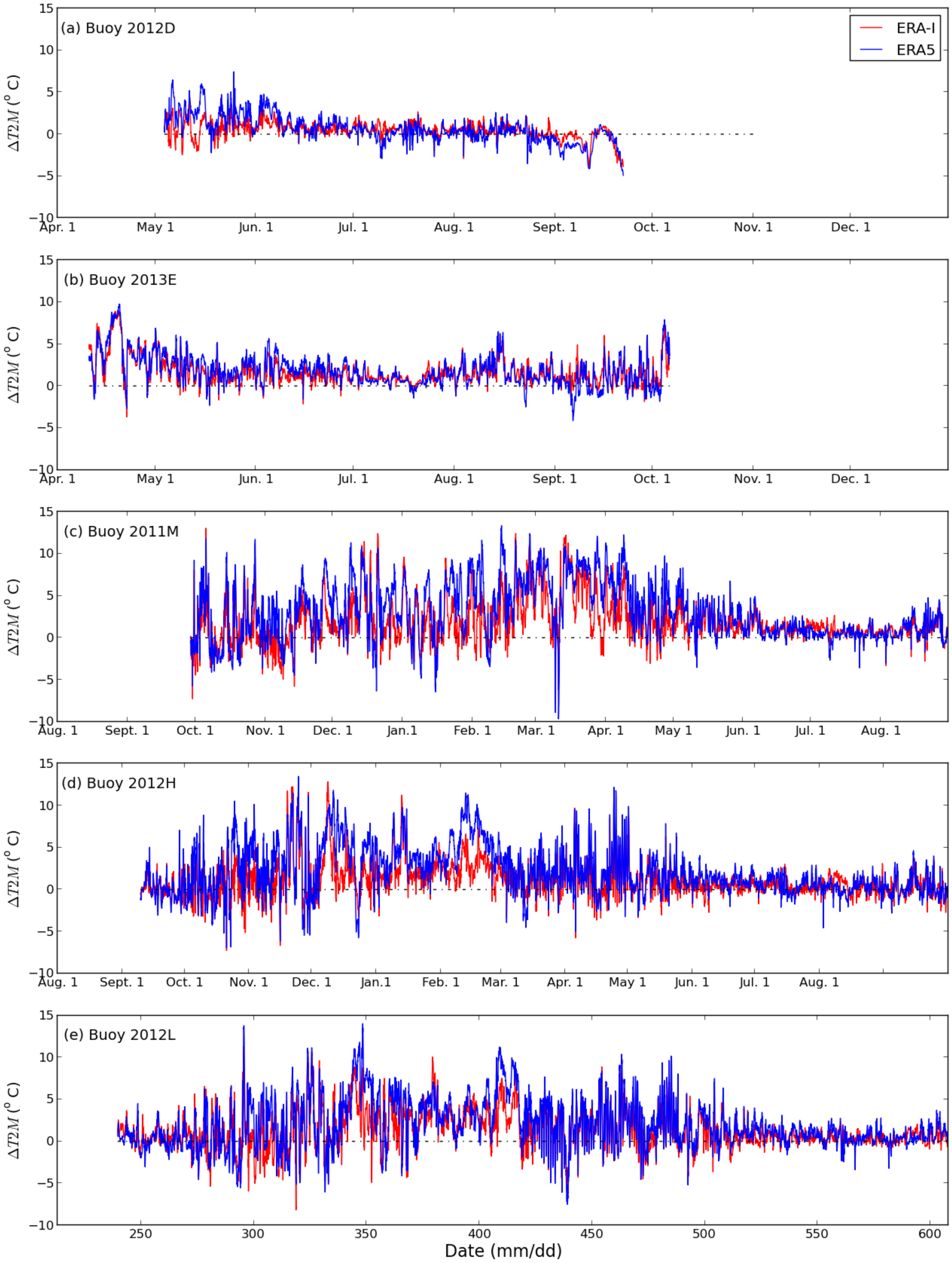

Figure 4. Variation of T2M differences between ERA5/ERA-I and buoys for (a) buoy 2012D, (b) buoy 2013E, (d) buoy 2011M, (d) buoy 2012H, and (e) buoy 2012L.

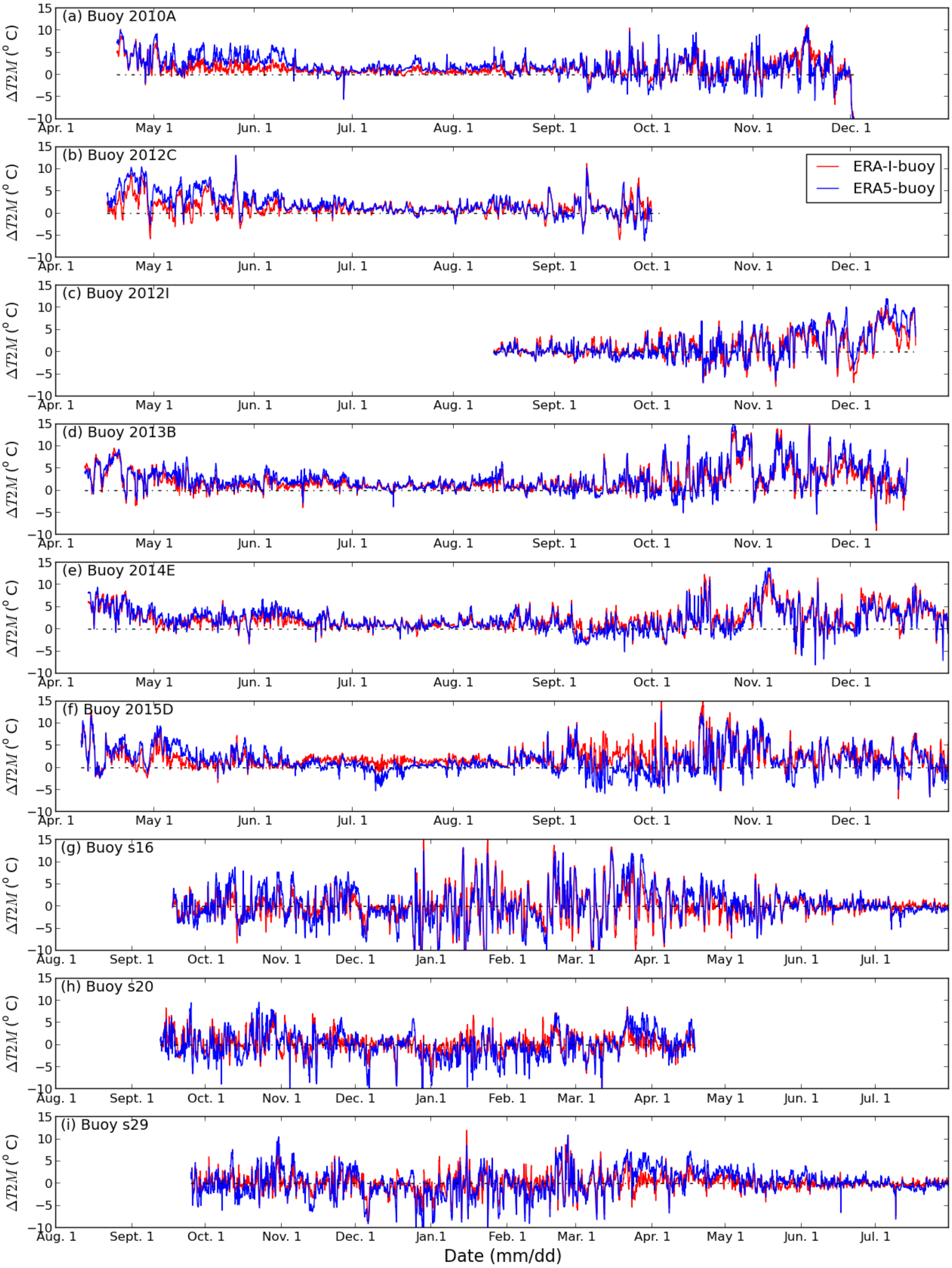

Figure S3. Variation of T2M differences between ERA5/ERA-I and buoys for buoys (a) 2010A, (b) 2012C, (c) 2012I, (d) 2013B, (e) 2014E, (f) 2015D, (g) s16, (h) s20, and (i) s29.

Figure 4: this is a good figure!

Thank you.

Figure 5: you don't need the second y-axis, just state in the legend that the dashed lines are snow. You should add the units to the label and legend. Why does this not include the buoys change in snow depth? As a second y-axis?! Or converted to cumulative precip and plotted on the same axis.

We have replotted the Figure 5 (now Figure 7) as mentioned in our response to your General comments. We now show the snow depth as snow water equivalent from buoys, as suggested. Please refer to our reply in the General comments part.

Figure 7: I don't think you need to show the FDD values as they don't mean much physically.

As mentioned in our response to your General comments, we made the FDD part more concise, and moved the figure to supplementary information (Figure S6).

Figure 8: why the weird staircase in buoy 2011M? Lower temporal resolution for some reason?

The staircase pattern is an artifact of the automated system that is being applied to archive the data. If the acoustic sounders fail and the temperature string is still working, the positions of the ice surface and bottom are determined from the temperature string (with much less accuracy than from the acoustic sounder). We have added one sentence to explain how the ice thicknesses were determined for the buoys on P3, L22-23 in section 2.1, and pointed out the reason for the staircase ice thickness for buoy 2011M and 2012J on P10, L31-33.