October 3, 2017
Dr. Masashi Niwano
Meteorological Research Institute
Japan Meteorological Agency
Tsukuba 305-0052, Japan
Phone: +81-29-853-8714
FAX: +81-29-855-6936
e-mail: mniwano@mri-jma.go.jp

Prof. Tedesco
Editor, The Cryosphere

Dear Prof. Tedesco:

Please find enclosed our responses to the all reviewers' comments as well as the revised marked-up manuscript entitled as "NHM-SMAP: Spatially and temporally high resolution non-hydrostatic atmospheric model coupled with detailed snow process model for Greenland Ice Sheet" by Masashi Niwano et al. [Paper # tc-2017-115] submitted to the journal The Cryosphere. We have revised the manuscript according to the all reviewers' comments. We believe these constructive comments and suggestions provided by the reviewers significantly improved the manuscript. We would like to thank you for obtaining such expert reviewers.

We hope that the revised manuscript is suitable for publication. We look forward to hearing from you.

Yours sincerely,

Masashi Niwano and co-authors

Reply to Reviewer#1

We sincerely appreciate the reviewer for taking the time to provide valuable comments and suggestions. Below we describe our responses (in blue text) point-by-point to each comment (in black text). In addition, we indicate revisions in the updated manuscript together with new line numbers.

General synopsis

This is a useful and original study of Greenland climate and surface mass balance conducted using a non-hydrostatic regional climate model. I would like to see some comparison of NHM-SMAP model output, for example as presented in Figures 9 and 10, with other RCM model output (e.g. MAR, RACMO, HIRHAM). The paper is generally well structured, written and illustrated, and should be publishable with relatively minor modifications. Citation of related work can be improved in places.

We highly appreciate for this positive evaluation. In the revised manuscript, we have included simulation results from MAR v3.5.2 forced by JRA-55. At present, there are many different points in model formulations and configurations of existing RCMs, namely, resolution, ice sheet mask, dynamic core of atmospheric model, albedo model, water percolation scheme for snow/firn, etc. Therefore, detailed model inter-comparison is beyond the scope of this paper; however, we do hope to perform such a comparison in the near future. Regarding the insufficiency of references, we have included all the references suggested by the reviewer in the revised manuscript.

Specific comments

p.2, l.35 Consider adding more recent relevant references, e.g. van den Broeke 2016 The Cryosphere, Hanna et al. 2013 Nature:

van den Broeke, M. R., Enderlin, E. M., Howat, I. M., Kuipers Munneke, P., Noël, B. P. Y., van de Berg, W. J., van Meijgaard, E., and Wouters, B.: On the recent contribution of the Greenland ice sheet to sea level change, The Cryosphere, 10, 1933-1946, doi:10.5194/tc-10-1933-2016, 2016.

Hanna, Edward and Navarro, Francisco J. and Pattyn, Frank and Domingues, Catia M. and Fettweis, Xavier and Ivins, Erik R. and Nicholls, Robert J. and Ritz, Catherine and Smith, Ben and Tulaczyk, Slawek and Whitehouse, Pippa L. and Jay Zwally, H. (2013) Ice-sheet mass balance and climate change. Nature, 498 (7452). pp. 51-59. ISSN: 0028-0836.

Thank you for the suggestion. We have added these important references in the updated manuscript. (P. 2, L. 35 - 36) p.2, l.66: Not just RCMs but also statistically-downscaled meteorological reanalysis data have been successfully used here (Hanna et al. 2005 & 2011, Wilton et al. 2017)

– please add these relevant references:

Hanna, E. and Huybrechts, P. and Janssens, I. and Cappelen, J. and Steffen, K. and Stenhens, A. (2005) Runoff and mass balance of the Greenland ice sheet: 1958-2003. Journal of Geophysical Research Atmospheres, 110 (13). ISSN: 2169-897X.

Hanna, E. and Huybrechts, P. and Cappelen, J. and Steffen, K. and Bales, R. C. and Burgess, E. and McConnell, J. R. and Steffensen, J. P. and Van Den Broeke, M. and Wake, L. and Bigg, G. and Griffiths, M. and Savas, D. (2011) Greenland Ice Sheet surface mass balance 1870 to 2010 based on Twentieth Century Reanalysis, and links with global climate forcing. Journal of Geophysical Research: Atmospheres, 116 (24). ISSN: 2169-897x.

Wilton, D. J. and Jowett, A. and Hanna, E. and Bigg, G. R. and Van Den Broeke, M. R. and Fettweis, X. and Huybrechts, P. (2017) High resolution (1 km) positive degree-day modelling of Greenland ice sheet surface mass balance, 1870-2012 using reanalysis data. Journal of Glaciology, 63 (237). pp. 176-193. ISSN:0022-1430.

We agree with this point. All the suggested papers have been listed up in the reference, and we have revised the sentence as follows:

"Several physically based regional climate models (RCMs) have been applied in the GrIS (e.g., MAR: Fettweis, 2007; RACMO2: Noël et al., 2015; Polar MM5: Box, 2013; and HIRHAM5: Langen et al., 2015) that have been found reliable in terms of reproducing current climate conditions (e.g., Fettweis, 2007; Box, 2013; Fausto et al., 2016; van den Broeke et al., 2016) and simulating realistic future climate change (e.g., Franco et al., 2013)."

----->

"Several physically based regional climate models (RCMs) (e.g., MAR: Fettweis, 2007; RACMO2: Noël et al., 2015; Polar MM5: Box, 2013; and HIRHAM5: Langen et al., 2015) and statistically-downscaled meteorological reanalysis data (Hanna et al., 2005, 2011; Wilton et al., 2017) have been applied in the GrIS that have been found reliable in terms of reproducing current climate conditions (e.g., Fettweis, 2017; Hanna et al., 2011; Box, 2013; Fausto et al., 2016; van den Broeke et al., 2016) and simulating realistic future climate change (e.g., Franco et al., 2013)." (P. 2, L. 63-68)

--- Please note that the above revised sentence has been improved from our initial response submitted to the discussion board. ---

p.3, ll.83-85: Consider emphasising more that a key advantage of using a nonhydostatic model is its ability to be run at much higher spatial resolutions («5 km) than hydrostatic models. Bearing the above in mind, was it considered to run the JMA-NHM at higher spatial resolutions than 5km (p.6, l.204)?

Thank you for the encouraging comment. To emphasize a key advantage of a non-hydrostati model more, we have revised the sentence as follows:
"In general, a high-resolution non-hydrostatic atmospheric model has the advantage of simulating detailed meso-scale cloud structures, unlike a traditional hydrostatic atmospheric model."
----->
"In general, a non-hydrostatic atmospheric model can be run at much higher horizontal resolution (less than 10km, the limit of validity of the hydrostatic approximation) than a hydrostatic atmospheric model. Accordingly, a high-resolution non-hydrostatic atmospheric model has the advantage of simulating detailed meso-scale cloud structures, unlike a traditional hydrostatic atmospheric model. In light of recent evolution of supercomputers, it is inevitable to perform dynamical downscaling with a very high horizontal resolution, which allows us to consider effects of complex terrain like the GrIS margin on the atmospheric field explicitly." (P. 3, L. 83-89)

Regarding the latter comment, the 5km horizontal resolution was selected considering computational costs in the supercomputer of Meteorological Research Institute (Fujitsu PRIMEHPC FX100 and PRIMERGY CX2550M1). Now, the described model configuration faces a performance limit of the supercomputer. At the end of Sect. 2.3.1, we have added the following comment:
"At present, the above-mentioned domain setting faces a limitation imposed by practical computational costs in the supercomputer of Meteorological Research Institute (Fujitsu PRIMEHPC FX100 and PRIMERGY CX2550M1)." (P. 6, L. 221-223)

p.6, l.218 "increased with altitude from 40 m NEAR the surface to: : :"

OK. Revised as suggested. (P. 7, L. 235)

p.7, l.234: "for PRODUCING daily weather forecasts: : :"

The sentence has been corrected as suggested. (P. 7, L. 251)

p.9, l.307: add that PROMICE data were also used for validating 1x1-km statistically downscaled SMB based on ERA-I reanalysis data (Wilton et al. 2017, reference as above).

Thank you for the comment. We have added the explanation as follows:
"Recently, SMB data from PROMICE were used for the validations of MAR (Fettweis et al., 2017),

and the 1km horizontal resolution GrIS SMB product statistically downscaled from the daily output of RACMO2.3 (Noël et al., 2016) and ERA-Interim (Wilton et al., 2017).” (P. 9, L. 327-330)

p.9, l.322 “were superior on average” – quantify by how much and say whether statistically significant.

OK. We have indicated differences in ME and RMSE from on-line and off-line simulations. In addition, significance of these differences are explained by utilizing the p-value. Now the updated sentence is as follows:
“Average ME and RMSE at all sites were improved for the on-line simulation by 1.4 °C ($p < 0.01$) and 0.7°C ($p < 0.1$), respectively.” (P. 10, L. 344-346)

p.9, l.324: “ME was WITHIN 2.3˚°C at all sites”.

Corrected as suggested. (P. 10, L. 347)

p.10, l.338: change comma to colon.

Corrected as suggested. (P. 10, L. 362)

p.10, l.354: “except for Summit” – why the difference there?

At present, we have no idea why the difference was made at Summit; however, it should be noted that ME and RMSE are still reasonable when they are compared against those obtained at other sites (Table S2). We have added the following explanation:
“Even at Summit, ME and RMSE were still reasonable when they were compared against those obtained at other sites (Table S2). The reason why R^2 at Summit was relatively low should be investigated in the future.” (P. 11, L. 380-382)

p.10, l.359 add the relevant reference Orr et al. (2005):
Orr, Andrew and Hanna, Edward and Hunt, Julian C. R. and Cappelen, John and Steffen, Konrad and Stephens, Ag (2005) Characteristics of stable flows over Southern Greenland. Pure and Applied Geophysics, 162 (10). pp. 1747-1778. ISSN: 0033-4553

Thank you for introducing the paper. The suggested reference has been added. (P. 11, L. 387) p.14, l.11 at end of sentence suggest to add “Moreover, Wilton et al. (2017) show generally favourable results from a 1x1-km statistical downscaling of reanalysis data, with results generally comparing well with MAR and RACMO RCM output.“.

OK. We have added the following sentence as suggested:
“Moreover, Wilton et al. (2017) showed generally favourable results from a 1km statistical downscaling of reanalysis data, with results generally comparing well with MAR and RACMO RCM output.” (P. 16, L. 574-576)

p.16, l.605 after “statistical downscaling or further dynamical downscaling”, add “to a higher spatial resolution than used here, e.g. 1 km (Noel et al. 2016, Wilton et al. 2017): : :”.

OK. The suggested explanation has been added. (P. 18, L. 681-684)

p.27, Table 3: Suggest giving mean values in new row at bottom of table.

Thank you for the constructive suggestion. We have added a new row indicating mean values. In addition, tables in the supplementary file has been updated in the same manner.

Reply to Reviewer#2

We sincerely appreciate the reviewer for taking the time to provide valuable comments and suggestions. Below we describe our responses (in blue text) point-by-point to each comment (in black text). In addition, we indicate revisions in the updated manuscript together with new line numbers.

Summary

This paper introduces a new regional climate model for use over the Greenland ice sheet. The scientific impact is modest, as a) the modelled period is relatively brief, b) there clearly are issues that need to be addressed and c) the model data are not used for improved process understanding. But I presume the authors will at a later stage start using the model for these purposes. The technical quality of the figures is good, as are readability and length (apart from the last section, see below).

Thank you for the comment. The main purpose of this paper is to present a new regional climate model for Greenland. Owing to constructive comments and suggestions provided by all the reviewers, we believe the scientific impact of the revised manuscript has been increased. Now, a long-term climate simulation by NHM-SMAP is ongoing. Obtained results will be presented in the future.

Major comments

l. 166: it is unclear what the physical basis is of a parameterization of ice albedo as a function of density. Ice has a near-constant density?

In the current model, ice albedo is set to 0.55 when surface density is 830 kg m^{-3}, and assumed to decrease into 0.45 that is assigned when surface density is 917 kg m^{-3}. The sentence has been revised as follows:

“The albedo of ice was calculated by a linear equation as a function of density and ranged from 0.55, the typical albedo of clean firn (Cuffey and Paterson, 2010), to 0.45, taken from the MAR model setting as explained by Alexander et al. (2014).”

----->

“The albedo of ice was calculated by a linear equation as a function of density and ranged from 0.55 for a surface density of 830 kg m^{-3}, the typical albedo of clean firn (Cuffey and Paterson, 2010), to 0.45 for a surface density of 917 kg m^{-3}, taken from the MAR model setting as explained by Alexander et al. (2014). ” (P. 5, L. 169-172)

Section 2.2.3 explains how drifting snow sublimation at 2 m is calculated. But what is done with this information? Is a vertical sublimation profile assumed to calculate column blowing snow sublimation? Is the moisture source included in the atmospheric moisture conservation equation, i.e. is the additional water vapour used to moisten the boundary layer? What happens to surface sublimation when drifting snow sublimation starts? Please provide details to answer these questions.

Thank you for the comment. We have included the following description:
"In NHM-SMAP, surface mass loss due to drifting snow sublimation is assumed by Eq. (5); however, it is not used to moisten the boundary layer in the current version, because an interaction between the atmosphere and the snow/firn/ice surface is performed through the medium of albedo and surface temperature as mentioned later in Sect. 2.3.4." (P. 6, L. 201-204)

l. 197: once drifting snow transport is calculated, the erosion can be simply obtained by taking the divergence of the transport. It is unclear why the authors claim that this is computationally too expensive? If it is not taken into account, the surface mass balance is locally not closed, this must at least be mentioned.

We agree with reviewer that this is an important point for a model that calculates GrIS SMB. We have revised the sentence as follows:
"Although it is ideal to calculate the erosion of drifting snow (redistribution of near-surface snow caused by drifting snow), it was neglected in NHM-SMAP because of computational costs."
----->
"Although it is ideal to calculate the erosion of drifting snow (redistribution of near-surface snow caused by drifting snow), tracking changes in physical conditions of snow particles (prognostic variables of SMAP, namely, snow grain size, grain shape, density, and so on) during a drifting snow event and redistributing them in an updated surface field demands substantial computational costs. Therefore, the current version of NHM-SMAP neglects this process, which implies that simulated SMB is not closed locally." (P. 6, L. 205-210)

l. 210: "Ice sheet area minimum" suggests that ice sheet mask is not constant in time?

Our ice sheet mask is constant in time. The original description might cause misunderstanding, therefore, it has been revised as follows:
"The ice sheet mask for the GrIS was based on Bamber et al. (2001) as updated by Shimada et al. (2016) from 2000 to 2014, including the ice sheet area minimum of summer 2012, on the basis of MODIS satellite images."

----->

“The ice sheet mask for the GrIS, which is constant in time, was based on Bamber et al. (2001) as updated by Shimada et al. (2016) on the basis of 2000 to 2014 MODIS satellite images.” (P. 6, L. 223 – P. 7, 225)

Section 3.2: How did the authors deal with the mismatch between SMB observation and model period?

We referred the metadata of PROMICE SMB data and comprehended observation period. The NHM-SMAP calculated SMB data at each PROMICE site were retrieved during the exact measurement period. It is mentioned even in the original manuscript (at the end of Sect. 3.2).

Fig. 3: There is a systematic and considerable underestimation of LWin of up to 50 W m-2, which should lead to too low surface temperature, yet the snow surface temperature is overestimated in the model. I cannot reconcile this?

In the original manuscript, we mentioned possible causes for the discrepancy in terms of only insufficiencies of the model. However, we think there is also a problem in the measurement data. In the revised manuscript, we have discussed the issue as follows:
“On the other hand, observation data for downward longwave radiant flux can also have error especially during the winter period due to riming, which may act to increase measured values. In SIGMA-A, measured 2m air temperature often decreased to about –40 °C during the 2013-2014 winter (Fig. 3a). Although such reductions in 2m air temperature during March and April 2014 were followed by significant reductions in downward longwave radiant flux (Fig. 3e), they did not synchronize in December 2013 and January 2014. These results suggest that observed downward longwave radiant flux especially during December 2013 and January 2014 were affected by riming and forced to increase. A reliable quality control technique for automatic downward longwave radiant flux measurements in the polar region should be developed in the future to perform not only model validation but also climate monitoring accurately.” (P. 12, L. 438–447)

In the summary and conclusions section, an additional summary regarding this issue has been added as follows:
“On the other hand, observation data for downward longwave radiant flux can also have error especially during the winter period due to riming, which might affect the evaluation.” (P. 18, L. 656-657)

During the revision, we performed additional data quality control for downward longwave radiant flux. What we performed is that rejecting such data as downward and upward longwave radiant fluxes agree exactly. This situation is caused when extreme riming occurs and these two properties are diagnosed only from sensor temperature. However, our discussion was not affected by the reassessment of measurement data.

The summary and conclusions section can be written up much more concisely: just list the main conclusions.

The first paragraph of the summary and conclusions section have been updated as follows:
"We developed the NHM-SMAP polar RCM, with 5km resolution and hourly output, to reduce uncertainties in SMB estimates for the GrIS. Combining JMA's operational non-hydrostatic atmospheric model JMA-NHM and the multi-layered physical snowpack model SMAP, it is an attempt to take advantage of both short-term detailed weather forecast models and long-term computationally stable climate models. Model output data from NHM-SMAP hold promise for assessing not only long-term climate change in the GrIS, but also detailed diurnal variations of meteorological, snow, firn, and ice conditions in the GrIS. We initialized the atmospheric profile every day by referring to JRA-55 (weather forecast mode) to minimize deviations between the JRA-55 and NHM-SMAP atmospheric fields, while simulating the physical states of snow/firn/ice without any initialization (climate simulation mode). The model, forced by the latest Japanese reanalysis data JRA-55, was evaluated in the GrIS during the 2011–2014 mass balance years using in situ data from the SIGMA, GC-Net, and PROMICE AWS networks, PROMICE SMB data, and ice core data from SIGMA-D and SE-Dome. After updating SMAP by incorporating physical processes for new (polar) snow density, ice albedo, and effects of drifting snow, we validated NHM-SMAP in terms of hourly 2m air temperature, 2m water vapor pressure, surface pressure, 10m wind speed, downward shortwave and longwave radiant fluxes, snow/firn/ice surface temperature and albedo, surface height change, daily melt area extent, and the GrIS accumulated SMB."
----->
"We developed the NHM-SMAP polar RCM, with 5km resolution and hourly output, to reduce uncertainties in SMB estimates for the GrIS. Combining JMA's operational non-hydrostatic atmospheric model JMA-NHM and the multi-layered physical snowpack model SMAP, it is an attempt to take advantage of both short-term detailed weather forecast models and long-term computationally stable climate models. The model, forced by the latest Japanese reanalysis data JRA-55, was evaluated in the GrIS during the 2011–2014 mass balance years using in situ data from the SIGMA, GC-Net, and PROMICE AWS networks, PROMICE SMB data, and ice core data from SIGMA-D and SE-Dome." (P. 17, L. 613-624)

Minor and textual comments

l. 167: clean firn -> clean ice

I checked Cuffey and Paterson (2010) again, and confirmed this description is correct. In the book, albedo for clean ice is recommended to be 0.35.

Figure 1: ice mask in Canadian Arctic looks funny.

It is true we did not examine ice mask in Canadian Arctic sufficiently, because we focus the GrIS SMB in the present study. In the revised manuscript, we have mentioned this as follows:
"In the Canadian Arctic Archipelago, considerations for details in the ice sheet mask were nod given in the present study, because we focused the GrIS SMB. Therefore, there is room for improvement on the modelled ice sheet mask, which is a future issue for NHM-SMAP." (P. 7, L. 228-230)

In connection with this point, we recognized that a resolution of Fig. 1 was not enough. Therefore, the quality of Fig. 1 has been improved in the revised manuscript.

l. 287: Why was downward longwave radiation not used from PROMICE stations?

Downward longwave radiation data from PROMICE stations are used even in the original manuscript. Model performance at each PROMICE station are indicated in Table S5. At GC-Net stations, downward longwave radiation data were not employed in the present study, because they were not measured directly during the study period.

l. 320: Why is T2m "the most important climate parameter"? Better to leave out.

Thank you for the comment. We have deleted the sentence as suggested. (P. 10, L. 342)

l. 473: surface melt -> surface melt extent

It is an important point. We have revised as suggested. (P. 14, L. 514)

l. 478: "were almost the same" This is not very scientific. Please quantify or leave out. The same is true for the discussion in lines 518-520, please provide numbers.

Regarding the former comment, we have revised the sentence as follows:
“The basic geographic patterns of accumulation and ablation simulated for the 2011–2012, 2012–2013, and 2013–2014 mass balance years (Fig. S1) were almost the same as the annual mean SMB map created by RACMO2.3 (Noël et al., 2016).”
----->
“The geographic patterns of accumulation and ablation simulated for the 2011–2012, 2012–2013, and 2013–2014 mass balance years simulated by NHM-SMAP are depicted in Fig. S2.” (P. 14, L. 521-523)

As for the latter comment, we revised the manuscript by referring to the MAR model data provided by Xavier Fettweis (Reviewer #3), and now the description has been updated as follows:
“van den Broeke et al. (2016) reported that in estimates by RACMO2.3, SMB for the GrIS reached its lowest value since 1958 in 2012, then increased greatly in 2013 and decreased slightly in 2014. Our model produced a similar sequence in those years, with accumulated SMBs at the end of each mass balance year of –23, 420, and 312 Gt year^{-1}, respectively (Fig. 9a).”
----->
“According to simulation results by MAR v3.5.2 forced by JRA-55 (Fettweis et al., 2017), the GrIS SMB during the 2011-2012 mass balance year was relatively low (147 Gt year^{-1}), then increased greatly in 2012-2013 (473 Gt year^{-1}) and decreased slightly in 2013-2014 (403 Gt year^{-1}). Our model, which tends to simulate lower SMB compared to MAR v3.5.2 that uses the bucket schemes with an irreducible water content of 8 %, produced a similar sequence in those years, with accumulated SMBs at the end of each mass balance year of –23, 420, and 312 Gt year^{-1}, respectively (Fig. 9a).” (P. 16, L. 587-593)

Reply to Xavier Fettweis (Reviewer#3)

We sincerely appreciate Xavier Fettweis for taking the time to review our paper and providing insightful comments and suggestions as well as the MAR model output data as reference information. Below we describe our responses (in blue text) point-by-point to each comment (in black text). In addition, we indicate revisions in the updated manuscript together with new line numbers.

This paper presents a new RCM based simulation over the Greenland ice sheet. While the scientific interest of this paper is generally poor, this "model validation" paper deserves to be published in TC and opens the door to future applications over the GrIS using a new RCM in addition to the wide commonly used RCMs family (MAR, RACMO, HIRHAM). In addition to the justified remarks from both other reviewers, I have additional remarks that should be resolved before publication if it is not a too big job for the authors.

Thank you for the comment. Thanks to insightful comments and suggestions provided by all the reviewers, we believe the manuscript has been improved and scientific quality of the revised paper has been increased.

pg2, line 67: site rather Fettweis et al. (2017) here

Revised as suggested. (P. 2, L. 67)

pg5, section 2.2.1: What is the sensibility of the model results to the fresh snowfall density? With MAR, the sensibility is very small and MAR uses a minimum snowfall density of 200kg/m3. 300kg/m3 is a bit high for me.

Thank you for the comment. In fact, NHM-SMAP's sensitivity to the fresh snowfall density has not been investigated yet. The reason why we used the parameterization by Lenaerts et al. (2012a) is simple: this is based on in-situ measurements in polar region. If smaller fresh snowfall density is set in NHM-SMAP, underestimation of snow surface height discussed in Sect. 4.5 can be solved; however, I think we don't have enough measurement-based information for fresh snowfall density to change the model scheme now.

MAJOR: pg 7, line 231: As the JRA-55 surface conditions are bad (Section 4.1, line 325), is an atmospheric spin-up of 6h enough to be independent of the initial near surface atmospheric conditions? How are the results sensitive to this spin-up time? For me, performing 48h long simulations by keeping only the last 24h will be more robust.

Please note that insufficient conditions of JRA-55 surface analysis was unraveled through the present study. In addition, it should be noted that an appropriate spin-up period has not been established yet. An appropriate spin-up period can be found by performing a large number of simulations. The reason why we employed 6h spin-up time in the present study is that it is a typical model configuration in Japan. However, we agree with the point that further consideration of an atmospheric spin-up time can be effective to improve the model performance. The 6h spin-up period might not be suitable in the GrIS, although the setting seems to be effective empirically in Japan. In Sect. 4.1, we have added the following discussion:
"This result in turn suggests that making every day atmospheric spin-up period (6h; Sect. 2.3.2) longer than 6h can improve the performance of NHM-SMAP. Finding an appropriate spin-up period in the GrIS is a future issue to be coped with." (P. 10, L. 349-352)

pg9, section 4.1: As SMAP seems to underestimates the ablation (see Fig 8), the statistics over summer (JJA) should be provided at least in supplementary material? Is the model too warm or too cold in summer?

Thank you for the constructive comment. In Sect. 4.7 entitled as "Surface mass balance", we have added the following discussion:
"As presented in Sect. 4.1, the on-line version of NHM-SMAP successfully reproduced 2m air temperature at SIGMA-A during summer. Because surface mass loss during the summer is affected by near-surface (2m) temperature, model performance in terms of simulating JJA 2m air temperature at each AWS on the GrIS were re-examined (Table S8). As indicated in the table, significant or systematic error were not found, and obtained ME and RMSE were well (around –0.2 and 2.1 °C, respectively). Therefore, ---" (P. 15, L. 560-564)

In connection with point, a description referring to tables in the supplementary file at the beginning of Sect. 4 has been modified as follows:
"(see Table 3 and supplementary Tables S1 to S7)"
----->
"(see Table 3 and supplementary Tables S1 to S8)" (P. 10, L. 338 - 339)

MAJOR: pg 10, line 341: If a RCM is totally free, it should be normally independent of the surface biases in the forcing fields. A too short spin-up time of 6h starting from too warm JRA-55 based surface conditions explains likely these biases because MARv3.5.2 forced by JRA-55 is colder in winter than MARv3.5.2 forced by ERA-Interim. Therefore, extending the spinup time should better resolve this bias than changing of forcing reanalysis. Finally, SMAP seems to underestimate LWD in winter but overestimates temperature? This is very strange?? This issue should be discussed in the paper.

We think that a RCM cannot be totally free, because RCM-simulated atmospheric field is generally constrained by a parent reanalysis data in lateral and upper boundaries of the RCM model domain.

By the way, we imagine that simulated atmospheric field can be "almost" independent of the forcing data if we employ the "climate simulation mode", where the atmosphere is initialized only at the beginning of the simulation period, as employed by MAR. It seems to us that the present NHM-SMAP model configuration called "weather forecast mode" that initializes the atmospheric profile every day by referring to the forcing data is affected strongly by the parent data compared to the climate simulation mode. Based on this consideration, we agree with the reviewer's point that extending the spin-up time can resolve the reported bias. We have added the following discussion:
"At the same time, extending the atmospheric spin-up period discussed above can also resolve the issue, because simulation results are expected to less susceptible to a parent reanalysis data." (P. 10, L. 366- 368)

In the summary and conclusions section, it is mentioned again as follows:
"At the same time, extending the atmospheric spin-up period (6h) can also resolve the issue, because simulation results are expected to less susceptible to a parent reanalysis data." (P. 17, L. 641-642)

Regarding the last comment (sorry, we did not answer this comment during the open discussion period), we already discussed on this issue even in the original manuscript as follows:
"Figure 3a displays a year of observed and modelled 2m air temperature at SIGMA-A, from 1 September 2013 to 31 August 2014. The observed seasonal cycle was well reproduced by NHM-SMAP (R^2 = 0.95; Table 3); however, overestimation of the model was especially evident during winter (November to March), when measured 2m air temperature sometimes reached below –30 °C; this characteristic was found at all sites. The scatterplot of measurements versus model simulations for the whole study period at SIGMA-A (Fig. 3b) also displays this tendency. A possible reason for this discrepancy is that JRA-55 overestimates the surface temperature. The JMA Climate Prediction Division (CPD), which operationally develops JRA-55 data, recognizes that JRA-55 tends to overestimate winter surface air temperature in the polar region owing to inadequate treatment of energy exchanges between the atmosphere and the snow/firn/ice surface, especially under very

stable atmospheric conditions: a failure that also affects the reproducibility of the surface inversion layer and results in underestimation of the lower tropospheric temperature (S. Kobayashi, personal communication).” (P. 10, L. 353-364)

Underestimation of the lower tropospheric temperature leads to underestimation of downward longwave radiant flux. Based on this consideration, we already mentioned the cause of underestimation of downward longwave radiant flux in the original manuscript as follows:
“This characteristic was also found at other sites. One possible reason for this discrepancy is that the parent JRA-55 underestimates lower tropospheric temperatures, especially during winter (see Sect. 4.1).” (P.12, L. 431-433)

pg 10, section 4.2 : I do not see the interest of showing here the ability of SMAP only to simulate a single wind event. Outputs from JRA-55 should be added in the comparison to show the interest of SMAP in respect to JRA-55. MARv3.5.2 (at a resolution of 20km) forced by JRA55 underestimates also this event by a factor of 10-15m/s. The interest of using a non-hydrostatic model at 5 km should be highlighted here.

Thank you for the comment. We have included 10m wind speed data from JRA-55 in Fig. 4 and added the following discussion:
“In the figure, 10m wind speed from the parent JRA-55 reanalysis with a horizontal resolution of TL319 (~55 km) is depicted together. Clearly, JRA-55 could not reproduce the strong wind event and an advantage of a high-resolution non-hydrostatic atmospheric model is successfully demonstrated.” (P. 11, L. 395-398)

In connection with this point, we thought horizontal resolution of JRA-55 should be mentioned in Sect. 2.3.2: “Dynamical downscaling of atmospheric field from reanalysis data with JMA-NHM”. Therefore, it has been described in Sect.2.3.2 as follows:
“Horizontal resolution of JRA-55 is TL319 (~55 km).” (P. 7, L. 237)

pg 12, lines 409-423: the fact that SMAP overestimates surface temperature but underestimates both LWD/SWD fluxes suggests that SMAP is likely too dependent of the forcing data. What about the latent and sensible heat fluxes? The authors suggests that near-surface snow density is likely too high. I am very sceptic about this explanation. The sensibility of the results to the near-surface snow density can be tested offline. For me, the problem comes from the JRA-55 fields which are too warm and which are used every day to reinitialise the SMAP atmospheric fields.

Thank you for the insightful comment. First of all, regarding the underestimation of downward longwave radiant flux, we think that observation data also has error that affects model evaluation significantly. At the end of Sect. 4.3, we have added the following discussion:
"On the other hand, observation data for downward longwave radiant flux can also have error especially during the winter period due to riming, which may act to increase measured values. In SIGMA-A, measured 2m air temperature often decreased to about –40 °C during the 2013-2014 winter (Fig. 3a). Although such reductions in 2m air temperature during March and April 2014 were followed by significant reductions in downward longwave radiant flux (Fig. 3e), they did not synchronize in December 2013 and January 2014. These results suggest that observed downward longwave radiant flux especially during December 2013 and January 2014 were affected by riming and forced to increase. A reliable quality control technique for automatic downward longwave radiant flux measurements in the polar region should be developed in the future to perform not only model validation but also climate monitoring accurately." (P. 12, L. 438- 447)

In the summary and conclusions section, an additional summary regarding this issue has been added as follows:
"On the other hand, observation data for downward longwave radiant flux can also have error especially during the winter period due to riming, which might affect the evaluation." (P. 18, L. 656 - 657)

During the revision, we performed additional data quality control for downward longwave radiant flux. What we performed is that rejecting such data as downward and upward longwave radiant fluxes agree exactly. This situation is caused when extreme riming occurs and these two properties are diagnosed only from sensor temperature. However, our discussion was not affected by the reassessment of measurement data.

Based on these, we now agree with the reviewer's point that the problem comes from the JRA-55 fields which are too warm and which are used every day to reinitialize the SMAP atmospheric fields. At the same time, overestimation of relatively low surface wind speeds (Sect. 4.2) might affect this problem, because it acts to increase sensible heat flux. As a result, we have revised the sentence as follows:
"One possible cause of the model's overestimation of surface temperature is overestimation of the near-surface snow density profile, which would increase the conductive heat flux to the surface (see Sect. 4.5)."
----->
"One possible cause of the model's overestimation of surface temperature is overestimation of the

surface wind speeds when they are relatively low (see Sect. 4.2), which acts to heat the surface through increases in sensible heat flux. Of course, overestimation of 2m temperature by the model (see Sect. 4.1) especially during winter (November to March) also may contribute to the error." (P. 13, L. 459-464)

Related to this revision, the following description in the summary and conclusions section has been revised as follows:
"A possible cause for this overestimate is overestimation of the near-surface density profile, as suggested by validation of snow surface height changes."
----->
"A possible cause for this overestimate is overestimation of the surface wind speeds when they are relatively low, which acts to heat the surface through increases in sensible heat flux. In addition, overestimation of 2m temperature by the model especially during winter (November to March) also may contribute to the error." (P. 18, L. 662-666)

pg 12, lines 424-439: it is true that MAR overestimates albedo but as it also overestimates SWD. Due to error compensations (as explained in Fettweis et al., 2017), the MAR surface fields are OK. Here, it is strange that SMAP overestimates temperatures but overestimates albedo and underestimates SWD and LWD.

In the original manuscript, we mentioned two possible reasons for the overestimation of albedo by NHM-SMAP as follows:
"The dark microbe-rich sediment called cryoconite significantly reduces the surface albedo in the ablation area (Takeuchi et al., 2014; Shimada et al., 2016). Therefore, future models should consider this process as well as the possibility that NHM-SMAP overestimates snowfall during the summer period. In any case, it is necessary to conduct in situ measurements in the ablation area to confirm what is happening in reality." (P. 13, L. 476-480)

In the revision process, we conducted additional model sensitivity tests where ice albedo is set to 0.2 following the suggestion by the reviewer, which is detailed below. The results from the sensitivity tests indicate that simulated SMB did not change significantly compared to the control RE setting (Fig. 8). Based on the result, we reached a conclusion that overestimation of surface albedo by NHM-SMAP can be attributed mainly to overestimates of snowfall. These results are mentioned in Sect. 4.7 entitled as "Surface mass balance", and they can also be found in this answer file (our answer to "pg 14, line 513:").

Pg 13, section 4.6 : the comparison with the melt extent is excellent! Adding here a 2D comparison (nbr of melt days in 2012 for example) should be interesting to evaluate if this agreement is also OK locally. The simulated total melt extent could be good for bad reasons and local overestimation/underestimation of melt can be compensated.

Thank you very much for the encouraging comment. In Fig. S1 of the supplementary file, we have added the 2D comparison figure. At the end of Sect. 4.6, we have added the following explanation regarding the figure:
"Figure S1, which shows observed and simulated total numbers of surface melt days in 2012, supports this argument." (P. 14, L. 514 - 515)

pg 13, line 479: A 2D comparison with other RCM based estimations (RACMO, MAR, ...) is needed here for me. The raw 20km MARv3.5.2 daily outputs forced by JRA55 are available here:
ftp://ftp.climato.be/fettweis/MARv3.5.2/Greenland/JRA-55_20km/
and could be used in this paper just by citing Fettweis et al. (2017).

Thank you for the suggestion and providing the data. We considered whether we should use other RCM based estimations or not, and decided to include simulation results by MAR v3.5.2 forced by the same reanalysis data JRA-55 as used in the present study. At the beginning of Sect. 4.7, we have indicated it as follows:

"In addition, simulated SMB data from MAR v3.5.2 forced by JRA-55 (Fettweis et al., 2017) were employed as reference information." (P. 14, L. 519-521)

At present, there are many different points in model formulations and configurations of MAR and NHM-SMAP, namely, resolution, ice sheet mask, dynamic core of atmospheric model, albedo model, water percolation scheme for snow/firn, etc. Therefore, detailed model inter-comparison should be beyond the scope of this paper; however, we do hope to perform such a comparison in the near future.

MAJOR: pg 14, line 507: MAR at 20km is generally able to resolve the ablation zone. The 5 km resolution used here is not an issue here to explain the systematic SMB overestimation in the ablation zone by SMAP. RACMO at 11km works also already very well. Significant biases in energy balance fluxes could explain the underestimation of ablation.

Thank you for the comment. We think the reason why MAR at 20km successfully resolves the

ablation area is the introduction of a sub-grid mask, which is not considered by the present version of NHM-SMAP. Based on this consideration, we added the following discussion:
"On the other hand, MAR v3.5.2 with a horizontal resolution of 20km is generally able to resolve the ablation zone well (Fettweis et al., 2017). A possible cause for this success can be attributed to the introduction of sub-grid mask, which is not employed by NHM-SMAP. It appears that statistical downscaling or further dynamical downscaling or introduction of sub-grid mask is inevitable to obtain more realistic SMB estimates." (P. 16, L. 576-580)

Also, in the final section, we have mentioned it again as follows:
"Moreover, statistical downscaling or further dynamical downscaling to a higher spatial resolution than used here, e.g. 1 km (Noel et al. 2016, Wilton et al. 2017) or introduction of sub-grid mask (Fettweis et al., 2017) may inevitably be required to improve the SMB estimates." (P. 18, L. 681 - 684)

pg 14, line 513: to test the problem of the overestimation of albedo in SMAP, an offline simulation using a bare ice albedo of 0.2 could be carried out here and results should be shown in Fig 8.

It is a very nice suggestion. We have performed the suggested model sensitivity tests and discussed the results as follows:
"According to the PROMICE data in the ablation area, ice albedo often decreases to around 0.2 during summer. Therefore, additional model sensitivity tests, where ice albedo is set to 0.2, were performed. Obtained results indicate that simulated SMB did not change significantly compared to the control Richards equation setting (Fig. 8), suggesting that overestimation of surface albedo by NHM-SMAP can be attributed mainly to overestimates of snowfall as pointed out in Sect. 4.4." (P. 15, L. 566 – P. 16, L. 570)

In accordance with this, Fig. 8 has been updated. In the original manuscript, we did not refer Fig. 8 explicitly, therefore, it has been referred at the beginning of Sect. 4.7 as follows:
"During the study period, 55 measurements were available, and comparison results are presented in Fig. 8." (P. 14, L. 518 - 519)

Accordingly, the following sentence in the original manuscript (P. 14, L. 512-513 in the original manuscript) has been removed:
"Moreover, it is imperative that we develop a realistic albedo model for high-density firn and ice that incorporates the effects of cryoconite." (P. 16, L. 581-582)

Also, the following sentence in the original manuscript (P. 16, L. 592-594 in the original manuscript) has been removed as well:

"This finding underscores the need to develop a realistic albedo model for high-density firn and ice that allows us to consider the effects of darkening of the GrIS by cryoconite and so on.", (P. 18, L. 668-669)

and the following sentences has been added in the revised manuscript instead:

"It was attributed to overestimation of snowfall." (P. 18, L. 669 - 670)

"Resolving overestimation of snowfall by the model is also necessary." (P. 18, L. 680 –681)

pg 14, line 522, explicit comparison with MAR or RACMO is needed here for me. RACMO or MAR time series could be added in Fig 9.

As mentioned above, we have included simulation results by MAR v3.5.2 forced by JRA-55. In the revised manuscript, we have compared the SMB data with the NHM-SMAP-simulated GrIS SMB in Fig. 10a. The related description are as follows:

"According to simulation results by MAR v3.5.2 forced by JRA-55 (Fettweis et al., 2017) that uses the bucket schemes with an irreducible water content of 8 %, the GrIS SMB during the 2011-2012 mass balance year was relatively low (147 Gt year^{-1}), then increased greatly in 2012-2013 (473 Gt year^{-1}) and decreased slightly in 2013-2014 (403 Gt year^{-1}). Our model, which tends to simulate lower SMB compared to MAR v3.5.2, produced a similar sequence in those years, with accumulated SMBs at the end of each mass balance year of –23, 420, and 312 Gt year^{-1}, respectively (Fig. 10a)." (P. 16, L. 587 - 593)

pg 15, lines 532-540: such sensitivity to the irreducible water content is also simulated by MAR which uses a value of 8%.

Thank you for the information. The provided information has been included in the revised manuscript as mentioned in the previous answer.

Reply to Leo van Kampenhout

We sincerely appreciate Leo van Kampenhout for taking the time to review our paper. Below we describe our responses (in blue text) point-by-point to each comment (in black text). In addition, we indicate revisions in the updated manuscript together with new line numbers.

I agree with Xavier Fettweis that this work would be a welcome addition to the literature and the wider RCM modelling community. Some questions came up while reading the manuscript, in particular about the spinup method and the effect of percolation.

Thank you for the comment. We agree with the reviewer's point that we should detail more about the model spin-up and the effect of percolation. Please check our answers below.

L 238-240: I searched Dumont et al. (2014) for their spin-up procedure, but failed to find information on this. Did the authors obtain the method details through personal communication?

I am sorry the original description was incorrect. We have revised the sentence as follows:
"The initial snow/firn/ice physical conditions for the entire GrIS on 1 September 2011 were prepared by performing a 30year spin-up of the NHM-SMAP model following the procedure of Dumont et al. (2014)."
----->
"The initial top 30 m snow/firn/ice physical conditions for the entire GrIS on 1 September 2011 were prepared by performing a 30year spin-up of the NHM-SMAP model." (P. 7, L. 257-258)

L 238-240: I was wondering whether 30 years is sufficient to get a 30-m snowpack into equilibrium with the climate. Was there any remaining drift in e.g. the bottom layer temperature? What climate years were used to forced the spinup?

First, before starting model spin-up, we attempted prepare realistic initial profiles for snow/firn/ice physical conditions in the GrIS as much as possible. Thanks to this, we did not encounter any drift at the end of model spin-up. The performed procedure to prepare the initial conditions before the model spin-up has been described in the revised manuscript as follows:
"Before performing the model spin-up, the initial profiles for snow/firn/ice physical conditions in the GrIS were given following the procedure presented by Lefebre et al. (2005) and properties for snow/firn microstructure (e.g., optically equivalent grain size and grain shape) were given from the firn core analysis at SIGMA-A (Yamaguchi et al., 2014) equally in the GrIS." (P. 7, L. 259 – P. 8, L.

262)

As for climate years used to force the model spin-up, we used the data during the 2010-2011 mass balance year. First, we performed JMA-NHM stand-alone simulations forced by JRA-55 during the period. Then, the simulation results for surface atmospheric conditions forced SMAP 30 times cyclically (off-line calculation). Of course, it is ideal to perform continuous (not cyclic) 30year spin-up; however, it was not realistic due to computational costs. In the revised manuscript, it is described as follows:
"From the state, surface atmospheric conditions from September 2010 to August 2011 simulated by JMA-NHM forced by JRA-55 were used to drive SMAP for 30 times cyclically." (P. 8, L. 262-264)

Reference:

Lefebre, F., Fettweis, X., Gallée, H., Van Ypersele, J.-P., Marbaix, P., Greuell, W., Calanca, P.: Evaluation of a high-resolution regional climate simulation over Greenland, Climate Dynamics, 25, 99, doi:10.1007/s00382-005-0005-8, 2005.

L 242: It reads like you started with zero snow depth at the beginning of the spinup period. The zero heat flux is then assumed at the bottom of the snow pack, not at 30 m, which is almost never reached? (which you mention in 245-246)

We "did not" start with zero snow depth at the beginning of the spin-up period as mentioned above. During the simulation period, the thickness of snow/firn/ice is always constant: 30 m. It is mentioned in the revised manuscript as follows:
"The thickness of snow/firn/ice is always set to constant (30 m) during the calculation. In case snow accumulation or ablation is simulated, the thickness of the bottom model layer is modified accordingly." (P. 7, L. 255 –257)

L484-485: Figure 10 shows that runoff is larger for larger IWC value (6%), so the "piping" effect must be dominated by something else. Otherwise, the 2%-bucket model would have produced the largest runoff value.

We agree with the reviewer's point that 2%-bucket model setting can heat snow/firn effectively, then result in earlier onset of melting, which can produce larger runoff. However, in the sensitivity tests, we did not consider feedbacks that have more than a year time-scale due to our test setting. In the revised manuscript, we have added the following explanations regarding the setting of model sensitivity tests that changed water percolation schemes:

“In the sensitivity tests, profiles for snow/firn/ice physical conditions were reset at the beginning of the 2011–2012, 2012–2013, and 2013–2014 mass balance years by referring to the simulation data from the on-line version of NHM-SMAP. It means that feedbacks, which have more than a year time-scale, are not considered.” (P. 15, L. 534-537)

L 497-502: The authors do not supply any proof of their statement that the formation of ice layers is the reason for the increased runoff. In particular, they do not present melt and refreezing as separate terms. After the formation of (sub-surface) ice layers, one expects the melt to stay roughly the same order of magnitude, yet see a drop in refreezing due to the added effect of lateral runoff.

On the other hand, an increase in runoff could also occur due to increased melt. The reasoning is that when you have higher IWC and more refreezing, warmer snow will result which leads to stronger metamorphism and larger grains that lower the albedo. The warm snow also will persist throughout winter and helps to bring snow to the melting point in spring. This behaviour is also seen in other models. It would benefit this paper if light could be shed on the exact processes that are dominant in this study.

Thank you for the insightful comments and suggestion. Following the suggestion, we have included a figure showing melt and refreeze rates, which are monitored in NHM-SMAP operationally. In the revised manuscript, it is discussed as follows:
“To confirm the discussion, the GrIS-area-integrated daily melt and refreeze rates were investigated (Fig. 9). In the figure, results for the 2011-2012 mass balance year are shown, whereas results for other mass balance years are depicted in Fig. S3. During the 2011-2012 mass balance year, simulated daily melt rates were almost the same among the results from Richards equation scheme and two bucket schemes (Fig. 9a); however, refreeze rates from the control Richards equation scheme were much lower compared to other results (Fig. 9b), which is an evidence for the above-mentioned more impermeable ice in the results from Richards equation scheme. The same characteristics could be found in other mass balance years (Fig. S3).” (P. 15, L. 549-556)

What we found are basically the same as the reviewer’s recognition.

NHM-SMAP: Spatially and temporally high resolution non-hydrostatic atmospheric model coupled with detailed snow process model for Greenland Ice Sheet

Masashi Niwano[1], Teruo Aoki[2, 1], Akihiro Hashimoto[1], Sumito Matoba[3], Satoru Yamaguchi[4], Tomonori Tanikawa[1], Koji Fujita[5], Akane Tsushima[6], Yoshinori Iizuka[3], Rigen Shimada[7], and Masahiro Hori[7]

[1]Meteorological Research Institute, Japan Meteorological Agency, Tsukuba, 305-0052 Japan
[2]Graduate School of Natural Science and Technology, Okayama University, Okayama, 700-8530 Japan
[3]Institute of Low Temperature Science, Hokkaido University, Sapporo, 060-0819 Japan
[4]Snow and Ice Research Center, National Research Institute for Earth Science and Disaster Resilience, Nagaoka, 940-0821 Japan
[5]Graduate School of Environmental Studies, Nagoya University, Nagoya, 464-8601 Japan
[6]Research Institute for Humanity and Nature, Kyoto, 603-8047 Japan
[7]Earth Observation Research Center, Japan Aerospace Exploration Agency, Tsukuba, 305-8505 Japan

Correspondence to: Masashi Niwano (mniwano@mri-jma.go.jp)

Abstract. To improve surface mass balance (SMB) estimates for the Greenland Ice Sheet (GrIS), we developed a 5km resolution regional climate model combining the Japan Meteorological Agency Non-Hydrostatic atmospheric Model and the Snow Metamorphism and Albedo Process model (NHM-SMAP) with an output interval of 1 h, forced by the Japanese 55year Reanalysis (JRA-55). We used in situ data to evaluate NHM-SMAP in the GrIS during the 2011–2014 mass balance years. We investigated two options for the lower boundary conditions of the atmosphere, an "off-line" configuration using snow/firn/ice albedo and surface temperature data from JRA-55 and an "on-line" configuration using values from SMAP. The on-line configuration improved model performance in simulating 2m air temperature, suggesting that the surface analysis provided by JRA-55 is inadequate for the GrIS and that SMAP results can better simulate snow/firn/ice physical conditions. It also reproduced the measured features of the GrIS climate, diurnal variations, and even a meso-scale strong wind event. In particular, it reproduced the GrIS surface melt area extent well. Sensitivity tests showed that the choice of calculation schemes for vertical water movement in snow and firn has an effect as great as 200 Gt year^{-1} in the GrIS-wide accumulated SMB estimates; a scheme based on the Richards equation provided the best performance.

1 Introduction

In the Greenland Ice Sheet (GrIS), the second largest terrestrial ice sheet, a significant loss of ice mass has been occurring since the early 1990s (e.g., Rignot et al., 2008; van den Broeke et al., 2009~~).~~; Hanna et al., 2013; van den Broeke et al., 2016). Changes in the ice sheet mass (its mass balance, MB) are controlled by surface mass balance (SMB) and ice discharge across the grounding line (D), i.e., MB = SMB – D. The SMB component is related mainly to meteorological conditions and denotes the sum of mass fluxes towards the ice surface (precipitation) and away from it (runoff, sublimation, and evaporation). The Intergovernmental Panel on Climate Change's Fifth Assessment Report (IPCC AR5) (Vaughan et al., 2013) pointed out that SMB has decreased and discharge has increased at almost the same rates since the early 1990s (van den Broeke et al., 2009), accounting for the accelerated mass loss (Rignot et al., 2011). However, more recently the situation has changed drastically as mass loss has continued to increase. Enderlin et al. (2014) attributed 84 % of the increase in the GrIS mass loss after 2009 to increased surface runoff, which highlights the growing importance of SMB (see also Andersen et al., 2015; van den Broeke et al., 2016). Therefore, today, in situ measurements are of rising importance for monitoring changes in SMB as well as surface meteorological conditions.

Much effort has gone into monitoring surface weather conditions and SMB on the GrIS with in situ measurements. Steffen and Box (2001) established the Greenland Climate Network (GC-Net) consisting of 18 surface automated weather stations (AWSs), distributed mainly in the accumulation area. Ahlstrøm et al. (2008) built another AWS network as part of the Programme for Monitoring of the Greenland Ice Sheet (PROMICE), with stations distributed mainly in the ablation area. van den Broeke et al. (2008) constructed an AWS network in the K-transect, a stake array along the 67°N parallel in the south-western GrIS. Aoki et al. (2014a) installed two AWSs, Snow Impurity and Glacial Microbe effects on abrupt warming in the Arctic (SIGMA)-A and SIGMA-B, which are currently in operation in the northwestern GrIS. Regarding in situ SMB measurements, Machguth et al. (2016) compiled a large number of historical stake measurement data with a unified format, although the observations do not cover the entire GrIS. To fill geographic gaps, climate models have been developed that are constrained and calibrated by these in situ measurements. Once the validity of these models is confirmed on the basis of the in situ data, output from the models can be used for analysis of ongoing environmental changes around the entire GrIS. These models also enable us to perform present and future climate simulations for the GrIS, including the effects of ice mass loss on global sea level rise (e.g., Rignot et al., 2011).

Several physically based regional climate models (RCMs) ~~have been applied in the GrIS~~ (e.g., MAR: Fettweis, 2007; RACMO2: Noël et al., 2015; Polar MM5: Box, 2013; and HIRHAM5: Langen et al., ~~2015)~~2015) and statistically-downscaled meteorological reanalysis data (Hanna et al., 2005, 2011; Wilton et al., 2017) have been applied in the GrIS that have been found reliable in terms of reproducing current climate conditions (e.g., Fettweis, ~~2007~~2017; Hanna et al., 2011; Box, 2013; Fausto et al., 2016; van den Broeke et al., 2016) and simulating realistic future climate change (e.g., Franco et al., 2013). Nevertheless, considerable discrepancies can be found among the SMB components simulated by these models (Vernon et al., 2013), and uncertainties in the calculated SMBs are large compared to the uncertainties in ice discharge (Enderlin et al., 2014; van den Broeke et al., 2016). Regarding this situation, van den Broeke et al. (2016) pointed out that advances are imperative in two areas: improving the physics of SMB models and enhancing their horizontal resolution. As for the first area, the authors noted that current models poorly represent the effects of snow/firn/ice darkening, vertical and horizontal flow of meltwater in firn or over ice lenses, and the effect of liquid water clouds on the surface energy balance as well as the resulting melt. Regarding the second area, the authors argued the necessity of statistical and dynamical downscaling from RCM outputs.

In the present study, we constructed a high-resolution polar RCM called Non-Hydrostatic atmospheric Model–Snow Metamorphism and Albedo Process (NHM-SMAP), composed of atmospheric and snowpack models developed by the Meteorological Research Institute, Japan. We employed the Japan Meteorological Agency (JMA)'s operational non-hydrostatic atmospheric model JMA-NHM (Saito et al., 2006), with a high horizontal resolution of 5 km, for dynamical downscaling. ~~In general~~In general, a non-hydrostatic atmospheric model can be run at much higher horizontal resolution (less than 10km, the limit of validity of the hydrostatic approximation) than a hydrostatic atmospheric model. Accordingly, a high-resolution non-hydrostatic atmospheric model has the advantage of simulating detailed meso-scale cloud structures, unlike a traditional hydrostatic atmospheric model. In light of recent evolution of supercomputers, it is inevitable to perform dynamical downscaling with a very high horizontal resolution, which allows us to consider effects of complex terrain like the GrIS margin on the atmospheric field explicitly. We also utilized the detailed physical snowpack model SMAP (Niwano et al., 2012, 2014), which features a physically based snow albedo model (Aoki et al., 2011) and a realistic vertical water movement scheme based on the Richards equation (Richards, 1931; Yamaguchi et al., 2012). Combining high-resolution detailed atmospheric and snow models is a computational challenge that has limited previous efforts of this type (e.g., Brun et al., 2011; Vionnet et al., 2014). The purpose of this study was to assess the performance of the NHM-SMAP polar RCM in reproducing current GrIS atmospheric and snow/firn/ice conditions by utilizing in situ measurements. The chosen study period, September 2011 to August 2014, includes the record surface melt event that occurred during summer 2012 (Nghiem et al., 2012; Tedesco et al., 2013; Hanna et al., 2014). Using the data, NHM-SMAP was evaluated from various aspects, where 1 hour interval model output data were employed. Typical output data from this kind of RCM have a temporal resolution of 6 h to 1 day (Cullather et al., 2016). Therefore, this study was an attempt to take advantage of both short-term detailed weather forecast models and long-term computationally stable climate models. The success of our attempt may make model output data from NHM-SMAP valuable for assessing not only long-term climate change in the GrIS but also detailed diurnal variations of the meteorological, snow, firn, and ice conditions in the GrIS.

The purposes of this paper are to describe the NHM-SMAP polar RCM and to demonstrate its capacity to reproduce current GrIS atmospheric and snow/firn/ice conditions by utilizing in situ measurements. Section 2 of this paper describes the NHM-SMAP model in detail, and the in situ measurement data for surface meteorology and SMB we used in this study are introduced in Sect. 3. Section 4 presents the results of our validation analysis and discusses their implications for the future direction of NHM-SMAP's applications. Finally, in Sect. 5 we summarize our conclusions.

2 Model descriptions

2.1 Atmospheric model JMA-NHM

JMA-NHM employs flux form equations in spherical curvilinear orthogonal coordinates as the governing basic equations. Saito et al. (2006) demonstrated that JMA-NHM outperforms the JMA's previous hydrostatic regional model in predictions of synoptic meteorological fields and quantitative forecasts of precipitation. Although JMA-NHM is used mainly for operational daily weather forecasts around Japan, the model can also be used for long-term climate simulations (Murata et al., 2015). Recently, JMA-NHM was applied to support a field expedition in the GrIS (Hashimoto et al., 2017), and the model setting used on that occasion was followed in this study. A double-moment bulk cloud microphysics scheme was used to predict both the mixing ratio and concentration of solid hydrometeors (cloud ice, snow, and graupel), and a single-moment scheme was used to predict the mixing ratio of liquid hydrometeors (cloud water and rain). In addition, ice crystal formation in the atmosphere was simulated by using an up-to-date formulation that depends on temperature. Following Hashimoto et al. (2007), we did not employ the ice-saturation adjustment scheme and the cumulus parameterization used in the original configuration. The turbulence closure boundary layer scheme was formulated following the improved Mellor-Yamada Level 3 (Nakanishi and Niino, 2006). For atmospheric radiation, the transfer function in longwave radiation was computed by a random model developed by Goody (1952), and shortwave radiation was computed by diagnosing the transfer function following Briegleb (1992).

2.2 Physical snowpack model SMAP

The multi-layered physical snowpack model SMAP was developed for the seasonal snowy areas of Japan by Niwano et al. (2012, 2014). SMAP calculates the temporal evolution of broadband snow albedos in the UV-visible, near-infrared, and shortwave spectra as well as the internal physical parameters of snowpack such as temperature, density, grain size, and grain shape. Because the model incorporates the physically based snow albedo model (PBSAM) developed by Aoki et al. (2011), it can assess effects of snow grain size and impurity concentration (black carbon and dust) on snow albedo explicitly in principle. SMAP calculates vertical water movement in snow and firn by employing the detailed Richards equation (Richards, 1931; Yamaguchi et al., 2012). SMAP is also equipped with a bucket scheme to calculate vertical water movement in snow and firn, in which liquid water exceeding the maximum prescribed water content descends to the adjacent lower layer (Niwano et al., 2012). Because a bucket scheme is used in most existing polar RCMs (Reijmer et al., 2012), we investigated whether the Richards equation scheme improves the GrIS SMB (see Sect. 4.7).

Niwano et al. (2015) applied SMAP to the SIGMA-A site (Aoki et al., 2014b), on the northwestern GrIS, and demonstrated that when forced by the measured surface meteorological data, the model reproduced the temporal evolution of the physical conditions in near-surface snow (Yamaguchi et al., 2014) during the record surface melt event of summer 2012 (Nghiem et al., 2012; Tedesco et al., 2013; Hanna et al., 2014). The authors modified the original model settings only for the effective thermal conductivity of snow and the surface roughness length for momentum. In this study, we started with the same model settings described by Niwano et al. (2015). Because this was the first attempt to perform year-round regional simulations of the GrIS with SMAP, we were obliged to make adjustments for three snow/firn/ice physical processes: new snow density (density of falling snow), ice albedo, and effects of drifting snow.

2.2.1 New snow density

Previous studies have suggested that new snow density in the polar region exceeds 300 kg m^{-3} (Greuell and Konzelmann, 1994; Lenaerts et al., 2012a), whereas new snow density in mid-latitudes is typically around 100 kg m^{-3} (e.g., Niwano et al., 2012). For this study, we used the following parameterization for new snow density developed by Lenaerts et al. (2012a) in Antarctica:

$$\rho_{\mathrm{new}} = A + BT_{\mathrm{sfc}} + CU_{10\mathrm{m}}, \tag{1}$$

where ρ_{new} is the new snow density (kg m^{-3}), T_{sfc} is the surface temperature (K), $U_{10\mathrm{m}}$ is the 10m wind speed (m s^{-1}), and the coefficients were set at A = 97.5 kg m^{-3}, B = 0.77 kg m^{-3} K^{-1}, and C = 4.49 kg s m^{-4}. As an additional condition, the minimum and maximum values of ρ_{new} were set at 300 and 350 kg m^{-3} following Lenaerts et al. (2012a).

2.2.2 Ice albedo

Although the PBSAM snow albedo component in SMAP allows us to simulate snow albedo realistically, its present version cannot be applied to an ice surface because the optically equivalent grain size of high-density ice, an important input parameter, cannot be defined and calculated by SMAP. In this study, we calculated the albedos of snow and firn with the PBSAM snow albedo component, defining firn as snow with density between 400 and 830 kg m^{-3} following Cuffey and Paterson (2010). The albedo of ice was calculated by a linear equation as a function of density and ranged from 0.55 for a surface density of 830 kg m^{-3}, the typical albedo of clean firn (Cuffey and Paterson, 2010), to 0.45 for a surface density of 917 kg m^{-3}, taken from the MAR model setting as explained by Alexander et al. (2014).

2.2.3 Effects of drifting snow

Sublimation of drifting snow is an important contributor to the GrIS SMB (Lenaerts et al., 2012b). In SMAP, the drifting snow condition is diagnosed on the basis of a mobility index M_{O}, which describes the potential for snow erosion of a given snow layer, and a driftability index S_{I}. Following Vionnet et al. (2012), M_{O} is calculated by

$$M_{\mathrm{O}} = \begin{cases} 0.34(0.75d - 0.5s + 0.5) + 0.66F(\rho) & \text{for dendritic case} \\ 0.34(-0.583g_{\mathrm{s}} - 0.833s + 0.833) + 0.66F(\rho) & \text{for non-dendritic case'} \end{cases} \tag{2}$$

where d is dendricity, s is sphericity, ρ is snow density, and g_{s} is geometric snow grain size (mm). Here d describes the remaining portion of the original snow grains in a snow layer, and s is the ratio of rounded versus angular snow grains (Brun et al., 1992). These two parameters are calculated by SMAP as explained by Niwano et al. (2012). F as an empirical function of density is written as

$$F(\rho) = [1.25 - 0.0042(max(50, \rho) - 50)]. \tag{3}$$

Using M_O, S_I is diagnosed from the equation proposed by Guyomarc'h and Merindol (1998):

$$S_I = -2.868e^{-0.085U} + 1 + M_O, \tag{4}$$

where U is the 2m wind speed (m s^{-1}), and the value of U when S_I becomes 0 indicates the threshold wind speed U_t for the occurrence of drifting snow. Once the onset of the drifting snow condition is simulated by SMAP, the drifting snow sublimation rate F_s (kg m^{-2} s^{-1}) at 2 m above the surface is calculated following Gordon et al. (2006):

$$F_s = D\left(\frac{T_0}{T_a}\right)^{\gamma} U_t \rho_a q_{si} (1 - R_{Hi}) \left(\frac{U}{U_t}\right)^{E}, \tag{5}$$

where T_a is air temperature (K), T_0 is 273.15 K, ρ_a is air density (kg m^{-3}), q_{si} is saturation specific humidity with respect to ice at temperature T_a (kg kg^{-1}), and R_{Hi} is relative humidity with respect to ice. The dimensionless constants are D = 0.0018, γ = 4, and E = 3.6. In NHM-SMAP, surface mass loss due to drifting snow sublimation is assumed by Eq. (5); however, it is not used to moisten the boundary layer in the current version, because an interaction between the atmosphere and the snow/firn/ice surface is performed through the medium of albedo and surface temperature as mentioned later in Sect. 2.3.4.

Although it is ideal to calculate the erosion of drifting snow (redistribution of near-surface snow caused by drifting snow), ~~it was neglected in NHM-SMAP because of computational costs.~~tracking changes in physical conditions of snow particles (prognostic variables of SMAP, namely, snow grain size, grain shape, density, and so on) during a drifting snow event and redistributing them in an updated surface field demands substantial computational costs. Therefore, the current version of NHM-SMAP neglects this process, which implies that simulated SMB is not closed locally. Lenaerts et al. (2012b) reported that the contribution of drifting snow erosion to SMB is negligible on the GrIS; however, it is locally important, especially in areas where topographic features induce strong divergence or convergence in the wind field.

2.3 NHM-SMAP coupling simulation procedure

2.3.1 Model domain and ice sheet mask

The 5km horizontal resolution JMA-NHM outputs hourly values of surface meteorological properties including precipitation (snow and rain are discriminated internally), 2m air temperature, 2m relative humidity with respect to water, 2m and 10m wind speed, surface pressure, downward shortwave and longwave radiant fluxes, and cloud fraction in the calculation domain shown in Fig. 1. The model domain consists of 450 × 550 horizontal grid cells, each cell characterized as land, sea, snow and ice, or sea ice. At present, the above-mentioned domain setting faces a limitation imposed by practical computational costs in the supercomputer of Meteorological Research Institute (Fujitsu PRIMEHPC FX100 and PRIMERGY CX2550M1). The ice sheet mask for the GrIS, which is constant in time, was based on

Bamber et al. (2001) as updated by Shimada et al. (2016) ~~from 2000 to 2014, including the ice sheet area minimum of summer 2012,~~ on the basis of 2000 to 2014 MODIS satellite images. As a result, the modelled area of the GrIS and peripheral glaciers was 1.807 × 10^6 km^2, which agrees well with the estimate of 1.801 ± 0.016 × 10^6 km^2 by Kargel et al. (2012). The GrIS surface elevation was taken from Bamber et al. (2001). In the Canadian Arctic Archipelago, considerations for details in the ice sheet mask were nod given in the present study, because we focused the GrIS SMB. Therefore, there is room for improvement on the modelled ice sheet mask, which is a future issue for NHM-SMAP.

2.3.2 Dynamical downscaling of atmospheric field from reanalysis data with JMA-NHM

We performed our high-resolution atmospheric calculation by using the dynamical downscaling approach. The model atmosphere used by JMA-NHM in this study had a top height of about 22 km and included 50 grid cells in the vertical direction based on terrain-following coordinates. The vertical grid spacing increased with altitude from 40 m ~~at~~near the surface to 886 m at the top of the atmosphere. We used JRA-55 (Kobayashi et al., 2015) for the upper, lower, and lateral boundary conditions of the atmosphere. Horizontal resolution of JRA-55 is TL319 (~55 km). Simmons and Poli (2015) reported that the near-surface and lower-tropospheric warming of the Arctic over the past 35 years is well reproduced by JRA-55, very much like the European Centre for Medium-Range Weather Forecasts (ECMWF) Interim reanalysis (ERA-Interim) data (Dee et al., 2011). Surface physical properties, including albedo and temperature of land, sea, and sea ice, were taken from JRA-55 as the bottom boundary conditions of the atmosphere. As for those surface physical properties of snow and ice, our two options were possible: it was given from JRA-55 or SMAP (see Sect. 2.3.4).

Although it is possible for JMA-NHM to perform long-term climate simulations in "climate simulation mode", where the atmosphere is initialized only at the beginning of the simulation period (Murata et al., 2015), in this study we used the "weather forecast mode", initializing the atmospheric profile every day by referring to JRA-55. The purpose of this approach was to prevent large deviations between the JRA-55 and NHM-SMAP atmospheric fields. Therefore, every day a 30h long simulation was carried out starting from 1800 UTC of the previous day, and the model outputs of the last 24 h were employed after discarding output from the initial 6h spin-up period. This is the same procedure developed by Hashimoto et al. (2017) for producing daily weather forecasts for the GrIS.

2.3.3 SMAP calculation forced by results from JMA-NHM

We used SMAP, forced by the calculated surface meteorological data from the JMA-NHM, to simulate the temporal evolution of the top 30 m of snow, firn, and ice from September 2011 to August 2014. ~~The initial~~The thickness of snow/firn/ice is always set to constant (30 m) in the model during the calculation. In case snow accumulation or ablation is simulated, the thickness of the bottom model layer is modified accordingly. The initial top 30 m snow/firn/ice physical conditions for the entire GrIS on 1 September 2011 were prepared by performing a 30year spin-up of the NHM-SMAP model ~~following the procedure of Dumont et al. (2014).~~. Before starting the model spin-up, the initial profiles for snow/firn/ice physical conditions in the GrIS were given following the procedure presented by Lefebre et al. (2005) and properties for snow/firn microstructure (e.g., optically equivalent grain size and grain shape) were given from the firn core analysis at SIGMA-A (Yamaguchi et al., 2014) equally in the GrIS. From the state, surface atmospheric conditions from September 2010 to August 2011 simulated by JMA-NHM forced by JRA-55 were used to drive SMAP for 30 times cyclically. We restricted the number of vertical model layers in the snow/firn/ice to 40 to limit computational costs. The vertical grid spacing increased from 1 cm at the surface to around 10 m at the bottom. We assumed zero heat flux at 30 m depth. For mass flux, runoff was calculated when meltwater or rain reached impermeable ice (density higher than 830 kg m^{-3}) and saturated the layer above the impermeable ice. A slush layer was not allowed to form, and the runoff mass was removed from the GrIS instantaneously. When water reached 30 m depth and could not be retained, it was forced to run off immediately; however, this situation was quite rare during the study period.

Although the PBSAM component of the model allowed us to consider effects of snow impurities such as black carbon and dust explicitly, the relevant data were not available at high temporal resolution for the study period; therefore, we assumed a pure snow condition. Aoki et al. (2014b) examined published concentrations of black carbon in near-surface snow in the GrIS and noted that most were less than several parts per billion by weight (ppbw). Reducing the albedo of snow by 0.01 requires 40 ppbw of black carbon in new snow and 10 ppbw in old melting snow (Warren and Wiscombe, 1980). We concluded that the measured concentrations of black carbon in the GrIS would not reduce albedo in snow, except possibly in old melting snow. Therefore, the pure snow assumption is probably reasonable in the accumulation area of the GrIS. However, recent darkening of the GrIS (Shimada et al., 2016; Tedesco et al., 2016) has commanded attention. This effect is discussed in Sect. 4.4 and Sect. 4.7.

2.3.4 Interaction between the atmosphere and snow/firn/ice

In this study, we examined two configurations of the NHM-SMAP coupled model for the lower boundary condition of the atmosphere, using snow/firn/ice albedo and surface temperature from JRA-55 or from SMAP (Sect. 2.3.2). The on-line configuration (SMAP) allowed us to simulate the interaction between the atmosphere and the surface whereas the off-line configuration (JRA-55) treated only the one-way supply of energy and mass from the atmosphere. Bellaire et al. (2017) has used the data obtained at GC-Net stations to demonstrate that the off-line version yields sufficiently accurate input data for the detailed snow process model SNOWPACK (Lehning et al., 2002) to reproduce the measured near-surface snow density profiles at GC-Net stations.

2.3.5 Surface mass balance

Using NHM-SMAP, we calculated SMB, in meters of water equivalent (m w.e.), by the equation

$$SMB = P - SU_{\mathrm{s}} - SU_{\mathrm{ds}} - RU, \tag{6}$$

where P is precipitation, SU_{s} is sublimation or evaporation from the surface, SU_{ds} is sublimation from drifting snow particles, and RU is runoff. As mentioned in Sect. 2.2.3, we neglected drifting snow erosion to reduce computational costs.

3 Observational data

3.1 Surface meteorology and surface melt area extent

To validate NHM-SMAP, we employed hourly surface meteorological data obtained with the AWSs of SIGMA (Aoki et al., 2014a; Niwano et al., 2015), GC-Net (Steffen and Box, 2001; Box and Rinke, 2003), and PROMICE (Ahlstrøm et al., 2008; van As et al., 2012), as listed in Table 1 and shown in Fig. 2a. The properties we sought to validate were 2m air temperature, 2m water vapor pressure, surface pressure, 10m wind speed, downward shortwave and longwave radiant fluxes, snow/firn/ice surface temperatures, surface albedo, and snow surface height change. Our selection of AWSs was based on the availability of high quality data in adequate quantities during the study period and the elevation difference between the AWS site and the topographic model in NHM-SMAP (Sect. 2.3.1). To compare the in situ measurements and the NHM-SMAP results, we used modelled data for the grid cell nearest to each AWS. Differences in elevation were not corrected in NHM-SMAP, although elevation differences greater than 200 m were not allowed. From GC-Net stations, only 2m air temperature, surface pressure, 10m wind speed, and downward shortwave radiant flux were taken. From PROMICE stations, all the properties except for surface height change were acquired, and SIGMA stations provided all the properties. Because the sensor heights changed over time depending on accumulation and ablation, we calculated the 2m air temperature, 2m water vapor pressure, and 10m wind speed from the measurements by using the flux profile calculation module of SMAP (Niwano et al., 2012). Erroneous values were rejected after visual inspection, and temporal gaps left by the rejected data were not filled by interpolation.

For the extent of the surface melt area in the GrIS, we used the daily composite of satellite data developed by Mote (2007, 2014). This dataset, which was created from measurements by the Special Sensor Microwave Imager/Sounder (SSMIS), offers a daily record of surface and near-surface melting on the GrIS with 25km horizontal resolution. Hanna et al. (2014) utilized this dataset to evaluate recent changes in the GrIS melt area.

3.2 Surface mass balance

The SMB of the GrIS calculated by NHM-SMAP for the study period was evaluated by using data provided by PROMICE (Machguth et al., 2016) as well as ice core data from the SIGMA-D (Matoba et al., 2015) and SE-Dome (Iizuka et al., 2015) drilling sites (Table 2 and Fig. 2b). Most of the PROMICE stations are in the ablation area, whereas SIGMA-D and SE-Dome are in the accumulation area. Recently, SMB data from PROMICE were used for the validations of MAR (Fettweis et al., 2017~~)~~), and the 1km horizontal resolution GrIS SMB product statistically downscaled from the daily output of RACMO2.3 (Noël et al., 2016) and ERA-Interim (Wilton et al., 2017). The validation sites were selected on the same basis as AWSs: data availability and an elevation difference less than 200 m between the site and the model. By employing the provided information for measurement periods at each site, the NHM-SMAP calculated SMB for each exact corresponding period were retrieved.

4 Model validation results and discussion

In this section we present validation results of the 5km resolution hourly NHM-SMAP output for the GrIS using in situ data obtained from September 2011 to August 2014. We include detailed information for mean error (ME; the average of the difference between simulated and observed values), root mean square error (RMSE), and coefficient of determination (R^2) to assess the model performance (see Table 3 and supplementary Tables S1 to ~~S7~~S8). Sections 4.1 to 4.5 refer to hourly data from measurements and model simulations unless otherwise specified. Dates and times are expressed in UTC.

4.1 2m air temperature, 2m water vapor pressure, and surface pressure

~~The most important climatic parameter for this kind of polar RCM is 2m air temperature.~~ Table 3 lists the model performance for 2m air temperature during the study period at each AWS depicted in Fig. 2a. ~~Clearly,~~Average ME~~, RMSE,~~ and ~~R^2~~RMSE at all sites were improved for the on-line simulation ~~were superior to those for the off-line simulation at almost all sites~~by 1.4 °C ($p < 0.01$) and 0.7°C ($p < 0.1$), respectively. Notable overestimates by the model (ME reached 6.6 °C at Summit, for example) were corrected in the on-line configuration (ME was ~~less than~~within 2.3 °C at all sites). These results suggest that the surface analysis provided by JRA-55 is of inadequate quality in the GrIS and that SMAP improves the results through the use of more realistic snow/firn/ice physical conditions. This result in turn suggests that making every day atmospheric spin-up period (6h; Sect. 2.3.2) longer than 6h can improve the performance of NHM-SMAP. Finding an appropriate spin-up period in the GrIS is a future issue to be coped with. The following discussion focuses on results from the on-line simulation.

Figure 3a displays a year of observed and modelled 2m air temperature at SIGMA-A, from 1 September 2013 to 31 August 2014. The observed seasonal cycle was well reproduced by NHM-SMAP (R^2 = 0.95; Table 3); however, overestimation of the model was especially evident during winter (November to March), when measured 2m air temperature sometimes reached below –30 °C; this characteristic was found at all sites. The scatterplot of measurements versus model simulations for the whole study period at SIGMA-A (Fig. 3b) also displays this tendency. A possible reason for this discrepancy is that JRA-55 overestimates the surface temperature. The JMA Climate Prediction Division (CPD), which operationally develops JRA-55 data, recognizes that JRA-55 tends to overestimate winter surface air temperature in the polar region owing to inadequate treatment of energy exchanges between the atmosphere and the snow/firn/ice surface, especially under very stable atmospheric conditions~~,~~; a failure that also affects the reproducibility of the surface inversion layer and results in underestimation of the lower tropospheric temperature (S. Kobayashi, personal communication). Further investigation of this issue would require conducting further NHM-SMAP simulations forced by other reanalysis datasets like ERA-Interim, as done by Fettweis et al. (2017), which was beyond the scope of this study. At the same time, extending the atmospheric spin-up period discussed above can also resolve the issue, because simulation results are expected to less susceptible to a parent reanalysis data.

Tables S1 and S2 indicate statistics for the model performance in terms of 2m water vapor pressure and surface pressure. To summarize, R^2 for both parameters was acceptably high (more than 0.84), and ME and RMSE were reasonable. Relatively large biases and RMSE as well as relatively low R^2 were found for 2m water vapor pressure at sites TAS_U, QAS_L, and QAS_U. This result suggests that NHM-

SMAP forced by JRA-55 cannot adequately reproduce absolute water content in the southeastern GrIS. According to Hanna et al. (2006), the southeastern GrIS is characterized by high accumulation rates attributed to prevailing easterly winds, frequent cyclogenesis in and around Fram Strait, and relatively high moisture availability when source air originates over a warm ocean. Stations TAS_U, QAS_L, and QAS_U are very close to the margin of our model domain (Fig. 1). Therefore, the use of a larger model domain that includes all of Svalbard may improve model results by resolving frequent cyclone activity in and around Fram Strait. Surface pressure was well simulated by NHM-SMAP, because R^2 was very close to 1.0 except for Summit. Even at Summit, ME and RMSE were still reasonable when they were compared against those obtained at other sites (Table S2). The reason why R^2 at Summit was relatively low should be investigated in the future. The slightly larger ME and RMSE for surface pressure found at SIGMA-B, SCO_U, QAS_L, QAS_A, and NUK_U can be attributed to relatively large elevation differences between the actual topography and the topographic model (–165, 176, 85, 104, and 85 m, respectively), as indicated in Table S2.

4.2 10m wind speed

Orr et al. (2005) and Moore et al. (2016) pointed out that topographic flow distortion commonly induces high-speed low-level winds in the southern GrIS including tip jets, barrier winds, and katabatic flows. They also noted that an atmospheric model of Greenland would need a horizontal resolution of about 15 km to characterize the impact of topography on the regional wind field and climate; however, even at this resolution, features of the wind field would be under-resolved. Therefore, we investigated the reproducibility of a strong wind event observed at the TAS_U site (Fig. 2a) during the study period, when a maximum 10m wind speed of 46.9 m s^{-1} was recorded at 1700 UTC on 27 April 2013. A comparison of measured and simulated data (Fig. 4a) shows that the 5km resolution NHM-SMAP successfully reproduced the strong wind event but underestimated its maximum wind speed by about 5 m s^{-1}. In the figure, 10m wind speed from the parent JRA-55 reanalysis with a horizontal resolution of TL319 (~55 km) is depicted together. Clearly, JRA-55 could not reproduce the strong wind event and an advantage of a high-resolution non-hydrostatic atmospheric model is successfully demonstrated. A comparison of measured and modelled 10m wind speeds at TAS_U during the whole study period indicates that the model tended to underestimate high wind speeds (>30 m s^{-1}) but overestimated relatively low wind speeds, resulting in ME, RMSE, and R^2 of 2.5 m s^{-1}, 4.3 m s^{-1}, and 0.68, respectively (Fig. 4b). At other sites, absolute values for ME and RMSE were smaller than those at TAS_U, and R^2 ranged widely between 0.13 (SCO_U) and 0.78 (KAN_U) (Table S3).

These results confirm that it is difficult for atmospheric models to reproduce surface wind fields in the southern GrIS. This problem may be solved by updating the boundary layer scheme (Sect. 2.1) and increasing the horizontal resolution. In addition, a simple treatment of the surface roughness length for momentum (Niwano et al., 2015) also may affect surface wind speed estimates, as suggested by Amory et al. (2015). NHM-SMAP can provide synoptic weather data during strong wind events. Figure 4c, depicting the estimated surface wind speed field at 1700 UTC on 27 April 2013, shows that strong wind speeds were simulated near the southeastern margin of the GrIS. This surface strong wind event corresponds to the Køge Bugt Fjord katabatic flow reported by Moore et al. (2016).

4.3 Downward shortwave and longwave radiant fluxes

The downward shortwave and longwave radiant fluxes are important elements of the GrIS surface energy balance. During 30 June to 14 July 2012, Niwano et al. (2015) visited SIGMA-A (Fig. 2a) and witnessed the record surface melt event (Nghiem et al., 2012; Tedesco et al., 2013; Hanna et al., 2014). They reported mainly clear sky conditions until 9 July and cloudy conditions with occasional heavy rainfall after 10 July. NHM-SMAP successfully reproduced the observed temporal evolution and diurnal variation of downward shortwave radiant flux at SIGMA-A from 1 to 15 July; however, it tended to underestimate slightly when clouds appeared (Fig. 3c). This tendency was typical during the whole study period, as shown by Fig. 3d and the ME value of –13.5 W m^{-2} listed in Table S4, although the signs of ME differ from place to place. RMSE ranged from 56.0 W m^{-2} (KPC_U) to 127.3 W m^{-2} (KAN_L) and was close to values reported by Ohtake et al. (2013) when the operational version of JMA-NHM was validated using hourly data from Japan, and relatively accurate RMSEs were obtained in the northern GrIS (Table S4). The underestimation in cloudy conditions may arise from causes in the cloud radiation scheme or in the reproducibility of cloud amounts and types by the model.

Although the tendencies of ME for downward shortwave radiant flux vary from place to place, ME for the downward longwave radiant flux had a similar tendency across the GrIS, ranging from –25.1 W m^{-2} at SIGMA-A to –10.8 W m^{-2} at KAN_M (Table S5). Underestimates of downward longwave radiant fluxes at SIGMA-A were especially great during winter (November to January when observed values reached less than about 200 W m^{-2}) in the record from 1 September 2013 to 31 August 2014 (Fig. 3e) and over the whole study period (Fig. 3f). This characteristic was also found at other sites. One possible reason for this discrepancy is that the parent JRA-55 underestimates lower tropospheric temperatures, especially during winter (see Sect. 4.1). In addition, uncertainty in the winter cloud amount, low-level liquid clouds (Bennartz et al., 2013), and thin clouds (Cox et al., 2014) may affect the results. Improving the model would require detailed in situ measurements of cloud amount, cloud type, and atmospheric profiles as well as intercomparisons against satellite remote sensing data like that of Van Tricht et al. (2016). A model intercomparison like that done by Inoue et al. (2006) would also aid deeper understanding of the limitations of current polar RCMs. On the other hand, observation data for downward longwave radiant flux can also have error especially during the winter period due to riming, which may act to increase measured values. In SIGMA-A, measured 2m air temperature often decreased to about –40 °C during the 2013-2014 winter (Fig. 3a). Although such reductions in 2m air temperature during March and April 2014 were followed by significant reductions in downward longwave radiant flux (Fig. 3e), they did not synchronize in December 2013 and January 2014. These results suggest that observed downward longwave radiant flux especially during December 2013 and January 2014 were affected by riming and forced to increase. A reliable quality control technique for automatic downward longwave radiant flux measurements in the polar region should be developed in the future to perform not only model validation but also climate monitoring accurately.

4.4 Snow/firn/ice surface temperature and albedo

We assessed the surface energy balance of the GrIS simulated by NHM-SMAP in terms of surface temperature and albedo. Measured and simulated snow surface temperature at SIGMA-A from 1

September 2013 to 31 August 2014 agreed well, especially from May to October; however, overestimates were obvious at temperatures below about –20 °C (Fig. 3g), much like the pattern for 2m temperature (Sect. 4.1). As listed in Table S6, the model overestimated surface temperature at all sites except NUK_U, where 2m temperature was also underestimated (Table 3). Therefore, the temporal evolution of simulated surface and 2m temperatures followed the same pattern. Both ME and RMSE for surface temperature were slightly larger than those for 2m temperature (Table 3); however, they are reasonable because they were almost the same as those obtained in Japan (Niwano et al., 2014). It is difficult to ascertain which physical process affected the model tendency because that would require us to investigate the complicated atmosphere–snow/firn/ice coupled system simulated by NHM-SMAP. One possible cause of the model's overestimation of surface temperature is overestimation of the ~~near-~~surface ~~snow density profile,~~wind speeds when they are relatively low (see Sect. 4.2), which ~~would increase the conductive heat flux~~acts to heat the surface ~~(see Sect. 4.5).~~through increases in sensible heat flux. Of course, overestimation of 2m temperature by the model (see Sect. 4.1) especially during winter (November to March) also may contribute to the error. For deeper insight, each physical scheme related to this problem should be investigated by stand-alone tests utilizing detailed in situ measurements.

NHM-SMAP could not adequately reproduce surface albedo. The model tended to overestimate surface albedo, especially in the ablation area (Fig. 5a). Similarly, the RMSE increased at lower surface elevations (Fig. 5b). The model performance was best at SIGMA-A, in the accumulation area, and worst at QAS_L in the ablation area, the most southerly station in this study (Table S7). ME and RMSE at these two stations during months of the study period when the sun appeared (Fig. 5c and 5d) show that model performance was uniformly good at SIGMA-A, covered with snow throughout the year, but both ME and RMSE suddenly increased after June at QAS_L. These results imply that our version of NHM-SMAP has difficulty simulating high-density firn and ice. Alexander et al. (2014) and Fettweis et al. (2017) reported that this is also the case for the MAR model. Tedesco et al. (2016) argued that the discrepancy between measured firn/ice albedo trends and trends modelled by MAR can be explained by the absence in MAR of processes associated with light-absorbing impurities. The dark microbe-rich sediment called cryoconite significantly reduces the surface albedo in the ablation area (Takeuchi et al., 2014; Shimada et al., 2016). Therefore, future models should consider this process as well as the possibility that NHM-SMAP overestimates snowfall during the summer period. In any case, it is necessary to conduct in situ measurements in the ablation area to confirm what is happening in reality.

4.5 Snow surface height

If a polar RCM can calculate changes in surface height realistically, it can be used to partition volume changes supported by satellite altimetry observations into mass changes related to SMB and ice dynamics (Kuipers Munneke et al., 2015). Therefore, we compared the modelled changes in hourly snow surface height against in situ measurements obtained at SIGMA-A and SIGMA-B. Because the SIGMA AWSs started operation in the summer of 2012 (Aoki et al., 2014a), comparisons were performed for the 2012–2013 and 2013–2014 mass balance years (September to August). On the whole, the model captured the trend of measured changes, but underestimations were apparent for both sites and years (Fig. 6). At SIGMA-A, ME and RMSE were respectively –0.19 and 0.21 m for 2012–2013 and –0.13 and 0.17 m for

2013–2014. At SIGMA-B, ME and RMSE were –0.24 and 0.26 m for 2012–2013 and –0.04 and 0.12 m for 2013–2014. These scores are still acceptable by comparison to the SMAP validation results for seasonal snowpack in Japan (Niwano et al., 2014). As discussed in Sect. 4.7, SMB at the SIGMA-D site, located near SIGMA-A and SIGMA-B, is well reproduced by the model. Therefore, the underestimation can be attributed mainly to overestimation of simulated snow density, as mentioned in Sect. 4.4. Schemes for new snow density and the viscosity coefficient of snow in the polar region may need to be upgraded by performing detailed laboratory experiments.

4.6 Melt area extent

The area of surface melt in the GrIS was extensive in the summer of 2012, setting a new record on 12 July 2012 (Nghiem et al., 2012; Tedesco et al., 2013; Hanna et al., 2014). At present, the melt area extent in the GrIS is commonly diagnosed from satellite data (Mote, 2007, 2014; Nghiem et al., 2012; Hall et al., 2013). A polar RCM that can simulate the melt area extent realistically would enable us to investigate atmospheric and snow/firn/ice physical factors controlling the melt area extent within the same RCM framework, as was done by Fettweis et al. (2011). We compared the simulated daily melt area extent with the data of Mote (2007, 2014) during 2012 and 2013.

The daily melt area extent simulated by NHM-SMAP was diagnosed from hourly snow/firn/ice surface temperature data and water content profiles. First, the daily maximum surface temperature was extracted at each grid point. If the value reached 0 °C and the top model layer contained water at the time when the maximum surface temperature was recorded, we considered the grid point to have experienced surface melt. Figure 7 shows that the simulated results matched the data well (R^2 was 0.97 and 0.94 for 2012 and 2013, respectively), and NHM-SMAP successfully reproduced the record melt event around 12 July 2012, at which time the simulated melt area extent reached 92.4 %. The following year was relatively cold, as suggested by the maximum observed melt area extent of 44 %, and the model successfully replicated the satellite-derived results. It appears that NHM-SMAP can reliably and consistently simulate surface melt ~~in the GrIS.~~extent in the GrIS. Figure S1, which shows observed and simulated total numbers of surface melt days in 2012, supports this argument.

4.7 Surface mass balance

We evaluated the simulated SMB for the GrIS by using the PROMICE stake measurements and the ice core data obtained at SIGMA-D and SE-Dome (Table 2 and Fig. 2b). During the study period, 55 measurements were available~~. The basic~~, and comparison results are presented in Fig. 8. In addition, simulated SMB data from MAR v3.5.2 forced by JRA-55 (Fettweis et al., 2017) were employed as reference information. The geographic patterns of accumulation and ablation simulated for the 2011–2012, 2012–2013, and 2013–2014 mass balance years ~~(Fig. S1) were almost the same as the annual mean SMB map created by RACMO2.3 (Noël et al., 2016).~~simulated by NHM-SMAP are depicted in Fig. S2.

The default version of NHM-SMAP employs the Richards equation to calculate vertical water movement in snow and firn. However, most polar RCMs employ a simpler scheme in which the maximum amount of water retained against gravity (irreducible water content) controls the vertical water movement (Reijmer et al., 2012). The irreducible water content is typically set at 2 % or 6 % of the pore volume, depending on the chosen modelling strategy. The lower of these values can induce more rapid transport of water towards lower layers, mimicking the piping process. To examine the adequacy of the Richards equation for GrIS SMB estimates, we performed sensitivity tests in which the Richards equation scheme was replaced by bucket schemes with irreducible water contents of 2 % and 6 %. The tests employed only the stand-alone SMAP simulations forced by the atmospheric field calculated by the on-line version of NHM-SMAP, which implies that interaction between the atmosphere and the snow/firn/ice was not considered. In the sensitivity tests, profiles for snow/firn/ice physical conditions were reset at the beginning of the 2011–2012, 2012–2013, and 2013–2014 mass balance years by referring to the simulation data from the on-line version of NHM-SMAP. It means that feedbacks, which have more than a year time-scale, are not considered. In the accumulation area where the observed SMB was positive, the simulated SMB agreed well with measurements during the study period regardless of the choice of vertical water movement scheme; however, the model did not capture large mass losses in which observed SMB reached values lower than –4 m water equivalent (m w.e.). The model tended to overestimate SMB in the lower part of the ablation area. In the default simulation, ME, RMSE, and R^2 were 0.75 m w.e., 1.07 m w.e., and 0.86, respectively. With the bucket scheme, these scores worsened slightly, to 0.82 m w.e., 1.12 m w.e., and 0.85 for the case of 6 % irreducible water content and to 0.95 m w.e., 1.26 m w.e., and 0.85 for the case of 2 % irreducible water content. The Richards equation generally allows more water retention than the bucket scheme (Yamaguchi et al., 2012), which may result in higher near-surface density. In turn, more impermeable ice can form near the surface and induce runoff from the near-surface layer. On the other hand, lower irreducible water content forces rapid transport of water towards lower layers as expected, which acts to prevent the formation of ice layers and thus surface mass loss. To confirm the discussion, the GrIS-area-integrated daily melt and refreeze rates were investigated (Fig. 9). In the figure, results for the 2011-2012 mass balance year are shown, whereas results for other mass balance years are depicted in Fig. S3. During the 2011-2012 mass balance year, simulated daily melt rates were almost the same among the results from Richards equation scheme and two bucket schemes (Fig. 9a); however, refreeze rates from the control Richards equation scheme were much lower compared to other results (Fig. 9b), which is an evidence for the above-mentioned more impermeable ice in the results from Richards equation scheme. The same characteristics could be found in other mass balance years (Fig. S3).

Although the Richards equation scheme contributed to improved SMB estimates by NHM-SMAP, the model still produced significant overestimates, especially in the ablation area. Deviations between the measurements and the default model simulation results became larger where the measured SMB was smaller. ~~A~~As presented in Sect. 4.1, the on-line version of NHM-SMAP successfully reproduced 2m air temperature at SIGMA-A during summer. Because surface mass loss during the summer is affected by near-surface (2m) temperature, model performance in terms of simulating JJA 2m air temperature at each AWS on the GrIS were re-examined (Table S8). As indicated in the table, significant or systematic error were not found, and obtained ME and RMSE were well (around –0.2 and 2.1 °C, respectively). Therefore, a possible cause is overestimation of surface albedo by NHM-SMAP, especially in the ablation area (Sect. ~~4.4).~~4.4). According to the PROMICE data in the ablation area, ice albedo often decreases to around 0.2 during summer. Therefore, additional model sensitivity tests, where ice albedo is set to 0.2, were performed. Obtained results indicate that simulated SMB did not change significantly compared to the control Richards equation setting (Fig. 8), suggesting that overestimation of surface albedo by NHM-SMAP can be attributed mainly to overestimates of snowfall as pointed out in Sect. 4.4. In addition, it is possible that even at 5km resolution, NHM-SMAP cannot resolve the complex topography in the ablation area. Recently, Noël et al. (2016) demonstrated that statistical downscaling of individual SMB components from 11km resolution RACMO2.3 to a 1km ice mask and topography (Howat et al., 2014) can improve SMB estimates owing to the correction of modelled surface elevations. Moreover, Wilton et al. (2017) showed generally favourable results from a 1km statistical downscaling of reanalysis data, with results generally comparing well with MAR and RACMO RCM output. On the other hand, MAR v3.5.2 with a horizontal resolution of 20km is generally able to resolve the ablation zone well (Fettweis et al., 2017). A possible cause for this success can be attributed to the introduction of sub-grid mask, which is not employed by NHM-SMAP. It appears that statistical downscaling or further dynamical downscaling or introduction of sub-grid mask is inevitable to obtain more realistic SMB estimates~~.~~ ~~Moreover, it is imperative that we develop a realistic albedo model for high-density firn and ice that incorporates the effects of cryoconite~~.

Using the SMB estimates from NHM-SMAP, we calculated the temporal evolution of accumulated SMB over the entire GrIS during the 2011–2012, 2012–2013, and 2013–2014 mass balance years. We set the area of the GrIS and peripheral glaciers at 1.807×10^6 km^2, as explained in Sect. 2.3.1. The 2011–2012 and 2012–2013 mass balance years present a strong contrast as warm and cold years, respectively. ~~van den Broeke~~According to simulation results by MAR v3.5.2 forced by JRA-55 (Fettweis et al~~. (2016) reported~~., 2017) that ~~in estimates by RACMO2.3, SMB for~~uses the bucket schemes with an irreducible water content of 8 %, the GrIS ~~reached its lowest value since 1958 in~~ SMB during the 2011-2012~~,~~ mass balance year was relatively low (147 Gt year^{-1}), then increased greatly in 2012-2013 (473 Gt year^{-1}) and decreased slightly in 2013-2014~~.~~ (403 Gt year^{-1}). Our model, which tends to simulate lower SMB compared to MAR v3.5.2, produced a similar sequence in those years, with accumulated SMBs at the end of each mass balance year of –23, 420, and 312 Gt year^{-1}, respectively (Fig. ~~9a~~10a). In each of these years, the differences in these estimates emerged after the beginning of June.

Figures ~~9b~~10b to ~~9e~~10e show the accumulated totals of each SMB component in Eq. (6) for the same three mass balance years. They make it clear that the differences in the yearly estimates can be attributed almost entirely to the differences in runoff amounts (Fig. ~~9c~~10c), the differences in P, SU_s, and SU_{ds} being relatively small. As mentioned, NHM-SMAP overestimated SMB especially in the ablation area, which implies that the runoff amount is still underestimated. Future studies should upgrade the model physics in the ways mentioned above, then clarify how much the current version overestimates SMB across the entire GrIS. At the same time, it is imperative to validate the simulations of each SMB component in Eq. (6). In a comparison of SMB components from four reanalysis datasets and the MAR model, Cullather et al. (2016) found that large variations exist for all of the SMB components.

In light of the importance of runoff amount for our SMB estimates, we again investigated the sensitivity of our SMB simulations to the three different vertical water movement schemes. The results clearly showed that the vertical water movement scheme made a notable difference in our GrIS-wide SMB estimates: for the relatively warm 2011–2012 mass balance year, the accumulated SMBs were –

23, 113, and 174 Gt year^{-1} for the default setting and the bucket schemes with irreducible water contents of 6 % and 2 %, respectively (Fig. ~~10a~~11a). Even in the other two relatively cold years, the SMB estimates deviated by as much as 100 Gt year^{-1} (Figs. ~~10b~~11b and ~~10c~~11c). Clearly, the percolation and retention of water in snow and firn plays an important role in estimates of the present-day SMB for the GrIS.

5 Summary and conclusions

We developed the NHM-SMAP polar RCM, with 5km resolution and hourly output, to reduce uncertainties in SMB estimates for the GrIS. Combining JMA's operational non-hydrostatic atmospheric model JMA-NHM and the multi-layered physical snowpack model SMAP, it is an attempt to take advantage of both short-term detailed weather forecast models and long-term computationally stable climate models. ~~Model output data from NHM-SMAP hold promise for assessing not only long-term climate change in the GrIS, but also detailed diurnal variations of meteorological, snow, firn, and ice conditions in the GrIS. We initialized the atmospheric profile every day by referring to JRA-55 (weather forecast mode) to minimize deviations between the JRA-55 and NHM-SMAP atmospheric fields, while simulating the physical states of snow/firn/ice without any initialization (climate simulation mode).~~ The model, forced by the latest Japanese reanalysis data JRA-55, was evaluated in the GrIS during the 2011–2014 mass balance years using in situ data from the SIGMA, GC-Net, and PROMICE AWS networks, PROMICE SMB data, and ice core data from SIGMA-D and SE-Dome. ~~After updating SMAP by incorporating physical processes for new (polar) snow density, ice albedo, and effects of drifting snow, we validated NHM-SMAP in terms of hourly 2m air temperature, 2m water vapor pressure, surface pressure, 10m wind speed, downward shortwave and longwave radiant fluxes, snow/firn/ice surface temperature and albedo, surface height change, daily melt area extent, and the GrIS accumulated SMB.~~

We first tested two options for the lower boundary conditions of the atmosphere. The off-line configuration used values for snow/firn/ice albedo and surface temperature from JRA-55, and the on-line configuration used values from SMAP calculations. The on-line version improved the model performance for 2m air temperature, suggesting that the surface analysis provided by JRA-55 is of inadequate quality, at least for the GrIS, and that SMAP simulates more realistic snow/firn/ice physical conditions. Therefore, we continued our investigation using only the on-line version of NHM-SMAP.

Although the on-line version of NHM-SMAP reproduced a realistic history of 2m air temperature, it produced slight overestimates, especially during winter. A possible cause is overestimation by JRA-55 of surface temperatures in the parent data. JRA-55 overestimates surface air temperature in the polar region and underestimates lower tropospheric air temperature, apparently from deficient treatment of energy exchanges between the atmosphere and the snow/firn/ice surface, especially under very stable atmospheric conditions. To confirm this reasoning would require NHM-SMAP simulations forced by other reanalysis datasets. At the same time, extending the atmospheric spin-up period (6h) can also resolve the issue, because simulation results are expected to less susceptible to a parent reanalysis data. Regarding 2m water vapor pressure, NHM-SMAP did not adequately reproduce absolute water content in the southeastern GrIS, and expanding the model domain to include all of Svalbard, where frequent cyclogenesis accompanies prevailing easterly winds, might improve this result. Surface pressure was simulated realistically. As for 10m wind speed, NHM-SMAP successfully reproduced a Køge Bugt Fjord katabatic flow event observed at station TAS_U on 27 April 2013. Downward shortwave and longwave radiant fluxes, which are important contributors for the GrIS surface energy balance, were also reproduced adequately. Although our RMSEs for downward shortwave radiant flux were almost the same as those reported for Japan with the operational version of JMA-NHM, NHM-SMAP produced greater underestimates when clouds were present. Possible causes for the error include the cloud radiation scheme and the reproducibility of cloud amount and cloud type. For downward longwave radiant flux, the model produced underestimates, especially during winter (November to January). A possible reason is underestimation of lower tropospheric temperature (especially during winter) by JRA-55, and results may also be affected by inadequate reproducibility of the winter cloud amount, low-level liquid clouds, and thin clouds. On the other hand, observation data for downward longwave radiant flux can also have error especially during the winter period due to riming, which might affect the evaluation. Detailed in situ measurements for cloud amount, type, and atmospheric profiles would be required to improve model performance for downward radiant fluxes.

We assessed the simulated surface energy balance in the GrIS in terms of surface temperature and albedo. The model generally overestimated surface temperatures of snow/firn/ice, although our ME and RMSE values were close to those obtained in Japan. A possible cause for this overestimate is overestimation of the ~~near-surface density profile, as suggested by validation of snow surface height changes.~~surface wind speeds when they are relatively low, which acts to heat the surface through increases in sensible heat flux. In addition, overestimation of 2m temperature by the model especially during winter (November to March) also may contribute to the error. The model overestimated the snow/firn/ice albedo, particularly in the ablation area, where both ME and RMSE suddenly increased after June. ~~This finding underscores the need to develop a realistic albedo model for high-density firn and ice that allows us to consider the effects of darkening of the GrIS by cryoconite and so on.~~It was attributed to overestimation of snowfall. Because surface temperature and albedo were reasonably well reproduced in the accumulation area, the model successfully simulated the GrIS melt area extent, including the record surface melt event during the warm summer of 2012 and the relatively cold year 2013.

In our assessment of the model's simulation of SMB, the ME, RMSE, and R^2 values during the study period were fairly good (0.75 m w.e., 1.07 m w.e., and 0.86, respectively). We performed additional sensitivity tests in which the Richards equation scheme to calculate vertical water movement in snow and firn was replaced by simple bucket schemes with irreducible water contents of 2 % and 6 %, demonstrating that the realistic Richards equation scheme contributed to the improvement in SMB estimates. However, the model still produced significant overestimates, especially in the ablation area. Improving this would require developing a realistic albedo model for high-density firn and ice. Resolving overestimation of snowfall by the model is also necessary. Moreover, statistical downscaling or further dynamical downscaling to a higher spatial resolution than used here, e.g. 1 km (Noël et al. 2016, Wilton et al. 2017) or introduction of sub-grid mask (Fettweis et al., 2017) may inevitably be required to improve the SMB estimates. The estimates of accumulated SMB for the entire GrIS were also affected by the choice of vertical water movement scheme, which resulted in differences as great as 200 Gt year^{-1} in our estimates. The process chosen to simulate water percolation and retention in snow and firn thus plays an important role in estimating SMB for the present-day GrIS.

6 Data availability

All of the NHM-SMAP model output data presented in this study are available upon request by contacting the corresponding author (Masashi Niwano, mniwano@mri-jma.go.jp).

Author contributions.

M. Niwano and A. Hashimoto developed the NHM-SMAP coupled system and performed numerical simulations. T. Aoki, S. Yamaguchi, K. Fujita, T. Tanikawa, S. Matoba, and Y. Iizuka contributed ideas for the model improvement. T. Aoki, S. Matoba, S. Yamaguchi, T. Tanikawa, K. Fujita, A. Tsushima, and M. Niwano prepared the SIGMA AWS data. S. Matoba and Y. Iizuka processed in situ SMB data from the SIGMA-D and SE-Dome ice cores. M. Niwano, R. Shimada, A. Hashimoto, T. Tanikawa, and M. Hori created the GrIS ice sheet mask used in this study. M. Niwano prepared the manuscript with contributions from all coauthors.

Acknowledgements.

We thank Tetsuhide Yamasaki for logistical and field support of our field measurements in the GrIS and Sakiko Daorana for her help during our stay in Greenland. We are grateful to Konrad Steffen (Swiss Federal Institute for Forest, Snow and Landscape Research WSL) for providing the GC-Net AWS data, Dirk van As (Geological Survey of Denmark and Greenland) for providing the PROMICE AWS and SMB data, ~~and~~ Thomas L. Mote as well as the National Snow & Ice Data Center for providing the satellite-derived GrIS melt area extent data~~.~~, Xavier Fettweis for providing the MAR model data. We thank Hiroshige Tsuguti, Nobuhiro Nagumo, and Syugo Hayashi of MRI for their help performing numerical calculations and post-processing with JMA-NHM with the supercomputer of MRI (Fujitsu PRIMEHPC FX100 and PRIMERGY CX2550M1). We would like to thank Xavier Fettweis, Leo van Kampenhout, and two anonymous reviewers for providing constructive comments and suggestions, which significantly improved the manuscript. This study was supported in part by (1) the Japan Society for the Promotion of Science through Grants-in-Aid for Scientific Research number JP16H01772 (SIGMA project), JP15H01733 (SACURA project), and JP17K12817, (2) the Japan Aerospace Exploration Agency through the Global Change Observation Mission—Climate (GCOM-C)/Second-generation GLobal Imager (SGLI) Mission, (3) the Ministry of the Environment of Japan through the Experimental Research Fund for Global Environment Conservation, and (4) the Institute of Low Temperature Science, Hokkaido University, through the Grant for Joint Research Program.

References

Ahlstrøm, A. P., Gravesen, P., Andersen, S. B., van As, D., Citterio, M., Fausto, R. S., Nielsen, S., Jepsen, H. F., Kristensen, S. S., Christensen, E. L., Stenseng, L., Forsberg, R., Hanson, S., and Petersen, D.:

A new programme for monitoring the mass loss of the Greenland ice sheet, Geol. Surv. Den. Green. Bull., 15, 61–64, 2008.

Alexander, P. M., Tedesco, M., Fettweis, X., van de Wal, R. S. W., Smeets, C. J. P. P., and van den Broeke, M. R.: Assessing spatio-temporal variability and trends in modelled and measured Greenland Ice Sheet albedo (2000–2013), The Cryosphere, 8, 2293-2312, doi:10.5194/tc-8-2293-2014, 2014.

Amory, C., Trouvilliez, A., Gallée, H., Favier, V., Naaim-Bouvet, F., Genthon, C., Agosta, C., Piard, L., and Bellot, H.: Comparison between observed and simulated aeolian snow mass fluxes in Adélie Land, East Antarctica, The Cryosphere, 9, 1373-1383, doi:10.5194/tc-9-1373-2015, 2015.

Andersen, M. L., Stenseng, L., Skourup, H., Colgan, W., Khan, S. A., Kristensen, S. S., Andersen, S. B., Box, J. E., Ahlstrøm, A. P., Fettweis, X., Forsberg, R.: Basin-scale partitioning of Greenland ice sheet mass balance components (2007–2011), Earth Planet. Sci. Lett., 409, 89–95, doi:10.1016/j.epsl.2014.10.015, 2015.

Aoki, T., Kuchiki, K., Niwano, M., Kodama, Y., Hosaka, M., and Tanaka, T.: Physically based snow albedo model for calculating broadband albedos and the solar heating profile in snowpack for general circulation models, J. Geophys. Res., 116, D11114, doi:10.1029/2010JD015507, 2011.

Aoki, T., Matoba, S., Uetake, J., Takeuchi, N., and Motoyama, H.: Field activities of the "Snow Impurity and Glacial Microbe effects on abrupt warming in the Arctic" (SIGMA) Project in Greenland in 2011-2013, Bull. Glaciol. Res., 32, 3-20, doi:10.5331/bgr.32.3, 2014a.

Aoki, T., Matoba, S., Yamaguchi, S., Tanikawa, T., Niwano, M., Kuchiki, K., Adachi, K., Uetake, J., Motoyama, H., and Hori, M.: Light-absorbing snow impurity concentrations measured on Northwest Greenland ice sheet in 2011 and 2012, Bull. Glaciol. Res., 32, 21-31, doi:10.5331/bgr.32.21, 2014b.

Bamber, J. L., Ekholm, S., and Krabill, W. B.: A new, high-resolution digital elevation model of Greenland fully validated with airborne laser altimeter data, J. Geophys. Res., 106(B4), 6733–6745, doi:10.1029/2000JB900365, 2001.

Bellaire, S., Proksch, M., Schneebeli, M., Niwano, M., and Steffen, K.: Measured and modeled snow cover properties across the Greenland Ice Sheet, The Cryosphere Discuss., doi:10.5194/tc-2017-55, in review, 2017.

Bennartz, R., Shupe, M. D., Turner, D. D., Walden, V. P., Steffen, K., Cox, C. J., Kulie, M. S., Miller, N. B., and Pettersen, C.: July 2012 Greenland melt extent enhanced by low-level liquid clouds, Nature, 496, 83–86, doi:10.1038/nature12002, 2013.

Box, J. E.: Greenland Ice Sheet Mass Balance Reconstruction. Part II: Surface Mass Balance (1840–2010), J. Climate, 26, 6974–6989, doi:10.1175/JCLI-D-12-00518.1, 2013.

Box, J. E. and Rinke, A.: Evaluation of Greenland ice sheet surface climate in the HIRHAM regional climate model using automatic weather station data, J. Climate, 16, 1302–1319, doi:10.1175/1520-0442(2003)16, 2003.

Briegleb, B. P.: Delta-Eddington approximation for Solar Radiation in the NCAR Community Climate Model, J. Geophys. Res., 97, 7603–7612, doi: 10.1029/92JD00291, 1992.

Brun, E., David, P., Sudul, M., and Brunot, G.: A numerical model to simulate snow-cover stratigraphy for operational avalanche forecasting, J. Glaciol., 38, 13–22, 1992.

Brun, E., Six, D., Picard, G., Vionnet, V., Arnaud, L., Bazile, E., Boone, A., Bouchard, A., Genthon, C., Guidard, V., Moigne, P. L., Rabier, F., and Seity, Y.: Snow/atmosphere coupled simulation at Dome C, Antarctica, J. Glaciol., 52, 721–736, 2011.

Cox, C. J., Walden, V. P., Compo, G. P., Rowe, P. M., Shupe, M. D., and Steffen, K.: Downwelling longwave flux over Summit, Greenland, 2010–2012: Analysis of surface-based observations and evaluation of ERA-Interim using wavelets, J. Geophys. Res. Atmos., 119, 12 317–12 337, doi:10.1002/2014JD021975, 2014.

Cuffey, K. and Paterson, W. S. B.: The Physics of Glaciers, Elsevier, Butterworth-Heineman, Burlington, MA, USA, 2010.

Cullather, R.I., Nowicki, S. M. J., Zhao, B., and Koenig, L. S.: A characterization of Greenland ice sheet surface melt and runoff in contemporary reanalyses and a regional climate model, Front. Earth Sci., 4:10., doi:10.3389/feart.2016.00010, 2016.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteorol. Soc., 137, 553–597, doi:10.1002/qj.828, 2011.

~~Dumont, M., Brun, E., Picard, G., Michou, M., Libois, Q., Petit, J.-R., Geyer, M., Morin, S., and Josse, B.: Contribution of light-absorbing impurities in snow to Greenland's darkening since 2009, Nat. Geosci., 7, 509–512, doi:10.1038/ngeo2180, 2014.~~

Enderlin, E. M., Howat, I. M., Jeong, S., Noh, M.-J., van Angelen, J. H., and van den Broeke, M. R.: An improved mass budget for the Greenland ice sheet, Geophys. Res. Lett., 41, 866–872, doi:10.1002/2013GL059010, 2014.

Fausto, R. S., van As, D., Box, J. E., Colgan, W., Langen, P. L., and Mottram, R. H.: The implication of nonradiative energy fluxes dominating Greenland ice sheet exceptional ablation area surface melt in 2012, Geophys. Res. Lett., 43, 2649–2658, 2016.

Fettweis, X.: Reconstruction of the 1979–2006 Greenland ice sheet surface mass balance using the regional climate model MAR, The Cryosphere, 1, 21-40, doi:10.5194/tc-1-21-2007, 2007.

Fettweis, X., Tedesco, M., van den Broeke, M., and Ettema, J.: Melting trends over the Greenland ice sheet (1958–2009) from spaceborne microwave data and regional climate models, The Cryosphere, 5, 359-375, doi:10.5194/tc-5-359-2011, 2011.

Fettweis, X., Box, J. E., Agosta, C., Amory, C., Kittel, C., Lang, C., van As, D., Machguth, H., and Gallée, H.: Reconstructions of the 1900–2015 Greenland ice sheet surface mass balance using the regional climate MAR model, The Cryosphere, 11, 1015-1033, doi:10.5194/tc-11-1015-2017, 2017.

Franco, B., Fettweis, X., and Erpicum, M.: Future projections of the Greenland ice sheet energy balance driving the surface melt, The Cryosphere, 7, 1-18, doi:10.5194/tc-7-1-2013, 2013.

Goody, R. M.: A statistical model for water vapour absorption, Quart. J. Roy. Meteor. Soc., 78, 165–169, doi: 10.1002/qj.49707833604, 1952.

Gordon, M., Simon, K., and Taylor, P. A.: On snow depth predictions with the Canadian land surface scheme including a parametrization of blowing snow sublimation, Atmosphere-Ocean, 44(3), 239-255, doi:10.3137/ao.440303, 2006.

Greuell, W. and Konzelmann, T.: Numerical modelling of the energy balance and the englacial temperature of the Greenland Ice Sheet. Calculations for the ETH-Camp location (West Greenland, 1155ma.s.l.), Global Planet. Change, 9, 91–114, doi:10.1016/0921-8181(94)90010-8, 1994.

Guyomarc'h, G. and Merindol, L.: Validation of an application for forecasting blowing snow, Ann. Glaciol., 26, 138–143, 1998.

Hall, D. K., Comiso, J. C., DiGirolamo, N. E., Shuman, C. A., Box, J. E., and Koenig, L. S.: Variability in the surface temperature and melt extent of the Greenland ice sheet from MODIS, Geophys. Res. Lett., 40, 2114–2120, 2013.

Hanna, E., Huybrechts, P., Janssens, I~~McConnell, J., Das, S~~., Cappelen, J., Steffen, K., and Stephens, A.: ~~Observed~~Runoff and ~~modeled~~mass balance of the Greenland ice sheet~~ snow accumulation,~~: 1958–2003, ~~and links with regional climate forcing, J. Climate, 19, 344–358, 2006~~J. Geophys. Res., 110, D13108, doi:10.1029/2004JD005641, 2005.

Hanna, E., McConnell, J., Das, S., Cappelen, J., and Stephens, A.: Observed and modeled Greenland ice sheet snow accumulation, 1958–2003, and links with regional climate forcing, J. Climate, 19, 344–358, 2006.

Hanna, E., Huybrechts, P., Cappelen, J., Steffen, K., Bales, R. C., Burgess, E., McConnell, J. R., Steffensen, J. P., Van den Broeke, M., Wake, L., Bigg, G., Griffiths, M., Savas, D.: Greenland Ice Sheet surface mass balance 1870 to 2010 based on Twentieth Century Reanalysis, and links with global climate forcing, J. Geophys. Res., 116, D24121, doi:10.1029/2011JD016387, 2011.

Hanna, E., Navarro, F. J., Pattyn, F., Domingues, C. M., Fettweis, X., Ivins, E. R., Nicholls, R. J., Ritz, C., Smith, B., Tulaczyk, S., Whitehouse, P. L., and Zwally, H. J.: Ice-sheet mass balance and climate change, Nature, 498, 51–59, doi:10.1038/nature12238, 2013.

Hanna, E., Fettweis, X., Mernild, S. H., Cappelen, J., Ribergaard, M. H., Shuman, C. A., Steffen, K., Wood, L., and Mote, T. L.: Atmospheric and oceanic climate forcing of the exceptional Greenland ice sheet surface melt in summer 2012, Int. J. Climatol., 34, 1022–1037, doi:10.1002/joc.3743, 2014.

Hashimoto A., Murakami, M., Kato, T., and Nakamura, M.: Evaluation of the influence of saturation adjustment with respect to ice on meso-scale model simulations for the case of 22 June, 2002, SOLA, 3, 085-088, doi:10.2151/sola.2007–022, 2007.

Hashimoto, A., Niwano, M., Aoki, T., Tsutaki, S., Sugiyama, S., Yamasaki, T., Iizuka, Y., Matoba, S.: Numerical weather prediction system based on JMA-NHM for field observation campaigns on the Greenland ice sheet, Low Temperature Science, 75, 91-104, doi:10.14943/lowtemsci.75.91, 2017.

Howat, I. M., Negrete, A., and Smith, B. E.: The Greenland Ice Mapping Project (GIMP) land classification and surface elevation data sets, The Cryosphere, 8, 1509-1518, doi:10.5194/tc-8-1509-2014, 2014.

Iizuka, Y., Matoba, S., Yamasaki, T., Oyabu, I., Kadota, M., and Aoki, T.: Glaciological and meteorological observations at the SE-Dome site, southeastern Greenland Ice Sheet, 34, 1-10, doi:10.5331/bgr.15R03, 2015.

Inoue, J., Liu, J., Pinto, J. O., and Curry, J. A.: Intercomparison of Arctic regional climate models: Modeling clouds and radiation for SHEBA in May 1998, J. Clim., 19, 4167–4178, doi:10.1175/JCLI3854.1, 2006.

Kargel, J. S., Ahlstrøm, A. P., Alley, R. B., Bamber, J. L., Benham, T. J., Box, J. E., Chen, C., Christoffersen, P., Citterio, M., Cogley, J. G., Jiskoot, H., Leonard, G. J., Morin, P., Scambos, T., Sheldon, T., and Willis, I.: Greenland's shrinking ice cover: "fast times" but not that fast, The Cryosphere, 6, 533-537, doi:10.5194/tc-6-533-2012, 2012.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K., Takahashi, K.: The JRA-55 reanalysis: General specifications and basic characteristics, J. Meteorol. Soc. Jpn., 93, doi:10.2151/jmsj.2015-001, 2015.

Kuipers Munneke, P., Ligtenberg, S. R. M., Noël, B. P. Y., Howat, I. M., Box, J. E., Mosley-Thompson, E., McConnell, J. R., Steffen, K., Harper, J. T., Das, S. B., and van den Broeke, M. R.: Elevation change of the Greenland Ice Sheet due to surface mass balance and firn processes, 1960–2014, The Cryosphere, 9, 2009-2025, doi:10.5194/tc-9-2009-2015, 2015.

Langen, P. L., Mottram, R. H., Christensen, J. H., Boberg, F., Rodehacke, C. B., Stendel, M., van As, D., Ahlstrøm, A. P., Mortensen, J., Rysgaard, S., Petersen, D., Svendsen, K. H., Aðalgeirsdóttir, G., and Cappelen, J.: Quantifying energy and mass fluxes controlling Godthåbsfjord freshwater input in a 5 km simulation (1991–2012), J. Climate, 28, 3694–3713. doi:10.1175/jcli-d-14-00271.1, 2015.

Lehning, M., Bartelt, P., Brown, B., Fierz, C., and Satyawali, P.: A physical SNOWPACK model for the Swiss avalanche warning. Part II: Snow microstructure, Cold Reg. Sci. Technol., 35, 147–167, doi:10.1016/S0165-232X(02)00073-3, 2002.

Lefebre, F., Fettweis, X., Gallée, H., Van Ypersele, J.-P., Marbaix, P., Greuell, W., Calanca, P.: Evaluation of a high-resolution regional climate simulation over Greenland, Climate Dynamics, 25, 99, doi:10.1007/s00382-005-0005-8, 2005.

Lenaerts, J. T. M., van den Broeke, M. R., Déry, S. J., van Meijgaard, E., van de Berg, W. J., Palm, S. P., and Sanz Rodrigo, J.: Regional climate modeling of drifting snow in Antarctica, Part I: Methods and model evaluation, J. Geophys. Res., 117, D05108, doi:10.1029/2011JD016145, 2012a.

Lenaerts, J. T. M., van den Broeke, M. R., van Angelen, J. H., van Meijgaard, E., and Déry, S. J.: Drifting snow climate of the Greenland ice sheet: a study with a regional climate model, The Cryosphere, 6, 891-899, doi:10.5194/tc-6-891-2012, 2012b.

Machguth, H., Thomsen, H. H., Weidick, A., Abermann, J., Ahlstrøm, A. P., Andersen, M. L., Andersen, S. B., Bjørk, A. A., Box, J. E., Braithwaite, R. J., Bøggild, C. E., Citterio, M., Clement, P., Colgan, W., Fausto, R. S., Gleie, K., Hasholt, B., Hynek, B., Knudsen, N. T., Larsen, S. H., Mernild, S., Oerlemans, J., Oerter, H., Olesen, O. B., Smeets, C. J. P. P., Steffen, K., Stober, M., Sugiyama, S., van As, D., van den Broeke, M. R., and van de Wal, R. S.: Greenland surface mass balance observations from the ice sheet ablation area and local glaciers, J. Glaciol., doi:10.1017/jog.2016.75, 2016.

Matoba, T., Motoyama, H., Fujita, K., Yamasaki, T., Minowa, M., Onuma, Y., Komuro, Y., Aoki, T., Yamaguchi, S., Sugiyama, S., and Enomoto, H.: Glaciological and meteorological observations at the SIGMA-D site, northwestern Greenland Ice Sheet, Bull. Glaciol. Res., 33, 7-14, doi:10.5331/bgr.33.7, 2015.

Moore, G. W. K., Bromwich, D. H., Wilson, A. B., Renfrew, I., and Bai, L.: Arctic System Reanalysis improvements in topographically forced winds near Greenland, Q.J.R. Meteorol. Soc., 142, 2033–2045, doi:10.1002/qj.2798, 2016.

Mote, T. L.: Greenland surface melt trends 1973–2007: evidence of a large increase in 2007, Geophys. Res. Lett., 34, L22507, doi:10.1029/2007GL031976, 2007.

Mote, T. L.: MEaSUREs Greenland Surface Melt Daily 25km EASE-Grid 2.0, Version 1, Boulder, Colorado, USA., NASA National Snow and Ice Data Center Distributed Active Archive Center, doi: doi:10.5067/MEASURES/CRYOSPHERE/nsidc-0533.001, 2014.

Murata, A., Sasaki, H., Kawase, H., Nosaka, M., Oh'izumi, M., Kato, T., Aoyagi, T., Shido, F., Hibino, K., Kanada, S., Suzuki-Parker, A., and Nagatomo, T.: Projection of future climate change over Japan in ensemble simulations with a high-resolution regional climate model, SOLA, 11, 90–94, doi:10.2151/sola.2015-022, 2015.

Nakanishi, M., and Niino, H.: An improved Mellor-Yamada level-3 model: Its numerical stability and application to a regional prediction of advection fog, Bound.-Layer Meteor., 119, 397-407, doi: 10.1007/s10546-005-9030-8, 2006.

Nghiem, S. V., Hall, D. K., Mote, T. L., Tedesco, M., Albert, M. R., Keegan, K., Shuman, C. A., DiGirolamo, N. E., and Neumann, G.: The extreme melt across the Greenland ice sheet in 2012, Geophys. Res. Lett., 39, L20502, doi:10.1029/2012GL053611, 2012.

Niwano, M., Aoki, T., Kuchiki, K., Hosaka, M., and Kodama, Y.: Snow Metamorphism and Albedo Process (SMAP) model for climate studies: Model validation using meteorological and snow impurity data measured at Sapporo, Japan, J. Geophys. Res., 117, F03008, doi:10.1029/2011JF002239, 2012.

Niwano, M., Aoki, T., Kuchiki, K., Hosaka, M., Kodama, Y., Yamaguchi, S., Motoyoshi, H., and Iwata, Y.: Evaluation of updated physical snowpack model SMAP, Bull. Glaciol. Res., 32, 65-78, doi:10.5331/bgr.32.65, 2014.

Niwano, M., Aoki, T., Matoba, S., Yamaguchi, S., Tanikawa, T., Kuchiki, K., and Motoyama, H.: Numerical simulation of extreme snowmelt observed at the SIGMA-A site, northwest Greenland, during summer 2012, The Cryosphere, 9, 971-988, doi:10.5194/tc-9-971-2015, 2015.

Noël, B., van de Berg, W. J., van Meijgaard, E., Kuipers Munneke, P., van de Wal, R. S. W., and van den Broeke, M. R.: Evaluation of the updated regional climate model RACMO2.3: summer snowfall impact on the Greenland Ice Sheet, The Cryosphere, 9, 1831-1844, doi:10.5194/tc-9-1831-2015, 2015.

Noël, B., van de Berg, W. J., Machguth, H., Lhermitte, S., Howat, I., Fettweis, X., and van den Broeke, M. R.: A daily, 1 km resolution data set of downscaled Greenland ice sheet surface mass balance (1958–2015), The Cryosphere, 10, 2361-2377, doi:10.5194/tc-10-2361-2016, 2016.

Ohtake, H., Shimose, K-I., Fonseca. Jr., J., Takashima, T., Oozeki, T., and Yamada, Y.: Accuracy of the solar irradiance forecasts of the Japan Meteorological Agency mesoscale model for the Kanto region, Japan, Solar Energy, 98, 138-152, doi:10.1016/j.solener.2012.10.007, 2013.

Orr, A., Hanna, E., Hunt, J. C., Cappelen, J., Steffen, K., Stephens, A. G.: Characteristics of stable flows over southern Greenland, Pure appl. geophys., 162, 1747, doi:10.1007/s00024-005-2691-x, 2005.

Reijmer, C. H., van den Broeke, M. R., Fettweis, X., Ettema, J., and Stap, L. B.: Refreezing on the Greenland ice sheet: a comparison of parameterizations, The Cryosphere, 6, 743-762, doi:10.5194/tc-6-743-2012, 2012.

Richards, L. A.: Capillary conduction of liquids through porous mediums. J. Appl. Phys., 1, 318-333, doi:10.1063/1.1745010, 1931.

Rignot, E., Box, J. E., Burgess, E., and Hanna, E.: Mass balance of the Greenland ice sheet from 1958 to 2007, Geophys. Res. Lett., 35, L20502, doi:10.1029/2008GL035417, 2008.

Rignot, E., Velicogna, I., van den Broeke, M. R., Monaghan, A., and Lenaerts, J.: Acceleration of the contribution of the Greenland and Antarctic ice sheets to sea level rise, Geophys. Res. Lett., 38, L05503, doi:10.1029/2011GL046583, 2011.

Saito, K., Fujita, T., Yamada, Y., Ishida, J., Kumagai, Y., Aranami, K., Ohmori, S., Nagasawa, R., Kumagai, S., Muroi, C., Kato, T., Eito, H., and Yamazaki, Y.: The operational JMA nonhydrostatic mesoscale model, Mon. Weather Rev., 134, 1266-1298, doi:10.1175/MWR3120.1, 2006.

Shimada, R., Takeuchi, N., and Aoki, T.: Inter-annual and geographical variations in the extent of bare ice and dark ice on the Greenland ice sheet derived from MODIS satellite images, Front. Earth Sci., 4:43., doi:10.3389/feart.2016.00043, 2016.

Simmons, A. J., and Poli, P.: Arctic warming in ERA-Interim and other reanalyses, Q. J. R. Meteorol. Soc., 141, 1147–1162, doi:10.1002/qj.2422, 2015.

Steffen, K. and Box, J. E.: Surface climatology of the Greenland ice sheet: Greenland Climate Network 1995–1999, J. Geophys. Res., 106, 33951–33964, 2001.

Takeuchi, N., Nagatsuka, N., Uetake, J., and Sshimada, R.: Spatial variations in impurities (cryoconite) on glaciers in northwest Greenland, Bull. Glaciol. Res., 32, 85-94, doi: 10.5331/bgr.32.85, 2014.

Tedesco, M., Fettweis, X., Mote, T., Wahr, J., Alexander, P., Box, J. E., and Wouters, B.: Evidence and analysis of 2012 Greenland records from spaceborne observations, a regional climate model and reanalysis data, The Cryosphere, 7, 615-630, doi:10.5194/tc-7-615-2013, 2013.

Tedesco, M., Doherty, S., Fettweis, X., Alexander, P., Jeyaratnam, J., and Stroeve, J.: The darkening of the Greenland ice sheet: trends, drivers, and projections (1981–2100), The Cryosphere, 10, 477-496, doi:10.5194/tc-10-477-2016, 2016.

van As, D., Hubbard, A. L., Hasholt, B., Mikkelsen, A. B., van den Broeke, M. R., and Fausto, R. S.: Large surface meltwater discharge from the Kangerlussuaq sector of the Greenland ice sheet during the record-warm year 2010 explained by detailed energy balance observations, The Cryosphere, 6, 199-209, doi:10.5194/tc-6-199-2012, 2012.

van den Broeke, M., Smeets, P., Ettema, J., van der Veen, C., van de Wal, R., and Oerlemans, J.: Partitioning of melt energy and meltwater fluxes in the ablation zone of the west Greenland ice sheet, The Cryosphere, 2, 179–189, doi:10.5194/tc-2-179-2008, 2008.

van den Broeke, M. R., Bamber, J., Ettema, J., Rignot, E., Schrama, E. J. O., van de Berg, W. J., van Meijgaard, E., Velicogna, I., and Wouters, B.: Partitioning recent Greenland mass loss, Science, 326, 984–986, doi: 10.1126/science.1178176, 2009.

van den Broeke, M. R., Enderlin, E. M., Howat, I. M., Kuipers Munneke, P., Noël, B. P. Y., van de Berg, W. J., van Meijgaard, E., and Wouters, B.: On the recent contribution of the Greenland ice sheet to sea level change, The Cryosphere, 10, 1933-1946, doi:10.5194/tc-10-1933-2016, 2016.

Van Tricht, K., Lhermitte, S., Lenaerts, J. T. M., Gorodetskaya, I. V., L/'Ecuyer, T. S., Noel, B., van den Broeke, M. R., Turner, D. D., and van Lipzig, N. P. M.: Clouds enhance Greenland ice sheet meltwater runoff, Nat. Commun., 7, 10266, doi:10.1038/ncomms10266, 2016.

Vaughan, D. G., Comiso, J. C., Allison, I., Carrasco, J., Kaser, G., Kwok, R., Mote, P., Murray, T., Paul, F., Ren, J., Rignot, E., Solomina, O., Steffen, K., and Zhang, T.: Observations: Cryosphere, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., 4, 317–382, Cambridge University Press, 2013.

Vernon, C. L., Bamber, J. L., Box, J. E., van den Broeke, M. R., Fettweis, X., Hanna, E., and Huybrechts, P.: Surface mass balance model intercomparison for the Greenland ice sheet, The Cryosphere, 7, 599-614, doi:10.5194/tc-7-599-2013, 2013.

Vionnet, V., Brun, E., Morin, S., Boone, A., Faroux, S., Le Moigne, P., Martin, E., and Willemet, J.-M.: The detailed snowpack scheme Crocus and its implementation in SURFEX v7.2, Geosci. Model Dev., 5, 773–791, doi:10.5194/gmd-5-773-2012, 2012.

Vionnet, V., Martin, E., Masson, V., Guyomarc'h, G., Naaim-Bouvet, F., Prokop, A., Durand, Y., and Lac, C.: Simulation of wind-induced snow transport and sublimation in alpine terrain using a fully coupled snowpack/atmosphere model, The Cryosphere, 8, 395-415, doi:10.5194/tc-8-395-2014, 2014.

Warren, S. G., and Wiscombe, W. J.: A model for the spectral albedo of snow. II: Snow containing atmospheric aerosols, J. Atmos. Sci., 37, 2734–2745, doi:10.1175/1520-0469(1980)037<2734:AMFTSA>2.0.CO;2, 1980.

Wilton, D., Jowett, A., Hanna, E., Bigg, G., Van den Broeke, M., Fettweis, X., and Huybrechts, P.: High resolution (1 km) positive degree-day modelling of Greenland ice sheet surface mass balance, 1870–2012 using reanalysis data, Journal of Glaciology, 63(237), 176-193. doi:10.1017/jog.2016.133, 2017.

Yamaguchi, S., Watanabe, K., Katsushima, T., Sato, A., and Kumakura, T.: Dependence of the water retention curve of snow on snow characteristics, Ann. Glaciol., 53(61), 6-12, doi:10.3189/2012AoG61A001, 2012.

Yamaguchi, S., Matoba, S., Yamazaki, T., Tsushima, A., Niwano, M., Tanikawa, T., and Aoki, T.: Glaciological observations in 2012 and 2013 at SIGMA-A site, Northwest Greenland, Bull. Glaciol. Res., 32, 95-105, doi:10.5331/bgr.32.95, 2014.

Table 1: Locations of observation sites for surface meteorology, including surface elevations measured on site (z_{obs}) and specified in NHM-SMAP (z_{model}).

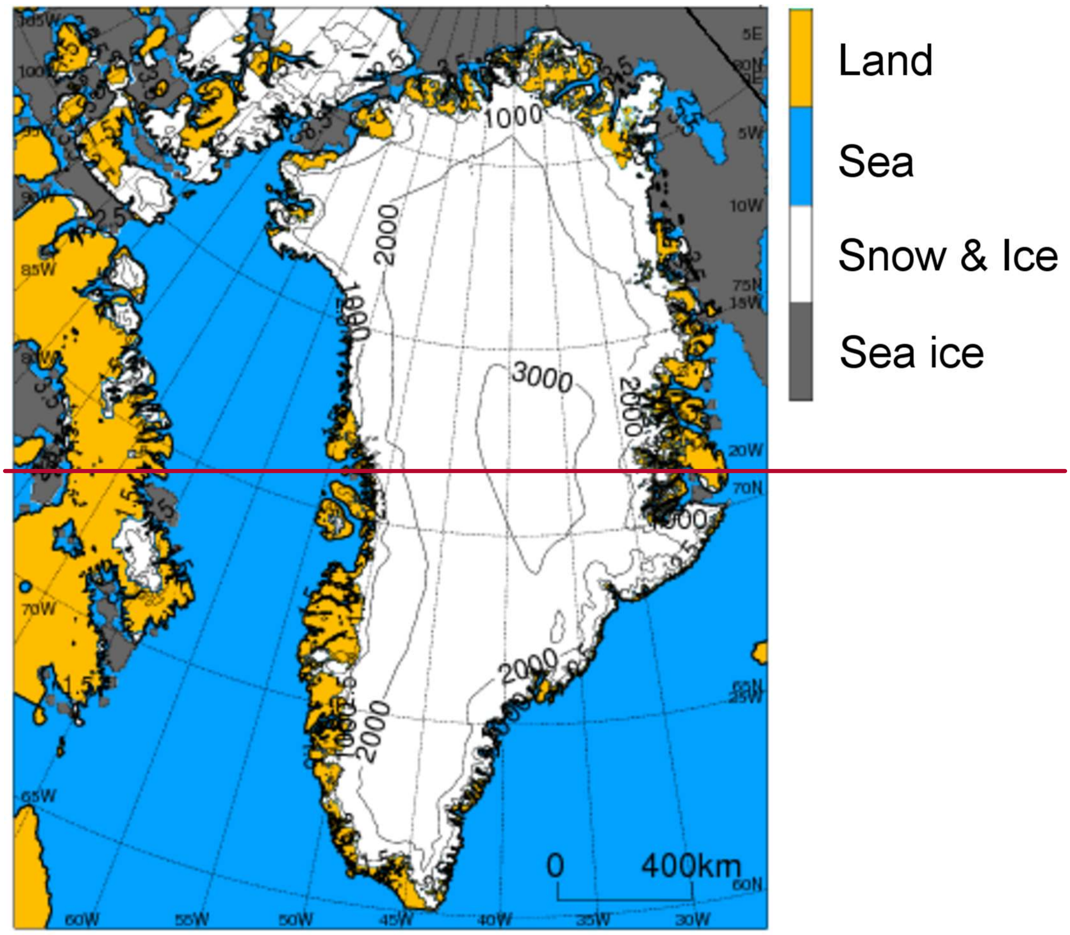
Sites	Latitude (°N)	Longitude (°E)	z_{obs} (m)	z_{model} (m)
SIGMA-A	78.05	–67.63	1490	1494
SIGMA-B	77.52	–69.06	944	779
Summit	72.58	–38.51	3208	3252
S-Dome	63.15	–44.82	2901	2921
KPC_U	79.83	–25.17	870	893
SCO_U	72.39	–27.24	980	1156
TAS_U	65.70	–38.87	570	571
QAS_L	61.03	–46.85	290	375
QAS_A	61.24	–46.73	1010	1114
NUK_L	64.48	–49.53	550	576
NUK_U	64.51	–49.27	1130	1215
NUK_N	64.95	–49.88	920	966
KAN_L	67.10	–49.95	680	606
KAN_M	67.07	–48.83	1270	1319
KAN_U	67.00	–47.02	1840	1860
UPE_L	72.89	–54.3	220	254
UPE_U	72.89	–53.57	940	1017

Table 2: Locations of observation sites for SMB, including the official ID for PROMICE sites and surface elevations measured on site (z_{obs}) and specified in NHM-SMAP (z_{model}).

Glacier names or sites	PROMICE ID	Latitude (°N)	Longitude (°E)	z_{obs} (m)	z_{model} (m)
Tuto Ramp	120_THU_L	76.4	–68.26	570	576
	120_THU_U	76.42	–68.14	770	583
Qaanaaq Ice Cap	126_Q05	77.52	–69.11	839	779
Kronprins Christian Land	170_KPC_U	79.83	–25.17	870	893
A.P. Olsen Ice Cap	220_11	74.66	–21.55	1132	1270
	220_12	74.65	–21.6	1226	1270
	220_13	74.66	–21.6	1271	1270
	220_14	74.68	–21.61	1334	1270
Violin Glacier	232_SCO_U	72.39	–27.26	1000	1156
Isertoq	270_TAS_L	65.64	–38.9	270	337
Qassimiut Ice Lobe	340_QAS_L	61.03	–46.85	310	375
	340_QAS_U	61.18	–46.82	890	894
Qamanarssup Sermia	414_NUK_L	64.48	–49.53	560	576
	414_NUK_U	64.5	–49.26	1140	1215
Kangilinnguata Sermia	416_NUK_N	64.95	–49.88	930	966
K-Transect	454_S4	67.1	–50.19	383	364
	454_S5	67.1	–50.09	490	473
	454_SHR	67.1	–49.94	710	606
	454_S6	67.08	–49.4	1010	1056
	454_S7	66.99	–49.15	1110	1136
	454_S8	67.01	–48.88	1260	1277
	454_S9	67.05	–48.25	1520	1525
	454_S10	67	–47.02	1850	1860
	454_KAN_L	67.1	–49.93	680	606
	454_KAN_M	67.07	–48.82	1270	1319
	454_KAN_U	67	–47.02	1850	1860
Upernavik	475_UPE_L	72.89	–54.29	230	254
	475_UPE_M	72.89	–53.53	980	1017
SIGMA-D		77.64	–59.12	2100	2097
SE-Dome		67.18	–36.37	3170	3031

Table 3: Model performance in simulating hourly 2m air temperature at each AWS on the GrIS (locations in Fig. 1). ME, mean error (average of the difference between simulated and observed values); RMSE, root mean square error; R^2, coefficient of determination.

Sites	Off-line configuration			On-line configuration			Number of observations
	ME (℃)	RMSE (℃)	R^2	ME (℃)	RMSE (℃)	R^2	
SIGMA-A	2.5	3.7	0.94	1.5	3.0	0.95	18998
SIGMA-B	2.8	3.4	0.97	2.3	2.9	0.97	18540
Summit	6.6	8.1	0.88	2.3	5.2	0.89	21137
S-Dome	1.9	3.4	0.91	0.7	2.8	0.92	15059
KPC_U	3.9	5.5	0.93	2.3	4.4	0.94	26139
SCO_U	2.8	4.6	0.86	0.9	3.9	0.85	25786
TAS_U	2.8	3.7	0.84	2.3	3.2	0.87	23263
QAS_L	1.1	2.3	0.89	0.4	2.0	0.90	23483
QAS_A	0.9	2.8	0.91	~~=~~-0.3	2.6	0.92	8679
NUK_L	1.2	2.8	0.92	0.3	2.1	0.94	21933
NUK_U	0.4	2.4	0.93	~~=~~-0.9	2.4	0.93	20908
NUK_N	1.2	2.6	0.92	0.2	2.1	0.94	19955
KAN_L	2.2	3.3	0.94	0.9	2.5	0.95	25518
KAN_M	2.2	3.6	0.93	0.3	2.7	0.94	21091
KAN_U	2.6	4.0	0.94	0.0	2.7	0.95	22925
UPE_L	2.1	3.8	0.91	1.4	3.5	0.91	25434
UPE_U	1.8	2.9	0.95	0.4	2.2	0.96	23036
Mean value	2.3	3.7	0.92	0.9	3.0	0.92	


Land
Sea
Snow & Ice
Sea ice
1000
2000
3000
0
400km
60W
55W
50W
45W
40W
35W
30W
70N
20W
10W
5W
60N
65W
70W
85W

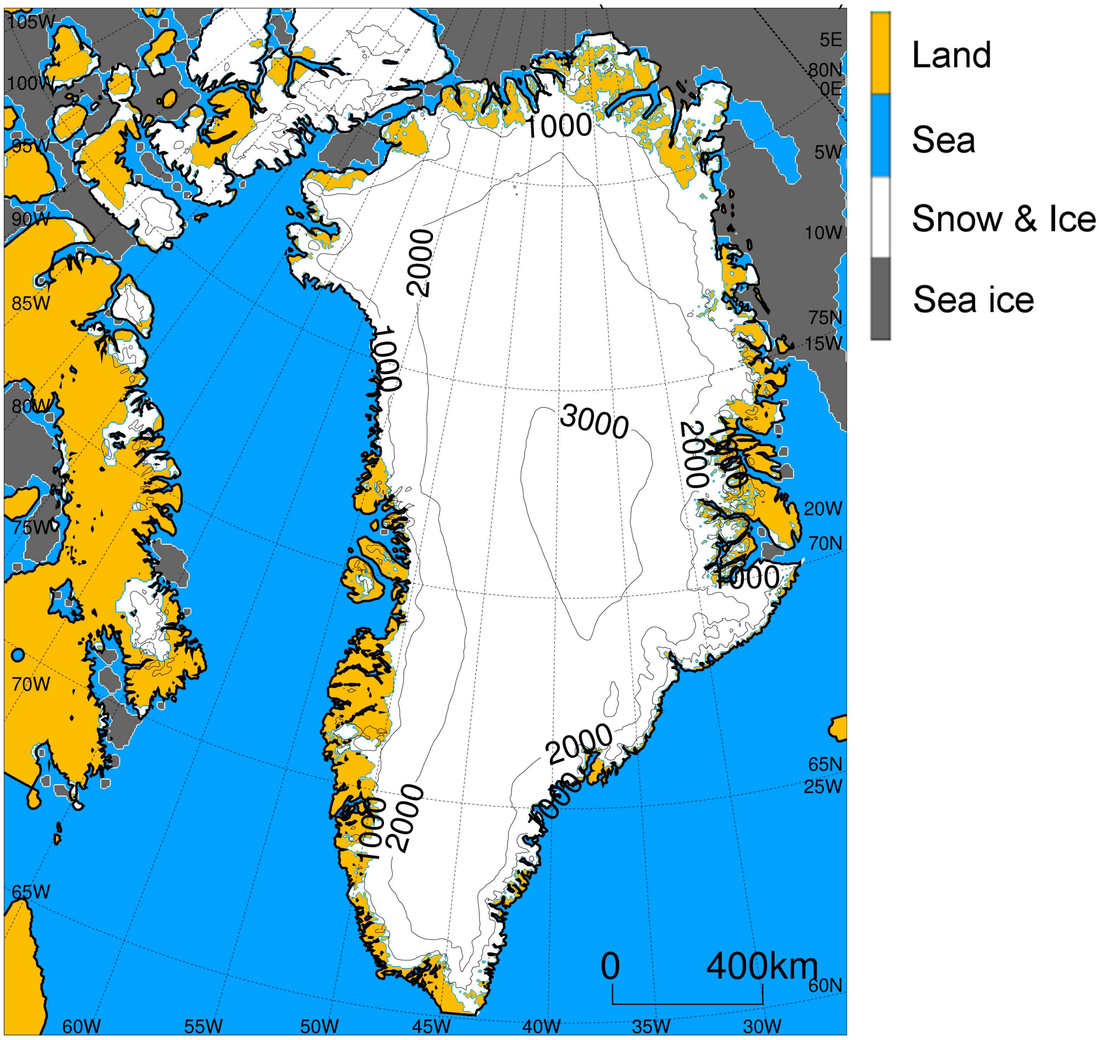


Figure 1: Model domain of NHM-SMAP used in this study showing surface types (colours). The sea ice pattern is depicted for 1 July 2012, and it changes from day to day. Contours on ice sheets and ice caps indicate surface elevation (contour interval 1000 m).

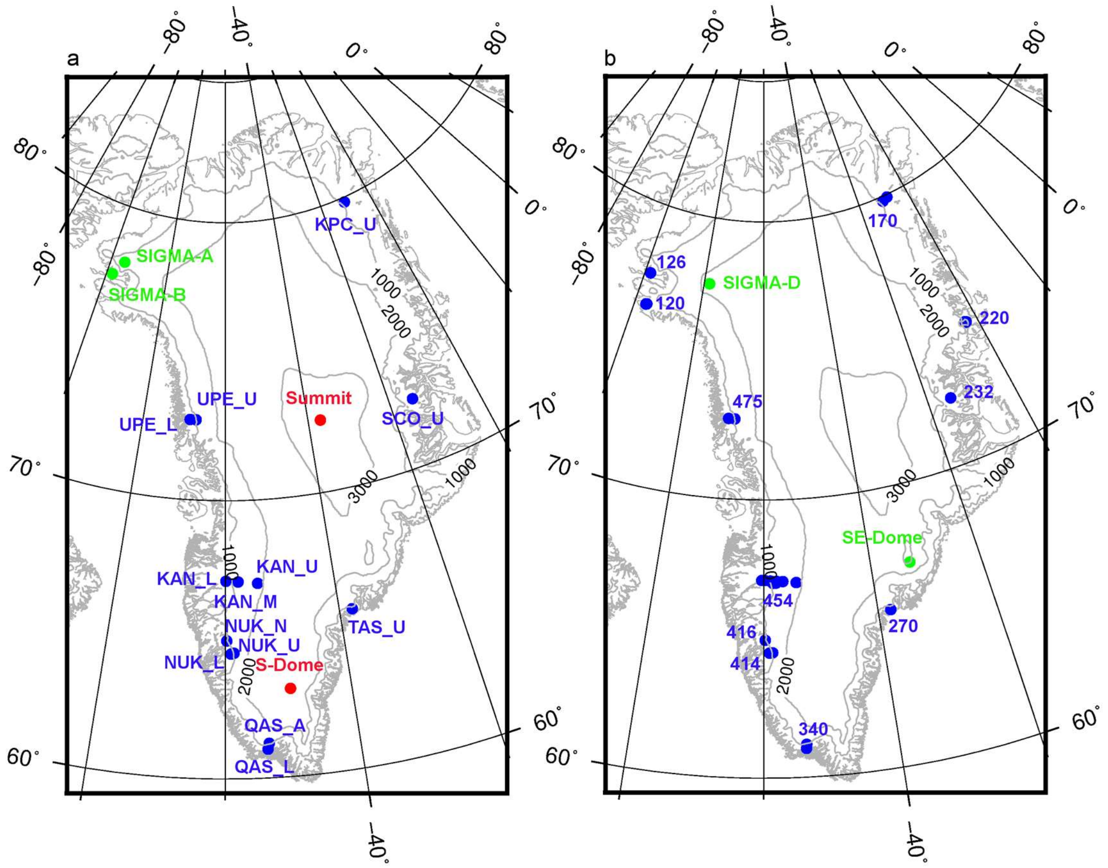
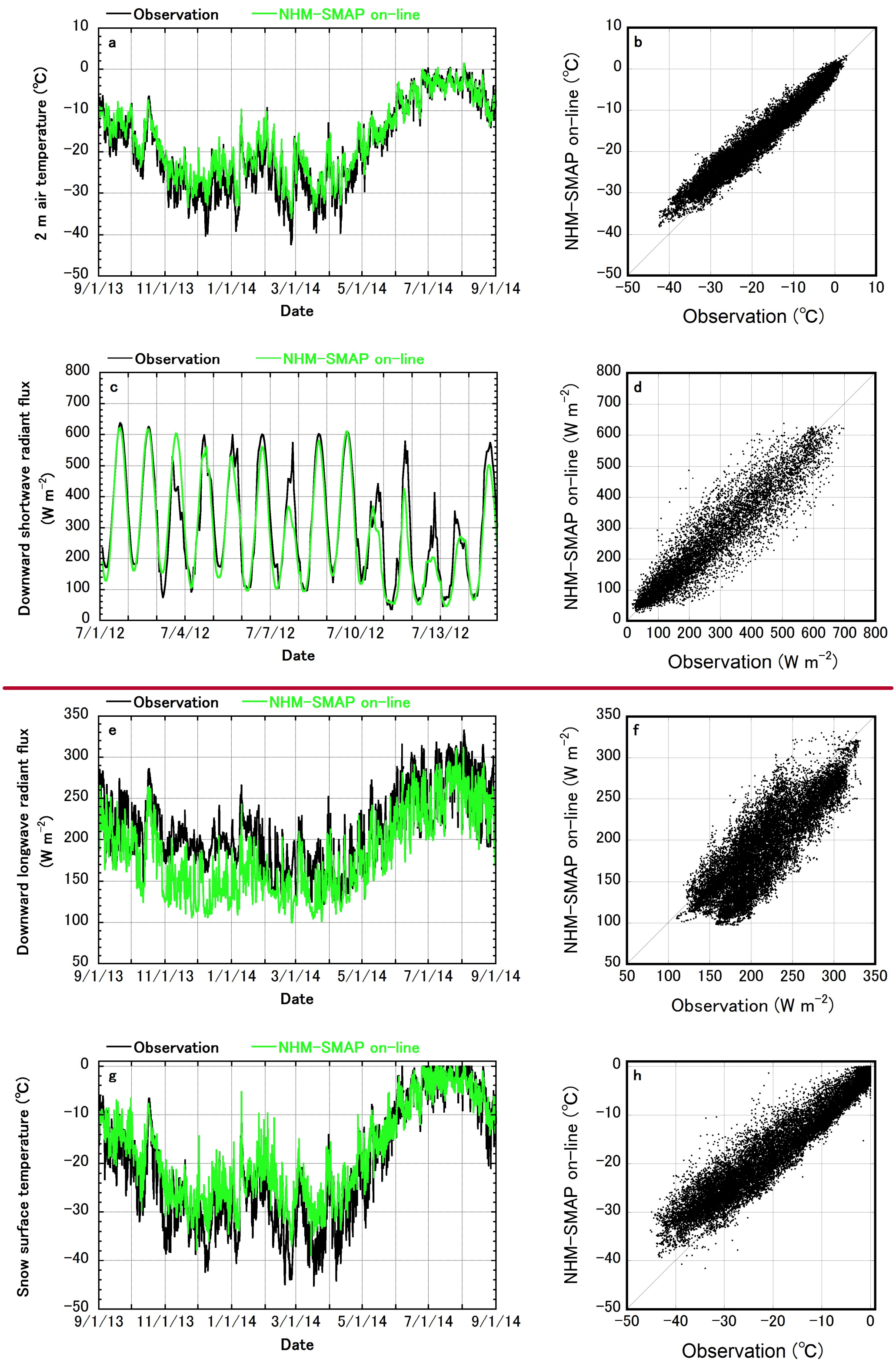


Figure 2: Locations of observation sites for (a) surface meteorology and (b) SMB. Green circles indicate SIGMA and Japanese sites, red circles denote GC-Net sites, and blue circles represent PROMICE sites. Contours on ice sheets and ice caps indicate surface elevation (contour interval 1000 m). All sites are listed in Tables 1 and 2. Site numbers in (b) identify specific glaciers and make up the first part of the PROMICE IDs listed in Table 2.


Observation
NHM-SMAP on-line
a
2 m air temperature (°C)
Date
9/1/13 11/1/13 1/1/14 3/1/14 5/1/14 7/1/14 9/1/14
b
NHM-SMAP on-line (°C)
Observation (°C)
Observation
NHM-SMAP on-line
c
Downward shortwave radiant flux (W m^{-2})
Date
7/1/12 7/4/12 7/7/12 7/10/12 7/13/12
d
NHM-SMAP on-line (W m^{-2})
Observation (W m^{-2})
Observation
NHM-SMAP on-line
e
Downward longwave radiant flux (W m^{-2})
Date
9/1/13 11/1/13 1/1/14 3/1/14 5/1/14 7/1/14 9/1/14
f
NHM-SMAP on-line (W m^{-2})
Observation (W m^{-2})
Observation
NHM-SMAP on-line
g
Snow surface temperature (°C)
Date
9/1/13 11/1/13 1/1/14 3/1/14 5/1/14 7/1/14 9/1/14
h
NHM-SMAP on-line (°C)
Observation (°C)

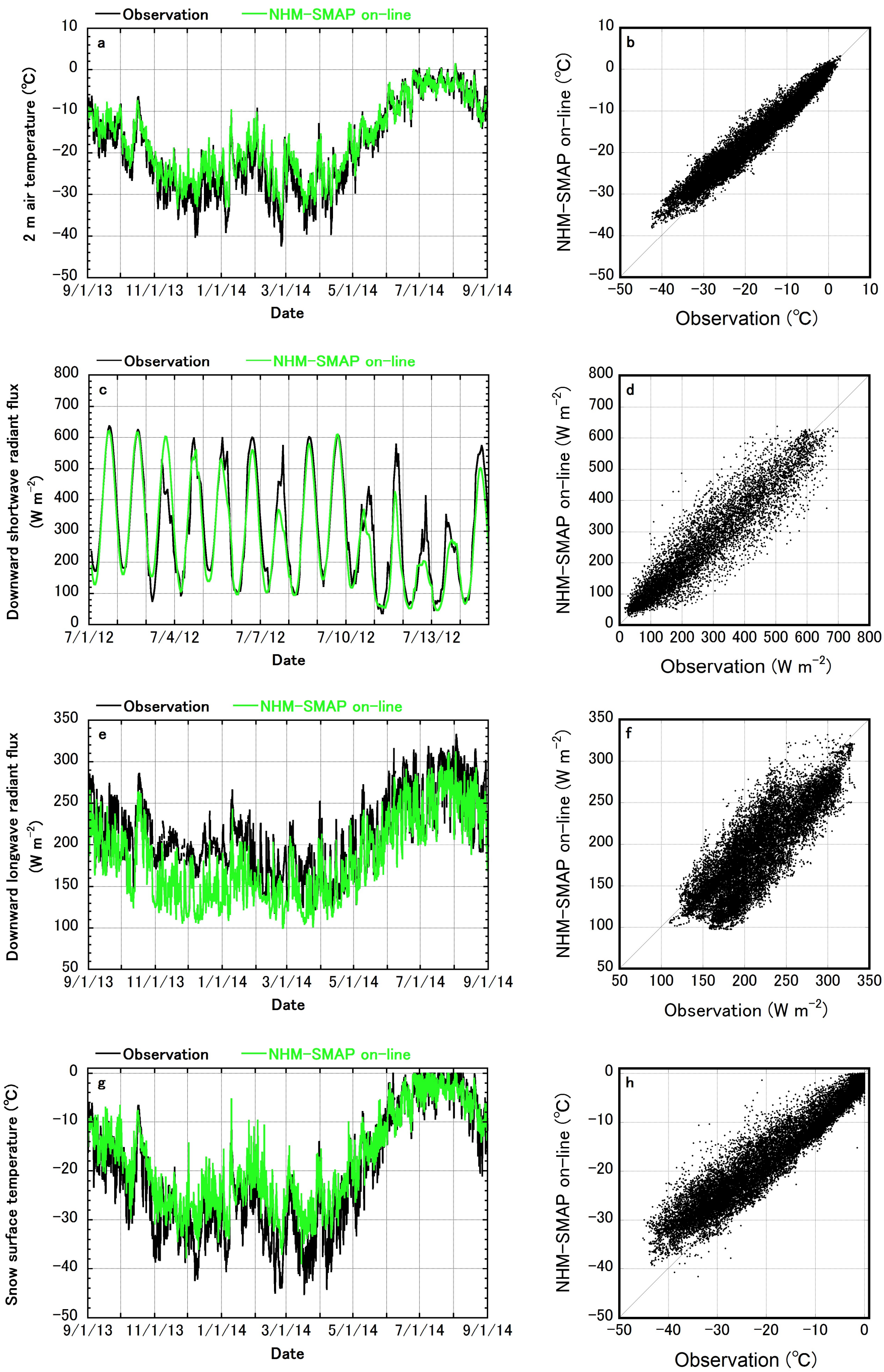
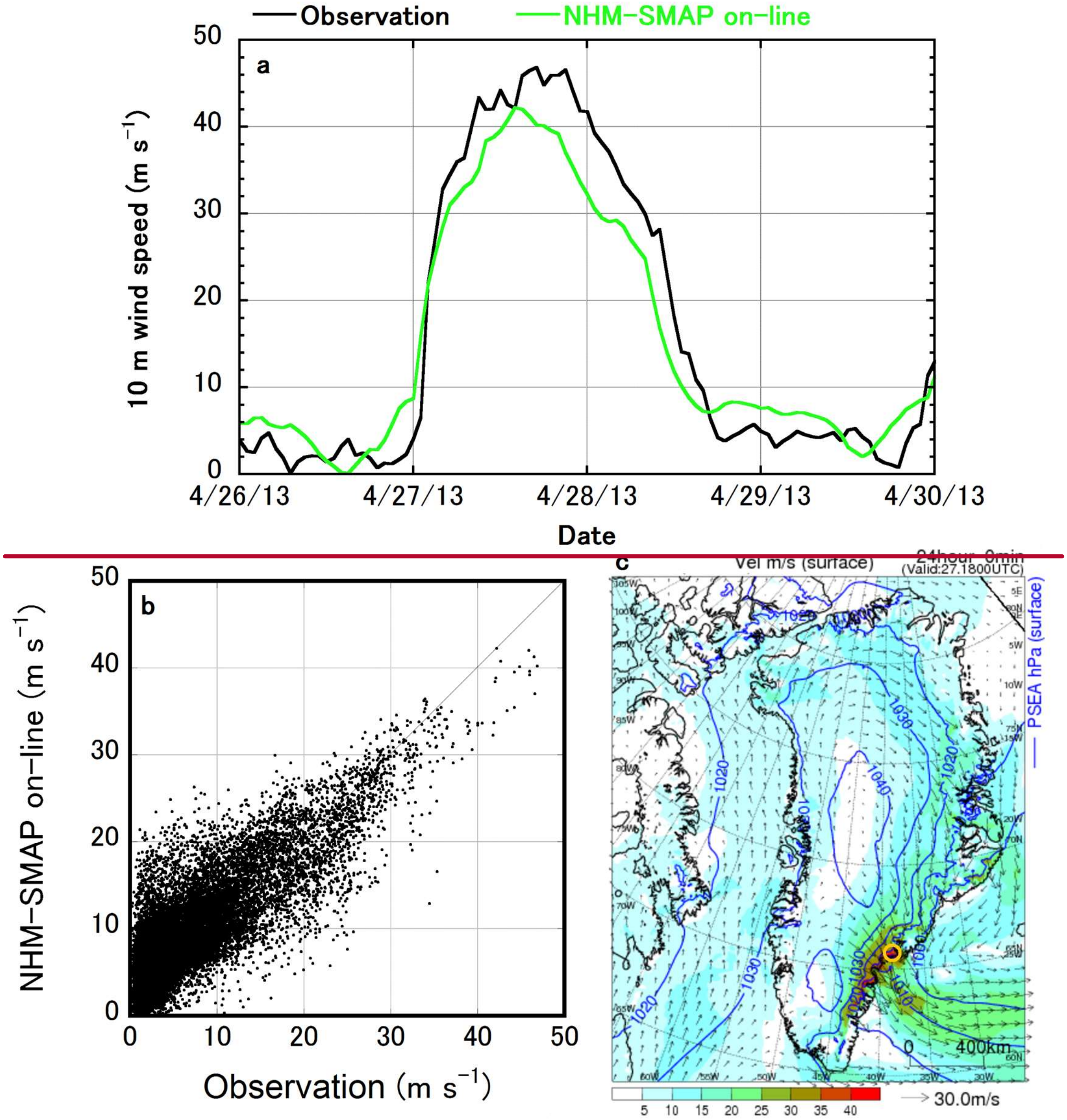


Figure 3: Model validation of hourly (a and b) 2m air temperature, (c and d) downward shortwave radiant flux, (e and f) downward longwave radiant flux, and (g and h) snow surface temperature at SIGMA-A. Target periods for the time series on the left are (a, e, and g) 1 September 2013 to 31 August 2014 and (c) 1–14 July 2012. Data for the scatterplots on the right are from the whole study period, 1 September 2011 to 31 August 2014.


Observation
NHM-SMAP on-line
a
10 m wind speed (m s^{-1})
4/26/13
4/27/13
4/28/13
4/29/13
4/30/13
Date
b
NHM-SMAP on-line (m s^{-1})
Observation (m s^{-1})
c
Vel m/s (surface)
24hour 0min
(Valid:27.1800UTC)
PSEA hPa (surface)
400km
30.0m/s

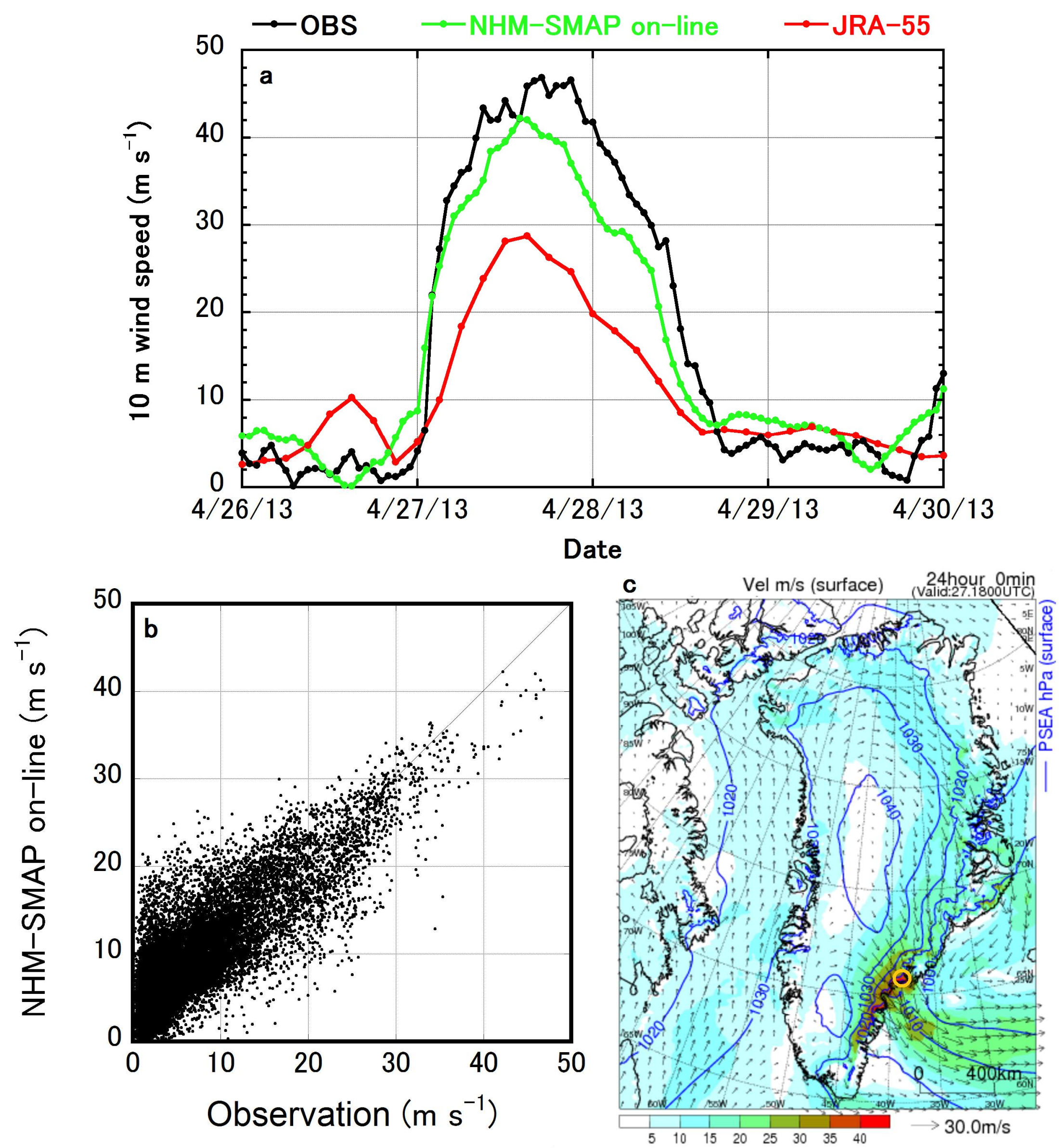
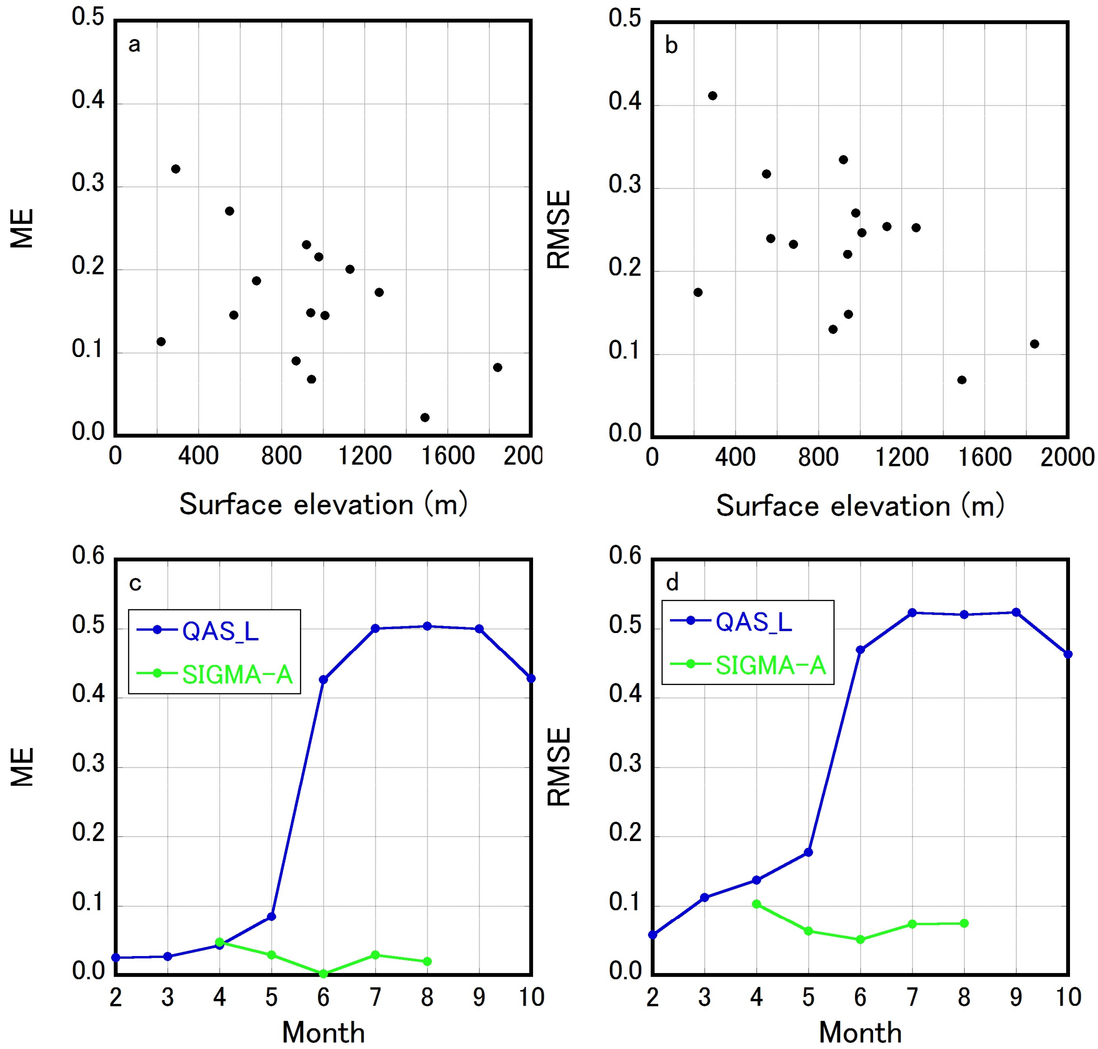
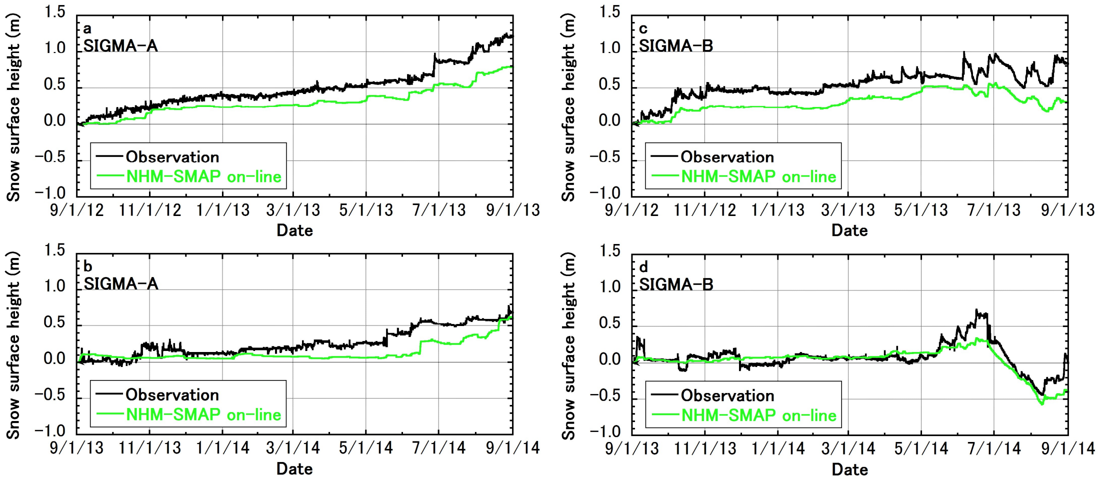


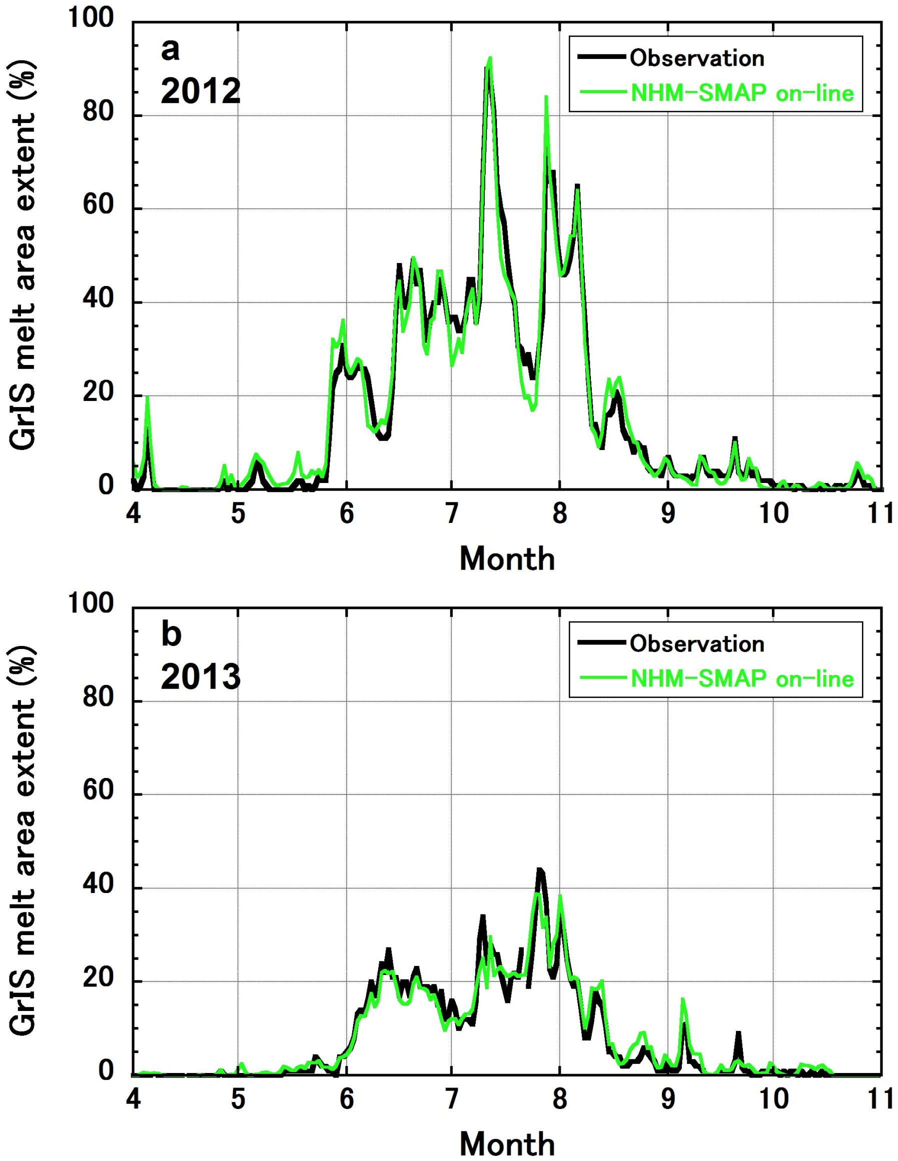
Figure 4: Model evaluation of hourly 10m wind speed at TAS_U. (a) Time series of observed and simulated 10m wind speed at TAS_U from 26 to 29 April 2013. 3 hour interval 10m wind speed from JRA-55 is depicted together. (b) Scatterplot of observed and simulated 10m wind speed at TAS_U during the study period. (c) Surface synoptic weather map for the model region at 1700 UTC on 27 April 2013 simulated by NHM-SMAP, showing surface wind speed (colour), surface wind vector (arrows), and sea level pressure (contours, at 10hPa intervals). Open yellow circle indicates the position of TAS_U.



1045 **Figure 5: Evaluation of the hourly snow/firn/ice albedo simulated at each AWS (Fig. 1 and Table S7). (a) Mean error (ME) and (b) root mean square error (RMSE) as a function of surface elevation. (c) Monthly changes in ME and (d) monthly changes in RMSE for simulated snow/firn/ice albedo at QAS_L (blue line) and SIGMA-A (green line) during months when the sun appears at each site.**



1050 **Figure 6: Time series of observed and simulated hourly snow surface height with respect to 1 September. (a) SIGMA-A, 2012–2013; (b) SIGMA-A, 2013–2014; (c) SIGMA-B, 2012–2013; (d) SIGMA-B, 2013–2014.**



1055 **Figure 7: Time series of observed and simulated daily GrIS melt area extent for (a) 2012 and (b) 2013. Observation data are from Mote (2014).**

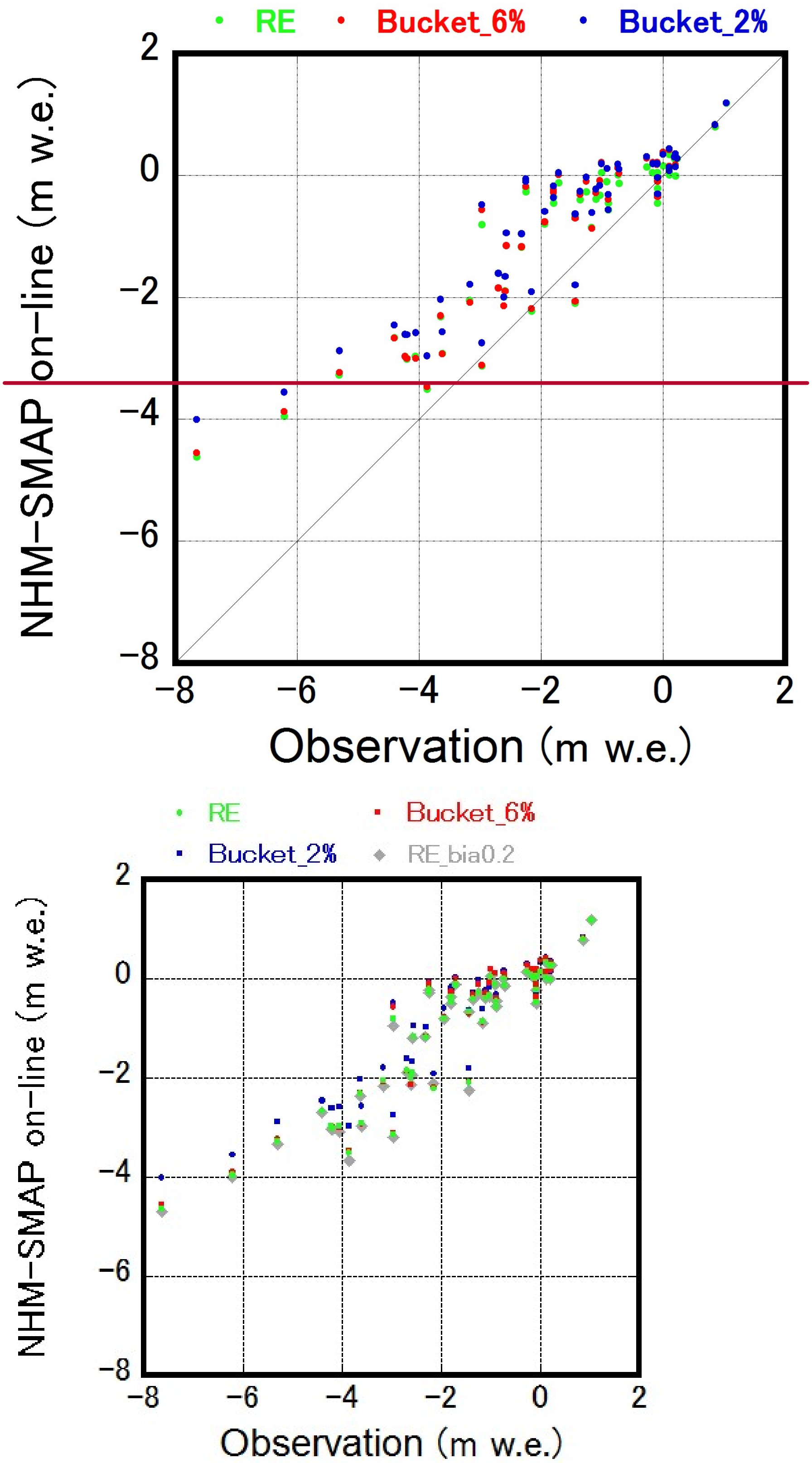
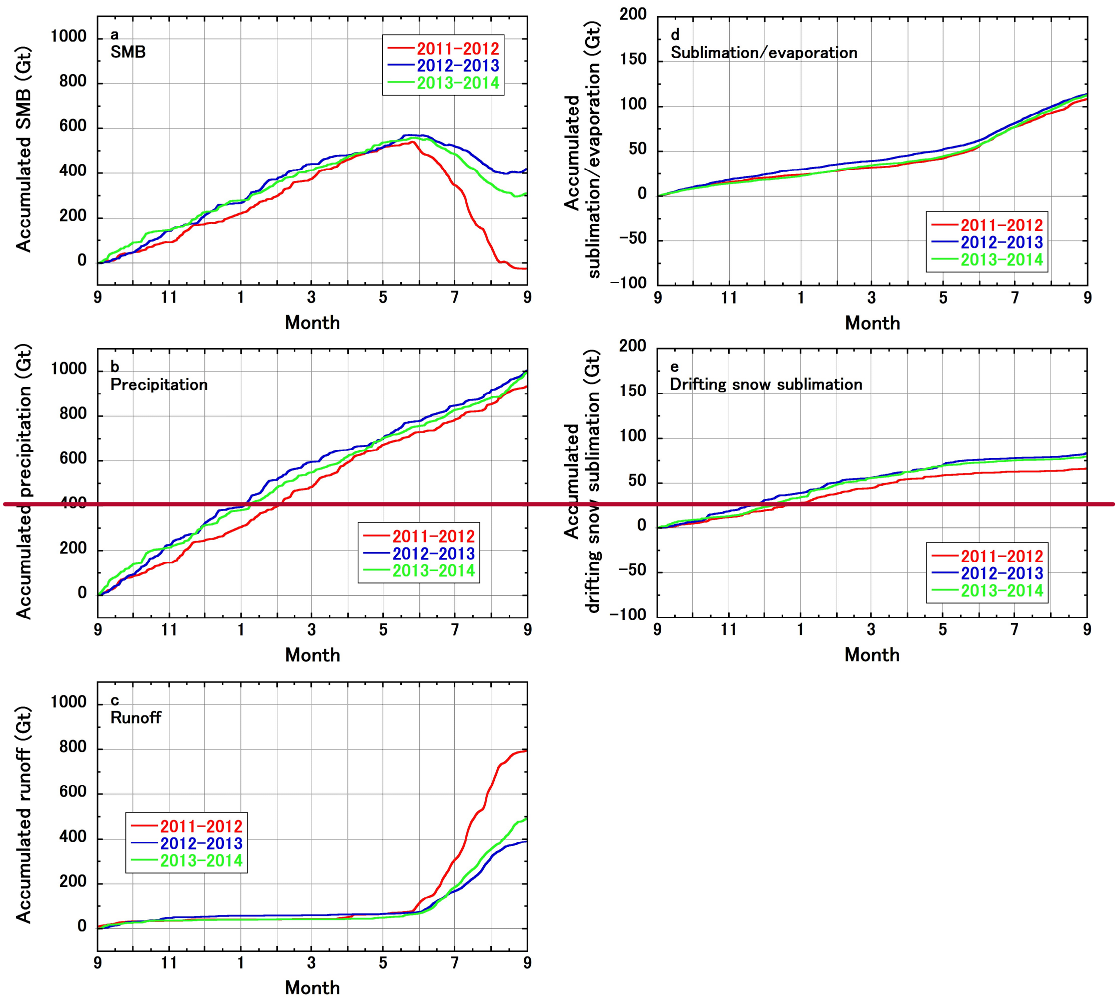
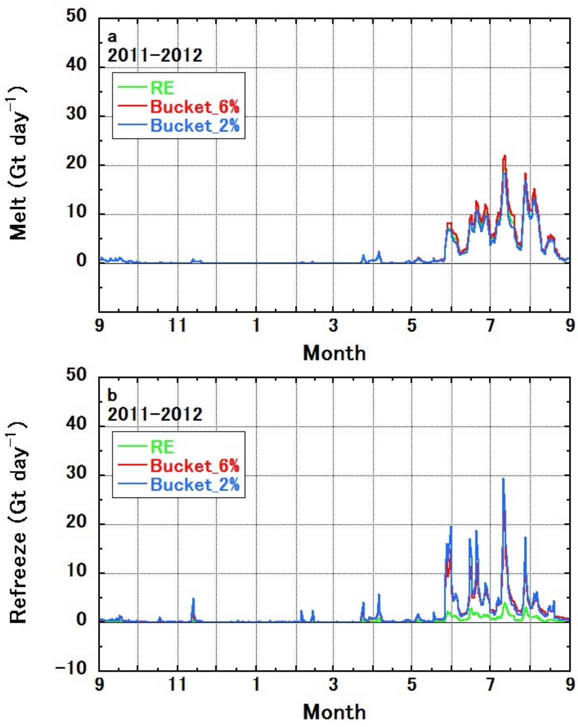


Figure 8: Scatterplot of observed and simulated SMBs during the study period. Observation data are from stake measurements compiled by PROMICE and ice core measurements from SIGMA-D and SE-Dome. RE indicates the default setting for vertical water movement in snow and firn based on the Richards equation; Bucket_6% and Bucket_2% are alternative settings based on simple bucket schemes with irreducible water contents of 6 % and 2 % of the pore volume~~.~~;

RE bia0.2 is another alternative setting, where bare ice albedo is set to 0.2, while other configuration is as same as RE.





~~Figure 9~~

Figure 9: Sensitivity to the choice of vertical water movement scheme of the simulated top 30m integrated (a) melt and (b) refreeze rates for the GrIS during the 2011-2012 mass balance year. RE indicates the default setting for vertical water movement in snow and firn based on the Richards

equation; Bucket 6% and Bucket 2% are alternative settings based on simple bucket schemes with irreducible water contents of 6 % and 2 % of the pore volume.

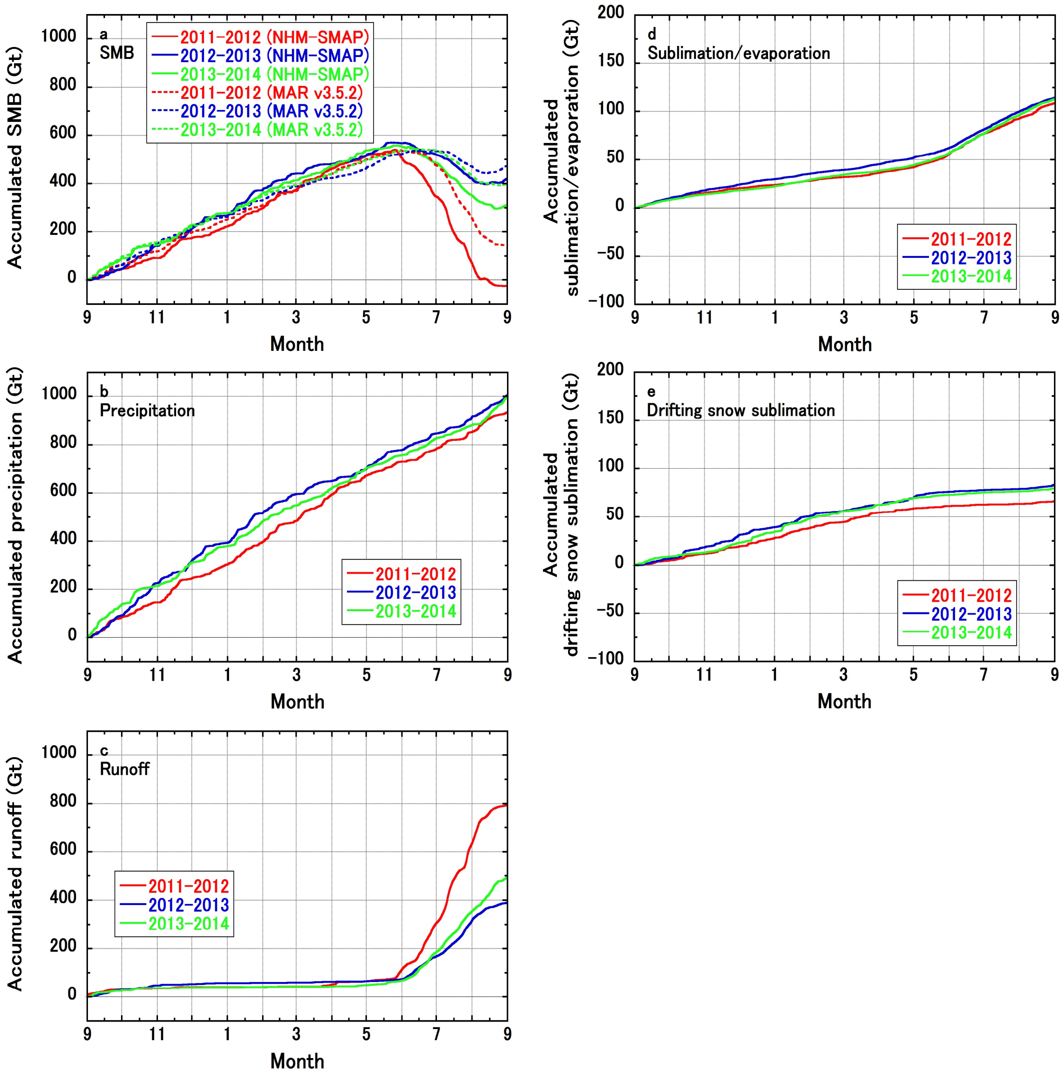
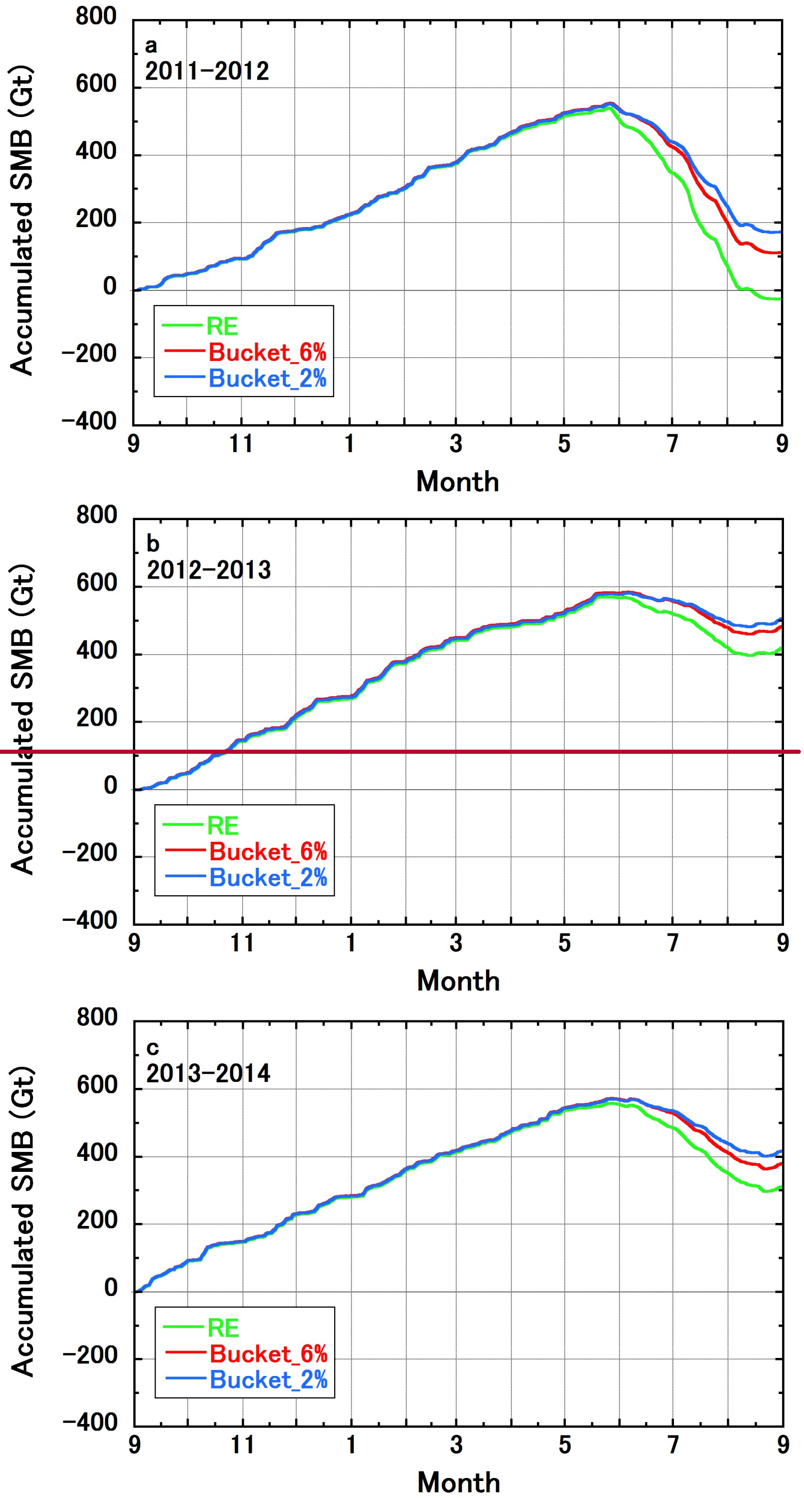


Figure 10: Seasonal evolution of accumulated (a) SMB, (b) precipitation, (c) runoff, (d) sublimation and evaporation from the surface, and (e) drifting snow sublimation over the GrIS with respect to 1 September, during the periods 2011–2012 (red), 2012–2013 (blue), and 2013–2014 (green). Note that the vertical scale differs between the left and right columns. All results are from the default setting for vertical water movement in snow and firn based on the Richards equation. Only for SMB, data from MAR v3.5.2 forced by JRA-55 are displayed together.


a
2011−2012
b
2012−2013
c
2013−2014
Accumulated SMB (Gt)
Month
RE
Bucket_6%
Bucket_2%
800
600
400
200
0
−200
−400
9
11
1
3
5
7
9

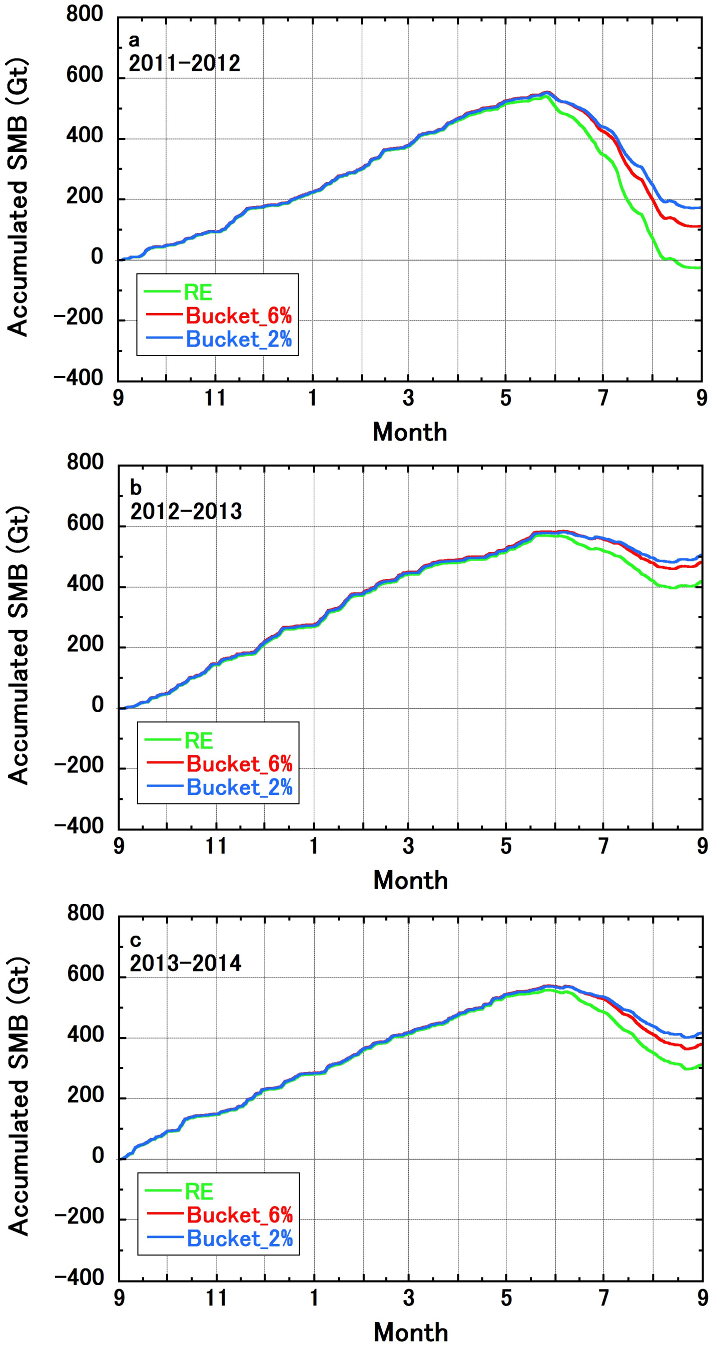


Figure ~~10~~11: Sensitivity to the choice of vertical water movement scheme of the simulated SMB for the GrIS during the (a) 2011–2012, (b) 2012–2013, and (c) 2013–2014 mass balance years. RE indicates the default setting for vertical water movement in snow and firn based on the Richards equation; Bucket_6% and Bucket_2% are alternative settings based on simple bucket schemes with irreducible water contents of 6 % and 2 % of the pore volume.

Table S1. Model performance in terms of simulating hourly 2 m water vapor pressure (in hPa) at each AWS on the GrIS (Figure 1). Note that the evaluation were conducted at only SIGMA and PROMICE sites. ME, RMSE, and R^2 are the mean error (the average of the difference between simulated values and observed values), and the coefficient of determination, respectively. Number of observations (OBS) employed for the comparison are also listed.

Sites	ME (hPa)	RMSE (hPa)	R^2	Number of observations
SIGMA-A	0.07	0.36	0.95	18998
SIGMA-B	0.21	0.48	0.94	18541
KPC_U	–0.01	0.44	0.95	26139
SCO_U	–0.16	0.62	0.90	25786
TAS_U	–0.33	0.76	0.84	23263
QAS_L	–0.53	0.88	0.89	23483
QAS_A	–0.42	0.77	0.89	8678
NUK_L	–0.23	0.67	0.92	21933
NUK_U	–0.30	0.63	0.92	20908
NUK_N	–0.23	0.56	0.93	19955
KAN_L	–0.02	0.52	0.94	25518
KAN_M	–0.15	0.59	0.92	20379
KAN_U	–0.05	0.46	0.93	22925
UPE_L	–0.27	0.69	0.92	25409
UPE_U	–0.27	0.56	0.95	23036
Mean value	–0.18	0.60	0.92	

Table S2. Model performance in terms of simulating hourly surface pressure (in hPa) at each AWS on the GrIS (Figure 1). Elevation differences between the reality and NHM-SMAP are indicated together.

Sites	ME (hPa)	RMSE (hPa)	R^2	Number of observations	Elevation difference (m)
SIGMA-A	–2.8	2.9	0.99	18998	4
SIGMA-B	17.4	17.4	0.99	18550	–165
Summit	–7.6	8.9	0.86	13064	44
S-Dome	–4.3	4.4	1.00	11161	20
KPC_U	–5.5	5.6	0.99	26304	23
SCO_U	–23.1	23.2	0.98	26249	176
TAS_U	–2.3	2.6	0.99	23330	1
QAS_L	–12.5	12.6	0.99	26302	85
QAS_A	–13.9	13.9	1.00	9267	104
NUK_L	–7.5	7.6	0.99	26296	26
NUK_U	–13.0	13.1	0.98	20933	85
NUK_N	–8.4	8.5	0.99	23570	46
KAN_L	5.6	5.7	0.99	26303	–74
KAN_M	–7.8	8.0	0.98	21208	49
KAN_U	–3.7	3.7	0.99	24084	20
UPE_L	–7.2	7.3	0.98	25743	34
UPE_U	–8.6	8.7	0.99	26300	77
Mean value	–8.0	8.9	0.98		

Table S3. Model performance in terms of simulating hourly 10 m wind speed (in m s^{-1}) at each AWS on the GrIS (Figure 1).

Sites	ME (m s^{-1})	RMSE (m s^{-1})	R^2	Number of observations
SIGMA-A	–0.5	2.6	0.40	17846
SIGMA-B	1.0	3.2	0.14	17851
Summit	–0.7	2.5	0.54	18825
S-Dome	–2.0	4.0	0.76	10624
KPC_U	0.4	1.7	0.65	25921
SCO_U	–0.2	2.3	0.13	25774
TAS_U	2.5	4.3	0.68	22977
QAS_L	0.2	2.8	0.51	23423
QAS_A	–0.6	2.5	0.59	8481
NUK_L	0.4	2.3	0.52	21808
NUK_U	2.2	3.2	0.64	20807
NUK_N	–0.3	2.4	0.65	19773
KAN_L	0.8	2.4	0.54	25432
KAN_M	–0.1	2.3	0.72	21047
KAN_U	–1.4	2.8	0.78	22660
UPE_L	1.3	3.1	0.44	25051
UPE_U	0.6	2.5	0.69	22906
Mean value	0.2	2.7	0.55	

Table S4. Model performance in terms of simulating hourly downward shortwave radiant flux (in W m^{-2}) at each AWS on the GrIS (Figure 1).

Sites	ME (W m^{-2})	RMSE (W m^{-2})	R^2	Number of observations
SIGMA-A	–13.5	60.2	0.86	8077
SIGMA-B	–9.4	72.6	0.80	8069
Summit	–9.1	75.9	0.88	10945
S-Dome	52.6	112.3	0.82	10556
KPC_U	–28.6	56.0	0.90	11443
SCO_U	0.6	69.0	0.88	10972
TAS_U	–9.6	88.9	0.81	8588
QAS_L	16.6	96.5	0.83	11229
QAS_A	–3.8	103.7	0.81	3962
NUK_L	2.2	90.8	0.83	8384
NUK_U	–10.5	82.8	0.87	8341
NUK_N	4.4	84.5	0.86	9534
KAN_L	–17.1	127.3	0.70	10837
KAN_M	–16.4	73.0	0.88	8510
KAN_U	–39.4	81.3	0.91	10467
UPE_L	–0.7	78.5	0.83	11007
UPE_U	–7.0	65.0	0.88	11061
Mean value	–5.2	83.4	0.84	

Table S5. Model performance in terms of simulating hourly downward longwave radiant flux (in W m^{-2}) at each AWS on the GrIS (Figure 1). Note that the evaluation were conducted at only SIGMA and PROMICE sites.

Sites	ME (W m^{-2})	RMSE (W m^{-2})	R^2	Number of observations
SIGMA-A	–24.3	36.6	0.71	18353
SIGMA-B	–14.4	31.6	0.72	18440
KPC_U	–14.3	28.3	0.74	26066
SCO_U	–17.0	28.3	0.78	26221
TAS_U	–20.5	32.7	0.66	23107
QAS_L	–19.8	30.2	0.80	26216
QAS_A	–21.4	32.5	0.76	9209
NUK_L	–21.7	32.0	0.80	21835
NUK_U	–13.6	28.6	0.78	20827
NUK_N	–21.3	15.0	0.77	23441
KAN_L	–13.0	28.1	0.76	26155
KAN_M	–10.7	28.5	0.75	21140
KAN_U	–11.7	29.8	0.71	23962
UPE_L	–22.2	35.8	0.72	25562
UPE_U	–13.9	29.8	0.77	26225
Mean value	–17.3	29.9	0.75	

Table S6. Model performance in terms of simulating hourly snow/firn/ice surface temperature (in ºC) at each AWS on the GrIS (Figure 1). Note that the evaluation were conducted at only SIGMA and PROMICE sites.

Sites	ME (°C)	RMSE (°C)	R^2	Number of observations
SIGMA-A	2.3	4.7	0.91	19007
SIGMA-B	3.2	4.9	0.91	18551
KPC_U	2.6	4.8	0.93	26139
SCO_U	1.1	4.3	0.82	26235
TAS_U	1.7	3.2	0.82	23316
QAS_L	0.4	2.2	0.87	26301
QAS_A	0.0	2.6	0.90	9264
NUK_L	0.4	2.7	0.88	21944
NUK_U	–0.3	2.7	0.90	20920
NUK_N	0.1	2.8	0.89	22793
KAN_L	1.1	3.2	0.90	26284
KAN_M	1.0	3.5	0.91	21184
KAN_U	0.9	3.3	0.93	24039
UPE_L	2.0	4.5	0.85	25747
UPE_U	1.0	3.3	0.92	26291
Mean value	1.2	3.5	0.89	

Table S7. Model performance in terms of simulating hourly snow and ice albedo at each AWS on the GrIS (Figure 1). Note that the evaluation were conducted at only SIGMA and PROMICE sites.

Sites	ME	RMSE	R^2	Number of observations
SIGMA-A	0.02	0.07	0.04	3150
SIGMA-B	0.07	0.15	0.06	3250
KPC_U	0.09	0.13	0.06	4451
SCO_U	0.22	0.27	0.09	5297
TAS_U	0.15	0.24	0.10	3627
QAS_L	0.32	0.41	0.12	6415
QAS_A	0.15	0.25	0.03	2252
NUK_L	0.27	0.32	0.13	4501
NUK_U	0.20	0.25	0.09	4752
NUK_N	0.23	0.33	0.12	5352
KAN_L	0.19	0.23	0.16	6003
KAN_M	0.17	0.25	0.12	4571
KAN_U	0.08	0.11	0.07	5967
UPE_L	0.11	0.17	0.19	5136
UPE_U	0.15	0.22	0.10	5243
Men value	0.16	0.23	0.10	

Table S8. Model performance (on-line version of NHM-SMAP) in terms of simulating JJA hourly 2m air temperature at each AWS on the GrIS (Figure 1).

Sites	ME (°C)	RMSE (°C)	R^2	Number of observations	Elevation (m)
SIGMA-A	–0.1	1.5	0.83	5894	1490
SIGMA-B	1.1	1.7	0.87	5446	944
Summit	–0.1	3.5	0.67	5772	3208
S-Dome	0.2	2.2	0.80	4521	2901
KPC_U	–1.2	2.0	0.79	6624	870
SCO_U	–1.7	2.6	0.57	6122	980
TAS_U	2.6	3.2	0.41	4414	570
QAS_L	1.4	2.3	0.45	4273	290
QAS_A	–1.4	2.3	0.48	1992	1010
NUK_L	0.1	1.7	0.53	5351	550
NUK_U	–0.7	2.0	0.67	5308	1130
NUK_N	–0.3	1.5	0.74	3227	920
KAN_L	0.1	1.2	0.76	5960	680
KAN_M	–0.8	1.9	0.80	5097	1270
KAN_U	–1.5	2.5	0.81	6618	1840
UPE_L	–0.1	1.7	0.62	6360	220
UPE_U	–0.6	1.4	0.87	5044	940
Mean value	–0.2	2.1	0.69		

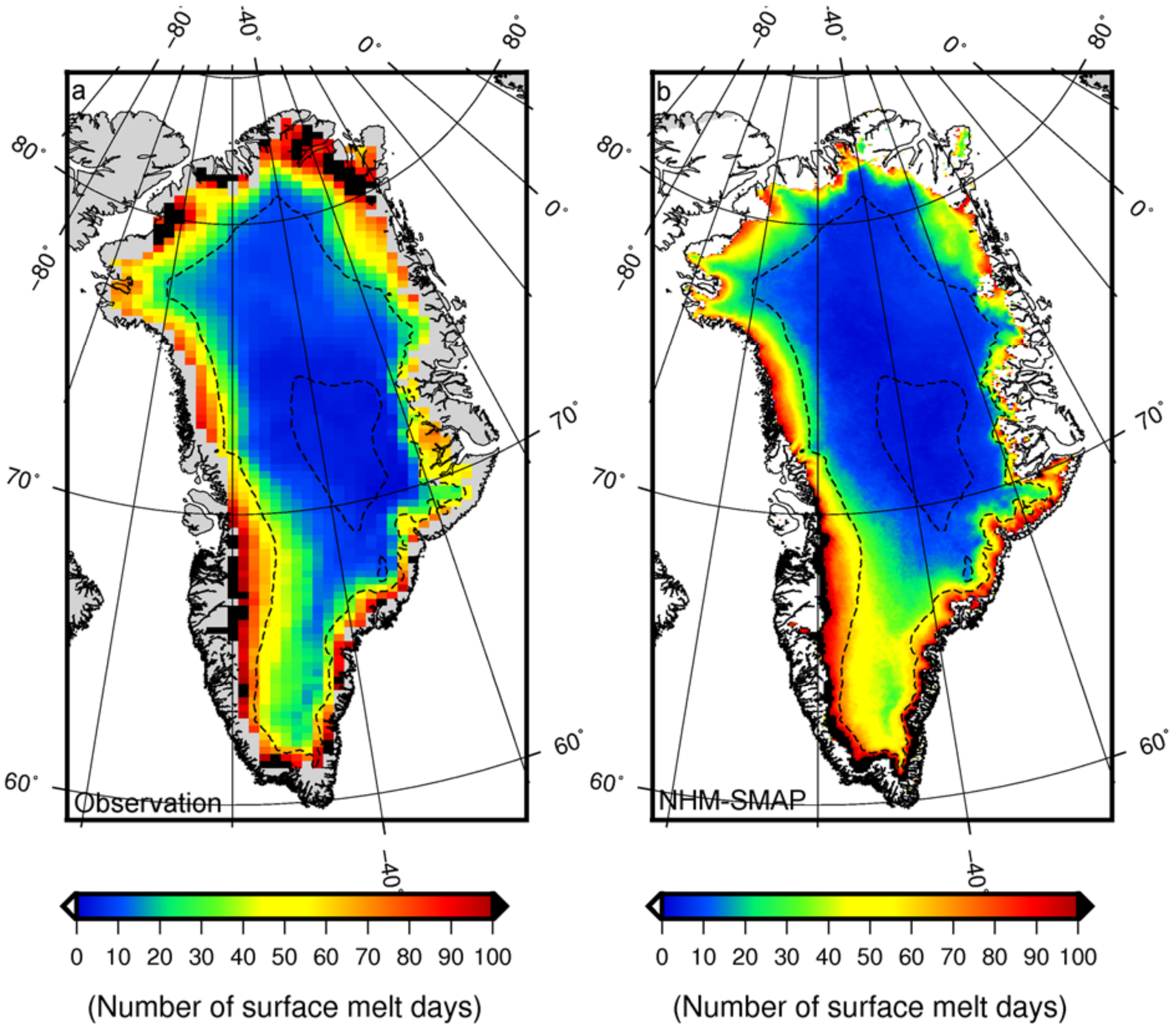


Figure S1: (a) Observed and (b) simulated numbers of the GrIS surface melt days in 2012. Observation data are from Mote (2014).

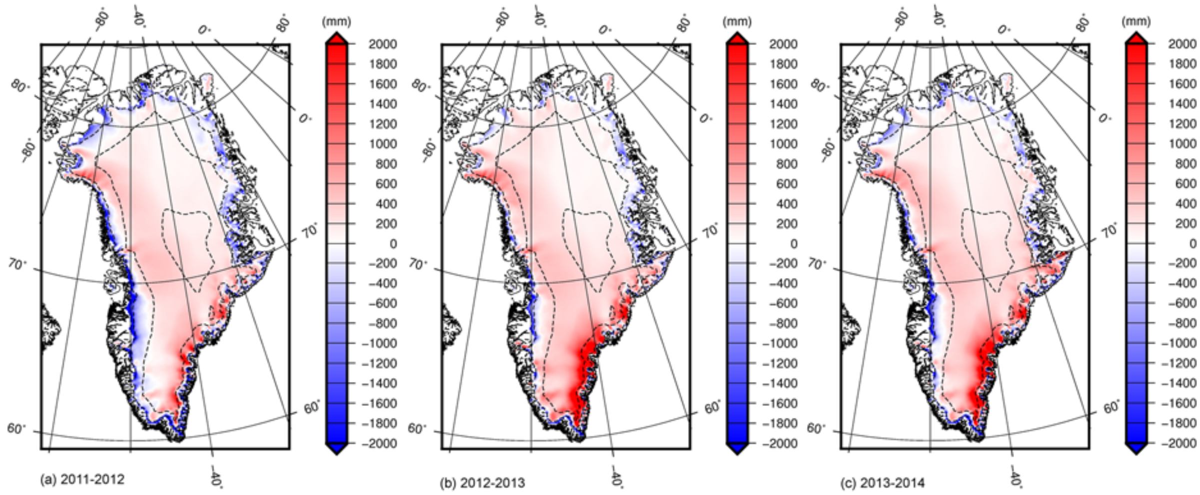


Figure S2: The NHM-SMAP simulated accumulated GrIS SMB (in mm) during the (a) 2011-2012, (b) 2012-2013, and (c) 2013-2014 mass balance years (September to August).

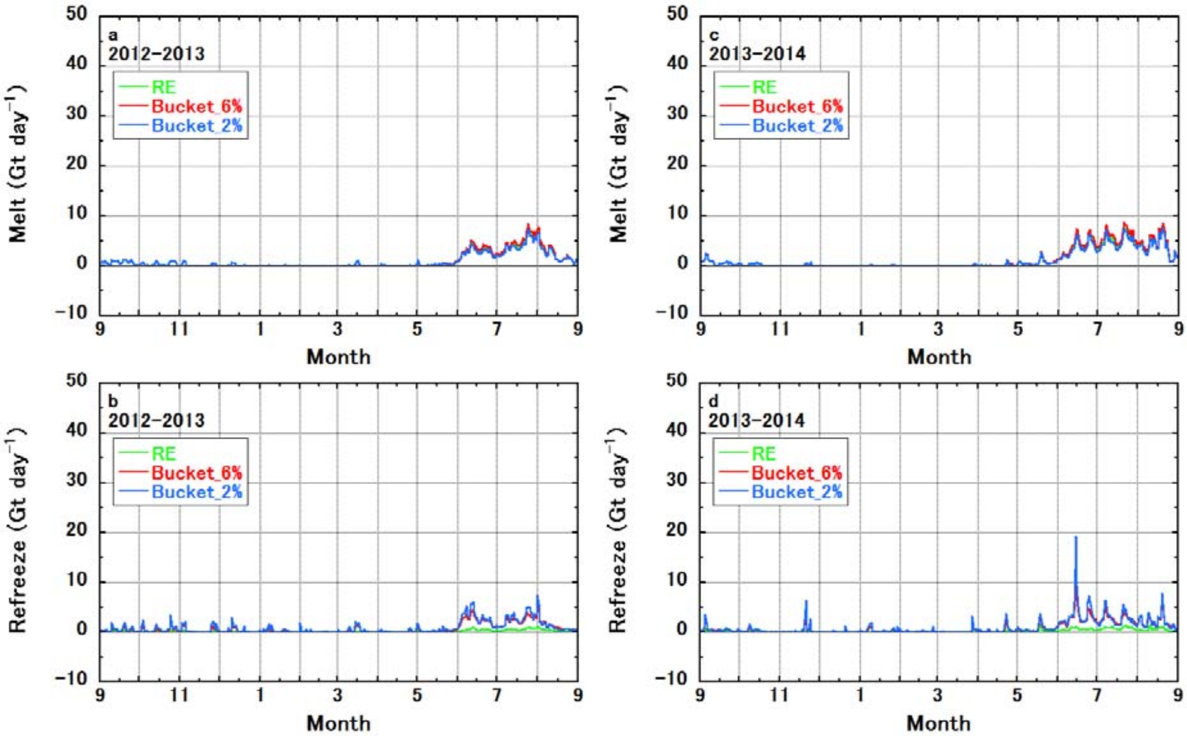


Figure S3: Sensitivity to the choice of vertical water movement scheme of the simulated top 30m integrated (a and c) melt and (b and d) refreeze for the GrIS during the (a and b) 2012-2013 and (c and d) 2013-2014 mass balance years. RE indicates the default setting for vertical water movement in snow and firn based on the Richards equation; Bucket_6% and Bucket_2% are alternative settings based on simple bucket schemes with irreducible water contents of 6 % and 2 % of the pore volume.