Geophysical Research Letters

RESEARCH LETTER
10.1029/2021GL094602

Key Points:

- A data assimilation method was used to generate a reanalysis estimates of the surface mass balance of the Greenland ice sheet along the K-transect stations
- A particle batch smoother technique was used to condition the prior estimates of surface mass balance on 16-day MODIS albedo
- Results show that the assimilation of albedo reduces the root mean square error of the surface mass balance estimates by 51%

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

M. Navari,
mahdi.navari@nasa.gov

Citation:

Navari, M., Margulis, S. A., Tedesco, M., Fettweis, X., & van de Wal, R. S. W. (2021). Reanalysis surface mass balance of the Greenland ice sheet along K-transect (2000–2014). *Geophysical Research Letters*, *48*, e2021GL094602. https://doi.org/10.1029/2021GL094602

Received 14 JUN 2021
Accepted 12 AUG 2021

Reanalysis Surface Mass Balance of the Greenland Ice Sheet Along K-Transect (2000–2014)

Mahdi Navari[1,2], Steven A. Margulis[3], Marco Tedesco[4,5], Xavier Fettweis[6], and Roderik S. W. van de Wal[7]

[1]University of Maryland Earth Systems Science Interdisciplinary Center and Hydrological Sciences Laboratory, Collage park, MD, USA, [2]NASA Goddard Space Flight Center, Greenbelt, MD, USA, [3]Department of Civil and Environmental Engineering, University of California, Los Angeles, Los Angeles, CA, USA, [4]Lamont Doherty Earth Observatory of Columbia University, Palisades, NY, USA, [5]NASA Goddard Institute for Space Studies, New York, NY, USA, [6]Department of Geography, SPHERES Research Unit, University of Liège, Liège, Belgium, [7]Department of Physical Geography, Institute for Marine and Atmospheric Research Utrecht, Utrecht University, Utrecht, The Netherlands

Abstract Accurate estimates of surface mass balance over the Greenland ice sheet (GrIS) would contribute to understanding the cause of recent changes and would help to better estimate the future contribution of the GrIS to sea-level rise. Given the limitations of in-situ measurements, modeling, and remote sensing, it is critical to explore the opportunity to merge the available data to better characterize the spatial and temporal variation of the GrIS surface mass balance (SMB). This work utilizes a particle batch smoother data assimilation technique that yields SMB estimates that benefit from the snow model Crocus and a 16-day albedo product derived from satellite remote sensing data. Comparison of the results against in-situ SMB measurements shows that the assimilation of the albedo product reduces the root mean square error of the posterior estimates of SMB by 51% and reduces bias by 95%.

Plain Language Summary Greenland ice sheet (GrIS) is losing mass through ice discharge from outlet glaciers and surface processes (e.g., meltwater runoff, sublimation, and evaporation). Recent studies suggest that meltwater runoff will be the dominant mass loss process over the GrIS in the future as it will increase under climate warming. Accurate estimates of the GrIS surface mass balance (SMB) are a critical objective, which, despite its importance, continues to contain large uncertainties from significant errors in forcing data as well as model errors. This work uses a data assimilation framework (which has not been used in estimation of the GrIS SMB) and a satellite-derived 16-day albedo product to produce a reanalysis estimates of SMB along the Kangerlussuaq transect (K-transect) stations in west Greenland. We used the K-transect in-situ SMB measurements to validate our results over the 2000–2014 hydrological year. The data assimilation technique (i.e., particle batch smoother) reduces the spatial root-mean-square error of SMB over the K-transect stations by 51% from 858 millimeter water equivalent (mmWE) to 423 mmWE and the bias in the estimates by 95%, from −70 to 3.5 mmWE. It was shown that this methodology has the potential to resolve the spatial variability of the surface processes along the K-transect stations and in particular of the bare ice surface albedo that is not resolved by the model at a resolution of 25 km (i.e., the model uses a constant bare ice albedo). The results suggest that the methodology can be applied over the entire GrIS using MODIS albedo observations to generate an improved reanalysis of SMB estimates.

1. Introduction and Background

The Greenland ice sheet (GrIS) has been the focus of climate studies due to its considerable impact on sea-level rise through net mass gain or loss (mass balance; MB). Surface mass balance (SMB) and ice dynamics are the two mechanisms that control the overall MB. The former includes mass loss/gain through surface processes, and the latter results from direct displacement of water by glacial ice and basal melt. GrIS MB observations show that mass loss has been accelerating over the last decade (e.g., Lenaerts et al., 2019; Mouginot et al., 2019). Fettweis et al. (2013, 2017) and Van Angelen et al. (2013) highlighted that meltwater runoff is the dominant mass loss process over the GrIS. Despite its importance, estimates of SMB over the GrIS contain significant uncertainties (e.g., Fettweis et al., 2020; Rignot et al., 2011; Vernon et al., 2013).

The focus of this study is to further evaluate the proposed methodology on better constraining SMB through surface processes through additional (indirect) observations and a data assimilation framework. Numerical models offer a key tool for quantifying the GrIS surface mass fluxes. However, modeling approaches are generally limited by uncertainty in meteorological forcing data and model errors, which directly propagate into the model results (e.g., Fettweis et al., 2020; Vernon et al., 2013). In-situ measurements such as the network of Kangerlussuaq transect (K-transect) stations in west Greenland (Smeets et al., 2018; Van de Wal et al., 2012) offer valuable information about SMB over the western ablation area. However, there is an extremely limited number of in situ SMB data in the rest of GrIS.

Using surface remote sensing data is a practical alternative to in-situ measurements. Given the spatial coverage, remote sensing data can partially address some of the issues related to the sparse in situ measurements. Unfortunately, surface mass fluxes such as precipitation, evaporation, and runoff cannot be directly measured through remote sensing. This makes the possibility of quantitatively characterizing the surface mass fluxes from remote sensing retrieval algorithms difficult. Despite this challenge, indirect or implicit relationships do exist between surface remote sensing data/products and surface mass fluxes. Surface remote sensing data such as albedo, contain valuable information about the GrIS surface features, including fresh snow, impurities in the snowpack, the presence of liquid water, the formation of supraglacial lakes, and the exposure of bare ice, among other things. However, albedo alone fails to provide quantitative information about the surface mass fluxes.

Given the limitation of different individual methodologies, we argue that using a reanalysis–type approach that merges relevant data streams from satellite observations with a physical model can significantly improve the spatially and temporally continuous estimates of SMB over the GrIS. Margulis et al. (2016) provided an example of such an approach to characterize thirty years of snow water equivalent (SWE) over the Sierra Nevada (United States). We adopted that methodology for a one-year proof of concept study (Navari et al., 2018) in the K-transect region in west Greenland. This work is a continuation of our previous work by extending the study period from 2000 to 2014 (this period is based on the availability of the MAR v3.5 data; Fettweis et al., 2017). In this study, we take advantage of the reanalysis approach to merge a 16-day albedo product with near-surface meteorological forcing data and a snow model knowing that albedo variability is highly correlated to the recent surface mass balance decrease (Riihelä et al., 2019). The reanalysis approach uses a fully Bayesian data assimilation framework to merge different data streams (Margulis et al., 2015, 2016).

The primary difference between the previous studies and the one presented herein includes the fact that the previous studies focused primarily on estimating SMB through remote sensing (e.g., Enderlin et al., 2014) or modeling tools (e.g., Fettweis et al., 2017; Noël et al., 2018; Sellevold et al., 2019) alone. Here, we use a data assimilation approach that does not rely on in-situ data and integrates satellite observations into a snow model. To-date, there have been very limited studies that used data assimilation to predict the GrIS MB/SMB components (e.g., Larour et al., 2014; Navari et al., 2016, 2018).

The methodology described below is for estimating the SMB along the K-transect stations, but the methodology is general and can be applied to larger domains where surface albedo measurements are available. The primary objective of this paper is to present the SMB reanalysis data set, its validation, and an illustration of its utility in the context of quantifying the current GrIS SMB. Methods and data are presented in Section 2. The results are presented in Section 3. Discussion, conclusions, and implications for future work extending to the entire GrIS are presented in Section 4.

2. Methods and Data

2.1. Application Domain

The selected area is located near the city of Kangerlussuaq, in west Greenland along an east-west transect (K-transect) close to the 67°N latitude. The transect has been equipped with different observation instruments including automatic weather stations and SMB stakes at different elevations. Seven SMB measurement stations (S4, S5, SHR, S6, S7, S8, and S9) are located in the ablation zone, and the last station (S10) is located in the percolation zone (Van de Wal et al., 2012). The station IDs and their spatial distribution within the MAR/Crocus 25 km^2 computational grid cells are shown in Figure 1b.

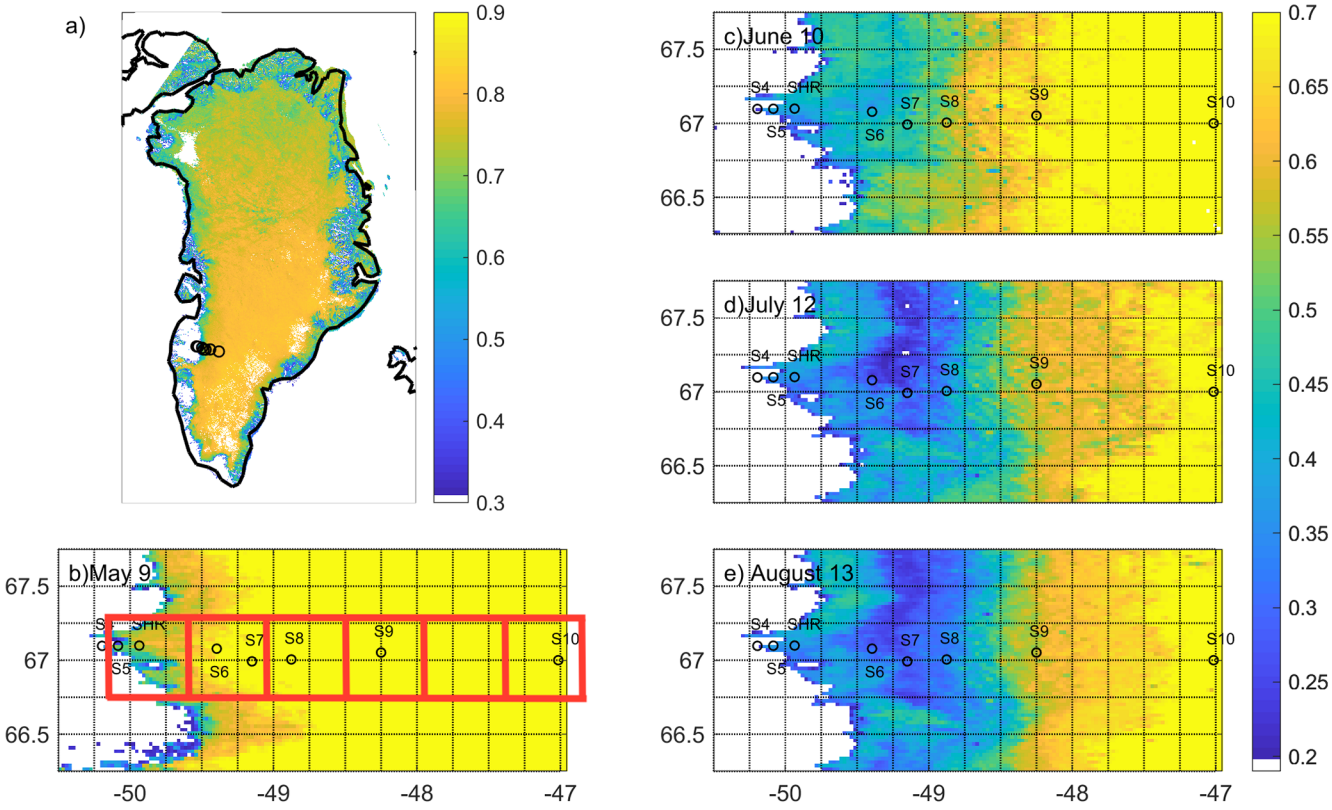

Figure 1. The 2010 16-day albedo map for (a) May 9 and a detailed map showing the location of K-transect stations, day of year (b) May 9, (c) June 10, (d) July 12, and (e) August 13. The black circles show the location of the K-transect stations S4, S5, SHR, S6, S7, S8, S9, and S10 respectively. The dark blue areas near S6–S8 in (d) and (e) represent the Greenland dark zone. The MAR/Crocus computational grid cell are shown in red in panel (b).

2.2. Reanalysis Methodology

The reanalysis method applied in this study consists of a particle batch smoother (PBS) data assimilation approach developed and validated in Margulis et al. (2015, 2016) for seasonal snowpack over the Sierra Nevada, and Navari et al. (2018) for a single year SMB along the K-transect. The snowpack model Crocus (Vionnet et al., 2012) is used to generate uncertain prior estimates of SMB and snow albedo over the study area. An effective way to significantly reduce model and observational errors and produce superior estimates of SMB fluxes is to constrain the snowpack model predictions with satellite observations through data assimilation methods.

There are many DA techniques to integrate model estimates with observations. Sequential DA methods such as Ensemble Kalman Filter (EnKF) Kumar et al. (2017), and Particle Filter (PF) Cluzet et al. (2021) are used for real-time application. Non-sequential DA methods such as ensemble batch smoother (EnBS) and PBS are usually used for reanalysis. Margulis et al. (2016) showed that PBS generally outperformed the EnBS because it provides optimal Bayesian conditioning. For a more detailed discussion about the methodology, see Text S1.

The PBS ingests prior estimates from the snow model Crocus and remotely sensed 16-day albedo observations. The approach generates an ensemble posterior SMB estimate over the study domain. The data assimilation step consists of using the 16-day albedo misfit and the 16-day albedo measurement error covariance to derive the posterior ensemble estimate through a likelihood function. In PBS, each replicate is a realization of the coupled states/fluxes/diagnostics that include the measured quantity (albedo) and the states of interest. The likelihood function simply increases the weight of the replicates (particles) that are closer to the observations and decreases the weights of the replicates that are further from the observations. The updated weights are a discrete approximation of the posterior PDF which can be used to directly estimate posterior statistics of the state variables. In other words, those replicates with "better albedo" realizations

will apply higher weights to all variables associated with that realization, including the state variables of interest (e.g., SMB). The PBS updates the states in a single step using all measurements in the assimilation window. The PBS approach characterizes the mean, median, variance, and interquartile range (IQR) of SMB. For a more detailed discussion about the methodology, see Text S1.

2.3. Data Sets Used

2.3.1. Forcing Data

The data assimilation framework in this study uses hourly near-surface meteorological forcing data (i.e., temperature, pressure, wind speed, longwave and shortwave radiation, precipitation, and humidity) and initial snow and ice profile from the regional climate model Modèle Atmosphérique Régional (MAR; Gallée & Duynkerke, 1997; Gallée & Schayes, 1994) to force the offline Crocus model. The regional climate model MAR is fully coupled with the snow physical model Crocus. However, the integration of the MAR model in an ensemble mode would require significant computational resources. Instead, in the reanalysis data assimilation system, we perturb MAR output meteorology (i.e., temperature, longwave, shortwave, and precipitation) to generate an ensemble of forcing that is passed to Crocus, which significantly reduces the computational cost.

2.3.2. Snow Model Crocus

In this study, we used a stand-alone version of Crocus (Crocus V24), which is different from the version coupled with MAR. The albedo module in this version is simple, therefore, we have updated the Crocus albedo module based on the MAR albedo as follows: The site-specific snow aging term was removed from the visible snow albedo formulation, the lower limit of snow albedo for all spectral ranges was set to 0.65, and the MAR snow and ice albedo transition equations were used to ensure a smooth transition between snow and ice albedo. The more recent version of Crocus is integrated into Meteo-France's SURFEX platform, and it is available online at https://opensource.umr-cnrm.fr/projects/snowtools_git/wiki.

2.3.3. Remotely Sensed Albedo Data

MODIS 16-day albedo (MCD43A, Version-6; Schaaf & Wang, 2015) is a high-quality combined product that uses both Terra and Aqua data to provide 500 m albedo. The product algorithm uses data from days 1 through to 16 to compute the albedo on day 9. We used the direct and diffuse fraction of shortwave irradiance based on the algorithm by Allen et al. (2006) to linearly combine black-sky albedo and white-sky albedo at local solar noon to obtain the true blue-sky albedo that is most consistent with the model-generated albedo. Note that we aggregated the raw albedo data to the model resolution of 25 km^2 and performed the data assimilation in the model spatial resolution. The quality of albedo products decreases where the solar zenith angle is higher than 70° (Stroeve et al., 2005, 2006). Therefore, all the observations corresponding to solar zenith angles larger than 70° were removed from the data set.

Figures 1b–1e show an illustration of the spatial and temporal evolution of albedo on May 9, June 10, July 12, and August 13, 2010, around the K-transect stations. As seen most clearly in panels (b) and (c), albedo gradually increases with elevation. Albedo is low along the margins of the ice sheet, where the first three K-transect stations (S4, S5, and SHR) are located and gradually increases eastward. The irregular variations (low albedo values) seen in panels (d) and (e) around −49° longitude (known as the "dark zone") are mainly driven by snow and ice impurities (e.g., Ryan et al., 2018; Wientjes et al., 2011). During May and June, fresh snow has covered the impurities allowing for a smooth east-west transition between high and low albedo. Temporal variations of albedo can also be seen, most notably between May 9 and July 12. With the advancing melt season, snow cover gradually disappears, and light-absorbing impurities significantly reduce the albedo. This GrIS dark zone increases the absorbed solar radiation and consequently enhances the magnitude of snow/ice melt. Hence, obtaining accurate estimates of snow/ice melt require incorporating this information into model estimates.

2.3.4. Verification Data

The data used for verification in this study are in situ SMB measurements obtained from eight stations (SMB stakes) of differing elevations along the K-transect. Since 1990, SMB measurements have been carried out

at these locations in western Greenland (Smeets et al., 2018; Van de Wal et al., 2012). This data set is the longest record of ground-based SMB measurements in Greenland. The measurements and adjustment of the stakes were conducted in late August, first days of September every year and all data have been quality controlled, reconstructed, and adjusted as needed.

2.3.5. A Quantitative Comparison of Forward Modeling Approaches to Characterize Spatial and Temporal Variation of the GrIS SMB

Forward modeling is a commonly used methodology for providing spatially and temporally continuous estimates of SMB over the GrIS (e.g., Fettweis et al., 2017; Noël et al., 2018; Sellevold et al., 2019). Such deterministic methodologies differ from the work presented herein in several important ways.

Offline forward modeling methodologies are generally limited by uncertainty in meteorological forcing data and subject to direct propagation of errors in those forcings and model errors, both of which impact the estimates of SMB over the GrIS (e.g., Vernon et al., 2013).

The novelty of the reanalysis data set presented herein lies in the use of a fully probabilistic Bayesian assimilation method and the use of information available in the satellite-derived albedo data. The Bayesian PBS method leverages the information content of a meso-scale atmospheric model (i.e., MAR) and the albedo to generate the best estimates of SMB, related fluxes, and their corresponding uncertainty.

The application and characterization described in this research are meant to be illustrative of potential future applications over the entire GrIS. The primary focus of this study was to use the SMB reanalysis data set to characterize the 15-year SMB in the K-transact in terms of its mean and interannual variability. We will focus on comparing reanalysis results with the in-situ measurements to explain basic changes in the SMB through space and time.

3. Results

The model setup and open-loop simulation are described in Navari et al. (2018). For this simulation, we used MAR v3.5 (Fettweis et al., 2017) outputs for offline model inputs and we also used a new set of albedo data (i.e., MCD43A, Version-6). The data assimilation was performed over the 15 years between 2000 and 2014. To illustrate how the PBS data assimilation methodology works, we first present results for a representative individual computational grid-cell and then present time series of SMB for all computational grid cells.

The in-situ measurements sites are located within six MAR grid 25 km^2 cells. Two grid cells contain two stations, three grid cells contain one station, and there is no station in one grid cell (Figure 1b). A preliminary analysis of in-situ SMB measurements (from 1990 to 2014) reveals that more than 60% of the total mass loss along K-transect stations occurs within the first computational grid cell from the edge of the ice sheet and more than 95% in the first two grid cells (see Text S2 and Table S1). Therefore, for illustrative purposes, the focus will be on the first computational grid cell (co-located with the S5 and SHR stations). S5 and SHR located at 6 and 14 km from the ice sheet margin, respectively.

Figure 2 illustrates the time series of the 16-day albedo data during each year for the selected computational grid cell for the study period (see Figures S2–S5 for other grid-cells). Note that PBS applied to each grid cell sequentially with an assimilation window of one year. The observations consist of the MODIS-derived 16-day albedo which are used to condition the prior Crocus estimates in order to produce the posterior estimates. The prior and posterior estimates are the 16-day albedo data from the Crocus snow model before and after the assimilation of satellite-derived albedo. The ensemble statistical information such as the IQR and median are used to evaluate the accuracy of the estimates.

In the selected grid cell, the fresh and dry snow accumulated over the bare ice during the previous accumulation season shows a very high albedo where the satellite-derived albedo remains around 0.75 until mid-April (Figure 2). During the spring (April, May), the snowpack ages and starts to melt, causing a decrease in albedo from 0.75 to around 0.40. During the summer (June–August), the observed albedo remains almost constant at around 0.4. During this period, incoming solar radiation contributes primarily to melting ice layers. The surface albedo increases again when a new accumulation season starts in September.

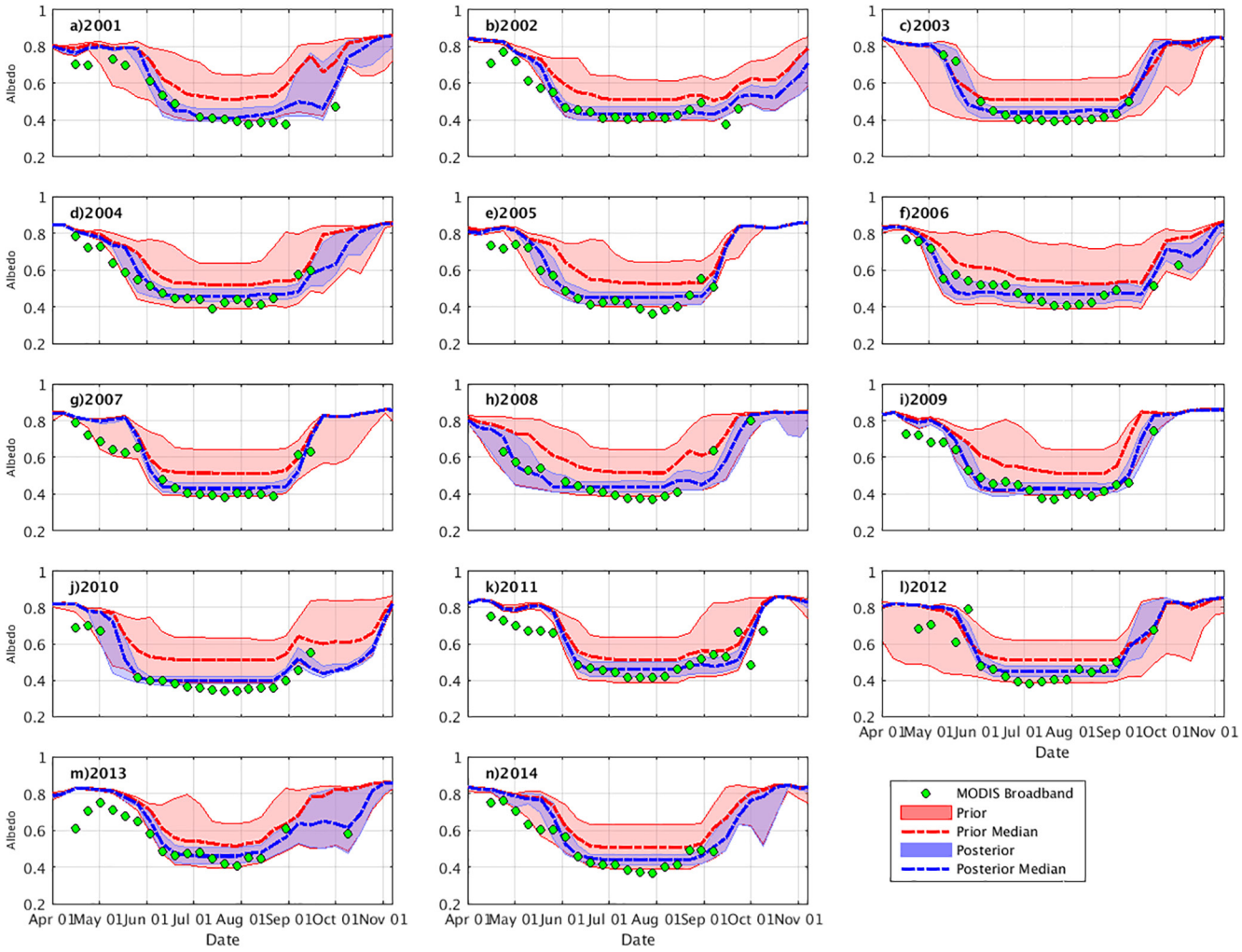

Figure 2. Time series of albedo (—) for the grid cell co-located with stations S5 and SHR for 2001 to 2014. The red and blue shaded areas represent the prior and the posterior uncertainty band (interquartile range [IQR]) and the red and blue dash lines represent the median of the prior and posterior, respectively. The green circles represent the satellite-derived 16-day albedo.

As shown in Figure 2 (all panels), the prior albedo estimates (red dashed lines) are higher than the satellite observations (green circles) over the entire simulation period, in agreement with Alexander et al. (2014). These figures clearly show that the model overestimates the (bare ice) albedo mainly due to inaccurate bare ice albedo parameterization, and a lack of parameterization of impurities. The generally higher albedo implies that the ice sheet surface reflects more incoming solar radiation into the atmosphere and consequently underestimates the snowmelt and ice melt. The ensemble spread of the prior estimates is very large (the IQR ranges from around 0.4 to 0.85) due to uncertain forcings.

The PBS uses the distance between the prior model estimates and observations to generate the posterior estimates. As shown in Figure 2, by construct, assimilation of satellite observation moves the posterior estimates (blue shaded areas and blue lines) toward the observations, and most of the observations are within the ensemble spread. The PBS accomplishes this by more heavily weighting ensemble members that more closely match the observations through the likelihood function embedded in the PBS formulation. The blue dashed lines in Figure 2 are more in agreement with the observations than are the prior estimates. The uncertainty of the posterior estimates is considerably smaller than that of the prior estimates due to the fact that the PBS reduces the weights of ensemble members that are far from the observations. Comparing the uncertainty (IQR) of the prior estimates with that of the posterior estimates in Figures 2 and S2–S5 shows that the PBS reduces the uncertainty of albedo estimates by about 70%. We also compared the ensemble

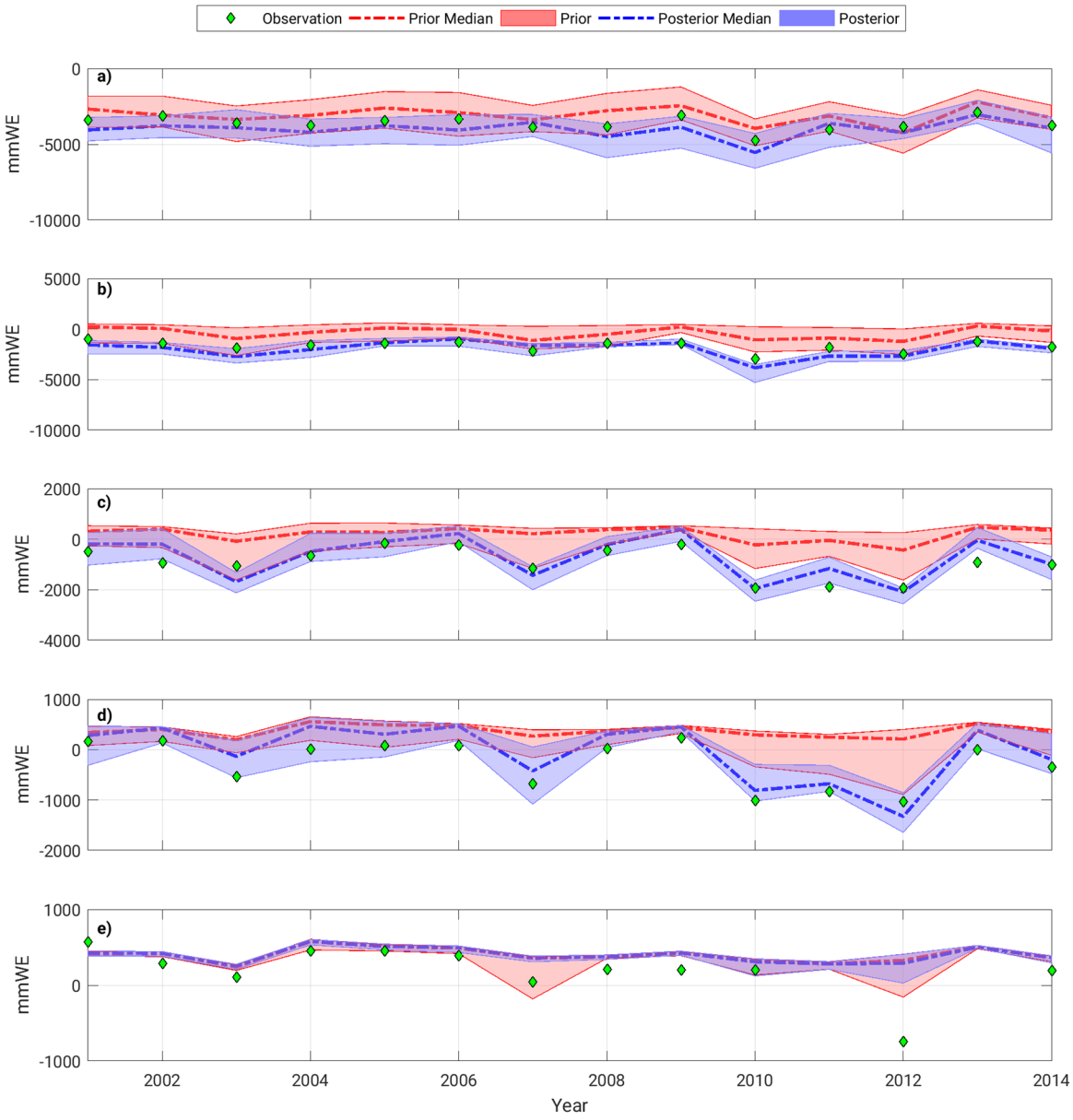

Figure 3. Time series of the surface mass balance (SMB) for grid-cell co-located with the (a) S5 and SHR, (b) S6 and S7, (c) S8, (d) S9, and (e) S10 stations along the K-transect. Values indicate year-over-year losses or gains in millimeter water equivalent (mmWE). The prior and posterior estimates are shown in red and blue, and in-situ SMB estimates along the K-transect are shown with green diamond.

spread of the posterior estimates and residuals of the measurements with respect to the posterior median for all computational grid cells. Results (not shown) indicate that residuals are within the posterior uncertainty.

Figure 3 shows the SMB for all model grid cells along the K-transect. To be able to compare the results with in-situ measurements at K-transect, we convert the time series of SMB to annual cumulative SMB which starts from September 1 of each year until August 31 of the following year (see Text S2 and Table S1). Positive SMB values indicate a net increase in mass, and negative values indicate net mass loss within a given year. The assimilation of 16-day albedo data results in an updated ensemble of SMB estimates that are more consistent with the observed albedo. As shown in Figure 3, the PBS generally improves the SMB: the posterior estimates move toward the independent SMB observations. Averaging over all grid cells, the PBS improves the SMB estimates by 51% for the simulation period of 2000–2014.

Although in many cases the posterior simulations appropriately match the K-transect observations, in some instances, the PBS estimates are inaccurate. For example, the posterior SMB values in Figure 3a are relatively smaller than observations (overestimating mass loss), and they are relatively larger than observations in Figure 3d (underestimating mass loss) (see Table S1 for quantitative values over the simulation period). By construct, the PBS attempts to fit the model albedo to the observation (Figure 2); however, whether this results in an improvement in SMB estimates depends on the accuracy of Crocus albedo module and how it is related to SMB and the other energy fluxes. Note that assimilation of albedo is not intended to solve the model errors that are not associated with albedo. While the albedo module has been updated to use the MAR albedo module, the prior simulation shows that there are considerable differences between the model and observations, which may be due in part to the inadequacy of the albedo module and the complex nonlinear relationship between albedo and SMB. Scatterplot of predicted albedo (prior and posterior) and MODIS 16-day albedo (see Figure S6) shows that assimilation of MODIS albedo reduces bias by shifting the median toward the observations and reduces the uncertainty as shown by the error bar. The scatter plots suggest that PBS significantly improves the albedo estimates in the grid cells with dominant melt processes (Figures S6a–S6h). Estimated SMB (both prior and posterior) and measured SMB in the grid cell co-located with the S10 station are close to each other. Since the model error is not significantly higher than the observation error, the PBS does not modify the prior estimates.

In the first computational grid cell, which is co-located with stations S5 and SHR, the PBS improves the posterior estimates of SMB by an average of 18.4% (Figure 3a) over the simulation period (2000–2014). Although the PBS significantly improved model albedo in this grid cell and posterior estimates of albedo are in good agreement with the observations (Figures 2h–2j), it overestimates the SMB in several years, for example, year 2008–2010 (Figure 3a). It is likely an indication that the nonlinear relationship between the albedo and SMB is biased toward overestimating of SMB when bare ice is exposed. In the grid cell co-located with stations S6 and S7 the PBS improves the SMB by 64% relative to the mean of independent in situ measurements (Figure 3b). In the grid cell co-located with stations S8 and S9. The PBS improves the posterior SMB of these two grid cells by 61% and 60%, respectively (Figures 3c and 3d). In the grid cell co-located with station S10 (in the percolation zone) the prior SMB simulation is very close to the in-situ observations (Figure 3e). The posterior simulation therefore shows similar result as the prior. Note that the temporal variation of albedo in this station during the melt period is very small (Figures 1b and 1c) and the difference between the observed albedo and the model estimated albedo (Figure S5) are significantly smaller over this snow-dominated grid cell than it is over grid cells with mixed snow and bare ice exposure. Therefore, there is limited information to be exploited by the PBS toward improving the SMB.

4. Discussion and Conclusion

The PBS data assimilation (reanalysis) method for estimating SMB using 15 years of MODIS albedo data was applied over the several MAR computational grid cells along K-transect stations. The main outcomes of this study are as follows:

The PBS considerably improves the posterior estimates of SMB over the prior estimates. While the prior simulations overestimate the albedo and consequently underestimate the surface mass balance, the posterior simulations result in albedo that closely match the satellite-derived albedo and provides posterior SMB estimates that are in good agreement with the in-situ SMB measurements. Indeed, the PBS reduces the RMSE of the SMB estimates by 51% and reduces bias by 95%. This is an encouraging result because the grid cells are distributed across different mass balance zones with different melt and accumulation processes.

Ensemble size and measurement error are two key inputs to the data assimilation methodology (Margulis et al., 2016). A sensitivity analysis (see Text S3) showed that by increasing the number of ensembles from 50 to 300 (computational cost increases by 600%) the RMSE of the SMB estimates improves by 8%. These results support the argument that an ensemble size of $N = 100$ replicates seems sufficient for SMB reanalysis in this study.

The measurement error indicates the degree of trust in the observation relative to the prior estimates. The sensitivity analysis on measurement error (see Text S4, Table S2, Figures S7 and S8) showed that the posterior ensemble collapses when using small measurement error. This means one ensemble member (the

best fit) gets weight 1 and the rest of the weights are very close to zero. Based on results from the sensitivity analysis a measurement error standard deviation of 0.2 provides the best results. Considering the fact that the dynamic range of albedo is between 0.8 for fresh snow to around 0.3 for bare ice, a measurement error of 0.2 is not unrealistic. In the process of averaging the raw albedo to the model scale, the known error of the MODIS albedo product may change. Moreover, it is inevitable that some of the "measurement error" is indeed model-measurement representation error. So, we are acknowledging the representativeness issues and we think inflating the measurement error could be attributed to this issue (see Text S5, Table S3, Figures S9–S14).

A close examination of the primary results shows that some ensemble members in the grid cells close to the ice sheet margin do not experience the exposure of bare ice during the summer. Therefore, time series of summertime albedo show bimodal distributions wherein the ensemble members experiencing bare ice exposure cluster around a model-defined ice albedo and other ensemble members cluster around a typical snow albedo value. These bimodal distributions result in the PBS failing to find a robust fit. To mitigate this issue and smooth transition between the replicates, each ensemble member uses a different, randomly assigned ice albedo between 0.35 and 0.65 instead of a constant ice albedo. This strategy results in intermediate albedo values when grid cells contain a combination of snow and ice during the summertime. A combination of a continuous transition between the ensemble members and higher measurement error help to prevent ensemble collapse and provides a robust fit to observations (Figure 2).

Results presented in this work are obtained by comparing the SMB of the grid cells with the in-situ SMB measurements at K-transect stations. As described above, there are two SMB stations in each of the first two model grid cells from the edge of the ice sheet along the K-transect and above 95% of the SMB along the K-transect happens in these two grid cells. Therefore, using observations from two in-situ stations over relatively flat topography with moderate slope can significantly reduce the error of evaluation (i.e., point scale verses grid cell scale). However, we acknowledge that in-situ point scale measurements may not be adequate for comparison against model grid cells of 25 km^2.

Results from the proof-of-concept study (Navari et al., 2018) and results presented in this work suggest that using the PBS framework is an alternative solution for improving estimates of SMB over the traditional deterministic modeling solutions or pure remote sensing solutions. The results also suggest that the methodology can be applied over the entire GrIS using MODIS albedo observations (improving in particular the bare ice albedo value used by the model) to generate an improved reanalysis of SMB estimates. It should also be highlighted that the reanalysis framework presented in this work is a general framework and could, therefore, be utilized over other domains such as the Antarctica ice sheet or mountainous glaciers. Finally, it is also important to note, that in view of the sensitivity of the model-based SMB estimations to the used bare ice albedo, a mean MODIS based bare ice albedo is notably used in RACMO in the aim of improving its SMB estimation (Noël et al., 2016).

Data Availability Statement

The MODIS data sets were downloaded from https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table. All preprocessing and data assimilation codes and evaluation data are available at https://www.hydroshare.org/resource/a8c5e72010064d3cb52522463fb42a9b/.

Acknowledgments
This work was partially sponsored by NASA through grant 80NSSC19K0304. Field work along the K-transect has been funded by the Dutch Polar Program through various grants.

References

Alexander, P. M., Tedesco, M., Fettweis, X., van de Wal, R. S. W., Smeets, C. J. P. P., & van de Broeke, M. R. (2014). Assessing spatiotemporal variability and trends in modelled and measured Greenland ice sheet albedo (2000–2013). *Cryosphere*, *8*, 2293–312. https://doi.org/10.5194/tc-8-2293-2014

Allen, R. G., Trezza, R., & Tasumi, M. (2006). Analytical integrated functions for daily solar radiation on slopes. *Agricultural and Forest Meteorology*, *139*(1), 55–73. https://doi.org/10.1016/j.agrformet.2006.05.012

Cluzet, B., Lafaysse, M., Cosme, E., Albergel, C., Meunier, L.-F., & Dumont, M. (2021). CrocO_v1.0: A particle filter to assimilate snowpack observations in a spatialised framework. *Geoscientific Model Development*, *14*, 1595–1614. https://doi.org/10.5194/gmd-14-1595-2021

Enderlin, E. M., Howat, I. M., Jeong, S., Noh, M.-J., van Angelen, J. H., & van den Broeke, M. R. (2014). An improved mass budget for the Greenland ice sheet. *Geophysical Research Letters*, *41*, 866–872. https://doi.org/10.1002/2013GL059010

Fettweis, X., Box, J. E., Agosta, C., Amory, C., Kittel, C., Lang, C., & Gallée, H. (2017). Reconstructions of the 1900-2015 Greenland ice sheet surface mass balance using the regional climate MAR model. *Cryosphere*, *11*(2), 1015–1033. https://doi.org/10.5194/tc-11-1015-2017

Fettweis, X., Franco, B., Tedesco, M., van Angelen, J. H., Lenaerts, J. T. M., van den Broeke, M. R., & Gallée, H. (2013). Estimating the Greenland ice sheet surface mass balance contribution to future sea level rise using the regional atmospheric climate model MAR. *The Cryosphere*, *7*(2), 469–489. https://doi.org/10.5194/tc-7-469-2013

Fettweis, X., Hofer, S., Krebs-Kanzow, U., Amory, C., Aoki, T., Berends, C. J., et al. (2020). GrSMBMIP: Intercomparison of the modelled 1980–2012 surface mass balance over the Greenland ice sheet. *The Cryosphere*, *14*, 3935–3958. https://doi.org/10.5194/tc-14-3935-2020

Gallée, H., & Duynkerke, P. (1997). Air-snow interactions and the surface energy and mass balance over the melting zone of west Greenland during the Greenland ice margin experiment. *Journal of Geophysical Research*, *102*(D12), 13813–13824. https://doi.org/10.1029/96JD03358

Gallée, H., & Schayes, G. (1994). Development of a three-dimensional meso-γ primitive equation model: Katabatic winds simulation in the area of Terra Nova Bay, Antarctica. *Monthly Weather Review*, *122*(4), 671–685. https://doi.org/10.1175/1520-0493(1994)122<0671:DOATDM>2.0.CO;2

Kumar, S. V., Dong, J., Peters-Lidard, C. D., Mocko, D. M., & Gomez, B. (2017). Role of forcing uncertainty and background model error characterization in snow data assimilation. *Hydrology and Earth System Sciences*, *21*, 2637–2647. https://doi.org/10.5194/hess-21-2637-2017

Larour, E., Utke, J., Csatho, B., Schenk, A., Seroussi, H., Morlighem, M., et al. (2014). Inferred basal friction and surface mass balance of the Northeast Greenland ice stream using data assimilation of ICESat (Ice Cloud and land Elevation Satellite) surface altimetry and ISSM (Ice Sheet System Model). *The Cryosphere*, *8*, 2335–2351. https://doi.org/10.5194/tc-8-2335-2014

Lenaerts, J. T. M., Medley, B., van den Broeke, M. R., & Wouters, B. (2019). Observing and modeling ice sheet surface mass balance. *Reviews of Geophysics*, *57*(2), 376–420. https://doi.org/10.1029/2018RG000622

Margulis, S. A., Cortés, G., Girotto, M., & Durand, M. (2016). A landsat-era Sierra Nevada snow reanalysis (1985-2015). *Journal of Hydrometeorology*, *17*(4), 1203–1221. https://doi.org/10.1175/JHM-D-15-0177.1

Margulis, S. A., Girotto, M., Cortés, G., & Durand, M. (2015). A particle batch smoother approach to snow water equivalent estimation. *Journal of Hydrometeorology*, *16*, 150504130725006. https://doi.org/10.1175/JHM-D-14-0177.1

Mouginot, J., Rignot, E., Bjørk, A. A., van den Broeke, M., Millan, R., Morlighem, M., et al. (2019). Forty-six years of Greenland ice sheet mass balance from 1972 to 2018. *Proceedings of the National Academy of Sciences*, *116*(19), 9239–9244. https://doi.org/10.1073/pnas.1904242116

Navari, M., Margulis, S. A., Bateni, S. M., Tedesco, M., Alexander, P., & Fettweis, X. (2016). Feasibility of improving a priori regional climate model estimates of Greenland ice sheet surface mass loss through assimilation of measured ice surface temperatures. *The Cryosphere*, *10*(1), 103–120. https://doi.org/10.5194/tc-10-103-2016

Navari, M., Margulis, S. A., Tedesco, M., Fettweis, X., & Alexander, P. M. (2018). Improving Greenland surface mass balance estimates through the assimilation of MODIS Albedo: A case study along the K-transect. *Geophysical Research Letters*, *45*(13), 6549–6556. https://doi.org/10.1029/2018GL078448

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

Noël, B., van de Berg, W. J., van Wessem, J. M., van Meijgaard, E., van As, D., Lenaerts, J. T. M., et al. (2018). Modelling the climate and surface mass balance of polar ice sheets using RACMO2 –Part 1: Greenland (1958–2016). *The Cryosphere*, *12*, 811–831. https://doi.org/10.5194/tc-12-811-2018

Rignot, E., Velicogna, I., van den Broeke, M. R., Monaghan, A., & Lenaerts, J. T. M. (2011). Acceleration of the contribution of the Greenland and Antarctic ice sheets to sea level rise. *Geophysical Research Letters*, *38*, L05503. https://doi.org/10.1029/2011GL046583

Riihelä, A., King, M. D., & Anttila, K. (2019). The surface albedo of the Greenland ice sheet between 1982 and 2015 from the CLARA-A2 dataset and its relationship to the ice sheet's surface mass balance. *The Cryosphere*, *13*, 2597–2614. https://doi.org/10.5194/tc-13-2597-2019

Ryan, J. C., Hubbard, A., Stibal, M., Irvine-Fynn, T. D., Cook, J., Smith, L. C., & Box, J. (2018). Dark zone of the Greenland ice sheet controlled by distributed biologically-active impurities. *Nature Communications*, *9*(1), 1–10. https://doi.org/10.1038/s41467-018-03353-2

Schaaf, C., & Wang, Z. (2015). *MCD43A4 MODIS/Terra+Aqua BRDF/Albedo Nadir BRDF Adjusted Ref Daily L3 Global - 500m V006* (Data set). NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MODIS/MCD43A4.006

Sellevold, R., Kampenhout, L., Lenaerts, J. T., Noël, B., Lipscomb, W., & Vizcaino, M. (2019). Surface mass balance downscaling through elevation classes in an earth system model: Analysis, evaluation and impacts on the simulated climate. *The Cryosphere Discussions*, *6*, 1–25. https://doi.org/10.5194/tc-2019-122

Smeets, P. C. J. P., Kuipers Munneke, P., van As, D., van den Broeke, M. R., Boot, W., Oerlemans, H., & van den Wal, R. S. W. (2018). The K-transect in west Greenland: Automatic weather station data (1993–2016). *Arctic Antarctic and Alpine Research*, *50*(1), S100002. https://doi.org/10.1080/15230430.2017.1420954

Stroeve, J., Box, J. E., Gao, F., Liang, S., Nolin, A., & Schaaf, C. (2005). Accuracy assessment of the MODIS 16-day albedo product for snow: Comparisons with Greenland in situ measurements. *Remote Sensing of Environment*, *94*(1), 46–60. https://doi.org/10.1016/j.rse.2004.09.001

Stroeve, J. C., Box, J. E., & Haran, T. (2006). Evaluation of the MODIS (MOD10A1) daily snow albedo product over the Greenland ice sheet. *Remote Sensing of Environment*, *105*(2), 155–171. https://doi.org/10.1016/j.rse.2006.06.009

van Angelen, J. H., van den Broeke, M. R., Wouters, B., & Lenaerts, J. T. M. (2013). Contemporary (1960–2012) evolution of the climate and surface mass balance of the Greenland ice sheet. *Surveys in Geophysics*, *35*(5), 1155–1174. https://doi.org/10.1007/s10712-013-9261-z

Van de Wal, R. S. W., Boot, W., Smeets, C. J. P. P., Snellen, H., van den Broeke, M. R., & Oerlemans, J. (2012). Twenty-one years of mass balance observations along the K-transect, West Greenland. *Earth System Science Data*, *4*(1), 31–35. https://doi.org/10.5194/essd-4-31-2012

Vernon, C. L., Bamber, J. L., Box, J. E., van den Broeke, M. R., Fettweis, X., Hanna, E., & Huybrechts, P. (2013). Surface mass balance model intercomparison for the Greenland ice sheet. *The Cryosphere*, *7*(2), 599–614. https://doi.org/10.5194/tc-7-599-2013

Vionnet, V., Brun, E., Morin, S., Boone, A., Faroux, S., Le Moigne, P., et al. (2012). The detailed snowpack scheme Crocus and its implementation in SURFEX v7.2, Geosci. *Model Dev*, *5*, 773–791. https://doi.org/10.5194/gmd-5-773-2012

Wientjes, I. G. M., Van de Wal, R. S. W., Reichart, G. J., Sluijs, A., & Oerlemans, J. (2011). Dust from the dark region in the western ablation zone of the Greenland ice sheet. *The Cryosphere*, *5*(3), 589–601. https://doi.org/10.5194/tc-5-589-2011

References From the Supporting Information

De Lannoy, G. J. M., Reichle, R. H., Houser, P. R., Arsenault, K. R., Verhoest, N. E. C., & Pauwels, V. R. N. (2010). Satellite-scale snow water equivalent assimilation into a high-resolution land surface model. *Journal of Hydrometeorology*, *11*(2), 352–369. https://doi.org/10.1175/2009JHM1192.1

Forman, B. A., & Margulis, S. A. (2010). Assimilation of multiresolution radiation products into a downwelling surface radiation model: 1. Prior ensemble implementation. *Journal of Geophysical Research*, *115*(D22), 1–14. https://doi.org/10.1029/2010JD013920

Lafaysse, M., Cluzet, B., Dumont, M., Lejeune, Y., Vionnet, V., & Morin, S. (2017). A multiphysical ensemble system of numerical snow modelling. *The Cryosphere*, *11*, 1173–1198. https://doi.org/10.5194/tc-11-1173-2017

Zhou, Y., McLaughlin, D., & Entekhabi, D. (2006). Assessing the performance of the ensemble Kalman filter for land surface data assimilation. *Monthly Weather Review*, *134*, 2128–2142. https://doi.org/10.1175/MWR3153.1