Response to Reviewer Comments on "Improving a priori regional climate model estimates of Greenland ice sheet surface mass loss through assimilation of measured ice surface temperatures" *by M.Navari et al.*

We would like to thank the reviewers for their constructive comments. We have tried our best to make the text as clear as possible based on the comments we have received from the reviewers.

The reviewer comments are provided below in black font and responses are provided in blue font. Sentences with quotation marks are proposed new text to be added to the revised manuscript and any references to page numbers and line numbers refer to the published manuscript (http://www.the-cryosphere-discuss.net/9/3205/2015/tcd-9-3205-2015.pdf).

Response to Reviewer 1

General comments

… . However, the paper is too long, as is the reference list. The paper is written in a rather technical style, which in combination with the length makes it a hard read for the non-specialist. If these and the issues raised below are addressed, the paper can be published in TC after which, what I believe are, relatively minor revisions.

Major comments

1. The introduction at places reads like a review article, which is also reflected in the amount of citations. Please select only the most relevant studies to cite, and also try to avoid duplicate citations (i.e. citing the same paper several times). To shorten the remainder of the paper and improve its readability, consider moving part of the methods to an appendix.

We have attempted to make the text more concise by removing any duplicate citations and shortening the introduction. The proposed revised Introduction is as follows:

> **"Introduction and Background**
> The Greenland ice sheet (GrIS) has recently experienced thinning of the marginal ice (e.g. Straneo et al. 2013, Khan et al., 2014), thickening of its interior (e.g. Johannessen et al., 2005; Fettweis, 2007), acceleration and increase in ice discharge from many of Greenland's outlet glaciers (e.g. Rignot et al., 2008; Wouters et al., 2013), and enhanced surface melt (e.g. Tedesco et al., 2013; Vernon et al., 2013). The melting of the GrIS due to increased temperature has the potential to affect deep ocean circulation, and sea level rise (Hanna et al., 2005; Fettweis et al., 2007; Tedesco 2007, Rahmstorf et al., 2015). While van Angelen et al. (2012) and Fettweis et al. (2013) predict that meltwater runoff will be the dominant mass loss process in the future due to the retreat of the tidewater glaciers above sea level; a recent study showing that the dynamic mass loss was reduced from 58% before 2005 to 32% for the period between 2009 and 2012 (Enderlin et al., 2014).
> Many studies (e.g. van de Wal et al., 2012) have taken advantage of in situ measurements to provide a direct point-scale estimate of the surface mass balance (SMB, i.e. the difference between accumulation and ablation). However, with these limited in situ measurements alone, large-scale mapping of the GrIS surface mass fluxes (i.e. precipitation, evaporation, sublimation,

condensation, and runoff) is impossible. The availability of remote sensing data and/or products has taken GrIS from a remote ‘‘data poor’’ region that is reliant mostly on sparse in situ measurements to a potentially ‘‘data rich’’ environment. In this regard, a key research objective is to better understand how such data can be optimally leveraged for quantitatively estimating the surface mass balance (SMB) and its associated fluxes.

Surface remote sensing data and products (i.e., surface or skin temperature, multi-frequency brightness temperature, and albedo) have been used to characterize various aspects of SMB such as snow melt, melt extent, melt duration, new snow, extreme melt events (e.g. Abdalati and Steffen, 1995; Tedesco et al., 2011; Box et al., 2012; Hall et al., 2013). However, the relationship between surface remote sensing data/products and surface mass fluxes are most often indirect and implicit. For example, ice surface temperature can be indicative of melt, but it fails to quantitatively estimate the volume of meltwater produced. More importantly, other surface mass fluxes such as evaporation, condensation, sublimation, and runoff cannot be directly quantified via remote sensing. This makes the possibility of quantitatively characterizing the surface mass fluxes from remote sensing retrieval algorithms difficult if not impossible. It can therefore be argued that the information content of remotely sensed data remains underutilized due to indirect and implicit links between the various data streams and surface mass fluxes.

Given the limitations of the observation-based methods, numerical models offer an alternative mechanism to quantify the GrIS surface mass fluxes. Several model-based approaches have been used to characterize the spatio-temporal variability of the GrIS surface mass fluxes in both historical and future contexts (e.g. Hanna et al., 2011, Box et al., 2006; Fettweis, 2011; Ettema et. al., 2009; Lewis and Smith, 2009; Vernon et al. 2013; Franco et al. 2013). Although the aforementioned methodologies have provided the ability to estimate the GrIS SMB and related fluxes, their estimates vary considerably, mainly due to the different physics parameterizations in the models and simplifying assumptions, the inherent uncertainty of each method, error in model and input data, and the length of data records (e.g. Rignot et al., 2011; Vernon et al., 2013; Smith et al., 2015). Therefore, it is imperative to design techniques that bridge the gap between different methods by merging relevant data streams with a physical model with the aim of better spatial-temporal characterization of the GrIS surface mass fluxes. In this study, we provide an example of taking advantage of information in the relevant data streams to provide a better spatial-temporal characterization of the model outputs (i.e., the GrIS surface mass fluxes). This can be done using a data assimilation approach which attempts to merge model estimates with measurements in an optimal way (Evensen, 2009).”

Regarding the Methods section, since this section explains the EnBS which is the main contribution of this work we believe it should be in the main body of the paper.

2. The fact that this paper presents a proof of concept should be reflected in the title ("e.g. ...feasibility of...") and also discussed in the main text.

A new title is proposed to reflect the study properly as following:

> **“Feasibility of improving a priori regional climate model estimates of Greenland ice sheet surface mass loss through assimilation of measured ice surface temperatures”**

3. Nothing is said about how data assimilation in general could help improve the forcing models; i.e. by assimilating vertical profiles of humidity (using for instance radio occultation) into an RCM, its

prediction of e.g. precipitation could also be improved. Could results be further enhanced by including MODIS albedo? Please comment.

The following sentences are proposed for addition to the text on p. 3215, l. 4:

> … (1989, 1992). "Assimilation of data into an RCM is another option for attempting to improve RCM fields (such as precipitation, for example), but beyond the scope of this work. The focus of this work is improving of surface mass fluxes using RCM outputs and assimilation of a surface remote sensing data stream." Furthermore, the use of a fully coupled MAR-CROCUS system to …

We chose to focus on the assimilation of a single remote sensing data stream for clarity and to better understand the potential improvement deriving from the use of the merging of assimilation techniques with RCM over Greenland. Future work will investigate assimilation of other data including albedo, passive microwave data, etc. This is indicated on p. 3232, l. 6, l. 14.

4. p. 3214, l. 6: "Surface and sub-surface melting (which ultimately contribute to runoff) are dictated by the evolving snow temperature driven by energy inputs." This statement is not formally correct. Surface melting is driven by the SEB imbalance once Ts reaches the melting point, after which it remains constant. So during melting, variations in Ts cannot be used to infer melt rate.

To clarify we propose the following edits:

> "The temporal evolution of snow temperature in a vertical snow column is constrained by the conservation of energy equation, i.e. (Brun et al. 1989):
>
> $$\frac{\partial(\rho c_p T)}{\partial t} = \frac{\partial^2(\kappa T)}{\partial z^2} + q \qquad (2)$$
>
> where ρ is the snow density, c_p is the snow heat capacity, T is the snow temperature at depth z and time t, and κ is the snow heat conductivity, and q represents a sink (melt) and source (refreezing). It is worth noting that Eq. (2) is valid for T<273.15K; any energy inputs that would raise the temperature beyond freezing instead contribute directly to melt. Equation (2) is subject to the surface energy balance as a boundary condition, which is the key driver of the snowpack energy budget."

5. Moreover, subsurface melting in a model is only possible when subsurface heat sources are allowed, such as the penetration of shortwave radiation; it is not clear whether this is the case in CROCUS. If so, please mention it; if not, subsurface melting cannot take place.

We propose adding a paragraph to the manuscript (p. 3215, l. 3) to explain the CROCUS model in more detail to address this comment:

> "CROCUS computes albedo and absorbed energy in each layer for three spectral bands (i.e. visible, and two near infrared bands). The capability of the model to partition the incident solar radiation between the layers allows melt occurs on multiple depths."

6. p. 3214, l. 10: Equation 2 misses a source term associated with the refreezing of percolating meltwater. It is true that the SEB determines the upper boundary condition (Ts) to force snow temperature, but in Eq. 3 melt is (I assume, because it is missing from the equation) incorporated in Qg, which is normally assigned to the subsurface conductive heat flux. Not explicitly including melt in the SEB equation is not logical, as Qg can still be nonzero under non-melting conditions. Please consider to reformulate the SEB equation.

Unfortunately, there was a typo in Equation 2. The corrected equation is:

$$\frac{\partial(\rho c_p T)}{\partial t} = \frac{\partial^2(\kappa T)}{\partial z^2} + q \qquad (2)$$

Please see response to Reviewer #1, major comment # 4.

We propose to correct the Equation 3 as follows to include melt energy as follows:

"$Q_M = R_s^{\downarrow}(1-\alpha) + R_l^{\downarrow} - R_l^{\uparrow} + Q_{SH} + Q_{LH} + Q_G$

where "Q_M is the melt energy, " $R_s^{\downarrow}$ is the downward shortwave radiation, α is the …"

Technical comments

1. Abstract, l. 7: "... there is considerable disparity between the results from different methodologies that need to be addressed." But these disparities are not addressed in this paper.

Considering this comment we propose editing the sentence as follows:

> "Though the estimates of the GrIS surface mass fluxes have improved significantly over the last decade, there is still considerable disparity between the results from different methodologies (e.g., Rae et al., 2012; Vernon et al., 2013,). Data assimilation approach can merge information from different methodologies in a consistent way to improve the GrIS surface mass fluxes. In this study, an Ensemble Batch Smoother data assimilation approach was developed to assess the feasibility of generating a reanalysis estimate of the GrIS surface mass fluxes via integrating remotely sensed ice surface temperature measurements with a regional climate model (a priori) estimate. "

2. p. 3208, l. 11: " While recent estimates..." The estimates listed are made for certain periods; it is now well known that over recent years surface processes have outpaced dynamical changes (e.g. Enderlin et al., Geophysical Research Letters, 2014).

The proposed change to the manuscript is as follows:

> "While van Angelen et al. (2012) and Fettweis et al. (2013) predict that meltwater runoff will be the dominant mass loss process in the future due to the retreat of the tidewater glaciers above sea level; a recent study showing that the dynamic mass loss was reduced from 58% before 2005 to 32% for the period between 2009 and 2012 (Enderlin et al., 2014)."

3. p. 3213, l. 6: Please consider reducing the amount of citations.

The proposed revised sentence is as follows:

> "The version of the model used here (i.e. MARv2) has been applied extensively over the GrIS and is described in more detail in previous studies (e.g., Lefebre et al., 2003; Fettweis et al., 2005)."

4. p. 3213: is MAR not also forced at the top by ERA-Interim? If so, please mention this.

MAR is forced at the top in the stratosphere but it is totally free in the troposphere. In p. 3213, l. 12 - l. 17 it is mentioned that:

The ERA-Interim reanalysis from the European Centre for Medium-Range Weather Forecasts (ECMWF) was used to initialize the MAR meteorological fields at the beginning of the simulation (1979) and to force the atmospheric lateral boundaries as well as the oceanic conditions (surface temperature and sea ice extent) every 6 h over 1979– 2010.

5. p. 3214, l. 6: what is meant by 'subsurface melting'? Is shortwave radiation allowed to penetrate the snow/ice? Otherwise, no heat sources would be available to enable subsurface melt in a model.

Please see response to Reviewer #1 major comment #5.

6. p. 3227, l. 19: "Sublimation and evaporation play an important role in the GrIS surface mass loss (Lenaerts et al., 2012) and after runoff are the main components of the GrIS SML." This is true for drifting snow sublimation, which was included in Lenaerts (2012) but not in CROCUS or MAR. Ordinary surface sublimation is typically three times smaller than surface and drifting snow sublimation together.

We propose replacing l. 19 - l. 20 with the following lines to address the comment:

> "Sublimation and evaporation play an important role in the GrIS surface mass loss. However, it should be noted that MAR and CROCUS estimate surface sublimation which is considerably smaller than drifting snow sublimation. Lenaerts et al. (2012) reported for the period 1960-2011 on average surface sublimation is responsible for 40% of total sublimation and drifting snow sublimation is responsible for another 60%."

7. p. 3218, l. 22: "uncertainty of precipitation estimates from different modeling frameworks are less than that of the other terms (Fettweis, 2007)". Vernon and others (2013) show that this is not true in general; moreover, large intermodel differences occur also in melting, refreezing....

We propose removing this sentence to be consistent with recent studies. We showed that data assimilation framework improves the estimates of surface mass fluxes using surface remote sensing data. In other words, independent of chosen model the data assimilation framework moves the model-estimated states and fluxes toward the true estimates (i.e. satellite measurements). Therefore intermodal variability will not considerably affect the data assimilation results.

Response to Reviewer 2

General

The title describes the content of the manuscript well, although it should be noted that the manuscript contains evaluation of the methods only. The manuscript does not contain an application of the method using real data. It focuses entirely on the results using synthetic data.

We propose editing the title to reflect that the paper contains an evaluation of the method (please see response to Reviewer #1, Major comment #2).

In general the manuscript is well written, some parts needs to be clarified. I've read it with interest although I was left with one major concern.

Major comment

1. My primary concern is that the synthetic truths used were, albeit outliers, results from the CROCUS model driven by adjusted MAR data. Hence, this synthetic truth is within the state space of trajectories accessible by CROCUS. It is by no means granted that the real trajectory of the surface state lies within this space reachable by CROCUS.
2. If not, one can assimilate, but it might possibly not help enough to approach the true state evolution. This is a concern for the energy balance (SEB) terms and temperature (Table 2), but posterior SEB and temperature estimates after assimilation with real satellite derived ice sheet temperature (IST) can at least be evaluated using, for example, GC-net data. However, runoff is much more dependent on hardly-to-evaluate model physics than the SEB and moreover runoff is very hard to evaluate. Hence, it will be extremely hard to assess the error and uncertainty in runoff with actual observations once real ISL is used. I expect the authors in that case to look at this paper, so the uncertainty estimates presented here matters. However, given that the synthetic true is a CROCUS state too, I don't buy the presented biases and RMSEs for runoff as a relevant number for test with true data.

Although it is not a full remedy for the problems sketched above, I request to authors to repeat the OSSE using SEB and SMB data from another RCM than MAR/CROCUS, e.g. HiRHAM or RACMO2. I know that the required high-temporal resolution data are not floating around but I guess the authors have the right connections to get these data.
This assessment can then presented in the added paragraph 5.4.
I know that this addition requires a significant effort, but I believe this would improve strongly the assessment of what could be expected from this method.

The reviewer raises a valid point, but considering the data assimilation approach developed in this paper is being applied to the GrIS for the first time and the focus is on basic proof-of-concept, we believe that it is well beyond the scope of this work to include other models to generate a synthetic truth for testing the assimilation. Setting up and using such models are non-trivial and in some cases the models are not open-source and therefore not availble for off-the-shelf use.

While we acknowledge that using the synthetic true from the state space trajectory of MAR-CROCUS might be somewhat optimistic, we were careful to choose outliers so that the true was significantly different from the nominal prior.

Additionally, model intercomparison (e.g., Vernon et al., 2013, Fettweis et al., 2013, Rae et al., 2012) shows considerable similarities (i.e. trend and features in the time series) between the results from well developed RCMs (e.g., MAR, RACMO2, PMM5) despite the differences in the integrated SMB values. It has also been shown that the surface mass fluxes from these models are highly correlated and the differences between the results are within the interannual variability of models. Therefore, it can be argued that the selected true using the other RCMs is likely to fall into the state space trajectory of MAR-CROCUS ensemble estimates. Moreover, while recently HIRHAM has been coupled with the land surface model MIKE-SHE (Larsen et al., 2014) we haven't found a validated application of this new version of the model in the Polar Regions. Furthermore, in section 5.3 "Sensitivity to the synthetic truth values" we showed that even for the extreme cases where the real true stats fall beyond the chosen values, the developed algorithm can be used to retrieve the true states. Therefore, we chose to use the synthetic truth from MAR rather than RACOM2 and HIRHAM.

To address the reviewer's concern, we propose adding a caveat to the manuscript by adding the following paragraph in the manuscript (p. 3221, l. 1) while we have already suggested such an effort could be done as future work (p. 3232, l. 5-7) and we also propose adding a paragraph to the text in p. 3231, l. 25 (please see response to Reviewer #2, Comments related to text parts #8).

> "In the OSSE system, traditionally the synthetic true ensemble is chosen from state space trajectory of the forward model (e.g., Crow and Van Loon, 2006; Durand and Margulis, 2006; Bateni et al., 2013). While an alternative approach could involve choosing the synthetic truth from the trajectory space of another well developed RCM model, running multiple RCM models to generate a synthetic truth is prohibitive."

Other concerns

1. Precipitation: If got it right, precipitation has been varied during the tests, but precipitation results are not discussed at all. It is not so easy to evaluate real-world precipitation but within the experiment design you can. Yes, IST has only a very weak link to precipitation but now precipitation remains a free variable to change, allowing taking very unrealistic values. Your figures should show that this deteriorating of results is not the case. After all, precipitation affects the SML through albedo and refreezing capacity. Precipitation must thus be added in Figure 3, 4 and 9, and, if you take this really seriously, discussed in a figure similar to figures 5 to 8.

MAR precipitation is perturbed around its nominal value to take into account uncertainty of the precipitation which is a base for the ensemble approach. This means in each realization CROCUS uses γ_j percent of the MAR nominal precipitation (equation 5a). Despite the extensive effort using different experimental designs, data assimilation framework used in this work was not able to update the precipitation robustly, therefore, the focus of the study shifted from estimating the surface mass balance (SMB) toward estimating the surface mass loss (SML) which to a large degree is independent of precipitation. In the other words, precipitation does not directly affect the SML fluxes in a sizeable way and the effects will be indirect through the albedo and energy fluxes due to precipitation. To take into account these indirect effects, we chose to perturb the precipitation instead of using nominal MAR precipitation. As the reviewer stated, precipitation is a free variable, however, to prevent unrealistic precipitation values we carefully perturbed the precipitation to represent the real uncertainty of the precipitation. Similar perturbation variables have been frequently reported in the literature (e.g. De Lannoy et al., 2012, and Girotto et al., 2014). In addition, Fettweis 2006 compare the MAR precipitation in 1990 with 12 coastal weather stations and reported a mean and standard deviation of 428 mm and 235 mm respectively. coefficient of variation (CV) of precipitation from this study is in close agreement with the value (i.e. CV=0.5) we used in perturbation framework. Therefore, we believe that perturbed precipitation represents a realistic uncertainty of the precipitation over the GrIS.
Figure 3, 4 and 9 compare the prior and posterior states and fluxes with the truth. But for precipitation we did not update the precipitation (i.e. the prior and posterior precipitation is the same); therefore, there is

no posterior result to be compared with the prior precipitation and adding precipitation to these figures does not provide any information about the DA process.

We propose adding the following note to the manuscript (p. 3231, l. 24) to make this clear:

> …the precipitation flux was not updated in this context "(i.e. the prior and posterior precipitation is the same)."

2. At the sideline, GRACE data could be helpful to constrain regional precipitation and runoff on monthly timescales and longer when the method is applied on real IST data.

This would be a very interesting subject for future work.
We propose editing the text (p. 3232, l. 14) to reflect the fact that GRACE data could be included in the list of future data that could be assimilated.

> The data assimilation framework is general and could also include the potential application of assimilation of passive microwave, albedo "and even Gravity Recovery and Climate Experiment (GRACE) data to further constrain GrIS SMB estimates."

3. Runoff: Runoff is not a simple direct result from surface processes; snowpack processes seriously adapt runoff. The manuscript tends to be over detailed, but a description how CROCUS models runoff and which subsurface processes are modeled in CROCUS is missing at all. This should be added.

In this work a bulk "surface" mass and energy balance for each pixel were computed for surface layers (about top 10 meters). We propose clarifying the definition of the "surface layer" in the manuscript (p. 3213, l. 24-25) as follows:

> The bulk surface mass balance for each model pixel " (i.e., integrated over the top ~10 meters of the ice sheet) can be written as:"

In addition, we propose adding the following paragraph (p. 3215, l. 3) to clarify how CROCUS handles runoff:

> "In CROCUS each snow layer in the snow column is treated as a reservoir with a maximum water holding capacity of 5% of the pore volume. When the liquid water content (LWC) exceeds the threshold, excess water moves toward the layer below and the process continues until the water reaches the bottom layer and generates runoff. In addition, CROCUS takes into account changes in LWC due to snow melt, refreezing, and evaporation during a model time step."

4. For example, I got the feeling that runoff is allowed in the predefined ablation zone but excluded elsewhere. Is such a prior assumption justifiable for a method like this?

We did not impose any condition in CROCUS, and runoff is a direct result of CROCUS integration using perturbed (prior, posterior) meteorological data. Here, the GrIS mass balance zones are presented for visualization purpose only.

Comments related to text parts

1. p. 3211 l. 16-19 & Figure 1: Why is the border between the dry snow zone and the percolation zone no straight border? Furthermore, these zones are not mentioned later, only a difference between the ablation zone and the accumulation zone is made. So why are you introducing the percolation zone?

Using MAR nominal surface air temperature will result a continuous border between the two zones. Here, to be consistent with the data assimilation framework in which all results are presented based on the mean (median) of the ensemble of the estimates, we used the ensemble mean annual surface air temperature to draw the border between the dry snow zone and the percolation zone. That is the reason the border is not a continuous straight line.

The definition of the three mass balance zones was presented for illustration only. The focus of the paper is on the ablation zone.

2. Paragraph 3.4: I'm missing quite a few things here:
1. Equation 2: is there no refreezing in the subsurface model? In case of yes (no refreezing), is this not a major model shortcoming? In case of no (there is refreezing), why is it absent as heat source?

Please see response to Reviewer #1, Major Comment #4

2. Add information how Q_{sh} and Q_{lh} are depending on T and U and surface properties.
What kind of meteorological principles are applied?

We propose adding the following paragraph to the text on page 3214, l. 21:

> "The sensible/latent heat fluxes are the heat exchange between the surface and overlaying air due to the temperature/water vapor gradient between the surface and the reference-level (i.e. meteorological forcing variables). The fluxes are also modulated by wind speed through a typical conductance term. The ground heat flux is driven by the temperature difference between the surface temperature and subsurface layers, hence highly affect the ice/snow melt and runoff. Sensible/latent heat fluxes reduces the surface temperature and have cooling effects; in contrast ground heat flux warms the surface via conducting energy into the underlying surface."

3. How is melt generated? Is there radiation penetration implemented, in that case melt could occur on multiple depths. Otherwise, melt is modeled only for the uppermost layer, isn't it?

Please see response to Reviewer #1, Major Comment #5

4. Concluding, add a brief description of the physics in the subsurface model of CROCUS relevant for runoff estimates. Grain shape evolution (which is in CROCUS) is in this context not very relevant, but the implementation of percolation, retention and refreezing is relevant because you are intending to estimate runoff.

The physics used by CROCUS has been explained in Brun et al 1989, 1992 and we have referred the reader to these two papers. However, we propose adding the following paragraph to the manuscript (p. 3215, l. 3) to address the reviewer concern:

> "CROCUS is a 1D energy balance model consisting of a thermodynamic module, a water balance module taking into account the refreezing of meltwater, a turbulent module, a snow metamorphism module, a snow/ice discretization module and an integrated surface albedo module. CROCUS computes albedo and absorbed energy in each layer for three spectral bands (i.e. visible, and two near infrared bands). The capability of the model to partition the incident solar radiation between the layers allows melt occurs on multiple depths. In CROCUS each snow layer in the snow column is treated as a reservoir with a maximum water holding capacity of 5% of the pore volume. When the liquid water content (LWC) exceeds the threshold, excess water moves toward the layer below and the process continues until the water reaches the bottom layer and generates runoff. In addition, CROCUS takes into account changes in LWC due to snow melt, refreezing, and evaporation during a model time step. " The physics of CROCUS and its validation are detailed in Brun et al. (1989, 1992)."

5. p. 3217, l. 21 – p. 3218, l. 15: I was able to follow and understand for long how the method is constructed, but the concept of multiplicative coefficient as the states to be estimated remains unclear for me given the current text. Assuming that I'm representative for the TC readers – although I'm afraid that many readers stop understanding the method at an earlier point – I ask to clarify this part. Introduce a figure or scheme or whatever you need, but make this clear.

We propose adding the following lines to the text (p. 3218, l. 3) to make the concept more clear:

> …multiplicative coefficient as the states to be estimated. "In other words, the multiplicative coefficients have been used to transfer the nominal MAR forcing into probabilistic space (i.e. prior and posterior forcings). The DA algorithm uses IST measurements to condition the probability density function (pdf) of the prior multiplicative coefficients to compute the posterior pdf of the multiplicative coefficients."

6. p. 3227, l. 17: The term improvement factor is misleading, result aren't up to a factor 400 times better. Given the definition it has the same dimension as the variable of interest, so improvement rate is better. If you would like to present it as factor, you could divide the prior errors by the posterior errors.

We propose replacing the usage of "factor" with "metric".

7. 5.1 - 5.2: Although strictly spoken not a SML term, I'm missing a discussion of modeled snow/ice melt energy. In the set-up of CROCUS, melt energy is not a component of the SEB although the frozen surface is bound to the freezing temperature. Also, melt can happen at some depth. So, melt energy is not fully a SEB term too.
Having said this, melt energy is in my view a very important term to evaluate if the SEB is correct for ablation processes. Now, runoff is evaluated only but runoff estimates includes the effect of subsurface processes on the initial melt water flux. Yes, where runoff peaks, melt and runoff are almost equal, but for most sites refreezing mitigates some of the melt. Subsurface processes in snow are still rather unknown and extremely hard to evaluate (Even in situ observations won't tell easily if your percolation/refreezing model is correct). So, if the melt energy is estimated correctly but the subsurface model is err, the runoff is wrong. Or vice versa, an incorrect subsurface model can correct wrong melt water energy into a correct runoff flux.
Therefore, add to 5.1 a discussion of the (vertically integrated (?)) melt energy is improved in the posterior estimates. Yes, I expect that these results largely coincide with the results obtained for runoff (subsurface parameters aren't varied as the variables in Eq. 5), but that's a false guaranty. The real

subsurface processes are not automatically equal as modeled in CROCUS, that's why I request a repeat of the procedure using SEB and SMB data from another RCM.

The snow/ice model CROCUS takes into account the surface and subsurface snow processes (please see response to Reviewer #1, Major Comment #4 and #5). In this study, we simulated approximately the top 10 meters of snow/ice. In this context, this represents the "surface" mass and energy balance via the vertically integrated states and fluxes within this top layers of the ice sheet.

We propose editing the manuscript (p. 3215, l. 22) to address the definition of surface mass and energy balance:

> … Fettweis (2006), the bottom boundary condition was modified for simulating "approximately the top 10 meters of the ice sheets. In this context, this represents the "surface" mass and energy balance via the vertically integrated states and fluxes within these top layers of the ice sheet." This method consists of …

Please also see response to Reviewer #2, other concerns Comment #3.

We propose adding the following lines to the manuscript (p. 3226, l. 6) to address the second part of comment (add to 5.1 a discussion …):

> "Therefore, using the improved surface energy terms to force CROCUS improves vertically integrated melt energy and enhances the estimates of the states and fluxes over the vertical snow/ice column."

8. Section 6: The conclusions should be extended with the results coming from the new paragraph 5.4, a brief discussion of precipitation and a discussion of the uncertainty due to the fact that for most results CROCUS has been used to obtain the synthetic truth. Yes, the paper is a successful proof of concept to improve CROCUS results with respect to synthetic CROCUS data, but not yet a proof of concept that CRUCUS results can be improved compared to real world or other arbitrary but sensible SEB and SMB data.

This comment is in line with Reviewer #2, Major comment #1 and #2. In addition, we propose adding the following lines to conclusion (p. 3231, l. 25).

> "However, it should be noted that, using MAR-CROCUS to generate the synthetic truth might lead to optimistic results since the truth is taken from the same model. Mitigation of this was attempted by using an outlier for the truth. An expensive alternative, but worth pursuing in future work, would be to use other RCM models to generate the synthetic truth. That said, it can be argued that using another model such as RACMO2 to generate the true realization will not significantly affect the results because the synthetic truth from RACMO2 is likely to fall within the ensemble spread of MAR-CROCUS trajectory. The main reasons for that are (1) the SMB fluxes from MAR and RACMO2 are highly correlated (Fettweis et al., 2013), (2) the trends of SMB fluxes from two models are very similar Vernon et al., (2013). Furthermore, sensitivity analysis shows that the proposed algorithm is able to retrieve the synthetic truth for the extreme cases where the real true stats fall beyond the chosen values."

Textual comments

1. p. 3207, l. 4 & l. 26: remove "unprecedented' because it is untrue on geological time scales.

Please see response to Reviewer #1, major comment # 1.

2. p. 3208, l.2: You could also add Johannessen et al, Science, 310 (2005).

Please see response to Reviewer #1, major comment # 1.

3. p. 3208, l. 22: it is not "difficult, if not impossible". It's simply impossible in my view.

Please see response to Reviewer #1, major comment # 1.

4. p. 3209, l. 11-13: Rephrase this a bit to make clearer that people haven't made use of the indirect or implicit information in remotely sensed data.

We propose to rearrange the paragraph to address this comment:

> "However, the relationship between surface remote sensing data/products and surface mass fluxes are most often indirect and implicit. For example, ice surface temperature can be indicative of melt, but it fails to quantitatively estimate the amount of melt. More importantly, other surface mass fluxes such as evaporation, condensation, sublimation, and runoff cannot be directly quantified via remote sensing. This indirect relationship makes the possibility of quantitatively characterizing the surface mass fluxes from remote sensing retrieval algorithms difficult if not impossible. It can therefore be argued that the information content of remotely sensed data remains underutilized due to indirect and implicit links between the various data streams and surface mass fluxes."

5. p. 3212, l. 10: maybe add: . . . IST, of all remote sensing products available, may contain the most information about physical processes. . .

We propose to edit the text as follows:

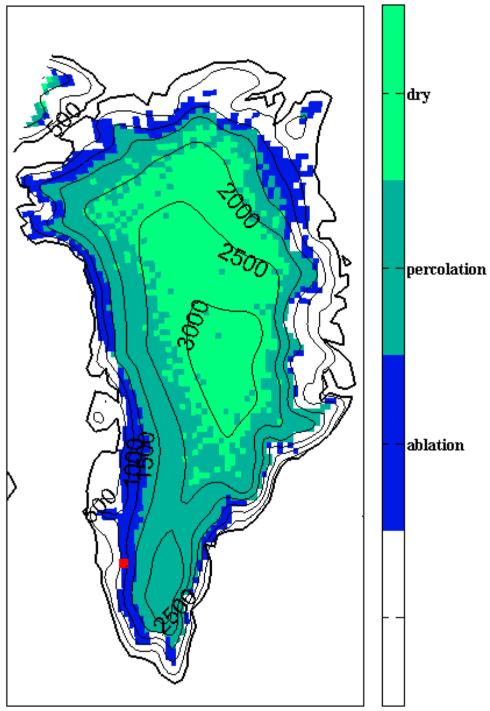
> These facts support the idea that clear-sky IST", of all remote sensing products available," may contain the most information about …

6. p. 3215, l.5-11: In MAR CROCUS is run online for a good reason. There is a feedback between the surface state and the atmospheric conditions (primarily through albedo).Is there any check that posterior energy fluxes are realistic given this atmospheric feedback?

Coupled MAR-CROCUS allows two-way interaction between the MAR and CROCUS. However, there is a tradeoff between the accuracy and computational cost. Since the use of a fully coupled MAR-CROCUS system in a data assimilation framework would be computationally prohibitive; in this work higher accuracy is compromised in favor of the computational cost which is justifiable since this is a proof of concept study.

7. p. 3223 l.8: Display this point in Figure 1.

The location of point added as a red square in figure we propose editing Figure 1 as follows:



8. p. 3225, l.6-7: It makes no sense to repeat data that is also in a Table.

We propose to edit the manuscript as follows to address the comment.

> "As can be seen for the entire simulation period, the mean bias (RMSE) of cumulative shortwave, longwave, PDD, and NDD are, respectively, 84% (70%), 82% (85%), 94% (71%), and 65% (86%) less than the mean bias (RMSE) of the prior estimates."

9. Table 1: P has likely also a time dimension. mm per day, year or second?

Precipitation unit is in [mm/hour]. Table 1 will be corrected to reflect this comment.

10. Table 3: Explain why the bias and RMSE in SML is much smaller than in runoff. Apparently the values of RMSE are derived for a subdomain. This should be clear from the text in 5.2 and the header of the table. If my assumption is not correct, explain this difference. (And precipitation should be added here as discussed above).

We propose adding the following note to the manuscript (p. 3229, l. 22)

… -54 mm (250 mm). “Note that runoff occurs in the ablation zone therefore the spatial mean bias and spatial RMSE for runoff were computed over the ablation zone. The spatial mean bias and spatial RMSE for sublimation, condensation, and SML were computed over the entire ice sheet.”

We propose to revise the header of the table 3 as follows:

Table 3: … “The spatial mean bias and the spatial RMSE for runoff were computed over the ablation zone and for the other surface mass fluxes were computed over the entire ice sheet. “

12. Figure 2: What is I.C.? I can’t find it in the text.

We propose adding the following lines to the text to address this comment.

p 3216, l.17

“Note that; $y_j(\tau = 0)$ represents the initial snow profile (IC: initial condition).”

p 3220, l.3

… The posterior forcings “and initial snow profile (I.C.) “ were used as inputs in CROCUS to estimate the posterior surface mass fluxes.

13. Figure 4c: extend the y-axis to 350 or even further until the bars aren’t clipped any more.

We propose editing the Figure 4 as follows: