Surface mass balance model intercomparison for the Greenland ice sheet

C. L. Vernon1, J. L. Bamber1, J. E. Box2, M. R. van den Broeke3, X. Fettweis4, E. Hanna5, and P. Huybrechts6

1School of Geographical Sciences, University of Bristol, Bristol, UK
2Ohio State University, Byrd Polar Research Center and Dept. of Geography, Columbus, Ohio, USA
3Institute for Marine and Atmospheric Research Utrecht, Utrecht University, Utrecht, the Netherlands
4Department of Geography, University of Liege, Liege, Belgium
5Department of Geography, University of Sheffield, Sheffield, UK
6Earth System Sciences and Department of Geography, Vrije Universiteit, Brussels, Belgium

Correspondence to: C. L. Vernon (chris.vernon@bristol.ac.uk)

Received: 31 July 2012 – Published in The Cryosphere Discuss.: 20 September 2012
Revised: 7 March 2013 – Accepted: 8 March 2013 – Published: 3 April 2013

Abstract. A number of high resolution reconstructions of the surface mass balance (SMB) of the Greenland ice sheet (GrIS) have been produced using global re-analyses data extending back to 1958. These reconstructions have been used in a variety of applications but little is known about their consistency with each other and the impact of the down-scaling method on the result. Here, we compare four reconstructions for the period 1960–2008 to assess the consistency in regional, seasonal and integrated SMB components. Total SMB estimates for the GrIS are in agreement within 34 % of the four model average when a common ice sheet mask is used. When models’ native land/ice/sea masks are used this spread increases to 57 %. Variation in the spread of components of SMB from their mean: runoff 42 % (29 % native masks), precipitation 20 % (24 % native masks), melt 38 % (74 % native masks), refreeze 83 % (142 % native masks) show, with the exception of refreeze, a similar level of agreement once a common mask is used. Previously noted differences in the models’ estimates are partially explained by ice sheet mask differences. Regionally there is less agreement, suggesting spatially compensating errors improve the integrated estimates.

Modelled SMB estimates are compared with in situ observations from the accumulation and ablation areas. Agreement is higher in the accumulation area than the ablation area suggesting relatively high uncertainty in the estimation of ablation processes. Since the mid-1990s each model estimates a decreasing annual SMB. A similar period of decreasing SMB is also estimated for the period 1960–1972. The earlier decrease is due to reduced precipitation with runoff remaining unchanged, however, the recent decrease is associated with increased precipitation, now more than compensated for by increased melt driven runoff. Additionally, in three of the four models the equilibrium line altitude has risen since the mid-1990s, reducing the accumulation area at a rate of approximately 60 000 km$^2$ per decade due to increased melting. Improving process representation requires further study but the use of a single accurate ice sheet mask is a logical way to reduce uncertainty among models.

1 Introduction

The Greenland ice sheet (GrIS) is the world’s second largest single ice mass representing approximately 7 m of sea level rise were it to melt entirely. Whilst it contains only an eighth of the mass of the Antarctic ice sheet, the GrIS is important to study since surface melting is already intense (estimated in excess of 1 m water equivalent) along its margins (Fettweis, 2007). Further, the susceptibility to Northern Hemisphere polar amplification makes this ice sheet particularly vulnerable to future climate change.

The mass balance of an ice sheet is its time-varying rate of change of mass and is an important measure of its “state of health”. It can be determined from the sum of two terms: the solid ice discharge to the ocean and the surface mass...
balance (SMB), with a minor contribution from basal melt (Huybrechts et al., 2011). The SMB is the sum of the surface mass gain by precipitation (solid ice and liquid) and water vapour deposition (collectively termed: accumulation) and mass loss from the surface by runoff and sublimation (collectively termed: ablation). For the ice sheet to be in mass balance, the ice lost through ice flow discharge from the ice sheet must be matched by mass gain from SMB. For the GrIS this was identified to be the case between 1971 and 1988 (Rignot et al., 2008); however, from the mid-1990s there has been an increase in both solid ice discharge and runoff, resulting in the GrIS losing mass at an accelerating rate (Velicogna, 2009; Rignot et al., 2011). These two increases attributed to ice dynamics and surface processes have contributed approximately equally to recent mass loss (van den Broeke et al., 2009).

Estimating the mass balance of the ice sheet from the sum of SMB and discharge is termed the mass budget or input-output method (Rignot et al., 2011) and the errors in this approach are generally dominated by uncertainties in SMB. Mass budget estimates have, in the past, used the SMB reconstructions that are discussed here (Rignot et al., 2008, 2011; van den Broeke et al., 2009) and relied on the error estimates from these studies. It is, however, difficult to determine errors in downscaling models and numerical models of the climate system in general (Stephenson et al., 2012), especially when, as is the case here, validation data are limited in space, time and accuracy. A model intercomparison exercise is useful, in such circumstances, to identify the degree of consistency in spatial and temporal (seasonal-decadal) patterns between the models and to determine whether the differences are consistent with the quoted errors. In some instances, a multi-member ensemble can have greater model skill than any individual reconstruction (Overland et al., 2011). Comparison with independent validation data can also provide insights into the behaviour of the different methods. Here, we examine four reconstructions, all driven, by a common forcing: The European Centre for Medium Range Weather Forecasting (ECMWF) re-analysis data. We compare seasonal, interannual and regional estimates of various components of the SMB to examine the impact of the downscaling methodology on the modelled parameter and to quantify the degree of consistency between the approaches. In this study, we do not attempt to propose a preferred approach or data set because (i) the independent validation data used here are inadequate for this purpose, and (ii) different components of the SMB, for different regions, may have different levels of skill between the approaches and it depends, therefore, on the metric of interest as to which is considered more reliable. Irrespective of this limitation, it should be stressed that excellent progress has been made over about the last decade in modelling the SMB of the GrIS. Surface processes are now captured, over the whole ice sheet, at length scales down to 5 km, and with time steps that accurately represent the diurnal cycle in temperature and radiation fluxes and the seasonal cycle of mass fluxes (van den Broeke et al., 2009). The reasonable agreement between three different approaches for determining the time-evolving mass trends for the ice sheet suggests that the SMB reconstruction used must be relatively robust in terms of temporal, and large scale regional, trends (Sasgen et al., 2012). Thus, although validation data for runoff and ablation processes (including refreezing) are limiting and limit our ability to test the simulations in the ablation zone, confidence in the quality and reliability of the reconstructions can be gained from inter-comparisons such as that undertaken by Sasgen et al. (2012). In addition, improvements to the models are being continuously made.

GrIS SMB has been estimated through the statistical downscaling of low resolution re-analysis data Hanna et al. (2011, 2005, 2002, 2008), and low resolution global climate model data (Huybrechts et al., 2004; Gregory and Huybrechts, 2006) through regional climate models (Box et al., 2004, 2009; Ettema et al., 2010a, b; Fettweis, 2007; Fettweis et al., 2008; Lucas-Picher et al., 2012) and also through the interpolation of in situ observations from ice cores, snow pits, stake measurements and automatic weather stations (Bales et al., 2009); however, these measurements are limited in both spatial and temporal coverage and are unable to separate components of SMB. The release of high-quality and consistent historical weather reanalysis data has made possible the modelling of the processes controlling SMB over the whole ice sheet. Temporal differences between these SMB estimations have been identified (Dahl-Jensen et al., 2009; Rae et al., 2012) yet remain to be explained. Here we consider four such reconstructions of GrIS SMB carried out using a common forcing: re-analysis data from ECMWF.

The purpose of this study is to elucidate the differences in SMB estimation arising from different models forced with common forcing data in order to assess their consistency between models and against their quoted errors. We compare these reconstructions, with a focus on ice sheet mask choice, annual time series and the seasonal cycle over the period 1960–2008 to each other and to in situ data where they are available.

1.1 Modelling approach

The annual specific SMB is defined as net accumulation minus net ablation during a year:

\[ \text{SMB} = P - \text{TMT} - R, \]  

where \( P \) = precipitation; \( \text{TMT} \) = turbulent moisture transport (evaporation, sublimation and deposition) and \( R \) = runoff.

Runoff is in turn defined as

\[ R = \text{Rain} + M - \text{RF} - \text{RE}, \]  

where \( M \) is melt, \( \text{RF} \) is refreeze and \( \text{RE} \) is melt and rain water retention.
All components are quoted in units of mm water equivalent. Specific SMB is integrated over the ice sheet to give the total SMB, given in Gt yr\(^{-1}\).

Each component of SMB can be estimated from the ECMWF global climatology re-analysis product known as ERA-40 (Uppala et al., 2005), a data assimilation product covering the period 1957–2002 produced by re-analysing historic observations using a consistent weather prediction model. After 2002 either ERA-INTERIM (Dee et al., 2011) or operational analysis is available up to the recent past. As ERA-40 is not available over the recent years, we need to use two different forcing products (first ERA-40 followed by ERA-INTERIM or operational analysis) and inhomogeneity can occur between the two data sets due to changes in the assimilation procedure or in the data availability over the study period (1960–2008). In addition the observational record for Greenland is relatively poor so ERA-40 is more strongly influenced by the ECMWF atmospheric forecast model used during data assimilation. ERA-40 uncertainty is therefore likely to be larger over Greenland than other areas.

With a resolution of \(\sim 110\ \text{km}\), ERA-40 is relatively coarse to be used directly to estimate GrIS SMB. SMB is dependent on smaller scales due primarily to the narrow ablation area (with a typical width of 10–100 km) but more generally the topographic length scale. One approach, employed by Hanna et al. (2002, 2005), downscales surface air temperature and precipitation to a higher resolution. We call this approach ECMWF-downscale (ECMWFd). A second approach, employed independently by Box et al. (2006, 2009), Fettweis (2007) and Ettema et al. (2010a, b) uses high resolution regional climate models (RCMs) as physically-based interpolators of the re-analysis climatology.

Concern has been raised about the suitability of ERA-40 and subsequent operational analysis for Arctic studies following the discovery of a discontinuity in mid- to lower troposphere air temperatures in 1997 (Screen and Simmonds, 2011). It remains to be seen what, if any, impact this result has on SMB modelling of the GrIS. Other reanalysis products are available and some have been used to force RCMs over the GrIS but, with the aim of eliminating an “external” source of variability in the analysis, only the four models detailed below that employ ECMWF re-analyses have been compared here.

2 Model and data description

2.1 Polar MM5

The Polar Pennsylvania State University (PSU) – National Center for Atmospheric Research (NCAR) Fifth-Generation Mesoscale Model (Polar MM5) is a regional climate model coupled to a surface snow model, with a horizontal resolution of 24 km. The realisation of the model used for this study is described in Box et al. (2009).

The model is forced at its boundaries every 6 h with ERA-40 to 2002, thereafter 12 hourly operation analysis data. The model is reinitialised every month. Tests indicated that the difference in the time interval of the forcing data did not introduce any inconsistency in the downscaled product. Additionally, no significant difference was seen between ERA-40 and operational analysis for overlapping periods (Box et al., 2009). Melting is modelled by an energy balance model (EBM) including model bias correction based on in situ data. The amount of melting, \(M\), is related to the amount of residual energy, \(Q_M\), as

\[
M = Q_M t (L\rho)^{-1},
\]

where \(t\) is time, \(\rho\) is ice density and \(L\) is the latent heat of fusion. Thus \(Q_M\) is calculated as

\[
Q_M = Q_N - (Q_H + Q_E + Q_G + Q_R),
\]

where \(Q_N\) is the net radiative flux, \(Q_H\) and \(Q_E\) are the turbulent sensible and latent heat fluxes, respectively, \(Q_G\) is the firn/ice conductive heat flux and \(Q_R\) is the sensible heat flux from rain. Runoff is calculated using the Pfeffer et al. (1991) meltwater retention and refreezing parameterisation with runoff occurring when surplus water is free to percolate downslope once firm is saturated. This model is limited in scope to a single season’s accumulation layer, so may overestimate runoff in warm years as meltwater is unable to percolate into the earlier year’s accumulation. A three hourly model output is integrated to produce monthly distributions of SMB components.

In the period 2000–2008 surface albedo observations from the NASA Terra platform (MODIS) sensor MOD10A1 product (Hall et al., 2011; Stroeve et al., 2006) were updated daily when grid cells were determined by the data product to be clear sky. In the 1981–1999 period, albedo from the 1981–2000 period based on AVHRR APP-x data (Key et al., 2006) is prescribed. Prior to 1981, multi-year (2000–2008) daily averaged MODIS MOD10A1 albedo data are prescribed.

2.2 RACMO

The Regional Atmosphere Climate Model version 2.1 (RACMO2.1) was developed by the Royal Netherlands Meteorological Institute (KNMI) (van Meijgaard et al., 2008). Adjustments specific to the Arctic environment have been made to the original model to produce RACMO2/GR. The regional climate model with a resolution of approximately 11 km, is forced at its boundaries by ERA-40 (followed by operational analysis after August 2002) every 6 h and evolves freely within the model domain. Sea surface temperature and sea ice open fraction are prescribed. The model is described by Ettema et al. (2010a, b).

In addition to the atmosphere model, an energy balance snow metamorphism model is required for the calculation

www.the-cryosphere.net/7/599/2013/
of SMB components (excluding precipitation). The RACMO atmosphere model is coupled to an energy balance model, described by Bougamont et al. (2005), to calculate $T$, the surface temperature and thus the melt energy, $M_E$:

$$M_E = SW \beta (1 - \alpha) + \varepsilon LW \alpha - \varepsilon_0 T^4 + Q_H + Q_E + Q_R,$$

where $M_E$ = melt energy, $SW$ and $LW$ are the downward fluxes of solar and longwave radiation, $\alpha = \text{albedo}$ and $\varepsilon = \text{surface emissivity}$. $T$ is used as the boundary condition for temperature modelling through the snow/ice layer. Albedo is a function of snow density and cloudiness.

The subsurface multi-layer snow model is based on the SOMARS model (Simulation Of glacier Surface Mass balance And Related Sub-surface processes) described by Greuell and Konzelmann (1994). Here temperature diffuses vertically through the column and surface melt and rainwater percolate downward. Refreezing increases subsurface temperatures and density, which cannot exceed the melting point or the density of ice. Remaining water after these constraints are met percolates to the next layer with a small proportion held by capillary forces. Upon reaching an impermeable ice layer, pore space is filled to form a slush layer and liquid runoff is generated from an exponential decay of slush content as a function of surface slope. This is the highest resolution RCM considered in this study and has been shown to generate significantly more precipitation and higher SMB, suggested to be a result of capturing high accumulation peaks (Ettema et al., 2009).

2.3 MAR

Modèle Atmosphérique Régional (MAR) is a coupled atmosphere–snow regional climate model with horizontal resolution of 25 km. The atmospheric part is described by Gallee and Schayes (1994) and the surface component, SISVAT (soil, ice, snow, vegetation atmosphere transport), by De Ridder and Gallee (1998). As with RACMO a multi-layered energy balance snow model is employed to estimate meltwater percolation, retention and refreezing, snow densification due to liquid refreezing and firn compaction. This model, called CROCUS, is notably described by Brun et al. (1992) and (Reijmer et al., 2012). Albedo is calculated differently to RACMO though. In addition to cloudiness and zenith angle, when snow depth exceeds 10 cm, albedo depends on the shape and size of the snow grains as described by CROCUS. However, for shallow snow cover, albedo varies linearly from snow to bare ice ($\alpha = 0.45$) (Lefebre et al., 2003). Refreeze data have not been provided for MAR therefore has been generated as follows:

Refreeze = Melt + Rain – Runoff. (6)

Liquid water retention and evaporation data are not provided so it is included in the generated refreeze series.

As with RACMO2/GR and PMM5, MAR was forced every 6 h with temperature, humidity and wind fields from ERA-40 and from 2002 onwards with ERA_INTERIM at the boundaries of its integration domain. Sea surface temperatures and sea ice cover were prescribed from the ECMWF re-analysis (Fettweis, 2007; Fettweis et al., 2005). However, unlike the PMM5 model, MAR is not recalibrated or corrected against in situ data. This iteration of MAR has notably been used in studies by Box et al. (2012), Reijmer et al. (2012) and Franco et al. (2012).

2.4 ECMWF-downscale

Unlike the previous three atmospheric models, the approach taken by Hanna et al. (2011, 2005, 2002, 2008) referred to here as ECMWFd does not model the atmosphere over Greenland. Instead ERA-40 and operational analysis surface air temperature, precipitation and surface latent heat flux are downscaling by bilinear interpolation from the original resolution of $1.125^\circ$ latitude $\times 1.125^\circ$ longitude to a $0.5^\circ \times 0.5^\circ$ grid. Evaporation when temperatures exceed 0°C and sublimation for $< 0^\circ$C are calculated from the latent heat of vaporisation and sublimation, respectively. To model runoff on the narrow GrIS margin these downscaling fields are averaged monthly and further downscaling onto a 5 km grid before runoff is calculated. A correction is applied to the surface air temperature (SAT) field based on lapse rates to compensate for the often several hundred metre elevation error present in the relatively low resolution ($\sim 110$ km) ERA-40.

Runoff and refreeze is calculated with a positive degree day (PDD) driven parameterisation after Janssens and Huybrechts (2000); therefore, correcting the SAT is important. Runoff is assumed to occur when melt exceeds a certain fraction of precipitation so the model depends on SAT and precipitation (Hanna et al., 2005). The fraction of precipitation falling as rain is calculated as proportional to the time fraction with SAT $> 1^\circ$C, with rainfall freezing if the snowpack is cold enough. Runoff only occurs once pores in the snowpack are filled to a certain density and removes the snow plus capillary water. Melt of the ice can only occur once the snow is removed (Hanna et al., 2005). This PDD approach does not require albedo to be estimated. Refreeze fields have not been provided for ECMWFd so refreeze has been generated per Eq. (6).

Downscaled output is validated by in situ data collected by the Danish Meteorological Institute’s (DMI) weather stations (Cappelen, 2011) and the Greenland Climate Network (GC-Net) of automatic weather stations (Steffen and Box, 2001). These data are used to derive the empirical lapse rates applied to the SAT field as it is downscaled to the 5 km grid. The calibration results in a good fit, with modelled SAT within 1 K of the observed for most weather stations, and larger deviations of $\pm 3$ K only occurring in a few locations (Hanna et al., 2005). The authors note that the downsampled temperatures for the runoff area are probably within several tenths of a degree of reality. One concern over this confidence is that the DMI data were assimilated into the
Table 1. Ice sheet mask area. Differences between each mask’s total area and the common mask are small, however, differences in the influential low elevation areas are significant.

<table>
<thead>
<tr>
<th>Mask</th>
<th>Total area [km$^2$]</th>
<th>Difference from common mask</th>
<th>Area below 500 m [km$^2$]</th>
<th>Difference from common mask</th>
<th>Area below 1000 m [km$^2$]</th>
<th>Difference from common mask</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common Mask</td>
<td>1 646 860</td>
<td></td>
<td>17 408</td>
<td></td>
<td>96 795</td>
<td></td>
</tr>
<tr>
<td>RACMO</td>
<td>1 761 630</td>
<td>+7 %</td>
<td>48 291</td>
<td>+64 %</td>
<td>166 536</td>
<td>+42 %</td>
</tr>
<tr>
<td>MAR</td>
<td>1 763 610</td>
<td>+7 %</td>
<td>56 513</td>
<td>+69 %</td>
<td>170 760</td>
<td>+43 %</td>
</tr>
<tr>
<td>ECMWFd</td>
<td>1 700 430</td>
<td>+3 %</td>
<td>32 632</td>
<td>+47 %</td>
<td>132 419</td>
<td>+27 %</td>
</tr>
<tr>
<td>PMM5</td>
<td>1 869 700</td>
<td>+12 %</td>
<td>90 571</td>
<td>+81 %</td>
<td>247 881</td>
<td>+61 %</td>
</tr>
</tbody>
</table>

Fig. 1. Four ice sheet masks. Areas in white are common to all masks, dark blue represents additional areas of each mask (a) RACMO, (b) MAR, (c) ECMWFd, (d) PMM5. Land area is shown in grey.

ERA-40 product so downscaled data are not independent from observations. The GC-Net data however remains independent. The ECMWFd version used here is described in Hanna et al. (2011). A further RCM for the GrIS, HIRHAM5 (Lucas-Picher et al., 2012), is not considered here as estimates are only available over the period 1989–2009, forced by ERA_INTERIM.

The statistical downscaling model, SnowModel (Mernild and Liston, 2012), is also not considered for this study as modelled data forced by the common ERA-40 used for the other models are not available.

2.5 Mask variation

The four models use different ice sheet masks (Fig. 1). For ease of comparison each data set is regridded (with resulting fractional grid boxes on the edge rounded to form a binary mask) to a common projection and 5 km grid, a process which introduces small (∼1%) changes from the native output. While the differences in area are relatively small with the largest ice sheet mask (PMM5, including all permanent ice) having an area 110% of the smallest (ECMWFd, only the contiguous ice sheet), the difference is important because some areas of the ice sheet are more affected by certain processes than others due primarily to their elevation. ECMWFd only models SMB on the ice sheet itself rather than marginal ice caps and glaciers. Low lying ice has a disproportional impact on the SMB and components of SMB (Hanna et al., 2005; Franco et al., 2012). The ice sheet area below 1000 m, well below the equilibrium line altitude (ELA), varies from 247 881 km$^2$ to 132 419 km$^2$ between models, a difference of 87%. The scale of this mask variation is described in Table 1.

To allow more meaningful model comparison in this work either a common mask, or a metric unaffected by mask variation such as the ELA is used for subsequent analysis. The common mask is not closer to reality; it is very likely smaller than the GrIS but is a common denominator for model comparison. However, this does not remove all variation as the different model resolution will affect how topography is represented.

PMM5 uniquely uses a fuzzy mask. In this case classification of the grid cells as permanent ice, land, ocean, and mixed “pixels” is made using 1.25 km resolution June–August NASA Moderate Resolution Imaging Spectroradiometer (MODIS) bands 1–4 and 6, cloud-free imagery from 2006. The surface is considered permanent ice if surface reflectance exceeds 0.3 and if the normalised difference vegetation index is less than 0.1. When the 1.25 km grid is
interpolated to 5 km using a “nearest neighbour” basis to quantify how much mixing of the grid cells by land takes place, it is possible to define a “fuzzy” mask that quantifies the mixing of land and ice using a value between 0 and 1. As such, selection of a mask threshold to represent the average case of permanent ice partially addresses the sub-grid issue while maintaining the ability to accurately determine the mass flux. Based on this classification, our best estimate of the permanent ice covered area tuned to match permanently glaciated areas reported by Kargel et al. (2011) corresponds with mask values greater than or equal to 0.587, resulting in an area of $1.824 \times 106$ km$^2$. Non-ice sheet grid cells are excluded from the reconstruction. There are 72,974 grid cells counted as glaciated. In all, 2,496 of these grid cells are isolated from the inland ice sheet, totalling 62,393 km$^2$.

### 2.6 Validation Data

Ice cores and snow pits are used to measure accumulation rates in the accumulation area. Bales et al. (2001, 2009) and Cogley (2004) have collated the available data. In the ablation area stake measurements are taken but due to the logistical difficulties of repeat visits there are fewer of this type of measurement. Here we only consider ablation area observations along the K-transect, lying at approximately 67° N in West Greenland (van de Wal et al., 2012).

Each observation provides an average yearly SMB over a several year period. After filtering temporally (1960–2008 period) and spatially (common mask) for the period and area under consideration, there are 101 observations above 1500 m in the accumulation area and five from the K-transect below 1500 m comparable with model estimates. In the accumulation area, there is good temporal coverage throughout the period; however, observations from the ablation area are only from 1990 onwards. The locations of the in situ data are illustrated in Fig. 2.

### 3 Results

#### 3.1 Impact of mask variation

Approximately a third of inter-model SMB variation can be explained by the large mask variation at low altitude (Fig. 3). Using a common mask did reduce the model differences for total SMB and most components of SMB. Figure 3 illustrates the annual spread between highest and lowest modelled SMB annual series on the common (141 Gt mean) and native (208 Gt mean) masks. There is a smaller spread, indicating less difference between modelled output on the common mask than the models’ individual masks. Moving to the common mask reduces mean inter-model spread by 67 Gt or 32%. The remaining 141 Gt spread is 34% of the four model mean annual SMB (411 Gt yr$^{-1}$) on the common mask over the 1960–2008 period. This compares to a spread of 57% of the mean when the models’ native masks are used due to a larger spread (208 Gt mean) of a smaller four model average annual SMB (363 Gt yr$^{-1}$). This observation hides some structure. On years with high SMB, RACMO and to a lesser extent MAR show higher SMB on their native larger masks than on the common mask, meaning the relatively low
Table 2. Average annual SMB on model’s native mask and the common mask.

<table>
<thead>
<tr>
<th>Model</th>
<th>SMB (Gt yr(^{-1})) 1960–2008 Native Mask</th>
<th>SMB (Gt yr(^{-1})) 1960–2008 Common Mask</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>RACMO</td>
<td>470</td>
<td>470</td>
<td>+0 %</td>
</tr>
<tr>
<td>MAR</td>
<td>401</td>
<td>432</td>
<td>+8 %</td>
</tr>
<tr>
<td>PMM5</td>
<td>301</td>
<td>404</td>
<td>+34 %</td>
</tr>
<tr>
<td>ECMWFd</td>
<td>279</td>
<td>340</td>
<td>+22 %</td>
</tr>
<tr>
<td>Mean</td>
<td>363</td>
<td>411</td>
<td>+13 %</td>
</tr>
</tbody>
</table>

altitude region, not included in the common mask, has a net positive SMB. This is not the case for PMM5 and ECMWFd where their native masks always produce lower SMB. This different response to the common mask increases the spread between models.

The spread is dominated by the high estimate of RACMO (470 Gt yr\(^{-1}\) on both native and common masks) and the low estimate of ECMWFd (279 and 340 Gt yr\(^{-1}\) on native and common masks, respectively; Table 3). The quoted uncertainties are 9 % for RACMO (Ettema et al., 2009) and 32 % for ECMWFd, the latter figure calculated from quoted uncertainties of 20 % for accumulation and 25 % for runoff (Hanna et al., 2011). Assuming the errors are uncorrelated the difference should not exceed the RMS error of the two, here 33 %. However, we find that the average difference between RACMO and ECMWFd is 191 Gt yr\(^{-1}\), 51 % of their mean, on their native masks and 130 Gt yr\(^{-1}\), 32 % of their mean on the common masks. On the native mask the estimates deviate more than 33 % implying underestimated errors for one or both results; however, using the common mask the deviations are within the quoted uncertainties.

Using the common mask reduces inter-model spread for melt from 74 % to 38 % of the mean, for refreeze from 142 % to 83 % of the mean, and for precipitation from 24 % to 20 % of the mean. For runoff, however, the common mask produces a larger spread of 41 % compared to 29 % on the models’ native masks.

The reduction from a model’s native mask to the common mask changes the ratio of low elevation area to high elevation. Processes are highly sensitive to temperature and therefore elevation so this change in ratio shifts the relative contribution of each process. The magnitude of the contribution is also related to the area over which it is integrated.

The differences in Table 2 relate to modelled areas that are lost when moving from native masks to common mask. Generally, the models estimate a higher (positive) SMB integrated over the common masks than on their native masks. This result is expected, despite the common mask being smaller, because the areas lost when using the common mask tend to be at low elevation, below the ELA in the ablation area. The common mask reduces the proportion of ablation area, making the accumulation area more significant to the integrated estimate. Since there is less agreement between the models’ representation of the ablation processes (melt, refreeze, runoff), reducing this area while maintaining the accumulation area decreases the variability between the models. The mask used for ECMWFd estimates is the smallest and therefore the closest to the common mask (Table 1) yet RACMO and MAR are the least affected by the change. This highlights the disproportionate impact of some regions of permanent ice cover.

3.2 SMB components

Annual time series of total SMB and SMB components (precipitation, runoff, melt and refreeze) for the common mask are presented in Fig. 4. There is large inter-annual variability.
Reasonable correlation (SMB can have significant impact on the period’s mean. There is period’s disagreement (Alley et al., 2010) in mass balance estimation based on observation (altimetry and gravimetry) over the last decade as studies are typically carried out over different periods. Relatively small differences in the study period can have significant impact on the period’s mean. There is reasonable correlation (SMB \( r = 0.75–0.91 \), precipitation = 0.75–0.93, runoff = 0.73–0.95, melt = 0.84–0.96, refreeze = 0.69–0.88) between the four models’ annual time series which is unsurprising due to the constraint provided by the common ERA-40 forcing. As indicated by the inter-model spread (Fig. 3), however, large variations in the absolute values remain, especially for refreeze.

The GrIS has undergone a climate forcing with surface air temperatures increasing over the study period (Box et al., 2011; Hanna et al., 2008). While the impact of this is evident in the annual time series, the models’ sensitivity to forcing since the mid-1990s is more apparent when viewed as a cumulative anomaly from a reference period. This approach also minimises the effects of biases present in the models. The period 1961–1990 has been used in a number of previous studies to represent a climatological mean when the ice sheet was close to balance (Hanna et al., 2005; Rignot et al., 2008) and we use the same approach for estimating anomalies.

Each model’s cumulative anomaly for SMB, precipitation and runoff is plotted in Fig. 5. These cumulative anomalies are important when used in mass budget estimates as differences in cumulative anomaly translate directly into differences in the mass imbalance calculated. Thus, using MAR would result in mass imbalance (from 2000–2010) that is about 1000 Gt larger than for ECMWFd and 500 Gt larger than for RACMO and PMM5. In other words, the trend is about 1000 Gt a\(^{-1}\) larger for MAR than ECMWFd due to a greater sensitivity in runoff during this period in MAR. The runoff increase over the 2000s results from the ablation area expansion as observed by the MODIS satellite product (Box et al., 2012). The current runoff increase is higher in MAR because this model uses a fixed bare ice albedo of 0.45, which is lower than other models. This results in larger bare ice areas, which have a greater impact on the MAR melt. On the other hand, the current bare ice area expansion could be overestimated in MAR although it compares well with the MODIS-based estimations (Tedesco et al., 2012). It is important to remember here that the four models are all responding to the same external forcing, yet their SMB sensitivity is markedly different.

The 1995–2008 SMB decrease in RACMO, MAR and PMM5 is comparable in magnitude to the change from 1960–1972. However, the earlier change appears to have resulted solely from a decrease in precipitation during a period when runoff for each reconstruction was stable. The SMB decrease is correlated with a decrease in precipitation. The recent decline by contrast is associated with increased precipitation and an even greater increase in runoff. This striking difference in recent behaviour compared with the previous period of reduced SMB demonstrates the importance of accurate modelling of runoff and its most significant components: melt and refreeze.

The magnitude of the SMB response from each model is different during this recent period (1995–2008) compared with the earlier period (1960–1972). There is greater agreement between the precipitation estimates than runoff estimates. The sensitivity of RACMO and MAR runoff estimates to the same re-analysis forcing is twice and three times, respectively, that of the similar PMM5 and ECMWFd. Since the mid-1990s the GrIS SMB, while remaining positive, shows a negative anomaly (Fig. 5). This decrease is different from a similar decrease seen in the period 1960–1972 in that it occurs during a period of increased precipitation where the earlier decline was driven by low precipitation. The recent decline in SMB is due to an increase in runoff.

### 3.3 Equilibrium line altitude

The area above the ELA, the accumulation area, is independent of mask differences but captures the net behaviour

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>15.8</td>
<td>19.3</td>
<td>20.9</td>
<td>7.2</td>
<td>-82.6 %</td>
<td>-101.3 %</td>
<td>-142.7 %</td>
<td>-180.4 %</td>
</tr>
<tr>
<td>B</td>
<td>28.3</td>
<td>24.9</td>
<td>52.6</td>
<td>46.2</td>
<td>-12.1 %</td>
<td>-77.0 %</td>
<td>-36.5 %</td>
<td>-40.7 %</td>
</tr>
<tr>
<td>C</td>
<td>56.2</td>
<td>68.0</td>
<td>83.0</td>
<td>67.5</td>
<td>6.2 %</td>
<td>-3.1 %</td>
<td>-11.7 %</td>
<td>-13.2 %</td>
</tr>
<tr>
<td>D</td>
<td>96.9</td>
<td>177</td>
<td>133</td>
<td>123</td>
<td>-9.0 %</td>
<td>-23.9 %</td>
<td>-36.8 %</td>
<td>-26.7 %</td>
</tr>
<tr>
<td>E</td>
<td>103</td>
<td>109</td>
<td>102</td>
<td>112</td>
<td>-31.1 %</td>
<td>-61.6 %</td>
<td>-76.5 %</td>
<td>-33.5 %</td>
</tr>
<tr>
<td>F</td>
<td>41.1</td>
<td>81.5</td>
<td>57.7</td>
<td>56.4</td>
<td>-80.2 %</td>
<td>-54.2 %</td>
<td>-86.4 %</td>
<td>-60.5 %</td>
</tr>
</tbody>
</table>
Table 4. Annual regional runoff (common mask); the mean over 1961–1990 period and the change over 1996–2008 period. Largest change highlighted in bold.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>180</td>
<td>133</td>
<td>168</td>
<td>135</td>
<td>141.0 %</td>
<td>131.2 %</td>
<td>78.1 %</td>
<td>153.7 %</td>
</tr>
<tr>
<td>A</td>
<td>17.1</td>
<td>10.2</td>
<td>11.3</td>
<td>22.1</td>
<td>66.8 %</td>
<td>131.2 %</td>
<td>78.1 %</td>
<td>153.7 %</td>
</tr>
<tr>
<td>B</td>
<td>29.7</td>
<td>18.1</td>
<td>8.5</td>
<td>15.4</td>
<td>36.5 %</td>
<td>85.0 %</td>
<td>103.1 %</td>
<td>63.9 %</td>
</tr>
<tr>
<td>C</td>
<td>6.9</td>
<td>5.5</td>
<td>3.9</td>
<td>1.4</td>
<td>87.9 %</td>
<td>67.3 %</td>
<td>96.0 %</td>
<td>153.7 %</td>
</tr>
<tr>
<td>D</td>
<td>23.9</td>
<td>23.4</td>
<td>33.8</td>
<td>19.5</td>
<td>-2.9 %</td>
<td>31.3 %</td>
<td>43.1 %</td>
<td>50.3 %</td>
</tr>
<tr>
<td>E</td>
<td>64.3</td>
<td>65.9</td>
<td>82.5</td>
<td>54.8</td>
<td>17.2 %</td>
<td>88.9 %</td>
<td>78.0 %</td>
<td>17.1 %</td>
</tr>
<tr>
<td>F</td>
<td>38.3</td>
<td>10.2</td>
<td>28.2</td>
<td>21.6</td>
<td>122.4 %</td>
<td>453.8 %</td>
<td>147.8 %</td>
<td>96.0 %</td>
</tr>
</tbody>
</table>

Fig. 5. SMB, precipitation and runoff cumulative anomalies from the 1961–1990 period for each model.

of accumulation processes. Above the ELA there is net accumulation. Below the ELA, net ablation prevails. During the 1961–1990 period, the modelled ELA and hence the area above the ELA was approximately stable in each of the four models and averaged: PMM5 \(1.49 \times 10^6 \text{ km}^2\), MAR \(1.52 \times 10^6 \text{ km}^2\), ECMWFd \(1.48 \times 10^6 \text{ km}^2\) and RACMO \(1.53 \times 10^6 \text{ km}^2\). Since the mid-1990s, however, PMM5, MAR and RACMO have a rising ELA and shrinking accumulation area at a rate of approximately 60 000 km\(^2\) per decade. ECMWFd, also with the lowest melt and associated lowest refreeze, does not show a significant difference between the two periods. The rise of the ELA is not due to decreased precipitation; precipitation has increased over this period. It is due to increased runoff, as a result of increased melt outweighing this increased precipitation (Mote, 2007). For MAR and RACMO this increased runoff is due to expansion of bare ice areas which significantly decreases the albedo (Tedesco et al., 2011). The satellite derived albedo product used by PMM5 should also capture this decrease. ECMWFd does not model or incorporate albedo variations.
3.4 Regional analysis

To examine the regional variations between the models, Greenland is divided into six regions (Fig. 6) based on groupings of drainage basins identified in Rignot et al. (2008). For each region two comparisons are made for SMB and runoff; first the average annual SMB (Fig. 7) and runoff (Fig. 8) during the 1961–1990 reference period and secondly the rate of change during the 1996–2008 period. Data for SMB and runoff are provided in Table 3 and Table 4, respectively. The first observation to note is that the rank order of the models varies from region to region. RACMO has the highest SMB in regions D and F, yet the lowest in region B. PMM5 estimates the highest SMB in region E, yet the lowest in A. Rank order also varies for runoff. MAR has lowest runoff in region B, yet the highest in D and E and PMM5 has the highest runoff in region A, yet lowest in C.

In four out of six regions (A, B, D and F), the highest model’s SMB estimate is approximately double the lowest, a much larger inter-model difference than seen over the whole ice sheet (Fig. 3). ECMWFd usually shows the smallest change in SMB over the 1996–2008 period. This reconstruction is the least sensitive to the common forcing. MAR and PMM5 usually show the largest response. These regional variations suggest spatially compensating errors are leading to the appearance of greater agreement over the whole ice sheet than the localised process modelling is able to reproduce. They imply that, at a basin scale, the errors in one or more model are significantly (up to a factor 2) larger than at the whole ice sheet scale. This also has important implications for the confidence given to mass budget estimates at a basin scale. Moreover, basins along the southwest and north coast are projected to have the highest sensitivity of SMB to increasing temperatures (Tedesco and Fettweis, 2012).

3.5 Seasonal cycle

The seasonal cycle describes each model’s sensitivity to a common seasonal forcing. Figure 9 illustrates the seasonal cycle for SMB and precipitation, runoff and melt components over both the 1961–1990 and 1996–2008 periods. For the 1961–1990 period, monthly SMB is generally positive, adding between 30 and 60 Gt to the ice sheet except for the peak melt period (June–August) where, despite increased precipitation, there is a net mass loss due to runoff. PMM5 appears to be an outlier with a smaller SMB seasonal cycle and including an increase in SMB in May caused by an earlier increase in precipitation than the other models. The precipitation seasonal cycle is greater in magnitude and earlier by several months. PMM5’s runoff seasonal cycle is a little over half the magnitude of the other models. This is likely to be due to the PMM5 snow model not being fully coupled to the atmosphere so neither melt-albedo feedback, nor expansion of dark, bare ice is explicitly considered. This...
period includes the dip and recovery in SMB between 1960 and 1980 due to changes in precipitation. However, this is not expected to affect the mean SMB or precipitation seasonal cycles.

During the 1996–2008 period each model estimates slightly increased SMB during the non-summer months of positive SMB with respect to the 1961–1990 period. This is due to increased precipitation, however, this increase is outweighed by reduced SMB during the summer months due to increased melt driven runoff. The magnitude of the response in this latter period does vary significantly between models. For the three months, JJA, total melt increased by: MAR 43 %, RACMO 38 %, ECMWFd 28 %, and PMM5 27 %. The two models with fully coupled snow models (RACMO and MAR) show the larger melt response, most probably because they allow for feedbacks between the atmosphere and surface. This highlights the limitation of not considering the melt-albedo feedback in a coupled approach.

3.6 Comparison with observation

Figure 10 compares in situ SMB observation with equivalent (location, period) modelled SMB estimates. Model estimates differ from observations by a larger amount in the ablation
Fig. 9. Seasonal cycle of SMB, precipitation, runoff and melt for each model averaged over the 1961–1990 and 1996–2008 periods.

Fig. 10. In situ SMB observations (purple) plotted in increasing SMB (rank) order along horizontal axis with the equivalent model estimated SMB in colour.
area (Table 6) than the accumulation area (Table 5). It should be noted that both ECMWFd and PMM5 have used in situ observations during their calibration. Those data have not been separated out from other observations so the comparison shown here is not fully independent for those two models. We would expect this to result in ECMWFd and PMM5 showing lower RMS errors than RACMO and MAR in the accumulation area.

In the accumulation area, MAR tends to overestimate SMB. In comparison to RACMO, this overestimation is due to increased precipitation in the interior of the ice sheet, leading to fewer bare ice pixels and higher albedo (Fettweis et al., 2011). ECMWFd significantly overestimates ablation along the low elevation K-transect due to higher melt driven runoff in this area.

4 Conclusions

The same climate reanalysis data have been used to force four different models in order to estimate Greenland ice sheet surface mass balance and its sub-components. First, we considered the effect of the ice sheet mask. Total surface mass balance estimates from the four models for the ice sheet show reasonable agreement once mapped onto the common mask. Applying the common mask to model output decreases the variation between models. However, part of this reduction may only be a result of reducing the amount of ablation area, where disagreement is larger, but leaving the accumulation area the same size. The largest inter-model variations remain for refreeze estimates. ECMWFd’s refreeze estimate deviates most from the other three models and therefore contributes most to the observed inter-model variation. ECMWFd’s low refreeze is partially attributable to its lower melt magnitude but the main difference with respect to the other models is the use of a degree-day approach instead of an energy balance model. Modelled surface mass balance uncertainty may be smaller than previously thought due to variations in the native mask, and so we recommend the use of an accurate common mask for future modelling work.

The differences between models both at the whole ice sheet (native masks), and more acutely at the basin scale, are larger than the combined RMS errors of the models, suggesting an underestimation of the model error. This has implications for uncertainty estimates in mass balance studies based on the mass budget or input-output method.

Next, we compared the physical differences between the models by component and by region now on the common mask. There are large regional inter-model variations, and, in particular, the rank order of the model outputs vary considerably regionally, suggesting a likely spatial compensation of errors. This compensation exaggerates the level of agreement seen over the whole ice sheet leading to the appearance of greater agreement than the localised process modelling is able to reproduce. The variations in ice sheet mask increased apparent differences in the modelled estimates, which can be partly alleviated by using a common mask. More work is needed to understand the origin of the regional differences between models as, whilst the current differences reduce when integrated over the whole ice sheet, this may not continue to be the case as the climate over the GrIS changes in the future.

The comparison of SMB components revealed a change in the physical drivers of ice sheet reduction. The decreasing annual SMB modelled since the mid-1990s is similar to the earlier period 1960–1972, however, where the earlier decrease is due to reduced precipitation with runoff remaining unchanged, the recent decrease is associated with increased precipitation, now more than compensated for by increased melt driven runoff. The ability to separate the components of SMB in this way is a strength of the modelling approach used here. We examined the response of the four models to the common forcing since the 1990s and also to the seasonal cycle and found marked differences between the models in terms of cumulative SMB anomaly and amplitude of seasonal cycle, with differences of up to factor 3 and 2, respectively.

Finally a comparison was made with in situ observation data. ECMWFd and PMM5 have used some in situ data during their development which is likely to explain the lower RMS errors seen in the accumulation area for these reconstructions. Modelled estimates differ from observations by a larger amount in the ablation area, suggesting melt and refreeze, with their strong albedo feedbacks, are more challenging to model than the precipitation downscaling.

Future work should focus on the physical reasons for the regional differences identified here, in order to inform future model development. The models investigated in this study are continuously under development and comparisons between simulations with new and improved physics are likely to reduce or ameliorate the discrepancies that we identified in this analysis.

Acknowledgements. We thank the editor and the two anonymous reviewers for their constructive comments, which have greatly improved this manuscript. Funding for this study was provided by a NERC studentship awarded to C. L. Vernon.

Edited by: G. H. Gudmundsson

References


van de Wal, R. S. W., Boot, W., Smeets, C. J. P. P., Snellen, H.,
van den Broeke, M. R., and Oerlemans, J.: Twenty-one years
of mass balance observations along the K-transect, West Green-

van den Broeke, M., Bamber, J., Ettema, J., Rignot, E., Schrama, E.,
van de Berg, W. J., van Meijgaard, E., Velicogna, I., and Wouters,
B.: Partitioning Recent Greenland Mass Loss, Science, 326, 984–

van Meijgarrd, E., van Ulft, L. H., Van De Berg, W. J., Bosveld, F.
KNMI regional atmospheric climate model RACMO version
2.1, KNMI, PO Box 201, 3730 AE De Bilt, Wilhelminalaan 10,
De Bilt, the Netherlands, 2008.

Velicogna, I.: Increasing rates of ice mass loss from the Green-