Dear editor,

our answer file is organized in the following way:

- list of changes
- detailed justification for our choice of the reanalysis product
- detailed answers to the points raised by M.Pelto
- detailed answers to the point raised by H.Rott
- detailed answer to the point raised by the 2nd reviewer

We resubmitted our article in a compact style, since we think that the discussion format is very complicated to handle and eventually produces unnecessary waste of paper, but of course we are also happy to resubmit in this format.

List of changes:

- subsection were introduced in section 3 to improve the readability
- two columns were added to Table 1 in which we compare average model ELAs with SLAs presented in de Angelis 2014
- a paragraph was added in which we discuss the ELA-SLA comparison (line 326ff)
- numbers for the inferred calving fluxes 1975-2000 were updated using a simulation based on the RGI catchments, which were in much better agreement with catchments originally employed in Rignot2003 then the catchments of Willis2012.
- It was emphasized in the text that the errors of the calving fluxes presented in Table 1 are not overall uncertainties of these values but “quantifiable uncertainties” and that only the comparison of the three columns is a way to get an idea about the total uncertainty of the calving fluxes.
- A subsection (3.6) was added in which we discuss in more detail the modeled surface mass balance on the Perito Moreno Glacier and compare our modeled surface mass balance profile with the surface mass balance profiles of two other studies (Figure 5).
- A subsection was added in which we discuss the possible sources of uncertainties of our surface mass balance simulations (3.7).
**Justification of the choice of the reanalysis product:**

The model runs were forced with data from the Reanalysis NCEP-NCAR (R1) product. These are the original reanalysis data and are based a frozen global data assimilation system as of 1995. The R1 data were used to both force the mesoscale model for the 7-year simulation period and also as an input for the statistical procedure used to extrapolate model results over the entire study period from 19XX – 2010.

As the reviewers point out, there are a number of alternative re-analysis data available nowadays and some, such as the ERA-40 product, may offer a slightly better performance in general at high latitudes over the southern hemisphere, as discussed in the Bromwich et al (2004) paper mentioned by the reviewer. However, it is also clear from the same paper that the performance of each reanalysis product varies quite considerably from place to place and there is no guarantee that the ERA-40 product has a better performance for all variables just upwind of the Southern Icefield (the relevant location for our study). In fact, as can be seen from the second figure in the Bromwich paper (figure 2), the R1 product, while worse in general, actually performs slightly better for sea-level-pressure at the Punta Arenas site, which is the closest to the SPI of all the sites used in their evaluation.

The R1 product assimilates only a very limited amount of satellite data (the assimilation system is frozen at 1995) and over the southern hemisphere oceans the data are mainly determined by radiosonde data and model 'first guess' fields. As far as we are aware the AMSU data that caused spurious trends in the Reanalysis-2 product that were documented in Nicolas and Bromwich (2010) are not assimilated by the R1 system. In fact, the insensitivity of R1 to the increasing availability of satellite data from the 90’s onwards was one of the reasons we choose it for this study. We believe that the R1 results upstream of the SPI are likely to be very strongly constrained by the radiosonde data available in Puerto Montt to the north of the SPI and Punta Arenas to the south. As such, while we cannot rule out the possibility of spurious trends in the R1 forcing data, we feel it is most likely that the precipitation variability simulated in the 90s is related to real variability in the zonal wind and moisture fluxes as measured by radiosonde. While it would be interesting to examine in more detail the origin of the accumulation trend over the SPI, we feel this to be beyond the scope of the our paper.

Reference:

Our answers to the points raised by M. Pelto in *italic*:

Ablation zone validation: Compare the results of this model for ablation to the basic well used degree day models for which there is local validation of the coefficients. For example, Stuefer et al (2007) noted mean degree-day factors during summer periods amount to 0.61-0.64 cm w.e. C-1 d-1 for ice ablation. These values are comparable to those of other glaciers in maritime regions (Hock, 2003). For snow they indicate typical maritime values of between 0.27 and 0.43 cm w.e. C-1 d-1. De Angelis (2013) modeled ablation using a degree-day factor of 0.65 cm w.e. C-1 d-1 for ice, and 0.35 cm w.e. C-1 d-1.

*We will compare our model result with the results of Stuefer et al (2007) and De Angelis (2014) in a mass balance profile plot. Generating degree day factors and comparing them to values found in literature, we think is not a very powerful quality assessment tool for our model's results.*

Precipitation trend verification: The authors identify a recent increase in precipitation that I cannot discern from their data. Aravena and Luckman (2009) identified the dominant spatial and temporal patterns of a network of 23 homogenous instrumental rainfall records of Southern South America but do not identify this increase. The same group of authors in reporting on NPI Figure 5 and 6 do not display this trend in precipitation (Schaefer et al, 2013). Garreaud et al (2013) examine Patagonia climate in detail and derive maximum precipitation of 9000 mm. They also identify no trend, though data ends in 2001. I am not arguing that the trend does not exist or that the modelled results for precipitation are not correct. However, without better comparison and verification the cited increase in precipitation is not demonstrated. The 8.36 m of average accumulated precipitation could also be compared to other model results such as Garreaud et al (2013).

*The problem of a comparison with the two cited studies is the different time span. The significant trend in average accumulation in our data of 0.043±0.009 m/year in 1975-2011 for example changes to 0.033±0.028 m/year for 1975-2000 which is not significant any more at the 5% level. Many of the stations with longer records of Table 2 in Schaefer et al (2013) show increasing trends (Coyhayque, Puerto Aysen, Bahia Murta), which however are not significant at the 5% level. Puerto Chacabuco, being less than 20 km away from Puerto Aysen, shows a negative trend. Figure 4 in the same paper shows that precipitation trends can spatially vary very much in this region. Garreaud et al (2013) do not compute the average annual precipitation over SPI.*

ELA-Balance Gradient verification: The paper does not present a balance gradient which is the standard graph for surface mass balance reporting by the WGMS. Since, there are lots of directly measured balance gradients, the range of possible gradients is well constrained. DeAnglis (2013) Figure 1 provides a range of balance gradients. Is the derived balance gradient from the model used here appropriate?

*Mass balance profile plots will be provided for Perito Moreno Glacier and compared to mb-profiles in literature.*

The ELA is a key measure of mass balance and WGMS plots the relationship between ELA and annual balance for each reporting glacier. The ELA can be approximately
observed using satellite imagery and hence can be used for verification. ELA is not mentioned in this paper. For a given year is the ELA correctly modelled? Barcaza et al (2009) use satellite imagery to report annual ELA for many years during the 1979-2003 period on NPI glaciers. Willis et al (2012) and Schaefer et al (2013) do use ELA observations for comparison on NPI, so maybe it was done for SPI as well. Table 6 in the latter mentioned paper provides just the comparison that would be ideal. DeAnglis (2013) identified the snowline for SPI glaciers using cloud free MODIS images. For the NPI Schaefer et al (2013) note that nearly all ELA’s obtained from the simulation are higher than the observed snowline altitudes at the end of the ablation season. This is the type of comparison that is important.

We compare the ELA’s obtained from our model to the observed snowline altitudes at the end of the ablation season by De Angelis (2014) in a new version of the paper.

Verification of overall surface mass balance: The authors assert that the mass balance loss is due to calving for SPI and the surface mass balance is somewhat positive. This implies that glaciers that are not calving should not be losing significant volume. An easy validation therefore is to compare what the mass balance of some non-calving glaciers is with the observed area and volume losses of recent studies. That noncalving glaciers have a positive mass balance does not fit with findings of Davies and Glasser (2012). Of course many of these glaciers are smaller and due not reach the highest elevations. Hence, without a specific validation by the authors this discrepancy is suggestive but not indicative of model issues. Glasser and Davies (2012) note that small annual rates of area loss increased dramatically after 2001 for mountain glaciers north of 52 S including the large icefields. For SPI they noted the fastest SPI loss since 1870 was from 2001-2011. Further for SPI though some calving outlet glaciers are shrinking rapidly in general, small, land-terminating glaciers are experiencing the highest loss. In Table 3 the land-terminating glaciers are shrinking at rates of 0.29% a−1 from 2001 to 2011, compared with 0.08% a−1 for calving glaciers. If the model does not generate negative balances for these land based glaciers that have been losing volume without calving, then the surface mass balance model must be adjusted.

Our model is producing negative surface mass balances for 182 of the 395 analysed catchments. Several of them are visible as small yellow-greenish patches in Figure 2c). However care has to be taken with the non-calving glacier classification, since formerly non-calving glaciers of the SPI have developed pro-glacial lakes now, similarly as it was documented for the NPI (Loriaux 2013). For example the three glaciers that were classified as non-calving in Rignot et al(2003), (Bravo,Frias and Olvidado) have all developed pro-glacial lakes now.

References:


Our answers to the points raised by H. Rott in italic:

Verification of precipitation and accumulation on the ice field is largely missing due to the lack of data. In particular accumulation is a main source of uncertainty in the mass balance computations. The three mass balance point values in accumulation areas in Fig. 3 underline this problem. There is no 1:1 correspondence (indicated by the line) in the accumulation area. Point 6 (ice core on Tyndall Glacier, Shiraiwa et al. 2002) covers only two years of accumulation, with annual accumulation in the two years differing by as much as 7 m w.e. This is not even adequate for verifying multi-annual mean accumulation at this single point, not to mention accumulation over the whole ice field. For the SMB simulations a large error should be assigned to the accumulation component of the simulations, in particular when stepping down to the scale of individual glaciers.

We agree that the verification of the modelled accumulation is difficult. In Figure 3 we compare modelled accumulation for the same time span as the measurements. The firn corn of Shiraiwa et al. (2002) did not even contain 2 entire years, which makes it difficult to infer inter-annual variability from these data. We are still happy about their measurements, since they prove that extremely high accumulation is taking place at some places on the SPI. Of course four point measurements of accumulation is very little for 12500 km² of ice. This is why we try to quantify the different components of the mass balance for the individual glacier catchments. We can get an idea about the uncertainties of the individual mass balance components at every glacier by comparing columns 1 and 2 to columns 3 in Table 1. We think that this is much more informative then inventing some arbitrarily high a priori uncertainty to the modelled accumulation.

Lacking details on simulation results for individual SMB components impairs comparisons with field measurements and with studies in other glacier regions. Mass balance profiles (specific MB in dependence of altitude) should be provided, e.g. for comparison with balance profiles by De Angelis (2014) and Stuefer et al. (2007).

Mass balance profile plots will be provided for Perito Moreno Glacier and compared to mass balance profiles presented in Stuefer et al (2007) and De Angelis (2014).

For the glaciers in Table 2 it would be useful adding the net balance values for ablation and accumulation areas. Stuefer et al. (2007) specify for Moreno Glacier numbers on net balance for accumulation area and ablation area, based on ice flux through at a gate below the equilibrium line and ablation measurements 1995 to 2003.

We choose a different approach here and compute flux gates at the tongue of the glaciers (columns 3 to 7 in Table 1). For our analysis (equation 1) no separation in net balances for the accumulation and ablation area is necessary. A comparison with the analysis of Stuefer et al. (2007) will be provided in specific mass balance profile plot (see above).

Late summer snow line (e.g. De Angelis, 2014, Section 2.5) would be useful for checking the SMB model performance near the equilibrium line.

We compare the ELA's obtained from our model to the observed snowline altitudes at the end of the ablation season by De Angelis (2014) in a new version of the paper.
There is an obvious mismatch between observed retreat of non-calving glaciers and multi-year trends in modelled SMB. Non-calving glaciers (in particular if small) are more directly linked to climate trends than calving glaciers. According to the increasing positive SMB trend in Fig. 4, the retreat of small glaciers should have stopped (or even turned over to advance) during recent years. Davies and Glasser (2012), however, show ongoing retreat of non-calving glaciers. Although the mass turnover of these glaciers is small compared to the calving glaciers of SPI, this seems to indicate some bias (overestimation of accumulation?) in the SMB model.

*Our model is producing negative surface mass balances for 182 of the 395 analysed catchments. Several of them are visible as small yellow-greenish patches in Figure 2c). However care has to be taken with the non-calving glacier classification, since formerly non-calving glaciers of the SPI have developed pro-glacial lakes now, similarly as it was documented for the NPI (Loriaux 2013). For example the three glaciers that were classified as non-calving in Rignot et al(2003), (Bravo,Frias and Olvidado) have all developed pro-glacial lakes now.*

Besides, one would expect that increase in accumulation is reflected in increase of surface height in level parts of the ice sheet (in areas with little motion). This has not been reported by geodetic data.

*The increase in accumulation could be cancelled out by an acceleration which has been observed at the tongues of several glaciers (see your statement below), which should probably propagate up to higher elevations as well.*

The relevance of computing calving fluxes using velocities of a single date for comparison with fluxes over multi-year periods (Table 1) is doubtful. The lack of information on calving cross sections further increases the uncertainty. Several of the main calving glaciers show strong temporal variations of calving velocity (e.g. Muto et al., 2013; Sakakibara et al., 2013). Comparisons of SMB inferred and velocity-based calving fluxes should better focus at a few glaciers where information on calving cross section is available (e.g. from bathymetric data, ice thickness, height above floating) and should account for multi-annual variations in velocity. Accurate data on retrieved calving fluxes would be important for checking the performance of inferred calving fluxes (and SMB).

*Again we agree with the reviewer in most of the points. However, we think that our analysis, although containing high uncertainties, is still valuable. Indeed, the observed acceleration of many glaciers is probably one of the reasons for the disagreement of the values presented in columns 1, 2 and column 3 in Table 1. Due to the apparent underestimation of current glacier velocities, column 3 in Table 1 can be considered as lower limit of possible calving fluxes for glaciers on the SPI whilst column 2, due to reasons discussed below, provides probably the upper estimate.*

Further issues:

Information should be provided on the data base and performance of statistical downscaling (mentioned on page 3120, line 13 ff). Statistical downscaling requires a representative observational data base. The only station data shown are precipitation data of three stations (not very close to SPI), each of which covers only a subset of the 35 years (Fig. 4).
Our statistical downscaling is not based on an observational data base, but on the base of a seven year simulation with the regional climate model Weather Research and Forecasting (WRF). The performance of the statistical downscaling can be judged by the correlations indicated in the text. For further details on the downscaling technique we refer to Schaefer et al. (2013).

The error estimate for the inferred calving fluxes (Table 1) should be revisited. At least for Moreno Glacier there is a consolidated number for 1995 – 2003 (0.36 Gt/yr, Stuefer et al.), whereas the SMB inferred calving flux for 2000-2011 is 4 times higher.

The errors denoted in column 1 and 2 of Table 1 are not meant as error estimated, but a priori quantified errors by the authors of the geodetic mass balances. This will be will pointed out better in a new version of the manuscript. We will also discuss in more detail the validity of this quantifications of errors.

The performance of the geodetic balances, based on differencing of DEMs retrieved from spaceborne sensors, is critical for estimating calving fluxes from SMB data. The authors use data published by now (only option anyway). Nevertheless, I want to bring forward some points that might be relevant for future work. Regarding the 1975-2000 Volume change, Rignot et al. (2003; Notes 15. and 16) explain that the 1975 DEM did not cover areas at elevations above 1200 m, whereas the SMB simulations extend over the whole ice field. For recent years, new evaluations of volume change based on single pass interferometry data of 2001 and 2012 (Abdel Jaber et al., 2013) indicate less mass depletion than data based on SRTM/optical DEM differencing (for which earlier versions agree better with SRTM-TanDEM-X differencing, both for NPI and SPI).

This is very interesting. Lower overall mass losses would imply lower inferred calving fluxes, which would bring the inferred calving fluxes (columns 1 and 2 Table 1) in better agreement with the calving fluxes estimated from front velocities (column 3 in Table 1). We are looking forward to see the detailed data of Jaber et al. (2013) published in a peer-reviewed journal (e.g. The Cryosphere).

References:


Major comments:

The method to determine the SPI SMB is sound, although very hard to evaluate (I follow the suggestions of the other comments posted in the discussion). However, the inferred calving rates are associated with too high uncertainties to present as such, since the uncertainties in the SMB fields, ice thickness, as well as the volume-mass conversion are high, but difficult to estimate. Therefore, I suggest deleting the part on the calving, or, at least, give it less weight in the paper and give more weight to the associated uncertainties. The large differences presented in Table 1, column 3 and 4, already indicate that the method is not working. This is not a critique to the method or to the authors, as this is the best available as yet, but I don’t think it can be presented in this form. The text on page 3127 and 3128 suggests that the authors also have strong doubts regarding the results, and try to collect all possible evidence why this might not be the case.

For more than the half of the analysed glaciers the method is working! Where it it not working we try to find out why it is not working. Is it due to errors in the geodetic balances, due to errors in estimating the calving fluxes from the velocity field or due to errors of the SMB model? We find this analysis interesting, which is the reason why we want to share it with the community.

An alternative could be to use GRACE data as a tool to evaluate the modeled SMB (e.g. looking at the seasonal amplitude) and/or the total mass balance.

The spatial resolution of GRACE data is far too low to make an analysis on glacier basin scale.

Does the model include a firn model, and if so, why are all the results presented as volumes, and not as mass? The authors could interpret the observed volume changes, and convert them to mass changes, using their model, which would be a great addition to the paper.

Mass balance values are mostly expressed as meters water equivalent (mweq) as it is common practice for glacier mass balances. Sometimes we changes to km^3 of ice per year, which for us is a natural unit for calving fluxes and which can be converted into Gigatons (Gt) by multiplying with the density of ice (mostly used is 0.9 Gt/km^3).

The model includes a very simple firn model in which snow turns into firn after one year and firn turns into ice after another year. No detailed layering or compaction is modelled. This, together with the unknown layering in the beginning of the modelling period impedes that model could contribute to the interpretation of the observed volume changes (in the geodetic balances).

P3120, L 9: Why did the authors choose for NCEP, and why backwards until 1975? Reanalyses on the southern hemisphere are known to perform very poorly before the satellite era (1979, see e.g. Bromwich et al., 2004), and NCEP appears to perform poorly even after 1979 in high southern latitudes (Nicolas and Bromwich, 2011).

For a detailed justification for the choice of the reanalysis product please refer to the section: Justification of the choice of the reanalysis product. We used these data backwards until 1975, because we wanted to compare our model results with the geodetic balance of Rignot et al. (2003) which spans the period 1975-2000. We know that reanalysis products have lower quality in the pre-satellite area. However the good correlations found between our downscaled climate data and the few available observations (see Schaefer et al., 2013) make us confident about the quality of our final downscaled climate data, which were driving the SMB model.
P 3121, L 4: why is this constant lapse rate used? In this moist environment, I expect strong temporal (i.e. seasonal) and spatial variability of lapse rate. Why is the lapse rate not taken directly from the NCEP output.

The principal lapse rate in the downscaling procedure is given by the 7 year WRF simulation, which determines the statistical downscaling to a 5 km resolution grid. Then, a further downscaling is effectuated inside of each 5 km^2 grid cell, where the constant lapse rates are applied for temperature and precipitation.

Naming conventions: The authors continuously switch between SMB, accumulation and mass balance. For instance, Figure 2c does not show glacier mass balance, but area-integrated glacier SMB. This should be considerably revised and improved in a potential revision.

Ok, we have revised it!

P3128: how is the potential change in SMB from volcanic heating calculated? More details should be added here.

In order to give a upper bound to a process like volcanic heating, we assume that an increased geothermal heat flux of 1000 mW/m^2 is evenly distributed below the SPI, which is roughly 15 times the average heat flux over the continental crust. Multiplying this additional heat flux with the number of seconds in a year and dividing this specific energy available in one year by the latent heat of fusion of water a and by the density of water we get a specific mass loss in mweq/year. This is standard calculation in mass balance modelling, so we are not sure if it has to be detailed much more in an experts journal. Editor?

Specific comments: (ok means we accept the proposed change, if not we comment why we do not accept)

P3118
L2: model cannot be validated, only evaluated. ok
L5: high: : : quantify we prefer to keep the abstract short, and hope that people that are interested in the numbers will read the paper, Coauthors?
L7: positive and has been increasing during the period 1975-2011 we prefer: “was positive and increasing”
P3119
L13: models. For the period 1975-2011, Rivera: : : ok
L25: by an increase of calving we prefer “losses by calving”
L26: in this paper ok
P3120
L3: As a first step ok
L4: one or two-way nesting? Specify resolutions of each domains. The nesting scheme was one-way involving three computational domains with spatial resolution of 45 km, 15 km and 5km.
P3119, line 19
L22: define NPI is defined P3119
L24: define correlations, of the linear fit, R of R2? R is the correlation and not the coefficient of adjustment (R²)!
P3121
L19: we present the annual mean incoming: : : ok
L20: a sharp west-east gradient ok
P3122
L6: this is unnessary information, this is a forcing and not a result it is both: result of the downscaling and a forcing of the smb model. We think it fits well here.
L14: mass balance = SMB !! ok
L14: mweq: define ok
L25: SMB values (please check the manuscript for these inconsistencies) ok
L26-27: this is information for in the figure caption. ok
P3123
L9: is sublimation accounted for in the model, and if yes, how? The physical process of sublimation is not modelled. However the model is a semi-empirical model with two tuning parameters for ablation and one tuning parameters for accumulation (Schaefer, 2013). So tuning the model's results to observations (Schaefer, 2013) should bring the modelled accumulation to value which in reality probably corresponds to accumulation minus sublimation.
L17: if this is not the case, you should not use accumulation, but precipitation in our case it should be accumulation (see above).
L19: albedo-melt feedback. Give a short explanation. ok
L22 and further: why not present these as area-integrated values? We think that specific values are suitable, because they allow an easy comparison between glaciers and ice caps of different size.. Area-integrated values are presented in P3124 line 28ff, when comparing SMB processes with calving fluxes.
P3127
L25: overestimate ok
P3128
L2: I would remove this sentence, or elaborate. This adds to the feeling that the authors doubt their own results.
This sentence introduces the analysis which follows in the next paragraph.
P3129
L7 and L12: I see two different numbers for the same process. It is the same process but time intervals vary: line 7: 1975-2011, line 12: 1975-2000 and 2000-2011
P3130
L4: rather then wind exposed peaks ok, we changed the last sentence to: Appropriate sites for accumulation measurements are smooth and flat areas in the central plateau of the accumulation area of the glaciers at elevations of about 1500 m.a.s.l., rather than wind exposed peaks and ridges where snow drift is dominating the accumulation patterns.

References:
