**Response to Referee #1 comments:**
**Review of "Multi-year analysis of distributed glacier mass balance modelling and equilibrium line altitude on King George Island, Antarctic Peninsula", by Falk *et al*. (tc-2017-232)**

Response to major (structural) points:
We thank the reviewer for the thorough evaluation of our manuscript. The paper includes complex data sets and analysis, that has affeted the structure of the paper. We realize that it impacted the readability and agree to restructure the manuscript according to the referee's comments. In our experience, the meteorological data processing and gap filling is crucial for the results of the modelling work. Thus, we would like to keep it in the main text, but we agree that it can be condensed, shortened and relocated when necessary. Thanks again for the comprehensive reading and detailed advice.

Response to specific points:

P1 Abstract, l13: What do you mean by "no drift"?
We mean the deviation of model results from observations over the five-year period. The model is started with initial grid conditions, and it is a significant result that model and observations of glacier mass balance do not drift apart over time. We will rephrase this sentence.

P2, l1: Quantify "large fraction".
ca. 18% (17.995), we changed "large fraction" to "18%"

P2, l26: Sentence starting "The seasonal variability…" needs clarification. SAM needs to be defined.
The Southern Annular Mode (SAM) is a low-frequency mode of atmospheric variability that describes the north-south movement of the westerly wind belt around Antarctica. In recent years, the SAM has shown high positive numbers during autumn-winter which is associated with a contraction of the Antarctic high pressure cell and the circumpolar low pressure trough. We added the definition to the manuscript.

P3, l12-16: Present as continuous text rather than a numbered list.
The numbered list is a different style that we have seen in other publications, but we don't object to the proposed change if this increases readability.

P3, l18: Give lat/lon for KGI. Refer to Fig 1 at this point (figure should ideally include a further map locating KGI with respect to the Antarctic Peninsula, South America, etc.).
Thanks for this advice. We can easily adapt the map to include the relative position of King George Island in the southern hemispherical context.

P3 l30: Delete "in" before "especially".
Done.

P4, section 3.1.1: Mark AWS and temperature sensor locations on Fig. 1. Need to reference figures 2 and 3 (photographs) in this section.
The AWS location is marked in the paper, although apparently not clearly enough. The additional temperature sensor locations are not marked since they were part of the former publication by Falk & Sala (2015), and air temperature lapse rates are discussed there in detail. We can add this information to the map.

P 4, l15: Insert “with” after “equipped”.
Done.

P5, l6+: Discussion of the effects of pyroclastic debris does not belong in this section describing the measurements – move to the results section. Figure 4 also belongs in the results section, not here.
We put this remark here to explain the observations of low albedo. But we can see your point and will move this part to the results section.

P7, l20: I don’t understand why you applied a 48-hour smoothing to the cloud observations after interpolating to 1-hourly data. Surely you should use these data at the highest temporal resolution available (to make them compatible with your other driving data)?
The linear interpolation between data points lead to a cloud cover curve that is very angular. This is not realistic and changes in cloud cover are more transitional. The smoothing was applied so that the resulting curve “looks” more realistic. It does not change the actual observational points much but the interpolated values. Cloud coverage at KGI in general is very high and the smoothing served only to shape the interpolated values to a less angular shape.

P8, l9: Surely m=Pa/(99*cos(ψ)) (assuming m is defined relative to 99 kPa)?
Yes, of course. This is a mistake, and we will correct it. Many thanks to the referee!

P8, eqn. (9): Should also include a term for reflected longwave radiation = (1-ε)*LW↓
The term for longwave radiation flux towards the surface (downward) is represented by equation (8). Longwave radiation is absorbed by the cloud/atmosphere and then it is emitted as longwave radiation again from the atmosphere. This is not the same process as reflection. The total downward longwave radiation flux is calculated by considering the atmosphere as a black body radiator with a certain body temperature.

P8, l25+: Give values for RMS differences between measured and modelled radiation components, as well as mean bias and r values. During winter Rn is typically in the range +/- 50 $Wm^{-2}$, the offset of 15.9 $Wm^{-2}$ apparent in eqn. (11) is really quite significant at this time of year.
The referee is correct in this comment. RMSE are <=10 W/m2. Cloud coverage is typical for the area around the South Shetlands, even in winter. The meteorological data gathered at the Station shows an average of less than 10 clear sky days in most of the winters. This is reflected in the net radiation. We did not want to extend the meteorological data section more than necessary, but we agree that this is an important point. We will therefore add the statistics of differences and RMSE’s.

P9, section 3.3: This section describes results so really belongs in section 4. (but needs to come before the section on model calibration). Include a table giving the elevations of the stakes.
This section comprises a detailed description of the glaciological observations and data time series. The last paragraph contains a comparison to the model output (p 11 ll 10-17), and indeed belongs to the results section. We included this paragraph here, since it is included in Fig. 8 to see the comparison of the model output for the stake locations to the observations. This is really interesting, but we agree that it belongs into the results section and will move the text part accordingly.

P10, l30, and Fig. 8: What do the broken lines on Fig. 8 signify? The figure caption should state the reference date from which CMB has been calculated (i.e. the start date of the calculation, where CMB=0 for all stakes). You say that PG04 is in the accumulation zone, but there are hardly any measurements shown from this site and above – why not show results from at least one stake that is clearly within the accumulation zone? The overall trend at PG05 looks pretty close to zero, suggesting that this site is more or less on the equilibrium line.
The broken lines represent an interpolation of the mass balance stake data points and should be

included in the legend. This is a mistake. We will update the figures accordingly.

P11, l10-17: Section 3.3 is concerned with surface mass balance observations. I think it is confusing to start talking about GMM results here before the GMM has been properly described.
The referee is right (as stated in the response above) and we will straighten this section.

P11, section 3.4: What is needed here is a section describing the GMM. Start by describing the model, then say how the model domain/catchment was set up and finish with a section on model validation against stake measurements.
The model description is included in section 3.5 (p 12 l 15 to p13 l 9), but the referee is right and the description should be BEFORE the description of model input and model calibration. We will move the model description to its own section as section 3.4, then input grids to the model as section 3.5, then calibration of the model as section 3.6. The current structure clearly reduces readability. Thanks a lot for making us aware of this flaw.

P13, l17+: How can you be certain that the model error results form an underestimate of accumulation rather than an overestimate of ablation?
The process that are not incorporated in the physical model are that of snow drift due to high wind speeds, and turbulence-driven snow deposition (also associated with high wind speeds). Refreezing processes are not included in the model physics since it was run in catchment configuration mode. These are all processes that are associated with accumulation of mass at the single grid points. Due to the high time resolution of the observations, we can clearly differentiate between different climatic periods, and thus feel confident enough to make this statement.

P13, l30: What do you mean by "a drift or disagreement … cannot be seen in the data"? Figs. 10 and 11 clearly show disagreement (hence drift) over the lower part of the catchment.
We discussed in the text the different processes leading to the difference between the model and the observations. Further, we discussed that the lowest stakes, PG09 and PG 19, are clearly subject to turbulence-driven snow deposition since located close to the glacier border and adjacent moraine. Periods of main differences are also associated with climatic conditions during late autumn and early winter. The overall behavior, though, shows no temporal accumulation of difference that would arise if model physics were not configured correctly. This can be properly seen in Fig. 8 and 9 that includes the spread of the GMM output for the stake locations to the MBS data time series.

P14, eqn 14: Not sure why you show this equation – you are only able to measure the surface mass balance components.
The goal was to put our results in the broader context of the mass balance, but we can adjust the manuscript in this point.

P14,l24: "coverage". Why does high cloud coverage imply less precipitation and low ablation?
The referee is right with this comment, and we did not intend this meaning. What we meant was that there was less precipitation (compared to other years) in form of rain leading to erosion of the snow and ice pack, therefore less ablation. High cloud coverage (meaning less global radiation) over the summer leads to less ablation due to less energy input to the surface.

P16, l6: Briefly explain how you calculated ELA from observations and model.
We interpolated the calculated net balance (bn) with a line, visually and by regression. The high variance of bn prevents an automated approach. The linear interpolation between data points close to the zero crossing is the most promising.

P17, l24: I don’t understand the sentence starting “If underneath the glacier is mountainous terrain…”. The time taken for the glacier to disappear after the accumulation zone disappears depends on the magnitude of the ablation and the thickness of the glacier.
We have to admit that our expression sounds a bit long-winded. Knowing the glacier surface elevation, the thickness of the glacier is determined by the bedrock. The magnitude of ablation depends on surface elevation. We mean that the time taken for the glacier to disappear after the accumulation zone disappears depends on the underlying bedrock, defining the thickness of the glacier and the magnitude of ablation.