Response to anonymous reviewer's 1 comments on tc-2016-93.

We thank the reviewer for her-his constructive and thorough review of the manuscript. We provide an updated version of the manuscript, written according to the reviewer's comments and suggestions, as well as those from the second reviewer.

We responded to all of the general and specifics comments below (reviewers comments are in black bold, our response in blue), and provide, when relevant, a reference to the corresponding position of the related changes in the text. We provide an edited version of the paper, including track changes (in red, added or modified text, and suppressed text), and a final version. Many issues were also raised by the other reviewer, we simply provide the same answers in these cases.

General comments

The paper presents an analysis of micro-meteorological data from two periods on the mountain glacier St Sorlin. These data provide the platform to assess the contribution of turbulent fluxes, and their uncertainties, to modelled glacier surface melt in the context of different weather types. The measured turbulence spectrum is presented and compared to theoretical predictions that form the basis of the widely used Monin-Obukhov similarity theory. Large deviations from theoretical predictions are found in both weakly and strongly turbulent conditions. The ability to correctly model the turbulent heat fluxes using bulk aerodynamic methods is discussed in the context of weather types and the characteristic turbulence regimes they experience. The authors find that despite the divergence of measured turbulence from theoretical predictions, simple schemes are able to model the melt to within the uncertainty in observed melt.

The paper presents useful and insightful data that support emerging conceptual models of the influence of outer layer turbulence on surface layer turbulent fluxes in glacial environments. The analysis of weather types and the associated characteristic valley circulation, wind speed and temperature profiles and TKE relationships are compelling and link well together. The analysis of turbulence spectra and co-spectra is clear and fits well into the progression of analysis.

The attempt to link these fine scales of turbulence to surface melt, through the calculation of fluxes in using the bulk aerodynamic (BA) method, is ambitious but makes sense in the conceptual framework of the paper. However, the formal links between each scale of analysis are not always well made and this reduces the confidence in the interpretations made. The large scope also makes the paper somewhat disjointed.

We agree with this comment and we considerably changed section 4 (Results) and section 5 (discussion), in order to better link the different scales of analysis. In section 4, Table 2 has been split in two tables. The first one is describing the main characteristics of the weather pattern based and the TKE based classifications in terms of meteorology, and the second one is containing the turbulent fluxes information for each class. This second table includes previously missing information about the regression coefficients between measured and modeled turbulent fluxes. Meteorology and wind regimes (4.1) are now described differently: first for weak forcing, then for strong forcing. The subsections in 4.1 have been removed. The TKE description (4.2.1) has been simplified and the related Fig. 6 has been removed. We profoundly changed the description of the turbulent fluxes (4.2.3), we removed Fig. 8 and added a figure comparing the measured and the modelled fluxes (Fig. 7 in the new version). The discussion about the turbulence in the surface layer (5.1) is now better framed, and we provide a better estimate of the error on the melt energy as derived with the sonic ranger (5.2).

Table 1 has been updated with more information about the sensor heights.

Minor changes have also been made in the Introduction, 2nd and 3rd paragraph know provide a better overview of the katabatic flow issues. In methods section (3), we slightly modified the TKE description (section 3.2, paragraph 1), we provide a better description of the roughness length derivation method (section 3.2, paragraph 2) and include the EC-based method for roughness length derivation (related additional results are discussed in section 5.1, 2nd paragraph). Description of the error derivation method has been improved and has now its own subsection, 3.3.

We hope the new version will be more straightforward to read.

At times analyses are presented that are not entirely relevant to the key points e.g. a time series of net turbulent heat fluxes for each period. Other analysis necessary to support the points being made are missing e.g. a direct comparison of hourly turbulent heat fluxes from EC and BA methods along with their uncertainties. Despite these omissions, the methods are well described and logical with some gaps noted in the specific comments.

We agree with this comment and consequently we removed the figure of the time series of the fluxes (Fig. 8 in former version) and replaced it by a direct comparison of the fluxes from the EC and the BA method (Fig. 7 in the new version). Thus, the comparison is only available for the 2006 campaign.

There are clearly some interesting interactions occurring between the magnitude of net turbulent fluxes and the quantity of melt, which deserve further elucidating. The EC method gives the lowest melt of all methods, yet the has the same or higher net turbulent heat fluxes as the other methods. Why do we see a breakdown of theoretical predictions, yet no impact on modelled melt?

Actually, our results show the net turbulent fluxes are higher in Strong Forcing than in Weak Forcing conditions. Focusing on *H* (*LE* is small or erratic due to large random errors) we show the BA method fluxes are lower than the EC method fluxes in Strong Forcing conditions. Considering the EC method is more reliable than the BA method, we consider the BA method underestimates fluxes in this cases. This effect is even more pronounced if you consider a subset of only large TKE: turbulent fluxes *H* are quite different with the BA and the EC (Table 3, new version). The results suggest that an effective roughness length (z_e, introduced in former studies) can be used as a tuning parameter to correct the BA method for this underestimation.

But, since the total melt results from turbulent exchanges **and** radiative exchanges, we are not sure that we clearly understand this comment. Relating turbulent fluxes to melt cannot be done without considering the whole surface energy balance, and especially the radiative balance which plays a key role on melt. The total melt can be low if the radiative balance is low, even with strong net turbulent fluxes, and *vice versa*. That explains why we do not find a clear relationship between turbulent fluxes magnitudes and melt. Actually, the radiative balance is controlled by the albedo. In section 5.2, the first sentence states: *"During both campaigns, the SEB was mainly controlled by large radiative fluxes, regardless of large-scale forcing, but the contribution of turbulent fluxes to the SEB was significant"*.

In section 5.2 which describes Fig. 9 (former version, now Fig. 8), we added: *"Yet, changes in net turbulent fluxes resulting from the choice of calculation method remained too small in comparison with the radiative balance to yield significant differences in the SEB estimates."* We hope the issue is clearer now.

Further description of the processes occurring at the hourly and daily scale are needed to formally link uncertainties in the calculation method for turbulent heat fluxes to melt. Along with this, one of the key points introduced is the errors in the calculation of the turbulent heat flux, yet this is poorly addressed in later analysis. This would put the paper in a much stronger position to comment on the

conditions in which the violation of current turbulent flux modelling methods will have an impact on melt.

Regarding error calculation, we apply here methods to estimate the random errors resulting mainly from instrumental uncertainties that has been developed on Zongo glacier in Litt et al., (2015). We only briefly describe these methods herein, to avoid having a too long paper, but we could add more details if required by the reviewers. We encourage the reader to check the above mentioned paper where all the error calculation processes are described in detail. As presented in section 4.2.3, and discussed in section 5.1, page 18 lines 3-7 (new version), the main result here is that the random errors cannot explain the differences between the EC and the BA methods when TKE is high. We suggest the difference is due to the inability of the BA method to capture part of the flux in non-stationary conditions. We inserted the paragraph about errors into a new subsection (3.3), where we improved the method explanation. We also improved the interpretation of the results, especially for the turbulent fluxes, in relation with the turbulence characteristics, in section 5.1.

The sub-setting of the analysis by weather types (i.e. Strong/ weak gradient wind forcing) is a particularly interesting approach. I find no issue in sub-setting the spectral analysis by TKE, rather than the weather types as for most of the other analysis. However, the discussion should focus around one or the other for clarity, rather than switching between the two. As it stands this switching is rather confusing.

Since the classification by weather patterns doesn’t always reliably work for describing turbulent conditions in the surface layer, we decided to focus on a characterization of the turbulence and turbulent fluxes using a TKE classification. Since the strong WP tend to select high TKE conditions and low WP, low TKE conditions, the TKE classification helps understanding the processes occurring in the WP-related classification. In the new version, we emphasized this point as much as possible throughout the text. In the new version, this point is evidenced in the results section 4.1, then most of the results are presented in terms of the TKE classification, and we finally discuss how that impacts the results of fluxes comparisons for the WP classification.

Specific Comments

P2 ln2. “englacial”- do you mean inside the glacier? Or rather in catchments with glaciers in them, which would be described “glacierised/ glacierized”.

We meant catchments with glaciers in them. We changed to “glacierized mountain catchments”.

P2 ln11. Correct date - Anderson (2010). Also note that recent work has indicated the contribution of turbulent heat fluxes to melt was very likely overstated in this paper (see Conway and Cullen, 2016), so it is, perhaps, not the best reference to use.

We removed the reference to Anderson and also to Gillet and Cullen. We added reference to Six et al., 2009, together with reference to Sicart et al., 2008, this must be sufficient to support the statement.

P2 ln20. The work of Denby and Greuell (2000) and associated papers needs to be addressed in the introduction as this has been a key paper justifying the use of the BA method over glacier surfaces.

We agree and the reference has been added. It seems that Denby and Greuell assessed the role of flux divergence, but associated papers (Denby, 1999; Denby and Smeets, 2000) do not mention effects of katabatic oscillations or non-stationnarity. Therefore, we consider this to still be an open question. To reflect this we modified the text of the introduction: *“Although numerical simulations have shown the*

BA method was reliable in estimating the surface fluxes in the presence of flux divergence below a wind-speed maximum (Denby and Greuell, 2000), the effects of non-stationnarity and outer-layer interactions remains poorly documented over mountain glaciers (Smeets et al., 1998, 2000)"

Added references:

Denby, B., Second-order modelling of turbulence in katabatic flows, Boundary-Layer Meteorology 92: 67–100, 1999.

Denby, B., and C.J.P.P. Smeets, Derivation of Turbulent Flux Profiles and Roughness Lengths from Katabatic Flow Dynamics, Journal of Applied Meteorology 39: 1601-1612, 2000.

P3 ln18. Were any corrections for tilting of the radiometer measurements necessary given the high melt rates? If so, please detail the procedure used to correct for tilting of the mast, or if corrections were not performed please comment on the effect of tilt on the radiation values.

Corrections were not performed, since tilting was not automatically measured. We assumed the potential correction on the radiometer remained low, since the sensor was rarely found to be out of alignment during each field visit (every 10 days), and it was levelled every time small deviations were noted.

P4 ln5. Please explain why 1-hour runs were chosen over standard 30-minute runs?

The choice was constrained by the parallel use of the off-glacier (on the moraine) meteorological station data for which sampling was set on 1-hour. Calculations of the fluxes have been done with 30 min based runs and did not changed the relative contribution of the fluxes to the SEB.

P5 ln2. "apparition" -> appearance

Corrected

P5 ln18. geopotential -> geopotential height

Corrected when necessary

P5 ln19. Please explain the procedure used to 'analyse and compare' each day of the study period to the WP, particularly if this is an objective or expert judgement procedure.

The weather pattern series has been kindly provided by "Electricite de France" (EDF), and we did not undergo this analysis ourselves. The paragraph provides a brief explanation of the whole method. For the specific step of analysis and comparison, each day, the shape of the observed geopotential height field over Europe is characterized by the geopotential height at 0h and 24h at the 700 hPa level and the 1000 hPa level for 110 grid points (a total of 440 points). This field is compared, in this 440 dimensions space, to the 8 geopotential height fields types which have been obtained by the statistical sorting of measured rain patterns in south-east France. The WP of the day is determined by finding the WP for which the geopotential height pattern is the nearest to the observed geopotential height field in the 440 dimension space, using the Teweles-Wobus score (Garavaglia et al., 2010).

Detailing this complex procedure is beyond the scope of this paper. We provide reference to the paper from Garavaglia et al. 2010. We modified the text to add a short description for the selection of the WP: *"Then, for any day (inside or outside the period used to characterize the decomposition, e.g. 1956-1996) the observed geopotential height field shape over Europe is characterized by the observed geopotential height at 0h and 24h, at the 700 hPa level and the 1000 hPa level for 110 grid points (a* *total of 440 points). This field is compared, in this 440 dimensions space, to the 8 geopotential height fields proposed by the rain-patterns analysis. The nearest of the 8 fields provides the WP."*

We added the following sentence at the end of the paragraph*: "Details of the procedure can be found in Garavaglia et al., 2010."*

P6 ln9. What do you mean by "bad" weather conditions? Please replace with a more descriptive comment.

We improved the text to be more specific. We meant rainy or freezing conditions which evidently disturb the good quality of the measurements.

P6 ln11. Table 3 -> Table 2

Corrected.

P6 ln15. More accepted acronyms are MOST or M-O theory, please use one of these throughout.

MOST is now used throughout.

P7 ln20. What runs were chosen for the analysis of roughness lengths, and how was stability taken into account? This is especially important given the low wind speed maximum observed. Also, why was the EC data not used to analyse roughness lengths?

The complete procedure, developed and presented in Sicart et al., 2014, is based on an iterative fitting of profiles between wind and temperature as described in Andreas et al., 2002. The selection of runs was based on a set of criteria, including neutrality (based on Richardson-bulk parameter analysis), quality of the fits (R^2>0.975), absence of a katabatic wind-speed maximum, and some others. We agree this is not clear in the text, so we included a statement referring to the Sicart et al., 2014 paper more explicitly: *"The method for roughnesses determination was inspired from Andreas (2002) and developed for the tropical Zongo glacier. It is detailed in Sicart et al. (2014)"*

We had not included the roughness values calculated with the EC system since we thought using only one point above the ground for this calculation was not reliable enough. Anyway, the dynamic roughness length had been evaluated with the EC system, by inverting the log-linear wind speed profile relationship. The values were dispersed but we found a mean value of z_0 around 0.02 m and z_t around 6.6×10^{-6} m when derived from the EC system. We now mention that in the new manuscript (page 8, line 17):

"We also evaluated the roughness lengths using the EC system and inverting equations 4 and 5, and selecting neutral runs. The median z_0 was 0.022 m, and the median z_t 6.6×10^{-6} m. We did not use these values to calculate fluxes through the BA method."

Interestingly, the factor $(\ln(z/z_0)\ln(z/z_t))^{-1}$, which is the denominator in the bulk formulation of the fluxes (assuming neutrality), calculated with the EC derived roughness lengths or the effective roughness is roughly identical (0.0165 and 0.0167). This shows that effective roughness lengths values can be used to compensate for the BA method underestimation of the fluxes. We now mention this results in the discussion section p18 line 6:

"This is supported by the values of the EC derived roughness lengths, considering they account for the additional mixing: the denominator $(\ln(z/z_0)\ \ln(z/z_t))^{-1}$ in the bulk formulation of the fluxes, calculated with the EC derived roughness lengths or the effective roughness length is roughly identical (0.0165 and 0.0167)."

P7 ln25. The important small scale topography referred to by Smeets and van den Broeke (2008a) is on horizontal scales 5 to 10m, which suggests that the evolution of topography on Saint-Sorlin on the order of 20-30 cm over a few metres should have an impact on the aerodynamic roughness lengths. Also note that in another paper (Smeets and van den Broeke, 2008b) the authors find these same hummocks have a profound effect on scalar transfer. Please change this sentence to accurately reflect the papers conclusions.

Thanks for highlighting this. We updated the text according to this comment, and added references, see text added: Actually the geometry of the gullies was probably not high enough (up to 10-30 cm cm) to stand out against the dispersion in the aerodynamic roughness measurements (roughness values were ranging between 10^{-1} and 10^{-5} m). We updated the text as follow, p 8 line 5:

"The results show a large scatter, about four orders of magnitude. The median values are z_0 = 0.001 m and z_t = 0.00001 m. The scatter can be attributed to poor accuracy of the temperature measurements leading to large random uncertainties (Sicart et al., 2014). The scatter was too large to observe significant changes in the measured roughness lengths during the 2009 campaign, in spite of snow falls or snow melt that uncovered the ice surface, or appearance of small gullies of about 0.1 − 0.3 m height variations on a few meters horizontal scale that could also have impacted the roughness lengths (Smeets and Van den Broeke, 2008a, b)."

P9 ln3. Please refrain from using parentheses to denote opposites and reword these sentences appropriately: i.e. "The symbols SW and LW stand for hourly mean shortwave and longwave radiation, respectively."

Done.

P9 ln10. Please justify the exclusion of heat from precipitation. This is only reasonable if the contribution of this flux can be shown to be negligible.

This is a common assumption over alpine glaciers, with significant mass turnover. Also, heat advected by precipitation is generally negligible because the temperature difference between the rain and the ice is low, the rain intensity is small, and anyway rain is rarely observed. See Paterson (1994), and Oerlemans (2001). We added references in the text.

P10 Section 4.1.1. It is not clear how this text describing the temporal progression of the meteorology is central to the paper and should be shortened to a few sentences. Likewise, section 4.1.3 should be shortened and included in here.

We significantly modified the text in this section 4.1 to take into account your comment together with the other reviewer comments. Meteorology and wind regimes (4.1) are now described differently: first for weak forcing, then for strong forcing. The subsections in 4.1 have been removed. The whole text is now much shorter and clearer, the whole section 4.1 included 42 lines, it now contains only 21.

P13 ln9. Please fix the use of the parentheses as per comment on p9 ln3.

Done

P13 ln12. Please introduce the acronym GQR. Also do you mean H_{EC} > 5 W m^{-2} given the sign convention used in the paper?

GQR (good quality runs) is now introduced in the "Data processing" section. Yes, H_{EC} > 5 W m^{-2} with our sign convention, we changed it.

P13 ln20. Is it reasonable to use z/L from the BA method when you show later that the BA method underestimates the sensible heat flux and therefore, is likely to incorrectly represent z/L? Also, please explain why z/L was calculated from the BA method and not directly from the EC measurements.

This was actually an editing mistake. We had used the z/L calculated from the Eddy correlation measurements. Corrected in the text.

P14 ln12 and further references: "Kaimal curve" -> "Kansas curve". This is more descriptive and is consistent with p6 ln31.
Corrected.

P14 Section 4.2.3 The first two paragraphs can be shortened to a few sentences as it is not clear how the temporal progression of the SEB is central to the paper. Conversely, further discussion of the results in Table 2 are needed, as is a presentation of a direct comparison of turbulent fluxes from the EC and BA methods. This would ideally take the form of scatter plots of hourly fluxes that include error bars. At the very least some descriptive statistics of the correlation between and spread within each (measured and modelled) flux are needed.

Following the reviewer comment, we removed the temporal description of the turbulent fluxes, and present a description of the fluxes with regard to the SF and WF conditions. We replaced the figure with a scatter plot of fluxes as measured by the EC and as modelled by the BA method. We updated the table 2 with the correlation coefficients and parameters of regressions. The text has been updated to include these results clearly.

P14 ln33 and further references. For clarity, acronyms for the two BA calculation methods need to be introduced earlier and used throughout, e.g. BA1, BA2 or BA_{pro} BA_{eff}. The descriptive names "the BA method based on the profile-derived roughness lengths" become confusing when comparing methods.

Done

P15 ln2. The analysis of errors in the EC and BA methods needs to be presented as it is the key link between the representation of turbulence mechanisms and surface melt. This should include uncertainties on the figures given in Table 2 as well as error bars on any hourly fluxes presented.

The errors are now included when relevant in the figures and in Table 3 (former Table 2). They are discussed in the updated text.

P15 ln9. Previsions -> conclusions?

In the Hogstrom paper, the authors provide a parameterization of the variances. That we called prevision. We changed the sentence to: *"[…] agrees with the **parameterization** of Hogstrom […]".*

P16 ln8. The discussion here about the TKE budget is not well framed and more theoretical background is needed before the results are presented. As it is, this first sentence is unclear and needs to be broken up and reworded.

We entirely removed the section regarding the TKE budget since experimental data is insufficient to support the discussion. Also figure 6 has been removed. We reframed the TKE budget discussion, where the references included are sufficient to support our conclusions. We hope the discussion is clearer as is.

P16 ln16. "The EC method probably accounted for this" I don't understand what you mean here. The measured EC fluxes account for all of the extra turbulence observed as they are based on the same data. Perhaps you mean there is extra turbulence not captured by the EC method, or that the EC was not entirely in the surface layer. Please clarify this in the text.

We are sorry for the confusion. We meant that the EC method fluxes included the contribution of the low frequencies, whereas the bulk method didn't. Text has been updated to clarify.

P17 ln8. The use of the word "probably" here and elsewhere (p16 ln 8, section 6) suggests the interpretations and conclusions reached may not be well founded. It would be better to frame the results in the context of a certain conceptual framework, noting where the results agree or disagree with this framework.

We understand the reviewer refers to section 5, discussion. We agree with the statement. We included various changes in the section, to provide stronger background and description of the expected turbulence characteristics in different cases, and how this is reflected in our data. Detailed changes can be monitored from the edited version of the manuscript.

P17 ln10. "were decorrelated from the surface fluxes". These analyses need to be shown in the paper.

Reference to the Table 3 (formerly Table 2), that is now updated with correlation coefficients and linear regression parameters, has been added into the text to support this statement. Also the new figure 7 illustrates more clearly this results.

P17 ln16. Where does the +/- 10% error on the surface height come from? Earlier an uncertainty of 0.1 m was stated. The uncertainty for each type is presumably some combination of the daily sum of melt for each type and the instrument uncertainty.

The 10 % estimate was a rough estimation of the error on the melt derived from the SR50. The error we set of 0.1 m is the error used as the height uncertainty in turbulent fluxes error calculation. This is set like this since the exact distance of the various instruments to the ground is initiated from manual measurements so it remains quite uncertain. But for the daily melt, the height differences between days are used, for which we assume the absolute height is not relevant, and so we use the traditional sensor error of 0.01 m on daily melt. Propagating this error to hourly melt energy gives an error of 40 Wm^{-2}. This is not far from, but slightly larger than the assumed 10%. We changed it in the graphs where we now use this 40 Wm^{-2} error.

P17 ln33. Why were the BA fluxes not validated against the EC fluxes you already have?

We removed the statement. We actually compare BA fluxes with the z_{eff} from Six et al. 2009.

P17 ln34. The effective roughness length is, in reality, smaller than the aerodynamic roughness length and larger than the scalar roughness lengths. Please change.

z_{eff} is set to 0.001 m, which is equal to z_0 but higher than z_t, z_q. We changed the text to *"an effective roughness length larger than the profile-derived dynamic or thermal roughnesses, respectively*"

P18 Section 5.3. It is good reflect on the limitations of the particular weather typing method used, but the authors need to comment on if these limitations have a real bearing on the analyses made. If they do have a significant bearing on the results, then perhaps a different method needs to be chosen.

The way the section was written probably didn't reflect the actual advantages of the WP decomposition we used. Actually, the Strong forcing/Weak forcing classification does select between distinct wind regimes and turbulent characteristics above the glacier, and as such is quite useful. There are some limitations for SEB studies in a more general sense. We modified the text, we added: *"can be useful to assess turbulent surface fluxes characteristics"* and *"It also provides a rough classification of the turbulence conditions in the glacier surface-layer, since SF and WF conditions are, in the glacier surface layer, likely associated with high and low TKE conditions, respectively. We show that the BA method compared differently with the EC method for different WP. More generally for SEB studies, we highlight some limitations."* . We removed *"is only partly adapted for SEB studies"*.

P19 ln20. 'Sublimation' -> perhaps 'evaporation' would be more appropriate given you state the surface was melting most of the period.

Agreed. Sentence changed to: "*[…] magnitude of negative latent heat fluxes, mainly evaporation […]"*

P19 ln28. "erratic discrepancies between EC fluxes and BA fluxes". These crucial elements of the analyses are not presented and need to be.

Done. The correlation analyses and regression equation are now included in Table 3. A reference to section "turbulent fluxes" has been added herein.

P20 ln2 to 8. These conclusions a very speculative considering no EC measurements are available and that both methods yield melt that is within the measurement uncertainty. Please revise.

The net turbulent fluxes for SF conditions are higher in 2009 than in 2006 for the same conditions. If we assume the BA method is biased, this likely means that the turbulent fluxes in 2009 are biased similarly as reported for the 2006 campaign. We only suggest the effective roughness can be used to scale the BA method fluxes. We have updated the text to reflect this.

P20 ln14. Please be more specific about which high latitude glaciers have large contributions of turbulent heat fluxes.

Good examples are the Storglaciaren (Sicart et al., 2008) in Norway, and glaciers in the Canadian Arctic (Braithwaite 1981). We added these references in the text.

Table 1. It would be useful to know the maximum and minimum height for each instrument, or at least the standard deviation from the mean height.

We included in the table the standard deviation of the heights, as well as the minimum and maximum height observed. There was a mistake in the mean heights for the Vaisala and the SR50, which has been corrected.

Table 1. The accuracy of the CNR1 is listed as 0.4% - what is the justification for this figure, given the nominal accuracy should be on the order of 5 to 10%.

This comes from a previous analysis we underwent in Litt et al., 2015.b. See appendix A of this paper:

"The Kipp & Zonen sensor notice provides an estimated error of 10 % in the irradiance (for daily sums). If we apply this as a random error in the previous equation, it yields an unrealistically large relative error in estimates of surface temperature, i.e. as large as 6–7 K, considering a melting surface temperature of 273.15 K. Random noise must be lower; analysis of SW inc and SW out measurements during the night, when they should be zero, or of LW out when melting is observed, and thus longwave emission intensity must be constant, yields typical SDs of 1 or 2 W m −2 . This is equivalent to a random

noise of 0.4 % around radiation measurements". This description has been adapted and included in the text in the new section 3.3.

Similarly, for the CSAT3, the accuracy is on the order of ± 2 % and ± 6% of the wind speed for attack angles of 5 and 20 degrees from horizontal. This equates to uncertainty on the order of 0.1 ms^{-1} or larger for typical wind speeds. A more careful justification of the accuracy values is needed in the text.

We actually don't calculate any error resulting from noise on measurements for the CSAT3. Literature shows that it is too small to have a consequent impact on the measured fluxes. Rather, we compute the uncertainty due to other sources which is larger. We use methods developed in Litt et al 2015b. See original text still present in section 3.3: *"For the EC method, we followed Litt et al. (2015a), assuming measurement uncertainties on wind speed and temperature were negligible and that most random errors were due to insufficient statistical sampling of the largest eddies (Vickers and Mahrt, 1997)."*

Table 2. Why were average wind speed, air temperature etc. from the glacier AWS not included here?

We just wanted to give an overall estimate of the general meteorological conditions around the glacier (and away from its thermal influence), while not overwhelming the reader with details.

Table 2. The units for q would be more simply expressed as g/kg. Also the values seem an order of magnitude too high – typical values would be 5 g/kg and here they are 50 g/kg. Please check the units of these values.

Thanks for the comment. They are actually expressed in g/kg but there's an order of magnitude mistake. The mistake has been corrected.

Table 2. The table would be much better split into two or more tables that can be inserted at the appropriate sections (data, results).

Table 2 has been split in 2 as suggested. The new table (Table 3) includes the fluxes results and correlations and regression coefficients between BA methods and the EC.

Also you state you do not analyse the "other forcing" category further, but present it here. It would be less confusing to exclude it entirely.

The other forcing column has been removed.

Similarly, the inclusion of both the weather types and the TKE bands is confusing and should be clarified.

Since the WP classification did not allow a complete selection of the different turbulent conditions, the classification on TKE has been preferred to characterize the turbulence. Important results come from this classification, so we kept it. We updated the text (see comments above, *"this point is evidenced in the results section 4.1, then most of the results are presented in terms of the TKE classification, and we finally discuss how that impact on the results for the WP classification."*) so that this is made clearer.

Figure 1. It would be good for this figure to be larger and for the photos to be in colour.

Figure is now larger and pictures now in color

Figure 3. This figure is very hard to read due to the hourly data used, small size and lack of grid lines. Perhaps either daily means of each flux can be presented, or the figures made substantially larger. Tick marks also need to correspond to some meaningful interval (rather than 1/6th of 20 days).

The Figure was made larger, interval of ticks were set to 1 day, dates appear every 7 days. Daily means are now shown on the plot, and in shaded gray we kept the hourly data.

Figure 7. Caption: (b) spectra of w, (c) co-spectra of w and theta

Changed

Figure 9. These error bars seem unrealistically small – especially given the large spread in H from an uncertainty in surface temperature of +/- 1 K. Further justification of these errors is needed.

Note that these are only errors on the mean turbulent fluxes. Since they are random errors, on average over all the available runs, the error is quite reduced, whereas it can be quite large for an individual measurement (see new figure 7). The error is calculated on the basis of an error of +/- 0.35 K on Ts (see Litt et al. 2015, and updated error methods section).

Figure 9. Please use a different shading for H+LE that is not the same as for melt.

Done

Figure 9. Please explain where the 10% error bar for melt comes from.

The 10 % was a rough estimate. Calculating an error of 0.01 m on day height changes yield an hourly error on the hourly available melt energy of 40 Wm^{-2}. This has been updated in the text, see specific comments above.

Figure 9. Please explain why the relatively magnitudes of melt (Figure 9) for each method differ markedly from those for the net turbulent heat flux (Table 2). e.g. EC has the highest net flux in weak forcing, but the lowest melt?

This is because the melt energy plot in fig 9 contains radiative fluxes (in gray) plus the net turbulent fluxes (in white). Furthermore the net turbulent fluxes (in *H+LE*) are larger in strong forcing, as shown in both table 2 (now 3) and figure 9.