We thank Reviewer 1 for this clarification. Please find hereafter our response to his/her remarks (between quotation marks) hereafter.

"I would like to thank the authors for their comments on my review. I am glad to see that the essence of my points came across, but would like to add a short note here for final clarification."

"(1) My intention with terminology was not to exaggerate (and I think I did not do so)."
The term exaggerated was referring to the examples used to show that there were flaws in the statistical procedure, because the reviewer was referring to coefficients from Figure 3a that we never described as significant.
Another example is when reviewer 1 wrote: *"However, for the most reliable case (blue; where the likelihood is high that comparable surface conditions were found at the AWS and the melting boxes, i.e. ice) the R2 is much higher for the energy balance model (0.84) than for the PDD approach (0.58). Assessing these two numbers as "similar" seems a questionable interpretation."*. We never compared these coefficients and never wrote they were similar.
The selected coefficients by Reviewer 1 referred to "dirty ice", and we only compared those for "clean ice" in the submitted paper (lines 18-20, page 2652): "*Surprisingly, the latter correlations were similar to those obtained between measured melting amounts and estimated values using a surface energy balance approach for clean ice* (Figure 3c)". Indeed, coefficients were 0.45 and 0.47 in the submitted Figure 3.
However, we agree that we have to include our statistical tests in the text.

"For example, if you submit to Nature or Science you are not allowed to use the term "significant" unless you refer to your statistical analysis. Hence, all I am asking is, whenever the authors employ statistical significance testing, to not use this term for meanings like "important" or "big" or "strong" or "evident". This will help to not confuse readers"
We totally agree with this point, and will not use this term if it is not appropriate.

"(2) By including uncertainty in the model I mean to provide a plus/minus range for results. Having an "idea" of uncertainty from sensitivity tests is useful but not enough in the context of the intended application."
Yes, this is exactly what we propose to include in a new version of the paper

"(3) My understanding was that energy balance studies of Antizana were made for a point in the ablation zone (i.e. not with a distributed model). If I am mistaken and the distributed modeling already exists, you should include the output from this, for example to check how well the simple model reproduces energy balance-based VBPs."
Getting comparison data from a distributed model may allow analyzing how melting processes vary with elevation, and how the energy balance fits with the measured VBPs. However, energy balance data will not give more accurate VBPs than the field measurements. The energy balance model will even be validated with the measured VBP. As a consequence, verifying that the PDD is correctly reproducing the energy balance-based VBPs would give almost the same results as our comparison with measured VBPs.

“(4) My suggestion to drive a simple model with monthly values implied to also use monthly data as input in the model construction phase. This is what I mean by "consistency".”
We agree with this comment on “consistency”. Nevertheless, please note that the physical basis of the model at a monthly time scale is even more speculative or “strange” than at a daily time scale because surface changes (for instance albedo variations) are occurring at a daily time scale on Antizana Glacier 15, with important non-linear feedbacks on energy fluxes and melting. Monthly values do not reproduce these effects. As a consequence, there is no reason to believe that a model would work better at a monthly than at a daily time scale. Nevertheless, it has already been observed frequently that monthly temperature and monthly ablation are significantly correlated on Antizana Galcier 15 (e.g., see Figure 11 in Francou et al. 2004). In our study we wanted to retrieve the potential physical causes of this relationship. This is exactly what reviewer 1 asked for. However, as suggested with the albedo feedback, we believe that the correct analysis is at a daily time scale, not at a monthly one.

“(5) I looked at the chapter of Gurgiser’s thesis. Without doubt using simple models in the tropics is of interest. Yet I could not find explicit support for PDD modeling in this chapter.”
We used this reference to show that other research groups try to define simple models based on temperature and precipitation, and intend to verify whether the degree-day models correctly work or not.
In the chapter of Gurgiser’s thesis, there is no explicit element suggesting that the PDD is not adapted, because the presented “degree-day” model is based on the application of one coefficient for both snow and ice. However, they validate the use of the ITGG model, with different albedos for snow and for ice, demonstrating that it is possible to correctly model the mass balance in the tropics with temperature and precipitation only. This is also what the PDD is expected to do.

“(6) Please note that with my review I did not draw a conclusion, but provide a recommendation for the editor. Again, due to the inherent problem of PDD modeling with regard to the tropics, I am hoping that my review is stimulating the authors to come up with alternative ideas for simplified mass-balance models of tropical glaciers.”
Critics are stimulating for improvement.

“A quick style suggestion: In response to reviewer letters, it would be helpful to dis-tinguish between reviewer comments and author reply more clearly, e.g. by different colors, font styles, or at least by line breaks. This makes it easier for editor, reviewers, and readers.”
Actually we differentiated our response with another typo, but this disappeared due to The Cryosphere response submission process in which the pdf is directly processed from a txt file. We apologize for this difficulty and will directly the initial version of our word file to the editor.

References:

Francou, B., Vuille, M., Favier, V., and Cáceres, B.: New evidence for an ENSO impact on low latitude glaciers: Antizana 15, Andes of Ecuador, 0°28'S. J. Geophys. Res., 109, D18106, doi:10.1029/2003JD004484, 2004.