## Supplementary information

## Inter-comparison of temperature sensors

In addition to instrumental (manufacturer) error, temperature sensor data are subject to measurement error. A particular issue with temperature sensors housed in naturally ventilated radiation shields is the possibility of overheating under strong insolation, low wind speed conditions or overcooling under calm night-time conditions (Georges and Kaser, 2002). To evaluate potential measurement errors with our instrument setup, two inter-comparison tests were conducted prior to fieldwork using all TinyTag thermistors, housed in MET20 or MET21 radiation shields.

MET20/21 radiation shields have a double-louvered design to block solar radiation, and errors under low wind (< 1 m s-1), high insolation conditions are +0.75 °C and +0.5 °C for the MET20 and MET21, respectively, compared with an aspirated shield (Campbell Scientific Inc., 2017). Our data indicate that wind speeds < 1 m s-1 are uncommon, particularly when insolation is high (see Figure 2). The first test was conducted indoors beneath ceiling windows, exposing the shields and sensors to insolation in the absence of any wind; conditions which are most likely to produce errors. The second inter-comparison test took place outdoors over snow covered ground, as previously reported by Shaw and others (2016). In each experiment, the instruments were placed as close together as physically possible and hence it could be assumed that differences in recorded temperatures between sensors were due to measurement errors.

In the first experiment (INT1), instantaneous temperature differences between individual sensors and the mean temperature of all remaining sensors were < 0.30°C (maximum standard deviation (std. dev.) of any individual sensor = 0.08°C) using a leave-one-out analysis. In the second, outdoor experiment (INT2), instantaneous differences between individual sensors and the mean of all sensors in a leave-one-out analysis were of the order of 0.10°C (maximum std. dev. = 0.14°C).

Statistically, there were no significant *Ta* differences (*p* < 0.01) in sensor temperature between the types of naturally ventilated radiation shields used (shield and sensor combinations for this study are given in Table 1). Furthermore, no statistically significant differences (*p* < 0.01) in *Ta* differences between sensors were found for different wind speed ranges during the outdoor inter-comparison test. For wind speeds of 1-2 m s-1, mean differences of the leave-one-out analyses were 0.15°C (std. dev. = 0.04°C) and for wind speeds of 2-5 m s-1, mean differences were found to be 0.16°C (std. dev. = 0.05°C). Based on the results of this comparative analysis, we interpret differences between individual sensors of >0.30°C as indicative of real temperature differences at different locations on Tsanteleina glacier. A more conservative assessment of differences involving the maximum combined shield and sensor error could be as much as +1.1°C (not ±1.1°C), though is not evident from these tests. It must be noted that the HMP45C sensor (for UWS) was not included in these same tests, though it agreed well with multiple Tinytag sensor observations (mean differences <0.2°C) when installed on the same station mast on Miage Glacier, Italy in 2014 (the setup for Shaw and others, 2016).