Review of "Effect of snow-covered ground albedo on the accuracy of air temperature measurements" by Musacchio et al.

Craig Smith (Referee)

Referee comment on "Effect of snow-covered ground albedo on the accuracy of air temperature measurements" by Chiara Musacchio et al., Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2021-63-RC1, 2021

The paper entitled "Effect of snow-covered ground albedo on the accuracy of air temperature measurements" by Musacchio et al. describes a combined laboratory and field experiment to establish the impact of reflected radiation due to the presence of snow cover on the apparent temperature observation. Although temperature sensors employ radiation shields to prevent radiative errors to the measurement, most designs considered only minimizing errors from incoming radiation and not the errors associated with reflected radiation from snow cover (especially during conditions where the snow cover is fresh during low wind and high solar radiation conditions). The authors, after examining instrument uncertainty in a controlled laboratory, set up a field experiment using pairs of sensors located over a snow-covered area and over an area where the snow in the area of influence has been cleared. The results showed that the temperature error due to reflected radiative heating varied by configuration, but could be as large as 3° C.

I thought that the experiment plan was well thought out and appeared to be quite robust. The analysis and reported results were interesting and useful. I felt that the leading sections were generally well written with some wording and grammar issues that need to be cleaned up. The methodology was largely easy to follow and the figures, with some minor suggested edits, are appropriate. However, I felt that the Discussion (Section 5) and Recommendations (Section 6) are quite disorganized, poorly worded, and require a substantial amount of work before this can be published. My major concerns are as follows:

1) Generally, Sections 5 and 6 need to be re-written so that they convey the main points more succinctly. These require a thorough proof-read for English wording and grammar. I tried to make some recommendations, but since this needs to be re-worked anyways, I'm not sure they are that helpful.

The sections well be rewritten in the revised manuscript according to the indications of the reviewer.

2) In Section 5, many of the main considerations (lines 345-358) were confusing. The bulleted list is a good idea but this needs to be more clear. See the annotated file for more specific issues. (ANDREA/CHIARA)

Thanks to the reviewer for pointing that out. Following his suggestions, the authors are going to revise these considerations in order to make them clearer and easier to understand

3) The discussion of the night time intercomparison in Section 5 (lines 359-361) doesn't really add much to the discussion and should probably be removed. You mention this also in the Section 4.4, stating twice (I think) in the paper that it is out of scope.

We agree with the reviewer and those sentences will be removed from the revised manuscript.

4) I thought that your discussion of your study limitations in Section 5 (lines 362-385) were too negative, somewhat detracting from the valid results. I think that you can point out your limitations and compromises with more brevity with a short discussion about how these may or may not have impacted your results, and thus improve the flow and readability of the section.

We thank the reviewer for this comment. We feel that limitations should be mentioned, but we agree with the reviewer that this section is somewhat too lengthy and will be reduced in the reviewed manuscript. FATTO

5) In Section 5, on line 388, you state "A complete correction curve in function of reflected radiation, wind speed for a specific instrument can be generated by users" but earlier you note that the relationship between temperature differences and reflected radiation are not always clear: line 329 "…the spread in values does not allow for a definition of a relation by fitting, nor it was possible to calculate a function between Δt_{air} and $\Delta Rad_{\mathrm{ref}}$." Maybe I misunderstood the messages, but it appears contradictory. (ANDREA/CHIARA)

Line 329 refers specifically to measurement results presented in the paper (fig 14) that didn't allow to find a fitting function for the relation between Δt and ΔRad.

At line 388, the purpose is to suggest to manufacturers/ end user to characterize their own instruments to evaluate the albedo effect as a function of reflected radiation, wind speed etc. to obtain a correction function.
Since there is no certainty that a 'complete' correction function can be calculated, also in the case of a single instrument, the level of approximation that can be achieved must be taken into account.
In this view authors agree that the sentence must be rewritten to be clearer and not be in contradiction with the results of the analysis presented. (ANDREA)

6) In Section 5, line 390, you compare your methods to Huwald et al., (2009), referencing the use of a sonic anemometer on several occasions. Unless I missed something, you don't really talk much about the Huwald et al. methods so the reference to the anemometer use is confusing.

As a matter of fact, we reference the sonic anemometer only once, just to emphasize the difference between our relative measurement approach and the one by Huwald which used a sonic anemometer as fixed temperature reference.
We agree with the referee that the work by Huwald can be described a little bit more in detail in the text and that will be done in the revised manuscript.

7) In Section 6 (Recommendations), it is good that you point out that recommendations have been submitted to WMO committees, but I would recommend that you summarize the recommendations based on what is found in this paper, perhaps in bullet form. If the purpose of the study is to make these recommendations, then they should be clear and concise, and not just quoted from another report. Also, pay particular attention to the redundancies in this section. (ANDREA/CHIARA)

Authors thanks the reviewer for the suggestion, in the revision of the manuscript, the "recommendations" in section 6 will be summarized and presented in a clearer and concise way. (ANDREA)

8) I would like to see some discussion in Section 6 as to why you think the different configurations are being impacted differently by reflected radiation, wind speed, etc. Can you make some general recommendations on shield/instrument design given your results, and if so, include these in the recommendations section? Further related to this, it would have been interesting to see a non-shielded pair of sensors in the experiment. Why was this not considered as a control? (ANDREA/CHIARA)

The main purpose of the paper is to quantify the albedo effect involving different configurations to obtain a result as general as possible. However, the analysis is still limited to some possible configurations and the aim of the authors is not to force the choice of a type of configuration. However, some recommendation should be added in the discussion given the result of the analysis.

In environmental air temperature measurement, as a general rule, the sensor should be in thermal equilibrium with air without any other source of heating.
A completely non-shielded sensor is subject to direct radiation which causes a certain overheating depending on material, shape, dimension ecc. so in our view there is little value in using a non-shielded sensor as a control.

9) Section 7 (Conclusions) can be written more concisely, quickly summarizing the experiment, major findings, and main recommendations. The content is generally ok but it is too wordy with some confusing statements that muddle the message.
More specific comments and suggested revisions are embedded in the attached annotated document.

We again agree with the reviewer and the conclusion section will be reduced in the revised manuscript, focusing more on the findings and recommendations.

Please also note the supplement to this comment:
https://amt.copernicus.org/preprints/amt-2021-63/amt-2021-63-RC1-supplement.pdf

All the comments will be addressed properly in the revised document.

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques Discussions

Effect of snow-covered ground albedo on the accuracy of air temperature measurements

Chiara Musacchio[1], Graziano Coppa[1], Gaber Begeš[2], Christina Hofstätter-Mohler[3], Laura Massano[4], Guido Nigrelli[5], Francesca Sanna[6], Andrea Merlone[1]

[1]Istituto Nazionale di Ricerca Metrologica, Strada delle Cacce 91, 10135 Torino, Italy
[2]Univerza v Ljubljani – Laboratorij za Metrologijo in Kakovost, Ljubljana, Slovenia
[3]Bundesamt für Eich- und Vermessungswesen , Wien, Austria
[4]Università degli Studi di Torino, Italy
[5]Istituto di Ricerca per la Protezione Idrogeologica - Consiglio Nazionale delle Ricerche, Strada delle Cacce 73, 10135 Torino, Italy
[6]Istituto per le Macchine Agricole e Movimento Terra - Consiglio Nazionale delle Ricerche, Strada delle Cacce 73, 10135 Torino, Italy

Correspondence to: Graziano Coppa (g.coppa@inrim.it)

Abstract. Solar radiation is one of the main factors introducing significant deviations between thermometers reading and true air temperature value. Techniques to protect the sensors from direct radiative influence have been adopted almost since the beginning of meteorological observations. Reflected radiation from snow-covered surface can also cause extra warming to thermometers hosted in solar shields, not always optimized to protect the sensors from this further backward radiative heat transfer. This phenomenon can cause errors in near-surface temperature data series, with relevant impact on the quality of data records. The study here presented experimentally evaluates the effect of albedo radiation from snow-covered surface, on the accuracy of air temperature measurements. The investigation is based on evaluating temperature differences between couples of identical instruments positioned above ground covered by natural vegetation, being one in snow-free conditions and the other above snow-covered surface, at the same time in the same site in close vicinity. The work involved a representative number of different typologies of sensors and shields from different manufactures. A mountain site with appropriate field conditions, offering long-lasting snow presence to maximize data availability, was selected to host the experiment. Quantities of influence such as relative humidity, wind speed and direction, solar radiation (direct and reflected) were constantly measured. The effect was evaluated to range up to more than 3 °C for some typologies of sensors. Full data analysis is here reported, together with complete results. This main scope of this work is to report on an experimental estimation and method to evaluate and include this effect as a component of uncertainty in temperature data series for near-surface stations above snowy areas.

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques Discussions

1 Introduction

The World Meteorological Organisation (WMO) Commission for Climatology and the Global Climate Observing System (GCOS) are recommending study and definition of reference grade networks, installations and methods, to generate top quality data for climate studies (GCOS, 2019). A key requirement for a station taking part in a reference network is a documented measurement traceability and understanding of the total measurement uncertainty (Thorne et al., 2018). Consistent uncertainty calculation needs complete knowledge of the measurement system, the sensors calibration uncertainty, the characteristics of the site and the effects of environmental parameters such as wind, solar radiation and precipitation. Among the numerous observed Essential Climate Variables (ECV), atmospheric air temperature measurements have been collected for one and a half centuries. Such data series are the basis of scientific knowledge on local and global climate trends (Camuffo and Jones, 2002). Land-based stations are equipped with different kinds of thermometers whose performances have been constantly improved. Today, top quality instruments involve platinum resistance sensors and high-level reading and recording electronics. Many efforts have also been made to minimise the effect of quantities of influence on measurement results, with the aim to reduce the associated errors and measurement uncertainty. Solar radiation is one of the main factors influencing the instruments, causing significant deviations between the sensors reading and the real air temperature. Techniques to protect the sensors have been adopted almost since the beginning of meteorological observations. Shields to avoid direct solar radiation reaching the sensing element have been developed, from Stevenson screens (Stevenson, 1864) to modern "pagodas" and naturally or mechanically ventilated solar shields. Recent intercomparisons were organized by WMO (Lacombe et al., 2011) to evaluate the performances and differences among the numerous solutions adopted by manufacturers. While the practical/technical features offered by these shields are now optimized and prescribed (WMO, 2012), rarely their capability to protect the thermometers from backward radiation, reflected by the ground, is fully evaluated, taken into account in measurements and data series or documented in datasheets. This is proportional to the so called "albedo" which is the ratio of reflected radiation with respect to the direct radiation received by the ground that, in case of snow cover, is increased up to 95 % (Barry and Blanken, 2016). Like the direct radiation, this reflected component can cause extra warming of instruments, introducing errors in near surface temperature data series, with relevant impact on detected maximum values and anomalies. Such instrumental errors have different magnitudes depending on the equipment, and the technical solutions adopted in manufacturing thermometers and shields. This phenomenon is relevant in monitoring mountain climate, where the duration of snow cover is high (Nigrelli et al., 2018).

Only few studies can be found in literature evaluating the effect of albedo of snow-covered land on temperature sensors: among them the most significant work is the experiment from (Huwald et al., 2009) based on a very different approach and limited to a single typology of sensor and screen.

The task of the present work is to observe, measure and quantify the effect of extra heating on different kind of instruments positioned above snow covered land, in terms of deviations of sensors readings from actual temperature values. This work is the result of a study made in field conditions following a metrological protocol and experimental method, defined and reported

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques Discussions

in a previous study (Musacchio et al., 2019). The investigation is addressed at the evaluation of relative difference between the readings of pairs of identical sensors protected by solar shields as provided by manufacturers. One pair is positioned above snow-covered surface, while the other above grass-covered ground, in the same site, at the same time under equal environmental and topoclimatic conditions.

The problem of albedo effect on air temperature instruments, here studied, can be included as a part of the general study on assessing data quality and uncertainty in near-surface air temperature measurements. This wider subject is now being analysed and discussed by the WMO expert teams of the Commission of Climatology (CCl) and in the Commission of Instruments and Methods of Observations (CIMO) and is a key aspect in the creation of the Climate Reference Networks for the Global Climate Observing System (GCOS). The complete knowledge and evaluation of uncertainty budget components on air temperature measurement is also included in the roadmaps of scientific activities of the Working Group for Environment of the *Comité Consultatif de Thermométrie* (CCT - Consultative Committee for Thermometry) of the *Bureau International des Poids et Mesures* (BIPM - International Bureau of Weights and Measure) (CCT, 2017)

The activities here reported have been carried out in the framework of the MeteoMet project (Merlone et al., 2015a, 2015b, 2018), a funded joint research initiative of the European Metrology Research Project (EMRP) grouping a wide consortium of National Institutes of Metrology (NMIs), research institutes, universities and National Meteorological Hydrological Services (NMHSs).

2 Theoretical study and experimental method

The present work includes theoretical studies, the draft of an experimental protocol, collection of representative typologies of instruments (thermometers with shields and auxiliary equipment), their laboratory characterization, field installation, on-site measurement and data analysis.

The experiment here presented follows the prescriptions and assumptions proposed in (Musacchio et al., 2018) where a measurement protocol is presented following a theoretical study on the influence of various parameters such as wind speed and direction, snowpack thickness, incident solar radiation, snow conditions and humidity on air temperature measurements above snow-covered ground. In the cited work, the authors also give guidelines on the experiment design and the evaluation of uncertainty components, as well as laboratory characterisations of instruments and the ways to take into account all identified quantities of influence, both instrumental and environmental. Based on these considerations, for the realization of the field experiment, a measurement protocol was prepared, giving prescriptions on:

- design of the experimental set-up and definition of site requirements;
- evaluation of the quantities of influence;
- sensors characterization in laboratory and in field;
- evaluation of uncertainty components.

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques Discussions
Open Access

2.1 Experimental set-up and site requirements

The “albedo effect” here investigated is the sensors overheating due to backward reflected radiation from snow and it is measured as differences of air temperature readings t_{air} between couples of identical sensors in identical shields, positioned one in a point *a*, over snow, the other in a second point *b*, snow-free area: this difference is here indicated as

$$\Delta t_{air} = t_{air}(a) - t_{air}(b) \quad (1)$$

and it includes all the corrections evaluated for each pair of sensors during the described laboratory and field characterisations. These two measurement points are arranged in close vicinity and on a flat surface, free from obstacles, thus exposed to the same topoclimatic conditions, but far enough to accommodate a significant area covered by snow on one point and a sufficient area (some meters of radius) with natural ground left free from snow on the other point. Readings from each pair of sensors are recorded by means of a unique data logger. The investigated effect is therefore the result of a relative analysis, involving identical instruments and single reading unit: this allows minimizing influencing factors and uncertainties. In between the two measurement points, other instruments are deployed to measure the quantities of influence, which took part in the analysis as uncertainty components.

The site hosting the experiment requires a number of specific features. It must be a free, flat surface of at least 50 m of diameter with a minimum presence of obstacles - as trees, buildings or roads in the surrounding area - and spatially uniform solar exposure during the daytime central hours. Besides the presence of snow, the original ground must be covered with natural low vegetation. Other characteristics are related to logistic aspects such as: electric power available throughout the winter, easy vehicular access for maintenance, no agricultural or sport activities, strictly reduced access to public and no presence of vehicles. The experimental site scheme is reported in Figure 1.

2.2 Quantities of influence

The main quantities of influence on the temperature measurements for the evaluation of the albedo effect must be constantly monitored during the experiment. The theoretical work cited identified these quantities as wind speed, air relative humidity and solar radiation. Global solar radiation (downward) and reflected solar radiation (upward) were measured in the same position of each temperature sensor. Snow thickness and conditions have also been monitored (see section 3.2).

2.3 Sensor characterization

Before starting the experimental activities, temperature sensors must be characterized in order to understand their behaviour in different situations. The experimental protocol prescribes two different characterization phases: in laboratory and in field conditions.

The laboratory characterization is needed to evaluate possible systematic differences between couples of sensors exposed to the same temperature under controlled conditions. ~~Being the investigation based only on~~ relative temperature differences among pairs of identical instruments, the sensors calibration is not necessary as no traceable absolute temperature

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques Discussions

measurements are required for the evaluation of the albedo effect in field. This avoids the inclusion of the calibration uncertainty in the overall uncertainty budget and makes the adoption of this procedure easier, also for users willing to make similar analysis without the calibration costs and time required. Laboratory controlled conditions also allow the evaluation of the sensors' stability, sensitivity and resolution of the readout.

Different systematics can arise when the sensors are deployed in the field, due to environmental factors. For this reason, an in-field characterization of the sensors is also needed to evaluate their behaviour in such conditions. Performing an estimation of the uncertainty components of on-site measurements is necessary to quantify the accuracy reached in the experiment. For more details, Musacchio et al., (2019) gives an in-depth description of the whole method, as well as its assumptions and prescriptions.

3. Experimental set-up, preliminary characterizations of site and instruments, uncertainty components

The experimental activity reported in the present work was carried out in the framework of MeteoMet project. In 2016 a call was opened for MeteoMet collaborators manufacturers, to take part in the experimental activities by sending couples of identical thermometers, with identical solar shields. Data loggers were also requested, one for each couple of instruments, in order to make the data available as recorded by the users. Different shapes and dimensions, mechanically aspirated or naturally ventilated shields were collected to have a range of such commonly used devices as broad as possible.

Six different instruments from four different producers were selected for the experiment as described in Table 1.

Additional sensors for the measurement of the quantities of influence were installed, including a cup-and-vane anemometer, a thermo-hygrometer (both positioned in the central measurement point of the experimental area) and two albedometers, one for each measurement point (Table 2). The air temperature reference value was the one measured in the central point, but it is not included in the evaluation of the differences among the pairs of sensors under test, nor it contributes to the uncertainty budget. This further air temperature value is recorded as another potential quantity of influence, in terms of further possible dependence of the temperature differences also on the temperature itself, in addition to the one investigated in laboratory and was evaluated to be a negligible effect.

3.1 Laboratory activities

Preliminary tests on the selected sensors were performed in laboratory for the characterization of the sensors and the complete system. (Figure 2).

This part of the work was performed in the new "Climate Data Quality Laboratory" of the Istituto di Ricerca per la Protezione Idrogeologica - Consiglio Nazionale delle Ricerche (IRPI-CNR). During this phase, a study of the different data loggers working principles was also made, together with the evaluation of best mounting solutions.

The activities started with a preliminary evaluation of the differences between couples of sensors readings, without the shields, in stable temperature conditions, to check for systematic values to be corrected. The sensors were then assembled in the shields, taking notes of both sensors and shield serial numbers, in order to keep the same sensor-shield group in the field experiment,

Atmospheric Measurement Techniques Discussions

and all the relative differences of temperature measurements of each pair of instruments, Δt_{instr}, were measured. The characterization was then performed in a controlled environment with slow temperature change, to keep into account possible effects, without being affected too much by the sensors' dynamics. Too rapid air temperature transients will in fact not be included in the final data analysis, since sensors dynamics can predominantly influence the trueness of the analysis (Burt and de Podesta, 2020). All sensors underwent the laboratory characterisation in order to obtain the information reported in table 3. Stability of the instruments was also tested in laboratory during a one-month continuous acquisition, to check for longer term drifts and potential maintenance required in field. No failures or significant effects were observed.

3.1.1 Laboratory

The laboratory controlled experimental conditions have been evaluated in the testing zone, using traceable reference sensors. Room temperature drift was found to be < 0.02 °C for one day and < 0.05 °C over one week. For time interval corresponding to data loggers' acquisition and recording times, the laboratory air temperature stability was evaluated as $u_{stab} = 1$ mK. The temperature homogeneity was measured and found to be $< 0.05\ °\mathrm{C} \cdot \mathrm{m}^{-1}$ in the laboratory measuring volume. The sensor pairs were positioned at a distance of about 20 cm, as a compromise between minimizing the gradient and avoiding mutual influences such as heating from the electronics or fan motors. The uncertainty due to the laboratory temperature homogeneity was therefore evaluated as $u_{hom} = 0.01$ °C.

The total uncertainty contribution due to laboratory conditions was evaluated as $u_{lab} = \sqrt{u_{stab}^2 + u_{hom}^2} = 0.01$ °C for all the sensors.

3.1.2 Instruments

The evaluation of possible systematic differences Δt_{instr}, among pairs of identical sensors kept at the same temperature (within the laboratory homogeneity uncertainty) was performed by repeated readings over several intervals of about one hour. As shown in Figure 3, all sensor pairs were found to have systematic differences Δt_{instr}, which have to be taken into account for the correction of field data. Associated uncertainty values are reported in Table 3. The repeatability of temperature differences Δt_{instr} contribute to the uncertainty budget with a component reported as $u_{\Delta t_{instr}}$.

These contributions u_{lab} and $u_{\Delta t_{instr}}$ are reported in the overall uncertainty Table 5.

Finally, a check to verify the drift of the sensors was performed at the end of the experiment, after the field campaign and exposure to meteorological conditions. In particular, the drift of Δt_{instr} was evaluated again in stable laboratory condition, by measuring the systematic differences among the pairs of thermometers when kept at the same controlled temperature. The drift was then evaluated as differences in the systematic differences measured before and after the field campaign: values were found to be of the same order of magnitude of the instruments noise. This is an expected result, since only high-performance temperature sensors have been selected, normally produced to guarantee top level stability in time and low drifts, to reduce

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques Discussions

maintenance and recalibrations by the users. The drift in the relative difference becomes therefore negligible for the duration of the experiment and no correction or uncertainty components have been included.

3.2 Measurement site and experiment set-up

Being the experiment performed in a mid-latitude region, a mountain site in the Alps was chosen, to assure the presence of snow cover throughout the winter, for collecting enough records for a statistically significant analysis of the investigated effect. The measurement site, selected to meet logistics and experimental requirements, was found in the municipality of Balme at 1410 m of elevation (45°18'9.31" N, 7°13'19.18" E), in the Ala Valley, northwest of Turin (Italy). (Figure 4, 5)

Only a 3-m wide local road with almost no traffic and a small unmanned building were present in the area, at more than 50 m from the measuring point. Coppa et al., (2021) performed a metrological quantification of the influences on air temperature measurements introduced by the proximity of roads, that revealed a significant effect only at closer distances (less than 50 m) and mainly at very low or even null values of incident radiation. Since for this experiment only records associated to sun radiation are relevant and the road is at more than 50 m, its presence was considered negligible. According to a similar experiment for the evaluation of the effect of buildings, (Garcia Izquierdo et al., in prep) a building of the size of the hut ~~there present~~ and at that distance causes no influence in air temperature records. Moreover, during the experiment set-up, great care has been put in order to place both measurement points at similar distances from each possible source of heat and disturbance: their potential influences, as evaluated by the mentioned works, affect both measurement points in the same way, thus cancelling out during relative differential evaluations.

The chosen area turned out to be a reasonable compromise between the necessity of an alpine location in terms of snow cover presence and duration, and the logistics of a heavily-instrumented research site.

The equipment was installed following the prescription of the protocol. The experimental scheme of Figure 1, based on the three described measurement points, was followed: the two external poles hosting the couples of identical shielded thermometers and the albedometers, the central one with the data loggers, the electric power connection and the auxiliary measurements of humidity, wind speed, wind direction and central air temperature (Figure 6). The two instruments of each couple were positioned in the same orientation and in case of asymmetric shapes, following manufacturers specifications (i.e., ventilation aperture facing North).

~~At~~ significant precipitation events, the snow was removed from a 5 m radius area centred in point *b* (Figure 7 and Figure 8 respectively show the site before and after the removal of snow); the site and instrument were constantly supervised and the meteorological conditions recorded. The weather conditions are fundamental to select, in terms of day and time, periods when the albedo effect, as defined in the model, can be better detected in its maximum values. The 5 m radius was decided as a compromise between maximising the snow-free area under the sensor and having the measurement points close enough to keep the assumption of homogeneity of local weather conditions. This radius could not be expanded because the third measurement point, i.e., the one carrying control and ancillary measurements, would fall in the snow-free area, while it was important that these measurements were representative of the natural state of the site.

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

As mentioned in section 2.2, the experimental protocol mandates an evaluation of snowpack thickness and snow conditions for a full understanding of the quantities of influence. Instruments have been positioned at 2 m from the ground and during the whole measurement campaign the snow thickness never surpassed 40 cm, thus keeping sensors ~~never less than~~ 1.5 m from the surface below (both above natural soil and snow-covered area). The measurement protocol included a recommendation to remove data in case of snow thickness surpassing 1 m, to avoid other effects (extra cooling, turbulences) to introduce errors or uncertainties and the measurement be made not according to prescribed height. As for snow conditions, they were also neglected because snow removal was performed immediately after a snowfall, thus all the differential measurements have been recorded in conditions of fresh snow.

3.3 Characterization of sensors on-site

The theoretical method assumption is that, under the same conditions of snow cover, the difference of air temperature measurements between the two sensors at position *a* and position *b* (Δt_{site}) is zero. In real conditions, to achieve this hypothetical assumption, environmental factors must be evaluated and corrections must be adopted. Non-symmetries can occur, for instance, in cases of variable wind direction and speed, asymmetric shadows or other non-homogeneous atmospheric or surface conditions, causing a non-null temperature difference among the same sensors over the same snow-covered surface. A specific preliminary measurement campaign was made on site, at first snow event, to evaluate such possible systematic temperature differences Δt_{site} among the couples of instruments the following considerations were taken in to account:

- data was recorded when snow was present below both the measurements points;
- data was selected during day time with incident radiation greater than zero;
- data was selected when the reflected radiation difference was zero.

Values of Δt_{site} were then used to correct the raw data recorded on site.

The repeatability of the difference Δt_{site} itself can change significantly during the experiment, due to seasonal changes in the site conditions, such as vegetation on trees, water level in rivers etc. All those contributions have been taken into account when evaluating the uncertainty, $u_{\Delta t_{site}}$ on Δt_{site}. Results of this characterization are presented in Table 4.

Events of asymmetric shadows, cast only over one of the two measurement points, were also identified as due to a mountain peak occasionally projecting its shadow during the period of shortest daytime (December to January): records associated to this shadowing effect were neglected form the data analysis, thus also from the evaluation of $u_{\Delta t_{site}}$.

3.4 Uncertainty budget

The overall uncertainty budget $u_{\Delta t_{air}}$ for the temperature differences Δt_{air} have been derived according to the Guide to the expression of Uncertainty in Measurement (GUM) (JCGM, 2008), from the instruments characteristics and experimental conditions. As reported above, no calibration uncertainty components are here introduced, since the measurand is a relative difference, which does not require absolute accuracy.

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques Discussions

The expression for the evaluation of overall uncertainty is defined as:

$$u_{\Delta t_{air}} = \sqrt{u_{res}^2 + u_{lab}^2 + u_{\Delta t_{instr}}^2 + u_{\Delta t_{site}}^2} \quad (2)$$

where:

- u_{res} is due to the resolution of instruments and data loggers as provided by manufacturers;
- u_{lab} is the component of uncertainty due to laboratory conditions and is composed by temperature uniformity and stability of the laboratory itself;
- $u_{\Delta t_{instr}}$ was evaluated during the laboratory testing of thermometers and is mainly ascribed to sensors short term stability and statistical contributions;
- $u_{\Delta t_{site}}$ is related to the non-ideal characteristic of the on-site conditions.

Table 5 summarizes the components of uncertainty with the expanded uncertainty $U_{\Delta t_{air}}$ reported with coverage factor $k = 2$ and confidence level of 95 %.

4. Data analysis

4.1 Data selection

The measurement campaign was performed between 8 September 2016 and 24 March 2017. The acquisition frequency for every sensor was set to 10 min. The parameters recorded in both points *a* and *b* were: air temperature, incident and reflected radiation. During the campaign, an operator constantly accessed the experimental site and marked the best days for the analysis, in terms of sunny days (maximum radiation conditions) after a snowfall (maximum heat backward reflection) when the snow below instruments at point *b* was recently removed (maximum expected differences). Snow was dug away on 4 days: 30 November ~~and~~ 22 December 2016, 20 January and 23 February 2017. Each time, snow was completely removed within the radius of 5 m, leaving the natural soil exposed. Salt was used each time to prevent the formation of ice, which would have changed the natural soil reflectivity, and to make snow removal easier and more complete. The preliminary data analysis was addressed to measurements ~~values~~ recorded in the days immediately after ~~the intervention of~~ snow removal from point *b*.

Preliminary results show~~n~~ that the albedo effect leads to larger Δt_{air} values during the central hours of those days with higher value of solar radiation and no wind. The effect was negligible or hidden under the general thermal noise and uncertainties in those days characterised by fog, clouds cover or wind. When the weather conditions are such to make the effect evident, one day records present a similar trend as the one showed in the example in Figure 9, with night-time differences close to zero and a noise coherent with the instrumental relative uncertainty. In daytime the effect becomes evident and different in its magnitude among the different typologies of sensors (A to F).

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Differences of incident radiation in the two measurement points have also been evaluated and taken into account, in order to exclude the cases when incident radiation differences in point *a* and *b* was significant and due for example to asymmetric shadow from clouds or from the occurrence of the mountain peak shadow as mentioned in section 3.3. Having already excluded those values, the incident radiation differences were mostly consistent within the instrument uncertainty. This uncertainty was evaluated to be 35 W·m^{-2} on the basis of sensors characteristics such as sensitivity, repeatability, resolution. Records of temperature differences have been included in the data analysis only when the associated radiation difference was within this uncertainty value. As expected, due to the vicinity of the two measurement points, only few records were excluded due to larger incident radiation differences.

On the other hand, reflected radiation in the two measurement points show very large differences due to the difference in reflectivity between snow-covered area and the natural ground on point *a* (Figure 10).

Since the study intends to evaluate the largest values of the investigated effect, a threshold on the difference of reflected radiation, $\Delta Rad_{\mathrm{ref}}$, equal to 200 W·m^{-2} was chosen to better select significant temperature differences ~~records~~. The threshold was chosen by observing that below this value the temperature differences between the two measuring points were distributed in a non-deterministic way, close to the combined measurement uncertainty, site effects and data noise.

In Figure 10, plots (a) and (b) show the reflected radiation recorded in position *a* and *b* during the entire period. Plot (c) shows the differences of the reflected radiation recorded in position with and without snow. The threshold value on the reflected radiation is applied.

On this subset, a further data selection is applied, by excluding the values of temperature differences among pairs of sensors that fall below the Δt_{site}. This is the reason why the total number of significant records are not the same for all pairs of instruments. The number of available data for each pair of instruments was found to be proportional to the amplitude of the albedo effect. This result is not surprising, since when the differences distribution is skewed towards larger values, it follows that more temperature differences are found above the Δt_{site} limit. This is clearly evidenced in the following Figure 12.

4.2 Albedo effect evidence from recorded data

As a preliminary analysis, records from the deployed instruments were initially considered all together, as a single set, to understand an overall meaningful effect. The plot in Figure 11 shows the distribution of Δt_{air} amplitude grouped in bins of 0.2 °C regardless the sensors typologies.

The most frequent values of Δt_{air} are found between 0 °C and 0.4 °C, with a significant number of records between 0.4 °C and 1.6 °C. The less populated classes are from 2 °C to 4 °C. Maximum Δt_{air} values ranged up to 3.8 °C while 95 % of the values were found to be within 2.4 °C, which can be considered the highest significant value for this specific experiment.

The analysis shows that no instrument is immune from the effect, resulting in different values of Δt_{air} depending on the different technical features. Records where then segregated according to the manufacturers and typologies as reported in the following plots (Figure 12).

As in the previous histogram (Figure 11), most records are concentrated between 0 °C and 2 °C.

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques
Discussions

Looking at each pair of sensors, it became clearer that Type B and Type F show a wider range of Δt_{air}. The temperature differences for Type A, Type C and Type D sensors were always under 1.5 °C, and almost all records were concentrated between 0 °C and 1 °C, while Δt_{air} for Type E stays under 2 °C. In some cases, Δt_{air} reached 3.1 °C for Type B and 3.8 °C for Type F. Table 6 summarises the maximum difference Δt_{air} ~~measured,~~ for each instrument type with the associated uncertainty. The expanded uncertainty $U_{\Delta t_{air}}$ is reported with coverage factor $k = 2$ and confidence level of 95%.

4.3 Wind speed and radiation effects

Further data analysis was addressed to evidence the relations between temperature differences and the main quantities of influence such as wind speed and radiation.

Figure 13 shows Δt_{air} values, as calculated in previous section, as a function of wind speed.

Values between 0 m s^{-1} and 5 m s^{-1} were observed. As expected, increasing wind significantly reduce the albedo effect due to air mixing in the sensor area and to the increase in heat dissipation by convection. For wind speed greater than 3 m s^{-1} the effect was clearly reduced.

In the same plot for each measurement value, grey scale is used to evidence the difference of reflected radiation, $\Delta Rad_{\mathrm{ref}}$, associated to each temperature difference calculated.

To better evidence the behaviour of albedo effect, Figure 14 shows values of Δt_{air} as a function of $\Delta Rad_{\mathrm{ref}}$. In the plots, a positive trend of Δt_{air} is apparent for type B and type F instruments, but the spread in values does not allow for a definition of a relation by fitting, nor it was possible to calculate a function between Δt_{air} and $\Delta Rad_{\mathrm{ref}}$.

4.4 Night-time reversed effect

Differences of air temperature Δt_{air} during the night were also recorded to evaluate possible reversed differences: air temperature in the snow-covered area could result colder than in the ground area without snow, in some nights. This effect was observed in almost absence of wind, in the early spring time, when the day temperature rises above 0 °C. In this case, the ground covered by snow keeps its temperature around 0 °C, due to the energy absorbed in the snow phase transition to liquid, while the land portion without snow can become warmer and emit some heat at night. This effect falls out of the scope of this investigation, since it was not related to the instruments, but is a thermodynamic phenomenon here reported for completeness. The maximum observed meaningful difference was of 0.6 °C with opposite sign to the one recorded in daytime.

5. Discussion

The analysis here presented shows that the backward reflected radiation from a snow-covered surface affects the reliability of meteorological thermometers by transferring extra heat. This effect results in a temperature increase here evaluated with respect to identical sensors positioned in the same site, at the same time, but placed over natural ground not covered by snow. Air

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

temperature records are therefore higher than the expected true value, being the latter the actual temperature of the atmospheric air in the site, which also takes into account the warming of the air due to snow presence.

The main considerations are here summarised:

- Some typologies of instruments resulted more influenced than others, with significant differences observed even over 3 °C;
- Out of the whole group of instruments, 95 % of the differences were found within 2.4 °C;
- Highest temperature differences between couples of identical instruments were found to significantly vary among the different types of shields;
- Forced ventilated shields differences were found falling in the lower mean values recorded, with naturally ventilated shields with helical shapes scoring the lowest difference values;
- The highest temperature differences between couples of identical instruments were found in conjunction with the maximum reflected radiation differences between the two positions, as expected;
- The wind has the effect of reducing the highest temperature differences between couples of identical instruments;
- The overall uncertainty on temperature differences in field conditions ranged between 0.1 °C and 0.4 °C in $k = 2$;
- The distribution of differences as a function of the reflected radiation was found, at first approximation, to be uniform.

This experiment also evaluated a minor reversed effect at night of maximum 0.6 °C where the sensors over the snow-removed area warmed more than those with the snow below. This effect and value are not within the scope of this investigation and is here reported just as an occurrence due to the artificial removal of the snow.

Three limitations are identified in this work: (a) a limited number of instruments available (b) short period of field measurement (c) non-perfect site. Some considerations are reported on how those limitations do not affect the overall value of the study:

(a) As reported, this work involved six couples of identical thermometers and shields, plus a complete set of instruments to measure ~~all~~ other quantities of influence. Although limited in number, the instruments selected covered ~~all the instrumental commercial solutions~~ of modern meteorological sensors with a reasonable balance of aspirated, naturally ventilated and new designs included. It is not expected that other studies involving this kind of instrumentation, could show very different values of the effect.

(b) Due to the costs of installation, maintenance, site activities, work in remote location, this study required a funded project. Project rules are such that results must be delivered during the three years of the project lifetime. For this reason, after the selection process, laboratory characterisation and field installation activities, only one winter season was left before the project conclusion. Nevertheless, the number of days with snow cover lasted from November to March, allowing to meet almost all meteorological conditions in the site, including radiation and wind variability. Repeating the experiment is surely beneficiary to apply more statistical analysis, although it is not expected that the

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques Discussions

differences among the instruments vary significantly. A valuable representativeness of the obtained results, with respect to atmospheric ranges and measurement conditions is here considered achieved.

(c) An appropriate site with easy access for maintenance, long lasting presence of snow but not excessive thickness, electric current to feed all instrumentation and ventilated screens (solar panels cannot supply enough power), maintenance and surveillance, closeness to have staff removing the snow at early occurrence is not an easy find. Moreover, the site needs to be flat, covered by natural soil, without significant shadows and slopes. Alpine valleys rarely allow such multitude of conditions. Thanks to an intensive site selection and the support of local Communities, a site was identified showing reasonable compromises among the required characteristics. Field measurements on possible effects of the site features (trees, small building, shadow) were in any case made, to select data and correct ~~from~~ systematic effects. Such effects resulted to be of the order of magnitude of the uncertainty and only for one couple of sensors they were taken into account in data analysis.

From the considerations above, the results here delivered are considered valid to understand the order of magnitude of the effect. This work also gives an example on how to evaluate this phenomenon and take it into account in terms of correction and associated uncertainty. A complete correction curve in function of reflected radiation, wind speed for a specific instrument can be generated by users, having available at least a couple of identical instruments and the auxiliary equipment here described.

In scientific literature or documented procedures, very few - close to none, are the examples of similar evaluations, methods or prescriptions to quantify the studied effect on near surface thermometers. The work mentioned in the introduction by Huwald et al., (2009), reaches the same conclusions that "Temperature errors decrease with decreasing solar radiation and increasing wind speed" and that this effect ranges in the order of degrees Celsius. With respect to the aforementioned study, the key difference of the work here presented, which we consider as an improvement for the analysis, is that here the effect is evaluated in a relative way, instead than assuming the readings of a sonic anemometer as reference. It is agreed that non-contact thermometry is immune from a number of quantities of effect, but the accuracy achieved by using anemometers as thermometers is not sufficient to let this instrument be a reference. The method here proposed can be adopted just by buying a second identical thermometer and shield, significantly reducing costs. The resulting uncertainties in detecting the effect are reduced with respect to comparing different systems and even different physical principles in measuring air temperature. Finally, in this analysis the investigation was extended to more different kind of sensors and shields, thus making the results representative of a wider typology of solutions adopted in meteorology.

Beside delivering the numerical results, the key output of this work is a methodology for evaluating a factor affecting temperature data in climatology (and meteorology) and give an example on how this can be implemented and adopted when selecting instruments as in the case of surface stations of climatological networks.

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques
Open Access
Discussions

6. Recommendations to users and manufacturers

One of the main tasks of the MeteoMet project was to give metrological support to the meteo-climatology community, including data users, station staff and manufacturers. A synthetic indication that summarise the outcomes of the work here presented has been presented at the WMO CIMO TECO 2018 (Musacchio et al., 2018) and sent to the WMO CIMO expert team on Observation In-Situ technologies.

Following the publication of the experimental method (Musacchio et al., 2019) indications on how to design and implement a field experiment, to evaluate the errors in temperature readings in thermometers positioned above snow-covered land have been prepared and sent to WMO CIMO task and expert teams on "Metrology", "Surface Measurements" and "Overall Uncertainties". The indication is summarised as it follows.

When instruments are positioned in sites where snow can occur, it is recommended that a specific analysis is performed to evaluate the bias in data records, due to extra heating from reflected radiation, causing errors in thermometers readings. To perform the study, two identical systems (thermometers and shield) must be installed in the vicinity, one positioned above a snow-covered area and one above an area where snow is removed. The procedure here reported and the cited protocol must be followed, involving the required auxiliary instrumentation to constantly record and monitor the environmental factors of influence. The differences between the temperature readings of the sensors should be recorded for at least one full snow season, to meet most of the meteorological conditions of the sites and evaluating the associated effects and factors of influence. A correction can then be generated in terms of a relationship between temperature corrections with respect to the reflected radiation, wind speed and air temperature. The associated uncertainty budget is then evaluated through the Gaussian propagation: components of uncertainty are calculated by field analysis of systematic differences in temperature and by knowledge of each involved instrument performance, including radiometers and anemometers.

The objective of the recommendation is to report and inform users and manufacturers of instruments to consider and possibly minimize the effect of reflected radiation from a snow-covered surface on near surface temperature data. The content is also relevant for the definition of data quality and instrument features by the GCOS and the WMO in promoting climatological reference stations. For high quality installations and climate reference stations, the analysis here presented can lead to improving the data quality by adding an evaluated correction and associated uncertainty. Manufacturers should also evaluate and declare this effect on their product datasheets and where possible adopt solutions to minimize it.

7. Conclusions

The study here presented was performed to evaluate the accuracy of near-surface air temperature data series, recorded by thermometers in solar shields positioned above snow. The work investigated this phenomenon in terms of extra heating transferred to sensors, causing a temperature difference in their measurements. The well-known effect of air temperature increase due to the snow albedo effect is not here considered, since this is a meteorological phenomenon: only the extra heating induced to the sensors and shields is here considered as a bias in record series. The study strictly followed an already published

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques
Discussions

method and its associated experimental protocol. The method designs the experimental site in terms of two measuring points, equipped with the same kind of identical sensors and shields or groups of different couples of sensors and shields as in this case, mounted 2 m above the ground level. At snow occurrence, snow is removed from one measurement point, leaving natural ground exposed, thus thermometers measuring actual air temperature; snow is left on the second point and the sensors are thus exposed to the backward reflected radiation that, especially during sunny days, warms the sensors introducing a temperature difference Δt_{air}. Preparing a theoretical method and coherent experimental protocol, based on deeply discussed aspects, made the installation and the data analysis a more robust process, with less adjustments due in course of the work. The study involved a representative number and different typologies of modern sensors and solar shields, including naturally ventilated, aspirated and helical shields. The instruments were provided as commercially offered by manufacturers, equipped with dedicated data loggers and measurements have been taken in the same conditions of use. The effect was evident for all the typologies of sensors with maximum values observed in absence of wind and at high reflected radiation conditions such as in sunny days with clean fresh snow. The amplitude of the maximum effect ranged from 1.2 °C up to 3.8 °C.

The method was validated by the experimental results and can be considered a procedure for further similar investigations involving other typologies of sensors. This process can be adopted by manufacturers to test and characterise their product as well as by station staff and data users to include this effect, correction and associated uncertainty to the records. A similar analysis should be performed when selecting instruments to take part in a climate reference network, such as the planned GCOS Surface Reference Network GSRN, for those stations positioned on sites with snow presence.

Finally, further work can be addressed to evaluating correction curves in the form of temperature difference relationship with reflected radiation and wind conditions. The calculation of a correction function requires longer time of field activities, to meet the wider range of atmospheric conditions as well as having more data available for statistical analysis by filtering significant data. The uncertainty budget associated to the curve will then be completed by including the statistical analysis and all components from the instruments involved: thermometers, anemometers, radiometers.

In a site where a high-quality installation is planned to be permanent, a study like this is recommended among the overall efforts to increase data quality and understand uncertainties in meteorological observations for climate.

Author contributions

Andrea Merlone, Chiara Musacchio and Graziano Coppa designed and run the experiment, with contributions by all co-authors. Laura Massano performed the data analysis with contributions by Chiara Musacchio. Chiara Musacchio prepared the manuscript with contributions from all co-authors.

Acknowledgments

The authors want to thank the Agenzia Regionale per la Protezione Ambientale del Piemonte, the Municipality of Balme and the "Les Montagnards" mountain hut for the valuable support, as well as the manufacturers that took part in the experiment by providing the instrumentation.

This work is being developed within the frame of the EMRP (European Metrology Research Programme) joint research project "METEOMET". The EMRP is jointly funded by the EMRP participating countries within EURAMET and the European Union.

References

Barry, R. and Blanken, P.: Microclimate and Local Climate, Cambridge University Press, Cambridge., 2016.

BIPM and Joint Committee For Guides In Metrology: Evaluation of measurement data — Guide to the expression of uncertainty in measurement. [online] Available from: http://www.bipm.org/en/publications/guides/gum.html, 2008.

Burt, S. and de Podesta, M.: Response times of meteorological air temperature sensors, Q. J. R. Meteorol. Soc., 146(731), 2789–2800, doi:10.1002/qj.3817, 2020.

Camuffo, D. and Jones, P.: Improved Understanding of Past Climatic Variability from Early Daily European Instrumental Sources, edited by D. Camuffo and P. Jones, Springer Netherlands, Dordrecht., 2002.

CCT: Strategy Document for Rolling Programme Development for 2018 to 2027. The Consultative Committee for Thermometry., 2017.

Coppa, G., Quarello, A., Steeneveld, G., Jandric, N. and Merlone, A.: Metrological evaluation of the effect of the presence of a road on near-surface air temperatures, Int. J. Climatol., in press, doi:10.1002/joc.7044, 2021.

GCOS: GCOS Surface Reference Network (GSRN): Justification, requirements, siting and instrumentation options. [online] Available from: https://library.wmo.int/doc_num.php?explnum_id=6261, 2019.

Huwald, H., Higgins, C. W., Boldi, M. O., Bou-Zeid, E., Lehning, M. and Parlange, M. B.: Albedo effect on radiative errors in air temperature measurements, Water Resour. Res., 45(8), 1–13, doi:10.1029/2008WR007600, 2009.

Lacombe, M., Bousri, D., Leroy, M. and Mezred, M.: Instruments and Observing Methods report No . 106 WMO Field Intercomparison of Thermometer Screens/Shields and Humidity Measuring Instruments., 2011.

Merlone, A., Lopardo, G., Sanna, F., Bell, S., Benyon, R., Bergerud, R. A. A., Bertiglia, F., Bojkovski, J., Böse, N., Brunet, M., Cappella, A., Coppa, G., del Campo, D., Dobre, M., Drnovsek, J., Ebert, V., Emardson, R., Fernicola, V., Flakiewicz, K., Gardiner, T., Garcia-Izquierdo, C., Georgin, E., Gilabert, A., Grykalowska, A., Grudniewicz, E., Heinonen, M., Holmsten, M., Hudoklin, D., Johansson, J., Kajastie, H., Kaykisizli, H., Klason, P., Kňazovická, L., Lakka, A., Kowal, A., Müller, H., Musacchio, C., Nwaboh, J., Pavlasek, P., Piccato, A., Pitre, L., de Podesta, M., Rasmussen, M. K., Sairanen, H., Smorgon, D., Sparasci, F., Strnad, R., Szmyrka- Grzebyk, A. and Underwood, R.: The MeteoMet project - metrology for meteorology:

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques Discussions

Challenges and results, Meteorol. Appl., 22(January), 820–829, doi:10.1002/met.1528, 2015a.

Merlone, A., Musacchio, C. and Sanna, F.: The Metrology for Meteorology Conference: MMC 2014, Meteorol. Appl., 22, 817–819, doi:10.1002/met.1548, 2015b.

Merlone, A., Sanna, F., Beges, G., Bell, S., Beltramino, G., Bojkovski, J., Brunet, M., del Campo, D., Castrillo, A., Chiodo, N., Uytun, A., Voldan, M., Colli, M., Coppa, G., Cuccaro, R., Dobre, M., Drnovsek, J., Ebert, V., Fernicola, V., Garcia-Benadí, A., Garcia-Izquierdo, C., Gardiner, T., Georgin, E., Gonzalez, A., Groselj, D., Heinonen, M., Hernandez, S., Högström, R., Hudoklin, D., Kalemci, M., Kowal, A., Lanza, L., Miao, P., Musacchio, C., Nielsen, J., Nogueras-Cervera, M., Oguz Aytekin, S., Pavlasek, P., de Podesta, M., Rasmussen, M. K., Del-Río-Fernández, J., Rosso, L., Sairanen, H., Salminen, J., Sestan, D., Šindelářová, L., Smorgon, D., Sparasci, F., Strnad, R., Underwood, R., Uytun, A. and Voldan, M.: The MeteoMet2 project—highlights and results, Meas. Sci. Technol., 29(2), 025802, doi:10.1088/1361-6501/aa99fc, 2018.

Musacchio, C., Massano, L., Coppa, G., Merlone, A. and Hofstaetter, C.: Effect of snow-reflected radiation in near surface air temperature, in The 2018 WMO/CIMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation (CIMO TECO-2018)2., 2018.

Musacchio, C., Coppa, G. and Merlone, A.: An experimental method for evaluation of the snow albedo effect on near-surface air temperature measurements, Meteorol. Appl., 26(1), 161–170, doi:10.1002/met.1756, 2019.

Nigrelli, G., Fratianni, S., Zampollo, A., Turconi, L. and Chiarle, M.: The altitudinal temperature lapse rates applied to high elevation rockfalls studies in the Western European Alps, Theor. Appl. Climatol., 131(3–4), 1479–1491, doi:10.1007/s00704-017-2066-0, 2018.

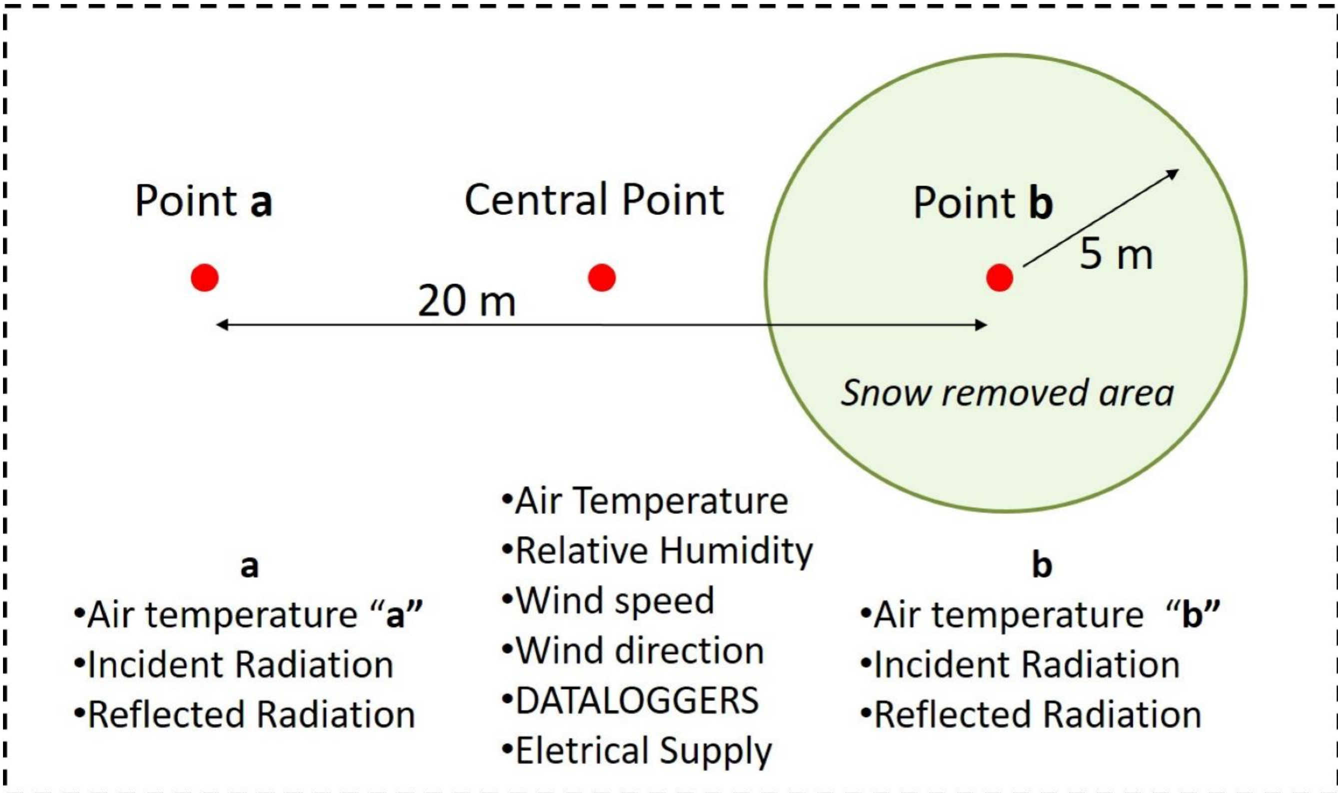
Stevenson, T. C. E.: New description of box for holding thermometers, J. Scottish Meteorol. Soc., 1, 122, 1864.

Thorne, P. W., Diamond, H. J., Goodison, B., Harrigan, S., Hausfather, Z., Ingleby, N. B., Jones, P. D., Lawrimore, J. H., Lister, D. H., Merlone, A., Oakley, T., Palecki, M., Peterson, T. C., de Podesta, M., Tassone, C., Venema, V. and Willett, K. M.: Towards a global land surface climate fiducial reference measurements network, Int. J. Climatol., 38(6), 2760–2774, doi:10.1002/joc.5458, 2018.

WMO: WMO/CIMO #8 Guide to meteorological instruments and methods of observation. [online] Available from: https://library.wmo.int/index.php?id=12407&lvl=notice_display, 2012.

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques
Discussions



525

Figure 1: Scheme of the installation area. Points “*a*” and “*b*” host the identical thermometers and shields of each couple. The central point hosts auxiliary equipment, data loggers and sensors for measuring quantities of influence.

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques
Discussions

Atmospheric Measurement Techniques Open Access EGU Discussions

530 **Figure 2: Part of the collected couples of instruments ready for the preliminary laboratory characterisation.**

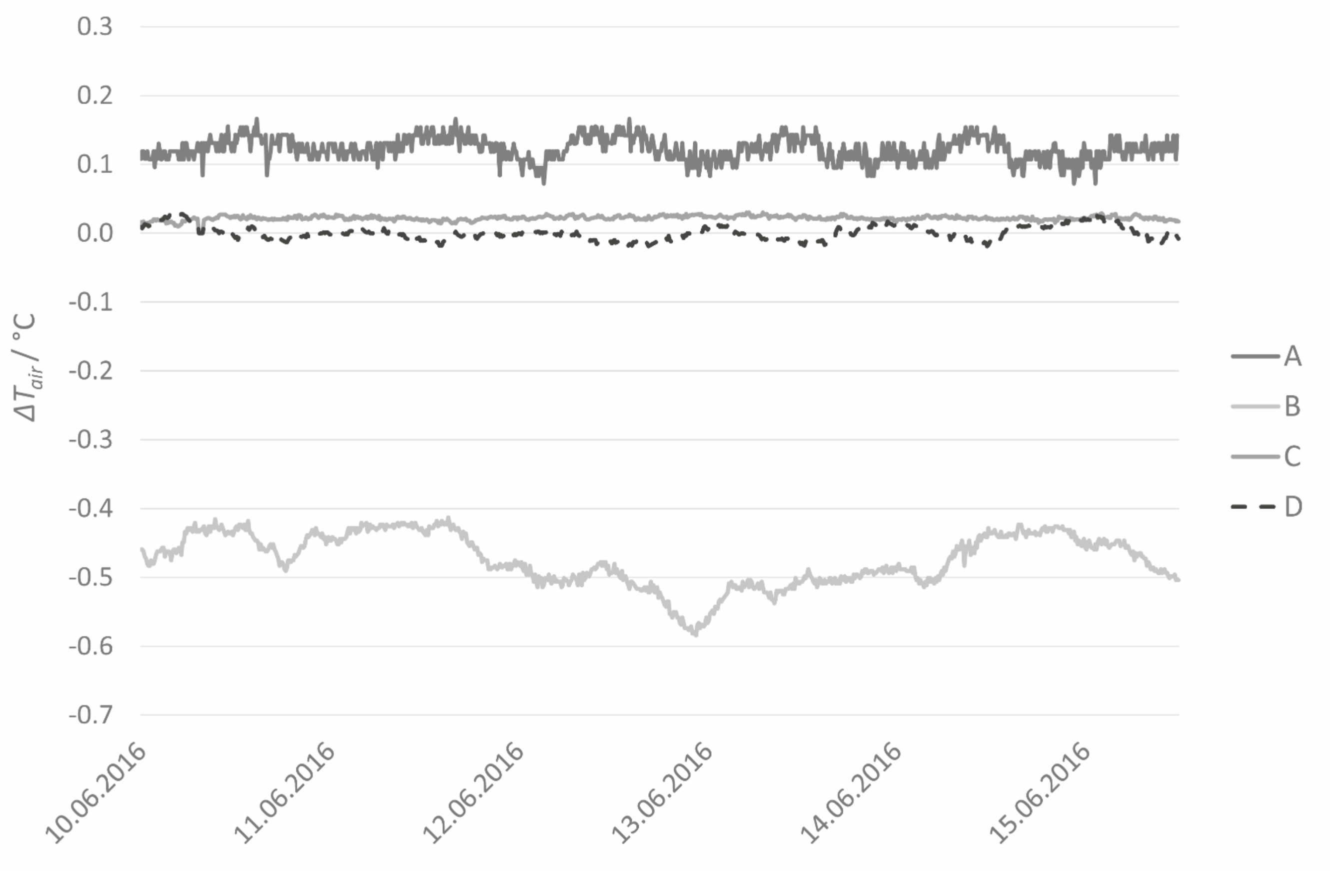


Figure 3: Example of laboratory characterisation for four couples of sensors. One-week acquisition of differences between the readings of the two sensors of the pair.

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques
Discussions

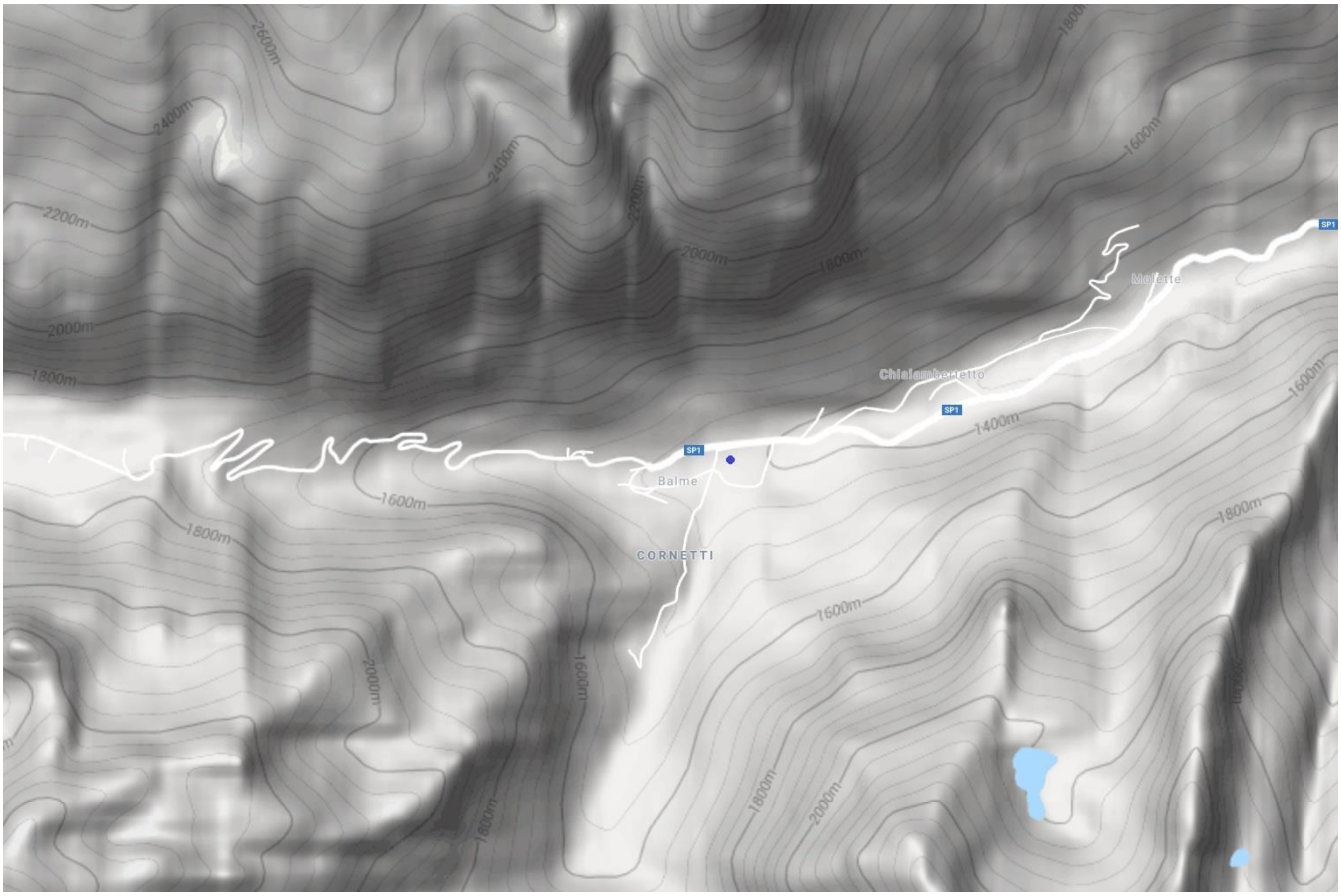


Figure 4: Google Relief map (©Google, 2017) of the area (approximately 2.5x1.5 km) surrounding the measurement site and its topography. The Ala valley is aligned in an East-West way: mountains close the valley from the North, while on the South a small lateral valley opens up the horizon to other high mountains. The measurement site is marked with a blue dot.

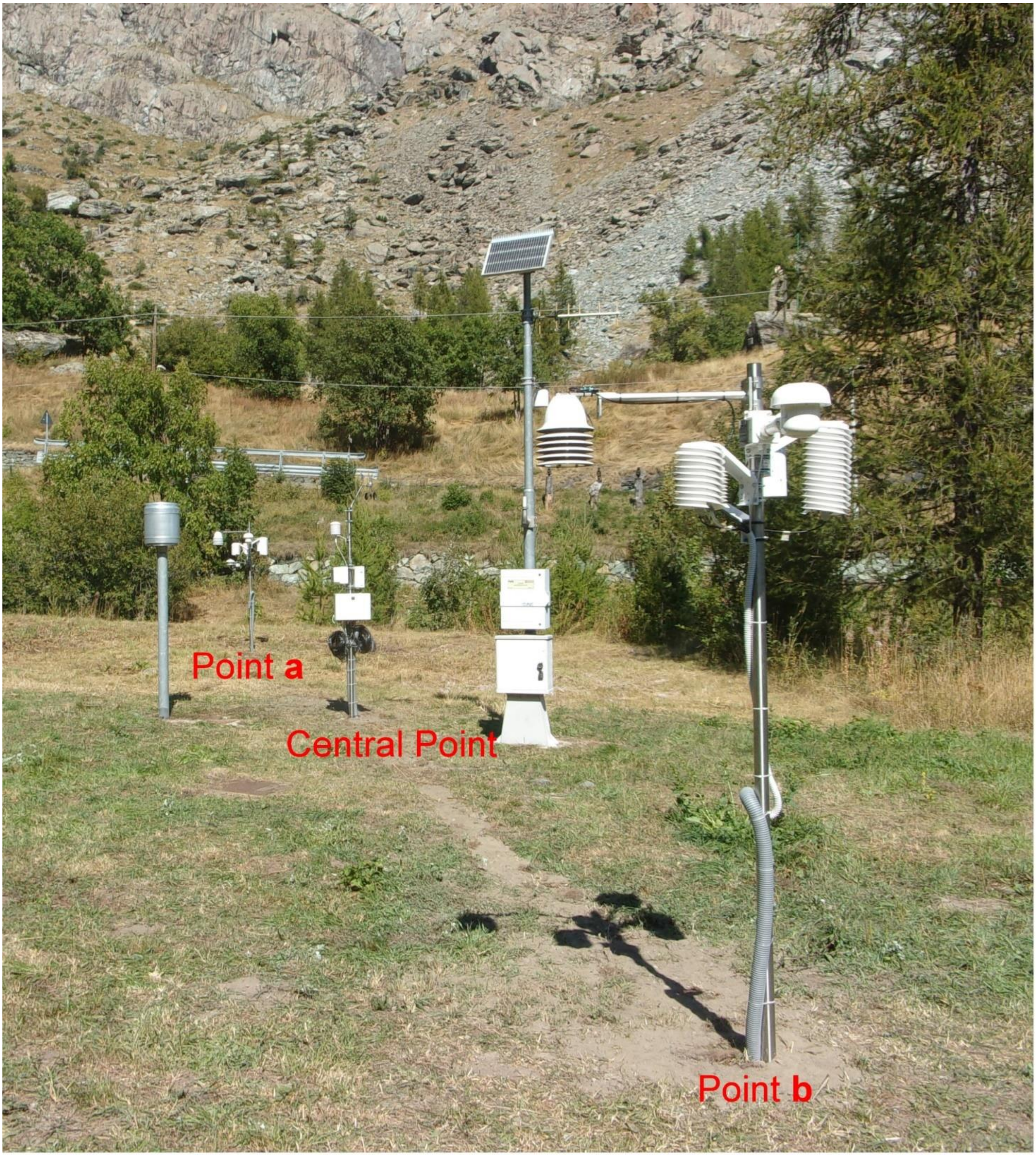
https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques Discussions

540 **Figure 5: Zoomed in (200x100 m) Google Earth (©Google 2017) picture of the experimental site. The approximate positions of the two measurement stations are marked by the yellow spots.**

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques
Discussions



545 **Figure 6: The experimental site in summer, after the installation of the instruments.**

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques
Discussions

Figure 7: The measurement site in its original, un-shovelled configuration. In the background position *a*, where snow will be left. In the foreground position *b*, with snow still to be removed.

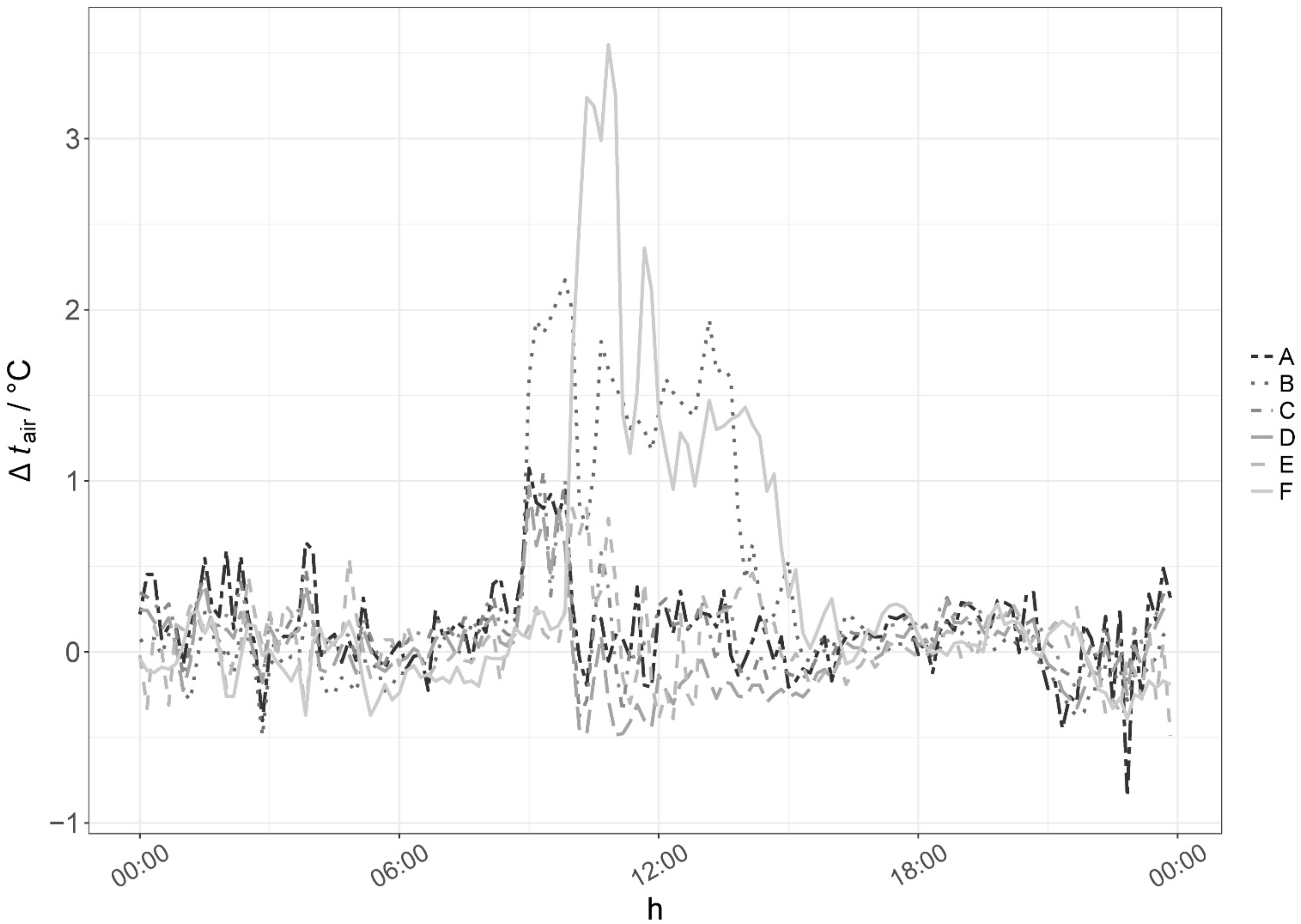
550

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Figure 8: Point *b* in snow-removed condition, after a snowfall.

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

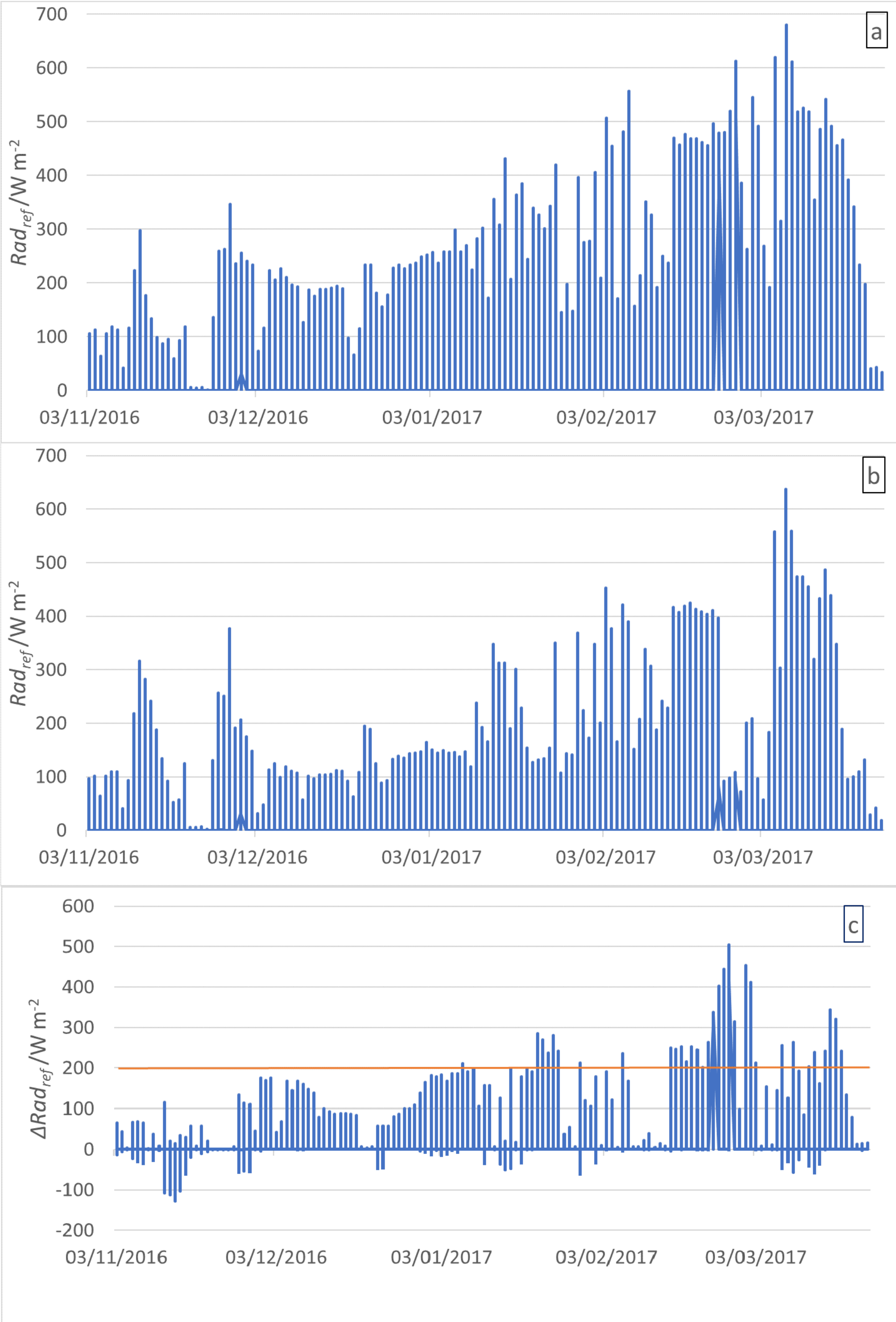
Atmospheric Measurement Techniques
Discussions



555 **Figure 9: A typical plot of a one-day-long acquisition, demonstrating the studied effect in terms of temperature differences Δt_{air} between each of the couples of identical sensors (25 February 2017). The day has been selected as a representative example, with snow removed some days before. Weather was mainly sunny, with maximum incident radiation of 700 W·m^{-2}, maximum reflected radiation of 500 W·m^{-2} in snow condition and less than 100 W·m^{-2} in the area where snow was removed. Hours are reported in local time (Central European Time - CET).**

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques
Discussions



560

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques
Discussions

Figure 10: Results of measured reflected radiation recorded in position *a* - sensor above snow (a), and *b* – sensor above natural ground (b) during the entire period of the experiment. Differences of reflect radiation recorded in position "*a*" and "*b*", ΔRad_{ref}, are shown in (c). The threshold chosen to better select the associated temperature differences are indicated by horizontal lines.

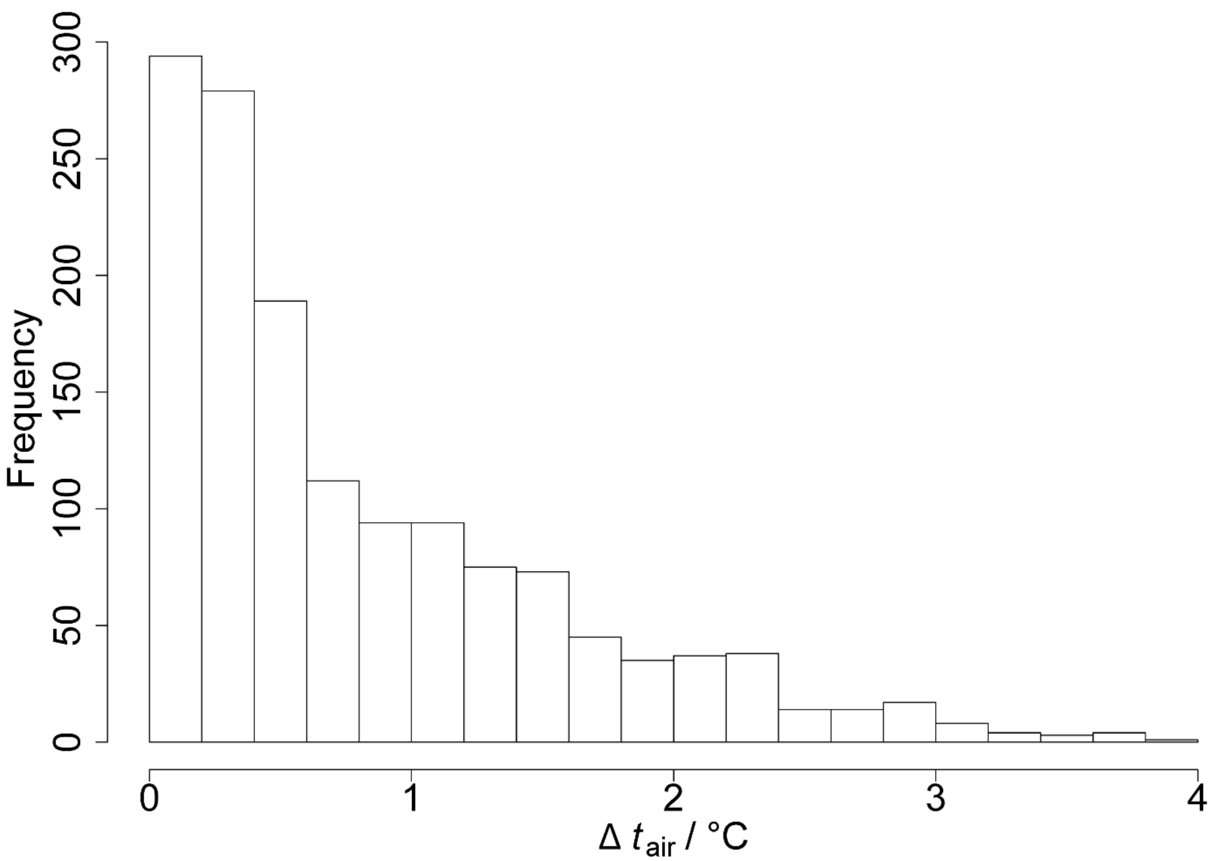
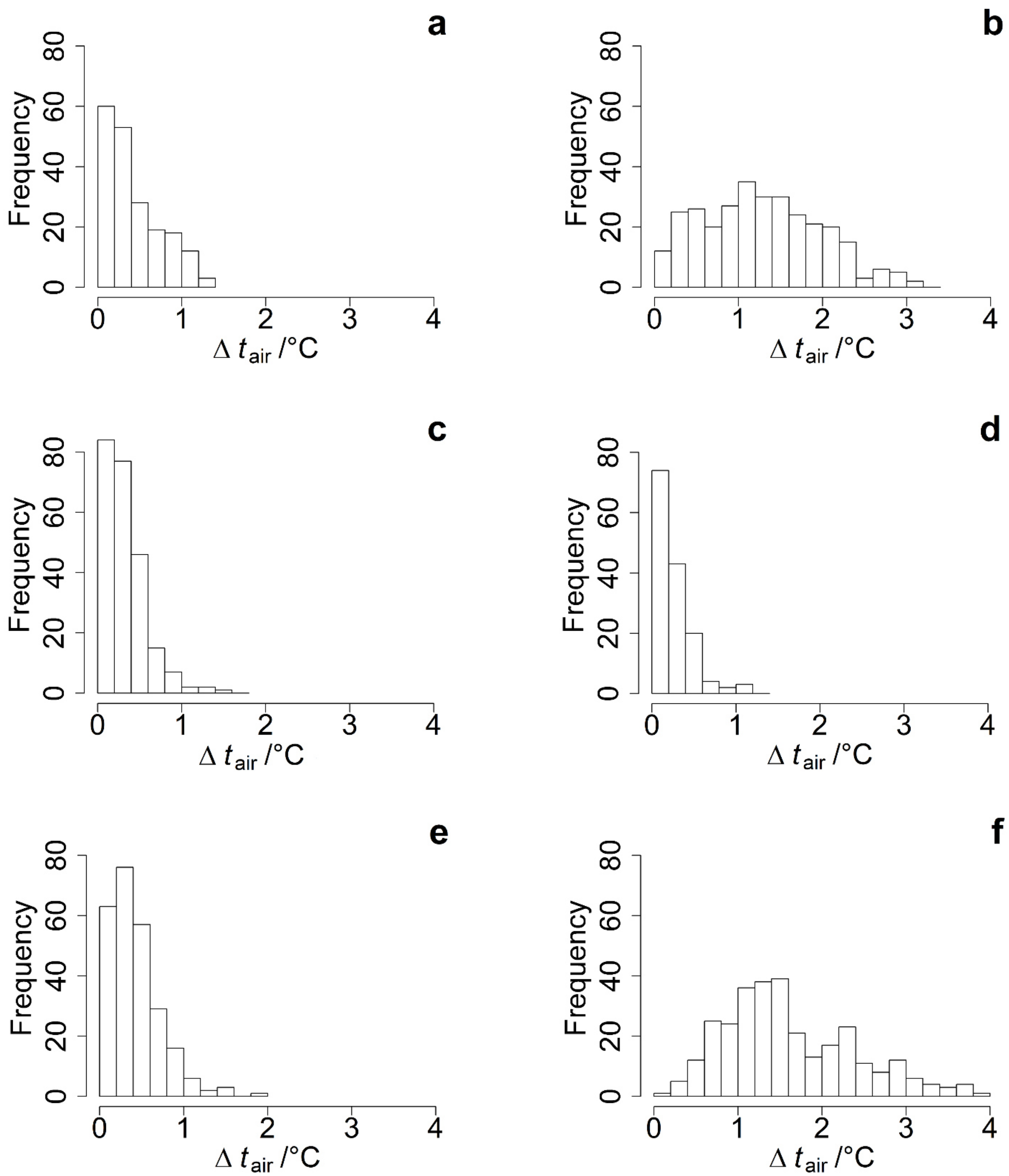


Figure 11: Frequency of temperature difference, Δt_{air}, considering all pairs of instruments, and the records that exceed the radiation difference threshold.

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques Discussions



https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques
Open Access
EGU
Discussions

Figure 12: Results of evaluation of Δt_{air} for each pair of sensors from different manufacturers. Instrument types are identified with letters from a to f. The histogram is divided in bins of 0.2 °C and the number of occurrences of Δt_{air} is shown for each instrument.

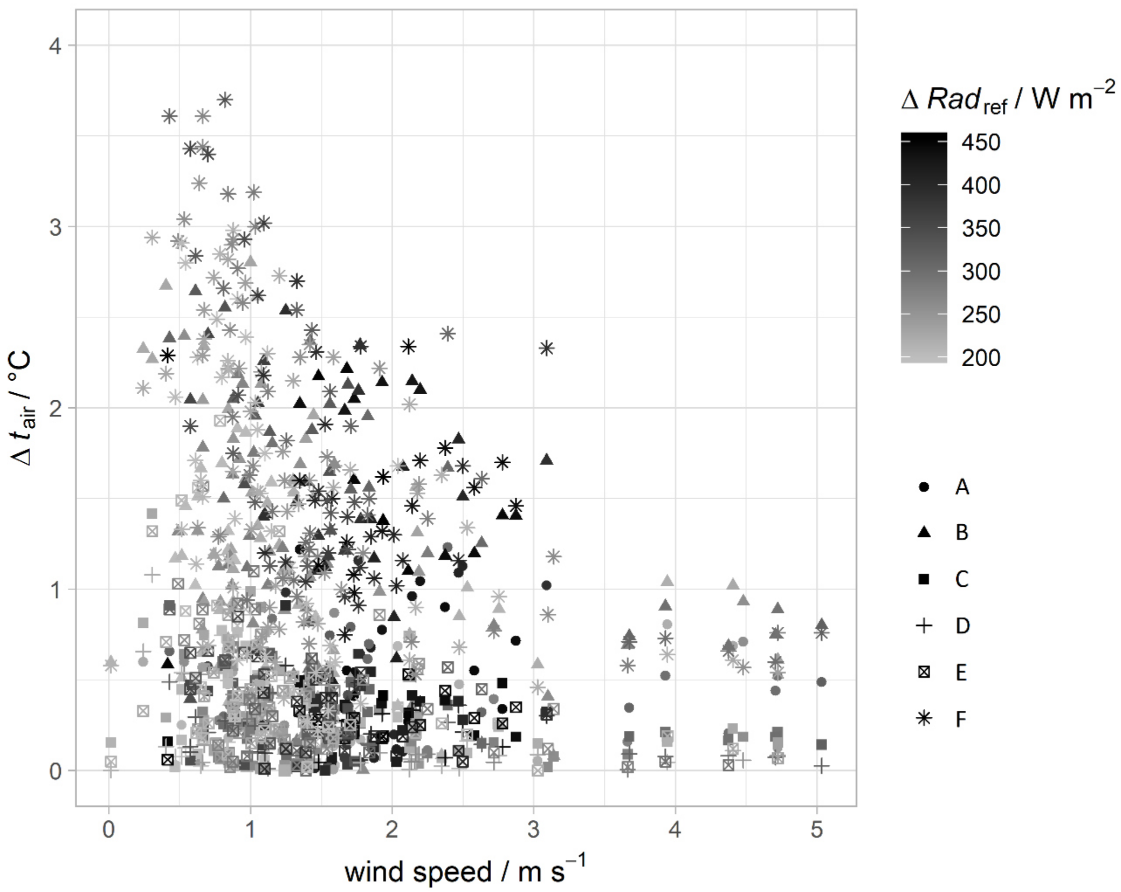
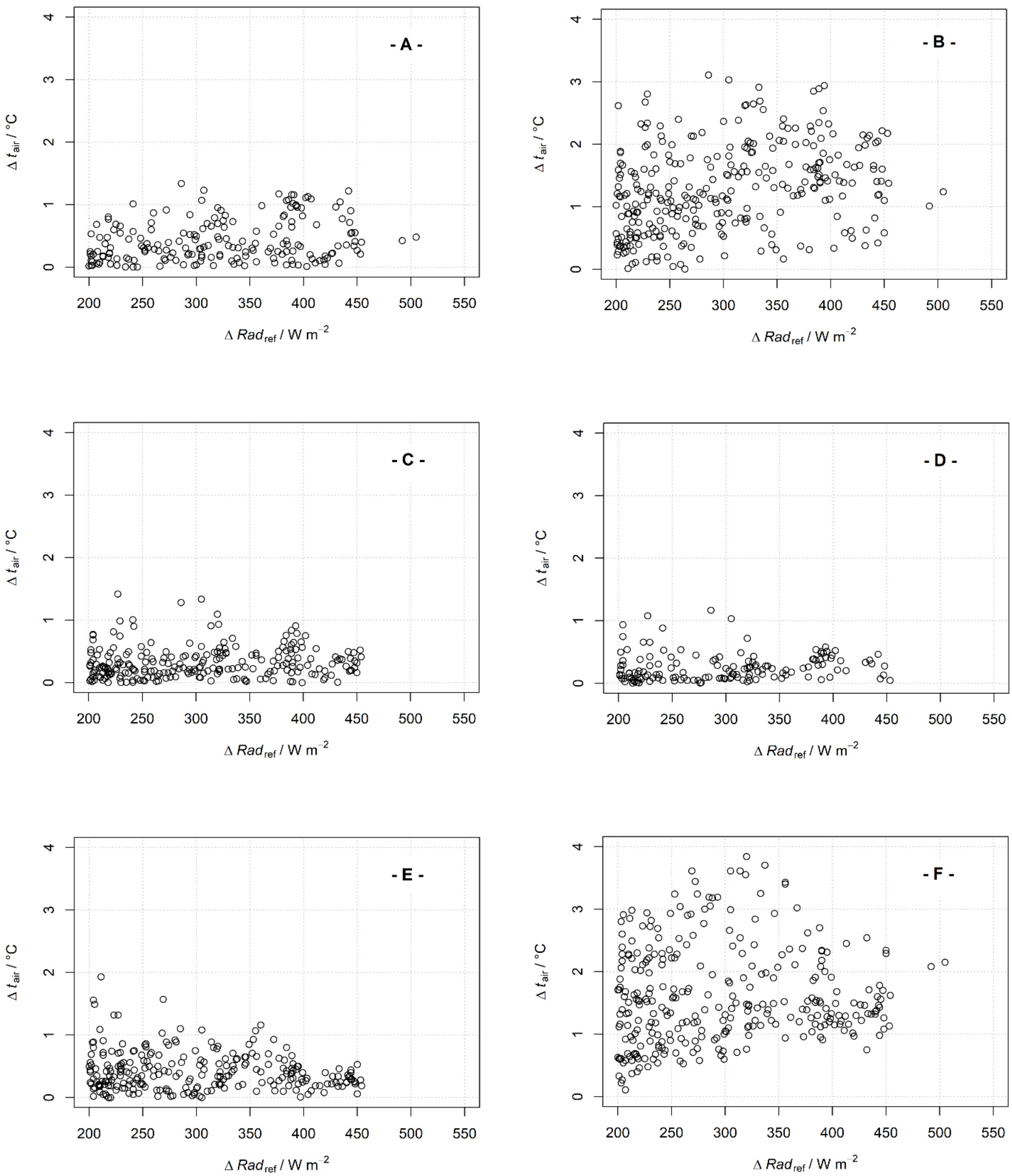


Figure 13: **Temperature difference Δt_{air} measured as a function of wind speed for all the instrument type A-F. Grey scale is used to evidence the value of the difference of reflected radiation, ΔRad_{ref},related to each value of Δt_{air} reported.**

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques
Discussions



https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Figure 14: Temperature differences Δt_{air} evaluated in the data analysis plotted as a function of reflected radiation difference between point a and b. Labels from A to F identify the instrument type.

580

Table 1. Selected air temperature measurement instruments and their main characteristics.

Instrument ID	**Sensor type**	**Shield type**
Type A	Pt100	Fan aspirated
Type B	Pt100	Passive
Type C	Thermo hygrometer	Passive helicoidal
Type D	Thermo hygrometer	Passive helicoidal
Type E	Pt100	Passive
Type F	Pt100	Passive

Table 2. Sensors used for measuring the quantities of influence and their positioning referred to the scheme of Figure 1.

Quantity	**Sensor type**	**positioning (see Fig. 1)**
Temperature and Relative Humidity	Pt100 class A and capacitor (thermo-hygrometer)	Central point
Wind	Cups and vane	Central point
Global incident Radiation	Thermopile (pyranometer)	Point a, facing up
Global reflected Radiation	Thermopile (pyranometer)	Point a, facing down

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques
Discussions

Global incident Radiation	Thermopile (pyranometer)	Point *b*, facing up
Global reflected Radiation	Thermopile (pyranometer)	Point *b*, facing down

Table 3. Results of the evaluation of Δt_{instr} and the associated uncertainties for each instrument type.

Sensor type	**Type A (°C)**	**Type B (°C)**	**Type C (°C)**	**Type D (°C)**	**Type E (°C)**	**Type F (°C)**
Δt_{instr}	0.12	-0.47	0.022	0.002	0.043	0.063
$u_{\Delta t_{instr}}$	0.05	0.09	0.015	0.026	0.035	0.067

Table 4. Results of the evaluation of Δt_{site} and the associated uncertainties for each instrument type.

Sensor type	**Type A (°C)**	**Type B (°C)**	**Type C (°C)**	**Type D (°C)**	**Type E (°C)**	**Type F (°C)**
Δt_{site}	0.02	0.17	0.12	0.10	0.08	0.11
$u_{\Delta t_{site}}$	0.02	0.17	0.11	0.11	0.09	0.09

Table 5. Contributions to the uncertainty budget evaluated in the laboratory and on-field characterization.

	Type A (°C)	**Type B (°C)**	**Type C (°C)**	**Type D (°C)**	**Type E (°C)**	**Type F (°C)**
u_{res}	0.012	0.003	0.001	0.001	0.01	0.01
u_{lab}	0.01	0.01	0.01	0.01	0.01	0.01

https://doi.org/10.5194/amt-2021-63
Preprint. Discussion started: 15 April 2021

Atmospheric Measurement Techniques Discussions

$u_{\Delta t_{instr}}$	0.05	0.09	0.015	0.026	0.035	0.067
$u_{\Delta t_{site}}$	0.02	0.17	0.11	0.11	0.09	0.09
$u_{\Delta t_{air}}$	0.05	0.19	0.11	0.11	0.10	0.11
$\boldsymbol{U_{\Delta t_{air}}}$ (k = 2)	**0.11**	**0.38**	**0.22**	**0.23**	**0.20**	**0.23**

Table 6. Maximum difference - $\boldsymbol{\Delta t_{air}}$ - measured, for each manufacturer on a significant number of events and with the associated uncertainty from Table 5. Values are rounded at first decimal and $\boldsymbol{U_{\Delta t_{air}}}$ rounded up according to normative (EA-4/02).

Instrument Type	Max diff. Δt_{air} (°C)	$U_{\Delta t_{air}}$ (°C) (k = 2)
A	1.4	0.1
B	3.1	0.4
C	1.4	0.3
D	1.2	0.3
E	1.9	0.2
F	3.8	0.3

Review of AMT-2021-63:

"Effect of snow-covered ground albedo on the accuracy of air temperature measurements"

by C. Musacchio, G. Coppa, G. Beges, C. Hofstätter-Mohler, L. Massano, G. Nigrelli, F. Sanna, A. Merlone

This paper investigates the influence of a snow cover on near-surface air temperature measurements which are generally subject to solar radiation-induced errors (even when sensors are protected in a shield), Such errors are amplified by a contribution of surface-reflected shortwave radiation, additionally heating sensor shields and sensors. To quantify this so-called albedo effect, several pairs of identical sensor/shield combinations are deployed at two neighboring locations, experiencing the same meteorological conditions. At one site the seasonal snow cover was unchanged, at the other snow has been removed after every snowfall such that these sensors were mounted above bare ground. Prior to the field experiment, the sensor pairs have been tested and characterized in controlled laboratory environments to determine relative errors and biases. Air temperature, shortwave radiation and wind speed data have been collected and analyzed to quantify the albedo effect on air temperature measurements by comparing temperature observation at the two sites. Observed temperature differences were related to the magnitude of incident solar radiation and wind speed to assess their relative importance. Albedo-induce errors were found to be as large as 3.8°C. The study follows certain WMO guidelines and proposes recommendations for measurement protocols to be applied on reference stations, e.g., in the context of GCOS.

The paper constitutes an important and interesting contribution to the improvement of measurement practices and the quantification and reduction of measurement errors, here specifically albedo-induced uncertainty and error in air temperature measurements. The study proposes an interesting experimental setup and obtains useful results from the acquired data. The paper is a perfect topical match for AMT.

We thank the reviewer for the kind words, which acknowledge the importance of our work.

Nevertheless, I have several concerns, suggestions, and questions which I would like to bring to the attention of the authors, detailed in the list of specific and sometimes more general comments below. It seems to me that the available data and results are partly underexploited and carry potential for more conclusive results. The manuscript needs substantial revision as indicated by the comments below. Integrating these suggested elements, clarifications, and discussions could strengthen this work and make it a more mature and consolidated paper. I am aware that I am raising a lot of points, but I hope that these comments are constructive suggestions and helpful for further improving this paper.

Thanks for the suggestions, comments, and points that help us improving the manuscript. All of them are addressed below:

1. The manuscript needs linguistic editing to improve spelling, grammar, and wording.

The manuscript has been thoroughly reviewed by English mother tongue collaborators.

2. The abstract should be more specific with respect to results and findings of the study.

The second part of the abstract has been revised, adding the main findings and more details on the results. Some text on recommendation has been removed, to keep word limits.

3. Field data have been acquired at 2m above ground. When the paper talks about "near-surface" air temperature measurements, what is be the upper limit for this term, and up to which height do you expect the albedo effect influencing air temperature measurements? E.g., sensor on a 10m mast?

For the experiment we relied on WMO/CIMO guide #8 to instruments and methods of observation, which prescribes that near-surface sensors should be placed between 1.25 and 2 m above ground. In the chosen site, snow cover higher than 2-1.25 = 0.75 m is extremely rare, therefore we stipulated that no mechanism for adjusting sensor height was necessary, as it was within specifications in all conditions. Temperature sensors at 10 m of height are not standard and not considered as "near-surface", therefore we did not estimate the contribution of albedo influencing such installations.

In order not to weigh down the text, these considerations have not been added. Only a quick reference to the definition of "near surface" has been added at the first occurrence in the text.

4. Related to comment #2, given a 5m radius of the snow-free surface and a 2m height of the sensors above ground, how does this relate to the field of view of the radiometer used which normally sees a hemisphere. In other words, what is the footprint the sensors (pyranometer and thermometer) see? Is a 10m diameter sufficiently large to neglect contributions from outside the circle?

The stated field of view of the pyranometer is 180° (but the effective one is a little less). In our configuration, a 1.5 m high pyranometer sees a diameter of 10 m at the ground under an angle of 146°, which we considered enough for our purposes. For trigonometric considerations, doubling the free radius under the sensor (thus, multiplying the area to be freed from snow fourfold) would have increased the field of view by mere 16°. This consideration has been added in the revised manuscript.

5. Section 2: I don't see much of a "theoretical study" in this section; is this statement related to previous work of the authors? Maybe consider revising the section title.

Yes, the "theoretical study" refers to Musacchio et al., 2019. The section title has been revised in the updated manuscript

6. Section 2.2 could briefly mention the influence of the identified quantities. In particular, the role of snow depth and the snow properties should be mentioned. Why is snow density not considered?

The influence of the identified quantities was more explicit in the previous work, and we felt it would have been unnecessary repetition. However, it has been added in the revised manuscript.

Snow density was not considered because literature works (like for instance Bohren and Beschta 1979) concluded that snowpack albedo is only weakly dependent upon it. This has been added to the revised manuscript.

7. Has the solar zenith angle been considered to have an impact? This is not only influencing the shortwave radiative heat flux but also the snow albedo (even significantly at large zenith angles).

These are really questions for the previous theoretical study (Musacchio et al 2019). However, zenith angle was not considered for two reasons: first, the site is in a narrow valley and sun's rays never come from large zenith angles, being blocked off by the orography of the site; second, while it is true that zenith angle influences the flux of reflected radiation, we only considered the ratio between global and reflected which is much less influenced by it: for instance Xiong et al (2015) show that, for high values of albedo (like those of snow), the dependence is almost flat; for smaller values, the dependence steepens after ~60° which is basically never achieved at our site.

8. P05L143: Why is the sensor at the central location considered the "reference" for the other two air temperature measurements if subsequently not considered in the analysis?

"Reference" in this instance means only that it was used only to have a second set of measurements of temperature to check if the other sensors were measuring correctly. Maybe the wording is misleading, so it has been changed in the revised manuscript.

9. Tables 1 and 2: Provide manufacturer, device model and sensor accuracy, resolution and response time information for temperature sensor and shield in the Tables. Additionally, describe the differences between sensor/shield combinations C and D, as well as E and F; this information is absent from the text and table. Which data logging devices have been used in this study?

Since the very beginning of the experiment, it was decided that no manufacturer and model would have been made explicit in the text, for various reasons: it was not intended as a competition, and we did not want to influence the market in any way. Makers have been informed on the results of their instruments, but this is an information we do not intend to make public.

Other information, like (stated) uncertainty, and resolution will be added to the Tables. Logging devices were the standard ones provided by each manufacturer, whose uncertainty was added in the final budget.

10. Figure 2 is not necessary and can be removed. The completed Tables 1 and 2 will be more useful.

We think Figure 2 is important, given that we are not going to explicitly state manufacturer and models of the sensors, because it immediately conveys the different sizes and shapes of the shields. Besides, it shows the laboratory characterization phase, which was an integral part of the experiment. We would therefore like to keep the picture in the revised manuscript.

11. Figure 3: what does the sign of the results indicate? Could be shown as absolute differences, to improve Figure clarity. Are all sensors sampled at the same frequency? Not clear for C and D; solid lines and color would help, also for attribution the sensors. Refer to the corresponding table for the letters in the legend or include this information in the caption or directly in the legend. Why are E and F not shown?

Negative signs only mean that the sensor that arbitrarily is going to be installed in the measurement point *a* measures lower values than the identical one which is going to be installed in the measurement point *b*. Redrawing the plot just to avoid negative values seems unnecessary to us: an explanation of the signs has been added to the caption.

Sensors are sampled at the same frequency (10 min), which is the frequency later used during the field experiment. This is better explained in the revised text.

All plots have been redrawn in colors, trying to maintain readability when printed in greyscales and for color blind persons.

E and F are not shown because they were added later in the batch, so for them we relied on manufacturer's characterization. This information has been added in the revised manuscript.

12. Figure 4: I suggest replacing the relief with a detailed local topographic map; add an inset showing a larger area for context. The blue dot is too small if this figure gets rescaled. Add a North flash and a scale. I would combine Figure 4 and 5 in one figure of two panels; maybe even removing Fig.4 and only giving the site coordinates.

Figure 4 (now 4a) has been replaced with a topographic map, taken from Piedmont Geoportal (credits are added), with an inset showing its position relative to Turin. Figure 5 has been retained as figure 4b.

13. Looking at Figure 5, I am surprised by the site selection; this is very heterogeneous terrain with buildings, trees, roads, streams, and complex topography nearby, all being elements that should be absent or far away according to the cited list of criteria. How do you justify the choice of this site given these constraints? The discussion section gives some explanation and justification for this choice. Nevertheless, my understanding of flat, open, homogeneous terrain is very different, and I would have selected a location with homogeneous fetch of at least a few 100m in all directions. I even fear that in wintertime sensors could experience shading from the tall trees South of the sites when the sun is low. Did the recent studies of Coppa et al. (2021) and Garcia Izquierdo et al. also consider potential effects of albedo in the hemispheric field of view of the sensors, influence of obstacles on local turbulence and sensible heat fluxes or terrain effects by topography and vegetation on longwave radiation?

We would have liked to have a flat, unobstructed, homogeneous 100-m field as well, but that's not what you find in the Italian Alps, and I suspect, in most other mountain settings as well. We had to work with what was available. The site was chosen because easily reachable, complete with electric power and other instruments, constantly managed and guarded. We judged that the influence of the cited obstacles was symmetrical on the two sites, and given that all the measurements we considered were relative, we stipulated they cancel out in the analysis.

The shading possibility by the trees was considered, but ruled out once in the field.

The cited studies did not investigate the effects mentioned by the reviewer, because it was not their scope: in these studies, vegetation was uniform (thus the albedo), topography was more forgiving. Local turbulence was investigated and found to be negligible with the sampling frequency used in both studies (10 min).

14. Figure 6: Where are the radiometers installed? I don't see them in the picture, and I would have expected them to point South, i.e., towards the observer to avoid shading from the other instruments.

The reviewer is correct in pointing this out. As a matter of fact, radiometers were not yet installed in that picture: they were installed later, together with sensors E and F which, as said, were added to the experiment at a later stage.

Unfortunately, we don't have detailed pictures that show all the three measurement points in summer configuration, with the radiometers. They can be spotted, with a little difficulty, in

Figures 7 (now 6(a)) and especially 8 (now 6(b)), as the horizontal lower bar extending to the right (South).

A sentence is added to the caption in Figure 6 (now Figure 5) clarifying this out.

15. I suggest combining Figures 7 and 8 in a single two-panel figure.

The pictures have been combined as Figure 6(a) and 6(b).

16. Section 3.2 lacks information on the experiment duration, time periods measured, sampling frequency.

All this information is present in section 4, where it seems more fit.

17. Section 2.2. raises expectations on how snow thickness and snow conditions have been measured with reference to Section 3.2, which then only states that any fresh snow has been removed, and the snow at site (a) has not been characterized. If no snow has been removed at site (a), the snow surface characteristics change over time influencing the snow albedo. How has the snow depth been measured, manually or with an automatic acoustic range finder? Not clear in the manuscript.

Snow depth has been measured as always lower than ~40 cm. While it is true that snow characteristics change over time, our measurement periods after snow removal lasted only few days, to prevent snow degradation from becoming important.

It is true that snow at site *a* was never touched, but since the analysis was always conducted after a snowfall, we assume that conditions of the snow were always with close-to-maximum albedo.

All this has been explained better in the revised version of the manuscript.

18. P08L232 (cf. also #12): Apart from topographic shading, shadows should actually be excluded based on the site selection criteria. And if an instrument is shaded by another sensor this should be identical for both stations. Otherwise, there are undesired perturbations from nearby objects or topography which should be removed. Also, vegetation, water level in rivers should not be a factor for a homogeneous site (see P08L240-241). It is not clear to me, how you account for these influences, e.g., the water level in the streams on both sides of the measurement field, and I have a hard time understanding the values presented in Table 4. I think this needs further explanation.

We agree with the reviewer that undesired perturbations from nearby objects or topography should not be a factor for homogeneous sites. In real conditions things always need compromises, to minimize such factors from one side and have logistical opportunities (access, power, maintenance) on the other. This and further explanations have been added to clarify the differences and uncertainty values in table 4.

19. Figure 9: Include the sensor type and shield in the figure legend. Y-axis label: put the units in [brackets] or (parenthesis), otherwise it could be misread as “delta_t per degree C”; this applies to most figures in the manuscript.

As for the request before, we do not think it is wise to clog plots with information that can be obtained elsewhere in the manuscript.

As for the units, some journals apply this convention in the plots. There is no prescription in the submission guidelines of this Journal on the format of plot label units. However, plots have been redrawn with label format updated, following reviewer's request.

20. P10L292: I understand the motivation for choosing a threshold value for the difference in reflected SW radiation between the two sites for identifying the largest albedo effect. However, it would be interesting to have a quantification of this effect over the continuum of differences in reflected radiation. Could this still be considered? Otherwise, include a reference to Figure 10 where this choice is graphically justified.

An attempt to include data also below this threshold cut limit was conducted, resulting in a large amount of data resulted in terms of temperature differences below 0.1 °C, thus extending the 0 °C – 0.2 °C range (first bar of graph in Fig. 10 – now Fig. 9). The resulting plot would decrease its graphical information in the higher and more important difference values, which result "compressed" thus less detailed. Moreover, below such threshold it was impossible to discriminate among the different kind of sensors and shields. This has been added to text, and more is added to the figure caption.

21. Figure 10: Which data is shown in this figure? What is the timestep on the x-axis? Are you showing instantaneous values of the 10min resolution data? If not, what exactly is shown here? If these are selected values excluding certain measurements not meeting the selection criteria, there should be data gaps (values cannot be equally spaced in time as the figure suggests). This information should be included in the caption. The graph looks like a bar plot; but why are there some diagonal lines in the graph? Panels (a) and (b) could be lumped using color. Why are there negative values in panel (c)? The reflected radiation over snow should always be larger than that over bare ground?! How do you explain the negative differences sometimes exceeding 100 W/m2?

Data shown in the plots are the whole 10-min dataset, not filtered. This info has been better explained in the caption.

The plots are not actually bar plots but line plots: this explains the diagonal lines that, depending on the pixel resolution of the image, sometimes produce this "moiré" effect. However, the plots have been re-drawn to reduce this effect.

Regarding color-coding essential information in the plots, we would like to reduce that to a minimum: as a color blind person, I have a lot of difficulties. Besides, putting together a) and b) plots has already been tried, with little success in terms of readability.

Negative values in panel c) are very few with regard to the whole dataset (~600 out of ~ 28000), and most of them are very small in magnitude: these are due to errors in the measurement of radiation being larger than the values themselves, just like shown in figure 8. Larger negative values are mostly confined to days around 14 November, even before the first snowfall event. We cannot be sure about what happened, but it was not related to snow. This explanation has been added to the caption as well.

22. Section 4.2 does not explicitly discuss the difference of passive vs. actively ventilated radiation shields which is a crucial point that definitely should be included in the discussion. This section also gives the misleading impression that the differences depend on the sensor type while the main cause is rather the sensor shield absorbing (incident and reflected) shortwave radiation. This point should be clarified.

By “sensor type” we always mean “combination of sensors and shield”, as the “types” in Table 1 clearly state. This has been added in the text to make it clearer.

As we only had 1 type of actively aspirated shields, it did not seem fair to draw general conclusions. Besides, it has never been the scope of the manuscript: this work was performed in order to test the method proposed in Musacchio et al (2019).

Given that this issue is addressed again by the reviewer in comment 25, it seems fit to discuss it there.

23. Figure 12: Add the sensor type and radiation shield information to the panel letter to facilitate reference.

We don’t want to repeat an information that’s already present in the manuscript, with the risk of clogging the plot. Please, refer to Table 1 for info.

24. Figure 13: Use color for better readability. I would also try to identify clusters to characterize different behavior of the available sensor/shield combinations.

As for the previous comment on colors, we would like to keep them in plots to a minimum.

However, in order to improve the readability of the plot, it has been divided into 6 panels, one for each sensor.

25. Figure 14 and Section 4.3: I am surprised that any possible difference between ventilated and passive shields is not at all discussed (apart from one sentence in Section 5). Figure 14 suggests that an aspirated sensor is not necessarily performing better than a non-ventilated, cf. A vs. D, respectively. From Figures 13 and 14 it can be concluded that the magnitude of the albedo effect is a combination of the influence of reflected shortwave radiation and wind speed; radiation alone does not explain the temperature differences (see Figure 14) and wind speed seems to dominate radiation effects. Sections 4.2 and 4.3 would largely benefit from such kind of a discussion. Additionally, a discussion on how the albedo-induced error in air temperature measurements compares to the magnitude of errors due to incident solar radiation heating of the shield and sensor may be another interesting point. Perhaps even identifying a ratio/relation between the two. And finally, what is the advantage of a helical shield in comparison to a standard multi-disk shield?

As mentioned in the answer to comment 22, having only 1 type of actively ventilated shields prevents us to draw general conclusions on the issue actively vs. passively ventilated shields.

We agree with the reviewer’s remark about active shields not necessarily being better than passive ones. However, we must keep in mind that, by comparing A and D systems, we are not only talking about shields but also about sensors. As a matter of fact, sensor A is a Pt100 while sensor D is a thermo-hygrometer: a conclusion about performance of the shield alone does not seem definitive.

In other similar experiments on obstacle effects on near-surface temperature measurements, it often emerges the fact that wind dominates radiation in terms of effect on temperature, for instance Coppa et al (2021). The uncertainties in temperature measurements due to incident solar radiation on shields have been investigated in several works, e.g. Erell et al (2005), Lopardo et al (2014), and a discussion on this has been added to the text.

Helical shields seem to maximize the air intake effectively cooling down the sensor inside better, with respect to other kinds of shields, but this is something to be investigated perhaps with a theoretical study. This has been added to the revised manuscript as well.

26. Section 4.4: Certainly an interesting point, but out of the scope of this study (as you already mention) since it is not related to albedo and solar radiation but rather to longwave radiation. You risk opening a can of worms here and I would remove this section completely (as well as the related lines 359-361).

Paragraph and related sentences are removed.

27. Have the instruments been monitored with an automatic camera to see occurrences of precipitation deposition on the sensors (especially the radiometers) or effects of riming? This way, spurious data can be flagged and removed.

No automatic camera was set up, because the costs would not be proportionate to the benefit. As a matter of fact, only periods after a snowfall (and subsequent snow clearing by us) where used for the analysis: in these cases, snow was always cleared from the sensors by hand.

28. A major point that really surprised me is that the albedo has not been plotted, analyzed, and exploited. Given that incoming and reflected shortwave measurements are available at both sites, I would have expected a plot and a detailed analysis of the albedo for the entire period. This is actually what the paper title suggests. Also, it is a pity that surface temperature has not been measured at the two sites which I consider would have been very useful for the analysis. IR surface temperature is an interesting indicator linked to longwave emission and possible sensible heat flux contributions. I strongly suggest including these elements, at least the albedo if IRT measurements are not available.

A full and detailed exploration of the relationship between the albedo and the error induced on the temperature sensors is out of the scope of the work; however, we agree with the reviewer that a plot and an analysis of the albedo can be of interest for the reader. Figure 8 has been added to the manuscript, showing the evolution of albedo measured in both sites.

As for the surface temperature, it is certainly something that we'll keep in mind should we repeat the experiment in the future.

29. The "recommendation" at the end of the manuscript suggests the generation of a correction for the albedo influence on near-surface air temperature measurements, but the paper does not propose such a correction for the study performed. Such a correction, maybe even shield-specific, would be of interest to the measurement community.

The reason why such correction was not calculated is already stated in the conclusion: it would need much more data and time, in order to cover all possible meteorological and geometrical combinations. Within the scope of our experiment, uncertainties would have been simply too high, with the risk of doing more harm than good.

30. This study quantifies the albedo effect on air temperature measurements which is the objective of the paper. This effect has been quantified to be as large as 3.8°C. For this value and as an example, the given conditions could be added, e.g., the corresponding incident and reflected SW radiation as well as the wind speed. It should be mentioned that these albedo-induced errors do not include radiative errors due to heating of the sensor shield from incident solar radiation; that error has to be added to determine a complete shortwave radiation-induced error on air temperature measurements. Since no

reference air temperature independent from radiation errors is available in this study, the total error due to heating of the sensor by solar radiation is not quantified which should be mentioned explicitly (one again in the conclusions).

As a matter of fact, some of the information on the conditions that generated the largest albedo effect is already available in Figure 12. However, it has been put in more evidence in the revised text.

The valid considerations raised by the reviewer have been added as well.

31. Almost throughout the manuscript, why do you call your results, analysis, tests etc. “preliminary”? e.g., L 134, 149, 154, 234, 273, 275, 304.

We are trying to repeat the experiment with other instruments and, hopefully, a longer time baseline which will give us more robust results. Since this is not mentioned in the manuscript, however, we remove all instances of the word “preliminary”.

All the following minor comments have been agreed to and addressed in the text, except when explicitly stated, with due explanation.

P01L019: Rather say “reflected radiation” instead of “albedo radiation”.

P02L052: It is global radiation, not only direct. Alternatively, say “incident solar radiation”. Also on L053.

P04L096: Remove “backward”.

P04L103: Instead of “some meters of radius”, better give the exact number.

We cannot be precise at this stage, since we are talking in general, not about our site. If someone wants to repeat the experiment and has a lot of space to use, it is useless for us to say “5 m” for instance. We replaced the vague statement with “at least”.

P04L104: Replace “unique” by “single”.

P05L135: Provide a few details about the project and the initiator/project owner.

They were already present in detail on P03L76, where they seem more fit, so it doesn’t look advisable to repeat them here.

P05L148: I suggest “Laboratory tests”.

P05L155: Replace “systematic values” with “systematic bias”.

P06L159-160: What is meant with “sensor dynamics”, the sensor response time?

Well, partly. Dynamics includes response time as one particular case, related to the step change of the measurand. Dynamics involves also particular responses like hysteresis, or the response of the sensor to the change of other influence quantities.

Tables 3 and 4: The table caption should include the term u_delta-t_instr

P06L161: The term u_delta-t_instr should be defined in the text, e.g., before the reference to Table 3.

P07L205: I would probably not call this site "heavily instrumented".

P07L206: Give a reference for the installation protocol.

The protocol is described in the parent paper Musacchio et al 2019. Added.

P07L215: To which model are you referring here?

The model described in Musacchio et al 2019. Added for clarity.

P08L237: Write "incident solar radiation" (to distinguish from incoming longwave radiation) which is > 0.

P09L261 and P11L318: "expanded uncertainty" should be defined and explained.

P09L267: Write "incident and reflected solar radiation"

P09L269: Rather say "(highest albedo)"

P09L270: "Snow was removed on…"

P09L271 and 272: Soil or short grass? This is important for albedo classification.

Soil. Salt was used to prevent water from freezing and help us during the snow removal; as a side effect, it also removed vegetation. Some dark green patches of grass remained, though, as visible in Figure 8.

P10L283 and throughout the paper: cross-references to Figures, Tables, Sections, Equations are capitalized.3

Caption Fig.10: "…is indicated by the red horizontal line."

Caption Fig.11: "…, for records exceeding the selected threshold for reflected shortwave radiation of 200 W/m2."

P12L356: Insufficient information regarding "k=2".

P15L442: "…exposed to the additional reflected solar radiation…"

Due to other reviewer's comment, that sentence has been removed altogether.

Throughout the manuscript: perhaps use "pairs of instruments" instead of "couples of instruments"

Author contributions of Beges, Hofstätter-Mohler, Nigrelli and Sanna are missing.

Their contribution is already recognized in the text: where their contribution was important, “with contribution by all co-authors” was already added.

Citation: https://doi.org/10.5194/amt-2021-63-RC3