



## Air temperature measurements using autonomous self-recording dataloggers in mountainous and snow covered areas

F. Navarro-Serrano <sup>a, b, \*</sup>, J.I. López-Moreno <sup>a</sup>, C. Azorin-Molina <sup>c</sup>, S. Buisán <sup>d</sup>, F. Domínguez-Castro <sup>a</sup>, A. Sanmiguel-Valladolid <sup>a</sup>, E. Alonso-González <sup>a</sup>, M. Khorchani <sup>a</sup>

<sup>a</sup> Department of Geoenvironmental Processes and Global Change, Pyrenean Institute of Ecology, CSIC, Campus de Aula Dei, Avenida de Montañana, P.O. Box 202, 50080 Zaragoza, Spain

<sup>b</sup> Department of Geography, University of Zaragoza, San Juan Bosco, 7, 50009 Zaragoza, Spain

<sup>c</sup> Regional Climate Group, Department of Earth Sciences, University of Gothenburg, Box 460, 405 30 Gothenburg, Sweden

<sup>d</sup> Regional AEMET (Spanish Meteorological Agency) Office in Aragón, Paseo del Canal, 17, 50009 Zaragoza, Spain

### ARTICLE INFO

#### Keywords:

Air temperature  
Temperature logger  
Radiation shield  
Snow  
Complex terrain  
SPICE (Solid Precipitation Intercomparison Experiment)

### ABSTRACT

High mountain areas are poorly represented by official weather observatories. It implies that new instruments must be evaluated over snow-covered and strongly insulated environments (i.e. mid-latitude mountain areas). We analyzed uncertainty sources over snow covered areas including: 1) temperature logger accuracy and bias of two widely used temperature sensors (Tinytag and iButton); 2) radiation shield performance under various radiation, snow, and wind conditions; 3) appropriate measurement height over snow covered ground; and 4) differences in air temperature measured among nearby devices over a horizontal band.

The major results showed the following. 1) Tinytag performance device (mean absolute error:  $MAE \approx 0.1-0.2^\circ\text{C}$  in relation to the reference thermistor) was superior to the iButton ( $MAE \approx 0.7^\circ\text{C}$ ), which was subject to operating errors. 2) Multi-plate radiation shield showed the best performance under all conditions ( $> 90\%$  samples has bias between  $\pm 0.5^\circ\text{C}$ ). The tube shield required wind ( $> 2.5\text{ m s}^{-1}$ ) for adequate performance, while the funnel shield required limited radiation ( $< 400\text{ W m}^{-2}$ ). Snow cover causes certain overheating. 3) Air temperatures were found to stabilize at 75–100 cm above the snow surface. Air temperature profile was more constant at night, showing a considerable cooling on near surface at midday. 4) Horizontal air temperature differences were larger at midday ( $0.5^\circ\text{C}$ ). These findings indicate that to minimize errors air temperature measurements over snow surfaces should be carried out using multi-plate radiation shields with high-end thermistors such as Tinytags, and be made at a minimum height above the snow covered ground.

### 1. Introduction

Air temperature affects many environmental and technical processes, and has been measured since the 17th century (Camuffo and Bertolin, 2012). Air temperature data are necessary for a variety of purposes, including to analyze crop yields, hydrological regimes, energy efficiency, biological processes (Hubbart, 2011) or to evaluate the impact of global change on economy of mountain areas (López-Moreno et al., 2018). Air temperature varies temporally and spatially, so well distributed measurement networks are needed. However, high mountain areas are usually poorly represented in official networks because of the complex topography, snow cover, and the low population density, all of which affect the installation and maintenance of stations

(Alonso-González et al., 2018; Barry, 1992). Moreover, the periodic presence of snow cover modifies the spatial distribution of air temperature (Navarro-Serrano et al., 2018; Rolland, 2003), and therefore the typical location of measurement stations, in valley bottoms, is inadequate. For these reasons the guidelines provided by the World Meteorological Organization (WMO, 2014) concerning the location, devices, shielding, and height above the ground for temperature measurement are very difficult to meet in mountainous areas. Thus, the absence of official measurement stations necessitates the use of portable, autonomous, easily installed, and self-recording temperature measurement devices (Nakamura and Mahrt, 2005), and their use for scientific and applied purposes has increased markedly in recent years (Bonnardot et al., 2012; Lundquist and Lott, 2008). However, a number of uncertainties need to be adequately quantified prior to analyzing data collected using such devices.

\* Corresponding author at: Pyrenean Institute of Ecology, Campus de Aula Dei, Avenida de Montañana, P.O. Box 202, Zaragoza 50080, Spain.  
Email address: [fnavarro@ipe.csic.es](mailto:fnavarro@ipe.csic.es) (F. Navarro-Serrano)

In terms of accuracy, temperature logger miniaturization is a positive development because of the associated reduction in shortwave radiative heating, as the small size favors convective flux (Fritschen and Gay, 1979; Richardson et al., 1999). Nevertheless, problems including bias and device failure and data loss for no apparent reason should be addressed by the comparison with reference measurements (Roznik and Alford, 2012; Wolaver and Sharp, 2007). Calibration process is common for scientific purposes (Shedekar et al., 2016). Several studies have analyzed and calibrated miniature temperature loggers in ice baths or other laboratory-controlled conditions (Hubbart et al., 2005; Imholt et al., 2013; Johnson et al., 2005; Yang et al., 2012), but rarely has this been done under field conditions (Lundquist and Huggett, 2008; Pagès et al., 2017).

Regardless of the device accuracy and quality, to obtain accurate measurements it is necessary to protect the device from shortwave and longwave radiation fluxes (Harrison, 2010; WMO, 2014), especially in environments having strong insolation, high albedo rates, and weak winds (Huwald et al., 2009; Nakamura and Mahrt, 2005). Direct exposure of devices to sunlight causes an increase in the measured air temperature (Nordli et al., 1997). The opposite occurs at night because of the effect of outgoing longwave radiation (Azorin-Molina and Azorin-Molina, 2008). Radiative heating problems increase in snow covered environments because of reflection of shortwave radiation from the snow (Arck and Scherer, 2001; Huwald et al., 2009), especially in areas having temporary snow cover, where nonsystematic heating occurs (Georges and Kaser, 2002; Lundquist and Huggett, 2008). The design of radiation shields, and their construction materials and geometry are not trivial factors, as they determine the protection from radiation and internal air flow efficiency. To ensure correct internal airflow, forced fan-aspirated radiation shields claim to provide continuous and regular airflow, and are preferred (Georges and Kaser, 2002; Nakamura and Mahrt, 2005). However, the energy requirements of the fan-aspirated system prevent its use in remote mountain areas. Thus, since the 19th century the Stevenson wooden double-louvre naturally-ventilated screen has generally been used by meteorological agencies (Brunet et al., 2006). Device miniaturization has occurred concurrently with the development of radiation shields and their design. The most commercialized small shield is the plastic multi-plate radiation shield, based on the design of Gill (1983), and has been generally used in official automatic weather station networks since the mid-1980s. A major disadvantage is that miniaturization of radiation shields has not resulted in a reduction in price, and generally the shields are more expensive than the miniature air temperature loggers. This is particularly serious for dense air temperature monitoring networks, because it drastically increases costs. Consequently, there has been a concerted search for non-commercial and cheaper radiation shields that achieve equilibrium between ventilation and radiation protection. These alternatives have usually been homemade, and generally have not been designed for snow covered and high-albedo environments, and as a result have produced measurements having substantial uncertainties (Hubbard et al., 2001). Although there have been studies of the performance of alternative radiation shields above grass ground or bare soil (Holden et al., 2013; Kurzeja, 2010; Richardson et al., 1999; Tarara and Hoheisel, 2007), few have been carried out above simulated high-albedo (Georges and Kaser, 2002; Hubbard et al., 2001) and snow covered surfaces (Huwald et al., 2009; Lundquist and Huggett, 2008). To avoid uncertainties associated with inappropriate uses of radiation shields, it is essential that shelter performance be analyzed.

Another uncertainty source that has rarely been taken into account in the miniature temperature logger network setting is the height of the device above the surface. The vertical temperature in the near-surface boundary layer is dependent on turbulent energy fluxes between the ground and the atmosphere (Anderson and Neff, 2008; Helgason and Pomeroy, 2012; Sokol et al., 2017), which cause significant variations

over distances of a few meters or centimeters, and even greater variations in snow covered environments (Halberstam and Schiedge, 1981). Temperature extremes occur in the bottom air layer, near the exchange surface (Campbell and Norman, 1998), but the amplitude decreases as the measurement height increases. It has been reported (WMO, 2014) that the optimal air temperature measurement height is between 125 and 200 cm above bare ground or natural vegetation. However, no guideline has been reported for snow covered environments, except that it should be constant among observatories, and studies of this issue are rare and limited (Anderson and Neff, 2008; Halberstam and Schiedge, 1981; Hanna et al., 2017). In addition, snow depth varies among seasons, introducing seasonal measurement uncertainty. Thus, it is important to analyze the height of the stabilization layer above snow covered ground. In this regard, Lundquist and Huggett (2008) installed devices at different heights, in function of the snow depth variability, from 2 to 10 m above the ground with the aim of to measure comparable air temperature. Thus, installation of a measurement device at an inappropriate height can cause inconsistencies among measurements, and horizontal and vertical air temperature differences can be mixed (Nakamura and Mahrt, 2005).

In addition to the uncertainty sources noted above, there are others that could affect the measurement reliability but are not so evident. The distribution of horizontal air temperature locally and regionally is mainly determined by elevation, orography, orientation, and land use (Lundquist and Cayan, 2007). However, other environmental factors can affect the air temperature, including sunshine, the presence of trees or buildings (Kurzeja, 2010), shrub height (Nakamura and Mahrt, 2005), the presence of water, and plant density. Consequently, it is extremely difficult to account for all the factors influencing temperature in designing a miniature temperature logger network, so device measurement differences in part result from interferences, and not just because of real horizontal differences in temperature. All these issues mean that chosen measurement points are rarely fully representative of the entire experimental plot. Thus, it is necessary to include a certain measure of uncertainty for mountainous areas that, in principle, we consider homogeneous as, in this case, an evergreen mountain forest.

The aim of this study was to analyze uncertainty sources that could affect air temperature measurements in mountainous and snow covered areas. Users of naturally-ventilated shielded miniature temperature loggers in snow covered and forested areas need to be aware of the magnitude and the effects of such uncertainties on air temperature measurements, particularly associated with inappropriate design of observational networks, and choice of measuring device and radiation shields. It is necessary to obtain precise and accurate measurements of air temperature, because these are indispensable in environmental applications. This is particularly the case in the context of global change, which seems to affect high mountain areas more rapidly (Morán-Tejeda et al., 2016). We undertook four experiments to assess uncertainty sources, using autonomous and commonly used temperature loggers and radiation shields:

- Experiment 1 assessed temperature logger accuracy in terms of consistency among several sensors, and bias compared with a reference air temperature device.
- Experiment 2 assessed the performance of different types of radiation shield over snow covered ground.
- Experiment 3 assessed the influence of measurement height above a snow covered surface.
- Experiment 4 assessed variation in air temperature measured by nearby devices in a similar elevation band.

The results of this study will improve the quality of temperature data for disciplines that depend on accurate information on snow covered and mountainous environments.

## 2. Study areas and material

### 2.1. Study areas

All experiments were carried out in Aragón (Spain). Experiment 1 was undertaken at Daroca and Canfranc-Station (Fig. 1). Daroca is a small town located on a high continental plateau (779 m a.s.l.; 41.11°N, 1.41°W). The climate in Daroca is cold and dry (average temperature: 6.5°C; precipitation: 65 mm from January to March), and snowfalls are occasional (12 snowfall days yr<sup>-1</sup>). Canfranc-Station is a valley-bottom town in the Spanish Pyrenees (1170 m a.s.l., 42.75°N, 0.52°W), near the French border. Its climate is cold and humid (average temperature: 3.7°C; precipitation: 417 mm from January to March), and snowfalls are common in winter (25 snowfall days yr<sup>-1</sup>), and snow cover is typical throughout winter and early spring. These locations facilitated the analysis under cold conditions of different climate impacts on measurement accuracy, and the influence of seasonal snowpack (more common at Canfranc-Station). Daroca and Canfranc-Station have meteorological stations managed by the Spanish Meteorological Agency (AEMET), and served as the reference temperature measurement for analyzing the accuracy and performance of miniature dataloggers.

Experiment 2 was carried out at Canfranc-Station (Fig. 1) because of the availability of a reference Stevenson screen, in addition to wind speed measures and the occurrence of contrasting continuous snow cover during winter and grass cover during spring. Canfranc-Station is relatively sheltered from the winds, but as a mountainous location, it can suffer strong winds in certain conditions.

Experiment 3 was carried out at Formigal-Sarrios (Fig. 1). Formigal-Sarrios is an experimental plot located in the high elevation Formigal Pyrenees ski resort (1820 m a.s.l.; 42.76°N, 0.39°W), near the French border. This plot is integrated within the SPICE Project (WMO Solid Precipitation Intercomparison Experiment). It is exposed to cold and humid air flows from the north, snowfalls are common and regular, and the snow depth can reach 2–3 m in optimal years (Buisán et al.,

2017). The climatic conditions result in continuous snow cover from late autumn to early summer. Formigal-Sarrios was chosen for this experiment because the main objective was to assess the vertical distribution of air temperature over snow covered areas. At this location the snow cover is continuous, and the snow depth enables assessment of temperature at various heights above the snow surface.

Experiment 4 was undertaken in Panticosa Valley (Fig. 1), which is a high Pyrenees valley (from 1600 to 3064 m a.s.l.; 42.76°N, 0.23°W) on the Gállego River. The valley is covered with a dense evergreen forest of *Pinus mugo* subsp. *uncinata*. Four experimental plots were established in various combinations of north-south slope and high-low elevation bands (Upper-North: 2195–2213 m a.s.l.; Bottom-North: 1742–1757 m a.s.l.; Upper-South: 2039–2053 m a.s.l.; Bottom-South: 1854–1882 m a.s.l.). Panticosa-Valley was chosen because its evergreen forest and the presence of snow during several months of the year. The potential to analyze the horizontal air temperature at four sites enabled comparison based on different exposures and elevational environments. These four sites had an approximate area of 1500–2000 square meters each.

### 2.2. Thermistors

Autonomous self-recording and price competitive temperature loggers are widely used in research in areas involving building construction (Fang et al., 2014), animal biology (Schofield et al., 2009), and vegetation (Brabyn et al., 2014; Measham et al., 2017), and also in climatic-meteorological applications (Dominguez-Villar et al., 2015; Imholt et al., 2013; Juliussen and Humlum, 2007; Lundquist and Lott, 2008; O'Neill and Christiansen, 2018). In the present study temperature measurements were made using three thermistor types: Thies PT100, Tinytag-Plus-2 (Tinytag), and Thermochron iButton (iButton).

The accuracy of the Thies PT100 thermistor is  $\pm 0.2^\circ\text{C}$  for high and low temperature daily extremes (Buisán et al., 2015), and its resolution is  $0.1^\circ\text{C}$  within the measurement range of  $-30^\circ$  to  $+70^\circ\text{C}$  (<https://www.thiesclima.com/pdf/en/Products/Temperature-Electrical-sensors/?art=965>; last accessed 20 March 2019). In Spain the Thies PT100 thermistors are managed by AEMET, and their use follows all WMO guidelines. Price range is \$600–700 (Thies Clima representative contact in Spain).

The Tinytag autonomous miniature and waterproof temperature logger model TGP-4017 (produced by Gemini DataLoggers UK Ltd.; Chichester, West Sussex, UK) is equipped with a 10K NTC thermistor. The manufacturer reports a precision error of  $< \pm 0.5^\circ\text{C}$  over a measurement range of  $-40^\circ\text{C}$  to  $+85^\circ\text{C}$  (<http://gemini2.assets.d3r.com/pdfs/original/3239-tgp-4017.pdf>; last accessed 20 March 2019) and a resolution of  $0.01^\circ\text{C}$ . It offers high storage capacity (up to 32,000 measurements) and is powered by a replaceable battery ( $\approx 3$  years of service). The Tinytag price range is \$150–200.

The iButton autonomous miniature temperature logger models DS-1922L and DS-1921G are produced by Dallas Semiconductors, Texas, USA (<https://datasheets.maximintegrated.com/en/ds/DS1922L-DS1922T.pdf>; last accessed 20 March 2019). The manufacturer reports an accuracy of  $< \pm 0.5^\circ\text{C}$  over a measurement range  $-10^\circ\text{C}$  to  $+65^\circ\text{C}$ , and a resolution of  $0.5^\circ\text{C}$ . The clock calendar is accurate to  $\pm 2$  min per month within a  $0$ – $40^\circ\text{C}$  range (Gasvoda et al., 2002). It is powered by a non-replaceable internal battery ( $\approx 10$  years of device service), and the price range is \$50–80 per device.

### 2.3. Radiation shields

In recent years, commercial naturally-ventilated plastic multi-plate radiation shields based on the design of Gill (1983) have been widely used (Hanna et al., 2017; Hofer et al., 2017; Makgose and Phillipus, 2015; Quéno and Bonnardot, 2014; Richardson et al., 1999). Because

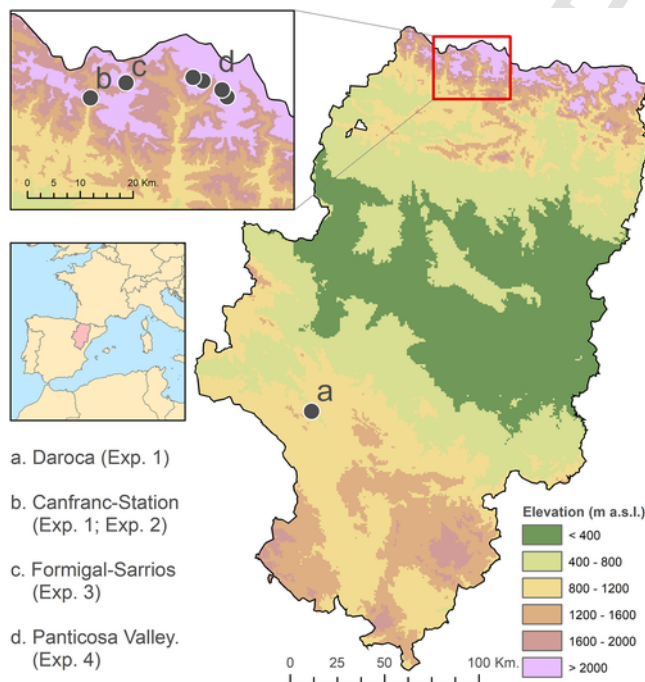


Fig. 1. Aragón (Spain) terrain map showing the location of the four experimental sites. Elevation is expressed in meters above sea level (m a.s.l.).

of the high price of commercial shields, non-commercial naturally-ventilated alternative radiation shields (e.g. cones, tubes, funnels, pagodas) have also been used in environmental research (Hubbart et al., 2007; Lundquist and Huggett, 2008; Pagès et al., 2017; Pepin and Kidd, 2006; Tarara and Hoheisel, 2007). In this study we used four naturally-ventilated radiation shields including the Stevenson and Datamate ACS-5050 shields, and a tube and a funnel shield.

The Stevenson double-louvered wooden screen (Fig. 2a) is the shield typically used by official meteorological agencies. It is made from wood that is painted white, to reflect direct solar radiation. The standard dimensions, colour, and materials for this screen have been published (WMO, 2014).

The Datamate ACS-5050 Weather Shield (Fig. 2b; <http://gemini2.assets.d3r.com/pdfs/original/1684-ac5-5050.pdf>; last accessed 20 March 2019) is produced by Gemini DataLoggers UK Ltd. (Chichester, West Sussex, UK). It is an 8-plate PVC radiation shield having 15 mm spaces between the plates, to enable air to circulate freely. Its dimensions (142 mm height  $\times$  198 mm diameter) provide space for two miniature dataloggers.

The tube shield (Fig. 2c) is a non-commercial radiation shield. It was built from white PVC tube (295 mm long  $\times$  100 mm diameter), following the design of Pepin and Kidd (2006), and was covered by reflective aluminum foil duct tape, based on the reports of Fuchs and Tanner (1965) and Holden et al. (2013). Based on the suggestion of Pepin et al. (2010), to avoid direct solar radiation it was oriented towards the north and sloped  $K$  degrees from the vertical ( $K = 90^\circ$  minus latitude, in degrees).

The funnel shield (Fig. 2d) is also a non-commercial bottom-opening radiation shield. Following the design of Hubbart (2011), it was built from two perforated and one non-perforated white PVC funnels (diameter: 130 mm) assembled vertically with a 10 mm gap between the funnels, providing a perforated area equivalent to 20% of the total.

#### 2.4. Environmental variables

Downwelling radiation was measured at Formigal-Sarrios using a Kipp & Zonen model CM11 pyranometer (<http://www.kippzonen.com/Product/13/CMP11-Pyranometer>, last accessed 20 March 2019). Wind speed was measured at a standard height of 10 m using a heated anemometer (Thies Clima; Göttingen, Germany) managed by AEMET at Canfranc-Station and Formigal-Sarrios. Snow surface presence information in Canfranc-Station and Daroca was derived from AEMET snow data, measured by manual operators by traditional poles. Snow depth was measured at Formigal-Sarrios using a SR50A acoustic snow depth measurement sensor ([https://www.eoas.ubc.ca/courses/atc303/Instruments/manuals/TBC\\_sonic\\_ranger.pdf](https://www.eoas.ubc.ca/courses/atc303/Instruments/manuals/TBC_sonic_ranger.pdf); last accessed 20 March 2019). Precipitation was measured at Formigal-Sarrios using an automatic weighing gauge (OTT Pluvio2 gauge; OTT Hydromet, Kempten,

Germany) with a 200 cm<sup>2</sup> orifice area and 1500 mm capacity. The gauge was used in reference configuration, and was shielded inside a DFIR (Double Fence Intercomparison Reference, single Alter, SA) fence (Buisán et al., 2017). For analysis, each day was separated into four periods from solar incoming radiation cycle, following the procedure of Tarara and Hoheisel (2007). The four periods were generated from downwelling global solar radiation hourly averages measured at the Formigal-Sarrios reference location during the entire 2017–2018 (November to May) snow season, and included: Period 1 “Night”: 20:01–04:00 UTC; Period 2 “Sunrise-Evening”: 04:01–07:00 and 17:01–20:00 UTC; Period 3 “Morning-Afternoon”: 07:01–10:00 and 14:01–17:00 UTC; Period 4 “Midday”: 10:01–14:00 UTC.

### 3. Experimental design and methods

#### 3.1. Experiment 1. Temperature logger accuracy, consistency among simultaneous measurements, and bias relative to a reference temperature

The Tinytag and iButton temperature loggers were compared with the Thies PT100 reference logger (calibrated annually at the AEMET calibration laboratory). The measurement periods were during winter, when snow can occur: 22 December 2016 to 13 February 2017 at Daroca, and 20 January 2017 to 1 March 2017 at Canfranc-Station. In total, 14 Tinytag (Daroca: 8; Canfranc-Station: 6) and 14 iButton (Daroca: 6; Canfranc-Station: 8) autonomous temperature loggers having 10-min sampling intervals were installed. The loggers were placed near the PT100 reference sensor, and shared a double-louvered Stevenson screen at a height of 150 cm.

Measurement differences ( $\Delta T$ , in  $^\circ\text{C}$ ) between the Tinytag and iButton temperature loggers and the Thies PT100 reference thermistor within the Stevenson reference radiation shield were calculated. Mean absolute error (MAE) from reference were calculated. Hourly and snow effects were analyzed and confidence intervals (0.95) were calculated, and the range between the coolest and hottest temperatures measured by the loggers were determined.

#### 3.2. Experiment 2. Performance of the tested radiation shields

It involved measurements made in the absence of a shield (no shield), and using the tube and funnel radiation shields and a commercial white multi-plate Datamate ACS-5050 shield. Comparisons were made with reference to measurements made using the Stevenson reference shield. The shields were mounted at 150 cm above the ground, and all measurements were logged using 5 autonomous Tinytag devices (10-min sampling interval) within each shielding situation. The paired data measurement period was from 11 January 2018 to 30 May 2018. Snow cover was present for 33% of days ( $n = 33,150$  samples). Solar radiation was measured at Formigal-Sarrios in a location comparable to

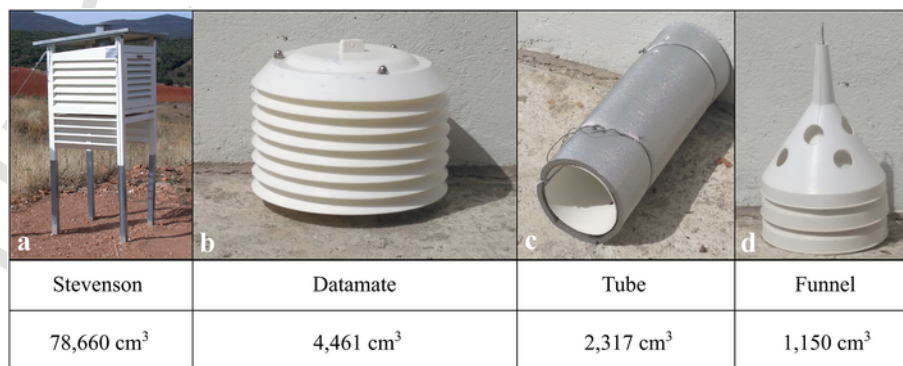


Fig. 2. The Stevenson, Datamate ACS-5050, tube and funnel radiation shields. Total volume is expressed in cubic centimeters (cm<sup>3</sup>).



that at Canfranc-Station in terms of orientation, exposure, and topography.

Differences (bias) in relation to measurement data derived using the Stevenson screen were expressed in degrees Celsius ( $\Delta T$ ). Wind speed was aggregated based on hourly averaged speeds. Downwelling global solar radiation ( $Wm^{-2}$ ), wind speed ( $ms^{-1}$ ) and snow depth (cm) effects on  $\Delta T$  were summarized and analyzed in relation to bias. The Wilcoxon-Mann-Whitney (Wilcoxon, 1945) test was used to analyze differences among shields and between snow- and grass-covered ground.

### 3.3. Experiment 3. Effect of sensor height above the snowpack on temperature measurements

This experiment was conducted at Formigal-Sarrios and involved comparison of hourly air temperature measurements made from 15 November 2017 to 30 May 2018 (during the snow season) at 6 heights above the land surface (100, 150, 200, 250, 300, and 350 cm) using Tinytags within naturally-aspirated Datamate ACS-5050 radiation shields mounted on north-facing wooden poles. Snow depth was measured simultaneously. Hourly mean wind speed was measured at a height of 10 m above bare ground. The mean hourly incident global solar radiation was measured 1 m away Tinytags. The 2017–2018 snow season at Formigal-Sarrios station lasted approximately 170 days ( $n = 24,510$ , 86.8% samples), from 30 November 2017 until 20 May 2018. The snow depth was  $> 100$  cm for 54.3% of the time,  $> 150$  cm for 32.4% of the time,  $> 200$  cm for 21.0% of the time, and  $> 250$  cm for 1.9% of the time; the greatest snow depth was 275 cm, on 11 April 2018 (18:00 UTC). No failures in measurement were detected. Measurements recorded from sensors completely covered by snow were not included in the analysis ( $n = 5134$ ; 18.2% samples).

Vertical air temperature profiles were analyzed and assessed in relation to environmental data (global solar radiation, wind speed, and precipitation). Vertical temperature stabilization was calculated from the lowest device height above the snow (not bare ground); the temperature difference relative to the highest device (always snow-free) was  $< \pm 0.1$  °C.

### 3.4. Experiment 4. Horizontal variability

The main objective of this experiment was to assess the horizontal variability of temperature over relatively homogeneous terrain (mountain forest), to determine the potential impact of random selection of the location used to characterize the temperature regimen for a given area, in this case within the 1742–2213 m elevation band. Air temperature was measured in four homogeneous plots (two north-facing and two south-facing) in terms of latitude, longitude, land use, slope, and elevation. In these cases, air temperature differences may be determined by non-identified factors, which have been wanted to quantify (e.g. specific-site sun hours, water presence, shrub height, proximity to sources of longwave transmission including rocks). Hourly paired data were logged from 5 September 2017 to 22 January 2018; at the end of the experiment the plots were covered by snow. Because of the very steep slopes it was difficult to find suitable measurement locations at equivalent elevations, but all were within a maximum difference of 28 m, imply a difference of  $0.13$  °C assuming a seasonal lapse rate of  $-0.47$  °C  $100m^{-1}$  (Navarro-Serrano et al., 2018). Measurements were made using Tinytags (5 in each plot, with a distance between loggers of around 15–20 m; 10 in each altitudinal band; 20 total) in Datamate ACS-5050 radiation shields. To avoid snow effects the loggers were installed at a height of 250–300 cm above the ground, and within an evergreen forested area.

Hourly absolute bias between the mean plot air temperature and specific devices was calculated and analyzed to determine the varia-

tions involved. ANOVA, Scheffe Post-Hoc test, and descriptive statistical analyses were made using the original absolute temperature data to determine variability, and the confidence intervals (0.95) in relation to the mean were calculated.

## 4. Results

### 4.1. Temperature logger accuracy and consistency

Eight (6) Tinytag and six (8) iButton devices were originally installed at the Canfranc-Station (Daroca). However, two of the iButtons (one each at Daroca and Canfranc-Station) failed and were discarded, and three additional iButtons at Canfranc-Station failed to save data during the last month of the experiment, reducing the expected number of measurements. The remaining data were analyzed.

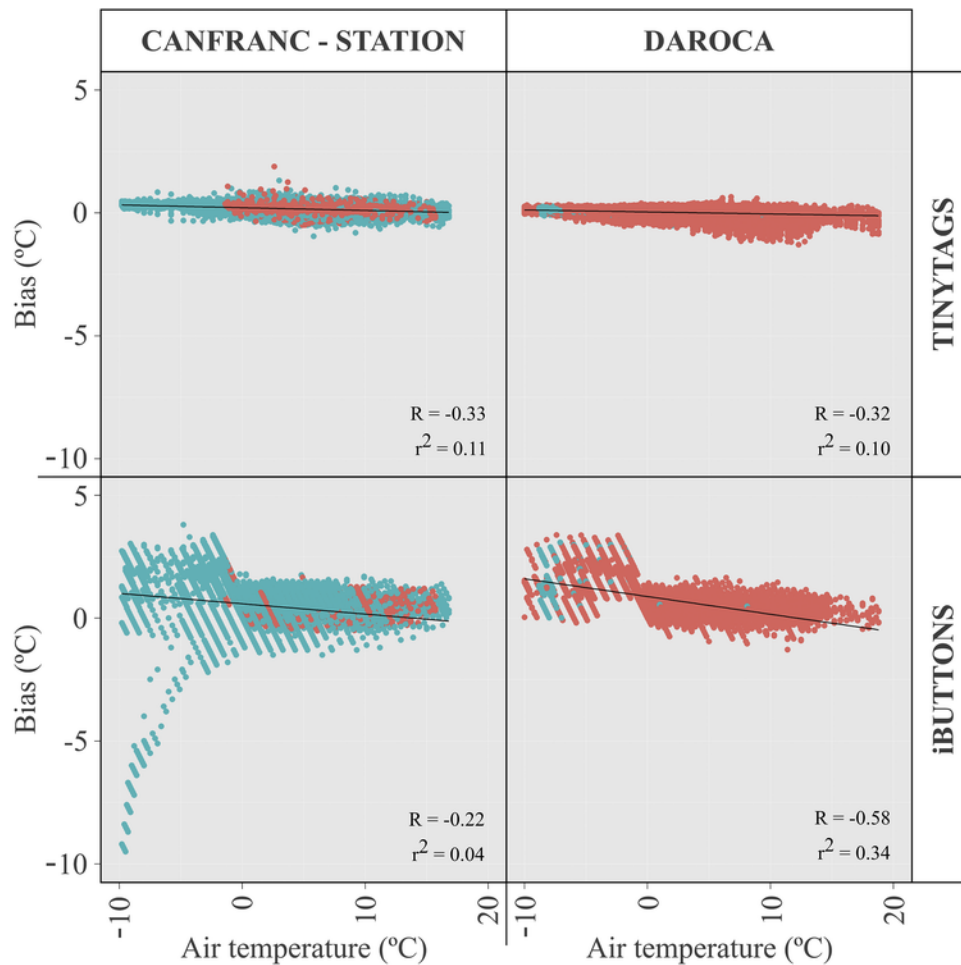
Based on analysis of MAE and confidence intervals, Table 1 shows that Tinytags had less bias than iButtons (especially at morning-afternoon and midday), and there was less bias at Daroca than at Canfranc-Station. Radiation effects were not equal between device types and locations. Thus, the MAE for Tinytags decreased and the confidence interval increased with increasing radiation, while for the iButtons there was only a decrease in the confidence interval at morning-afternoon and midday. Fig. 3 shows the variation between the Tinytag and iButton data and the reference Thies PT100 data. This indicates that bias for almost all Tinytag measurements was  $\pm 0.3$  to  $\pm 0.4$  °C, especially at Daroca, while for the iButtons  $> 20\%$  of measurements showed  $\pm 1$  °C deviation.

Fig. 3 presents that was no clear effect of the presence of snowpack on the observed bias. However, differences were found between the biases of the Tinytags and iButtons, as shown on Table 1. iButton records below  $0$  °C shows a noticeable scattering. There was a slight and negative Pearson correlation (non-significant) between the bias and the ob-

**Table 1**

MAE (°C) and confidence interval (0.95) for Tinytag and iButton bias relative to the Thies PT100 reference thermistor. Mean range (°C) between devices by hourly periods and location.

Hourly period (UTC h)	Tinytags	iButtons	Location
1. Night (21–04)	MAE: $0.223 \pm 0.009$ Range: 0.19 (n: 11520)	MAE: $0.703 \pm 0.106$ Range: 1.21 (n: 10244)	Canfranc-Station
2. Sunrise-Evening (05–07 & 18–20)	MAE: $0.226 \pm 0.009$ Range: 0.20 (n: 8640)	MAE: $0.687 \pm 0.100$ Range: 1.22 (n: 7668)	
3. Morning-Afternoon (08–10 & 15–17)	MAE: $0.196 \pm 0.011$ Range: 0.27 (n: 8640)	MAE: $0.669 \pm 0.090$ Range: 1.16 (n: 7668)	
4. Midday (11–14)	MAE: $0.146 \pm 0.011$ Range: 0.40 (n: 5760)	MAE: $0.631 \pm 0.056$ Range: 1.00 (n: 5112)	
1. Night (21–04)	MAE: $0.085 \pm 0.005$ Range: 0.18 (n: 20352)	MAE: $0.767 \pm 0.074$ Range: 0.77 (n: 17808)	Daroca
2. Sunrise-Evening (05–07 & 18–20)	MAE: $0.090 \pm 0.006$ Range: 0.21 (n: 15264)	MAE: $0.653 \pm 0.062$ Range: 0.75 (n: 13356)	
3. Morning-Afternoon (08–10 & 15–17)	MAE: $0.096 \pm 0.008$ Range: 0.25 (n: 15243)	MAE: $0.633 \pm 0.053$ Range: 0.71 (n: 13335)	
4. Midday (11–14)	MAE: $0.166 \pm 0.015$ Range: 0.29 (n: 10176)	MAE: $0.530 \pm 0.037$ Range: 0.64 (n: 8904)	



Location / Device		± 0.1°C	± 0.2°C	± 0.3°C	± 0.4°C	± 0.5°C	± 1°C
Canfranc Station	Tinytags	16.6 %	49.0 %	83.8 %	96.9 %	99.3 %	100 %
	iButtons	9.0 %	19.1 %	28.6 %	36.9 %	46.0 %	82.7 %
Daroca	Tinytags	59.7 %	89.4 %	96.4 %	98.2 %	99.1 %	100 %
	iButtons	9.2 %	18.9 %	30.9 %	42.9 %	54.2 %	79.2 %

**Fig. 3.** Tinytag and iButton bias relative to the Thies PT100 reference air temperature. Days with snow on the ground (turquoise) and snow-free days (red) are distinguished for the two locations. Global linear fit is expressed by black line. The table shows the % device bias relative to the reference Thies PT100 data. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

served temperature (for temperatures  $>0^{\circ}\text{C}$  the bias rarely exceeded  $+0.5^{\circ}\text{C}$  for the Tinytags and  $+2^{\circ}\text{C}$  for the iButtons). At Canfranc-Station there was an extreme and short negative bias when one iButton DS-1921G device measured  $-19^{\circ}\text{C}$  while the Thies reference thermistor measured  $-9.7^{\circ}\text{C}$ , although the device returned to normal functioning 4 h later (Fig. 4).

#### 4.2. Performance of radiation shields

Table 2 shows temperature bias in relation to a Stevenson-shielded Tinytag. All snow-grass and radiation shield combinations were found to differ significantly ( $p < .05$ ; Wilcoxon-Mann-Whitney test). Under all conditions the Datamate radiation shield showed the smallest  $\Delta T$ , with  $>90\%$  of measurements between  $\pm 0.5^{\circ}\text{C}$  (Percentile<sub>10</sub>-0.33 and

$P_{90}$  0.18), indistinctly over snow and grass. As expected, measurements made in the absence of a shield were the worst in all situations ( $P_{10}$ -1.37 and  $P_{90}$  1.64), with  $<40\%$  and  $<25\%$  of samples between  $\pm 0.5^{\circ}\text{C}$  over snow and grass, respectively; however, the percentage of samples between  $\pm 2^{\circ}\text{C}$  were similar between the tube and funnel shields. In all situations the funnel shield showed less uncertainty than the tube shield, reaching approximately 70% of samples between  $\pm 0.5^{\circ}\text{C}$  ( $P_{10}$ -0.65; and  $P_{90}$  0.82), while the tube shield ( $P_{10}$ -0.94 and  $P_{90}$  1.16) was 10% and 20% of samples below in the snow and grass situations, respectively. The slightly better whole shield performances over snow than grass indicates the need for specific radiation-wind analysis.

Fig. 5 shows the mean  $\Delta T$  for combinations of ground surface, radiation, and wind speed for each shielding condition. An increase in  $\Delta T$  was common under high solar radiation levels and with weak winds.

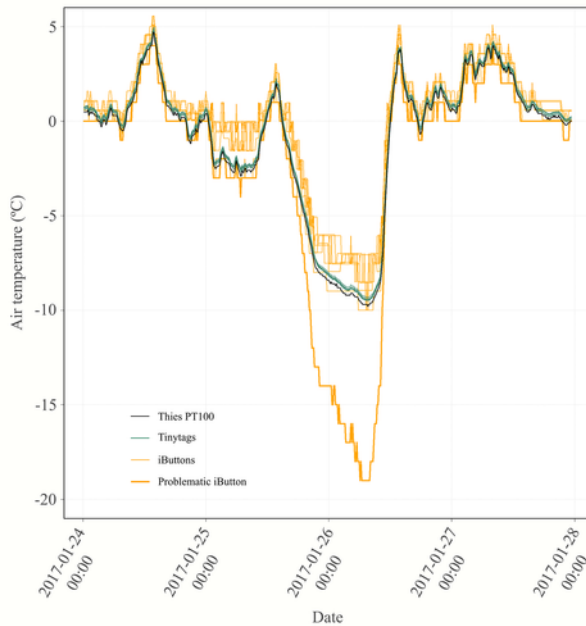


Fig. 4. Episode (between 24 January 2017 and 27 January 2017) of extreme decoupling of temperature measurement between one iButton device (thick orange) and the Thies PT100 reference thermistor (black), Tinytags (green), and other iButtons (orange). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Nevertheless, the magnitude of  $\Delta T$  varied among shielding conditions and the cover type/radiation/wind speed combinations. However, the Datamate  $\Delta T$  was almost equilibrate between all combinations (generally between  $\pm 0.5$  of the mean  $\Delta T$ ), although there was general slightly overcooling with grass cover conditions. In the absence of a shield (no shield) the  $\Delta T$  was extremely positively correlated with radiation (Pearson's correlation;  $r = 0.72$ ). The tube and funnel shields behaved differently. The tube shield performance was more related to wind speed, while for the funnel it was more related to radiation. Above snow cover (Fig. 5, right side) the environmental effects were variable. The mean no shield  $\Delta T$  suffers a relevant overcooling under null radiation conditions (night), and extreme overheating in strong radiation conditions, almost independent of wind speed. The tube shield performance was more related to wind speed than to radiation, and performed well under strong radiation conditions if the wind speed was  $> 2.5 \text{ m s}^{-1}$ . The funnel shield performance was more correlated with radiation, and performed well under zero to weak wind conditions if radiation was  $< 400 \text{ W m}^{-2}$ . The Datamate performance was constant below different conditions, and no significant differences were found between snow and grass ground cover situations (always between  $\pm 0.5 \Delta T$ ). Above grass cover (Fig. 5, left side), environmental effects were similar, but some differences were found. The performance of the funnel improved significantly, with a generalized decrease in  $\Delta T$  for

the majority of combinations, except under null radiation. The performance of the tube shield was slightly worse, especially under weak wind conditions. Frequency distribution of wind-radiation combinations is shown in Supplementary Material (Fig. 1).

#### 4.3. Effect of sensor height above the snowpack

Fig. 6 shows the most representative hour of each hourly period. It shows that there was a cooling effect on sensors located near the snow surface ( $< 100 \text{ cm}$  above the surface) that followed a near logarithmic function, especially at 12 and 16 UTC. This pattern weakened between late 19 and 2 UTC, when the vertical temperature distribution was more stable and homogeneous. Moreover, the variation decreased with time from midday to night. Maximum cooling generally occurred under calm or weak wind conditions (red colour in Fig. 6), indicating an influence of wind on the vertical temperature profile. Generally, temperature stabilized at  $\approx 75\text{--}100 \text{ cm}$  above the snow surface, with clear stabilization above this height, particularly between 12 and 16 UTC. A more constant air temperature profile was found under rainy/snowy conditions (Supplementary Material. Fig. 2), and greater differences occurred when the weather was stable. To analyze the height at which stabilization occurred (representing the temperature stabilization height above snow cover), we used the lowest device height above snow cover where a difference  $< \pm 0.1^\circ \text{C}$  in relation to the highest device was found (Fig. 7).

#### 4.4. Horizontal variability

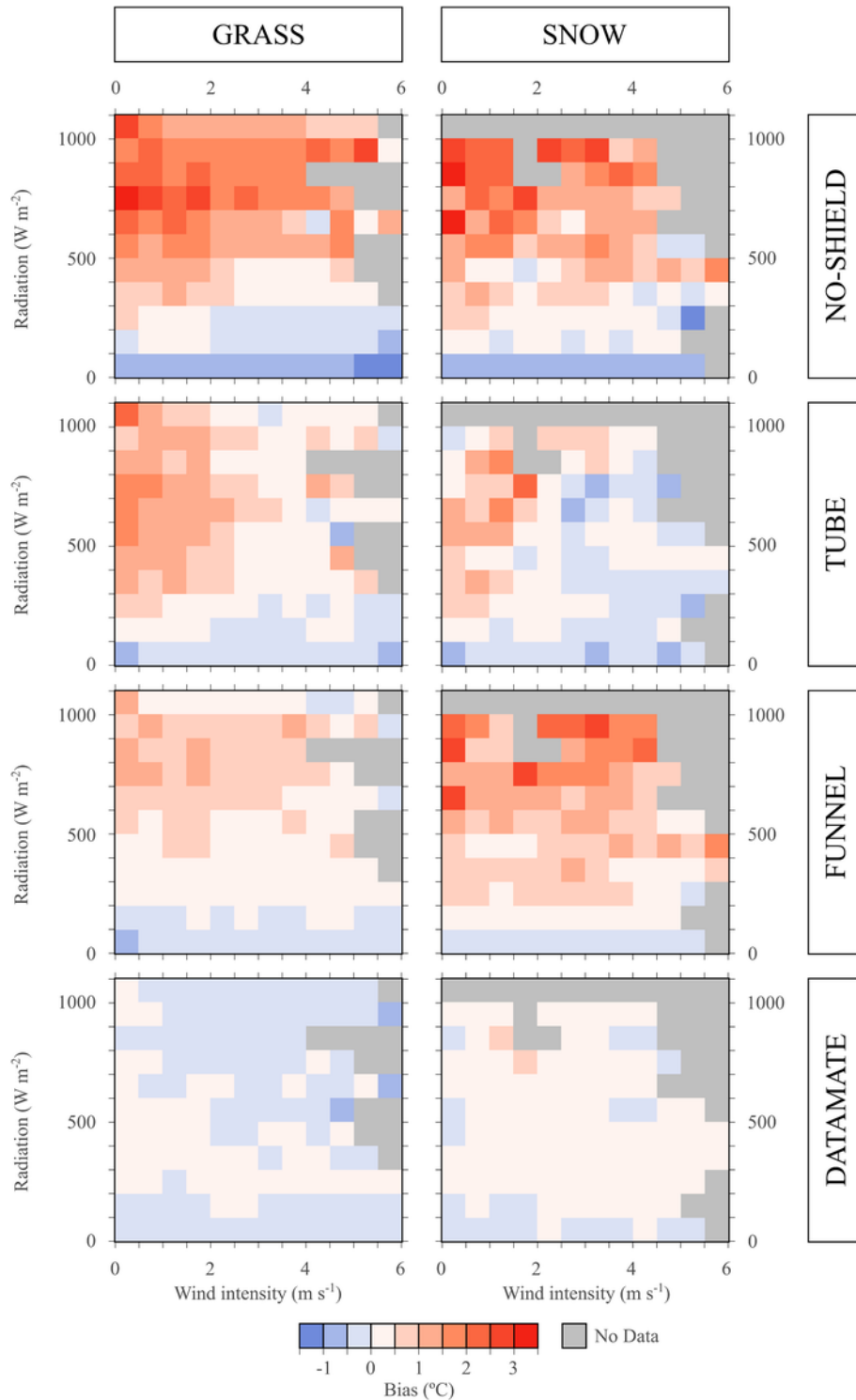
Fig. 8a shows an evident reduction in bias at night ( $\approx 0.2^\circ \text{C}$ ), while at midday the mean absolute bias increased to produce the largest differences (peaks of  $0.6\text{--}0.8^\circ \text{C}$ ). Differences during the night (22–04 UTC) were small and constant ( $< \pm 0.5^\circ \text{C}$  for 89.6% of the cases). Fig. 8b shows the mean hourly confidence interval (0.95) in each site, derived by averaging the confidence interval for specific devices. This figure shows an evident decrease in the confidence interval at night, and maximum rates between 8 and 17 UTC, particularly at midday. The general behavior approximated the daily global radiation curve. The Upper-South site showed less uncertainty than the other sites, especially at night. Similar behavior among the other sites was found, despite a slight increase in uncertainty for the Upper-North site during the night. The maximum mean hourly confidence interval was small ( $\pm 0.10^\circ \text{C}$ ). ANOVA and Scheffe Post-hoc test analysis of the original absolute temperature data enabled identification of those devices that behaved differently relative to the sensors in the same plot ( $p < .05$ ), including: Upper-North, 1 case; Bottom-North: 1 case; Upper-South: No cases; Bottom-South: 1 case. The ANOVA results were consistent with Fig. 8b, which shows both night-day phases.

### 5. Discussion

Experimental results (in particular Experiment 1) showed that Thermochron iButton devices were subject to read/load measurement prob-

Table 2  
Temperature bias (% of total observations) under different shielding conditions for snow and grass ground covers.

Condition		$\pm 0.1^\circ \text{C}$	$\pm 0.2^\circ \text{C}$	$\pm 0.3^\circ \text{C}$	$\pm 0.4^\circ \text{C}$	$\pm 0.5^\circ \text{C}$	$\pm 1^\circ \text{C}$	$\pm 2^\circ \text{C}$
Snow	No-shield	5.9	13.8	23.2	32.6	39.9	68.9	95.3
	Tube	8.4	21.3	36.3	47.6	58.7	87.0	98.4
	Funnel	17.9	37.0	52.3	61.3	68.5	87.4	97.9
	Datamate	50.0	74.3	85.7	91.3	94.5	99.5	99.8
Grass	No-shield	3.1	7.1	12.2	18.7	25.4	56.7	89.4
	Tube	5.7	12.7	23.0	35.3	46.2	77.8	94.1
	Funnel	11.8	26.6	42.0	55.5	66.9	90.9	98.0
	Datamate	39.8	64.0	79.5	87.0	91.6	98.7	99.9

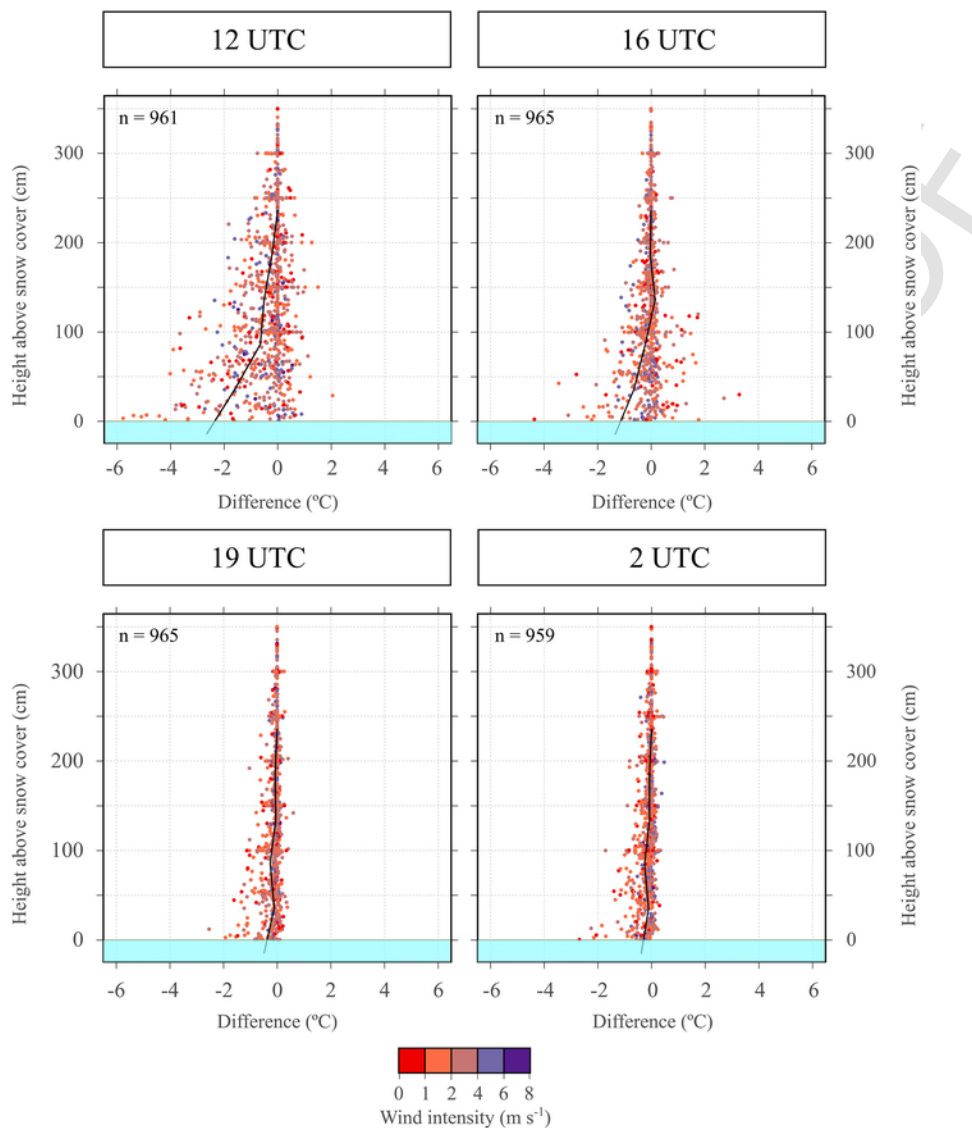


**Fig. 5.** Performance of radiation shields under grass (left) and snow (right) cover, and under differing wind speed ( $\text{m s}^{-1}$ ) and shortwave radiation ( $\text{W m}^{-2}$ ) conditions. 1) No shield, 2) Tube Shield, 3) Funnel Shield, 4) Datamate ACS-5050 shield. Bias is expressed in  $^{\circ}\text{C}$ , calculated in relation to measurements made using the Stevenson shield. Positive values imply a relative overheating effect, and negative values indicate a cooling effect.

lems that lead to total data loss. This issue has been reported in previous studies (Roznik and Alford, 2012; Wolaver and Sharp, 2007), and can be attributed to occasional operational problems resulting from poor water protection, despite iButton devices being submerged in water in some studies (Angilletta and Krochmal, 2003). In addition, we found some isolated measurements made using iButtons produced biases of up to  $10^{\circ}\text{C}$  compared with the reference thermistor. The Tiny-

tag devices were not subject to data loss, and provided reliable air temperature estimates compared with the reference over the temperature range. This is consistent with the results of Imholt et al. (2013), Yang et al. (2012), and the manufacturer's specifications. Substantial biases (bias  $> \pm 0.5^{\circ}\text{C}$ ; defined by Hubbard et al., 2005) in measurements using Tinytag sensors only affected 0.9% of measurements at Daroca and 0.7% at Canfranc-Station (compared with 45.8% and 54% using iBut-



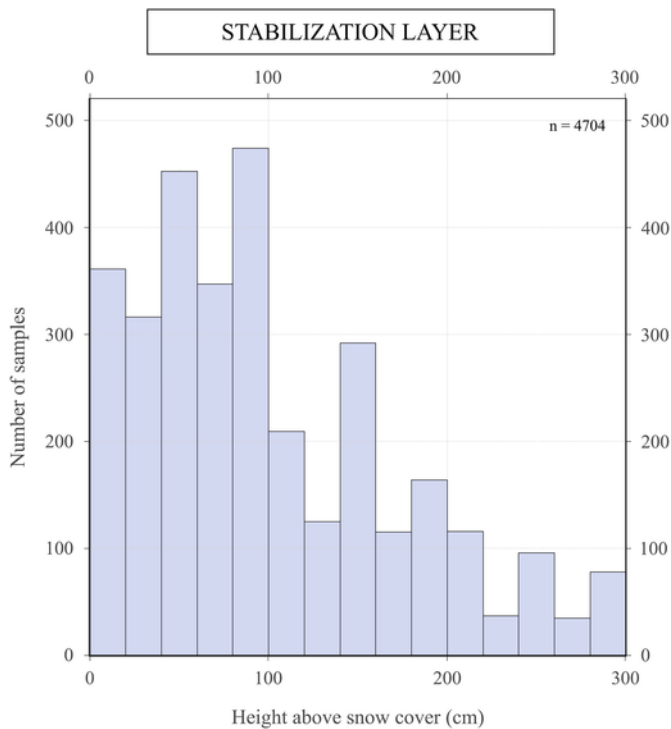


**Fig. 6.** Vertical temperature profile at 12, 16, 19, and 02 UTC. The black line shows the mean for the 6 devices. Bias (X axis) was calculated in relation to the highest device (no. 6). The height (Y axis) was calculated dynamically in relation to the snow depth present. Wind speed is represented from strong (purple) to weak (red) winds. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ton devices, respectively). The performance of the iButton devices in the present study are consistent with those reported by Hubbard et al. (2005), who noted that these devices worked properly at air temperatures  $> 0^{\circ}\text{C}$ , but bias and uncertainty increased below this temperature. In previous studies (Hubbard et al., 2005; Lundquist and Huggett, 2008) iButton devices were calibrated in ice baths, so the problematic frozen conditions were not checked. This issue is clearly relevant for applications involving use of these sensors in sub-zero environments. A slight negative correlation was found between bias and observed air temperature for both the Tinytag and iButton sensors, in agreement with the findings of Johnson et al. (2005).

Experiment 2 shows that air temperature variation relative to measurements made using the Stevenson screen was less above snow than above grass. However, this does not show that snow was the factor responsible, as analysis of radiation-wind effects showed that variations were larger above snow covered surfaces than above grass when radiation and wind conditions were constant. Therefore, radiation and wind are the major factors influencing the performance of radiation shields, as reported previously (Buisán et al., 2015; Georges and Kaser, 2002; Holden et al., 2013; Huwald et al., 2009). The results showed that un-

der varying radiation, wind, and snow conditions the commercial multi-plate Datamate ACS5050 shield provided the best performance compared with the other simpler and more economic alternatives. The Datamate radiation shield is based on Gill's multi-plate models (Gill, 1983), but has been substantially improved over the intervening years. The outstanding performance of multi-plate radiation shields has been previously reported. For example, Payne (1987) indicated a bias range from  $+0.2$  to  $+0.3^{\circ}\text{C}$  for winds  $> 2\text{ m s}^{-1}$ , Hubbard et al. (2005) reported a bias generally ranging from  $+0.5$  to  $+1.0^{\circ}\text{C}$ , and Martínez-Ibarra et al. (2010) showed that these radiation shields provide greater protection than the Stevenson screen with less temperature variation, which is consistent with our findings. Our results showed that the bias was  $\pm 0.5^{\circ}\text{C}$  above grass, and  $0$ – $0.5^{\circ}\text{C}$  above snow. The non-commercial tube and funnel shields, and the absence of a shield, provided variable results depending on the environmental conditions, as the measurements were differently affected by radiation and wind speed. Generally, the tube, funnel, and no shield results were consistent with those of Georges and Kaser (2002), who reported a significant decrease in bias for wind speeds  $> 3.5$ – $4\text{ m s}^{-1}$ . Based on the  $\pm 1^{\circ}\text{C}$  threshold proposed by Tarara and Hoheisel (2007), the no shield data were only

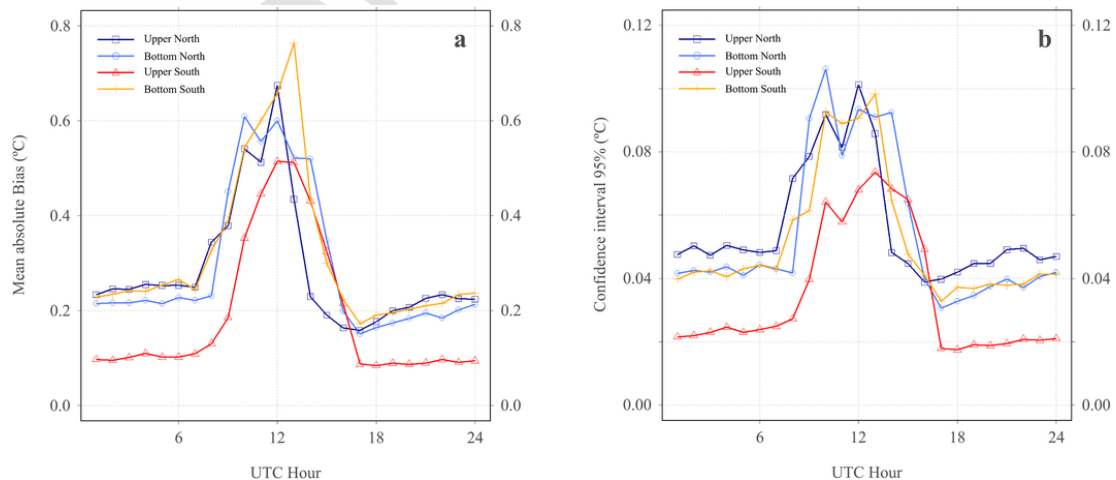


**Fig. 7.** Stabilization layer frequency calculated as the lowest device height where a bias  $< \pm 0.1^\circ\text{C}$  in relation to device no. 6 was measured. The height (X axis) was calculated dynamically in relation to snow depth.

slightly valid at mean night temperatures (zero radiation). Tube shield data were when wind intensity was  $> 3\text{ms}^{-1}$ , funnel shield data were when radiation was  $< 700\text{W m}^{-2}$  ( $400\text{--}500\text{W m}^{-2}$  with snow presence), and Datamate data were in all situations. The absence of a shield provided the worst conditions, with 95% measurements showing variation of  $\pm 2^\circ\text{C}$ , although the bias was relatively low at night (slight cooling). Unshielded devices (no protection and maximum ventilation) were influenced mainly by incident radiation and snow albedo effects. Thus, unshielded devices used in previous studies (e.g., Shine, 2003; Sternberg, 2011) may have been affected by daytime overheating and underestimation at night (Azorin-Molina and Azorin-Molina, 2008). The funnel design provided little protection from incident radiation (large positive bias even with strong winds). Because of its open bottom design, the performance of devices using the funnel shield was

similar to that where no shield was used above high albedo snow surfaces, consistent with the findings of Nakamura and Mahrt (2005) and Tarara and Hoheisel (2007), and better above grass. The funnel design causes overheating that cannot be dissipated by natural ventilation, so better results have been found at low radiation levels (Ribeiro da Cunha, 2015) or avoiding the central hour of the day or in little insulated slopes or forest areas (Lundquist and Huggett, 2008). The tube shield does not favor internal ventilation, so large positive biases occurred even at low radiation levels, with a deficient performance for wind speeds  $< 4\text{ms}^{-1}$  above grass and snow. However, the performance for wind speeds  $> 4\text{ms}^{-1}$  was much better than for the funnel design, especially above snow covered surfaces. In any case, the performance of tube shield is partially dependent on the latitude of measurements, since the inclination depends on it and ventilation is easier when the tube is horizontal, and because albedo effect is more effective overheating when tube is vertical. This implies that its performance would be better in low latitudes than in high ones.

The results of Experiment 3 showed a radiation effect on vertical hourly temperature profiles, with an increasing of variability at midday and a progressive decreasing at night. In addition, this increasing of variability for measurements made close to the snow surface; extreme examples involved deviations of  $-6^\circ\text{C}$  in relation to highest measurement, especially at midday. Hanna et al. (2017) indicated that  $\pm 4^\circ\text{C}$  differences can occur at midday at heights of 1–2m above the snow surface. Midday results showed generalized cooling with decreasing height of devices above the snow, which can be explained by heat exchange between air and the frozen snow surface during daylight hours. Gustavsson (1995) also found that snow can cause cooling of the near surface layer air, and that this is more pronounced at low wind speeds. We found small differences among measurements recorded at night at different heights above the snow. On an hourly basis our results were not always consistent with those of previous studies. For example, Halberstam and Schieldge (1981) proposed that night generally favors air stabilization and near surface air cooling relative to higher air layers, because of air subduction and stratification caused by the cold snow surface. That study also reported a constant diurnal air temperature profile, in contrast with our measurements. However, differences may be determined by differences in climatic conditions, where Pyrenees generally rounds isotherm  $0^\circ\text{C}$  during winter. In addition, Experiment 3 encompassed a complete snow season of 7 months, compared with 3 days in the study of Halberstam and Schieldge (1981) and 1 week in the study of Hanna et al. (2017). Although the optimal is to measure temperature data at a constant height above the snow surface along snow season, remote areas and autonomous self-recording data-



**Fig. 8.** a) Mean hourly absolute temperature bias for the four sites. b) Experimental plots hourly air temperature confidence interval (0.95), based on temperature differences between devices and plot average.

loggers make it difficult to carry it out. For this, Experiment 3 analyzes temperature stabilization. Stabilization boundary analysis has showed that stabilization occurs at 75–100 cm above the snow surface. This suggests that devices should be placed at heights > 100 cm above the maximum expected snow depth. Lundquist and Cayan (2007), Lundquist and Huggett (2008) and Hubbart et al. (2007) anticipated this issue and overcame the problem by installing their devices at varying heights (2–10 m) above the snow surface. Parallel snow depth measurements during field studies can be very useful for estimating data quality.

Experiment 4 showed that there was a significant radiative effect on the horizontal variability of air temperature within forested mountain plots, with two clear temporal periods (daytime and night). Uncertainty was similar among all plots, with uncertainty constant during the night and increasing during daylight hours, especially at midday. The temperature at random selected locations showed a temporal behavior related to shortwave radiation (early morning-late afternoon, morning, afternoon, midday). However, individual devices occasionally produced significantly different measurements. The largest divergences among sensors took place at midday, when solar radiation fluxes were stronger. Thus, despite the use of radiation shields in this experiment, when strong shortwave radiation fluxes occurred, other environmental factors including the presence of rocks, water, or trees produced longwave radiation (Gustavsson, 1995; Holden et al., 2013; Lundquist and Huggett, 2008) that could have affected the measurements made by individual devices. The smallest variations were recorded in the Upper-South site, in agreement with ANOVA analysis, which indicated that measurements made by devices in this plot were consistent. However, in the other plots there were always some devices producing significantly different measurements, and so increased uncertainty. The more dense forest in the Upper-South plot may have reduced measurement variability, because of reduced downwelling of shortwave radiation. In conclusion, diurnal variations may have been because of different shortwave radiation exposures (i.e. forest vs. cleared land; local topographic shadowing) or the proximity to longwave transmitters (i.e. rock surfaces). Night variations were probably mainly to because of the proximity of longwave transmitters (i.e. forest vs. cleared land; rock surfaces; Lundquist and Huggett, 2008). These results suggest the need for analysis of horizontal temperature variability on flat and open areas in future investigations.

## 6. Conclusions

The major findings of this research can be summarized as follows:

- Autonomous self-recording dataloggers are an appropriate and easy option for measuring air temperature in snow-covered and mountainous areas. Tinytag devices showed a robust and constant performance in relation to a reference sensor, whereas the iButton devices showed unstable biases (especially at temperatures < 0 °C), and were subject to unexpected read-save errors. Thus, the use of Tinytag devices is recommended. In the event that iButton devices are used, it is necessary to place multiple devices (2 or more) in parallel.
- Protection of temperature measuring devices from solar radiation is necessary to ensure the quality of air temperature data. Use of the traditional double-louvre naturally-ventilated Stevenson screen is not necessary, as there are other reliable screens available. The commercial multi-plate Datamate ACS5050 shield provided the best protection over all radiation, wind, and snow conditions. The tube and funnel shields showed variable performance, and their limitations must be considered: the tube shield is more useful in snowy and windy environments, while the funnel shield is more useful in the snow-free and low wind areas). Overall, use of commercial plate radiation shields, as the Datamate ACS5050, provide much better results than

other more inexpensive alternatives. Only low radiation and windy conditions seemed to be favorable for use of the inexpensive shielding alternatives.

- The air temperature vertical profile showed stabilization at 75–100 cm above the snow cover surface. Thus, prior to device installation it is important to know the expected snow depth, and to install devices above this level by > 100 cm. Different temporal performances were found including: i) intense superficial air cooling during daylight hours, especially at midday, possibly because of energy fluxes from air to cold snow surfaces; ii) stabilization at night, when all devices involved in the experiment were within (always < 350 cm) the zone of environmental cooling, favored by topography.
- Horizontal air temperature variability analysis showed that measurement variation increased at midday, simultaneously with the increase in solar radiation. Solar radiation enhances physical characteristics of the materials (absorption, emission, reflection), so affecting the temperature measurements.

More research is needed to assess the performance of other devices (dataloggers and radiation shields) used in mountain environments. However, our horizontal variability analysis is novel, and can help explain the “noise” effect found when measuring air temperatures in mountain and snowy environments.

## Acknowledgments

This study was funded by the research projects “El papel de la nieve en la hidrología de la península ibérica y su respuesta a procesos de cambio global-HIDROIBERNIEVE-CGL2017-82216-R” and CLIMPY “Characterization of the evolution of climate and provision of information for adaptation in the Pyrenees” (FEDER-POCTEFA). We thank the Spanish Meteorological Agency (AEMET) and observers Y. Jimenez and S. Campo for the permissions to use official observatories and for the meteorological data used in this study. Navarro-Serrano, F. and Sanmiguel-Vallelado, A. are granted with a pre-doctoral FPU grant (Spanish Ministry of Education, Culture and Sports). Alonso-González, E. and Khorchani, M. are granted with a pre-doctoral FPI grant (Spanish Ministry of Economy and Competitiveness). The authors thank all colleagues who helped in the development of the experiments. Raw data is freely downloadable at <https://zenodo.org/record/1495572>. The authors wish to acknowledge anonymous reviewers for their detailed and helpful comments to the original manuscript.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.atmosres.2019.03.034>.

## References

- Alonso-González, E., Ignacio López-Moreno, J., Gascoin, S., García-Valdecasas Ojeda, M., Sanmiguel-Vallelado, A., Navarro-Serrano, F., Revuelto, J., Ceballos, A., Esteban-Parra, M.J., Essery, R., 2018. Daily gridded datasets of snow depth and snow water equivalent for the Iberian Peninsula from 1980 to 2014. *Earth Syst. Sci. Data* 10. <https://doi.org/10.5194/essd-10-303-2018>.
- Anderson, P., Neff, W., 2008. Boundary layer physics over snow and ice. *Atmos. Chem. Phys.* 8, 3563–3582. <https://doi.org/10.5194/acp-8-3563-2008>.
- Angilletta, M., Krochmal, A., 2003. The thermochron: a truly miniature and inexpensive temperature logger. *Herpetol. Rev.* 34 (1), 31–32.
- Arck, M., Scherer, D., 2001. A physically based method for correcting temperature data measured by naturally ventilated sensors over snow. *J. Glaciol.* 47, 665–670. <https://doi.org/10.3189/172756501781831774>.
- Azorin-Molina, C., Azorin-Molina, J.C., 2008. Estudio experimental del efecto artificial introducido por la garita meteorológica en la temperatura mínima del aire. In: Sigro, J., Brunet, M., Aguilar, E. (Eds.), *Cambio Climático Regional y Sus Impactos*. 2, Spanish Climatological Society (AEC), Tarragona (Spain), pp. 25–35.
- Barry, R., 1992. *Mountain Weather and Climate*, 2nd ed Psychology Press.

- Bonnardot, V., Carey, V., Madelin, M., Cautenet, S., Coetzee, Z., Quénol, H., 2012. Spatial variability of Night temperatures at a fine scale over the Stellenbosch Wine District, South Africa. *J. Int. Sci. Vigne Vin* 46, 1–13.
- Brabyn, L., Zawar-Reza, P., Stichbury, G., Cary, C., Storey, B., Laughlin, D., Katurji, M., 2014. Accuracy assessment of land surface temperature retrievals from Landsat 7 ETM+ in the Dry Valleys of Antarctica using iButton temperature loggers and weather station data. *Environ. Monit. Assess.* 186, 2619–2628. <https://doi.org/10.1007/s10661-013-3565-9>.
- Brunet, M., Saladié, O., Jones, P., Sigró, J., Aguilar, E., Moberg, A., Lister, D., Walther, A., Lopez, D., Almaraz, C., 2006. The development of a new dataset of Spanish Daily Adjusted Temperature Series (SDATS) (1850–2003). *Int. J. Climatol.* 26, 1777–1802. <https://doi.org/10.1002/joc.1338>.
- Buisán, S., Azorín-Molina, C., Jimenez, Y., 2015. Impact of two different sized Stevenson screens on air temperature measurements. *Int. J. Climatol.* 35, 4408–4416. <https://doi.org/10.1002/joc.4287>.
- Buisán, S., Earle, M.E., Collado, J.L., Kochendorfer, J., Alastrué, J., Wolff, M., Smith, C.D., López-Moreno, J.I., 2017. Assessment of snowfall accumulation underestimation by tipping bucket gauges in the Spanish operational network. *Atmos. Meas. Tech.* 10, 1079–1091. <https://doi.org/10.5194/amt-10-1079-2017>.
- Campbell, G., Norman, J., 1998. *Introduction to Environmental Biophysics*. Springer.
- Camuffo, D., Bertolin, C., 2012. The earliest temperature observations in the world: the Medici Network (1654–1670). *Clim. Chang.* 111, 335–363. <https://doi.org/10.1007/s10584-011-0142-5>.
- Dominguez-Villar, D., Lojen, S., Krllec, K., Baker, A., Fairchild, I., 2015. Is global warming affecting cave temperatures? Experimental and model data from a paradigmatic case study. *Clim. Dyn.* 45, 569–581. <https://doi.org/10.1007/s00382-014-2226-1>.
- Fang, Z., Li, N., Li, B., Luo, G., Huang, Y., 2014. The effect of building envelope insulation on cooling energy consumption in summer. *Energy Build.* 77, 197–205. <https://doi.org/10.1016/j.enbuild.2014.03.030>.
- Fritschen, L., Gay, L., 1979. *Environmental Instrumentation*. Springer, New York.
- Fuchs, M., Tanner, C., 1965. Radiation shields for air temperature thermometers. *J. Appl. Meteorol.* 4, 544–547.
- Gasvoda, D., Tinus, R., Burr, K., Trent, A., 2002. Monitoring the temperature of tree seedlings with the thermochron iButton data logger. In: *USDA For. Serv. Technol. Dev. Program, Timber Tech Tips*, 2400, March 2002.
- Georges, C., Kaser, G., 2002. Ventilated and unventilated air temperature measurements for glacier-climate studies on a tropical high mountain site. *J. Geophys. Res.* 107, 4775. <https://doi.org/10.1029/2002JD002503>.
- Gill, G., 1983. Comparison Testing of Selected Naturally Ventilated Solar Radiation Shields. NOAA, Bay St. Louis.
- Gustavsson, T., 1995. A study of air and road-surface temperature variations during clear windy nights. *Int. J. Climatol.* 15, 919–932. <https://doi.org/10.1002/joc.3370150806>.
- Halberstam, I., Schieldge, J., 1981. Anomalous behavior of the atmospheric surface layer over a melting snowpack. *J. Appl. Meteorol.* 20, 255–265. [https://doi.org/10.1175/1520-0450\(1981\)020<0255:ABOTAS>2.0.CO;2](https://doi.org/10.1175/1520-0450(1981)020<0255:ABOTAS>2.0.CO;2).
- Hanna, E., Mernild, S., Yde, J., Villiers, S., 2017. Surface Air temperature fluctuations and lapse rates on Olivares Gamma Glacier, Rio Olivares Basin, Central Chile, from a novel meteorological sensor network. *Adv. Meteorol.* 17, 1–15. <https://doi.org/10.1155/2017/6581537>.
- Harrison, R.G., 2010. Natural ventilation effects on temperatures within Stevenson screens. *Q. J. R. Meteorol. Soc.* 136, 253–259. <https://doi.org/10.1002/qj.537>.
- Helgason, W., Pomeroy, J., 2012. Characteristics of the near-surface boundary layer within a mountain valley during winter. *J. Appl. Meteorol. Climatol.* 51, 583–597. <https://doi.org/10.1175/JAMC-D-11-058.1>.
- Hofer, I., Hart, A., Martín-Vega, D., Hall, M., 2017. Optimising crime scene temperature collection for forensic entomology casework. *Forensic Sci. Int.* 270, 129–138. <https://doi.org/10.1016/j.forsciint.2016.11.019>.
- Holden, Z., Klene, A., Keefe, R., Moisen, G., 2013. Design and evaluation of an inexpensive radiation shield for monitoring surface air temperatures. *Agric. For. Meteorol.* 180, 281–286. <https://doi.org/10.1016/j.agrformet.2013.06.011>.
- Hubbard, K., Lin, X., Walter-Shea, E., 2001. The effectiveness of the ASOS, MMTS, Gill, and CRS air temperature radiation shields\*. *J. Atmos. Ocean. Technol.* 18, 851–864. [https://doi.org/10.1175/1520-0426\(2001\)018<0851:TEOTAM>2.0.CO;2](https://doi.org/10.1175/1520-0426(2001)018<0851:TEOTAM>2.0.CO;2).
- Hubbart, J., 2011. An inexpensive alternative solar radiation shield for ambient air temperature micro-sensors. *J. Nat. Environ. Sci.* 2, 9–14.
- Hubbart, J., Link, T., Campbell, C., Cobos, D., 2005. Evaluation of a low-cost temperature measurement system for environmental applications. *Hydrol. Process.* 19, 1517–1523. <https://doi.org/10.1002/hyp.5861>.
- Hubbart, J., Kavanagh, K., Pangle, R., Link, T., Schotzko, A., 2007. Cold air drainage and modeled nocturnal leaf water potential in complex forested terrain. *Tree Physiol.* 27, 631–639. <https://doi.org/10.1093/treephys/27.4.631>.
- Huwald, H., Higgins, C., Boldi, M., Bou-Zeid, E., Lehning, M., Parlange, M., 2009. Albedo effect on radiative errors in air temperature measurements. *Water Resour. Res.* 45, <https://doi.org/10.1029/2008WR007600>.
- Imholt, C., Soulsby, C., Malcolm, I., Hrachowitz, M., Gibbins, C., Langan, S., Tetzlaff, D., 2013. Influence of scale on thermal characteristics in a large montane River Basin. *River Res. Appl.* 29, 403–419. <https://doi.org/10.1002/rra.1608>.
- Johnson, A., Boer, B., Woessner, W., Stanford, J., Poole, G., Thomas, S., O'Daniel, S., 2005. Evaluation of an inexpensive small-diameter temperature logger for documenting ground water-river interactions. *Groundw. Monit. Remediat.* 25, 68–74. <https://doi.org/10.1111/j.1745-6592.2005.00049.x>.
- Juliussen, H., Humlum, O., 2007. Towards a TTOP ground temperature model for mountainous terrain in central-eastern Norway. *Permafrost. Periglac. Process.* 18, 161–184. <https://doi.org/10.1002/ppp.586>.
- Kurzeja, R., 2010. Accurate temperature measurements in a naturally-aspirated radiation shield. *Bound. Layer Meteorol.* 134, 181–193. <https://doi.org/10.1007/s10546-009-9430-2>.
- López-Moreno, J.I., Navarro-Serrano, F., Azorín-Molina, C., Sánchez-Navarrete, P., Alonso-González, E., Rico, I., Morán-Tejeda, E., Buisán, S., Revuelto, J., Pons, M., Vicente-Serrano, S.M., 2018. Air and wet bulb temperature lapse rates and their impact on snowmaking in a Pyrenean ski resort. *Theor. Appl. Climatol.* <https://doi.org/10.1007/s00704-018-2448-y>.
- Lundquist, J., Cayan, D., 2007. Surface temperature patterns in complex terrain: Daily variations and long-term change in the Central Sierra Nevada, California. *J. Geophys. Res.* 112, D11124. <https://doi.org/10.1029/2006JD007561>.
- Lundquist, J., Huggett, B., 2008. Evergreen trees as inexpensive radiation shields for temperature sensors. *Water Resour. Res.* 44, <https://doi.org/10.1029/2008WR006979>.
- Lundquist, J., Lott, F., 2008. Using inexpensive temperature sensors to monitor the duration and heterogeneity of snow-covered areas. *Water Resour. Res.* 44, <https://doi.org/10.1029/2008WR007035>.
- Makgose, M., Phillipus, C., 2015. Effect of plant growth regulators on growth, yield, and quality of sweet pepper plants grown hydroponically. *Hort Sci.* 50, 383–386.
- Martínez-Ibarra, E., Azorín-Molina, C., Bañón, M., Olcina Cantos, J., Estrela, M.J., Gil-Olcina, A., 2010. Intercomparación de las temperaturas extremas en tres tipos de garita meteorológica: Montsouris, Stevenson y Young.
- Measham, P., Darbyshire, R., Turpin, S., Murphy-White, S., 2017. Complexity in chill calculations: a case study in cherries. *Sci. Hortic. (Amsterdam)* 216, 134–140. <https://doi.org/10.1016/j.scienta.2017.01.006>.
- Morán-Tejeda, E., Bazo, J., López-Moreno, J.I., Aguilar, E., Azorín-Molina, C., Sanchez-Lorenzo, A., Martínez, R., Nieto, J.J., Mejía, R., Martín-Hernández, N., Vicente-Serrano, S.M., 2016. Climate trends and variability in Ecuador (1966–2011). *Int. J. Climatol.* 36, 3839–3855. <https://doi.org/10.1002/joc.4597>.
- Nakamura, R., Mahrt, L., 2005. Air temperature measurement errors in naturally ventilated radiation shields. *J. Atmos. Ocean. Technol.* 22, 1046–1058. <https://doi.org/10.1175/JTECH1762.1>.
- Navarro-Serrano, F., López-Moreno, J.I., Azorín-Molina, C., Alonso-González, E., Tomás-Burguerra, M., Sanmiguel-Vallado, A., Revuelto, J., Vicente-Serrano, S.M., 2018. Estimation of near-surface air temperature lapse rates over continental Spain and its mountain areas. *Int. J. Climatol.* <https://doi.org/10.1002/joc.5497>.
- Nordli, P., Alexandersson, H., Frich, P., Førland, E., Heino, R., Jónsson, T., Tuomenvirta, H., Tveito, O., 1997. The effect of radiation screens on Nordic time series of mean temperature. *Int. J. Climatol.* 17, 1667–1681. [https://doi.org/10.1002/\(SICI\)1097-0088\(199712\)17:15<1667::AID-JOC221>3.0.CO;2-D](https://doi.org/10.1002/(SICI)1097-0088(199712)17:15<1667::AID-JOC221>3.0.CO;2-D).
- O'Neill, H., Christiansen, H., 2018. Detection of ice wedge cracking in permafrost using miniature accelerometers. *J. Geophys. Res. Earth Surf.* 123, 642–657. <https://doi.org/10.1002/2017JF004343>.
- Pagès, M., Pepin, N., Miró, J.R., 2017. Measurement and modelling of temperature cold pools in the Cerdanya valley (Pyrenees), Spain. *Meteorol. Appl.* 24, 290–302. <https://doi.org/10.1002/met.1630>.
- Payne, R.E., 1987. *Technical Report: Air Temperature Shield Test*.
- Pepin, N., Kidd, D., 2006. Spatial temperature variation in the Eastern Pyrenees. *Weather* 61, 300–310. <https://doi.org/10.1256/wea.106.06>.
- Pepin, N., Duane, W., Hardy, D., 2010. The montane circulation on Kilimanjaro, Tanzania and its relevance for the summit ice fields: comparison of surface mountain climate with equivalent reanalysis parameters. *Glob. Planet. Change* 74, 61–75. <https://doi.org/10.1016/j.gloplacha.2010.08.001>.
- Quénol, H., Bonnardot, V., 2014. A multi-scale climatic analysis of viticultural terroirs in the context of climate change: the TERADCLIM project. *J. Int. Sci. Vigne Vin* 25–34.
- Ribeiro da Cunha, A., 2015. Evaluation of measurement errors of temperature and relative humidity from HOBO data logger under different conditions of exposure to solar radiation. *Environ. Monit. Assess.* 187, 236. <https://doi.org/10.1007/s10661-015-4458-x>.
- Richardson, S., Brock, F., Semmer, S., Jirak, C., 1999. Minimizing errors associated with multiplate radiation shields. *J. Atmos. Ocean. Technol.* 16, 1862–1872. [https://doi.org/10.1175/1520-0426\(1999\)016<1862:MEAWMR>2.0.CO;2](https://doi.org/10.1175/1520-0426(1999)016<1862:MEAWMR>2.0.CO;2).
- Rolland, C., 2003. Spatial and seasonal variations of air temperature lapse rates in alpine regions. *J. Clim.* 16, 1032–1046. [https://doi.org/10.1175/1520-0442\(2003\)016<1032:SASVOA>2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)016<1032:SASVOA>2.0.CO;2).
- Roznik, E., Alfrod, R., 2012. Does waterproofing Thermochron iButton dataloggers influence temperature readings?. *J. Therm. Biol.* 37, 260–264. <https://doi.org/10.1016/j.jtherbio.2012.02.004>.
- Schofield, G., Bishop, C., Katselidis, K., Dimopoulos, P., Pantis, J., Hays, G., 2009. Microhabitat selection by sea turtles in a dynamic thermal marine environment. *J. Anim. Ecol.* 78, 14–21. <https://doi.org/10.1111/j.1365-2656.2008.01454.x>.
- Shedekar, V.S., King, K.W., Fauser, N.R., Soboyejo, A.B.O., Harmel, R.D., Brown, L.C., 2016. Assessment of measurement errors and dynamic calibration methods for three different tipping bucket rain gauges. *Atmos. Res.* 178–179, 445–458. <https://doi.org/10.1016/j.atmosres.2016.04.016>.
- Sokol, Z., Bližňák, V., Sedláč, P., Zacharov, P., Pešice, P., Škuthan, M., 2017. Ensemble forecasts of road surface temperatures. *Atmos. Res.* 187, 33–41. <https://doi.org/10.1016/j.atmosres.2016.12.010>.
- Tarara, J., Hoheisel, G., 2007. Low-cost shielding to minimize radiation errors of temperature sensors in the field. *Hort Sci.* 42, 1372–1379.
- Wilcoxon, F., 1945. Individual comparisons by ranking methods. *Biom. Bull.* 1, 80. <https://doi.org/10.2307/3001968>.
- WMO, 2014. *Guide to Meteorological Instruments and Methods of Observation*, (Geneva, Switzerland).
- Wolaver, B., Sharp, J., 2007. Thermochron iButton: limitation of this inexpensive and small-diameter temperature logger. *Ground Water Monit. Remediat.* 27 (3), 127–128.
- Yang, Z., Hanna, E., Callaghan, T., Jonasson, C., 2012. How can meteorological observations and microclimate simulations improve understanding of 1913–2010 climate change around Abisko, Swedish Lapland?. *Meteorol. Appl.* 19, 454–463. <https://doi.org/10.1002/met.276>.