Statistical analysis and clusterization of environmental parameters measurements from co-located sensors

Statistical analysis and clusterization of environmental parameters measurements from co-located sensors

G. Coppa[1], M. Soncini[2], F.R. Pennecchi[1]

1 Istituto Nazionale di Ricerca Metrologica, Torino, IT
2 Università di Torino, Torino, IT

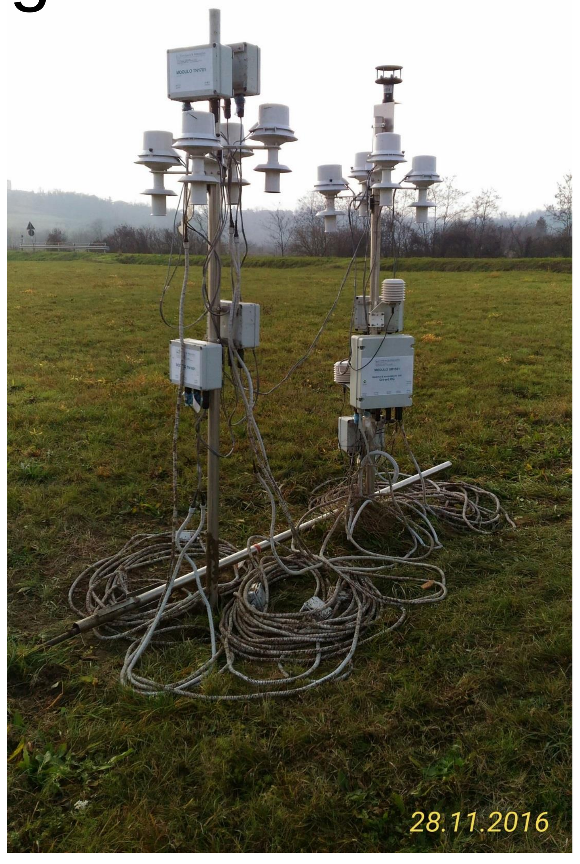
Abstract

One of the most difficult things in metrology for the environment is the **evaluation of measurement uncertainties**. **Air temperature sensors are influenced by many external factors**, such as solar radiation – both direct and diffuse – humidity, wind speed and direction. Laboratory calibration is necessary but not sufficient for a comprehensive understanding of the behaviour of air temperature sensors while subject to the wide variety of environmental conditions they meet in normal operation.

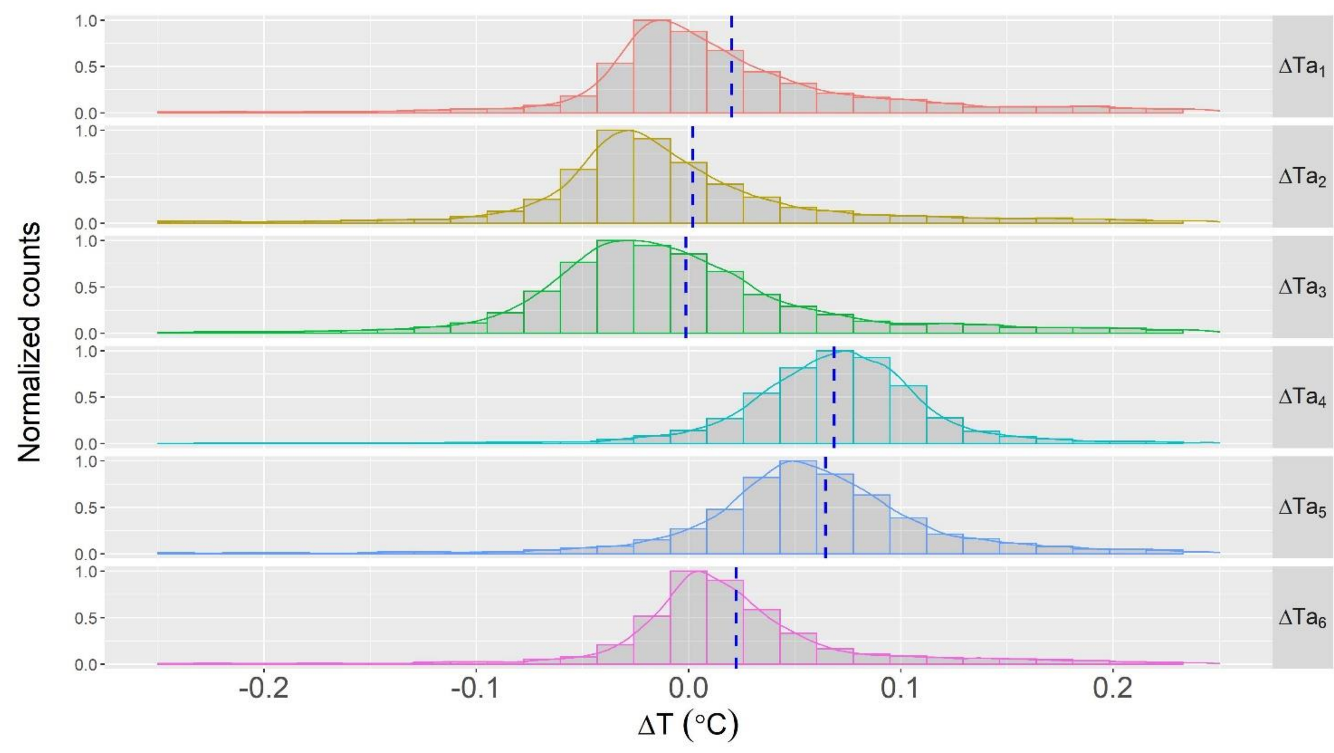
This work follows up the one by **Coppa et al., (2021),** where the evaluation of uncertainties of commercially available air temperature sensors and shields was computed by a dedicated experiment, where 7 sensors were co-located for 2 months. In that work, **uncertainties were evaluated simply by computing standard deviations of temperature differences between each sensor and one taken as a reference**.

The road siting effect evaluation experiment

Sensor array during normal phase of the experiment.



Co-located sensors on the experimental field site.



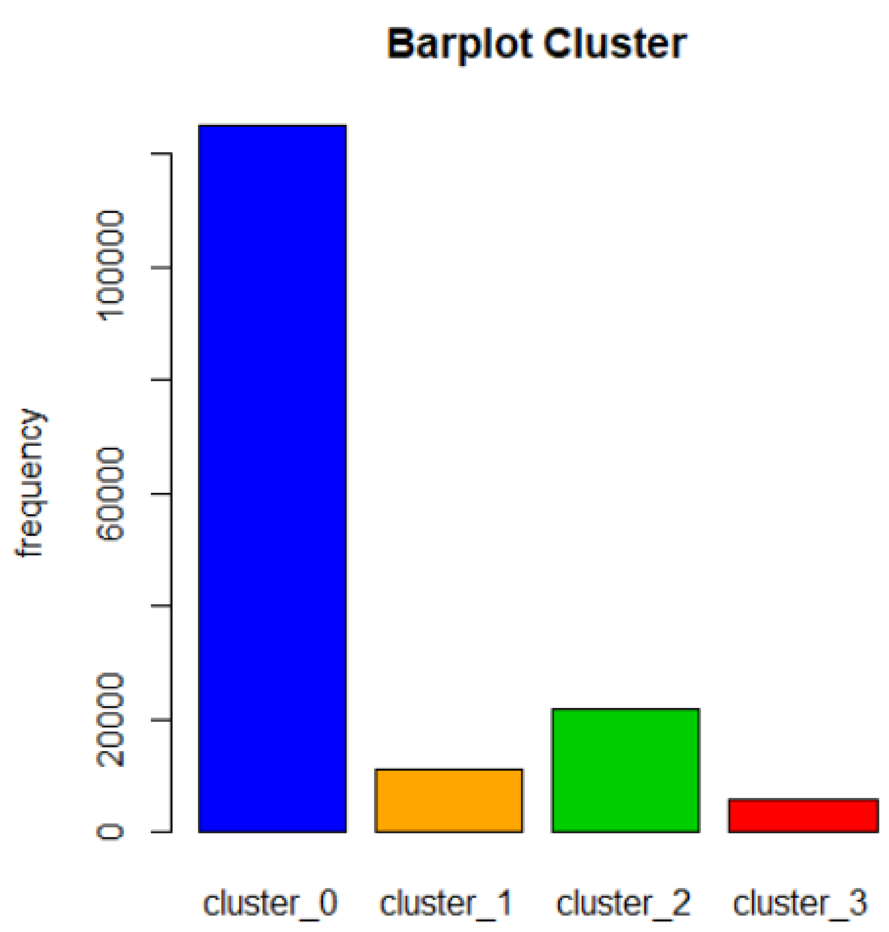
Distributions of temperature differences between each sensor and a reference during co-location of the measurement stations. Blue vertical dashed lines represent means.

Coppa, Quarello, Steeneveld, Jandric and Merlone, 2021, *IntJClim*, 41: 3705– 3724 DOI: 10.1002/joc.7044

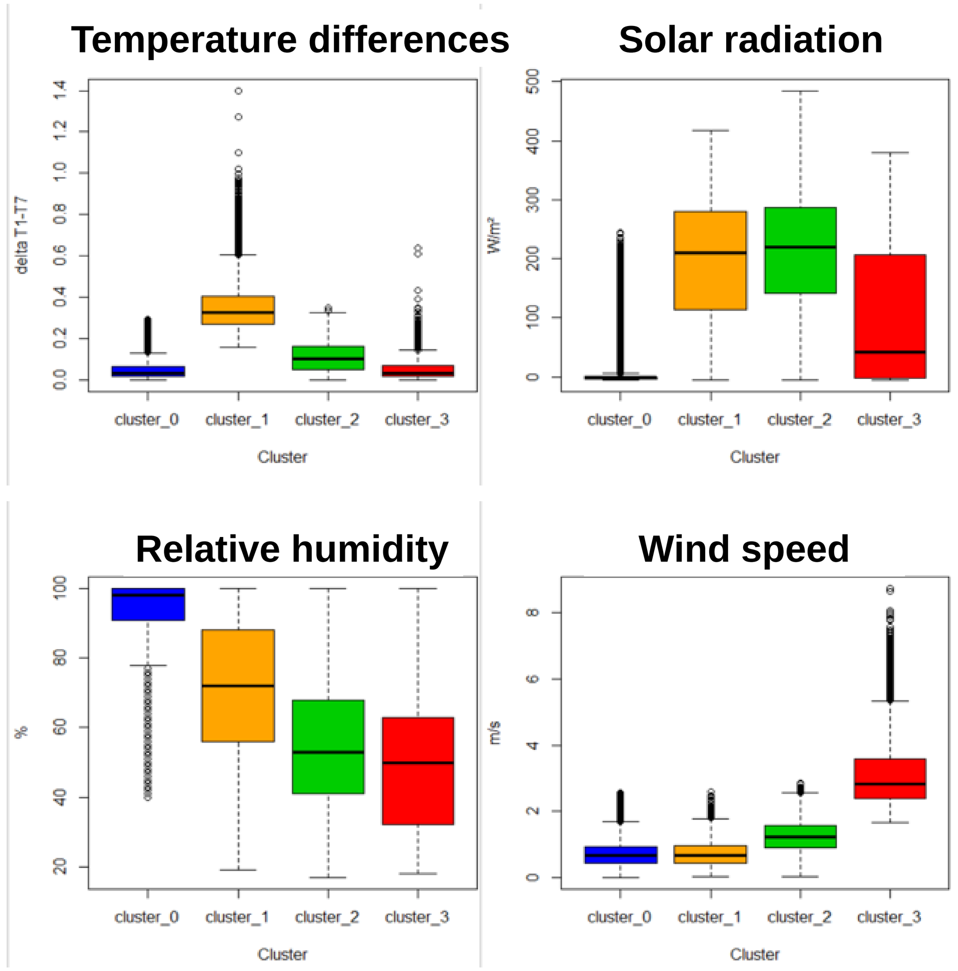
The present work provides a more refined uncertainty evaluation by a statistical method, through **machine learning and clusterization**, in order to **evaluate uncertainties in dependence on the environmental conditions.**

Temperature differences during the co-location experiment have been run through an **unsupervised k-means algorithm** that revealed several clusters: in particular, **temperature differences are lower, thus leading to smaller uncertainty**, under two very different environmental conditions: the first, **at nights with very low winds**; the second **during high-radiation and high-wind days**.

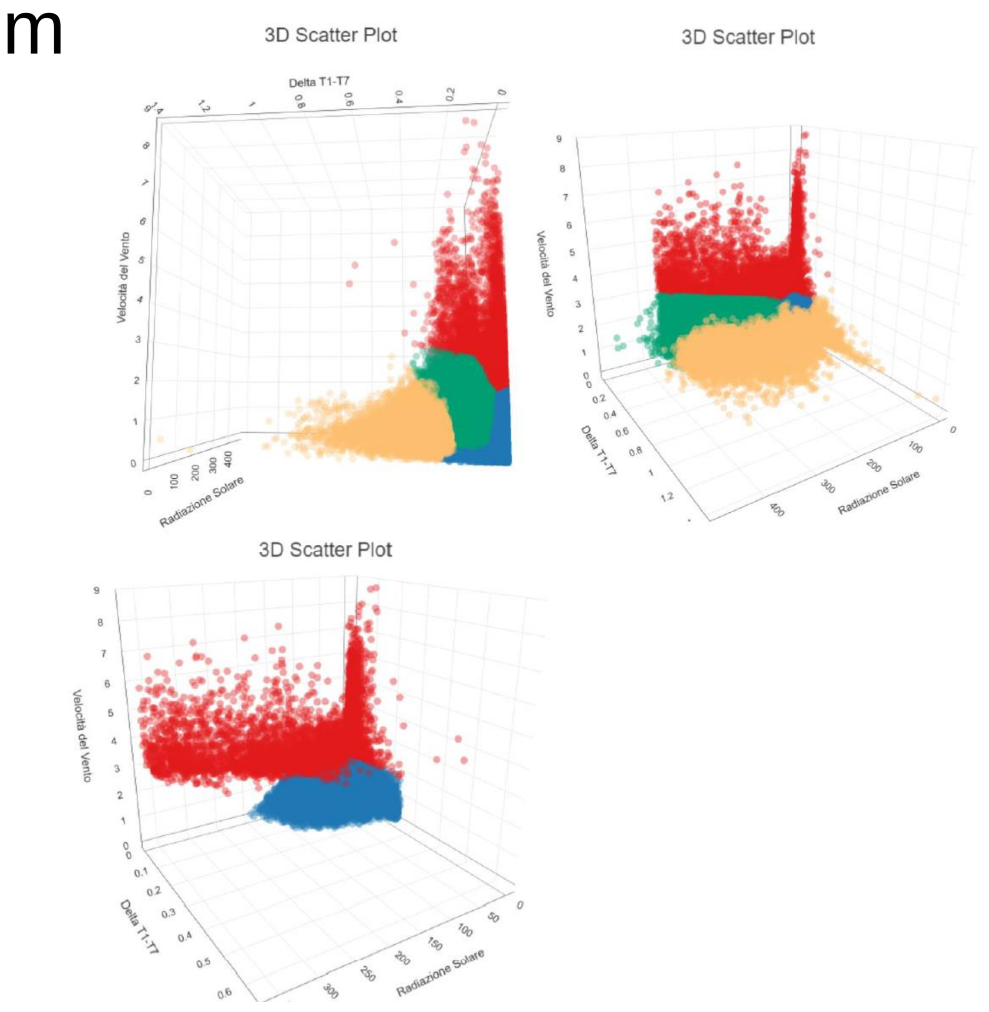
Unsupervised k-means clustering algorithm



Numerosity of the meteorological observations in the 4 clusters identified by the k-means algorithm.



Distributions of meteorological variables in the 4 clusters identified by the k-means algorithm.



Representation of 4 clusters in a simplified 3D space of the meteorological parameters.

Temperature differences can be corrected by subtracting the **cluster centroid** value from each observation, while **uncertainties can be taken** as the **standard deviations of each cluster distributions**, then associated to each sensor+shield system depending on the specific environmental conditions (evaluated by ancillary measurements). This methodology **reduces the overall uncertainty** and **improves the level of detail** about sensors behaviour knowledge.

	ΔT (°C)	Rad (W/m²)	RH (%)	Wind speed (m/s)
Cluster 0	0.047	7	94.0	0.71
Cluster 1	0.352	192	71.0	0.73
Cluster 2	0.109	207	54.1	1.24
Cluster 3	0.051	106	49.9	3.11

Cluster centroids of the 4 different parameters.

	$\sigma_{\Delta T}$ (°C)	Variation
Cluster 0	0.045	43%
Cluster 1	0.120	115%
Cluster 2	0.070	67%
Cluster 3	0.055	53%

Standard deviations of ΔT measurements and variation over previous evaluation.

Conclusions

- **A way to refine the evaluation of measurement uncertainties in air temperature, is presented.**
- **A clustering algorithm to group together measurements taken in similar conditions is used**
- **Centroids as bias corrections**
- **Standard deviations of cluster as uncertainties**
- **Uncertainties for the largest part of measurements are reduced**

Future work

- **Extend to longer time periods (whole year)**
- **Explore other clustering algorithms**

Mail to g.coppa@inrim.it