Review of The AntAWS dataset: a compilation of Antarctic automatic weather station observations
Author(s): Yetang Wang, Xueying Zhang, Wentao Ning, Matthew A. Lazzara, Minghu Ding, Carleen H. Reijmer, Paul C. J. P. Smeets, Paolo Grigioni, Elizabeth R. Thomas, Zhaosheng Zhai, Yuqi Sun, and Shugui Hou
MS No.: essd-2022-241

Christophe Genthon, Laboratoire de Météorologie Dynamique, Paris, France

This paper presents a welcome compilation and tentative unification of data from automatic weather stations (AWS) in Antarctica.

Much of meteorological Antarctica would be essentially unknown if it was not for for the data provided by networks of AWS. Such networks have been developed and deployed by different groups, such as the AMRC, the University of Utrecht, etc, with largely similar instruments and methods but little homogeneity in the way the data are quality controlled and distributed. Yetang Wang and colleague's work is timely and definitely useful. Along with the dataset itself and associated metadata, they provide some statistics of the Antarctic meteorology and climate from AWS.

Of course one can (and should?) find that there is room for improvement. I see 2 points which need to be at least reported, possibly improved to the extent that information is available that is not yet reported in the data set.

One issue is the height of the instruments above the snow surface. The authors report that the AMRC standard height is 3 m but in practice, because of snow accumulation and infrequent visits, the height is quite variable and mostly unknown. This is an issue for temperature in areas where surface based temperature inversions can be strong (high plateau in winter). This is more particularly an issue with wind speed which varies strongly with height in the first few meters above the surface, decreasing 0 at the surface even with the strongest winds a few meters above. A few AWS (an increasing number?) have snow height variations measurements using acoustic depth gauge. This is an interesting information by itself as it measures snow accumulation which relates to snow fall and other accumulation processes such as snow erosion and blowing snow. In the first place, this is an essential information to adequately exploit the wind speed information. Unfortunately, because of snow height uncertainly, the consistency of Wang et al.'s data set is unwarrented for wind speed. This should reported.

An other issue is with the humidity. Most practical humidity sensors for AWS use Vaisala's Humicap capacitive sensor. The humicap is calibrated to report the relative humidity with respect to liquid water even below 0°C. The authors report that the relative humidity is often close to 100%, which suggest that they have converted humidity with respect to liquid into relative humidity with respect to ice. They should be clear which relative humidity they report and which conversion formulas they use. Also, it is known that in Antarctica, even near the surface, the relative humidity with respect to ice often reaches well over 100%, but few sensors can measure humidity above 100% and generally not those operated on AWS. There are recent publications in ESSD reporting this and distributing corresponding data for the high antarctic plateau. This should be reported here, and thus the fact that the database is biased low with respect to humidity, at least on the antarctic plateau.

The concluding remarks could offer some recommendations as to how to improve the AWS network. Of course we want more of these but funds are limited : which regions should be prioritized for more AWS deployment? It could be recommended that snow height sensors be systematically implemented and snow height data provided along with the meteorological data. One could also recommend to use mechanically ventilated radiation shields. Of course, power for ventilation is an issue but radiation biases occur in summer when solar power is available.

Specific comments:

Line 33: these are fairly outdated references. There surely are more recent references e.g. from the more recent IPCC reports

Line 41: I beleive this is 1958. There was no IGY in 2007. There was an international polar year started in 2007 but certainly not 50 staffed stations established then.

Lines 63-64: Why are those other AWS left aside? One major virtue of the work presented here is the efforts made to collate, harmonize and consistently distribute data which are otherwise scattered here and there. Why leave aside some data known to exist?

Line 106 and further: CR1000 is a device, not a series. It is a datalogger and should be presented as such, as this is the way the manufacturer Campbell Sci presents it. Campbell Sci should show as the manufacturer.

Line 110: Verify with BAS but initially (circa 200s), BAS made their own data loggers. They shifted to CR1000 later on.

Lines 118-119: hard to understand: is this a tripod or a mast? In fact most long term AWS are on masts, e.g. AMRC’s.

Line 137: This is the problem, nominal height, possibly known at deployment and after visits but most of the time it is fully unknown unless the AWS is equiped with an ADG which is generally not the case.

Lines 139-140: Sorry but this is this is a ridiculous estimation of the error. Eisen et al. is about long term mean snow accumulation, and they report accumulations up to and more than 1 m / year in some places in Antarctica. Surely the height uncertainty issue is less where accumulation is less, e.g. on the high plateau, but this uncertainty is first a matter of mean accumulation and servicing frequency.

Table 1: Any information here on where temperature reports may benefit aspirated radiation shielding to avoid radiation biases?

Also in table 1: I am a bit confused with the term “impeller”. Vane manufacturer R. M. Young, for instance, call it “propeller”

Still Table 1, BAS is reported using HMP155 resistance probe for relative humidity. HMP155 actually uses the Humicap capacitive sensor. The temperature report from HMP155 uses platinum resistance to report temperature, not humidity.

Lines 178-179: please provide internet links for consistency with other sources of information. Otherwise, should this be “personal communication”?

Figure 4: In the data processing step, should this be “flagging” rather than interpolating?

Lines 209-211: concerning the 3-hour time step: are the data instant measures every 3 hours , or averaged over 3 hours? Is this consistent across datasets? How do you average wind direction?

Lines 232-232: did / could you check that no mechanical ventilation is used before blacklisting low wind cases? This is probably mostly the case, but should there be some valid reports by low wind speed thanks to mechanical ventilation?

Line 239: the 3-value criteria should probably also include that the 3 values are homogeneously distributed during the day, otherwise the a time-of-day bias is likely is summer when temperature strongly varies with sun elevation.

Line 252: again, any indication that some temperature reports may benefit mechanical ventilation?

Line 262: in fact, Dome C and Concordia are one and the same site, if not necessarily the same AWS. No wonder they show the same extremes. I suggest keep only Dome C here.

Section 4.3: please mention the relative humidity issues raised above here: sensors report RH with respect to liquid, data and must be converted to get RH with respect to ice; and the sensors used on AWS cannot report supersaturation, which is frequent on the high antarctic plateau – the humidity data are thus biased low there.

Section 4.4: please mention that poorly known instrument height above the snow surface affects the data quality / consistency. Still, the time evolution of wind speed with time is an important information, but the modulus is not well known and not consistent in the dataset.

Figures S1, S2, S3, S4: mention that there is no color code, colors are used to improve readability?