Answer to Anonymous Referee #2 RC4, for AMT-2018-179

Title: Evaluation of Windsond S1H2 performance in Kumasi during the 2016 DACCIWA field campaign

Authors: Geoffrey E.Q. Bessardon, Kwabena Fosu-Amankwah, Anders Petersson, Barbara J. Brooks

We thank the anonymous referee for the helpful comments. We are responding to all the comments of referee #2 in this document and we have prepared a revised manuscripts where changes are marked and removed parts as follows: blue correspond to reviewer #1 RC2, red reviewer #1 RC3, green reviewer #2 RC4. In the following, comments of the referee #2 are given in bold and italic with our responses in normal font.

This paper presents an evaluation of a relatively new low-cost radiosonde system against a well-established and widely used radiosonde based on measurements performed in June and July 2016 during a field campaign in Ghana - Western Africa. The low-cost radiosondes were recovered by the operators and reused up to 8 times, which allows the authors to analyse a relatively high number of ascends. It is shown that under "simple" atmospheric conditions temperature, humidity and pressure measured by both systems compare reasonably well, but as soon as larger vertical inhomogeneities occur the low-cost radiosonde suffers from slow sensor response and hysteresis. GPS-derived wind from the low-cost system is of very bad quality.

Unfortunately, the paper suffers from several weaknesses starting by the design of the measurements, missing technical information, lack of measurements under laboratory conditions and a very limited analysis of the data. The authors miss to cite and discuss relevant literature e.g . Legain et al. 2013 doi:10.5194/amtd-6-3339-2013 and Nash et al, 2010 WMO Report No. 107 Instruments and Observations. The weather situation is not sufficiently discussed and taken into account. Overall it seems to me that the paper is a kind of side product produced with minimal effort.
I think that the paper will not warrant publication as long as a mayor revision is done which addresses the following comments.

We would like to thank the reviewer for the references to the relevant literature we have missed and the their comments concerning the contextualisation of the work. This work was made in the context of the DACCIWA field campaign were the Windsond S1H2 was being integrated into a large scale scientific sounding programme for the first time. This sonde has never been used in this manner before hence we took the opportunity presented by the DACCIWA field campaign to compare the radiosonde with a proven system: providing benchmarking for the interpretation S1H2 sonde data obtained as part of DACCIWA. The experimental design was limited by the needs of the field campaign and the resources available but despite these limitations it did allow the identification of a number of issues with the Windsond S1H2 and to feed these back to the manufacturer to help the development of a reusable sonde system that is easier to use than the system presented by Legain et al., 2013.

The conclusion drawn in this paper is by no means a definitive conclusion on the Sparv Embedded system but a list of recommendation for development as well as recommendation for the future users of the DACCIWA field campaign data and as such is still usefull for the community.

Specific comments:2

Page 2 The first section is a marketing analysis which is mostly irrelevant if you want to discuss a reusable low cost sonde that is limited to 6000 m altitude. Sonde costs are fixed - price differences for launches in different regions depend on logistics and local labour.

Answer: Agreed, following reviewer #1 comments we have noticed that the first section is confusing so we added "This rough estimate varies regionally as the price of labour, helium and balloons and is not the same around the globe. Yet operational costs are a significant investment in countries with limited resources."

We have introduced Legain et al., 2013 system and discussed the limitations of the system for the development of an operational network using this sonde: "Re-usable sondes have been introduced for the first time by Legain, et al., 2013 which modified a Vaisala sonde to enclose it in a cage which is tied to a couple of balloon. The caged allowed the balloon to detached at a desired altitude and slowly descend before recovery. Despite this system has shown successful results in pressure temperature and humidity, and recovery rate it does not asses the effect of the cage and the two balloons on the obtained wind profile. The sonde modification required makes the use of this system more complex and can be an obstacle towards a global use of the system, this shows that re-usable sonde technologies are still a work in progress where manufacturers can develop their own solutions."

If the sounding program had the objective to evaluate the Windsond performance already from the beginning please explain the following: 1) Why is there only one tandem flight reaching higher altitudes performed 2) Why are all low altitude intercomparison flights performed only at 0600 and not distributed over day and night or at least over the launch times shown on figure 2. As the sondes were recovered no significant additional costs would have been created. 3) Why are the RS42 and Windsond not tied together for the low altitude intercomparison flights – the resulting spacial difference makes it impossible to separate instrument errors from atmospheric variability.

Answer: The goal of the DACCIWA ground field campaign was to provide a high-quality comprehensive dataset for processes studies, in particular interactions between low-level clouds (LLCs) and boundary layer conditions. The DACCIWA radio sounding program was then designed to complete these main objectives. The Vaisala RS42s were launched to provide synoptic observations during the campaign with complementary synoptic measurements during IOPs. The Windsond S1H2s were launched to provide more frequent boundary layer sounding during DACCIWA IOPs, to observe the evolution of the LLCs, and associated phenomena such as the Nocturnal Low-Level Jet (NLLJ) The frequent radiosounding program thus focussed on night-time measurements as detailed on figure 2.

As the S1H2 was never used in the context of a field campaign we had to control the quality of the S1H2 in order to facilitate the interpretation of the data recorded by the S1H2. This performance assessment had however to be done without impacting on the main objectives. We agree that only one sounding reaching higher altitudes is not statistically sufficient to assess the S1H2 performances. Launches in which an S1H2 and RS92 are tapped together by default result in the loss of the sonde and so would compromise the completion of the Windsond objectives. We were planning to perform more intercomparison flights toward the end of the campaign unfortunately, as quoted in section 6.2 the radio receiver has been damaged during the campaign preventing us from completing this final objective.

Due to limited human resources (5 scientists separated into 2 teams performing 12 hours shifts on-site to run the whole Kumasi supersite instrumentation), distributed flights over day and night were not possible, thus to comply with the DACCIWA objective the frequent radiosounding program focussed on night-time measurements as detailed on figure 2. It was advised by our Ghanaian partners to avoid going out of the supersite at night to avoid encounters with tropical wilderness and limit the robbery risk. For these reasons, sonde recovery took place after sunrise, thus, to limit the time between the launch of the sonde and the recovery, all the frequent radiosounding flights took place in the last part of the night and the intercomparison flights at 0600 UTC.

According to Sparv Embedded, the Windsond S1H2 system is reusable, requires a smaller balloon and less helium and can receive multiple sonde, for the first utilisation of the sonde these features have to be tested. The S1H2 sonde were launched using manufactuer recommendations to evaluate the performance of the sonde in regular use. At low wind speeds the signal-to-noise ratio in the wind speed measurement is worse, so reducing the ascent speed can adversely affect the wind speed accuracy. For these reasons the sonde were not tied together for the low-altitude comparison flights. We agree that this experimental design does not allow us to quantitatively asses the Windsond S1H2 performance, however, this design allows us to qualitatively asses the system where the wind speed and direction issues have been confirmed and evaluate an eventual data alteration trend through use.

Eleven RS41-SG have been launched simultaneously with two S1H2 during the re-use evaluation, so twenty-two S1H2 flights have been compared to eleven RS41-SG in the boundary layer. This method allowed us more data comparison and also less travel time to search the sondes as the two S1H2 launched simultaneously where landing in the same area.

Page 3 Please us UTC or LT but not AM / PM Is Fig. 2 really needed ? Please explain what you mean with simultaneous launched (see above). Please give more information about the calibration of the Windsond. Do sondes have individual factory calibrations stored on the sonde or does the manufacturer rely on the quality of its sensors only ? How is the multi sonde reception realized – please give details on the receiver technology. Please use Kelvin instead of _C for Accuracy and Resolution in Table 2

Answer: Following reviewer #1 recommendation all the times are now changed to UTC. Individual calibration is done by the sensor manufacturers, thus stored in the sensors, multi sonde reception is realized by time-division multiplexing. As requested temperatures are now expressed in Kelvin in Table 2.

Table 3-5 Anders Petersson is affiliated to the manufacturer of Windsond. You should be able to give detailed information about the sensors used in Windsond and their performance instead of "not available (to be assessed)". Is the given value for pressure accuracy valid for Vaisala or Windsond? Why is the Wind speed accuracy relative to the wind speed?

Answer: The cell alignment for the table 2 leads to a confusion for the pressure accuracy, the given accuracy value is only valid for Windsond, the Vasaila value is defined as combined uncertaincy and reproducibility. At low-speed the signal-to-noise ratio in the wind speed is worse, thus the dependency of wind speed accuracy relative to wind speed.

Page 5: Pressure sections: Please include in the discussion the results of the WMO radiosonde intercomparison 2010 about direct pressure measurements vs. derived pressure.

Answer: We have added: "The WMO radiosonde intercomparison experiment 2010 showed that pressure measurement derived from geopotential heights and radiosonde measurements of temperature and relative humidity profile were very reproducible and suitable for all radiosounding operations for system where GPS system are set up correctly which includes the Vaisala system. This shows that the Vaisala derived pressure is a reliable reference to assess the Windsond pressure sensor, and the Windsond cost can be lowered by removing the pressure sensor in future version of the Windsond system depending on its GPS system accuracy."

Page 6: Please explain uncorrected data vs data correction for all parameters for the Windsond. What was the procedure to "adjust" ground pressure altitude and temperature? Wow large were these "adjustments" and why was this not done for humidity? I am still astonished that only one tandem flight to higher altitudes was performed! A larger number of such flights under different weather situations as well as during day and night would have improved the evaluation significantly. The flight was in 2016 and not 2006. Since all flights were performed during night or early morning radiative effects cannot be evaluated. Experimental design needs to be explained in more detail.
What was the length of the line connecting the sondes to the balloon? How did you tape the sondes together? Is it excluded that waste heat of one sonde influenced the other? Why did you set the Windsond acquisition to 3 seconds - according to table 1 the measurement cycle is 1s for both sondes. Please give details about the weather situation.

Answer: As mentioned in section 4 the Windsond S1H2 firmware has a single operational mode and produces uncorrected data, the only correction applied was to simply differentiate the ground value from the ground-based instrumentation and apply this difference to the profile. The altitude correction was in the [-10; +10] m range, pressure correction was [-2;+2] hPa, and temperature [-2;+2] range. In the stable nocturnal boundary layer, surrounding vegetation is expected to affect the local humidity values while having a limited effect on the temperature, for this reason humidity correction were not performed.

We agree that a larger number of flights would have improved the evaluation, however, the goal of this evaluation was to assess the quality of the Windsond data in context of the DACCIWA data analysis framework and also to address limitations of the Windsond system in the tested configuration. The radiative effect could not be be evaluated but needs to be evaluated for future use of the system. The flight was in 2016 we have corrected the typo.

The line connecting the sonde to the balloon was 20 m and the sonde were taped together making sure that the temperature and humidity sensors of each sonde were not interfering or influenced each other. At the time of the test, 1 s sampling rate was not compatible with the version of the firmware tested this is now available in the new version.

5.1.2 Should be renamed to Signal processing for low altitudes – Boundary layer higght was not detected - I would expect a boundary layer height around 100 m at the launch time of the sonde rising up to 1500 m during the day in this region during the monsoon period.

Answer: Agreed the title has been modified

Page 7: Can you explain why you have chosen different ascent rates and non-attached sondes for evaluating the reproducibility? I can't see any sense in this procedure since natural atmospheric variation will be at least in the same range as the instrument error. Profile comparison – It would be nice to have a profile plot if you do profile comparisons! Instead of showing scatter plots it would make much more sense to plot vertical profiles of PTH as well as wind for both sondes with an additional profile showing the vertical profile of the difference (Vaisala – Windsond) for each parameter together with the accuracy as stated by the manufacturer's datasheet. This would allow a meteorological interpretation. How do you measure cloud top temperature above the cloud top – The RS41-SG sensors are detecting the cloud top temperature and humidity before the S1H2 : : :..????

Answer: As mentioned in our answer concerning the page 2 comment we have chosen a different ascent rate in order to test the accuracy of the sonde by following constructor recommendation, test the multisonde reception and recovery system, and to increase the number of sonde tested.

We have chosen to directly compare each variable as the altitude error on the Windsond S1H2 would have superposed on each sensor error and not each sensor performance. Moreover, the Vaisala system does not have a pressure sensor and the pressure is calculated by the MW41 as detailed in section 3.3 while the Windsond has a pressure sensor so the profiles as a function of pressure would display data as a function of a calculated variable in one hand to a measured variable on the other hand. As discussed in the WMO intercomparison, basic raw data are to diagnose problems with a radiosonde during evaluation. We could also display both sonde data as a function of the Vaisala altitude, but this will involve modifying the shape of the S1H2 profile to fit in the Vaisala altitude profile and consequently not display the actual profile obtained with the S1H2 system. We consequently think that displaying the data of the Vaiala sonde as the function of the

Windsond data is the best way to assess the performance of each sensor without interference from other sensor errors.

The structure of this sentence was confusing and has been changed to: “For both temperature and relative humidity, the RS41-SG sensors are detecting the sudden temperature and humidity changes consecutive of the top of a cloud before the S1H2 sensors”

Page 8: Change reply time to response time
The atmosphere is characterized by vertical inhomogeneities, inversions and clouds – radiosondes therefore have to have sensors with low response time and neglectable hysteresis – if this is not the case the sonde is simply not suitable as radiosonde – or only for nice weather well mixed cloud free boundary layer.

Answer: Following reviewer #1 comments answer time has been replaced by response time.

The current response time limitation is the weakness of the system for boundary layer applications. In small-scale, Sparv Embedded uses temperature and relative humidity sensors with a better response time, but currently, the cost is high in the context of radiosondes. Lowering the per-unit cost would take a sizeable investment in the production process to automate assembly and calibration.

Page 9: More recent versions of the Windsond firmware certainly correct the altitude bias - have you checked this? Is it possible to reprocess the measurements to verify? To me it is not shown that newer firmware versions correct the altitude bias. The conclusions are too favourable – Windsond cannot handle inhomogeneities due to the high response time of the sensors, GPS derived wind error is far above the 5% error given by the manufacturer and to my opinion useless. It is not shown that at least the altitude correction in the latest versions of Windsond improve the systematic altitude error. As the WMO intercomparison results and the Vaisala sonde show pressure sensors are not needed any more for radiosondes – the “robust performance“ of the pressure sensor us unfortunately only of minor importance.

Answer: We agree the word certainly was too favourable for something we have not tested, this has been corrected to probably. We have added the statement in the conclusion : “These limitations make the deployment of an operational network using this system under the tested configuration impossible.”

The WMO intercomparison shows that pressure sensor are not needed anymore for radiosondes in situation where GPS radiosondes are set up correctly which is not the case of the Windsond. Thus the evaluation of the pressure sensor is important to assess the use of the Windsond S1H2 data in a meteorological context such as tephigrams.

Page 10:
The experimental design shows several weaknesses – as already addressed the fact that the sondes were not tied together during the ascends makes it nearly impossible to

separate instrument error and atmospheric variance. I would recommend to test each sonde prior nest launch instead of a simple visual inspection.

Answer: In the morning stable boundary layer horizontal variations are small so the instrument error would still be significant compared to the atmospheric variance. We agree, however, that atmospheric variance will add some noise to the error recorded between sondes. Despite these weakness this design allows us to show that there is no clear trend in data alteration consecutive with sonde re-use during the experiment.

We agree that testing sonde prior the next launch should be the standard, however limited ressources made a detailed inspection impossible. This standard is unfortunately not always respected as Legain et al,. 2013 also relaunched a large number of sonde immediately after recovery.

I would strongly recommend to perform additional measurements with a larger number of sondes under laboratory conditions to determine sensor accuracy and inertia over a wide range of temperature and humidity and to compare the results to the sondes datasheets first. Reproducibility can also better be tested in a combination of repeated tandem flights and climate chamber measurements – this would allow the separation of sensor degradation and atmospheric influence in real atmospheric conditions.

Answer: These are interesting comments, but these go beyond the scope of this paper, here we test the performance of the sonde during a field campaign which is similar to Legain et al., 2013 where the system was tested during 2 field campaigns. The sonde recovery system cannot be tested in an atmospheric chamber as well as the different natural hazard encountered in a rough environment such as West Africa. Despite the different limitations of the experimental design we have been able to identify some limitations of the system especially for the GPS system which are worth publishing to provide indication to Sparv Embedded and future users of the system.

Table 6 is unreadable – it extends 4 pages – please consider a condensed way of presentation.

Answer: Agreed, we have substituted table 6 with figure 8 that condenses all the information of table 6

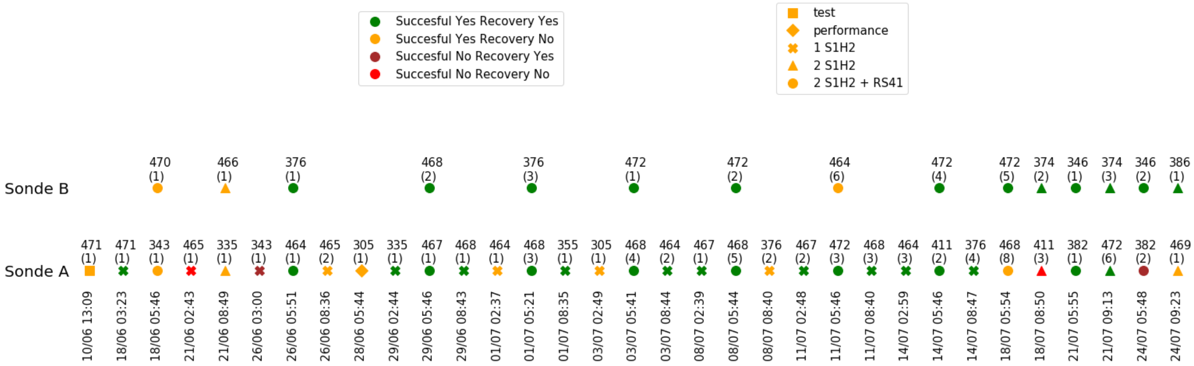

Figure 1 Timeline listing sounding time in UTC, the shapes indicate the corresponding number of radiosonde S1H2 launched (test denotes the test sonde, performance denotes the S1H2 launched taped to an RS41-SG, +RS41 denotes simultaneous launched with the Kumasi Agromet supersite), the sonde id with the number of time the sonde has been used under brackets, the colors indicates flight result and the recovery result.

Page 11:
Please give the percentage of unsuccessful flights and flights with sondes that did not cut off. Is the number of data from sondes that did not cut off large enough to do a representative evaluation for altitudes between 650 and 1000 m?
It is nearly impossible to separate the different markers in Fig. 8. Maybe separated figures would help.

Answer: Only 3 flights with failed cut-off were launched simultaneously with an RS41 so an evaluation between 650 and 1000 m would not lead to a statistically representative evaluation of the Windsond system. This evaluation would also be beyond the scope of this evaluation which focusses on how well the low-level cloud and low-level jets are represented in this study.

We agree that on figure 8 the markers are nearly impossible to separate, but separated figure would increase the number of figures without leading to more interesting conclusions.

As you have a large number of flights over several days available I would recommend to do not only a statistical analysis based on scatter plots and regressions but also a more meteorological where you create classes of different weather situations e.g. with and without low level clouds and analyse the behaviour of the sondes along the vertical profile.

Answer: During the DACCIWA field campaign the low-level cloud was a recurrent feature and only one IOP night was identified without a low-level cloud so a meteorological climatology based analysis would not be relevant.

Fig. 9 – why do you use lines to connect the markers?

Answer: We decided to use lines to connect the marker to help the reader to see that there is no real trend between sonde usage and data alteration.

Page 12: A check of the sonde sensors before reusing it should be the standard procedure – see my comment to page 10.

Answer: Agreed this should be a standard but field campaign constraints can limit the time on the sonde check to avoid the radiosonding program to compromise other instruments objectives.

A system for low altitude rapid soundings using high quality radiosondes was already introduced and tested by Legain et al. 2013. The questions to me is if a low cost and unfortunately low-quality system like the Windsond really makes sense with all the weaknesses we have seen in your evaluation especially - considering the fact that higher quality sondes can also be recovered and reused so that the cost difference between the sondes gets even less important. Please discuss!

Answer: As remarked by the referee #1 our conclusions shows that the Windsond S1H2 is a work in progress and the result presented in this manuscript are introducing a new instrument in the early stage of its development. The system presented by Legain et al 2013 does not asses the consequences of the drag generated by the protection cage on the measured wind profile. Moreover, the modification applied to the sonde system required qualified personal and can limit the generalised used of this sonde.

A key point is the Windsond system does not requires any modification or complex balloon system so if that system become accurate enough it will provide a easy to use solution in countries with limited to deploy a radiosounding network utilising the more accurate but more expensive sondes. The multi sonde capability is also another key point for field campaigns were multiple shallow sounding are required. An example of this application is the VORTEX-SE project, where Penn State University released 24 sondes at the same time to study winds around storm supercells and might release as many as 100 at a time in the next season. This is a unique feature of Windsond for dense measurements (http://windsond.com/swarmsonde-is-in-the-news/).

Please change longer answer time to response time!

Change from answer time to response time have been made following remarks from reviewer #1.

It would be nice to know if newer Windsond firmware really has corrected the problems with GPS derived wind and pressure. I therefore strongly encourage you to perform the further performance evaluations and include their results into a revised version of your manuscript.

Answer: We agree that it would be interesting but this analysis would go beyond the scope of this paper and would require more time and capabilities to perform this study.