AR2 Reply

General comments:

- The title of the manuscript is a "stand-alone calibration approach". But why is an in-flight calibration actually necessary? What is the robustness, especially compared to dedicated calibration measurements? In general, the calibration flights that are presented are not well described. A time series of wind speeds during the calibration would be necessary to judge stationarity of the flow during calibration. What are the limits of the calibration approach in terms of wind speed and turbulence conditions? What are the expected uncertainties?

There are different approaches to calibrate a multicopter for wind measurements. Using a wind tunnel to map the tilt of the copter at different flow speeds is an option. Still, of course, it requires a wind tunnel, which is very expensive, and sometimes it has a long waiting list for availability. Moreover, for medium size copters (like the S900 we used), one would need a vast test chamber to avoid disturbances from the wall, especially at low speeds. Another option is to hover next to a sonic anemometer out in the field. It might be a very accurate way to build a model since one already has a precise reference (the sonic). However, it is impossible to control the atmospheric wind. Therefore creating a model covering all possible wind speeds would be a very complex task.
With the method proposed in this study, it is possible to map the complete range of flight speeds at a very low cost with comparable results in terms of accuracy. (I added some lines to clarify these points in the introduction and Section 2.2.5.)

Unfortunately, a time series of the wind speed at 50 m altitude is unavailable at the moment of the calibration flights. However, we built our method so that no other instrument is needed. We can estimate the wind speed with equations 28 and 29 for each forward and backward flight. The duration of a single forward and backward path is approximately 2.5 minutes. We believe the average wind can be considered stationary over this time window, and small-scale turbulence can be filtered out by taking the mean value of the tilt angle. (See also comment below about surface wind speeds).

The limit in terms of average wind speed for the calibration flights is when the copter can not hold its GS anymore under the effect of the external wind. In our case, the S900 can fly safely up to 20 ms-1. Then our limit in terms of wind speed is 6ms-1 if we want to calibrate up to 14ms-1. This calculation does not consider possible gusts, but it is a basic example to explain the concept. For turbulence, it is difficult to identify a precise threshold. It is more a matter of flight safety and whether the autopilot can stabilize the flight under gusty conditions.

The main uncertainties of our method lie in the steady equilibrium hypothesis made in the model formulation. We assume that the copter tilt angle represents the actual wind speed at any time, but this neglects any UAS transient. (Added section 5.6).

- How does this calibration approach differ from a calibration approach that is described for the PX4 autopilot:https://docs.px4.io/v1.9.0/en/advanced_config/tuning_the_ecl_ekf.html#mc_wind_estimation_using_drag ?

The calibration proposed by this library does not take into account wind disturbances. Moreover, after the complete procedure, the ballistic coefficient is set to be a single value. So, it is assumed that this value does not change with the Reynolds number or tilt angle. Our dataset shows that the terms that compose the ballistic coefficient change quite significantly over the range of flight speed of the multicopter.

- *Does the sphere really create rotationally symmetric flow around the multicopter? The rotors are a very important part of the aerodynamic features of the UAS and probably have more effect than the frame shape. I would at least have expected a graph showing the calibration error versus flow direction. There is probably not enough data there for this purpose, but I do not think that any conclusive statement can be made without such tests.*

We recently performed a test in order to prove this. The result is now shown in section 5.1.

Specific comments:

p.1, l. 7: *isotropy the right word? I do not think it should be used for the shape, but for the attributes of an object / material. Example: a perfect wooden shpere is anisotropic, because its material properties like shear stresses etc are different depending on the grain direction.*
We changed isotropy with symmetry.

p.1, l.17: *sonic anemometers do not need wind vanes, because they provide at least the two-dimensional wind vector.*
Yes, we agree, we changed the sentence, the order of the words was wrong and caused confusion.

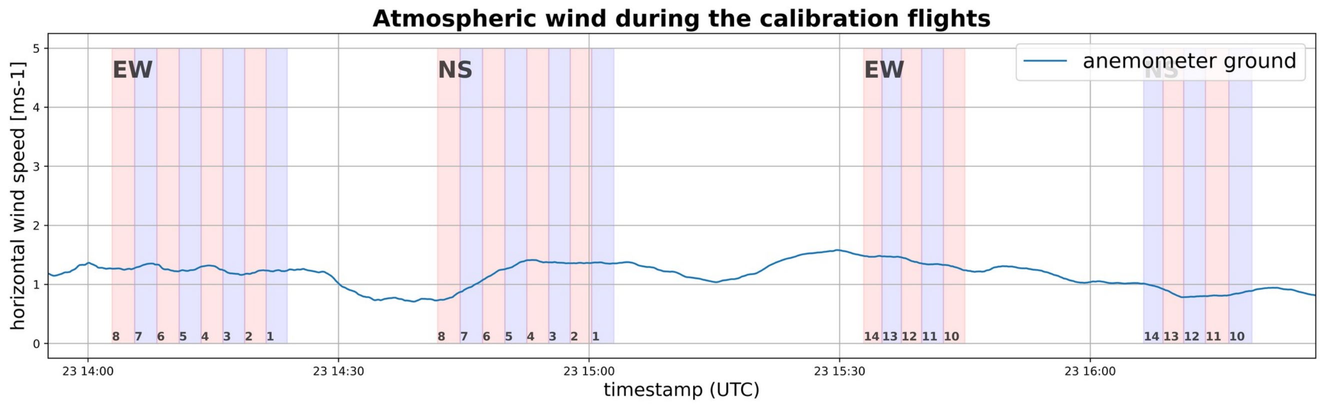
p.2, l.47: *"linear behaviour" I am not sure what is meant by a linear behaviour of the parameter and I do not believe that Wetz et al. do assume a linear behaviour of the drag equation.*
They do not assume linear behavior of the drag equation, however they model the UAS drag coefficient using a linear model: Cd A = Cd0 A0 + cp * tilt. Instead of “parameter” we will use again “coefficient” to be clearer.

p.2, l.50: *I would prefer "symmetry" over isotropy, see above.*
Changed.

p.2, l.56: *multicopter*
We changed all multi-copter to multicopter throughout the paper.

p.7, l.165: *PosHold mode is not explained.*
We added a short explanation similar to what mentioned in the official ArduPilot documentation.

p.8, l.186: *Were surface wind speeds measured? If so, where and how?*
We had an anemometer placed close to the ground station and measured for the whole time of the missions. The picture below shows the 10min average wind speed and includes the mission sequence.



From this picture the wind is constant during the single forward and backward leg for each GS.

p.9, l.206f: *"disturb the UAS with the same magnitude in an opposite way..." This assumes that wind is constant during the flight, but during an afternoon with low wind speeds, radiation input and probably turbulence, this assumption can be violated significantly.*
As shown in the picture in the previous comment, the wind at the ground is constant along the same GS forward and backward leg. We are not assuming a constant wind during the whole 20 minutes flight but only along a single forward and backward pair of legs. A forward and backward test lasts approximately 2.5 minutes, as seen in Figure 5, which allows assuming a constant wind speed during this time window.

Fig.4: *There seems to be no scale or information about the position and extent of the map.*
We updated the figure adding the lat lon ticks.

p.11, l.222: *what is 8-1 ms^{-1}?*
It is the range of velocities mapped during the first mission, we have changed the content of the parentheses to be clearer.

p.12, l.239f: *I do not understand this statement "around 10 for 1 ms^{-1}". 10 what?*
It refers to the Ca coefficient. The coefficient has no unit, being dimesionless.

p.13, l.247f: *"wind to be constant". It would be good if you showed this with independent measurements. Turbulence occurs on a wide range of scales and can significantly disturb scales within the flight time.*
See comments **p.8, l.186, p.9, l.206f**

Table 3: *I am much concerned about the comparison of ERA 5 data with a grid size of 9km and local measurements in quite complex, heterogeneous terrain as in Poltringen. Can it really be expected to match?*
Our intention is that the ERA 5 data gives an overview about the synoptic situation including the wind speed and direction during the measurements flights. In addition, the terrain around Poltringen is fairly flat without obstacles around it, so we assume the ERA5 wind gives a solid estimate of the wind situation in Poltringen.

p.14, l.269: *but is ERA5 a good comparison?*
We do not want to use the ERA5 data to compare the copter measurements directly, but more as an indicator that we had indeed a low wind situation at the beginning while higher mean wind for the last two missions (see also comment above).

p.14, l.278 and p.15, l.291: What does this RMSE represent?
It is the RMSE value between the fit function and the original data. We opted for RMSE instead of the more common R squared since the drag coefficient model has been fitted by using a non-linear function.

p.16, l.306: *every sensor can only resolve signals up to half of the sampling frequency according to the Nyquist theory.*
The sentence was unclear, the point here is that we have a reference that provides valid data up to 10Hz which is exactly the same frequency our autopilot logs data. We changed the sentence to be clearer.

p.16, l.313: *but the time lag depends on the wind speed. Has this been considered?*
The time lag here has been calculated for each hovering section. It represents an averaged time lag over the whole time of hovering next to the sonic. In some cases, the wind speed changes sensibly; however,

we consider the usage of this lag conservative enough in terms of data quality. In other words, if we would calculate the real time lag for each instant, we would most likely get a better result in terms of RMSE and not worse.

Sec.4.1: *As the text says, turbulence is out of scope of the work. I that context I wonder what the section and Fig. 9 add to the goals of the manuscript. Maybe they can be removed. If it is only used to show that there is noise above 0.2 Hz and data needs to be rejected above that frequency, this could be mentioned early in the data processing.*

These plots were extremely useful for us in the post-processing procedure, since only thanks to this analysis we could correctly chose the resampling frequency for our data. Therefore we suggest to keep this Section and Figures.

p.20, l.356ff: *I have doubts about these opinions and theories. I think they should not be presented without evidence, because it could confuse readers. A simple explanation for a plateau followed by a steep decay often is a noise level in the dynamics (e.g. by vibrations) and low-pass filters in the senors at higher frequencies (here the Kalman Filter of the angle solution).*

This section has been removed since we have not enough data to support our statements

p.20, l.370: *It is unclear how MBE and RMSE are calculated. What are the considered averaging periods. What is the cause of the MBE?*

The MBE and RMSE reported in table 5 are calculated using all the data we gathered while hovering close to the sonic. Of course, the parameters are computed on the 0.2Hz resampled data. Also, as mentioned in the text, the RMSE is calculated by subtracting the MBE between the two signals.

The cause of MBE is challenging to identify. Also, previous literature does not explain why a bias in the real environment is still present after proper calibration. Regarding the direct model, any change in multicopter payload, or air density might cause this bias. On the other hand, for the Drag coefficient model, the bias caused by the air density is not a problem anymore. In contrast, the one caused by the mass is still partially an issue. In the end, the slightly higher bias for the Ca model might result from a non-perfect fit of the Ca points by the exponential decay.

Sect. 5.5: *I do not understand why there is a subsection "vertical velocity and mass" and then subsections for each of those terms on the same level.*

We changed the structure of the last sections so that they are not more on the same level.