

Reviewer:1

We thank Reviewer 1 for his/her constructive comments.
Our responses are given below in red.

Responses to the major comments:

1. The authors conducted two verification of a case study and a statistical examination. The case study showed that the RASTA assimilation improved wind and rainfall fields, and the 3-h assimilation looked best. On the other hand, the statistical examination for the entire domain in Figure 1 illustrated that the RASTA assimilation mostly had no impact even on the wind field, and only the 1-h assimilation has some skill in rainfall forecasts. Since the RASTA data is limited in cloud region, the assimilation impact is also limited in time and space. I suggest that the statistical examination is re-conducted over a limited area, for instance, the Figure 2 area, or convective-system-related area, or RASTA-related area, since the inconsistent results between the case study and the statistical examination makes the readers confused.

The authors are grateful to Reviewer 1 for his comment because now we have a better consistency between the case study and the statistical examination. The statistical examination has been re conducted over a RASTA-limited area. This area contains the aircraft flight path $\pm 0.5^\circ$ both in longitude and latitude. The RASTA-limited validation domain is larger than the exact flight path because the increments are advected as the forecast term increases. The text has been modified in section 6.1.

The comparison against conventional observations indicates similar results (see section 6.1 and Figure 9 of the revised version): generally the impact is slightly negative to slightly positive. Besides, the differences are less than 0.5 m/s, so the impact is neutral.

The methodology employed to compute the scores against rain gauge measurements has been modified. In the RASTA-limited validation area, observations and model outputs are first averaged in boxes of $0.25^\circ \times 0.25^\circ$, and then concatenated over the 35 assimilation cases. Bootstrap confidence intervals are calculated with these new sets of observations/model outputs. To avoid the spin-up problem, the first hour of rainfall accumulation has also been removed from the calculations.

- The new results are more consistent with the case study: the best scores are reached with the largest assimilation windows (2h or 3h) and the most significant differences appear with the RASTA_3h and RASTA_2h experiments.
- Generally, the impact is slightly positive to neutral. The use of the smallest assimilation window leads to the most neutral impact, which is also consistent with the IOP7a case study.
- In the previous version of the paper, the differences between the CTRL experiment and the RASTA experiments appeared above approximately 25 mm. Now in Figure 9 (Figure 10 in the revised version) we can see differences above 10 mm.
- Figure 9 has been modified (Figure 10 in the revised version), together with the text in sections 6.2, 7 and in the abstract.

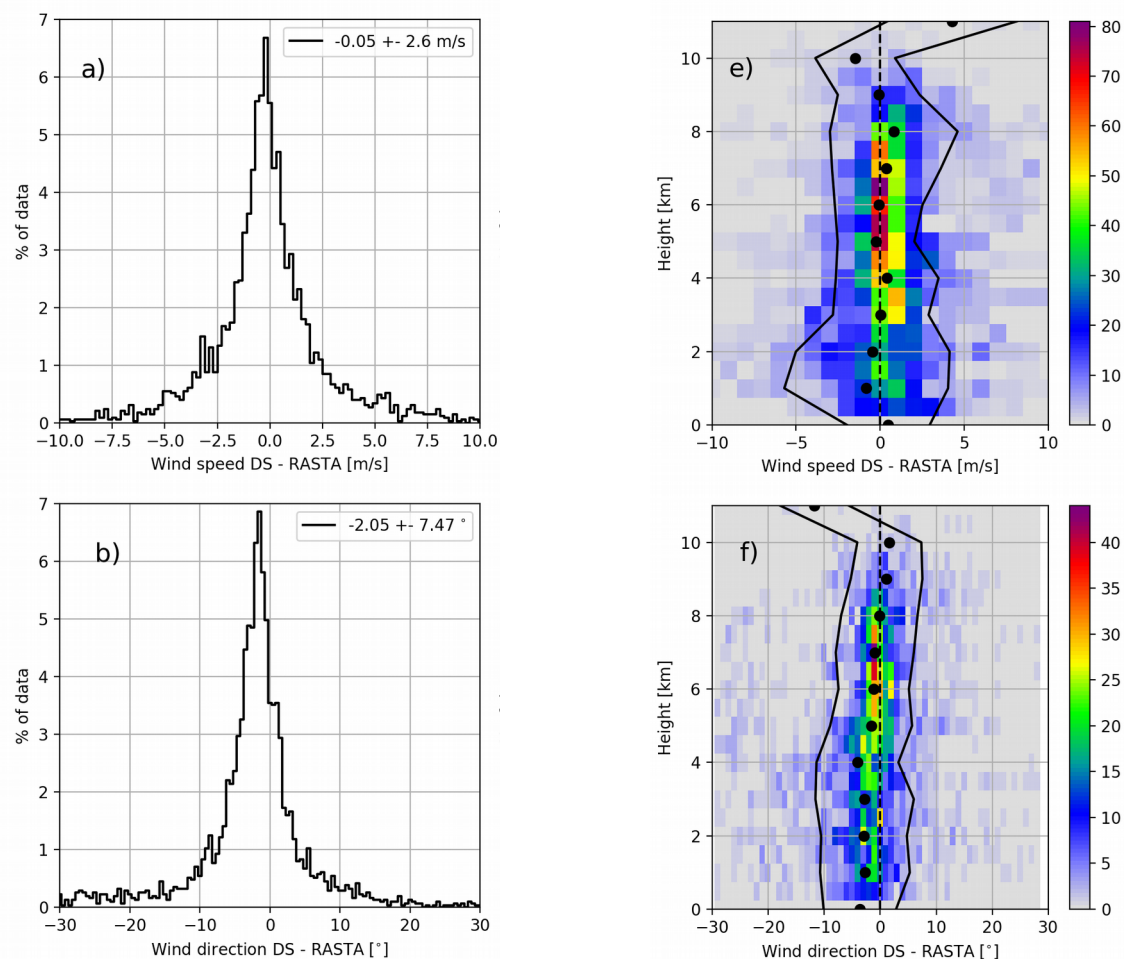
2. I agree that Figure 6 implies a spin-up problem in forecast. For the reason of the spin-up, I doubt that the observational error of RASTA use in the present study would be smaller than appropriate value because it is the same with that of radiosondes. The error should be larger since RASTA includes much more sources of errors than radiosondes include.

We are not sure that the observation error should be larger than the one used for radiosondes.

Indeed, RASTA wind data during the HyMeX-SOP1 field campaign have been compared against ground-based Doppler radar by Bousquet et al. (2016). Results of their study show that “The low values of the bias error suggest that errors are close to Multiple-Doppler wind synthesis and should remain comprised between 1 and 1.5m/s” (see section 3.2, page 93).

These values are smaller than the radiosonde ones (between 1.8 and 2.52 m/s). We added these values in section 4.2. “Bousquet et al. (2016) demonstrated that the bias error of RASTA wind data is comprised between 1 and 1.5 ms^{-1} . In this study, it has been decided to use the same observation error as the one used for radiosondes, which increases with the altitude (from $\approx 1.8 \text{ ms}^{-1}$ at 900 hPa to $\approx 2.52 \text{ ms}^{-1}$ at 200 hPa).”

RASTA wind data have also been evaluated during the NAWDEX field campaign which occurred in Iceland (<http://www.pa.op.dlr.de/nawdex/>). In the following figures, RASTA wind retrieval were compared against radiosonde measurements. These Figures demonstrate that the observational error for RASTA wind data is of the same order of magnitude as that of radiosondes.



Bousquet, O. , Delanoë, J. and Bielli, S. (2016), Evaluation of 3D wind observations inferred from the analysis of airborne and ground-based radars during HyMeX SOP-1. Q.J.R. Meteorol. Soc., 142: 86-94. doi:[10.1002/qj.2710](https://doi.org/10.1002/qj.2710)

3. I suggest that Figure 5, and explanations for Figures 4 and 5 will be modified. The increment in Fig. 5A is reflected the flight path of all observations, thus, all data points assimilated in this 3-h window should be presented in Fig. 5A. Moreover, the 1-h, 2-h, and 3-h assimilation window experiments include the observation until 0630, 0700, and 0730 UTC, respectively (L10 P7). I think that this different time limitations create the difference between panels in Figure 4 unlike the authors explanation on overpasses (L6-17 P8). Please exam and discuss this point of view.

Figure 5 (now Figure 6) has been modified: All the data that are assimilated in the RASTA_3h experiment are now shown. The explanations have also been modified. Fig 5a is first described in section 5.2:

“Figure 5A represents the wind speed increments at approximately 4 km of altitude (model level 30) between the RASTA_3h and the CTRL analysis. Wind directions are also indicated by the green (resp. black) arrows for the CTRL (resp. RASTA_3h) analysis. The data points assimilated in the RASTA_3h experiment until 07:30 UTC are also represented by the black data points.”

Then, Fig.5 B-D are explained at the beginning of section 5.2: *“Figure 5 (panels B to D) represents the wind speed differences of the RASTA_3 h 1-, 2- and 3-h forecasts and the CTRL ones. At each forecast term, the black data points indicate the different RASTA locations which are available during a 1-h time window centred on the forecast time (forecast term $\pm$ 30 minutes).”*

Reviewer 1 is correct, the different time limitations explain the differences in wind and direction in Figure 4 (now Figure 5). We added this explanation in the text.

4. It is amazing for me that RASTA_3h in Figure 4 improved the wind field even at the end of the assimilation window because the experiment did not employ FGAT. Since the RASTA data only exist in cloud area, 3 hours seems too long to assimilate the data appropriately. I understand that this is the motivation of the authors to conduct three experiments. If they use FGAT, the 3-h experiment may significantly improve the result. I recommend the authors to conduct the FGAT experiment additionally if possible.

Reviewer 1 is right, FGAT is a way to improve the handling of the time dimension in a 3D-Var scheme as it allows to compute the innovations (i.e. the observation-guess differences) at the time of the observations for different times during the assimilation window. For the AROME model, the FGAT option has been evaluated by Brousseau (2012) for moving platforms, but without any positive improvement in the subsequent forecasts (Brousseau et al. 2016, section 2). For observations from static platforms, the 3DVar without FGAT only uses the observations performed at the middle of the assimilation window. The FGAT option allows to estimate innovations for sub-hourly data from the same instrument at the same location. More observations are assimilated, but the 3D-Var minimisation, without time dimension, uses these several innovations at the middle of the assimilation window. This leads to an averaging and a smoothing effect on these observations and a loss of information on the temporal details, which is not desirable in a convective DA system. Therefore, in this study we decided to use conventional 3DVar to assimilate all the different kinds of observations in the same way.

Brousseau, 2012: Propagation of observed information into the AROME data assimilation and atmospheric model, PhD thesis, Université de Toulouse III – Paul Sabatier

Brousseau, P. , Seity, Y. , Ricard, D. and Léger, J. (2016), Improvement of the forecast of convective activity from the AROME-France system. Q.J.R. Meteorol. Soc., 142: 2231-2243. doi:[10.1002/qj.2822](https://doi.org/10.1002/qj.2822)

5. The authors used the median value of observations in a grid box (L16 P5) for thinning. If observational data distribute followed the Gaussian PDF and their number are large enough, the median and the mean values are the same. Usually “super observations” are made by the “mean” method in order to reduce representativeness errors and avoid noises. Therefore, the authors should explain why they adopt the “median” method instead of the mean.

The two approaches have been tested by the authors. After the data processing described in section 2.1 (whose description has been enhanced in the revised version of the paper), some spurious data were still occasionally present. Using a median filter, instead of the mean filter, helps to reduce the weight that these spurious observations can have when we calculate the “Super-observations”. Besides, a median filter is also employed by Bousquet et al. (2016) and by Tabary et al. (2006) to calculate the “super observations” of ground-based radar Doppler velocity observations

Tabary, P., F. Guibert, L. Perier, and J. Parent-du-Chatelet, 2006: [An Operational Triple-PRT Doppler Scheme for the French Radar Network](https://doi.org/10.1175/JTECH1923.1). *J. Atmos. Oceanic Technol.*, **23**, 1645–1656, <https://doi.org/10.1175/JTECH1923.1>

Bousquet, O. , Delanoë, J. and Bielli, S. (2016), Evaluation of 3D wind observations inferred from the analysis of airborne and ground-based radars during HyMeX SOP-1. *Q.J.R. Meteorol. Soc.*, 142: 86-94. doi:[10.1002/qj.2710](https://doi.org/10.1002/qj.2710)

6. English needs to be proofread by professional native speaker(s) with scientific background

The revised manuscript has been carefully copy-edited for English. Together with the copy-editing standard service offered by Copernicus, we believe that the English should be sufficiently polished in the final version of our manuscript.

Responses to the minor comments:

L23 P1: “To fill the gap in clear air condition” I suggest the authors to refer the following articles, because wind observations in clear air can be also provided by Doppler lidars (air-born and ground-based), and clear air echoes (insects) by Doppler radars.

[Ground-based lidar] Kawabata, T., H. Iwai, H. Seko, Y. Shoji, K. Saito, S. Ishii, and K. Mizutani, 2014: Cloud-Resolving 4D-Var Assimilation of Doppler Wind Lidar Data on a Meso-Gamma-Scale Convective System. *Mon. Wea. Rev.*, 142, 4484–4498, doi: 10.1175/MWR-D-13-00362.1.

[Air-born lidar] Weissmann, M., R. H. Langland, C. Cardinali, P. M. Pauley, and S. Rahm, 2012: Influence of airborne Doppler wind lidar profiles near Typhoon Sinlaku on ECMWF and NOGAPS forecasts. *Quart. J. Roy. Meteor. Soc.*, 138, 118–130, doi:10.1002/qj.896.

[Clear air echoes] Kawabata, T., H. Seko, K. Saito, T. Kuroda, K. Tamiya, T. Tsuyuki, Y. Honda, and Y. Wakazuki, 2007: An assimilation and forecasting experiment of the Nerima heavy rainfall with a cloud-resolving nonhydrostatic 4-dimensional variational data assimilation system. *J. Meteor. Soc. Japan*, 85, 255–276, doi:10.2151/jmsj.85.255.

The authors are grateful to Reviewer 1 for these references. We now refer to the suggested articles from L 23: “In clear air condition, wind observations can be provided by insect-derived Doppler radar measurements (Kawabata et al., 2007; Rennie et al., 2011) or by Doppler lidars (Weissmann et al., 2012; Kawabata et al., 2014).”

L19 P2: “has never been investigated” I did not understand what thing has never been investigated in the following “vertical profiles from Doppler W-band radar”. “vertical profiles from Doppler radar”? “W-band radar”? “vertical profiles” by W-band radar? (“horizontal” winds have been done)? Please clarify.

We meant “vertical profiles by W-band radar”. This has been rectified in the text (Doppler has been removed in the sentence).

L30 P2: “first” This is the same with the above. What is the first?

“First” means the assimilation of wind profiles measured by Doppler W-band radar. Since Doppler is redundant with “wind profiles”, we removed Doppler.

L8 P3: “HyMeX-SOP1” What is this? Spell out it and add explanation.

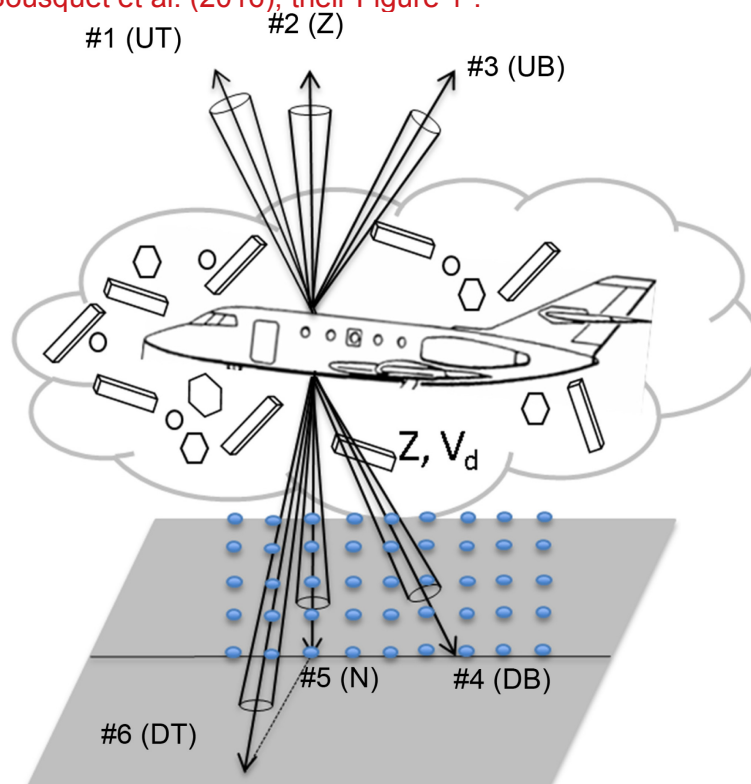
HyMeX (Hydrological cycle in the Mediterranean Experiment) aims at a better understanding, quantification and modelling of the hydrological cycle in the Mediterranean, with emphasis on the predictability and evolution of extreme weather events (Drobinski et al., 2014). The HyMeX first Special Observing Period (HyMeX-SOP1, Ducrocq et al., 2014) took place during a 2-month period during the autumn 2012. The main goal of the HyMeX-SOP1 was to document the heavy rainfall and flashflood events which regularly affect the mediterranean area.

We added some informations L35, page 2 about the HyMeX-SOP1. For further explanations, the reader can refer to Ducrocq et al., (2014).

L28 P3: “six Cassegrain antennas” How do these six antennas observe three directions above and below the aircraft? Add explanation and, if possible, a schematic figure.

RASTA configuration during the HyMeX-SOP1 is given by Bousquet et al. (2016) (Figure 1). The radar is equipped with 6 antennas pointing either upward (antennas 1-3) or downward (antennas 4-6). Labels 1-6 refer to the ‘upward transverse’ (UT), ‘zenith’ (Z), ‘upward backward’ (UB), ‘downward backward’ (DB), ‘nadir’ (N) and ‘downward transverse’ (DT) antennas, respectively.

In the text, we added “A schematic figure of RASTA configuration during the HyMeX-SOP1 is given by Bousquet et al. (2016), their Figure 1”.



L30 P3: “unambiguous distance” “unambiguous velocity” What are these? Observational range and available observations? But, in Figure 4, we see larger observations than 7.8 m/s.

Unambiguous distance is maximum range and unambiguous velocity is Nyquist velocity. The text has been rectified.

We see observations larger than 7.8 m/s because an unfolding algorithm is applied to the Doppler velocities. The algorithm is explained by Bousquet et al. (2016): “Radial velocities are processed by first removing the projection of aircraft ground speed along the six radar beams. Doppler observations are then unfolded by using an in situ wind sensor as a reference for the first valid gate and by applying a gate-to-gate correction for the next ones.”

We added more informations about RASTA data processing in section 2.1

L22 P4: “2.5 km x 2.5 km” Modify it to 2.5 x 2.5 km” and add the number of horizontal grids or the horizontal size of the domain.

The authors modified it, and added “It has 948 * 628 horizontal grid points, which is equivalent to a horizontal size of 2370*1570 km².”

L25 P4: “specially designed” What is the special in this study? Please clarify.

Please see comment n°2 of Reviewer 2. The AROME-WMED model was specifically designed for the HyMeX-SOP1 field campaign to support the instrument deployment. It is dedicated to the heavy precipitation events which regularly occur in the autumn.

The major differences between AROME and AROME-Wmed are:

- the AROME-Wmed domain has been extended and centred on our area of interest (the northwestern Mediterranean area)
- The background error covariance matrix
- The number of assimilated observations in the southern part of the domain.

We added more explanations in section 3.1

L7 P5: “GPS” Spell out it. GPS stands for Global Positioning System operated by U.S.A.. I guess the authors use other navigation satellite systems like Galileo and GLONASS. In this case, GPS should be replaced by “GNSS” (Global Navigation Satellite System).

The authors replaced GPS by GNSS.

L17-19 P5: “When the aircraft – removed from the interpolation.” It is hard to understand the situation and removed data. Did the authors remove the data only outside the grid box or the whole profile of the data? It should be better to show a schematic figure of the aircraft with the six radar antennas, and wind profiles in and out the grid boxes.

Only the data that are outside the grid box are removed. We added a schematic figure (Figure 2 in the new version) to explain this sentence.

L29 P6 and L10 P10 I suggest that the title of Section 5 and 6 as well as the examinations are named as “the case study” and “the statistical study” instead of IOP7a and HyMeX SOP1, respectively.

The titles of Sections 5 and 6 have been replaced by “Results on the case study” and “Statistical study”

L30, L31, L34 P9: “the maximum rainfall” Please show the exact maximum values in each experiment, not approximated values.

The exact maximum values are now shown in each panel.

L31-31 P10: “small number” From Figure 8, the numbers of observations are several thousands. These are not “small”.

We meant the number of observations in the area of interest. Following the suggestions of the major comment N1, Figure 8 (Figure 9 in the revised version) has been changed. The examinations are only conducted in RASTA-limited area (see major comment #1). Therefore, we only have hundreds of observations to evaluate the 3-h forecasts. Generally, the impact is neutral.

Figure 1 Add the explanation on the red box.

We added the following explanation in the caption: "The area surrounding the IOP7a case study is indicated by the red box."

Figure 3 It is helpful for the readers if the authors add the information on flight level in this figure, for instance, by changing the size of circles as height, or by replacing the circles with triangles or rectangular or cross-marks as height.

To add an information on flight level, a circle has been set if the aircraft is below an altitude of 4km, a square if the altitude is between 4 and 6 km, a star if the aircraft is between 6 and 8 km, and a triangle if the aircraft is above 8 km. This new information is now written in the figure caption (now Figure 4).

Figure 4 Add (a), (b), (c) and etc. or figure titles to each panel to refer it easier.

We modified figure 4 (Figure 5 in the revised version) by adding A to a E for the wind speed and F to J for the wind direction. We now refer to the to A, B, etc. in the text.

Figure 7 Add the maximum rainfall amount values to each panel.

The maximum rainfall amount values are now displayed in each panel (now Figure 8).

Figure 8 I suggest that Figure 8 will be illustrated by the difference between CTRL and others, not each profile, in addition to the examination on the limited area (see the major comment)

Figure 8 has been modified (now Figure 9). It now illustrates the differences between the standard deviation of (OBS - CTRL) and the standard deviation of (OBS - Rasta experiment) on the Rasta-limited area. The text has also been modified in section 6.1. The results are still neutral

Reviewer 2:

We thank Reviewer #2 for his/her constructive comments.

Our responses are given below in red.

Page 2, line 16: "...at each kilometer levels..." Please explain better what this means?

We replaced "at each kilometer levels" with "a high vertical resolution".

Page 2, line 35: What is different compared to the operational version? Just a brief, short explanation would be good.

The AROME-WMED model was specifically designed for the HyMeX-SOP1 field campaign to support the instrument deployment. The main differences between AROME and AROME-Wmed are:

- The domain: the AROME-Wmed domain is centred on the area of interest (northwestern Mediterranean area)
- The Background error covariance matrix, which has been calculated during an autumnal period in October 2010 characterised by heavy rainfall events.
- The number of assimilated observations: AROME-Wmed assimilates more observations in the southern part of the domain.

We added "specifically designed for the HyMeX-SOP1" in the introduction section (L6, P3). We also added the above informations in section 3.1

Page 3, line 30: 7.8 m/s is a rather low value for the unambiguous velocity. In fact this is one of the main challenges in Doppler wind assimilation. A de-aliasing, or unfolding algorithm can work fine for unfolding once but what if the wind speed is high enough to fold twice? Then it will be much more uncertain. Are the authors confident with the algorithm used and/or that there are no wind speed above this limit? Another complicating factor is that the aircraft is moving and this also needs to be taken into account in the unfolding. Perhaps it is outside the scope of the paper to discuss this in detail but a brief discussion about this is necessary since it is crucial when using the data.

In situ wind measurements at flight altitude are used to check the number of foldings for the first valid gate and then by applying a gate to gate correction for the next ones. The authors are confident with the algorithm because most of the time RASTA was collecting data in cloudy areas. In addition to that the combination of the three non colinear beams is used to verify potential unfolding issues as the retrieval would be locally inconsistent.

The exact speed of the aircraft and the pointing angles allow one to rigorously determine the component related to the aircraft's movement. The algorithm is further described by Bousquet et al. (2016). We added some more information in section 2.1.

Bousquet, O. , Delanoë, J. and Bielli, S. (2016), Evaluation of 3D wind observations inferred from the analysis of airborne and ground-based radars during HyMeX SOP-1. Q.J.R. Meteorol. Soc., 142: 86-94. doi:[10.1002/qj.2710](https://doi.org/10.1002/qj.2710)

Page 5, line 21: "every 3 time steps" What does this mean?

First, super-observations are calculated using the median filter (median value of all data available along the aircraft track within a box of 2.5 km length between the two half model levels surrounding each model level). After this median filter, a thinning is applied to these super-observations. It has been decided to select one super-observation out of three.

We modified the text in section 4.1 by:

"After this pre-processing, to satisfy assumptions about observation error covariances, which are supposed to be $0\sim m^2/s^2$, a thinning is applied to RASTA wind "super-observations". One super-observation out of three is then assimilated, which is equivalent to approximately one observation every 5~km to 9~km depending on the aircraft speed."

Page 5, lines 25-28: It could also be so that “important” data is collected late (or early) in a longer assimilation window if the aircraft e.g. flies into a convective cell. In studies like this it would also be very beneficial to run with FGAT (First Guess at Appropriate Time) or even better using a 4D-Var assimilation scheme.

Reviewer 2 is right, FGAT is a way to improve the handling of the time dimension in a 3D-Var scheme as it allows to compute the innovations (i.e. the observation-guess differences) at the time of the observations for different times during the assimilation window. For the AROME model, the FGAT option has been evaluated by Brousseau (2012) for moving platforms, but without any positive improvement in the subsequent forecasts (Brousseau et al. 2016, section 2). For observations from static platforms, the 3DVar without FGAT only uses the observations performed at the middle of the assimilation window. The FGAT option allows to estimate innovations for sub-hourly data from the same instrument at the same location. More observations are assimilated, but the 3D-Var minimisation, without time dimension, uses these several innovations at the middle of the assimilation window. This leads to an averaging and a smoothing effect on these observations and a loss of information on the temporal details, which is not desirable in a convective DA system. Therefore, in this study we decided to use conventional 3DVar to assimilate all the different kinds of observations in the same way.

The 4DVar assimilation scheme is numerically too costly for the AROME model (Brousseau et al. 2016).

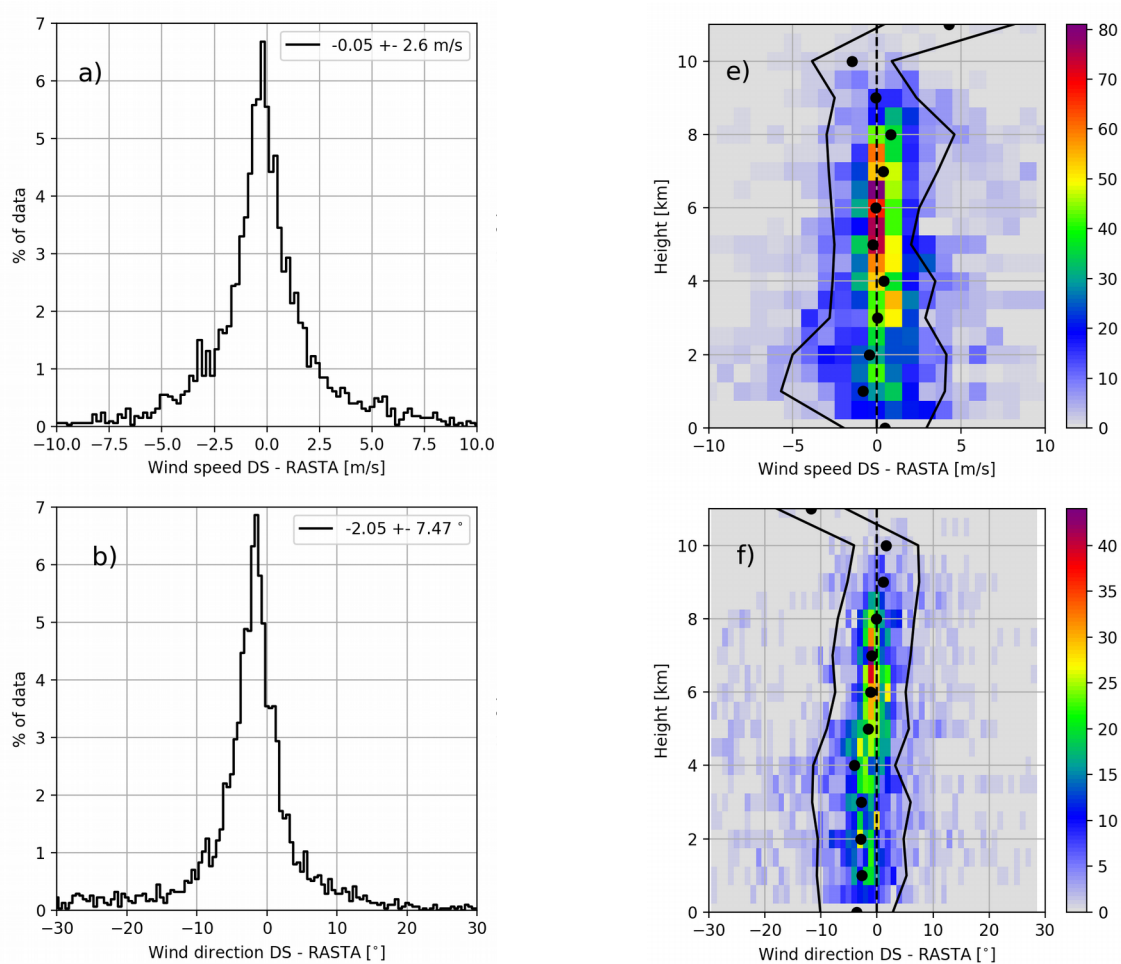
Brousseau, 2012: Propagation of observed information into the AROME data assimilation and atmospheric model, PhD thesis, Université de Toulouse III – Paul Sabatier

Brousseau, P. , Seity, Y. , Ricard, D. and Léger, J. (2016), Improvement of the forecast of convective activity from the AROME-France system. Q.J.R. Meteorol. Soc., 142: 2231-2243. doi:[10.1002/qj.2822](https://doi.org/10.1002/qj.2822)

Page 6, line 10: The same observation error as radiosondes. Isn't this a bit optimistic?

RASTA wind data during the HyMeX-SOP1 field campaign have been compared against ground-based Doppler radars by Bousquet et al. (2016). Results of their study show that “The low values of the bias error suggest that errors are close to Multiple-Doppler wind synthesis and should remain comprised between 1 and 1.5m/s” (see section 3.2, page 93). These values are smaller than the radiosonde ones (between 1.8 and 2.52m/s). We added these values in section 4.2. *“Bousquet et al. (2016) demonstrated that the bias error of RASTA wind data is comprised between 1 and 1.5 ms⁻¹. In this study, it has been decided to use the same observation error as the one used for radiosondes, which increases with the altitude (from ≈ 1.8 ms⁻¹ at 900 hPa to ≈ 2.52 ms⁻¹ at 200 hPa).”*

RASTA wind data have also been evaluated during the NAWDEX field campaign which occurred in Iceland (<http://www.pa.op.dlr.de/nawdex/>). In the following figures, RASTA wind retrievals were compared against radiosonde measurements. These Figures demonstrate that the observational error for RASTA wind data is of the same order of magnitude as that of radiosondes.



Bousquet, O. , Delanoë, J. and Bielli, S. (2016), Evaluation of 3D wind observations inferred from the analysis of airborne and ground-based radars during HyMeX SOP-1. Q.J.R. Meteorol. Soc., 142: 86-94. doi:[10.1002/qj.2710](https://doi.org/10.1002/qj.2710)

Page 6 lines 11-13: Is there any other quality control applied to the observations? If yes, what and how. If no, why not?

There is no other quality control applied to the observations. We did not apply any other quality control because after visual inspection, we did not see any remaining spurious observations.

Page 6, lines 20-24: The last two sentences in the paragraph is really hard to follow. Please re-write to make it more clear.

The text has been modified. We hope it is now clearer:

“A larger assimilation window results in assimilating data more frequently, but the time lag between the observation time and the analysis time is greater than one hour. On the other hand, a smaller assimilation window constrains the number of analyses to those for which the observations are valid near the analysis time. Therefore, the percentage of analyses in which RASTA wind data were assimilated decreases with the length of the assimilation window from 9.5% in the RASTA_3h experiment to 7.2% in the RASTA_1h experiment. Finally, the last column of Table 1 represents the percentage of RASTA wind data which were assimilated among the total number of assimilated data (conventional, GNSS, radar, satellite, RASTA, etc.) over the entire AROME-WMed domain (represented in Figure 1).

This percentage is quite small because of the already dense observing network used in AROME-WMed.”

Page 7, table 1: In the column of assimilated data it says Conventional and Conventional + RASTA. Are only conventional observations assimilated apart from the RASTA observations? In section 3.2 there are many more observations mentioned that are not considered to be conventional, e.g. GNSS and satellite data.

*All the observations which are mentioned in section 3.2 are assimilated.
Table 1 has been rectified.*

Page 7, lines 10-14: The data collection starts at 06:10 and the analysis time chosen to study is 06:00. This means that the 3 hour window only will be a 1.5 hour window. Is there no better example where one can find a data collection more centered around the analysis time. Why not show an example from 09:00? Then the data collection will also be skewed but there will at least be data available on both sides of the analysis time.

For the case study, we looked for a situation in which the Falcon 20 aircraft was collecting data around the location where the increments are advected as the forecast range increases. Such a configuration only happens twice during the entire HYMEX-SOP1 period: the IOP7a case study at the analysis time of 06 UTC, and the 2012-10-11 case study at the analysis time of 18:00 UTC, in which the observation time starts at 17:50 UTC.

There are other examples in which RASTA observations are more centred around the analysis time, for example at the same date at 09:00 UTC. However, there are no observations available after 9:30 UTC to validate the subsequent forecasts. Therefore, we decided to only show the 06:00 UTC analysis results because RASTA observations are available at 07, 08 and 09 UTC to validate the forecasts.

Page 7, line 20: “...expected to improve the forecast...” Is this really the case? It depends on how the data is introduced, observation errors and how the model performed without the data.

The authors agree with reviewer 2. We wrote this sentence because the increments are advected at approximately the same place of which the rainfall event took place. Therefore, the impact of the assimilation of RASTA data is expected to have a noticeable impact on the rainfall event. This is the main reason why this case study has been selected. However, at this stage, it is too early to talk about any possible improvement.

We replaced “expected to improve the forecasts” with “expected to have an impact on the forecasts”.

Section 5.2: It would be interesting to see the same case in a cycled period. If the cycled run builds up its own “climate” could the results be even better?

The results of this case study in the cycled run (starting from September 24) are indeed better for the CTRL and for the different RASTA experiment cycled runs. However, in this section we want to demonstrate the impact of the assimilation of RASTA wind data at a specific analysis time. Therefore, to disentangle the benefits brought by the cycling effect from the impact of the assimilation of RASTA wind data, results are shown for the non-cycled experiment runs for the case study.

In the statistical study, results are shown for the cycled experiment runs.

Page 8, lines 19-25: Please explain figure 5 better.

The explanations have been modified. Fig 5a (now 6a) is first described in section 5.2:

“Figure 6A represents the wind speed increments at approximately 4 km of altitude (model level 30) between the RASTA_3h and the CTRL analysis. Wind directions are also indicated by the green (resp. black) arrows for the CTRL (resp. RASTA_3h) analysis. The data points assimilated in the RASTA_3h experiment until 07:30 UTC are also represented by the black data points.”

Then, Fig.6 B-D are explained at the beginning of section 5.2: *“Figure 6 (panels B to D) represents the wind speed differences of the RASTA_3 h 1-, 2- and 3-h forecasts and the CTRL ones. At each forecast term, the black data points indicate the different RASTA locations which are available during a 1-h time window centred on the forecast time (forecast term $\pm$ 30 minutes).”*

We hope it is now clearer.

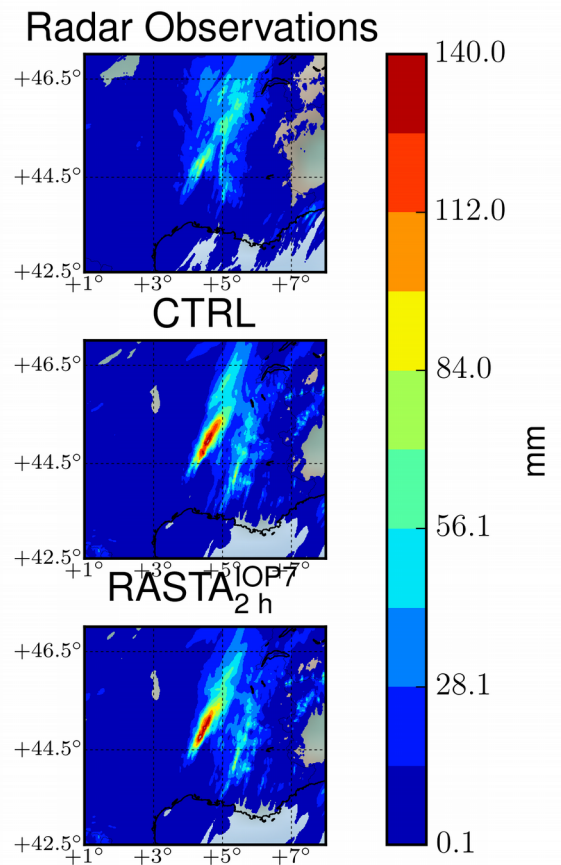
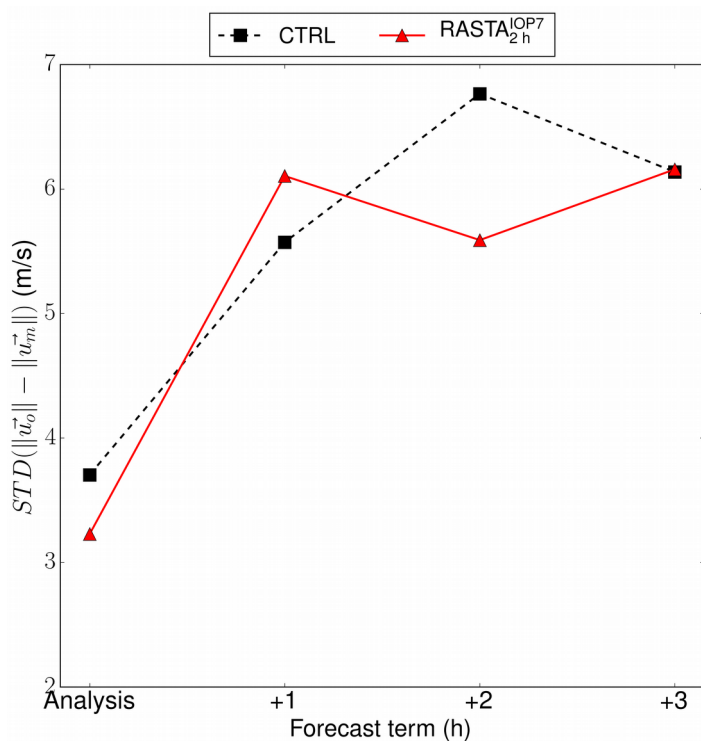
Page 9, line 9: The observations assimilated are not from $\pm$ 30 minutes from 06 UTC. They are from +30 minutes. Right?

Yes they are from +30 minutes. We rectified the text.

Page 9 and figure 6: This is a typical behavior when observations are assimilated with a too small observation error. The analysis is adjusted to fit the RASTA observations too much but as soon the model starts running it adjust itself to its own more comfortable state. The analysis will look very good, especially compared to RASTA observations, but there will be a spinup to the model state as seen in the figure. Why not run the same experiment with different observation errors too see if that can reduce the spinup and improve the forecasts, not only the analysis?

We ran the same experiment with a larger observation error of 6 m/s for the IOP7 case study (with an assimilation window of 2h). Results are displayed in the following figures.

The comparison against RASTA wind data (Figure on the left) shows that the spinup is not reduced and the 3-h forecast is not improved. Besides, the comparison against radar observations (Figure on the right) indicate that the assimilation of RASTA data with a larger observation error doesn't improve at all the 12-h accumulated rainfall.



Page 9, line 28: What forecast lengths are used for the 12 hour accumulation? Interesting to know in view of the above discussion about spinup.

We used the 12 hour rainfall forecasts for the comparisons in Figure 7 (between 06:00 UTC and 18:00 UTC). However, we agree that there is a spinup problem; and hence the first hour of rainfall accumulation should not be taken into consideration for the calculation. Therefore, we now compare the 11-h accumulated rainfall forecasts between 07:00 UTC and 18:00 UTC in Figure 7 (now Figure 8). The results are similar. We rectified the text in section 5.4.

In the statistical study, we also removed the 1st hour to calculate the scores.

Page 10, line 18: Calculations are only performed over the 35 runs (I assume that this means analysis times) with RASTA observations. Why?

As pointed out by Reviewer 1 in his/her major comment #1, the impact of the assimilation of RASTA wind data is limited in space and in time. To maximise our chances to see an impact of the assimilation of RASTA data on the scores, we constrained the number of assimilation times to only those for which RASTA wind data were assimilated. This is the reason why the scores are only calculated over the 35 analysis times in which it was possible to assimilate RASTA data. For the same reason, the scores are now calculated over a RASTA-limited area. This area contains the aircraft flight path +/- 0.5° both in longitude and latitude.

Page 13, line 11: Again quality control is mentioned and that it is important. What quality control was applied to the data assimilated here (see also comment above)?

The quality control applied to the observations depends on the first guess departure ($|Obs - Guess|$). This quality control depends on the observation error and on the background error. Since the observation error increases with the altitude, the quality control also depends on the altitude.