Dear AMT Editors and Reviewers,

Thanks for starting the review process and for the precious comments and suggestions provided. In what follows we list the comments expressed by the Anonymous Referee #1 (paragraphs written in blue color), provide our answers (paragraphs written in black color) to the raised points and explain how we addressed each point in the revised manuscript.

Comment on amt-2022-61 from Anonymous Referee #1
Referee comment on "3D cloud envelope and cloud development velocity from simulated CLOUD/C3IEL stereo images" by Paolo Dandini et al., Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2022-61-RC1, 2022

This work proposes a method to estimate cloud envelope and cloud development velocities based on high resolution satellite train imagery. Because the planned satellite train is not launched yet, the work relies on simulated image data to show the proof of concept. Although the simulated data are perfect in the sense that noise is not accounted for, they are helpful to compare against the ground truth (that will not be available in the case of real data) and therefore show the potential. Before publication, I suggest that the following points are addressed:

Line 12-14: "An independent method based on optimizing the superposition of the cloud top, issued from the atmospheric research model, allows to obtain a ground estimate for the velocity from two consecutive acquisitions." I do not understand this statement, what is meant by superposition in this context? Maybe it is more clear in the text but it just makes the abstract confusing.

Authors' reply :
For more clarity, in the abstract, lines 12-14 (pp. 1) (lines 15-16 – pp. 1 in the revised manuscript), we replace the following sentence :
«An independent method based on optimizing the superposition of the cloud top, issued from the atmospheric research model, allows to obtain a ground estimate for the velocity from two consecutive acquisitions. »
with the following one :
«An independent method based on minimizing the RMSE for a continuous horizontal shift of the cloud top, issued from the atmospheric research model, allows to obtain a ground estimate of the velocity from two consecutive acquisitions.»

Comment on amt-2022-61 from Anonymous Referee #1
Similar for the lines 14-17, please considering restating these lines to make it easier for the reader to understand.
Authors' reply:
For more clarity, in the abstract, lines 14-17 (pp.1) (lines 18-21 – pp. 1 in the revised manuscript), we replace the following sentence:
«The distribution of retrieved velocity and ground estimate exhibits small biases but significant discrepancy in terms of distribution width. Furthermore, the average velocities derived from the mean altitude from ground for a cluster of localized cloud features identified over several acquisitions, both in the simulated images and in the physical point cloud, are in good agreement.»

with the following one:
«The mean values of the distributions of the stereo and ground velocities exhibit small biases. The width of the distributions is significantly different with higher distribution width for the stereo retrieved velocity. An alternative way to derive an average velocity over 200s, which relies on

tracking clusters of points via image feature matching over several acquisitions, was also implemented and tested. For each cluster of points, mean stereo and ground positions were derived every 20s over 200s. The mean stereo and ground velocities, obtained as the slope of the line of best fit to the mean positions, are in good agreement.»

Comment on amt-2022-61 from Anonymous Referee #1
Line 56, Do these error bands differ by magnitude of the velocity and the height? If so, is it better to give a percentage?
Authors' reply:
The error assessment was done in the work by Horvath and Davis (Horvath and Davis, 2001a) who demonstrate the feasibility of the retrieval of cloud top and winds from MISR. The errors were determined on simulated data and reported as absolute value. We prefer to report the errors as the authors of the work did. However, we correct the sentence by referencing the work, from the same authors, that presents the first retrievals from actual data (Horvath and Davis, 2001b).

We then correct the following sentence, lines 55-56 (pp. 3) (lines 64-66 – pp. 3 in the revised manuscript):
«The first retrievals from actual data (Horvath and Davis, 2001) were consistent with the prelaunch error estimates of ±3 m/s and ±400 m for winds and heights, respectively.»
by replacing it with the following one:
«The first retrievals from actual data (Horvath and Davis, 2001b) were consistent with the prelaunch error estimates (Horvath and Davis, 2001a) of ±3 m/s and ±400 m for horizontal winds and heights, respectively.»
We also add the reference to the work of Horvath and Davis (Horvath and Davis, 2001b) in the edited manuscript.

We also add after the following sentence, lines 56 (pp. 3) (lines 66 – pp. 3 in the revised manuscript):
"These retrievals were obtained for the first time from the polar orbiting spacecraft Terra."
what follows:
"Only recently, Mitra et al. (2021) provided the first evaluation of the Terra Level 2 cloud top height (CTH) retrievals against the Cloud-Aerosol Transport System (CATS) Lidar CTHs, with uncertainties of −280 ± 370 m."

and then correct the following phrase:
"The main limitations of their method is the fact that vertical cloud motion is neglected and a constant horizontal cloud advection over the domain is assumed which under intense convection, for instance, may lead to unreliable retrieved winds."
as follows:
"The main limitations of the method of Horvath and Davis is the fact that vertical cloud motion is neglected and a constant horizontal cloud advection over the domain is assumed which under intense convection, for instance, may lead to unreliable retrieved winds."

Consequently, we also add the work by Mitra et al. (2021) to our list of references.

Comment on amt-2022-61 from Anonymous Referee #1
Line 75, Please be specific, magnitude acceleration in what? Computation time?
Authors' reply:
We rephrase the following phrase, lines 74-75 (pp. 3) (lines 85-86 – pp. 3 in the revised manuscript):
«Sde-Chen et al. (2021) devised a neural network for spaceborne 3D cloud CT, leading to 5-order of magnitude acceleration, relative to Levis et al., 2015.»

by rewriting:
«Sde-Chen et al. (2021) devised a neural network for spaceborne 3D cloud CT, leading to a significant reduction in terms of run-time, relatively to Levis et al., 2015.»

Comment on amt-2022-61 from Anonymous Referee #1
Line 101, Do I understand correctly that the image will not be taking images continuously but will start when triggered and for 200 seconds only. If so, please make this explicit and explain what will trigger the image capture event.
Authors' reply:
Line 101, that is correct. The cameras will not be taking images continuously. When triggered, the cameras will take two or three simultaneous images every 20 seconds during 200 seconds (11 acquisitions=11 pairs or triplets of simultaneous images). While no dynamic triggering is foreseen, the acquisition sequence will be scheduled at selected latitudes, depending on the season and on climatology, where and when clouds are more likely to be observed. Moreover, the acquisition schedule will be periodically (two/three times a year) updated to target the observation of convective cloud scenes and, when possible, to achieve co-location with ground observations. However, with such acquisition strategy, measurements over clear skies will not be avoidable. As for the synchronization of the image capture event from the different cameras, the pulse per second (PPS) signal from the GNSS receiver will allow to achieve atomic-clock accuracy with no need of communication between satellites.

For more clarity we correct the phrase, lines 102-105 (pp. 4) (lines 113-116 – pp. 4 in the revised manuscript):
"The observational strategy for the imagers will consist in multi-angular observations of a given cloud scene during 200 s with instantaneous stereoscopic pairs or triplets captured every 20 s (11 acquisitions A1-A11 see Fig. 1) corresponding to the life time of cloud perturbations at small scale."
as follows:
"The observational strategy for the imagers will consist in multi-angular observations of a given cloud scene during 200 s with instantaneous (not continuous) stereoscopic pairs or triplets captured every 20 s (11 acquisitions A1-A11 - see Fig. 1) corresponding to the life time of cloud perturbations at small scale. 3 or 4 sequences of acquisition, each of the duration of 200s, will be acquired per orbit. The image capture event will not be triggered dynamically but scheduled at specific latitudes, depending on the season and on climatology, where and when clouds are more likely to be observed. This schedule will also be tuned periodically, two/three times a year, to maximize the chance of observing convective cloud scenes and to achieve co-localized measurements with ground observations, when possible. As for the synchronization of the image capture event from the different cameras, the pulse per second (PPS) signal from the GNSS receiver will allow to achieve atomic-clock accuracy with no need of communication between satellites."

For more clarity we also add in the caption of Figure 1:
"Instantaneous (not continuous) stereoscopic pairs or triplets captured every 20 s over 200 s, that is 11 acquisitions A1-A11."

Comment on amt-2022-61 from Anonymous Referee #1
Line 110, How accurately can the satellite positions can be retained? That is what are the bounds of change in baseline? Is this likely to impact the retrieval quality?

Authors' reply:
Satellite positions are assumed to be known with good accuracy thanks to the GNSS receiver on-board. Satellite relative position mainly depends on the frequency of the operations of adjustment, coupled with solar activity, that for C3IEL are targeted to guarantee a bound of change of about 50

km, at worst. Rather than simulating the effect of "small" baseline changes to get insights into the feature matching quality and reconstruction uncertainty, we used two baselines, 300 km and 600 km, and found no significant difference in terms of stereo-retrieval, as Figure 10 shows. Our view is that the pointing error for an orbit of 600 km affects the line-of-sight (LOS) much more severely than a small error (as small as around 10 meters) on the satellite position. However, as for the work here presented neither orientation nor position errors were accounted for. This should be done in the future with the aim of decoupling cloud motion from the contribution of AOCS (Attitude and Orbit Control System) errors.

Lines 443-445 (pp. 30) (lines 501-503 – pp. 32 in the revised manuscript), we rewrite the following sentence:
“However, these sources of error have not been taken into account to carry out this work that instead is based on realistic but perfect images, "perfect" in that neither radiometric nor attitude errors have been accounted for as out of the scope of this paper.”
as follows:
“However, these sources of error and likewise the satellite position error, have not been taken into account to carry out this work that instead is based on realistic but perfect images, "perfect" in that neither radiometric nor attitude errors have been accounted for as out of the scope of this paper.”

We also add in the conclusions (last paragraph) that:
“However, in the future, to generalize our results, we plan to test our methods for other cloud types, scenes and solar geometries. This will be done by taking into account radiometric noise and image distortion as well as satellite orientation and position errors. This will allow to quantify the degradation of the results here obtained for “perfect simulations”.”

We also add in the introduction of section 3 (lines 165-167 – pp. 9 in the revised manuscript).
“This second simulation is thus more realistic than the first one and in the future will allow to account for image distortion and satellite orientation error.”

Line 200 (pp. 12) (line 273 – pp. 16 in the revised manuscript), after:
“In this way, the sphericity of the orbit and the orientation of the satellites are accounted for.”
we add:
“In this work, satellite position and orientation are assumed to be known exactly. However, this is not true and the corresponding errors are expected to deteriorate the results here presented. We will be able to test such statement, in the future, once these sources of error will have been modeled for each camera, by exploiting the combined use of Euclidium and 3DMCPOL.”

Comment on amt-2022-61 from Anonymous Referee #1
Line 140, is 22.5 km the location from a reference point? Please clarify.
Authors’ reply:
For more clarity, we rewrite the following sentence, lines 140-141 (pp. 9) (lines 221, 222 – pp. 13 in the revised manuscript):
«Fig. 4 shows the vertical section, at 22.5 km, of total water content, as well as vertical and horizontal wind components.»
as follows:
(Notice that Figure 4 becomes Figure 6 in the revised manuscript!)
«Fig. 6 shows the vertical section, at 22 km (6 km from the origin of the y axis located at 16 km, see Fig. 6g,6h), of total water content, as well as vertical and horizontal wind components.»
We also edit the caption of Figure 6:
“Deep convective cumulus physical properties. Fig. 6a, 6b, 6c: vertical section (y=22.5 km) of cloud total water content, vertical and horizontal wind components, respectively - Fig. 6d, 6e, 6f: vertical section (y = 22.5 km) of total extinction coefficient, cloud phase and effective radius,

respectively. Value of 31 (dark red) is associated to voxels where the mean ice phase function is used - Fig. 6g, 6h: liquid optical thickness and total optical thickness (liquid + ice), respectively."
as follows:
"Deep convective cumulus physical properties. Fig. 6a, 6b, 6c: vertical section (y=22 km, that is 6 km from the origin of the y axis located at 16 km) of cloud total water content, vertical and horizontal wind components, respectively - Fig. 6d, 6e, 6f: vertical section (y = 22 km) of total extinction coefficient, cloud phase and effective radius, respectively. Value of 31 (dark red) is associated to voxels where the mean ice phase function is used - Fig. 6g, 6h: liquid optical thickness and total optical thickness (liquid + ice), respectively."

Comment on amt-2022-61 from Anonymous Referee #1
Why do Shallow cumulus clouds have better resolution? Shouldn't it be the opposite?
Authors' reply:
The spatial resolution depends on the configuration of the atmospheric research models. As these simulations are highly time demanding, we used a deep convective cloud case, modeled via Meso-NH, from previous work (Strauss et al, 2019).
The atmospheric research model SAM used to simulate the trade wind cumulus was instead configured with a resolution of 20m.

Comment on amt-2022-61 from Anonymous Referee #1
Lines 327-333, is there a specific reason for merging 2xtwo-view instead of using three-view when data from Sats 1-2-3 are used?
Authors' reply:
As we stated in the manuscript, the working principle of the s2p algorithm is based on stereo matching. Stereo Matching and Structure from Motion (SfM) are currently the predominant methods to derive geometric structures from satellite images (de Franchis et al., 2014; Zhang et al.,2019). However, while SfM reconstructs information of multiple images, Stereo Matching is restricted to single image pairs. SfM based approaches are inherently better suited to process large (unstructured) image sets such as multi-date satellite imagery and could be tested in future works. We clearly stated (lines 328-333, pp. 21) (lines 374-379 – pp. 22 in the revised manuscript) that:
«It is important to emphasize that the s2p algorithm uses two-view stereo at a time and then merges these independent two-view stereo reconstructions into a single reconstruction. This is contrary to full multi-view stereo methods (e.g., which use the whole set of three-views simultaneously). Multi-view methods are widely used in computer vision due to the advantages they bring over the two-view stereo (Zhang et al. 2019). Using full multi-view stereo methods might lead to different results in terms of 3D reconstruction via three cameras, namely that the 3D cloud envelope retrieval can be more accurate and lead to more detected points, than when using only two views.».

We rephrase the last sentence of sec. 4.1, lines 331-333 (pp. 21) (lines 379-382 – pp. 22 in the revised manuscript) as follows:
"Using full multi-view stereo methods might lead to different results in terms of 3D reconstruction via three cameras, namely that the 3D cloud envelope retrieval can be more accurate and lead to more detected points compared to when using two views."

Comment on amt-2022-61 from Anonymous Referee #1
Figure 10, Do I understand correctly from the figure that no points are retrieved with Sats 1-3 and Sats 1-2-3 scenarios in A3-5 and A8-9 views? If so, how can you say that none of the configurations outperform? Also, what is thee reason for the skewed-towards-A9/A10-views distribution of error in z in Figure 10.b?
Authors' reply:

Figure 10: For more clarity we repeated the calculations and accordingly updated figure 10 that now includes calculations for all scenarios. The conclusions are the same as previously, that is none of the configurations outperforms the others.
However, by repeating the calculations, we notice that the mean Y error for configurations 1-3 becomes now 5 m/s whereas previously it was about 15 m/s. We then correct the sentence, lines 357, 358 (pp. 24) (lines 406, 407 – pp. 25 in the revised manuscript):
«The mean difference (its absolute value) along z is less than 25 m while it is less than about 5 m and 15 m along x and y, respectively.»
as follows:
«The mean difference (its absolute value) along z is less than 25 m while it is less than about 5 m along x and y.»
and then we add:
"Such values can partly be ascribed to the stereo-opacity bias, associated to low extinction near the cloud top, as discussed by Mitra et al. (2021)."

In the abstract we also correct, lines 9-11 (pp. 1) (lines 12-14 – pp. 1 in the revised manuscript):
"The accuracy of the retrieval of cloud topography is quantified in terms of RMSE and bias that are respectively, less than 25 m and 15 m for the horizontal components and less than 40 m and 25 m for the vertical component."
as follows:
"The accuracy of the retrieval of cloud topography is quantified in terms of RMSE and bias that are respectively, less than 25 m and 5 m for the horizontal components and less than 40 m and 25 m for the vertical component."

Concerning the skewed distribution of the error along z for the A9-A11 views, we add the following sentence, line 359, (pp. 24) (lines 406, 407 – pp. 25 in the revised manuscript) after 'The mean difference (its absolute value) along z is less than 25 m while it is less than about 5 m along x and y.':
"The skewed distribution of the error in Figure 10b, for the views A9-A11, may be due to the fact that fewer cloud features are visible as the clouds are less illuminated by the sun, with larger portions of the cloud field shaded, as it can be seen from Figure 5g, 5h and 5i."

Still with respect to the calculations concerning the cumulus cloud, but this time with regard to Figure 9b, for better clarity and to improve on points visibility, we replace the three figures at the bottom of Figure 9b showing the stereo and ground truth cumulus clouds and the corresponding M3C2 cloud-to-cloud distance.

Comment on amt-2022-61 from Anonymous Referee #1
Equation 10, why are the subscripts suddenly 9 and 11?
Authors' reply:
Equation 10, we replaced 9 and 11 with 5 and 6, respectively.

Comment on amt-2022-61 from Anonymous Referee #1
Line 418, please explain why a dual mode caused by a possible "divergence of the cloud top in the center right part" would not show in the GE distribution?
Authors' reply:
This diverging effect appears in both the retrieved velocity and the GE velocity although it is more pronounced in the former. The absence of a dual mode in the GE distribution is due to the fact that the M3C2 method underestimates the actual distance vector especially when cloud development does not occur along the local normal as, for instance, at the top of an eddy.
Lines (418, 419 – pp. 28) (lines 473-476 – pp. 30 in the revised manuscript), we rephrase the sentence :

“The double modes in the retrieved Vy histogram, not present in the GE, could be associated to the divergence of the very cloud top in the center right part of the image.”
as follows:
“The double modes in the retrieved Vy histogram, which could be associated to the divergence of the very cloud top in the center right part of the image, are not present in the GE distribution. Although a hint of cloud divergence is also visible in the ground estimate of Vy, the double modes are smoothed out because of the underestimation of the distance vector associated to the use of the M3C2 metric.”