Answers to reviewer 2:

General comments: The paper is of interest and should be published after revision. Taking the actual contents of the paper the reader gets the impression that this is a report on Aldis GBT examinations, rather than on Euclid. Moreover, a large fraction of the reported information is already known and published. Nevertheless, the compilation of different aspects related to DE and LA is useful and gives an excellent overview. Still, there is dominant reporting about GBT, but very little on the entire Euclid area. Since this discrepancy is not readily remedied, the title should be changed.

The GBT data is used as a ground truth reference to validate the performance on a given location of the network. We further compare GBT results to results from video and E-field measurements on different locations within the network. We show that the DE/LA results from the GBT are similar to DE results from video and E-field measurements (downward flashes) on different locations in the network and conclude those results are also applicable to other regions of the network with similar sensor baselines. This procedure is similar to what is done in the U.S. for the NLDN and other networks using triggered lightning as a ground truth.

The paper uses very heavily internal measuring and analysis information from Vaisala (the producer and owner of both HD and SW of the described networks), although the name is not mentioned. The impression of having some company influence could be countered by considering the scientific principle to place the described results in a more general context, namely by mentioning alternative principles and other lightning detection networks that yield similar results related to DE and LA.

Thanks for this comment – we will mention the manufacturer of the network in the final paper. We agree that other networks may have similar performance but to the best of our knowledge no other network analysis in Europe was done with independent ground truth data. Further this first part of the companion paper should help to support the reader of the second part in understanding the changes in performance during the analyzed time period.

Specific comments: 1. p.5326 Line 9 and ch. 5.1. It is stated that the media LA is in the range of 100m; it should be made clear that this value does not apply to the total network but only to some smaller areas, and that the Euclid-typical error is listed as 500m (p.5334 line27).

Thanks for this comment – we will clarify this in the document. We agree that the LA is variable in the network. The statement on page 5334 line 27 gives the worse median LA in the network. With the video and E-field measurements we show that large regions of the network exhibit a much better LA.

2. the authors should give the scatter, especially the upper half-width of the error distribution, for the data that is derived from CHI2 analysis and does not refer to GBT.

Thanks for this comment – we will add in the final paper also the 95% value so that the complete error distribution is defined.

3. the median LA results from the well-known error ellipse. A quantitative result implies assumption about the data error in the least-squares fit. The authors should explain what errors they assume and whether these errors are the same in the entire network.

The calculation of the error ellipse is solely based on the assumption that the time error and angle errors are based on a Gaussian distribution. The parameters used for the calculation namely the standard deviations of the time and angle error are determined for each individual sensor on a regular base. It was shown with GBT data that the error ellipse is a good measure for LA [Diendorfer et al. 2014]
We will add a sentence in the paper to make this clear.

4. p.5327 line 6. It should be clarified here that rocket-triggered lightning is not used in the present paper.

We mention in the paper p.5327 line 11 that we use approach (A) and (C) and not rocket triggered lightning.

5. ch. 5.1 and Fig.5. Many improvements of the LA are mentioned, but all of them refer to the TOA locating technique. The authors should explain to what extent direction finding (DF) affects the LA. It is well known that almost 50% of all locations from Vaisala networks are produced with 2 and 3 sensor reports; thus, TOA is not effective and DF is dominant. As a result, since DF is intrinsically inaccurate, large errors and many outliers are present. The authors should show a graph that displays the number of sensors used for a locating process, preferably for two areas with small and large baselines, where TOA is more or less dominant, respectively.

We agree that 2 and 3 sensor solution have a lower LA compared to a 10 sensor solution. The validated LA based on ground truth data includes of course also 2 and 3 sensor solutions and therefore the resulting LA describes the performance of the network as it is.

6. ch.5.1 p.5334 line 27. It is stated that the LA of the entire network is around 500m. A value like this is not new and has always been stated by Vaisala and in Vaisala-related analysis work by scientists, at least since 10 years. The authors may explain this discrepancy.

We mention "better than 500m in the majority of the network" and we do not see any discrepancy.

7. Fig. 5 must have a better caption. The dimension should be given (km). The legend gives the lower value of an error interval (zero!).

Thanks for the comment - the unit is missing and we will adapt also the colors and start with 0.1km accuracy.

8. ch.5.2.1 DE around GBT. It is stated that the DE for 2 kA strokes at GBT is 70%. How many sensors contribute to this type of stroke locating?

We mention that for strokes greater than 2 kA the DE is 70% and for those strokes the number of contributing sensors varies in a wide range.

9. ch.5.2.2 and Tab.3; DE in the network. A paper on performance of Euclid should give more information about the DE. This chapter is much too short. At least a DE map should be given for the entire area, along with proper definition of DE

2

In fact this is an important issue to be discussed. Or focus in this paper is to show the performance analysis we did based on ground truth data. Regarding DE there exists no way to determine DE from the network data only. A spatial variation of the flash DE is always based on a so called DE model. Because of our focus on ground truth data validation we decided not to include this type of DE representation.
Model based DE is still idealized as it does not consider any temporary sensor outage due to communication or hardware problems.

10. ch.5.3. and Fig. 8/9 on peak current. In other publications much smaller errors are claimed. Here, it is surprising that peak-current errors as large as 50% seem to be frequent. The authors should give an explanation.

Only a small fraction of the strokes exhibit an error >=50%. It is commonly agreed that the largest part of the observed variation is a result of a variation of the return stroke speed which is assume to be constant in the peak field to peak current conversion.