Answer to Referee 3

Manuscript: "Monitoring of the Eyjafjallajökull volcanic aerosol plume over the Iberian Peninsula by means of four EARLINET lidar stations" (acp-2011-863)

The paper reports lidar vertical resolved observations of the volcanic aerosol over the Iberian Peninsula. These observations are interesting because cover a region poorly investigated concerning the Eyja eruption. On the other hand, lidar measurements are really valuable for studying the vertical stratification of the volcanic cloud.
Nevertheless, the presentation is confusing and results are not well presented. Data presentation should be strongly improved for the publication on ACP. In the present shape, reader could doubt about some data reliability.
These are the main questionable points:

- **For sun-photometer data, authors do not use always data of the best available quality level.**
- **The performances of the lidar systems in terms of retrievable properties and related errors are not reported and are not available in the referenced literature for some stations.**
- **Lidar profiles are shown in such a way that readers cannot judge their reliability (see details reported in the following).**
- **Questionable differences are reported for almost near stations in terms of intensive optical properties. These differences should be explained and discussed in more details.**
- **Discussion of optical properties is often confusing**

I suggest a careful revision of the paper for taking the most from the data and giving them the right visibility.
>>>>>> A careful revision of the paper has been made. All the changes are indicate sin bolt in the revised manuscript. A point-by-point answer is also provided below.

Specific comments
Page29684 Line 7:
air traffic was blocked over Southern Europe also in other periods. Maybe you refer only to Iberian Peninsula.
>>>>>> The end of the sentence was replaced by "…and in southeastern Europe in May 2010.".

Page29684 Line 9:
Here there is a precise quantification of the cancelled flights and passengers involved. This kind of statement requires a reference. Otherwise authors should avoid this level of precision.

>>>>>> This is general information from the press. The sentence has been replaced by the following less quantitative statement "Several thousands of flights, which affected at several millions of passengers, were canceled.".

Page 29685 Line5-8

Cryptic sentence. Re-write it in a more readable way. About its content: what is the source of this information? Please include references. Or are these already results of the paper? If it is the case, this is not the right place to report it.

>>>>>> The first part of the sentence has been rephrased. Those results are being published in the work by Pappalardo et al. (2012) in this special issue. The reference has been added.

Page 29685 Line 15:

Authors state that these observations are interesting for testing models at boundary conditions but they say this was already done. This is misleading and gives the impression that this paper is completely useless. On the other hand, is the work reported in Molero et al. still in progress and will lead to the publication on a peer reviewed journal beyond the cited proceeding?

>>>>>> The study by Molero et al. presents a simple comparison between model predictions and lidar vertical profiles and does pretend to test models at boundary conditions. This work has not been published in a peer-reviewed journal and at the moment it is not in progress. The sentence has been totally rephrased as "In a preliminary study Molero et al. (2010) made a simple comparison between model predictions such as EURAD (EURopean Air Pollution Dispersion, http://www.eurad.uni-koeln.de/index_e.html) and FLEXPART (http://transport.nilu.no/flexpart) and vertical profiles obtained by four lidar systems over the IP.".

Page 29687, Lines 1-2:

How are the performances of the systems? In table 1 you reported the characteristics of systems but without any details about their observational capabilities. The cited abstract is not available online so it is impossible to the readers understand the quality of the systems. It is essential for the reliability of presented results that some numbers are provided: covered altitude ranges, absence of instrumental errors, typical performances and statistical errors.

"systems…were satisfactorily compared": the authors should at least report the most important results about the inter-comparison campaign as the altitude region where the systems are reliable and their degree of accuracy at different altitudes level for aerosol backscatter and extinction coefficient.

>>>>>> In the frame of the VI EU-FP EARLINET-ASOS project, a Quality Assurance programme certifies the accuracy and the temporal stability of the lidar systems through internal Quality Assurance & Quality Check tests, and hardware and software

intercomparisons. Reports about the quality assurance of the software and about the internal quality checkups have been provided as deliverables of the project. With respect to the results of the hardware intercomparison campaigns, a report has been published in March 2011 (Freudenthaler et al., 2011). Unfortunately this report is accessible via web only through the intranet of the project. It states in page 49: "The results of the campaign can be considered satisfactory. The campaign allowed to check the performance of the systems and when they were not fully satisfactory, the reasons of the failure were understood and the way to solve them were defined". This reference has been added in the revised manuscript. The reference of Molero et al. has been deleted.
With respect to the following information:

- covered altitude ranges: see Sicard et al. (2009). A line has been inserted in Table 1.absence of instrumental errors: see Sicard et al. (2009)
- typical performances and statistical errors: see Sicard et al. (2009)
- altitude region where the systems are reliable: see Sicard et al. (2009). A line has been inserted in Table 1.
- their degree of accuracy at different altitudes level for aerosol backscatter and extinction coefficient: some hints of information is given in Sicard et al. (2009). However this is a very difficult number to provide. It depends extensively on the system itself (wavelength, power of the laser, etc…) and on the atmosphere and aerosol loading.

Page 29687, Line 6:
Authors state that "All nighttime measurements were also inverted using the Raman lidar inversion algorithm". What does it mean? Authors have two different inversions for each night-time measurements? And which is at the end used for the overall analysis?
>>>>>> Yes the nighttime measurements have been inverted with both the elastic and the Raman inversion algorithms. The elastic inversion has been performed systematically in order to get a uniform set of AOT to plot in Figure 3 (Section 3). In the rest of the paper (Section 4) when the measurement is during nighttime the Raman inversion is presented and when the measurement is during daytime the elastic inversion (the only one available) is presented. At the beginning of Section 3 and 4 it is clearly indicated which inversion product is used.

Page 29687, Line 11:
For Madrid, sun-photometer instrument is located at 250 km far from the lidar station. Taking into account that Madrid is a big city, probably large differences should be expected at these distances. The appropriateness of combining lidar and sun-photometer data for so distant observational points should be discussed.
>>>>>> A similar comment was made by Referee 1. Here is our answer.
The authors are aware of the lack of representativeness of AOT data obtained at Caceres respect to the Madrid area, both due to the distance (Cáceres is 250 km away from

Madrid) and the different conditions of a small town (Cáceres: 100.000 inhab.) respect to a larger metropolitan area (Madrid: 5 million inhab.). Anyhow, the authors decided to show the Cáceres data for completeness of Figure 2, as this is the closest AERONET station to Madrid. Originally, the decision to use data from this station was based on previous studies of long-range-transport events, namely Saharan dust intrusions, where PM10 concentrations from several Iberian Peninsula EMEP stations correlated reasonably well (Escudero et al., 2006). Of course, such long-range-transport events are more homogeneous spatially than volcanic events, so the decision is somehow compromised. Differences between the AERONET station during the volcanic event have been published (Toledano, 2012), although the center of the Iberian Peninsula, where Madrid is situated, is not well represented. A sentence has been added and references given in Section 2.2 saying that the Cáceres AOT is a good indicator of long-range transport aerosols.

Page 29687, Line 20:
Angstrom exponent are also provided within AERONET products. Why do you calculate them from AOT? Are there any difference?
>>>>>> We apologize for this erroneous information. The Angstrom exponent presented in the paper is the one given by AERONET. The last paragraph of Section 2.2 has been deleted.

Page 29687, Line 21:
Why level 2 data are not available for those stations after 1.5 years? Indeed I checked the current database of AERONET and Barcelona and Granada has level 2 data for May 2010 and they should be used instead of level 1.5.
>>>>>> Level 2.0 from AERONET at Granada and Barcelona has been released while the paper was in review. We have used those data in the revised manuscript and clarified in the paper that the highest AERONET level available was used: 2.0 for Madrid, Granada and Barcelona and 1.5 for Evora.

Page 29688, Line 18:
Authors should explain why they use 1200 UTC as representative for the whole day, if it is the case. Otherwise they should better explain that backtrajectories are reported as example but of course they are checked case by case accordingly to the measurement timing.
>>>>>> The backtrajectories have been checked case by case at each station. A couple of sentences have been added in the text to clarify this point.

Page 29688, Line 13:

Figure 3 is not readable in term of RSCS structures and features are not visible at all. Nevertheless there are strange different colors reported for Madrid. Also AOT and its components are not well distinguishable into these plots.

>>>>>> The same comment was made by Referee 2.

There are several things to take into account if we want to maintain Figure 3 as it is now: we are dealing with 4 days of measurements going from clear to cloudy skies and from volcanic aerosol layers of 200-m thickness and AOT < 0.005 and layers of ~2-km thickness and AOT on the order of 0.1, in average. Putting in evidence all the volcanic aerosol layers in Fig. 3 would require a different colorbar for each measurement. However the chronological plots of the RSCS as they are shown in Fig. 3 are represented with a single colorbar per station for the sake of clarity and for comparison purposes, so that a compromise has to be made in the selection of the colorbar. We have chosen to lose some information at the beginning of the event when the layers are very thin and to enhance the contrast starting on 6 May. In the revised manuscript Fig. 3 has been enlarged to fulfill an A4-size sheet and its resolution/quality has been improved.

If the referees prefer, the authors could change the format of Fig. 3 that would result in a loss of information that the authors would like to avoid. It would be possible to make two panels and zooming in the 2 most intensive periods of the event, e.g. Évora-Madrid on 6 May and Évora-Madrid-Granada on 7 May. But again the authors would prefer not to lose the continuity of the plots as they are presented now in the current manuscript.

As fas as the AOT and its components are concerned, several options have been tested to improve the plot: making the symbols smaller, change the scale of the AOT and the Angstrom exponent, etc. But none of them was satisfactory. We finally realized that, except on 5 May where the total and the fine mode AOT overlap and are not distinguishable, the sun-photometer products are clearly legible during 6, 7 and May.

Page 29689, Lines 21-23:

Some of the involved stations have Raman capabilities. Why authors do not reported really measured AOT? The 50 sr assumptions should be discussed and justified. Is this assumption in agreement with your results reported in the Section 4?

>>>>>> There are several reasons:

- Because of the difficulties to invert the Raman signals in the very thin layers of volcanic aerosols, the set of good quality Raman-inverted profiles is poor. If we plotted only the AOT calculated from those profiles Fig. 4 would be made of very few and spotty points.
- In Fig. 4 we want to show the same quantity. The only way to plot the AOT calculated with a common method for day and nighttime measurements is to use the elastic algorithm and a constant lidar ratio.

The assumption of 50 sr, as a first approximation, has been justified in the text with the sentence: “This value is a good approximation in view of the lidar ratios found in Ansmann et al. (2010) and the results presented in Section 4 of this paper.”.

Page 29688, Line 27:
Typo error. Add daily mean of the
>>>>>> Corrected.

Page29690, Table2:
It is reported in table 2 that wide volcanic layers are observed while in the abstract it is reported that "The volcanic aerosol layers observed over the IP were rather thin (< 1000 m)"
>>>>>> The layers observed in Evora, Granada and Barcelona are usually thin. However some thick layers are observed punctually at the beginning of the period in Evora and Madrid. The abstract has been completed with the following sentence: "However, in some cases at the beginning of the period the thickness of those layers reached several kilometers in Évora and Madrid.".

Page29690, Line 12:
What does it mean here "Taking into account the low values of AOTVA found, those uncertainties are reasonable"? Reasonable for what?
>>>>>> The word "reasonable" has been replaced by "acceptable".

Page29690, Line 20:
This is also supported by the low AOT _VA
>>>>>> This is correct.

Page29690, Line 23:
Did the thermal low persist also during the following days? Is this in agreement with layers as observed by Madrid stations and reported in figure 4?
>>>>>> After looking at the record of the surface level temperature (between 10 and 18ºC during daytime) and pressure in Madrid in the period 6 – 8 May our conclusion is that the thermal activity was not enough to form a thermal low. This explanation has been eliminated from the manuscript.

Page29690, Line 28:
Please report here that you are referring to Madrid
>>>>>> Corrected.

Page 29691, Line2-3:
Madrid reported observations are not linked in any way to the rest of the paragraph. Please comment on them.

>>>>>> The sentence has been completed as: “In Madrid the first VA layers detected have an lower than 0.016 which indicates extremely optically thin layers given their large vertical extension.”.

Page 29691, Line 12-14:
Authors report about a new layer, but it is not visible in the RSCS in the current version.
>>>>>> The new layer is perfectly visible in Fig. 3. Please look at the Evora plot, on 6 May in the middle between 12 and 00 UTC a layer appears (red/yellow spot) at about 5 km.

Page 29691, Line17:
Is the cirrus discarded in VA AOT calculation? Maybe the strong AOT in VA reported in figure 4 is related to this case.
>>>>>> Yes, all the cases where VA are mixed with clouds have been discarded from the analysis.
The strong AOT_VA reported in Fig. 4 is in the first hours of 7 May. The corresponding RSCS in Fig. 3 is in darker colors compared to the measurements of the second half of 6 May and therefore also indicates an optically thicker layer.

Page 29693, Line 10:
Raman inversion is limited but not the others, rephrase it in order to avoid giving the impression that during night-time nothing can be provided.
>>>>>> The sentence focusing on the nighttime Raman inversion has been deleted and the sentence before has been reformulated as “The fact that the lofted VA layers of interest in this study were optically very thin (in average AOT_VA < 0.020 over all stations and over the whole period) induced lidar signals with very low signal-to-noise ratios which made the inversions, and in particular the nighttime Raman inversions, quite difficult. ”.

Page 29693, Line 11:
Are these cases selected also because of interesting situations or are these the only profiles you have?
>>>>>> This question was also asked by Referee 1. The 3 cases shown in the paper are not the only good quality inversions available. There are a selection made out of the good quality inversion based on a chronological criterion during the most intense period (7-8 May) and showing profiles from all stations (for Barcelona there were no other choice than showing the profile on 8 May at 1600 UTC). The number of cases was limited to 3 in order not to make the paper too long.

Page 29693, Last line:
You should motivate the choice of lidar ratio

>>>>>> Please refer to the justification on the assumption of the 50 sr value given earlier. The sentence has been completed as "In this section AOT_VA was calculated as the integration of the extinction profiles when Raman measurements were available and otherwise as the integration of the backscatter profiles multiplied by a constant lidar ratio of 50 sr in a first approximation (see Section 3.2).".

Page 29693, Figure 5:
This figure should be significantly improved. Profiles should cover the calibration range too and the low troposphere (above the overlap region, of course) to clearly show the quality of reported data. Moreover the high lidar ratio values observed over Madrid should be discussed in more details by the authors. Error bar on lidar ratio is very high (about 50%) and it is clear looking the figure that S in the central part of the layer is around 60sr.
A layer has been identified and its boundary reported for the Evora profile. What about the Madrid one? In the calculation of mean lidar ratio authors used also the values around 20 sr obtained at about 7 km asl?
>>>>>> A figure 5 is added as supplement to the answer to the referee. There we have plotted 1) the Madrid backscatter profile from above the overlap region and up to 7 km, 2) the Madrid extinction profile from above the overlap region and up to 7 km without smoothing (dark green) and with smoothing in the VA region (light green), 3) the Madrid lidar ratio profile in the VA region calculated with the smoothed extinction profile. This figure is intended to check the quality of the reported profiles.
The boundary of the VA layer in Madrid is identified between 4.2 and 6.8 km by looking at the backscatter profile.
The mean lidar ratio provided in Fig. 5 includes the whole profile between 4.2 to 6.8 km.
As for Fig. 6 and 7, for the sake of clarity of the figure Évora profiles are not represented above 4.0 km and those of Madrid below 4.0 km.

How do authors explain the different behavior of Angstrom exponents with the altitudes? Extinction related Angstrom exponent suggests that particles below 3 km are larger than those at upper levels. But backscatter related Angstrom exponents show exactly the opposite.
>>>>>> In general extinction and backscatter coefficients are not correlated so that a variation of one of them does not imply a variation for the other. The resolution with which both coefficients are retrieved is different in the Raman inversion algorithm. The extinction retrieval requires long intervals of integration that imply a considerable loss of resolution while the backscatter retrieval, once the extinction retrieved, is performed with the original resolution of the lidar signal. The spatial integration needed for the extinction retrieval can lead to profiles with misleading regions depending on how the integration intervals have been seleted.
As a rule of thumb, we suggest to look at the values in the middle of the Angstrom profile (but it is also valid for extinction and the lidar ratio) and not at the boundaries.

What about other wavelength backscatters for Madrid? It could provide indication about the size of the particle if compared to the Evora measurements.
>>>>>> At this period (May 2010) Madrid did not have other wavelengths than 532 and 607 nm.

Error in figure4 caption: the 532/1064 is backscatter and not extinction angstrom
>>>>>> Corrected.

Page 29694, lines 12-end:
It is very difficult to follow this discussion. Authors report that lidar ratio values are rather small compared to recent studies (Ansmann et al., 2010; Mona et al., 2011; Wiegner et al., 2011), but then refer to Mona et al., for justifying it. In that paper it is reported: the lidar ratio increasing with
RH from about 40 sr at 20% as RH up to about 70 sr at RH of about 70% …..suggest the presence, besides sulfates aerosols, of some ash affected by the aging through the European continent I would expect a completely different situation over Iberian Peninsula: in this case there should be not contamination with continental aerosol, maybe with marine aerosol. This could justify the smaller lidar ratio values.
>>>>>> We refer to Mona et al. to justify the lidar ratio dependence on relative humidity: the lidar ratio decreases when the relative humidity decreases. This is a general result independent from the mixing of the volcanic aerosols.
As suggested by the referee, we have made clear in the text that, by looking at the backtrajectories, no contamination is expected with anthropogenic aerorols but that possible mixing with sea salt might explain the lower values of lidar ratio observed.

Page29695, Line 26:
This is a different phase of the eruption so that completely different particles in terms of depolarization could be present. However here you just want to say that it is reasonable you have a mixture of ash and no-ash and this is clear from the dep value as itself also without any comparison.
>>>>>> The referee is true. The discussion has been rewritten mostly by changing the order of the sentence.

Page29695, Line 20:
I tried to understand something about these lidar ratio values and their ratio. But also this part is difficult to read. Authors report "according to Mona et al. (2011) a ratio of lidar ratio values around 1 suggests the presence of both aerosol types: non-ash and aged ash." I found in Mona et al that "the lidar ratio increasing with RH from about 40 sr at 20% as RH up to about 70 sr at RH of about 70% and the ratio of lidar ratio values

below 1 suggest the presence, besides sulfates aerosols, of some ash affected by the aging through the European continent." and "High Suv, particle linear depolarization ratio increasing with RH and values of the ratio of lidar ratios greater than 1 suggest the presence of volcanic sulfates/continental mixed aerosol." I found no information about values around 1.

>>>>>> On this specific matter of ratio of lidar ratio the authors are not totally confident with the results from Mona et al. (2011) which have not been found in any other work and have decided to delete their discussion on this topic.

Page 29696, Line 1:

There are big differences in observations reported for Germany and Italy, consisting respectively of mainly ash and mixed ash/no-ash particles. Again authors are providing a confusing and misleading discussion of their results. Please try to be more clear and precise.

>>>>>> Section 4 has been revised entirely. Many parts and confusing words have been modified. In relation with the depolarization ratio, the different plumes observed in Germany (ash) and in Italy (ash and non-ash particles) have been clearly stated.

Figure 6-7:

Same comment of figure 4 applies also to these figures: it is not good to show profiles extending only over 1-2 km of altitude range. Why the Barcelona profile do not extend, for example, up to the beginning of Granada ones? If Barcelona profile is not reliable above 3 km how it was calibrated?

>>>>>> The Barcelona profiles have been extended up to the beginning of the Granada ones and down to the minimum height scale. And the Granada profiles have been extended up to 3.4 km, the maximum height scale of the figure. For the sake of clarity of the figure Granada profiles are not represented below 3.0 km and those of Barcelona above 3.0 km.

Page 29698, Line 8:

As authors pointed out mass concentration is an important parameter and there is a strong request from no-scientific community to have this kind of information and conversion factors between optical properties and mass concentration. However, the volcanic plume analyzed by Tesche t al., 2011 is completely different from what observed over Iberian Peninsula: different eruptive phase, different transport path, different potential mixtures. Authors should at least comment on the uncertainty related to the pure ash and pure non-ash particle depolarization ratios assumed values. The same applies also for density and S values.

>>>>>> The reference to Tesche et al. (2011) has been deleted following a recommendation from Referee 1. The method is the one described by Tesche et al. (2009) and apply to ash and non-ash particles as in Ansmann et al. (2011).

A paragraph has been added in Section 4.2 where the uncertainty of the mass concentration is estimated. All the numbers reported for mass densities, the lidar ratios, and the mean extinction-to-mass conversion factors are taken from the literature. The uncertainty for the total aerosol backscatter coefficient has been calculated like in Section 3.2 and converted to ash and non-ash backscatter coefficients following the method by Tesche et al. (2009). The resulting mass concentration uncertainty is 45-50 %. It is higher than in Ansmann et al. (2011) mostly because the uncertainty in the retrieval of the total aerosol backscatter coefficient is higher.

Page 29702, Line 20:
Yes, but how large is the systematic error on backscatter? Which are the S and calibration contribution?
>>>>>> The systematic errors on backscatter are on the order of 10, 6 and 5 % at 355, 532 and 1064 nm, respectively (Sicard et al., 2009). The contribution of an erroneous backscatter coefficient at the calibration height is less than 1 %.
The authors do not think that the present paper requires such precision taking into account the reference of Sicard et al. (2009).

Page 29703, Line 9:
As above, how large is the statistical error on backscatter? Furthermore I do not understand the need for maximizing the error in particular the statistical one. You could calculate it precisely. I do not see the rationale for this Appendix, if at the end the error sources are not quantified.
>>>>>> One of the major comments of the editors of ACPD justifying the rejection of the first submission of this paper was the lack of error budget given the small AOT measured. The authors have made an effort to quantify those errors.
The statistical error on backscatter is calculated and provided for all profiles of backscatter (this is a common product of EARLINET). Following the referee's advice the statistical error on the AOT_VA has been calculated precisely instead of maximized following the new equations (A3) and (A4) given in the Appendix. The ranges in which the error on AOT_VA varies are roughly the same as those obtained by maximizing the error on the backscatter. This is due to 2 reasons: 1) the AOT layers are thin which limits the variation of the backscatter and its associated error, and 2) the error on the backscatter is nearly constant in the VA layers.