Response to Referee #1

We wish to thank the reviewer for his or her appreciation of our work and for the careful review. We address here the points raised by the reviewer:

Figures: we will address this issue partly by improving the figure files that we submit to ACP and partly by working with typesetters to ensure that figures are not compressed excessively to small parts of the page, when the paper is edited in final format.

We have now improved the images and increased the fonts on the axis labels. Our aim is to have Figures 5, 6, 8, 10 and 11 on a single page each, but this will have to be negotiated with production at a later stage.

P. 31747 L. 22. Thanks for this suggestion. We will amend the plots to show the reference ranges used.

We have now added horizontal lines indicating where reference is taken for each profile in Figure 4.

P. 31749 L. 1-3. Thanks for this feedback. We believe that the figure was reproduced too small and we will work with typesetters to have it printed on a full A4 page, possibly improving the symbol/character size if needed to make it more readable.

P. 31754 L. 8-13. Thanks for spotting this error. It is correct that the plume predicted by ECMWF-MACC is too high. We will therefore correct this description.

We have now reworded this description (see lines 358-361 of the new version of the manuscript).

P. 31755 L. 2. The MetUM actually displays a large gradient of extinction coefficient along the track, with very large values on the left of the graph and smaller values to the right. We will reword our description to better match the model plots.

We have now reworded this description (see lines 375-379 of the new version of the manuscript).

P. 31757 L. 12. Ok.

Now updated (line 439).

P. 31757 last paragraph. The reviewer suggests a more in-depth discussion of the model comparisons: How good do the models need to be? What is the aim of the model simulations? Why are we seeing some differences with the observations? Have we learned anything useful to improve the models?

We believe this comment to be very valuable. Highlighting the potential sources of error and areas of improvement would give more insight and benefit the article. We shall address this in a revised version and we give here a first outline of our thoughts.

To address the first question: How good do the models need to be?
This is very much related to the purpose of the simulations. The MetUM SAMBBA limited area model (LAM) wasset up specifically to support the SAMBBA field campaign and the primary purpose was to facilitate flight planning. The purpose of the ECMWF-MACC simulations is somewhat different. This is an operational global model, with forecast charts made available publically on the web on a daily basis, and for which specially zoomed charts can be requested for campaign support. In both cases the simulations are judged to be useful if they provide some skill in predicting the typical vertical distribution of aerosol, the regional distribution and day-to-day variability of aerosol loadings. The results in this paper clearly show some skill in simulating these aspects, even if the fine detail is not always captured. In this sense the simulations have proved to be useful and have served their purpose well.

The LAM simulations were a first attempt at generating forecasts of biomass burning aerosols with the CLASSIC prognostic aerosol scheme, and the LAM simulations provide an opportunity to test this potential advance in the Met Office's operational NWP atmospheric composition modelling capability. Therefore, in this paper we are looking at what benefits the prognostic treatment of biomass burning aerosols offer over an aerosol climatology. The fact that the MetUM can predict regional and vertical variations with some skill is very satisfying and important to document. Aerosol schemes can be sensitive to the host atmospheric model and its configuration (e.g. the grid-resolution, the representation of dynamics and other atmospheric processes). Therefore, it is important to evaluate the scheme with some detailed observations to see if the simulated spatial patterns are realistic when run at high resolution. Some aspects of the LAM aerosol simulations were also evaluated by Kolusu et al. (ACP 2015). They showed that the regional distribution and magnitude of AODs agree well with observations. The current study adds the evaluation of the vertical profile to this assessment. A first objective of our article is thus to highlight the strengths and weaknesses of the models in predicting the vertical structure, which is usually considered a weak point.The fact that biomass burning aerosol emissions needed tuning up by a factor of 1.7 in the LAM, and by 3.4 in MACC is also an interesting result to inform future model development.

A second application of these forecasts is investigation of the impacts of the prognostic BBA on the simulated meteorology. Kolusu et al. 2015 have investigated this with the MetUM, and found large impacts of the biomass burning aerosol on the radiation balance, improvements in forecasts of temperature and RH, and have jighlighted important changes in the representation of the regional hydrological cycle. The vertical profile of the aerosols is important in this respect, for its associated changes to the heating profile.

A third application of the forecasts is the prediction of air quality in Brazil the neighbouring countries, although this was not a focus of the SAMBBA campaign.

For both models, there are various sources of errors and the possible reasons as to why the simulated aerosol does not always provide a good match to the observed profiles:

- Biomass burning aerosol sources as initialized from GFAS: location, intensity, injection height. Note that other fire emissions datasets are also available (e.g. GFED3, FINN1), and their relative strengths and weaknesses are not fully known.
- Vertical transport and turbulent diffusion. The models don't take into account the impact of fire on localized convection or the formation of pyrocumulus clouds. A better representation of these processes is needed.
- Horizontal transport: uncertainties in the underlying NWP model
- Model resolution always places a limit on the representation of atmospheric processes and transport.

Both models use the same biomass burning aerosol emission sources allowing us to concentrate our interpretation of model-model differences on other aspects. For instance, the MetUM forecasts the thin plume observed on flight B733 at 2 km (between 50-200km along the flight track) while ECMWF does not, and the evidence points to the model vertical resolution as a plausible cause.
In the ECMWF-MACC model, injection heights are simulated interactively from a plume rise model and this led to some improvements in the vertical profile of aerosol in MACC for flights B741, B742 and B746 (a separate paper on this topic is in preparation): the height of the plume layers is closer to the observations. For flight B742, the MACC simulation is able to simulate two distinct smoke layers that were observed when using injection heights derived from the plume-rise model; otherwise it only simulates one broader layer.

In summary, this study illustrates the application of the observations to modeling applications and opens the door to a series of further studies. Indeed we are already planning to use it in an evaluation of the GFAS inventory (Remy et al, submitted to ACPD) and an evaluation of the UKCA-MODE aerosol scheme in the Met Office climate model (Johnson et al, in preparation).

This discussion has now been added in the conclusions (lines 472-518).

Response to Referee #2

We wish to thank the reviewer for his or her positive comment on this paper and his or her advice on how to improve it. In what follows, we address the points suggested by the reviewer:

Introduction: Amazonia vs. Southern Amazonia: we will reword the introduction accordingly.

Now reworded. See lines 21, 37, and 59 of revised manuscript.

Methodology – choice of reference value: the reviewer asks how the reference in the far range of the lidar profiles is chosen and why. This has been done using a new and non-standard method, which is dictated by the nadir-viewing geometry of the experiment (as opposed to zenith-viewing which is usually used with ground-based instruments). This method was described in a previous paper (Marenco, 2013), and a few examples using actual and simulated data were shown in that paper to justify the use of this method. Basically, the problem is that the forward solution in this geometry and at this wavelength is extremely unstable, to the point that it can't be used when an aerosol layer is deep and the extinction is large. The main concepts of the method are summarised in the manuscript (section 3.2), and the full detail can be found in Marenco (2013). Marenco et al (2014) also illustrates the application of the method in a comparison with CALIPSO retrievals, whereas in the present paper we use it on a larger scale. The method makes use of the slope method for a first estimate of the extinction coefficient at the reference interval. This is equivalent to the assumption of a well-mixed layer, but in that paper it is demonstrated that the method gives reasonable results if the reference is chosen within an aerosol layer avoiding its boundaries. The reviewer also asks to better specify the 50% uncertainty set for the reference value. This number cannot be evaluated from the measurements and needs therefore to be assumed. We have chosen a 50% statistical error (1-sigma) to indicate the possibility that we have a very large misjudgement of the reference value. Note that this is equivalent already to a very large total error (3-sigma=150%). As the results show, fortunately, however large this error it quickly decreases to very small values when moving inwards towards the lidar.

Some text has been added earlier in section 3 to clarify the reader on what approach is used for data inversion method (see lines 148-153 of revised). This was already contained in the paper but possibly too late in the reading (section 3.2).

Methodology – choice of reference value: the reviewer asks if we use a priori information (e.g. surface observations) to set the reference value. The answer is "no": lidar data only are used, and it cannot be otherwise because the aerosol is variable in time and space, the airplane travels thousands of km, and few ground-based sites are available in this region. In Marenco et al (2014), however, the reviewer can find a comparison of measurements with the unconstrained method and measurements constrained with AERONET: the constrained measurements fall within the estimated error range. The reviewer also asks if we “change everything until in clear air regions the Rayleigh value is obtained”. The answer to this second question is as follows: (1) the reference value is selected independently of Rayleigh scattering; (2) in the first iteration (see section 3.1), the lidar ratio only (not the reference value) is varied until Rayleigh scattering is matched in clear air; (3) then a second iteration (section 3.2) is done using the selected reference and the average lidar ratio that was found in the first iteration.

In the added text at line2 150-151 we now specify that the reference is set based on the lidar signals themselves. Moreover, at lines 153-155 we recall the 2014 paper that contains a comparison with AERONET-constrained retrievals.

Comparisons of the method with the Manaus lidar and CALIPSO: we unfortunately have not done coordinated measurements with the ground-based lidar, but we have indeed a comparison with CALIPSO that can be found in Marenco et al (2014). In the revised version of the paper, we shall remind readers of this previous use of the inversion method.

At lines 153-155 we recall the 2014 paper that contains a comparison with CALIPSO.

P. 31742 L. 5-18: Thanks. We shall add the reference suggested and mention ice-smoke interactions.

Now added (lines 57-58).

P. 31745 L. 3: Both photon counting and analogue are used in the Leosphere lidar, and they are glued at pre-processing. Gluing is done by choosing an overlap area based on photon-counting thresholds, and normalising the signals to each other in that area.

Now specified at lines 125-126.

P. 31749 L. 10: We are unsure what is the doubt expressed by the reviewer. For each vertical profile, we use the maximum of the extinction coefficient (the peak extinction) and the columnar aerosol optical depth (AOD) and then use the formula indicated in the paper to evaluate the layer depth. We cannot find the typesetting error suggested. This definition of layer depth yields quite stable results (i.e. less noisy than FWHM).

Please see line 234. The formula should read as follows:

$$\sqrt{2} \; * \text{(AOD)} / \text{(peak extinction)}$$

P. 31754 L. 11: This description will have to be reformulated as suggested by referee #1. In any case, if a plume has small horizontal or vertical dimensions and the model resolution is coarser, then it is more likely to be captured at the nearmost gridpoint than if the resolution is high.

This sentence is now reworded, as referee #1 spotted it to be not well describing the plotted data. This sentence has now disappeared (lines 358-361).

Figure 12: When preparing the new version of the manuscript, we will consider the possibility of moving the discussion earlier than in the conclusions. Thanks for this suggestion.

This has now been moved to section 5 (lines 406-411).

Technical suggestions: fully agreed.

Typesetting of 'ff' and 'ffi': we leave this for the production team.

Figures: We have now improved the images and increased the fonts on the axis labels. Our aim is to have Figures 5, 6, 8, 10 and 11 on a single page each, but this will have to be negotiated with production at a later stage.

Many thanks for your time!