We thank the referee for his constructive comments and respond in the following. Below, bold texts are the original referee comments.

1. This is a good review of wildfire plume injection height state-of-the-art. I found some minor technical issues, some grammatical issues, and a few significant omissions, all enumerated below. In my opinion, this is worthy of publication in ACP after revisions are made.

2. Introduction, P3, lines 9-10. You have four inventories mentioned, and only three references, so the use of "respectively" is ambiguous here. Also, the following reference might be worth including in the discussion: Kukkonen et al., Applicability of an integrated plume rise model for the dispersion from wild-land fires. Geosci. Model Dev., 7, 2663–2681, 2014, doi:10.5194/gmd-7-2663-2014

The list of inventories was corrected and Kukkonen et al 2014 included in the discussion (Section 5.2.2).
"... emphasis the importance of correct atmospheric profile as already mentioned by Kahn et al. (2007) or Kukkonen et al. (2014)."

3. Section 2, P5, line 21. I think you mean: " ... and therefore even though such large fires... "

Thanks Corrected.

4. Section 2, P5, lines 23-26. Since this is a review paper, another example is Mims et al. (IEEE Trans. Geosci. Remt. Sens. 48, 25-35, 2009).

This reference was added.
" ...average fire (Chen et al., 2009). Other evidences show that fires from agricultural or grassland vegetation type (usually less intense than those from boreal forest) can also generate plumes reaching the FT. Amiridis et al. (2010) show that half of the agricultural fires they observed over eastern Europe for the period 2006-08 reach heights above the PBL. Mims et al. (2010) report that 26% of the 27 Australian grassland fires they studied with various stereo-height retrieval algorithms rose above the PBL. In summary,..."

5. Section 2, P6, last paragraph. You could mention here something about the temporal and spatial resolutions that might be appropriate for coupling to CTMs. For example, a CTM might not resolve hour-to-hour differences in an individual smoke plume, even a large one.

We add in the last paragraph of page 4 a sentence on the spatial resolution of the parameterization and the host CTM.
"... atmosphere, where they can fully interact with ambient atmospheric circulation. In a recent study on fire emission transport Gonzi et al. (2015) use the GEOS-Chem CTM with a 2^o x 2.5^ohorizontal resolution and 47σ levels forming a vertical stretched mesh with a resolution of 150 − 200 m near the Planetary Boundary Layer (PBL). Since at these resolutions we cannot resolve the plume dynamics (100 m;Trentmann et al., 2006), parameterizations are therefore required to represent these 'smoke plume injection heights'. The aim of this manuscript is ..."

6. Section 3, P8, lines 12-15. You might note here that Pyrocumulus convection ismuch more common than pyro-cumulonimbus events. Another interesting and relatively recent reference is Peterson et al., BAMS 2014, "The 2013 Rim Fire: Implications for Predicting Extreme Fire Spread, Pyroconvection, and Smoke Emissions," doi:http://dx.doi.org/10.1175/BAMS-D-14-00060.1. And I recall Kahn et al. (2007) emphasize the role of entrainment as well as atmospheric stability structure.

The role of entrainment was mentioned in the list of processes affecting the plume dynamics. It is now on Page 7 line 22.

The work of Peterson et al.(2015) was also added to the discussion, in particular to emphasise the importance of ambient moisture in the plume development. See Section 3.

- "... vapour entrained into the plume from the combustion zone (water is a primary combustion product) and/or from the ambient fresh air (Freitas et al., 2007; Peterson et al., 2015)."
- "the ambient mid-level moisture (Peterson et al., 2015) and/or the presence of an approaching cold front (Fromm et al., 2010; Dirksen et al., 2009; Luderer et al., 2006; Peterson et al., 2015)."

7. Section 4.1.1, P10, lines 5-6. Another advantage of CALIOP is sensitivity to low-AOD aerosol layers aloft, even downwind from the source; it tends to retrieve higher plume elevations than the stereo-height or the thermal-band methods due to this sensitivity.

This was mention in the comparison between MISR and CALIOP in section 4.1.2. See Page 12 line 20-21 in the new manuscript.

8. Section 4.1.1, P10, line 20. Were there many cases where CALIPO had collocateddata over actual active fire areas?

The work of Amidiris et al 2010 is using collocated MODIS and CALIOP data. We are not aware of any other collocated data.

9. Section 4.1.2, P11, line 15. I think the Terra and A-Train are separated by about 3hours local time (10:30 AM vs. 1:30 PM), not 4.

Thanks, corrected.

10. Section 4.1.2, P11, line 18. MISR acquires 9 images at multiple angles (not 8). Also on P12, line 1

Thanks, corrected.

11. Section 4.1.2, Figure 4. I've seen many MINX stereo height images on their web site that are prettier than this one. Since this seems to be an arbitrary example, you might pick a more attractive one.

The picture was modified, and a nicer stereo height retrieval from Nelson et al. (2013) was inserted.

12. Section 4.1.2, last paragraph. I don't think the ATSR technique can retrieve or compensate for cloud motion (due to wind) between the multiple views (I recall they only have two views). More generally, are there any papers that use the split-window, CO2-slicing, or other thermal IR techniques for smoke plume heights? I know these are used for volcanic plumes. Also, proxy techniques for estimating smoke plume height, e.g., using CO sounding, might be discussed here.

A comment on the ATSR technique was added to specify that no wind correction was available.

“... (AATSR). Unlike MINX, M6 is not able to correct for the plume displacement induced by the ambient wind. However, it was estimated that such correction would lie in the error of the M6 algorithm (D. Fisher personal communication, 2015). M6 was applied...”

The split-window technique is efficient in volcanic plume as volcanoes are usually emitted for a long time, making long plumes with observable thin tail. In case of wildfire, the diurnal cycle of the fire activity induces more variable plumes, and therefore shorter observable tails that less likely to performed in the split-window algorithm. We do not know any attempt to use it for fire plume height retrieval.

CO profiles were use by Gonzi and Palmer, 2010. This is already referenced in the manuscript and appears in Section 2.

We do not know any studies using CO_2 slicing for the estimation of fire plume height.

13. Section 4.2, P13, last 3 lines. I think the LCL generally refers to a property of the ambient atmosphere. The plume must reach its own lifting condensation level to release latent heat, but this level might not be the same as that of the surrounding atmosphere. You might want to make that clear to readers.

We made the sentence clearer.

“The latent heat, which provides energy to the plume is a secondary source, can only be trigger if the plume reaches its LCL altitude. This LCL altitude can be different from the atmospheric LCL as water content and temperature profiles in plume usually differ from the ambient conditions.”

14. Section 4.2, P14, lines 5-7. The factor of 5 range for beta represents the range of flaming to smoldering in practice, but I’m not sure what you mean by a “constant partition,” as the phases within a fire can change rapidly.

The sentence was made clearer. We were discussing modelled fire and not fire in general.
“...the convective heat flux. In large scale model (> 100 m) wildfires are usually represented with a constant partition of convective and radiative energy emission (Trentmann et al., 2006; Freitas et al., 2007) with a ratio β of convection to radiation ranging from 1 to 5, so that Q_c is related to the radiative heat flux Q_r , $Q_c = \beta Q_r$. The values of β are essentially based on experimental studies performed at small scale (Freeborn et al., 2008; McCarter and Broido, 1965), and their applications to large scale remain uncertain. In a model sensitivity ...”

15. Section 4.2, P14, lines 7-11. In addition to the value of beta, wouldn’t this depend on the stability structure of the atmosphere, and maybe also the horizontal wind shear and conditions for entrainment?

We agree with the reviewer that other factors affect pyro-convection mechanisms. However in the sensitivity analysis of Luderer et al. (2006) we mentioned, the authors consider among other effects, the particular effect of the beta parameter (ratio convection to radiation), and this is what we refer to here.

16. Section 4.2, P14, last two lines. Maybe say "the best available option ... " to allow that someone might come up with another idea.
Thanks,edited.

17. Starting on P14, I noticed many minor grammatical errors, missing articles, in- correct word usage, etc. To take just a few examples: "In the later work [should be "latter"], the author use [should be "authors"] ... " And: P16, lines 7-8. Should be: " ... the different types of existing plume rise models. In particular ... the use of a plume rise model ... which aims to derive a smoke emission system for air quality models such as ... Here, we limit our review ... originally built ... " And P18, lines 4-5: " ... is that ambient shear a the sub-grid level is not represented. This certainly over-predicts ... more sensitive to ... " P 19, line 24: " ... does not differ from ... " P22, line 1. " ... the use of PRMv0 improves the results, when compared with ... " p22, line17. " ... of these studies emphasize ... " There are many more, so you might want to have the text copy edited.
Thanks. Edited. We thank the reviewer for pointing out the grammar errors. We revised thoroughly the text and corrected them.

18. Section 5.2.1, P17, line 19. Do you mean: " ... turbulent heat flux ... "
The sentence was modify to "but rather evaluate turbulent fluxes produced by the temperature anomaly …". In Rio et al.,(2010) Pyro-EDMF evaluates turbulent fluxes for the total water, the liquid potential temperature and the CO2 concentration, not only turbulent heat flux.

19. Section 5.2.1, P17, line 22. The convective heat flux is modeled as a fraction of the model flux averaged over the host model grid cell, or do you mean something else?
This sentence was made clearer. "The fire is considered as a sub-grid effect and its CHF is modelled as a fraction of the surface sensible heat flux averaged over the host model grid cell."

20. Section 5.2.2, P19, paragraph 2. Both Val Martin et al. (2010) and Kahn et al. (2007) indicate that: "smoke injected above the BL tends to accumulate in layers of relative stability." Would this be worth mentioning somewhere here?
Thanks for the comment. This was added in section 3 in the physical description of fire plumes. "(iii) the ambient atmospheric stratification which acts on the buoyancy of the initial updraft (Kahn et. al. 2008) and also on the later level of the detrained smoke as smoke injected above the PBL tends to accumulate in layers of relative stability (Kahn et. al. 2007 and Val Martin et. al. 2010)"

21. Section 5.2.2, P20, lines 1-3. Didn't Val Martin et al. (2013) test a range of fire- area and energy flux estimates using the PRMv1 model, and conclude the same thing (underestimation of high plumes)? It might be worth mentioning which methods did better.
We modified the plot and added the same analysis as done in Val Martin et al., (2012) so that comparison could be made between the performance of the Sofiev model and the PRMv1 model used in Val Martin et al.,(2012).

22. Section 5.2.2, Figure 7b. This figure is over-plotted too much to be especially useful. You might consider a density plot instead.

Thanks. The Figure was redrawn and extra info added to compare with the work of Val Martin et al.,(2012).

23. Section 5.3, P23, lines 1-5. I'm not sure the non-uniqueness of the classes is that large an issue, provided the underlying plume height data set is extensive enough to be partitioned based on regional properties (e.g., surface vegetation type) and environmental conditions. Also, some of the uncertainty in FRP for this application might be reduced or eliminated by the Dozier approach discussed in the next paragraph. Plus, see Note #20 above. In summary, this section could conclude with a position on using the statistical-model approach, with better-constrained FRP and maybe some adjustment based on the atmospheric structure.

If for a given fire event, the input information (i.e., FRP, stability and biome) belong to more than one pre-determined categories, it would be difficult to determine an injection height for this fire without a large uncertainty. We tried to make this clearer in the text.

I think we agree with the reviewer comment. The section concludes on the attractiveness of the approach, but emphasizes the crucial need of good input information for the determination of the categories, i.e. better-constrained FRP and more accurate adjustments based on the atmospheric structure.

24. Section 6, P24, line 21. Do you mean "solve" here – which would exclude statistical approaches – or "account for" – which could include such methods?

"account for", this was modified. Thanks.

25. Section 6, P24, line 27. I think you mean "Despite the demonstrated diurnal bias of the MISR-derived plume heights ... " That is, the issue is with MISR's lack of diurnal sampling, not with the stereo technique itself.

Right. It was modified.

26. Section 5, Captions for Figures 9, 10, and 11. I think Nelson et al. (2013) might be a better reference here, as it is more current.

This was modified.

27. Section 5, P26, top. Another example, for an even larger fire, is Figure 14 in Kahn et al. (2007).

A reference to this event was added in the discussion.

"...and a close inspection of the MODIS hot pixels shows that most of the fire pixels do not seem to be part of the plume dynamics (see Figure11). An even more extreme scenario is shown in Figure 14 of Kahn et al. (2007). In these particularly extreme fire cases, it seems that all fire..."