Response to Anonymous Referee #1

We would like to thank the referee for their review of this manuscript and their useful comments which have helped to improve the paper. Below we provide our response to the comments. Text in blue refers to text that has been added to or adjusted in the manuscript.

1) In Section 3.2.1 the methodology for the comparison between FRP-PIXEL product and other geostationary fire products is described. For FIR and WFABBA products, the less conservative classes of fire detections are excluded by the analysis. (For WFABBA, only filtered fire detections have been used in the analysis. This product has also different classes of outputs (Processed, Saturated, Cloud Contaminated, High Probability, Medium Probability, Low Probability). Have all of them been included in the analysis?). This comparison analysis shows that in general the FRP-PIXEL product generates a much higher number of fires detections with respect to the other geostationary fire products. In the light of this, do the authors think that it would be of interest to include in this comparative analysis also the less conservative detections for the other satellite fire products? If not related to the exclusion of the less probable classes of detection, what are, according to the authors, the main reasons of the differences observed with the other active fire products derived from the same Meteosat SEVIRI observations.

As suggested by the reviewer, we have updated the comparison between the different SEVIRI active fire datasets and now include all four variations of the WFABBA dataset. These are the inclusion of all active fire detections, all fire detections WFABBA 'filtered' dataset (where SEVIRI fire pixels that area only detected once during 24 hrs are removed) and WFABBA 'filtered' detections keeping only the higher probability fires (WFABBA flags 0 to 3) and high and medium probability fires (WFABBA flags 0 to 4). Figure 5 (in the manuscript; and shown below) has been updated to show the diurnal cycle of fire pixel detections which now includes the different variations of the active fire datasets. The full WFABBA dataset provides a marginally greater number of active fire detections than the filtered WFABBA dataset using all detections irrespective of detection confidence. Both of these datasets detects fewer active fire pixels than even the LSA SAF FRP-PIXEL dataset screened to only include pixels with an FRP >50 MW (which are generally the high confidence detections). Table 2 in the manuscript (and shown below) has been adjusted to include this new analysis :

Table 2 : Summary of active fire pixel detection errors of omission and commission of the four SEVIRI-derived active fire products explored herein (LSA SAF FRP-PIXEL product; Wooster *et al.*, 2015, WF-ABBA; Prins *et al.,* 1998, Fire Detection and Monitoring - FDeM; Amraoui *et al.,* 2010, and FIR Active Fire Monitoring; Joro *et al.,* 2008). Data were collected over the LSA SAF southern Africa geographic region during August 2014, when fire activity is widespread in this area. The MODIS active fire products (MOD14 and MYD14; Giglio *et al.*, 2003) acted as the independent data source for the comparison.

	FRP-PIXEL	WFABBA	WFABBA	WFABBA	WFABBA	FDeM	FIR
		All detection	Filtered	Filtered (flags 0-4)	Filtered (flags 0-3)		
SEVIRI fire pixels at coincident MODIS overpasses	33414	15610	13008	9736	8832	7664	7151
SEVIRI fire pixels detected by MODIS	29037	14521	12284	9369	8496	7260	6730
Commission error (%)	13	7	6	4	4	5	6
Omission error (%)	77	82	84	87	88	92	95

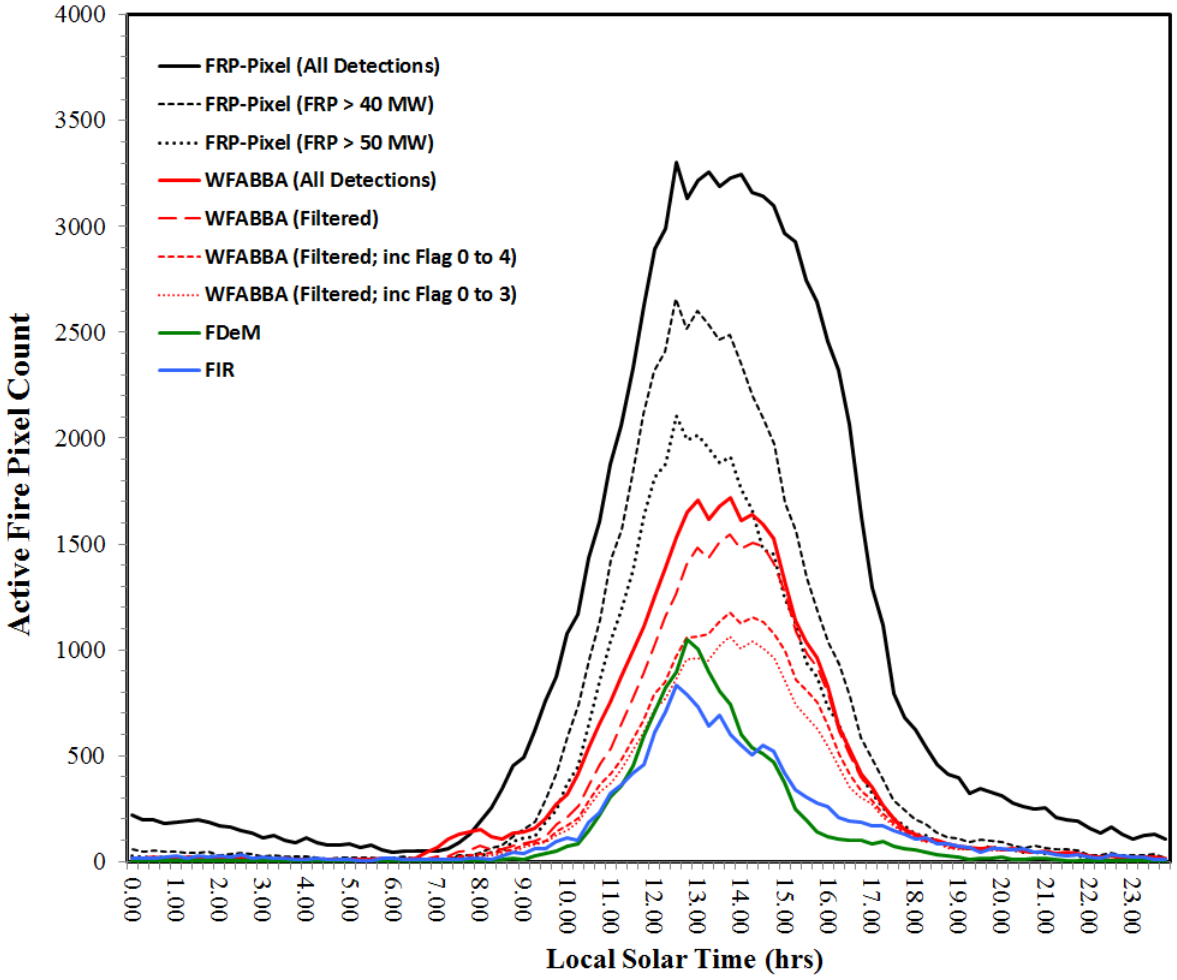

Figure 5: Diurnal cycle of active fire detections made by the four SEVIRI active fire products discussed herein over the LSA SAF southern Africa geographic region (Figure 1) on a single day (30th August 2014). The products are the LSA SAF FRP-PIXEL product (Wooster *et al.*, 2015), Wildfire-ABBA (WFABBA; Prins *et al.,* 1998), Fire Detection and Monitoring (FDeM; Amraoui *et al.,* 2010) and Active Fire Monitoring (FIR; Joro *et al.,* 2008). All confirmed active fire detections made in each product are included here for completeness, and results are shown in terms of the local solar time of detection. For the FRP-PIXEL product, three active fire time-series are shown; 1) all detections, and those only those detections from fire pixels

with FRP magnitudes 2) >40 MW and 3) >50 MW since it is known that increasing undercounting of active fire pixels occurs around these limits (Roberts and Wooster, 2008). For the WFABBA active fire detections, four versions of the dataset are included 1) all active fire detections, 2) the WFABBA 'filtered' detections where active fire pixels only detected once during 24 hrs are removed; and the WFABBA filtered detections keeping only 3) the high probability fires (flags 0 to 3) and 4) high and medium probability fires (flags 0 to 4). The LSA SAF FRP-PIXEL product detects a total of 89781 active fire pixels over the day which reduces to 53561 and 39461 when restricted to fire pixels with FRP magnitudes >40 MW and >50 MW respectively. For the WFABBA detections, the total number of active fire detections is 35759, the WFABBA filtered dataset contains 35759 detections which reduces to 30751 and 23957 when low and medium probability fire detections are removed. The FDeM and FIR detect only 13477 and 14645 active fire pixels respectively.

If not related to the exclusion of the less probable classes of detection, what are, according to the authors, the main reasons of the differences observed with the other active fire products derived from the same Meteosat SEVIRI observations?

As the revised Figure 5 indicates, the exclusion of fire pixel detections which are deemed to have a lower detection confidence reduces the total number of detections significantly. Differences in the number of active fire detections between active fire datasets are also the result of the inputs to the algorithm (e.g. cloud and land cover masks), pre-processing (e.g. atmospheric correction) and the algorithm itself (e.g. the type and value of the thresholds applied to discriminate fire affected pixels from non-fire affected pixels). The majority of the performance difference is very likely due to the detection algorithm methods and thresholds, but other factors are important. For example, Freeborn *et al.* (2014) found that the sensitivity of the SEVIRI cloud mask accounted for 30% of the LSA SAF FRP-PIXEL products omission rate when compared against the MODIS active fire product. It is difficult to exactly partition the cause of the differences in the number of fire pixel detections between the various SEVIRI active fire products, in part because with the exception of the LSA SAF FRP-PIXEL product no additional information is provided with the other products other than that of the detected active fire pixels. Therefore, we are unable to determine the effect of algorithm threshold difference on the numbers of detected fire pixels verses, for example, the different cloud mask used. Nevertheless, it is clear from the results that the LSA SAF FRP-PIXEL product is by far the most sensitive to the presence of active fires.

As a demonstration of the different algorithm sensitivities, we further analysed the per-pixel FRP frequency-magnitude distribution of each SEVIRI active fire product using one month of SEVIRI observations (August 2014). The FDeM and FIR products don't provide FRP estimates for detected active fire pixels, whilst the WFABBA algorithm uses a different approach (Dozier method) to measure FRP. Therefore, for consistency, we used the FRP values from the FRP-PIXEL product that are coincident with MODIS active fire detections as the basis for the assessment, and compared only active fire detections from the FRP-PIXEL, FDeM, FIR and WFABBA datasets that are spatially and temporally coincident with these in the analysis. This approach does not account for active fire detections present in the non FRP-PIXEL products but not in the FRP-PIXEL active fire detections. However, these are small in number since the FRP-PIXEL product delivers by far the greatest number of active fire counts.

Results, shown in Figure 1 below, indicate that the FRP-PIXEL product detects a greater

number of low FRP pixels compared to the other products, and can detect fires with an FRP are low as 30 MW with confidence. The "pins" mark the point at which there is a decline in algorithm fire detection performance, i.e. when the algorithm starts to become obviously weaker at discriminating the thermal radiance emitted from small and/or lower intensity fires. This occurs below around 50 MW for the filtered WFABBA product, around 80 - 100 MW for the FIR and FDEM products respectively. The percentage of pixels from each dataset which are not coincident with a FRP-PIXEL and MODIS active fire detection are 21% (WFABBA), 29% (FDeM) and 19% (FIR). Because this analysis only includes detections coincident with the FRP-PIXEL product and with MODIS fire pixels it includes only ~70 - 80% of the pixels present in all of the analysed SEVIRI active fire datasets, so the results do not fully represent these products and we do not include the analysis in the manuscript. However, it is broadly indicative of performance and we include it here for the benefit of the review and for completeness.

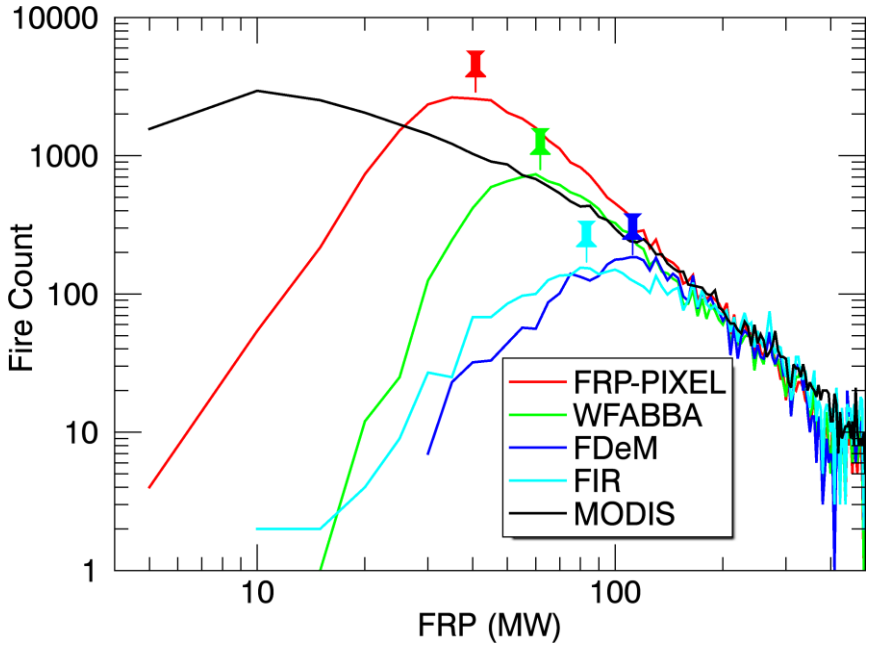

Figure 1. Frequency-magnitude distributions constructed from coincident active fire pixels detected by the FRP-PIXEL, WFABBA, FDeM and FIR products that are coincident with MODIS active fire detections over the southern Africa region during August 2014. The lower breakpoint of the distribution of each product (shown as a pin), coincides with the decline in each active fire detection product performance as the thermal radiance emitted from small and/or lower intensity fires cannot be distinguished from the background signal in each product. The FRP values at the lower breakpoint are 30 MW, 50MW, 100 MW, 80 MW for FRP-PIXEL, WFABBA, FDeM and FIR respectively.

2) In Section 5.2.2 and in Table 4 it is not clear which enhancement factor has been used to adjust the bottom-up aerosol emission estimates to those observed in top-down inventories.

We noticed an unnecessary adjustment factor in Equation 3 (β=1 in our study) and an unnecessary subscript in Equation 4, that describe the calculation of emissions from the

Peloponnese fires. For the Peloponnese region, Table 4 contains the emissions factors (η_s) for a number of gas and aerosol species for land cover type (*l*). In fact, the extratropical forest of this island is the only landcover type used in this case study. A constant (α) is used to adjust the bottom-up aerosol emissions to those observed via top-down inventories. In the manuscript this was given a subscript *l* (α_l) when, given the constant landcover type a global constant of 3.1 for aerosol emissions and 1 for gaseous smoke constituent was applied. This enhancement makes our calculation consistent with Ichoku and Kaufman *et al.* (2005) and just 10% lower than Kaiser *et al*. (2012). The unnecessary subscript l has been removed:

$$\eta_s = \alpha(s) \times \kappa_l(s) \qquad (4)$$

where $\alpha = 3.1$ for aerosol emissions and $\alpha = 1$ for gaseous smoke constituents. To make this clearer in the manuscript, the paragraph discussing Equation 4 has been adjusted to :

$$\eta_s = \alpha(s) \times \kappa_l(s) \qquad (4)$$

where κ_l is the land cover (l) specific emissions factor for species s and α is a constant which is used to adjust bottom-up aerosol emissions estimates to those observed in top-down inventories. A regionally varying bias occurs between bottom-up derived aerosol emissions and MODIS AOD measurements, requiring the former to be adjusted when being used in air quality or climate model simulations (Peterenko *et al.*, 2012). Yang *et al.* (2011) also found smoke emissions ($PM_{2.5}$) derived using the bottom-up approach was underestimated by a factor of three when compared to MODIS AOD retrievals. Kaiser *et al.* (2012) recommend a global aerosol enhancement by a factor of 3.4 as first-order correction. These values are also broadly consistent with differences of up to a factor of three found by Ichoku and Kaufmann (2005) using satellite observations of FRP and AOD compared to measurements of $c \times \kappa_l(s)$ derived from laboratory measurements. Here, we estimate emissions of organic matter and black carbon in exact agreement with Ichoku and Kaufmann (2005) by enhancing their emission factors for Andreae and Merlet (2001) with a factor of 3.1. According to the GFEDv3 land cover dataset, also used for our calculations in GFAS (Kaiser *et al.,* 2012), the fire affected region of Greece is classed as extratropical forest and the emitted species and relevant emissions factors are given in Table 4.

3) In Section 5.2.2 (pg.15939 line 9) the choice of releasing the smoke emissions in the lowest atmospheric level has not been discussed. Given the magnitude of the modelled fires, how much the authors think, the missing information of the plume penetration above the Planetary Boundary Layer, could have impacted the simulation of the smoke plume evolution?

It is likely that releasing the calculated smoke emissions into the lowest atmospheric layer, rather than at higher altitudes, does have an impact on the modelled concentration and transport. For example, Leung *et al.* (2007) and Guan *et al*. (2008) found that the inclusion

of plume injection height information in atmospheric transport models led a reduction in surface CO concentration around the source, since a greater proportion of the smoke emissions were lofted above the planetary boundary layer. However, actually estimating the correct smoke injection height to use in a particular fire simulation is a topic of much current debate and research, and is not yet solved. Even if for a particular event we could know the correct height from independent observations, this is not the case for the vast majority of fires that the Copernicus Atmosphere Service (CAMS) has to model.

Where it is actually attempted, the parameterisation of smoke plume injection height is currently typically achieved using direct EO measurements from stereo-imagery or lidar based methods, and sometimes through using empirical or deterministic models. Paugam *et al*. (2015a) for example provide a review of current methods used to estimate smoke plume injection height. Typically these depend on both the fire behaviour and ambient atmospheric situations. During the Peloponnese fires studied here, Lui *et al.* (2009) have provided an estimate of 2.5 km for the height of the plume closest to the wildfires using MISR observations acquired on the morning of 26th August. From Figure 9a in our manuscript, it is evident that the fire emitted FRP on the morning of the 26th August was high, but only around half that seen on the 25th August. This, coupled with morning image acquisition of MISR, when fire activity is typically less intense, suggests that the injection heights during the Peloponnese wildfires are actually likely to exceed those derived by Lui *et al.* (2009) from MISR; particularly during periods when the fire activity was most intense. However, currently we do not know the true injection height on the different days and times of the Peloponnese fires, and nor is this information available for other fires modelled in the Copernicus Atmosphere Service (CAMS) to which our study is a demonstration. One area of research, discussed by Sofiev *et al*. (2012) and Paugam *et al.* (2015b) is the use of FRP measurements within plume rise models in order to provide estimates of plume injection height over a wider range of fires. To reflect this, the following text has been added to the manuscript (Section 5.2.2) :

Smoke emissions from the Peloponnese fires were calculated using Equations 3 and 4, along with the emissions factors given in Table 4. The smoke emissions must be injected into the atmosphere at a particular height, or distribution of heights, and such injection height assumptions can have implications for the resulting spatio-temporal distribution of the emitted species. Leung *et al.* (2007) and Guan *et al*. (2008) demonstrated that use of more detailed plume injection height assumptions resulted in a reduction in near surface CO concentrations, since more plumes were assumed to be lofted above the boundary layer. Paugam *et al.* (2015a) provided a recent review of approaches to estimate smoke plume injection height, including the methods of Sofiev *et al.* (2012) and Paugam *et al.* (2015b) that use FRP measurements to characterise wildfire thermal properties related to plume rise. This research remains at a relatively early stage, but it appears that FRP measures may indeed have a role to play in characterising smoke plume injection height as well as the rate of emission of chemical and aerosol species. Here we retained the commonly used assumption that the calculated smoke emissions are injected into the lowest atmospheric level, since this is generally what has been

assumed in the series of MACC projects thus far (Kaiser *et al.,* 2012). The CAMS is anticipated to use injections heights from Paugam *et al.* (2015b) in the future.

Guan, H., Chatfield, R., B., Freitas, S. R., Bergstrom, R. W., Longo, K. M. (2008) Modeling the effect of plume-rise on the transport of carbon monoxide over Africa with NCAR CAM. *Atmospheric Chemistry and Physics*. 8. 6801–6812.

Leung ,F-Y, T., Logan, J. A., Park, R., Hyer, E., Kasischke, E., Streets, S and Yurganov, L. (2007) Impacts of enhanced biomass burning in the boreal forests in 1998 on tropospheric chemistry and the sensitivity of model results to the injection height of emissions. *Journal of Geophysical Research*. 112. D10313, doi:10.1029/2006JD008132

Liu, Y., Kahn, R. A., Chaloulakou, A. and Koutrakis, P. (2009) Analysis of the impact of the forest fires in August 2007 on air quality of Athens using multi-sensor aerosol remote sensing data, meteorology and surface observations. *Atmospheric Environment*. 43. 3310-3318.

Paugam, R., Wooster, M., Freitas, S. R. and Val Martin, M. (2015a) A review of approaches to estimate wildfire plume injection height within large scale atmospheric chemical transport models – Part 1. *Atmospheric Chemistry and Physics Discuss*ions. 15. 9767-9813.

Paugam, R., Wooster, M., Atherton, J., Freitas, S. R., Schultz, M. G. and Kaiser, J. W. (2015b) Development and optimization of a wildfire plume rise model based on remote sensing data inputs – Part 2. *Atmospheric Chemistry and Physics Discuss*ions. 15. 9815-9895.

Sofiev, M., Ermakova, T., and Vankevich, R. (2012) Evaluation of the smoke-injection height from wild-land fires using remote sensing data. *Atmospheric Chemistry and Physics*. 12. 1995-2006. doi:10.5194/acp-12-1995-2012.

4) From Section 5.3.1 SEVIRI saturation seems to be a major limitation of the FRPPIXEL product in describing the 2007 Greek fire episodes. Do the authors think that including MODIS-FRP derived emissions in the description of the selected fire episode could help to understand the impact that SEVIRI saturation has in underestimating the magnitude of the studied fire emissions?

The 2007 Peloponnese "mega fires" were very large and extremely intense - the greatest fire event recorded in Greece since the satellite era began we believe. Due to this, and the closely spaced nature of the fires which meant that many could be burning within a single SEVIRI pixel or group of pixels, the MWIR channel saturated on a number of occasions. Analysis shows that a maximum of 23% of the detected SEVIRI fire pixels were saturated in a single timeslot, on the day when the wildfires were at their most intense. Analysis of the total FRP record from the Peloponnese wildfires indicates that, on average, when SEVIRI and MODIS viewed the fires simultaneously SEVIRI measured 58% of the FRP measured by MODIS. Between the 24th and 26th August, when the wildfires were most intense (Figure 9a), MODIS made 13 overpasses and in total during co-incident observations SEVIRI measured 39% of the total FRP measured by MODIS.

The MODIS instrument offers the advantage of providing mostly unsaturated FRP observations, but is only capable of providing intermittent temporal sampling and thus estimating FRE from MODIS' FRP observations is not trivial. Baldassarre *et al*. (2015) report far better simulations of the Antalya (2008) fire in Turkey using the SEVIRI FRP-PIXEL product than with MODIS-derived FRP data, despite some saturation of the SEVIRI FRP observations. Estimating daily fire radiative energy (FRE) and fuel consumption using the temporally intermittent MODIS FRP measurements has been conducted in a number of different ways. For example, one approach is via temporal integration of the daily FRP using an assumed diurnal fire cycle model (e.g. Kaiser *et al.,* 2012; Vermote *et al*., 2009). Freeborn *et al.* (2011) provide a further method based on derived conversion coefficients, but all such approaches are often best suited to coarse spatial (e.g. 1° × 1° grid cells) and/or temporal resolution (e.g. 1 day, 1 week) derivations, not derivations for single fire events. This is in part evident from Figure 9a in the manuscript, where a fire diurnal cycle is evident for the Peloponeese fires but one that may not be described by an assumed diurnal cycle model (e.g. a modified Gaussian as assumed by some of the above MODIS-based methods). A coarser regional scales, such assumed fire diurnal cycles maybe more realistic fits to the true nature of fire activity. It is also the case that the FRP-to-FRE methods described by Vermote *et al.* (2009), and Freeborn *et al.* (2011), actually tend to underestimate SEVIRI FRE-derived fuel consumption estimates - when the latter have been adjusted to account for the low-spatial resolution bias of SEVIRI (Roberts and Wooster, 2008; Freeborn *et al.,* 2011).

Following the approach used in the Global Fire Assimilation System (GFAS) of the Copernicus Atmosphere Service (CAS; Kaiser *et al*., 2012), we have adjusted the "raw" dry matter (DM) combustion estimates obtained using our remotely-sensed FRP measures by a land cover specific coefficient (Equation 2 in the manuscript) such that the totals are in line with those provided by the Global Fire Emissions Database (GFED, v3.1). This will aid somewhat aid accounting for the impact of underestimation of FRP caused by sensor saturation, and we have added the following text to Section 5.3.1 of the manuscript to indicate that the SEVIRI estimates are affected by saturation :

Between the 24th and 26th August, when the wildfires were most intense, MODIS made 13 overpasses and SEVIRI measured 39% of the total FRP measured by MODIS. This demonstrates the massive scale of these fires, particularly given that the SEVIRI's pixel area over the region is ~ 14 km^2.

Due to the aforementioned SEVIRI MWIR channel saturation, the SEVIRI FRP-derived fuel consumption estimate is considered a minimum estimate.

Freeborn, P. H., Wooster, M. J., Roberts, G. (2011) Addressing the spatiotemporal sampling design of MODIS to provide estimates of the fire radiative energy emitted from Africa. *Remote Sensing of Environment*. 115. 2. 475 – 489

Kaiser. J. W., Heil, A., Andrae, M. O., Benedettie, A., Chubarova, N., Jones, L., Morcrette, J-J., Razinger, M., Schultz, M. G., Suttie, M. and van der Werf, G., R. (2012) Biomass burning emissions estimates with a global fire assimilation system based on observed fire radiative power. *Biogeosciences*. 9. 5125-5142. doi:10.5194/bg-9-5125-2012

Vermote, E., Ellicott, E., Dubovik, O., Lapyonok, T., Chin, M., Giglio, G. and Roberts, G. (2009) An approach to estimate global biomass burning emissions of Organic and Black Carbon from MODIS Fire Radiative Power. *Journal of Geophysical Research*. 114. D18205. doi:10.1029/2008JD011188.

Minor comment: Page 15921, Line 11. ": : : further from the Meteosat sub-satellite point (SSP) : : :"

The manuscript has been altered accordingly