Author Response to Anonymous Referee #1

The authors would like to thank Referee #1 for the thorough review of our manuscript. The general as well as the specific comments will help to improve the quality of the paper. In the following we list the referee comments together with our response.

General Comments:

#1: During discussion of radiative forcing (RF), it is important to note that this forcing is from direct aerosol impacts only (I believe). You touch on this on page 6717, lines 5 -10 with: "the RF represents exclusively the radiative perturbation that is introduced by wildfire emissions […] regardless of anthropogenic sources." However, I believe it would be beneficial to the reader if it was made clear that the RF being reported was from direct interference with radiation, and not indirect impacts on clouds or on semi -direct impacts on atmospheric warming.

We agree that a more detailed explanation of the RF is required to clarify the results presented in Chapter 5. Therefore, page 6717 lines 4-6 will be replaced by the following statement:

"Here, the RF represents exclusively the radiative perturbation that is introduced by wildfire emissions (BC, OC, SO_2), while anthropogenic emissions are kept constant. The radiative perturbation which is attributed to direct aerosol-radiation interference is referred to as "clear sky" RF; the RF which also includes indirect and semi-direct effects due to aerosol-cloud interaction is referred to as "total sky" RF. Aerosol-induced changes in atmospheric temperature profiles are implicitly included in both RF parameters, but due to our nudging towards reanalysis data every six hours, they are partly suppressed."

#2: It is not well described in the manuscript why the "SURFACE" simulations actual emit aerosols into the first two layers of the model, and not just the surface layer. It seems that if you wanted the more "extreme" lower - boundary you might just emit aerosols into the lowest layer. Probably a sentence or two justifying this decision would clear up this confusion.

There are two independent reasons why we chose to distribute wildfire emissions in simulation 'SURFACE' into the first two model layers, not just the surface layer. First, in ECHAM6-HAM2 all anthropogenic emissions from industrial sources are injected into the lowest as well as the second-lowest model layer in order to realistically account for the concentrated heat and emission release at a certain stack height (see Stier et al., 2005 and Dentener et al., 2006). Similarly, a large fraction of wildfires includes crown burning of trees representing an emission release comparable to industrial emission sources. Therefore we think that wildfire emission release should be treated similar as emission release from industrial sources even for the SURFACE scenario.

Second, preliminary test runs prior to this study had shown that a very intense wildfire emission release concentrated at one specific model layer may result in an ECHAM6-HAM2 model collapse. These instabilities might be attributed to radiative imbalance, but we didn't further investigate the issues as the problem disappeared when wildfire emissions were distributed into at least two model layers. We will add a short statement in the description section of simulation SURFACE in the revised manuscript.

Specific Comments:

#1: Page 6700, line 5, "Does the […] matter on the global scale" – what do the authors mean by matter? Perhaps this should say something like: "Does the […] enhance, dampen or change the sign of the globally averaged climate response" (?)

We admit that our statement is lacking precision and we will directly apply your suggested change in the revised manuscript.

#2: Page 6700, line 8. As above, what is meant by "is appropriate for […]"?

We will replace the sentence "What degree of […] is appropriate for global climate modeling?" by "What degree of complexity in plume height parametrization is required to capture the emission height impact on aerosol long-range transport and atmospheric radiation in global climate models?"

#3: It may be too late to change this, but is there a reason the authors chose "Aerosol Optical Thickness" instead of "Aerosol Optical Depth"? The MODIS product (and in general the aerosol community) seems to prefer "AOD".

We are aware of this inconsistency in terminology, but we think that AOD and AOT are still both coequally used in current literature. A Web of Science search yields 5,317 entries for AOT and 10,646 for AOD, for google scholar it's 205,000 for AOT versus 191,000 for AOD. We prefer to use AOT, because in our opinion the term *thickness* implies an integrated measure (which AOT measured by MODIS and AERONET represents), whereas the term *depth* is usually applied for specific layers or vertical locations, not necessarily integrated measurements of specific quantities. Furthermore, also on the NASA webpage, AOD and AOT are likewise used for the description of the MODIS product, see http://earthobservatory.nasa.gov/GlobalMaps/view.php?d1=MODAL2_M_AER_OD .

In the revised manuscript, we will replace the introduction of "Aerosol Optical Thickness (AOT)" by "Aerosol Optical Thickness (AOT, also referred to as Aerosol Optical Depth, AOD)". However, we will consistently keep AOT everywhere else in the manuscript.

#4: Page 6702 -6703: This is an interesting problem; namely, why the emissions dataset needs to be multiplied by 3.4 to produce reasonable AOD in the GCM, as discussed extensively in Kaiser et al. (2012). Tosca et al. (2013), referenced later in the manuscript, encountered a similar problem, though the multiplier for CESM simulations was closer to 2.0. Randerson et al. (2012) postulated that part of this problem may be due to an under - representation of small fires in global emissions datasets, but this does not seem to address the underlying issue of why (most all) GCMs produce low biases in AOD

with reasonable emissions data input. I wonder if the authors have any insight on this problem, and whether it may be worth mentioning?

It would be highly desirable to further investigate this general emission estimate problem in a separate study. Basically, the problem could arise from an underestimation in emission fluxes, an overestimation of removal rates, shortcoming in the model representation of aerosol micro-physical properties, or a combination of these. Several studies (e.g. Schwarz et al., 2013, Hardenberg et al., 2012) have demonstrated the limited skills of aerosol-climate models to realistically simulate BC transport and removal. However, as atmospheric aerosol lifetimes do match observations fairly well, shortcomings in transport and removal processes cannot be expected to be the only reason for this problem. While the 3.4 factor is required for wildfire *aerosol* emissions, it is not required for *trace gas* emissions when using the same model setup of ECHAM6 just with an extension of a trace gas module (personal communication with Martin Schultz, Institute for Energy and Climate Research, Juelich). Therefore, we assume that emission factors, which convert biomass burned into aerosol emissions, are one major source of this bias. GFAS as well as GFED emission estimates are based on vegetation species emission factors. These emission factors in turn are based on a very limited number of experimental case studies, see e.g. Akagi et al., 2011 and references therein. A comprehensive revision of these emission factors based on up-to-date experimental measurement techniques could contribute to an improvement in all satellite-derived emission inventories as well as an assessment of the aerosol microphysics representation in the model. We will briefly mention these ideas in the revised manuscript.

#5: Page 6703, lines 10 -15: Why do the authors choose to inject aerosols into the bottom two layers of the model for the SURFACE simulations, rather than just the lowest layer? As described above, this should probably be clarified.

See general comment #2.

#6: Page 6709, line 13: Are these S. Hemi. changes positive or negative?

These changes are negative, see Fig. 2 (e). We will add this information in the revised manuscript.

#7: Page 6719, lines 10 -15: The wording here is a bit hard to follow. Tosca et al. (2013) do not calculate a 'true' surface RF. What they do calculate is the net change in surface shortwave due to fire aerosol emissions. Their calculations are therefore a response, not a true RF. They do, however, calculate TOA RF (as the authors mention). The way this sentence (line 14) is worded is confusing; it may be helpful to add "However" to the beginning of "In contrast to our study […]"

We tried to improve this text passage (page 6719, lines 10-15) in the revised manuscript as follows: "Tosca et al., 2013 compared a simulation based on GFEDv3 wildfire emissions to a zero wildfire emission control run to estimate the net change in surface shortwave fluxes in the Community Earth System Model (CESM). The authors only considered a prescribed wildfire emission release at the surface. The difference in net short-wave fluxes at the surface was found to be $-1.3 \pm 0.2\,Wm^{-2}$

leading to a surface cooling of $-0.13 \pm 0.01\,Wm^{-2}$. However, in contrast to our study the sign of the TOA RF was positive ($+0.18 \pm 0.10\,Wm^{-2}$).

#8: Page 6721, lines 1 -3: Why do the authors conclude that a 5 -25% change in deposition rates represents only "limited sensitivity." To me, 25% seems to be a reasonably large change.

We agree that this statement might me misleading, particularly in the conclusions section. According to your comment, we dropped 'only' and 'even' in the mentioned text passage and replaced 'limited sensitivity' by 'moderate sensitivity'.

However, the studies by Hardenberg et al., 2012 (ACP) and Bourgeois and Bey, 2011 (JGR) showed that the ECHAM5-HAM1 model bias in polar BC deposition rates was substantially larger than the changes introduced by the emission heights found in this study. The changes from ECHAM5-HAM1 to ECHAM6-HAM2 (see Zhang et al., 2012) cannot be assumed to compensate these biases. Therefore, on page 6712, line 25, we will additionally mention the results of Hardenberg et al., 2012 (ACP) and Bourgeois and Bey, 2011 (JGR) in order to relate the plume height sensitivity of BC deposition rates to the presumably larger general model bias in aerosol long-range transport.

#9: Page 6721, lines 15 -23: As I understand it, the calculated TOA RF is for all fire, most of which is probably natural; a back of the envelope guess might assume that 40 -50% of global fires are anthropogenic. Since you compare your modeled RF to the IPCC anthropogenic forcing of 0.9Wm -2, it would be worth mentioning that your calculated RF values would be cut approximately in half (?) if we consider only anthropogenic fire contributions.

We admit that referring to the IPCC anthropogenic forcing might be misleading in this context. As you mentioned, the human impact on wildfire emissions in the past and its contribution to the total anthropogenic aerosol RF is difficult to quantify. We agree that the IPCC anthropogenic aerosol RF does not represent a wise reference point. Therefore, we will exchange page 6721, lines 21-23, sentence "These changes in RF are small […]" by

"These changes in TOA RF are small compared to the spread of the overall wildfire emission RF in other state-of-the-art climate models (-0.3 to $+ 0.2\,Wm^{-2}$)."