We appreciate the comments from the Reviewers. Below is our response provided in **bold**.

Anonymous Referee #1

This paper explores the sensitivity of black carbon (BC) aerosol radiative forcing (RF) to the complexity of the radiation code (2-stream versus 4- to 16-stream). As such, it nicely extends and illustrates the conclusions of recent multi-model radiation code inter-comparisons. It is well written and should be accepted to ACP with very minor revisions, which I describe below:

1. I am somewhat confused by exactly how BC RF is calculated. It seems you use aerosol present-day (PD) and pre-industrial (PI) aerosol distributions from AeroCom Phase II? However, optical properties and meteorological data (i.e. clouds) come from OsloCTM2 and ECMWF, respectively? Perhaps some rewording could be done to make this a bit more clear in Section 2. Further, when calculating BC RF, you seem to imply in Section 3.2 that BC RF is calculated when other scattering aerosol types are included, and I"m not sure how this is done; some elaboration would be appreciated.

Response: The description has been clarified both connection between OsloCTM2 and AeroCom in relation to ECMWF and that other aerosol components are included. The following text is now is section 2:
'In the global simulations we used meteorological data from ECMWF, and specified aerosol optical properties (Myhre et al., 2009a) and aerosol distribution from the OsloCTM2 chemical transport model (Skeie et al., 2011). To study the impact of the radiation code on global mean RF of BC, input fields and results from OsloCTM2 part of AeroCom Phase II for several aerosol components were used.'

2. This point in Section 3.2 about how the BC RF changes much more in the presence of scattering aerosols (more so then when alone) is very interesting and could be highlighted a bit more in the abstract and conclusions.

Response: The following text is included in the conclusions:
The underestimation for BC is largest in the presence of scattering components and the same applies to gases with solar absorption. However, under clear sky condition only gases in the lower levels of the troposphere with solar absorption in UV and visible region where Rayleigh scattering is strong causes underestimation of similar magnitude as for BC.

3. Did you look at other aerosol types in terms of RF at all? It would be interesting to mention the effect of the radiative transfer code complexity on other aerosol types as well.

Response: We have added the following sentence: 'For pure scattering aerosols 2-stream simulations varies with solar zenith angle (see Randles et al. (2013)) and surface albedo compared to 8-stream simulations, but on a global mean 5% stronger negative RF for anthropogenic sulphate aerosols.'

4. Figure 1; it would be useful to print the global mean on the plot.

Response: We would like to avoid global mean of the ratio of the 2-stream and 8-stream simulations. Global mean numbers of 2-stream and 8-stream simulations are given in the text.

5. Figure 2: The legend is confusing. What is meant by “All eff”? and “BC only”? The thinnest lines are difficult to see. There are also 8 things in the legend but only 6 lines on the plot (a). Perhaps make the lines in (b) a different color to distinguish them from lines in (a).

Response: The figure and caption have been updated, and are now hopefully clearer. (There were 8 lines in the plot, as stated, however two of them are almost overlapping – as can be seen from the ratio in panel b.) The new caption reads:

Figure 2: (a) BC RF normalized by abundance, as a function of altitude. Solid lines: 8-stream simulations. Dashed lines: 2-stream simulations. Colors represent all sky and clear sky conditions, and whether a full atmospheric simulation including Rayleigh scattering, water vapour and background aerosols was performed (“Full sim.”), or if BC was the only radiatively active agent (“BC only”). (b) Ratio of 2-stream to 8-stream simulation results, for the four cases shown in panel (a).

O. Boucher (Referee)

The manuscript is short, well written and presents some interesting results. The authors present a convincing explanation for the BC radiative forcing discrepancy they observe between 2-stream and multi-stream RT codes. The manuscript being short, it nevertheless raises a number of questions, which once answered by the authors, should improve the manuscript further and warrant pulication.

The authors use 8-stream and 16-stream configurations of their multi-stream RT code. However I could not find a comparison between these two versions. It seems that the 8-stream configuration is used for Figs. 1 and 2, while the 16-stream configuration is used for Fig. 3. It would be good to know the minimum number of stream that is required to get a stable solution. In other words, does the BC RF converge and how fast at low surface albedo when the number of streams is increased.

Response: We have replaced the 16-stream simulations by 8-stream simulations in Fig 3. The following text is included in section 3.3:
'The agreement between 8-stream and even higher number of streams such as 16-stream simulations is generally within 1%, except for very small absolute RF values. Simulations with 4-streams are generally close to 8-stream simulations.'

If the authors' explanation for what they observe for BC is correct, then shouldn't that apply also to gaseous absorption in the solar spectrum? Wouldn't gaseous absorption by O_3, H_2O, CO_2, CH_4, ... also be underestimated by 2-stream RT codes? A lot of the gaseous absorption occurs in the near-infrared where scattering is less, and O3 absorbs everything where it absorbs in th UV, so the effect should be less. But H2O has many absorption bands below 1 _m. Maybe the authors should look into this or at least mention it. This could have some implications for the water vapour feedback in climate models (e.g. in polar regions, where surface albedo is large, T change is amplified and q change per unit of T change is large).
I wonder if the effect found on an absorbing aerosol (i.e. BC) is not associated with a countereffect on scattering aerosols (e.g. sulfate). Does the radiative effect of scattering aerosols depend on the number of streams used?

Response: We have added the following text into the manuscript where additional simulations have been performed for gases absorbing in the solar spectrum in section 3.2:
'For pure scattering aerosols 2-stream simulations varies with solar zenith angle (see Randles et al. (2013)) and albedo compared to 8-stream simulations, but on a global mean 5% stronger negative RF. '

Further the following has been added to section 4:

'The underestimation for BC is largest in the presence of scattering components. This also applies to gases with solar absorption. However, under clear sky condition, underestimation of a similar magnitude to BC will only be caused by gases with solar absorption in UV and visible region where Rayleigh scattering is strong. Thus ozone in the lower troposphere is the only gas that is substantially influenced by the number of streams in the radiative transfer simulations. For a global increase in water vapour by 20% in the lowest 1-2 km of the atmosphere, the difference between 2-stream and 8-stream simulations is found to be less than 1%.'

Finally I wonder how general the results here are. The authors mention that most climate models use two-stream approximations with the delta-M method, but do not substantiate their findings. I would be surprised if all two-stream models use the same approximation to truncate the forward peak of the aerosol phase function. Isn't there some spread in the Stier et al. and Randles et al. papers among the two-stream models?

Response: We have modified the sentence somewhat from 'are around' to 'could be up to 10%' since it is somewhat difficult to quantify how many of the models have used a standard 2-stream setup. With some very few exceptions the 2-stream simulations in Randles et al. shows a consistent pattern with underestimations compared to multi-stream simulations. In Stier et al. other factors overwhelm this 10% factor such as surface albedo, clouds, water vapour. For a better quantification of the underestimation in Randles et al. we have added the following sentence.

'The clear sky results for selected profiles and solar zenith angles in Randles et al. (2013) showed an average model underestimation between 12 and 15% compared to benchmark model simulations.'

How does the two-stream RT code used here compare with the other two-stream RT code used in Stier et al. and Randles et al.? I think the statement on lines 6–10 of page 26179 should be more substantiated or qualified.

Response: As mentioned above the following sentence has been added: 'The clear sky results for selected profiles and solar zenith angles in Randles et al. (2013) showed an average model underestimation between 12 and 15% compared to benchmark model simulations'

The caption for Figure 1 needs to be revised. It looks like a ratio rather than a relative difference. Same for Figure 2b.

Response: Thanks, corrected.

Anonymous Referee #3

General comment:
This paper illustrates that the radiative forcing of BC in climate models is sensitive to the number of streams in the radiative transfer model. It shows that the 2-stream approximation led to a 10% underestimation of BC RF compared to executing the radiation code with 8-streams. The low bias is more severe over low albedo surfaces and under cloud-free conditions. The physical mechanism for this bias is fully explained and explored in the paper. The authors show that the low bias arises because 2-stream models forward scatter too much radiation leading to lower mean path length of solar radiation through the atmosphere, and hence reduces the opportunities for absorption. The result will presumably apply to any absorber, BC, aerosol, gas or otherwise so I believe it is most likely a generalised result. Though the bias is not large (_10%) it is nonetheless worth quantifying as it helps to explain discrepancies identified in an earlier study by Randales et al. (2013). In this way the present manuscript helps to complete the story / understandings gained from recent radiative forcing model intercomparisons.
It is also useful information for climate model developers to have a paper quantifying the magnitude of this error associated with the 2-stream approximation. The paper is nice and short and really was quite straighforward and interesting to read. I therefore consider the paper to be of significant scientific merit and to be appropriate for publication in ACP with minor revisions.

Specific comments:
I am not sure I agree with the final statement of the conclusions "Even so, radiation schemes in global models should be improved to provide more accurate calculation of present and future radiative forcing due to BC". Uncertainty in BC FR is really dominated by problems in simulating the BC abundance and optical properties. Also, increasing the number of streams in a radiative transfer model would increase the overall CPU and complexity of a climate model and so this may not be justifiable at this stage of climate model development. My interpretation is that the inclusion of multiple stream radiation is shown to improve the accuracy of BC RF accuracy and that the benefits of multi-stream radiation are certainly "worth considering" in aerosol-radiation and climate studies.

Response: We have modified the sentence to reflect that this could be considered. It may be other possibilities for improvements than using multi-stream radiative transfer schemes.

In the conclusions could the authors be a little more specific about the size of errors in BC RF stemming from BC abundance, BC optical properties, and BC forcing efficiency. For example, how does the error associated with the 2-stream aproximation weigh up against problems in getting the vertical distribution right or other factors that affect the efficiency of BC absorption in the atmosphere. It is maybe asking too much from a short paper but a little extra information I think would help to weigh this up in terms of model development priorities.

Response: The following is added: Burden of BC and the normalized RF has a standard deviation of the order of 50% relative to mean values for the 15 global aerosol models in AeroCom Phase 2 (Myhre et al., 2013).

It would be quite useful to know how much additional cost (CPU) is consumed by opting for multi-stream over 2-stream. I know this would be different from model to model but even so, I think this information would help give an idea of the affordability of 8-streams. It would be nice to see the headline results in a table (global mean RF of BC, for clear and all-sky conditions showing the 2 and 8 streams, and then the ratio of 2-stream/8-stream). Maybe would be interesting to put the TOA SW fluxes in the table too to highlight any change in TOA balance from switching between 2 and 8 streams. In section 3.2 the authors have done something to keep the albedo of the clouds the same or at least to keep the TOA reflected SW flux the same in the 2-stream and 8-stream cases. Can the authors explain how they did this? I am not sure this is the best thing to do here. After all, if a climate model is upgraded to 8-stream then the change in cloud reflectance will occur and I think we really want to know the overall impact including any changes to scattering by clouds.

Response: The additional computer time between our 2-stream and 8-stream simulations is relatively small. However, this is not very relevant for climate model simulations where specific 2-stream approaches are adopted unlike our multi-stream code which is flexible in terms of number of streams. We have added that the modifications to clouds is due to constraint by measurements with the following addition: 'and close to measured fluxes'

In figure 3 the results switch now to comparing 2-stream with 16-streams rather than 2 versus 8. Was there a reason for this switch? It'd seem better maintain the same 8-stream configuration throughout the results.

Response: All results now in 8-streams.

The wording and presentation of results could be improved in a few places to make it a little clearer and easier to follow (see minor suggestions below).
Minor suggestions for clarity of text / figures:
1. Abstract line 5. Replace "absorption by BC" with "absorption by BC in the atmosphere".
Response: Modified as suggested.

2. Abstract line 6. Replace "RF by 10%" with "the positive RF of BC by 10%".
Response: Modified as suggested.

3. Abstract line 9. Replace "at high surface albedo" with "over high surface albedo".
Response: Modified as suggested.

4. Introduction, line 16. Replace "mainfold" with "many".
Response: Modified as suggested.

5. Introduction, page 2, line 6, "Further, one of the radiative transfer codes...". Can the authors say which one it was and/or what kind of radiative transfer method it had. Are the authors in this study using a similar or same radiative transfer to the Randales et al. study?
Response: The following included: These two codes where denoted as number 3 and 4, respectively in Randles et al. (2013) and used in the current work

6. Section 2. Can the authors be a little more specific on how the aerosol fields were generated. It comes across as a bit confusing and vague in the way it is worded. Can they explain what aerosol species are included in the simulation other than BC.
Response: More details added, see response to Reviewer 1.

7. Section 3.1, page 2, line 13. Replace "The radiative effect of clouds" with "The albedo of clouds".
Response: Modified as suggested.

8. Section 3.2, page 3, line 13. Replace "enhances" with "increases".
Response: Modified as suggested.

9. Section 3.3, page 3, line 12. Replace "the importance multiple" with "the importance of multiple".
Response: Modified as suggested.

10. Figure 1. Would be good to be more explicit in the description of the figure caption that the data shown is the ratio of BC RF with 2-stream / BC RF with 8-stream. Why not also put a title on the figures "Clear-sky" "All-sky".

Response: Caption modified stating that it is ratio.
11. Figure 2. There are maybe too many lines on here, or at least it is not totally clear what the "BC only" and "All eff" refer to. The BC only case is maybe a distraction from the main point so maybe this could be omitted and simply explained in the text. The caption also has an awkward string "(a). (b)" in it. Maybe put the "(a)" at the beginning of the sentence.

Response: The figure and caption have been updated, and are now hopefully clearer. The new caption reads:

Figure 2: (a) BC RF normalized by abundance, as a function of altitude. Solid lines: 8-stream simulations. Dashed lines: 2-stream simulations. Colors represent all sky and clear sky conditions, and whether a full atmospheric simulation including Rayleigh scattering, water vapour and background aerosols was performed ("Full sim."), or if BC was the only radiatively active agent ("BC only"). (b) Ratio of 2-stream to 8-stream simulation results, for the four cases shown in panel (a).

12. Figure 3. Could the authors produce these results with higher resolution of surface albedo intervals? The lines look like they are going only in steps of 0.1 but the results seem very sensitive in the region where suface albedo is between 0 - 0.2 and I would guess are not at all linear as the lines presently suggest.
Response: Fig 3 has been improved with a factor of 2 higher resolution than in the submitted version.