We would like to first thank the anonymous reviewer for the suggestions to improve our manuscript. We address the reviewer's comments below. The original comments are in **bold font** and our responses are in *italic font*.

This manuscript presents an analysis of airborne in situ data collected during a series of flights over western Europe. This is an interesting and useful analysis that investigates the sensitivity of direct radiative forcing to variations in parameters. That said, there are some issues with the analysis methodology that require modifications to the manuscript. The description of the techniques used is not complete, thorough or clear, and some assumptions are made that are not explained adequately. For this reason I suggest minor revisions to the manuscript.

Below are my primary concerns. Page numbers refer to the "printer-friendly" pdf version.

1) p. 3 line 21. Why is the un-humidified nephelometer "assumed" to represent dry conditions? The RH is typically directly measured (albeit not very well) within the TSI nephelometers. And lacking an RH measurement, one could calculate it knowing the temperature change between the ambient conditions and the instrument conditions. An f(RH) value calculated when the "dry" condition is actually at 50% RH, for example, could be different by >30% from a true value measured with a dry scattering measured at RH<10%. What is the sensitivity of the analysis to this assumption?

The measurements of the standard nephelometer are assumed to be applicable to "dry" aerosol, although the sample is not actively dried during the flights, since the sample is at lower than ambient relative humidity due to the effect of heat from the nephelometer lamp and electronics, the dynamic heating through deceleration of the input flow which reaches the instrument and the increased temperature of the sample lines compared to ambient air. Thus, due to the dehydrating nature of sampling on the aircraft, the standard nephelometer will measure aerosol scattering at low RH (<40%), which is sufficient to dry out many atmospheric aerosols (Osborne et al., 2007; Haywood et al., 2008).

Since the standard nephelometer (assumed to measure "dry" aerosol") is operated in series with a "wet-nephelometer", the hygroscopic scattering growth factor, f(RH), can be estimated as the ratio of the scattering coefficient measured in the wet-nephelometer to the scattering coefficient measured by the standard nephelometer, and plotted as a

function of RH to obtain a hygroscopicity curve for each flight of the campaign. To minimise the uncertainty due to the unknown RH of the sample in the dry nephelometer, only sections of SLRs where the RH measured in the dry nephelometer is less than 30% are used. The "measured" scattering for "ambient" aerosol at a given relative humidity is then derived by increasing the scattering from the nephelometer by the growth factor indicated by the hygroscopicity curve.

Highwood et al. (2012) found that, since the ambient relative humidity during the EUCAARI-LONGREX campaign was generally below 70%, f(RH) was relatively modest and therefore, its influence on the scattering and/or single scattering albedo was quite small (differences in the average values compared to the dry values being well within the uncertainty bounds on either quantity).

Therefore, the uncertainties that include our calculated values of the aerosol radiative effect related to the way in which aerosols are represented in the ES96 radiative transfer model are going to be mainly due to other assumptions rather than to the assumption that the measurements of the standard nephelometer are applicable to "dry" aerosol. Thus, since aerosols in the ES96 radiative transfer model, which was used in our estimations of the aerosol radiative effect during EUCAARI-LONGREX, are represented by their optical properties and vertical profile, in terms of mass mixing ratio, we have only tested the sensitivity of the calculated aerosol radiative effect to the main assumptions (aerosol composition, size distribution and/or single scattering albedo) made to obtain these.

References:

Haywood, J., Bush, M., Abel, S., Claxton, B., Coe, H., Crosier, J., Harrison, M., Macpherson, B., Naylor, M. and Osborne, S.: Prediction of visibility and aerosol within the operational Met Office Unified Model. II: Validation of model performance using observational data. Q.J.R. Meteorol. Soc., 134: 1817–1832. doi: 10.1002/qj.275, 2008.

Highwood, E.J., Northway, M.J., McMeeking, G.R., Morgan, W.T., Liu, D., Osborne, S., Bower, K., Coe, H., Ryder, C., and Williams, P.: Aerosol scattering and absorption during the EUCAARI-LONGREX flights of the Facility for Airborne Atmospheric Measurements (FAAM) BAe-146: can measurements and models agree? Atmos. Chem. Phys., 12, 7251–7267, doi:10.5194/acp-12-7251-2012, 2012.

Osborne, S. R., Haywood, J. M. and Bellouin, N.: In situ and remote-sensing measurements of the mean microphysical and optical properties of industrial pollution aerosol during ADRIEX. Q.J.R. Meteorol. Soc., 133: 17–32. doi: 10.1002/qj.92, 2007.

2) p. 3 lines 23-26. This description of the calculating the ambient extinction is hard to understand. Reading Highwood et al. (2012) and Esteve et al. (2014) it appears that the humidified nephelometer is scanned through a range of RH values and a parameterized f(RH) equation fitted to the data. Then the ambient extinction is calculated using the parameterized curve and the ambient RH. Is this correct? Is it done for every second of flight? Please describe in more detail, but succinctly, how ambient extinction is calculated. What is the sensitivity of your analysis to the "dry" f(RH) value?

During the EUCAARI-LONGREX campaign, a "dry" and "wet" nephelometer system was operated aboard the BAe-146 aircraft. The sample flow through the wet-neph is humidified to a set value between 45 and 95 %. During a straight level run (SLR), the humidity is either cycled through a range between these values, or set at a fixed high level. This allowed us to plot f(RH) (defined as the ratio of the scattering coefficient measured in the wet-neph to the scattering coefficient measured by the dry-neph) as a function of RH, and to fit the data to Model 2 from Kotchenruther et al. (1999)

$$f(RH) = \sigma_{s,d}\left(1 + a\left(\frac{RH}{100}\right)^{b}\right)$$

where $\sigma_{s,d}$, a, and b are fitting parameters to the data. To minimise the uncertainty due to the unknown RH of the sample in the dry nephelometer, only sections of SLRs where the RH measured in the dry nephelometer is less than 30% are used to estimate the growth factors for scattering. Moreover, the data is averaged into 2% mean RH bins to minimise also the high variability due to small scale variatons in f(RH), and RH is assessed and used to scale the 2% mean RH bins to account for particle loss between the nephelometers and any zero offset issues. Therefore, a hygroscopicity curve is obtained for each flight of the campaign. The scattering for "ambient" aerosol at a given relative humidity is then derived by increasing the scattering from the nephelometer by the growth factor indicated by the hygroscopicity curve.

As suggested by the reviewer, a more detailed, but succinctly, explanation of this procedure has been included in the manuscript:

"A second TSI 3563 integrating nephelometer ("wet-nephelometer") is operated with a humidified RH between 45 and 95% in series with the first nephelometer (e.g. Haywood et al., 2008). During a straight level run (SLR), the humidity is either cycled through a range between these values, or set at a fixed high level. This allows us to estimate the hygroscopic scattering growth factor, f(RH), as the ratio of the scattering coefficient measured in the wet-nephelometer to the scattering coefficient measured by the standard nephelometer, and this can be plotted as a function of RH to obtain a hygroscopicity curve for each flight of the campaign. To minimise the uncertainty due to the unknown RH of the sample in the "dry nephelometer", only sections of SLRs where the RH measured in the "dry nephelometer" is less than 30% are used. The "measured" scattering for "ambient" aerosol at a given relative humidity is then derived by increasing the scattering from the nephelometer by the growth factor indicated by the hygroscopicity curve. A more detailed description of this process can be found in Highwood et al. (2012)."

In order to minimise the uncertainty due to the unknown RH of the sample in the "dry nephelometer", we have only used sections of SLRs where the RH measured in the "dry nephelometer" is less than 30% to estimate the growth factors for scattering. Moreover, Highwood et al. (2012) found that, since the ambient relative humidity during the EUCAARI-LONGREX campaign was generally below 70%, f(RH) was relatively modest and therefore, its influence on the scattering and/or single scattering albedo was quite small (differences in the average values compared to the dry values being well within the uncertainty bounds on either quantity). Therefore, our calculated values of the aerosol radiative effect are more affected by the uncertainties due to the way in which aerosols are represented in the ES96 radiative transfer model than by our estimation of the "ambient" aerosol. Thus, since aerosols in the ES96 radiative transfer model, which was used in our estimations of the aerosol radiative effect during EUCAARI-LONGREX, are represented by their optical properties and vertical profile, in terms of mass mixing ratio, we have only tested the sensitivity of the calculated aerosol radiative effect to the main assumptions (aerosol composition, size distribution and/or single scattering albedo) made to obtain these.

References:

Haywood, J., Bush, M., Abel, S., Claxton, B., Coe, H., Crosier, J., Harrison, M., Macpherson, B., Naylor, M. and Osborne, S.: Prediction of visibility and aerosol within the operational Met Office Unified Model. II: Validation of model performance using observational data. Q.J.R. Meteorol. Soc., 134: 1817–1832. doi: 10.1002/qj.275, 2008.

Highwood, E.J., Northway, M.J., McMeeking, G.R., Morgan, W.T., Liu, D., Osborne, S., Bower, K., Coe, H., Ryder, C., and Williams, P.: Aerosol scattering and absorption during the EUCAARI-LONGREX flights of the Facility for Airborne Atmospheric Measurements (FAAM) BAe-146: can measurements and models agree? Atmos. Chem. Phys., 12, 7251–7267, doi:10.5194/acp-12-7251-2012, 2012.

Kotchenruther, R.A., Hobbs, P.V., and Heff, D.A.: Humidifcation factors for atmospheric aerosols off the mid-Atlantic coast of the United States. J. Geophys. Res., 104 (D2), 2239–2251, doi:10.1029/98JD01751, 1999.

3) p. 4 lines 2-3. Why is the under-wing PCASP instrument assumed to measure a dry size distribution? This instrument is under the wing specifically to minimize perturbation (via inlet losses and thermodynamic changes) to the ambient aerosol. What is the sensitivity of your results to this assumption?

The PCASP dries the aerosol to some degree through deceleration of the sample air, interaction with dry sheath air, heating by the de-icing heaters and by the electrical components associated with the probe (Strapp et al., 1992). Since Strapp et al. (1992) suggest that in laboratory tests with the de-icing heaters switched on aerosol particles are completely dried out, it is standard practice to operate the PCASP of the FAAM BAe-146 aircraft with the de-icing heaters switched on throughout the measurements, so we can assume that we are measuring the dry aerosol size distribution.

We have changed the sentence about the dry aerosol size distribution:

"Considering the dehydrating nature of the PCASP (Strapp et al., 1992), this is taken to be the dry aerosol size distribution."

A new reference has been added to the References section as well:

"Strapp, J.W., Leaitch, W.R., and Liu, P.S.K.: Hydrated and Dried Aerosol-Size-Distribution Measurements from the Particle Measuring Systems FSSP-300 Probe and the Deiced PCASP-100X Probe. J. Atmos. Oceanic Technol., 9, 548-555, doi: 10.1175/1520-0426(1992)009<0548:HADASD>2.0.CO;2, 1992."

References:

Strapp, J.W., Leaitch, W.R., and Liu, P.S.K.: Hydrated and Dried Aerosol-Size-Distribution Measurements from the Particle Measuring Systems FSSP-300 Probe and the Deiced PCASP-100X Probe. J. Atmos. Oceanic Technol., 9, 548-555, doi: 10.1175/1520-0426(1992)009<0548:HADASD>2.0.CO;2, 1992.

4) p. 4 lines 2-3. What is the error in the size distribution associated with using calibration particles of a fixed refractive index (latex beads? ammonium sulfate?) when the atmospheric aerosol has a different refractive index (which can be calculated from the AMS measurements)? If this is described in detail in an earlier publication, at least summarize the results here.

Rosenberg et al. (2012) described two methods for calibrating optical particle counters (OPCs), such as the PCASP, based on the principle that an OPC measures an electrical pulse height which is related to a particle's scattering cross section. They also provided a method, which is based on a probability density function, to modify the OPC bin boundaries when the scattering properties of the measured particles are different to those of the calibration particles due to differences in refractive index or shape.

Calibrating the PCASP using these methods showed that the bin centers are systematically higher than those reported by the manufacturer using the refractive index for polystyrene latex spheres as in the manufacturer's specification, and this calibration may change by up to 20% when routine maintenance is carried out. However, the change in the calibration over time is less than 5% when calibrations are performed before and after projects with a duration of ~1 month.

We have included a new sentence about this:

"We have taken into consideration the known difference in bin sizing from the calibration latex spheres based on our estimate of the refractive indices of the aerosol as described in Rosenberg et al. (2012). The drift of this offset has been shown to be small over the duration of a campaign."

A new reference has been added to the References section as well:

"Rosenberg, P. D., Dean, A. R., Williams, P. I., Dorsey, J. R., Minikin, A., Pickering, M. A., and Petzold, A.: Particle sizing calibration with refractive index correction for light scattering optical particle counters and impacts upon PCASP and CDP data collected during the Fennec campaign, Atmos. Meas. Tech., 5, 1147-1163, doi:10.5194/amt-5-1147-2012, 2012."

References:

Rosenberg, P. D., Dean, A. R., Williams, P. I., Dorsey, J. R., Minikin, A., Pickering, M. A., and Petzold, A.: Particle sizing calibration with refractive index correction for light scattering optical particle counters and impacts upon PCASP and CDP data collected during the Fennec campaign, Atmos. Meas. Tech., 5, 1147-1163, doi:10.5194/amt-5-1147-2012, 2012.

5) p. 4 lines 26-29. Why is the aerosol mass profile calculated from the scattering measurements and an (assumed?) mass extinction efficiency? The measurements include direct observations of aerosol mass (submicron at least) from the AMS and SP2 instruments. Why not use these measured mass values directly?

Although the instrumentation aboard the FAAM Bae-146 aircraft during the EUCAARI-LONGREX campaign included an Aerodyne Time-of-Flight Aerosol Mass Spectrometer (ToF-AMS), which measured the total non-refractory aerosol mass (e.g. organics, nitrate, sulphate, ammonium and chloride), and a Single Particle Soot Photometer (SP2), which measured the mass of sub-micron refractory black carbon, we can't use these measurements of the aerosol composition to estimate the vertical aerosol profile in terms of mass mixing ratio that we need in our simulations with the ES96 radiative transfer model because the vertical profiles during the flights were relatively short compared to the sampling frequency of these two instruments, and thus we would have very few data points in our measured aerosol vertical profile.

Therefore, we need to use other measurements available during the flights in order to specify the vertical aerosol profile in terms of mass mixing ratio, and for this we need to convert first the vertical profile of the aerosol scattering measured by the FAAM BAe-146 aircraft to a vertical profile of the extinction coefficient by dividing by the single scattering albedo measured at an appropriate altitude (closest SLR to the vertical profile), and then the extinction coefficient is converted to a mass mixing ratio using the mass extinction coefficient (which is calculated with a Mie scattering code within a

framework that includes the measured composition and dry size distribution for the closest SLRs to each vertical profile and grows the aerosol according to parameterisation of growth factors from literature).

6) p. 6 line 6. Is this calculation for TOA? Please be clear

At the beginning of Section 4, we introduce the concept of shortwave aerosol radiative effect (or aerosol radiative effect, ARE, as we call it throughout the manuscript), which is the change in net shortwave irradiance due to the presence of aerosols, ARE = NET^{aer} – NET^{clr}, where NET=SWD-SWU. This definition is valid for both the aerosol radiative effect at the surface (SFC) and the top of the atmosphere (TOA).

Since this is a general definition, we haven't changed this sentence. However, as suggested by the reviewer, we have made clear later on our manuscript that our calculations of the aerosol radiative effect are for the surface and the top of the atmosphere:

"Having some confidence in our representation of the aged aerosol from the success of radiative closure obtained in Section 3, we would like to estimate the aerosol radiative effect at the surface (SFC) and the top of the atmosphere (TOA) for this and other profiles from the EUCAARI-LONGREX campaign."

"Table 3 shows the "ambient" aerosol optical depth at 550 nm (AOD) from the aircraft profiles and "dry" single scattering albedo (SSA) and specific aerosol extinction (kext) at 550 nm from the closest SLR to each profile derived from the measurements of the aerosol scattering and absorption made by the FAAM BAe-146 aircraft, together with the diurnally averaged values of the aerosol radiative effect (ARE) and radiative efficiency (RE) at the surface and the TOA over Europe during EUCAARI-LONGREX."

7) p. 8 line 15. It is not clear how the choices of ranges of the various input parameters for the sensitivity calculations were made. Here it states, "The sensitivity to the various assumptions made in the radiative transfer model is estimated by repeating our calculations of the aerosol radiative effect using a different assumption, and then comparing the new results with the original values and calculating the difference between them." What were the ranges of input parameters, and how were they chosen? How many different perturbations were considered? Is this done by hand-chosing a few values, or was there a comprehensive Monte Carlo simulation done? Do the perturbations tested

represent the measured variation in that parameter, or just the uncertainty in the measurement, as is suggested by Table 2? What does it mean to vary the "size distribution"? Is the number varied, the mean diameter, or the standard deviation? What is the sensitivity to each of these components of the size distribution function? This part of the manuscript is the heart of the analysis and the methodology needs to be much clearer. If at all possible, I recommend that the sensitivity study be separated into portions due to experimental uncertainties and measured (geophysical) variability. This would be very interesting–how well can we determine the direct effect (measurement uncertainty) and how much does it vary (geophysical variation)?

In our calculations of the aerosol radiative effect we have made various assumptions, mostly related to the way in which aerosols and other parameters are represented in the ES96 radiative transfer model. Table 1 shows a summary of these. In our original calculations, we used what we think is the best representation we can use of those parameters considering the in-situ meteorological and aerosol measurements obtained by the FAAM Bae-146 aircraft during the EUCAARI-LONGREX campaign. In order to analyse the sensitivity of the calculated aerosol radiative effect to the various assumptions made in the radiative transfer model, we have repeated our calculations using a different, but plausible, assumption, and then comparing the new results with the original values and calculating the difference between them. Table 1 shows a summary of the alternative assumptions made for our analysis. This type of analysis has previously been done in Esteve et al. (2014).

Table 1. Assumptions made about the way in which aerosols and other parameters are represented in the ES96 radiative transfer model to calculate the aerosol radiative effect.

	ORIGINAL	***TEST***
ES96 spectral resolution	*220 bands (0.2 - 10 µm)*	*6 bands (0.2 - 10 µm)*
aerosol composition & size distribution	*measurements for closest SLRs to each vertical profile*	*flight-mean values*
surface albedo	*MERRA reanalysis*	*values from pyranometer measurements during SLRs close to the surface (not always available)*
single scattering albedo	*measurements for closest SLRs to each vertical profile*	*flight-mean values*
hygroscopicity of organic aerosol	*Brooks et al. (2004) (depends on the ambient RH)*	*Gysel et al. (2007) (doesn't depend on the ambient RH)*

As suggested by the reviewer, the analysis of the sensitivity of the calculated aerosol radiative effect has been separated into two different parts in Section 5:

"The values of the aerosol radiative effect calculated in Section 4 include uncertainties due to the modelling methods used in the radiative transfer model and the spatial variations in the aerosol.

...

5.1 Sensitivity to modelling methods

The ES96 model covers the range 0.2 - 10 µm, but its spectral resolution is flexible, with differing resolutions typically being used for off-line calculations (220 bands, as in the previous sections of this paper) and within NWP and climate models (more typically 6 bands). Fig. S1 in the Supplement shows the spectral dependence of the aerosol extinction, the single scattering albedo and the phase function over the ranges 0.2 - 10 µm. To test the sensitivity of the calculated aerosol radiative effect to the spectral resolution of the information used in ES96, we have repeated our calculations using the spectral resolution used for the HadGEM climate model and the Met Office global forecast model. The changes in the calculated ARE (SP_RES in Figure 5) compared to

our original higher spectral resolution version are large: an increase of ~17% on average and up to 20 – 21% at the surface and the TOA.

The sensitivity to the surface albedo is tested here by repeating our calculations of the aerosol radiative effect with values of the surface albedo from pyranometer measurements of upwelling and downwelling irradiance during SLRs close to the surface whenever it was possible, instead of using values from the MERRA reanalysis. The change in the calculated ARE (SFC_ALB in Figure 5) is either an increase or a decrease, and at the surface is relatively low, 6% on average and up to 12%. However, the calculated ARE at the TOA is much more sensitive to the surface albedo, ~20% on average and up to ~65%.

Esteve et al. (2014) found that one of the largest sources of uncertainty in the calculated scattering is the assumed hygroscopic growth factor for organic aerosol. To test the sensitivity of the calculated aerosol radiative effect to the hygroscopic growth factor for organic aerosol used in the Mie calculations of the aerosol optical properties included in ES96, we have repeated our calculations using the value from Gysel et al. (2007), which is independent of the ambient relative humidity, instead of the hygroscopic growth factor from Brooks et al. (2004). The change in the calculated ARE is similar at the surface and TOA (GF_OC in Figure 5) being an increase of ~5%.

5.2 Sensitivity to spatial variations in aerosol

Aerosol optical properties are included in the ES96 radiative transfer model by calculating them with a Mie scattering code based on their composition and size distribution (Esteve et al., 2014). To test the relative importance of representing the spatial variability of aerosol optical properties, we have repeated our calculations using flight-mean aerosol compositions and size distributions instead of the measured ones for the closest SLRs to each vertical profile of EUCAARI-LONGREX used here. The change in the calculated ARE for a particular SLR when using local vs flight-mean values (AVG_AER in Figure 5) is an increase of ~7% on average and up to 12% at the surface and the TOA, which agrees with the change in the calculated aerosol optical properties of ~7% on average and up to 14%.

The ES96 radiative transfer model requires the vertical aerosol profile in terms of mass mixing ratio as an input. This is derived from the measured scattering profile but this calculation requires inputs of SSA and mass extinction coefficient. To test the sensitivity

of the calculated aerosol radiative effect to the single scattering albedo used in this conversion we have repeated our calculations using flight-mean values of the SSA instead of those matching the closest SLRs to each vertical profile of EUCAARI-LONGREX (resulting in SSA differences of 2.6% on average but up to 11%). (Note that the optical properties of aerosol applied in the radiative transfer calculations are not changed here, only the column mass of aerosol is changed). The change in the calculated ARE (SSA_AER in Figure 5) is a decrease of ~2% on average and up to 5% at the surface and the TOA. The change in the SSA is buffered by other factors in the model."

A sentence in the Conclusions has been changed as well:

"We also quantified here the uncertainties in our calculations due to the modelling methods used in the radiative transfer model and the spatial variations in the aerosol."

References:

Esteve, A.R., Highwood, E.J., Morgan, W.T., Allen, G., Coe, H., Grainger, R.G., Brown, P., and Szpek, K.: A study on the sensitivities of simulated aerosol optical properties to composition and size distribution using airborne measurements, Atmos. Environ., 89, 517–524, doi:10.1016/j.atmosenv.2014.02.063, 2014.