We thank reviewer 2 for their comments and suggestions that have provided a useful perspective on the study. We specifically address their points below.

This paper deals with a new lightning parameterization to be used in global models. The authors use the ERA reanalysis model that supplies the best estimate of global large scale files every 6 hours. The authors compare their new ice flux parameterization to other commonly used parameterizations in the literature.

I have a number of major comments, and a few minor comments:

1) All lightning parameterizations depend on two factors. First, the reliability of the link between some cloud parameter (height, precip, ice flux, etc.) and lightning in reality; and second, the ability of the model to reproduce correctly these cloud parameters. We may have a perfect relationship between CTH and lightning in the field, but if the model or reanalysis does not produce the correct CTH, the lightning will be incorrectly predicted. On the other hand, if the cloud parameters are correctly predicted, it is possible the parameterization is not exact, leading to erroneous results. The authors do not address this point at all. How well does the reanalysis predict CTH, ice flux, convective precip, etc.?? Before knowing this it is impossible to say which parameterization is better or worse. Is one better than the other because the parameterization is better, or because the reanalysis does a better job predicting the cloud parameter? How do we know?

RESPONSE. The reviewer raises a pertinent issue regarding model parametrisations in general. In order to estimate lightning we are required to use cloud variables that are still under development themselves and contain unquantified errors. We agree that there are two aspects to the errors in forming a parametrisation: 1) The errors in input data, and 2) the accuracy of the relationship between a given meteorological variable and lightning. Over such large scales it is difficult to provide estimates of which input variables are best produced but we have provided a substantial evaluation synthesis to section 2.1 which acknowledges the error source and outlines the existing knowledge of some of the variables. Several of our results regarding existing parametrisations, CTH specifically, are not unexpected and have been found in other studies which suggests that the evaluation is a reasonable indication of the relative performance of the parametrisations.

State-of-the-art reanalysis data is the best input available over regions as large as the tropics and for the range of input parametrisations needed to evaluate model performance. However, several known issues exist with the data will affect the performance of parametrisations during the evaluation regardless of the correctness of their relationship with lightning. Broadly, it can be assumed that where observations are less dense there will be less accuracy, e.g. over Africa and the oceans.

There have been specific studies, both from ECMWF and independently, to assess the performance of the reanalysis. Dee et al. (2011) provides an in-depth evaluation of ERA-Interim with respect to observations and improvements upon its predecessor, ERA-40. Several improvements, with reference to International Satellite Cloud Climatology Project (ISCCP) observations, have been made to the representation of clouds. Those relevant here are:

- Reduced tropical cloud cover by 5-15% resulting from an improved hydrological cycle and the introduction of ice supersaturation which delays the formation of ice clouds (Tompkins et al., 2007).
- Increased tropical land cloud cover by 20-30% resulting from increased high cloud from improved deep convective triggering and additional low cloud from a new boundary-layer scheme.

There are few other studies directly evaluating the cloud properties as observations of clouds have their own large uncertainties. However, Schreier et al. (2014) and Ahlgrimm and Köhler (2010) have studied trade cumulus clouds represented in ERA-Interim. These are not directly related to deep convective clouds but at least can hint at some of the differences between the reanalysis and observations. A main finding was that the population of trade cumulus is overestimated by ERA-Interim while the cloud fraction was underestimated by ERA-Interim. Meanwhile, the cloud top was biased high by about 500m.

Much more research has been done on the evaluation of precipitation. Dee et al. (2011) identified the hydrological cycle as a weakness in ERA-40 and substantial improvements were made by ERA-Interim. Both the mean daily precipitation rate, compared to the Global Precipitation Climatology Project (GPCP), and the mean total column water vapour, compared to microwave imager satellite retrievals, have improved. There are still biases remaining over the tropical oceans, specifically around the western pacific and southeast Asia where the precipitation rate is up to 5 mm day^{-1} greater in some parts. The time series of precipitation rate over total land performs well while during the 2007-2011 period there is an overestimation by approximately 0.4 mm day^{-1} over total ocean when compared to GPCP.

Many other studies have looked at precipitation compared to observations and other reanalysis sets. Generally ERA-Interim was found to perform well. The applicable findings regarding biases in precipitation were:

- Overestimation of small and medium precipitation but underestimation of high amounts compared to rain gauge data in the tropical Pacific (Pfeifroth et al., 2013) and to the Tropical Rainfall Measuring Mission Satellite in the tropics (Kim et al., 2013).
- When compared to rain gauge data, ERA-Interim was found to outperform satellite and other reanalyses in Australia but not in southeast Asia (Peña Arancibia et al., 2013).
- In a study of southern African precipitation it was found that reanalysis data did a reasonable job in representing the seasonal precipitation cycle (Zhang et al., 2013). However, the best preforming reanalysis product was used to suggest that ERA-Interim

may have excess moisture convergence and therefore rainfall in the eastern equatorial Atlantic. It was also suggested that ERA-Interim may have a northward biased ITCZ in the Indian Ocean.

- A study into representation of global monsoon precipitation found that reanalysis performed reasonably with ERA-Interim having the most skill (Lin et al., 2014). However, it did note that none of the reanalyses identified the increasing tendency in northern African.

While the errors in variables such as updraught mass flux remain unknown we can assume that ERA-Interim has remaining problems with the hydrological cycle over the western pacific and southeast Asia and that this is likely to affect all input variables. This will be considered when drawing conclusions from the evaluation.

We have included a summary of this information in section 2.1.

In addition, every model and reanalysis will have its own meteorology, own cloud parameterization, own physics. So how can we say that if a parameterization is better in the ERA reanalysis, it will also be better in any other GCM? And if we cannot reach such a conclusion from this study, how could other researchers benefit from this study? Can these results be transferred to another climate model? If anything it would be interesting to compare the ICEFLUX parameterization in many different models to see how variable or stable to lightning distributions will be. If the results vary enormously from model to model (due to the different cloud parameterizations) then what is your conclusion?

RESPONSE. We fully support the testing of the parametrisations in a range of models and are actively engaging with modelling groups to do so; some preliminary investigation using the UKCA chemistry-climate model suggests that many results from this paper are transferable. The conclusions from this paper give evidence that a more physically based parametrisation can work on a large scale and may be preferable to existing parametrisations. Its development provides the opportunity to benefit from future improvements in modelled cloud microphysics and dynamics. In addition several CTMs use ERA-I as input so these models benefit directly from the comparison.

2) The parameterization itself is not clear. How does the ERA reanalysis determine the ice mass at -25C? As you know water can exist down to -40C as supercooled water. This is also important for the electrification of clouds. When does water become ice in the model? What is the threshold? What fraction of the water at -25C is liquid? How sensitive are the results to this threshold? This needs to be addressed if you are proposing for others to use this parameterization.

RESPONSE. The parametrisation is designed to use the 440hPa in accordance with the ISCCP definition of deep convection. We clarify that we have not used the -25℃ isotherm but that we give the value as an estimate of an average atmosphere. We have altered phrasing to be clear the meaning of -25℃. We agree that a description of th e ERA-I cloud scheme is important for other modellers to interpret our conclusions. We have included an additional paragraph in section 2.1 as also recommended by reviewer 1. We are not in a position to test the sensitivity of the ERA-I cloud scheme as the model is not directly accessible. However, we have tested the sensitivity of the choice of pressure level by finding relationships at different pressures – these have been included in the discussion. We find that the choice of pressure level in the ICEFLUX parametrisation has a decreasing correlation over land with higher pressures but only slight increases in correlation with lower pressures. There is little effect on ocean correlation.

Furthermore, the parameterization should be developed using instantaneous observations of ice flux (model) and lightning (LIS). The orbit data from LIS is available and the observations of lightning (within 90sec) are available for individual clouds. You should take the ice flux for the same pixels in the reanalysis for building your parameterization. On the other hand, if you want to use only the monthly mean lightning and ice flux, then you should later only calculate monthly mean lightning using the reanalysis. You cannot calculate the lightning every 6 hours when the parameterization was build using monthly mean values.

RESPONSE. This is a fair comment and a valid and alternative approach to the one we have taken would have been to use instantaneous measurements of LIS and match them to ERA-I timesteps. This would require ERA-I to match up convective events with high accuracy. However, we note that a 6 hourly timestep is unlikely to represent individual measurements by LIS. The approach we have taken coarsens the temporal and spatial resolutions in order to look at larger and longer timescale features. Due to the functional form of ICEFLUX as a linear relationship, averaging 6 hourly data to monthly and applying the linear ICEFLUX parametrisation is equivalent to applying the parameterisation to 6 hourly and averaging the flashrate to monthly. Hence in this case our monthly linear parametrisation is applicable to 6 hourly data. However, the existing parametrisations do not have this characteristic and since they were developed on hourly timescales, we have performed our comparison at the 6-hourly time-scale which is also similar to scales used by CTMs.

Your parameterization is also sensitive to the cloud cover in the model. How well does the reanalysis simulate cloud cover? Have you compared the convective cloud cover with satellite data? If the cloud cover is wrong, the whole relationship is wrong. And cloud cover is one of those sensitive parameters that are highly variable from GCM to GCM. Do you want your parameterization to be so sensitive to the model used?

RESPONSE. It is true that the parametrisation is sensitive to cloud cover but this has been included to shift the focus from basing a parametrisation on the gridcell to the clouds that are in the gridcell. We consider this an important move to a physically based approach, so have included it.

There is still large uncertainty in both modelling and observations of cloud so at present it is not possible to conclusively determine the accuracy of cloud features. However, it is sensible to think that the use of cloud cover will be appropriate with an accurate cloud input.

You have also used the Price and Rind (1993) relationship to estimate total lightning for the parameterizations compared to the ICEFLUX method. This includes information about the CTH which may be in error in the ERA reanalysis. Hence, you may be adding problematic parameters to the other parameterizations by introducing the p factor in their calculation. So it is very difficult to know what is causing the differences between the methods.

RESPONSE. For a lightning parametrisation to be used in simulating lightning emissions it must provide an estimate of total lightning, which the “p factor” (the ratio of cloud to ground flashes) provides. Parametrisations that do not estimate total lightning must be adjusted by a factor which in large-scale models is done using cold cloud depth. Therefore, we believe that including this is a fair and useful comparison for a parametrisation’s accuracy in modelling total lightning. However we have added an extra sentence after equation 2 to note that the p factor will introduce errors associated with cold cloud depth.

3) You need to know that while the LIS data is the best global lightning data set to date, there are still many problems with these data. First, the satellite only samples a fraction of the true lightning, and only for 90sec per storm. Second, we know of problems over South America due to the South Atlantic Anomaly (SAA). Third, the diurnal cycle is problematic when looking at individual months of data. Hence, some of the differences in the comparisons with LIS may be also due to sampling problems of LIS (especially in the SAA region).

RESPONSE. Thank you for highlighting this information regarding LIS data. We have included discussion of LIS difficulties regarding the SAA in section 6. With our evaluation approach, issues pertaining to individual month biases in LIS data are greatly reduced by using a 5-year climatology.

I would recommend looking at other regional lightning networks that give large scale lightning data (eg. UK ATD, WWLLN, LINET, STARNET) continuously in space and time, even if mainly from CG lightning. This would also allow you to develop a better parameterization using ERA reanalysis ice flux vs, regional lightning over 6 hours.

RESPONSE. We agree that an alternative approach would be to use ground-based data. This approach has been used in other studies such as Meijer et al., (2001). However, such an approach, either has rather limited spatial or temporal coverage, or large variability in detection efficiency over large areas such as the tropics. We chose our approach in order to be able to develop a parametrisation valid at the large-scale. We would like to see the evaluation of the parametrisation tested against other methods for detecting lightning. LIS has in the past provided this evaluation tool. Our study will hopefully be able to feed into possible future evaluation studies using the upcoming geostationary satellites measuring lightning over previously less understood areas such as Central Africa (Meteosat Third Generation from 2018) and South America (GOES-R from 2016).

Minor comments:

Line 56: "more-or-less". How else would you get charge separation? I think this is accepted, not more-or-less.

RESPONSE. We agree this phrasing is confusing and have removed it.

Line 59: heavier

RESPONSE. Corrected.

Line 78, 92: the resolution of ∼75km is still quite coarse to get any microphysical information out of these clouds. How do you determine the cloud cover in the box? How do determine the fraction of supercooled drops vs. ice?

RESPONSE. The ERA-I cloud parametrisation is described in Tiedtke, (1993). Cloud condensate is generated through transport, detrainment from convection and stratiform formation (e.g. large-scale lifting of moist air or radiative cooling). Cloud condensate is destroyed through evaporation and precipitation. The phase of cloud condensate is diagnosed according to a temperature-dependent function with all liquid phase for temperatures warmer than 0℃, an increasing fraction of ice in the mixed-phase temperature region between 0℃ and -23° C and all ice phase for temperatures colder than -23℃. This has been included in the paper wit h an ERA-I description paragraph in section 2.1 as also suggested by reviewer 1.

Line 124: estimates

RESPONSE. Corrected.

Line 127: that a 5-year

RESPONSE. Corrected.

Line 154, 207: One the one had you claim the CTH is problematic in the reanalysis, while here you use CTH to correct for total lightning. Seems a contradiction.

RESPONSE. This has been discussed above as essential for a comparison of parametrisations but as stated above we have added a further sentence that notes the sensitivity to simulated cold cloud depth.

Line 166, 170: Is not 1.09 a 109% increase?

RESPONSE. The factor of 1.09 is a multiplicative factor so corresponds to an increase of 0.09 times the original flash rate. A change of phrasing has been made for clarity.

Line 262: How sensitive is the parameterization to the cloud cover determination?

RESPONSE. Unfortunately, this feature cannot be tested with ERA-I dataset but we have done the alternative sensitivity tests of the pressure level, described in section 6, to inform future studies.

Line 301: Until you can simulate the cloud microphysics, GCMs will not be able to differentiate between graupel and ice.

RESPONSE. Chemistry transport models are applied to global-scale down to cloud-scale and on some scales cloud microphysics are simulated. Our statement here was to suggest that our findings can feed into parametrisations being developed on these scales now and, as resolution undoubtedly improves in the future, larger scales as well.

Line 316, Fig. 4: I see more that 3 cells that are not stippled. Please explain.

RESPONSE. Our cell numbers were with respect to the "underestimated" region we were discussing. We have updated the figure and caption to highlight the region for clarity. Within this region there are 3 cells not stippled.

Line 336: I suggest using monthly data, if this is how the parameterization was developed.

RESPONSE. As discussed above, for ICEFLUX the order of averaging will not have an effect due the linear form of the iceflux parameteriation. However, for the comparison of existing parametrisations it is important to use the 6 hourly resolution.

Line 376: don't forget LIS has problems over SAA

RESPONSE. LIS issues over the SAA region are now mentioned in the data description section and the discussion section. We make the particular point that evaluating model performance against LIS over South America cannot draw robust conclusions but future satellite instruments may further understanding in this area in years to come.

In conclusion, it is not clear to me from this paper that the ICEFLUX parameterization is better that any of the other previous parameterizations. The authors have not convinced this reviewer that their methodology is better. And it is not clear whether the mismatch with the LIS data is due to the parameterizations or the reanalysis products used. Hence, it is difficult to know the usefulness of this study for others in the field. I think this paper needs major revisions and additions before it can be considered for publication.

RESPONSE. The method here is presented as a means to develop a lightning parametrisation which is closely linked with the non-inductive charging mechanism and which can be successfully used within models. It succeeds in offering a proof-of-concept that such a parametrisation can be applied which compares favourably to existing parametrisations. It clearly remains to be seen how this will transfer to other models but in any case the method has highlighted some important points regarding a good correlation over Central Africa despite the underestimation and the possible usefulness of applying fluxes on a cloud mean opposed to as a gridcell mean basis. The inclusion of the ERA-Interim literature on biases and the LIS SAA difficulties has allowed us now to comment more on the robustness of evaluation in particular regions. The SAA in particular reduces the robustness of evaluation over South America, however, in doing the evaluation it at least has been established how the parametrisations behave on this continent which can inform future evaluations which may use a different product to LIS, unaffected by the SAA.

We are pleased to note that the other two reviewers are satisfied that we have presented a novel, physically-based parametrisation for lightning using state-of-the-art theory which has the potential for use by the modelling community. Cloud ice flux is integral to lightning generation and is undoubtedly affected by the changing climate. We see it as an important step for this feature to be introduced into models in order to increase confidence in projected lightning emissions.