# Using cloud ice flux to parametrise large-scale lightning

**D. L. Finney**[1]**, R. M. Doherty**[1]**, O. Wild**[2]**, H. Huntrieser**[3]**, H. C. Pumphrey**[1]**, and A. M. Blyth**[4]

[1]School of GeoSciences, The University of Edinburgh, Edinburgh, UK
[2]Lancaster Environment Centre, Lancaster University, Lancaster, UK
[3]Institut für Physik der Atmosphäre, Deutsches Zentrum für Luft- und Raumfahrt (DLR),
Oberpfaffenhofen, Germany
[4]National Centre for Atmospheric Science, University of Leeds, Leeds, UK

Correspondence to: D. L. Finney (d.finney@ed.ac.uk)

**Abstract**

Lightning is an important natural source of nitrogen oxide especially in the middle and upper troposphere. Hence, it is essential to represent lightning in chemistry transport and coupled chemistry-climate models. Using ERA-Interim meteorological reanalysis data we compare the lightning flash density distributions produced using several existing lightning parametrisations, as well as a new parametrisation developed on the basis of upward cloud ice flux at 440 hPa. The use of ice flux forms a link to the non-inductive charging mechanism of thunderstorms. Spatial and temporal distributions of lightning flash density are compared to tropical and subtropical observations for 2007–2011 from the Lightning Imaging Sensor (LIS) on the Tropical Rainfall Measuring Mission satellite. The well-used lightning flash parametrisation based on cloud-top height has large biases but the derived annual total flash density has a better spatial correlation with the LIS observations than other existing parametrisations. A comparison of flash density simulated by the different schemes shows that the cloud-top height parametrisation has many more instances of moderate flash densities and fewer low and high extremes compared to the other parametrisations. Other studies in the literature have shown that this feature of the cloud-top height parametrisation is in contrast to lightning observations over certain regions. Our new ice flux parametrisation shows a clear improvement over all the existing parametrisations with lower root mean square errors and better spatial correlations with the observations for distributions of annual total, and seasonal and interannual variations. The greatest improvement with the new parametrisation is a more realistic representation of the zonal distribution with a better balance between tropical and subtropical lightning flash estimates. The new parametrisation is appropriate for testing in chemistry transport and chemistry-climate models that use a lightning parametrisation.

## 1 Introduction

Lightning is always occurring somewhere on Earth with an average of 46 flashes every second (Cecil et al., 2012). Every flash has enormous quantities of energy and can extend over tens of kilometres which allows the dissociation of nitrogen ($N_2$) and oxygen ($O_2$) molecules in the air. The dissociation products combine to form reactive nitric oxide (NO) which quickly oxidises to $NO_2$, and an equilibrium between NO and $NO_2$ is reached, and together they are known as $NO_x$. Air is predominantly detrained in the upper anvil levels of the thunderstorm thereby providing the principal natural source of these ozone precursors to the middle and upper troposphere (Grewe, 2007). In total, lightning is estimated to contribute approximately 10 % of the global $NO_x$ source (Schumann and Huntrieser, 2007). Lightning has a large spatial variability as well as a seasonal cycle and interannual variability. As an important but highly-variable source of $NO_x$ driven by meteorological processes, both chemistry transport models and coupled chemistry-climate models require parametrisations of lightning.

The first stage of a parametrisation is to estimate the large-scale distribution of flashes. Previous investigations have found several empirical relationships between lightning and convective variables including relationships based on cloud-top height (Price and Rind, 1992), updraught mass flux (Grewe et al., 2001; Allen and Pickering, 2002) and convective precipitation (Meijer et al., 2001; Allen and Pickering, 2002). The cloud-top height parametrisation is the most widely, almost universally, used but this is not considered ideal because it lacks a direct, physical link with the charging mechanism and because it has a fifth-power relationship for land which introduces large errors for any model bias in cloud-top height (Allen and Pickering, 2002; Tost et al., 2007).

Satellite observations of lightning have enabled useful testing of the ability of parametrisations to reproduce the large-scale distribution (e.g. Tost et al., 2007). The Lightning Imaging Sensor has good quality measurements of lightning for over a decade which allow model comparison over longer climatological periods. These most recent satellite observations lie between $\pm 38°$ latitude. Bond et al. (2002) estimates that 76–85 % of all global lightning

occurs within this region. Therefore, there is scope for using several years of observations to look at how well the parametrisations match the various statistical features of a lightning climatology. How the parametrisations differ with respect to their input variables, functional form, and their strengths and weaknesses may guide development of new parametrisations.

Atmospheric reanalysis data provide the closest representation of global meteorological conditions maintaining a spatially complete and coherent record. These type of data are used to drive chemistry transport and "nudge" global climate models towards real conditions. By using reanalysis data offline several parametrisations can be directly compared to the lightning observations.

As well as large-scale data enabling a top-down approach to evaluation and development, much work has been done with storm-scale models and field campaigns which offer insight for bottom-up development. Charge separation is necessary for the production of lightning in thunderstorms and occurs via the non-inductive charging mechanism (Reynolds et al., 1957; Latham et al., 2004). This postulates that light ice crystals in clouds that rise on convective updraughts collide with heavier, falling graupel and in doing so the two particle types become oppositely charged. The result is net accumulation of opposite charge in different parts of the thundercloud. This has been shown to be a realistic theory through a combination of laboratory, field measurement and satellite studies (Williams, 1989; Blyth et al., 2001; Petersen et al., 2005; Saunders, 2008; Deierling et al., 2008; Liu et al., 2012).

Global climate models are still at the early stages of representing large-scale distributions of ice in clouds. However, development is on-going with satellite and field measurements helping to form a picture of the current distributions of cloud ice (Waliser et al., 2009; Stith et al., 2014). The objective of this study is to test the usefulness of the current state of cloud ice modelling within a lightning parametrisation. It introduces a parametrisation that is more physically based and tests it against existing parametrisations.

The next two sections describe the data and existing parametrisations to be evaluated. Section 4 explains the development of a simple cloud-ice based parametrisation. Section 5 evaluates the climatological performance of all the parametrisations. This is followed by a discussion and conclusions.

4

## 2  Data description

### 2.1  ECMWF ERA-Interim

The European Centre for Medium-Range Weather Forecasting (ECMWF) provides the ERA-Interim global atmospheric reanalysis data product (Dee et al., 2011). ERA-Interim spans from 1989 to near present. The dynamical core is based on a T255 spectral grid which can be interpolated to a regular 0.75° lat-lon grid. In the vertical, a hybrid sigma-pressure grid is used with 60 levels up to 0.1 hPa. Some variables, such as updraught mass flux, are only archived as forecast data on 6 and 12 h steps initialised at 00:00 and 12:00 UT. While analyses exist for some other variables used here, such as temperature, only the forecast type is used to maintain consistency. There is also a distinction between accumulated (e.g. updraught mass flux) and instantaneous (e.g. temperature) variables. The accumulated variables have been divided by 6 h to obtain an average over the period and are used in combination with the instantaneous variables at the end of the 6 h period - a necessary approximation given the output data available.

A selection of variables have been used as input to lightning parametrisations: surface pressure, temperature, cloud cover, specific cloud ice water content, convective precipitation, updraught mass flux and updraught detrainment rate. Processing of the raw data allowed the formation of 6 hourly data for cloud-top height, cold cloud depth, convective precipitation, updraught mass flux at 440 hPa and upward cloud ice flux at 440 hPa on a 0.75° regular grid. The use of each of these variables is explained in Sects. 3 and 4.

Cloud-top height was taken as the highest level containing a non-zero updraught detrainment rate. This definition follows that used in the TM5 model which also uses ECMWF reanalysis data (P. Le Sager, 2012, personal communication). The cold cloud depth was calculated as the difference between cloud-top height and the interpolated height of the 0 °C isotherm. Updraught mass flux was interpolated to the 440 hPa level (typically about 6 km and −25 °C), as were cloud cover and specific cloud ice water content, which are used along with the updraught mass flux to calculate the upward cloud ice flux at 440 hPa as described in Sect. 4.

5

The cloud parametrisation in the ECMWF model version used for ERA-Interim is based on Tiedtke (1993) with moisture-related prognostic variables for humidity, cloud condensate and fractional cloud cover. Sources and sinks describe the major generation and destruction processes of cloud and precipitation including direct detrainment from parametrised convection. The phase of the cloud condensate is diagnosed according to a temperature-dependent function with all liquid phase for temperatures warmer than $0\,^{\circ}C$, an increasing fraction of ice in the mixed-phase temperature region between $0\,^{\circ}C$ and $-23\,^{\circ}C$, and all ice phase for temperatures colder than $-23\,^{\circ}C$. The parametrisation of convection is based on Tiedtke (1989) and detrains directly into the prognostic humidity, condensate and fractional cloud cover variables following the same temperature dependent phase function as above, and provides a direct link between the convection and stratiform cloud schemes.

State-of-the-art reanalysis data is the best input available over regions as large as the tropics/subtropics and for the range of input parametrisations needed to evaluate model performance. However, several known issues exist with the data that 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.

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 improved tropical cloud cover resulting from an improved hydrological cycle and the introduction of ice supersaturation which delays the formation of ice clouds (Tompkins et al., 2007), and from improved deep convective triggering and 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

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) found that 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 from ERA-40 to ERA-Interim. There are still biases remaining over the tropical oceans, specifically around the western pacific and southeast Asia where the precipitation rate is upto 5 mm day$^{-1}$ greater in some parts. The time series of precipitation rate over total land performs well whilst during the 2007-2011 period there is an overestimation by approximately 0.4 mm day$^{-1}$ over total ocean when compared to GPCP. Many independent studies have looked at precipitation compared to observations and other reanalysis sets (Kim et al., 2013; Peña Arancibia et al., 2013; Pfeifroth et al., 2013; Zhang et al., 2013; Lin et al., 2014). Generally ERA-Interim was found to perform well. Most notably though was further confirmation of biases in southeast Asia (Kim et al., 2013), and the 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).

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.

## 2.2 Lightning Imaging Sensor

The Lightning Imaging Sensor (LIS) is a lightning detection instrument aboard the Tropical Rainfall Measurement Mission (TRMM) satellite (Boccippio et al., 2002; Cecil et al., 2012). Measurements have been made since 1998 but the satellite received an orbit boost in 2001 which resulted in a larger field of view and slightly longer sampling duration. Lightning is detected by pulses of illumination in the 777.4 nm atomic oxygen multiplet above background levels. TRMM is in a low earth orbit, has coverage between $\pm38°$ latitude and views each

surface location for $\sim 90\,\mathrm{s}$ with more time spent viewing the edges of its latitudinal coverage. Over the course of 99 days, LIS samples the full diurnal cycle twice for each location (Cecil et al., 2012). Its spatial resolution is approximately 5 km. Detection efficiency ranges between 69 % at local noon to 88 % at midnight (Cecil et al., 2012). The Optical Transient Detector (OTD) is a similar instrument which is now obsolete but provided a broader latitudinal coverage of $\pm 75°$ until 2000 (Christian et al., 2003). OTD is not heavily used in this study but it contributes to the product used to determine the total global flash rate. As with all low-orbit satellites, the accuracy of OTD and LIS reduces within the South Atlantic Anomaly (SAA) which reduces the robustness of the evaluation within this region. This point is elaborated upon in the discussion.

Several products are produced by the NASA Marshall Space Flight Center lightning team using LIS data which are described fully in Cecil et al. (2012). The LIS/OTD low resolution full climatology (LRFC) total flash count is used here to scale all lightning models to the same global annual total. The main product, used throughout this paper, is the LIS/OTD low resolution monthly time series (LRMTS). The LRMTS provides one flash rate density per month on a 2.5° regular lat-lon grid for every month between May 1995 to present; post-2000 it contains data from the LIS instrument only.

It is useful to determine the number of years necessary to produce a lightning climatology. Using 10 years (2002–2011) of LRMTS data as the "true" climatology, different numbers of years are compared to determine their ability in representing that "true" 10 year climatology. Figure 1 shows some example plots from the approach with the 10 year average annual total spatial distribution with differences to the 5 (2007–2011) and 2 (2010–2011) year sets, along with standard deviations for the three cases. It demonstrates that using 2 year averages would not be appropriate for evaluating this climatological period of lightning in terms of the spatial distribution or the interannual variability, but that 5 years is representative.

For each set of years, significance tests were applied to each grid cell of average annual total spatial distribution as compared to the decade to determine which grid cell estimates diverged from the decadal climatology. An additional comparison was made using the spatial distribution of the standard deviation of annual totals to ensure there were no grid cells

8

where the spread of annual totals was being overestimated or underestimated. It was found that at a 5 year set was needed to satisfy these tests. Hence, the years 2007–2011 have been chosen to evaluate the lightning parametrisations in Sect. 5.

A lightning parametrisation based on upward ice flux is developed using the LRMTS product in Sect. 4. To reduce the bias that may occur by using the same data for development and evaluation, a year within 2002–2006 was chosen to develop the parametrisation which was most different from the 5 year evaluation set. The LRMTS average annual total spatial distribution was calculated for the 5 year climatology and each individual year. The sum of the absolute differences between the 5 year climatology and any given year was used as a metric for the difference. The equation for this metric is:

$$d_{\text{year}} = \frac{1}{A} \sum_{i=\text{cells}} a_i |f_{i,\text{year}} - f_{i,5}| \tag{1}$$

where $d$ is the total area-averaged, absolute difference in flash density between the annual mean of any year in the range 2002–2006 and the climatological mean of the 5 years, 2007–2011. On the right of the equation, $A$ is the total area, $i$ loops over spatial grid cells, $a$ is the area of a grid cell and $f$ is the flash rate density. The difference, $d_{\text{year}}$, was greatest for 2002.

## 3 Existing parametrisations

Four existing parametrisations have been chosen for testing with ERA-Interim data. These chosen parametrisations include the commonly used cloud-top height scheme, along with three others which use different input variables and functional forms.

Lightning flashes can be classified in different ways and in chemistry models they are typically separated into cloud-to-ground and intra-cloud types as these have different emissions. Some of the parametrisations have been developed to calculate total flashes and some to calculate cloud-to-ground flashes. The LIS satellite instrument measures total flashes, it does not discriminate between flash types and therefore, where necessary,

parametrisation outputs for cloud-to-ground flashes are adjusted to represent total flashes by dividing by the proportion of total flashes that are cloud-to-ground, $p$. The ratio is determined by a fourth order polynomial based on cold cloud depth as found by Price and Rind (1993):

$$p = \frac{1}{64.09 - 36.54D + 7.493D^2 - 0.648D^3 + 0.021D^4} \tag{2}$$

where $D$ is the depth of cloud above $0\,^\circ\mathrm{C}$. In addition, a minimum depth of 5.5 km is required for any flashes to occur (Price and Rind, 1993).

This approach is required to make like-for-like comparisons of parametrisations and is important for their use in estimating lightning emissions. However, it will introduce the error associated with determining cold cloud depth into parametrisations which include the use of Eq. (2).

Some parametrisations include scaling equations to account for different model spatial resolutions. However, it is found here that none of these scalings produce the correct magnitude for the total global flash rate as estimated by LIS. This problem has been raised in other studies. In particular, Tost et al. (2007) shows that scaling factors can vary by three orders of magnitude depending on the input from different convective schemes. Here, the same convective scheme is used throughout so variation between parametrisations is partly due to the use of different input variables and partly because the parametrisations were developed using different scales and regions. In this study, the global flash rate has been calculated from the LIS LRFC product to be 44 fl. s$^{-1}$ ("fl." is used throughout to abbreviate "flashes"). Only the average global annual total of the 5 year climatology for each parametrisation has been scaled. Spatial, seasonal and interannual distributions are produced using the parametrisations. On this basis, the additional global annual total scaling factors are 0.05, 1.39, 0.32 and 0.70 for the cloud-top height, updraught mass flux, convective precipitation (polynomial), and convective precipitation (linear) parametrisations of Sect. 3.1–3.4 respectively, and 1.09 for the new cloud ice flux scheme of Sect. 4. These additional scaling factors are smaller, and in the case of the cloud-top height parametrisation much smaller, than previously stated scalings (Tost et al., 2007; Murray et al., 2012). Much of this is ex-

10

pected to be related to a greater spatial resolution than those used in Tost et al. (2007) and Murray et al. (2012), and the 6 hour temporal resolution that this study is based upon. The increase of 9 % needed for the newly developed parametrisation is due to the combination of: the use of different years for forming and evaluating the parametrisation and, the fitting of a parametrisation over anything less than the full globe is likely to mean that it does not predict exactly the same LIS/OTD global magnitude. Lightning scalings need to be more frequently discussed in future studies so that a clear picture of their dependencies can emerge.

In the title to each subsection a label is shown which is used to refer to the parametrisation throughout the paper. For example the following cloud-top height based parametrisation will be referred to as CTH.

## 3.1 Cloud-top height (CTH)

A commonly used proxy for lightning flash density is cloud-top height as proposed by Price and Rind (1992). Price and Rind built on theories developed by Vonnegut (1963) and Williams (1985) using storm measurements and satellite data to form the following parametrisation:

$$F_l = 3.44 \times 10^{-5} H^{4.9} \tag{3}$$
$$F_o = 6.2 \times 10^{-4} H^{1.73} \tag{4}$$

where $F$ is the total flash frequency (fl. min$^{-1}$), $H$ is the cloud-top height (km) and subscripts l and o are for land and ocean respectively. The separation between land and ocean is used to incorporate the difference in updraught velocity over the two surface types. In cases of a cloud depth less than 5 km the flash value was set to zero. The use of 5 km is based on the range of data used to develop the relationship in Price and Rind (1992). Note that Price and Rind (1994) developed an equation to translate the above equations to varying model resolutions. The equation used to calculate the scaling factor is:

$$C = 0.97241 e^{0.048203R} \tag{5}$$

11

where $R$ is the product of longitude and latitude resolution (degrees$^2$) and $C$ is a multiplication factor applied to Eqs. (3) and (4). In this study the scaling factor is applied to the initial flash estimates on the regular $0.75\,^\circ$ grid. The scaling factor used is 0.9992. While this scaling has been included for consistency with the parametrisation, it is clear that at resolutions used in this study and higher resolutions that the scaling has very little impact. As discussed above, an additional scaling factor of 0.05 was applied to match the LIS global total flash rate.

### 3.2 Updraught Mass Flux (MFLUX)

A parametrisation based on updraught mass flux at $\sim 440\,\text{hPa}$ was obtained by Allen and Pickering (2002). The choice of 440 hPa is based upon definitions of deep convective clouds in the International Satellite Cloud Climatology Project (ISCCP) (Rossow et al., 1996). In this parametrisation no distinction is made between land and ocean locations. The equation is:

$$F = \frac{\Delta x \Delta y}{A}(a + bM + cM^2 + dM^3 + eM^4) \tag{6}$$

where $F$ is the flash frequency of cloud-to-ground flashes (fl. min$^{-1}$), $M$ is the updraught mass flux at 440 hPa (kg m$^{-2}$ min$^{-1}$), $\Delta x\,\Delta y$ is the area of a grid cell and $A$ is the area of a $2.0\,^\circ \times 2.5\,^\circ$ box centred at $30\,^\circ$ N. The polynomial coefficients $a$–$e$ have the respective values of $-2.34 \times 10^{-2}$, $3.08 \times 10^{-1}$, $-7.19 \times 10^{-1}$, $5.23 \times 10^{-1}$ and $-3.71 \times 10^{-2}$.

Equation (6) only estimates cloud-to-ground flashes and is therefore divided by $p$ in Eq. (2). Following the condition on Eq. (2), cases where the depth is less than 5.5 km are set to zero. The use of areas in this equation is again an approach to account for varying horizontal resolutions. As with the cloud-top height parametrisation, the scaling grid box area is based on that of a regular $0.75\,^\circ$ grid. Limitations exist on the values of mass flux such that $0 < M < 9.6\,\text{kg m}^{-2}\,\text{min}^{-1}$; all values outside this range are set to their nearest acceptable value in the range.

### 3.3 Convective Precipitation (polynomial) (CPPOLY)

A parametrization based on convective precipitation is also presented by Allen and Pickering (2002). There are separate polynomial expansions for land and ocean,

$$F_i = \frac{\Delta x \Delta y}{A}(a_i + b_i P + c_i P^2 + d_i P^3 + e_i P^4) \tag{7}$$

where $F$ is the flash frequency of cloud-to-ground flashes (fl. $\text{min}^{-1}$), $P_i$ is the daily grid cell convective precipitation ($\text{mm day}^{-1}$) during the time step for grid cell type $i$: where for land $i = \text{l}$ and for ocean $i = \text{o}$. The polynomial coefficients $a_\text{l}$–$e_\text{l}$ have the respective values of $3.75 \times 10^{-2}$, $-4.76 \times 10^{-2}$, $5.41 \times 10^{-3}$, $3.21 \times 10^{-4}$ and $-2.93 \times 10^{-6}$. The polynomial coefficients $a_\text{o}$–$e_\text{o}$ have the respective values of $5.23 \times 10^{-2}$, $-4.80 \times 10^{-2}$, $5.45 \times 10^{-3}$, $3.68 \times 10^{-5}$ and $-2.42 \times 10^{-7}$.

Equation (7) only estimates cloud-to-ground flashes and is therefore divided by $p$ in Eq. (2). Following the condition on Eq. (2), cases where the depth is less than 5.5 km are set to zero. The use of area is the same as for Eq. (6). Limitations exist on the values of convective precipitation such that $7 < P < 90 \, \text{mm day}^{-1}$; all values outside this range are set to their nearest acceptable value in the range.

### 3.4 Convective Precipitation (linear) (CPLIN)

An alternative parametrisation based on convective precipitation which uses a linear relationship is proposed by Meijer et al. (2001):

$$\overline{F} = 14700\text{cp} + 1.7 \tag{8}$$

where $\overline{F}$ is the mean number of flashes and cp is the convective precipitation (m). Under this scheme ocean flashes are 10 times less than calculated by Eq. (8) based on findings that convection over oceans is 10 times less efficient at generating lightning (Levy II et al., 1996; Boersma et al., 2005). Equation (8) only estimates cloud-to-ground flashes and is therefore divided by $p$ in Eq. (2).

## 4  A new ice flux based parametrisation (ICEFLUX)

The non-inductive charging mechanism is widely accepted as the primary means for charge separation and therefore lightning generation (Barthe and Pinty, 2007; Saunders, 2008). However, only indirectly related convective characteristics have so far been introduced into large-scale lightning parametrisations. Improved representation of cloud ice in models now allows the implementation of another aspect of the theory, the upward flux of ice crystals. Deierling et al. (2008) have shown that the upward ice flux displays a strong linear correlation with lightning flashes in 11 observed US storms.

The direct implementation of the fitted equation in Deierling et al. (2008) for non-precipitating ice (i.e. upward ice crystal) mass flux above $-5\,^\circ\mathrm{C}$ ($\mathrm{kg\,s^{-1}}$) was explored but it was found that anomalously high flash densities would be estimated along mid-latitude storm tracks. The bias could be due to underlying meteorology but also may be attributable to the form of the ice flux variable. In this study we use cloud fraction in the grid cell to propose an alternative measure of ice flux in storms which is related to the intensity of the flux ($\mathrm{kg\,m^2\,s^{-1}}$) as opposed to the mass of ice alone.

There have been past comparisons of the ECMWF ice water content product to satellite measurements of cloud ice content (Li et al., 2007; Waliser et al., 2009; Wu et al., 2009; Delanoë et al., 2011). They show that while the ice content may be underestimated, there is at least good spatial agreement between ECMWF and the satellite measurements. The ERA-Interim specific cloud ice water content product is an estimate of the non-precipitating ice (i.e. suspended ice crystals) in the grid cell. ERA-Interim also contains updraught mass flux and fractional cloud cover.

As with the parametrisations of Allen and Pickering (2002) and as defined by the International Satellite Cloud Climatology Project (Rossow et al., 1996), the 440 hPa level is used as a pressure level representative of fluxes in deep convective clouds. An estimate for upward cloud ice flux at 440 hPa, $\phi_{\text{ice}}$ ($\mathrm{kg_{ice}\,m_{cloud}^{-2}\,s^{-1}}$), for each 6 hourly time step has been

calculated using the following equation:

$$\phi_{\text{ice}} = \frac{q \times \Phi_{\text{mass}}}{c} \tag{9}$$

where $q$ is specific cloud ice water content at 440 hPa ($\text{kg}_{\text{ice}}\,\text{kg}_{\text{air}}^{-1}$), $\Phi_{\text{mass}}$ is the updraught mass flux at 440 hPa ($\text{kg}_{\text{air}}\,\text{m}_{\text{cell}}^{-2}\,\text{s}^{-1}$) and $c$ is the fractional cloud cover at 440 hPa ($\text{m}_{\text{cloud}}^2\,\text{m}_{\text{cell}}^{-2}$). Upward ice flux was set to zero for instances where $c < 0.01\,\text{m}_{\text{cloud}}^2\text{m}_{\text{cell}}^{-2}$. The relationship between this newly formed variable and lightning is explored below.

### 4.1 The upward ice flux–lightning relationship

To develop a relationship between lightning and upward ice flux, the ice flux produced using Eq. (9) is compared to the lightning flash density of the LRMTS product. As described in Sect. 2.2, the year 2002 has been chosen as the training year as it is most different in terms of the spatial distribution to the lightning climatology of years used in later sections. To compare to LRMTS, upward ice flux values between $\pm 38°$ latitude were averaged to monthly values and interpolated to the LIS grid. A scatter plot of all monthly cell values are shown in Fig. 2a and b for land and ocean regimes respectively.

Lower levels of lightning over ocean have been attributed to weaker updraught strengths within ocean storms (Xu and Zipser, 2012). Many parametrisations are unable to predict these lower oceanic levels of lightning. Likewise, it is necessary to separate ocean and land regimes in the ICEFLUX parametrisation since ocean flash densities for a given upward ice flux were $\sim \frac{1}{7}$ of the land flash densities, as can be seen in Fig. 2. This difference between land and ocean regimes is not quite as large but is of the same order of magnitude as the differences in existing parametrisations. The equations of the best fit lines in Fig. 2a and b are:

$$f_{\text{l}} = 6.58 \times 10^{-7} \phi_{\text{ice}} \tag{10}$$
$$f_{\text{o}} = 9.08 \times 10^{-8} \phi_{\text{ice}} \tag{11}$$

15

where $f_l$ and $f_o$ are the flash density (fl. m$_{cell}^{-2}$ s$^{-1}$) of land and ocean respectively.

The best fit equations use only one parameter, the slope of the regression. Other fits were tested including a two parameter linear fit, polynomial fits and power fits. The single- and 2 parameter-linear fits produced the best results. The intercepts from the two-parameter linear fit for both land and ocean were small and positive. A positive intercept within the modelling environment results in erroneous flashes in time steps which contain no upward ice flux. Since the intercepts are small, there is little change to the fit if only a single-parameter fit is used. Furthermore, an intercept at the origin remains consistent with the non-inductive charging mechanism as charging would not be expected in cases of zero upward ice flux. These are the justifications for the single-parameter linear functional form.

## 4.2 Application of the ICEFLUX relationship

Clearly there are shortcomings of the upward ice flux relationship when applied over such a large region. The correlation over the ocean is poor, $r = 0.25$, but this was also found when comparing cloud-top height against flash density, shown in Fig. 2d. The land correlation is stronger at $r = 0.63$ but there are persistent deviations from the best fit. ERA-Interim reanalysis data, while being the best spatially complete representation of reality, is not equivalent to observations. Where observations are sparse, as over Africa and the oceans, there could be large errors. It does, however, offer a bridge between observation and modelling studies; providing the opportunity to compare model behaviour to measurements of lightning. While the correlation is only 0.06 greater than the cloud-top height variable, the use of upward ice flux is also a step towards a more physically based lightning parametrisation. This is also a step towards the possible future inclusion of a downward graupel flux, the other component of ice collisions, offering potential improvements in lightning estimation.

The regions and months for which continental lightning would have a large underestimation (points with $f > 1.8 \times 10^{-12}$ fl. m$_{cell}^{-2}$ s$^{-1}$, $\phi < 1.8 \times 10^{-6}$ kg$_{ice}$ m$_{cloud}^{-2}$ s$^{-1}$) or overestimation ($f < 1.0 \times 10^{-12}$ fl. m$_{cell}^{-2}$ s$^{-1}$, $\phi > 2.5 \times 10^{-6}$ kg$_{ice}$ m$_{cloud}^{-2}$ s$^{-1}$) are shown in Fig. 3. These regions are highlighted in Fig. 2 as light blue and red respectively. There are large portions of Central Africa and northwest India where flashes will be underestimated, a problem found

16

in other studies (Tost et al., 2007). To explain flash density differences between continental regions, Williams and Sátori (2004) explore a novel concept by describing meteorology in the Amazon as more oceanic in nature than that in Africa. This may suggest that fits of continental lightning are an average of different convective regimes with Central Africa representing the continental extreme thereby explaining its underestimation in the continental fit.

In addition to the full-region scatter plot in Fig. 2, relationships have been found for each grid cell individually using the twelve monthly data points. This uses the same data points but splits them so that each grid cell can be studied separately. Gradients and significance for the underestimated portion of Central Africa are shown in Fig. 4. Only 3 grid cells out of 32 corresponding to the Central African blue region in Fig. 3 do not have significant correlation between flash density and upward ice flux. As demonstrated in Fig. 4, the reason for underestimation of the overall fit in Central Africa is not because a linear model does not apply but because the gradient is steeper, and the relationship between upward cloud ice and lightning is stronger. Gradients are up to three times greater than the full-region relationship in Eq. (10). This could be accounted for using regional gradient lookup tables but the focus of this study is to explore a process-based parametrisation that is globally applicable. A lookup table would not allow consistent study of time periods with meteorology that is substantially different than present-day conditions.

Figure 3 shows adjacent grid cells with overestimation and underestimation in northeast India which suggest a lightning peak in that area that is misplaced to the east. There is some overestimation near the southern Andes and underestimation in Argentina and southern Brazil. This is a region where LIS has lower accuracy due to the SAA which makes it difficult to draw significant conclusions about the region. Figure 3 also shows which months contain model over- and under-estimation of flash density. The underestimation occurs throughout the year but with least in January and February and the highest levels between August and November. Overestimation of lightning occurs over much fewer regions and decreases gradually through the year.

17

### 4.3 Robustness on the 6 hourly timescale

Due to the temporal resolution of LIS data products it has been most appropriate to develop the ICEFLUX parametrisation using monthly data. However, in chemistry transport and chemistry-climate models the temporal resolution is on the order of an hour. To check that the parametrisation behaves reasonably when applied at these temporal scales, 6 hourly ECMWF data is used. A histogram of 6 hourly flash densities in the year 2011 using the 5 parametrisations is shown in Fig. 5. All the tested parametrisations had approximately 95 % of instances less than 0.075 fl. $km^{-2}$ 6 $h^{-1}$.

Wong et al. (2013) used hourly values from Earth Networks Total Lightning Network observations and shows that the CTH parametrisation produces fewer low and high flash frequencies compared to the observations. In another study for the US by Allen and Pickering (2002), National Lightning Detection Network and Long Range Flash Network observations at four locations for June 1997 were used to compare CTH, MFLUX and CPPOLY estimates of lightning. They found that CTH did not pick out the variability in flash rates as the model did not accurately represent the variability in cloud-top height. MFLUX and CPPOLY generally produced much more realistic distributions than CTH but at one location (Carlsbad, New Mexico) the instances of high flash rates are greatly underestimated. This was attributed to the inability of the model to represent the North American monsoon.

Problems with the CTH frequency distribution are possibly a result of modelled convection or representation of cloud-top height within models, opposed to a failure of the relationship of the cloud-top height to lightning. A smaller scale study measuring frequency distributions of cloud-top height and lightning would be required for confirmation. Given that this is a further study to find discrepancies in the lightning frequency distribution of the cloud-top height parametrisation, it may be important to determine how this affects the chemistry associated with lightning emissions.

In Fig. 5 ICEFLUX is qualitatively more similar to MFLUX and CPPOLY than CTH but gives fewer high flash densities. Given that it lies between existing parametrisations and has a similar distribution, ICEFLUX is considered appropriate to apply at 6 hourly time scales.

18

## 5   Evaluation of the large-scale lightning parametrisations

The years 2007–2011 have been chosen to evaluate the performance of five different lightning parametrisations against LIS observations. As explained in Sect. 2.2, these five years provide a good estimate of the lightning climatology within the tropics and subtropics. There is a small ($+1.0 \times 10^{-5}$ fl. km$^{-2}$ day$^{-1}$), significant ($p = 0.02$) trend in the deseasonalised, monthly, global total LIS flashes over the 5 years. The trend is at least an order of magnitude smaller than the statistics used here so the findings are considered to represent lightning behaviour independent of climate change. However, that is not to say that the small trend found over this short time period describes any long-term trends in lightning activity.

The parametrisations have been applied to the 6 hourly, 0.75$^\circ$ resolution ERA-Interim data to estimate flash density. For comparison to LIS measurements the parametrisation flash density is then averaged to monthly values, re-gridded to the 2.5$^\circ$ LIS grid, scaled to the same global total and the LIS viewing region of $\pm 38^\circ$ latitude selected.

The climatological average annual total flash density for the observations and parametrisations is shown in Fig. 6. These results show that all parametrisations underestimate flash density over Central Africa compared to LIS satellite measurements suggesting that either an important component of lightning generation is missing from all parametrisations or the underlying meteorology data is biasing the results of parametrisations. In addition, the ocean flash density distribution of all parametrisations is focused heavily along the Inter-Tropical Convergence Zone whereas LIS measurements do not show such a focus. The zonal and meridional average distributions are shown in Fig. 7 to demonstrate that there are significant changes in the zonal distribution and ratios of lightning in the tropical chimneys. ICEFLUX has greatly improved on the ratio of tropical to subtropical lightning compared to the parametrisations as can be seen in the zonal mean plot. The meridional mean plot highlights the underestimation in Africa (0–30$^\circ$ E) and the overestimation over the Americas (60–90$^\circ$ W) and Asia (90–120$^\circ$ E) of all parametrisations. ICEFLUX has made definite improvements within the latter two regions. Table 1 shows the spatial correlations and errors of average annual total between LIS measurements and each of the five parametrisations.

19

By far the best spatial correlation with the LIS measurements over this period is with ICE-FLUX, $r = 0.77$. The ICEFLUX parametrisation also shows the lowest root mean square error (RMSE) of 4.65 which is almost half of the RMSE of CTH at 8.61.

CTH shows very large flash densities over northern South America and southeast Asia while having low flash densities in most subtropical locations. ICEFLUX is spread much more evenly zonally and meridionally (Fig. 7). MFLUX has a very different distribution to the other schemes. It shows high flash densities in southern South America and lower values elsewhere. Out of all the parametrisations it best estimates the activity in the southern US. This is to be expected as the parametrisation was developed using US data. As the only parametrisation not to distinguish between ocean and land it is unsurprising to see the overestimation in the ocean. CPPOLY also shows very high ocean flash densities. CPLIN is qualitatively similar to CTH but with a smaller land-ocean contrast.

CTH has a reasonable correlation with the LIS observations but large errors while CPLIN shows a similar correlation but reduced errors. MFLUX and CPPOLY have very poor spatial correlations with LIS. ICEFLUX has a good correlation and low errors.

Figure 8 shows the climatological annual cycle over the northern and southern tropical and subtropical regions. In both hemispheres we can see a delay of the peak month by $\sim 1$ month for all parametrisations. Statistics regarding timing of the peak month in each grid cell are shown in Table 1. The statistics give a more precise measure of the delay by finding the difference in peak month for each grid cell in the LIS viewing region. They show that on average the peak month in each grid cell is shifted by 0.16 of a month, with the parametrisations ranging in their delay from 0.09 to 0.24 of a month. The average absolute difference has been calculated as it will better represent the total bias in the distribution of peak month. It shows that ICEFLUX performs best and MFLUX the worst for this metric. Interestingly with ICEFLUX there is a larger average delay in peak month but lower overall error in peak month compared to the other parametrisations.

As well as the delay of the peak month, Fig. 8 shows there are biases in the magnitude. In the Southern Hemisphere all parametrisations except ICEFLUX overestimate the magnitude of flashes. In the Northern Hemisphere the magnitude is well produced by CTH,

CPPOLY and CPLIN and underestimated by ICEFLUX and MFLUX. An inspection of the spatial distributions for July has shown that CTH, CPPOLY and CPLIN achieve the correct Northern Hemisphere flash density magnitude although large overestimation occurs over India and southeast Asia and underestimation occurs in other areas. ICEFLUX does not contain the same overestimation but does underestimate lightning activity in West Africa leading to the overall underestimation in the Northern Hemisphere peak. Another important issue is that none of the parametrisations establish the difference in total lightning between the Northern and Southern Hemispheres seen in LIS measurements.

Figure 9 depicts the climatological seasonal peak-to-peak difference (difference of the minimum monthly value and the maximum monthly value during a year), and correlations and root mean square errors are given in Table 1. This metric brings together the spatial and temporal variation in lightning. It highlights where inter-seasonal variation is large and therefore areas which can be dominant regions of lighting activity even if they are not so prominent when considering yearly totals alone. The notable features in the observations that differ from the annual total plots are the low seasonal variation at the Central African location of maximum annual total and increased importance of northern India and North America.

CTH and ICEFLUX provide a reasonable distribution around Central Africa but have large biases elsewhere particularly in Asia and the US. ICEFLUX also underestimates the seasonal variation in West Africa. Ocean seasonal variations are underestimated by CTH and overestimated by ICEFLUX. MFLUX and CPPOLY both produce too much inter-seasonal variation over the oceans. MFLUX overestimates the seasonal variation in South America. CPLIN, as with other metrics, is qualitatively similar to CTH. The correlations and errors have the same ranking of ability as for the annual totals, with ICEFLUX consistently performing well.

Figure 10 shows the average change in lightning activity between consecutive years during the 5 year sample, and correlations and errors are given in Table 1. We use this metric to study the interannual peak-to-peak difference of the parametrisations and LIS measurements. Unlike other results so far, the spatial distribution is more heterogeneous with large

21

differences in interannual variation between neighbouring cells. For example, the interannual variation in Central African grid cells differs by over an order of magnitude even though all cells display the same high annual total flash density. The ocean, unsurprisingly, shows lower interannual change compared to the land. The northeast and northwest of India are the two regions that stand out as having the greatest interannual variation which is expected to be related to monsoonal variability. Cecil et al. (2012) discussed lightning activity in northeast India in depth as it contains the greatest monthly flash density measured by LIS. The study of Cecil et al. (2012) and this study support the significance of India with respect to the interannual and seasonal peak-to-peak differences of global lightning distributions.

All schemes have a low correlation of interannual peak-to-peak difference with the LIS measurements and fail to pick out northern India as having the greatest variation. Instead the parametrisations find that southeast Asia has the greatest variation. Out of all the lightning statistics studied, the interannual peak-to-peak difference is least well simulated across all parametrisations. This will be in part due to the underlying meteorology. Despite these difficulties ICEFLUX has made some improvement on the abilities of the other schemes in its ability to simulate interannual peak-to-peak difference of lightning flash density.

## 6 Discussion

This study compares one new and four existing lightning parametrisations using 6 hourly meteorological data. Other studies have compared some of the same existing parametrisations used here. Tost et al. (2007) and Murray et al. (2012) give correlations for CTH, MFLUX and CPPOLY and reach the same conclusions as in this study. CTH has a reasonable correlation whereas MFLUX and CPPOLY have poor correlations with CPPOLY slightly better. Barthe et al. (2010) compared CTH, two updraught-based parametrisations and three ice and ice flux based parametrisations in cloud-resolving model simulation for two storms of different types. Most parametrisations had some success for particular storms and particular features with none standing out above the rest as best overall. This is contrary to our larger scale findings which suggest that an ice flux based parametrisation success-

fully captures many large-scale features compared to the parametrisations based only on convective characteristics. A difference between the two studies is the nature of the upward ice flux variable; an intensive property was used in our case whilst an extensive property was used in the case of Barthe et al. (2010) i.e. mass per area per time was used in our case opposed to only mass per time. The use of areal density provides a better measure of intensity of ice movement in the grid cell whereas mass alone would have high values even if there is a high amount of cloud ice in a grid cell but rising slowly. As discussed in Sect. 4, this appears to be an important choice when including ice flux into the modelling environment.

By looking at several years of lightning satellite measurements this study has been able to quantify the annual total, seasonal and interannual behaviour of lightning across the tropics and subtropics. In line with other studies, Central Africa stands out as the most important feature with the greatest annual total lightning flashes. However, when considering seasonal and interannual spatial distributions the subtropics are just as important. India shows the greatest seasonal and interannual variation. There is substantial evidence linking these variations in lightning activity to monsoon seasons (Kumar and Kamra, 2012; Pawar et al., 2012; Chaudhuri and Middey, 2013; Penki and Kamra, 2013). The ability of models to represent the monsoon as well as the links between the monsoon and lightning are important to consider when studying lightning in India.

In some of the parametrisations, most notably CTH, there is a clear bias towards the tropics which is not evident in the LIS measurements. While a significant portion of global lightning activity occurs in the tropics, demonstrated by the global annual peak located on the equator, the next most active regions are in the subtropics. CTH exhibits this tropical bias due to its foundation in cloud-top height which is limited by tropopause height, since tropopause height reduces away from the equator. ICEFLUX goes a long way to addressing this issue with the incorporation of updraught mass flux. Updraught mass flux on its own is typically not enough to provide a robust parametrisation – at least not with the formulation here. An alternative parametrisation by Grewe et al. (2001) exists which incorporates up-

draught mass flux into Eq. 3. It was tested by Tost et al. (2007) and performed similarly to the other existing parametrisations being evaluated in this study.

To aid in understanding the sensitivities of the ICEFLUX parametrisation, some additional relationships using different independent variables have been found using the same approach as for Fig. 2. These used the same LIS data but compared against: 1) Updraught mass flux only at 440 hPa, 2) Upward cloud ice flux but at 540 hPa, and 3) Upward cloud ice flux at 340 hPa. The correlations over land were found to be 0.51, 0.47 and 0.65, respectively. Correlations over ocean varied by no more than 0.03. This suggests that the upward cloud ice variable significantly improves upon the correlation of updraught mass flux alone over land. The sensitivity to pressure is also demonstrated with higher pressures reducing the correlation but lower pressures maintaining a similar level of correlation. The gradients found for the 340 hPa level are $7.33 \times 10^{-7}$ fl. $m_{cell}^{-2}/kg_{ice}\, m_{cloud}^{-2}$ over land and $1.30 \times 10^{-7}$ fl. $m_{cell}^{-2}/kg_{ice}\, m_{cloud}^{-2}$ over ocean. These are similar to those at the 440 hPa level though larger. The gradients for the 540 hPa level are larger again which suggests the 440 hPa level gradients are sitting close to a minima in the gradient values. The significance of such a minima is not obvious and further investigation at this point does not seem necessary since only slight gains can be made by arbitrarily optimising the pressure level.

A long standing problem when parametrising lightning has been overestimation in the Amazon Basin and underestimation in Central Africa. Complicating the study of this problem and evaluation of parametrisations in South America is the SAA which reduces the accuracy of LIS. So whilst this study has found that the ICEFLUX parametrisation reduces the Amazon bias, though not solving it, a robust conclusion cannot be drawn about lightning in the region. Regarding Central Africa, the good correlation but higher sensitivity to ice flux than elsewhere suggests additional factors are involved in the charging process. It is hoped that through presenting the results, they act as extra information to inform possible future satellite-based studies using the upcoming GOES-R and Meteosat Third Generation lightning detectors. These detectors, by operating continuously, over the Americas and Europe and Africa on a geostationary satellite unaffected by the SAA will provide a new and more robust perspective on lightning activity in these regions.

Error characteristics in both seasonal variation and annual total appear to depend very much on the functional form of the parametrisation. The power law form of CTH leads to large errors where biases in lightning flashes exist as demonstrated by it having the largest errors in the annual total spatial distribution. The polynomial forms of CPPOLY and MFLUX display less coherence between neighbouring grid cells, especially in the seasonal peak-to-peak difference plots, compared to the linear forms of CPLIN and ICEFLUX. The effect of functional form on errors is worth remembering when applying a parametrisation, which may have been developed for a specific region, on the global scale.

All the parametrisations have an average delay across all the grid cells in the seasonal peak of approximately 3–7 days. This may be a consequence of inaccuracies in the reanalysis data, or the smoothing and averaging of LIS measurements which uses a 99 day and $7.5° \times 7.5°$ boxcar moving average (Cecil et al., 2012). It could also be related to the occurrence of lightning in relation to other features of storms. It has been discussed in some papers that lightning peak months precede the rainfall peak months in monsoonal regions (Chaudhuri and Middey, 2013; Penki and Kamra, 2013). This may be relevant to variables such as convective rain and cloud-top height, however, one would expect the use of upward ice flux to begin to correct the delay. On the contrary, the ICEFLUX parametrisation shows the greatest delay bias although the lowest absolute bias. This suggests it could be an issue with the input meteorology or the need for an extension of the model to include a graupel flux, as in Deierling et al. (2008), to fully account for the seasonal cycle. Cloud Resolving Models would provide an appropriate means of testing the importance of graupel and other features of ice-based parametrisations as they include more explicit representations of cloud parameters. They also can be used in combination with field measurements to study the applicability of the ICEFLUX parametrisation on smaller scales which is vital in determining how widely the parametrisation can be used.

One other factor that may be important for explaining the variance in flash frequency density is geographical differences in flash characteristics. Recent research using a variety of data types has demonstrated different characteristics between land and ocean flashes such that ocean flashes are more energetic, powerful or longer (Hutchins et al., 2013; Said

et al., 2013; Peterson and Chuntao, 2013; Beirle et al., 2014). If this is the case then using the charging theory alone may not account for locations where there are fewer, more energetic flashes opposed more, less energetic flashes. The characteristics of lightning may also hold important information regarding the variance in emissions from lightning. Future development of lightning parametrisations should consider if flash characteristics contribute to observed lightning distributions.

## 7 Conclusions

A large-scale lightning parametrisation based on upward ice flux at 440 hPa (ICEFLUX) closely connected with the non-inductive charging mechanism has been developed here. While its development highlighted the challenge of forming a parametrisation for lightning over large scales, it showed no weaknesses that are not already inherent in existing parametrisations and which are, in part, due to the modelled input meteorology especially over southeast Asia and the western Pacific. Its evaluation compared to satellite observations demonstrated several improvements on existing parametrisations regarding the large-scale spatial features of lightning in the tropics and subtropics. Underestimation of Central African lightning remains but it has usefully been shown that linear relationships apply in this region, however the flash rate here is highly sensitive to the upward ice flux compared to the rest of the tropics and subtropics.

The evaluation applied five different lightning parametrisations to the ECMWF ERA-Interim reanalysis, four well-known and one newly developed, and compared their five year climatologies to Lightning Imaging Sensor (LIS) satellite measurements for the same period. The new ICEFLUX parametrisation showed the highest correlation and lowest bias for the spatial distributions of three properties: average annual total lightning density, and average seasonal and interannual peak-to-peak differences. It also represented well the annual cycle of lightning in the southern tropics and subtropics but underestimated in the northern tropics and subtropics principally due to a low bias in West Africa.

26

The Price and Rind (1992) parametrisation based on cloud-top height (CTH) had reasonable correlations with the spatial distributions and the least delay in the annual peak. However, it showed large biases in the zonal average distribution of lightning. The large biases were attributed to functional form which exacerbates any regional biases in the parametrisation.

The convective precipitation based linear parametrisation (CPLIN) of Meijer et al. (2001) was qualitatively similar to CTH for all studied metrics. Two polynomial parametrisations based on convective precipitation (CPPOLY) and updraught mass flux (MFLUX) by Allen and Pickering (2002) were tested but found to perform poorly for the metrics and ERA-Interim meteorological input used here.

The simple ICEFLUX parametrisation more closely linked to the charging theory has been developed which now requires testing online in chemistry transport models to ensure its applicability for simulating NO emissions. The sensitivity of the chemistry to the different lightning features discussed in the study such as the seasonal variation will also be studied in future work. Results obtained indicate potential of the parametrisation but rely on the convective scheme used in the ECMWF model. Therefore, use of the parametrisation by different modelling groups and other evaluation studies is needed to confirm the merits presented here. While individual chemistry-climate models are needed to confirm the wider use of the parametrisation, the conclusions presented here are directly applicable to chemistry transport model simulations performed using ERA-Interim meteorological input.

The upward ice flux parametrisation is presented as a means to explore the importance of cloud ice while models are still in the process of improving their cloud ice schemes. In future field campaigns it would be helpful to estimate the areal coverage of the storm along with the ice flow rate which can then be combined to give ice flux in units of $kg_{ice} \, m_{cloud}^{-2} \, s^{-1}$. This may aid the formation of a global parametrisation based on storm observation data rather than reanalysis data as was necessary in this study. Even with this reanalysis approach, good improvements are made on the existing parametrisations compared here. Furthermore, there may be the opportunity in future, as models of cloud ice develop, that a downward

graupel flux can be combined with the upward ice flux to represent both aspects of the non-inductive charging mechanism.

## References

Ahlgrimm, M. and Köhler, M.: Evaluation of trade cumulus in the ECMWF model with observations from CALIPSO, Mon. Weather Rev., 138, 3071–3083, doi:10.1175/2010MWR3320.1, 2010.

Allen, D. J. and Pickering, K. E.: Evaluation of lightning flash rate parametrizations for use in a global chemical transport model, J. Geophys. Res., 107, 4711, doi:10.1029/2002JD002066, 2002.

Barthe, C. and Pinty, J.-P.: Simulation of electrified storms with comparison of the charge structure and lightning efficiency, J. Geophys. Res., 112, D19204, doi:10.1029/2006JD008241, 2007.

Barthe, C., Deierling, W., and Barth, M. C.: Estimation of total lightning from various storm parameters: A cloud-resolving model study, J. Geophys. Res., 115, D24202, doi:10.1029/2010JD014405, 2010.

Beirle, S., Koshak, W., Blakeslee, R., and Wagner, T.: Global patterns of lightning properties derived by OTD and LIS, Nat. Hazards Earth Syst. Sci. Discuss., 2, 2765–2787, doi:10.5194/nhessd-2-2765-2014, 2014.

Blyth, A. M., Christian Jr, H. J., Driscoll, K., Gadian, A. M., and Latham, J.: Determination of ice precipitation rates and thunderstorm anvil ice contents from satellite observations of lightning, Atmos. Res., 59–60, 217–229, 2001.

Boccippio, D. J., Koshak, W. J., and Blakeslee, R. J.: Performance assessment of the Optical Transient Detector and Lightning Imaging Sensor. Part I: Predicted diurnal variability, J. Atmos. Ocean. Tech., 19, 1318–1332, 2002.

Boersma, K. F., Eskes, H. J., Meijer, E. W., and Kelder, H. M.: Estimates of lightning $NO_x$ production from GOME satellite observations, Atmos. Chem. Phys., 5, 2311–2331, doi:10.5194/acp-5-2311-2005, 2005.

Bond, D. W., Steiger, S., Zhang, R., Tie, X., and Orville, R. E.: The importance of $NO_x$ production by lightning in the tropics, Atmos. Environ., 36, 1509–1519, doi:10.1016/S1352-2310(01)00553-2, 2002.

Cecil, D. J., Buechler, D. E., and Blakeslee, R. J.: Gridded lightning climatology from TRMM-LIS and OTD: Dataset description, Atmos. Res., 135–136, 404–414, doi:10.1016/j.atmosres.2012.06.028, 2012.

Chaudhuri, S. and Middey, A.: Disparity in the characteristic of thunderstorms and associated lightning activities over dissimilar terrains, Meteorol. Atmos. Phys., 119, 151–161, doi:10.1007/s00703-012-0226-4, 2013.

Christian, H. J., Blakeslee, R. J., Boccippio, D. J., Boeck, W. L., Buechler, D. E., Driscoll, K. T., Goodman, S. J., Hall, J. M., Koshak, W. J., Mach, D. M., and Stewart, M. F.: Global frequency and distribution of lightning as observed from space by the Optical Transient Detector, J. Geophys. Res., 108, 4005, doi:10.1029/2002JD002347, 2003.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Holm, E. V., Isaksen, L., Kallberg, P., Kohler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteor. Soc., 137, 553–597, 2011.

Deierling, W., Petersen, W. A., Latham, J., Ellis, S., and Christian, H. J.: The relationship between lightning activity and ice fluxes in thunderstorms, J. Geophys. Res., 113, D15210, doi:10.1029/2007JD009700, 2008.

Delanoë, J., Hogan, R. J., Forbes, R. M., Bodas-Salcedo, A., and Stein, T. H. M.: Evaluation of ice cloud representation in the ECMWF and UK Met Office models using CloudSat and CALIPSO data, Q. J. Roy. Meteor. Soc., 137, 2064–2078, doi:10.1002/qj.882, 2011.

Grewe, V.: Impact of climate variability on tropospheric ozone, Sci. Total Environ., 374, 167–181, doi:10.1016/j.scitotenv.2007.01.032, 2007.

Grewe, V., Brunner, D., Dameris, M., Grenfell, J., Hein, R., Shindell, D., and Staehelin, J.: Origin and variability of upper tropospheric nitrogen oxides and ozone at northern mid-latitudes, Atmos. Environ., 35, 3421–3433, doi:10.1016/S1352-2310(01)00134-0, 2001.

Hutchins, M. L., Holzworth, R. H., Virts, K. S., Wallace, J. M., and Heckman, S.: Radiated VLF energy differences of land and oceanic lightning, Geophys. Res. Lett., 40, 1–5, doi:10.1002/grl.50406, 2013.

Kim, P. S., Jacob, D. J., Liu, X., Warner, J. X., Yang, K., Chance, K., Thouret, V., and Nedelec, P.: Global ozone–CO correlations from OMI and AIRS: constraints on tropospheric ozone sources, Atmos. Chem. Phys., 13, 9321–9335, doi:10.5194/acp-13-9321-2013, 2013.

Kumar, P. R. and Kamra, A.: Variability of lightning activity in South/Southeast Asia during 1997–98 and 2002–03 El Nino/La Nina events, Atmos. Res., 118, 84–102, doi:10.1016/j.atmosres.2012.06.004, 2012.

Latham, J., Blyth, A. M., Christian, H. J., Deierling, W., and Gadian, A. M.: Determination of precipitation rates and yields from lightning measurements, J. Hydrol., 288, 13–19, doi:10.1016/j.jhydrol.2003.11.009, 2004.

Levy II, H., Moxim, W. J., and Kasibhatla, P. S.: A global three-dimensional time-dependent lightning source of tropospheric $NO_x$, J. Geophys. Res., 101, 22911–22922, 1996.

Li, J.-L., Jiang, J. H., Waliser, D. E., and Tompkins, A. M.: Assessing consistency between EOS MLS and ECMWF analyzed and forecast estimates of cloud ice, Geophys. Res. Lett., 34, L08701, doi:10.1029/2006GL029022, 2007.

Lin, R., Zhou, T., and Qian, Y.: Evaluation of global monsoon precipitation changes based on five reanalysis datasets, J. Climate, 27, 1271–1289, doi:10.1175/JCLI-D-13-00215.1, 2014.

Liu, C., Cecil, D. J., Zipser, E. J., Kronfeld, K., and Robertson, R.: Relationships between lightning flash rates and radar reflectivity vertical structures in thunderstorms over the tropics and subtropics, J. Geophys. Res., 117, D06212, doi:10.1029/2011JD017123, 2012.

Meijer, E., van Velthoven, P., Brunner, D., Huntrieser, H., and Kelder, H.: Improvement and evaluation of the parametrisation of nitrogen oxide production by lightning, Phys. Chem. Earth C, 26, 577–583, doi:10.1016/S1464-1917(01)00050-2, 2001.

Murray, L. T., Jacob, D. J., Logan, J. A., Hudman, R. C., and Koshak, W. J.: Optimized regional and interannual variability of lightning in a global chemical transport model constrained by LIS/OTD satellite data, J. Geophys. Res., 117, D20307, doi:10.1029/2012JD017934, 2012.

Pawar, V., Pawar, S. D., Beig, G., and Sahu, S. K.: Effect of lightning activity on surface $NO_x$ and $O_3$ over a tropical station during premonsoon and monsoon seasons, J. Geophys. Res., 117, D05310, doi:10.1029/2011JD016930, 2012.

Peña Arancibia, J. L., van Dijk, A. I. J. M., Renzullo, L. J., and Mulligan, M.: Evaluation of precipitation estimation accuracy in reanalyses, satellite products, and an ensemble method for regions in Australia and south and east Asia, J. Hydrometeorol., 14, 1323–1333, doi:10.1175/JHM-D-12-0132.1, 2013.

Penki, R. K. and Kamra, A. K.: The lightning activity associated with the dry and moist convections in the Himalayan Regions, J. Geophys. Res.-Atmos., 118, 6246–6258, doi:10.1002/jgrd.50499, 2013.

Petersen, W. A., Chistian, H. J., and Rutledge, S. A.: TRMM observations of the global relationship between ice water content and lightning, Geophys. Res. Lett., 32, L14819, doi:10.1029/2005GL023236, 2005.

Peterson, M. and Chuntao, L.: Characteristics of lightning flashes with exceptional illuminated areas, durations, and optical powers and surrounding storm properties in the tropics and inner subtropics, J. Geophys. Res.-Atmos., 118, 11,727–11,740, doi:10.1002/jgrd.50715, 2013.

Pfeifroth, U., Mueller, R., and Ahrens, B.: Evaluation of satellite-based and reanalysis precipitation data in the tropical Pacific, J. Appl. Meteor. Climatol., 52, 634–644, doi:10.1175/JAMC-D-12-049.1, 2013.

Price, C. and Rind, D.: A simple lightning parametrization for calculating global lightning distributions, J. Geophys. Res., 97, 9919–9933, doi:10.1029/92JD00719,1992.

Price, C. and Rind, D.: What determines the cloud-to-ground lightning fraction in thunderstorms?, Geophys. Res. Lett., 20, 463–466, doi:10.1029/93GL00226,1993.

Price, C. and Rind, D.: Modeling global lightning distributions in a general circulation model, Mon. Weather Rev., 122, 1930–1939, 1994.

Reynolds, S. E., Brook, M., and Gourley, M. F.: Thunderstorm charge separation, J. Meteorol., 14, 426–436, 1957.

Rossow, W. B., Walker, A. W., Beuschel, D. E., and Roiter, M. D.: International Satellite Cloud Climatology Project (ISCCP) documentation of new cloud data sets, Tech. Rep. January, World Meteorological Organisation, WMO/TD 737, Geneva, 1996.

Said, R. K., Cohen, M. B., and Inan, U. S.: Highly intense lightning over the oceans: Estimated peak currents from global GLD360 observations, J. Geophys. Res.-Atmos., 118, 6905–6915, doi:10.1002/jgrd.50508, 2013.

Saunders, C.: Charge Separation Mechanisms in Clouds, Space Sci. Rev., 137, 335–353, doi:10.1007/s11214-008-9345-0, 2008.

Schreier, M. M., Kahn, B. H., Sušelj, K., Karlsson, J., Ou, S. C., Yue, Q., and Nasiri, S. L.: Atmospheric parameters in a subtropical cloud regime transition derived by AIRS and MODIS: observed statistical variability compared to ERA-Interim, Atmos. Chem. Phys., 14, 3573–3587, doi:10.5194/acp-14-3573-2014, 2014.

Schumann, U. and Huntrieser, H.: The global lightning-induced nitrogen oxides source, Atmos. Chem. Phys., 7, 3823–3907, doi:10.5194/acp-7-3823-2007, 2007.

Stith, J. L., Avallone, L. M., Bansemer, A., Basarab, B., Dorsi, S. W., Fuchs, B., Lawson, R. P., Rogers, D. C., Rutledge, S., and Toohey, D. W.: Ice particles in the upper anvil regions of midlatitude continental thunderstorms: the case for frozen-drop aggregates, Atmos. Chem. Phys., 14, 1973–1985, doi:10.5194/acp-14-1973-2014, 2014.

Tiedtke, M.: A comprehensive mass flux scheme for cumulus parameterization in large-scale models, Mon. Weather Rev., 117, 1779–1800, 1989.

Tiedtke, M.: Representation of clouds in large-scale models, Mon. Weather Rev., 121, 3040–3061, 1993.

Tompkins, A. M., Gierens, K., and Rädel, G.: Ice supersaturation in the ECMWF integrated forecast, Q. J. Roy. Meteor. Soc., 133, 53–63, doi:10.1002/qj.14, 2007.

Tost, H., Jöckel, P., and Lelieveld, J.: Lightning and convection parameterisations – uncertainties in global modelling, Atmos. Chem. Phys., 7, 4553–4568, doi:10.5194/acp-7-4553-2007, 2007.

Vonnegut, B.: Some facts and speculations concerning the origin and role of thunderstorm electricity, Meteorol. Monogr., 5, 224–241, 1963.

Waliser, D. E., Li, J.-L. F., Woods, C. P., Austin, R. T., Bacmeister, J., Chern, J., Del Genio, A., Jiang, J. H., Kuang, Z., Meng, H., Minnis, P., Platnick, S., Rossow, W. B., Stephens, G. L., Sun-Mack, S., Tao, W.-K., Tompkins, A. M., Vane, D. G., Walker, C., and Wu, D.: Cloud ice: a climate model challenge with signs and expectations of progress, J. Geophys. Res., 114, D00A21, doi:10.1029/2008JD010015, 2009.

Williams, E., and Sátori, G.: Lightning, thermodynamic and hydrological comparison of the two tropical continental chimneys, J. Atmos. Sol.-Terr. Phys., 66, 1213–1231, doi:10.1016/j.jastp.2004.05.015, 2004.

Williams, E. R.: Large-scale charge separation in thunderclouds, J. Geophys. Res., 90, 6013–6025, doi:10.1029/JD090iD04p06013, 1985.

Williams, E. R.: The tripole structure of thunderstorms, J. Geophys. Res., 94, 13151–13167, doi:10.1029/JD094iD11p13151, 1989.

Wong, J., Barth, M. C., and Noone, D.: Evaluating a lightning parameterization based on cloud-top height for mesoscale numerical model simulations, Geosci. Model Dev., 6, 429–443, doi:10.5194/gmd-6-429-2013, 2013.

Wu, D. L., Austin, R. T., Deng, M., Durden, S. L., Heymsfield, A. J., Jiang, J. H., Lambert, A., Li, J., Livesey, N. J., McFarquhar, G. M., Pittman, J. V., Stephens, G. L., Tanelli, S., Vane, D. G., and Waliser, D. E.: Comparisons of global cloud ice from MLS, CloudSat, and correlative data sets, J. Geophys. Res., 114, D00A24, doi:10.1029/2008JD009946, 2009.

Xu, W. and Zipser, E. J.: Properties of deep convection in tropical continental, monsoon, and oceanic rainfall regimes, Geophys. Res. Lett., 39, L07802, doi:10.1029/2012GL051242, 2012.

Zhang, Q., Körnich, H., and Holmgren, K.: How well do reanalyses represent the southern African precipitation?, Clim. Dynam., 40, 951–962, doi:10.1007/s00382-012-1423-z, 2013.

**Table 1.** Statistics of annual total spatial distribution, peak timing, and interseasonal and interannual spatial distributions. For the spatial distributions, the correlation ($r$) and root mean square error (RMSE) are given.

| Parameterisation | Annual total | | Annual peak month | | Seasonal variation | | Interannual variation | |
|---|---|---|---|---|---|---|---|---|
| | $r$ | RMSE (fl. km$^{-2}$ yr$^{-1}$) | Mean bias (month) | Mean absolute bias (month) | $r$ | RMSE (fl. km$^{-2}$ day$^{-1}$) | $r$ | RMSE (fl. km$^{-2}$ day$^{-1}$) |
| CTH | 0.62 | 9.10 | 0.09 | 1.63 | 0.70 | 0.03 | 0.38 | 0.56 |
| ICEFLUX | 0.77 | 4.53 | 0.24 | 1.62 | 0.78 | 0.02 | 0.47 | 0.41 |
| MFLUX | 0.36 | 8.46 | 0.18 | 1.74 | 0.32 | 0.05 | 0.20 | 1.30 |
| CPPOLY | 0.37 | 7.74 | 0.14 | 1.71 | 0.38 | 0.04 | 0.09 | 0.95 |
| CPLIN | 0.62 | 7.13 | 0.13 | 1.65 | 0.69 | 0.03 | 0.34 | 0.52 |
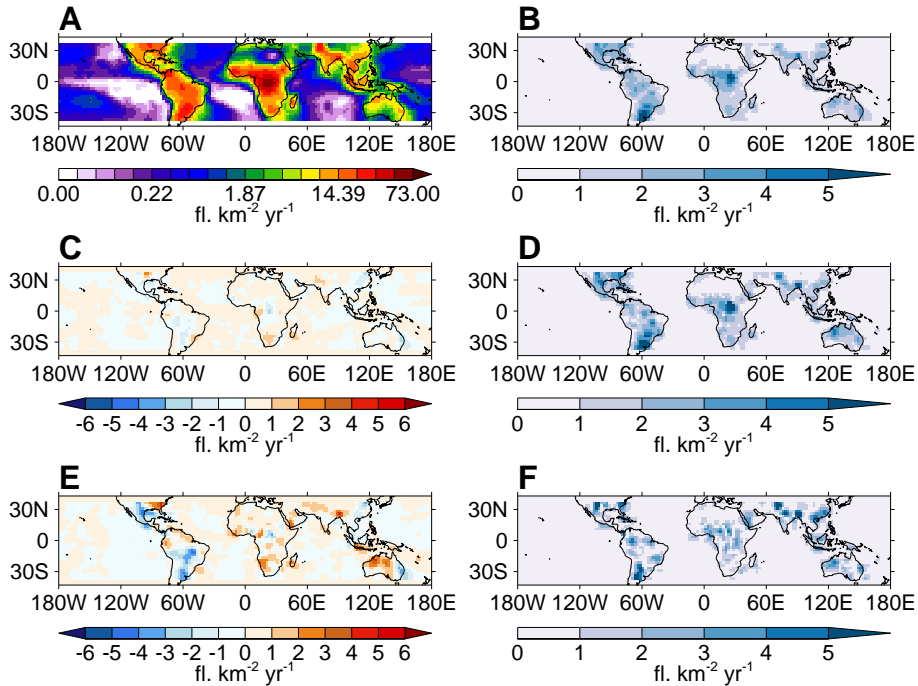
**Figure 1.** Average annual total LIS flash density spatial distributions of **(A)**, the 10 year climatology (2002–2011) and differences between the **(C)**, 10 year and 5 year (2007–2011) and **(E)**, 10 year and 2 year (2010–2011) climatologies. **B**, **D** and **F** show the standard deviations of annual LIS totals for each of the 10 year, 5 year, and 2 year climatologies respectively.
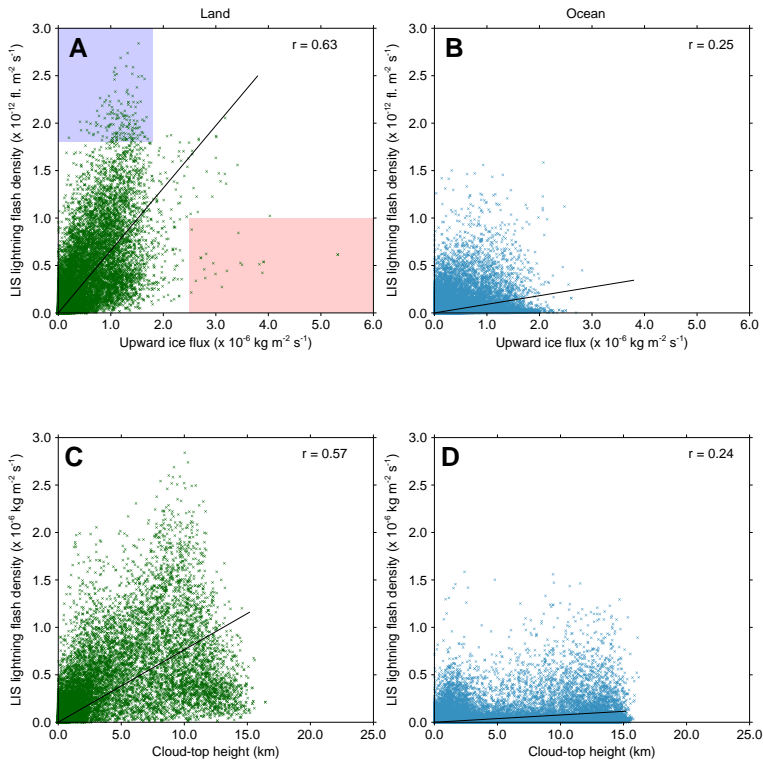
**Figure 2.** Scatter plots of upward ice flux at 440 hPa formed from ERA-Interim reanalysis against LIS flash density. Shown in **(A)** are land grid cells and **(B)** are ocean grid cells. Also shown is the cloud-top height formed from ERA-Interim reanalysis against LIS flash density over **(C)**, land and **(D)**, ocean. Each point represents the monthly average of each variable for a grid cell in the range $\pm 38°$ latitude. The scatter points highlighted in **(A)** are used in Fig. 3 for studying under- (light blue) and over- (light red) estimation of this regression. All scatter points, even within the highlighted regions, were used to create the linear regression.
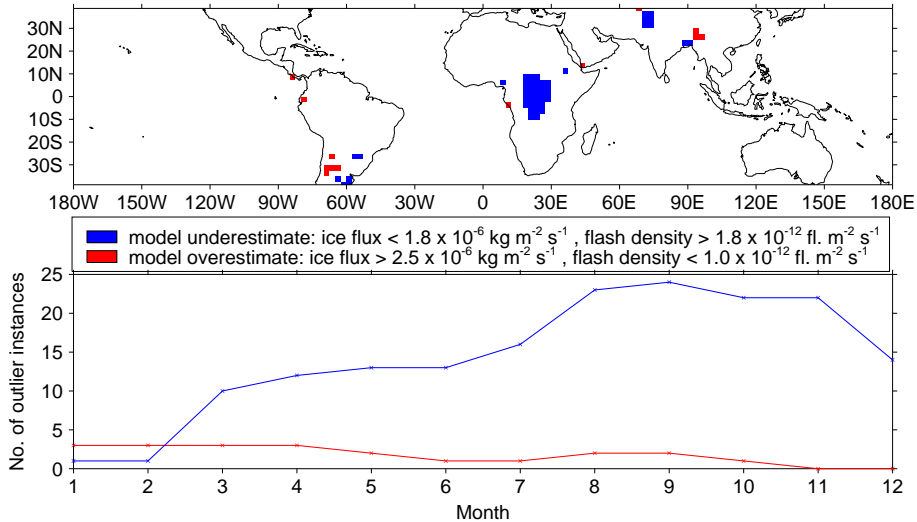
**Figure 3.** Continental regions that would be under- or over-estimated with the proposed ice flux parametrisation. Underestimation in blue and overestimation in red. Any coloured cell in the figure contains at least one month of large under- or over-estimation. Shown bottom is the number of cells of under- or over-estimation in each month. Under- and over- estimation are defined as scatter points in Fig. 2 in the axis ranges of y-axis $> 1.8 \times 10^{-12}$ fl. $m_{cell}^{-2}\, s^{-1}$, x-axis $< 1.0 \times 10^{-6}$ kg$_{ice}$ $m_{cloud}^{-2}\, s^{-1}$ and y-axis $< 1.8 \times 10^{-12}$ fl. $m_{cell}^{-2}\, s^{-1}$, x-axis $> 2.5 \times 10^{-6}$ kg$_{ice}$ $m_{cloud}^{-2}\, s^{-1}$ respectively. The scatter points used to produce this plot are highlighted as the light-coloured regions in Fig. 2.
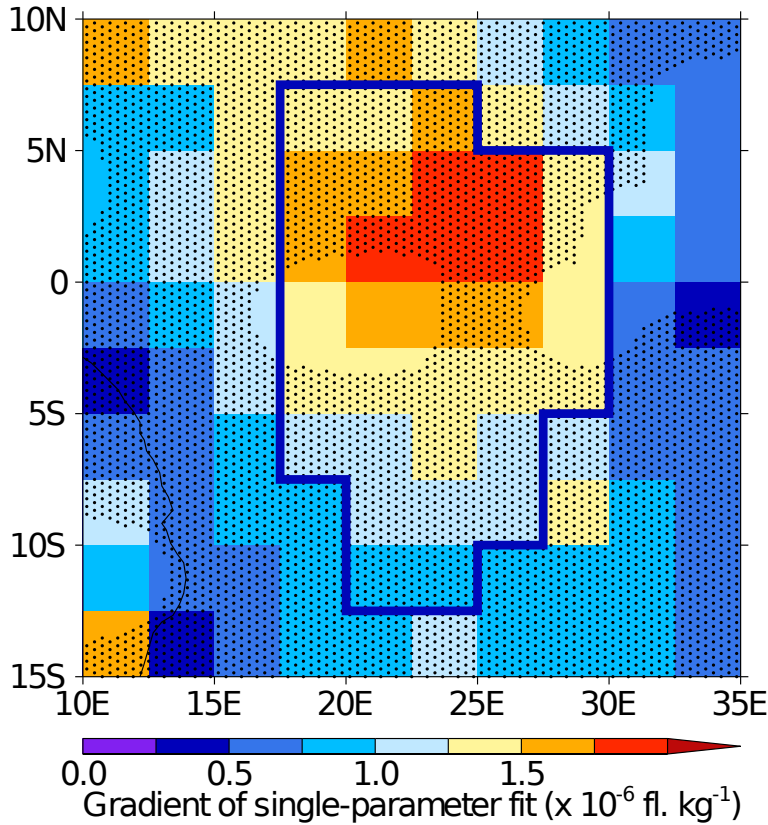
**Figure 4.** The gradients from single-parameter fitting of individual grid cells in Central Africa. The stippling shows grid cells with correlations significant at the 95 % level. Grid cell fits are made using the 12 monthly points for the cell in the year 2002. The solid blue line outlines the model underestimated cells from the blue region in Fig. 4.
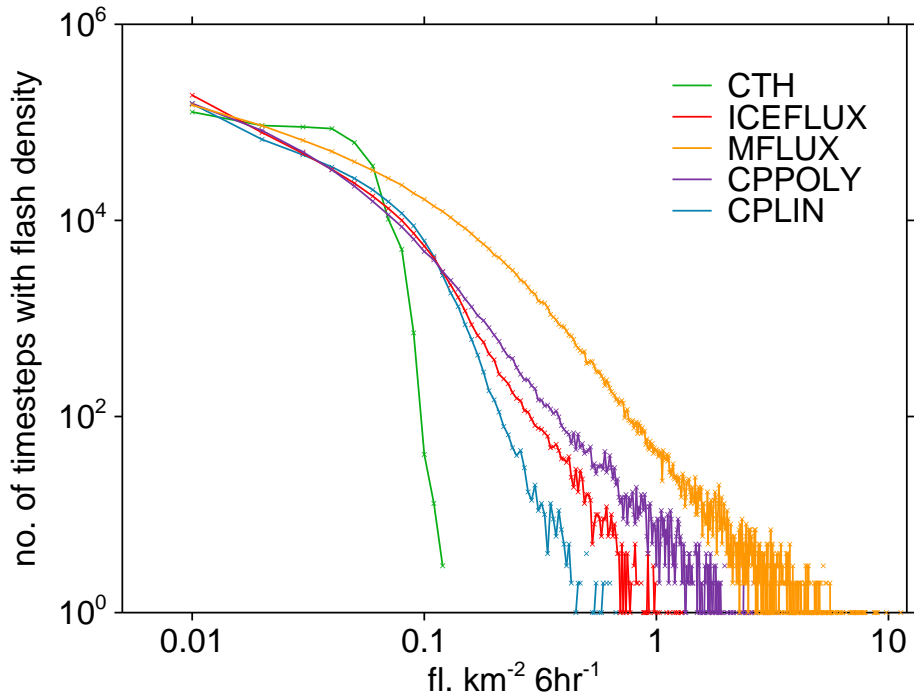
**Figure 5.** Histograms of 6 hourly flash density in the year 2011 for each parametrisation. Binsize is 0.02 fl. km$^{-2}$ 6 h$^{-1}$. The total number of time steps represented by each curve is the same. Grid cells from the full global region are used ($\pm 90°$ latitude).
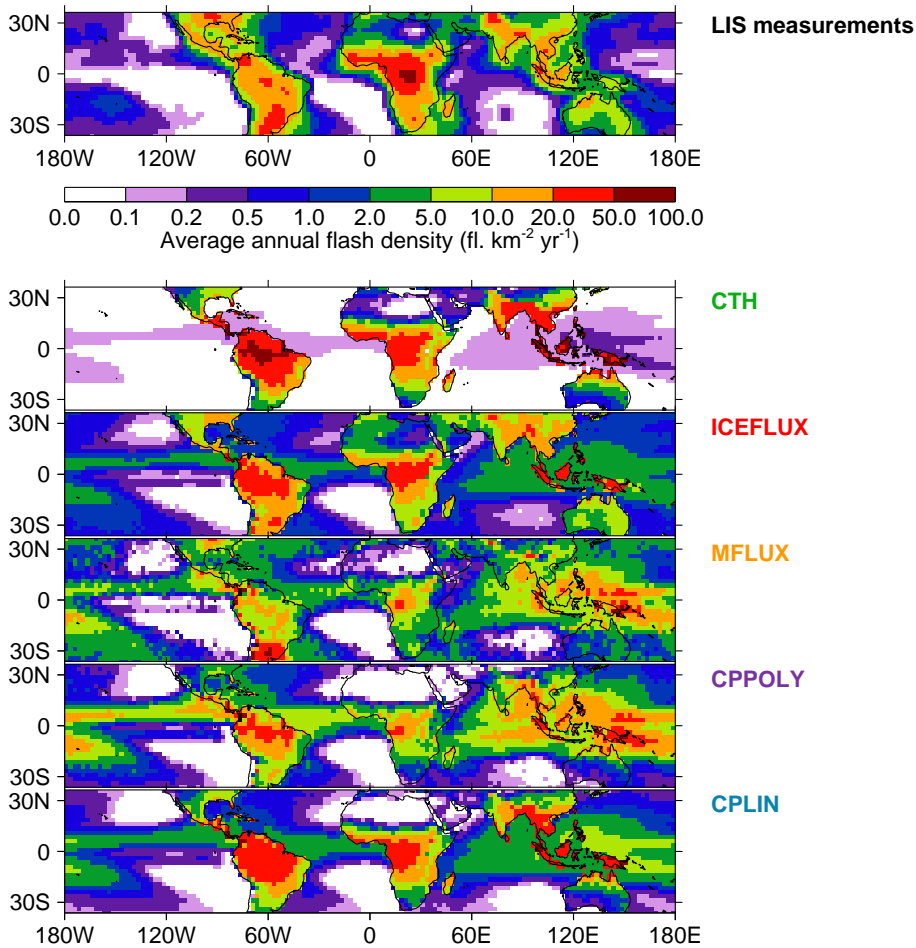
**Figure 6.** Five year climatological annual total flash density (2007–2011). Results are shown for the LIS measurements and the five parametrisations.
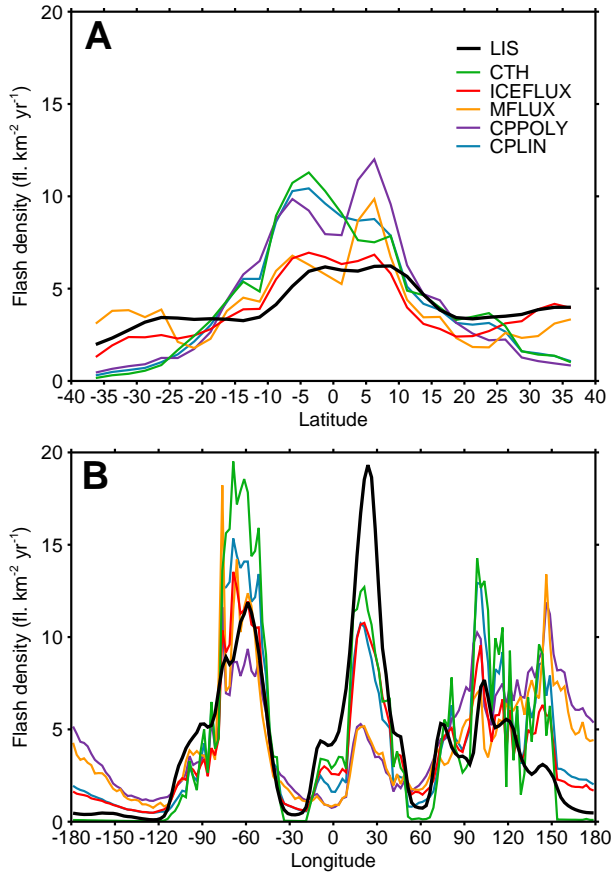
40

**Figure 7.** Five year climatological **(A)**, zonal and **(B)**, meridional average flash density distribution (2007–2011). The meridional average is only taken within the LIS viewing region of $\pm 38°$ latitude. Results are shown for the LIS measurements and the five parametrisations.
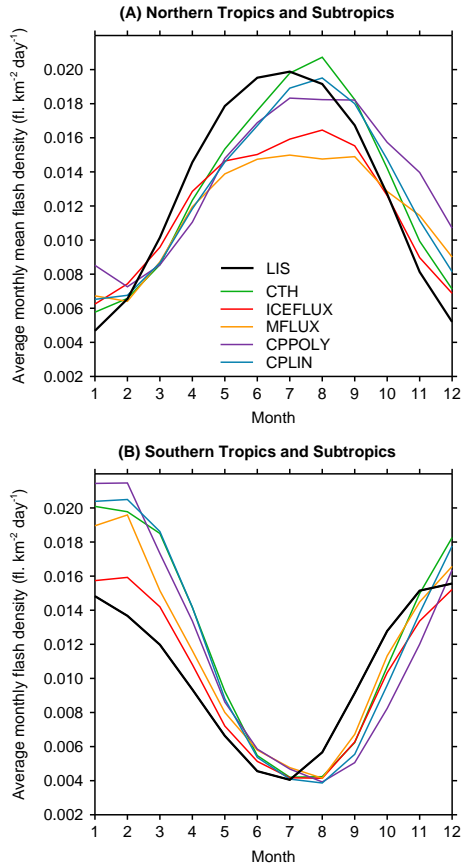
**Figure 8.** Five year climatological annual cycle of flash density for 2007–2011 for **(A)**, the northern region (0 °–38 ° N) and **(B)**, the southern region (38 ° S–0 °). Results are shown for the LIS measurements and the five parameterisations.
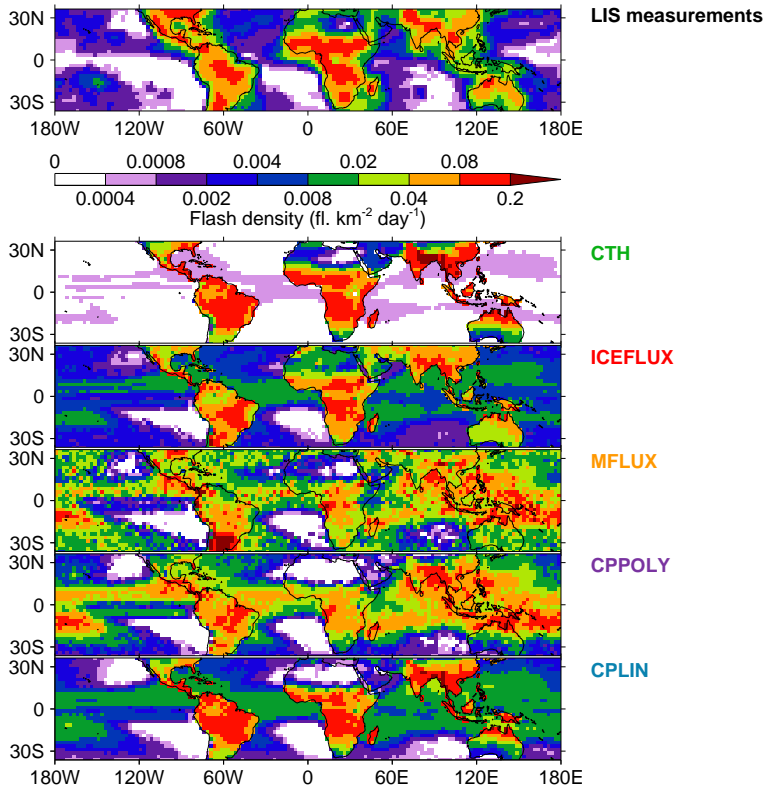
42

**Figure 9.** Five year climatological seasonal peak-to-peak difference spatial distribution of flash density for 2007–2011. The seasonal peak-to-peak difference is the difference between the minimum monthly value and the maximum monthly value. Results are shown for the LIS measurements and the five parametrisations.
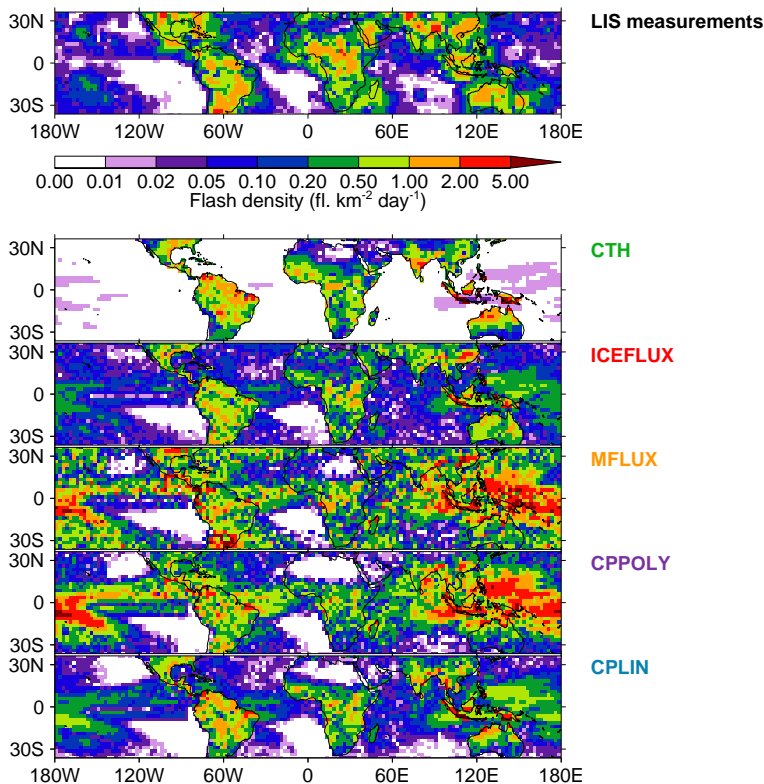
43

**Figure 10.** Five year climatological interannual peak-to-peak difference spatial distribution of flash density for 2007–2011. The interannual peak-to-peak difference has been calculated as the average difference between consecutive years over the five year period. Results are shown for the LIS measurements and the five parametrisations.

44