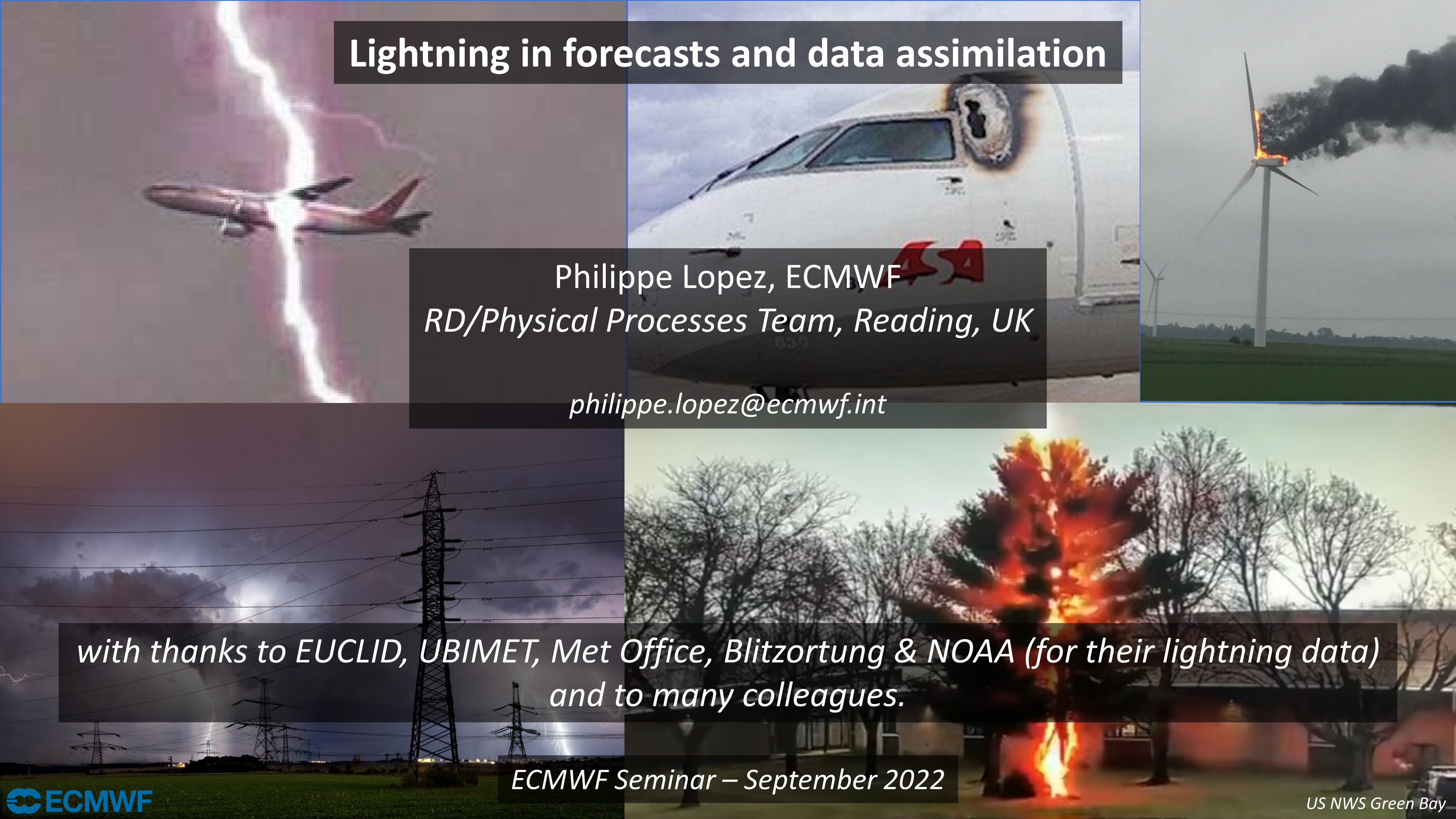
Lightning in forecasts and data assimilation



US NWS Green Bay

Philippe Lopez, ECMWF

RD/Physical Processes Team, Reading, UK

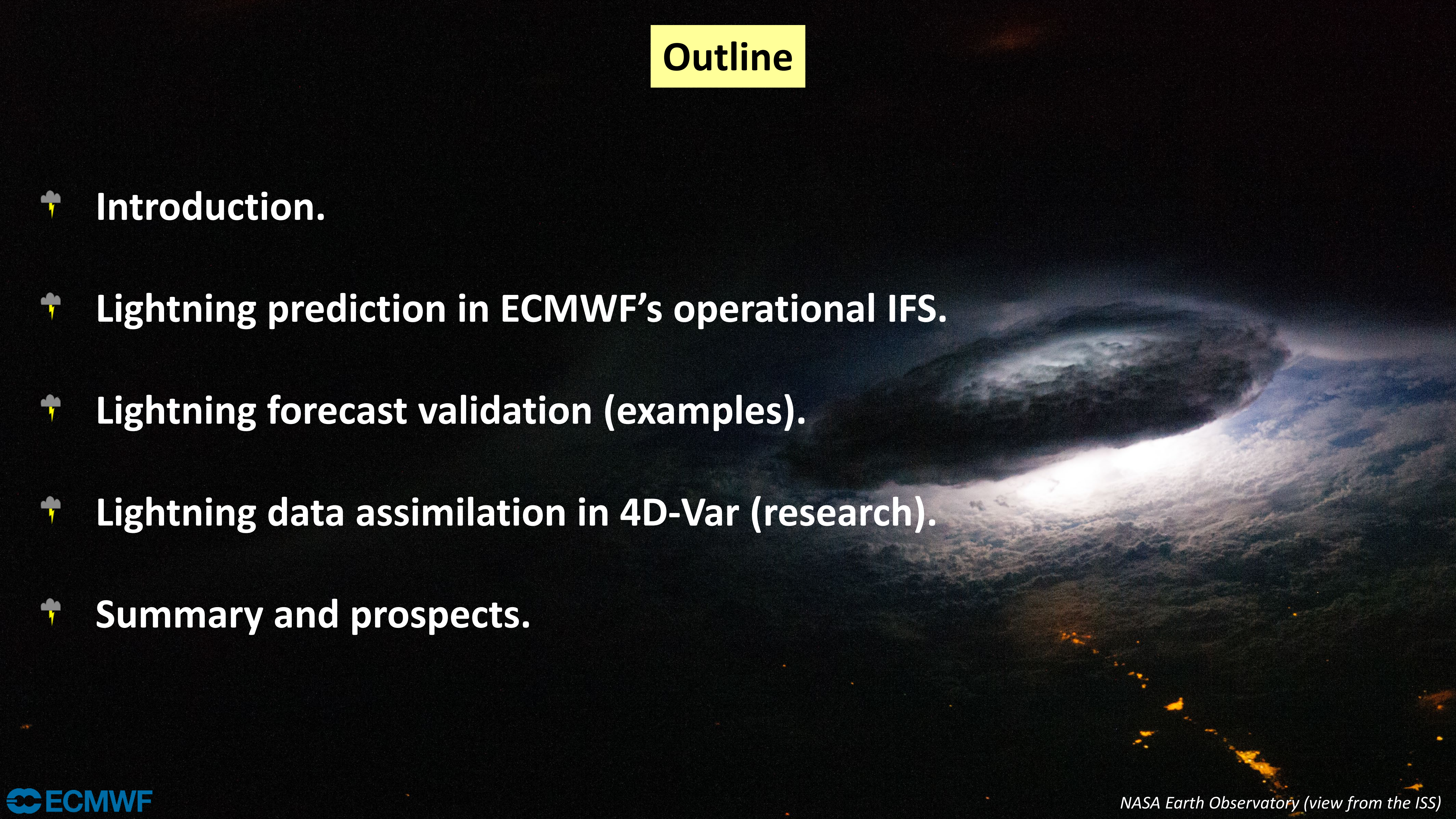
philippe.lopez@ecmwf.int

with thanks to EUCLID, UBIMET, Met Office, Blitzortung & NOAA (for their lightning data) and to many colleagues.

ECMWF Seminar – September 2022

Outline

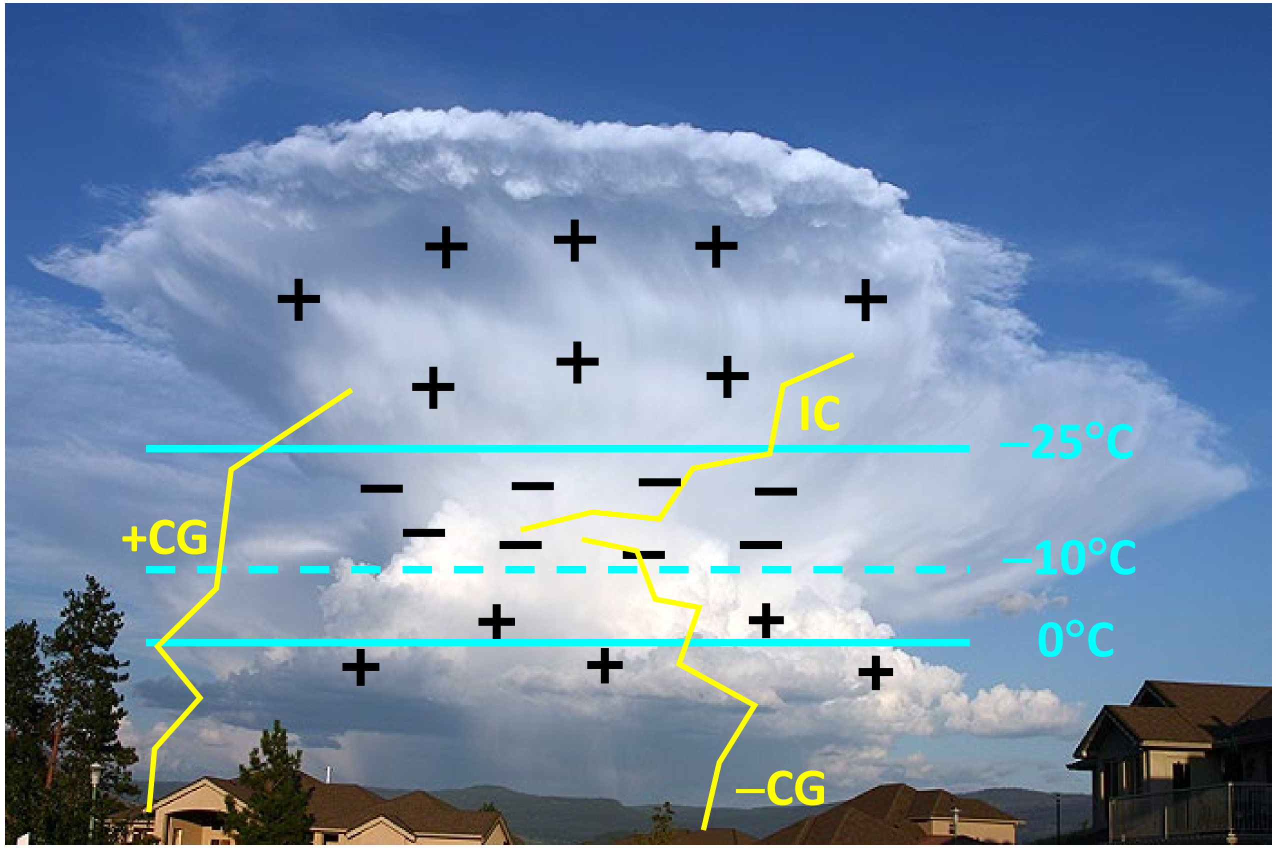
- Introduction.
- Lightning prediction in ECMWF’s operational IFS.
- Lightning forecast validation (examples).
- Lightning data assimilation in 4D-Var (research).
- Summary and prospects.



NASA Earth Observatory (view from the ISS)

Cloud electrification and lightning generation

* **Charge structure in a typical thunderstorm (idealized):**



Electric charge separation occurs through collisions between different types of hydrometeors (hail, graupel, snow, cloud ice and liquid water), with high relative velocities (favoured by strong mixing associated with intense convection).

-> Build-up of the electric field E.

-> Discharge occurs when E > 100-200 kV m^{-1}.

-> Lightning flashes (and thunder).

Occurrence

\+

IC = Intra/Inter-Cloud flash.

–CG = Cloud-to-Ground flash with negative charge transferred to ground.

+CG = Cloud-to-Ground flash with positive charge transferred to ground.

\-

Peak current (typ.)

few kA

30 kA

200 kA

Lightning observations

Lightning (or its direct effects) can be observed using:

*** Ground-based sensors which measure either electromagnetic emissions (sferics) at VLF, LF or VHF (remotely) or variations in the electric field (locally).**

- **VLF and LF sensors → mostly CG + strongest IC events (long range).**
- **VHF mapping arrays → both CG and IC events in 3D (short range).**

Examples of networks of VLF/LF sensors:

- **Global: GLD360 (Vaisala), ENTLN (Earth Networks), WWLLN (Univ. Washington).**
- **Europe: EUCLID (Europe), ATDnet (Met Office, UK), Météorage, UBIMET LDS.**

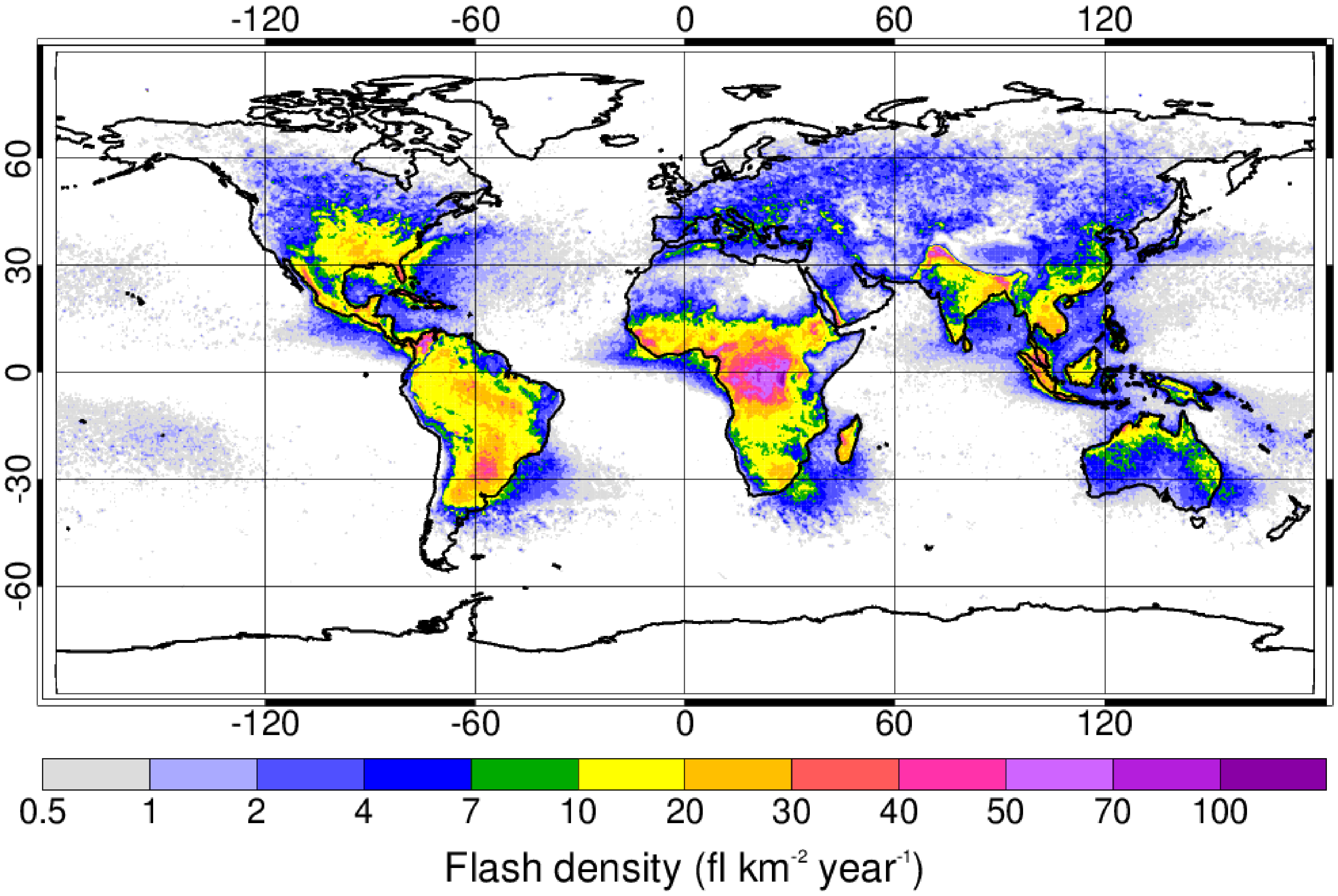
*** Space-borne imagers which detect CG & IC lightning optical signals ($\lambda \approx 777$ nm):**

- Low Earth Orbit: OTD (Optical Transient Detector, 1995-2000);
LIS (Lightning Imaging Sensor on board TRMM, 1998-2013; ISS, 2017→).

- Geostationary satellites: GLM on board GOES-16, GOES-17 & GOES-18 (NOAA, 2017→);
FY-4A/LMI (CMA; 2016→);
MTG/LI (EUMETSAT; 2023?).

Lightning climatology (satellite-based)

Annual mean lightning flash densities from LIS/OTD (1995-2010; Cecil et al. 2014):



Global mean = 2.86 flashes km^{-2} year^{-1} ≈ 46 flashes s^{-1}.

ECMWF

ECMWF's parameterization predicts total (CG+IC) lightning flash densities from a set of predictors diagnosed from the convection scheme of the IFS:

$$f_T = 37.5\, Q_R \sqrt{CAPE}\left[min(z_{base}, 1.8)\right]^2$$

where

$$Q_R = \int_{z_{0^\circ C}}^{z_{-25^\circ C}} q_{graup}\,(q_{cond} + q_{snow})\,\bar{\rho}\, dz$$

Proxy for charging rate (collisions btw. hydrometeors)

with

$$q_{graup} = \frac{\beta\, P_f}{\bar{\rho}\, V_{graup}}$$

graupel content [kg kg^{-1}]

← graupel fall velocity set to 3.0 m s^{-1}

and

$$q_{snow} = \frac{(1-\beta)\, P_f}{\bar{\rho}\, V_{snow}}$$

snow content [kg kg^{-1}]

← snow fall velocity set to 0.5 m s^{-1}

$CAPE$ = convective available potential energy [J kg^{-1}] → Proxy for updraft strength
P_f = convective frozen precipitation flux [kg m^{-2} s^{-1}]
z_{base} = convective cloud base height [km] → Proxy for updraft size
q_{cond} = convective cloud condensate content [kg kg^{-1}]
β = 0.7 over land and 0.45 over ocean (graupel/snow partitioning).

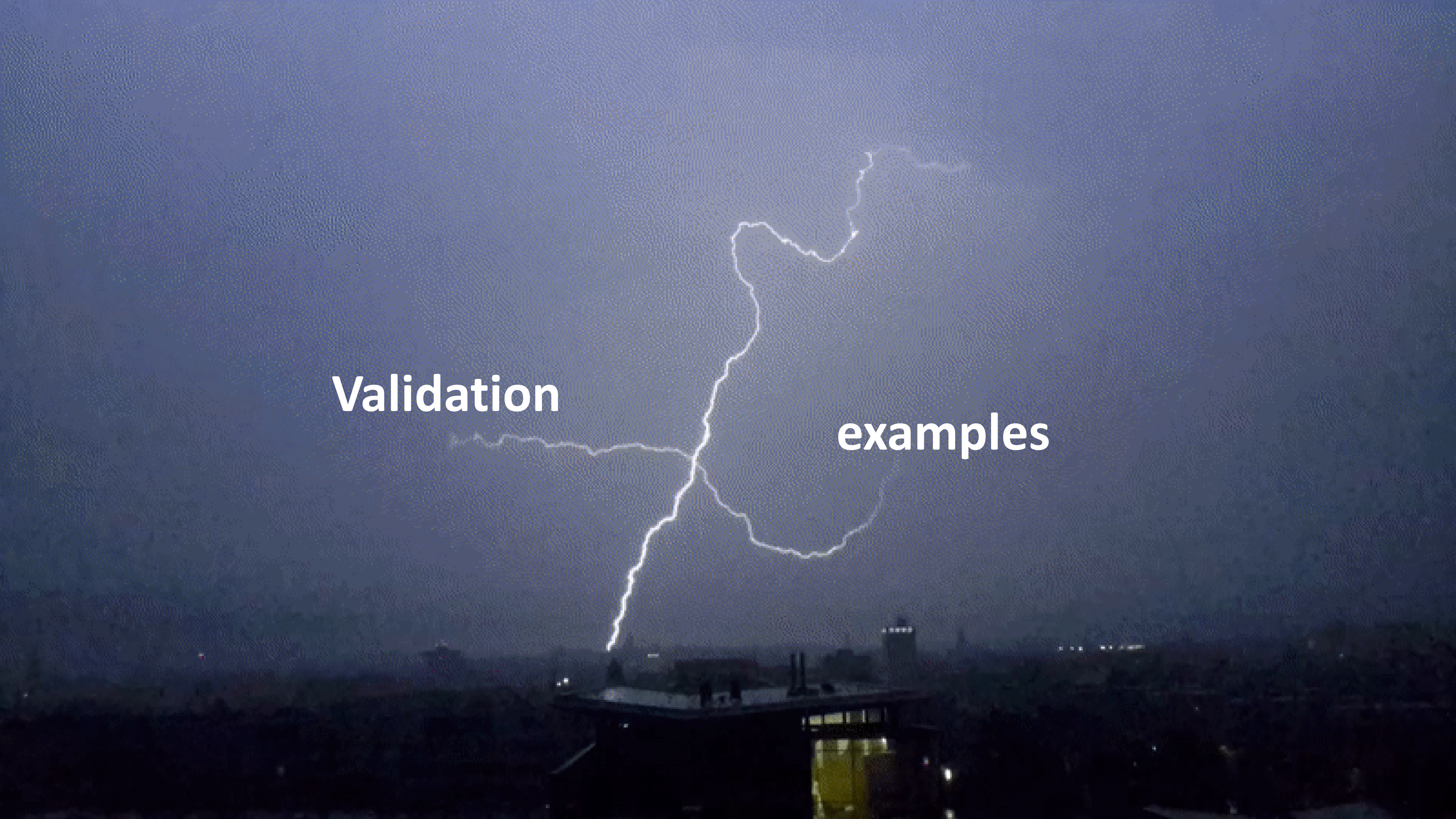
Lopez 2016, MWR

The lightning parameterization of the IFS

- **The parameterization became operational in both deterministic (9-km resolution) and ensemble (18-km resolution) forecasts on 7 June 2018.**

- **It outputs total lightning flash densities that are both “instantaneous” (over a model time step) and averaged over 1, 3 and 6 hours (all expressed in flashes/km^2/day).**

- **In addition to severe weather prediction, it is also being used to forecast:**
  - **lightning-triggered wildfires (Coughlan *et al.*, 2021),**
  - **atmospheric NOx emissions from lightning (CAMS chemistry model).**

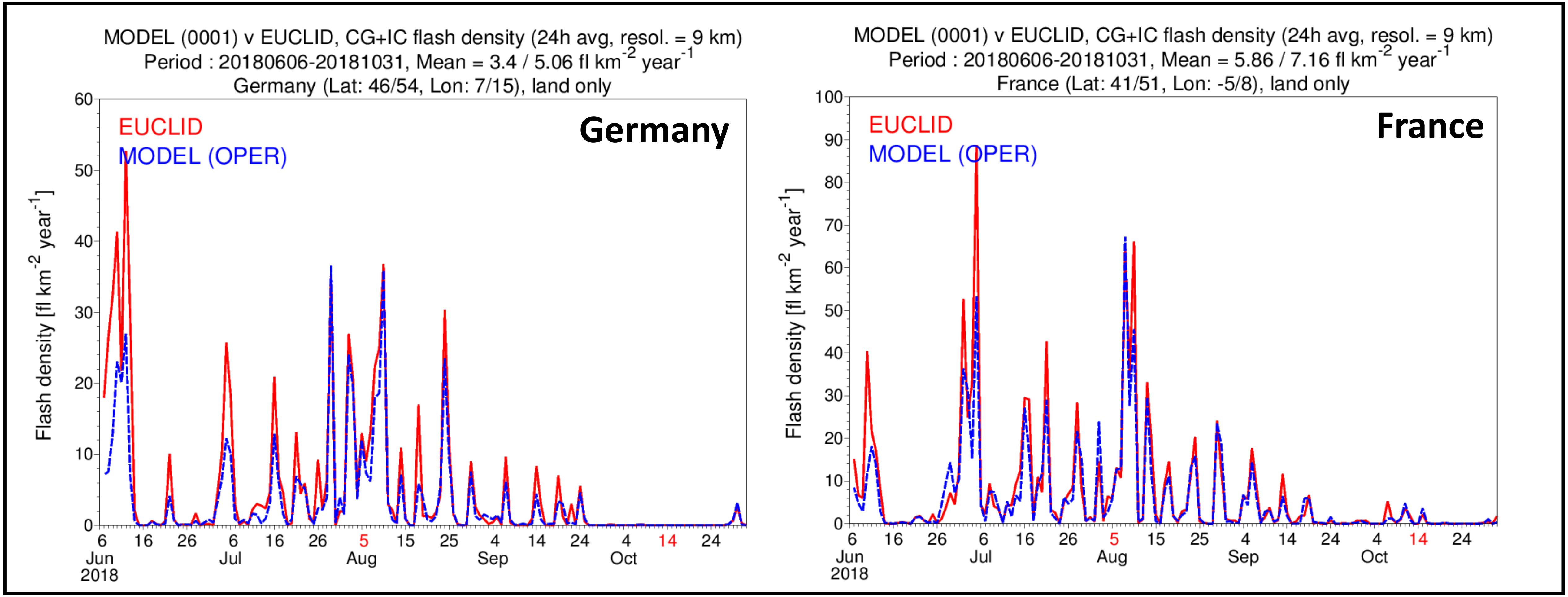


Validation

examples

Comparison of ECMWF MODEL with EUCLID (lightning flash densities)

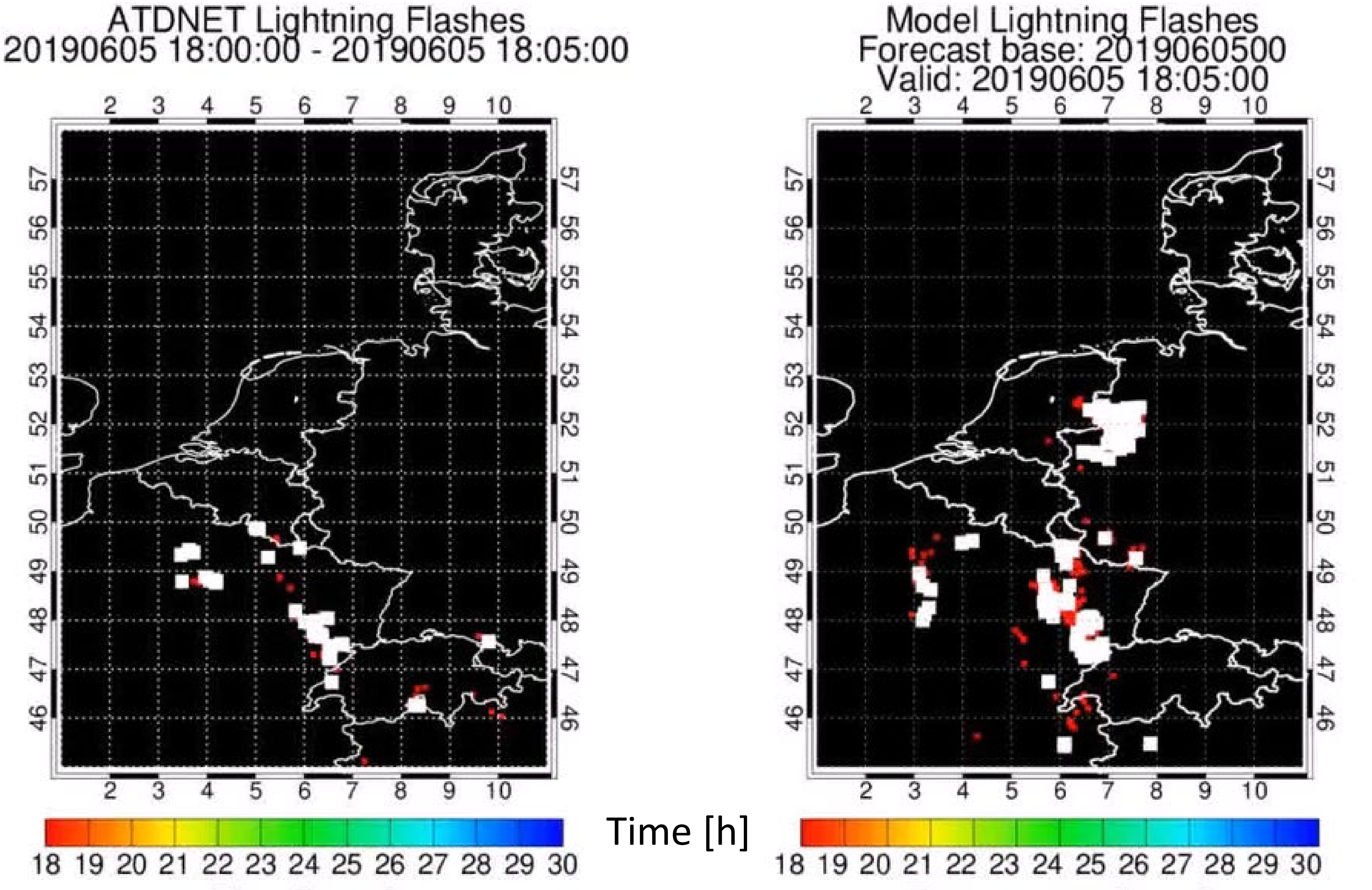
Time series of daily mean flash densities over various European land subdomains during the period 6 Jun-31 Oct 2018: ECMWF model (blue; 9 km) against EUCLID observations (red).



Comparison of model with ATDnet lightning flashes

12h animation of 2-mn flash data starting from 5 June 2018 at 12Z.

9-km resol. L137 model forecast: +18h to +30h range.



ATDNET Lightning Flashes
20190605 18:00:00 - 20190605 18:05:00

Model Lightning Flashes
Forecast base: 2019060500
Valid: 20190605 18:05:00

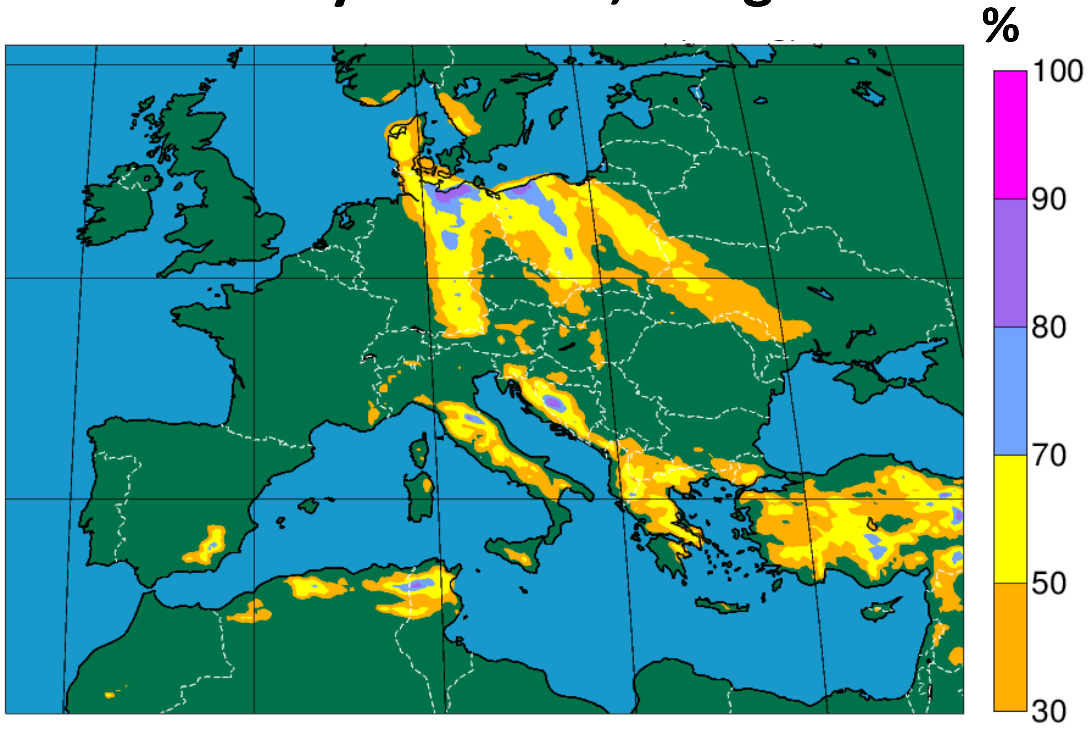
Model flashes were randomly generated to match the simulated flash densities.

Ensemble forecasts can be used to deal with the random and discrete nature of lightning.

ECMWF ensemble forecast

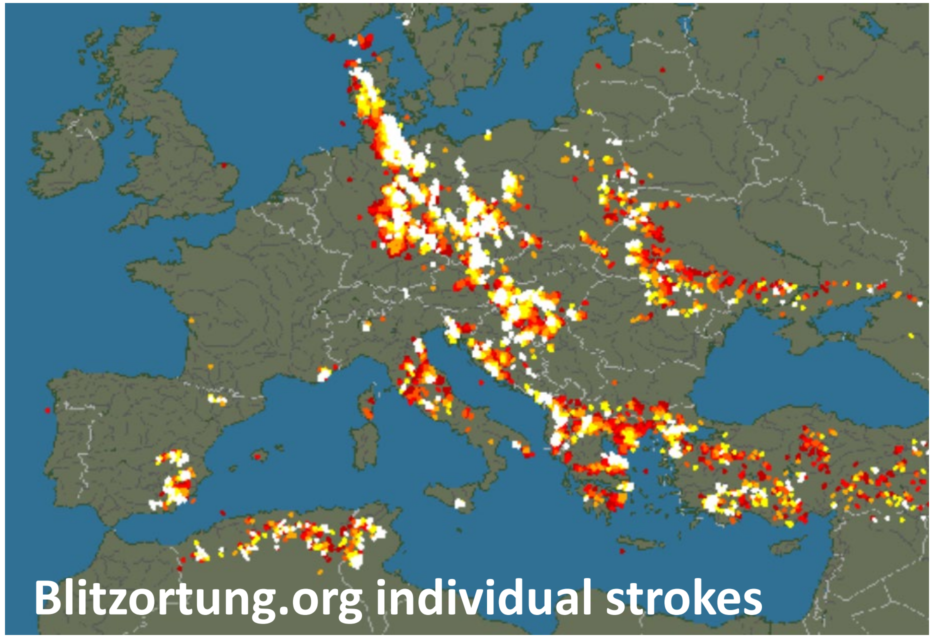
Probability[flash density > 0.1 fl/100km^2/h]

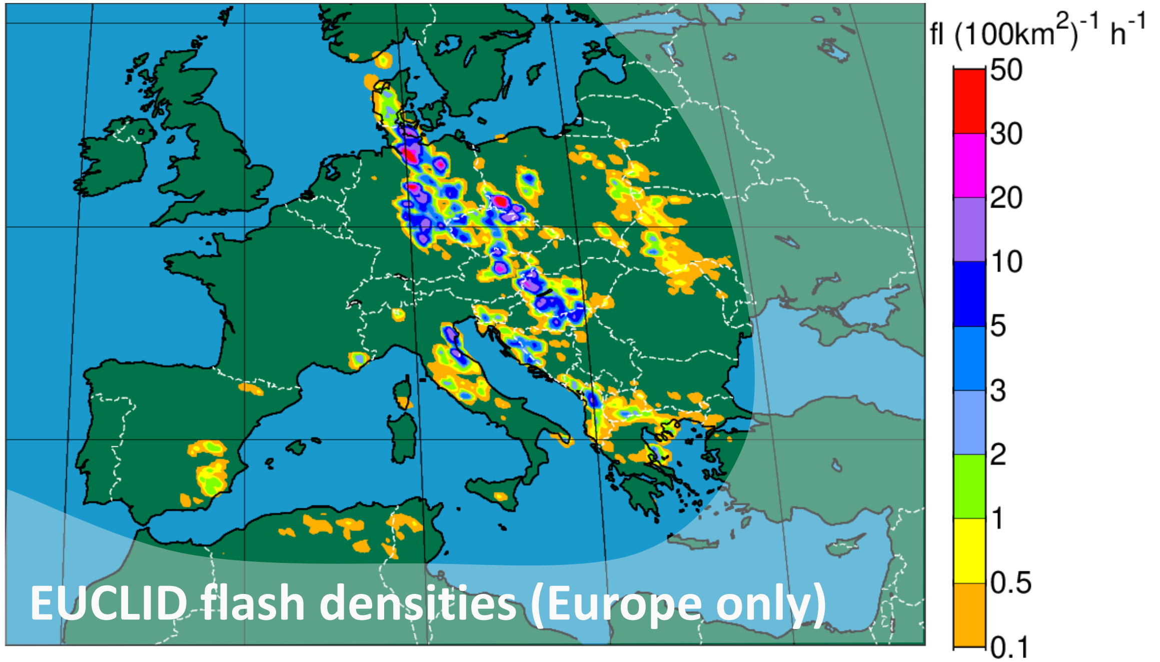
FC Base: 8 May 2018 00Z, Range: +60 to +63h.



→ Ensemble lightning forecasts can offer useful guidance to forecasters up to day 3 (in mid-latitude regions).

Ground-based obs., 10 May 2018 12-15Z


Blitzortung.org individual strokes

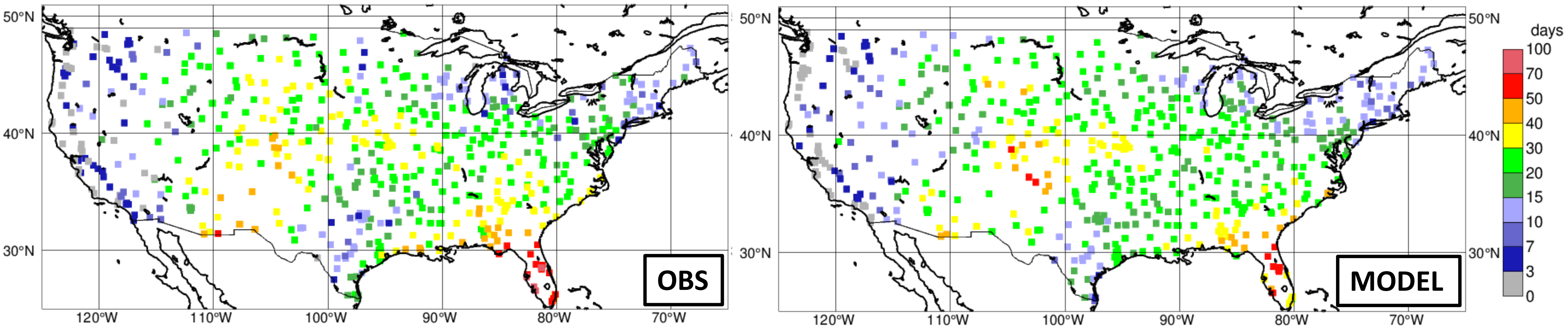

EUCLID flash densities (Europe only)

Validation of thunder days (keraunic levels)

Thunder days can be estimated from the model by counting the number of days for which lightning flash density exceeds 2 /100 km^2/day (empirical threshold).

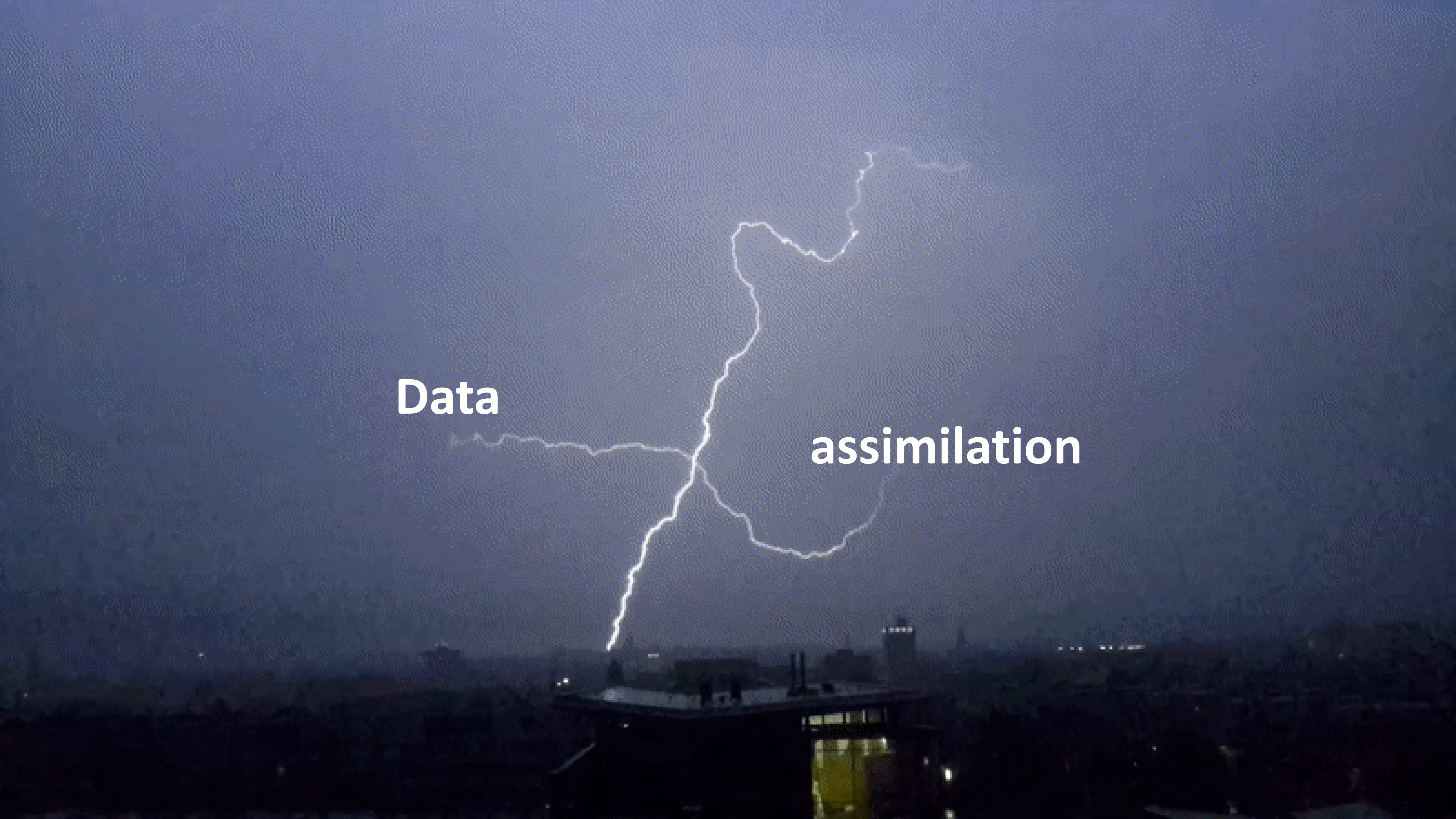
These numbers can then be compared with ground-based observations from both human observers (sound) and automatic stations equipped with lightning sensors (electrical field + optical detection).

Example: Validation against WBAN and ASOS data over the USA:



Thunder days: observed (top) and from model (bottom) in summer 2015.

→ Fairly good agreement: Bias = –0.74 day (–3%); R = 0.827.

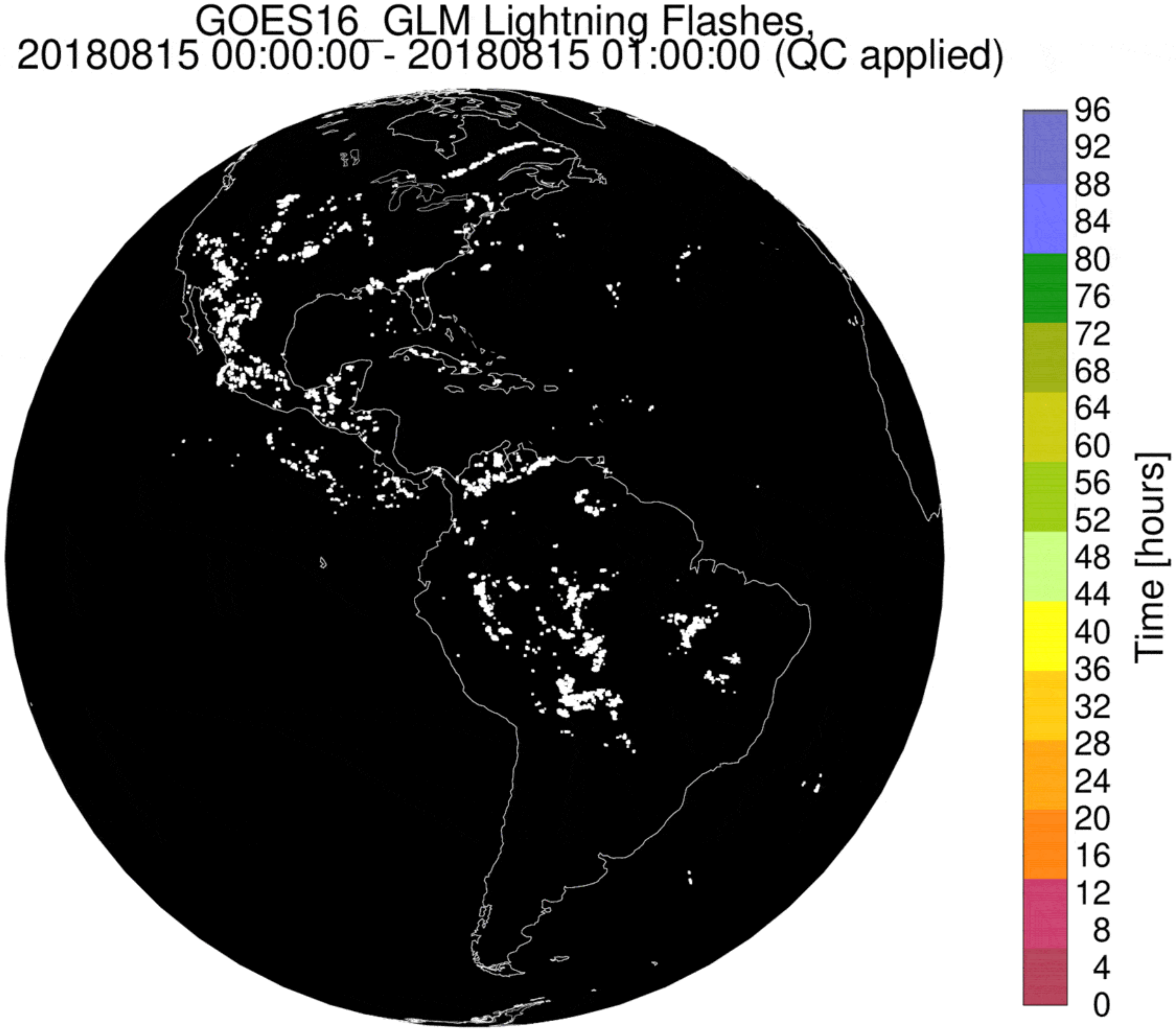


Data assimilation

GOES-16 GLM lightning observations:

- The Geostationary Lightning Mapper (GLM) on board the new NOAA GOES-16 and 17 satellites provides continuous full-disk lightning observations at 8 km resolution (nadir) and in quasi-real-time.

- Lightning pulses are detected through their optical signature in the 777.4 nm oxygen band (lightning peak emission).



Animation of GOES-16 GLM lightning flashes over 4 days.

ECMWF

Assimilation of GOES-16 GLM lightning flash densities: Main set-up.

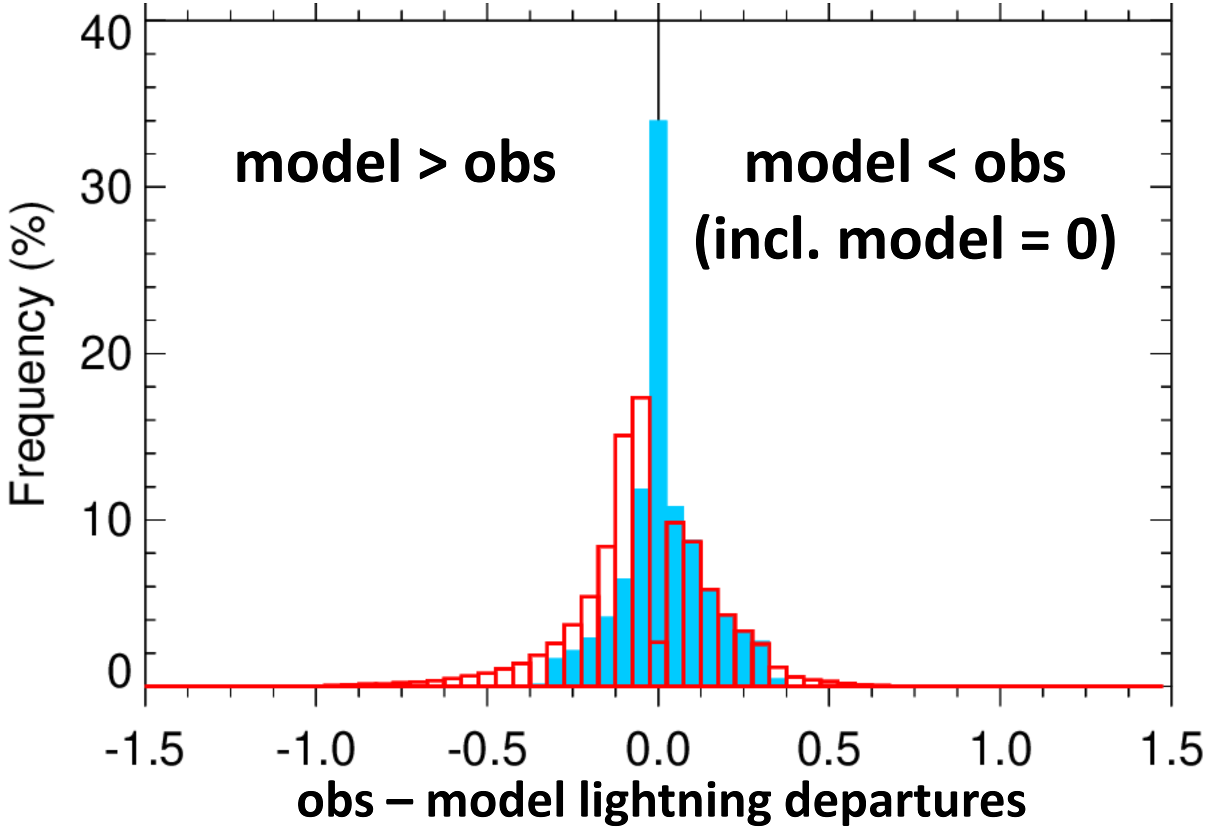
- **Method:**
  - **Direct 4D-Var (like all other observations already assimilated),**
  - **12-hour assimilation window.**

- **Quantity to be assimilated:**
  - **Lightning flash density,**
  - **Averaged over 6 hours (to reduce effects of non-linearities),**
  - **Logarithmic transform applied prior to assimilation (more Gaussian departures).**

- **Lightning observations can provide a direct constraint on convective precipitation within the 4D-Var minimization process (much more difficult to obtain when using precipitation observations, which can be large-scale or convective).**

4D-Var assimilation of GOES-16 GLM lightning flash densities: First long experiment.

Histograms of obs–model lightning departures, before and after assimilation:



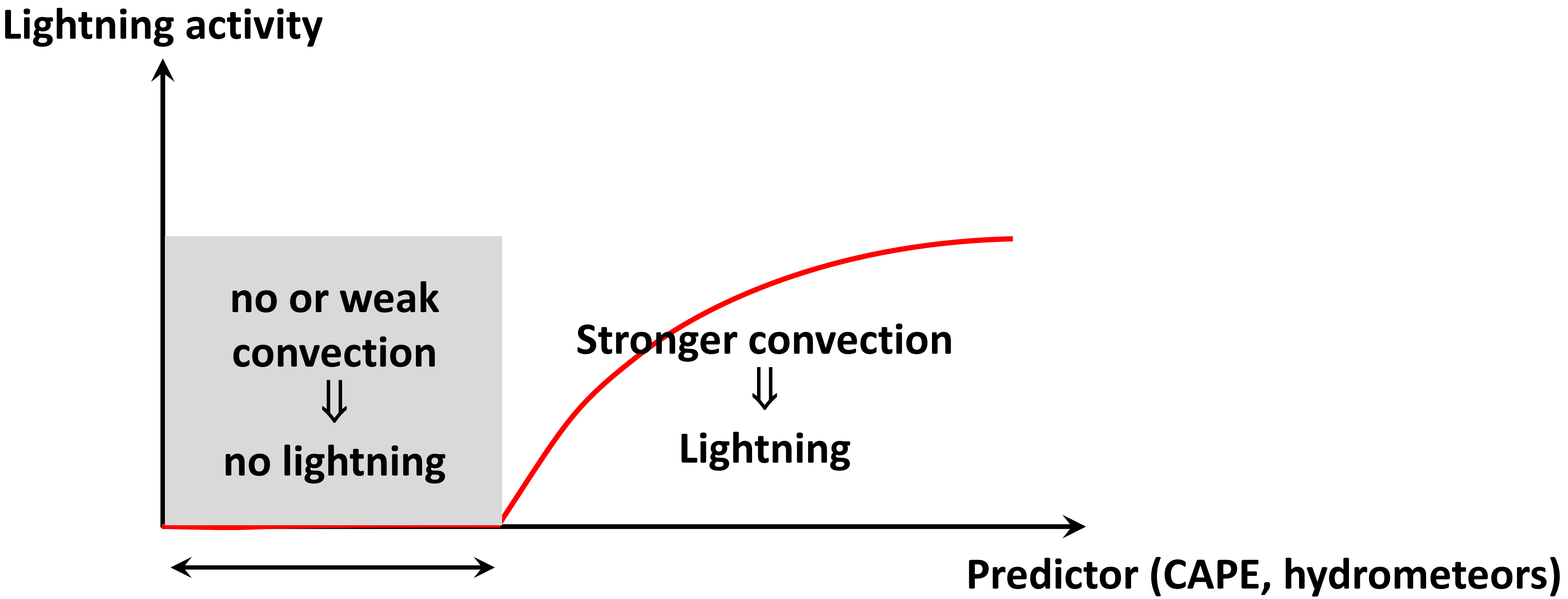
Before assimilation

After assimilation

4D-Var experiment
with GOES-16 GLM
lightning obs.
Jun-Aug 2018
(28-km resol.)

- ✓ Histogram of (obs – model) departures becomes narrower after assimilation → good.
- × However, noticeable asymmetry between (obs > model) and (obs < model) cases: it is usually easier to decrease model lightning than to increase it.

4D-Var assimilation of GOES-16 GLM lightning flash densities: The “zero issue”.



- If the model trajectory (= linearization point for 4D-Var) is lightning-free, any co-located obs with lightning will have no impact in the 4D-Var analysis (no sensitivity in grey-shaded zone).
- How to address this issue in the 4D-Var framework remains unclear.

Summary and prospects

Until now:

- Lately, there has been a growing interest worldwide in predicting lightning activity and assimilating lightning observations in NWP systems.
- Operational prediction of lightning flash densities at ECMWF started in June 2018.
- 4D-Var assimilation of GOES-16 GLM lightning flash densities is being tested (research).

Prospects:

- Improve the lightning parameterization especially at km-scale (e.g, prognostic graupel).
- In lightning data assimilation, the “zero issue” will need to be overcome (esp. in 4D-Var).
- Extend lightning assimilation to GOES-17/18 GLM (Pacific) and MTG-LI (avail. 2023?) and maybe to ground-based networks (despite their more variable detection efficiency).

Thank you!

References: (Ctrl + click to follow links)

Coughlan, R., Di Giuseppe, F, Vitolo, C, Barnard, C, Lopez, P, Drusch, M., 2021: Using machine learning to predict fire-ignition occurrences from lightning forecasts. *Meteorol. Appl.*, 28:e1973. https://doi.org/10.1002/met.1973

Lopez, P., 2020: Quality Control for GOES Geostationary Lightning Mapper Level-2 flash products, *ECMWF Technical Memorandum 872,* 15 pages, available at https://www.ecmwf.int/en/elibrary/19804-quality-control-goes-geostationary-lightning-mapper-level-2-flash-products

Lopez, P., 2018: Promising results for lightning predictions, *ECMWF Newsletter 155, Spring 2018*, 14-19.

Lopez, P., 2016: A lightning parameterization for the ECMWF Integrated Forecasting System, *Monthly Weather Review*, **144**, 3057-3075.

Lightning parameterization implementation in ECMWF's IFS (model version 45R1, as of 2018): https://www.ecmwf.int/en/elibrary/18714-part-iv-physical-processes

Extras

Cloud electrification and lightning production (1)

Convective clouds become electrified through the interactions between hydrometeors (different types and fall velocities).

* Two main mechanisms:

Non-inductive:

- Colliding graupel and snow/cloud ice particles gain opposite charges.
- Charge polarity function of T (reversal around –10°C), liquid water content and relative humidity.
- Required to initiate storm electrification.

Inductive:

- Requires the pre-existence of an ambient electric field $E > \sim 10$ kV m^{-1} to polarize hydrometeors prior to their collisions.
- Maximum efficiency between frozen particles and super-cooled droplets.

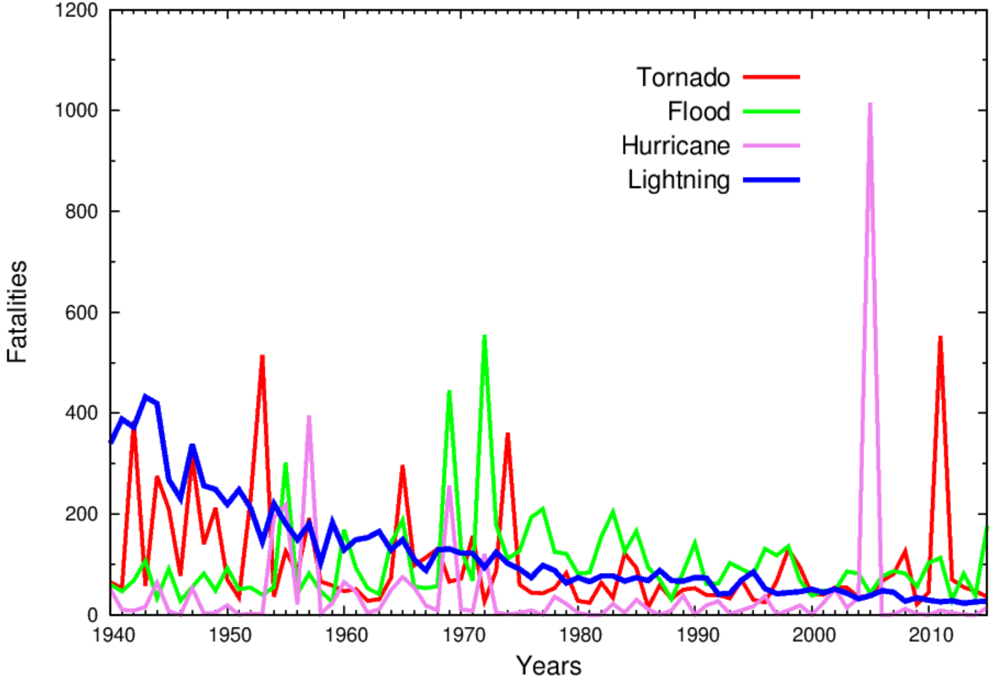
* Charge separation due to fall speed differences → build-up of the electric field.

* Lightning discharges occur when $E > 100$-200 kV m^{-1} (typically).

Impact of lightning on human activities

Lightning can impact various human activities:

- Power supply (outages).
- Air traffic operations (in-flight and at airports).
- Forestry (wildfires, esp. after droughts).
- Buildings (structural and electrical damages).
- Health: injuries/fatalities in humans (and cattle).



Weather fatalities in the United States (1940-2015; data from NOAA)

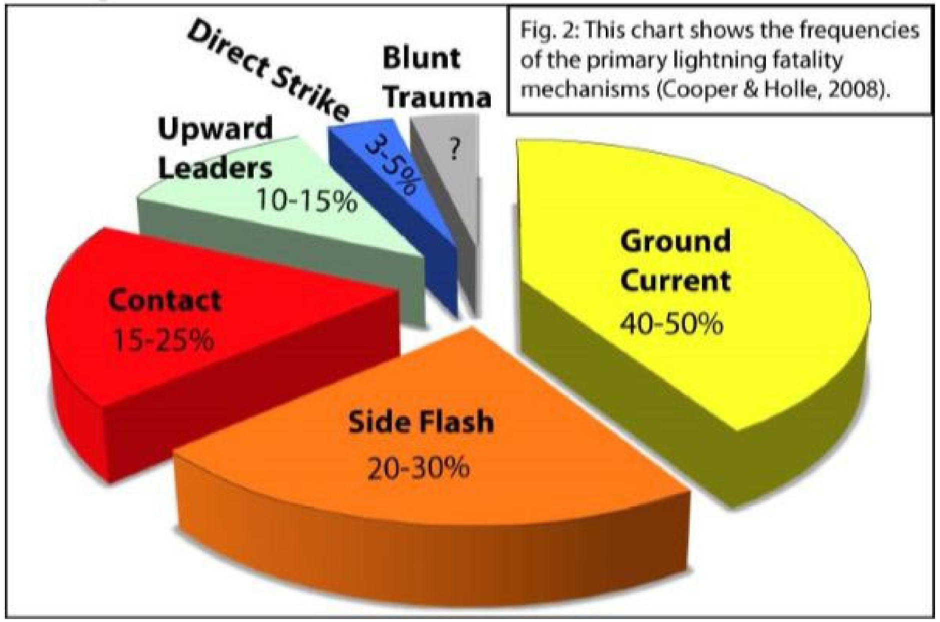
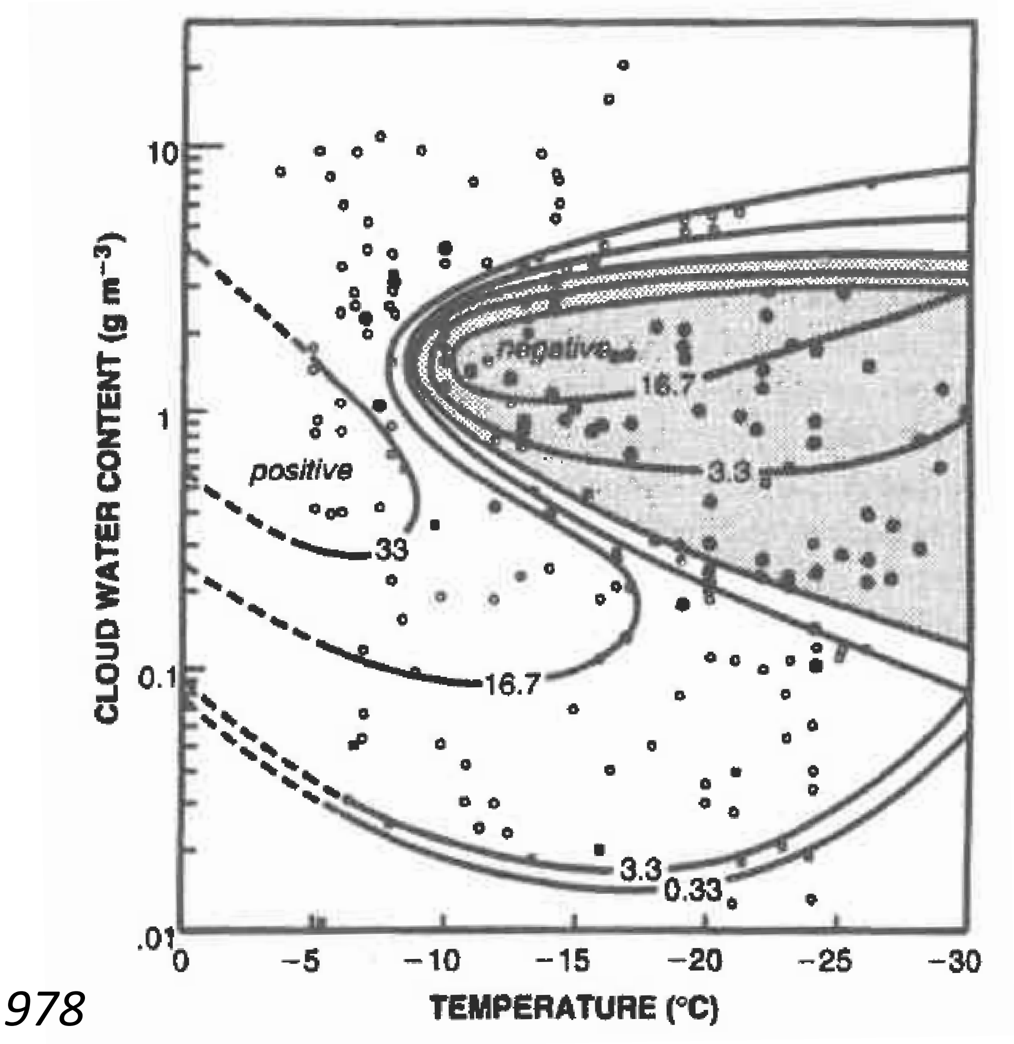


Fig. 2: This chart shows the frequencies of the primary lightning fatality mechanisms (Cooper & Holle, 2008).

Causes of lightning fatalities in USA

Effect of temperature and cloud liquid water content on charge separation

Charge gained by rimed graupel = f(T,q_{liq})



From Takahashi 1978

Existing lightning parameterizations

Over the past decades, various lightning parameterizations with different levels of complexity have been proposed for NWP applications:

- **Simple formulae:**
  - ☞ **based one or two convective outputs: cloud top height, precipitation rate or mass-flux.**
  - ☞ **used in global models.**

  e.g. Price and Rind (1994), Meijer et al. (2001), Grewe et al. (2001), Allen and Prickering (2002), McCaul et al. (2009)), Lopez (2016).

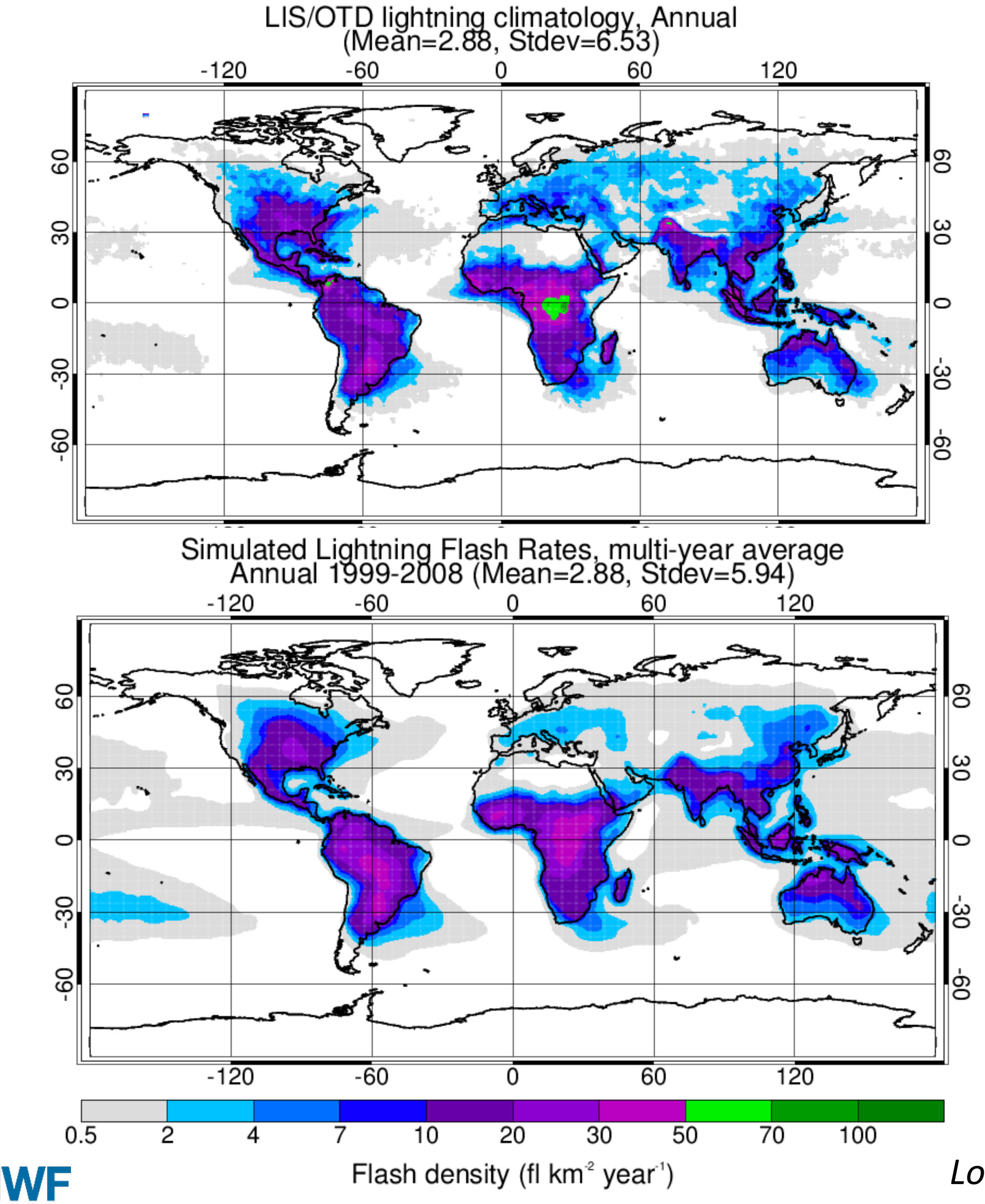
- **Bulk electrification schemes:**
  - ☞ **explicit computations of charge distribution and electric field.**
  - ☞ **require detailed microphysics (hail, graupel).**
  - ☞ **used in limited-area models.**

  e.g. Mansell et al. (2005).

- **Explicit schemes:**
  - ☞ **describe the propagation of individual flashes (stochastic approach).**
  - ☞ **require very detailed microphysics.**
  - ☞ **used in cloud-resolving models (very expensive!).**

  e.g. Mansell et al. (2002), Barthe and Pinty (2005).

New lightning parameterization versus LIS/OTD climatology



LIS/OTD climatology (1995-2010)

Mean (CG+IC) lightning flash densities (flashes/km²/year)

ECMWF model “climatology” (1999-2008)

80-km resolution, L137

10 × 1 year

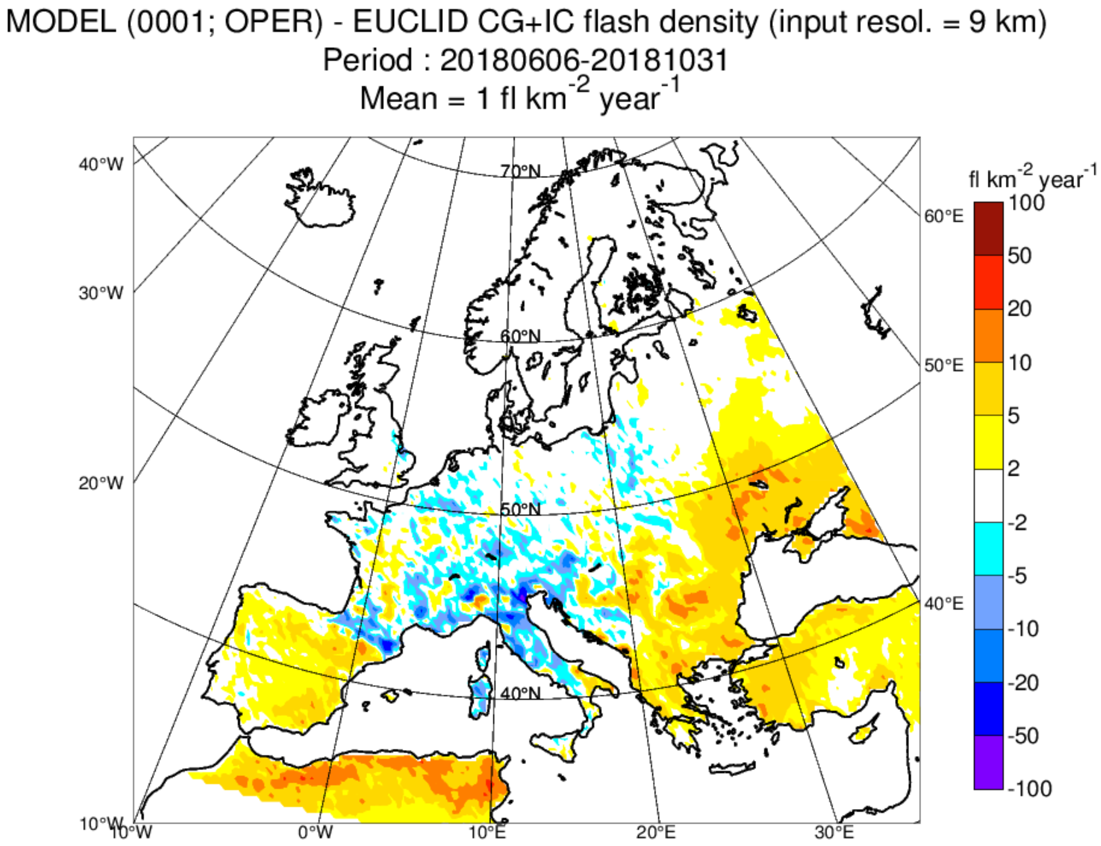
ECMWF

Lopez, Monthly Weather Review, 2016

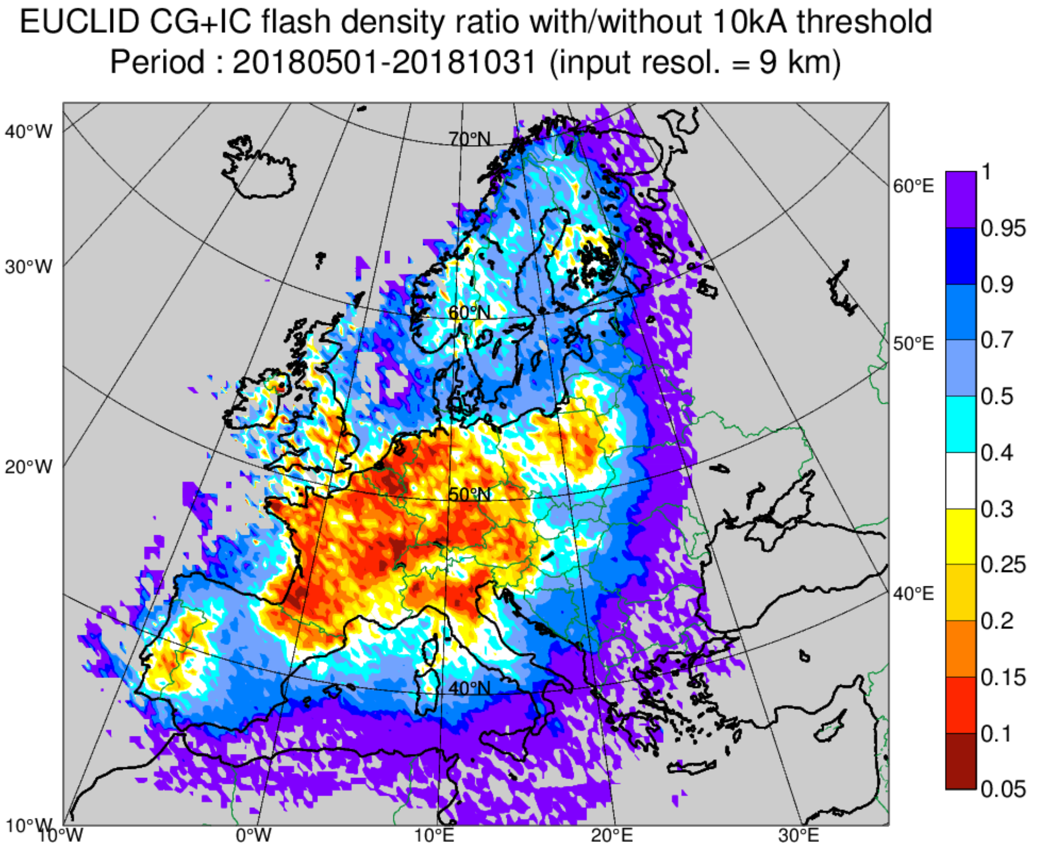
Comparison of ECMWF MODEL with EUCLID ground-based network

Left: Oper 24h ECMWF forecasts vs EUCLID over period: 6 Jun - 31 Oct 2018 over Europe.

Right: EUCLID network's detection efficiency inferred using 10kA/0kA peak current thresholds in the clustering algorithm used to compute flash densities from strokes.


MODEL (0001; OPER) - EUCLID CG+IC flash density (input resol. = 9 km)
Period : 20180606-20181031
Mean = 1 fl km^{-2} $year^{-1}$

ECMWF MODEL – EUCLID mean difference

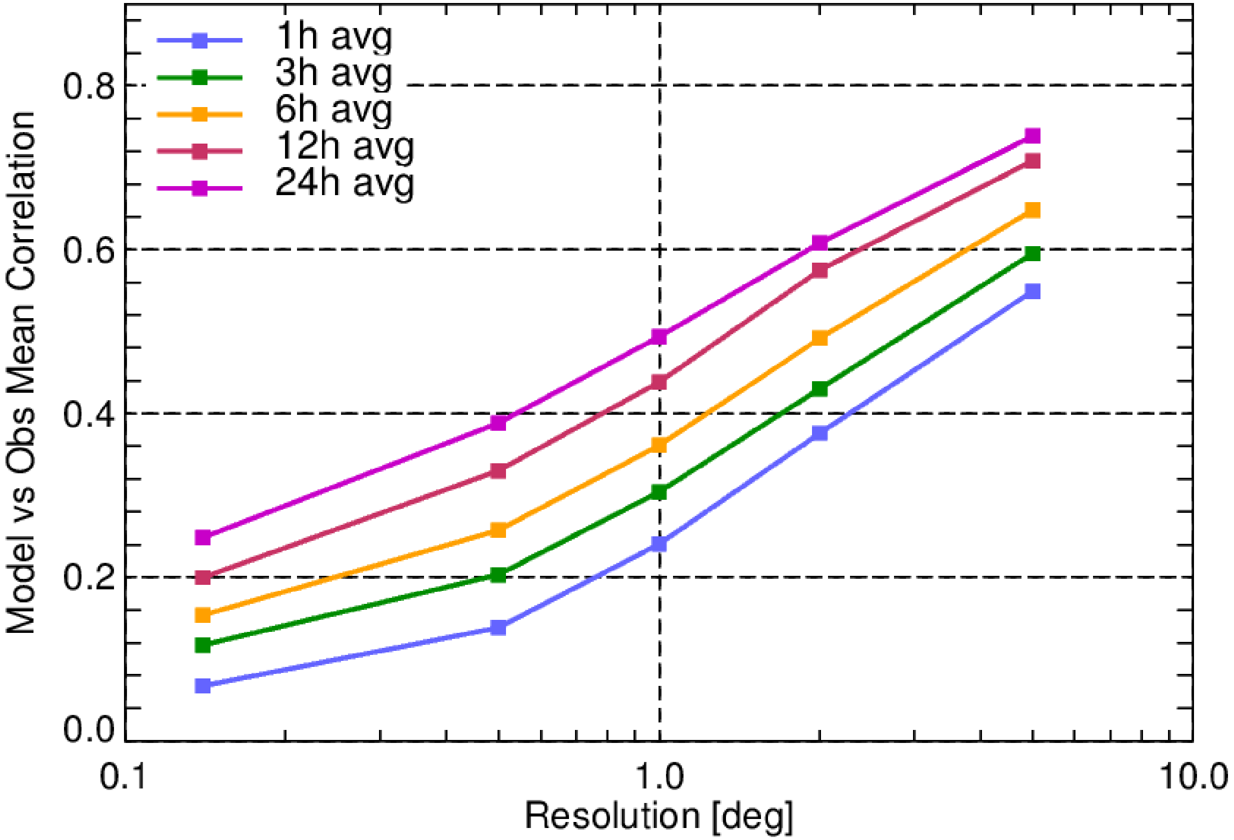

EUCLID CG+IC flash density ratio with/without 10kA threshold
Period : 20180501-20181031 (input resol. = 9 km)

EUCLID ratio: $f_T(I_{pk} \geq 10kA) / f_T(I_{pk} \geq 0kA)$

ECMWF model vs UBIMET LDS observations

Mean correlations (between maps of flash density) for various averaging scales in time and space.

Based on 0-24h forecasts (16-km resol.) over Europe in summer 2015.

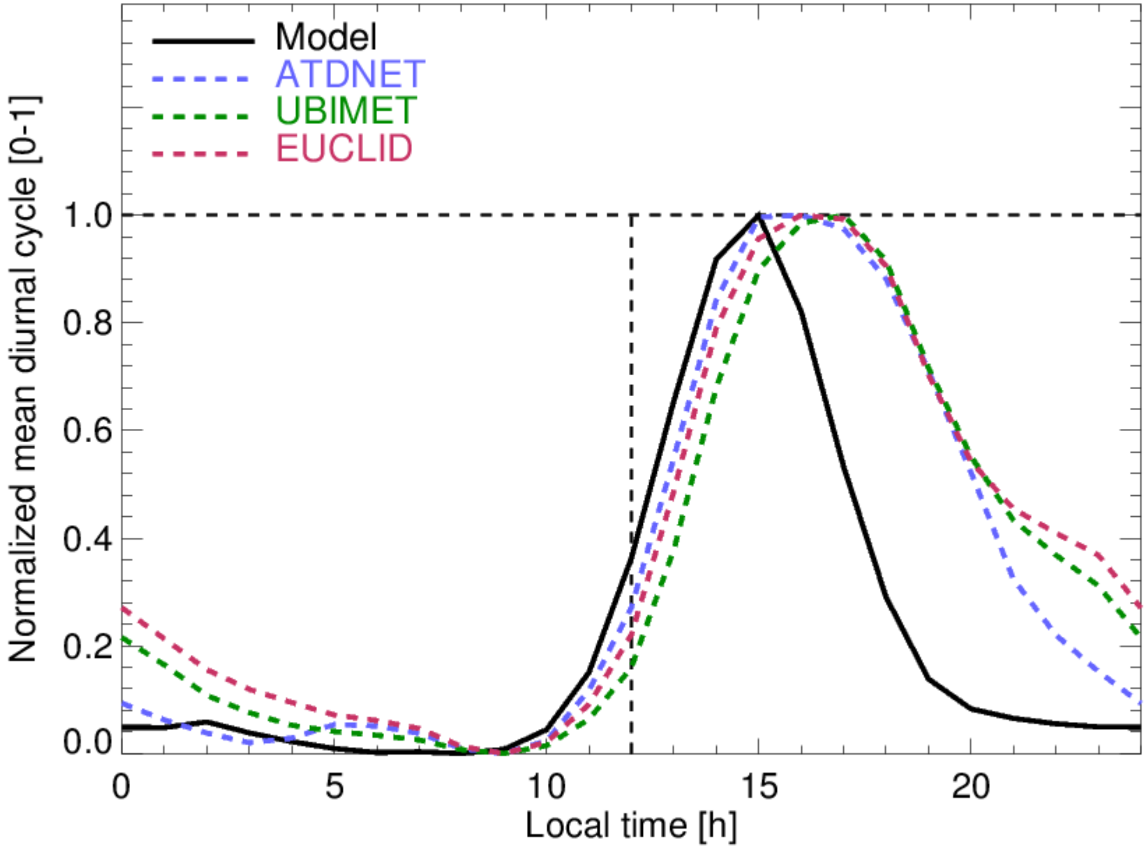


Model and observations correlation improves for wider temporal and spatial scales.

ECMWF model vs various ground-based lightning networks

Diurnal cycle of mean flash densities (normalized by amplitude).

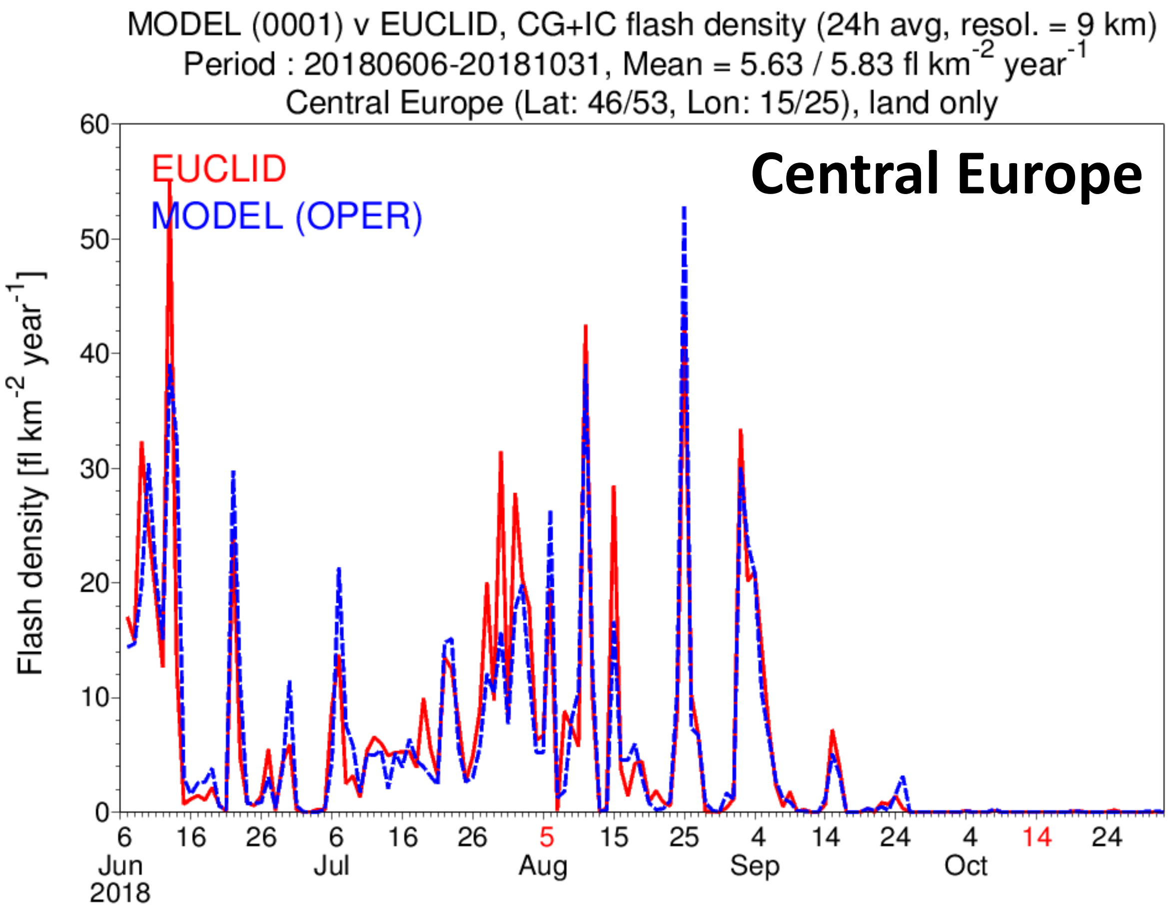
Based on 0-24h forecasts (16-km resol.) over Europe in summer 2015.



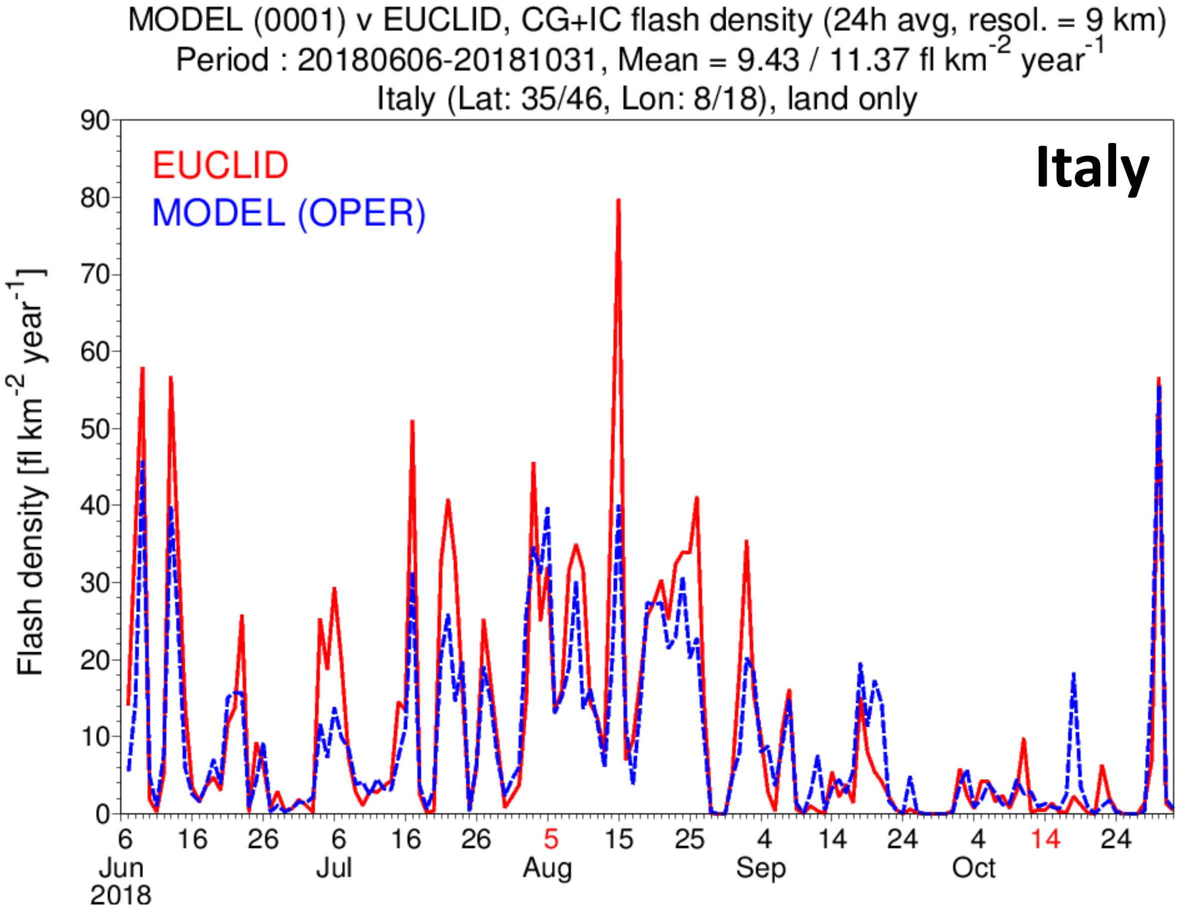
Predicted lightning declines too early in the afternoon.
→ Consistent with previous studies focusing on precipitation.

Comparison of ECMWF MODEL with EUCLID (lightning flash densities)

Time series of daily mean flash densities over various European land subdomains during the period 6 Jun-31 Oct 2018: ECMWF model (blue; 9 km) against EUCLID observations (red).



MODEL (0001) v EUCLID, CG+IC flash density (24h avg, resol. = 9 km)
Period : 20180606-20181031, Mean = 5.63 / 5.83 fl km^{-2} year^{-1}
Central Europe (Lat: 46/53, Lon: 15/25), land only

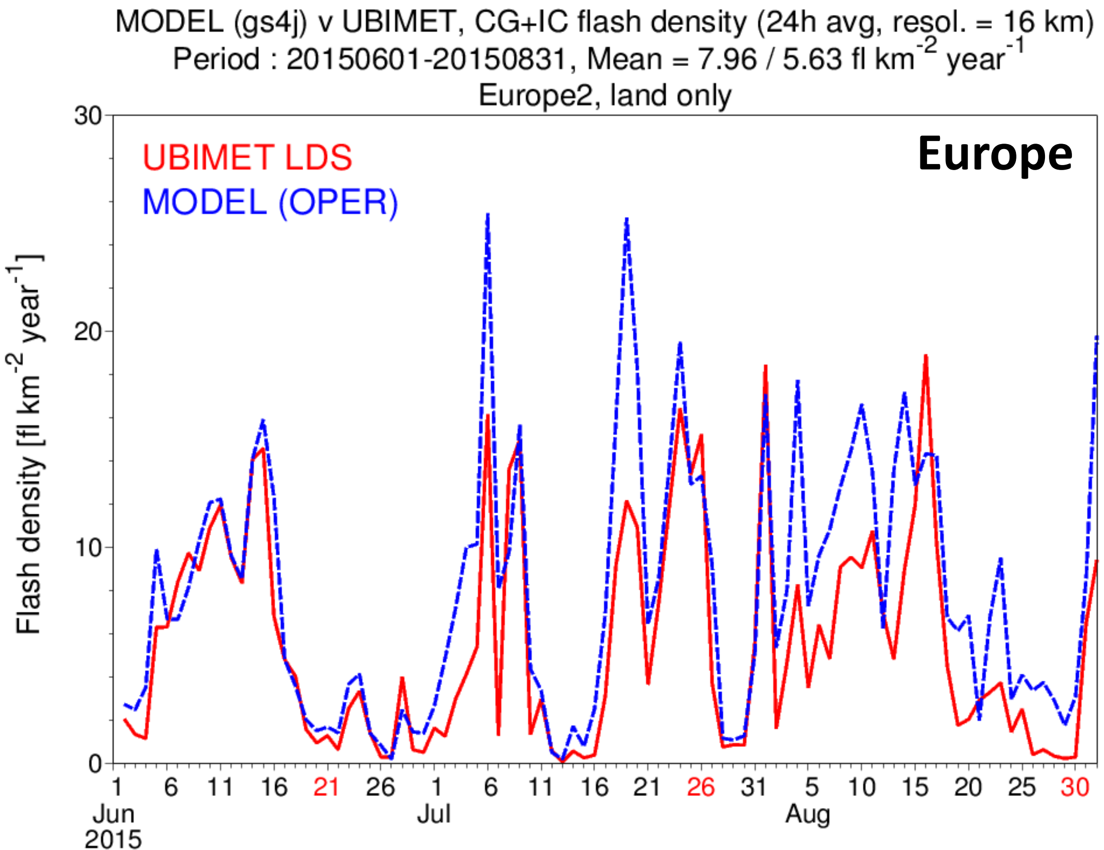


MODEL (0001) v EUCLID, CG+IC flash density (24h avg, resol. = 9 km)
Period : 20180606-20181031, Mean = 9.43 / 11.37 fl km^{-2} year^{-1}
Italy (Lat: 35/46, Lon: 8/18), land only

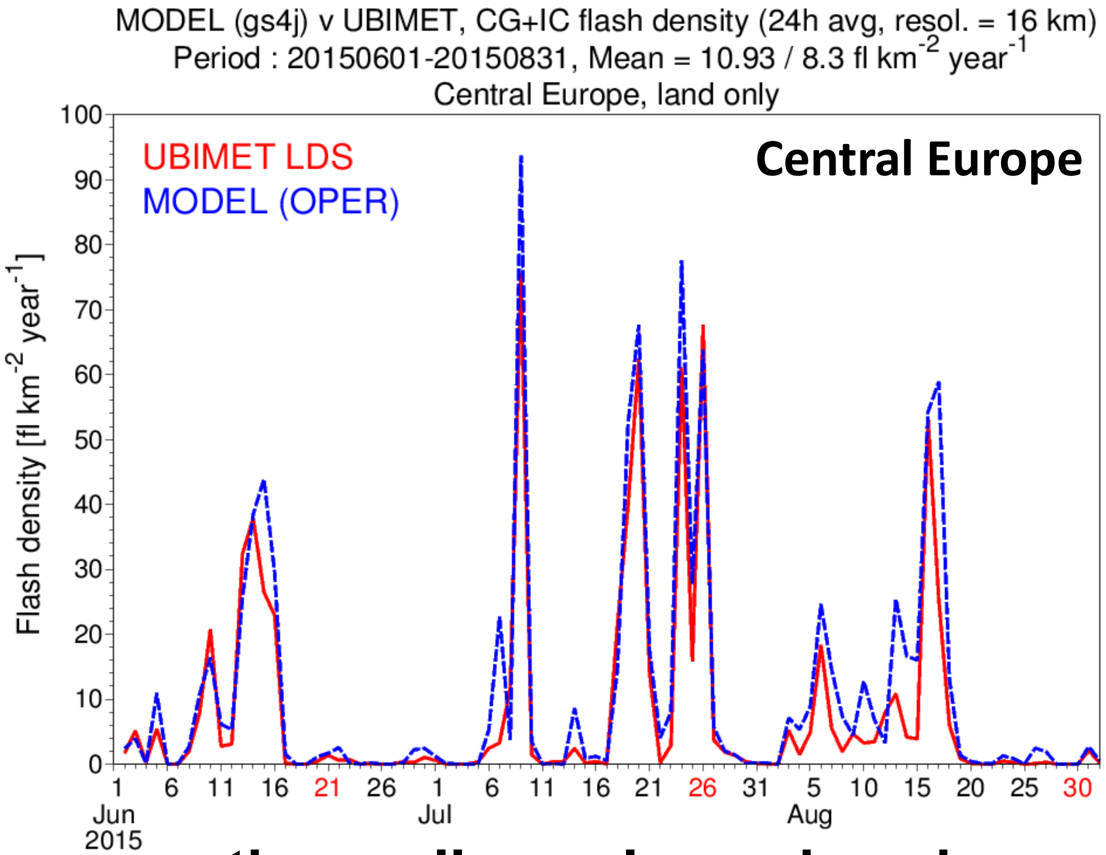
ECMWF model vs UBIMET LDS observations.

Time evolution of daily average lightning flash densities.

Based on 24h forecasts (16 km res.) over Europe in summer 2015.



MODEL (gs4j) v UBIMET, CG+IC flash density (24h avg, resol. = 16 km)
Period : 20150601-20150831, Mean = 7.96 / 5.63 fl km^{-2} year^{-1}
Europe2, land only

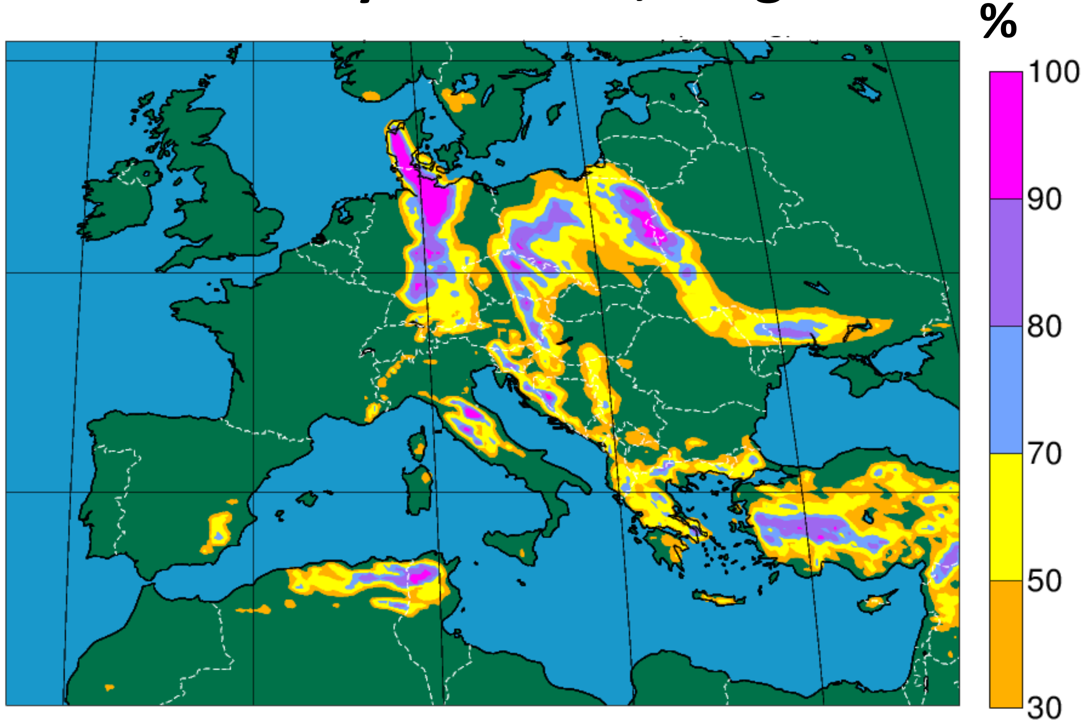


MODEL (gs4j) v UBIMET, CG+IC flash density (24h avg, resol. = 16 km)
Period : 20150601-20150831, Mean = 10.93 / 8.3 fl km^{-2} year^{-1}
Central Europe, land only

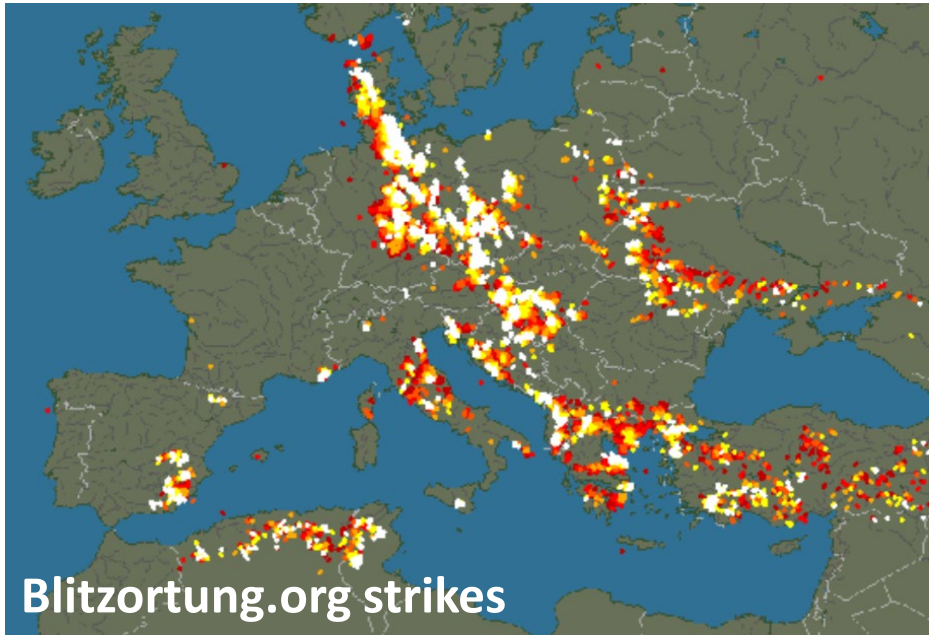
Model and observed daily variations agree rather well over large domains.

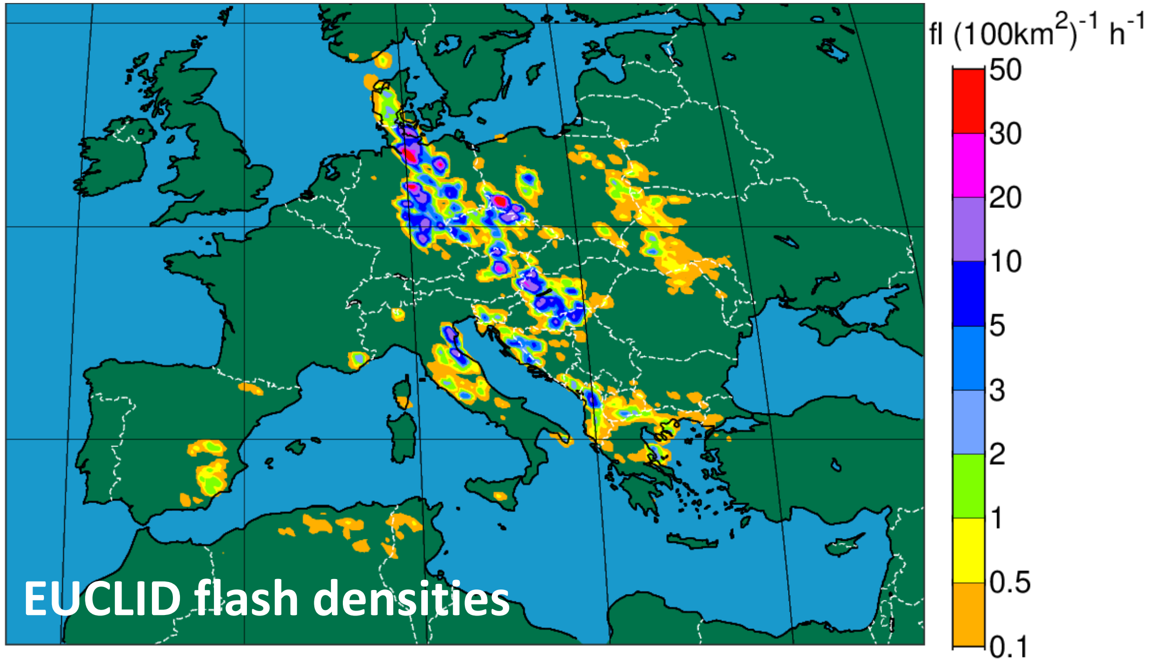
The ensemble forecast approach is particularly adequate to deal with the random and discrete nature of lightning.

Example: ECMWF ensemble forecast
Prob[flash density > 0.1 fl/100km²/h]
FC base: 10 May 2018 00Z, range: +12 → +15h



Observations, 10 May 2018 15Z

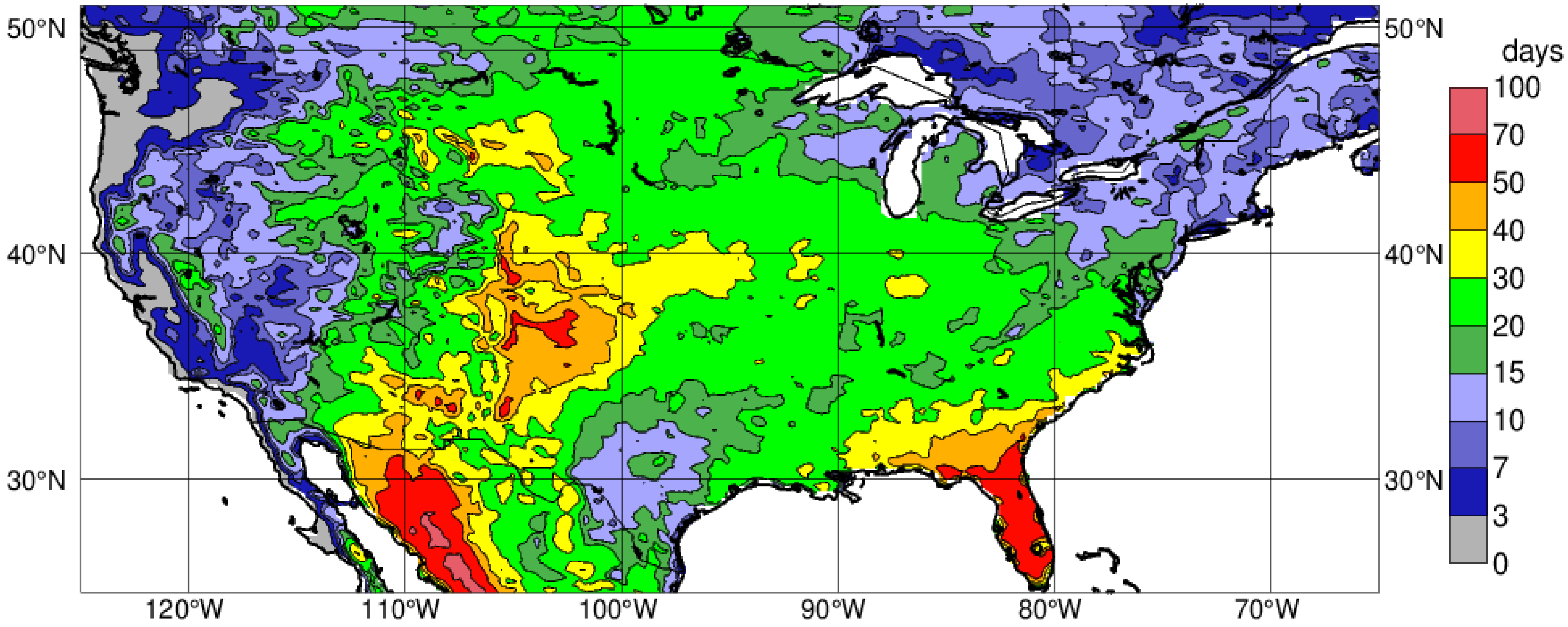


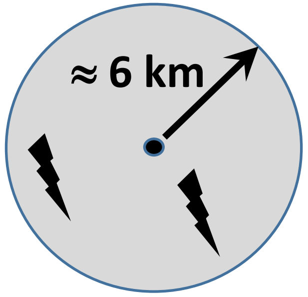


Validation of keraunic levels (1)

The number of days with thunder at a given point is estimated from the model by counting the number of days where total-lightning flash density exceeds an (empirical) threshold of 2 flashes/100 km^2/day.

Estimated total number of days with thunder from a series of 24h forecasts over summer 2015 (31 km resolution, 137 vert. levels, cy43r1).

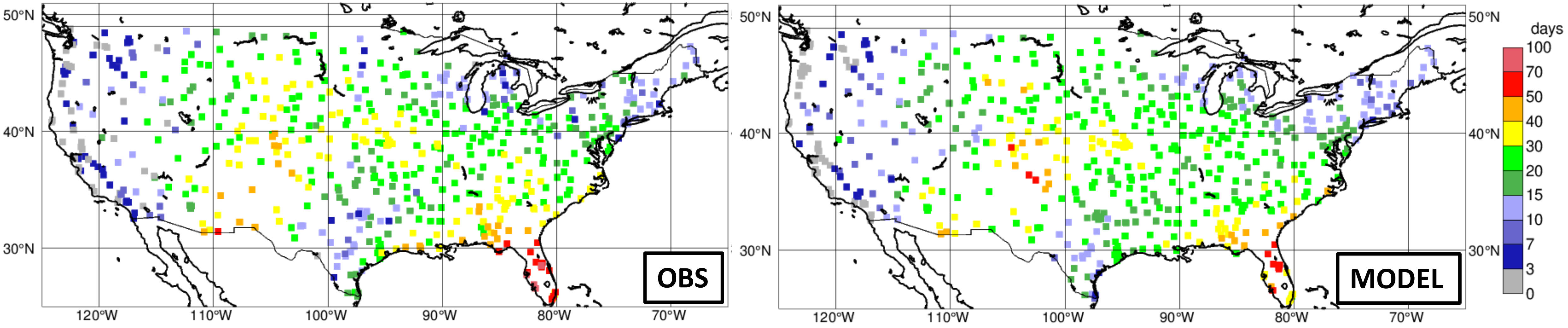




Validation of keraunic levels (2)

Keraunic levels estimated from the model can be compared with ground-based observations from human observers (sound) and automatic stations equipped with lightning sensors (electrical field + optical detection).

→ WBAN and ASOS data over the USA (extracted from the GHCN global dataset).



Total number of days with thunder: observed (top) and from model (bottom) in summer 2015.

→ Fairly good agreement: Bias = –0.74 day (–3%); R = 0.827.

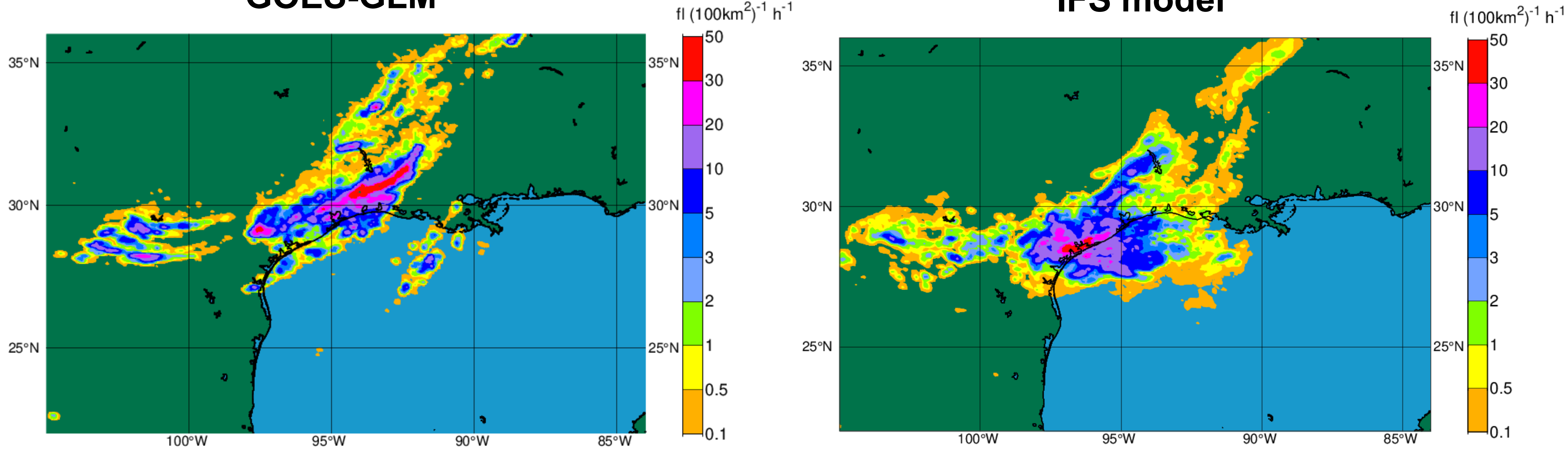
Lightning parameterization: Moving to km-scale resolution.

The original lightning parameterization was based on outputs from the convection scheme of the IFS.

It had to be revised to work at km-scale resolution, when the convection scheme is switched off (i.e., resolved convection) → Use resolved hydrometeors instead.

Lightning flash densities from GOES-GLM observations against IFS FC00Z+6h at 9 km.

1 October 2018 (over Texas)



GOES-GLM

IFS model

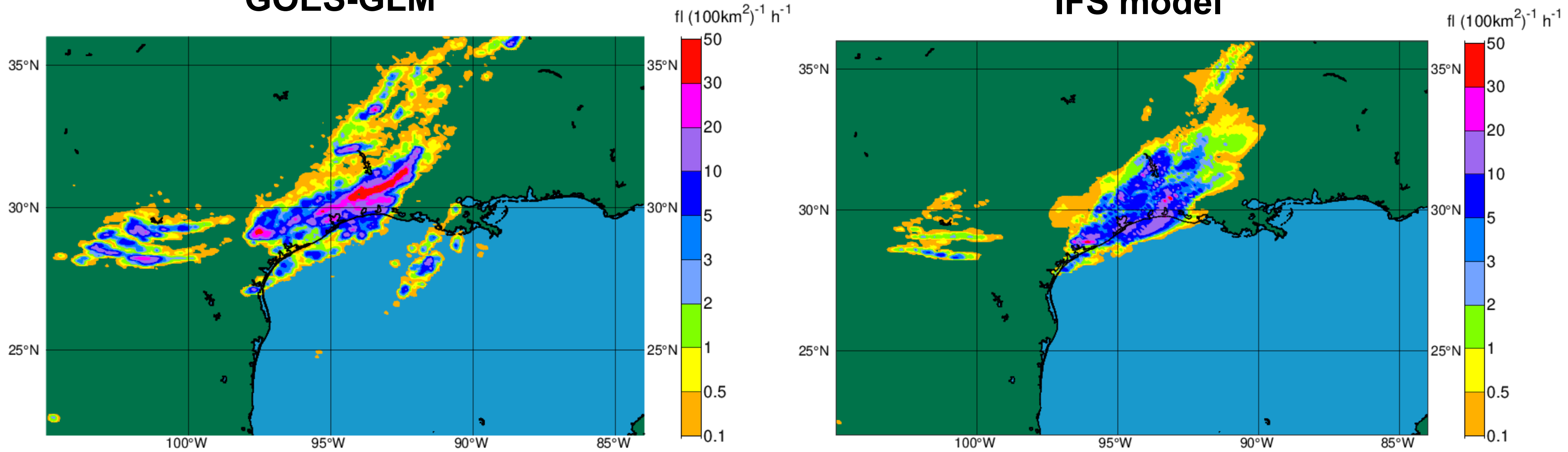
Lightning parameterization: Moving to km-scale resolution.

The original lightning parameterization was based on outputs from the convection scheme of the IFS.

It had to be revised to work at km-scale resolution, when the convection scheme is switched off (i.e., resolved convection) → Use resolved hydrometeors instead.

Lightning flash densities from GOES-GLM observations against IFS FC00Z+6h at 2.5 km.

1 October 2018 (over Texas)

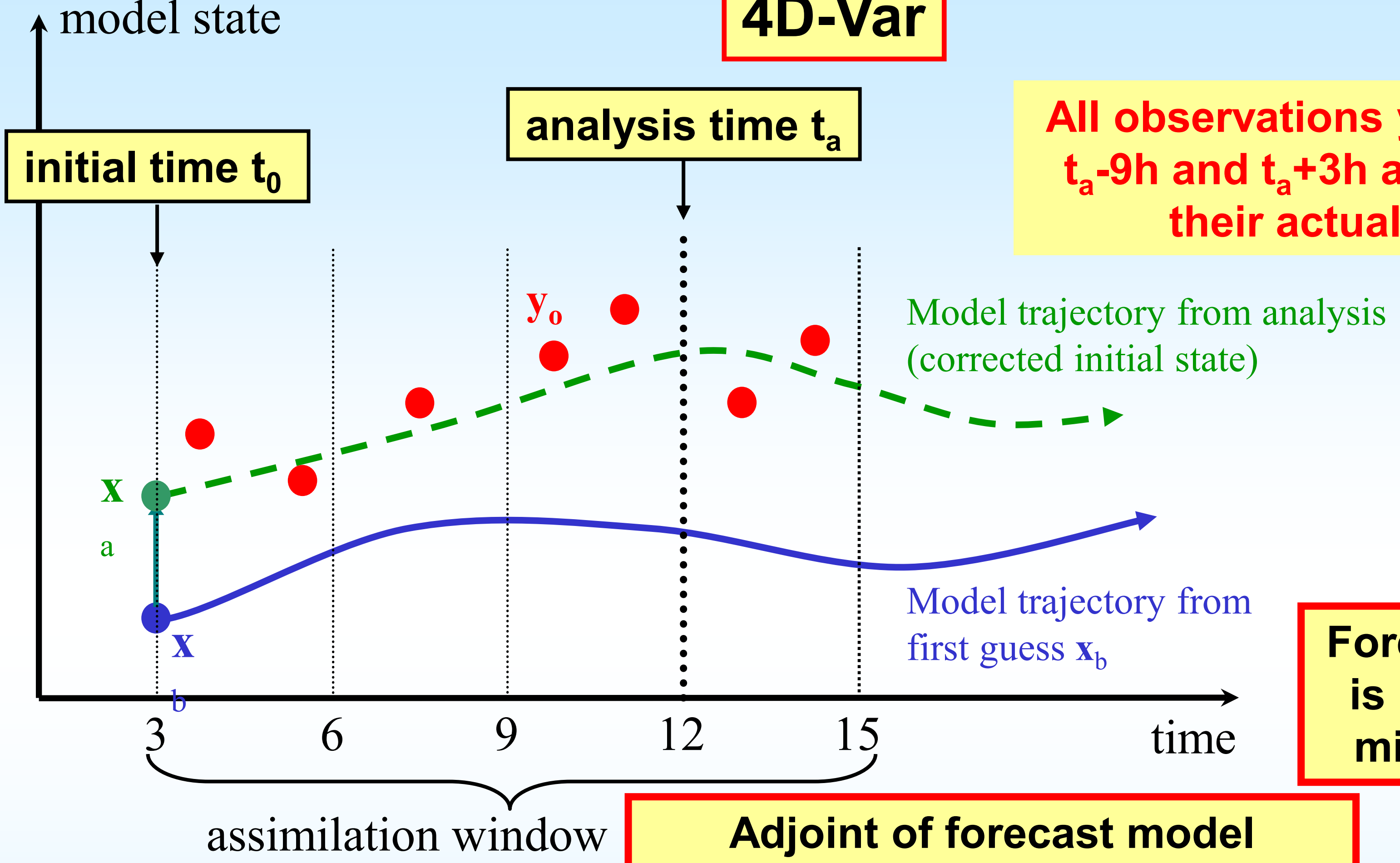


GOES-GLM

IFS model

However, proper retuning at such high resolution will be needed, once month-long experimentation becomes more affordable.

4D-Var



All observations y_o between t_a-9h and t_a+3h are valid at their actual time

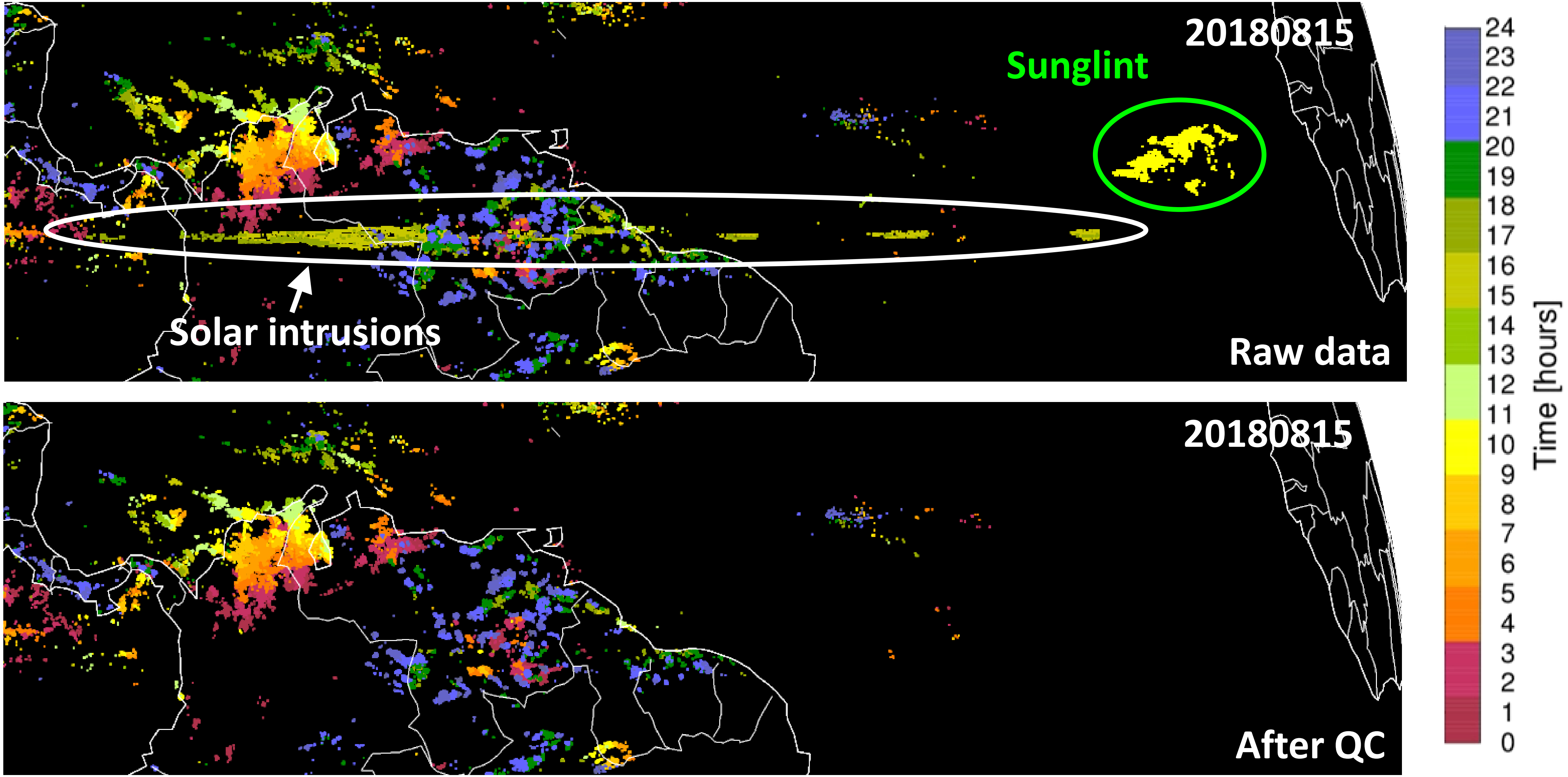
Forecast model is involved in minimization

Adjoint of forecast model with simplified linearized physics

$$\min J = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_b) + \frac{1}{2}\sum_{i=0}\left(H_i(M[\mathbf{x}_0])) - \mathbf{y}_{oi}\right)^T \mathbf{R}_i^{-1}\left(H_i(M[\mathbf{x}_0])) - \mathbf{y}_{oi}\right)$$

$$\Leftrightarrow \nabla_{\mathbf{x}} J = \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_b) + \sum_{i=0}^{n} \boxed{\mathbf{M}^T[t_i,t_0]}\,\mathbf{H}_i^T \mathbf{R}_i^{-1}\left(H_i(M[\mathbf{x}_0])) - \mathbf{y}_{oi}\right) = 0$$

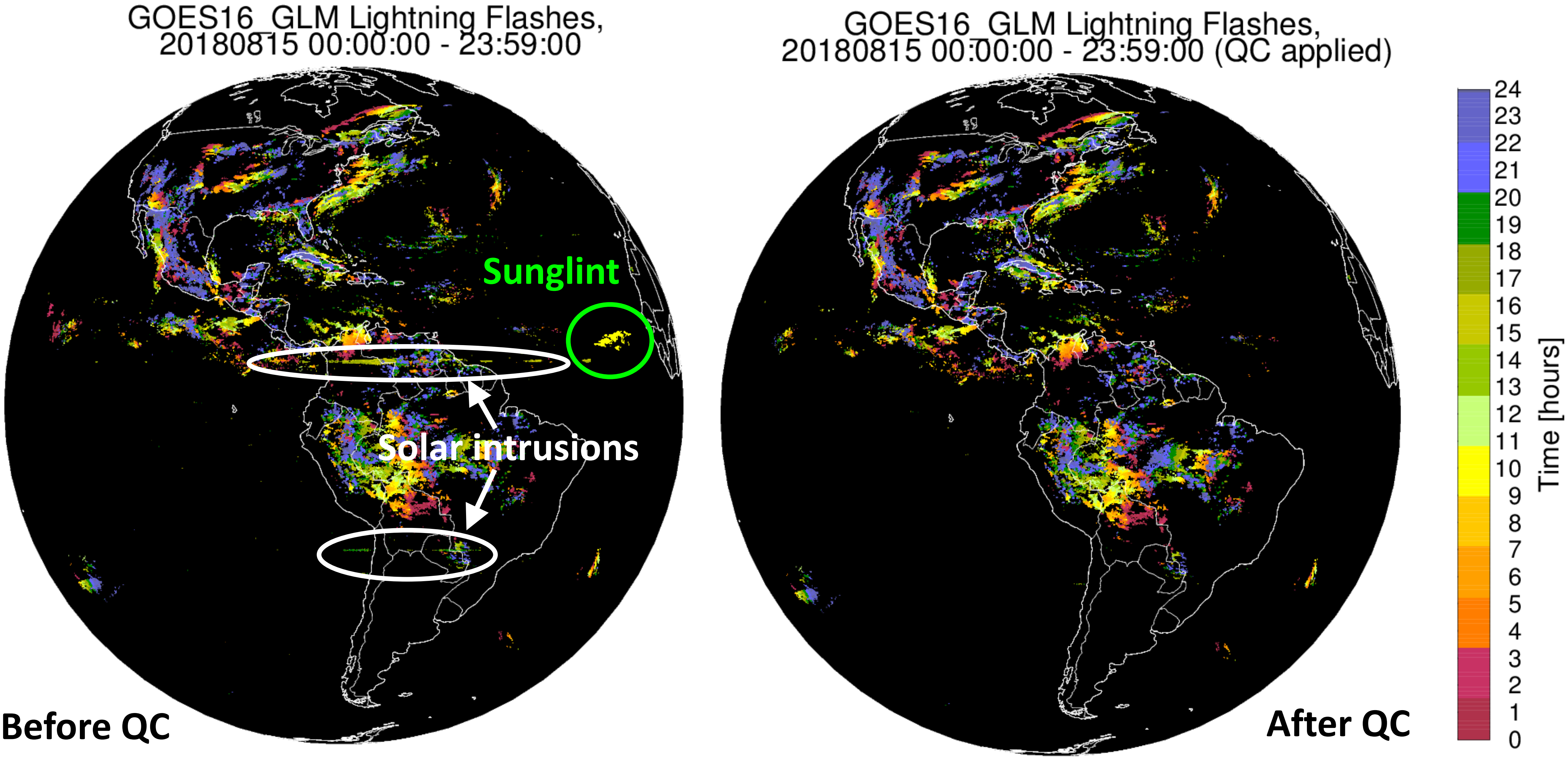
GOES-16 GLM flash data: Quality control (example; zoom over South America)



Lopez 2021, ECMWF Tech Memo 872

ECMWF

GOES-16 GLM flash data: Quality Control (example)



GOES16_GLM Lightning Flashes, 20180815 00:00:00 - 23:59:00

GOES16_GLM Lightning Flashes, 20180815 00:00:00 - 23:59:00 (QC applied)

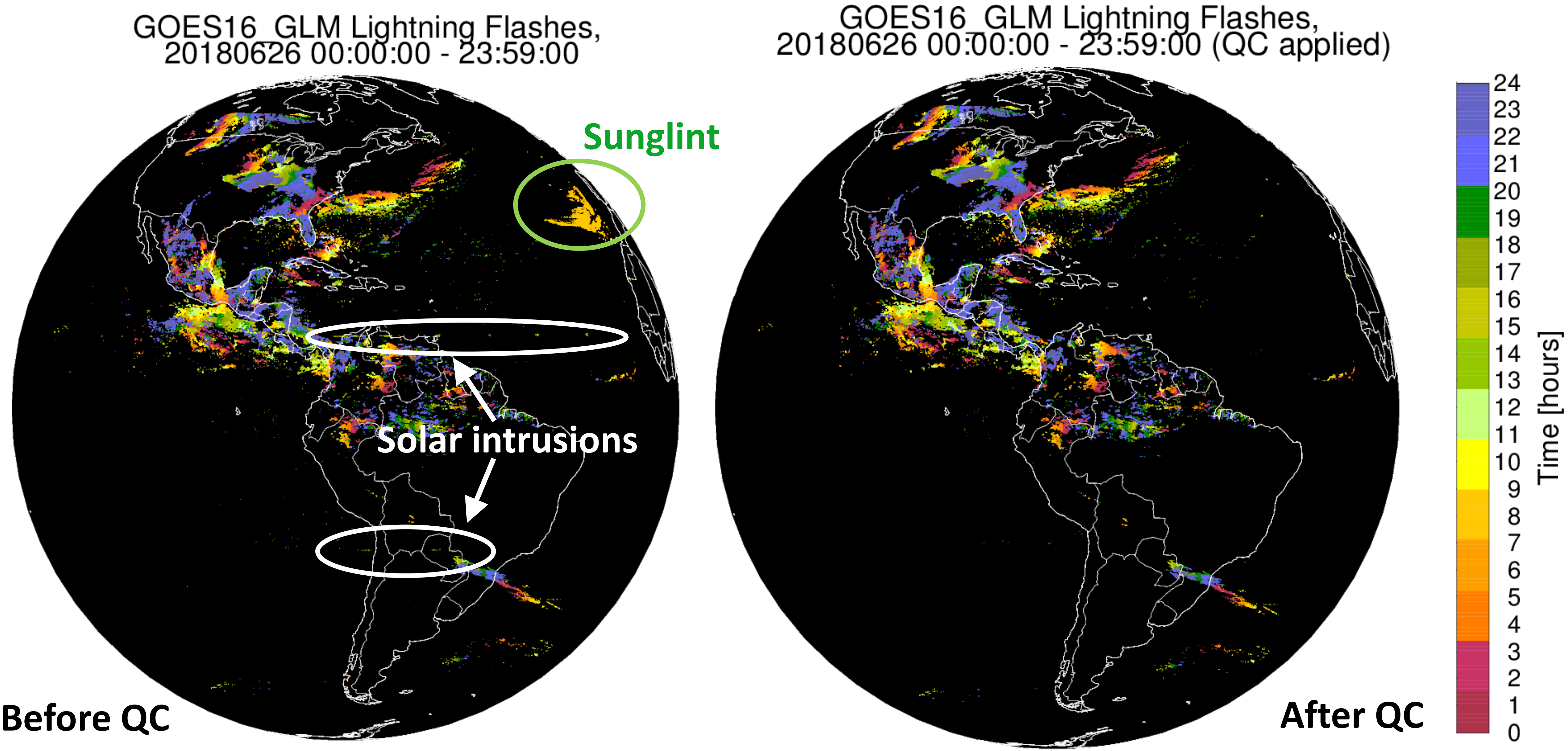
Sunglint

Solar intrusions

Before QC

After QC

GOES-16 GLM flash data: Quality Control (example 1)



4D-Var assimilation of GOES-16 GLM lightning flash densities: Quality control.

- Homemade quality control of the GLM flash product had to be developed:

Features to be removed	Screening method
Spurious flashes caused by sunglint	Remove all flashes inside sunglint region, throughout day
Persistent isolated lines of flashes (solar intrusion)	Convolution with line-identifying kernel
Flashes organized in short-lived regularly-spaced patterns (~ SSP noon; solar intrusion)	Convolution with comb-shaped function
Isolated flashes (e.g. due to detector noise, jitter)	Time and space criterion (±2 hr, ±80 km)

- Most technical developments needed to assimilate lightning obs have been made in the IFS (CY46R1):
  - include flash detection efficiency (75 to 88%, as a function of solar zenith angle);
  - averaging of obs over 6 hours and onto the model grid (outer loop);
  - obs quality control and screening;
  - new obs operator (incl. tangent-linear and adjoint);
  - logarithmic transform applied to flash density (more Gaussian distributions).

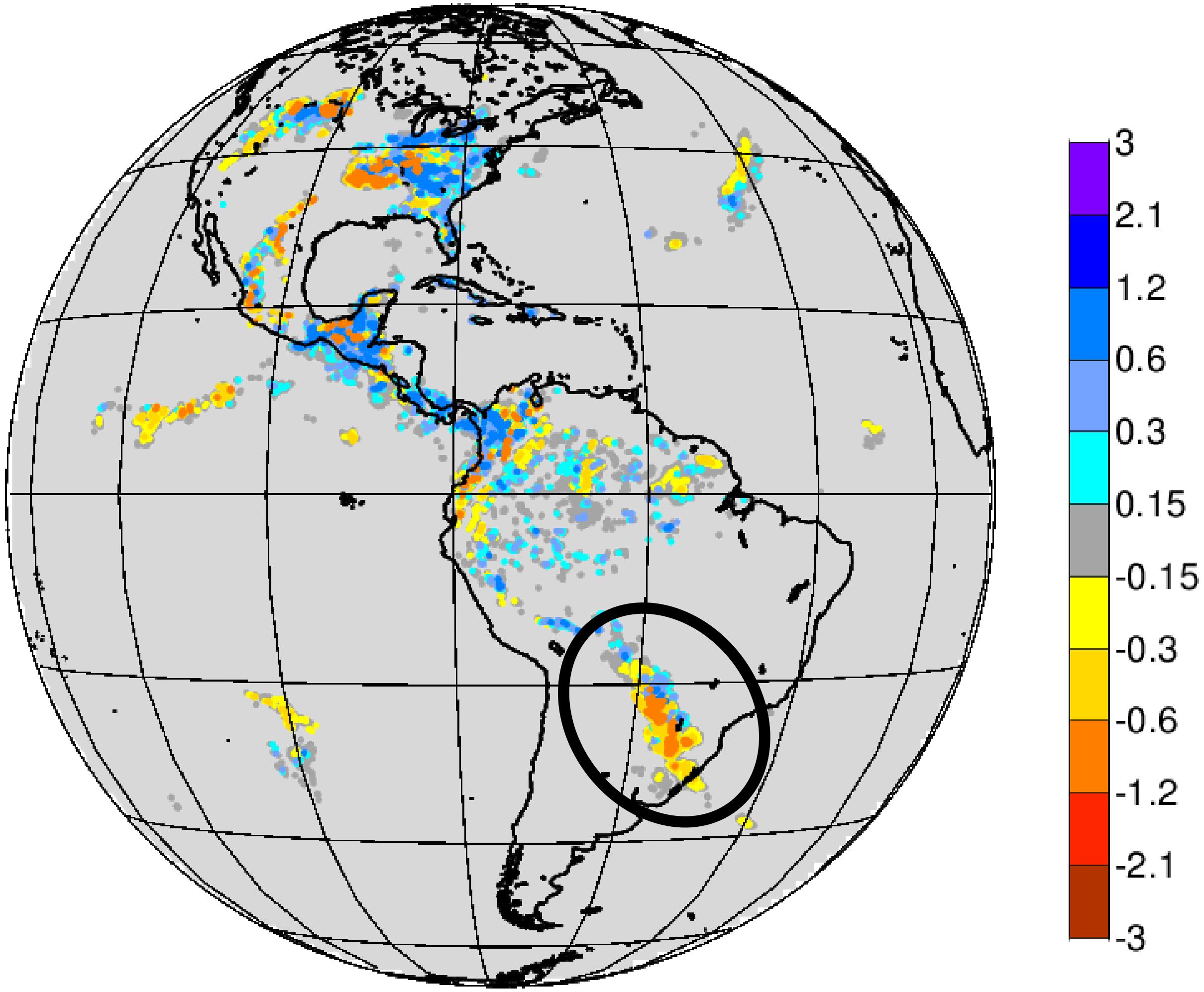
No bias correction used for the moment.

Lopez 2021, ECMWF Tech Memo 872

Assimilation of GOES-16 GLM lightning flash densities: Single 4D-Var cycle.

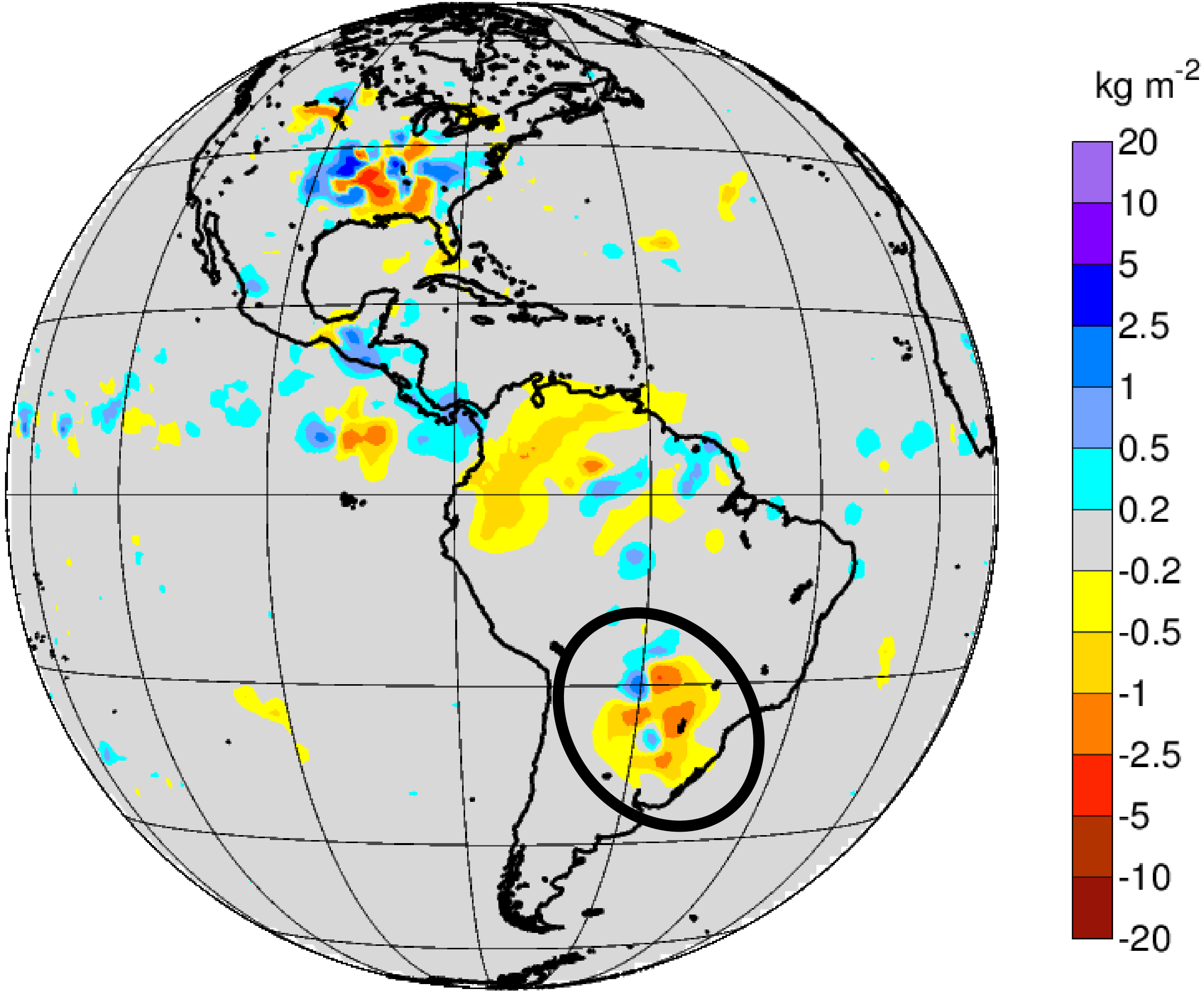
Single 4D-Var cycle (28-km resol., 137 lev.) using 6h-avg flash density on 1 Jun 2018 at 00Z.
All operational observations also assimilated.

Background lightning departures (before assim.)



TCWV analysis increments <u>due to lightning obs.</u>

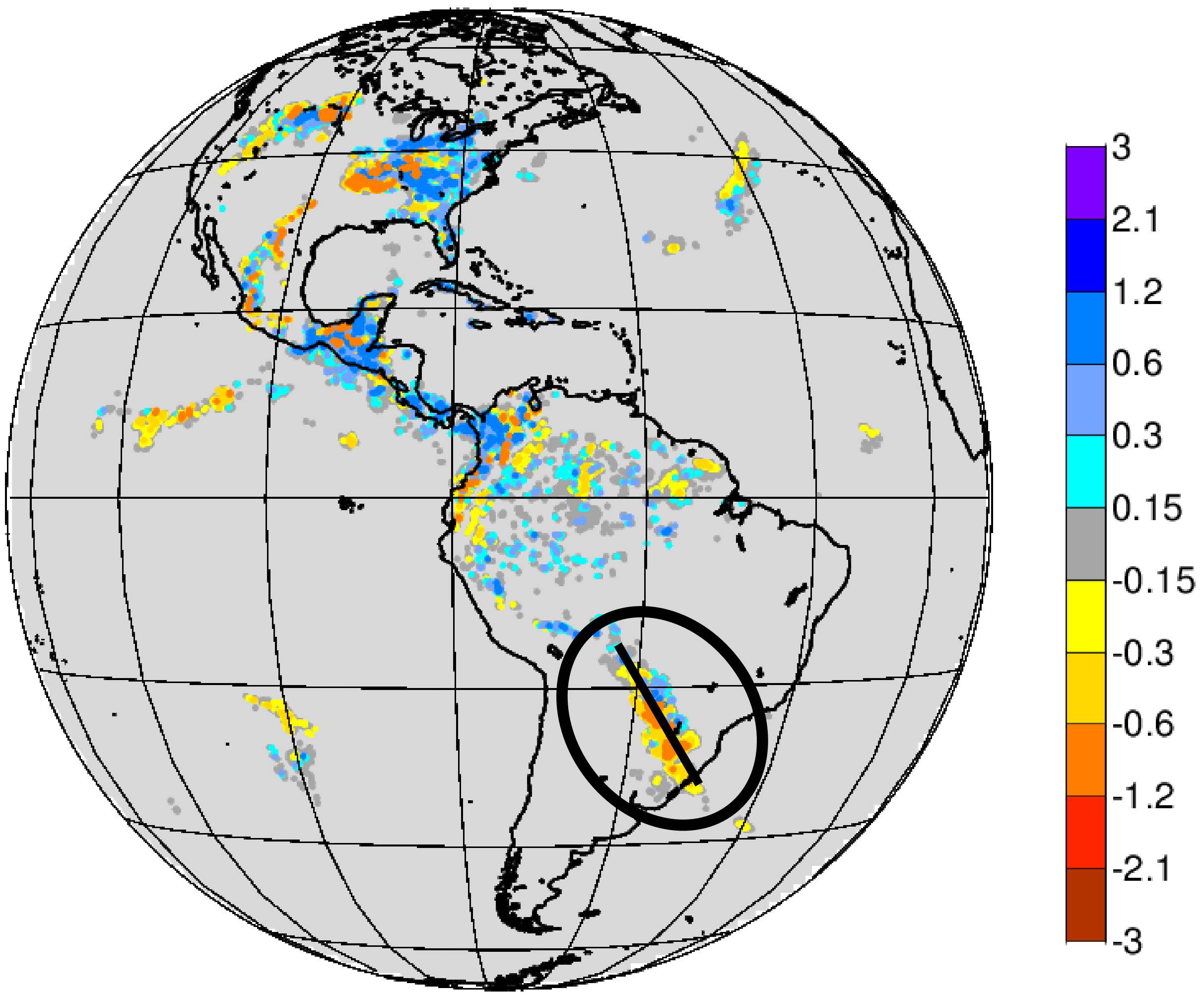
Mean = -0.03 kg m^{-2}



4D-Var assimilation of GOES-16 GLM lightning flash densities: First cycle.

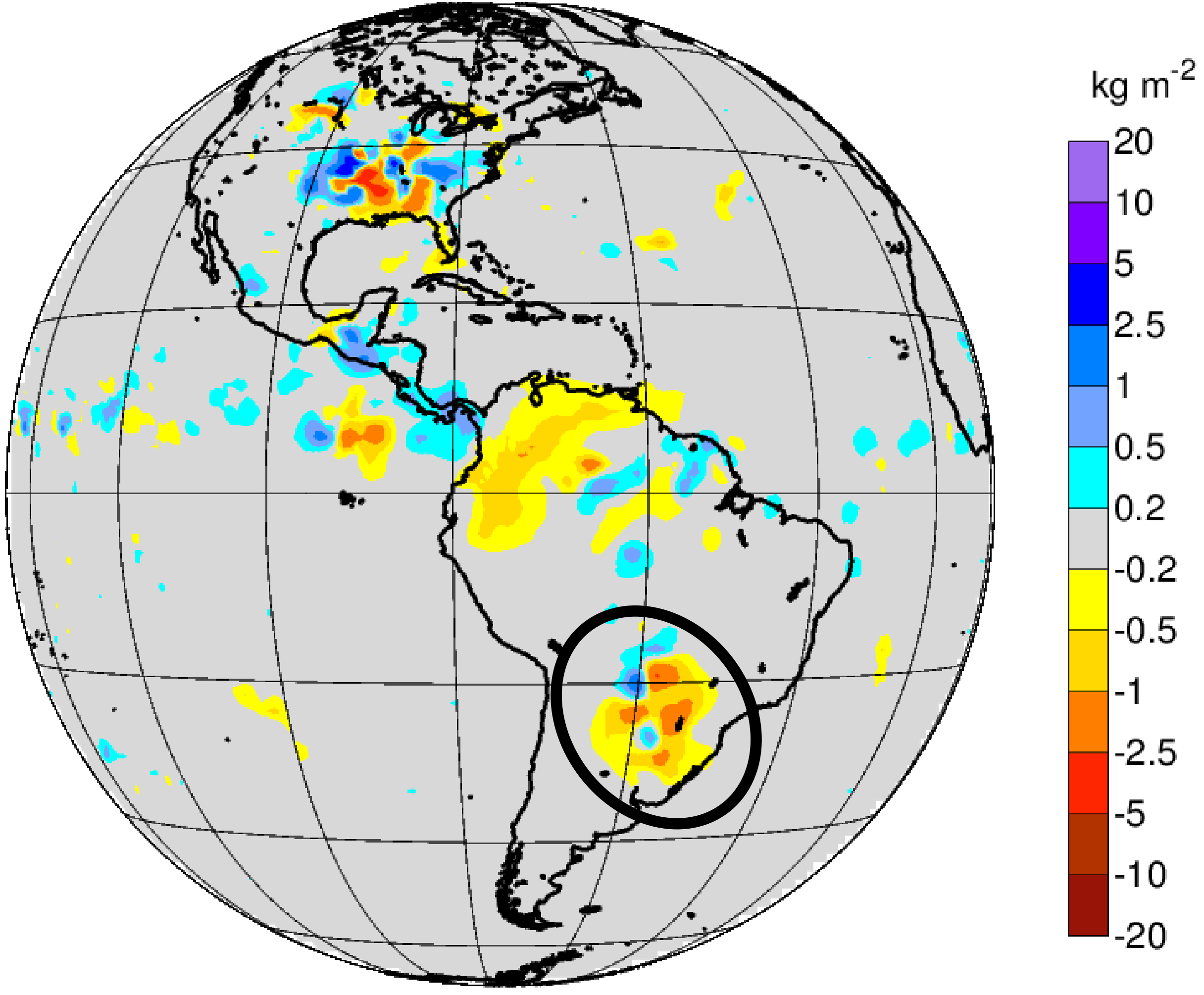
Single 4D-Var cycle (28-km resol., 137 lev.) using log$^{(2)}$[6h-avg flash density] (no bias corr.) on 1 Jun 2018 at 00Z. All operational observations also assimilated.

Background lightning departures (before assim.)



TCWV analysis increments due to lightning obs.

Mean = -0.03 kg m^{-2}



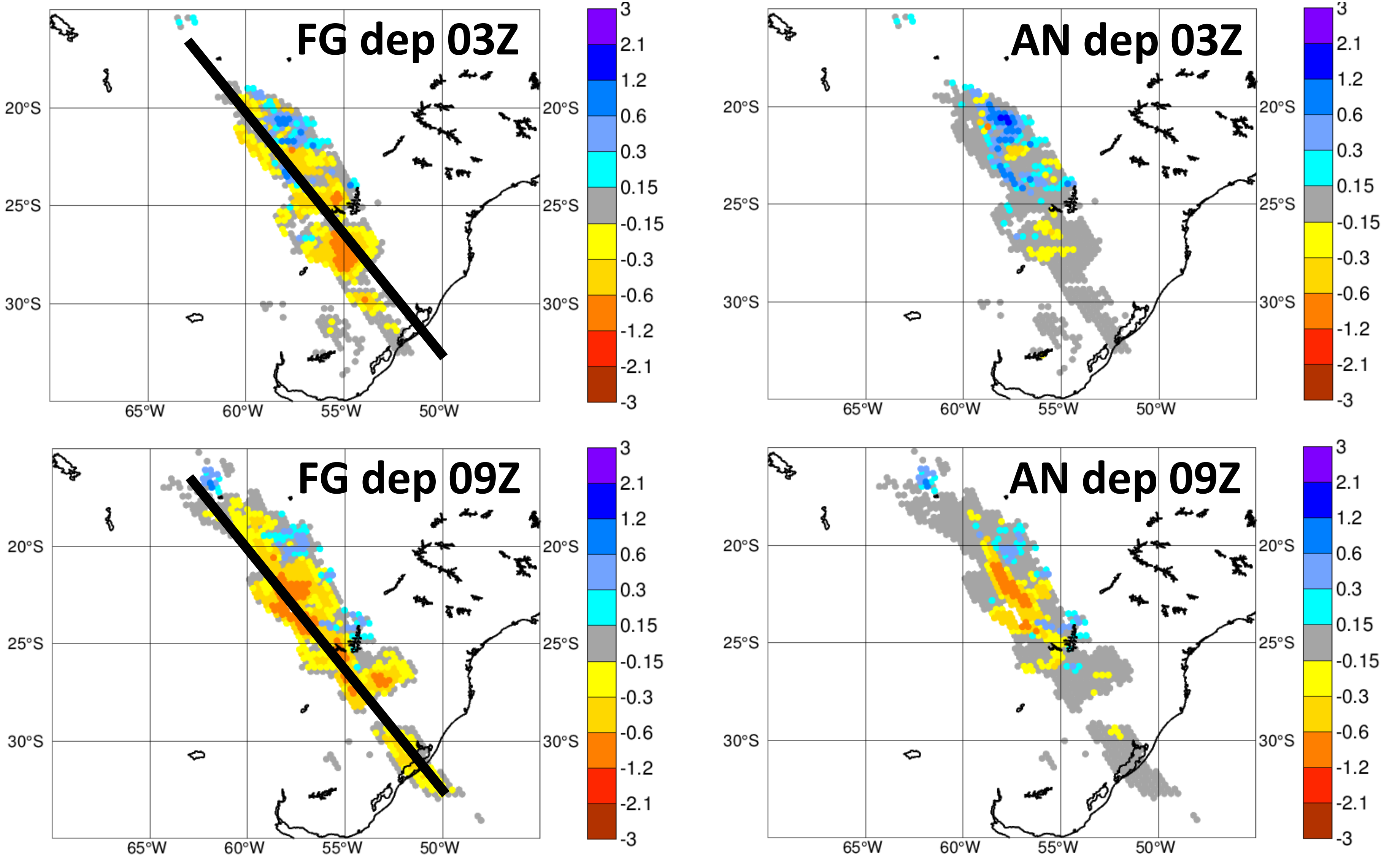
4D-Var assimilation of GOES-16 GLM lightning flash densities: First cycle.

Single 4D-Var cycle (28-km resol., 137 lev.) using log$^{(2)}$[6h-avg flash density] (no bias corr.).

Lightning obs–model departures

Zoom
South Brazil-Uruguay
on 1 June 2018

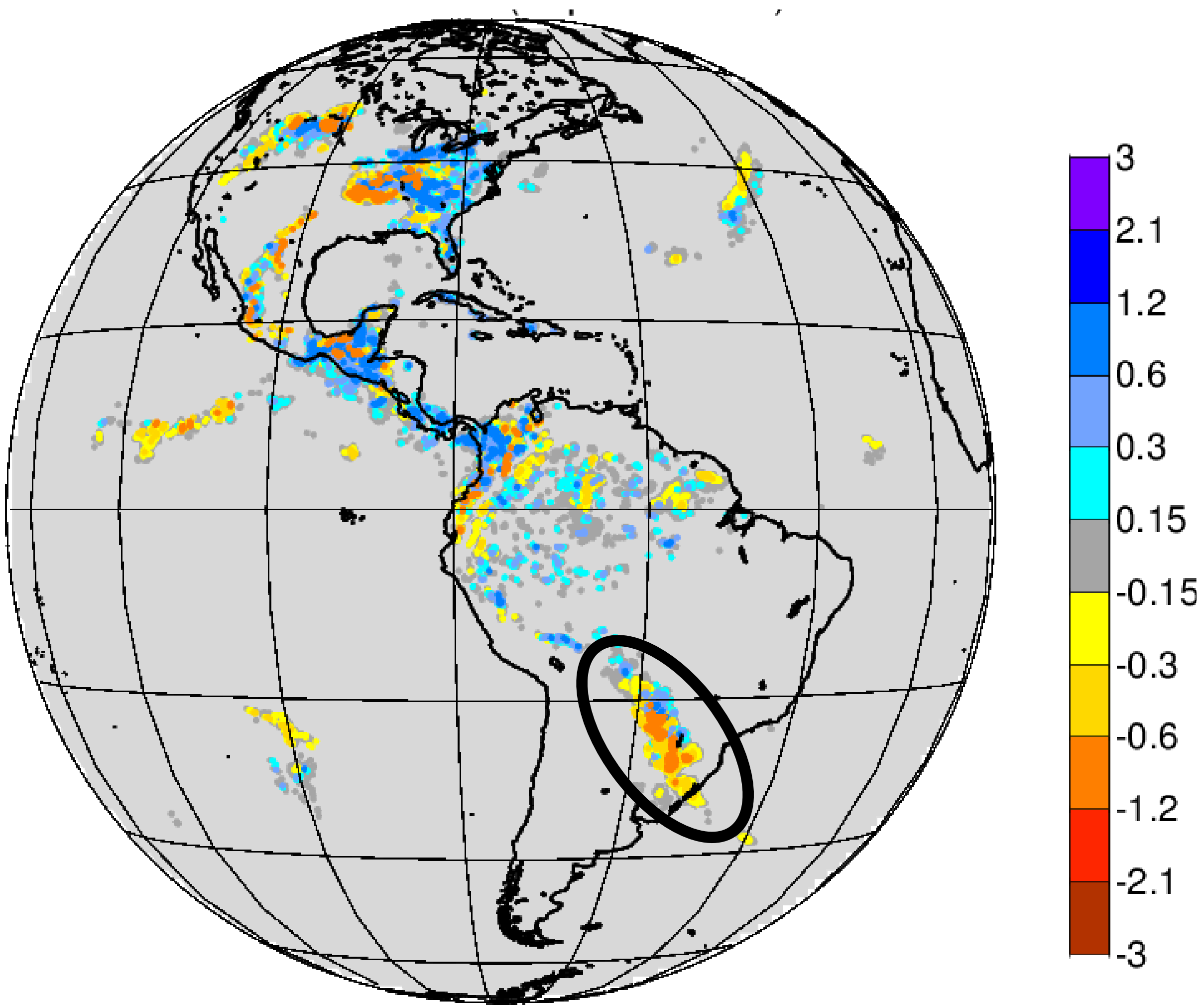
(2 time slots within
4D-Var 12h window)



GOES-16 GLM lightning flash density assimilation: First attempt.

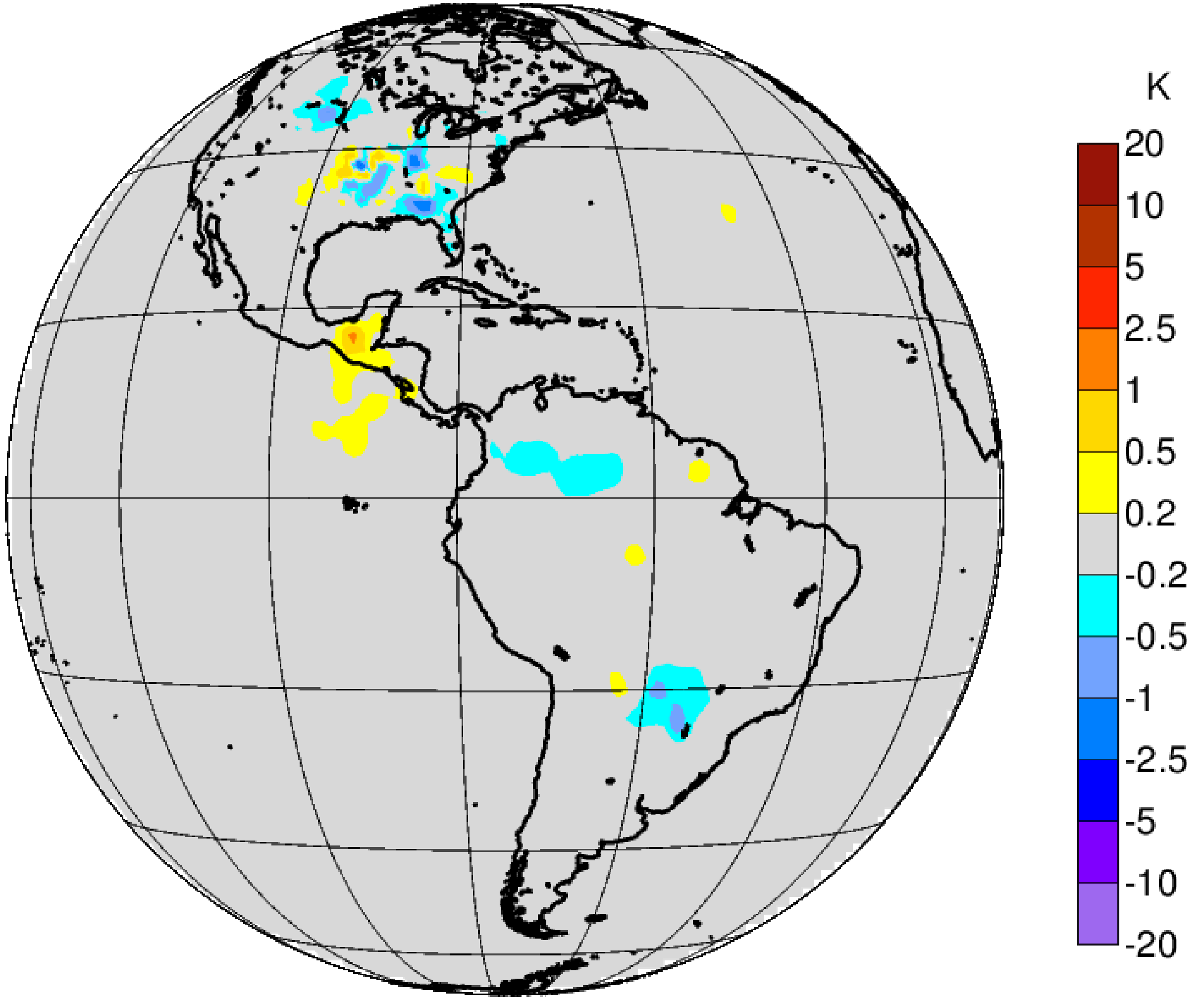
Single 4D-Var cycle (28-km resol., 137 lev.) using log$^{(2)}$[6h-avg flash density] (no bias corr.) on 1 Jun 2018 at 00Z. All operational observations also assimilated.

Background lightning departures

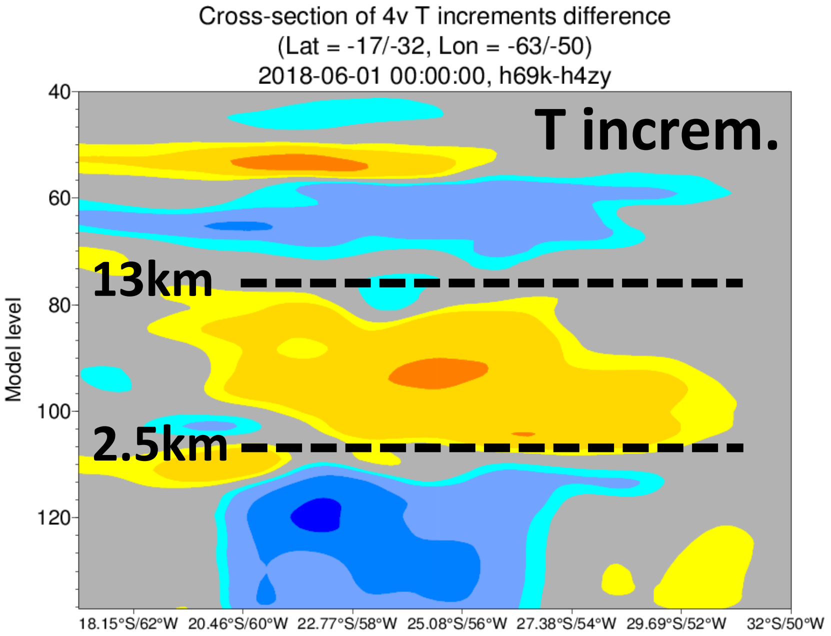


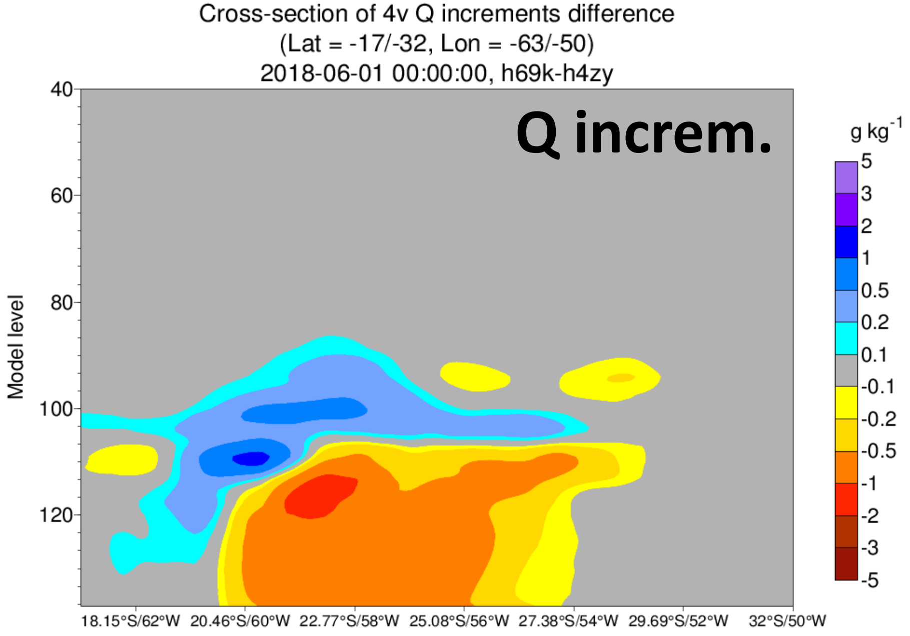
T analysis increments due to lightning obs.

Mean = 0 K

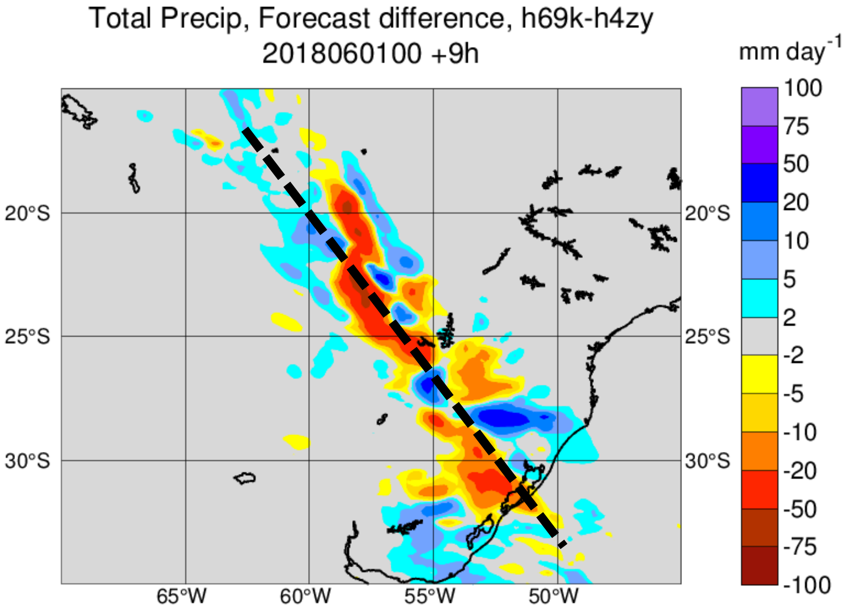


4D-Var assimilation of GOES-16 GLM lightning flash densities: First cycle.


Cross-section of 4v T increments difference
(Lat = -17/-32, Lon = -63/-50)
2018-06-01 00:00:00, h69k-h4zy


Cross-section of 4v Q increments difference
(Lat = -17/-32, Lon = -63/-50)
2018-06-01 00:00:00, h69k-h4zy

Cross-sections of
T & Q analysis increm.
due to lightning obs
(South Brazil-Uruguay).

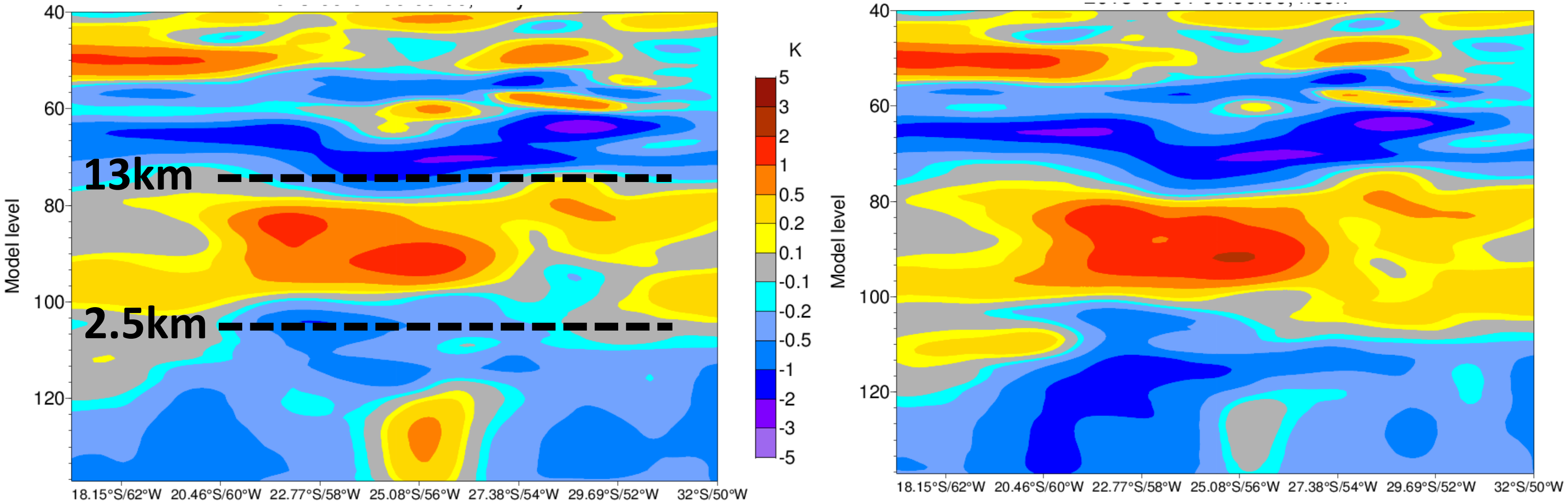

Total Precip, Forecast difference, h69k-h4zy
2018060100 +9h

← Impact on 9h total precipitation forecast
(South Brazil-Uruguay).

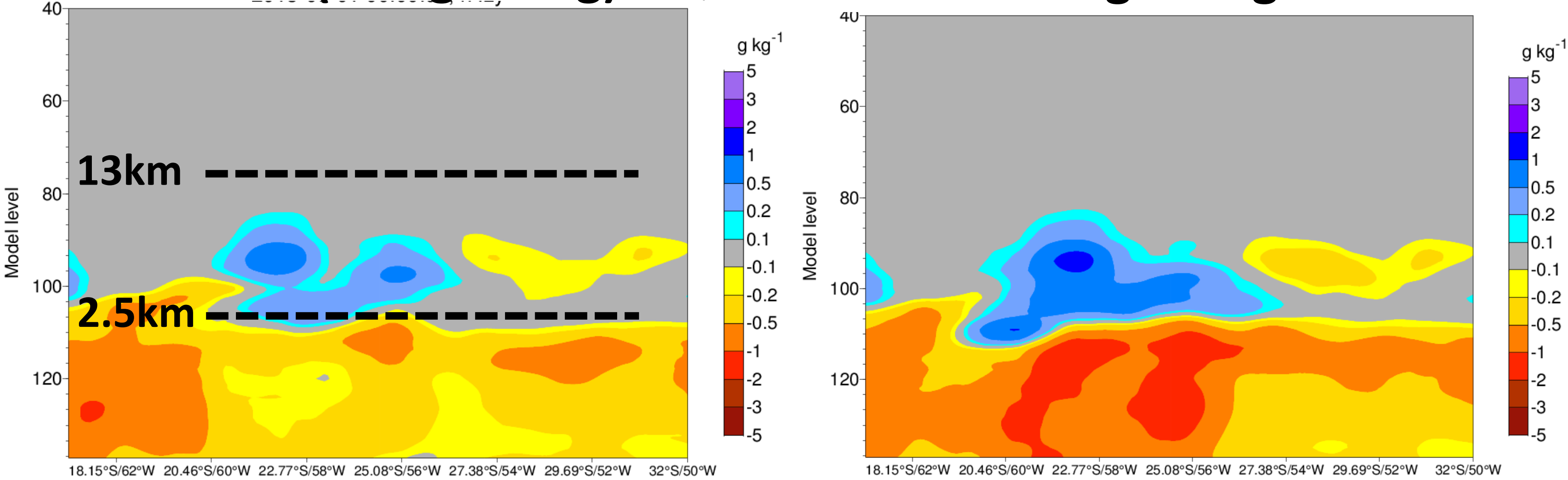
→ All these changes make sense to reduce lightning in the model.

4D-Var assimilation of GOES-16 GLM lightning flash densities: First cycle.

Control (no lightning) T increm. With lightning assim.



Control (no lightning) Q increm. With lightning assim.



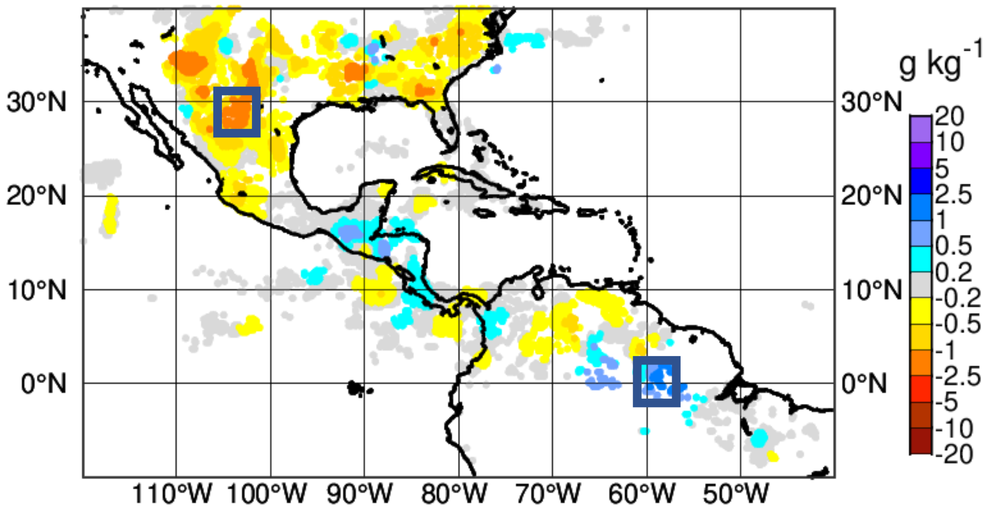
Cross-sections of
T and Q
analysis increments

→ Increments due to
lightning assimilation
are consistent with
or strengthen
those due to all other obs.

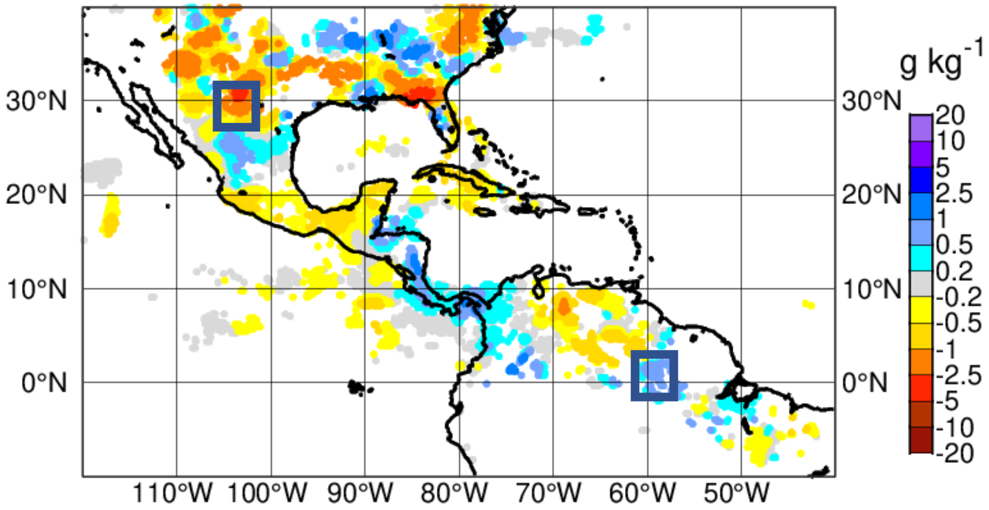
GOES-GLM-lightning only vs CTRL assimilation experiments: 4D-Var increments.

4D-Var humidity increments for assimilation cycle on 8 July 2019 at 00Z (28-km resol., 137 levels).

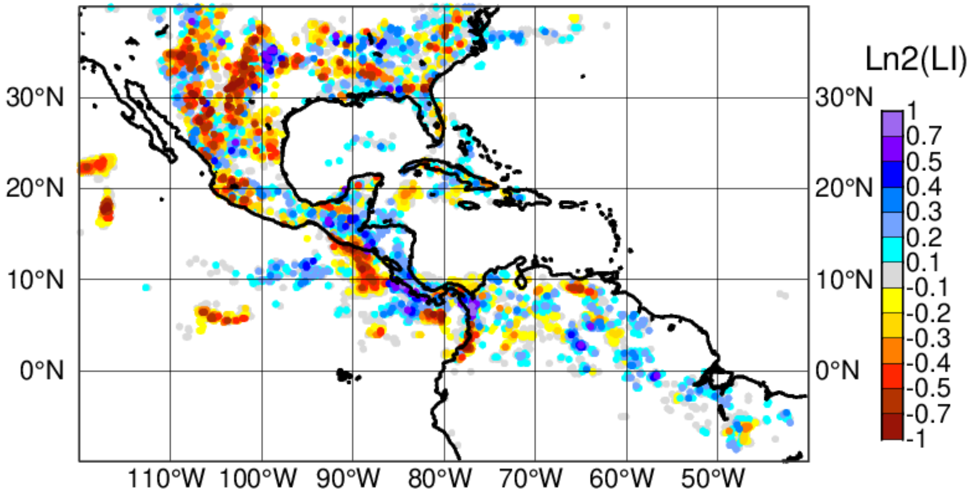
Q incr., Lev 137, GLM-only (hd5m)



Q incr., Lev 137, CTRL (hd5n)



Lightning FG depart. (hd5m)



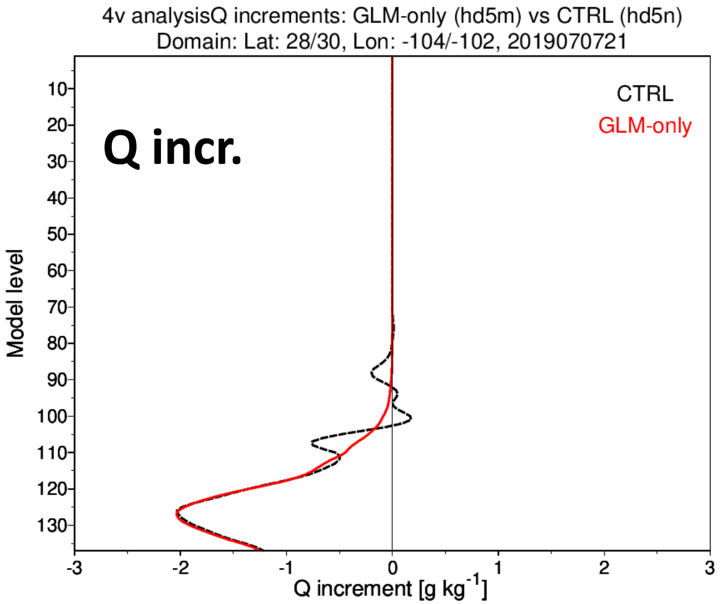
→ Humidity increments from GLM obs and from all other observations shown reasonable level of consistency in the lower troposphere where convective sensitivities are the strongest.

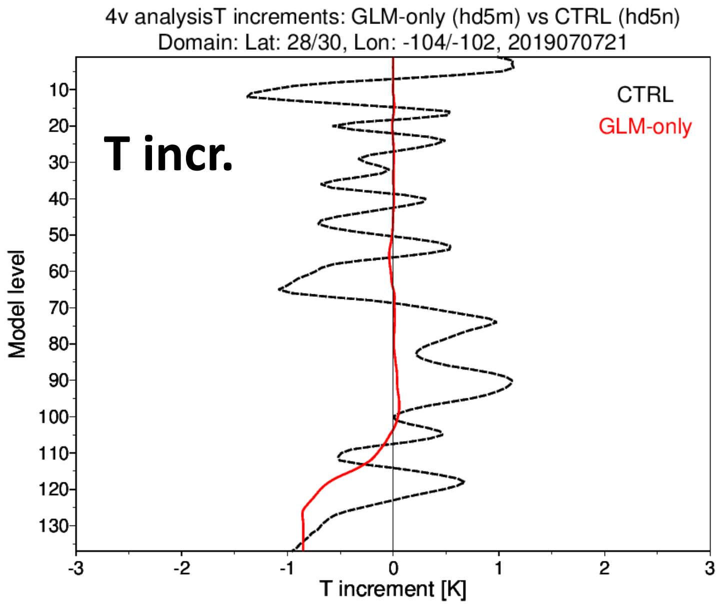
ECMWF

GOES-GLM-lightning only vs CTRL assimilation experiments: 4D-Var increments.

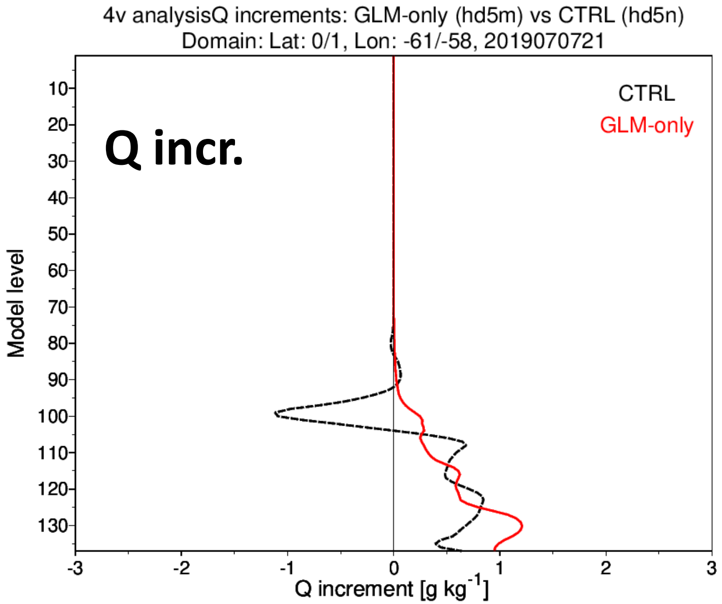
Vertical profiles of 4D-Var increments at two selected locations (with positive/negative departures):

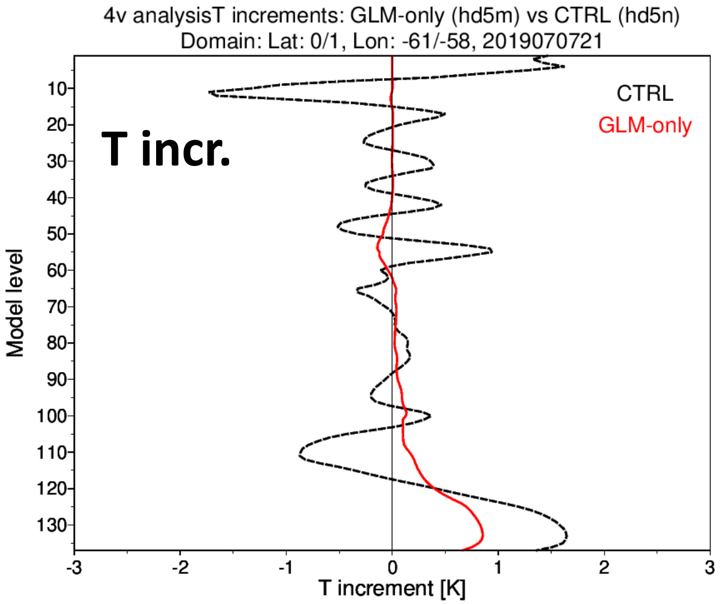
Mexico (model > GLM)


4v analysisQ increments: GLM-only (hd5m) vs CTRL (hd5n)
Domain: Lat: 28/30, Lon: -104/-102, 2019070721


4v analysisT increments: GLM-only (hd5m) vs CTRL (hd5n)
Domain: Lat: 28/30, Lon: -104/-102, 2019070721

Amazon (model < GLM)

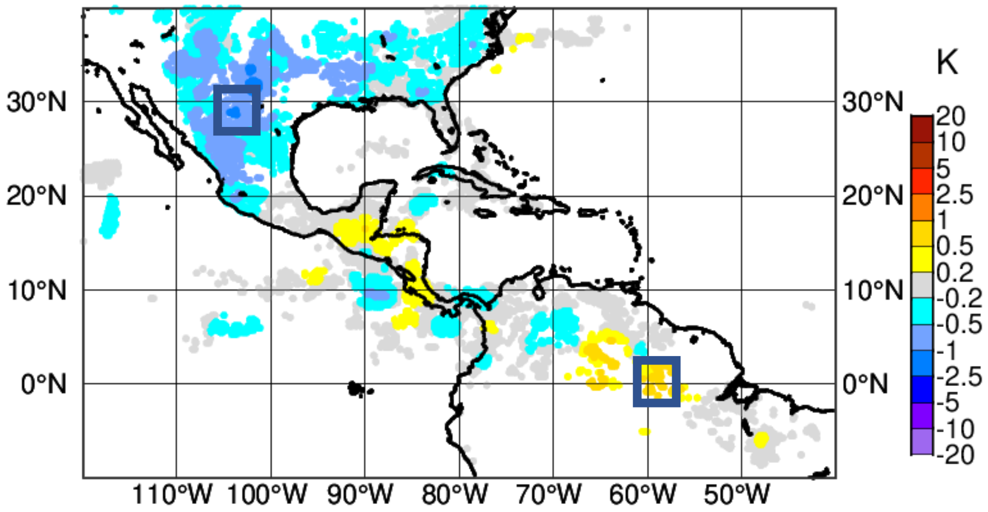

4v analysisQ increments: GLM-only (hd5m) vs CTRL (hd5n)
Domain: Lat: 0/1, Lon: -61/-58, 2019070721


4v analysisT increments: GLM-only (hd5m) vs CTRL (hd5n)
Domain: Lat: 0/1, Lon: -61/-58, 2019070721

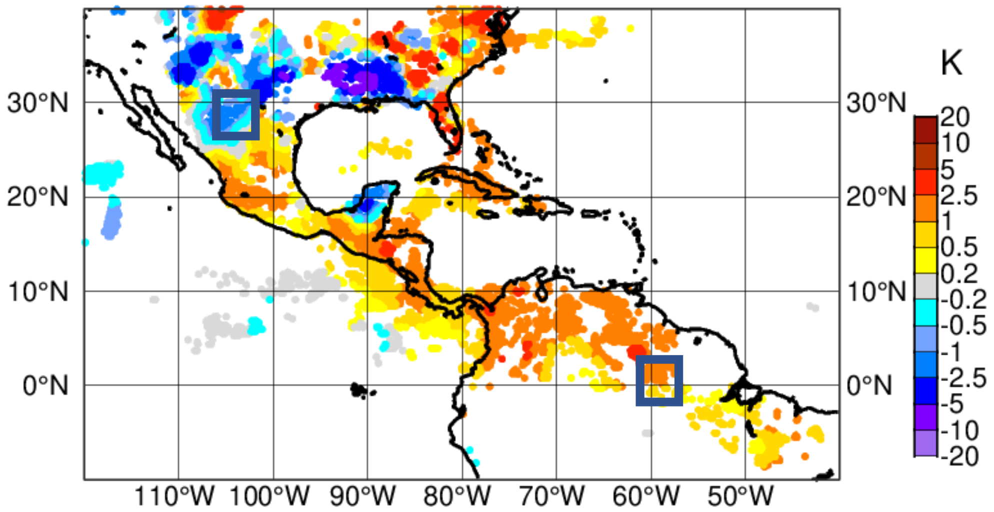
GOES-GLM-lightning only vs CTRL assimilation experiments: 4D-Var increments.

4D-Var temperature increments for assimilation cycle on 8 July 2019 at 00Z (28-km resol., 137 levels).

T incr., Lev 137, GLM-only (hd5m)



T incr., Lev 137, CTRL (hd5n)



→ Temperature increments from GLM obs and from all other observations seem much less consistent.

One possible reason for this: in CTRL, there is no constraint on how the increments are produced (i.e. via large-scale condensation or convection, Which have very different sensitivities!).

Lightning FG depart. (hd5m)