Assimilation of Meteosat Third Generation (MTG) Lightning Imager (LI) ~~observations~~ pseudo-observations in AROME-France – Proof of Concept

Felix Erdmann[1,2], Olivier Caumont[1,3,*], and Eric Defer[4,*]

[1]CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France
[2]Royal Meteorological Institute, 3 av Circulaire, 1180 Brussels, Belgium
[3]Météo-France, Direction des opérations pour la prévision, Toulouse, France
[4]Laboratoire d'Aérologie, 14 avenue Edouard Belin, 31400 Toulouse, France
[*]These authors contributed equally to this work.

Correspondence: Felix Erdmann (felix.erdmann@meteo.be)

Abstract. This study develops a Lightning Data Assimilation (LDA) scheme for the regional, convection-permitting NWP model AROME-France. The LDA scheme intends to assimilate total lightning, i.e., cloud-to-ground (CG) and inter- and intra-cloud (IC), of the future Meteosat Third Generation (MTG) Lightning Imager (LI). MTG-LI proxy data are created and Flash Extent Density (FED) fields are derived. An FED forward observation operator (FFO) is trained based on modeled, column integrated graupel mass from 24 storm days in 2018. The FFO is successfully verified for 2 independent storm days. With the FFO, the LDA adapts a 1-dimensional Bayesian (1DBay) retrieval followed by a 3-dimensional variational (3DVar) assimilation approach that is currently run operationally in AROME-France for radar reflectivity data. The 1DBay retrieval derives relative humidity profiles from the background by comparing the FED observations to the FED inferred from the background. Retrieved relative humidity profiles are assimilated as sounding data. The evaluation of the LDA comprises different LDA experiments and four case studies. It is found that all LDA experiments can increase the background integrated water vapor (IWV) in regions where the observed FED exceeds the FED inferred from AROME-France outputs. In addition, IWV can be reduced where spurious FED is modeled. A qualitative analysis of 6-hour accumulated rainfall fields reveals that the LDA is capable of locating and initiating some local precipitation fields better than a radar data assimilation (RDA) experiment. However, the LDA also leads to too high rainfall accumulations at some locations. Fractions Skill Scores (FSSs) of 6-hour accumulated rainfall are overall similar for the developed LDA and RDA experiments. An approach aiming at mitigating effects due to differences in the optical extents of lightning flashes and the area of the corresponding cloud was developed and included in the LDA, however, it does not always improve the FSS.

1 Introduction

Convective weather phenomena such as thunderstorms threaten the society by producing severe weather and related impacts, e.g., flash floods, large hail, tornadoes, and strong winds. Cloud electrification and subsequent lightning discharges are caused by interactions of different ice particles inside convective clouds. The process makes lightning an effective tracer of deep

convection. The new type of lightning locating systems (LLSs) on geostationary (GEO) satellites allows for continuous large-scale monitoring of total, i.e., cloud-to-ground (CG) and inter/intra-cloud (IC), lightning activity. The Geostationary Lightning Mappers (GLMs) on the GOES-R series satellites cover the Americas and the adjacent oceans (Goodman et al., 2013). A

25 similar instrument, the Meteosat Third Generation (MTG) Lightning Imager (LI), will be launched in the 2022 time-frame to monitor lightning over among others Europe, Africa, and wide parts of the Atlantic Ocean (Dobber and Grandell, 2014). ~~GEO lightning sensors provide continuous high-frequency observations of the total lightning activity over a wide domain.~~ High-quality GEO data is also available in data-sparse regions with limited access, such as over oceans and mountainous terrain as well as countries without radar networks. GEO lightning data may become important for numerical weather prediction (NWP),

30 potentially improving the initial state of the model and the accuracy of the predicted storm location, timing, and intensity. ~~The lightning~~ Lightning data assimilation (LDA) addresses this objective.

However, LDA is not trivial. Operational models, e.g., the Application of Research to Operations at Mesoscale (AROME) model (Brousseau et al., 2016), the regional operational model of Météo-France, do not include any explicit representation of lightning. Other difficulties of lightning data assimilation include the ephemeral nature of lightning relative to the life cycle

35 of ~~thunderstorm~~ thunderstorms and time-space shifts between observations and the background, which are usual at the storm scale, challenge up-to-date data assimilation systems (zero-spread, zero-gradient problems) (e.g., Janisková and Lopez, 2013). Background errors are not Gaussian, and lightning is more related to lower-impact variables such as hydrometeor content and vertical velocity than to thermodynamic variables. Despite ~~the difficulties, former~~ those difficulties, previous studies have shown the benefits of LDA for forecasts of convection and related phenomena.

40 Early LDA studies used NWP models that parametrize convection. Lightning data controlled the activation of the convective parametrization scheme (CPS). Often nudging techniques were applied to modify the model humidity and wind fields (e.g., Papadopoulos et al., 2005; Mansell et al., 2007; Lagouvardos et al., 2013; Giannaros et al., 2016), or latent heat (e.g., Pessi and Businger, 2009). These studies reported positive effects of LDA on the score of precipitation forecasts, especially for the location and quantity of heavy rain.

45 When NWP models resolve convection (<3 km horizontal resolution), so-called lightning proxies can be used to relate NWP output to lightning observations. Some ~~former~~ previous studies suggested and tested cloud top height, cloud top pressure, and cold cloud depth as predictors to estimate the lightning density ~~(e.g., Price and Rind, 1992, 1993; Wong et al., 2013; Giannaros et al., 2015;~~ (e.g., Price and Rind, 1992, 1993; Allen and Pickering, 2002; Wong et al., 2013; Giannaros et al., 2015; Karagiannidis et al., 2019).

50 Recent development corroborates using cloud dynamics or ice hydrometeors, e.g., graupel, for reliable lightning proxies. Indeed, field campaigns as those reported by Deierling et al. (2008) found a strong correlation between lightning activity and ice mass fluxes. Deierling and Petersen (2008) showed a robust relationship between total lightning rate and updraft characteristics, especially the updraft volume. Buiat et al. (2017) analyzed CloudSat Cloud Profiling Radar (CPR) to relate cloud ice water content (IWC) and effective radius (ER) to LINET (Betz et al., 2009) lightning strokes. They found that high

55 IWC and high ER are favorable to CG strokes. Graupel contributes to both high IWC and high ER. Strong on-going updraughts can cause high IWC. Hence, the findings agree with the studies of Deierling et al. (2008); Deierling and Petersen (2008).

Various lightning proxies have been tested in NWP studies. McCaul et al. (2009) ~~brought forward two lightning proxies based on~~ related domain-wide peaks of proxies based ice-phase hydrometeor fields, one related to upward fluxes and the second related to vertically integrated ~~content, in the WRF model. Relating domain-wide statistics of peaks in the proxy fields~~ ice content, to peaks of the flash rate density, and suggested a blended solution using both proxies~~is suggested to reproduce both temporal variability and areal coverage of the lightning observations~~. Yair et al. (2010) introduced a lightning potential index (LPI) that is calculated using the simulated grid-scale vertical velocity and simulated hydrometeor mass mixing ratios of liquid water, cloud ice, snow, and graupel. Barthe et al. (2010) concluded that none of precipitation ice mass, ice water path, ice mass flux product, updraft volume, maximum vertical velocity, and cloud top height could predict the lightning flash rates and trends well for both of their two cases. ~~Formenton et al. (2013) confirms~~ However, more recent studies of Formenton et al. (2013); Federico et al. (2014); Bovalo et al. (2019) confirm a key role of graupel in the cloud electrification ~~, as described in the non-inductive charging mechanism, in their 1D model. Federico et al. (2014) simulate lightning based on graupel carrying negative charge and ice carry positive charge. Bovalo et al. (2019) recommended graupel mass as most reliable proxy for lightning location and peak flash rate after analyzing eight test cases. Their cell scale approach resulted in more robust relationships than a domain scale approach~~and use ice and /or graupel content for lightning simulation and proxy definition.

With the lightning proxies, lightning observation operators are created to compare the model output to the lightning observations and assimilate the lightning data. Most of the recent LDA schemes use lightning operators based on ice hydrometeors or updraft characteristics, e.g., graupel mass or updraft volume. In the following, a brief overview of LDA techniques is presented, with the focus on 3-dimensional variational (3DVar) LDA and studies to assimilate GLM data.

Fierro et al. (2012) put forward a widely used (e.g., Dixon et al., 2016; Wang et al., 2017a; Federico et al., 2017; Wang et al., 2017b) concept of lightning proxy calculating water vapor mixing ratio from simulated graupel mass and observed flash rates. They then increased the water vapor mixing ratio in the $0\,^\circ\mathrm{C}$ to $-20\,^\circ\mathrm{C}$ layer where lightning was observed and the relative humidity (RH) of the background was less than $81\,\%$. Precipitating convection was better correlated with observed reflectivity fields for the LDA than for the control experiment. The method was further tested by Fierro et al. (2014) using 10-minute Flash Extent Density (FED). Lynn et al. (2015); Marchand and Fuelberg (2014) refined the nudging technique of Fierro et al. (2012) to increase the virtual temperature perturbation and favor static instability or by warming the source layer of the convective updraft, respectively. Lynn et al. (2015) also introduced an extension to suppress spurious convection. ~~Qie et al. (2014) constructed empirical formulas between total lightning and graupel, ice, and snow mixing ratios based on thunderstorms over northern China. Lightning data were used by Wang et al. (2014) to generate~~ Other nudging-based LDA make use of ice-phase particle mixing ratios (e.g., Qie et al., 2014; Wang et al., 2018) and pseudo radar reflectivity ~~which was subsequently assimilated using the physical initialization method to adjust model variables. Wang et al. (2018) used time-lagged ensemble forecasts to retrieve q_g from lightning rates~~(Wang et al., 2014).

Mansell (2014); Allen et al. (2016) assimilated synthetic GLM total lightning represented as FED using an Ensemble Kalman Filter (EnKF)~~with positive effects on predicted spurious convection, updraft locations, and reflectivity~~. Mansell (2014) used both the simulated flash rate and a linear relationship between total lightning and graupel volume as observation operators.

Allen et al. (2016) recommend a linear best fit operator based on graupel mass and graupel volume. In preparation of a hybrid variational-ensemble LDA technique for GLM, Apodaca et al. (2014) assimilated World Wide Lightning Location Network (WWLLN) data in WRF. However, the use of WWLLN as a GLM proxy is debatable as WWLLN mainly detects CG flashes, whereas GLM detects total lightning.

A 3DVar LDA was applied by Wang et al. (2017b) using the relationship given by Fierro et al. (2012) and assimilated pseudo RH profiles in the form of sounding data. They noted increased forecasts skills in general, but also that the method still needs improvement to suppress spurious convection. Fierro et al. (2016) replaced their previous nudging by a 3DVar LDA with 10-minute FED. They found that the Radar Data Assimilation (RDA) yielded better forecasts of convective cells during the first $30\,\mathrm{min}$ of the forecast, while the LDA gave better storm structures one hour into the forecast. The combination of both RDA and LDA provided the highest forecast skill. Fierro et al. (2019) tested the 3DVar LDA technique developed for ground-based data with GLM total lightning observations. It adjusts water vapor mass mixing ratio (q_v) in regions of lightning by setting RH to $95\,\%$ in a layer of $3\,\mathrm{km}$ above the LCL if the background RH is less than $95\,\%$. Both LDA and RDA improved the short-term accumulated precipitation and radar reflectivity composite. Hu et al. (2020) adopted the technique of Fierro and found, as also previously reported, a wet bias in the model that increased with the forecast time. The method still misses a suppression of spurious convection in regions without lightning. They also conducted a layer depth sensitivity study with similar results for adding q_v in layers of $2\,\mathrm{km}$ to $10\,\mathrm{km}$ depth. Kong et al. (2020) present an LDA of real GLM data in an EnKF framework. FED at $10\,\mathrm{km}$ pixel resolution is assimilated using both graupel mass and graupel volume-based observation operators, with positive results on the convection forecasts. LDA of Chinese satellite Fengyun-4 (FY4) lightning data was realized by Liu et al. (2020) through creating pseudo-RH profiles and by Chen et al. (2020) through retrieving maximum proxy-reflectivity and finally pseudo-reflectivity profiles from the lightning data.

The main goal of this work is to develop a 3DVar LDA for MTG-LI data to improve analyses and forecasts, especially in convective situations. Pseudo-observations of the MTG-LI are generated to form the lightning data base (Section 3). A novel
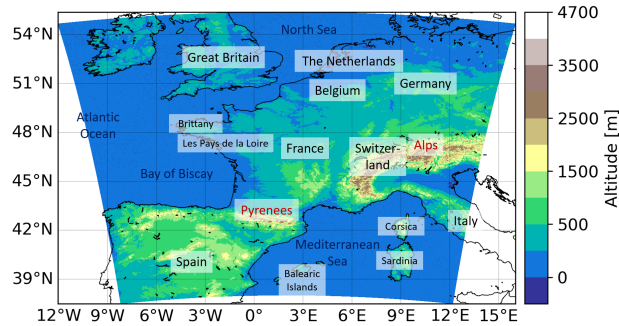
Figure 1. AROME-France physical domain and model topography. The AROME-France horizontal grid is equidistant at $1.3\,\mathrm{km}$ resolution.

LDA scheme is developed for the regional, operational, convective-scale model AROME-France of Météo-France. ~~Different assimilation approaches including the presence and absence of radar reflectivity , Doppler wind velocity, and FED in the set of assimilated observations~~ Various approaches in which FED is assimilated together with or without radar reflectivity and Doppler wind are compared to assess the added value of the developed LDA. Recent Var LDAs using GLM data, e.g., Fierro et al. (2019); Hu et al. (2020) cannot suppress spurious convection. Our LDA of GEO lightning data should promote convection if needed and also suppress spurious convection.

The NWP model configuration and lightning data as used for this work are briefly explained in Sections 2 and 3, respectively. Section 4 ~~reviews lightning observation operators and~~ introduces the lightning observation operator developed for this study. Then, our LDA method and the model experiments are explained in Sections 5 and 6. Section 7 describe the AROME-France analysis resulting from LDA. One of four case studies is detailed using different assimilation experiments, ~~among others~~ including the new LDA. Finally, Section 9 concludes this work.

2 Model configuration

AROME-France is the convective-scale, limited-area model run operationally by Météo-France since 2008 (Seity et al., 2011). It provides 36 to 42-hour forecasts five times a day (00 UTC, 03 UTC, 06 UTC, 12 UTC, 18 UTC). After an update in 2015, the model grid comprises 1440×1536 grid points in the horizontal with uniform $1.3\,\mathrm{km}$ horizontal resolution. The physical model domain and model topography are shown in Figure 1. In the vertical, the lowest model level is situated at around $5\,\mathrm{m}$ above ground. Each column ~~reaches up to the highest level at~~ extends up to pressure level $10\,\mathrm{hPa}$. The vertical resolution is refined homogeneously from top to bottom by a factor of 1.5. In total, 90 vertical levels (33 levels below $2000\,\mathrm{m}$) are computed. Model time steps equal $50\,\mathrm{s}$. Model dynamics are non-hydrostatic, semi-implicit, and semi-Lagrangian. Lateral boundary conditions (LBCs) are extracted from the global model ARPEGE (Bouyssel et al., 2022). AROME-France features a Davies relaxation (Davies, 1976) coupling and ARPEGE synchronization. The initial conditions rely on a 3DVar data assimilation technique (~~section~~ Section 5).

Deep convection is expected to be mostly resolved on the model grid (Fischer et al., 2018). Parametrization of sub-grid
scale shallow convection is based on Pergaud et al. (2009). AROME-France uses a mixed-phase microphysical scheme with
riming processes and graupel (Seity et al., 2011). In particular, the microphysics scheme of AROME-France separates five
prognostic hydrometeor variables, that are specific contents of precipitating species rain (q_r), snow (q_s), and graupel (q_g) and
the two non-precipitating species ice crystals (q_i) and cloud droplets (q_c). In addition, the water vapor specific content q_v (also
termed specific humidity) is computed. Hail is assumed to behave as large graupel particles. Overall, more than 25 processes
are parametrized into the microphysics scheme (Lascaux et al., 2006). AROME-France physics include a 1-dimensional (1D)
turbulence parametrization as combination of a prognostic turbulent kinetic energy (TKE) equation with a diagnostic mixing
length. An Externalized Surface (SURFEX) scheme and the European Center for Medium-Range Weather Forecasts (ECMWF)
radiation parametrization are other components of AROME-France model physics. Details can be found in Seity et al. (2011);
Brousseau et al. (2016).

3   Lightning data

This work adapts the GEO lightning pseudo-observation generator as developed by Erdmann et al. (2022). It was trained us-
ing low frequency (LF) ground-based National Lightning Detection Network (NLDN) records collected over the South-East
US. The LF ground-based lightning observations in this study are provided by the French network Meteorage (Schulz et al.,
2016; Pédeboy, 2015) as input. Meteorage locates total lightning with a ~~separation of~~ discrimination between CG strokes and
IC pulses. They are clustered to the flash level data using the same method as Erdmann et al. (2020), i.e., a spatiotemporal
clustering with criteria of $20\,\mathrm{km}$ and $0.4\,\mathrm{s}$. To verify Meteorage data as suitable input to the generator, Meteorage and NLDN
observations were compared to space-borne ISS-LIS data as common reference (Erdmann, 2020): The flash detection effiency
(DE) of NLDN relative to ISS-LIS was about $76\,\%$. ISS-LIS detected about $59\,\%$ of the NLDN flashes. In France, Meteor-
age flash DE relative to ISS-LIS was about $83\,\%$ while ISS-LIS detected about $57\,\%$ of the Meteorage flashes. It should be
mentioned that the low earth orbit, optical ISS-LIS observes lightning in a different way than the ground-based LF networks,
with different detection efficiencies for CG and IC flashes, respectively (e.g., Erdmann et al., 2020; Zhang et al., 2019). The
statistical evaluation of general flash characteristics revealed that NLDN and Meteorage flashes feature similar distributions of
flash extent (average of about $7.0\,\mathrm{km}$), flash duration (average of about $0.2\,\mathrm{s}$), stroke/pulse number per flash (average of about
3.6). Meteorage, with shorter baseline distance than NLDN, recorded on average higher LF currents for the occurring flashes
(average of $9.5\,\mathrm{kA}$ and $8.3\,\mathrm{kA}$ for Meteorage and NLDN, respectively). In all cases, flashes with large extent, long duration, ~~as
well as flashes with~~ or with a high number of ~~events and~~ ISS-LIS events or LF strokes/pulses ~~for ISS-LIS and the LF network,
respectively,~~ were more likely detected by ~~the other LLS~~ both LLSs than the small, short duration flashes. Overall, Meteorage
was validated as a suitable ~~source for the~~ input to the GEO lightning pseudo-observation generator. Pseudo-observations of
MTG-LI are then generated on a regular latitude-longitude grid with average pixel resolution of $7\,\mathrm{km}$, which approximates
the expected MTG-LI resolution over France (personal communication Bartolomeo Viticchiè, 2020). FEDs are calculated per
5-minute interval that can then be summed as needed for the assimilation. This work uses a 10-minute interval of FED data in
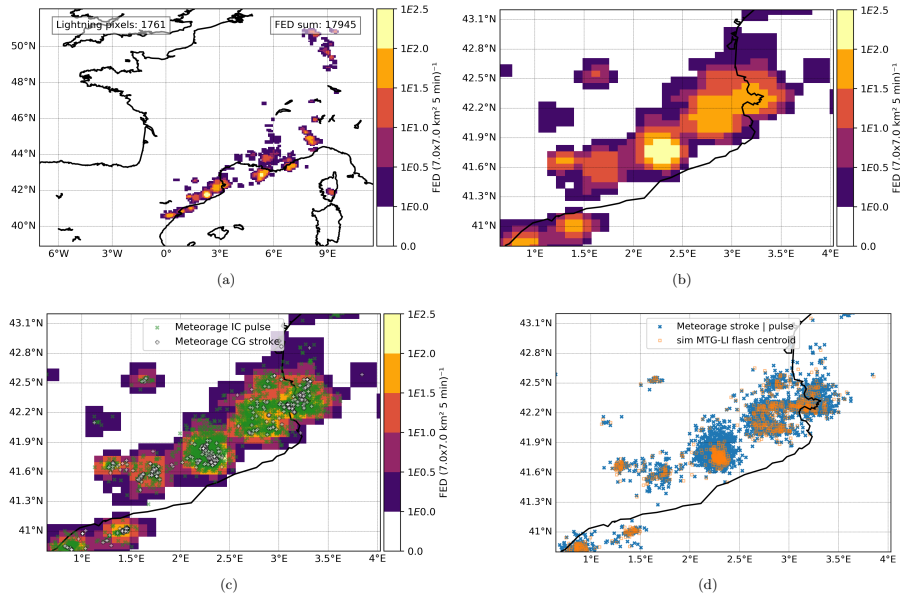
Figure 2. Simulated MTG-LI FED in the entire domain (a) and zoomed on the region of the maximum FED value (b) on $7\,\mathrm{km} \times 7\,\mathrm{km}$ grid. The Meteorage strokes and pulses that were used to generate the MTG-LI flashes are superimposed on the FED in (c). Simulated MTG-LI flash centroids and the corresponding Meteorage strokes+pulses in (d). Example for the period from 09 Aug. 2018, 13:55 UTC to 14:00 UTC.

the LDA that is centered at the time of the analysis (as done by Fierro et al., 2016; Hu et al., 2020). The short period around the analysis reduces displacement errors in the analysis. The domain is limited to $40°$N to $51°$N and $5.5°$W to $10°$E, which is inside the AROME-France physical model domain. The pseudo MTG-LI FED is referred to as FED observation hereafter to avoid confusion with the pseudo-observations (POs) created by the 1-dimensional Bayesian (1DBay) retrieval.

An example of simulated MTG-LI is provided in Figure 2(a) for the entire study domain, and Figure 2(b) zoomed on the FED maximum on 09 Aug. 2018, 13:55 UTC to 14:00 UTC. The regular latitude-longitude grid of the FED is obvious ~~. Figure 2 accumulates data of 5 min, and~~ in Figure 2(b). Throughout the domain, the FED pixels have a ~~constant size of about 0.0630° latitude and 0.0896° longitude that corresponds to an average pixel size of~~ size of roughly 7 km times 7 km~~in the study domain~~.

Figure 2(c) ~~shows~~ superimposes the zoomed-in simulated FED and ~~adds~~ the Meteorage CG strokes and IC pulses that were used as input ~~for~~ to the GEO lightning pseudo-observation generator. The vast majority of strokes and pulses lies within the area of ~~strictly positive~~ non-zero FED. Figure 2(d) illustrates that the simulated MTG-LI flash centroids are situated within the corresponding Meteorage stroke and pulse distribution.

4   Lightning observation operator

~~In consequence of the former studies (see Section 1), graupel mass appears to be~~ Previous studies suggest that graupel is a reliable proxy for lightning (see Section 1). Our observation operator is trained ~~based on an observed relation~~ using the

relationship found between MTG-LI FED observations and ~~1-hour forecasts of AROME-France~~ simulated graupel mass (m_g) from the 1-hour forecast of AROME-France, for 24 days in 2018 ~~, i.e., (~~2 days ~~each month~~per month). FED time periods of 10 min are used ~~(like, e.g., Fierro et al., 2016; Hu et al., 2020)~~ (e.g., as in Fierro et al., 2016; Hu et al., 2020) centered at the

200 corresponding time of the AROME-France analysis. The m_g profile is extracted from the AROME-France grid point closest to the FED pixel center. Thus, one specific m_g profile is related to one FED value. This gives an equal count of FED and column integrated m_g values that are used to train and validate our observation operator. Following Deierling et al. (2008), ~~column integrated graupel mass is taken for~~ m_g is taken from layers where the temperature was below $-5\,°C$. As a reminder, AROME-France graupel mass combines graupel and hail. All FED and column graupel mass values are further processed

205 as climatological distributions (see Figure 3) ~~without~~ regardless of observation location and time. This approach is different from Deierling et al. (2008); Barthe et al. (2010) who used storm-based relationships of case studies, and from McCaul et al. (2009) who used the domain-wide peak values. Pixel-to-pixel m_g and FED were barely correlated in our study, with about 0.09 Pearson correlation coefficient, likely as a consequence of a typical displacement of convection in the model by more than the FED spatial resolution of $7\,km$. The second approach with domain-wide peak values was not further tested as it reduces the

210 sample size for a regression analysis drastically. Our observation operator uses a simple linear regression between observed FED and simulated column graupel mass. It is unbiased by definition. Our approach optimizes both the slope factor and the y-intercept of the regression, whereas, e.g., McCaul et al. (2009) only used a proportionality between FED and a proxy.

Figure 3 presents the analyzed linear relationship between FED and column integrated graupel mass m_g~~that are calculated per 7km × 7km pixel whereby~~. The functions are obtained after sorting both the FED and m_g ~~uses a single model grid

215 point closest to the FED pixel center~~data individually. Paired datapoints in Figure 3 are geographically independent values as described in Combarnous et al. (2022). It should be noted that all pixels with either FED equal to zero or m_g equal to zero are removed from the data. The observation operator represents the cases when lightning was actually observed. The training data (24 days) results are shown in Figure 3(a) and the results for 06 and 07 October 2018 as validation period can be seen in Figure 3(b). The Pearson correlation coefficient equals 0.97 for the training data and 0.92 for the validation data (0.96

220 combined).

Analyzing the training data, the linear relationship fits the majority of the data well. A discrepancy is identified for the largest graupel mass and FED values (Figure 3). Here, the observation operator tends to underestimate the FED for a given graupel mass. The high values (of both FED and graupel mass) are rare relative to the lower values as indicated by the colored pixels in Figure 3(a). Hence, this observed discrepancy litte affects the Pearson correlation coefficient. It should still be considered that

225 high FED values are systematically underestimated. It is further noted that the y-intercept is negative, meaning that statistically a certain mass of graupel is required to ~~get the first~~ initiate lightning. This result is well in accordance with the widely accepted non-inductive charging as main electrification process in extratropical storms.

The validation data (Figure 3(b)) ~~follow roughly~~ roughly follow the regression line inferred from the training data. One can see, however, that the slope is smaller than that of the black regression line, i.e., ~~the~~ observed values of FED are always lower

230 for a given m_g than the training data implies. The number of FED-m_g-pairs is significantly lower for the 48-hour validation case than for the 24-day training dataset. Furthermore, the validation data comprises one single meteorological situation, while
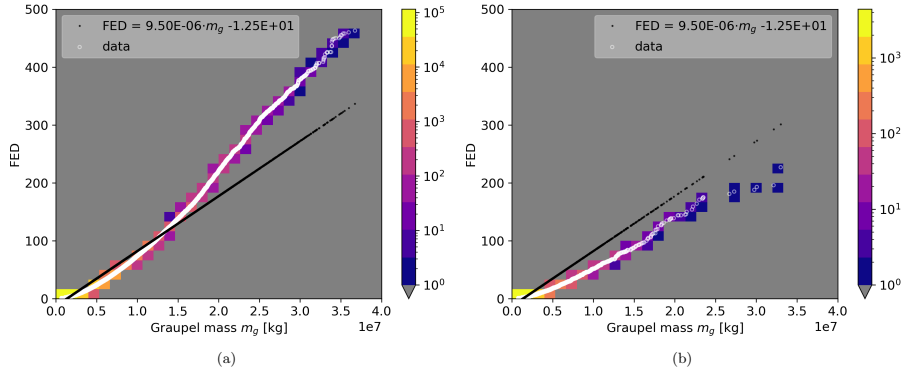
Figure 3. The observed FED, per $10\,\mathrm{min}$ and on $7\,\mathrm{km} \times 7\,\mathrm{km}$ pixels, climatological distribution versus the climatological distribution of AROME-France graupel mass m_g at model grid points closest to each FED observation. Grid points with any of FED or m_g equal to zero are not considered. (a) shows the training of a linear regression for 24 days in 2018, and (b) shows the results of a validation for independent data of 2 additional days in 2018. Colors indicate the number of samples, the white ~~points~~ circles plot the data points, and black ~~points~~ dots applied the m_g values in the linear regression equation (see Figure legend).

the training data includes several weather situations in different seasons. The observed discrepancy is considered during the evaluation of results.

5   1DBay+3DVar assimilation method

235   ~~The observation operator described in the previous section is used in the new lightning data assimilation method~~FED cannot be assimilated directly since lightning is not predicted by AROME-France. The previously described observation operator relates the FED observations to the prognostic variables of the NWP model and allows for comparing background and observations. The final step prior to the assimilation creates pseudo observations that can be ingested in the 3DVar system. Unlike Fierro et al. (2019); Hu et al. (2020) and others who used an empirical method to adjust moisture in thunderstorms, the expected
240   water vapor PO is retrieved for each model grid column (applying Bayes' Theorem) in a 1DBay approach. This allows to i) replace the humidity field in spurious convection areas with that of their convection-free environment, and thus in principle abolish the wet bias that results from data assimilation techniques that only consider the occurrence of lightning (e.g., Fierro et al., 2012, 2019; Hu et al., 2020), and ii) make use of the FED value to modulate the humidity field in observed lightning areas by leveraging the model's ability to create consistent, flow-dependent humidity and graupel profiles near the observation.
245   Technically speaking, synthetic profiles are created and assimilated as sounding data.

5.1   1DBay retrieval

All FED values are transformed to units of dB as $10 \cdot \log_{10}(\mathrm{FED}/(7\,\mathrm{km} \times 7\,\mathrm{km}\, 10\,\mathrm{min})^{-1})$ to account for the large range of scales, referred to as dBFED. FED equal to zero is transformed to dBFED of $-10\,\mathrm{dB}$. In general, dBFED below $0\,\mathrm{dB}$ means

linear FED less than 1^1 $(7\,\mathrm{km} \times 7\,\mathrm{km}\ 10\,\mathrm{min})^{-1}$, i.e., no lightning activity. The 1DBay retrieval of POs of relative humidity
RH, $y_{\mathrm{po}}^{\mathrm{RH}}$, is defined as

$$y_{\mathrm{po}}^{\mathrm{RH}} = \sum_i x_i^{\mathrm{RH}} \frac{W_i}{\sum_j W_j} \tag{1}$$

with the weights W_i for each point as

$$W_i = \exp\left\{ -\frac{1}{2} \frac{[\mathrm{FED} - \mathrm{FFO}(x_i)]^2}{\sigma_o^2} \frac{[\mathrm{dBFED} - \mathrm{dBFFO}(x_i)]^2}{\sigma_o^2} \right\} \tag{2}$$

where i, j are counters for the grid points within the defined area, referred to as vicinity, ~~and~~ FFO means the FED forward
operator as specific observation operator. ~~FFO(x_i)~~, and σ_o the standard deviation of the 1DBay retrieval. dBFFO(x_i) then
defines the simulated AROME-France ~~FED~~dBFED from using dB units for the output of the FFO.

Equations (1) and (2) can be used with different model variables or diagnostics, e.g., the sensitivity test for σ_o as described
in the following used x_i^{FED} rather than x_i^{RH}.

Each RH PO is the best estimation from a weighted linear combination of RH profiles taken from the model background in
the vicinity of the FED observation.

The method was proposed by Caumont et al. (2010), used in the operational model AROME-France (Wattrelot et al., 2014),
and applied by Borderies et al. (2019) for radar reflectivity data, and Duruisseau et al. (2019) for microwave radiances. Here,
weights are calculated based on differences between observed and simulated FED. The 1DBay retrieves the best ~~estimation~~
estimate of RH at the center of each observed FED pixel from the background using Bayes' Theorem. It is expected that the
model can predict a quantity similar to the observation within a certain area in most cases. The vicinity is initially fixed to
a square of $160\,\mathrm{km} \times 160\,\mathrm{km}$ centered at the observation. This size follows the suggestion of Borderies et al. (2018, 2019).
However, the FED pixels (about $7\,\mathrm{km}$) are significantly larger than the radar reflectivity pixels (about $1\,\mathrm{km}$) and a given vicinity
contains less FED than radar datapoints to retrieve the expected profile. Hence, the vicinity is adapted to the specifications of
the FED data. 1DBay vicinities of $160\,\mathrm{km}$, $320\,\mathrm{km}$, and $500\,\mathrm{km}$ were tested regarding the RMSE between observed FED and
retrieved FED (not shown). It was found that larger vicinities give lower overall RMSE as there are more background profiles
to feed the retrieval and to find grid points that are similar to the FED observation. In consequence, all LDA experiments
presented in the following use a 1DBay vicinity of $500\,\mathrm{km}$.

The database comprises model forecasts at the observation time and in the vicinity of the observation location. ~~The covariance
matrix of observation and observation operator is chosen diagonal with a trace of $n \cdot \sigma_o^2$, where n is the vertical dimension of
the observations and equals 1 for the 2D variable FED~~Since FED has no vertical dimension, the covariance matrix is reduced
to a single value of variance, σ_o^2, which is assumed to be constant. A small σ_o means that the retrieval favors columns with
values close to the FED observation. This can produce accurate retrievals; however, no retrieval can be created when the differ-
ence between all simulated FED in the vicinity and the observed FED is large relative to σ_o (see Equation (2) where weights

---

[1] Non-integer FED values are possible when applying the observation operator.
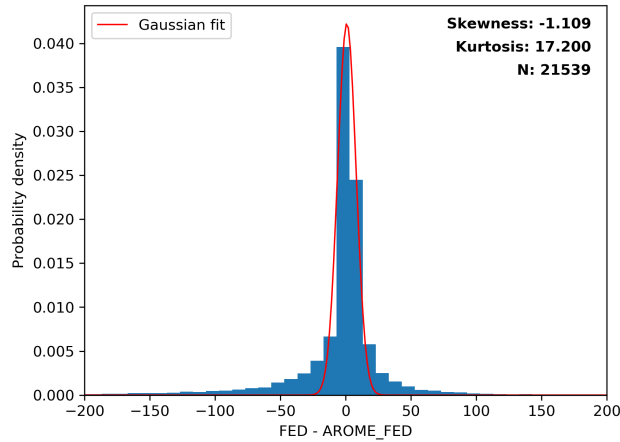
Figure 4. Innovations of FED observation minus AROME_FED for the 24-hour assimialtion cylces on 08 Aug. 2018. Skewness, kurtosis, and the sample size N are given. The red line shows the Gaussian fit on the PDF of the innovations.

~~approximate~~ approach 0 in that case). Large values of σ_o cause smoothing over all grid points in the vicinity of the observa-
280   tion. In that case, the likelihood to retrieve POs is high, at the cost of a less accurate retrieval potentially independent of the
observation.

An assumption of the Bayesian retrieval is that differences between the observation and the model background, referred to as
innovations, are Gaussian distributed. Figure 4 shows the PDF of innovations (dBFED observation minus AROME_dBFED)
for the 24-hour assimilation cylces on 08 Aug. 2018. The distribution is bell-shaped and centered at 0. It is symmetric and the
285   skewness is close to 0. Although the kurtosis is higher than for a classical Gaussian distribution, this distribution of innovations
sufficiently fulfills the assumption of the 1DBay method to justify its use in this study.

The value for the standard deviation σ_o of the observation and observation operator is inferred from a sensitivity study. It
aims at minimizing the root mean square error (RMSE) between observed and retrieved pseudo FED for the 24 training days
also used in Section 4. Pixel-to-pixel RMSE is used for all non-zero FED pixels and 10-minute ~~FED~~ intervals centered at each
290   full hour. The pseudo FED is computed from joint use of the observation operator and the 1DBay retrieval. ~~All FED values areF~~
~~transformed to units of dB as $10 \cdot \log_{10}(\text{FED}/(7\,\text{km} \times 7\,\text{km}\,10\,\text{min})^{-1})$ to account for the large range of scales, referred to as~~
~~dBFED. FED equal to zero is transformed to dBFED of $-10\,\text{dB}$. In general, dBFED below $0\,\text{dB}$ means linear FED less than~~
~~$1\,(7\,\text{km} \times 7\,\text{km}\,10\,\text{min})^{-12}$, i.e., no lightning activity.~~ Figure 5(a) shows the curve of the RMSE between dBFED observation
(from pseudo MTG-LI) and retrieved pseudo dBFED from the 1DBay retrieval for different σ_o. The RMSE between AROME-
295   France background dBFED (AROME_dBFED), that was obtained from the observation operator without additional 1DBay,
is shown as reference. Including the 1DBay retrieval (blue curve in Figure 5(a)) produces pseudo dBFED much closer to the
dBFED observation, i.e., lower RMSE, than the AROME_dBFED (orange curve in Figure 5(a)). The minimum RMSE for the
retrieval is found at σ_o of $2.0\,\text{dB}$. This value is used for retrieving the pseudo-RH profiles in the following.

---

[2] ~~Non-integer FED values are possible when applying the observation operator.~~
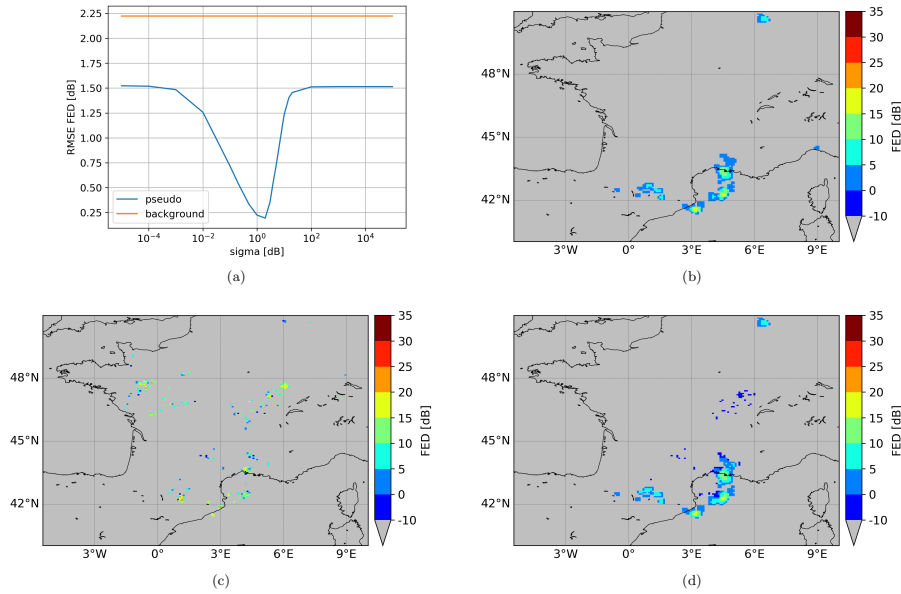
Figure 5. (a) The sensitivity test for standard deviation σ_o of the 1DBay retrieval as inferred from the pixel-to-pixel RMSE between the dBFED observation and the AROME_dBFED (background) as well as the 1DBay retrieved pseudo dBFED (pseudo). 07 Oct. 2018, 00:00 UTC, case with the MTG-LI dBFED observation (b), the model background AROME_dBFED (c), and the 1DBay retrieved pseudo dBFED with a σ_o of 2 (d). Grey color means no lightning~~in general~~.

Case studies of fields of observed, background, and retrieved dBFED are conducted to visualize the effect of σ_o on the pseudo
300   dBFED. ~~One example of a~~ The example of 07 Oct. 2018, 00:00 UTC ~~case~~ is presented in Figure 5(b-d). The dBFED observation in Figure 5(b) shows lightning activity mainly in the south center of the domain. ~~AROME_dBFED (Figure 5c) of the background exhibits positive~~ Figure 5(c) indicates positive background dBFED in the region of ~~dBFED observation~~lightning observations, however, also widespread near the center of the domain and over western France. The ~~2D distribution~~ map of the 1DBay retrieved pseudo dBFED in Figure 5(d) demonstrates that the method effectively reduces the spurious dBFED in these
305   regions (retrieved dBFED below 0, grey and marine blue). In some regions, the 1DBay retrieves ~~profiles~~ pseudo dBFED to completely remove the spurious lightning, e.g., over western France and for some grid points over Switzerland. Spurious high values of AROME_dBFED near the center of the domain (Figure 5c) are effectively decreased to negative dBFED (marine blue in Figure 5d) ~~meaning eventually~~ in other words no lightning activity, similar to the observation in that region (Figure 5b). Furthermore, the pseudo dBFED values and areas of positive pseudo dBFED in the south center of the domain closely match
310   the dBFED observation. The observed, ~~shallow dBFED in northwestern Italy, to the north of Corsica,~~ weak dBFED over the Gulf of Genoa could not be reproduced by the 1DBay retrieval as the background does not provide any positive dBFED in the ~~vicinity around this regional area~~retrieval vicinity.

An additional method, the humidity adjustment (HA), is applied if lightning (i.e., positive FED) is observed but all background FED values in the vicinity are zero. In this case, there is no estimated profile at this point and all W_i in the vicinity

equal zero. The HA is also applied if all background dBFED values within the vicinity of the FED observation are at least $10\,\text{dB}$ smaller than the observed dBFED value. Although the sum of W_i can become greater than zero in this case, the retrieval would generate a profile that is too dry with respect to the FED observation. To produce RH POs, the layer between lifted condensation level (LCL) and $13\,\text{km}$ is saturated (i.e., RH set to $100\,\%$) at all levels where the modeled RH is less than $100\,\%$. This is conceptually similar to the method of Fierro et al. (2019), among others. However, the HA is only applied for few pseudo-RH profiles where the 1DBay method did not retrieve POs. In example, the HA generates 1.6% of almost 20,000 assimilated profiles during 08 Aug. 2018, the study case that is detailed below.

Another case with Equations (1) and (2) equal to 0 for all weights may occur if the observed FED equals 0 and the background FED is positive for all grid points in its vicinity. This behavior was observed for the initial 160-km vicinity, but not within the 320-km and 500-km vicinity. Thus, if the vicinity has a sufficient size, there are background grid points without lightning activity that can be used in the 1DBay retrieval. If, however, the specified vicinity were too small, one would need to artificially remove spurious FED, i.e., convection and humidity, from the model.

Eventually, no pseudo-RH profiles are created if both the observed FED and the closest (in space) AROME-France background FED equal zero.

It should be mentioned that the 1DBay retrieval method was initially developed and applied to retrieve humidity and cloud profiles from passive and active remote sensing data (e.g., Olson et al., 1996; Kummerow et al., 2001). Caumont et al. (2010) brought this approach forward by restricting the 1DBay method to use model profiles at the forecast time and in the neighborhood of the observation. This approach was successfully applied by Wattrelot et al. (2014); Borderies et al. (2019) for radar reflectivity assimilation in AROME-France. However, a 1DBay has not yet been used to retrieve humidity profiles from FED data. Whereas an FED greater than zero is always related to the presence of graupel and thus a RH profile with a cloud, an FED equal to zero does not necessarily mean a location without cloud coverage. It is the same problem which is faced for radar reflectivity, but presumably more marked than for the centimeter wavelength radars (Caumont et al., 2010) and even more than for the millimeter wavelength radars which are even more sensitive to the small hydrometeors (Borderies et al., 2019). In addition, FED is a 2D variable without vertical extent while the 1DBay retrieves vertical RH profiles. The use of integrated column graupel mass as proxy addresses the latter aspect by converting the 3D AROME-France outputs into a 2D variable comparable to FED. RH-profiles are assimilated since assimilating hydrometeor contents when VAR is able to update these variables often results in poor performances because the cross correlations with key variables such as temperature and humidity are poorly represented in climatological $\mathbf{B}$ matrices. To mitigate this effect, some recent studies assimilate humidity along with hydrometeor contents (see, e.g., Wang et al., 2013; Do et al., 2022, for radar reflectivity).

5.2 3DVar assimilation

The retrieved pseudo-RH profiles are assimilated as sounding data in the 3DVar assimilation system of AROME-France. AROME-France uses a one-hour assimilation window. The short assimilation cycles aim at partially overcoming the missing temporal dimension and at allowing the assimilation of more high-frequency observations that can improve the initial conditions especially on the convective scale. AROME-France operationally assimilates surface (e.g., ground-stations, ships, buoys)

and aircraft measurements, Global Positioning System (GPS) Zenith Tropospheric Delay (ZTD) data, satellite brightness tem-
peratures of several polar orbiting satellites and from Meteosat Second Generation (MSG) SEVIRI, satellite-based atmospheric
motion vectors, and radar velocity and reflectivity data (Seity et al., 2011; Brousseau et al., 2016). The control variables are
temperature, specific humidity, surface pressure, and horizontal wind components. The 3DVar system minimizes ~~a classical~~ the
3DVar cost function J of the state vector x:

$$J(x) = \frac{1}{2}(x - x_b)^T \mathbf{B}^{-1}(x - x_b) + \frac{1}{2}[y_o - H(x)]^T \mathbf{R}^{-1}[y_o - H(x)] \tag{3}$$

with the state vector of the background x_b, the observation vector y_o, the observation operator H, and the observation error
covariance matrix $\mathbf{R}$. The climatological background error covariance matrix $\mathbf{B}$ is inferred from offline AROME-France en-
semble assimilation as a multivariate set of calculations for the control variable covariances and cross-covariances (Brousseau
et al., 2014).

6   ~~Model simulation experiments~~Experimental set-up

This section evaluates the effect of LDA relative to RDA. Since the application of the 1DBay retrieval for FED data constitutes
a new approach, different experiments of AROME-France with respect to LDA are conducted. All experiments are initiated at
0000 UTC and run for a forecast period of 30 hours until 0600 UTC the following day.

Table 1 lists the six different assimilation experiments. First, there is a control experiment used as reference without RDA,
without Doppler wind velocity assimilation, and without LDA. It is called *CTRL*. The second experiment includes the use of
radar data and is similar to the current operational AROME-France. It is referred to as *RDA*. All RDA experiments shown here
assimilate both reflectivity and Doppler wind velocity.

While analyzing first LDA experiments, it was evident that the use of all FED observations ~~leads~~ led to wrong results,
i.e., the changes to the AROME-France background humidity contradicted the assimilated pseudo-RH profiles (not shown).
This behavior is known to occur because the $\mathbf{R}$ matrix is diagonal whereas observation error cross-correlations are actually
present, which leads to sub-optimal solutions (e.g., Rabier, 2006). It could be mitigated by reducing the number of assimilated
observations (*thinning* Järvinen and Undén, 1997). ~~This work thinned FED~~ In this work, FED data are thinned by a factor of 2
in latitude and longitude direction, i.e., 1 ~~of~~ in 4 observations are assimilated, so that no observations of adjacent FED pixels
are used. Tests revealed that this thinning was sufficient to eliminate the FED observation error correlations. For comparison,
1 ~~of~~ in 64 radar observations, with a higher horizontal resolution than ~~this~~ our FED data, are assimilated in AROME-France
(Michel, 2018). The thinning, thus, prevents from assimilating observations with correlated observation errors~~contradicting~~,
which would contradict the assumption of a diagonal $\mathbf{R}$ matrix. The resulting LDA without RDA is labeled *LDA*.

The second LDA experiment without RDA ,*LDAnC*, adds a so-called *noCloud*-filter. This *noCloud*-filter is utilized for
locations where the observed FED equals zero but the AROME_FED exhibits lightning activity. Then, the distance to the
closest positive FED observation, d_{FED}, is computed. If d_{FED} remains within $21\,\mathrm{km}$, i.e., a maximum of three FED pixels,
it is assumed that the profile is still situated within the same thundercloud being responsible for observed FED greater than

Table 1. Simulation and assimilation techniques of the different AROME-France experiments~~conducted~~. They differ by the use of radar data assimilation (RDA) and lightning data assimilation (LDA), as well as the application of the *noCloud*(nC)-filter as described in the text.

Experiment	Assimilation	*noCloud*-filter
R (CTRL)	Reference (control)	no
RDA	CTRL + RDA	no
LDA	CTRL + LDA	no
LDAnC	CTRL + LDA	yes
RDA_LDA	CTRL + RDA + LDA	no
RDA_LDAnC	CTRL + RDA + LDA	yes

zero. In this case, the RH profile of the background is kept to avoid reducing the RH if observed FED equals zero but the location is likely associated with a cloud. In the case where d_{FED} exceeds $21\,\mathrm{km}$, AROME-France profiles within $21\,\mathrm{km}$ are not considered in the 1DBay as they ~~might belong to the spurious thunderclouds modeled but not observed~~probably belong to spurious simulated thunderclouds. Hence, the *noCloud*-filter should help to effectively reduce ~~the background humidity for~~
385 background humidity in spurious convection.

The two remaining experiments combine both LDA methods with RDA. They are labeled *RDA_LDA* and *RDA_LDAnc*, respectively.

All model experiments are initiated one day prior to the start of the model forecast. During the first 24 hours, the 3DVar assimilation system of AROME-France creates 23 analyses (there is no analysis for the time of initialization). ~~Therefore,~~
390 ~~reference~~ ~~experiment, the~~ The reference, LDA, and RDA experiments are conducted for 23 hours ~~in advance of~~ prior to the evaluated analysis. This time period ~~is chosen as convection was observed within~~ has been chosen because convection was continuously observed inside the model domain~~continuously~~. In addition, the long assimilation period allows AROME-France to ~~well ingest the~~ efficiently ingest all available observations.

7   The AROME-France analysis and forecast assimilating FED observations

395 This section describes the LDA ability to modify and update the model background. The following sections briefly introduce the four test cases, detail one case, and discuss i) the LDA effects on the AROME-France background and ii) AROME-France rainfall forecasts for this selected case. AROME-France simulations are mainly analyzed for the first 12 ~~forecast hours since the most effects~~ hours of the forecast since the strongest impacts of RDA and LDA are expected during these forecast hours, as seen in Fierro et al. (2019), for example~~, and the~~. The influence of the lateral boundary conditions becomes predominant after
400 12 hours (Vié et al., 2011).

(a)   (b)
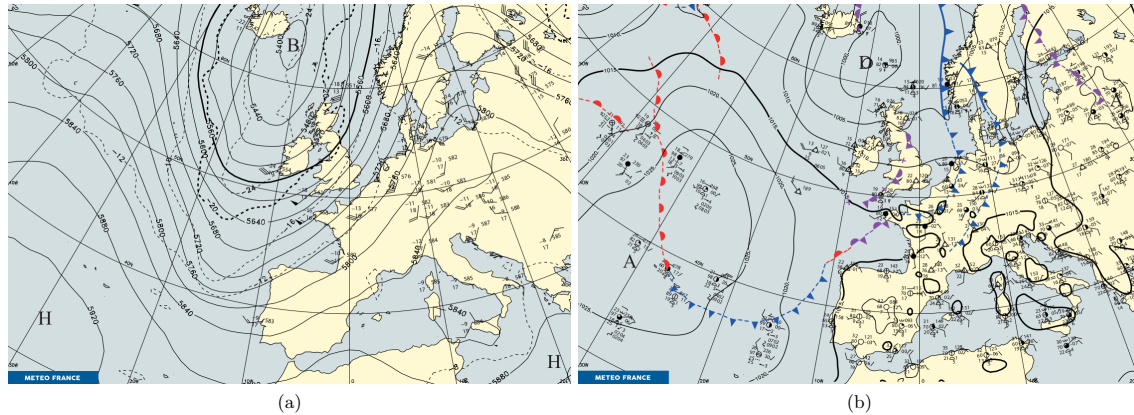
Figure 6. (a) Météo-France 500 hPa analysis of geopotential height (in geopotential metres, gpm, solid lines) and temperature (in °C, dashed lines) and (b) Météo-France surface analysis on 08 Aug. 2018, 12:00 UTC. Surface (altitude) low-pressure centres are indicated by "D" ("B") and high-pressure centres by "A" ("H"). Maps from the Météo-France daily weather bulletin archive https://donneespubliques.meteofrance.fr/?fond=produit&id_produit=129&id_rubrique=52 (accessed 19 October 2022).

7.1   Case 1: 08 to 10 August 2018

Strong convection and lightning activity was observed over France on 07 Aug., and the electrical cells were still active over northern France and Belgium during the early hours of 08 Aug. Moist and warm air created unstable conditions over southeastern France where lightning was observed during the entire day. Whereas a dominant southwest flow was present aloft, winds in lower levels were calm but caused slight warm air advection. The 2-m temperatures reached up to $33\,°C$ while the dew point temperatures exceeded $20\,°C$ over southwestern France. CAPE values far beyond $1000\,\mathrm{J\,kg^{-1}}$ indicated potential for widespread deep convection and thunderstorm activity. Figure 6(a) and (b) show the 500 hPa map and surface analysis, respectively, on 08 Aug. 2018, 12:00 UTC. The amplifying trough over the Atlantic is evident, and the major cold front was about to enter the study domain at this time. In the evening of 08 Aug., the ~~major cold front associated with the amplifying trough entered the study domain~~ ~~from the west~~trough that was situated over the Atlantic amplified and the associated cold front started to impact the weather in the domain. Strong convective cells formed over Spain and propagated all the way over France to Belgium and western Germany following the air mass boundary and the eastward movement of the trough. The storms left the study domain in the afternoon of 09 Aug. At that time, new cells formed over southeastern France and remained quasi-stationary until the end of the day. ~~Only~~ Then, only few cells developed ~~within the domain, i.e.,~~ over the northern Mediterranean Sea ~~, thereafter and~~ until the end of the forecast period. The 30-hour forecast starts 09 Aug. 2018, 00:00 UTC, preceded by a 24-hour period with hourly data assimilation.

7.1.1   AROME-France analysis

Details to create the AROME-France analysis are presented in Figure 7 for ~~the model~~ experiment *LDAnC* on 08 Aug. 2018, 23:00 UTC. The first row of the figure (a-c) includes (a) the AROME_dBFED~~(a) ,~~, (b) the 1DBay dBFED~~(b), and the dBFED observation~~, and (c) the dBFED observation. AROME-France could not predict most of the lightning activity over the Bay of Biscay and southwestern France (a and c; see Figure 1 for the locations of the regions of interest). The region of the highest observed dBFED is correctly predicted, however, with underestimated dBFED values (Figure 7 a,b versus c). The 1DBay dBFED (Figure 7b) shows positive dBFED patterns that match the observed dBFED (Figure 7c) despite the fragmented dBFED in the AROME-France background (Figure 7a). Hence, Figure 7(g) shows ~~added IWV~~ that IWV is added to the background in ~~the~~ regions where the AROME-France background underestimates the dBFED. ~~The~~ Over southern France, the marine blue spots in Figure 7(b) indicate that the 1DBay ~~dBFED (Figure 7b) method retrieves profiles southern France(marine blue) to reduce~~ retrieval succeeds in reducing the spurious AROME_dBFED (Figure 7a versus c)~~, implying that~~. Hence, IWV should be reduced here by the assimilation. Consequently, the PO IWV (Figure 7e) clearly decreases compared to the background IWV (Figure 7d) as shown ~~with the difference between PO IWV and background IWV (Figure 7~~in Figure 7(g). In addition, the LDA experiment reduces the model IWV at rare locations where the AROME_dBFED exceeded the observed dBFED, e.g., due to a slight spatial shift of the local dBFED maxima at about $44.3°$N and $0.8°$E (compare Figure 7a,c and g at this location).

The 1DBay method aims at finding the profiles that ~~should lead to an analysis with the expected increase and decrease of IWV relative to the observed and background dBFED values~~are physically consistent with FED observations in terms of relative humidity. Indeed, the analysis (Figure 7f) adds humidity to the background over the southwestern regions and for the cell over eastern France~~, as it~~. It can also be seen in the difference between analysis IWV and background IWV (Figure 7i) where LDA reduces model IWV where spurious lightning activity was predicted. In addition, Figure 7(h) ~~shows in addition~~ displays the IWV as output of the 1DBay retrieval method where no background IWV is included. ~~At pixels with retrieved pseudo-RH profiles(positive values in panel h), the 1DBay-only IWV equals the PO IWV for most pixels. As IWV integrates the water vapor content per column, this finding~~ It is intended that the 1DBay retrieves complete profiles, whereas the HA only adds humidity to certain layers. A smaller IWV from pure PO (Figure 7(h)) than from background plus PO (Figure 7(e)) means that the HA method was used. The comparison of Figure 7(e) and (h) shows the same IWV values for most points with PO. This result means that the HA is rarely used, and the preferred 1DBay method~~retrieves mostly complete vertical profiles. At pixels where 1DBay-only IWV values in Figure 7(h) would be smaller than the corresponding PO IWV values (Figure 7e)the HA method was used.~~, that retrieves flow consistent PO profiles, is mostly used to get the POs.

Figure 8 illustrates RH vertical structure on the $44°$N latitude between $4°$W et $5°$E at the same time as fields in Figure 7 are taken. AROME-France ~~background IWV (d), PO IWV merged with the background where no profiles where retrieved (e), the IWV~~ background RH (a) and the difference of analysis minus background (b) are included. The latitude was chosen as the analysis increases and decreases the background IWV here (Figure 7i). The latitude also exhibits regions where AROME-France over- and underestimated the FED (Figure 7a,c). In Figure 8(b), higher RH in the analysis than the background (positive values) from $4°$W to $2°$W (-4 to -2 in the figure) corresponds well with Figure 7(g, i) where humidity as evidenced by IWV was added
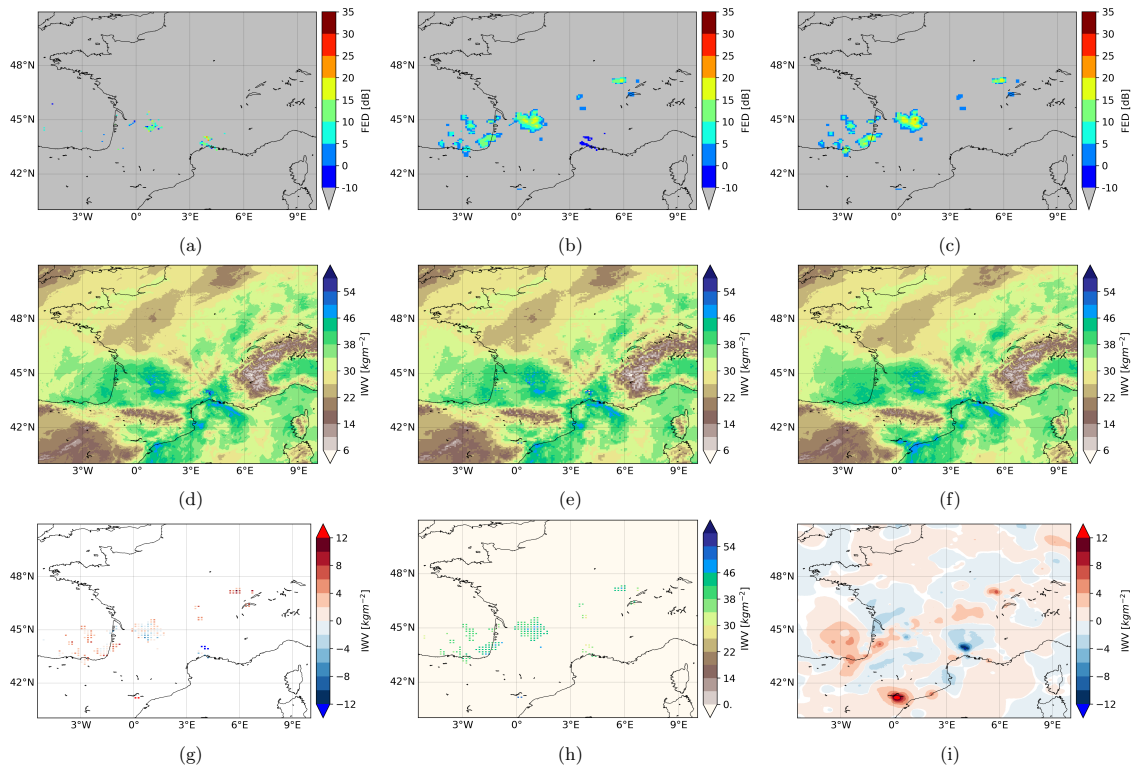
Figure 7. (a) The background AROME_dBFED, (b) the 1DBay retrieved dBFED, (c) the MTG-LI dBFED observation, (d) the AROME-France background IWV, (e) PO IWV merged with the background where no profiles were retrieved, (f) the IWV of the analysis, (g) the difference between the PO IWV minus the background IWV, and (h) the 1DBay-only IWV including only points where RH was retrieved, and (i) the difference in IWV of analysis and background. Results for 08 Aug. 2018, 23:00 UTC and model experiment *LDAnC*.

to the background since AROME_FED was lower than the FED observation. RH is reduced in columns near 4°E where spurious AROME_FED was found (Figure 7a,c). Figure 8 also shows that the RH is not changed when observed FED was higher than AROME_FED but the background RH was at least $100\,\%$ (e.g., altitudes up to $4\,\mathrm{km}$ and from 4°W to 2°W). All in all, the vertical cross-section shows that changes induced by the LDA result in physically consistent analysis in the vertical
455   struture.

7.1.2   AROME-France rainfall forecasts

The *Antilope* rainfall accumulation (RA) combines the data of the operational radar network of Météo-France and ground-based rain gauges on a $0.01°$ resolution grid at each hour (Laurantin, 2008, 2013). RA maps of for the first 6 hours of the forecast are shown in Figure 9a to d for *Antilope*(a)and the experiments *RDA Antilope* and experiments (b), *LDA CTRL*, (c)and *LDAnCRDA*,
460   (d)*LDA*, and (e)*LDAnC*.
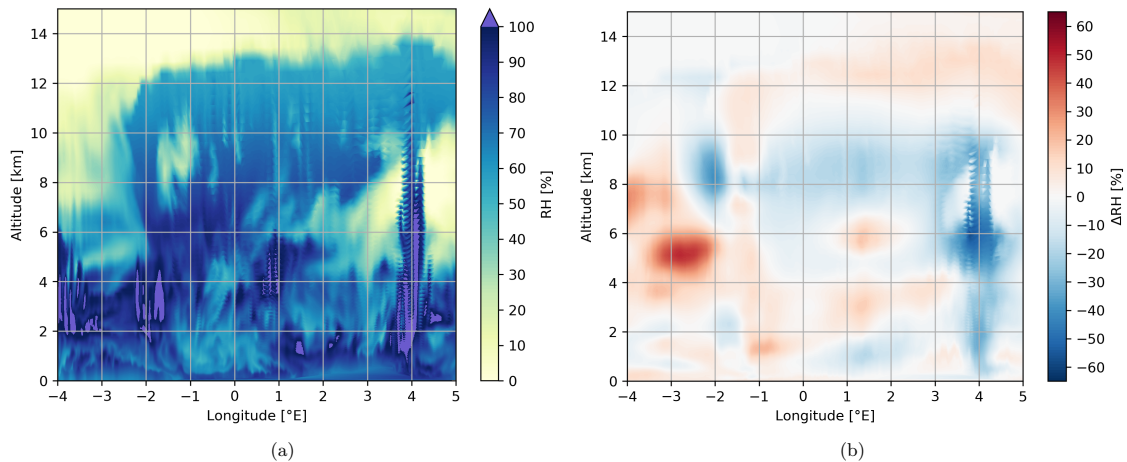
Figure 8. ~~Vertical cross-sections on 44°N of~~ the analysis (~~f~~a) ~~, the difference between the PO IWV minus~~ the AROME background ~~IWV (g),~~ ~~and the pure PO IWV including only points where~~ RH ~~was retrieved~~ and (~~h~~b) ~~, and~~ the difference ~~in IWV~~ of analysis ~~and~~ minus background ~~(i)~~RH. Results for 08 Aug. 2018, 23:00 UTC and model experiment *LDAnC*.

Three major thunderstorm tracks, labeled 1 to 3 from north to south hereafter, ~~produced~~ produce the bands of high RAs over the northwestern part of the domain (Figure 9a). The local maximum with very high RA up to 150 mm per 6 hours over southern France ~~was~~ is caused by a quasi-stationary thunderstorm development. ~~All three shown~~ The AROME-France experiments (Figure 9~~b-d) predicted~~ b-e) predict the RA of track 1 relatively well. ~~The~~ Experiment *CTRL* (Figure 9b) and the

465   two LDA experiments ~~underestimated~~ (Figure 9d,e) underestimate the area of high RA in the northwest of the domain for track 1. Experiments *RDA* (Figure 9c) and *LDA* ~~overestimated~~ (Figure 9d) overestimates the maximum RA over this region by about 20 mm and 5 mm, respectively. Track 2 is best predicted by experiment *LDA*~~(Figure 9c)~~, with good agreement in maximum RA and area of high RAs to the observation (Figure 9a). Experiment *LDAnC* (Figure 9~~d) could also predict~~ e) also predicts the extent of the RAs related to track 2, but ~~underestimated~~ underestimates the RA amounts especially for the

470   northern part of the track, i.e., most recent storm location to the prediction time. ~~The~~ Experiment *CTRL* and the experiment using RDA without LDA poorly ~~predicts RA~~ predict the high RAs related to track 2. It is not ~~totally~~ clear whether the storm ~~was~~ is modeled at all or placed too far south~~, however, the high RAs of track 2 were barely predicted~~. RAs of track 3 ~~were~~ are similarly predicted by the ~~three shown~~ RDA and LDA AROME-France experiments that all underestimate the RA of the southern part, i.e., during the beginning of the forecast period, and overestimate the RA of the northern part of the track, i.e., the

475   most recent RA. Experiment *CTRL* predicts the southern part of track 3 arguably better than the RDA and LDA experiments. 6-hour RAs related to the local thunderstorm over southern France ~~were~~ are best predicted by *CTRL* and *RDA*. Both LDA experiments somewhat ~~underestimated~~ underestimate the local RA maximum and the area of high RAs. In addition, the LDA experiment produces a spurious, local high RA cluster over northeastern Spain.
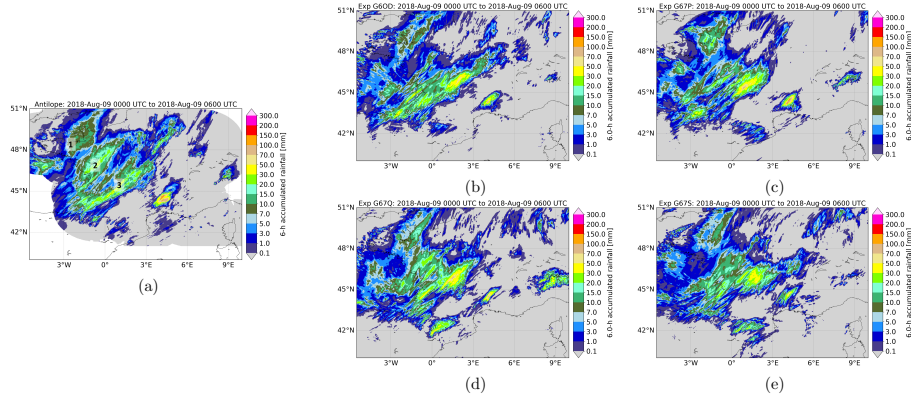
Figure 9. 6-hour rainfall accumulations (RAs) for the period until 09 Aug. 2018, 06:00 UTC, i.e., the first 6 ~~forecast~~ hours of the 30-hour forecast. ~~RA observations~~ (a) RA observations, the AROME-France experiments *RDA* (b) *CTRL*, *LDA* (c) ~~and *LDAnC*~~ *RDA*, (d) *LDA*, and (e) *LDAnC*. Three main thunderstorm tracks are labeled with black numbers 1 to 3 in a.

7.1.3 ~~Evaluating the skill of AROME-France rainfall forecats~~

480 ~~This section compares the predicted 6-hour RAs of different AROME-France experiments (Table 1) and the observations. RA observations are compared to the AROME-France output. The forecast skill is quantified calculating Fractions Skill Scores (FSSs). The FSS was introduced by Roberts and Lean (2008): The FSS can be calculated as skill score from the mean squared error (MSE) for the observed and forecast fraction $O_{(n)}$ and $M_{(n)}$, respectively, from a neighborhood of length n as~~

$$\mathrm{FSS}_{(n)} = 1 - \frac{\mathrm{MSE}_{(n)}}{\mathrm{MSE}_{(n)ref}}$$

with

$$\mathrm{MSE}_{(n)} = \frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \left[O_{(n)i,j} - M_{(n)i,j} \right]^2$$

$$\mathrm{MSE}_{(n)ref} = \frac{1}{N_x N_y} \left[\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} O_{(n)i,j}^2 + \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} M_{(n)i,j}^2 \right]$$

and

$$O_{(n)i,j} = \frac{1}{n^2} \sum_{k=1}^{n} \sum_{l=1}^{n} I_O \left[i+k-1-\frac{n-1}{2}, j+l-1-\frac{n-1}{2} \right]^2$$

$$M_{(n)i,j} = \frac{1}{n^2} \sum_{k=1}^{n} \sum_{l=1}^{n} I_M \left[i+k-1-\frac{n-1}{2}, j+l-1-\frac{n-1}{2} \right]^2$$

485 ~~with the observation and forecast binary fields I_O and I_M that equal 0 if the field value is smaller than the threshold, and 1 in all other cases. Our implementation uses the fast calculation of FSS in Python as proposed by Faggian et al. (2015). The FSS~~

is calculated hourly for 6-hour RAs with a sliding 6-hour time window for the 30-hour forecast period. Forecasts are initiated at 00:00 UTC. It should be noted that the first FSS value of the time series includes the last hour of the assimilation window, and ends at 05:00 UTC of the corresponding day.

490 Here, 6-hour RA thresholds of 0.1 mm, 1.0 mm, and 10.0 mm are used to represent different RA categories (weak, moderate, and strong rain). An FSS neighborhood of 0.5° is used. The figure shows the FSSs for the entire forecast period of 30 hours for the August case.

FSS of 0.8 to 0.9 for the RA thresholds 0.1 mm and 1.0 mm (**??**a and b, respectively) indicate that regions with precipitation were equally well identified by all six conducted AROME-France experiments. The LDA without RDA and without the
495 *noCloud*-filter,*LDA*, clearly outperforms all other experiments for the high RA threshold and during the first 10 hours of the forecast (**??**c). *RDA* exhibits the lowest FSS during the first 6 hours of the forecast due to the lack of high RAs of track 2 over *les Pays de la Loire* (see also Figure 9). The experiment *LDA* exhibits the highest FSS for the first 6 to 7 hours of the forecast demonstrating the high importance of lightning data for this case. Only after 12 hours of the forecast, the combination of RDA and LDA, *RDA_LDA*, has the highest FSS so RDA gets more important the longer the forecast period.

500 FSS for 6-hour RAs calculated hourly for a sliding window during the 30-hour forecast initiated 09 Aug., 00:00 UTC. The colors indicate the model experiments as defined in Section 6. 3 different RA thresholds are used, 0.1 mm (a), 1.0 mm (b), and 10.0 mm (c). The size of the domain used to calculate the FSS equals 0.5°.

This case study ~~leads to the conclusion that the~~ implies that LDA without both RDA and the *noCloud*-filter of FED has the highest ~~FSS to predict the heavy precipitationamounts~~ Fraction Skill Score (FSS, not shown, see Section 8) for predicting heavy
505 precipitation. In that it differs from the other three case studies where the *noCloud*-filter ~~mostly~~ improves the LDA experiment skill. A future analysis might detail why the RDA and LDA with *noCloud*-filter have lower skill than the *LDA*. Such an analysis is beyond the present proof-of-concept scope.

7.2 Cases 2, 3, and 4

All case studies are summarized in Table 2. Besides the detailed case 1, the other case studies analalyzed a cyclone with
510 mainly frontal precipitation in autumn (case 2), shallow convection in ~~cold early spring~~ winter (case 3), and widespread deep convection in late spring (case 4). The evaluation of these three cases revealed that the RDA and LDA improve the RA forecast skill for convection within the study domain over western Europe. ~~The combination of RDA and LDA exhibits always skill between RDA and LDA while once~~ Once *RDA* and once *LDAnC* ~~provided~~ provide the best RA forecast. ~~The best trade-off is found for the experiments using both~~ Experiments combining RDA and LDA exhibit FSS values between *RDA* and *LDA*, thus,
515 mean the best trade-off overall. In the future, a coupled assimilation of RDA and LDA could be tested. For example, PO RH profiles could be retrieved from a weighted product that includes both radar reflectivity and FED thereby adapting the 1DBay.

In situations, where convection from the Atlantic Ocean is advected into the study domain from the west, i.e., case 3, RDA and LDA could not improve the control run. Both RDA and LDA currently rely on ground-based observations with limited coverage over the Atlantic ocean. It is expected that the real MTG-LI observations will help to predict cloud systems forming

Table 2. Summary of the 4 case studies.

Case	Dates (2018)	Weather conditions	Main precipitation source
1	08 to 10 Aug	instability, warm and moist low levels	deep convection, cold front
2	06 to 08 Oct	dissipating cyclone	cold front
3	01 to 03 Feb	trough and baroclinic wave	shallow convection, cold front
4	27 to 29 May	unstable conditions, low level moisture	deep convection, mesoscale system

over the ocean with higher efficiency than the LDA using the generated MTG-LI observations derived from ground-based observations.

8 Quantitative skill of AROME-France RA forecasts

This section compares the predicted 6-hour RAs of different AROME-France experiments (Table 1) and the observations in a statistical way. The forecast skill is quantified calculating Fractions Skill Scores (FSSs). The FSS was introduced by Roberts and Lean (2008): The FSS can be calculated as skill score from the mean squared error (MSE) for the observed and forecast fraction $O_{(n)}$ and $M_{(n)}$, respectively, from a neighborhood of length n as

$$\text{FSS}_{(n)} = 1 - \frac{\text{MSE}_{(n)}}{\text{MSE}_{(n)ref}}$$

(4)

with

$$\text{MSE}_{(n)} = \frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \left[O_{(n)i,j} - M_{(n)i,j} \right]^2$$

$$\text{MSE}_{(n)ref} = \frac{1}{N_x N_y} \left[\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} O_{(n)i,j}^2 + \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} M_{(n)i,j}^2 \right]$$

and

$$O_{(n)i,j} = \frac{1}{n^2} \sum_{k=1}^{n} \sum_{l=1}^{n} I_O \left[i + k - 1 - \frac{n-1}{2}, j + l - 1 - \frac{n-1}{2} \right]$$

$$M_{(n)i,j} = \frac{1}{n^2} \sum_{k=1}^{n} \sum_{l=1}^{n} I_M \left[i + k - 1 - \frac{n-1}{2}, j + l - 1 - \frac{n-1}{2} \right]$$

with the observation and forecast binary fields I_O and I_M that equal 0 if the field value is smaller than the threshold, and 1 otherwise. Our implementation uses the fast calculation of FSS in Python as proposed by Faggian et al. (2015). The FSS is calculated hourly for 6-hour RAs with a sliding 6-hour time window for the 30-hour forecast period. Forecasts are initiated at 00:00 UTC. The FSS can be generalized to yield an average score if the numerator and denominator are averaged separately
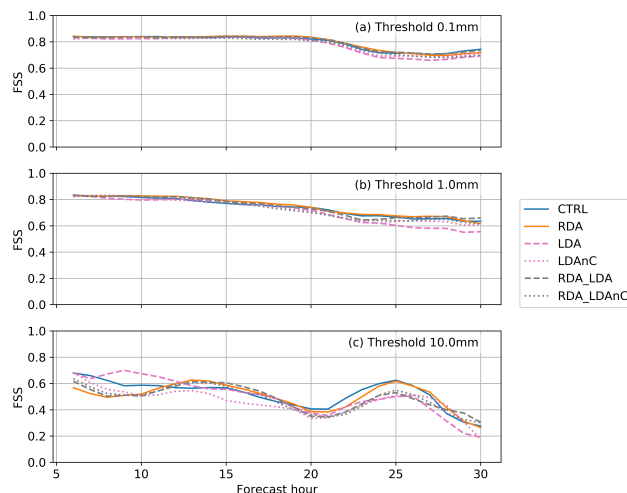
Figure 10. FSS average of 6-hour RAs calculated hourly for a sliding window during the 30-hour forecast initiated 00:00 UTC for the four study cases. The colors and line style indicate the model experiments as defined in Section 6. 3 different RA thresholds are used, (a) 0.1 mm , (b) 1.0 mm, and (c) 10.0 mm. The size of the domain used to calculate the FSS equals 0.5°.

and then the FSS is calculated (Faggian et al., 2015). This allows to achieve overall FSSs including forecasts of all four case studies.

Here, 6-hour RA thresholds of 0.1 mm, 1.0 mm, and 10.0 mm are used to represent different RA categories. An FSS neighborhood of 0.5° is used. Figure 10 shows the FSSs for the entire forecast period of 30 hours and combination of the four cases.

FSS of 0.8 to 0.9 during the first 15 hours of the forecast for the RA thresholds 0.1 mm and 1.0 mm (Figure 10a and b, respectively) indicate that regions with precipitation were equally well identified by all six AROME-France experiments. RDA or LDA effects on the FSS diminish beyond 12 hours of the forecast as the effect of the boundary conditions becomes predominant. The LDA without RDA and without the *noCloud*-filter (*LDA*) gains the highest FSS for the high RA threshold and during the first 12 hours of the forecast (Figure 10c). This finding demonstrates the high potential of LDA in AROME-France. The *noCloud*-filter cannot always improve the LDA. *RDA* exhibits the lowest FSS during the first 6 hours of the forecast due to a low skill during the Aug case where high RAs of storm track 2 over *les Pays de la Loire* were not predicted (see also Figure 9). The combination of LDA and RDA (*RDA_LDA*) gives FSS between the skill of RDA and LDA.

This proof-of-concept study uses four cases, a rather low amount of data. Usually, scores are calculated over a several months for evaluating whether a new method improves an existing NWP model. FSS curves appeared noisy for high thresholds and short accumulation periods (not shown) as a result. The main conclusion here is that there is likely not enough data to show significant differences, however, the encouraging point being that the effect of the assimilation is neutral.

9 Conclusions

The objective of this work is to design ~~of~~ an assimilation technique for the upcoming MTG-LI data for the regional, convection-permitting model AROME-France. To date, AROME-France applies a 3DVar assimilation system. A tailored 1DBay+3DVar assimilation technique (Caumont et al., 2010) is used to assimilate pseudo MTG-LI flash extent density (FED) in AROME-France. A similar assimilation technique is currently used operationally for radar reflectivity data assimilation in AROME-France, but has not yet been tested for ~~the~~ LDA.

This work first generated MTG-LI data that are used to create the FED observations (Erdmann et al., 2022). Then, an observation operator for FED is developed based on a linear, climatological relationship between observed FED and the column integrated AROME-France graupel mass, m_g, above the $-5\,^\circ$C isotherm ~~(m_g; as suggested by Deierling et al., 2008)~~. The operator is trained for 24 days in 2018, and validated for 2 independent days in 2018. Pearson correlation coefficients of 0.97 and 0.92 for the training and validation data, respectively, reveal a very strong relation between the distributions of observed FED and model m_g. Nevertheless, the observation operator systematically overestimates the FED for m_g values greater than 1.5×10^7 kg per AROME-France grid cell of $1.3\,\text{km} \times 1.3\,\text{km}$. More sophisticated observation operators are currently tested (Combarnous et al., 2022) but have not been included in this work yet.

The observation operator is then used to compare AROME-France-derived background FED (AROME_FED) to the FED observations. The 1DBay method identifies the best estimation of the FED from the background to in turn create pseudo-observations (POs) of RH-profiles based on both the FED observation and the AROME-France background fields. As background profiles are processed, the 1DBay method maintains model physics and flow characteristics.

The PO RH profiles add humidity to the AROME-France background where the observed FED exceeds the AROME_FED. It is further found that the 1DBay retrieval leads to reduction of humidity where the observed FED equals zero and the AROME_FED is positive, i.e., in regions of spurious convection with substantial m_g. Hence, the LDA technique improves the AROME-France background humidity. It is capable of both promoting convection in regions with lightning and suppressing spurious convection. This was successfully verified in four case studies where the new LDA technique provided similar skill than the operational RDA in AROME-France.

FED exhibits the highest values near the convective core of a thunderstorm and the lightning activity does not always cover the entire cloud size. In fact, zero FED exists at cloudy locations. In order to address the specific nature of FED data, the 1DBay retrieval method is adapted. In detail, ~~the vicinity used~~ a wider vicinity is used to identify vertical profiles in the 1DBay method ~~is expanded~~ and a so-called *noCloud*-filter is introduced. First results reveal that this adaption of the method can help to more effectively reduce the background humidity in regions of spurious convection, and to avoid a reduction of the background humidity if the profile occurs at the location of a cloud. Nonetheless, one of the four case studies revealed more skill in the LDA without the *noCloud*-filter than the LDA using this modification. ~~The authors thus~~ No lightning locating systems detects all flashes. The *noCloud*-filter is a first approach trying to overcome this issue by keeping the background humidity constant when lightning was correctly simulated but not observed. The authors encourage further research on how the specifications of FED data can best be addressed by the LDA scheme, i.e., through correlations between lightning locations and cloud cover.

In addition, this study found that thinning of the FED observations (on a grid of resolution of 7 km) was necessary to avoid the effect of correlations between the observation errors that violates the assumptions of a diagonal observation error matrix in
585 the AROME-France 3DVar data assimilation system (not shown).

Finally, ~~6-hour forecasts of RA fields~~ forecasts of 6-hour rain accumulations are evaluated through FSS analysis. The developed LDA scheme can in general compete with the established ~~radar data assimilation (RDA). Especially for the higher precipitation threshold of 10 mm per 6 hours, the LDAs exhibit higher FSSs during the first hours of the 30-hour forecast for two of four case studies~~RDA, an encouraging result for further testing and development of the LDA in AROME-France. Posi-
590 tive effects on the forecast of rainfall by both RDA and LDA are found mainly for the high precipitation threshold and during the first 9 to 12 hours of the forecast. Longer forecast times show small spread in FSS between control run, RDA, and LDA for the three RA thresholds indicating that the assimilated radar and FED data do not significant effect the model forecast after 12 hours. The FSS of the combination of RDA and LDA indicate in most cases skill between the RDA and the LDA. ~~In particular, the combination does not gain the highest skill, meaning that either an LDA without RDA or an RDA without LDA can lead to~~
595 ~~more accurate rainfall forecasts.~~ The combined RDA-LDA approach ~~does, however, provide~~ provides the best trade-off ~~for the four different cases studied overall~~in general sense and with respect to the assimilated observations. The case of 02 February 2018 was unique in that neither RDA nor LDA improved the FSS of the control run. This case was significantly influenced by weather phenomena over the Atlantic Ocean, with limited radar and Meteorage coverage. Real MTG-LI data will provide more accurate lightning records over the Atlantic than the synthetic MTG-LI data based on Meteorage observations used in
600 this study, and the LDA performance might improve in such situations.

It could be further studied whether the radar reflectivity and LDA can be coupled. A coupled approach may help to overcome some issues explained for zero FED at cloudy locations in general or at least for precipitating clouds, and, at the same time benefit from the successful promotion and suppression of convection by both RDA and LDA. As the LDA ~~explorered~~ explored here is based on satellite observations, such a coupled approach will also benefit ~~to~~ forecasts over regions with limited radar
605 coverage. The specific improvement over regions such as oceans and mountain ranges has not been studied here since our pseudo MTG-LI observations are based on ground-based Meteorage records. A study addressing in particular the skill over such regions would be of great interest as soon as real MTG-LI data become available.

Author contributions. FE wrote the paper text and created the figures. OC and ED were involved in content creation. All contributed to article reviews. OC and FE developed the assimilation technique.

610 *Competing interests.* The authors declare that they have no conflict of interest.

Acknowledgements. This work was a part of FE's PhD and as such funded by CNES and Météo-France. This article is funded by the SOLID (CNES) project and the EXAEDRE project (ANR-16-CE04-0005).

References

Allen, B. J., Mansell, E. R., Dowell, D. C., and Deierling, W.: Assimilation of Pseudo-GLM Data Using the Ensemble Kalman Filter, Monthly Weather Review, 144, 3465–3486, https://doi.org/10.1175/MWR-D-16-0117.1, 2016.

Allen, D. J. and Pickering, K. E.: Evaluation of lightning flash rate parameterizations for use in a global chemical transport model, Journal of Geophysical Research: Atmospheres, 107, ACH 15–1–ACH 15–21, https://doi.org/https://doi.org/10.1029/2002JD002066, 2002.

Apodaca, K., Zupanski, M., DeMaria, M., Knaff, J. A., and Grasso, L. D.: Development of a hybrid variational-ensemble data assimilation technique for observed lightning tested in a mesoscale model, Nonlin. Processes Geophys., 21, 1027–1041, https://doi.org/10.5194/npg-21-1027-2014, 2014.

Barthe, C., Deierling, W., and Barth, M. C.: Estimation of total lightning from various storm parameters: A cloud-resolving model study, Journal of Geophysical Research: Atmospheres, 115, https://doi.org/10.1029/2010JD014405, 2010.

Betz, H. D., Schmidt, K., Laroche, P., Blanchet, P., Oettinger, W. P., Defer, E., Dziewit, Z., and Konarski, J.: LINET—An international lightning detection network in Europe, Atmospheric Research, 91, 564 – 573, https://doi.org/https://doi.org/10.1016/j.atmosres.2008.06.012, 13th International Conference on Atmospheric Electricity, 2009.

Borderies, M., Caumont, O., Augros, C., Bresson, E., Delanoë, J., Ducrocq, V., Fourrié, N., Bastard, T. L., and Nuret, M.: Simulation of W-band radar reflectivity for model validation and data assimilation, Quarterly Journal of the Royal Meteorological Society, 144, 391–403, https://doi.org/10.1002/qj.3210, 2018.

Borderies, M., Caumont, O., Delanoë, J., Ducrocq, V., Fourrié, N., and Marquet, P.: Impact of airborne cloud radar reflectivity data assimilation on kilometre-scale numerical weather prediction analyses and forecasts of heavy precipitation events, Natural Hazards and Earth System Sciences, 19, 907–926, https://doi.org/10.5194/nhess-19-907-2019, 2019.

Bouyssel, F., Berre, L., Bénichou, H., Chambon, P., Girardot, N., Guidard, V., Loo, C., Mahfouf, J.-F., Moll, P., Payan, C., and Raspaud, D.: The 2020 Global Operational NWP Data Assimilation System at Météo-France, Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications, IV, 645–664, https://doi.org/10.1007/978-3-030-77722-7_25, 2022.

Bovalo, C., Barthe, C., and Pinty, J.-P.: Examining relationships between cloud-resolving model parameters and total flash rates to generate lightning density maps, Quarterly Journal of the Royal Meteorological Society, 145, 1250–1266, https://doi.org/10.1002/qj.3494, 2019.

Brousseau, P., Desroziers, G., Bouttier, F., and Chapnik, B.: A posteriori diagnostics of the impact of observations on the AROME-France convective-scale data assimilation system, Quarterly Journal of the Royal Meteorological Society, 140, 982–994, https://doi.org/10.1002/qj.2179, 2014.

Brousseau, P., Seity, Y., Ricard, D., and Léger, J.: Improvement of the forecast of convective activity from the AROME-France system, Quarterly Journal of the Royal Meteorological Society, 142, 2231–2243, https://doi.org/10.1002/qj.2822, 2016.

Buiat, M., Porcu, F., and Dietrich, S.: Observing relationships between lightning and cloud profiles by means of a satellite-borne cloud radar, Atmospheric Measurement Techniques, 10, 221, https://doi.org/10.5194/amt-10-221-2017, 2017.

Caumont, O., Ducrocq, V., Wattrelot, E., Jaubert, G., and Pradier-Vabre, S.: 1D+3DVar assimilation of radar reflectivity data: a proof of concept, Tellus A: Dynamic Meteorology and Oceanography, 62, 173–187, https://doi.org/10.1111/j.1600-0870.2009.00430.x, 2010.

Chen, Y., Yu, Z., Han, W., He, J., and Chen, M.: Case Study of a Retrieval Method of 3D Proxy Reflectivity from FY-4A Lightning Data and Its Impact on the Assimilation and Forecasting for Severe Rainfall Storms, Remote Sensing, 12, 1165, https://doi.org/10.3390/rs12071165, 2020.

Combarnous, P., Erdmann, F., Caumont, O., Defer, E., and Martet, M.: An observation operator for geostationary lightning imager data assimilation in the French storm-scale numerical weather prediction system AROME, Natural Hazards and Earth System Sciences Discussions, 2022, 1–30, https://doi.org/10.5194/nhess-2022-39, 2022.

Davies, H. C.: A lateral boundary formulation for multi-level prediction models, Quarterly Journal of the Royal Meteorological Society, 102, 405–418, https://doi.org/10.1002/qj.49710243210, 1976.

Deierling, W. and Petersen, W. A.: Total lightning activity as an indicator of updraft characteristics, Journal of Geophysical Research, 113, D16 210, https://doi.org/10.1029/2007JD009598, 2008.

Deierling, W., Petersen, W. A., Lathman, J., Ellis, S., and Christian, H. J.: The relationship between lightning activity and ice fluxes in thunderstorms, Journal of Geophysical Research, 113, D15 210, https://doi.org/10.1029/2007JD009700, 2008.

Dixon, K., Mass, C., Hakim, G., and Holzworth, R.: The Impact of Lightning Data Assimilation on Deterministic and Ensemble Forecasts of Convective Events, Journal of Atmospheric and Oceanic Technology, 33, https://doi.org/10.1175/JTECH-D-15-0188.1, 2016.

Do, P.-N., Chung, K.-S., Lin, P.-L., Ke, C.-Y., and Ellis, S. M.: Assimilating Retrieved Water Vapor and Radar Data from NCAR S-PolKa: Performance and Validation Using Real Cases, Monthly Weather Review, 150, 1177 – 1199, https://doi.org/https://doi.org/10.1175/MWR-D-21-0292.1, 2022.

Dobber, M. and Grandell, J.: Meteosat Third Generation (MTG) Lightning Imager (LI) Instrument Performance and Calibration from User Perspective, in Proceedings of the 23rd Conference on Characterization and Radiometric Calibration for Remote Sensing (CALCON), 11-14 August 2014, Utah State University, Logan, Utah, USA. 13 pages, 2014.

Duruisseau, F., Chambon, P., Wattrelot, E., Barreyat, M., and Mahfouf, J.-F.: Assimilating cloudy and rainy microwave observations from SAPHIR on board Megha Tropiques within the ARPEGE global model, Quarterly Journal of the Royal Meteorological Society, 145, 620–641, https://doi.org/10.1002/qj.3456, 2019.

Erdmann, F.: Préparation à l'utilisation des observations de l'imageur d'éclairs de Météosat Troisième Génération pour la prévision numérique à courte échéance (Preparation for the use of Meteosat Third Generation Lightning Imager observations in short-term numerical weather prediction), Ph.D. thesis, Université Toulouse 3 – Paul Sabatier, Toulouse, France, http://thesesups.ups-tlse.fr/4947/, 2020.

Erdmann, F., Defer, E., Caumont, O., Blakeslee, R. J., Pédeboy, S., and Coquillat, S.: Concurrent satellite and ground-based lightning observations from the Optical Lightning Imaging Sensor (ISS-LIS), the low-frequency network Meteorage and the SAETTA Lightning Mapping Array (LMA) in the northwestern Mediterranean region, Atmospheric Measurement Techniques, 13, 853–875, https://doi.org/10.5194/amt-13-853-2020, 2020.

Erdmann, F., Caumont, O., and Defer, E.: A geostationary lightning pseudo-observation generator utilizing low frequency ground-based lightning observations, Journal of Atmospheric and Oceanic Technology, 39, 3 – 30, https://doi.org/10.1175/JTECH-D-20-0160.1, 2022.

Faggian, N., Roux, B., Steinle, P., and Ebert, B.: Fast calculation of the Fractions Skill Score, Mausam, 66, 457–466, 2015.

Federico, S., Avolio, E., Petracca, M., Panegrossi, G., Sanò, P., Casella, D., and Dietrich, S.: Simulating lightning into the RAMS model: implementation and preliminary results, Natural Hazards and Earth System Sciences, 14, 2933–2950, https://doi.org/10.5194/nhess-14-2933-2014, 2014.

Federico, S., Panegrossi, G., Petracca, M., Transerici, C., and Dietrich, S.: Impact of the assimilation of lightning data on the precipitation forecast at different forecast ranges, Advances in Science and Research, 14, 187–194, https://doi.org/10.5194/asr-14-187-2017, 2017.

685   Fierro, A. O., Mansell, E. R., Ziegler, C. L., , and MacGorman, D. R.: Application of a Lightning Data Assimilation Technique in the WRF-ARW Model at Cloud-Resolving Scales for the Tornado Outbreak of 24 May 2011, Monthly Weather Review, 140, 2609–2627, https://doi.org/10.1175/MWR-D-11-00299.1, 2012.

Fierro, A. O., Gao, J., Ziegler, C. L., Mansell, E. R., MacGorman, D. R., and Dembrek, S. R.: Evaluation of a Cloud-Scale Lightning Data Assimilation Technique and a 3DVAR Method for the Analysis and Short-Term Forecast of the 29 June 2012 Derecho Event, Monthly

690   Weather Review, 142, 183–202, https://doi.org/10.1175/MWR-D-13-00142.1, 2014.

Fierro, A. O., Gao, J., Ziegler, C. L., Calhoun, K. M., Mansell, E. R., and MacGorman, D. R.: Assimilation of Flash Extent Data in the Variational Framework at Convection-Allowing Scales: Proof-of-Concept and Evaluation for the Short-Term Forecast of the 24 May 2011 Tornado Outbreak, Monthly Weather Review, 144, 4373–4393, https://doi.org/10.1175/MWR-D-16-0053.1, 2016.

Fierro, A. O., Wang, Y., Gao, J., and Mansell, E. R.: Variational Assimilation of Radar Data and GLM Lightning-Derived Water Vapor for

695   the Short-Term Forecasts of High-Impact Convective Events, Monthly Weather Review, 147, 4045–4069, https://doi.org/10.1175/MWR-D-18-0421.1, 2019.

Fischer, C., Bouyssel, F., Brousseau, P., El Khatib, R., Pottier, P., Seity, Y., Wattrelot, E., and Joly, A.: Les modèles opérationnels de prévision numérique à aire limitée de Météo-France, La Météorologie, 100, 18–28, https://doi.org/10.4267/2042/65139, numéro Spécial Anniversaire 25 ans, 2018.

700   Formenton, M., Panegrossi, G., Casella, D., Dietrich, S., Mugnai, A., Sano, P., Di Paolo, F., Betz, H.-D., Price, C., and Yair, Y.: Using a cloud electrification model to study relationships between lightning activity and cloud microphysical structure, Natural Hazards and Earth System Sciences, 13, 1085–1104, https://doi.org/10.5194/nhess-13-1085-2013, 2013.

Giannaros, T. M., Kotroni, V., and Lagouvardos, K.: Predicting lightning activity in Greece with the Weather Research and Forecasting (WRF) model, Atmospheric Research, 156, 1 – 13, https://doi.org/https://doi.org/10.1016/j.atmosres.2014.12.009, 2015.

705   Giannaros, T. M., Kotroni, V., and Lagouvardos, K.: WRF-LTNGDA: A lightning data assimilation technique implemented in the WRF model for improving precipitation forecasts, Environmental Modelling & Software, 76, 54 – 68, https://doi.org/https://doi.org/10.1016/j.envsoft.2015.11.017, 2016.

Goodman, S. J., Blakeslee, R. J., Koshak, W. J., Mach, D., Bailey, J., Buechler, D., Carey, L., Schultz, C., Bateman, M., McCaul, E., and Stano, G.: The GOES-R Geostationary Lightning Mapper (GLM), Atmospheric Research, 125-126, 34 – 49,

710   https://doi.org/https://doi.org/10.1016/j.atmosres.2013.01.006, 2013.

Hu, J., Fierro, A. O., Wang, Y., Gao, J., and Mansell, E. R.: Exploring the Assimilation of GLM-Derived Water Vapor Mass in a Cycled 3DVAR Framework for the Short-Term Forecasts of High-Impact Convective Events, Monthly Weather Review, 148, 1005–1028, https://doi.org/10.1175/MWR-D-19-0198.1, 2020.

Janisková, M. and Lopez, P.: Linearized Physics for Data Assimilation at ECMWF, in: Data Assimilation for Atmospheric, Oceanic and

715   Hydrologic Applications, edited by Park, S. and Xu, L., vol. 2, Springer, Berlin, Heidelberg, https://doi.org/10.1007/978-3-642-35088-7_11, 2013.

Järvinen, H. and Undén, P.: Observation screening and background quality control in the ECMWF 3D-Var data assimilation system, Tech. rep. ECMWF, https://doi.org/10.21957/lyd3q81/, 1997.

Karagiannidis, A., Lagouvardos, K., Lykoudis, S., Kotroni, V., Giannaros, T., and Betz, H.-D.: Modeling lightning density using cloud top

720   parameters, Atmospheric Research, 222, 163 – 171, https://doi.org/https://doi.org/10.1016/j.atmosres.2019.02.013, 2019.

Kong, R., Xue, M., Fierro, A. O., Jung, Y., Liu, C., Mansell, E. R., and MacGorman, D. R.: Assimilation of GOES-R Geostationary Lightning Mapper Flash Extent Density Data in GSI EnKF for the Analysis and Short-Term Forecast of a Mesoscale Convective System, Monthly Weather Review, 148, 2111–2133, https://doi.org/10.1175/MWR-D-19-0192.1, 2020.

Kummerow, C., Hong, Y., Olson, W. S., Yang, S., Adler, R. F., McCollum, J., Ferraro, R., Petty, G., Shin, D.-B., and Wilheit, T. T.: The Evolution of the Goddard Profiling Algorithm (GPROF) for Rainfall Estimation from Passive Microwave Sensors, Journal of Applied Meteorology, 40, 1801–1820, https://doi.org/10.1175/1520-0450(2001)040<1801:TEOTGP>2.0.CO;2, 2001.

Lagouvardos, K., Kotroni, V., Defer, E., and Bousquet, O.: Study of a heavy precipitation event over southern France, in the frame of HYMEX project: Observational analysis and model results using assimilation of lightning, Atmospheric Research, 134, 45 – 55, https://doi.org/https://doi.org/10.1016/j.atmosres.2013.07.003, 2013.

Lascaux, F., Richard, E., and Pinty, J.-P.: Numerical simulations of three different MAP IOPs and the associated microphysical processes, Quarterly Journal of the Royal Meteorological Society, 132, 1907–1926, https://doi.org/10.1256/qj.05.197, 2006.

Laurantin, O.: ANTILOPE: hourly rainfall analysis merging radar and rain gauge data, Grenoble and Autrans, France, 10-15 March 2008, 2008.

Laurantin, O.: ANTILOPE: hourly rainfall analysis over France merging radar and rain gauges data, edited by: Leijnse, H. and Uijlenhoet, R., KNMI, Ede-Wageningen, the Netherlands, 30 June to 3 July 2013, 2013.

Liu, P., Yang, Y., Gao, J., Wang, Y., and Wang, C.: An Approach for Assimilating FY4 Lightning and Cloud Top Height Data Using 3DVAR, Frontiers in Earth Science, 8, 288, https://doi.org/10.3389/feart.2020.00288, 2020.

Lynn, B. H., Kelman, G., and Ellrod, G.: An Evaluation of the Efficacy of Using Observed Lightning to Improve Convective Lightning Forecasts, Weather and Forecasting, 30, 405–423, https://doi.org/10.1175/WAF-D-13-00028.1, 2015.

Mansell, E. R.: Storm-Scale Ensemble Kalman Filter Assimilation of Total Lightning Flash-Extent Data, Monthly Weather Review, 142, 3683–3695, https://doi.org/10.1175/MWR-D-14-00061.1, 2014.

Mansell, E. R., Ziegler, C. L., and MacGorman, D. R.: A Lightning Data Assimilation Technique for Mesoscale Forecast Models, Monthly Weather Review, 135, 1732–1748, https://doi.org/10.1175/MWR3387.1, 2007.

Marchand, M. R. and Fuelberg, H. E.: Assimilation of Lightning Data Using a Nudging Method Involving Low-Level Warming, Monthly Weather Review, 142, 4850–4871, https://doi.org/10.1175/MWR-D-14-00076.1, 2014.

McCaul, Eugene W., J., Goodman, S. J., LaCasse, K. M., and Cecil, D. J.: Forecasting Lightning Threat Using Cloud-Resolving Model Simulations, Weather and Forecasting, 24, 709–729, https://doi.org/10.1175/2008WAF2222152.1, 2009.

Michel, Y.: Revisiting Fisher's approach to the handling of horizontal spatial correlations of observation errors in a variational framework, Quarterly Journal of the Royal Meteorological Society, 144, 2011–2025, https://doi.org/https://doi.org/10.1002/qj.3249, 2018.

Olson, W. S., Kummerow, C. D., Heymsfield, G. M., and Giglio, L.: A Method for Combined PassiveActive Microwave Retrievals of Cloud and Precipitation Profiles, Journal of Applied Meteorology, 35, 1763–1789, https://doi.org/10.1175/1520-0450(1996)035<1763:AMFCPM>2.0.CO;2, 1996.

Papadopoulos, A., Chronis, T., and Anagnostou, E.: Improving Convective Precipitation Forecasting through Assimilation of Regional Lightning Measurements in a Mesoscale Model, Monthly Weather Review, 133, 1961–1977, https://doi.org/10.1175/MWR2957.1, 2005.

Pergaud, J., Masson, V., Malardel, S., and Couvreux, F.: A Parameterization of Dry Thermals and Shallow Cumuli for Mesoscale Numerical Weather Prediction, Boundary-Layer Meteorology, 132, 83–106, https://doi.org/10.1007/s10546-009-9388-0, 2009.

Pessi, A. T. and Businger, S.: The Impact of Lightning Data Assimilation on a Winter Storm Simulation over the North Pacific Ocean, Monthly Weather Review, 137, 3177–3195, https://doi.org/10.1175/2009MWR2765.1, 2009.

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

Price, C. and Rind, D.: What determines the cloud-to-ground lightning fraction in thunderstorms?, Geophysical Research Letters, 20, 463–466, https://doi.org/10.1029/93GL00226, 1993.

Pédeboy, S.: Analysis of the French lightning locating system location accuracy, pp. 337–341, https://doi.org/10.1109/SIPDA.2015.7339299, balneário Camboriú, Brazil, 28th Sept. to 2nd Oct. 2015, 2015.

Qie, X., Zhu, R., Yuan, T., Wu, X., Li, W., and Liu, D.: Application of total-lightning data assimilation in a mesoscale convective system based on the WRF model, Atmospheric Research, 145-146, 255 – 266, https://doi.org/https://doi.org/10.1016/j.atmosres.2014.04.012, 2014.

Rabier, F.: Importance of Data: A Meteorological Perspective, in: Ocean Weather Forecasting, edited by Chassignet, E. and Verron, J., chap. 12, pp. 343–360, Springer, Dordrecht, https://doi.org/10.1007/1-4020-4028-8_12, 2006.

Roberts, N. and Lean, H.: Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events, Monthly Weather Review - MON WEATHER REV, 136, 78–97, https://doi.org/10.1175/2007MWR2123.1, 2008.

Schulz, W., Diendorfer, G., Pedeboy, S., and Poelman, D. R.: The European lightning location system EUCLID – Part 1: Performance analysis and validation, Nat. Hazards Earth Syst. Sci., 16, 595–605, https://doi.org/10.5194/nhess-16-595-2016, 2016.

Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., and Masson, V.: The AROME-France Convective-Scale Operational Model, Monthly Weather Review, 139, 976–991, https://doi.org/10.1175/2010MWR3425.1, 2011.

Vié, B., Nuissier, O., and Ducrocq, V.: Cloud-Resolving Ensemble Simulations of Mediterranean Heavy Precipitating Events: Uncertainty on Initial Conditions and Lateral Boundary Conditions, Monthly Weather Review, 139, 403 – 423, https://doi.org/10.1175/2010MWR3487.1, 2011.

Wang, H., Sun, J., Fan, S., and Huang, X.-Y.: Indirect Assimilation of Radar Reflectivity with WRF 3D-Var and Its Impact on Prediction of Four Summertime Convective Events, Journal of Applied Meteorology and Climatology, 52, 889 – 902, https://doi.org/https://doi.org/10.1175/JAMC-D-12-0120.1, 2013.

Wang, H., Liu, Y., Cheng, W. Y. Y., Zhao, T., Xu, M., Liu, Y., Shen, S., Calhoun, K. M., and Fierro, A. O.: Improving Lightning and Precipitation Prediction of Severe Convection Using Lightning Data Assimilation With NCAR WRF-RTFDDA, Journal of Geophysical Research: Atmospheres, 122, 12,296–12,316, https://doi.org/10.1002/2017JD027340, 2017a.

Wang, H., Liu, Y., Zhao, T., Liu, Y., Xu, M., Shen, S., Jiang, Y., Yang, H., and Feng, S.: Continuous Assimilation of Lightning Data Using Time-Lagged Ensembles for a Convection-Allowing Numerical Weather Prediction Model, Journal of Geophysical Research: Atmospheres, 123, 9652–9673, https://doi.org/10.1029/2018JD028494, 2018.

Wang, Y., Yang, Y., and Wang, C.: Improving forecasting of strong convection by assimilating cloud-to-ground lightning data using the physical initialization method, Atmospheric Research, 150, 31 – 41, https://doi.org/https://doi.org/10.1016/j.atmosres.2014.06.017, 2014.

Wang, Y., Yang, Y., Liu, D., Zhang, D., Yao, W., and Wang, C.: A Case Study of Assimilating Lightning-Proxy Relative Humidity with WRF-3DVAR, Atmosphere, 8, 20, https://doi.org/10.3390/atmos8030055, 2017b.

Wattrelot, E., Caumont, O., and Mahfouf, J.-F.: Operational Implementation of the 1D+3D-Var Assimilation Method of Radar Reflectivity Data in the AROME Model, Monthly Weather Review, 142, 1852–1873, https://doi.org/10.1175/MWR-D-13-00230.1, 2014.

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

Yair, Y., Lynn, B., Price, C., Kotroni, V., Lagouvardos, K., Morin, E., Mugnai, A., and Llasat, M. d. C.: Predicting the potential for lightning activity in Mediterranean storms based on the Weather Research and Forecasting (WRF) model dynamic and microphysical fields, Journal of Geophysical Research: Atmospheres, 115, https://doi.org/10.1029/2008JD010868, 2010.

Zhang, D., Cummins, K. L., Bitzer, P. M., and Koshak, W. J.: Evaluation of the Performance Characteristics of the Lightning Imaging Sensor, Journal of Atmospheric and Oceanic Technology, 36, 1015–1030, https://doi.org/10.1175/JTECH-D-18-0173.1, 2019.

800