

Review of manuscript egusphere-2022-637 by Erdmann *et al.* (2022)

Summary and recommendation:

First, the authors introduce a previously developed method to generate pseudo MTG-LI spaceborne lightning observations from real-life lightning measurements from the NLDN network over the US. Here, this method is applied using observations from the Meteorage ground-based detection network over France. Then, they present a simple lightning observation operator using a linear regression between observed lightning flash extent densities and corresponding graupel mass forecast by the AROME NWP model. Their two-step lightning data assimilation technique, in which output relative humidity profiles from a 1D Bayesian retrieval are subsequently assimilated in 3D-Var, is described. They note that a similar approach is already being used operationally to assimilate weather radar reflectivities in the AROME-France system. A novel treatment of situations with spurious lightning in the model is introduced and included in their experimentation. One case study is then presented in detail, which illustrates the reasonable performance of the 1DBay+3D-Var lightning data assimilation method, both in terms of the improvement of the analysis itself, but also of the subsequent 30h-range precipitation forecast. The respective benefits of assimilating lightning and radar reflectivity observations are also assessed.

This manuscript therefore offers a valuable new contribution to recent efforts towards the assimilation of lightning observations, which is clearly very challenging, but might also bring sizeable improvements in the quality of numerical weather forecasts, especially in regions of intense convection. Overall, the text is rather well written and the figures provide appropriate information (even though larger size at the printing stage will certainly be beneficial). However, I would recommend that a few important issues (listed as major comments) be addressed before the manuscript can be accepted for publication. Although categorized as "**major revision**", I believe addressing my concerns should not be too difficult, since additional experimentation should not be needed.

Major comments:

- [1] Unless the authors decide to show some plot or table summarizing the performance of their method for all four cases, section 7.2 and Table 2 do not seem relevant. In particular, the detailed description of the three extra cases in Table 2 is not needed. If the authors decide to focus on Case 1 only, the main conclusions (esp. at lines 475, 482 and 497) should be changed accordingly.
- [2] Since actual MTG-LI observations are not yet available, I would recommend to modify the title of the paper to "Assimilation of Meteosat Third Generation (MTG) Lightning Imager (LI) **pseudo**-observations in AROME-France - Proof of Concept".
- [3] The introduction, although interesting, feels rather lengthy. Even though previous studies do deserve to be cited, some references are cited with excessive detail. As it stands, the introduction could almost make a review paper in itself. I would therefore recommend that the authors try to shorten their review of past literature.
- [4] In section 3, even though the goal here is to show that NLDN and Meteorage perform in a similar way, the authors should explicitly underline the fact that ground-based networks, on the one hand, and ISS-LIS, on the other hand, are likely to have quite different detection efficiencies for CG and IC flashes, respectively.
- [5] To define their lightning observation operator, the authors compute a linear regression from the statistical distributions of observed FED and simulated graupel mass values. Their aim is to overcome

the issue of frequent point-to-point mismatches between observations and model. It is also mentioned that MTG-LI FED observations are paired with the closest graupel mass values. The authors should clarify whether they specified a maximum distance in both space and time, when searching which model point should be paired with a given observation. If no maximum distance was applied, their pairing approach might be questionable in some situations.

- [6] The flash detection efficiency of ground-based lightning detection networks, such as Meteorage, is well below 100%, especially for IC flashes. This implies that a no-lightning observation does not necessarily guarantee that (weak) lightning activity was not present in the real world. I believe the assimilation of such ambiguous observation to correct a lightning-active model background might lead to a degradation in the analysis. Could the authors clarify whether/how this possible ambiguity of no-lightning observations is/could be handled in their method?
- [7] Could the authors comment on the choice of RH as the sole variable of their 1DBay procedure? How does their approach address the ambiguity caused by RH's dependence on both temperature and specific humidity? Indeed, the same RH value can be obtained for very different temperature and specific humidity values.
- [8] Which types of observations are assimilated in experiment CTRL?
- [9] It would be very interesting to see a comparison between two (appropriately located) vertical cross-sections of 3D-Var analysis increments in temperature and specific humidity, the first one from CTRL, and the second one for LDA-CTRL. This would help to check the consistency between the increments due to the lightning assimilation (alone) and those produced by all standard observations (as assimilated in CTRL). In other respects, maps of IWV increments alone do not give any hint on whether the vertical structure of 3D-Var increments is realistic.
- [10] In section 7.1.2, a plot showing the RA maps for experiment CTRL should be added to help assess the relative improvements brought by RDA and LDA.
- [11] Figure 7: Unless article length is really an issue, it would be very helpful to add a figure similar to Figure 7, but showing the differences between the FSS of each experiment and the FSS of CTRL. The curves in Figure 7 are sometimes hard to distinguish, especially for the two lowest RA thresholds.

Scientific comments/questions:

- [1] Somewhere in the text, it would seem relevant to underline the pros and cons of the proposed two-step (1DBay + 3D-Var) assimilation technique, compared to the direct 3D-Var (or 4D-Var) assimilation of lightning observations.
- [2] Equation 2: The formulation of the weights assumes that observation and model errors follow Gaussian distributions. Is this assumption valid for FED data?
- [3] line 248: Here, the authors could simply state that "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." There is no need to mention the trace.
- [4] line 254: Is the optimal value of σ_o expected to be sensitive to the size of the "vicinity" (currently set to 500 km) (index i in Equation 1)?
- [5] line 286: Could the authors give a rough estimate of the percentage of profiles where their Humidity Adjustment needs to be applied?

- [6] line 332: Why would the thinning be applied in the latitude direction only?
- [7] Equation 4, lines 418-419: In the definition of O and M, the brackets ($[...]^2$) should not be squared (see Robert and Lean 2008). This is probably a typo.
- [8] line 429: "FSS between 0.8 and 0.9 [...]". This is only true during the first 15 hours of the forecast.
- [9] line 501: Could the authors expand a little on their idea of RDA-LDA coupling?

Minor comments:

Please note: The authors do not need to provide a formal reply to any of the following minor comments.

General remark: The American spelling used throughout the manuscript might not be suitable for a European journal (please check with the Editor).

- * line 26: Remove "among others".
- * lines 26-27: The sentence "GEO lightning sensors [...]" should be removed, since it just repeats what was written on lines 22-23.
- * line 30: Remove article "The" (end of line).
- * line 32: Add "However, LDA [...]" .
- * line 38: "Despite those difficulties [...]" .
- * line 38 and 46: "former" → "previous" .
- * line 45: Another noteworthy reference would be Allen and Pickering (2002)[doi:10.1029/2002JD002066].
- * line 65: "Federico et al. (2014) simulated" (past tense, for consistency).
- * line 66: "the most reliable proxy [...]" (add article).
- * line 67:"cell-scale" and "domain-scale" (hyphen).
- * line 98: "one hour into the forecast" .
- * line 113: At which vertical levels was this RH-correction applied? Only in the troposphere?
- * line 121: I would suggest to rephrase as follows: "Various approaches in which FED is assimilated together with radar reflectivity and/or Doppler wind [...]" .
- * line 126: Remove "reviews lightning observation operators and" .
- * line 128 "among others" → "including" .
- * line 132: Why not at 15 and 21h UTC? (just out of curiosity).
- * line 135: "extends up to pressure level 10 hPa" .
- * line 155: "separation of" → "discrimination between" .
- * lines 164-165: The sentence "In all cases [...]" ought to be rephrased, to make it clearer.
- * line 166: "as a suitable input to the GEO [...]" .
- * line 167: Add a comma after "7 km" .
- * line 173: The acronym "1DBay" should be defined here, in the main text.

- * line 175: "[...] is obvious in Figure 2(b)".
- * line 176: "Figure 2 accumulates data of 5 min" sounds redundant, since the authors already mentioned that the period shown in Figure 2 is 13:55-14:00 UTC.
- * lines 176-177: It would seem enough to write: "Throughout the domain, the FED pixels have a size of roughly 7 km x 7 km."
- * line 179: "strictly positive" → "non-zero".
- * line 178: Rephrasing suggestion: "Figure 2(c) superimposes the zoomed-in simulated FED and the Meteorage [...] used as input to the GEO lightning [...]"
- * line 183: I would suggest: "Previous studies suggest that graupel is a reliable [...]"
- * line 184: May I suggest: "[...] is trained using the relationship found between MTG-LI FED observations and simulated graupel mass (m_g) from hourly forecasts with AROME-France, for 24 days in 2018 (2 days per month)"
- * lines 185-186: "(e.g., as in Fierro et al. 2016 [...]"
- * line 190: "without" → "regardless of".
- * Figure 3 caption: "white points" and "black points" → "white dots" and "black dots".
- * line 209: "to get the first lightning" → "to initiate lightning".
- * line 211: "roughly follow".
- * line 212: "the values of FED are always [...]" → "observed values of FED are always [...]"
- * line 229: σ_o should be defined for the first time right after Equation 2.
- * line 237: "estimation" → "estimate".
- * line 251: "approximate 0" → "approach 0".
- * line 258: "areF" → "are".
- * line 260: Please check the exponent applied to "(7km x 7km 10min)" in the text.
- * line 267: "The example of 07 Oct. 2018 00:00 UTC [...]"
- * line 268: The whole sentence starting by "AROME_dBFED (Figure 4c) [...]" is somewhat awkward. It ought to be rephrased and maybe split into two.
- * line 269: "[...] in the region of positive dBFED observations [...]"
- * line 270: "The 2D distribution of the 1DBay [...]" → "The map of the 1DBay [...]"
- * line 272: "[...] 1DBay retrieves profiles [...]": are we really dealing with profiles, given that the 1DBay seems to be applied to FED values here?
- * line 274: "meaning eventually no lightning activity," → "in other words no lightning activity,"
- * line 276: "shallow dBFED in northwestern Italy, to the north of Corsica," → "weak dBFED over the Gulf of Genoa" ?
- * line 278: "in the vicinity around this regional area" → "in this region" ?
- * Figure 4: Remove "in general" at the end of the caption ?
- * line 285: "of among others Fierro et al. (2019)" → "of Fierro et al. (2019), among others".
- * line 290: "If, however, the **specified** vicinity **were** too small, one would need to **artificially** remove [...]"
- * line 294: "has been initially" → "was initially".

- * line 304: "by converting the 3D AROME-France **outputs into**".
- * line 308: "allowing **the** assimilation".
- * line 320: I would suggest to rename this section as "Experimental set-up".
- * line 328: "[...] the use of all FED observations **led** to wrong results [...]".
- * Table 1: I would suggest to modify the caption: "Set-up of AROME-France assimilation and simulation experiments: [...]".
- * line 332: "Here, FED data are thinned by a factor of 2 [...], i.e., 1 in 4 observations".
- * line 334: "1 in 64 radar observations".
- * line 335: "than our FED data".
- * line 336: "The thinning, thus, prevents from assimilating observations with correlated observation errors , **which would otherwise contradict** the assumption of a diagonal R matrix."
- * line 344: "as they probably belong to spurious simulated thunderclouds."
- * line 345: "reduce background humidity in spurious convection."
- * line 349: "The reference, LDA and RDA experiments [...] prior to the evaluated analysis".
- * line 351: "This time period has been chosen because convection was continuously observed inside the model domain." ?
- * line 352: "to efficiently ingest all available observations".
- * line 356: "the first 12 hours of the forecast".
- * line 357: "the most effects" → "the strongest impacts".
- * line 357: This sentence should probably be split into two, for clarity.
- * line 366: "the amplifying trough": which trough is referred to, here?
- * line 370: "thereafter and until the end of the forecast period": this bit of sentence does not sound right.
- * line 373: "for experiment LDAnC".
- * line 373: "Figure 5(g) shows that IWV is added to the background in regions where [...]".
- * line 380: Is there a missing word in the sentence: "The 1DBay dBFED (Figure 5b) [...]" ?
- * line 382: "as shown in Figure 5g." (more concise).
- * Figure 5 caption:
 - I believe that each panel number should appear before each corresponding description. Please check (also in main text).
 - Typo: "where no profiles **were** retrieved".
 - Should not panel (h) rather be described as "1DBay-only IWV", to be consistent with main text?
- * line 386: Suggestion: "The 1DBay method aims at finding the profiles that lead to analysis increments which improve the fit of simulated dBFED values to observations."
- * lines 387-389: Splitting this long sentence into two shorter ones would make it more readable.
- * line 389: A full stop seems to be missing near the end of this line.
- * line 390: "In addition, Figure 5(h) displays [...]" .

- * Section 7.1.2: Use present tense, when describing the rain validation from the experiments, throughout this section. Also, is there any reference available for Antilope?
- * line 396: Remove "ground-based".
- * line 397: "RA maps **for** the first 6 hours **of the forecast**".
- * line 398: Remove "the" (before experiments).
- * Figure 6 caption: "the first 6 hours of the 30h-range forecast."
- * line 408: "however" does not seem to be the right word here (rather simply: "and" ?).
- * line 415: Add "in a statistical way" at the end of the first sentence. Remove the second sentence, which is redundant.
- * line 421: "and 1 otherwise."
- * line 430: Remove "conducted".
- * line 434: Add a comma between "forecast" and "demonstrating".
- * line 434, second sentence: Suggestion: "The combination of LDA and RDA only becomes beneficial beyond the 12-hour forecast range".
- * line 436: Remove "the" before "LDA".
- * line 437: "to predict the heavy precipitation amounts" → "for predicting heavy precipitation".
- * line 437 (end): "mostly improves" → "greatly improves" ?
- * line 441: "the **other** case studies **analyzed** [...]".
- * line 442: Meteorologically, 1-3 Feb cannot be described as an "early spring" situation.
- * line 444: The sentence "The combination of RDA and LDA exhibits [...]" should be rephrased for better clarity.
- * line 454: Remove "of" in " [...] to design of an assimilation [...]".
- * line 456: Please add "pseudo" before "MTG-LI flash extent density".
- * line 458: Remove "the" before "LDA".
- * line 461: The reference given in brackets is not needed in the conclusions.
- * line 479: Missing "to" after "the vicinity used"? Is "expanded" the right word here?
- * line 484: "i.e., **through** correlations [...]".
- * line 488: "Finally, forecasts of 6-hour rain accumulations [...]" ?
- * line 489: "the established RDA".
- * line 503: Typo: "explored".
- * line 504: "also benefit forecasts [...]".
- * line 505: "our **pseudo** MTG-LI observations".