Reply on RC2

Comment 1:

Line 49: It will be possible to apply the authors' correction method to FY-4C LMI data, but not in real-time since it relies on other lightning data. Is such a post-processed dataset planned? More generally, will lessons learned from FY-4A mitigate the need for such a method due to improved operational algorithms for FY-4C?

Response:

We thank the reviewer for this forward-looking question.

Applicability of the correction method to FY-4C and planned post-processed dataset:

Yes, our correction method can in principle be applied to FY-4C LMI data, because the fundamental sensor design and the geolocation error sources (e.g., CTH parallax, thermal deformation, and platform jitter) are similar between FY-4A and FY-4C, although the latter includes hardware upgrades. However, as the reviewer correctly notes, our method relies on ground-based reference data (WWLLN) and therefore cannot be applied in real time. We are planning to produce a post-processed corrected dataset for FY-4C once a sufficient amount of coincident LMI and WWLLN data becomes available (e.g., after the first boreal summer of FY-4C operation). This dataset would be released in a similar format to the FY-4A product presented in this study.

However, it should be noted that FY-4A is positioned at 104.7°E, whereas FY-4C is positioned at 133°E. This means that the Lightning Mapping Imager (LMI) onboard FY-4C primarily observes ocean areas rather than land. Consequently, for lightning correction over coastal and near-ocean regions, the proposed method (using WWLLN as reference) remains applicable, similar to the approach used for FY-4A. For deep ocean areas where WWLLN has limited station coverage and lower detection reliability, the correction strategy may need to be adjusted. In such areas, alternative reference datasets such as the Lightning Imaging Sensor (LIS) onboard the Tropical Rainfall Measuring Mission (TRMM) could be considered as a replacement for WWLLN, given that LIS offers high-quality lightning detection over ocean regions. A multi-source data fusion approach that integrates information from different satellite platforms (e.g., LIS, WWLLN) may provide a more comprehensive correction capability for FY-4C LMI observations over the vast ocean. This will be an important direction for our future work.

Lessons learned and potential mitigation for FY-4C operational algorithms:

We fully agree that the experience gained from FY-4A should help improve the operational calibration and navigation algorithms for FY-4C. The systematic biases we identified (e.g., the diurnal thermal deformation signature and the CCD-array-dependent longitudinal offset) provide specific targets for hardware and software improvements. It is likely that FY-4C will

benefit from better thermal control, more accurate onboard calibration, and refined geometric correction models, which may reduce the magnitude of residual errors compared to FY-4A. Nevertheless, we believe that certain residual errors (e.g., random platform jitter and inter-detector inconsistencies) are difficult to completely eliminate by operational algorithms alone. Therefore, our empirical post-processing method may still be valuable for FY-4C to achieve the highest possible geolocation accuracy for scientific applications, especially for climate-quality datasets. We will add a brief discussion of this point in the Conclusions or a future outlook section of the revised manuscript.

Comment 2:

Line 75: How are thermal distortions (forced by the diurnal cycle) related to launch-induced payload displacement? Presumably those would occur even in the absence of any launch effects.

Response:

We thank the reviewer for this clarifying question. We would like to distinguish between two different types of displacement:

Launch-induced payload displacement: a permanent, static mechanical offset caused by the large acceleration during rocket launch. On-orbit thermal distortion: a periodic, dynamic deformation caused by diurnal variations in solar radiation flux after the satellite enters its operational orbit. Even in the absence of any launch effects, thermal distortion would still occur. FY-4A is in a geostationary orbit at an altitude of approximately 35,800 km. Due to Earth's shadow, the satellite is only blocked from solar radiation for about 2 hours per day; for the rest of the time it is exposed to sunlight. As the Earth rotates, the angle between the LMI viewing direction and the solar direction changes continuously, causing a periodic variation in the solar radiation flux entering the instrument. This in turn induces a diurnal pattern of thermal deformation on the CCD array. The geometric relationship is schematically illustrated in Fig. 1.

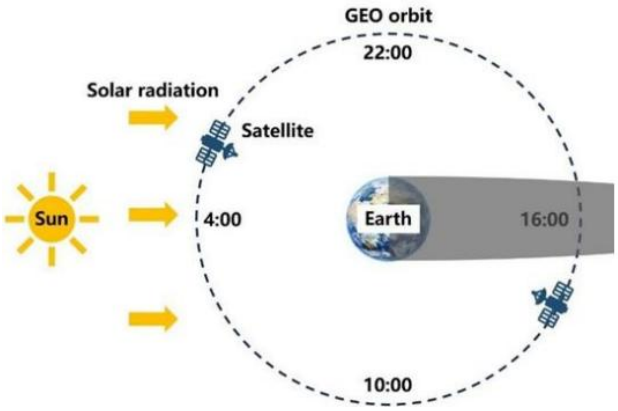

Figure 1 The schematic diagram illustrating the impact of solar radiation on FY-4A LMI thermal deformation. The time is in UTC. Figure from Zhang et.al., (2026).

https://doi.org/10.3390/rs18010163

Therefore, in our study we have considered both the static, launch-induced displacement and the periodic residual biases caused by on-orbit thermal deformation, which are the primary focus of our empirical correction. These two error sources are not contradictory; together they constitute the complete picture of LMI geolocation errors.

Comment 3:

Line 129: what were the other "ground-based lightning observations" used by Fan et al. (2018) for comparison with WWLLN?

Response:

Thank you for the question. According to Fan et al. (2018), the "other ground-based lightning observations" used for comparison with WWLLN were primarily the Cloud-to-Ground Lightning Location System (CGLLS) developed by the State Grid Corporation of China, for the period 2013–2015. CGLLS is designed to detect cloud-to-ground strokes, and its data were used to evaluate the detection efficiency and location accuracy of WWLLN over the mid-southern Tibetan Plateau. The quantitative results show that the average spatial separation between matched WWLLN and CGLLS strokes in the mid-southern Tibetan Plateau was 9.97 km, and that between matched WWLLN and LIS flashes over the entire Tibetan Plateau was 10.93 km. Furthermore, the detection efficiency of WWLLN increased markedly with stroke peak current; over the mid-southern Tibetan Plateau, the DE of WWLLN for CG lightning was 9.37% and for total lightning was 2.58%. We will add this description in the revised manuscript.

Comment 4:

Line 131: "of [WWLLN is] around 10 km"

Response:

Thank you for pointing this out. We have corrected Line 131 from "of around 10 km" to "of WWLLN is around 10 km" in the revised manuscript.

Comment 5:

Line 141: please provide a reference for "marginal effect analysis"

Response:

Thank you for pointing this out. In the revised manuscript (Section 3.1), we will add the following statement with appropriate references:

"Currently, the selection of spatiotemporal matching thresholds between different lightning detection systems mainly relies on sensitivity analysis combined with the marginal effect principle (Zhu, 2018; Thompson et al., 2014)."

Comment 6:

Line 179: "the weight assigned to the k-th data point" is defined as omega sub k with an overbar, but the formula also has a tilde above the overbar. In the definition of Eq. 2, there is an additional notation that adds a superscript "raw" to omega. Please clarify how each of these symbols are related to one another, and their possibly varied role in the spline fitting process. For example, is Eq. 2 only a starting estimate for omega?

Response:

We thank the reviewer for the careful reading and for pointing out the ambiguity in the notation. In the revised manuscript, we will clarify the relationship between the different weight symbols as follows:

$\widetilde{\varpi}_k$(appearing in Eq. 1, with a tilde above the overbar) is the final weight used in the fidelity term of the spline fitting. It is obtained by normalizing ϖ_k^{raw} to a reasonable range (0.1–1.0) and applying additional constraints, such as downweighting or excluding intervals with extremely small standard deviations, and those manually flagged as outliers.

ϖ_k^{raw}(defined in Eq. 2) is the raw weight computed solely from the number of samples N_k and the standard deviation σ_k. It represents an initial reliability estimate for each 10-minute interval.

We will add the above detailed explanation in the equation section of the revised manuscript.

Comment 7:

Section 3.1: matching of WWLLN strokes to LMI events is described here. I would have expected matching to group centroids. In the case of the a bright optical pulse comprised of many pixels, does the authors' fitting procedure only include those events that are in the 3s and 30 km space/time window? Are the events outside this window simply discarded in the fitting analysis?

Response:

We thank the reviewer for raising this important question. Initially, we also considered using LMI groups for matching, because by definition a group is the centroid of all events within a single frame that are spatially adjacent (<16.5 km), representing the center of a lightning optical footprint. However, through experimental comparisons, we found that using a single centroid (group) to represent lightning location is not always accurate. As shown in Fig. 2, when compared with ground-based lightning location network data (e.g., BLNET) and radar composite reflectivity, the full spatial distribution of events actually better represents the lightning optical footprint than the group centroid, and the overall geolocation deviation of LMI lightning is more clearly visible. Furthermore, using groups would introduce additional residual errors associated with the clustering algorithm itself. Therefore, we chose events (the most fundamental detection unit, a single pixel in a single frame) as the matching basis.

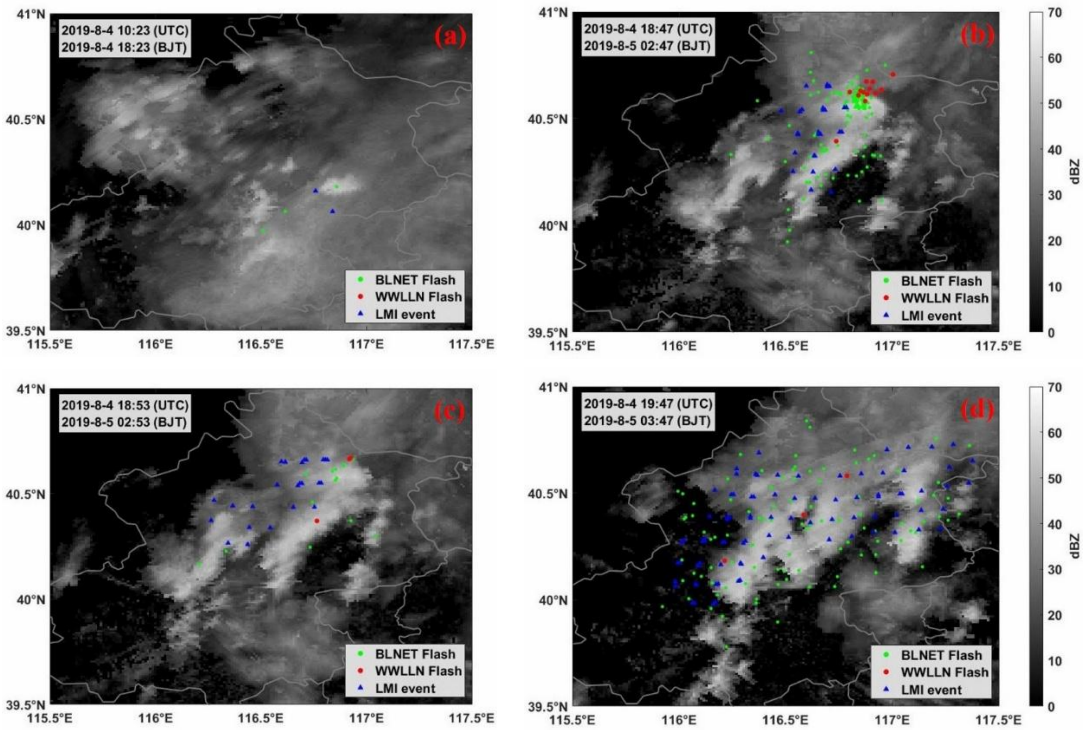

Figure 2: LMI event deviation at different stages of a severe convective event over Beijing on 4 August 2019.

Regarding the matching window: a bright optical pulse composed of many pixels will produce multiple LMI events within a single frame. In our procedure, each LMI event is matched independently to WWLLN strokes within the 3 s temporal window and 30 km spatial window. A single WWLLN stroke may be matched to multiple LMI events if they fall within the thresholds. We do not enforce one-to-one correspondence; instead, our approach aligns the spatial "footprints" of the two datasets.

Events outside the [3 s, 30 km] window are excluded from the fitting analysis. However, based on the sensitivity analysis, the marginal gain beyond these thresholds is very small, and including distant events would increase noise and false associations without significantly improving the curve fitting. As long as a sufficient number of valid matches are retained within the window, the discarded events do not adversely affect the robustness of the fitted correction curves. We will clarify these points in the revised manuscript.

Comment 8:

Please mark the four subregions shown in Fig. 4 on Fig. 3

Response:

We thank the reviewer for this constructive suggestion. In the revised manuscript, we will add a reference schematic diagram (as shown in the new Fig. 3) that indicates the locations of the four subregions presented in Fig. 4 on the full subregion division map (original Fig. 3). This

will help readers quickly identify the spatial positions of the selected subregions. We will update the figure caption and the relevant text accordingly.

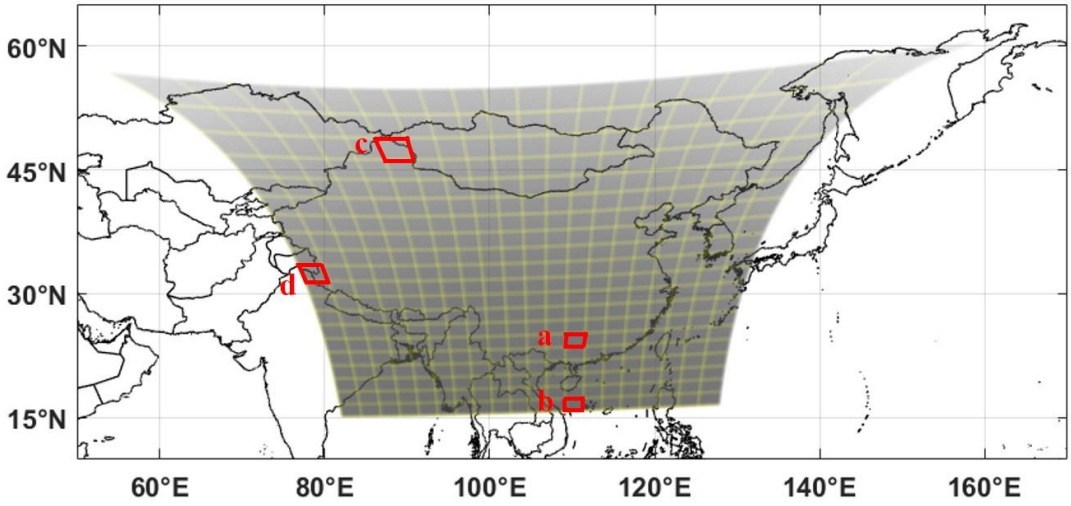

Figure 3 Schematic diagram of the locations of the four study case subregions.

Comment 9:

Line 243: I don't see an east-to-west propagation of the error pattern. To me it looks more like a steady amplification of regions that always have a relative maximum of errors.

Response:

We thank the reviewer for the keen observation. The reviewer's point is very accurate. The wording "east-to-west propagation" in our original manuscript is indeed not precise and may be misleading. In fact, the thermal deformation influence on the LMI CCD array exhibits a pattern of gradual expansion or amplification from the southeast toward the entire domain, rather than horizontal propagation from east to west. Therefore, we will replace "propagation" with "expansion" to more accurately describe the spatial evolution of the error pattern. We will correct this wording in the revised manuscript and adjust the related text accordingly. We appreciate the reviewer's valuable comment.

Comment 10:

Line 289: Typically, more samples should lead to improved location, but the conclusion here is that a low number of samples leads to better fits. Why? Does this suggest a shortcoming in some part of the fitting approach?

Response:

We thank the reviewer for this insightful question. The observed phenomenon is indeed interesting and deserves clarification.

In lightning-dense regions (e.g., areas dominated by large mesoscale convective systems), lightning activity is widespread over hundreds of kilometers, and different convective cells may overlap in space and time. Under our relatively broad matching thresholds (3 s, 30 km),

such widespread activity inevitably introduces mismatches from non-corresponding convective cells into the matching pool, leading to increased noise and larger standard deviations in the deviation estimates. Consequently, the fitted curves may appear less smooth. In contrast, in lightning-sparse regions (e.g., Xinjiang, Mongolia), lightning events are typically isolated, associated with weak, localized convection, and spatially concentrated. Under the same broad matching thresholds, the risk of mismatching across different convective cells is much lower. The matched events are therefore highly consistent in space, resulting in smaller standard deviations and visually "better" fits (i.e., smoother curves). However, this does not mean that the fit is more reliable. The sample size is very small, and the statistical representativeness and extrapolation capability of the fitted curve are actually poor.

Therefore, the statement in our original manuscript that "regions with fewer lightning events yield better fits" needs to be revised, as it is not accurate. A smoother fitted curve in low-sample regions merely indicates local smoothness, not greater overall reliability. True reliability still depends on sample size. We will clarify this distinction in the revised manuscript to avoid misunderstanding. Meanwhile, for regions with extremely limited sample sizes (e.g., Xinjiang and Mongolia), we ultimately choose not to apply the empirical correction to avoid the risk of insufficient data.

Comment 11:

Line 291: I'm not convinced that the large errors in the NE are due to the large ground footprint of the detector elements. If that were the case, the NW part of the domain should exhibit the same pattern.

Response:

We thank the reviewer for pointing out this logical inconsistency. Our original statement attributing the large errors in the northeastern part solely to the large ground footprint of the detector elements is indeed insufficient.

First, it should be noted that the left half of the LMI CCD array experienced a permanent mechanical offset due to the large acceleration during rocket launch, breaking the original geometric symmetry. Consequently, the northwest and northeast regions are no longer mirror images of each other, and a simple pixel-footprint comparison between the two regions is not valid.

Second, regarding lightning activity characteristics: in the northwest (e.g., Xinjiang), lightning activity is sparse, convection is weak, and lightning occurrences are relatively concentrated, resulting in fewer matched LMI – WWLLN events but lower matching noise. In contrast, the northeast exhibits relatively frequent lightning activity with active convection,

and lightning is widely distributed and scattered. Under our broad matching thresholds, this widespread activity tends to introduce more mismatches, leading to higher matching noise. It is important to emphasize that this does not mean the geolocation performance in the northeast is necessarily better than that in the northwest. On the contrary, the larger sample size in the northeast makes its statistical results more stable and reliable, whereas the northwest suffers from high statistical uncertainty due to insufficient samples. Furthermore, the standard deviation of the residuals in the northeast is indeed consistent with the increasing ground footprint of LMI pixels from ~7.8 km at nadir to >20 km toward the edge of the field of view, indicating that pixel projection geometry does contribute to the larger errors.
We will add the above more detailed analysis in the revised manuscript. We appreciate the reviewer's careful reading and constructive criticism.

Comment 12:
Line 300: While I don't object to distribution of a CSV file format, which is easily understood and machine readable (if lacking in terms of metadata), it is not true that NetCDF files will be larger. Naively reading the 20190406.csv dataset (49.2 MB) with pandas and writing it back to NetCDF with xarray was a bit smaller (48.7 MB), even naively using 64 bit ints and floats for all columns. However, the file can easily be made much smaller. Year, month, day, hour, minute and second can all be stored in much fewer than 64 bits, and packing the floats as either 16 or 32 bit floats, or as ints using scale_factor and add_offset, will save further space. Finally, internal zlib compression can also be used on NetCDF files and will be transparently decoded by the library. Finally, the authors might also consider using the CF metadata standard to encode time as a single variable.

Response:
We thank the reviewer for the very professional and detailed suggestions. The following improvements to the dataset:

Time format: The original separate columns for year, month, day, hour, minute, and second have been consolidated into a single ISO 8601 string column (e.g., 2019-03-31 00:00:01.512), with millisecond precision.

Numeric precision: Longitude, latitude, and radiance values are now stored with two decimal places as 32-bit floats. Cloud top height is stored as an integer (16-bit, in meters), which significantly reduces file size.

File formats: Both CSV and NetCDF formats are provided. The NetCDF files use internal zlib

compression (level 9) and appropriate data type reduction (e.g., 32-bit integers with scale/offset for time, 16-bit floats or 32-bit floats for geolocation). The resulting NetCDF files are approximately 60% smaller than the original CSV files, while preserving complete metadata and readability. CSV remains available as a lightweight, human-readable option for quick inspection and simple processing.

Metadata: The NetCDF files comply with the Climate and Forecast (CF) metadata conventions, including variable long names, units, and coordinate reference information, ensuring discoverability and interoperability.

We have updated the data availability section and the attribute descriptions in Appendix A (Table A1) accordingly. We greatly appreciate the reviewer's constructive and expert suggestions, which have substantially improved the usability and reusability of the dataset..