The 3rd AGILE Terrestrial Gamma Ray Flash Catalog. Part I: Association to Lightning Sferics

A. Lindanger1, M. Marisaldi1, M. Maiorana1, D. Sarria1, K. Albrechtsen1, N. Østgaard1, M. Galli1, A. Ursi4, C. Labanti2, M. Tavani4, C. Pittori5,6, and F. Verrecchia5,6

1Birkeland Centre for Space Science, Department of Physics and Technology, University of Bergen, Bergen, Norway, 2INAF-OAS Bologna, Bologna, Italy, 3ENEA, via Martiri di Monte Sole 4, Bologna, Italy, 4INAF-IAPS Roma, Rome, Italy, 5Space Science Data Center - Agenzia Spaziale Italiana, Roma, Italy, 6INAF - Osservatorio Astronomico di Roma, Monte Porzio Catone, Roma, Italy

Abstract We present a complete and systematic search for terrestrial gamma-ray flashes (TGFs), detected by AGILE, that are associated with radio sferics detected by the World Wide Lightning Location Network (WWLLN) in the period February 2009 to September 2018. The search algorithms and characteristics of these new TGFs will be presented and discussed. The number of WWLLN identified (WI) TGFs shows that previous TGF selection criteria needs to be reviewed as they do not identify all the WI TGFs in the data set. In this analysis we confirm with an independent data set that WI TGFs tend to have shorter time duration than TGFs without a WWLLN match. TGFs occurs more often on coastal and ocean regions compared to the distribution of lightning activity. Several multipulse TGFs were identified and their WWLLN match are always associated with the last gamma-ray pulse. We also present the first Terrestrial Electron Beam detected by AGILE. This data set together with the TGF sample identified by selection criteria (companion paper Maiorana et al., 2020) constitute the 3rd AGILE TGF catalog.

Part I: Association to Lightning Sferics

1. Introduction

Terrestrial gamma-ray flashes (TGFs) are sub-millisecond bursts of energetic photons produced in the Earth’s atmosphere and are associated with lightning flashes (Dwyer, 2012). They were first observed by the BATSE instrument onboard the Compton Gamma-ray Observatory in 1991 and were reported by Fishman et al. (1994). Since then, TGFs have been detected by RHESSI (Albrechtsen et al., 2019; Grefenstette et al., 2009; Gjesteland et al., 2012; Østgaard et al., 2015; Smith et al., 2005), Fermi (Briggs et al., 2013; Roberts et al., 2018), AGILE (Maiorana et al., 2010, 2014, 2015), and BeppoSAX (Ursi et al., 2017). The newest addition to space missions capable of detecting TGFs is the Atmospheric-Space Interactions Monitor (ASIM) mission, which became operational in June 2018 (Neubert et al., 2019).

In the first era of TGF detection, gamma-ray data from satellites provided the main insight into the physics of TGFs. However, to expand our knowledge supporting data from ground-based stations are necessary. As TGFs are associated with lightning flashes, the geolocation of the lightning associated to the TGF provides the source location of the TGF. To perform spectral analysis of a TGF, the production location is vital to correctly model the propagation of photons through the atmosphere reaching the satellite.

There are two main methods to identify TGFs in gamma-ray data. The gamma-ray data can be filtered by search algorithms, using some selection criteria (SC) to find the TGFs. These SC are not trivial to decide and are a trade-off between a clean sample of fewer but bright TGFs with low contamination of false TGFs, and a large sample containing more faint TGFs with risk for contamination from false TGFs. After the TGFs are identified with SC, they can be correlated with lightning flashes detected by ground stations. The other method is to start with lightning measurements and look at gamma-ray data detected by the satellite at the time of lightning. This method is able to identify weaker TGFs but is limited by the efficiency of the lightning network. The last method has previously been performed on RHESSI (Albrechtsen et al., 2019; Østgaard et al., 2015; Smith et al., 2016) and Fermi data (McTague et al., 2015), using ground-based lightning data.

In 2015, the AGILE TGF detection rate increased with one order of magnitude after the anticoincidence (AC) shield acting on the onboard mini-calorimeter (MCAL) was disabled (Marisaldi et al., 2015).
This has led to an intensive search for TGFs in the AGILE data, to further populate the AGILE TGF catalog.

For the first time, a complete and systematic search for TGFs based on time correlation with the World Wide Lightning Location Network (WWLLN) has been applied to all available MCAL data up to September 2018. The search for TGFs based only on WWLLN data has no bias from SC, other than the WWLLN efficiency itself. As the WWLLN detection efficiency is limited (Abarca et al., 2010; Bürgesser, 2017; Rudlosky & Shea, 2013), not all TGFs can be identified in the satellite data using time correlation with WWLLN data. Therefore, a WWLLN identified (WI) TGF sample can make the basis for deciding appropriate new SC to identify TGFs without associated WWLLN detections. These new SC are discussed in the companion paper by Maiorana et al. (2020), hereafter referred to as M20.

In the following sections, we present the data sets and the methods used in this analysis, the results from the different data periods of AGILE, a discussion of the findings, and a summary and conclusions section at the end.

2. Data Sets

The MCAL onboard AGILE includes 30 independent self-triggering CsI(Tl) scintillator bars with an energy range ∼0.35–100 MeV (Labanti et al., 2009). One should note that MCAL is triggered; for example, data are stored in the internal memory and sent to ground only if a significant excess of counts is detected in the specific time window and does not run in continuous data acquisition mode. Therefore, only a fraction of the measured data is downloaded to telemetry. Additional information on the instrument performance is included in M20.

We divide the MCAL data into four different subsets characterized by the absolute timing accuracy and if the AC shield is active on the instrument or not. A summary of the data sets is shown in Table 1.

The AC-ON data set spans from 28 February 2009 to 23 March 2015. In this period, the AC shield is active on MCAL and strongly limit the detection of brief duration events (Marisaldi et al., 2014) such as TGFs. The onboard absolute timing accuracy is on ∼2 μs level.

The reference (REF) data set spans from 23 March 2015 to 30 June 2015. The AC shield is disabled for MCAL from the onset of this period, resulting in an enhanced TGF detection rate (Marisaldi et al., 2015).

The DRIFT data set spans from 1 July 2015 to 31 December 2017. An issue with the onboard GPS caused a degradation of the AGILE μs timing performance and the absolute time accuracy started “drifting”. This drift is a systematic offset in time that remains constant for periods between days and weeks, and then suddenly “jumps” to a different value. The AGILE team performed a time correction of the offset using housekeeping data, resulting in an absolute timing accuracy of several tens of milliseconds.

The 3D-FIX data set spans from 17 January 2018 to 30 September 2018. After the DRIFT period, the GPS partly recovered and the AGILE team is able to perform a time correction down to ∼2 μs level, thus restoring the original timing accuracy of the instrument.

Lightning data are obtained from the WWLLN (Rodger et al., 2009) and provide time and geolocation by detecting very low frequency sferics produced by lightning flashes. Although many different sensors and lightning detection networks have been used in TGF studies (Marshall et al., 2013), WWLLN has become a standard choice and a benchmark following the work by Connaughton et al. (2010, 2013).

The timing uncertainty of WWLLN depends on the location uncertainty. Abarca et al. (2010) found an average location error of 4.03 km in the northward direction and an error of 4.98 km in the westward direction.

| Name     | Date                        | AC shield | Timing accuracy |
|----------|-----------------------------|-----------|-----------------|
| AC-ON    | 28 February 2009 to 23 March 2015 | On        | ∼2 μs           |
| REF      | 23 March 2015 to 30 June 2015 | Off       | ∼2 μs           |
| DRIFT    | 01 July 2015 to 31 December 2017 | Off    | Several tens of ms |
| 3D-FIX   | 17 January 2018 to 30 September 2018 | Off        | ∼2 μs           |
Hutchins et al. (2012) report that in 2011, WWLLN located 61% of the strokes within 5 km. Østgaard et al. (2013) assumed a time and location accuracy of 45 μs and 15 km, respectively. In this work, we assume a location uncertainty of 15 km as well.

3. Method
In this paper, two methods are used to search for TGFs. Stacking analysis, and the so-called “search for clusters analysis.” Both methods are based on time correlation with WWLLN detections.

3.1. Stacking Analysis
For every WWLLN detection within 1,000 km from the subsatellite point, a MCAL photon list is built according to equation (1), which is the time difference between the WWLLN detection and the list of counts detected by MCAL, corrected for propagation time. These photon lists are then superposed and binned. The motivation for this is to identify the few photons associated to lightning that would be indistinguishable from background without WWLLN association.

\[ \delta t = \text{time}_{\text{MCAL}} - \text{time}_{\text{propagation}} - \text{time}_{\text{WWLLN}}. \]  

The 1,000-km limit, which we will call the TGF field of view (FOV), is selected because most TGFs are detected within ~500 km from the subsatellite point, and very few farther away than 800 km, as it will be shown in section 4, consistent with earlier results (Collier et al., 2011; Cummer et al., 2005; Marisaldi et al., 2019). The propagation time of the photons is calculated assuming a production altitude of 15 km, which is a fair assumption based on modeling results from Dwyer and Smith (2005) and the expectation that TGFs detected from space are preferentially produced close to thunderstorm cloud tops.

3.2. Search for Clusters Analysis
As stacking analysis relies on timing accuracy on few hundred of microseconds or better to be effective, a new “search for clusters analysis” (SFC) was developed and implemented due to the decreased timing accuracy in the DRIFT period. The method was then also applied to the REF and 3D-FIX periods. The idea is to look at light curves and only keep histogram bars with large counts per 100 μs, assuming that a TGF is seen as an increased flux of counts, or a “cluster of photons” in a light curve. This is possible as the relative timing accuracy of MCAL count timestamps is still at microseconds level and the absolute timing offset is on tens of milliseconds level. The signal is defined to be a time window close to the WWLLN detection, and the background signal is defined to be some seconds before the WWLLN detection, well separated from the signal. The algorithm is described below and in Figure 1.

1. Build a light curve of MCAL photon data for each WWLLN detection for both signal and background time intervals.
2. Build the distribution of number of counts per 100 μs bin for each light curve.
3. Superpose the distributions of counts per 100 μs bin found in Point 2.
4. Set a threshold based on signal vs. background.
5. Define events with counts per 100 μs bin higher than the selected threshold as TGF candidates.

Based on the distribution of counts per 100 μs in Figure 1c, we select a threshold of x counts per 100 μs. The fraction of false positive TGFs is estimated by taking the integral from the selected threshold to 20 for background and signal and divide the background integral by the signal integral. The threshold is a trade-off between the number of TGF candidates and the fraction of false positive TGFs. Note that a plot equivalent to Figure 1c is shown with real data in Figures 3c and 5c, while a plot equivalent to Figure 1d is shown with real data in Figure 4c.

3.3. TGF Candidates Processing
After the TGF candidates are identified, each TGF is fitted by a Gaussian function on top of a constant background by means of an unbinned maximum likelihood technique. The intensity, peak time, and duration of each TGF are extracted from the parameters of the fit. The TGF duration is calculated as

\[ t_{50} = 1.349\sigma, \]

where \( \sigma \) is the standard deviation from the Gaussian fit, and \( t_{50} \) is the central time interval including 50% of the counts. This method is described in Marisaldi et al. (2015).
Figure 1. Illustration of the search for clusters analysis. (a) A single light curve associated with a WWLLN detection. (b) Distribution of number of counts per time bin for the single light curve. (c) Superposed distributions of number of counts per time bin associated to WWLLN detections in black and superposed distribution of background in red. A threshold is indicated by a dotted line. (d) The time difference between events with counts per time bin over the selected threshold, and WWLLN detections, corrected for propagation time.

If a TGF is associated with several WWLLN detections within the $\delta t$ range of the light curve, the algorithm will report the same TGF several times, depending on the number of associated WWLLN detections. We remove the duplicates by keeping the entry corresponding to the closest WWLLN detection in time.

We also determine if the TGFs occurs over land, coast, or ocean. Like the First Fermi-GBM TGF catalog (Roberts et al., 2018), we use the pre-calculated distance to shore file (https://www.soest.hawaii.edu/wessel/gshhg/) that is provided from the Global Self-consistent, Hierarchical, High-resolution Geography Database (Wessel & Smith, 1996). We define the coast as ±150 km from the shoreline.

4. Results

In this section, the results of the analysis for the different data sets are presented. The number of TGF candidates and the expected fraction of spurious signal, incorrectly identified as TGFs, are shown in Table 2. The fraction of incorrectly identified TGFs is estimated by taking the integral from the selected threshold to 20 for background and signal and divide the background integral by the signal integral. The threshold is 4 for REF and 5 for 3D-FIX.

| Name  | #TGFs | Fraction of incorrectly identified TGFs |
|-------|-------|----------------------------------------|
| AC-ON | 0     | n/a                                    |
| REF   | 111   | 0.03                                   |
| DRIFT | 310–1,294 | n/a                             |
| 3D-FIX | 171  | 0.01                                  |

Table 2: Number of TGFs Associated to WWLLN and Expected Background Contamination.

Abbreviations: TGFs, terrestrial gamma-ray ashes; WWLLN, World Wide Lightning Location Network.
4.1. AC-ON Period

A stack plot of the MCAL data for 440,234 WWLLN detections is shown in Figure 2a. There is no significant peak visible at $\delta t \approx 0$ confirming that the AC shield is suppressing the TGF signal as suggested in Marisaldi et al. (2014).

As most TGF-WWLLN matches occur within few hundred kilometers from the subsatellite point, a new stack plot, shown in Figure 2b, was calculated considering only WWLLN detections within 300 km from the subsatellite point and MCAL counts below 100 MeV. The distance and energy limit enhances the signal-to-noise...
The SFC analysis is not applied to the AC-ON data set, as the AC-shield suppresses any clusters associated to WWLLN detections.

4.2. REF Period

Figure 3a shows the stack plot for 9754 WWLLN detections in the REF period. The significance of the peak is \( \sim 27 \) standard deviations and shows a clear time correlation between counts in MCAL and WWLLN detections. If we remove known TGFs identified by Marisaldi et al. (2015), we obtain a 16 sigma peak showing that not all TGF-WWLLN matches are found by the SC described in Marisaldi et al. (2015).

The energy of the counts in Figure 3a is shown in Figure 3b. The peak at \( \delta t \approx 0 \) indicates the energy range of the photons associated with the TGFs. Note that the energy spectrum is not corrected for background and instrumental effects.

The SFC analysis is applied to the data to identify the TGFs. Based on the distribution of counts per 100 \( \mu s \) bin shown in Figure 3c, we select a threshold of 4 counts per 100 \( \mu s \). This identifies a total of 111 WI TGFs. The expected contamination of false positive TGFs is 3%. If we exclude the 111 WI TGFs from the stack plot in Figure 3a, we obtain no peak at \( \delta t \approx 0 \). Thus, the SFC analysis identifies all significant TGF-WWLLN matches in the REF period data set.

Figure 3d shows the distance between the subsatellite point and the associated WWLLN detection for the WI TGFs in black, and all WWLLN detections in the TGF FOV in red. The distance bin size is chosen so that the ratio as TGFs far from the subsatellite point are weaker due to atmospheric attenuation, and counts above 100 MeV are expected mostly to be cosmic rays. Note that a peak of four standard deviations is obtained at \( \delta t \approx -0.5 \) ms. This could be due to the detection of the first counts associated to TGFs before the AC shield suppresses the counts. However, this peak is farther from \( \delta t = 0 \) than expected; therefore, we cannot draw firm conclusions on it.

The SFC analysis is not applied to the AC-ON data set, as the AC-shield suppresses any clusters associated to WWLLN detections.
Figure 5. 3D-FIX period. (a) Stackplot of MCAL counts for 42,243 WWLLN detections within 1,000 km from subsatellite point. (b) Energy of the counts in (a) detected by MCAL in the full energy band 0.35–100 MeV. The unit of the color scale is counts per 100 µs per energy bin. (c) Distribution of counts per 100 µs time bin for the 3D-FIX period. The signal includes counts with $|\delta t| \leq 500$ µs and is normalized per 1 ms per 42,243 WWLLN detection. The background includes counts with $|\delta t|$ between $-2.7$ and $-2.1$ s and is normalized per 0.6 s per 47,966 WWLLN detection. (d) Number of TGFs, as a function of distance from the subsatellite point to WWLLN detection, and all the WWLLN detections in the TGF FOV. Each distribution is normalized to 1.

circular area, corresponding to each bin, is constant and equal to 125,664 km$^2$. In agreement with Cummer et al. (2005), Collier et al. (2011), and Marisaldi et al. (2019), most TGFs are detected within $\sim 500$ km from the subsatellite point, and very few TGF are detected farther away than 800 km. The red WWLLN distribution is not flat due to the nonconstant latitude distribution of the WWLLN detections and AGILE’s orbital inclination angle.

4.3. DRIFT Period

Figure 4a shows the stack plot for the DRIFT period with $\delta t$ range of ±100 ms and bin size of 1 ms. As the absolute time accuracy is not on microsecond level, data do not show a peak at $\delta t \approx 0$. The SFC analysis is applied in order to improve the signal-to-noise ratio, and all events with $\delta t$ within ±100 ms and a threshold of 7 counts per 100 µs is defined as TGF candidates. We choose 7 counts following the same approach as REF and 3D-FIX but the threshold needed to be higher due to the decreased absolute timing accuracy. This identifies 1,294 WI TGF candidates shown in Figures 4b, 4c, and 4d in black color. However, there is evidence of a flat component in Figure 4b in the WI TGFs. TGFs with distance longer than $\sim 500$ km from the subsatellite point can be real TGFs, but we do not expect many of them. The flat component indicates the contamination of false events and suggests that at least 33% of the WI TGF candidates are false TGFs.

To further enhance the signal-to-noise ratio, we compared the WI TGF candidates from the SFC analysis with the TGFs identified with SC in M20. There are 310 joint TGFs (selection criteria satisfied and WWLLN identified, SW hereafter) and these are shown in green color in Figures 4b, 4c, and 4d. (See figure 2 in the companion paper M20 for a graphical overview of SC, SW, and WI TGFs.) The SW TGFs further enhance the signal-to-noise ratio as seen in Figure 4b, where the number of TGF-WWLLN matches close to the subsatellite point are relatively higher for the SW TGFs, and lower for distances far from the subsatellite point, compared to the WI TGF candidates. Also in Figure 4c, we see a higher peak and lower background component for the SW TGFs, compared to the WI TGF candidates.
From Figure 4c, we see that the absolute timing accuracy in the DRIFT period is $\sim 30$ ms, with an offset of approximately $-17$ ms. However, the offset varies with time. That is evident in Figure 4d, which shows the time difference between the TGF and the WWLLN detection, corrected for propagation time, as a function of time between March 2016 and May 2017. “Contact” is the orbit number counting from the first orbit of AGILE, where one orbit is $\sim 94$ min. Considering only the SW TGFs, we see that between contact 46,000 and 48,000, the timing uncertainty seems to be approximately $\pm 25$ ms, and between contact 49,200 and 49,800, there seems to be a linear negative slope, indicating a constantly decreasing offset. After contact $\sim 51,000$ the absolute timing uncertainty is approximately $\pm 100$ ms.

### 4.4. 3D-FIX Period

Figure 5 is the result from the same analysis as for the REF period, applied to the 3D-FIX period. These plots show the same characteristics as in Figure 3. Figure 5a has a peak of $\sim 20$ standard deviations at $\delta t \approx 0$. Based on the signal-background ratio in Figure 5c, we selected a threshold of 5 counts per 100 $\mu$s and identified a total of 171 TGFs. The expected contamination of false TGFs is 1%. If we exclude the 171 WI TGFs from the stack plot in Figure 5a, we obtain no peak at $\delta t \approx 0$, indicating that the SFC analysis identifies all the WI TGFs in this data set, also. The energy peak in Figure 5b is less bright compared to Figure 3b in the REF period because the noise has slightly increased. This is the reason why we cannot set the threshold to 4 as in the REF period. The differences between the REF and 3D-FIX periods in terms of significance of the TGF peak, energies, and TGF detection rate can be explained by the combination of seasonal variability (different data span) and increased instrumental noise.

### 5. Discussion

The stacking analysis for the REF period reveals that the selection criteria in Marisaldi et al. (2015) needs to be reviewed as the SC did not identify all the WI TGF found by the SFC analysis. This is discussed further in the companion paper by M20.

Due to the large timing uncertainty in the DRIFT period, we are dependent on selection criteria to identify the TGFs. Therefore, we exclude the DRIFT period from the discussion, except in section 5.1, to prevent introducing a bias due to selection criteria and to keep the sample purely based on WI TGFs. In the following sections, we discuss the absolute timing accuracy of AGILE, the duration of WI TGFs, multipulse TGFs, local time and geographical distributions, and the first Terrestrial Electron Beam (TEB) detected by AGILE.
5.1. Diagnostics of AGILE Timing Accuracy by TGF-WWLLN Correlation

After the issue with the onboard GPS at the start of the DRIFT period, the AGILE team performed a time correction procedure using housekeeping data. The SFC analysis provides an independent assessment of the goodness and effectiveness of this time correction. Due to the low absolute timing uncertainties of WWLLN, it is possible to use TGF-WWLLN correlations to correct the onboard satellite clock. This has previously been done on RHESSI data (Mezentsev et al., 2016). The TGF-WWLLN correlations evidenced previously unidentified trends in the DRIFT period, impossible to identify by housekeeping data measurements (Figures 4c and 4d). In the 3D-FIX period, the authors of this paper identified a systematic time offset of 4 ms, constant with time. This time offset correction is already implemented in the gamma-ray data processing pipeline by the AGILE team, prior to the complete data analysis resulting in this paper. In this regard, checking AGILE data against WWLLN data provides a valuable, independent diagnostics of the AGILE timing accuracy.

5.2. Duration of TGFs and the Rate of Association with WWLLN

In Figure 6a, we see that the $t_{50}$ distribution of WI TGFs peaks in the range 20–40 μs. This is consistent with the predictions by Connaughton et al. (2013) showing that a TGF will produce a radio signal with a peak spectral energy density at 10 kHz (which is similar to lightning, and where WWLLN is optimized for), when $t_{50} = 21.5$ μs.

The $t_{50}$ distribution of the TGFs identified with the SC in M20, and the $t_{50}$ distribution of the TGFs in SC that are not WI are also shown. These samples are biased towards longer duration with respect to the WI sample. Figure 6b shows the fraction SW/SC for AGILE TGFs and Fermi-GBM TGFs (from figure 3 in Connaughton et al., 2013). It is hard to do a quantitative comparison of the distributions as the instruments are different, the orbital inclination are different, as well as the data span and the efficiency of WWLLN over different geographical regions. However, the trend from the two missions are compatible, where brief duration TGFs have a higher fraction of WWLLN matches.

5.3. Multipulse TGFs

Mezentsev et al. (2016) observed 16 multipulse TGFs out of 314 TGF-WWLLN matches in RHESSI data. In these 16 multipulse TGFs, the WWLLN detection is always associated with the last TGF peak. Motivated by this finding, we manually checked the 284 TGFs in the REF- and 3D-FIX period looking for multipulse TGFs. We identified seven multipulse TGFs shown in Figures 7a to 7g. The first (a) multipulse TGF is already reported in Mezentsev et al. (2016). TGF (b) has a small time separation between pulses, suggesting that it might not be a multipulse event. However, the two candidate pulses exhibit different spectral characteristics. The second pulse is much softer than the first pulse suggesting spectral evolution with time. The third multipulse TGF (c) has a rather weak first peak. The other multipulse TGFs (d–g) are clearly multipulse TGFs with ~0.5 ms between each peak. The multipulse TGFs observed by AGILE confirms, with an independent data set the findings of Mezentsev et al. (2016), that the WWLLN detection is always associated with the last pulse in a multipulse TGF. We found no multipulse TGFs with WWLLN associated with the first pulse. Note that sometimes WWLLN detect the same lightning flash twice, as seen in Figure 7e. The distance between the locations of the “double” lightning detections are within the uncertainties of WWLLN suggesting that the stroke responsible for the two detections is the same.

5.4. Local Time and Geographical Distribution

Figure 8a shows the local time distribution of AGILE WI TGFs compared with TGFs with a WWLLN association from the first Fermi-GBM TGF catalog (Roberts et al., 2018). The position of Fermi is restricted to the same latitude band as AGILE, giving a total of 142 Fermi-GBM TGFs with a WWLLN match. Figure 8b shows the longitudinal distributions of AGILE WI TGFs and Fermi-GBM TGFs with a WWLLN match. As in Figure 8a, Fermi is restricted to ±2.5° latitude. The distributions show a consistent behavior.

Like in Albrechtsen et al. (2019), we investigated whether the TGFs are located over land, ocean, or coast. We define the coast as ±150 km from the shoreline, and the simultaneous WWLLN matches are used to estimate the production origin of the TGFs. In Table 3, the number and percentage of TGFs detected over ocean, coast and land are shown. We also calculated the number and percentage of WWLLN detections within the TGF FOV in the REF and 3D-FIX period. This parameter basically accounts for both satellite exposure time and WWLLN efficiency over the different regions. Note that the WWLLN percentage distribution is not uniform like in Albrechtsen et al. (2019) where WWLLN detections below RHESSI is approximately
Figure 7. Light curve and energy vs. time scatter plot for the seven multipulse TGFs found in the REF and 3D-FIX periods. The bin size is 50 μs. The WWLLN detections, corrected for propagation time, is indicated as a green vertical line.
one third each over ocean, coast, and land. This may be explained by the much large orbital inclination of RHESSI compared to AGILE. It is evident that TGFs detected by AGILE, like both Fermi-GBM and RHESSI (Albrechtsen et al., 2019; Roberts et al., 2018), are mostly detected over coastal regions. If TGFs follow exactly the lightning distribution, the probability of having a TGF at the coast would be 0.51. If we consider whether a TGF is produced at the coast or not as a binomial process with probability of success 0.51, the probability of having 184 successes out of 282 trials is in the order of $10^{-7}$. Doing the same calculation for land and ocean gives a probability of $\sim 10^{-5}$ and 0.02, respectively. This shows that the ocean, land, coast distribution of TGFs does not follow the ocean, land, coast distribution of lightning detected by WWLLN. TGF production occurs relatively more often $\pm 150$ km from the coastline. This is also evident in Figure 9, where the AGILE WI TGFs, and TGFs with WWLLN matches from Fermi-GBM, are plotted together with the LIS 0.1 Degree Very High Resolution Gridded Lightning Full Climatology (VHRFC) dataset. The VHRFC dataset is gridded lightning rate density from the Lightning Imaging Sensor (LIS) from 1998 to 2013 (Albrecht et al., 2016). The color bar indicates flash rate density with unit flashes per square km per year. Fermi-GBM is not restricted to $\pm 2.5^\circ$ latitude in this figure. The three maps corresponds to each of the three lightning chimneys in Figure 8b. Figure 9a shows mainly Colombia and Ecuador, Figure 9b shows West and Central Africa, and Figure 9c shows the Borneo-Sumatra regions. The lack of TGFs south of Colombia is due to the South Atlantic Anomaly. We see that the TGFs follow the lightning activity, but clusters more on the coast than over ocean and land. Where there is high lightning activity on land, like in the Congo basin, more TGFs are observed. These results contribute to the discussion of the geographical asymmetry in the TGF/lightning ratio, already addressed in Smith et al. (2010), Fuschino et al. (2011), Briggs et al. (2013), and recently discussed in Fabró et al. (2019) specially concerning the physical characteristics of thunderstorms over Africa.

Table 3

|                      | #TGFs | %TGFs | #WWLLN   | %WWLLN  |
|----------------------|-------|-------|----------|---------|
| Ocean                | 49    | 17.4% | 483 498  | 21%     |
| Coast                | 184   | 65.2% | 1 155 716| 51%     |
| Land                 | 49    | 17.4% | 634 793  | 28%     |
| Total                | 282   | 100%  | 2 274 007| 100%    |
5.5. Terrestrial Electron Beam

High-energy photons of the TGF will interact with the atmosphere and produce secondary electrons and positrons. A fraction of the secondary electrons and positrons produced above 30–40 km can reach high enough altitudes (above ~120 km) where they stop interacting significantly with the atmosphere (Sarria et al., 2015). The geomagnetic field will then guide the motion of the electrons and positrons so that they stay together, forming a TEB. The TEBs were first described in Dwyer et al. (2008). Spacecraft detecting TGFs, such as BATSE, RHESSI, Fermi, BeppoSAX and ASIM, also detect TEBs (Dwyer et al., 2008; Roberts et al., 2018; Sarria et al., 2019; Ursi et al., 2017).

Here we present the first high confidence TEB identified in AGILE data. The TEB took place over the Indian ocean on 6 April 2018 20:51:50.404601 UTC. There is no lightning activity detected by WWLLN within 1 s of the TEB in a radius of 1,000 km from the subsatellite point, except two WWLLN matches close to the southern footpoint of the magnetic field line that intercepts AGILE at the moment of observation, ~733 km from the subsatellite point. We pick the closest WWLLN match to the magnetic footpoint as the most likely sferic associated with the TGF producing the TEB.

In Figure 10, lightning activity ±1 minute around the time of the TEB, the location of AGILE’s subsatellite point, and the magnetic field line and footpoints are shown. The southern magnetic footpoint is located at ~8.60° latitude and 68.80° longitude. The subsatellite point of AGILE is located at ~2.04° latitude and 68.08° longitude and AGILE has an altitude of 462.6 km. In Figure 10a, it is evident that there is no WWLLN detections directly below AGILE, but there is a cluster of WWLLN detections at the southern magnetic footpoint. The WWLLN match associated with the TEB are found in this cluster. The WWLLN match occurs 0.087 ms before the start of TEB, corrected for propagation time, assuming a straight line.
Figure 10. Two maps of the TEB event. (a) Map of WWLLN detections ± 1 minute around the TEB, AGILE’s subsatellite point, and the magnetic field line and footpoints. (b) Zoomed version of the southern magnetic footpoint. The 15-km uncertainty of the WWLLN match is indicated as a green circle.

Figure 11 shows the detected energy spectra and time histogram of the TEB, together with a simulation of the TEB. The TEB is simulated with a Geant4 based code assuming a TGF with a RREA energy spectrum proportional to $1/E \cdot \exp(-E/7.3 \text{ MeV})$, up to 40 MeV. The Geant4 code is the same as in Sarria et al. (2019). The source TGF is initiated at the WWLLN match with an assumed production altitude of 15 km. The photons are beamed upward with a Gaussian distribution with $\sigma_\theta = 30^\circ$. The Earth’s magnetic field is obtained using the IGRF-12 model (Thébault et al., 2015) and the atmosphere composition is obtained using the NRLMSISE-00 model (Picone et al., 2002). Geant4 includes all the relevant processes physical processes of photon, electron and positron transportation (Compton scattering, pair production, Rayleigh scattering, photoelectric absorption electrons inelastic and elastic scatterings, bremsstrahlung, position annihilation). The photons, electrons, and positrons reaching satellite altitude are saved including their energy and momentum information. The output from this simulation is used as input to the AGILE mass model, which simulates...
AGILE’s detector response to the particles, in terms of energy spectrum and arrival time. The energy spectrum accounts for energy resolution of the instrument. The proper attitude of the satellite at the TEB time is accounted for. This is shown in blue color in Figure 11. To evaluate the compatibility between the simulation and the measurement, we performed a $\chi^2$ test. The reduced $\chi^2$ value is 1.45 for the energy spectrum and the critical value for compatibility is 1.69 (14 degrees of freedom). The reduced $\chi^2$ value is 1.56 for the time histogram and the critical value for compatibility is 1.94 (8 degrees of freedom and the start of the TEB is a free parameter.). Figure 11 shows the resulting energy spectrum (a) and light curves (b), and the results of the $\chi^2$ tests. In both cases, the simulation is compatible with the measurement.

Given the observed WWLLN detection and the magnetic field line configuration, as well as consistency of simulated and measured spectra and time profile, we conclude that the event on 6 April 2018 20:51:50.404601 UTC is the first observation of a TEB detected by AGILE. The detection of TEBs is difficult for AGILE given the amount of screening material surrounding the detector, as pointed out in the comparative study by Sarria et al. (2017).

6. Summary and Conclusions

In this paper, we analyzed more than 9 years of AGILE gamma-ray data, searching for TGFs correlating with lightning strokes detected by WWLLN. We confirm that the AC shield suppresses the TGF signal in the AC-ON period, as suggested in Marisaldi et al. (2014). We have also selected a total of 282 WWLLN identified (WI) TGFs with high absolute timing accuracy (REF and 3D-FIX period), and 310 WI TGFs with low absolute timing accuracy, which also is satisfied by selection criteria (DRIFT period).

The search for clusters (SFC) method proved successful in identifying all significant TGFs with a WWLLN match where the absolute timing accuracy of AGILE is high. These WI TGFs provided a basis for improvements in selection criteria discussed in the companion paper M20. In the DRIFT period, when AGILE experienced absolute timing issues, the SFC method together with the selection criteria provided a very useful TGF-WWLLN data set capable of independent diagnostics of the AGILE timing accuracy. In the 3D-FIX period, the authors also identified a constant systematic offset of 4 ms.

The analysis of the WI TGF sample can be summarized as follows:

1. The duration of TGFs and the rate of association with WWLLN is assessed and shows, in agreement with Connaughton et al. (2013), that brief duration TGFs have a higher fraction of WWLLN matches than longer duration TGFs.
2. Seven multipulse TGFs detected by AGILE confirms the findings of Mezentsev et al. (2016) that a WWLLN detection associated with a multipulse TGF is always associated with the last pulse. No counterexamples were found.
3. The local time and geographical longitude distributions of WI TGFs, detected by AGILE, is consistent with TGFs with a WWLLN match detected by Fermi GBM.
4. The ocean, land, coast distribution of TGFs does not follow the ocean, land, coast distribution of lightning detected by WWLLN. TGF production occurs relatively more often ±150 km from the coastline.
5. The first TEB detected by AGILE is identified with a WWLLN detection close to the magnetic footprint of the satellite. Measured and simulated energy spectra and time profiles are consistent.

A catalog of the TGFs, including the TEB, from the REF, DRIFT, and 3D-FIX period are available online for the scientific community (www.ssdc.asi.it/mcal3tgfcat).

References

Abarca, S. F., Corbosiero, K. L., & Galanneau Jr., T. J. (2010). An evaluation of the Worldwide Lightning Location Network (WWLLN) using the national lightning detection network (NLDN) as ground truth. *Journal of Geophysical Research, 115*, D18206. https://doi.org/10.1002/2013JD019622

Albrecht, R. I., Goodman, S., Buechler, D., Blakeslee, R. J., & Christian, H. (2016). L1S 0.1 degree very high resolution gridded lightning full climatology (vhrfc) [vhrfc_lis_frd] dataset available online from the NASA Global Hydrology Center DAAC, Huntsville, Alabama, U.S.A. https://ghrc.nasa.gov/hydro/details/lavhrfc

Albrechtsen, K. H., Østgaard, N., Berge, N., & Gjesteland, T. (2019). Observationally weak TGFs in the RHESSI data. *Journal of Geophysical Research: Atmospheres*, 124, 287–298. https://doi.org/10.1029/2018JD029272

Briggs, M. S., Xiong, S., Connaughton, V., Tierney, D., Fitzpatrick, G., Foley, S., et al. (2013). Terrestrial gamma-ray flashes in the fermi era: Improved observations and analysis methods. *Journal of Geophysical Research: Space Physics, 118*, 3805–3830. https://doi.org/10.1002/2013JA019125
Sarria, D., Kochkin, P., Østgaard, N., Lehtinen, N., Mezentsev, A., Marisaldi, M., et al. (2019). The first terrestrial electron beam observed by the atmosphere-space interactions monitor. *Journal of Geophysical Research: Space Physics*, 124, 10,497–10,511. https://doi.org/10.1029/2019JA027071

Sarria, D., Lebrun, F., Blelly, P.-L., Chipaux, R., Laurent, P., Sauvaud, J.-A., et al. (2017). Taranis XGRE and IDEE detection capability of terrestrial gamma-ray flashes and associated electron beams. *Geoscientific Instrumentation, Methods and Data Systems*, 6(2), 239–256. https://doi.org/10.5194/gi-6-239-2017

Smith, D. M., Buzbee, P., Kelley, N. A., Infanger, A., Holzworth, R. H., & Dwyer, J. R. (2016). The rarity of terrestrial gamma-ray flashes: 2. RHESSI stacking analysis. *Journal of Geophysical Research: Atmospheres*, 121, 11,382–11,404. https://doi.org/10.1002/2016JD025395

Smith, D. M., Hazelton, B. J., Grefenstette, B. W., Dwyer, J. R., Holzworth, R. H., & Lay, E. H. (2010). Terrestrial gamma ray flashes correlated to storm phase and tropopause height. *Journal of Geophysical Research*, 115, A00E49. https://doi.org/10.1029/2009JA014853

Smith, D. M., Lopez, L. I., Lin, R. P., & Barrington-Leigh, C. (2005). Terrestrial gamma-ray flashes observed up to 20 MeV. *Science*, 307, 1085–1088. https://doi.org/10.1126/science.1107466

Thébault, E., Finlay, C. C., Beggan, C. D., Aubert, J., Barrois, O., et al. (2015). International geomagnetic reference field: The 12th generation. *Earth, Planets and Space*, 67(1), 79. https://doi.org/10.1186/s40623-015-0228-9

Ursi, A., Guidorzi, C., Marisaldi, M., Sarria, D., & Frontera, F. (2017). Terrestrial gamma-ray flashes in the BeppoSAX data archive. *Journal of Atmospheric and Solar-Terrestrial Physics*, 156, 50–56. https://doi.org/10.1016/j.jastp.2017.02.014

Wessel, P., & Smith, WalterH. F. (1996). A global, self-consistent, hierarchical, high-resolution shoreline database. *Journal of Geophysical Research*, 101(B4), 8741–8743. https://doi.org/10.1029/96JB00104