Title: RELAMPAGO - Electric Field Mill (EFM) Data Documentation

Version 1.0, October 12, 2019

Authors:
PI: Wiebke Deierling
Co-PI: Robert Marshall
André Antunes de Sá
Austin Sousa

Contact Information:
Robert Marshall
3775 Discovery Drive
CCAR, 429 UCB
Boulder, CO 80303
email: robert.marshall@colorado.edu

Wiebke Deierling
3775 Discovery Drive
CCAR, 429 UCB
Boulder, CO 80303
email:wiebke.deierling@colorado.edu

1 Introduction

This document provides a description of the data product associated with the Electric Field Mill (EFM) array deployed from 1 November to 15 December 2018 during the RELAMPAGO field campaign in Argentina. The document also provides a brief description of the EFM instrument, its deployment, and the data processing used to generate the data product.

2 Instrument Description

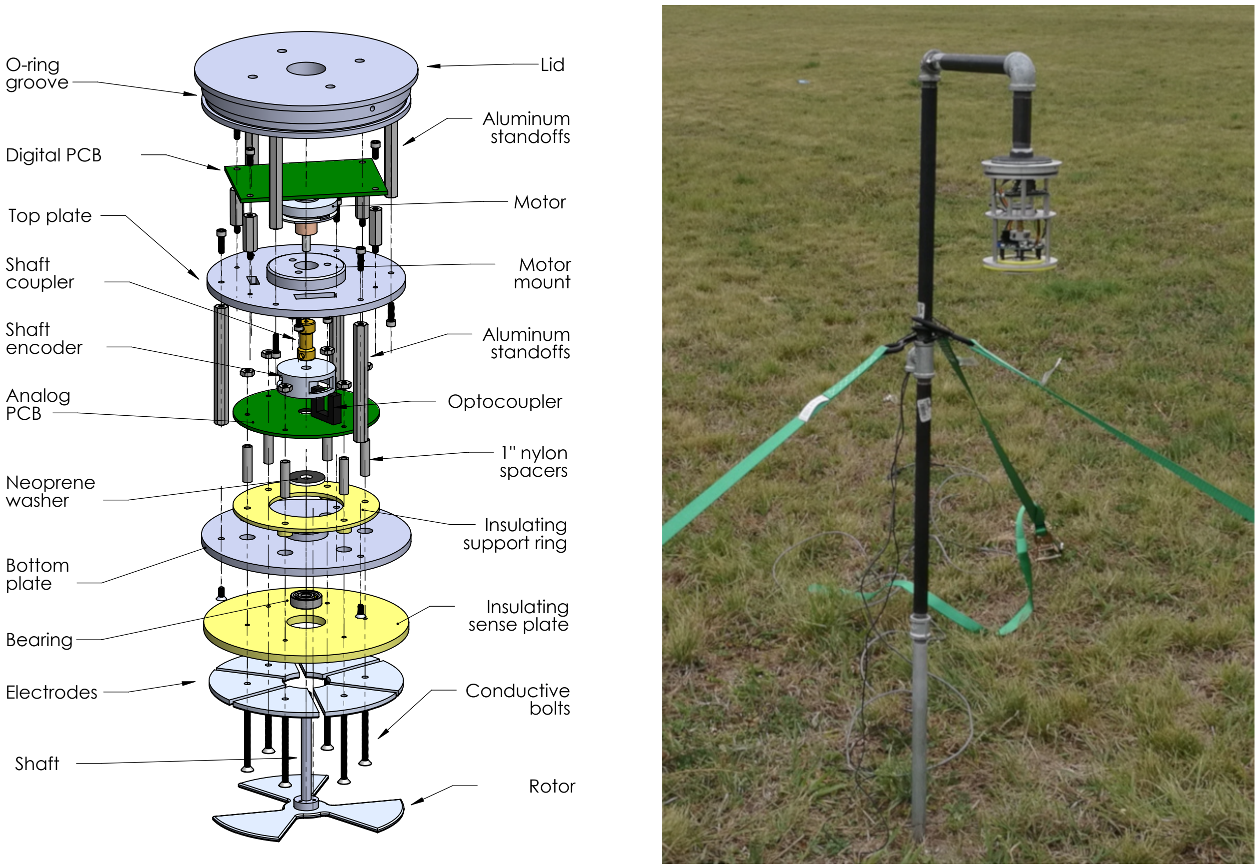

Figure 1: Exploded diagram of the EFM instrument without casing (left), and the instrument as deployed in Pilar, Argentina, also without its casing (right).

An array of eight Electric Field Mills (EFMs) were deployed during the RELAMPAGO field campaign (Fig. 1). These instruments measure the local, ambient in-situ electric field with an effective sample rate around 100 Hz, ranging from a typical 100 V/m field in the absence of weather to tens of kV/m with thunderstorm electrification overhead. Nearby lightning discharges can also be detected from spikes in the ambient electric field.
The EFM instrument operates by moving a grounded shield plate over electrode sense plates, alternately exposing the sense plates to the electric field and then shielding them. This alternation is done continuously using a grounded spinning rotor to cover one group of

plates while uncovering another group (Fig. 2). Each electrode in a group of plates is connected to each other but isolated from the plates of the other out-of-phase group. When the sense plates are exposed to the electric field, they collect free electrons according to:

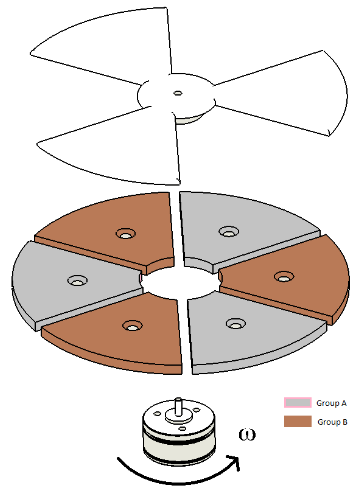

Figure 2: Driving of rotor over plates. The rotor (top) is grounded; the sense plates are connected in sets of three, noted by the two colored groups.

$$q(t) = \epsilon_0 E A(t) \tag{1}$$

where $q(t)$ is the time-varying charge accumulation, E is the background electric field, $A(t)$ is the time-varying exposed area of the plates, and ϵ_0 is the permittivity of free space. For a continuous, uniformly spinning rotor, $A(t)$ can be described by a sinusoid as the electrodes are shielded and exposed. Equation 1 then becomes:

$$q(t) = \epsilon_0 E A_0 \sin(3\omega t) \tag{2}$$

where A_0 is the nominal area of the set of plates, ω is the angular velocity of the rotor as it spins over the electrodes and the factor of three arises from there being two groups of three parallel-connected plates, which are alternately exposed and shielded three times for every rotation of the motor. In order to fulfill the constant speed requirement by the sinusoid description of $A(t)$, the shaft motor's rotational speed is actively managed in software using a PID controller.

The effective current induced in the electrodes can be determined by taking the time derivative of equation 2:

$$I(t) = \epsilon_0 E A_0 3\omega \cos(3\omega t) \tag{3}$$

With negligible effects by design, the voltage measured at each plate, $V(t)$, is proportional to the induced current through a simple resistance parameter, R. The voltage signals of each group of plates are added constructively through an amplifier stage, combining the 180° out-of-phase group A and B signals. The resulting voltage sinusoid is thus directly proportional to the ambient electric field, E as given by:

$$E = \frac{V(t)}{R}\frac{\epsilon_0 A_0 3\omega}{\cos(3\omega t)} \tag{4}$$

Polarity of the ambient electric field is inferred from the instantaneous phase of the measured sinusoid coupled to the corresponding position of the grounding shaft as either grounding electrode group A or group B.

The EFM instrument features a Teensy 3.6 microcontroller daughterboard for primary system control, data acquisition and storage into a micro SD card; and the Adafruit Ultimate GPS module for clock synchronization and data time tagging.

3 Network

The array of EFMs was deployed to cover the Cordoba region centered around the city of Pilar during RELAMPAGO (Fig. 3). Table 1 provides the location information on each of these sites.

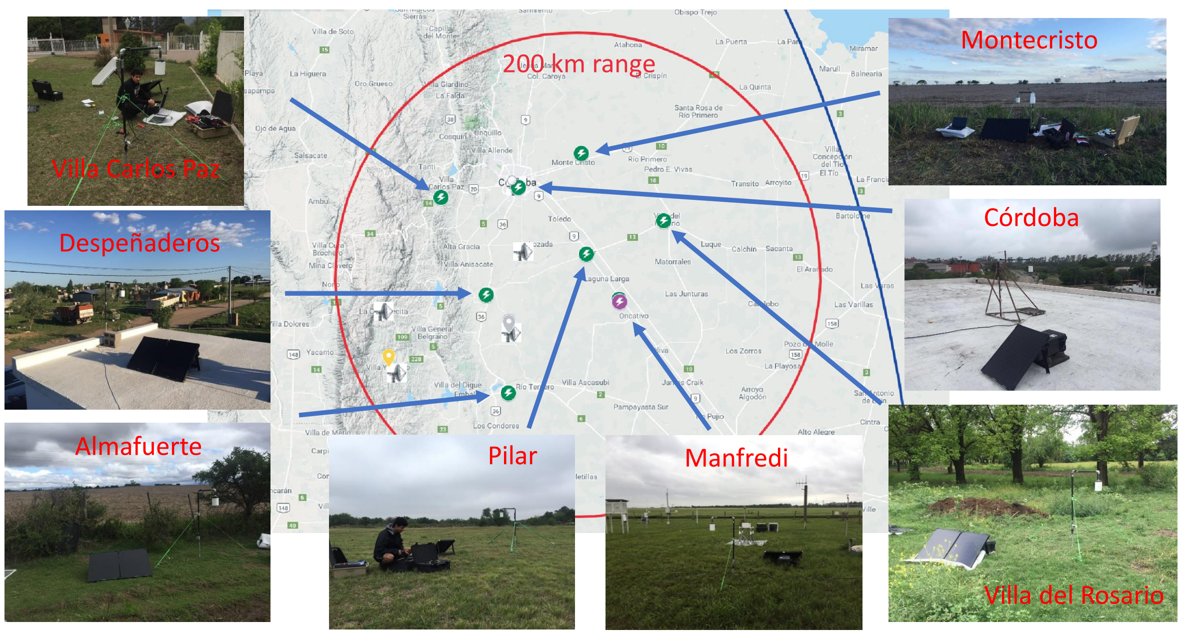

Figure 3: RELAMPAGO site map of the EFM sensors in Argentina.

4 Data Coverage

Once deployed, all instruments were programmed to acquire data continuously throughout the campaign, powered by a battery charged by solar panels or by wall power. Due to extensive power outages, the receivers have periods where no data was captured. The Montecristo, Almafuerte, and Despeñaderos stations had technical issues. No useful data could be recovered from those sites. Fig. 4 below presents the data coverage available from EFMs in RELAMPAGO.

Table 1: Table providing the deployed location of the EFM instruments. Montecristo and Despeñaderos stations are omitted from the table given that no data was recovered from these sites.

City	Ser. #	Latitude (°)	Longitude (°)
Cordoba	EFM011	-31.438438	-64.192988
Manfredi	EFM004	-31.857117	-63.748657
Pilar	EFM006	-31.667211	-63.882683
Vila Carlos Paz	EFM008	-31.475205	-64.526288
Vila del Rosario	EFM002	-31.5588116	-63.565553

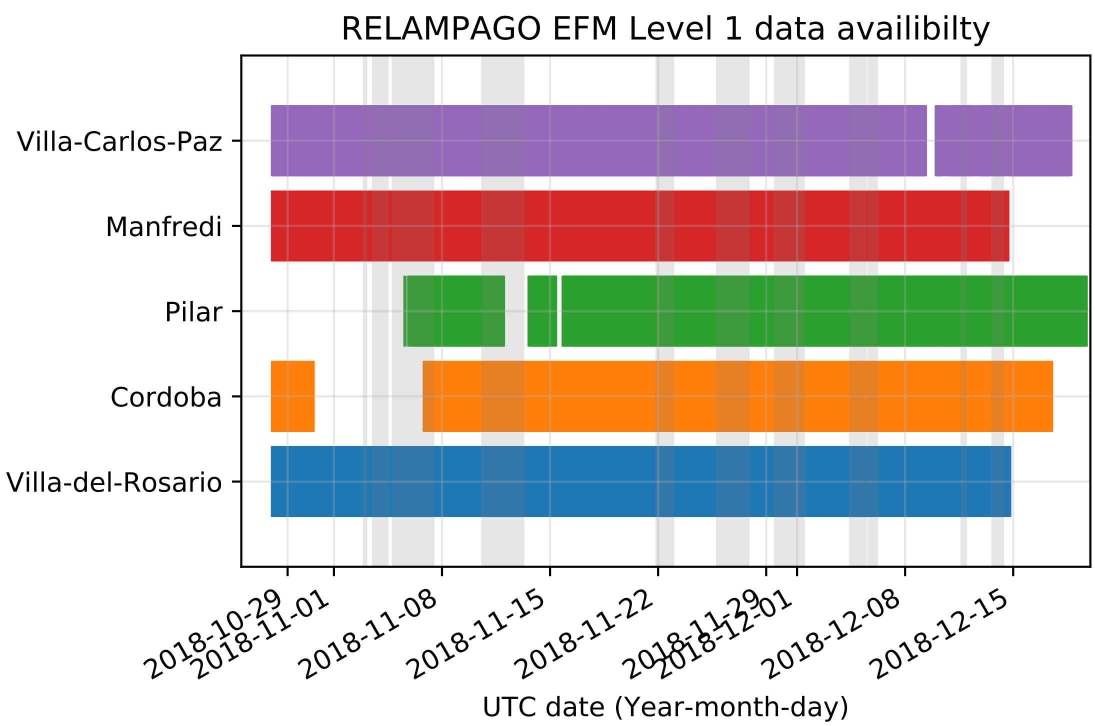

Figure 4: Data coverage during RELAMPAGO for each of the EFMs. Shaded gray regions represent an Intensive Observing Period (IOP) for the campaign

5 Data Processing

From Eq. 4, the ambient electric field can be described as:

$$\begin{aligned} |E| &= \frac{V_0}{R}\epsilon_0 A_0 3\omega, \\ E &= P_\mathrm{f}\frac{V_0}{R}\epsilon_0 A_0 3\omega, \\ E &= K_\mathrm{m}(V_0 P_\mathrm{f})V_0, \end{aligned} \tag{5}$$

where V_0 is the instantaneous voltage amplitude of the measured signal, P_f is a polarity factor that recovers polarity information for the electric field and K_m is a calibration map that translates the instantaneous voltage into a calibrated electric field.
The signal post-processing consists on the extraction of the electrode signal amplitude, V_0, the computation of the E-field polarity, P_f, and the mapping of the post-processed signal into an E-field measurement through a calibration map, K_m.

5.1 Instantaneous Amplitude Extraction The raw signal is bandpass filtered around the expected signal frequency (100±1 Hz), thus removing slow linear trends in the data as well as spurious signals not related to the E-field. Once the data is cleaned, the signal's instantaneous amplitude and phase are extracted using a Hilbert transform.

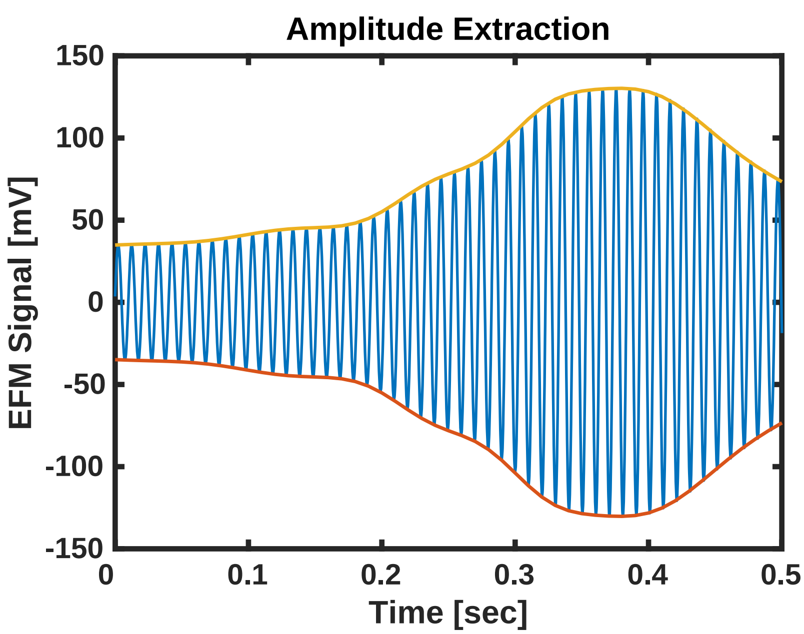

Figure 5: Plot of the digitized voltage readings of the electrode plates, $V(t)$, after bandpass filtering, with the extracted amplitude shown in red and yellow.

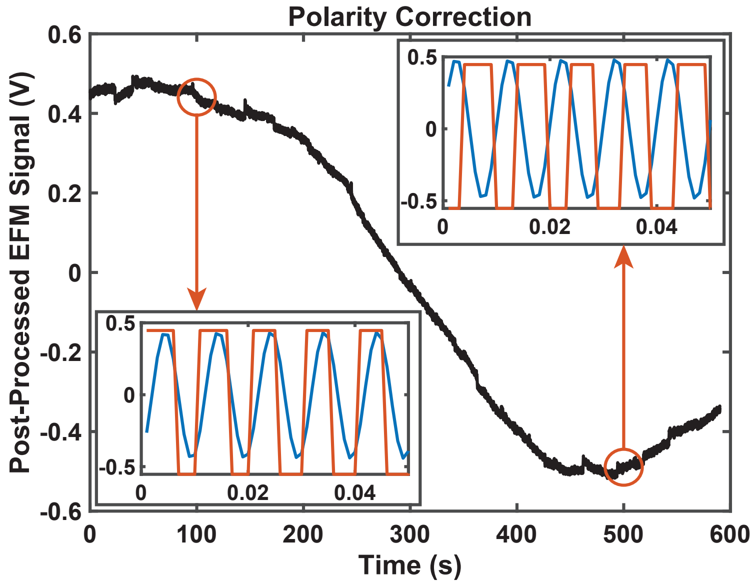

Figure 6: Plot of the post-processed EFM signal after polarity correction but before the calibration map. Two inserts are overlaid, displaying the raw and optical encoder signals under a positive ambient electric field (left), and a negative ambient electric field (right).

5.2 Polarity Computation The sign of the ambient electric field manifests itself as a 180° phase shift of the voltage and current functions with respect to time. The electrode voltage measurement, $V(t)$, by itself, has an absolute phase ambiguity and does not contain enough information for deriving the ambient electric field polarity. In order to solve this ambiguity, a polarity factor, P_{f}, that changes between ± 1 can be derived from the angular position of the rotor, indicating which set of electrodes is covered.

The rotor polarity is measured via an opto-mechanical encoder wheel, which returns a square wave signal corresponding to which set of electrodes are covered. This signal is bandpass filtered and processed through a Hilbert transform in the same manner as the electrode signal. The polarity of the electrode signal is then computed by taking the difference between the phase angles of each signal, which yields approximately 0° for in-phase signals, and approximately 180° for out-of-phase signals. The polarity signal is then thresholded to ± 1, separated by 90°.

The decoded signal is computed by multiplying the electrode amplitude and polarity signals, and downsampling the resulting signal to 100 Hz. Fig. 6 shows a comparison of the post-processed signal along with the electrode and opto-encoder signals.

5.3 Laboratory Calibration Calibration of the EFM is necessary to map the voltage readings on the electrode plates into a meaningful E-field measurements, and also to provide a better understanding of the instrument's linearity and sensitivity not captured in the theoretical electric field measurement equation (Eq. 5).

Calibration curves are recorded using a calibration table, consisting of two parallel metal plates, separated by ~ 30 cm. A hole is cut in the bottom plate, to allow the field mill's sensing electrodes to sit flush, mimicking a ground-installed, upward facing EFM. A controlled electric field is induced by applying a voltage across the two plates.

The electric field induced by the calibration table is first measured using a Campbell Scientific CS-110 electric field mill; A range of measurements are taken for plate voltages between ± 5 kV, which are used as a ground-truth relationship between plate voltage and electric field.
Each EFM is then placed in the calibration table, and the plate voltage swept. Using the previously recorded ground-truth measurements, a calibration map is then computed, which relates the processed EFM signal voltage to the electric field to be measured.
Fig. 7 shows the calibration curves used within the campaign. Each mill is approximately linear, with some nonlinearities arising at very high electric fields. An affine fit to the calibration data reveals a DC offset of $\sim$ 100 V/m, which may drift across the campaign.

5.4 Site Correction Electric field measurements are highly susceptible to environmental changes, such as ground conductivity, instrument location, and mounting apparatus. An ideal measurement of the electric field can be made by placing a field mill in a small hole, facing upward, so the sense plates sit flush with the ground, thus minimizing any impact the sensor itself may have on the electric field. However, such an installation is rarely practical. A site correction, C_s, is applied to the measured E-field if the mill is mounted in a different configuration, e.g., pointing downwards, which can greatly reduce its sensitivity, or moving it up from the ground, which can increase its sensitivity. The site correction is a constant and can be found for any mounted EFM by comparing its measured E-field to a second EFM pointing upwards and flush with the ground. The site correction should stay constant for similar mounting configurations, but can change depending on the site's vegetation and topology.
The field mills used in the RELAMPAGO campaign are mounted from a steel pipe approximately 1 meter above the ground, facing downward to protect against water ingress. The relative gain in electric field measurement is assumed to be linear; we can then compute a site correction C_s for each field mill.

$$E_\mathrm{true} = C_\mathrm{s} \cdot E_\mathrm{cal} + C_\mathrm{offset} \tag{6}$$

Where E_cal is the processed, table-calibrated measurement, and C_s is a site-specific scaling constant, depending primarily on installation height.
At deployment, short data segments were recorded simultaneously, side by side, for each CU EFM with a reference Campbell Scientific CS-110 EFM.
The site correction for the CS-110 EFM was computed after the RELAMPAGO campaign, using data from a CU EFM and CS-110 first side-by-side, and then again with the CS-110 mounted flush with the ground. The correction for the CS-110 mounting position was taken to be $C_\mathrm{s,CS110} = 0.233 \pm 0.03$.
Two site corrections were computed for each EFM – one, using the in-field data at the beginning of the campaign, and again after the campaign in Colorado. Fig. 8 shows the computed site corrections.
The selected C_s value varies for each site, as described below:

- In-field calibration data is used for Pilar and Villa Carlos Paz.
- The post-campaign site correction is used for Villa del Rosario, due to low electric field activity during the installation day.

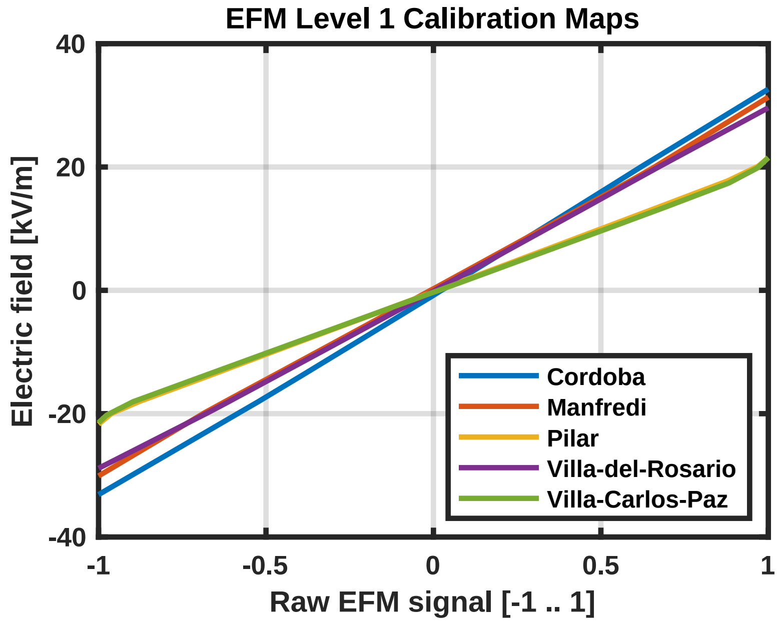

Figure 7: Calibration curves for each site, as computed using the calibration table.

- In-field calibration data is used for Cordoba, despite low activity, as the site correction must account for the rooftop-mounted EFM, with the reference CS-110 on the ground.
- Manfredi uses an average of the in-field and post-campaign site corrections, due to low electric field activity during installation, and an error in the polarity signal.

6 Timing Errors

A timing ambiguity on the order of ~ 45 seconds was discovered in post-campaign processing. The error has been attributed to a firmware fault present on all deployed EFMs, but corrected in EFM011 (Cordoba), which resulted in recurrent skipped data samples. The EFM software keeps an internal index of acquired samples, which is also used as the address for writing any acquired sample into the raw data file. This internal index is incremented once with each sample acquired, expected at 1 kHz. However, due to an interrupt race condition, there is a significant probability that the 1 kHz data sampling interrupt might be skipped, giving rise to missing samples. With each missing sample, the next acquired sample is erroneously placed where the previous sample should have been, i.e., the new sample will be recorded 1 ms into the past.

After testing of the deployed EFMs post-campaign, it was found that the missing samples cause an effective clock lag of a maximum of ~ 70 seconds before the clock is synchronized to a GPS clock, which occurs every 30 minutes. The period between GPS corrections normally starts with the clock lag, initially at zero, monotonically increasing in sporadic steps to a maximum lag, before being corrected by the next GPS update. When the GPS correction occurs, it forces the sample index to skip forward to present time, leaving a trail of unwritten samples for the skipped indices. This trail of unwritten samples are easily

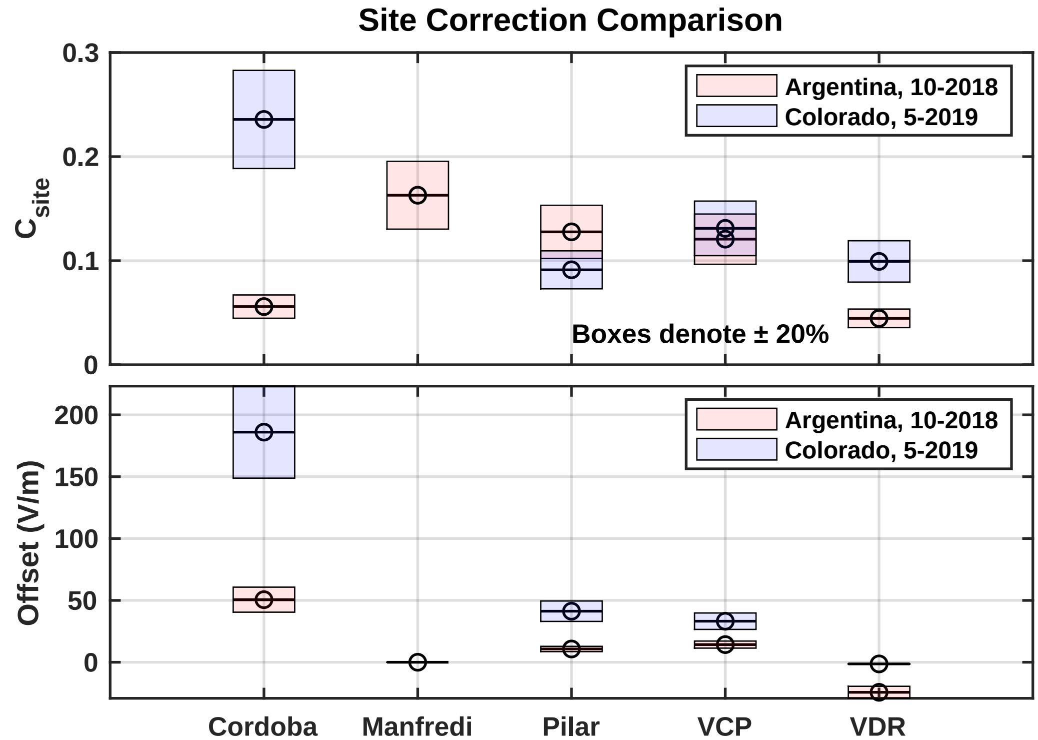

Figure 8: Computed site corrections (C_s) and offsets, for each site, as computed at the beginning of the campaign, and again using the same EFMs in Colorado.

identified in the raw data, and can be used to indicate the maximum clock error expected in the preceding 30-minute period of data.
In the released Level 1 data, the unwritten samples are represented as NaN values, and a maximum clock error measure, derived from the unwritten samples, is presented as a separate variable. Timing errors are at a minimum immediately following a series of unwritten samples, and then accumulated until the next section of unwritten samples, which appear every 30 minutes.
Fig. 9 shows a histogram of the maximum timing error within the dataset.

7 Data Format

Level 1 and Level 2 data are being made available on the EOL RELAMPAGO data webpage.

7.1 Level 1: The Level 1 data product is packaged in NetCDF 4 “Classic Mode” containers. Each hourly NetCDF file follows the name convention:

sitename_YYYY-MM-DDTHH.nc

where *sitename* is the EFM field site sites ("Cordoba", "Manfredi", "Pilar", "Villa-del-Rosario", or "Villa-Carlos-Paz"); and *YYYY*, *MM*, *DD* and *HH* corresponds to the year, month, day and hour, respectively, in UTC, of the data covered in the file.
Each file contains the following data and attributes:

- **E_field**: The calibrated electric field, given in V/m.

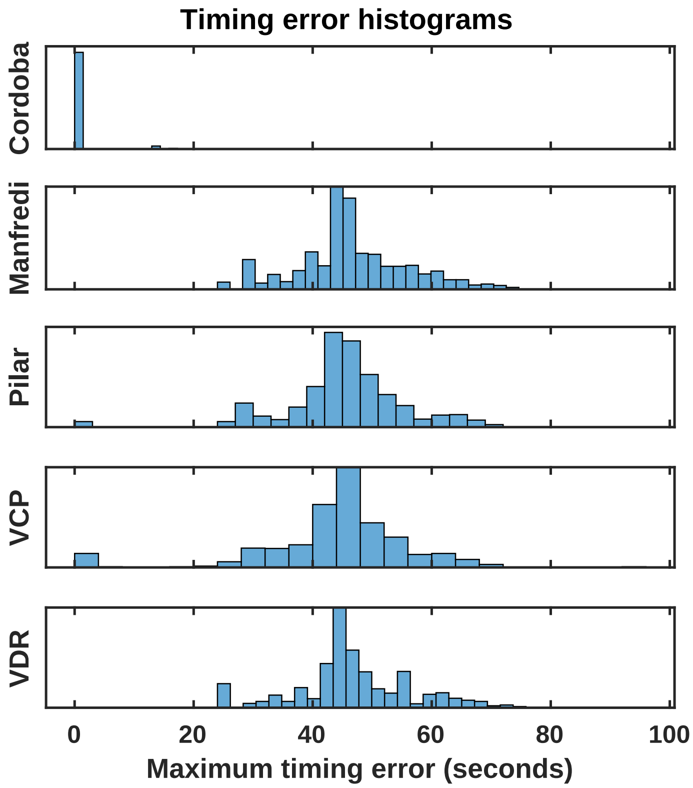

Figure 9: Histograms of the maximum timing error within each 1-hour file, for each site. Distributions are consistent for Manfredi, Pilar, Villa Carlos Paz, and Villa del Rosario, with ~ 45 seconds expected per file; Timing errors are substantially improved at Cordoba.

- **start_time**: The file start time, given as an iso-formatted string, referenced to UTC; *yyyy-mm-ddThh:mm:ss+000*;
- **SAMPLE_RATE**: The processed data cadence; 100 Hz for level 1 data.
- **E_saturation**: The electric field intensity upon which the instrument will saturate.
- **site_gain** and **site_offset**: The calibration gain and offset, corresponding to the local site correction. These parameters have already been applied to the **E_field** data.
- **latitude** and **longitude**: The site latitude and longitude, in decimal-formatted geographic coordinates.
- **max_timing_error**: An estimate of the maximum timing error within the file, derived from the total number of missing samples within the file.

The following instrument-related metadata are also included:

- **creation_date**: The date and time of file processing
- **calibration_file**: The laboratory calibration file used in processing
- **firmware_version**: The version number corresponding to the software onboard the EFM