



Severe convective storms' reproduction: empirical analysis from the marked self-exciting point processes point of view

Giada Lo Galbo¹ · Marcello Chiodi¹

Received: 9 October 2023 / Revised: 19 December 2023 / Accepted: 20 December 2023
© The Author(s) 2024

Abstract

The paper focuses on the evaluation of hailstorms' and thunderstorms winds' events in the United States of America, in the period from 1996 to 2022, under the marked spatio-temporal self-exciting point processes point of view. The aim of the present article is the assessment and description of the spatio-temporal spontaneous and reproducing activity of severe hailstorms' and thunderstorms winds' processes. The present application shows how the spatio-temporal pattern is well-fitted and clearly explainable, according to the flexible semi-parametric ETAS model fitting.

Keywords Declustering method · Epidemic Type-Aftershocks Sequence (ETAS) model · Marked spatio-temporal point process · Semi-parametric modeling approach · Severe convective storms

1 Introduction

Severe storms' events, such as hailstorms and thunderstorms winds, have significant impacts on human activities and infrastructures. Namely, their increasing occurrence in space and in time has a significant impact in terms of economic damages and heavy losses (Kossin et al. 2017). As stated by the Insurance Information Institute (2021), among severe storms' types, hailstorms and thunderstorms winds result to be the most dangerous events, in terms of insurance losses (Gao and Shi 2022), as well as life losses (Kunkel et al. 2013). Hailstorms and thunderstorms winds' events

Handling Editor: Luiz Duczmal.

✉ Giada Lo Galbo
giada.logalbo@unipa.it
Marcello Chiodi
marcello.chiodi@unipa.it

¹ Department of Economics, Business and Statistics, University of Palermo, Viale delle Scienze, 90128 Palermo, PA, Italy

are described as *severe local storms* and are classified as severe convective ones (Brooks 2013). Severe convective storms earn their designation based on the magnitudes of their grasps. A convective storm related to a hailstorm's event is considered severe if the hail diameter extension (National Center for Environmental Information (NCEI) of the National Oceanic and Atmospheric Administration (NOAA) of the United States of America (USA) 2022b) is higher than 2.50 cm \approx 0.98'' (inches), while that related to a gust wind's event is considered severe if its wind speed is higher than 95.00 km/h \approx 0.03 km/s (Kunkel et al. 2013). Both thunderstorms winds' and hailstorms processes are sensible to climate change (Raupach et al. 2021) as their relationship with environmental, meteorological and climatological covariates (Brown and Murphy 1996; Brooks et al. 2003) has been investigated through regression techniques, such as Principal Component Regression (PCR), Partial Least Squares Regression (PLSR—Eccel et al. 2011), Poisson counting regression models (Allen et al. 2015) and intensity estimation (Gensini and Allen 2018).

The spatio-temporal dynamics and the occurrence of severe storms' events are complex and require sophisticated modeling techniques (Sobash et al. 2011; Dehshiri and Firoozabadi 2023; Fortuin et al. 2022), that are not even enough satisfactory to describe their spatio-temporal structural and reproduction characteristics. Spatio-temporal dynamics of hailstorms have recently been evaluated according to replicated spatial point processes' models (Gao and Shi 2022), and that of thunderstorms winds have been evaluated through temporal Poisson renewal processes (Li 2000). Furthermore, extreme environmental hazards' processes involving the most violent gusty winds, i.e., tornadoes, are already evaluated from the spatio-temporal point processes point of view (González et al. 2014, 2019; Valente and Laurini 2020).

Spatio-temporal point processes' models have been used to understand the mechanism of the spatio-temporal evolution of phenomena (González et al. 2016), in specialized fields, such as epidemiology (Quesada et al. 2017) and seismology (Adelfio and Chiodi 2020). However, recently, their application has been developed in other fields such as ecology/biology (Soriano-Redondo et al. 2019), social sciences (Tang et al. 2022), finance (Adelfio et al. 2020). Spatio-temporal point processes allow the description of a point pattern's first-order characteristics, related to the spatio-temporal locations or in asymmetric relationships with the available environmental covariates, and its second-order characteristics. The latter involves the spatio-temporal interactions or interpoint distances. In summary, point processes allow for the recognition of the so-called generating process, i.e. the underlying spatio-temporal point process (Diggle 2006a, b; Illian et al. 2007).

An important issue in the assessment of the dynamics of thunderstorms winds and hailstorms processes relates to their mark variables, i.e. their extension and wind speed, respectively (Gao and Shi 2022). Models for marked spatio-temporal point patterns are suitable for the evaluation of spatio-temporal environmental hazards' processes (Mateu and Ignaccolo 2015), such as for seismological phenomena (Adelfio and Chiodi 2020). Our interest is to evaluate the historical processes of severe hailstorms and thunderstorms winds that occurred in the USA, from 1996 to 2022. More precisely, we aim to evaluate the dynamics of hailstorms and thunderstorms winds using stochastic point processes theory, and according to the marked

spatio-temporal self-exciting point processes approach. As well as many other natural phenomena, thunderstorms winds and hailstorms processes are characterized by an additional random variable indicating the power of occurrence at each spatio-temporal location. This power is measured in terms of the so-called mark, which represents the phenomenon's driving force. Marked self-exciting point processes, as well as log-Gaussian Cox processes for the unmarked case and patterns where the driving force is represented by the random intensity function, are suitable models for assessing the underlying processes of marked clustered patterns. Besides the aggregated nature of the data, they require an additional assumption regarding the real-valued driving random variable, or mark, that makes them *marked*. Each marked self-exciting process has a spatio-temporal distribution (and the associated spatio-temporal conditional intensity function) which depends on the mark distribution. Marked self-exciting point processes assume that the process self-excites with a higher frequency of offspring events conditional to a high value of the mark; i.e. the higher the mark value, the higher the *offspring* productivity, in the immediate nearest neighbourhood, according to some specific spatial and temporal reproduction laws. As well as seismic sequences, both hailstorms' and thunderstorms winds' after-sequences could be produced starting from spatio-temporal decay laws. The subject of the analysis is of great importance and allows for the explanation, as well as the possible forecast, of catastrophic events, damages and hazards resulting from severe convective storms, using the dataset provided by the National Center for Environmental Information (NCEI) of the National Oceanic and Atmospheric Administration (NOAA) of the USA (2022a).

This work aims to understand the spatio-temporal dynamics of severe hailstorms and thunderstorms winds coherently with stochastic processes' theory. We are here interested in the evaluation of thunderstorms winds' and hailstorms' processes in terms of *Windowed ETAS models* (Nicolis et al. 2015), over several temporal (in particular yearly) sub-patterns.

In order to fit and attempt to explain the spatio-temporal data related to the historical severe storms in the USA, according to the marked self-exciting spatio-temporal point processes theory, the formulation of the extended Hawkes (1971)—the Epidemic Type-Aftershocks Sequence (ETAS—Ogata 1988)—model is considered.

The paper is organized as follows: after this brief introduction, Sect. 2 introduces the point process framework related to spatio-temporal point processes and the ETAS model; Sect. 3 shows the case study, the exploratory analysis and the model fitting results; Sect. 4 provides the interpretation of ETAS model fitting; and finally, Sect. 5 provides a brief discussion and some concluding remarks on methodology and future developments.

2 Spatio-temporal point processes and conditional intensity function (CIF)

Let X be a marked spatio-temporal point process, so a random countable subset $X \in \mathbb{R}^2 \times \mathbb{R}^+ \times \mathcal{M}$, where a point $(\mathbf{u}, t, m) \in X$ corresponds to an event at $\mathbf{u} \in \mathbb{R}^2$ occurring at time $t \in \mathbb{R}^+$, with attached mark value $m \in \mathbb{R}^+$. A marked point pattern,

which is a realization of a marked spatio-temporal point process X , is a finite set of distinct points of size $n > 0$, $\{\mathbf{u}_i, t_i, m_i\}_{i=1}^n$, within a bounded spatio-temporal region $W \times T \times \mathcal{M} \subset \mathbb{R}^2 \times \mathbb{R}^+ \times \mathbb{R}^+$, with area $|W| > 0$, length $|T| > 0$ and marking structure $\mathcal{M} \in \mathbb{R}^+$ (Daley et al. 2003). Let N be a counting measure, such that $N(A \times B)$ denotes the number of (unmarked) points of a set $(A \times B) \cap X$, with $A \subseteq W$ and $B \subseteq T$.

First- and second-order properties, representing the point process characteristics associated with the mean and the covariance functions, respectively, are described by the corresponding first, $\lambda(\mathbf{u}, t, m), \forall \{(\mathbf{u}, t, m)\} \in X$, and second-order, $\lambda^{(2)}((\mathbf{u}, t, m), (\mathbf{v}, l, m')), \forall \{(\mathbf{u}, t, m), (\mathbf{u}, l, m')\} \in X$, intensity functions. A refined version (Illian et al. 2007) of the intensity function, $\lambda_\theta(\mathbf{u}, t, m)$, is the so-called conditional intensity function, $\lambda(\mathbf{u}, t, m | \mathcal{F})$, where \mathcal{F} indicates the *filtration*, given by a sequence of sub σ -algebras to which the evaluation of the pattern is subject (each time, whenever a new event occurs). The conditional intensity function is a stochastic process itself and completely characterizes the associated spatio-temporal point pattern (Adelfio and Chiodi 2014; Daley et al. 2003). If the filtration is represented by the past history of the process, including the marks: $\mathcal{H}_t = \{(\mathbf{u}_i, t_i, m_i) : t_i < t\}$, then the conditional intensity function is defined as:

$$\lambda(\mathbf{u}, t, m | \mathcal{H}_t) = J_{\mathcal{M}}(m)\lambda(\mathbf{u}, t | \mathcal{H}_t) \tag{1}$$

where $J_{\mathcal{M}}(m)$ is a probability density function describing the mark distribution; i.e., the mark density distribution function, which describes the expected number of offsprings of an event of magnitude m , where $m \geq m_0$; $\lambda(\mathbf{u}, t | \mathcal{H}_t)$ is the conditional intensity function of the unmarked spatio-temporal point pattern; \mathcal{H}_t is the past history of the marked point pattern (Ogata 1988; Stindl and Chen 2022; Zhuang et al. 2002). No inference about the mark density distribution function is provided in the paper.

Under the assumption of underlying marked self-exciting spatio-temporal point process, the conditional intensity function of the corresponding unmarked spatio-temporal point pattern, $\lambda(\mathbf{u}, t | \mathcal{H}_t)$, is defined as:

$$\lambda(\mathbf{u}, t | \mathcal{H}_t) = \lim_{\Delta \mathbf{u}, \Delta t \rightarrow 0} \frac{\mathbb{E}[N([\mathbf{u}, \mathbf{u} + \Delta \mathbf{u}] \times [t, t + \Delta t]) | \mathcal{H}_t]}{\Delta \mathbf{u} \Delta t} \tag{2}$$

where $\mathbb{E}[\cdot]$ is the expectation operator; $\Delta \mathbf{u}$ and Δt are spatial and temporal infinitesimal increments; \mathcal{H}_t is the spatio-temporal occurrence history of the process up to time t . This is also a conditional expectation, and represents the history-dependent (conditional) probability that an event occurs in the volume $\{[\mathbf{u}, \mathbf{u} + \Delta \mathbf{u}] \times [t, t + \Delta t]\}$, around the location $(\mathbf{u}, t) \in X$ and up to time t (Adelfio and Ogata 2010).

2.1 The Epidemic Type-Aftershocks Sequences (ETAS) model

A branching process is a Markov process in which each individual in the n th generation produces random numbers of individuals in the $(n + 1)$ th generation,

according to a probability distribution function which is repeated from individual to individual. Branching processes are generally used to model reproduction phenomena (Adelfio and Chiodi 2014).

The conditional intensity function of a spatio-temporal branching model corresponds to the sum between a term indicating spontaneous or background activity and a term indicating induced or triggered activity. The triggered component evaluates the interaction activity with the events characterizing the past history of the process, and the parametric conditional intensity function is:

$$\lambda_{\theta}(\mathbf{u}, t | \mathcal{H}_t) = \mu f(\mathbf{u}) + \tau_{\phi}(\mathbf{u}, t) \quad \text{with: } \tau_{\phi}(\mathbf{u}, t) = \sum_{t_j < t} v_{\phi}(\mathbf{u} - \mathbf{u}_j, t - t_j | m_j) \quad (3)$$

where $\theta = (\mu, \phi)$ is the vector of parameters of the background (μ) and the induced (ϕ) intensity, respectively; $f(\mathbf{u})$ is the spatial density; $\tau_{\phi}(\mathbf{u}, t)$ is the conditional triggered intensity to the mark value attached to the corresponding point. Under the assumption of spatio-temporal separability, the latter, $\forall j : t_j < t$, can be further factorized (Stindl and Chen 2023; Molkenhain et al. 2022) as the product between the spatial occurrence rate, $g(\mathbf{u} - \mathbf{u}_j)$, the temporal occurrence rate, $t(t - t_j)$, and a boost function, which expresses the proportionality with the shifted/truncated exponential law associated to the magnitude distribution, $\kappa(m_j)$, i.e.:

$$v_{\phi}(\mathbf{u} - \mathbf{u}_j, t - t_j | m_j) = f(t - t_j)g(\mathbf{u} - \mathbf{u}_j)\kappa(m_j) \quad (4)$$

An Epidemic Type-Aftershocks Sequence (ETAS—Ogata 1988) model, which can be considered as an extension of the Hawkes (1971) model, assumes a branching process of descendants (Adelfio et al. 2006), and is defined according to a parametric conditional intensity function which is expressed as:

$$\lambda_{\theta}(\mathbf{u}, t | \mathcal{H}_t) = \mu f(\mathbf{u}) + \kappa_0 \sum_{t_j < t} \frac{e^{\alpha(m_j - m_0)} \{(\mathbf{u} - \mathbf{u}_j)^2 + d\}^{-q}}{(t - t_j + c)^p} \quad (5)$$

where the triggered component results from the product between three components, which are those listed below:

- the Modified Omori Law (Omori 1895; Utsu 1961), representing the density of aftershocks in time, defining the occurrence rate of aftershocks at time t , following the earthquake of time t_j , which is represented by a non-stationary Poisson process (Ogata 1983), and which is expressed as:

$$f(t - t_j) = \frac{\kappa_0}{(t - t_j + c)^p} \quad \text{with: } t > t_j \quad (6)$$

- where c and p are the temporal displacement and decay parameters; κ_0 is the normalizing constant linked to the aftershock productivity;
- the density of aftershocks in space, associated with the spatio-temporal location of the same event,

$$g(\mathbf{u} - \mathbf{u}_j) = \frac{1}{[(\mathbf{u} - \mathbf{u}_j)^2 + d]^q} \tag{7}$$

where d and q are the spatial displacement and decay parameters, respectively;
 – the boost function, $\kappa(m_j)$, which indicates the mark density distribution function and is proportional to the assumed shifted exponential density $J_{\mathcal{M}}(m) = \beta \exp\{-\beta(m_j - m_0)\}$ in (3), as follows:

$$\kappa(m_j) = \exp\{\alpha(m_j - m_0)\} \propto \beta \exp\{-\beta(m_j - m_0)\} \quad \text{with: } m_0 \leq m_j \tag{8}$$

where: m_j represents the magnitude of the j th event; $\beta = b \log(10)$ is the rate value, which is a function of the b parameter describing the Gutenberg–Richter recurrence law (Gutenberg and Richter 1944—as detailed in Sect. 3.4); m_0 is the completeness magnitude threshold; α describes the expected number of offsprings generated by a single event.

The modified Omori Law in (6) describes the decay of the number of aftershocks after a mainshock event. The decay rates of aftershocks, including the modified Omori Law, are both described by power law decay functions.

The log-likelihood function to be maximized is expressed as:

$$\ell_X(\boldsymbol{\theta}; X) = \log [\mathcal{L}_X(\boldsymbol{\theta}; X)] = \sum_{i=1}^n \log [\lambda_{\theta}(\mathbf{u}_i, t_i | \mathcal{H}_i)] - \int_0^T \int_W \lambda_{\theta}(\mathbf{u}, t | \mathcal{H}_i) d\mathbf{u} dt \tag{9}$$

2.2 Parameters estimation

The set of parameters $\boldsymbol{\theta} = \{\mu, \kappa_0, c, p, \alpha, d, q\}$ is generally estimated through the maximum likelihood (ML) approach (Ogata 1988) or by implementing the expectation-maximization (EM) algorithm to maximize the expected complete data log-likelihood function (Veen and Schoenberg 2008). The background spatial intensity, $f(\cdot)$, is generally estimated through non-parametric methods (Chiodi and Adelfio 2008), involving kernel-based techniques. Namely, the latter can be estimated through a weighted Gaussian Kernel estimator (Sheather 2004; Silverman 1981), whose formulation is:

$$\hat{f}_{\Sigma}(\mathbf{u}) = \frac{\sum_{i=1}^n \rho_i \mathcal{K}(\mathbf{u} - \mathbf{u}_i, \Sigma)}{\sum_{i=1}^n \rho_i} \tag{10}$$

where the weight ρ_i is the estimated probability that the event (\mathbf{u}_i, t_i) belongs to the background component, and is given by:

$$\rho_i = \frac{\hat{\mu} \hat{f}_{\Sigma}(\mathbf{u}_i)}{\hat{\lambda}_{\theta}(\mathbf{u}_i, t_i | \mathcal{H}_i)} \quad (11)$$

where: $\hat{\mu}$ is the estimated temporal background intensity; $\hat{f}_{\Sigma}(\mathbf{u}_i)$ is the Gaussian Kernel estimate of the i th point; $\hat{\lambda}_{\theta}(\mathbf{u}_i, t_i | \mathcal{H}_i)$ is the estimated conditional intensity value of the i th point, all estimated by the ETAS model.

For the evaluation of the bandwidth, the classical Silverman's *rule of thumb* can be used, as well as the cross-validation one proposed in Adelfio and Ogata (2010), or the semi-parametric, simultaneous and alternated Forward Likelihood-based Predictive (FLP) approach, proposed by Chiodi and Adelfio (2017), which maximizes the *predictive information*. For further details about the FLP approach, see Chiodi and Adelfio (2008) and Adelfio et al. (2013).

3 Results

3.1 Data

The *Storm Events Database* (National Center for Environmental Information (NCEI) of the National Oceanic and Atmospheric Administration (NOAA) of the USA 2022a) is an integrated database containing information on severe weather events that occurred in the USA, from 1950 to 2023. For each severe event, a variety of variables are collected, including severe event type, magnitude (severity), magnitude type, spatio-temporal location, start and end time, event azimuth, direct and indirect impacts, such as the cost of damage to property and crops. Events are recorded if they are of sufficient intensity to cause loss of life, injury, significant property damage and/or disruption of commerce; or if they are considered rare and unusual enough to attract media attention; or if they are associated with other meteorological or climatic events; or if they are consistent with concurrently recorded maximum or minimum temperature values or precipitation amounts. Preliminary information and data are included in the Storm Prediction Center (2023) reports, during the first 120 days from the day of occurrence.

The corresponding documentation is prepared and published by the National Weather Service (1999). The recorded severe occurrences cover 40 types of meteorological and climatic events, including severe convective storms and, in particular, thunderstorms winds and hailstorms. The latter are the most frequent and spatio-temporally dynamic (Kelly et al. 1985) storms in the USA's history of natural disasters, and they are the only two types of severe weather events that have their own magnitude accompanying them.

In this application, the listed variables are considered for each severe convective storm event:

- Spatial location, related to longitude and latitude values, recorded in decimal degrees and converted into km;

- Temporal location, related to the exact occurrence's hour and day;
- Magnitude value, i.e., hail extension for hailstorms, recorded and analysed in inches, wind speed for thunderstorms winds, recorded in MPH and analysed in km/s.

The total number of recorded hailstorms' events, from 1996 to 2022, is 288, 126. The total number of recorded thunderstorms winds' events, from 2003 to 2022, is 240, 040. They are yearly analysed because of the lack of knowledge about the generating processes of both severe convective storms, i.e., as a first understanding of the hypothesized dynamics of both processes.

The marked spatio-temporal distributions of hailstorms' (green, Fig. 1a) and thunderstorms winds' (blue, Fig. 1b) selected events, that occurred within the USA territory, are represented in Fig. 1.

In Fig. 1, events are represented according to their spatio-temporal locations. Spatial (longitude and latitude) and temporal locations are represented in km and days, respectively. The time domain corresponds to the temporal distance, in days, from the occurrence of the first event. Each point is represented with size and color proportional to the recorded magnitude value, represented by hail extension in inches (for hailstorms, Fig. 1a) and thunderstorms winds' speed in km/s (for thunderstorms winds, Fig. 1b): the higher the recorded magnitude value, the darker and larger the point, for both spatio-temporal patterns.

Only events belonging to the complete catalogues—according to the completeness magnitude thresholds values, as detailed in Sect. 3.4—are shown, i.e., only hailstorms whose hail extension is greater than 0.8 inches and thunderstorms winds with wind speed greater than 1.35 km/s. The total number of hailstorms and thunderstorms winds above the corresponding magnitude threshold, is 188, 075 and 163, 783, respectively. Both patterns are yearly analysed from the marked self-exciting point processes' point of view. In particular, the occurrence of hailstorms is analysed for each year ranging from 1996 to 2022; the occurrence of thunderstorms

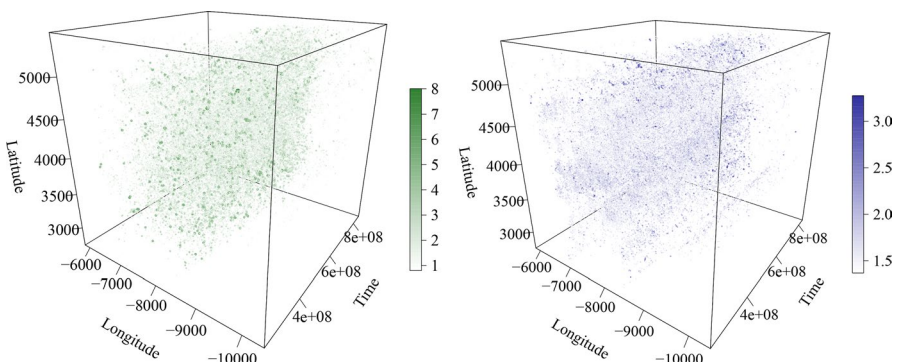
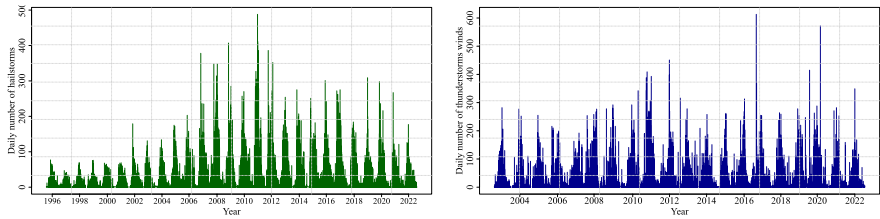
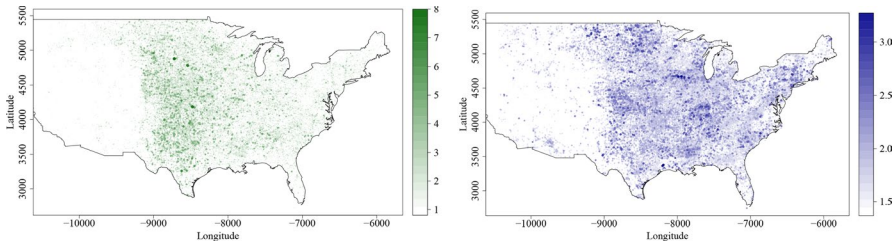


Fig. 1 Marked spatio-temporal distribution of hailstorms' (a) and thunderstorms winds' (b) selected events



(a) Temporal distribution of the total number of daily recorded hailstorms' events (b) Temporal distribution of the total number of daily recorded thunderstorms winds' events

Fig. 2 Temporal distribution of the total number of daily recorded hailstorms' (a) and thunderstorms winds' (b) events



(a) Spatial distribution of wind speed (km/s) (b) Spatial distribution of hail extension (inches)

Fig. 3 Spatial distribution of hail extension (a) and wind speed (b)

winds is yearly analysed for the years that go from 2003 to 2022 (due to the lack of occurrence for the years 1996/2002).

3.2 Exploratory analysis

Figure 2 shows the temporal distribution of the total number of daily recorded hailstorms' (green, Fig. 2a) and thunderstorms winds' (blue, Fig. 2b) events.

As shown in Fig. 2, both patterns follow a seasonal yearly pattern, and a higher number of events is recorded during mid-year. The highest number of severe hailstorms (Fig. 2a) is 488, which occurred on March 21, 2011. More than 400 severe hailstorms are observed on April 9 and 10, 2009, and on May, 21/22 and 24/26, 2011. In 2011, the higher severe hailstorm occurrence with hail extension greater than 0.8 inches are recorded. The highest number of severe thunderstorms winds (Fig. 2b) is 721, which occurred on June 29, 2012. More than 500 severe thunderstorms winds are recorded on 29/30 June, 2012 and on February 28 and March 1, 2017. Finally, at least one severe convective storm is observed within the USA, per type, between January 1st, 2003 and December 31st, 2022.

Figure 3 shows the spatial distributions of the magnitudes, and in particular of the hail extensions, in inches (green, Fig. 3a), and the wind speeds, in km/s (blue, Fig. 3b); the longitude and latitude coordinates' values are shown in km, and the higher the recorded magnitude value, the darker and larger the associated point.

As shown in Fig. 3, both hailstorms' and thunderstorms winds' patterns are mainly detected in the areas from the Central to the Northern-East. As shown in Fig. 3a, events with the highest hail extension, ranging from 5 to 8 inches, are recorded in the Central areas, and other events with higher hail extension are observed in the Northern-East areas. Some aggregations, as well as a dominating general trend towards spatial heterogeneity and sparsity, conditionally to the extension, are observed. As shown in Fig. 3b, thunderstorms winds' events with the highest wind speeds are grouped in well-defined areas, i.e. nearest to each other if with higher wind speed values. Otherwise, thunderstorms winds' pattern shows a dominating general trend towards spatial aggregation, conditionally to wind speed.

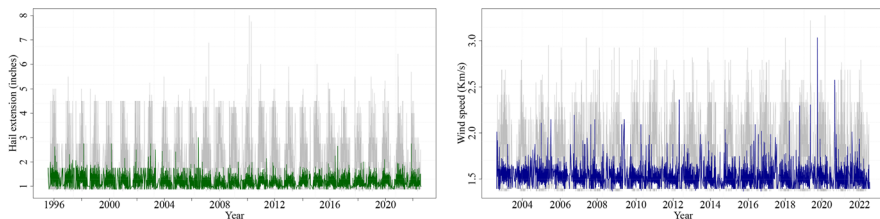
Figure 4 shows the average daily distributions of hail extension (green, Fig. 4a) and wind speed (blue, Fig. 4b) over the completeness magnitude threshold values, with associated daily minimum and maximum values (grey).

As shown in Fig. 4, seasonal yearly and grouping behaviour is observed in both hailstorms' and thunderstorms winds' average magnitude values, conditionally to the magnitude values. For each year, during 1th and 4th quarters and during 2nd and 3th quarters, respectively, the lowest and highest average daily magnitude values are observed. As shown in Fig. 4a, the average daily hail extension mainly ranges between 0.8 and 2, and a whole decreasing trend (from 2007 to 2019), which becomes increasing (from 2020 to 2022) is observed. As shown in Fig. 4a, the average wind speeds mainly range between 1.35 and 2.20, though some peaks over 2.20 km/s are observed in 2003, 2005, from 2009 to 2012 and from 2017 to 2022. A general increasing trend, as well as increasing variability, from 2003 to 2022, is observed.

3.3 ETAS models identification

The aggregation behaviour of the yearly spatio-temporal patterns is evaluated according to the empirical unweighted K -functions (Ripley 1976; Gabriel and Diggle 2009), whose formulation is:

$$K(r, h) = 2\pi \int_0^h \int_0^r \frac{\lambda^{(2)}((\mathbf{u}, t), (\mathbf{v}, l))}{\lambda(\mathbf{u}, t)\lambda(\mathbf{v}, l)} drdh \quad r, h > 0 \quad (12)$$



(a) Average (green), minimum and maximum (grey) daily distributions of hail extensions, in inches (b) Average (blue), minimum and maximum (grey) daily distributions of wind speeds, in km/s

Fig. 4 Average, minimum and maximum daily distribution of hail extension (inches, **a**) and wind speeds (km/s, **b**) patterns

where: $r = \|\mathbf{u} - \mathbf{v}\|$, is the Euclidean distance (or distance in norm L_2) between spatial locations \mathbf{u} and \mathbf{v} ; $h = |t - l|$ is the Manhattan distance (or distance in norm L_1) between temporal locations t and l . Under *Complete Spatial and Temporal Randomness (CSTR)*, the second-order function takes values as that describing the spatio-temporal homogeneous Poisson process: $K(r, h) = 2\pi hr^2$. The unbiased weighted spatio-temporal K -functions' estimator, $\hat{K}_I(r, h)$, is:

$$\hat{K}_I(r, h) = \frac{|W||T|}{n(n-1)} \sum_{(\mathbf{u}, t) \in X} \sum_{(\mathbf{v}, l) \in X \setminus \{(\mathbf{u}, t)\}} \frac{\mathbb{1}[\|\mathbf{u} - \mathbf{v}\| \leq r, |t - l| \leq h]}{\hat{\lambda}(\mathbf{u}, t)\hat{\lambda}(\mathbf{v}, l)w((\mathbf{u}, t), (\mathbf{v}, l))} \tag{13}$$

where: $\hat{\lambda}(\mathbf{u}, t)$ and $\hat{\lambda}(\mathbf{v}, l)$ are the inhomogeneous intensity values evaluated by the model, at points (\mathbf{u}, t) and (\mathbf{v}, l) ; $w((\mathbf{u}, t), (\mathbf{v}, l))$ is a correction weight which adjusts for possible biases due to the spatio-temporal domain definition (Baddeley and Turner 2000; Gabriel et al. 2013); $\mathbb{1}[\dots]$ is the indicator operator, which takes a value of 1, if the statement \dots is true and 0 otherwise. Any positive deviation from *CSTR* indicates spatio-temporal clustering with respect to the spatio-temporal Poisson distribution.

Finally, the chosen correction method is the translation (Ohser 1983) one, which allows amendment with respect to events lying further away from spatial and temporal bounded domains, with respect to those assessed for the event (Gabriel and Diggle 2009).

It is important to highlight that LISTA functions, such as the K and pair correlation functions, are also commonly used for the assessment of point processes' models' fitting. In this case, weighted second-order functions have to be computed under the assumption of *known first-order intensity function*, according to that estimated by the model. Further information on weighted second-order diagnostic methods can be found in Adelfio and Schoenberg (2008), Adelfio and Chiodi (2008) and Adelfio (2007).

Figure 5 shows the unweighted empirical spatio-temporal K -functions calculated for each yearly sub-pattern under translation edge correction, for hailstorms' and thunderstorms winds' patterns. In Fig. 5 green and blue surfaces are the empirical unweighted K -functions of the yearly hailstorms' (green, Fig. 5a) and thunderstorms winds' (blue, Fig. 5b) sub-patterns, respectively, whose intensity function estimates are computed under the assumption of spatio-temporal homogeneity ($\hat{\lambda}(\mathbf{u}, t) = \hat{\lambda}(\mathbf{v}, l) \equiv \hat{\lambda} = \frac{N(W \times T)}{|W||T|}$); black surfaces are the theoretical K -functions, computed under *CSTR* and grey surfaces are upper and lower 95% envelopes obtained with 39 Montecarlo simulations within the spatio-temporal domain where patterns are observed.

As shown in Fig. 5, yearly thunderstorms winds' patterns (Fig. 5a) are the most aggregated and less variable, while hailstorms' patterns (Fig. 5b) are the least aggregated and most heterogeneous, along the years of observation. Once we have proved the aggregation behavior of yearly sub-patterns for each severe convective storm, we will try to evaluate them according to the marked self-exciting point processes' model fitting.

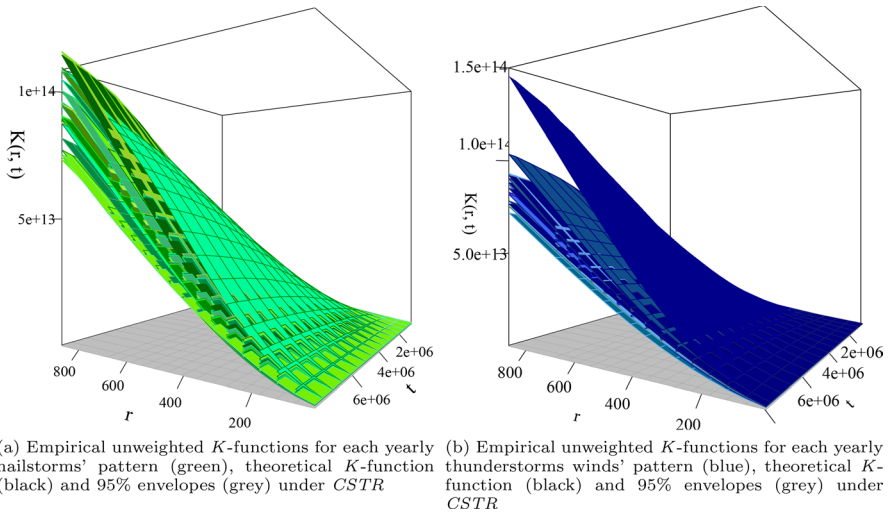


Fig. 5 Empirical unweighted K -functions for each yearly hailstorms' (a) and thunderstorms winds' (b) pattern, theoretical K -function (black) and 95% envelopes (grey) under $CSTR$

The proposed application for the evaluation of severe convective storms' underlying processes is based on a declustering technique for clustered processes. The whole pattern is hypothesized to be a superposition of a spatio-temporal Poisson (background) and a clustered (triggered) point pattern. For the resulting pattern to be declustered, the definition of two intensity functions—the former for the background pattern and the latter for the triggered one—is needed. We decide to assume a self-exciting spatio-temporal underlying generating process, subject to a driving random variable (the magnitude).

3.4 Empirical choice of the magnitude threshold

Based on the ETAS model fitting, an important parameter to be chosen in advance is the completeness magnitude threshold, m_0 , which is assumed to be the minimum value above which all events are included in the catalogue; i.e. the minimum value assumed by the model for the magnitude distribution. The completeness magnitude threshold is linked to the frequency-magnitude distribution (Ishimoto 1939), which for seismic processes is shown to be well approximated by an exponential function, and in particular the Gutenberg–Richter law (Gutenberg and Richter 1944), which is expressed as:

$$\log_{10}(N_m) = a - bm \quad (14)$$

where N_m is the cumulative number of earthquakes with a magnitude value greater than $m \in \mathcal{M}$; $\log_{10}(\cdot)$ is the logarithm with base 10; a and b are constants. The completeness magnitude threshold is the value of the maximum frequency at which events are observed.

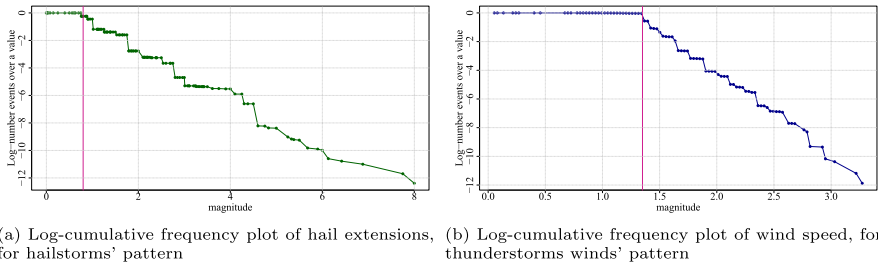


Fig. 6 Log-cumulative frequency plot of hail extension (a) and wind speed (b), with corresponding completeness magnitude threshold's values lines

The log-cumulative frequencies plot, with corresponding completeness magnitude threshold values (in violet) of hailstorms' (in green, Fig. 6a) and thunderstorms winds' (in blue, Fig. 6b) patterns is shown in Fig. 6.

Both values of the completeness magnitude threshold are graphically identified, through the log-cumulative frequency plots. For each plot, the former is identified as a point for which there is a strong change in the line slope. The completeness magnitude threshold values are $m_0 = 0.80$ inches, for hailstorms' pattern, and $m_0 = 1.35$, for thunderstorms winds' pattern.

3.5 Yearly models fitting

We fit an ETAS model for each year of observation, from 1996 to 2022 for the hailstorms pattern, and from 2003 to 2022 for the thunderstorms winds' pattern. The temporal parameters are estimated by setting up the day on which the events occur, while the spatial parameters refer to longitude and latitude values recorded in kilometers. We tried to consider a linear predictor within which we inserted a linear predictor containing trigonometric functions (such as $\sin(\cdot)$ and $\cos(\cdot)$ functions), to capture the daily within-year seasonality effect (clearly shown in Subsect. 3.2), but the estimated parameters for the same covariates are not statistically significant. In this paper, we used the R algorithms in the `etasFLP` package (Chiodi and Adelfio 2023). For parameter estimation and bandwidth lengths evaluation an Expectation-Maximization (EM) type algorithm—with a weighting step—is used. Namely, during the Expectation step, the bandwidth lengths are evaluated, using Silverman's *rule of thumb* and according to the set of weights [each computed with the formulation in (11)] attributed to each point; during the Maximization step, the log-likelihood function (9) is maximized with a Newton-type algorithm, to obtain the set of parameter estimates. The corresponding stop criterion is defined according to some convergence criterions. For the details on the used algorithm, see: Chiodi and Adelfio (2017).

The results of the ETAS model fitting are shown in Figs. 7 and 8, for marked spatio-temporal hailstorms' and thunderstorms winds' patterns, respectively. In Figs. 7 and 8, the black lines represent the values of the parameter estimates, and the colored lines above and below are the associated \pm estimated standard errors.

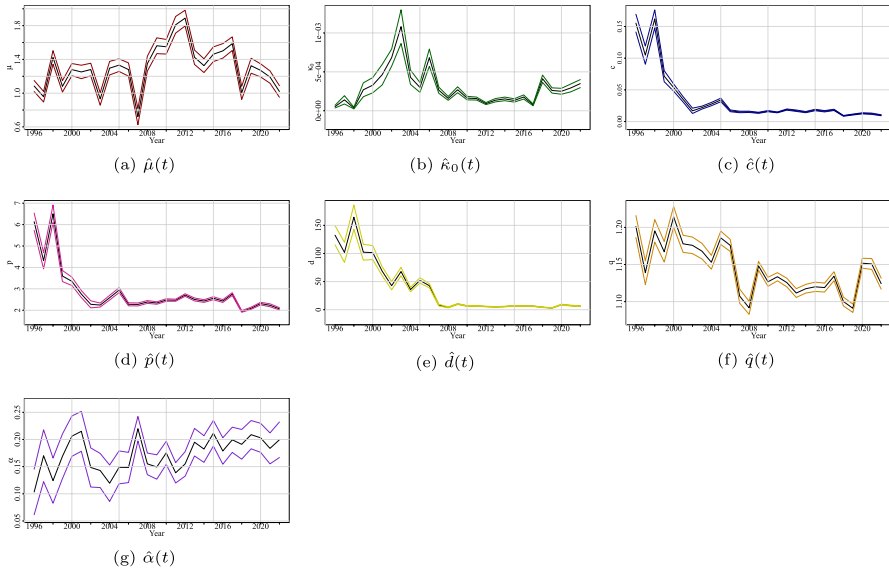


Fig. 7 Yearly parameter estimates of the ETAS models for hailstorms' patterns

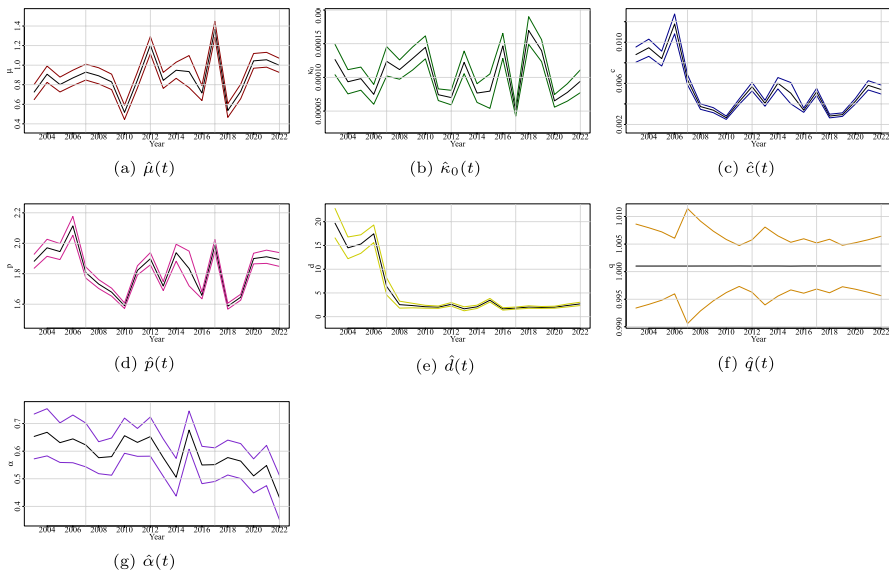


Fig. 8 Yearly parameter estimates of the ETAS models for thunderstorms winds' patterns

According to the results shown in Fig. 7, the of background events (Fig. 7a) varies between 0.5 and 2 (according to the estimated minimum and maximum values, in 2007 and 2012) each day. The aftershocks productivity parameter (Fig. 7b) estimate is the highest in 2003 and the lowest in 1998. Several peaks are observed, in

2003, 2006 and 2009, where the background component of the patterns recorded the highest amount of offsprings generation. The spatial (Fig. 7e) and temporal (Fig. 7c) displacement parameter estimates, as well as the spatial (Fig. 7f) and the temporal (Fig. 7d) decay parameter estimates, follow an overall decreasing trend, i.e. the pattern’s ability to reproduce itself decreases over years. According to the estimated parameters of the offspring generation (Fig. 7g) over the years of observation, the hail extension value has an increasingly positive effect on the spatio-temporal reproductivity of the phenomenon, although its estimated uncertainty is the highest.

As shown in Fig. 8, the daily rate of background events (Fig. 8a) varies between 0.5 and 1 (according to the estimated minimum and maximum values, in 2010 and 2017); the aftershock productivity parameter (Fig. 8b) follows a variable trend, i.e., peaks are continuously observed, as well as for the temporal displacement (Fig. 8d) parameter, according to which the temporal decay is highly variable along the years of observation. The spatial decay rate (Fig. 2a) follows a constant and mainly uncertain trend, and the estimated spatial displacement parameter’s value (Fig. 8c) slowly decreases until it maintains a constant trend. Finally, the estimated magnitude’s offspring generation’s (Fig. 8g) parameters follow a decreasing trend over the years of observation, i.e., thunderstorms winds’ speed has a decreasing effect on the spatio-temporal reproductivity of the phenomenon.

Finally, Fig. 9 shows the yearly number of events included within the catalogues (Fig. 9a, c) and the estimated proportions of background events (Fig. 9b, d) by the yearly ETAS models, for hailstorms’ (first row) and thunderstorms winds’ (second row) patterns.

As shown in Fig. 9, the number of recorded severe (with high magnitude values) events increases until 2011, for both patterns (Fig. 9a, c), and decreases in the following years. Regarding the yearly proportions of background events for the hailstorms’



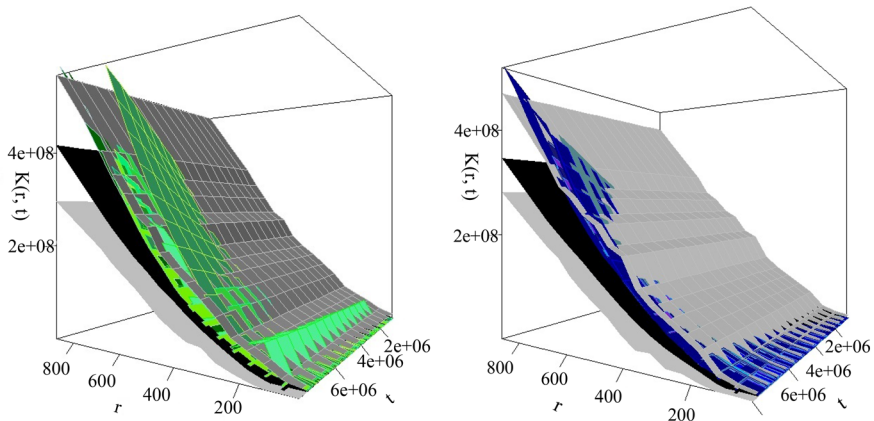
Fig. 9 Total number of events and estimated proportions of background events by each yearly ETAS model, for hailstorms’ (a, b) and thunderstorms winds’ (c, d) patterns

patterns (Fig. 9b): the maximum number of estimated mainshocks occurs in 1998, when 16% of events are classified as mainshocks, and linearly decreases until 2008, when 4% of events are classified as mainshocks. After 2008, the estimated proportion of mainshock events linearly increases until 2022. About the yearly proportions of background events for the thunderstorms winds' patterns (Fig. 9d): the maximum number of estimated mainshocks occurs in 2022, when $\approx 6\%$ of events are classified as mainshocks. The lowest number of mainshock events is estimated at 2010, and for the remaining years, it varies between 2.3% and 5.5%. For both severe convective storms' sub-patterns, the lower the number of severe convective storms, the higher the proportions of events classified as mainshocks, by the ETAS models.

3.6 Second-order diagnostics

The evaluation of the model fitting, in terms of residual analysis, is carried out by computing the yearly weighted K -functions, with the estimated conditional intensity functions by the fitted ETAS models, under translation edge correction. Figure 10 shows the yearly weighted K -functions (the black and grey surfaces are the same used in Fig. 5), for each yearly hailstorms (green, Fig. 10a) and thunderstorms winds (blue, Fig. 10b) pattern.

As shown in Fig. 10, most of the weighted surfaces are completely covered by the 95% envelopes around the theoretical K -function under $CSTR$, and some are close to them. As shown in Fig. 10a, one surface is above the 95% envelopes, and refers to the year 2008, for the associated severe hailstorms yearly process; similarly, as shown in Fig. 10b, one surface is above the 95% envelopes and refers to the year



(a) Weighted K -functions for each yearly hailstorms' pattern (green), computed under the assumption of known conditional intensity function by the fitted ETAS models, theoretical K -function (black) and 95% envelopes (grey) under $CSTR$

(b) Weighted K -functions for each yearly thunderstorms winds' pattern (blue), computed under the assumption of known conditional intensity function by the fitted ETAS models, theoretical K -function (black) and 95% envelopes (grey) under $CSTR$

Fig. 10 Weighted K -functions for each yearly hailstorms (a) and thunderstorms winds' (b) pattern, computed under the assumption of known conditional intensity function by the fitted ETAS models, theoretical K -function and 95% envelopes (grey) under $CSTR$

2011, for the corresponding severe thunderstorms winds yearly process. In both cases, there are some patterns that have not been fully captured by the fitted ETAS models (for the same years, and the associated patterns, the analysis should be further deepened).

4 Interpretation of the ETAS models' results

According to the results of the fitted ETAS models, the analysed hailstorms patterns recorded within the USA, between 1996 and 2022, are described according to the following interpretation:

- the number of severe hailstorms increases over time, and at least 1 hailstorm daily occurs, from 1996 to 2022, in the USA;
- the proportion of mainshock events, with respect to the detected severe events (events estimated as *parents* of a group of offspring events), decreases with time, i.e. all minor events are produced according to the spatio-temporal decay of the mainshock events, whose effect spatio-temporally lasts with an ever decreasing (and almost constant) displacement;
- the hail extension has an increasingly positive and strong effect on the offspring generation of small grasps;
- the higher the occurrence of daily mainshock events, the higher the aftershocks productivity effects.

while about the analysed thunderstorms winds' patterns recorded within the USA, between 2003 and 2022, we have:

- the proportion of mainshock events tends to be low and increases in 2022, when the number of severe convective (thunder)storms is the lowest recorded within the observation period;
- the higher the number of mainshock events, the higher the aftershock productivity rate;
- a decreasing temporal displacement effect along the years of observation is associated with a variable temporal decay, i.e. the mainshock effects on the generation of offspring contract and expand with some variability along the years of observation;
- the decreasing spatial displacement effect is always associated with a constant and approximately equal to 1 spatial decay, i.e. there is no variability around the estimated spatial decay, despite the decreasing trend of the spatial displacement;
- the wind speed of thunderstorms winds has a decreasing positive effect on the spatio-temporal phenomenon reproductivity.

5 Conclusions

In this paper, we focused on the use of stochastic process theory to evaluate the spatio-temporal behaviour of historical severe convective storms that occurred in the USA, starting from the available *Storm Events Database* (National Center for Environmental Information (NCEI) of the National Oceanic and Atmospheric Administration (NOAA) of the USA 2022a). We considered the hailstorms' and the thunderstorms winds' processes and evaluated their spatio-temporal aggregation behaviour from the marked self-exciting point processes point of view.

Our aim was to find a way to capture the yearly spatio-temporal behaviours, i.e. by dividing the temporal domain into sub-intervals, each associated with the corresponding year. We chose the temporal *Windowed ETAS model* approach, so that we could evaluate the behaviour of the sub-processes for each year of observation. We aimed to fit yearly spatio-temporal ETAS models for our first understanding of the most frequent severe convective storm processes (related to hailstorms and thunderstorms winds) that occurred in the USA (Kelly et al. 1985), between 1996 and 2022, from the marked self-exciting point processes point of view.

We evaluated the yearly patterns' behaviours and discovered how processes tend to spatio-temporally aggregate in all years of observations, conditionally to the magnitude values: hailstorms' pattern showed an aggregation behaviour mainly observed in the Central areas of the USA and along the days of the year (for each year), as well as a general trend towards sparsity and heterogeneity of the spatial hail extension within the whole territory; thunderstorms winds' pattern showed a spatial aggregation behaviour mainly observed in several sub-areas located from the Central to the Northern-East areas of the USA territory, as well as the highest temporal variability along each year of observation. For both patterns, given the aggregated nature of the data, it seemed reasonable to assume spatio-temporal self-exciting underlying processes. We then computed the yearly local indicators of spatio-temporal association (LISTA—Siino et al. 2018), i.e. the unweighted spatio-temporal K -functions. Taking into account the just-evaluated spatio-temporal aggregation behaviour of the yearly patterns, we fitted several yearly spatio-temporal ETAS models, such that yearly patterns were evaluated as self-exciting types. The yearly ETAS models were fitted according to the choice of the completeness magnitude threshold, i.e. by selecting only the events that make the catalogue complete. The complete magnitude threshold values were graphically assessed, according to the magnitude-frequency distributions plots (describing an exponential Richter-Gutenberg decay law). The chosen completeness magnitude threshold values were $m_0 = 0.80$ inches, for hail extension, and $m_0 = 1.35$ km/s, for thunderstorms winds' speed. Beyond these values, the catalogue should be complete, i.e. without missing events. The number of selected severe convective storms increased until 2011, and then slowly decreased in the successive years.

For all years of observation, the estimated parameters are always significant and we noticed how in some cases they follow a coherent (increasing or decreasing) long-term trend, or how they show a variable, jagged and (in some cases) uncertain trend.

For hailstorms' pattern, after 2008, a decrease in the spatial and temporal displacements and an associated decrease in the spatial and temporal decay laws is estimated, along the years of observation, so the spatio-temporal reproduction effects around the parents' centers, in terms of spatio-temporal distances, tend to reduce, as well as the aftershock productivity (and the number of offspring events) tends to be subject to the occurred mainshocks at highest spatio-temporal distances. The recorded magnitude values tend to have an increasing positive effect on the offspring generation, while the marginal aftershocks productivity parameter is estimated to have an increasing value, until 2010, to gradually decrease until 2017, and again to slowly increase along the remaining years of observation. The average number of days between two mainshocks tends to vary between the minimum temporal interval, corresponding to a half of day, and a maximum temporal interval of two days. For thunderstorms winds' patterns, a decrease in the spatial displacement is observed, as well as a tendency of the temporal displacement and decay laws estimated parameters, to assume lower and lower values, along the years of observation. For thunderstorms winds' patterns, the spatial decay law estimate is approximately equal to 1, and at a higher number of mainshocks events, a higher aftershocks productivity is always associated. Finally, for thunderstorms winds' patterns, a tendentially decreasing trend in the magnitude effect is estimated along the years of observation.

Finally, the estimated proportions of mainshock events tend to decrease for the hailstorms' patterns: a mainshocks percentage between 4% and 16% is always observed, and those for thunderstorms winds' patterns tend to vary along the years, until 2022, when the same assumes the highest value.

In further applications, we will consider the influence of covariates, as well as atmospheric and environmental variables, and eventually, several and different declustering techniques could be experimented.

Appendix

The statistical analysis is carried out through the statistical software R (R Core Team 2016). The packages used are `plot3D` (Soetaert 2014) for the graphical representations, `stpp` (Gabriel et al. 2013) for the exploratory and the inferential second-diagnostics, `etasFLP` (Chiodi and Adelfio 2023) for the ETAS model fitting, the first-order diagnostics and the spatial intensity function's estimates. The last updating of the main library (`etasFLP`) was carried out on September 2023.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10651-023-00593-4>.

Acknowledgements We would like to acknowledge the editors, for the necessary advice, and the anonymous reviewers, for improving the quality of this paper. The authors would like to express their appreciation for the detailed and useful comments. The first author gratefully acknowledges the financial support

for her doctoral study provided by the “Programma Operativo della Regione Siciliana, PR F.S.E. + Sicilia 2021/2027” (code: CUP G71I22001190006).

Author contributions Authors contributed equally to this work.

Funding Open access funding provided by Università degli Studi di Palermo within the CRUI-CARE Agreement.

Data availability The datasets analysed during the current study are freely available in the NOAA Institutional Repository (<https://repository.library.noaa.gov/>). Supplementary materials contain the graphical representations of the estimated background spatial intensity functions, which was carried out through the R package *etasFLP* (Chiodi and Adelfio 2023).

Code availability The code is available upon request from the authors.

Declarations

Conflict of interest The authors declare that they have no conflict of interest nor competing interests.

Ethical approval Not applicable.

Consent to participate Not applicable.

Consent for publication All authors have read and agreed to the published version of the manuscript.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adelfio G (2007) Statistical models diagnostics for earthquakes data. In: Atti del Convegno Intermedio della Società Italiana di Statistica (SIS) dal titolo “Rischio e previsione”. pp 29–30
- Adelfio G, Chiodi M (2008) Second-order diagnostics for space-time point processes with application to seismic events. *Environmetrics*. <https://doi.org/10.1002/env.961>
- Adelfio G, Chiodi M (2014) Space-time point processes semi-parametric estimation with predictive measure information. In: Proceedings of the METMA VII and GRASPA14 conference. Torino (IT), pp 10–12
- Adelfio G, Chiodi M (2020) Including covariates in a space-time point process with application to seismicity. *Stat Methods Appl* 30(3):947–971. <https://doi.org/10.1007/s10260-020-00543-5>
- Adelfio G, Ogata Y (2010) Hybrid kernel estimates of space-time earthquake occurrence rates using the Epidemic-Type Aftershock Sequence model. *Ann Inst Stat Math* 62:127–143
- Adelfio G, Schoenberg FP (2008) Point process diagnostics based on weighted second-order statistics and their asymptotic properties. *Ann Inst Stat Math* 61(4):929–948. <https://doi.org/10.1007/s10463-008-0177-1>
- Adelfio G, Chiodi M, De Luca L et al (2006) Earthquakes clustering based on maximum likelihood estimation of point process conditional intensity function. In: Riunione scientifica della società italiana di statistica, vol 1. Cleup, pp 79–82

- Adelfio G, Chiodi M et al (2013) Mixed estimation technique in semi-parametric space-time point processes for earthquake description. In: Proceedings of the 28th international workshop on statistical modelling. Istituto Poligrafico Europeo, 1, pp 65–70
- Adelfio G, Agosto A, Chiodi M et al (2020) Financial contagion through space-time point processes. *Stat Methods Appl* 30(2):665–688. <https://doi.org/10.1007/s10260-020-00538-2>
- Allen JT, Tippett MK, Sobel AH (2015) An empirical model relating U.S. monthly hail occurrence to large-scale meteorological environment. *J Adv Model Earth Syst* 7(1):226–243. <https://doi.org/10.1002/2014ms000397>
- Baddeley A, Turner R (2000) Practical maximum pseudolikelihood for spatial point patterns. *Aust N Z J Stat* 42(3):283–322. <https://doi.org/10.1111/1467-842x.00128>
- Brooks H (2013) Severe thunderstorms and climate change. *Atmos Res* 123:129–138. <https://doi.org/10.1016/j.atmosres.2012.04.002>
- Brooks HE, Lee JW, Craven JP (2003) The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data. *Atmos Res* 67–68:73–94. [https://doi.org/10.1016/s0169-8095\(03\)00045-0](https://doi.org/10.1016/s0169-8095(03)00045-0)
- Brown BG, Murphy A (1996) Verification of aircraft icing forecasts: the use of standard measures and meteorological covariates. In: Preprints, 13th conference on probability and statistics in the atmospheric sciences, San Francisco, CA. American Meteorological Society, p 252
- Chiodi M, Adelfio G (2017) Mixed non-parametric and parametric estimation techniques in R Package `etasFLP` for earthquakes' description. *J Stat Softw*. <https://doi.org/10.18637/jss.v076.i03>
- Chiodi M, Adelfio G (2023) `etasFLP`: mixed FLP and ML estimation of ETAS space-time point processes for earthquake description. R package version 2.2.2
- Chiodi M, Adelfio G et al (2008) Semiparametric estimation of conditional intensity functions for space-time processes. In: Scientific meeting of Italian Statistical Society, Cosenza
- Daley DJ, Vere-Jones D et al (2003) An introduction to the theory of point processes. Volume I: elementary theory and methods. Springer, New York
- Dehshiri SSH, Firoozabadi B (2023) A multi-objective framework to select numerical options in air quality prediction models: a case study on dust storm modeling. *Sci Total Environ* 863:160681. <https://doi.org/10.1016/j.scitotenv.2022.160681>
- Diggle PJ (2006a) Spatio-temporal point processes: methods and applications. *Monogr Stat Appl Probab* 107:1
- Diggle PJ (2006b) Spatio-temporal point processes, partial likelihood, foot and mouth disease. *Stat Methods Med Res* 15(4):325–336
- Eccel E, Cau P, Riemann-Campe K et al (2011) Quantitative hail monitoring in an alpine area: 35-year climatology and links with atmospheric variables. *Int J Climatol* 32(4):503–517. <https://doi.org/10.1002/joc.2291>
- Fortuin CC, Montes CR, Vogt JT et al (2022) Predicting risks of tornado and severe thunderstorm damage to southeastern U.S. forests. *Landsc Ecol* 37(7):1905–1919. <https://doi.org/10.1007/s10980-022-01451-7>
- Gabriel E, Diggle PJ (2009) Second-order analysis of inhomogeneous spatio-temporal point process data. *Stat Neerl* 63(1):43–51. <https://doi.org/10.1111/j.1467-9574.2008.00407.x>
- Gabriel E, Rowlingson BS, Diggle PJ (2013) `stpp`: an R package for plotting, simulating and analyzing spatio-temporal point patterns. *J Stat Softw* 53:1–29
- Gao L, Shi P (2022) Leveraging high-resolution weather information to predict hail damage claims: a spatial point process for replicated point patterns. *Insur Math Econ* 107:161–179. <https://doi.org/10.1016/j.insmatheco.2022.08.006>
- Gensini VA, Allen JT (2018) U.S. hail frequency and the global wind oscillation. *Geophys Res Lett* 45(3):1611–1620. <https://doi.org/10.1002/2017gl076822>
- González JA, Hahn U, Mateu J (2014) Analysis of spatio-temporal point patterns with replication. GRASPA Working Papers
- González JA, Rodríguez-Cortés FJ, Cronie O et al (2016) Spatio-temporal point process statistics: a review. *Spat Stat* 18:505–544. <https://doi.org/10.1016/j.spasta.2016.10.002>
- González JA, Hahn U, Mateu J (2019) Analysis of tornado reports through replicated spatiotemporal point patterns. *J R Stat Soc Ser C Appl Stat* 69(1):3–23. <https://doi.org/10.1111/rssc.12375>
- Gutenberg B, Richter CF (1944) Frequency of earthquakes in California. *Bull Seismol Soc Am* 34(4):185–188
- Hawkes AG (1971) Point spectra of some mutually exciting point processes. *J R Stat Soc Ser B (Methodol)* 33(3):438–443. <https://doi.org/10.1111/j.2517-6161.1971.tb01530.x>

- Illian J, Penttinen A, Stoyan H et al (2007) Statistical analysis and modelling of spatial point patterns. Wiley, Hoboken. <https://doi.org/10.1002/9780470725160>
- Insurance Information Institute (2021) Facts + statistics: homeowners and renters insurance. <https://www.iii.org/fact-statistic/facts-statistics-homeowners-and-renters-insurance>
- Ishimoto M (1939) Observations of earthquakes registered with the microseismograph constructed recently. Bull Earthq Res Inst 17:443–478
- Kelly DL, Schaefer JT, Doswell CA III (1985) Climatology of nontornadic severe thunderstorm events in the United States. Mon Weather Rev 113(11):1997–2014
- Kossin JP, Hall T, Knutson T et al (2017) Chapter 9: extreme storms. Climate science special report: fourth national climate assessment, volume I. Technical report. <https://doi.org/10.7930/j07s7kxx>
- Kunkel KE, Karl TR, Brooks H et al (2013) Monitoring and understanding trends in extreme storms: state of knowledge. Bull Am Meteorol Soc 94(4):499–514. <https://doi.org/10.1175/bams-d-11-00262.1>
- Li C (2000) A stochastic model of severe thunderstorms for transmission line design. Probab Eng Mech 15(4):359–364
- Mateu J, Ignaccolo R (2015) Spatio-temporal stochastic modelling of environmental hazards. Spat Stat 14:115–118. <https://doi.org/10.1016/j.spasta.2015.09.001>
- Molkenhain C, Donner C, Reich S et al (2022) GP-ETAS: semiparametric Bayesian inference for the spatio-temporal epidemic type aftershock sequence model. Stat Comput. <https://doi.org/10.1007/s11222-022-10085-3>
- National Center for Environmental Information (NCEI) of the National Oceanic and Atmospheric Administration (NOAA) of the United States of America (2022a) Severe storm event details. <https://www.ncdc.noaa.gov/stormevents/>
- National Center for Environmental Information (NCEI) of the National Oceanic and Atmospheric Administration (NOAA) of the United States of America (2022b) Storm data FAQ page. <https://www.ncdc.noaa.gov/stormevents/faq.jsp>
- National Weather Service (1999) <https://www.loc.gov/item/lcwaN0015050/>
- Nicolis O, Chiodi M, Adelfio G (2015) Windowed ETAS models with application to the Chilean seismic catalogs. Spat Stat 14:151–165. <https://doi.org/10.1016/j.spasta.2015.05.006>
- Ogata Y (1983) Estimation of the parameters in the modified Omori formula for aftershock frequencies by the maximum likelihood procedure. J Phys Earth 31(2):115–124. <https://doi.org/10.4294/jpe1952.31.115>
- Ogata Y (1988) Statistical models for earthquake occurrences and residual analysis for point processes. J Am Stat Assoc 83(401):9–27. <https://doi.org/10.1080/01621459.1988.10478560>
- Ohser J (1983) On estimators for the reduced second moment measure of point processes. Ser Stat 14(1):63–71. <https://doi.org/10.1080/02331888308801687>
- Omori F (1895) On the after-shocks of earthquakes. PhD Thesis, The University of Tokyo
- Quesada JA, Melchor I, Nolasco A (2017) Point process methods in epidemiology: application to the analysis of human immunodeficiency virus/acquired immunodeficiency syndrome mortality in urban areas. Geospat Health. <https://doi.org/10.4081/gh.2017.506>
- R Core Team (2016) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
- Raupach TH, Martius O, Allen JT et al (2021) The effects of climate change on hailstorms. Nat Rev Earth Environ 2(3):213–226
- Ripley BD (1976) The second-order analysis of stationary point processes. J Appl Probab 13(2):255–266. <https://doi.org/10.2307/3212829>
- Sheather SJ (2004) Density estimation. Stat Sci. <https://doi.org/10.1214/088342304000000297>
- Siino M, D'Alessandro A, Adelfio G et al (2018) Multiscale processes to describe the Eastern Sicily Seismic Sequences. Ann Geophys. <https://doi.org/10.4401/ag-7688>
- Silverman BW (1981) Using kernel density estimates to investigate multimodality. J R Stat Soc Ser B (Methodol) 43(1):97–99. <https://doi.org/10.1111/j.2517-6161.1981.tb01155.x>
- Sobash RA, Kain JS, Bright DR et al (2011) Probabilistic forecast guidance for severe thunderstorms based on the identification of extreme phenomena in convection-allowing model forecasts. Weather Forecast 26(5):714–728. <https://doi.org/10.1175/waf-d-10-05046.1>
- Soetaert K (2014) `plot3D`: tools for plotting 3-D and 2-D data. R package version. pp 10–2
- Soriano-Redondo A, Jones-Todd CM, Bearhop S et al (2019) Understanding species distribution in dynamic populations: a new approach using spatio-temporal point process models. Ecography 42(6):1092–1102. <https://doi.org/10.1111/ecog.03771>

- Stindl T, Chen F (2022) Spatiotemporal ETAS model with a renewal main-shock arrival process. *J R Stat Soc Ser C Appl Stat* 71(5):1356–1380. <https://doi.org/10.1111/rssc.12579>
- Stindl T, Chen F (2023) Stochastic declustering of earthquakes with the spatiotemporal renewal ETAS model. *Ann Appl Stat*. <https://doi.org/10.1214/23-aoas1756>
- Storm Prediction Center (2023) Maps, graphics, and data page. <https://www.spc.noaa.gov/wcm/>
- Tang J, Zhao C, Liu F et al (2022) Analyzing travel destinations distribution using large-scaled GPS trajectories: a spatio-temporal Log-Gaussian Cox process. *Physica A* 599:127305. <https://doi.org/10.1016/j.physa.2022.127305>
- Utsu T (1961) A statistical study on the occurrence of aftershocks. *Geophys Mag* 30:521–605
- Valente F, Laurini M (2020) Tornado occurrences in the United States: a spatio-temporal point process approach. *Econometrics* 8(2):25. <https://doi.org/10.3390/econometrics8020025>
- Veen A, Schoenberg FP (2008) Estimation of space–time branching process models in seismology using an EM–type algorithm. *J Am Stat Assoc* 103(482):614–624. <https://doi.org/10.1198/016214508000000148>
- Zhuang J, Ogata Y, Vere-Jones D (2002) Stochastic declustering of space-time earthquake occurrences. *J Am Stat Assoc* 97(458):369–380. <https://doi.org/10.1198/016214502760046925>