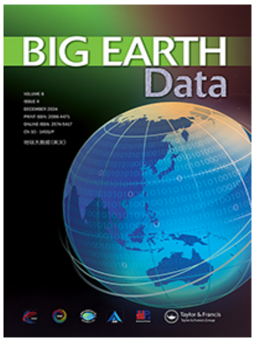

Big Earth Data

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tbed20

Enhancing satellite-based emergency mapping: Identifying wildfires through geo-social media analysis

Sebastian Schmidt, Monika Friedemann, David Hanny, Bernd Resch, Torsten Riedlinger & Martin Mühlbauer

To cite this article: Sebastian Schmidt, Monika Friedemann, David Hanny, Bernd Resch, Torsten Riedlinger & Martin Mühlbauer (30 Jan 2025): Enhancing satellite-based emergency mapping: Identifying wildfires through geo-social media analysis, Big Earth Data, DOI: 10.1080/20964471.2025.2454526

To link to this article: https://doi.org/10.1080/20964471.2025.2454526

Published online: 30 Jan 2025.

BIG EARTH DATA
https://doi.org/10.1080/20964471.2025.2454526

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Enhancing satellite-based emergency mapping: Identifying wildfires through geo-social media analysis

Sebastian Schmidt [a,b], Monika Friedemann [c], David Hanny [a,b], Bernd Resch [a,b,d], Torsten Riedlinger [c] and Martin Mühlbauer [c]

[a]Department of Geoinformatics - Z_GIS, University of Salzburg, Salzburg, Austria; [b]Geosocial Artificial Intelligence, IT:U Interdisciplinary Transformation University Austria, Linz, Austria; [c]Geo-Risks and Civil Security, German Aerospace Center (DLR), Weßling, Germany; [d]Center for Geographic Analysis, Harvard University, Cambridge, MA, USA

ABSTRACT

When a disaster emerges, timely acquisition of information is crucial for a rapid situation assessment. Although automation in the standard satellite-based emergency mapping workflow has been advanced, delays still occur at crucial steps. In order to speed up the provision of satellite-based crisis products to emergency managers, this paper proposes a geo-social media-based approach that detects disaster events based on the spatio-temporal analysis of georeferenced, disaster-related Tweets. The proposed methodology is validated on the basis of two use cases: wildfires in Chile and British Columbia. The results show the general ability of Twitter to forecast events several days in advance, at least for the Chile use case. However, there are large spatial differences, as there is a correlation between population density and the reliability of Twitter data. Consequently, only few meaningful alerts could be generated for British Columbia, an area with very low population numbers.

ARTICLE HISTORY
Received 1 August 2024
Accepted 23 December 2024

KEYWORDS
Disaster alerts; satellite-based emergency mapping; Twitter; RoBERTa

1. Introduction

Satellite-based emergency mapping (SEM) services such as the Copernicus Emergency Management Service (CEMS) Rapid Mapping (European Commission, 2023) provide geospatial information on demand in support of disaster management activities before, during, or immediately following a disaster (Voigt et al., 2016). Recent undertakings to accelerate SEM workflows in terms of delivery timeliness of crisis information by utilising early warning systems have proven effective (Wania et al., 2021). However, the process remains user-driven at crucial steps, thus inducing significant delays in tasking, data analysis and information provision.

CONTACT Sebastian Schmidt sebastian.schmidt@plus.ac.at Department of Geoinformatics - Z_GIS, University of Salzburg, Austria

While the automation of data analysis processes has been extensively researched, significant time delays still exist between the initial early warning and the SEM process activation – which is manually performed by an authorised user – and between the activation and satellite tasking. Considering these delays, initial uncertainties affect both users and SEM providers during disasters. Additionally, requesting authorities must be aware of the event location to define the AOI for a SEM activation. Therefore, this paper investigates possible automation enhancements to the initial steps of a SEM activation and proposes a concept for utilising disaster-related data from geo-social media to improve the timely provision of relevant EO-based crisis information.

We develop a methodology based on a spatio-temporal analysis of georeferenced, disaster-related Tweets. For this purpose, we take into account the difference in the relative number of disaster-related Tweets within an area compared to a baseline from the previous year. For the semantic classification of Tweets, we rely on a Robustly Optimized BERT Pre-training Approach (RoBERTa) model that was fine-tuned with the help of an active learning approach (Hanny et al., 2024). Our focus lies on two different 2023 wildfires in British Columbia and Chile, respectively. In doing so, we test the portability of our proposed methodology to find out for which regions a social media-based methodology for satellite tasking might be suitable. The AOIs differ in terms of their geographical structure (i.e. population density, topography), but also with regard to the language spoken. For this evaluation, we compare our results to official alerts and burnt areas derived from satellite imagery, i.e. post-event wildfire footprints. Consequently, we aim to answer the following research questions:

- **RQ1**: Is the analysis of geo-social media data suitable for the timely identification of areas affected by a wildfire and thus for improving SEM?
- **RQ2**: How do alerts derived from geo-social media posts differ spatially from official alerts or Sentinel-3 derived burnt areas?
- **RQ3**: Does the social media-based methodology work equally well for different geographical regions?

2. Related works

Today, SEM processes usually start with a user-driven activation, typically performed by an official authority in need of satellite-based crisis information. Upon activation, the state-of -the-art process follows the steps 1) tasking of on-demand satellites, 2) image acquisition, 3) image delivery, and 4) map product delivery including image analysis (Wania et al., 2021). Step 1 is only necessary for EO satellites which do not collect data on a permanent basis and instead have to be tasked. Examples are commercial Very High Resolution (VHR) optical satellites (e.g. WorldView-3) and radar satellites (e.g. TerraSAR-X).

Most attention has been given to optimising the delay produced in the last step by the time- and labour-intense manual and semi-automatic visual image interpretation. This has seen ongoing improvements, such as fully automated burnt area derivation (Knopp et al., 2020; Nolde et al., 2020). Such a real-time processor, which derives burnt areas from medium-resolution images (e.g. Sentinel-3), has recently been integrated into DLR's ZKI Fire Monitoring System (https://services.zki.dlr.de/fire). Similarly, Ajmar et al. (2019) propose to include early warning alerts like those of GDACS to reduce acquisition times by up to one day (GDACS, 2023a, 2023b). While such approaches shorten the time for triggering, they are limited to specific natural events, and in some cases are not commonly accessible.

Geo-social media data has already been used in some studies as an alternative data source to detect events induced by natural hazards (Hasan et al., 2019; Saeed et al., 2019) or even create "disaster alerts". Twitter has even already been used as a data source in GDACS since 2011 (Stollberg & De Groeve, 2012), albeit with a rather simplistic, solely keyword-filtering-based methodology. Nevertheless, relevant information on earthquake-related building collapses can be detected with this approach within the first half an hour of an event. Similarly, Shah et al. (2021) propose thresholds for alert generation based on the ratio of disaster-related, keyword-filtered Tweets in an AOI, starting at a threshold of 10%. More complex approaches for event detection based on machine learning algorithms have also been developed. For example, Rezaei et al. (2023) propose a semi-supervised framework based on an HAN. Havas and Resch (2021) use LDA to cluster Tweets in semantic topics for real-time monitoring of natural disasters. An extended LDA-based method that includes the location and time of a post and is conducted by performing similarity joins from a social media stream is presented by Zhou and Chen (2014). Azlan et al. (2020) compare KNN, SVM and NB for application in event detection. Pennington et al. (2022) develop a CNN-based system that detects landslide imagery from a keyword-filtered, live Twitter stream. Similarly, de Bruijn et al. (2019) use BERT to create a database of flood events from social media posts. Pinto et al. (2023) create wildfire probability heatmaps in Portugal from Twitter data by using a fine-tuned BERT model and employing location extraction with SpaCy. Li et al. (2022) propose a multimodal graph message propagation network based on a Graph Neural Network (GNN) that simultaneously considers text and images. While some existing approaches use georeferenced social media data, none of them explicitly compares their results to official alerts or satellite-derived information products.

There is a large body of literature that deals with the semantic classification of social media posts, especially in relation to natural hazards (e.g Resch et al. (2018); Huang et al. (2018); Parimala et al. (2021); Chae et al. (2012)). In this context, research dedicated to wildfires has been published, e.g. on the spatio-temporal distribution of wildfire-related Tweets in Israel (Zohar et al., 2023) and California (Wang et al., 2016). Common methods used in these studies include LDA (Havas & Resch, 2021), CNN (Huang et al., 2020) and BERT (Adwaith et al., 2022; Madichetty et al., 2021). Some studies also try to assign disaster-related posts to certain categories of relevance for disaster management (e.g. de Albuquerque et al. (2015); Powers et al. (2023); Blomeier et al. (2024)) or sentiments (e.g. Lever and Arcucci (2022)). An active learning based approach, such as the one used in this paper, has so far rarely been implemented for the classification of Twitter data (Cyril et al., 2021; Paul et al., 2023).

In general, there have been a few research approaches that explicitly combine social media analysis with remote sensing data. To map urban sprawl in Tanzania, Shao et al. (2021) combine user locations from Twitter data and Landsat 7 imagery. Karasov et al. (2022) investigate the demand for selected cultural ecosystem services (e.g. wildlife watching) in Estonia by classifying Landsat 8 data and geotagged photos from Flickr. Similarly, Lingua et al. (2023) use Flickr imagery and Landsat imagery to assess forest recreation in British Columbia. Vaz et al. (2019) use MODIS and Sentinel-2 data in combination with Flickr and Wikiloc images for the mapping of non-native tree species. By fusing social media data to remote sensing-based features via a residual CNN- and Long Short-Term Memory (LSTM)-based neural network, Cao et al. (2020) propose to improve land use classification. In the context of natural hazards and disasters, Huang et al. (2018) use Normalised Difference Water Index (NDWI) values from post-disaster satellite images to assign reliability scores to georeferenced, disaster-related Tweets and calculate inundation probability maps. Similarly, Rosser et al. (2017) combine Flickr data, Landsat-8 imagery and a Digital Elevation Model (DEM) to calculate similar flood probability maps. Zhong et al. (2023) propose fusing information derived from remote sensing, Weibo and mobile phone data for tracking wildfires based on a knowledge graph. Ahmad et al. (2019) use geo-location information in multi-modal social media data (text, images and videos) about disasters to retrieve and link it with Google Earth data for visualisation purposes. Boulton et al. (2021) find strong spatio-temporal correlations between the occurrence of keyword-filtered Tweets and MODIS-derived wildfires in some US states, without measuring the precision of their results. Bischke et al. (2016) enrich remote sensing data by classifying corresponding texts and images from social media and visualising them simultaneously.

There have also already been some studies on using social media data in connection with the acquisition of remote sensing images. Cervone et al. (2016) propose to employ a keyword filtering approach to identify disaster-related Tweets and subsequently spatial hot spots. However, their analysis focuses on flood events and does not include an evaluation of the speed of their approach. In their approach, Yang et al. (2022) similarly focus on flood events and Chinese language social media posts. Mühlbauer et al. (2024) present a method to speed up the SEM process by determining AOI from web data acquired by national or international entities, such as GDACS or the German Meteorological Service (DWD). Based on their proposal, we assess the applicability of geo-social media as a possible extension of their work. We are not aware of any study that uses geo-social media data to detect wildfires and, at the same time, compares the spatio-temporal accuracy of their results to official alerts or information derived from satellite imagery.

3. Methodology

In the following section, we will provide an overview of our data and methodological approach. We extracted data for our use cases from the official Twitter Application Programming Interface (API). We then trained a BERT-based model to classify these posts into two categories: disaster-related and unrelated. Subsequently, we generated alerts based on the spatially aggregated relative number of disaster-related Tweets. For this, we proposed three different methods that collate this Twitter activity against a baseline from the previous year. Figure 1 shows a schematic overview of our methodological approach. Lastly, we compared the output of the alert generation to spatio-temporal information derived from satellite imagery and official alerts.

3.1. Use cases

In the following, we will briefly describe the two use cases of our paper. We deliberately chose two AOI that differ in terms of their population structure and topography. The AOI

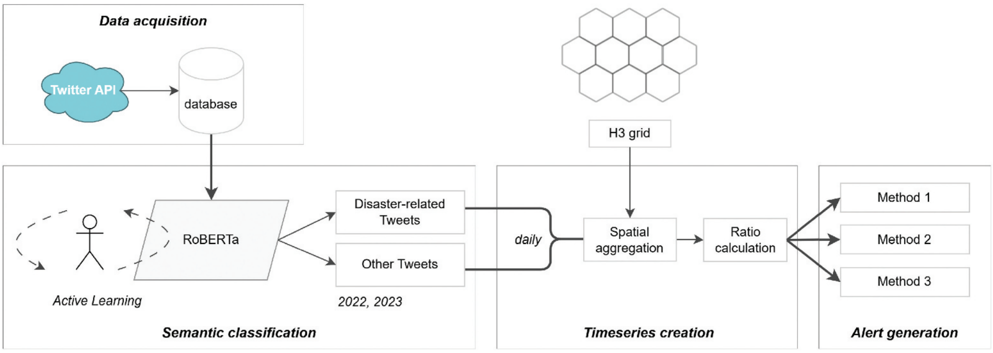

Figure 1. Overview of methodology.

in Central Chile is characterised by several large cities including Concepción (one of the largest Chilean agglomerations), Talca, and Los Ángeles, as well as several, rather clearly demarcated mountain ranges and valley. The AOI in British Columbia is a very rural area with large forests and more vast, continuous plains. Figure 2 shows the population distribution of our AOIs. We accessed the underlying data from the Global Human Settlement Layer (European Commission - Joint Research Centre, 2015), which has a spatial resolution of 1 km. We used the tool “Zonal statistics” in QGIS 3.32.2 to calculate the total population per grid cell.

3.1.1. 2023 Chile forest fires

In early February 2023, Central Chile was heavily affected by forest fires. Red alerts were declared by the Chilean National Disaster Prevention and Response Service (SENAPRED)

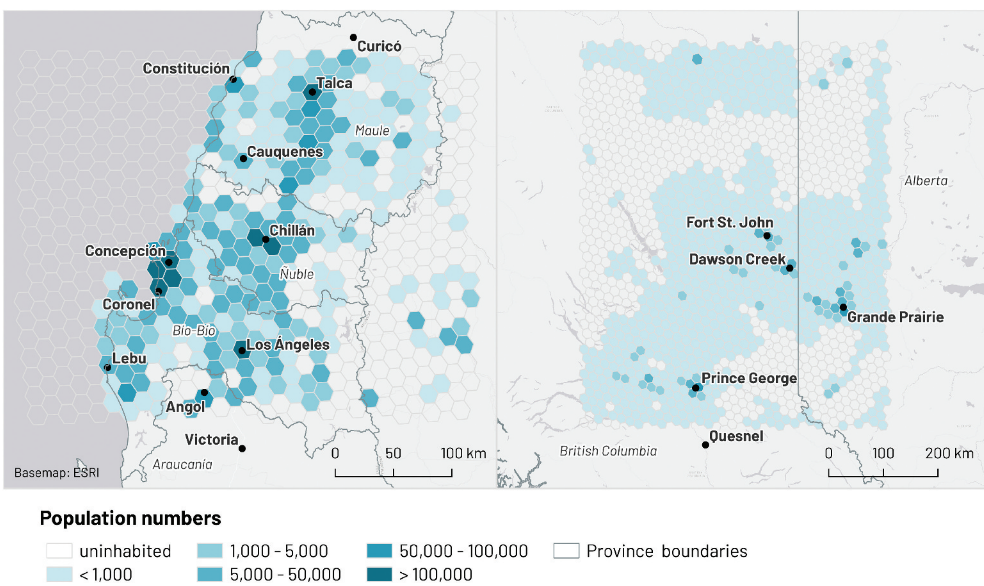

Figure 2. Population density of AOIs.

for the region of Ñuble on 2 February 20:18 UTC (17:18 local time), for the commune of Nacimiento in the Biobío region on 3 February 00:24 UTC (2 February 21:24 local time), and for the region of La Araucanía on 3 February at 22:09 UTC (19:09 local time). On 5 February, Chile requested support from UCPM member and participating states to mitigate the effects of the destructive fires. The CEMS Rapid Mapping was activated by the ERCC on the same day at 20:28 UTC (EMSR647) (European Commission, 2023) in support of operations in the affected areas. GDACS issued a first red alert on 14 February at 07:17:40 UTC (WF 1,012,119). In support of local authorities, DLR provided iteratively and automatically updated burnt area research products for the region of Ñuble starting from 3 February 15:22 UTC, for the commune of Nacimiento from 3 February 14:46 UTC and for the Region of La Araucanía from 1 February at 15:38 UTC through the ZKI Fire Monitoring System (available https://services.zki.dlr.de/fire). Figure 5 features the location of wildfire footprints, while Figure 3 shows a timeline of the alerts.

3.1.2. 2023 British Columbia forest fires

Beginning in March 2023, and with increased intensity starting in June, Canada was affected by a record-setting series of wildfires. As the worst wildfire season in recorded Canadian and North American history, eleven provinces and territories were affected, with large fires in Alberta, Nova Scotia, Ontario, and Québec. In northeastern British Columbia (BC), the Donnie Creek Wildfire became the single largest wildfire in BC history on June 18. By June 24, the fire was burning over an area of more than 5,648 square kilometres. The BC Fire Service discovered the wildfire on May 13 at 08:58:26 UTC (01:58:26 local time) and issued a first evacuation alert at 23:00:00 UTC (16:00:00 local time) (https://prrd.bc.ca/donnie-creek-tommy-lakes-evacuation-alert/). On May 19 a first map with the fire perimeters was published (G80280). CEMS rapid mapping was not activated and GDACS created a first green alert for the event on 23 June at 12:01:04 UTC (WF 1,015,007). Figure 6 features the location of wildfire footprints, while Figure 4 shows a timeline of the alerts.

3.2. Data

The social media platform Twitter (now: X) provides data access through various API endpoints. We retrieved Tweets using both the REST and the streaming API of Twitter, via which georeferenced data can be accessed. In our data collection approach, we followed Havas et al. (2021) and Schmidt et al. (2023). In a first

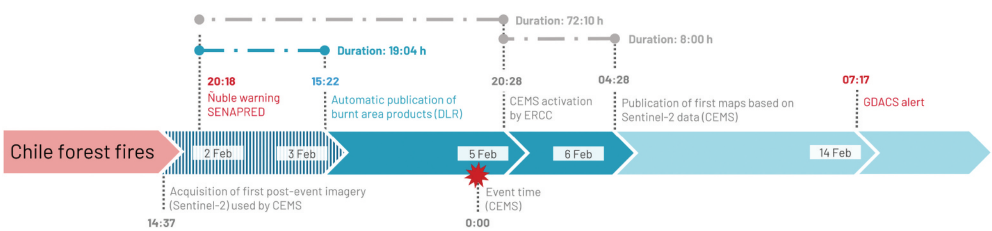

Figure 3. Timeline of public alerts and Eo-based crisis information provisions during forest fires in the region of Ñuble, Chile.

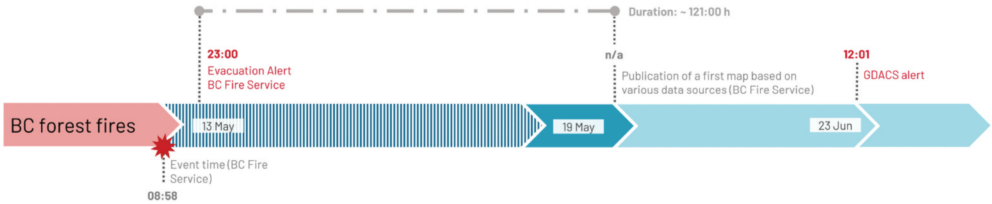

Figure 4. Timeline of public alerts and eo-based crisis information provisions during forest fires in the region of British Columbia.

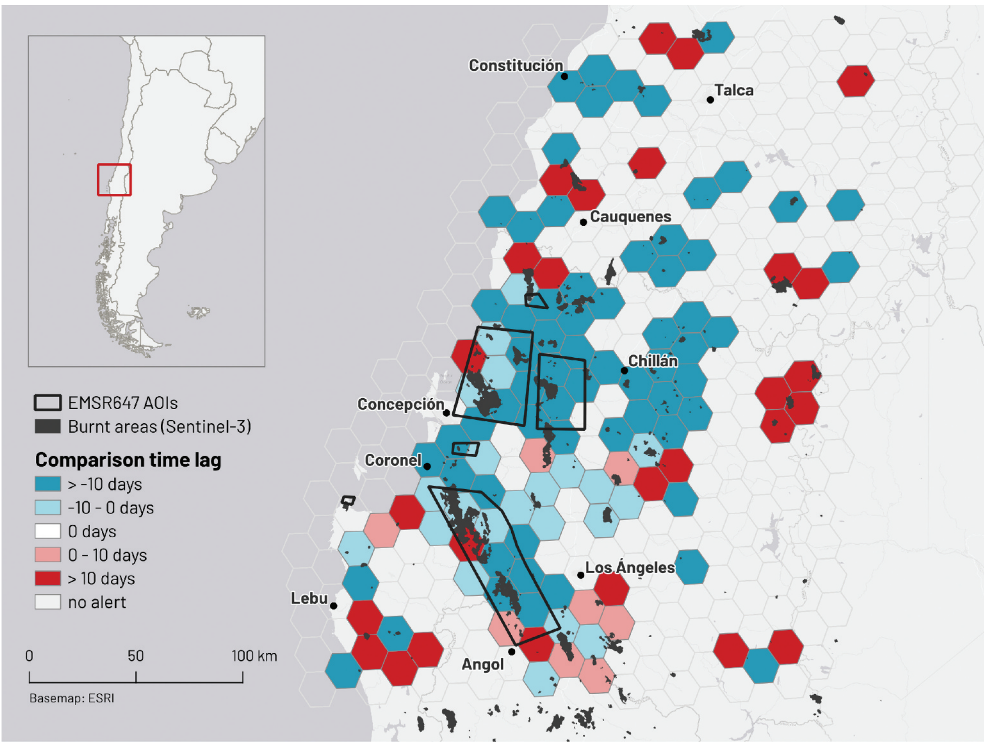

Figure 5. Evaluation of Chile alerts (Method 3). The temporal difference between the Twitter-based alerts and the burnt area derivation based on Sentinel-3 imagery is shown in five categories.

step, we collected global georeferenced data in 2022 and 2023. Then, we employed spatial and temporal filtering, i.e. for our Chile use case we extracted data for the Biobío, Ñuble and Araucanía regions and the first half of 2023. Semantic filtering was handled by the Disaster-RoBERTa model presented in Section 3.3. Table 1 shows the number of categorised Tweets for each of our use cases. To be able to calculate a baseline for our event detection, we also extracted data for the same timeframe and region in the previous year (2022) and performed the same analysis.

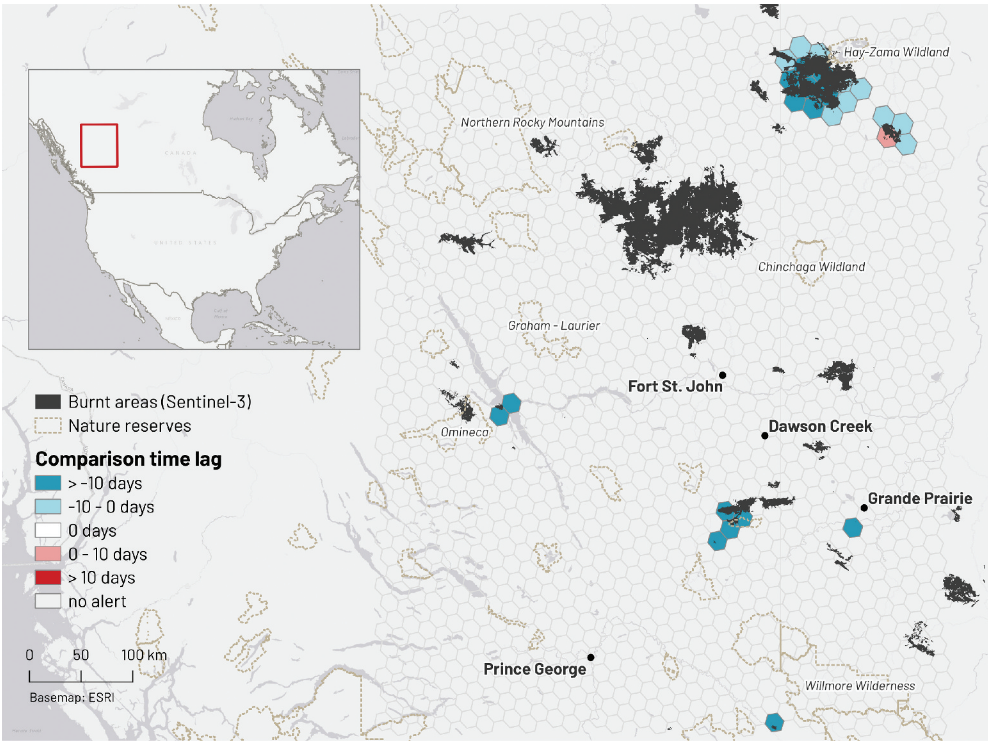

Figure 6. Evaluation of British Columbia alerts (Method 3). The temporal difference between the Twitter-based alerts and the burnt area derivation based on Sentinel-3 imagery is shown in five categories.

Table 1. Twitter data for use cases.

Use case	Timeframe	Population	Number of Tweets	
			all	disaster-related
2023 British Columbia fires	01.05.2023–30.06.2023	524,887	101,358	2,516
2023 Chile fires	01.01.2023–29.06.2023	3,464,883	2,282,470	97,846

3.3. Disaster-relatedness classification

To categorise Tweets as "related" or "unrelated" to a natural hazard and disaster, we used the multilingual *Twitter-XLM-RoBERTa-base* model (Barbieri et al., 2022) fine-tuned on labelled data as presented in Hanny et al. (2024). The base model was pre-trained on 198 M multilingual Tweets and fine-tuned using 179,391 Tweets labelled as "related" or "unrelated" to a disaster from the CrisisLexT6 (Olteanu et al., 2014) and CrisisLexT26 datasets (Olteanu et al., 2015), which reflect a broad collection of Tweets gathered during crises situations. Subsequently, the model was fine-tuned further using an active learning strategy where a total of 200 Tweets regarding the 2021 Ahr Valley floods in Germany and the 2023 Chile wildfires was labelled and returned to the model.

Table 2 compares the classification accuracy of the fine-tuned RoBERTa model against a keyword filtering approach as described in Hanny et al. (2024). The fine-tuned RoBERTa approach performed significantly better, achieving an accuracy score

Table 2. Accuracy of keyword filtering and the fine-tuned RoBERTa model presented by Hanny et al. (2024) for a binary disaster-relatedness classification task using different crisis tweet datasets. The fine-tuned RoBERTa model performed notably better across all test datasets.

Data	Keyword filtering	Fine-tuned RoBERTa
CrisisLex T6/T26	0.70	**0.94**
2021 Ahr Valley Floods	**0.96**	**0.96**
2023 Chile wildfires	0.71	**0.80**

of 0.94 versus 0.7 for a test dataset compromised of 44,848 Tweets from CrisisLex. The authors also evaluated the model on (1) 192 Tweets regarding the 2021 Ahr Valley flood in Germany where it performed on par with keyword filtering, and (2) 364 Tweets from the 2023 Chile wildfires. For the latter dataset, the fine-tuned RoBERTa model achieved an accuracy of 0.80 versus 0.71 for keyword filtering. In general, the training data of the model included multiple natural disasters including floods, forest fires, earthquakes, landslides and tropical storms.

3.4. Event detection

To detect specific events caused by natural hazards, we performed a daily spatial aggregation of our geo-social media data. For this, we used Uber's H3 grid, which provides a regular hexagonal, hierarchical grid of the entire planet (Uber, 2023). We opted for a hexagonal grid format as it fits the nearest neighbour logic for data aggregation less ambiguously than other approaches (Birch et al., 2007). We worked with grid level 5 (average edge length: 9.85 km) for both use cases. We chose this approach to achieve spatially comparable results. Since most of the geometries for Tweets are provided as polygons, i.e. bounding boxes of "places" specified by the user, we converted them to their centroids for the aggregation. For each grid cell, we then counted the number of disaster-related Tweets and all Tweets. Additionally, we calculated the ratio between both of these indicators to reduce the impact of the general population distribution, since considering only the pure number of disaster-related Tweets would have led to a bias towards larger agglomerations. Based on this data, we created a time series with daily interval. However, in our calculations we also considered a moving window for the previous 3 days to smooth the data a little, since disaster-related social media data can be quite noisy, i.e. contain a lot of unnecessary information. We only considered grid cells where the number of Tweets was higher than the previous year's median count to reduce the occurrence of outliers.

Based on the aforementioned ratio, i.e. number of disaster-related Tweets per grid cell in the given timeframe normalised by the overall amount of Tweets, we performed a spatial hot spot analysis using the Getis-Ord Gi* statistic (cf. Formula 1):

$$G_i^* = \frac{\sum_{j=1}^{n} w_{i,j}x_j - \bar{X}\sum_{j=1}^{n} w_{i,j}}{S\sqrt{\frac{\left[n\sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j}\right)^2\right]}{n-1}}} \tag{1}$$

Where x_j is the attribute value of feature j; $w_{i,j}$ is the spatial weight between feature i and j; n the sample size; $\bar{X}$ the mean value and S the standard deviation of the dataset (Ord & Getis, 1995). As the relative amount of disaster-related Tweets per grid cell, x_j had values between 0 and 1. As our data was aggregated on a regular, hexagonal grid, we considered the Queen contiguity as the most sensible approach for the definition of our spatial neighbourhood, i.e. spatial weights matrix. It only considers directly adjacent observations, i.e. sharing a vertex, as relevant for the calculation.

Overall, we proposed and implemented three different methods to detect events from geo-social media data:

- **Method 1**: An alert is generated if the amount of disaster-related Tweets within a cell is significantly higher than in the previous year (cf. Algorithm 1). For this method, the ratio of disaster-related posts is analysed, which must be two standard deviations higher than the average value of the baseline. As this ratio is usually very small, one standard deviation, i.e. the variation around the mean, was not enough to detect unusual social media activity. Additionally, only grid cells are analysed where the number of Tweets is higher than the median of the baseline, as we assumed that social media activity would be considerably higher than usual in the event of a disaster (Wang et al., 2021).

Algorithm 1: Event Detection - Method 1

```
Input: Count of disaster-related and all Tweets per H3 grid cell for 2022 (baseline) and 2023
Result: Disaster alert
Calculate ratio (disaster-related/all Tweets)
if Tweet count > median baseline count then
if Ratio > mean baseline ratio + two standard deviations then
Generate disaster alert;
end
end
```

- Method 2: An alert is generated if the amount of disaster-related Tweets within a cell is significantly higher than in the previous year (cf. Method 1) and there is at least one neighbouring cell to which this also applies (cf. Algorithm 2).

Algorithm 2: Event Detection - Method 2

```
Input: Count of disaster-related and all Tweets per H3 grid cell for 2022 (baseline) and 2023
Result: Disaster alert
Calculate ratio (disaster-related/all Tweets)
if Tweet count > median baseline count then
if Ratio > mean baseline ratio + two standard deviations then
if At least one neighbouring cell: Ratio > mean baseline ratio + two standard deviations then
Generate disaster alert;
end
end
end
```

- Method 3: An alert is generated if the results of the spatial hot spot analysis identify a cell as a highly significant hot spot (cf. Algorithm 3). For this, a p-value ≤ 0.1 is used as a threshold for statistical significance.

As with the other methods, only grid cells with considerably higher social media activity than usual are selected for the analysis.

Algorithm 3: Event Detection - Method 3
Input: Count of disaster-related and all Tweets per H3 grid cell for 2022 (baseline) and 2023
Result: Disaster alert
Calculate Getis-Ord Gi* hot spot analysis based on ratio (disaster-related/all Tweets)
if *Tweet count > median baseline count* **then**
if *z_score ≥ 1.65 and p_value ≤ 0.1* **then**
Generate disaster alert;
end
end

3.5. Evaluation

For the evaluation of our methodology, we compared the results of our event detection with official alerts and information derived from satellites. For an overview of official alerts for our two use cases, see Table 3. It should be noted that GDACS alerts are not static, but are updated in the course of a disaster. As these changes are overwritten, the final update that we were able to retrieve from the GDACS event page and analyse retrospectively was not identical to the very first alert. This information is provided in a XML alert message file which, in the case of GDACS, the DWD and most public alert providers, is based on the Common Alerting Protocol (CAP) (OASIS, 2010). For Chile, the latest CAP update with the ID 3 (CAP "currentepisodeid" parameter) was sent on 15 February at 00:00:00 UTC (CAP "sent" parameter). For the Donnie Creek event, the last update with the ID 10 was provided on 01 July at 00:00:00 UTC. While some providers use the CAP structure to include AOI geometries, the evaluated services, at least for our use cases, only included point coordinates or very rough bounding boxes.

Burnt areas were identified using a modular processing chain for automated burnt area derivation from optical Sentinel-2/3 satellite imagery as well as Aqua/Terra MODIS data, following the fully automated burnt area derivation developed by Nolde et al. (2020). This processor monitors burnt areas in near-real time and the satellite-derived products are made available through the ZKI Fire Monitoring System (Angermann et al., 2024). It uses the red and NIR bands to calculate the NDVI and, depending on the data source, optionally SWIR information to calculate the NBR. The generated Donnie Creek Wildfire dataset covering the affected region for summer 2023 was derived from Sentinel-3/Sentinel-2 imagery.

For each grid cell for which there was an event derived from the satellite data, we then calculated the difference between its timestamp and the date of our social media-based alerts in days. Additionally, we also derived a confusion matrix to check quantitatively how many areas were rightly identified and missed by our approach. For this, we

Table 3. Comparison of official alerts for use cases.

Use case	GDACS	CEMS	National alert
2023 Chile fires	CAP (14.02.2023)	EMSR647	Only textual alert by SENAPRED
2023 British Columbia fires	CAP (30.06.2023)	-	Information by BC Wildfire Service

converted both the social media alerts and satellite-based footprints into boolean values and checked their co-occurrence.

4. Results

In this section, we describe the results for our respective use cases. For this, we conducted a quantitative comparison between the social media-derived alerts and burnt areas based on Sentinel-3 imagery, as detailed in Section 3.5.

4.1. 2023 Chile forest fires

Figure 5 shows a comparison of the time lag between the Twitter-based alerts and burnt areas derived from Sentinel-3 imagery. It can be seen that the social media alerts were mostly simultaneous or even earlier than the satellite-based detection. This held particularly true for the Valle Central region (i.e. the longitude-parallel area between Talca and Los Ángeles) which is an area with comparatively high population density. There are, however, also several regions where the alert was very much delayed, particularly alongside the Andes mountain range in the East. The coastal region surrounding Concepción, the capital city, received early alerts, while the region close to Lebu and Angol had many late alerts. In other parts of the AOI that experienced wildfires, no alert was generated within the entire timeframe. However, this was only the case for very few regions and mostly very minor events. For larger agglomerations (e.g. Concepción, Los Ángeles, Talca), no alerts were generated, although most disaster-related Tweets originated from these urban centres. Since the methodology only considered the relative amount of disaster-related Tweets, this corresponded to the absence of wildfire footprints close to large urban centres. Overall, 399 alerts were generated for Method 1 over the entire timeframe. Since Method 2 only considered areas where Method 1 generated alerts in the respective cell and at least one of the neighbours, the number reduced to 138 alerts. For Method 3, 2,542 alerts were derived.

In this use case, the areas covered by official alerts were very roughly delineated. It can be seen that they include the largest burnt areas, but omitted many smaller forest fires. Most of these, however, overlapped with areas where social media-based alerts were generated at some point within the timeframe. A GDACS alert for the AOI was generated on 14 February, which was several days after the outbreak of major wildfires in the region. This is due to the fact that GDACS alerts focus on sudden-onset disasters such as earthquakes and possible subsequent tsunamis, flash floods, and volcanic eruptions. Nevertheless, GDACS also generates alerts for forest fire events which include automatic estimates and risk analysis provided by the JRC. For 80 grid cells in our AOI, our Twitter-based methodology generated an alert before 14 February. This represented 65.6% of all grid cells for which we could generate an alert. However, it should be repeated that GDACS alerts are not static, i.e. not ideal for spatio-temporal comparisons. For this reason, we mainly focused on the satellite-derived information for evaluation purposes.

Table 4 shows a confusion matrix between our social media-based alerts and the events detected from Sentinel-3 data. To achieve this, we compared the two datasets at the same timestamp. Based on the values of both datasets (0 = no alert, 1 = alert for social media; 0 = no burnt area, 1 = burnt area for satellite data), we calculated various model

Table 4. Confusion matrix for Chile use case (class 0 = no alert, class 1 = alert). Precision, recall, and F1-score are standard statistical metrics to evaluate model performance. Support is the number of observations per category.

Method	Class	Precision	Recall	F1-score	Support
Method 1	0	1.00	0.75	0.86	83,778
	1	0.01	0.52	0.02	399
	Accuracy				0.75
	Macro avg	0.50	0.63	0.44	84,177
	Weighted avg	0.99	0.75	0.85	84,177
Method 2	0	1.00	0.75	0.86	84,039
	1	0.00	0.74	0.01	138
	Accuracy				0.75
	Macro avg	0.50	0.75	0.43	84,177
	Weighted avg	1.00	0.75	0.86	84,177
Method 3	0	0.98	0.76	0.85	81,635
	1	0.06	0.48	0.10	2,542
	Accuracy				0.75
	Macro avg	0.52	0.62	0.48	84,177
	Weighted avg	0.95	0.75	0.83	84,177

evaluation metrics based on the distribution of True Positives, False Positives, True Negatives and False Negatives. While the resulting values of precision, recall, and F1-score seem high on first glance, it should be noted that an overall accuracy as a quality measure is relatively misleading here, as a large proportion of the data points had no event and were therefore not categorised in either the social media data or the satellite images. Overall, the recall of our geo-social media alerts was considerably higher than their precision for all methods, i.e. a lot of the relevant disaster-affected areas were identified. Method 3 had the highest F1 score. The meaningfulness of this confusion matrix is discussed in Section 5.1. Nevertheless, we wanted to include a purely quantitative evaluation of our methods.

4.2. 2023 British Columbia forest fires

Although the AOI was much larger than for the Chile wildfires, only very few areas were actually detected for the British Columbia use case. As shown in Figure 6, most of these areas were identified significantly earlier via geo-social media than satellite data. However, the time lag was mostly larger than 10 days. The largest wild fire event, the Donnie Creek Wildfire, which was located north of Fort St. John, was not detected at all based on Twitter data. The only major wildfire that was recognised well was one that affected the region south of the Hay-Zama Lakes Wildland Provincial Park, which is located in the northern most part of our AOI. Overall, only two alerts were generated for Method 1 over the entire timeframe. Unsurprisingly, Method 2, which required alerts in neighbouring cells, did not provide a single alert. For Method 3, 550 alerts were derived in total. Since there was no CEMS alert for this AOI, the corresponding polygons from Figure 6 could not be added to this visualisation. The BC Wildfire Service released some quite detailed fire perimeters for the Donnie Creek Wildfire on May 19, 2023. However, no overlap was found between our Twitter-based alerts and the reported fire perimeter.

Table 5 shows a confusion matrix between our social media-based alerts and the events detected from Sentinel-3 data. For this, once again, a comparison was conducted at the same timestamp. As described above, the high overall accuracy

Table 5. Confusion matrix for British Columbia use case (class 0 = no alert, class 1 = alert). Precision, recall, and F1-score are standard statistical metrics to evaluate model performance. Support is the number of observations per category.

Method	Class	Precision	Recall	F1-score	Support
Method 1	0	1.00	0.87	0.93	86,910
	1	0.00	0.00	0.00	2
	Accuracy				0.87
	Macro avg	0.50	0.44	0.47	86,912
	Weighted avg	1.00	0.87	0.93	86,912
Method 2	0	1.00	0.87	0.93	86,912
	1	0.00	0.00	0.00	0
	Accuracy				0.87
	Macro avg	0.50	0.44	0.47	86,912
	Weighted avg	1.00	0.87	0.93	86,912
Method 3	0	0.99	0.87	0.93	86,362
	1	0.01	0.22	0.02	550
	Accuracy				0.87
	Macro avg	0.50	0.55	0.47	86,912
	Weighted avg	0.99	0.87	0.92	86,912

of the models is rather misleading. Since Method 1 generated almost no alerts and Method 2 none at all, the precision and recall of 0.00 were hardly surprising. For Method 3, a similar pattern to the Chile use case could be observed, i.e. higher recall than precision for areas with event, although the recall of 0.22 was significantly lower.

5. Discussion

In the following, we discuss the results of our two case studies, before addressing some limitations of our proposed approach.

5.1. Discussion of results

As described in Section 4, the results for our two use cases differed greatly. We found that the social media-based alerts for Chile closely represented the actual events (i.e. with a reasonable anticipation). This held particularly true for the Valle Central area and larger events close to the city of Concepción. Smaller events, especially along the sparsely populated Andes mountain range, were unsurprisingly less likely to be identified. On the other hand, the results for British Columbia were not particularly meaningful. The most important event, the Donnie Creek Wildfire, was not identified at all. Only one event in the north-east of our AOI could be detected in advance based on social media data.

To achieve adequate data coverage, i.e. a significant amount of Tweets per spatial unit, we decided to aggregate our Twitter data on the level 5 h3 grid. Therefore, our social media-based grid cells were coarser than both CEMS (for Chile) or Sentinel-3 information. However, the aim of our study was not the exact spatial localisation of events, but the identification of possibilities to accelerate the SEM process. Satellites used in the SEM process chain generally have a swath width that also covers somewhat coarser AOI (between 120 km for SPOT-5 and 2,330 km for MODIS Aqua/Terra). Consequently, our timely geo-social media-based alerts should be well-suited as an early detection mechanism in the SEM process, given their spatial resolution.

We found that the scarcity of social media data in affected regions can be seen as a potential limitation of our approach (Yang et al., 2022). This was especially true for the British Columbia use case, where the population density was very low and consequently there were only very few Tweets. In other regions of the world where Twitter is not popular, similar limitations are to be expected. To analyse this more quantitatively, we compared the population figures in the grid cells with alerts with the other cells, finding that areas with higher population densities were more likely to be identified. For the British Columbia use case, we found a Point-biserial correlation coefficient of 0.25935 (p: 0.0) between an alert (Method 3) and the population numbers. For the Chile use case, this correlation was much lower with 0.08468 (p: 0.0). This could be related to the greater proximity of unpopulated areas to cities in Central Chile, such as more accessible and thus more popular nature parks.

In our study, we proposed three different methods. The first two methods, that were mainly based on the comparison of the ratio of disaster-related Tweets to all Tweets to a baseline from the previous year, generated a much smaller number of alerts. This was particularly noteworthy for the British Columbia use case, where the methods barely generated any alerts, although the wildfires were among the most devastating ever in the region. This was also confirmed by the generally lower performance with regard to precision and recall of Method 1 and Method 2 in both use cases. Furthermore, the results of Method 1 and Method 2 clustered much more around the big agglomeration of Concepción in the Chile use case, while affected regions closer to the Andes were more frequently identified by Method 1. Consequently, we evaluated Method 3 to be a more suitable approach.

In evaluating our results, we realised that a purely quantitative assessment as a comparison between our social media-based alerts and "ground-truth" from satellites was difficult conceptually. Merely comparing whether the same areas were detected at the same time (i.e. whether people are talking about the forest fire on social media at the time of the event), as we did it in Tables 4 and 5, does not fully meet the paper's goal of early detection. One solution to this would be to incorporate a fixed time lag for this evaluation. However, the definition of such a threshold is always associated with a certain degree of arbitrariness.

5.2. Discussion of methodology

The methodology presented in this paper is to be understood as a conceptualisation and a proof-of-concept that a social media-based triggering of remote sensing imagery acquisition could be sensible. The implementation of more refined statistical methods, mainly for event detection from time series, might be needed to transfer this concept into practise.

A major issue of geo-social media analysis is often the quality of georeferences. Most of the georeferenced Tweets only receive their geometries by "place" mentions (e.g. cities, neighbourhoods) specified by the user, which are represented as bounding boxes of different sizes. In our Chile use case, about 55% of the Tweets had polygon geometries smaller than a H3 grid cell. The need to convert these polygons into point data for unequivocal spatial aggregation introduced a certain degree of spatial blurring into our analysis. Since we considered this study a proof-of-concept, we did not filter out Tweets with coarser geometries in order not to reduce the size of our dataset. However, such an additional filtering would be advisable for a real-world application of our method, for which a use case-specific crawling of data should also be implemented.

Unfortunately, the same data quality restriction apply to many other social media platforms that mostly do not provide concrete georeferences in their data. Geocoding from textual content is an option (Serere et al., 2023), however, cannot provide the same spatial accuracy in many cases. The inclusion of other data sources, such as ground-sensor networks, might also be a valuable addition to our methodology. However, their often limited spatial availability would pose a major obstacle.

Another limitation of our social media analyses can be attributed to the semantic content of posts. While we were able to demonstrate generally good performance for the model we used to derive disaster-relatedness, other influencing factors, such as bot-generated content, can impact the reliability of results. However, there are indications that bot-generated content is not as frequent in georeferenced Tweets (Edry et al., 2021). A more fine-grained semantic analysis, e.g. by identifying specific topics within disaster-related Tweets, could further improve our approach.

Time series analysis is a quite wide field with a plethora of potential methods. For future research, more complex methods could be tested, such as the Pruned Exact Linear Time (PELT) algorithm (Killick et al., 2012) or piece-wise linear regression (Valsamis et al., 2019) for change point detection. Explicit spatio-temporal methods developed for other data sources could perhaps also be adapted (e.g. You et al. 2022; Anders et al. 2020). We decided to use the double standard deviation as a threshold for Method 1 and Method 2 based on several tests for our use case data. As with most threshold-based approaches, this decision is arbitrary in some way. Another aspect that can lead to distortions in the results is generally the selection of parameters. In our case, this mainly related to the choice of H3 resolution and the number of previous days taken into account for smoothing the time series. The latter might introduce autocorrelation into the analysis, since each smoothed value is dependent on previous values. While this creates rather gradual trends with less noise, it can also cause distortions (e.g. delays of signals) (Guerrero et al., 2018). In our case, this could lead either to alerts that are too late or to alerts that persist for an unnecessarily long time. As only the previous days are taken into account when smoothing, it should not lead to premature alerts. Nevertheless, to determine the optimal time lag, future studies could employ approaches such as the ACF (Box, 2008). We decided to use an uniform H3 resolution for comparability purposes and chose level 5, as we assumed that there would still be a considerable amount of Tweets per time step and grid cell at this level of aggregation.

For this first conceptualisation, we defined the baseline of Tweets as the values per region from the previous year. In principle, however, it would make more sense to consider a longer time series in order to mitigate the influence of individual disasters. However, we did not want to incur the additional, significant computational effort this would entail. As wildfire-related Tweets were also clustered in the larger Chilean cities in 2022, it can be assumed that this influenced the lack of alerts in vicinity to some large urban areas (e.g. Concepción, Los Ángeles, Talca). A baseline calculated over multiple years could possibly have reduced this effect. However, these regions were also typically not directly affected by wildfires, i.e. the relative amount of disaster-related Tweets was lower than in more rural zones in both years. Some of the social media-based alerts had a premature offset of more than 10 days to the satellite-based detections and were even outside of the actual fire season. These could be interpreted as “false alarms”, assuming that the timing of the respective events by the satellites was correct. These unwanted deviations could possibly be minimised by carrying out the baseline calculation over a longer period.

As stated in the introduction, satellite data acquisition planning usually constitutes a manual, time-consuming search and coordination process in a SEM workflow. Based on the results of this study, we will consider incorporating AOI automatically extracted from geo-social media and other open data (e.g. public alerts) (Mühlbauer et al., 2024) in combination with automatically processed satellite acquisition plans for an improved identification of possible and upcoming satellite data acquisitions. However, we realised that this approach only seems suitable for regions with a considerable activity on Twitter. Analysing the global distribution and density of geo-referenced Tweets would be a useful next step to identify regions of the world where our proposed approach could produce reliable results.

This paper shows the high relevance of geo-social media data in the disaster management context, which is why the availability of such data will remain of utmost importance. However, since the takeover of Twitter by Elon Musk in late 2022, access to Twitter data for research has been increasingly restricted. Additionally, there might be changes in the user profiles and discourses on the platform that can affect the quality of research based on Twitter data (Schmidt et al., 2023). The application of the methodology developed here to other data sources (e.g. Facebook, Telegram, YouTube) should therefore be explored. However, the limited spatial resolution of georeferences provided by these platforms will probably be a methodological challenge.

6. Conclusion

This paper presented a methodology to generate alerts for wildfires from Twitter data. There are three potential use cases for this information: Firstly, these alerts could be used to define and refine the AOI used for activating the SEM process. Secondly, AOI automatically extracted from geo-social media could help authorities in identifying for which regions on-demand satellites could be tasked actively. Thirdly, it is also conceivable that this data can be used to augment satellite-based crisis information, and vice versa.

We were able to show that georeferenced Twitter data can provide information on a wildfire simultaneously or even earlier than official alerts or information derived from fixed-orbit satellites like Sentinel-3, at least for some use cases and regions (**RQ1**). For regions with a high population density and Twitter activity, geo-social media can thus be seen as a suitable data source to issue early warnings or potential triggers within the SEM process. With regard to **RQ2**, we can conclude that the spatial resolution of the official alerts differed heavily. CEMS or GDACS alerts were mostly very coarse, mainly referring to large bounding boxes, or did not exist at all, as in our British Columbia use case. While the social media-derived information also has some imprecision, its spatial resolution can theoretically be more specific than official alerts and, most importantly, sufficient for acquiring fitting imagery. The sufficient availability of Twitter data is the major limitation of our approach, since it is highly dependent on the presence of georeferenced Tweets in the respective region. This was particularly evident for the British Columbia use case, as there was hardly any Twitter activity in the sparsely populated areas that were heavily affected by forest fires. We must therefore conclude that the social media-based methodology does not work equally well for different geographical regions (**RQ3**).

In future studies, the methodology presented in this paper should also be applied to other types of disasters triggered by natural hazards. Floods would be particularly suitable for this, as they often affect populated areas and are also typically longer-lasting events with a clear onset and progression. In this case, a more quantitative comparison with official alerts would also be more straightforward, since in many countries flood warnings are usually issued timely by national and international meteorological and hydrological organisations and services. Examples are the German Cross-state Flood Portal (https://www.hochwasserzentralen.de/en/LHP) as well as the Early Warning Dissemination System (https://meteoalarm.org/en/live/MeteoAlarm) that aggregates and accessibly provides awareness information from 38 European national meteorological and hydrological services. In particular for countries without sufficient flood warning infrastructure, the applicability of our approach should be evaluated.

Acknowledgements

We would like to thank Michael Nolde (German Aerospace Center) for providing Sentinel-3 derived burnt area data.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This project has received funding from the European Commission - European Union under HORIZON EUROPE (HORIZON Research and Innovation Actions) under grant agreement 101093003 (HORIZON-CL4-2022-DATA-01-01). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union - European Commission. Neither the European Commission nor the European Union can be held responsible for them.

Notes on contributors

Sebastian Schmidt is a PhD student and researcher in the Department of Geoinformatics – Z_GIS at the University of Salzburg. He holds a master's degree in Geography and a bachelor's degree in Romance philology from the University of Heidelberg. His research focuses on the analysis of geo-social media and corporate website data for environmental questions.

Monika Friedemann is a researcher at DLR's German Remote Sensing Data Center (DFD) in the department "Geo-Risks and Civil Security". She holds a diploma (equivalent to Master) in Media Informatics from the University of Munich (LMU). Her research and development activities focus on frontend design and development of remote-sensing-based environmental and crisis information systems, situation assessment and decision support.

David Hanny is a PhD student and researcher in the Geosocial Artificial Intelligence lab at the Interdisciplinary Transformation University Austria. He holds a master's degree in Data Science from the University of Salzburg and a BA in journalism from the FH Vienna. His current work focuses on the multimodal analysis of geo-referenced social sensing data.

Bernd Resch is a Professor and head of the Geosocial Artificial Intelligence lab at the Interdisciplinary Transformation University Austria, as well as a Visiting Scholar at Harvard University. His research interest revolves around understanding cities as complex systems through analysing a variety of digital data sources, focusing on developing machine learning algorithms to analyse human-generated data like social media posts and physiological measurements from wearable sensors.

Torsten Riedlinger is a senior scientist at DLR's German Remote Sensing Data Center (DFD), with background in remote sensing, disaster management support, analysis of geo-risks, decision support and early warning. He holds a PhD in Applied Geoscience from the University of Würzburg in Germany, is deputy head of DFD's department "Geo-risks and Civil security" and is leading a research team on the development of environmental and crisis information systems.

Martin Mühlbauer is a system engineer at the German Aerospace Center. His work focuses on warning systems and the analysis of heterogeneous data from the World Wide Web. He holds a diploma in geography from Ludwig Maximilian University, Munich, Germany.

ORCID

Sebastian Schmidt http://orcid.org/0000-0003-1912-9771
Monika Friedemann http://orcid.org/0000-0001-7383-1196
David Hanny http://orcid.org/0009-0004-8017-0786
Bernd Resch http://orcid.org/0000-0002-2233-6926
Torsten Riedlinger http://orcid.org/0000-0003-3836-614X
Martin Mühlbauer http://orcid.org/0000-0003-3849-1143

Data availability statement

The data that support the findings of this study are available from the corresponding author, S.S., upon reasonable request.

References

Adwaith, D., Abishake, A. K., Raghul, S. V., & Sivasankar, E. (2022). Enhancing multimodal disaster tweet classification using state-of-the-art deep learning networks. *Multimedia Tools & Applications*, *81*(13), 18483–18501. https://doi.org/10.1007/s11042-022-12217-3

Ahmad, K., Pogorelov, K., Riegler, M., Conci, N., & Halvorsen, P. (2019). Social media and satellites: Disaster event detection, linking and summarization. *Multimedia Tools & Applications*, *78*(3), 2837–2875. https://doi.org/10.1007/s11042-018-5982-9

Ajmar, A., Annunziato, A., Boccardo, P., Giulio Tonolo, F., & Wania, A. (2019). Tsunami modeling and satellite-based emergency mapping: Workflow integration opportunities. *Geosciences*, *9*(7), 314. https://doi.org/10.3390/geosciences9070314

Anders, K., Winiwarter, L., Lindenbergh, R., Williams, J. G., Vos, S. E., & Höfle, B. (2020). 4D objects-by-change: Spatiotemporal segmentation of geomorphic surface change from LiDAR time series. *ISPRS Journal of Photogrammetry & Remote Sensing*, *159*, 352–363. https://doi.org/10.1016/j.isprsjprs.2019.11.025

Angermann, L. D., Nolde, M., Mühlbauer, M., Böck, M., Heinen, T., & Riedlinger, T. (2024). ZKI fire monitoring service of DLR. *ISCRAM 2024*. https://elib.dlr.de/205289/1/ISCRAM2024_Poster_DLR_ZKI_Fire_Monitoring.pdf

Azlan, F. A., Ahmad, A., Yussof, S., & Ghapar, A. A. (2020). Analyzing algorithms to detect disaster events using social media. *2020 8th International Conference on Information Technology and Multimedia (ICIMU)* (pp. 384–389). IEEE, Selangor, Malaysia.

Barbieri, F., Espinosa Anke, L., & Camacho-Collados, J. (2022). XLM-T: Multilingual language models in Twitter for sentiment analysis and beyond. *Proceedings of the Thirteenth Language Resources and Evaluation Conference* (pp. 258–266). European Language Resources Association, Marseille, France.

Birch, C. P., Oom, S. P., & Beecham, J. A. (2007). Rectangular and hexagonal grids used for observation, experiment and simulation in ecology. *Ecological Modelling*, *206*(3–4), 347–359. https://doi.org/10.1016/j.ecolmodel.2007.03.041

Bischke, B., Borth, D., Schulze, C., & Dengel, A. (2016). Contextual enrichment of remote-sensed events with social media streams. *Proceedings of the 24th ACM international conference on Multimedia* (pp. 1077–1081). ACM, Amsterdam The Netherlands.

Blomeier, E., Schmidt, S., & Resch, B. (2024). Drowning in the information flood: Machine-learning-based relevance classification of flood-related tweets for disaster management. *Information*, *15* (3), 149. https://doi.org/10.3390/info15030149

Boulton, C., Shotton, H., & Williams, H. (2021). Using social media to detect and locate wildfires. *Proceedings of the International AAAI Conference on Web & Social Media*, *10*(2), 178–186. https://doi.org/10.1609/icwsm.v10i2.14850

Box, G. E. P. (2008). *Time series analysis: Forecasting and control. Wiley series in probability and statistics* (4th ed.). J. Wiley & Sons.

Cao, R., Tu, W., Yang, C., Li, Q., Liu, J., Zhu, J., Zhang, Q., Li, Q., & Qiu, G. (2020). Deep learning-based remote and social sensing data fusion for urban region function recognition. *ISPRS Journal of Photogrammetry & Remote Sensing*, *163*, 82–97. https://doi.org/10.1016/j.isprsjprs.2020.02.014

Cervone, G., Sava, E., Huang, Q., Schnebele, E., Harrison, J., & Waters, N. (2016). Using Twitter for tasking remote-sensing data collection and damage assessment: 2013 Boulder flood case study. *International Journal of Remote Sensing*, *37*(1), 100–124. https://doi.org/10.1080/01431161.2015.1117684

Chae, J., Thom, D., Bosch, H., Jang, Y., Maciejewski, R., Ebert, D. S., & Ertl, T. (2012). Spatiotemporal social media analytics for abnormal event detection and examination using seasonal-trend decomposition. *2012 IEEE Conference on Visual Analytics Science and Technology (VAST)* (pp. 143–152). IEEE, Seattle, WA, USA.

Cyril, C. P. D., Beulah, J. R., Subramani, N., Mohan, P., Harshavardhan, A., & Sivabalaselvamani, D. (2021). An automated learning model for sentiment analysis and data classification of twitter data using balanced CA-SVM. *Concurrent Engineering*, *29*(4), 386–395. https://doi.org/10.1177/1063293X211031485

de Albuquerque, J. P., Herfort, B., Brenning, A., & Zipf, A. (2015). A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management. *International Journal of Geographical Information Science*, *29*(4), 667–689. Publisher: Taylor & Francis. https://doi.org/10.1080/13658816.2014.996567

de Bruijn, J. A., de Moel, H., Jongman, B., de Ruiter, M. C., Wagemaker, J., & Aerts, J. C. (2019). A global database of historic and real-time flood events based on social media. *Scientific Data*, *6*(1). https://doi.org/10.1038/s41597-019-0326-9

Edry, T., Maani, N., Sykora, M., Elayan, S., Hswen, Y., Wolf, M., Rinaldi, F., Galea, S., & Gruebner, O. (2021). Real-time geospatial surveillance of localized emotional stress responses to COVID-19: A proof of concept analysis. *Health & Place*, *70*, 102598. https://doi.org/10.1016/j.healthplace.2021.102598

European Commission (2023). Copernicus Emergency Management Service - mapping https://emergency.copernicus.eu/mapping .

European Commission - Joint Research Centre. (2015). *GHS-POP R2015A - GHS population grid, derived from GPW4, multitemporal (1975, 1990, 2000, 2015)* https://data.jrc.ec.europa.eu/dataset/jrc-ghsl-ghs_pop_gpw4_globe_r2015a .

GDACS. (2023a). *Global disaster alert and coordination system* https://www.gdacs.org.

GDACS. (2023b). *Global disaster alert and coordination system - models - earthquakes*.

Guerrero, V. M., Cortés Toto, D., & Reyes Cervantes, H. J. (2018). Effect of autocorrelation when estimating the trend of a time series via penalized least squares with controlled smoothness. *Statistical Methods & Applications*, *27*(1), 109–130. https://doi.org/10.1007/s10260-017-0389-8

Hanny, D., Schmidt, S., & Resch, B. (2024). Active learning for identifying disaster-related tweets: A comparison with keyword filtering and generic fine-tuning. In K. Arai (Ed.), *Intelligent systems and applications* (Vol. 1066, pp. 126–142). Springer Nature Switzerland. Series Title: Lecture Notes in Networks and Systems.

Hasan, M., Orgun, M. A., & Schwitter, R. (2019). Real-time event detection from the Twitter data stream using the TwitterNews+ framework. *Information Processing & Management*, *56*(3), 1146–1165. https://doi.org/10.1016/j.ipm.2018.03.001

Havas, C., & Resch, B. (2021). Portability of semantic and spatial–temporal machine learning methods to analyse social media for near-real-time disaster monitoring. *Natural Hazards*, *108* (3), 2939–2969. https://doi.org/10.1007/s11069-021-04808-4

Havas, C., Wendlinger, L., Stier, J., Julka, S., Krieger, V., Ferner, C., Petutschnig, A., Granitzer, M., Wegenkittl, S., & Resch, B. (2021). Spatio-temporal machine learning analysis of social media data and refugee movement statistics. *ISPRS International Journal of Geo-Information*, *10*(8), 498. https://doi.org/10.3390/ijgi10080498

Huang, X., Li, Z., Wang, C., & Ning, H. (2020). Identifying disaster related social media for rapid response: A visual-textual fused CNN architecture. *International Journal of Digital Earth*, *13*(9), 1017–1039. https://doi.org/10.1080/17538947.2019.1633425

Huang, X., Wang, C., & Li, Z. (2018). A near real-time flood-mapping approach by integrating social media and post-event satellite imagery. *Annals of GIS*, *24*(2), 113–123. https://doi.org/10.1080/19475683.2018.1450787

Karasov, O., Heremans, S., Külvik, M., Domnich, A., Burdun, I., Kull, A., Helm, A., & Uuemaa, E. (2022). Beyond land cover: How integrated remote sensing and social media data analysis facilitates assessment of cultural ecosystem services. *Ecosystem Services*, *53*, 101391. https://doi.org/10.1016/j.ecoser.2021.101391

Killick, R., Fearnhead, P., & Eckley, I. A. (2012). Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association*, *107*(500), 1590–1598. arXiv: 1101.1438 [q-bio, stat]. https://doi.org/10.1080/01621459.2012.737745

Knopp, L., Wieland, M., Rättich, M., & Martinis, S. (2020). A deep learning approach for burned area segmentation with Sentinel-2 data. *Remote Sensing*, *12*(15), 2422. https://doi.org/10.3390/rs12152422

Lever, J., & Arcucci, R. (2022). Sentimental wildfire: A social-physics machine learning model for wildfire nowcasting. *Journal of Computational Social Science*, *5*(2), 1427–1465. https://doi.org/10.1007/s42001-022-00174-8

Li, J., Wang, Y., & Li, W. (2022). MGMP: Multimodal graph message propagation network for event detection. In B. Þór Jónsson, C. Gurrin, M.-T. Tran, D.-T. Dang-Nguyen, A.-M.-C. Hu, B. Huynh Thi Thanh, & B. Huet (Eds.), *MultiMedia modeling* (Vol. 13141, pp. 141–153). Springer International Publishing. Series Title: Lecture Notes in Computer Science.

Lingua, F., Coops, N. C., & Griess, V. C. (2023). Assessing forest recreational potential from social media data and remote sensing technologies data. *Ecological Indicators*, *149*, 110165. https://doi.org/10.1016/j.ecolind.2023.110165

Madichetty, S., Muthukumarasamy, S., & Jayadev, P. (2021). Multi-modal classification of Twitter data during disasters for humanitarian response. *Journal of Ambient Intelligence and Humanized Computing*, *12*(11), 10223–10237. https://doi.org/10.1007/s12652-020-02791-5

Mühlbauer, M., Friedemann, M., Roll, J., Riedlinger, T., Henkel, F., Angermann, L., Böck, M., Kaminski, T., & Barginda, K. (2024). Improved satellite-based emergency mapping through automated triggering of processes. *Proceedings of the 21st ISCRAM Conference -Münster, GermanyMay 2024*. Münster, Germany.

Nolde, M., Plank, S., & Riedlinger, T. (2020). An adaptive and extensible system for satellite-based, large scale burnt area monitoring in near-real time. *Remote Sensing*, *12*(13), 2162. https://doi.org/10.3390/rs12132162

OASIS. (2010). *Common alerting protocol version 1.2*. OASIS (Organization for the Advancement of Structured Information Standards). https://docs.oasis-open.org/emergency/cap/v1.2/CAP-v1.2-os.html.

Olteanu, A., Castillo, C., Diaz, F., & Vieweg, S. (2014). CrisisLex: A lexicon for collecting and filtering microblogged communications in crises. *Proceedings of the International AAAI Conference on Web & Social Media*, *8*(1), 376–385. https://doi.org/10.1609/icwsm.v8i1.14538

Olteanu, A., Vieweg, S., & Castillo, C. (2015). What to expect when the unexpected happens: Social media communications across crises. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW '15 (pp. 994–1009). Association for Computing Machinery, New York, NY, USA.

Ord, J. K., & Getis, A. (1995). Local spatial autocorrelation statistics: Distributional issues and an application. *Geographical Analysis*, *27*(4), 286–306. https://doi.org/10.1111/j.1538-4632.1995.tb00912.x

Parimala, M., Swarna Priya, R. M., Praveen Kumar Reddy, M., Lal Chowdhary, C., Kumar Poluru, R., & Khan, S. (2021). Spatiotemporal-based sentiment analysis on tweets for risk assessment of event using deep learning approach. *Software: Practice & Experience*, *51*(3), 550–570. https://doi.org/10.1002/spe.2851

Paul, N. R., Sahoo, D., & Balabantaray, R. C. (2023). Classification of crisis-related data on Twitter using a deep learning-based framework. *Multimedia Tools & Applications*, *82*(6), 8921–8941. https://doi.org/10.1007/s11042-022-12183-w

Pennington, C. V., Bossu, R., Ofli, F., Imran, M., Qazi, U., Roch, J., & Banks, V. J. (2022). A near-real-time global landslide incident reporting tool demonstrator using social media and artificial intelligence. *International Journal of Disaster Risk Reduction*, *77*, 103089. https://doi.org/10.1016/j.ijdrr.2022.103089

Pinto, J. C., Gonçalo Oliveira, H., Cardoso, A., & Silva, C. (2023). Generating wildfire heat maps with Twitter and BERT. In P. Quaresma, D. Camacho, H. Yin, T. Gonçalves, V. Julian, & A. J. Tallón-Ballesteros (Eds.), *Intelligent Data Engineering and Automated Learning - IDEAL 2023* (Vol. 14404, pp. 82–94). Springer.

Powers, C. J., Devaraj, A., Ashqeen, K., Dontula, A., Joshi, A., Shenoy, J., & Murthy, D. (2023). Using artificial intelligence to identify emergency messages on social media during a natural disaster: A deep learning approach. *International Journal of Information Management Data Insights*, *3*(1), 100164. https://doi.org/10.1016/j.jjimei.2023.100164

Resch, B., Usländer, F., & Havas, C. (2018). Combining machine-learning topic models and spatiotemporal analysis of social media data for disaster footprint and damage assessment. *Cartography and Geographic Information Science*, *45*(4), 362–376. https://doi.org/10.1080/15230406.2017.1356242

Rezaei, Z., Eslami, B., Amini, M. A., & Eslami, M. (2023). Event detection in Twitter by deep learning classification and multi label clustering virtual backbone formation. *Evolutionary Intelligence*, *16*(3), 833–847. https://doi.org/10.1007/s12065-021-00696-6

Rosser, J. F., Leibovici, D. G., & Jackson, M. J. (2017). Rapid flood inundation mapping using social media, remote sensing and topographic data. *Natural Hazards*, *87*(1), 103–120. https://doi.org/10.1007/s11069-017-2755-0

Saeed, Z., Abbasi, R. A., Maqbool, O., Sadaf, A., Razzak, I., Daud, A., Aljohani, N. R., & Xu, G. (2019). What's happening around the world? A survey and framework on event detection techniques on Twitter. *Journal of Grid Computing*, *17*(2), 279–312. https://doi.org/10.1007/s10723-019-09482-2

Schmidt, S., Zorenböhmer, C., Arifi, D., & Resch, B. (2023). Polarity-based sentiment analysis of georeferenced tweets related to the 2022 Twitter acquisition. *Information*, *14*(2), 71. https://doi.org/10.3390/info14020071

Serere, H. N., Kanilmaz, U. N., Ketineni, S., & Resch, B. (2023). A comparative study of geocoder performance on unstructured tweet locations. *GI_Forum*, *1*, 110–117. https://doi.org/10.1553/giscience2023_01_s110

Shah, S. A., Yahia, S. B., McBride, K., Jamil, A., & Draheim, D. (2021). Twitter streaming data analytics for disaster alerts. *2021 2nd International Informatics and Software Engineering Conference (IISEC)* (pp. 1–6). IEEE, Ankara, Turkey.

Shao, Z., Sumari, N. S., Portnov, A., Ujoh, F., Musakwa, W., & Mandela, P. J. (2021). Urban sprawl and its impact on sustainable urban development: A combination of remote sensing and social media data. *Geo-Spatial Information Science*, *24*(2), 241–255. https://doi.org/10.1080/10095020.2020.1787800

Stollberg, B., & De Groeve, T. (2012). The use of social media within the global disaster alert and coordination system (GDACS). *Proceedings of the 21st International Conference on World Wide Web* (pp. 703–706). ACM, Lyon France.

Uber. (2023). *H3: A hexagonal hierarchical geospatial indexing system* https://github.com/uber/h3.

Valsamis, E. M., Husband, H., & Chan, G. K.-W. (2019). Segmented linear regression modelling of time-series of binary variables in healthcare. *Computational & Mathematical Methods in Medicine*, *2019*, 1–7. https://doi.org/10.1155/2019/3478598

Vaz, A. S., Gonçalves, J. F., Pereira, P., Santarém, F., Vicente, J. R., & Honrado, J. P. (2019). Earth observation and social media: Evaluating the spatiotemporal contribution of non-native trees to cultural ecosystem services. *Remote Sensing of Environment*, *230*, 111193. https://doi.org/10.1016/j.rse.2019.05.012

Voigt, S., Giulio-Tonolo, F., Lyons, J., Kučera, J., Jones, B., Schneiderhan, T., Platzeck, G., Kaku, K., Hazarika, M. K., Czaran, L., Li, S., Pedersen, W., James, G. K., Proy, C., Muthike, D. M., Bequignon, J., & Guha-Sapir, D. (2016). Global trends in satellite-based emergency mapping. *Science (New York, NY)*, *353*(6296), 247–252. https://doi.org/10.1126/science.aad8728

Wang, K., Lam, N. S. N., Zou, L., & Mihunov, V. (2021). Twitter use in hurricane Isaac and its implications for disaster resilience. *ISPRS International Journal of Geo-Information*, *10*(3), 116. https://doi.org/10.3390/ijgi10030116

Wang, Z., Ye, X., & Tsou, M.-H. (2016). Spatial, temporal, and content analysis of Twitter for wildfire hazards. *Natural Hazards*, *83*(1), 523–540. https://doi.org/10.1007/s11069-016-2329-6

Wania, A., Joubert-Boitat, I., Dottori, F., Kalas, M., & Salamon, P. (2021). Increasing timeliness of satellite-based flood mapping using early warning systems in the copernicus emergency management service. *Remote Sensing*, *13*(11), 2114. Number: 2114. https://doi.org/10.3390/rs13112114

Yang, T., Xie, J., Li, G., Zhang, L., Mou, N., Wang, H., Zhang, X., & Wang, X. (2022). Extracting disaster-related location information through social media to assist remote sensing for disaster analysis: The case of the flood disaster in the Yangtze River Basin in China in 2020. *Remote Sensing*, *14*(5), 1199. https://doi.org/10.3390/rs14051199

You, Y., Zhang, L., Tao, P., Liu, S., & Chen, L. (2022). Spatiotemporal transformer neural network for time-series forecasting. *Entropy*, *24*(11), 1651. https://doi.org/10.3390/e24111651

Zhong, W., Mei, X., Niu, F., Fan, X., Ou, S., & Zhong, S. (2023). Fusing social media, remote sensing, and fire dynamics to track wildland-urban interface fire. *Remote Sensing*, *15*(15), 3842. https://doi.org/10.3390/rs15153842

Zhou, X., & Chen, L. (2014). Event detection over Twitter social media streams. *The VLDB Journal*, *23*(3), 381–400. https://doi.org/10.1007/s00778-013-0320-3

Zohar, M., Genossar, B., Avny, R., Tessler, N., & Gal, A. (2023). Spatiotemporal analysis in high resolution of tweets associated with the November 2016 wildfire in Haifa (Israel). *International Journal of Disaster Risk Reduction*, *92*, 103720. https://doi.org/10.1016/j.ijdrr.2023.103720