FloDusTA: Saudi Tweets Dataset for Flood, Dust Storm, and Traffic Accident Events

Btool Hamoui, Mourad Mars, Khaled Almotairi

College of Computer and Information System, Common First Year Deanship
Umm AL-Qura University, Saudi Arabia
s43680523@st.uqu.edu.sa, {msmars, khmotairi}@ uqu.edu.sa

Abstract

The rise of social media platforms makes it a valuable information source of recent events and users' perspective towards them. Twitter has been one of the most important communication platforms in recent years. Event detection, one of the information extraction aspects, involves identifying specified types of events in the text. Detecting events from tweets can help to predict real-world events precisely. A serious challenge that faces Arabic event detection is the lack of Arabic datasets that can be exploited in detecting events. This paper will describe FloDusTA, which is a dataset of tweets that we have built for the purpose of developing an event detection system. The dataset contains tweets written in both Modern Standard Arabic and Saudi dialect. The process of building the dataset starting from tweets collection to annotation by human annotators will be present. The tweets are labeled with four labels: flood, dust storm, traffic accident, and non-event. The dataset was tested for classification and the result was strongly encouraging.

Keywords: Arabic tweets, Saudi dialect, Twitter, Event detection, Classification.

1. Introduction

In recent years, social media platforms in the Arabic region have been evolving rapidly. Twitter provides an easy form of communication that enables users to share information about their activities, opinions, feelings, and views about a wide variety of social events. It has been a great platform to disseminate events as they happen and released immediately, even before they are announced in traditional media.

Tweets' contents have become major sources for extracting information about real-world events. Critical events such as violence, disasters, fires, and traffic accidents that need emergency awareness require an extreme effort to detect and track. Twitter users' posts have been utilized as data provider to detect high-risk events with their locations, such as earthquakes (Sakaki et al., 2010), Traffic incidents (Gu et al., 2016) and floods (Arthur et al., 2018). An earlier work done by Sakaki et al. (2010) predicted and detected the location of an earthquake in Japan more quickly than the Japan Meteorological Agency. (Gu et al., 2016) identified five categories of traffic incidents in the city of Pittsburgh and Philadelphia (USA) using twitter data. A recent study by Arthur et al. (2018) utilized tweets to locate and detect flood in the UK.

Recently, event detection has been considered an active area of researches due to the widespread availability of data in social media. However, researches about event detection on Twitter applying it on Arabic is hampered by the lack of datasets that could be used to design and develop an event detection system. Until now, the dataset of (Almerekhi et al., 2016) and (Alhelbawy et al., 2016) are the only published Arabic datasets for event detection purposes that are freely available for research.

To detect an event in the Arabic region, constructing a dataset of Arabic events is mandatory. Leveraging Twitter popularity in Saudi Arabia, we aim to build a dataset containing tweets written in both Modern Standard Arabic (MSA) and Saudi dialect to detect flood, dust storm, and traffic accidents. We focus on the flood, dust storm, and traffic accident events according to their significant influence on human life and economy in Saudi Arabia (Youssef et al, 2015; Karagulian et al.,2019; Mansuri et al., 2015). To the best of our knowledge, this is the first publicly available Arabic dataset for the aim of detecting flood, dust storm, and traffic accident events.

In this paper, the main contributions are:

- We describe an Arabic dataset of Saudi event tweets **FloDusTA**: **Flo**od, **Dus**t Storm, **T**raffic **A**ccident Saudi Event dataset. The dataset will be publicly available for the research community[1].
- A preliminary set of experiments were conducted to establish a baseline for future work on building an event detection system.

The rest of this paper is organized as follows. Section 2 reviews the related works. Section 3 describes how tweets were collected and the cleaning and filtering that were deployed to extract a dataset of Saudi event tweets. In Section 4 we explain the annotation process in detail. In Section 5 the experiments are illustrated. Finally, we conclude and discuss future work.

2. Related work

Researches conducted for event detection system utilizing Arabic tweets have emerged in 2015, with the goal of detecting a disruptive event occurred in Abu Dhabi (Alsaedi and Pete ,2015). The dataset contained 1 Million tweets collected by streaming API in 2013. Their dataset was labeled by three annotators, and it was not made publicly-available. EveTAR, Event-centric Test Collection of Arabic Tweets, is the first available Arabic dataset designed for event detection task (Almerekhi et al.,2016). They also crawled tweets using Twitter streaming API for a month in January 2015. Each event in their dataset is relevant to an event founded in Wikipedia's Current Events Portal (WCEP) during that month. They collected 590 million tweets labeled by three annotators. A dataset of Arabic violence tweets that was constructed by Alhelbawy et al. (2016) consisted of 20,151 tweets. The tweets labeled with eight classes (HRA, political opinion, accident, crime, violence and other). The definition of tweets classes was reviewed by experts in the Minority Rights Group (MRG).

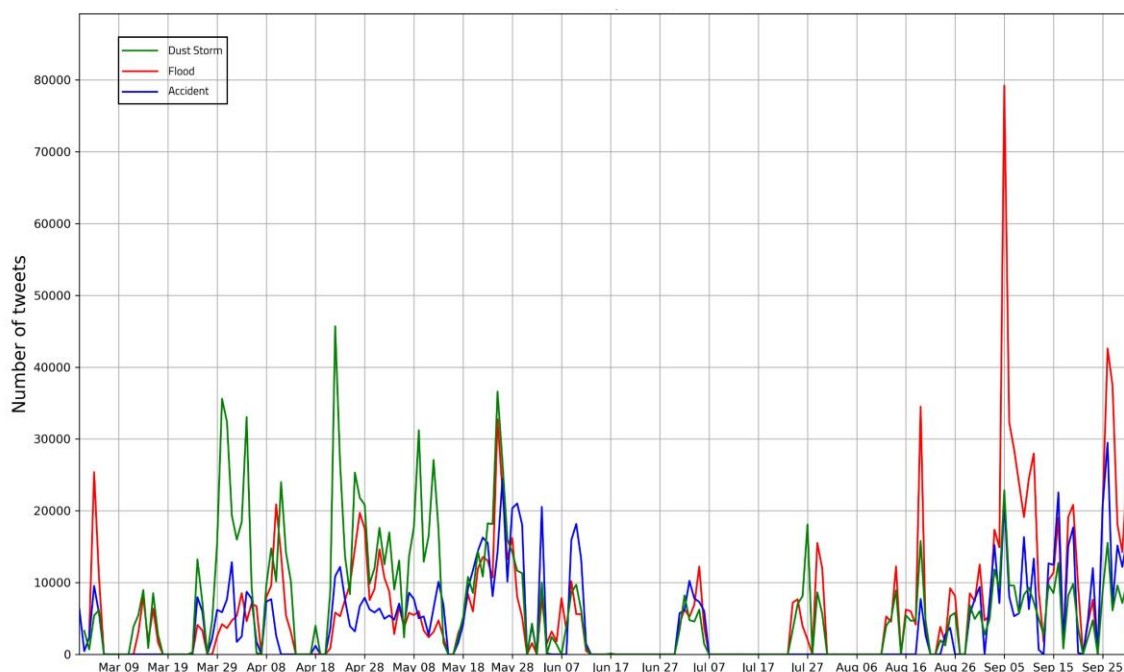[1] https://github.com/BatoolHamawi/FloDusTA

Figure1: Tweets collection per day between 1Mar 2018 to 30 Sep 2018

The labeling task in all the previous researches (Alsaedi and Pete ,2015; Almerekhi et al.,2016; Alhelbawy et al.,2016) was performed by crowdsourcing platform[2]. Both (Almerekhi et al.,2016; Alhelbawy et al.,2016) tweets datasets are publicly available for research purposes in the form of tweets IDs, and annotations. The datasets built in the previous researches focus on identifying events tweets without taking into consideration the events temporal indication mentioned in the tweets. Our dataset differs in that we distinguished the events time period and linked it with the type of events classes. We designed our dataset to be appropriate with the properties of flood, dust storm, and traffic accidents events by identify the event's occurring period. The labeling process with two stages annotations is performed by three Arabic native speakers. To the best of our knowledge, none of the works in the literature addressed the dust storm event as one of the event types to identify. The frequent occurrences of the dust storm event in Saudi Arabia inspired us to target it as an event to detect; since the annual occurring of dust storms ranges between 100 to 150 in a year (Albugami et al.,2019).

3.   Event Tweets Collection

In this section, the dataset building phase is divided into three main steps. In the first step, we collect tweets by keywords. In the second step, we apply certain cleaning and filtering on the collected tweets, then we perform two stages annotations in the labeling process.

3.1   Tweet Collection

Tweets were collected for seven months period, between March and September of 2018 from Twitter Streaming API, using searching terms. For each event type: flood, dust storm, and traffic accident, there were keywords list prepared. Figure1 illustrates the time-series of each event tweets collected per day. The collected tweets for all three

events consist of around 3.6 Million tweets (3,644,838). The tweets which were filtered by keywords for each event type: flood, dust storm, and traffic accident have reached about 1.3 Million (1,338,498), 1.3 Million (1,385,967), and 920,373 respectively.

3.2   Tweet Filtering

To construct a good quality dataset, tweets cleaning is applied to remove noise from the data. As our target goal is to detect an event written in Arabic language, non-Arabic tweets were filtered out. That is, we excluded the tweets that are written in other languages, since many of Arabic users' speakers may post tweets in different languages. Furthermore, the retweets were also removed to eliminate duplicated tweets. Since tweets were collected for each event separately, and there is a correlation between the events, any discovered duplicated tweets were filtered out as well. After cleaning was applied, we ended up with 894,277 tweets for all three events.

After the above filtering applied, we need to construct and prepare the collected tweets for event annotation. As our goal is to involve only the events occurred inside Saudi Arabia that were written either in modern standard Arabic (MSA), or in Saudi dialect. To this aim, several filters were performed on the collected tweets to exclude undesirable tweets. The tweets filters based on user location filter, hashtag filter, country filter, and time zone filter are described as follow:

User Location Filter: In the data collected, about 60% of tweets founded with the user location. Previous Arabic Datasets (Mubarak et al., 2014; Al-Twairesh et al., 2017) has been used to filter on user location for building dialect datasets. The user location field is user entry which could be written with a misspelling, either in Arabic or in English,

---

[2] http://www.crowdflower.com

using formal or informal writings. Accordingly, we filtered on three prepared lists of manually chosen locations that describe a place inside Saudi Arabia (city, street, province, etc.), whether it is written in Arabic or in English.

Hashtag Filter: The tweets were filtered on three manually chosen hashtag lists, all hashtags included in lists describe a place inside Saudi Arabia for example
(#حايل_الان, #غبار_القصيم).

Country Filter: Users in their Twitter profile can specify their country by selecting it from place dropdown list to select the geographic location. We found the country filed is useful for filtering based on. The tweets were filtered on the country name that had set to (Kingdom of Saudi Arabia, المملكة العربية السعودية).

Time zone Filter: The time zone describes the local time of a country and it was used as an indicator of location (Schulz et al.,2013). We filtered tweets on time zone that had set to Riyadh, as it was also used by Arthur et al (2018) in their approach, and by Al-Twairesh et al. (2017) to filter Saudi tweets.

The tweets resulted from the filtering step reached (85,894), (95,024), (39,270) for each event type, flood, dust storm, and traffic accident respectively. Then, to prepare the data for annotation, the tweets that have less than three words had been removed from the filtered tweets, to avoid the ambiguity and the difficulties which may cause confusion to the annotators in the labeling phase.

4. Annotation

In this phase, the data is filtered and cleaned. It is now ready for annotation, we asked three volunteer annotators to manually label a sample contains 9000 tweets which were randomly chosen. We randomly selected 3000 tweets from the filtered tweets that resulted from each event. By observation, we noticed that the tweet content of each event, have four classes: immediate, historical, warning and irrelevant, same classes that were identified by Arthur et al. (2018). Examples of classes are illustrated in table 1. The annotation process was performed in two stages. The annotators will determine whether the tweet is: immediate, historical, warning or irrelevant, then they will identify the type of event: flood, dust storm, or traffic accident. Thus, annotators will assign two types of labels per tweet. We call these labels, period labels (PL) and event labels (EL). To ensure the accuracy and reliability of labels, a set of instructions prepared with examples were given to the

Immediate	
Flood	Arabic: "سيل وادي الخرمة الان"
	Translation: Flood in Khormah valley now.
Traffic Accident	Arabic: "الان حادث مروري مروع على طريق #بيشه #الرين #الرياض بالقرب من الطريق السريع".
	Translation: Now a terrible traffic accident on #Bisha #Rin #Riyadh near the highway.
Dust Storm	Arabic: "موجة غبار الآن على محافظة #شرورة "
	Translation: Dust storm now in Sharoura province
Historical	
Flood	Arabic: " سيل الواسطة أمس قريب المغرب "
	Translation: Wasta flood yesterday near to Maghrib prayer.
Dust Storm	Arabic: "امس كان غبار ورياح شديدة جدا بعد العشاء "
	Translation: Yesterday was very strong dust and wind after dinner.
Traffic Accident	Arabic: " توفي شخص وأصيب 5 آخرون في حادثة مرورية مروعة وقعت مساء أمس الأول بين مركبتين اصطدمتا وجهاً لوجه على طريق بيشة "
	Translation: One person died and five others injured in a terrible traffic accident yesterday evening between two vehicles colliding head-on on the Bisha road
Warning	
Flood	Arabic: " الأرصاد تحذر من أمطار قد تؤدي إلى سيول على مكة المكرمة والمشاعر المقدسة"
	Translation: "Meteorological" warns of rain leading to floods on Makkah and Holy Masha'aer
Dust Storm	Arabic: "الارصاد تحذر من موجة غبارية الليلة على #الرياض و #المجمعة وألي حولها"
	Translation: #Metrological warns of a dusty wave tonight on #Riyadh and #Majma and all around
Traffic Accident	Arabic: " محطة الحرمين بعد اكسترا قبلها مطب جديد انتبهوا صار عنده عدة حوادث "
	Translation: Al-Haramain station after Extra and before the station there is a new bump Be aware there has been several accidents.
Irrelevant	
Non-event	Arabic: "يارب الشوط الثاني سيل من الاهداف"
	Translation: O Lord, let the second half has a flood of goals
	Arabic:"حادث مأساوي في #جدة .. طفل نزل إلى مسبح المدرسة فقضى غرقا"
	Translation: Tragic accident in # Jeddah .. A child came down to the school swimming pool and died drowned
	Arabic:"كلامك لا يشق له غبار"
	Translation: Your words do not shove by dust

Table 1: Tweet examples

annotators to follow. We have one and half hour session with annotators to instruct them with guidance and the examples given. The tweets were given to the annotators in an excel file which facilitates the annotation task.

4.1 Annotation guidance

The annotators are Arabic/Saudi native speakers. They were provided with labels description and guidance with examples explaining both the (PL) and the (EL) labels at the beginning of the annotation tasks. Table 2 summarizes the labels descriptions provided to annotators.

The guidance of labeling tweets presented to annotators is as follow:

(1) Significance: The event in the tweet content should be serious or significant; a joke, questions or poem, will be labeled irrelevant.
(2) Ambiguity: If the event in tweet content is not clear, don't try to predict the event, it is irrelevant.
(3) If a tweet content involves two events indications, dust storm, and accident, or flood and accident, where the cause of the accident is flood or dust storm; then the label is a traffic accident.
(4) If a tweet content involves flood and dust storm events, then the label is flood, because usually, the rain will follow dust storm.

The above labels description, guidance, and an additional list of key instructions were explained in detail along with tweets examples during annotation sessions.

4.2 Inter-annotator agreement

A random 3000 tweets sample was extracted from each collection of events: flood, dust storm, and traffic accident resulted with 9000 tweets for all. For each tweet, we provided two information (id, text). Each tweet is labeled with two labels: period label (PL), event label (EL). The text of each tweet is labeled by three annotators; we choose

three to avoid conflicts in annotation through majority voting. The annotators will assign two sub-labels for each tweet. The final majority label to be considered is the one that at least two annotators chose. After all tweets were labeled by the three annotators, each tweet will have one of the following labels: immediate-flood, immediate-dust storm, immediate-accident, historical-flood, historical-dust storm, historical-accident, warning-flood, warning-dust storm, warning-accident, and irrelevant. Due to the event type nature: flood, dust storm, and traffic accident that requires real-time detection, we adopted what was presented in Arthur et al. (2018) to focus on immediate classes. Therefore, each tweet will have one of these classes: flood, dust storm, and traffic accident which is immediate and a non-event. The non-event class includes (historical-flood, historical-dust storm, historical-accident, warning-flood, warning-dust storm, warning-accident, and irrelevant). To choose one label for each tweet based on labels assigned by annotators, we replaced historical and warning labels to a non-event. Then, the final majority label was obtained. We kept the labels (historical and warning) aside before replacement for future work.

To prove the reliability of the annotations, the inter-annotator agreement was measured using Fleiss Kappa measurement (Fleiss, 1971) which is used when the data item is labeled by more than two annotators. If the annotators assign the same labels for a tweet, this means that they agreed on the same decision about the tweet class, indicating that the guidance and examples were useful since the annotators share a similar understanding. The Fleiss kappa for 9000 tweets was 0.88 which is considered an almost perfect according to (Landis and Koch, 1977).

The result of all three collections of tweets annotations showed very high agreement between annotators; only two tweets were not agreed upon by all three annotators. The

Label		Description
Period (PL)	Immediate	An event that has occurred today, now, or it is still occurring.
		An event without time information mentioned in tweet.
	Historical	An event that has started and ended in the past.
	Warning	An alert, of future event will be occur.
		An expectation of event occurrence is live possibility.
	Irrelevant	Tweet content does not belong to any type of event, flood, dust storm or traffic Accident.
Event (EL)	Flood	An explicit indication in tweet content of flood event.
		An implicit indication in tweet content about the flood, such as flowing water after raining, or drowning streets and cars after raining, which mean that a flood occurred.
		Heavy rains included as flood event, as it caused severe floods, decreases the visibility and hinders traffic.
	Dust Storm	An explicit indication in tweet content of sand or dust storm event.
	Traffic Accident	An explicit indication tweet content of traffic accident event. including six categories, vehicle-vehicle collision, vehicle-fixed object (e.g., tree, or street lighting column) collision, vehicle-pedestrian collision, vehicle-animal collision, overturning of vehicle, and out-of-control accidents.
		An implicit indication in tweet content meaning that an accident happened in the road.
	Non-event	Tweet content does not belong to any type of event, flood, dust storm or traffic Accident.

Table 2: Labels description

not agreed tweets excluded from the dataset. Table 3 illustrate the statistics of the tweet dataset.

Labels	Non-event	flood	Dust storm	Traffic accident
No. tweets	4,708	1,556	1,348	1,386
Total	8,998			
Kappa	0.88			

Table 3: FloDusTA statistics

4.3 Labels Distribution

As mentioned before, tweets were collected separately for each event type, nevertheless, it might be found a flood event in dust storm collection or in accident collection and vice versa. According to the correlation between events, we try to solve this issue by annotation guidance with explaining supported by examples given to annotators. After annotators finished the annotation task, we compared between majority labels and found the following. In flood collection, four tweets were agreed by annotators as dust storm and only one tweet agreed as a traffic accident. Similarly, in dust storm collection, four tweets were agreed by annotators as flood and only one tweet was agreed as a traffic accident, while just two tweets, one in each, received a completely different annotation by the annotators. Besides, in traffic accident collection, the majority labels resulted with traffic accident events only, giving proof that the guidance was helpful and understandable. The disagreement took place mostly within dust storm collection, about 18% of tweets were agreed by two annotators, while it was about 9% and 8% of tweets within traffic accident and flood collections. Table 4 illustrates the distribution of the labels per event collection.

Collections \ Labels	Flood collection	Dust storm collection	Traffic accident collection
Flood	1,552	4	0
Dust storm	4	1,344	0
Traffic accident	1	1	1,384
Non-event	1,442	1,652	1,614

Table 4: Labels per event collection

Examples of annotated tweets founded in flood collection with dust storm and traffic accident labels:

Tweet (Example 1):

‎#انقلاب_سيارة في #سيول #حائل واصابة 13 من عائلة #السعودية

- Translation: #Car_overturned in #flood #Hail and 13 injured from a family #Saudi
- Annotation: Two annotators labeled this tweet as (immediate-accident) and one annotator labeled it as (immediate-flood).

Tweet (Example 2): ‎#امطار_مكه النواريه عندنا غبار وعج

- Translation: #Rain_Makkah Alnwariah we have dust and storm.

- Annotation: Three annotators labeled this tweet as (immediate-dust storm).

Examples of annotated tweets founded in dust storm collection with flood and traffic accident labels:

Tweet (Example 3): ‎أجواء بدأت بعاصفة رملية انقلب فيها بيت الشعر رأساً على عقب وانتهت بأمطار غزيرة والحمدلله

- Translation: The weather started with sand storm in which the house was turned upside and down and ended with heavy rain.
- Annotation: Two annotators labeled this tweet as (immediate-flood), and one annotator labeled it as (immediate-dust storm).

Tweet (Example 4): ‎موجة غبار تتسبب بحادثين على طريق رنية

- Translation: Dust storm causes two accidents on Rinnieh road.
- Annotation: Three annotators labeled this tweet as (immediate-accident).

In tweet case (1), the overall main context of tweet was reported about one of the traffic accident categories, overturning of a car and 13 injured where the reason for overturning is flooding. The majority label was suitable for the tweet and compatible with the guidance provided. In tweet case (2), the tweet is written in Saudi dialect. The hashtag talks about rain in Makkah while the tweet tells about dust and storm occurs. This tweet case represents the high noise founded in tweets, illustrating the difficulties faced during the annotation process. In tweet case (3) dust storm occurred first, then heavy rain, as a consequence, happened after. The majority label was suitable for the tweet and compatible with the guidance provided. In tweet case (4) two accidents happened because of dust storm. The label of tweet was agreed by three annotators as a traffic accident. Indeed, the description and guidance provided to annotators were well prepared and simplified, resulting in a high agreement in most cases.

5. Experiment

We built a supervised classifier model on **FloDusTA**. We performed multi-class classification using Support Vector Machine (SVM), the most used algorithm in text classification. SVM basically is a two-class binary classifier, finds the hyperplane that differentiates the two classes with largest margin. Subsequently, the binary SVM is modified to handle the multiclass classification problem. By building one SVM for each pair of classes, SVMs are trained to distinguish the samples of one class from the samples of another class resulting in k(k-1)/2 classes, which called "one-against-one approach" (Hsu and Lin, 2002). We conducted our experiment employing one-against-one approach and evaluated the prediction performance using 10-fold cross validation to obtain more accurate estimation of real model performance.

Before we train our classifier model, according to a noisy and informal form of tweet text, the following pre-processing steps were applied:

- Filtering: removing all punctuation marks, brackets, hyphens, numbers, symbols, URL links, mentions, and non-Arabic words.

- Tokenization: we split each tweet based on whitespaces.
- Normalization applied to unify the shape of Arabic letters: the letters "Allef"(أ) which has different forms in Arabic (إ آ أ ا), converted into (ا), and the letter "ta'a " "ة" converted to "ه". Moreover, to avoid noisy in the text, "tashkil" the vowel diacritics in Arabic removed from tweet words.
- Stop-words removal: removing unnecessary words such as "هذا" ,"و" and "من" which do not carry much specific information in the context of event detection.

Each tokenized word will be transformed into a form that can be used in the classification process by modelling in vector form by using term frequency and TF-IDF (Term Frequency – Inverse Document Frequency) which was used as a feature to train the classifier. To evaluate the SVM model, **FloDusTA** was divided into 10-fold cross-validation, one for testing and the rest nine for training the model and evaluating it on the test fold. The evaluation result will be the average of all 10 tests. We use precision, recall, and f1-score as evaluation metrics. Table 5 shows the mean scores for all the precision, recall, and F1 score. The results in table 5 demonstrate that TF-IDF achieves better f1-score, recall, precision compared with the term frequency.

Features	Precision	Recall	F1-score
Term frequency	0.866	0.857	0.858
TF-IDF	0.887	0.881	0.881

Table 5: Evaluation results

6. Conclusion

In this paper, we introduce **FloDusTA**, Saudi floods, dust storms, and traffic accident events dataset, containing both (MSA) and Saudi dialect written tweets. Our main aim is producing an Arabic dataset to be used for training and testing event detection systems. We describe the process of building the dataset, collection, filtering, and labeling. The dataset was manually annotated by three volunteers and the inter-annotator agreement was calculated using Fleiss kappa measurement. The dataset has an almost perfect inter-annotator agreement. The result of the experiment conducted on the **FloDusTA** as well as inter-annotator agreement showed high quality, and hence high potential of utilizing the dataset for training event detection system. **FloDusTA** will allow us to experiment a variety of different classification algorithms with different steps.

7. Bibliographical References

Arthur, R., Boulton, C. A., Shotton, H., and Williams, H. T. (2018). Social sensing of floods in the UK. *PloS one, 13*(1), e0189327.

Albugami, S., Palmer, S., Cinnamon, J., and Meersmans, J. (2019). Spatial and Temporal Variations in the Incidence of Dust Storms in Saudi Arabia Revealed from In Situ Observations. *Geosciences, 9(4)*, 162.

Alhelbawy, A., Massimo, P., and Kruschwitz, U. (2016). Towards a corpus of violence acts in arabic social media. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)* (pp. 1627-1631).

Almerekhi, H., Hasanain, M., & Elsayed, T. (2016). Evetar: A new test collection for event detection in arabic tweets. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval* (pp. 689-692). ACM.

Alsaedi, N., & Burnap, P. (2015). Arabic event detection in social media. In *International Conference on Intelligent Text Processing and Computational Linguistics* (pp. 384-401). Springer, Cham.

Al-Twairesh, N., Al-Khalifa, H., Al-Salman, A., and Al-Ohali, Y. (2017). AraSenTi-Tweet: A Corpus for Arabic Sentiment Analysis of Saudi Tweets. *Procedia Computer Science, 117*, 63–72.

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5), 378–382.

Gu, Y., Qian, Z. (S.), and Chen, F. (2016). From Twitter to detector: Real-time traffic incident detection using social media data. *Transportation Research Part C: Emerging Technologies, 67,* 321–342.

Hsu, C.-W., and Lin, C.-J. (2002). A comparison of methods for multiclass support vector machines. *IEEE Transactions on Neural Networks, 13(2),* 415–425.

Karagulian, F., Temimi, M., Ghebreyesus, D., Weston, M., Kondapalli, N. K., Valappil, V. K., … Abdooli, A. A. (2019). Analysis of a severe dust storm and its impact on air quality conditions using WRF-Chem modeling, satellite imagery, and ground observations. *Air Quality, Atmosphere & Health, 12*(4), 453–470.

Landis, J. R., and Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics, 33*(1), 159-174.

Mansuri, F., Al-Zalabani, A., Zalat, M., and Qabshawi, R. (2015). Road safety and road traffic accidents in Saudi Arabia. A systematic review of existing evidence. *Saudi Medical Journal, 36*(4), 418–424.

Mubarak, H., and Darwish, K. (2014). Using Twitter to Collect a Multi-Dialectal Corpus of Arabic. *Proceedings of the EMNLP 2014 Workshop on Arabic Natural Language Processing (ANLP)* (pp. 1-7).

Sakaki, T., Okazaki, M., and Matsuo, Y. (2010, April). Earthquake shakes Twitter users: real-time event detection by social sensors. In *Proceedings of the 19th international conference on World wide web* (pp. 851-860). ACM.

Schulz, A., Hadjakos, A., Paulheim, H., Nachtwey, J., and Mühlhäuser, M. (2013, June). A multi-indicator approach for geolocalization of tweets. In *Seventh international AAAI conference on weblogs and social media.*

Youssef, A. M., Sefry, S. A., Pradhan, B., and Alfadail, E. A. (2015). Analysis on causes of flash flood in Jeddah city (Kingdom of Saudi Arabia) of 2009 and 2011 using multi-sensor remote sensing data and GIS. *Geomatics, Natural Hazards and Risk, 7*(3), 1018–1042.