POLITECNICO
MILANO 1863

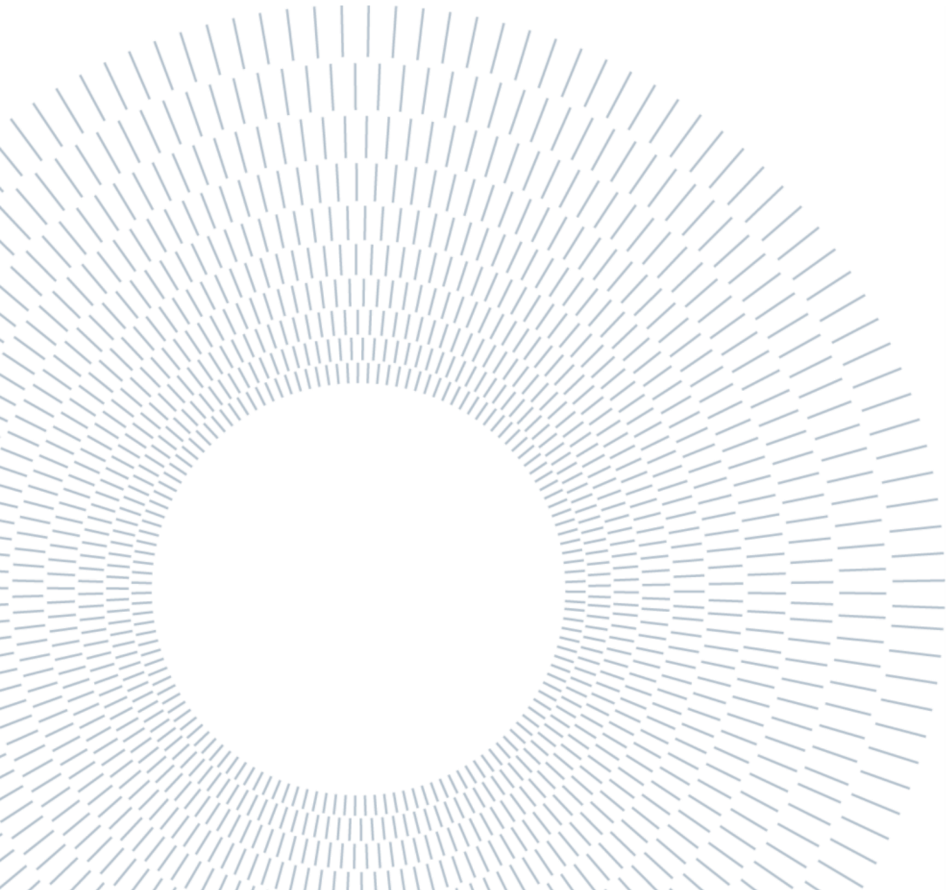
SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Analyzing drought impacts through social media and geophysical parameters

TESI DI LAUREA MAGISTRALE IN COMPUTER SCIENCE AND ENGINEERING - INGEGNERIA INFORMATICA

Author: **Violeta Marić**

Student ID:	970715
Advisor:	Prof. Barbara Pernici
Co-advisor:	Carlo Bono
Academic Year:	2022-23



Abstract

To be able to understand events like natural disasters, one of the potential starting points can be gathering information relative to their causes, characteristics, and impacts. Besides official traditional sources, there is a possible improvement from information coming from social media in the aspects of early warning and situational awareness, as well as assessment of the damage suffered by the affected area. In Italy, 2022 was revealed as a year of climatic extremes, in fact, the driest since 1800 according to studies. Unlike for example floods and earthquakes that express sudden and dramatic effects, drought impacts develop slowly over time and that is why they are often underestimated. This study addresses the problem that society, research, and emergency responders are facing due to a lack of data, public awareness, and interdisciplinary collaboration. The approach adopted is to leverage multilingual BERT model for the classification of tweets into multiple drought impact categories. It is an automated way of processing a large volume of data to give more insights into the distribution of various effects, trends, and patterns to effectively guide drought management. This case study focuses on worsening drought conditions in Italy, having as a goal to verify whether social media data can be considered valuable support to traditional sources. The tweet analysis results show a significant rise in the baseline rate of tweets in 2022, in comparison to the previous two years, which further adds to the significance of social media for raising awareness. Moreover, it was discovered that drought has a major effect on water scarcity which underlines the importance of water in drought conditions for the water cycle and ecosystem in general. Consequentially, agriculture suffers impacts from water scarcity since it is a key component for crop production. The findings indicate that the method can facilitate the identification of key impact indicators, as well as the assessment of the severity and extent of damage suffered. The study can contribute to the analysis of drought effects proving the distribution over the categories of impact. Ultimately, it evaluates the engagement of the public and verifies the possibility of using Twitter posts as an additional source of information to be included in drought management.

Keywords: disaster management, social media analysis, natural language processing, classification of drought impacts

Abstract in lingua italiana

Per poter comprendere eventi come i disastri naturali, uno dei potenziali punti di partenza può essere la raccolta di informazioni relative alle loro cause, caratteristiche e impatti. Oltre alle fonti ufficiali tradizionali, è possibile migliorare le informazioni provenienti dai social media per quanto riguarda l'allerta precoce e la consapevolezza della situazione, nonché la valutazione dei danni subiti dall'area colpita. In Italia, il 2022 si è rivelato un anno climaticamente estremo, di fatto il più secco dal 1800 secondo gli studi. A differenza, ad esempio, di alluvioni e terremoti che esprimono effetti improvvisi e drammatici, gli impatti della siccità si sviluppano lentamente nel tempo ed è per questo che spesso vengono sottovalutati. Questo studio affronta il problema che la società, la ricerca e i soccorritori si trovano ad affrontare a causa della mancanza di dati, di consapevolezza pubblica e di collaborazione interdisciplinare. L'approccio adottato è quello di sfruttare il modello BERT multilingual per la classificazione dei tweet in più categorie di impatto della siccità. Si tratta di un modo automatizzato di elaborare un grande volume di dati per fornire maggiori informazioni sulla distribuzione dei vari effetti, tendenze e modelli per guidare efficacemente la gestione della siccità. Lo studio si concentra sul peggioramento delle condizioni di siccità in Italia, con l'obiettivo di verificare se i dati dei social media possono essere considerati un valido supporto alle fonti tradizionali. I risultati dell'analisi dei tweet mostrano un aumento significativo del tasso di base dei tweet nel 2022, rispetto ai due anni precedenti, il che aggiunge ulteriore importanza ai social media per la sensibilizzazione. Inoltre, si è scoperto che la siccità ha un effetto importante sulla scarsità d'acqua, il che sottolinea l'importanza dell'acqua in condizioni di siccità per il ciclo idrico e l'ecosistema in generale. Di conseguenza, l'agricoltura subisce l'impatto della scarsità d'acqua, poiché è una componente fondamentale per la produzione di colture. I risultati indicano che il metodo può facilitare l'identificazione dei principali indicatori di impatto, nonché la valutazione della gravità e dell'entità dei danni subiti. Lo studio può contribuire all'analisi degli effetti della siccità, dimostrando la distribuzione delle categorie di impatto. Infine, valuta il coinvolgimento del pubblico e verifica la possibilità di utilizzare i post di Twitter come ulteriore fonte di informazioni da includere nella gestione della siccità.

Parole chiave: gestione delle catastrofi, analisi dei social media, elaborazione del linguaggio naturale, classificazione degli impatti della siccità

Acknowledgements

I would like to thank my family and friends for their love and support. Without them, this journey would not have been possible.

I am also grateful to my professor and supervisor for their feedback and guidance during this research.

Contents

List of Figures

List of Tables

List of Pictures

List of Listings

Introduction

Each year, millions of people are affected by natural disasters. Climate change is contributing to the frequency and intensity of some of it, leading to devastating effects. These effects are seen on infrastructure, economy, environment, and the most valuable resource - human lives. The inadequate infrastructure, the lack of necessary means and poverty are making the situation worse, and even developed countries are not bypassed. Proactive disaster management becomes a high priority requirement to deal with each phase that those events are going through. The crucial point is to understand the causes and impacts to be able to build resilient societies accordingly. Interpreting natural disasters properly requires a multidisciplinary approach, which consists of bringing together the information gathered from scientific, social, and environmental perspectives. Such a scenario typically involves multiple stakeholders that rely on various sources of information to guide the management. Traditional sources that include official reports, sensor networks or meteorological agencies are often not enough. They often provide one-way information flow from authorities to the public, neglecting the possibility to receive real-time feedback, address concerns, and gather information from affected individuals.

To fill the gaps that missing information during disaster is carrying along, citizen observations spread through social media are widely considered to be a promising source of relevant information, and many studies propose new methods to tap this source. In the study [1] the opportunities of spreading information via social media to fill the gaps are assessed as well as the risks of including them in the traditional acquisition process. In the field of impact assessment and model verification, the study concludes that excluding social media may prevent some impacts from being revealed and further estimated. On the other side, the risk coming from including social media to gather disaster information is that the impact assessment may be imprecise if data is too noisy. Additionally, it reveals major use cases for social media analysis including impact assessment, recruiting citizen volunteers, supporting weakly institutionalized areas, and narrowing surveillance areas. Social media could be included as a tool to mitigate the weaknesses of traditional sources for specific data needs in incidents and application scenarios. However, the essential point becomes the information coordination across various channels and ensuring consistent data flow can be complex.

Although social media is suggested to supplement existing on-the-ground efforts, particularly when in person needs assessments are hindered, it cannot provide a complete or representative picture of extreme weather events, especially in low-resource environments where people may not have access to technology [2]. For overcoming well known language and cultural barriers, tailor messages should reach all affected populations effectively. It has been also anticipated that these methodologies can be helpful for policymakers, disaster relief organizations, researchers and members of civil society who wish to leverage an additional tool to better understand the impacts of extreme weather events and how to focus their efforts. As identified in the study [1], the causalities between events and corresponding sub-events as well as assessment of response, recovery, and mitigation effort require further research.

The damage may be visible instantaneously, or only after months or even years. If the impacts are developing gradually over an extended period, it can pose challenges to their identification and severity prediction. Drought is one type of natural disaster that faces challenges mentioned and requires a lot of effort in assessing the impacts and damage suffered. The causalities of drought are a product of both the physical magnitude of the hazard and the ability to manage the potential disaster losses, including the systematic efforts to reduce exposure (prevention) and lessen vulnerability (mitigation) of people, livelihoods, and services [3]. It is one of the most complex natural phenomena, that is hard to quantify and manage, and has multiple and severe social and economic impacts. The pattern of various drought-related characteristics and impacts worldwide should reflect multiple aspects of drought, ranging from quantification of drought hazard and vulnerability of water resources systems - to measures of preparedness to face future droughts [4]. Another important motivation to target drought within this study was that the cumulative effect that drought has and the possibility to exacerbate the impacts over time, renders them more challenging to be addressed.

To quantify the severity and the extent of drought conditions, drought indicators are used by experts worldwide. Usually, they are grouped into meteorological, agriculture, hydrological, and composite based on the type of drought types. Existing longitudinal multi-sectoral datasets are limited in spatiotemporal homogeneity and scope, resulting in fragmented datasets [5]. There have been attempts to calibrate drought indicators to its various impacts and to connect indices to its impacts to proactively prepare people for it and mitigate drought. Drought can aggravate the impacts of other natural disasters therefore it is important to guide an effective disaster management strategy to reduce the risk of triggering another emergency. Having said that, it becomes essential to focus on guiding the response and decreasing future risks by understanding the challenges this natural disaster poses and providing appropriate solutions.

This work will investigate whether and to what extent social media can truly be beneficial in drought case to help overcome the lack of information that the researchers and emergency responders are facing. By cross-referencing information from multiple sources, it will try to evaluate user credibility during emergency events and confirm its perimeter. Furthermore, the study addresses some of the questions that were raised due to the lack of existing connections between an event and its effects. The contribution is also in helping the assessment of situational awareness reached during the detected natural disaster and guiding the response to perform the actions in categories where it left damage. Given that drought progression is slow and spatially extensive, an interesting set of questions arise, such as how the usage of Twitter by a large population may change during the development of a major drought alongside or how the changing usage facilitates drought detection [6]. To capture existing drought conditions, the work will consider assessing drought impacts, response, and its recovery, from Twitter user's perspective. In this way, it is easier to note viral trends and dynamics, as it provides more insights through user engagement metrics and other metadata.

Getting information about affected areas and assessing the severity of impacts is possible by applying some data analysis techniques to automatically extract insights and patterns from the data. They try to analyze the textual content of social media posts allowing identifying prevalent topics and critical terms associated with specific impacts or needs. User-generated data can fill information gaps by combining human intelligence together with artificial intelligence improving disaster management capabilities for labor-intensive manual tasks through employing digital volunteers [7].

A variety of different intelligence models can be trained for specific tasks [5] having as an overall goal to extract information or learn patterns and relationships from sources with large volumes of data, as social media platforms are. Recent advances in deep learning technologies have been applied to design disaster classification models [8]. To save significant computational resource and time that the training a model can require, researchers and practitioners can benefit from transfer learning, and apply the knowledge gained from solving one problem to another relatable problem. In that case, the pre-trained model can be applied to a similar task or a different dataset. Transformers-based language models are widely used in natural disasters domain where they demonstrate outstanding performance [9]–[11]. Recently, BERT has become broadly adopted due to its strong performance in short message processing

[10], and for pre-training on massive amount of unlabeled text data that enabled it to learn rich and contextualized representations of words.

The performance of pre-trained models is examined in various studies on the topic of natural disasters from high level point of view. However, in the area of drought, the research is more restricted in providing assessment of its impacts and possibly offer more insights into categories of impacts that require more attention and further assistance. Even though the demanding topic of drought is studied broadly, there were only a few research studies that dealt with social media generated data [6], [12].

The innovative approach adopted in this study is the combination of transformer-based language model trained on social media data and traditional sources for drought impact analysis. Additionally, unlike existing studies that examined drought through social media, this study case is concentrated on Europe, more specifically, on Italy. The aim of the study was to give its contribution in leveraging pre-trained models to try to classify drought relevant tweets into different impact groups and help in that way decision making nowadays, when drought conditions are reaching peak values. The choice to use BERT as representative of transformers-based language for this case study, was made due to the state-of-the-art results it showed in various natural language processing tasks, some of them also focused on natural disasters domain.

The study focused on the drought situation in Italy, mainly during 2021 and 2022, a period characterized by the anticipation of reaching peak values of drought severity. The specific time frame is set since the aim was to gain insights into the extent and impact of the drought conditions experienced in that period from Twitter users' point of view. The rationale behind is to have indications on potential implications for various sectors such as agriculture, economy, water supply, and societal well-being. The analysis will encompass multiple data sources, including official reports, social media data, and other relevant information, to provide a comprehensive understanding of the drought situation and its impacts in Italy during these critical years.

The tweets to use in this experiment are retrieved using keyword-based extraction for the years 2020, 2021 and 2022. Since social media contains an abundance of informal language and noise, with often lack any consistent formatting, the need for performing cleaning tasks is indisputable. Irrelevant parts of tweet content were neglected to not aggravate the performance of the classifier model. Taking into consideration the absence of domain-specific annotations, it has been proceeded with manual labeling of tweets related to drought. During annotation, each tweet was assigned to one or more categories of impacts. Alternatively, if the tweet has not been assigned to any of categories, it further indicated the irrelevance of the content.

The analysis then assesses the performances multilingual version of BERT in the area of multilabel text classification, while testing and experimenting with fine-tuning strategies to find the model architecture that fit the best this study case. The evaluation of multilingual BERT for drought-related tweets is proceeded by employing two methods:

1. applying the model trained on news media data on tweet dataset
2. using tweet datasets for both training and evaluation

Ultimately, the results obtained from the two approaches are review and evaluated with consultation of traditional sources and geophysical indicators, to gain more comprehensive perspective on drought conditions in Italy. The aim was to check whether surges of interest peaks were matching the traditional sources reports. Furthermore, the research tries to correlate the model results with the reports information to verify if they are synchronized and indicate the categories affected the most from drought and consequently guide the assistance in that direction. The main target was to try to provide useful information from Twitter analysis that can be further incorporated in standard decision-making procedures in drought management.

The motivation for this work to concentrate on new approaches to solve some of the problems that arise when dealing with effective drought management is explained by its constant need to be monitored and assessed, which is still not effectively managed, since slow onset characteristics and highly intertwined effects of drought [5]. The drought events tend to be longer-lasting and more severe, particularly in Mediterranean area [13].

The thesis is structured as follows:

Chapter 1 describes the state of the art, explaining in more detail the terms of drought, its indicators and impacts as well as disaster communication and social media usage in disaster management. It further defines tweet analysis, and natural language processing models commonly used for social media interpretation. The DIR dataset overview is also reported since it will be used further in the process.

Chapter 2 describes in more detail the scenario of drought conditions taking place in Italy during 2021 and 2022. Statistical overview of indicators measured for Italian territory are provided together with the reports from official institutions dealing with drought.

Chapter 3 explains the process of extraction and analysis of drought relevant content from Twitter. It reports data pre-processing tasks used for noise removal and content labelling activity performed on tweets to categorize them into categories of drought impacts.

Chapter 4 specifies the general BERT model and its multilingual version. Multilingual BERT is then trained with different approaches to find the most suitable which will be then applied on tweet dataset.

Chapter 5 provides overall results of the study from patterns observed in tweet datasets to the results achieved from the model. It further compares them with traditional data collected at the end of second chapter.

Chapter 6 concludes the work with observations and remarks acclaimed and proposes possible future development.

1. State of the art

1.1 About the drought

Unlike sudden weather events such as hurricanes, tornadoes, and thunderstorms, it is often difficult to pinpoint when drought has started or when it has ended. The initial effects of drought may be difficult to identify right away, so it may take weeks or months to determine it. Drought may last for weeks, months, or even years. Sometimes, drought conditions can exist for a decade or more in a region. The longer drought lasts, the greater the harmful effects it has on people [3], [13], [14].

Another difficulty is that drought means different things in different regions. A drought is defined depending on the average amount of precipitation that an area is accustomed to receiving.

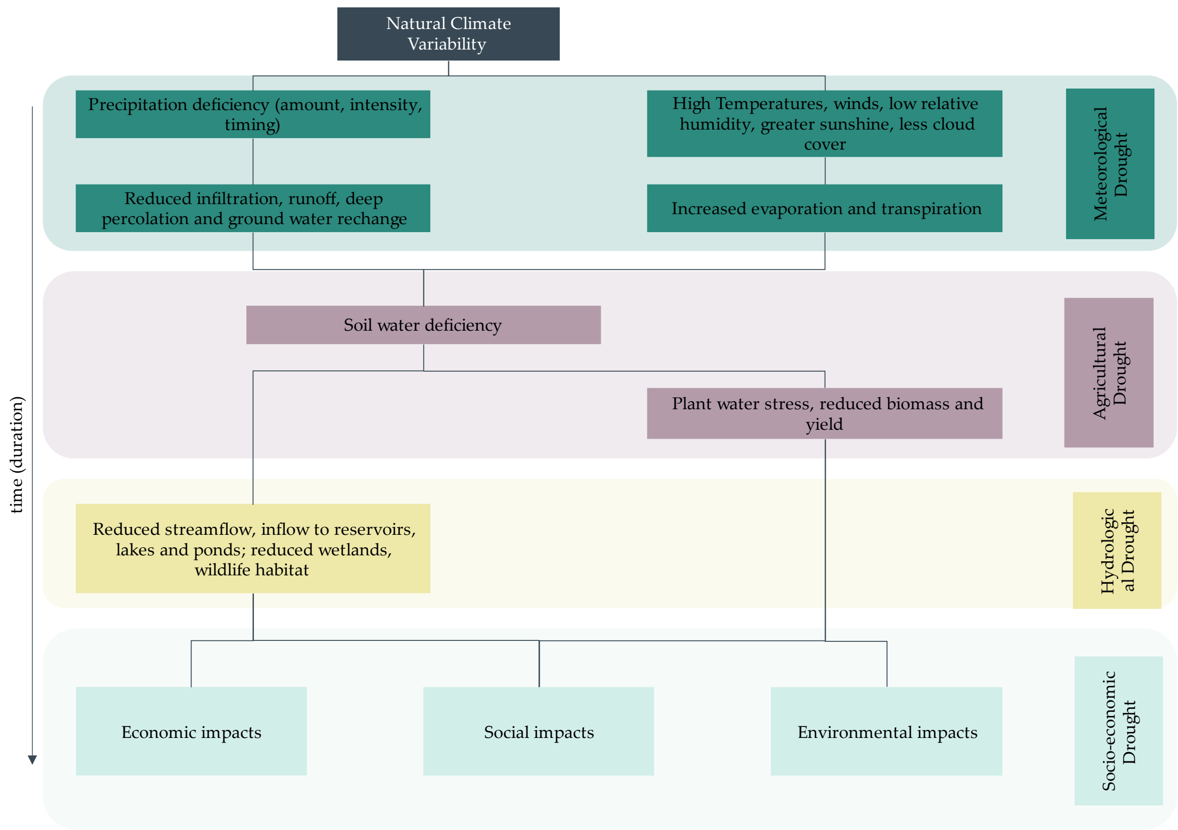
Drought definitions vary, depending on the variable used to describe the drought [15]. Hence, drought definitions can be classified into different categories in the following way:

(i) meteorological drought - a lack of precipitation over a region for a period of time

(ii) hydrological drought - a period with inadequate surface and subsurface water resources for established water uses of a given water resources management system

(iii) agricultural drought - a period with declining soil moisture and consequent crop failure without any reference to surface water resources

(iv) socio-economic drought - failure of water resources systems to meet water demands and thus associating droughts with supply of and demand for an economic good (water)



Picture 1.1: Classification of different types of droughts[1]

Because different types of droughts, such as meteorological, agricultural, and hydrological, have varied socio-economic consequences (*Picture 1.1*), a single physically measurable drought parameter for all these scenarios is not attainable [16].

When talking about reaction and response to drought, it often involves the taking of immediate decisions and actions by the government and the disaster organizations concerning the people and natural resources of the threatened region, to prevent widespread deaths of both human and animals, when a drought strikes. This precedes the pragmatic efforts to be taken, that constitute drought response. Drought response measures can be immediate, short-term, or long-term. They include all efforts such as assisting or intervening, at the time of, or soon after a drought occurrence, with the aim of preserving life and providing the basic needs of drought victims [17].

In the past decades, drought has been a recurrent feature of the European climate, with striking drought events in both high- and low-rainfall areas and in any season. From 2006 to 2010, 15% of the territory and 17% of the EU population was affected by drought on an annual basis, causing considerable damages and economic losses. [13].

[1] Source: National Drought Mitigation Center (NDMC)
`https://drought.unl.edu/Education/DroughtIn-depth/TypesofDrought.aspx`

There has been conducted various research on the topic of drought, trying to address some of the questions that arise while dealing with active drought management [3], [6], [18], [19].

Since the drought risk is increasing globally and the effective regulation of prevention and adaptation measures depends on drought hazard magnitude and its distribution for the future, the authors of the study [3] underline the necessity to progress on global initiatives that mitigate their impacts in the whole carbon cycle by late-century. They focused on assessment of drought hazard, as characterized by the likelihood of severe precipitation deficits in contemporary and future climates. The results show that there is a greater propensity to severe droughts in the Mediterranean ecosystems by the end of the century, which may be especially pertinent when specific impacts on human activities are taken into consideration, such as livestock farming, agricultural yields and household subsistence.

One of the studies [19] is investigating the relationships between the hydro-meteorological indicators and drought impacts trying to overcome the biggest drawback of data scarcity and complexity of it. The idea was to predict multi-category drought impacts and connect a drought indicator to the text-based impacts collected from the Drought Impact Reporter (DIR). The possibility is provided to appraise drought impacts using hydro-meteorological indicators with the proposed framework in the United States, which could help drought risk management by giving additional information and improving the updating frequency of drought impacts. The results revealed that the rules guiding the predictions of model comply with domain expertise knowledge around the role that some indicators play around drought impacts.

1.2 Drought indices

Droughts are fundamentally characterized in three dimensions: severity, duration, and spatial distribution [20].

By quantifying severity levels and declaring drought's start and end, drought indices currently aid in a variety of operations including drought early warning and monitoring and contingency planning. Given their variety and ongoing development, it is crucial to provide a comprehensive overview of available drought indices that highlights their difference and examines the trend in their development [20].

The nature of drought indices reflects different events and conditions; they can reflect the climate dryness anomalies (mainly based on precipitation) or correspond to delayed agricultural and hydrological impacts such as soil moisture loss or lowered reservoir levels. Using this relatively simple methodology, drought indices have developed into the primary tool for communicating drought levels among involved entities [20].

The available drought indices reflect the variability in perceptions about drought. This includes the basic definition of drought, which changes among different applications.

For example, agricultural drought primarily focuses on absent soil moisture content, while hydrological drought examines the lagged effects of precipitation deficiency on various water features [20].

The study [17] finds it crucial to have indices, models and tools that help in assessing vulnerability to drought, provide reliable information and understanding about the occurrence of drought to develop suitable measures to mitigate drought.

By quantifying severity levels and declaring drought's start and end, drought indices currently aid in a variety of operations including drought forecasting, declaring drought levels, contingency planning, and impact assessment. In addition to the variability in the types and applications of droughts the dissociation of drought indices with drought impacts has prompted calls for aggregate drought indices to cover more aspects and applications [20].

The following are the brief descriptions of two relevant ones that are going to be referenced in the work and then used further during the analysis of drought impacts and comparison with the results obtained.

1.2.1 SPI (Standard Precipitation Index)

The SPI was introduced as a measure of the precipitation deficit that is uniquely related to probability. It can be calculated for any accumulation timescale, usually from monthly precipitation observations, and is typically expressed as SPI-n, where n is the number of months of accumulation [21].

SPI is using monthly time series of rainfall data preferably (30 years or more) to compute its value. SPI measures the standardized departure that the observed precipitation on a given time scale deviates from the long-term average precipitation. Based on SPI values, meteorological droughts are classified into different levels of severity, as shown in *Table 1.1*.

SPI value	**Drought Category**
2.0 < SPI <= MAX	Extremely wet
1.5 < SPI <= 2.0	Very wet
1.0 < SPI <= 1.5	Moderately wet
-1.0 < SPI <= 1.0	Normal precipitation
-1.5 < SPI <= -1.0	Moderately dry
-2.0 < SPI <= -1.5	Very dry
MIN <= SPI <= -2.0	Extremely dry

Table 1.1: SPI classification scheme used in [22]

It can be computed at any number of time scales, from 1 month to 48 months or longer (typically up to 24). SPI values for 3 months or less might be useful for basic drought monitoring, values for 6 months or less for monitoring agricultural impacts and values for 12 months or longer for hydrological impacts. Drought events are indicated when the results of SPI, for whichever timescale is being investigated, become continuously negative and reach a value of −1. The drought event is considered ongoing until SPI reaches a value of 0 [23].

The greatest strength of SPI is that it has a wide application range and uses precipitation data only which makes it very easy to use and calculate. However, with precipitation as the only input, SPI is deficient when accounting for the temperature component, as well as damage of cultivated and natural ecosystems, and increasing evapotranspiration and water stress. This drawback can make it more difficult to compare events of similar SPI values but different temperature scenarios [24], [25]. However, The World Meteorological Organization is highlighting SPI as a starting point for meteorological drought monitoring [23].

Another weakness of using SPI only is that it does not consider the annual precipitation variability in estimating drought intensity. In general intensity values computed from the cumulative sum of normalized precipitation deficits during a dry season can be as extreme as during a rainy season. Since the timing of consecutive precipitation deficits relative to the local hydrologic cycle has more impact on the natural ecosystem and human activities than the seasonal or annual precipitation totals [3].

In the study [19] the authors-built framework to predict multi-category drought impacts and connected SPI to the text-based effects received from the Drought Impact Reporter (DIR). The possibility is provided to appraise impacts using hydro-meteorological indicators with the proposed framework in the United States, which could help drought risk management by giving additional information and improving the updating frequency of drought impacts. The results reveal that SPI-12 had the greatest impacts on the society and public health model, and that negative SPI-6 and SPI12 values might better explain the occurrence of drought impacts on society and public health more so than other indicators.

1.2.2 FAPAR (Fraction of Absorbed Photosynthetically Active Radiation)

This index is used for detecting and monitoring the impacts of agricultural drought on the growth and productivity of vegetation, especially plant water stress, implemented in the Copernicus European Drought Observatory (EDO). Its basic characteristics are reported in *Table 1.2*.

Variable	Temporal scale	Spatial scale	Coverage
FAPAR	10 days (= 1 dekad)	1 km	Europe

Table 1.2: Detailed description of FAPAR variable

FAPAR value	Interpretation
2.0 < FAPAR <= MAX	
1.5 < FAPAR <= 2.0	Photosynthetic activity higher than normal
1.0 < FAPAR <= 1.5	
-1.0 < FAPAR <= 1.0	Near normal conditions
-1.5 < FAPAR <= -1.0	
-2.0 < FAPAR <= -1.5	Photosynthetic activity lower than normal
MIN <= FAPAR <= -2.0	

Table 1.3: FAPAR classification scheme used in [26]

Table 1.3 presents the classification of FAPAR into different categories based on its value. Plant water stress caused by drought affects the capacity of agricultural crops and natural vegetation to intercept solar radiation, thereby reducing vegetation growth rate. FAPAR is a biophysical variable, derived from satellite observations, that represents the fraction of incident solar radiation that is absorbed by land vegetation for photosynthesis. Its values and their anomalies have been shown to be good indicators for detecting and assessing drought impacts on plant canopies, such as agricultural crops and natural vegetation, and thus provide information that is potentially useful for water and agricultural management purposes [26].

For every 10-day period (starting from January 2001), the FAPAR anomalies are computed as follows:

$$FAPAR\ anomaly_t = \frac{X_t - X}{\delta}$$

where X_t is the FAPAR of the 10-day period t of the current year, $\bar{X}$ is the long-term average FAPAR and δ is the standard deviation, both calculated for the same 10-day period *t* using the available time series [26].

It must be considered that this indicator shows variations in the vegetation health and/or cover though, which may be due to rainfall or soil moisture deficits, but may also be due to other stress factors, such as plant diseases. Therefore, this indicator must be used jointly with other indicators giving information on the deficit of rainfall and / or soil moisture, to determine if the variation in the vegetation response (FAPAR) is linked with a drought event or not [26].

1.3 Drought impacts

The recent disasters in developing and developed countries, and the personal hardships have underscored the exposure and vulnerability of all societies to drought [3]. Assessing the impacts of a drought involves a comprehensive evaluation of its effects on various aspects of life. Acquiring a better understanding of drought impacts itself becomes increasingly vital under a warming climate. Drought researchers find a difficulty in assembling fully representative quantitative drought impact data, because impact data either tend to be narrowly focused on single-sector results such as crop yield or exist across collections of painstakingly assembled, qualitative, text-based reports that lend themselves to statistical analysis as presence–absence data [18]. Traditional drought indices describe mainly biophysical variables and not impacts on social, economic, and environmental systems. Improving insight into other types of drought impacts can strengthen societal resilience to more severe droughts under a warming climate. A primary challenge of similar studies is that the input data sets for calculating indices fall short of capturing the diversity of drought impacts in social and economic sectors such as water supply and small business. Therefore, most current drought indices would underestimate the severity and scope of drought impacts in the human dimension due to the limitation of data [12].

Disaster management literature commonly distinguishes between rapid-onset disasters such as storm surges and earthquakes, which cause immediate loss and disruption, and slow-onset events, notably drought. Determining the start of a drought can be tricky. Unlike many natural hazards that bring about sudden and dramatic results—such as earthquakes, tornadoes, and hurricanes—the onset of a drought can be gradual and subtle. It can take weeks, months, or even years for the full effects of long-term inadequate rainfall to become apparent [13], [14]. Because the full effects of a drought can develop slowly over time, impacts can be underestimated.

Of all extreme weather types, droughts have one of the largest impacts on society and the environment. In fact, it is ranked second in the type of natural phenomena associated with billion dollars weather disaster during the past years. It is estimated that in EU countries the number of people affected by drought has increased by 20% over the last decades [27].

However, drought management in most parts of the world is still reactive, responding to drought after impacts have occurred. This approach, commonly referred to as crisis management, is known to be untimely, poorly coordinated and disintegrated. Moreover, the provision of drought relief or assistance to those most affected has been shown to decrease socioeconomic capabilities to face future drought episodes by reducing self-reliance and increasing dependence on government and donor organizations [28].

The approach for monitoring drought anomalies and some of its impacts is hardly comparable. Scientists should engage more closely with the affected parties (farmers and water-reliant sectors) end-users in a multi-disciplinary manner to establish a holistic drought monitoring system. This strategy would help by identifying the current weak links and suggesting future mitigation strategies. Consequently, it is necessary to appraise drought disasters by incorporating climate information, environmental and economic implications of drought in the study area and the surrounding environments, this will help in identifying the contributing factors and the actual impacts of its occurrences in the region [16].

Drought risk is lower for remote regions, such as tundra and tropical forests, and higher for populated areas and regions extensively exploited for crop production and livestock farming, such as South-Central Asia, Southeast of South America, Central Europe and Southeast of the United States. As climate change projections foresee an increase of drought frequency and intensity for these regions, then there is an aggravated risk for global food security and potential for civil conflict in the medium-to long term [28]. While drought is a naturally occurring part of the weather cycle and cannot be prevented, human activity can influence the effects that drought has on a region.

Assessing and reducing vulnerability to impacts affecting the environment, society, and the economy for regions beyond the local scale, spanning political and sectoral boundaries, requires systematic and detailed data regarding impacts [29].

To reduce the damage from drought, it is crucial to characterize droughts. Drought characterization enables operations such as drought early warning and drought risk analysis, which allow improved preparation and contingency planning [20].

1.3.1 Classification of the drought impacts

Impacts of natural hazards can be classified as direct or indirect on a high level. Examples of direct impacts are the physical destruction of buildings, infrastructures, crops, or other natural resources. In this context, drought can cause loss of life, reduce crop yields and limit public water supply. Indirect impacts are related to the indirect consequences of the destruction of natural resources. These include temporary rural unemployment or business interruption. In extreme cases, droughts may result in malnutrition, starvation, and disease in the more vulnerable countries [13]. Even if most of the drought impacts are indirect, they can propagate quickly through the economic system, also affecting regions far from the origin.

While large economic impacts of droughts are most relevant in wealthy industrialized nations, its social impacts are particularly severe in food-deficit countries with high dependence on subsistence agriculture and primary sector activities. In such cases, drought events combined with poor governance and poorly functioning market systems, oppressive policies, and intermittent or insufficient food aid, has historically led to food insecurity, famine, human conflicts and widespread mortality [28].

The results from a study [29] show that impacts on agriculture and public water supply dominate the collection of drought impact reports for most countries and for all major drought events since the 1970s, while the number and relative fractions of reported impacts in other sectors can vary regionally 15 and from event to event. The data also shows that reported impacts have increased over time as more media and website information has become available and environmental awareness has increased.

Droughts impact both surface and groundwater resources and can lead to reduced water supply, deteriorated water quality, crop failure, reduced range productivity, diminished power generation, disturbed riparian habitats, and suspended recreation activities, as well as affect a host of economic and social activities [15].

Categories are agriculture, energy, plants & wildlife, society & public health, water supply & quality, business & industry, fire, relief, response & restrictions, and tourism & recreation [12].

The impacts of categories are reported next, to understand better the complex relationship they have with each other.

1.3.2 Agriculture

The primary direct economic impact of drought in the agricultural sector is crop failure and pasture losses. Increasing demand for food and environmental stressors are some of the most challenging problems that human societies face today, and these have encouraged new studies to examine drought impacts on food production [16]. In recent decades, significant progress has been made in sustaining global food production. Nonetheless, feeding 9.8 billion people by 2050 would be a challenge, particularly in drought-prone and arid regions of the developing world. Food production shocks (i.e., unexpected losses and increases in price) have been more common in all sectors including food industries during the last five decades. Extreme weather causes half of these shocks, with disproportionate effects on countries with little coping capability, such as farmers' ability to diversify food production or governments' ability to import food or provide insurance. Understanding the interactions between drought and food security is critical for policymakers and stakeholders to develop adaptation policies that effectively reduce the effects of drought on agricultural production and increase societal resilience to future drought-induced emergencies, all while meeting competing demands and enhancing environmental sustainability.

Most specialty crops (such as fruits, vegetables, tree nuts, and medicinal herbs) are more vulnerable to drought than field crops and have a higher value per unit of land/water. They may therefore represent a higher risk for experiencing economic loss in drought if the crop water demand exceeds water supply. The costs are often passed on to consumers through increased prices and/or they may be offset through government disaster assistance programs.

Indirect impacts of drought in the sector can include reduced supplies to downstream industries, such as food processors, and reduced demand for inputs, such as fertilizer and farm labor. The non-market impacts of production losses include mental health strain on farmers [30].

Assessing agricultural drought and its potential impacts on food security in vulnerable regions is very crucial, especially in drought-prone areas. The implications of agricultural droughts on food supplies may be quantified, which helps policymakers make more sustainable agricultural decisions. It necessitates a thorough evaluation of the relationships between spatiotemporal drought fluctuations, farming systems, irrigation effects, and water resource availability. Various techniques of dealing with such issues have been reported. Survey methodology, for example, is useful for gathering first-hand information on how the drought has affected crop production and how farmers have reacted to drought [16].

Since most agricultural regions show high infrastructural vulnerability to drought, then regional adaptation to climate change may begin through implementing and fostering the widespread use of irrigation and rainwater harvesting systems. In this context, reduction in drought risk may also benefit from diversifying regional economies on different sectors of activity and reducing the dependence of their GDP on agriculture [28]. On the other hand, although new irrigation techniques have increased the amount of land that can be used for farming, they have also increased farmers' dependence on water [14].

1.3.3 Economy & Industry

Higher temperatures that often coexist with drought can impact roads, airport runways, and rail lines. During drought conditions that result in low water levels on rivers and other waterways, port and water-borne transportation operations may be limited due to a reduction in available routes and cargo-carrying capacity, resulting in increased transportation costs. Any increased shipping costs may be passed on to consumers, resulting in higher retail prices for goods [13], [30].

Airport runways are also vulnerable to extreme heat, which can cause asphalt to soften and deteriorate. Some airplanes themselves cannot fly in extremely high temperatures. It can also cause rail lines to buckle ("sun kinks"), causing derailments. When water supplies are depleted in drought, the ground can sink as more groundwater is removed. This affects infrastructure, including roads, buildings, and water pipes, and can lead to the formation of sinkholes.

A study was conducted in 2020 to estimate the total damage caused by agricultural droughts in the Italian economy. As a result, it was obtained that it can range from 0.01–0.10% of Italian GDP, that is, from approximately €0.55 to €1.75 billion. These damages concentrate but extend beyond the agricultural sector, with substantial identified impacts on food industry manufacturing and wholesale and trade services [31].

1.3.4 Energy

Drought and water scarcity present unique challenges for the energy sector. All sources of energy require water in their production processes, and energy is required to extract, convey, and deliver water.

As hydroelectricity production is related to the amount of water stored in the upper reservoirs, the production level can be lower during a drought. When water levels in reservoirs become low, the force of water pressure required to turn hydro turbine blades is reduced, which affects productivity. Peak demands for electricity then need to be satisfied by other means available in the short term (e.g., gas turbines). The number of losses depends on hydroelectricity infrastructures and drought severity [13].

High temperatures that often accompany and exacerbate drought affect the energy supply chain, reduce biofuel feedstocks, and increase the risk of wildfire, which can impact energy infrastructure.

Thermoelectric power plants use steam turbines to generate electricity using a variety of fuel sources. Large amounts of water are needed to generate steam and for cooling. Drought conditions can result in reduced plant efficiency and generation capacity and can also impact the supply chain for coal, natural gas, biofuel, and nuclear fuel.

During droughts, hydraulic fracturing (or fracking) and fuel refining operations can require alternative water supplies or may be forced to temporarily shut down. Shutdowns can increase costs, which in turn can raise consumer prices.

1.3.5 Tourism & recreation

Drought impacts the tourism and recreation sectors both directly and indirectly and affects the sectors during all seasons. Since many activities in the tourism sector are water-related, droughts can bring critical losses. Lower water levels or snowpack affects the availability of recreational activities and associated tourism, and a resulting loss of revenue can severely impact supply chains and the economy locally, regionally, and potentially nationally [30]. Drought directly affects snow sports, such as skiing and snowmobiling, and activities conducted on rivers and lakes, such as boating, rafting, canoeing, fishing, and swimming, due to reductions in snowpack and streamflow. In addition, activities such as biking, hiking, and camping also rely on sufficient water. Drought conditions can result in shortened or shifted seasons for these activities [30].

Reduced revenues in the tourism sector can negatively impact the livelihood of communities and the many small outdoor recreation businesses that have limited resources to manage the financial burden of drought. This, in turn, impacts the mental health of small business owners, staff, and communities [30].

1.3.6 Fire

The relationship between drought and fire is complex. The timing, intensity, and frequency of drought events have divergent impacts on fuel flammability and fire behavior. Drought, combined with unusually warm temperatures, very low precipitation and with dried out (and more flammable) vegetation, can result in decreased snowpack and streamflow, increased evaporative demand, dry soils, and large-scale tree deaths, which results in increased potential for large wildfires.

According to a recent study published in the scientific journal Nature, these intense heat waves are causing irreversible effects also on European forests, generating a vicious circle which limits the emission absorption capacity. An element not to be underestimated is that taking into account the fact that 41 million km2 of forests (equal to about 30% of the surface of the earth) play a fundamental role in the global cycle of carbon, absorbing 33% of the carbon emissions with a crucial function in the mitigation of climate changes.[32].

Due to lack of precipitation, in some instances, drought reduces the potential for wildfires by reducing the amount of vegetation available to burn [30]. There are also actions that individual homeowners can take to create a defensible space, an area around a building/property in which vegetation, debris, and other types of combustible fuels have been treated, cleared, or reduced to slow the spread of fire to and from the building.

When wildfire hits in drought-stricken areas, watersheds and reservoirs can be further impacted by ash and debris flows, water treatment facilities may shut down with damage or loss of power, crops can be destroyed, and smoke can affect animal and human health.

1.3.7 Plants & Wildlife

Drought affects the environment in many ways. Plants and animals depend on water, and under drought conditions their food supply can shrink, and their habitat can be damaged. Sometimes the damage is only temporary and their habitat and food supply return to normal when the drought is over. On the other side, if drought impacts on the environment last longer time it may lead to permanent land degradation. In general, the local and regional biodiversity may be highly affected by drought events [30]

Trees and other plants have adapted to withstand the effects of drought through various survival methods. Some plants (such as grasses) will slow their growth or turn brown to conserve water. Trees can drop their leaves earlier in the season to prevent losing water through the leaf surface. However, if drought conditions persist, much vegetation will die [14].

However, many organisms cannot adapt to drought conditions, and the environmental effects of extended, unusual periods of low precipitation can be severe. Negative impacts include damage to habitats, loss of biodiversity, soil erosion, and an increased risk from wildfires [14].

1.3.8 Society & Public Health

Welfare changes experienced by human beings should be accounted for in the measures of the socioeconomic impacts of drought. The social impacts of drought can affect people's health and safety, cause conflicts between people when water restrictions are required, and may result in changes in lifestyle [13]. Drought can cause significant human health outcomes that can challenge public health departments, emergency managers, and healthcare providers. Drought can lead to decreased water quantity and quality, increased incidence of illness or disease, increased mortality rates, and adverse mental health outcomes as livelihoods are challenged.

Drought can increase the risk of disease. For example, drier conditions can increase reproduction of a fungus found in soils and lead to the disease coccidioidomycosis, or Valley fever. It can also irritate bronchial passages and lungs and exacerbate chronic heart and lung conditions due to wildfires and dust storms. Moreover, drought and its economic consequences can lead to increased mental health impacts, including mood disorders, substance abuse, domestic violence, and suicide.

1.3.9 Water Supply

Drought conditions impact water supplies by decreasing supply and increasing demand for various usages (industrial, agriculture or residential use). Direct impacts of droughts on surface waters include reduced river flows and reservoir levels. Significant decreases in aquifer levels are the main expression of drought impacts on groundwater. These decreases can also be accompanied by an increased risk of sea intrusion and eutrophication. Due to the low oxygen concentration in shallow rivers and lakes which experienced eutrophication, fish mortality increased [13]. Drought can result in significant operational impacts to water utilities, including a loss of water supply and poor source water quality. When water supplies are depleted in drought, increased use of groundwater can cause the ground to sink (called land subsidence). Subsidence damages infrastructure and can cause sinkholes.

As anticipated, the depletion of water availability in soils causes significant declines in crops and livestock productivity. In addition, surface and groundwater supplies may decline during drought, affecting water availability and increasing costs to access water for crops or forage irrigation and watering livestock. With a return to normal precipitation, soil moisture typically recovers long before surface and groundwater supplies are replenished.

1.4 Disaster communication and management

Communication is a core component of disaster planning, response, and recovery. Effective disaster communication may prevent a disaster or lessen its impact, whereas ineffective disaster communication may cause a disaster or make its effects worse. Disaster communication typically has been thought of as occurring principally via the mass media and it generally consists of disaster warning messages and news coverage of disasters. Indeed, several works based in disaster management field performed the analysis or model training on news articles [5], [12], [19], [33]. However, mass mediated coverage of disasters is limited since it normally involves messages created by a single source and disseminated to large audiences, with little opportunity for audience response and participation [34]. Furthermore, it fails to capture the characteristics that newer forms of technology contain.

The communication of information about natural hazard risks to the public is a difficult task for decision makers. Research suggests that newer forms of technology present useful options for building disaster resilience [35]. However, how effectively these newer forms of media can be used to inform populations of the potential hazard risks in their community remains unclear. An analysis has been conducted to assess two predominant routes of dissemination for risk information: older traditional media and newer social media sources. This analysis revealed that age predicts the use or preference for kind of risk information source. Given population dynamics and the study's results about different preferences by age, it was concluded that effective risk communication will require multiple communication in the future to ensure that all groups are informed about potential hazards.

1.4.1 Social media for disaster management

To have a new lens into the physical world through the eyes of the social network [36], social sensing has been proposed in the literature as a term for describing the gathering of information from humans using crowdsourcing, human-connected devices (mobiles, etc.) and/or by extracting information from social media, with the goal of mining social signals to gather situation awareness and support decision making [37].

Even if citizen-generated data are considered a nontraditional data source, it can be beneficial to use them as a complement to the official data sources, that are often costly to produce in terms of both time and resources [37]. Compared to traditional media, web based social media technologies are characterized by greater capacity, dependability, and interactivity, each of which may be advantageous for disaster communication [34]. The main issue in this case is the quality of the collected data [37].

While satellite imagery is still a valuable source for disaster management, it can be introduced only with a certain delay that varies from at least 48-72h. Information products can be improved through complementing them with user-generated data like social media posts or crowdsourced data. The advantage of these new kinds of data is that they are continuously produced in a timely fashion because users actively participate throughout an event and share related information [7].

In recent publications, new data sources like social networks were added to disaster management workflows providing additional information for emergency managers [7], [8], [34], [38].

Before, during, and after major disasters, coordination of emergency response is an enormous problem due to the number of individuals and organizations involved in the response, issues with the interoperability of technology, impacts of the disaster on technologies used for communications, problems with adequate information sharing, and the lack of pre-existing social networks in place to support community response, among others [39].

The evolution of new communication technologies, such as social media technologies, offers more opportunity to bypass drawbacks having a one-way mediated communication, and gives a possibility also to audience to participate in it. As a result, the promise of richer disaster communication via social media has captured the attention of disaster communicators. In the study [34] social media uses were examined across three disaster phases: pre-event, event, and post-event. For pre-event, the cases identified were providing and receiving disaster preparedness information and warnings, for the event itself, sending and receiving requests for help and informing others about one's own condition. Uses of social media considered for both event and post-event, were documenting disasters, raising awareness of the event, identifying ways to assist, and expressing emotions and concerns. The cases that belong exclusively to post-event were discussing socio-political and scientific causes and implications of and responsibility for events. From the use cases listed, the study suggests the ways that social media can be employed to inform and improve disaster operations.

A comprehensive review about social media communication in disaster events was published with a framework demonstrating various entities that utilize and produce disaster social media content. Real-time crisis maps were provided based on social media data. The social media data were geocoded with locations from gazetteer, street map and user-generated data and reached remarkable results in comparison to post-event impact assessments from national civil protection authorities [7].

Starting from 2016, a 3-year long research project was conducted with an aim of demonstrating the technical and operational feasibility of the integration of social media analysis and crowdsourced information within Copernicus Emergency Management Service (EMS). The purpose of a new component that was developed as prototype was to improve the timeliness and accuracy of geo-spatial information provided to Civil Protection authorities, on a 24/7 basis, during the overall crisis management cycle and, particularly, in the first hours immediately after the event. As a result, it was achieved an early confirmation of alerts from running Early Warning Systems as well as first rapid impact assessment from the field [40].

The analyses of large volumes of data with a wide variety of information can be further improved leveraging Citizen Science methodologies and tools, that further engages members of the public, often in collaboration with or under the direction of professional scientists and scientific institutions [38]. The results suggest that a flexible approach to tools composition and configuration can support a timely setup of an analysis project by citizen scientists, especially in case of emergencies in unexpected locations.

1.5 Twitter analysis

The content can appear from a variety of social media platforms, the most popular of which is Twitter. As a matter of fact, Twitter is regarded as the 'most useful social media, tool', particularly for natural disasters [11]. The reasons are laying in its characteristics that make him stand out from the others. Twitter has a fixed format of the post with a limited number of characters allowed. Moreover, the usage of hashtags allows easy categorization and searchability of relatable information, rendering possible for users to follow and contribute to specific conversation about the topic. The thing that is making all listed features indeed advantages is the fact that Twitter is widely adopted by news, emergency responders and public generally.

The results of the study [41] that investigated Twitter usage during a hurricane event in 2012, confirms it is a highly valuable source of disaster-related information particularly during the power outage. With a substantial increase in the number of tweets and unique users during Hurricane Sandy, a large number of posts contained firsthand information about the hurricane showing the intensity of the event in real-time. The findings provide insights into the choice of keywords and sentiments and identifying the influential actors at different stages of disasters. The connectivity of the influencers and their followers on Twitter plays a vital role in information sharing and dissemination throughout the hurricane. These connections can provide an effective vehicle for emergency managers towards establishing better bi-directional communication during disasters. Another important finding of the study indicates that Twitter users generally receive emergency information from various sources at higher rates than non-Twitter users. This further affirms the important role that Twitter has in raising situational awareness. Unfortunately, improving the situational awareness of decision makers proved to not be sufficient to make informed decisions [42]. Practitioners put more emphasis on "identifying actionable information" as a primary information needed from data collection efforts [43].

Some climatologists and other expert drought observers have speculated about the value that monitoring of Twitter indeed has for #drought and related hashtags [18]. In that study, the relationships between the rate of tweeting using #drought and related hashtags is statistically examined and confirms that both are statistically significant predictors of #drought tweets in US. Moreover, it is concluded that #drought tweets can be one more metric to consider, as a real addition to quantifiable drought impact data.

The process of collecting information from any social media source introduces many challenges [36]:

(i) ambiguity in a sentence - important information is often presented implicitly, and Natural Language Processing (NLP) technologies currently rely heavily on surface processing

(ii) the importance of context - the current state-of-the-art considers informal text from social media an impoverished alternative to formal text. Innovative techniques on implicit and morphed information extraction are required to handle imprecise language

(iii) discourse ambiguity – for deep understanding of the text, it is necessary to know, for example, which sentence is the topic sentence, the elaboration or contrast of another sentence, and the temporal structure between sentences (in a story line)

(iv) expressions of fuzziness and vagueness – instead of using precise terms, human sources typically describe objects using fuzzy terms

Innovative approaches and technologies are suggesting the solutions that would overcome the problems listed. Continuous improvements of Natural Language Processing (NLP) techniques are addressing handling implicit information, analyzing discourse structure, and coping with fuzzy expressions.

1.6 NLP techniques

Although the data generated by social media platforms such as Twitter are ubiquitous, extracting useful and relevant information becomes highly demanding due to their enormous volume and velocity. Automatic disaster detection and classification models are showed to be feasible solutions [11].

By leveraging artificial intelligence (AI) technologies, more and more research works have been proposed for disaster response with social media data [8]. There is a study [44] dealing with Twitter-based natural disaster analysis overcoming the limitations on the number of supported languages, lack of sentiment analysis, regional restrictions, and lack of end-to-end automation. Their proposed solution uses a live Twitter feed to obtain natural disaster-related tweets in 110 supported languages. The system automatically executes AI-based translation, sentiment analysis, and automated K-Means algorithm to automatically discovered the similarities and detect which disaster types generated higher user engagement metrics.

Combined tweet-wise pre-trained neural networks and unsupervised semantic clustering were used in the study [45] to enhance the generalization capability of pre-trained models and to identify potentially crisis-related content. The analysis of cluster purities indicates that the method can isolate and identify crisis- and event-related Twitter content.

To deal with obstacles that analyzing social media brings, and improve existing approaches, transformers-based language models showed to dominate natural language applications with outstanding state-of-the-art performance [9]. This model introduced a self-attention mechanism that allows to capture dependencies between words in a sentence without relying on sequential processing [46].

In recent times, BERT has become widely adopted due to its strong performance in short message processing [10], and it was assessed in several studies in natural disaster domain in general [10], [11], [47]. There was an attempt to use a type of BERT model for crisis detection and classification purposes [47]. The authors have developed their own solution called CrisisBERT. Having DistilBERT as its base pre-trained model, CrisisBERT is aiming at both crisis detection and classification and shows promising results across accuracy and F1 scores.

However, in the area of drought specifically, there were only a few research studies that dealt with social media generated data [6], [12]. In the study [12] authors perform training of BERT base version using news media articles and validating drought relevant tweets from United States having as a validation parameter category-specific keyword label.

The study is leveraging multilingual BERT instead to classify drought relevant tweets in Italian into different impact groups. Another innovative approach adopted for this case study is the combination of classification model results with traditional sources for obtaining a wider picture on drought situation in Italy. The ultimate goal of the study is to help decision making and resource allocation nowadays, when drought conditions are reaching peak values.

Since in the research news media dataset used has been provided by external source [12], its characteristics are reported below.

1.7 DIR Dataset

The DIR dataset was developed and maintained by National Drought Mitigation Center (NDMC)[2] and the authors of the paper cited [12] provided data for this study. It contains 14.178 records of drought impacts data collected from news media. After the extraction, the data was manually classified by the experts from NDMC to 9 categories based on the text of the news itself. For the previous study, as well as for this one, it was used mainly for training the models, to additionally fine-tune the models for each of the tasks. Due to observed imbalance between the classes, *Energy*, *Business & Industry* and *Tourism & Recreation*, with the least number of tweets have been aggregated into 1 more general one as suggested by [12], to reduce the effects of imbalanced dataset that can lower the performance of the model trained. It was then proceeded with the pre-processing of data by removing HTML tags, URLs, and accented characters. Moving forward, all the contractions were expanded, and the special characters were removed. The ultimate dataset comprises of raw and processed

[2] https://drought.unl.edu/

title, row and processed description, a column for each category, and other columns not relevant for this study.

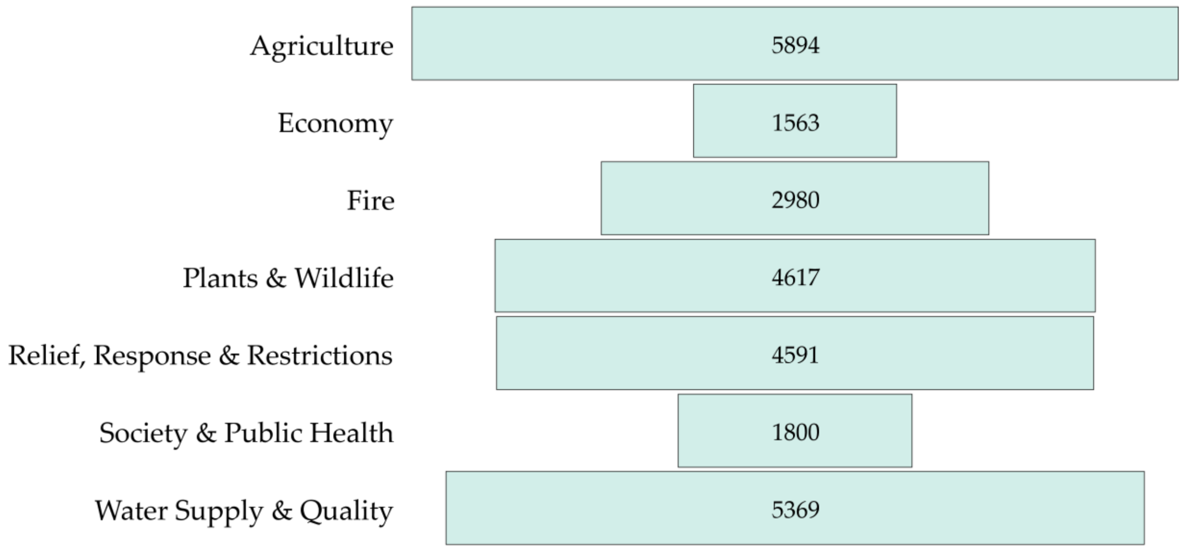


Figure 1.1: Category distribution of DIR data

2. Drought in Italy 2021-2022

This chapter reports severe drought that has affected areas of northern Italy and the Po River since December 2021. Following are brief overview of official institutions to monitor drought conditions, indicators measured and statistical reports that can help assessing of socioeconomic implications, as well as its effects on agriculture, industries, and local communities.

2.1 Organizations and institutions for drought monitoring in Italy

In the territory of Italy, there are several organizations and research institutions that are monitoring drought situation in the country through traditional drought indexes or others information that assist and contribute having a complete picture. There exist environmental agencies for each region as well that take care of managing issues in the area of interest and surroundings. These sides preferably collaborate with each other and exchange findings that may be crucial in drought conditions to possibly meliorate the process of assessing the impacts or improve mass assistance. The following are previews of some of them whose information was effective to estimate the conducted work in this study area.

- European Drought Observatory (EDO) - a platform for monitoring, assessing, and providing information on drought conditions across Europe. It offers drought-relevant information from monitoring stage, through assessment and forecasting, as well as detailed reports and tools to display and analyze the information or make maps of indicators derived from different data sources (e.g., precipitation measurements, satellite measurements, modelled soil moisture content) [49].
- National Research Council (CNR) [50] - the largest public research institution in Italy, it conducts research across a wide range of disciplines, including natural sciences, engineering, life sciences, social sciences, and humanities. It operates numerous research institutes and laboratories throughout Italy, each specializing in specific areas of study.
- The Institute of BioEconomy – (CNR – IBE) - includes the activities that use the bio-renewable resources of the terrestrial biosphere to produce food, materials

and energy, therefore includes the comparison of primary production (agriculture, forests, fishing), as well as the industrial sectors of use and transformation of resources, the agri-food sector, the wood sector, part of the chemical industry, biotechnology and industry. It also created a system to provide a semi-automatic, detailed, timely and comprehensive operational service in area of drought monitoring to support decision makers, water authorities, researchers, and general stakeholders [24].

- The Institute of Atmospheric Sciences and Climate – (CNR – ISAC) -provids integrated scientific understanding of the atmosphere, the ocean and their processes, by means of a multidisciplinary approach which combines scientific and technological skills in meteorology, climate, atmospheric dynamics and composition, Earth observations; it develops basic research, theoretical, experimental and numerical, and modelling work together with impact evaluation.
- Italian National Institute of Statistics (ISTAT) - main producer of official statistics in the service of citizens and policymakers. It operates in complete independence and continuous interaction with the academic and scientific communities. It provides the most up-to-date scientific standards, to develop detailed knowledge of Italy's environmental, economic and social dimensions at various levels of geographical detail and to assist all members of society (citizens, administrators etc.) in decision-making processes [51].

2.2 Geophysical data from official sources

In Italy, as reported by ISAC data, 2022 was the driest year since 1.800 with a deficit of 30% at the end of period. The deficit rises to 40% for the north of the country, which had 11 out of 12 months rainfall below the average. In first 5 months of 2022, in Italy, it rained 46% less than average from last 30 years (see *Figure 2.1*) [32]

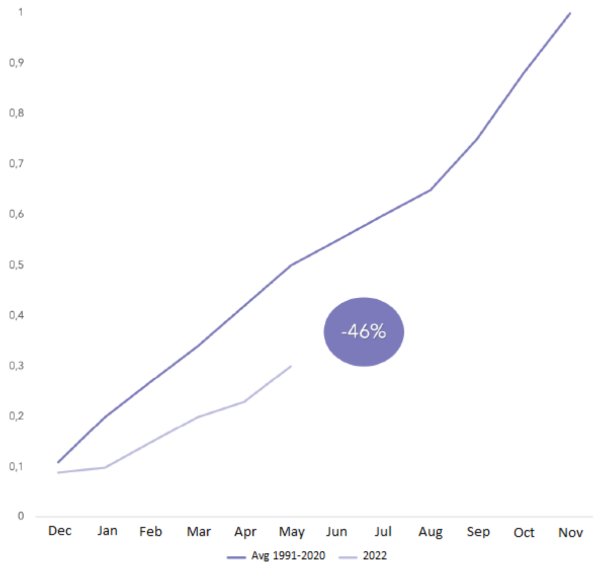


Figure 2.1: Precipitation long-time average compared to 2022

By analysing on of the indices (SPI), it has been revealed that SPI-3 showed rapidly evolving severe dry conditions, while when considering SPI over a longer period (SPI-12), it showed a progressive worsening over the same period (see *Picture 2.1*) [52].

Picture 2.1: Standardized Precipitation Indexes (SPI-3, SPI-12) displaying long term drought evolution in Piedmont from March 2020 [52]

The lack of precipitation lasted during the whole year, with several regions affected at the end of December 2022 as shown on *Figure 2.2*.

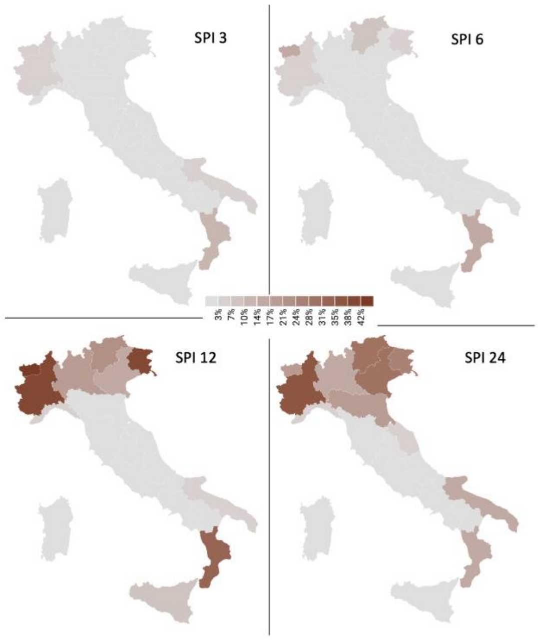


Figure 2.2: Region status of drought conditions in December 2022

Looking at anomaly of maximal temperature, which is used for detecting and characterizing periods of extreme-temperature (heat wave) anomalies, significant variations from the baseline are visible in the months of July, August, and October

2022. It is important to mention that heat and cold wave indices might not be exactly comparable with similar indicators that are available at national or regional levels, as the methods used for the measurements are different [53].

The figure presents of maximum temperature anomaly for each month of 2022, on the territory of Italy. The measurements are collected from EDO database that is recording data on daily basis. Monthly means are then calculated and normalized in range from -3°C to +3°C (see *Figure 2.3*).

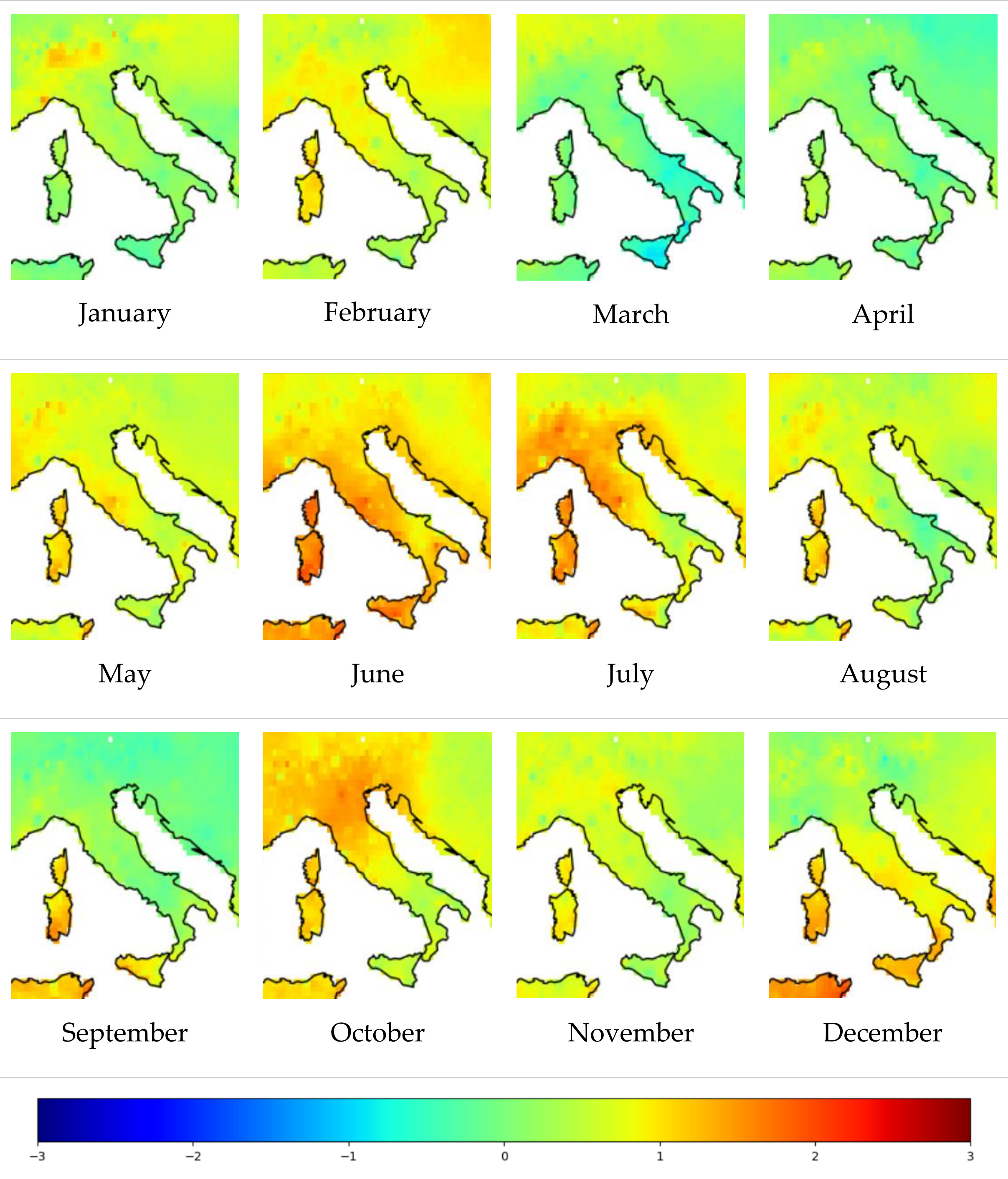
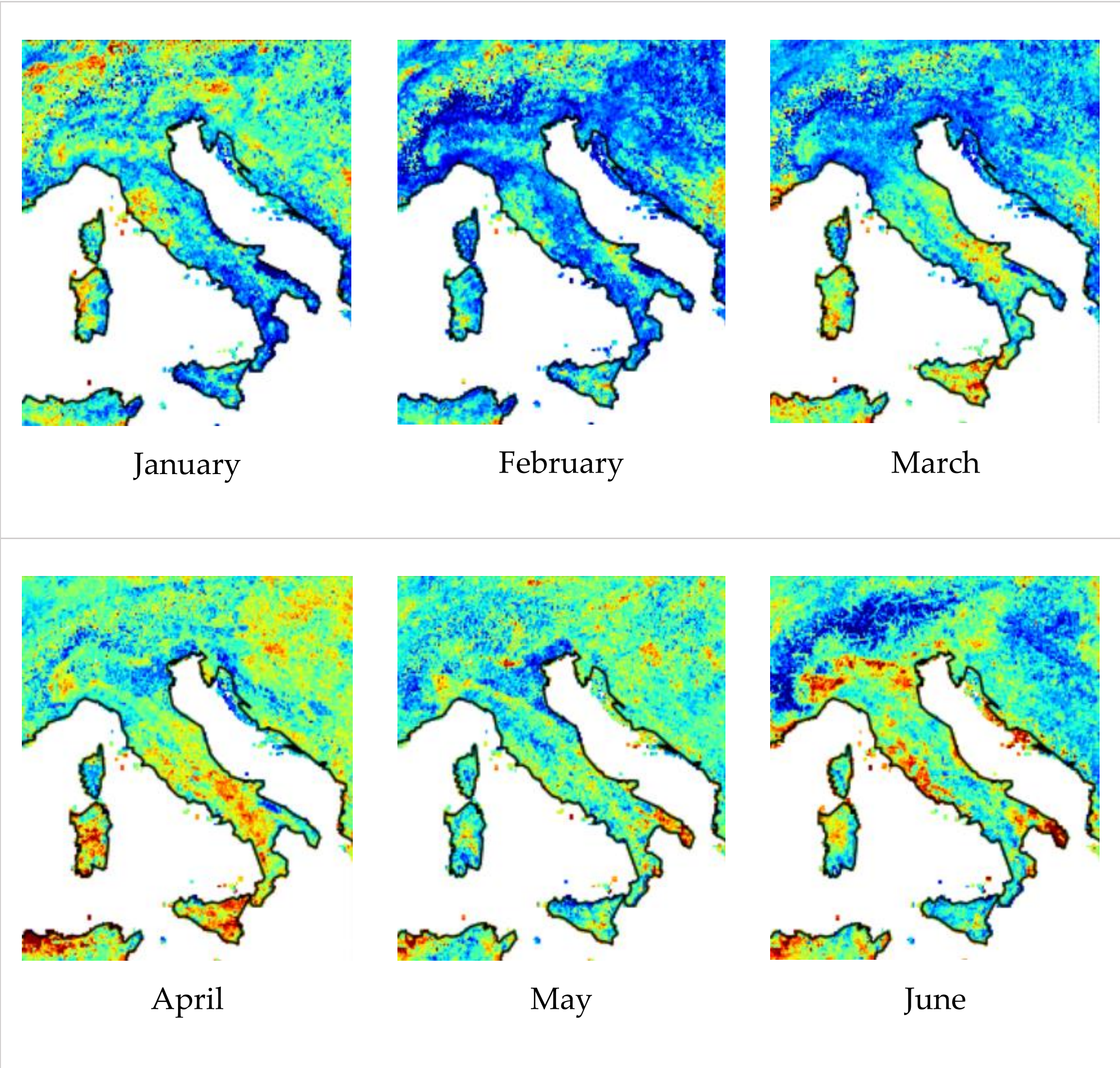


Figure 2.3: Maximum temperature anomaly in 2022 (if T_{max} is above the 90th percentile daily threshold for at least 3 consecutive days)

As far as Italy is concerned, ISTAT has detected in the provincial capitals an increase of temperature by 1.2C° compared to average 1971-2000 [32].

During the period preceding the core part of the growing season, plants have lower photosynthetic activity and lower water needs, so temperature becomes the main constraint for photosynthetic activity. The mild temperatures observed over most of northern Italy during winter caused increased FAPAR values photosynthetic activity. The computed indicator for 10-days intervals is loaded from EDO database and transformed to monthly data, that are visualized in *Figure 2.4*. More information regarding the calculation available in *Appendix – Anomaly indices information processing.*



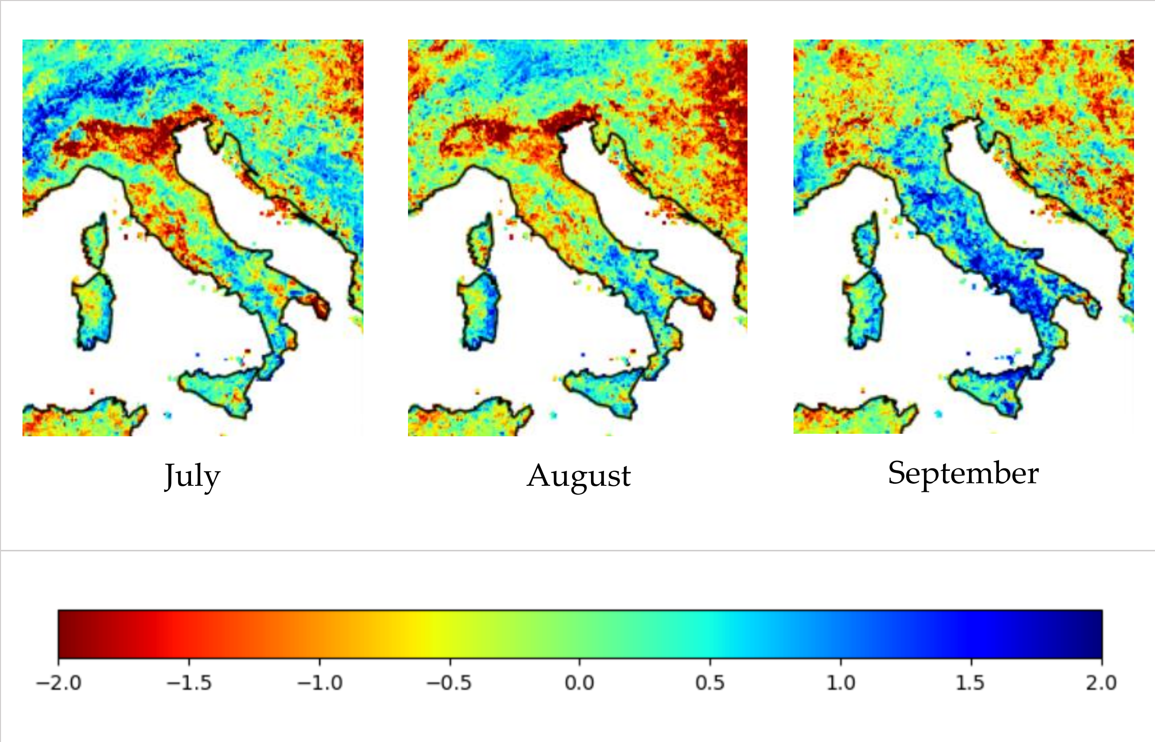


Figure 2.4: FAPAR anomaly for 2022 until September

The significantly lower FAPAR values are visible in July and August, especially in the north area, but also in the central and south parts. FAPAR values and their anomalies have been shown to be good indicators for detecting and assessing drought impacts on plant canopies, such as agricultural crops and natural vegetation [54] and thus provide information that is potentially useful for water and agricultural management purposes [26].

2.3 Drought reports

From the report EDO had published for the month of March 2022, severe drought has been affecting areas of northern Italy and the Po River since December 2021. At that point, it has already started impacting the stored water volume for energy production in the Italian hydropower system reaching the minimal historical value since 1970. During the winter, warmer temperature also had contributed to the reduced snow accumulation, and after poor winter rains and dismal Alpine snowfalls across northern Italy failed to replenish reserves, the arrival of spring has brought worries of more suffering following the drought of 2022. Inadequate water storage is a big problem since Italy receives an average rainfall of 301 billion cubic meters a year but collects and redistributes only 11% of it [55].

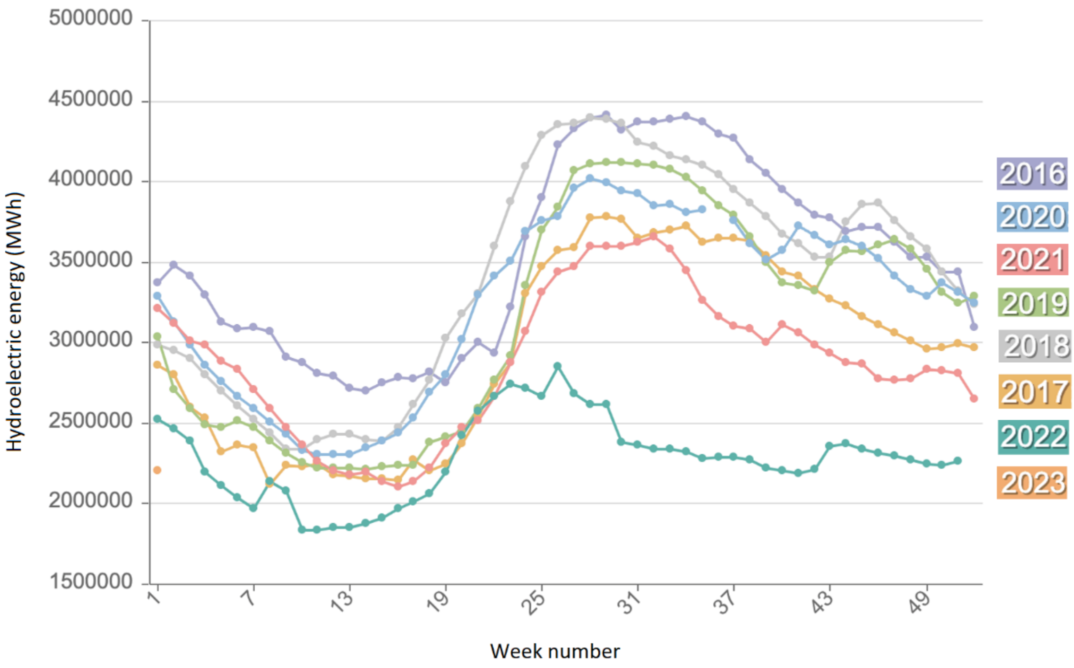
The strong instability of the international markets of agricultural raw materials and energy products was amplified partially by drought, that characterized the entire year affecting the volumes and quality of many crops. In 2022, agricultural production is reduced by 0.7%. Coldiretti, Italy's main farmer's association, stated that the sector had lost some 6 billion euros ($6.6 billion) in 2022 and predicted 300,000 businesses would lose more if the drought does not end. [55].After the crisis brought by the pandemic of the virus, the recovery of the secondary activities was held back by the sustained rise in input prices and by drought. Climatic events have conditioned the productions, with low spring temperatures, exceptional heat waves in the period summer and almost total absence of precipitation and a hot and dry climate that lasted for most of the year in many areas of the country. It was notable the increase in the prices of the products sold and of the products used. The was a significant decrement of oil (-17%) and cereal (-10,4%) production mainly, as shown in *Table 2.1* [56].

aggregates	**year 2022**	**% volume change 2022/2021**	**% price change 2022/2021**	**% value change 2022/2021**
cereals	6,427.9	-10.4	39.9	25.3
industrial plants	1049.7	-4.5	13.5	8.5
forage plants	2,662.7	-5.5	40.3	32.6
fresh vegetables	9,125.2	-1.8	21.2	19.0
flowers and plants	3,099.8	1.1	10.2	11.4
potatoes	696.3	-1.9	9.4	7.3
fruit	5,712.3	6.8	6.5	13.7
wines	8,844.4	0.1	10.1	10.2
olive oil	1,555.4	-17.0	8.4	-10.0
other vegetable products	331.2	0.3	9.2	9.5
vegetable production (1-10)	**39,504.8**	**-2.2**	**17.8**	**15.2**

Table 2.1: Italian Agriculture for Year 2022 in millions of current euros, percentage values [56]

The drought influenced the strong reduction in produced hydroelectric energy by about -40%. The impact report showed energy storage being affected from the middle of 2021, but the exceptionality of 2022 in comparison to the previous 6 years is evident.

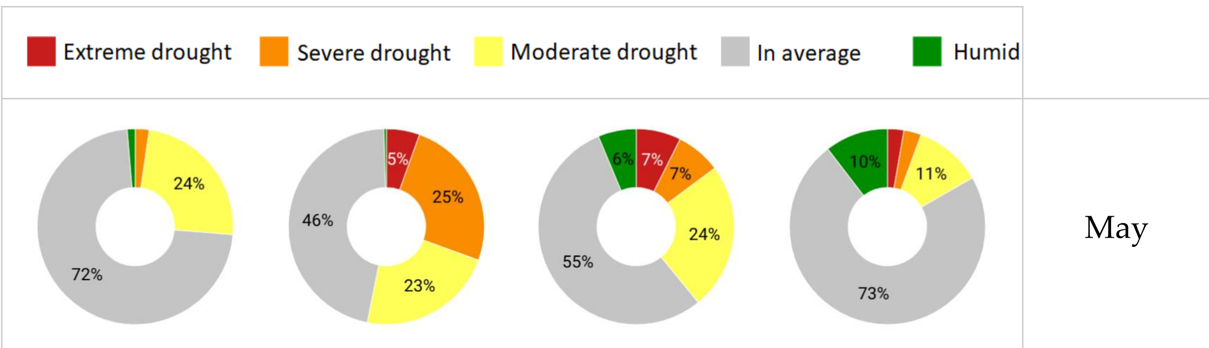
Unfortunately, the first week measurements of 2023 are not encouraging neither (see *Picture 2.2*) [25].



Picture 2.2: Hydroelectric energy levels through years [57]

Already in March were predicted the possible future problems with water derivation, due to the critical situation at the Po River Delta for the sea water intrusion. Even though in March was stated that winter crops are in normal conditions, in April both winter and summer crops were in sub-optimal position and water stress had reduced the yield potential. Extending the analysis on precipitation, a drought emergency has been officially declared in 5 regions. Water restrictions were introduced as the situation reached the highest level of drought severity [52], [58], [59].

The percentage of the population exposed to severe-extreme drought for December is lower than the data for the previous month, even if over the long-term overall percentage remains around 38%. Among the impacts that drought and thermal anomalies have caused, population exposed to drought at the end of 2022 is depicted with *Figure 2.5*:



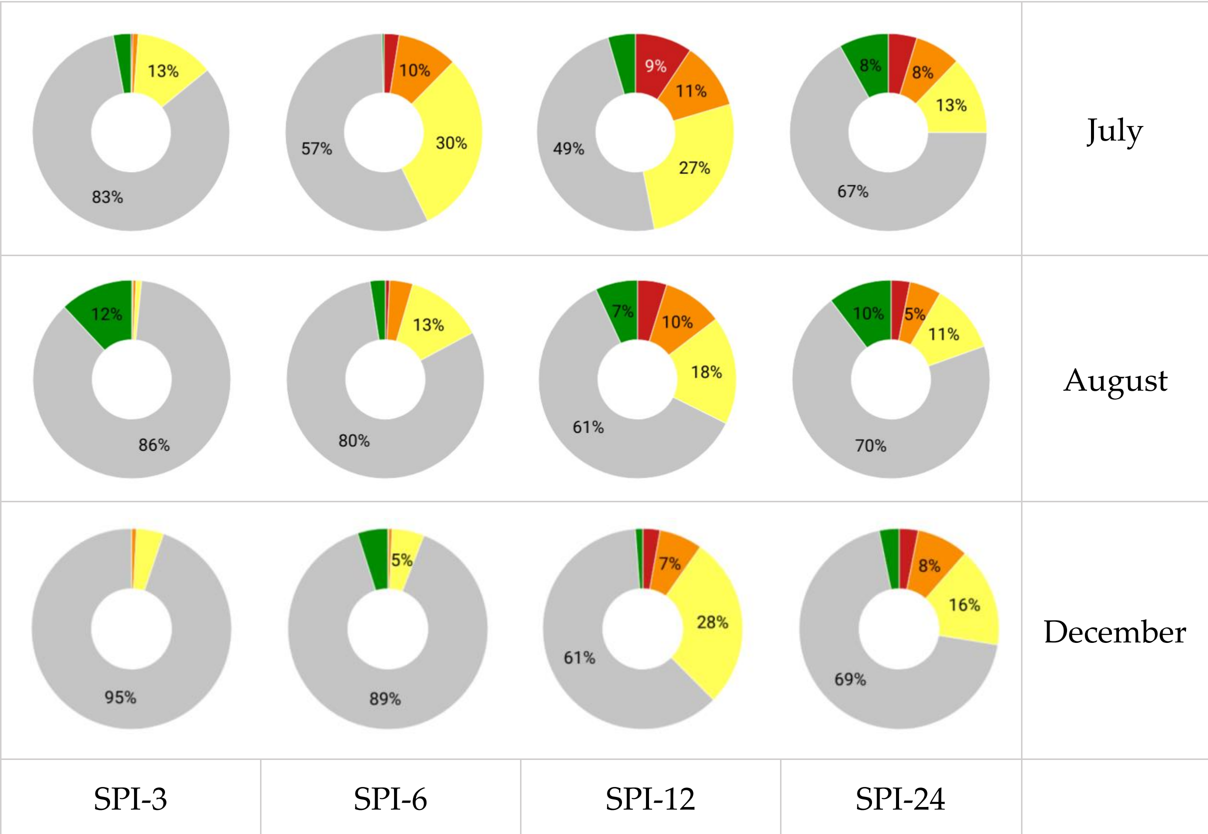


Figure 2.5: Population exposed to drought for different periods [25]

Even though from 11 to 17 August 2022, precipitation events finally occurred in many areas of Europe the Po River Basin Authority confirmed the current classification at the highest level of drought severity [60].

With the start of the new summer season yet another drought alarm returns and forcefully to flip. This time the data appears particularly alarming with some areas of the Peninsula, especially the North, who are living through the worst crisis of the last 70 years [32].

3. Analysis of drought-related tweets

Twitter as a free social networking site is easily accessible by a vast majority of the population. Users broadcast on it short posts – tweets that can contain text, videos, photos, or links. They are permanent, searchable, and public. Twitter data can support research not only from the social, but from engineering sciences too. Mainly because of previously mentioned reasons, it has been decided to perform tweet extraction for the period of interest, to better understand which problems the people are facing and gather more information on the topic of drought and its impacts.

3.1 Twitter API

Using Twitter's Developer Platform, it is possible to harness the power of Twitter's open, global, real-time, and historical platform within applications. The platform provides tools, resources, data, and API products for you to integrate, and expand Twitter's impact through research, solutions and more [61].

The Twitter API is a set of programmatic endpoints that can be used to understand or build the conversation on Twitter. Over the years, the Twitter API has grown by adding additional levels of access for developers and academic researchers to be able to scale their access to enhance and research the public conversation. The newest released, Twitter API v2, brought new endpoints, advanced metrics, and filters to use. However, Twitter API endpoints recommended to stick to are managing the tweets (posting), searching tweets and user lookup. To get access to Twitter API, it is necessary to create a developer account, after which the credentials for the application will be provided.

3.2 Keyword Based Tweet Extraction

The Twitter API allows several endpoints that grant interaction with different aspects of the Twitter platform. Searching directly for keywords and hashtags or recording the live stream of tweets containing those can be one of them, often used as a starting point for researchers [62]. Geolocation is another frequently employed feature that can be useful for retrieving tweets from an area affected by a disaster. However, in this specific case, it would miss important information that could come from a source outside the area, such as help providers or news sources [62]. Adding the fact that only a small fraction of tweets is actually geo-tagged makes this option not considerable in this case.

In this work, the solution code for interaction with Twitter API is written in Python programming language. The library chosen to use was *twarc*, as it is widely used by researchers, journalists, and organizations for collecting and preserving tweets. To access the API, one should use the provided credentials. After performing the configuration, it is possible to use the endpoint needed for tweets extraction.

As the focus is set specifically on the drought situation in Italy, all the tweets are extracted in Italian language. Keyword extraction is commonly used to retrieve the key information from the content and in our case, it helps further identifying main point of the tweet itself.

To capture all the possible, direct or indirect impacts of the drought, the only word Twitter was searched by was '*siccita*' or '*siccità*', which are Italian translations of the word '*drought*'. Since the objective was to understand better how the drought situation was developing during previous few years, tweets were extracted for three consecutive years (2020, 2021, 2022). The attributes of interest were date, the content of the tweet, and additional information such as user engaging metric for each one.

There were overall 121,254 retrieved results that contained drought as a keyword. Results divided by year are shown below in *Table 2.1*.

Year	Samples
2020	8224
2021	9285
2022	103745

Table 3.1: Number of retrieved samples per year

It should be noted that the numbers in *Table 2.1* represent raw number of tweets extracted without any processing and duplicates removal task performed.

3.3 Data pre-processing task

Twitter content often poses challenges for analysis because data contain a lot of noise, which makes more difficult to isolate signals from it. In cases like this, data pre-processing is crucial to improve the quality of the input data and to make it more suitable for the specific task that will be performed afterwards. Before proceeding with training of the model, it is desirable to clean the data and remove all missing values, outliers, inconsistent formatting, and duplicates.

In this case study, the data consists of tweets that have incorporated irrelevant or redundant information. To improve the dataset of raw tweets retrieved and prepare it for next steps, some pre-processing tasks are implemented as reported in the *Algorithm 1*.

Sequence 1 Data processing task

#the sequence of functions performed
for *tweet* in *tweets* **do**
  remove_URLs(*tweet*)
  remove_html_tags(*tweet*)
  remove_mentions(*tweet*)
  remove_special_characters(*tweet*)
  tweet.lower()
  remove_stopwords(*tweet*) *#for Italian language*
end for
#data preprocessing finished

Additionally, on tweet contents that are fetched after indicated steps another function is applied to remove duplicates. In this way, all insignificant information is filtered out, making it more suitable for analysis. The final datasets passed as an input to the model contain the following number of samples (see *Table 2.2*).

Year	Samples after cleaning
2020	6949
2021	8313
2022	85478

Table 3.2: Number of samples for each year after pre-processing task

The significant change is notable for 2022, where more than 18.000 tweets are removed, due to missing values and duplicates discharge. Afterwards, tokenization is applied to break the whole content down into individual units and to provide even more insights into each of datasets and make foundations for further analysis.

3.4 Labelling task

Natural disaster related datasets available for further use would usually contain classification on the types of different disaster event [47]. Drought specific datasets available online were containing mainly news articles [5], [12], [19], [33] that can fail to capture the characteristics that Twitter posts have. Due to lack of publicly available tweet datasets that contain specific annotations in drought assessment field, it has been decided to perform manual labelling of drought-related tweets, to have dataset tailored to this specific case study.

The tweets to label were extracted in the similar manner as explain in *3.2 Keyword Based Tweet Extraction*, retrieving them for the year 2019 with only date and the content of the tweet as the most relevant attributes in this case. The categories of drought impacts considered are the ones used in DIR dataset after the aggregation of three categories with the least number of tweets. The classes of impacts were defined by the experts from NDMC which adds on the certainty to perform division in this way.

As an initial step, DIR dataset was examined in detail to have a better understanding into each of drought category content characteristics, and to obtain basic competence to manage to perform manual classification task in acceptable way.

3.4.1 Classification conventions and distribution

To each tweet, 7 values have been assigned. The values considered were 0 (a tweet does not belong to a certain category) or 1 (a tweet belongs to a certain category) and were associated to each of 7 categories. Alternatively, if all 7 values are 0, this tweet sample is further considered as irrelevant. Conventions used during the labelling task to assign a specific tweet to one category are listed below. Classification conventions and some examples are reported in *Appendix – Classification conventions.*

The labelling task has been finished with 6233 labelled tweets. However, due to the imbalance of the classes (around 15% relevant and 85% irrelevant), the negative cases were reduced by randomly choosing samples from the dataset to have better ratio and possibly improve the model training. Finally, the dataset that was fed afterwards to the classifier counts 1898 tweets - 925 relevant (assigned to at least one category) and 973 irrelevant. The partitioned relevant tweets are represented in *Figure 3.1*.

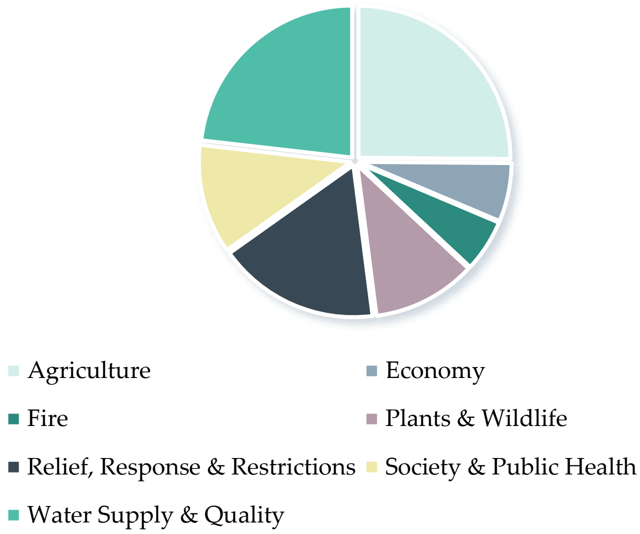


Figure 3.1: Labelled tweets distribution of categories they belong to

After having initial insights on data that it will be worked on in forthcoming steps, the speculation of having noisy data was confirmed through the results.

4. Applying multilingual BERT model on drought-related tweets

To get valuable insights into public opinion on specific events, topics of discussion and user behaviour, tweet analysis plays a significant role. Considering the large volume of information shared on social media, it is essential to not only effectively monitor it, but also categorize it in a way that can provide better situational awareness to emergency responders [63]. From these information, it is possible to effectively monitor disasters or emergencies, and help organizations or authorities respond providing timely assistance during the crisis. Using models that have already been trained on large-scale datasets could save significant time and resources, and knowledge obtained can be leveraged for specific tasks.

Encoder-based models are best suited for tasks that require an understanding of the entirety of the provided context. Such tasks include classification, named entity recognition and question answering [10]. BERT is considered having state-of-the-art performance on various NLP tasks [64]. With just one additional output layer it is possible to fine-tune it. BERT model is first initialized with the pre-trained parameters, and all of them can fine-tuned using labelled data from the downstream tasks. Each downstream task has separate fine-tuned models, even though they are initialized with the same pre-trained parameters.

In this work, the aim is getting more insights from tweets coming from the area of Italy, using Twitter content for both, training and assessing. In this way, tweet specific features can be represented better.

The chapter describes in detail how multilingual BERT model can be used and further fine-tuned for multilabel classification problem on tweet content. The model in focus is multilingual version of BERT. During fine tuning, two approaches are adopted, and their performances evaluated. At the end, the model that yielded better achievements is applied on social media generated content, in order to show how general-purpose language representation model act in task specific situation. A secondary goal was confronting the results of the two datasets in English and Italian language and how model act in those cases.

This phase of the study was mainly experimental, due to faced limitations of more specific documentation on the methodology and common practices used for multilabel classification with BERT model on tweet content.

4.1 BERT model preview

BERT is language representation model that can pretrain bidirectional representation from unlabelled text by conditioning on both, left and right context in all layers. It overcomes the limitations that unidirectional standard language models have when it comes to fine-tuning question answering tasks. The previous can be alleviated by using Masked Language Model (MLM) which enables to pretrain a deep bidirectional transformer [64]. Additionally, it has been used also Next Sentence Prediction (NSP) that allows BERT to learn relationships between sentences by predicting if the next sentence in a pair is the true next or not.

BERT has two phases, pre-training, and fine-tuning. During pre-training, the model is trained on unlabeled data from BooksCorpus (800M words) and Wikipedia (2500M words). For finetuning, it has been initialized with the pre-trained parameters, and then the parameters are fine-tuned using labeled data from the downstream tasks. Each downstream task has separate fine-tuned models, even though they are initialized with the same pre-trained parameters. The architecture of the model is composed of multi-layer bidirectional transformer encoder [64].

The model is initially implemented in two sizes: $\mathrm{BERT}_{\mathrm{BASE}}$ and $\mathrm{BERT}_{\mathrm{LARGE}}$ with the parameters shown below (see *Table 4.1*).

parameters	$\mathbf{BERT}_{\mathbf{BASE}}$	$\mathbf{BERT}_{\mathbf{LARGE}}$
layers (transformer blocks)	12	24
hidden size	768	1024
self-attention heads	12	16
number of parameters	110M	340M

Table 4.1: Reported the architecture of two different sizes of BERT model

These models are primarily aimed at being fine-tuned on tasks that use the whole sentence (potentially masked) to make decisions, such as sequence classification, token classification or question answering and not for the generation of the text. With a help of a dataset of labelled sentences it is possible to train a standard classifier using the features produced by the BERT models as inputs [65].

Multilingual BERT is issued by the same authors team [64] in November 2018 as a single language model pre-trained on the concatenation of monolingual Wikipedia corpora from 104 languages. BERT enables a very straightforward approach to zero-shot cross-lingual model transfer: fine-tune the model using task-specific supervised training data from one language, and evaluate that task in a possibly different language, thus allowing us to observe the ways in which the model generalizes information across languages [66]. The training set was machine translated, using the translations provided by XNLI [67]. The entire Wikipedia dump for each language (excluding user and talk pages) was taken as the training data for each language [68]. The languages with a larger Wikipedia are under-sampled and the ones with lower resources are oversampled. For languages like Chinese, Japanese Kanji and Korean Hanja that do not have space, a CJK Unicode block is added around every character. This specific model card was written by Hugging Face team [65].

There are two multilingual models currently available, cased and uncased version, in BERTBASE size [68].

4.2 Tokenizer

The input representation of BERT is able to represent both a single sentence and a pair of sentences (e.g., Question, Answers) in one token sequence. For tokenization, it was used a 110k shared WordPiece vocabulary. The word counts are weighted the same way as the data, so low-resource languages are upweighted by some factor. The first token of every sequence is a special classification token ([CLS]). The final hidden state corresponding to this token is used as the aggregate sequence representation for classification tasks. Another special token is used ([SEP]) to separate the sentences.

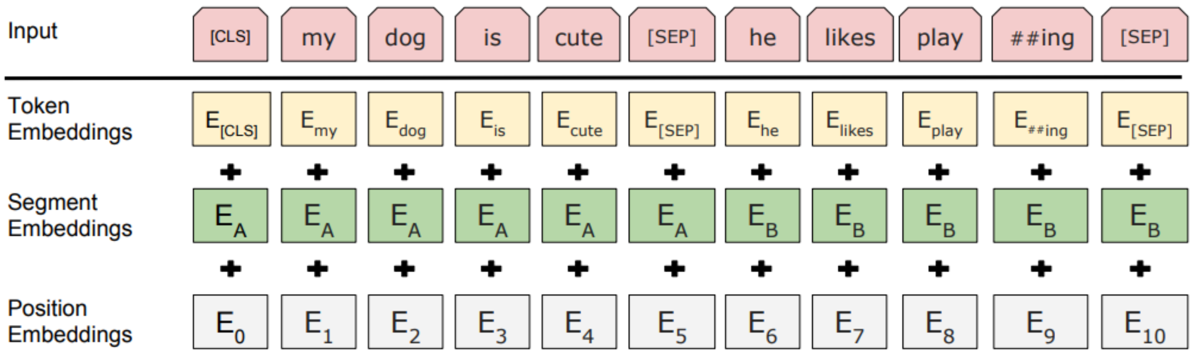


Figure 4.1: BERT input representation [64]

4.3 Importing BERT model

As anticipated, model and tokenizer are used from Hugging Face [65] open-source platform. Below it is represented how BERT multilingual model can imported and used. To use this model, it is necessary to load the tokenizer first:

```
from transformers import BertTokenizer
tokenizer = BertTokenizer.from_pretrained('bert-base- multilingual-cased ', do_lower
_case=True)
```

Listing 4.1: Importing BERT tokenizer from *HuggingFace*

Now, when tokenizer has been imported, the drought relevant dataset is fed to tokenizer to encode the content, which will transform the text into a format suitable for input to a model like BERT. The `max_length` variable has been previously calculated based on the tweet dataset fed as input to the model.

```
for tweet in tweets:
    inputs = tokenizer.encode_plus(
        tweet,  # sequence to be encoded
        add_special_tokens=True,  # to encode the sequences with the special tokens
        padding='max_length',  # activates and controls padding
        truncation=True,  # activates and controls truncation
        max_length=max_length,  # controls the maximum length to use
        return_attention_mask=True,  # whether to return the attention mask
        return_tensors='pt'  # to return PyTorch torch.Tensor objects
    )
    #list of token ids to be fed to a model.
    input_ids.append(inputs['input_ids'])
    #list of indices specifying which tokens should be considered to by the model
    attention_masks.append(inputs['attention_mask'])
```

Listing 4.2: Encoding the content that will be passed as input to the classifier

The obtained encoded dataset is divided into 3 sets for training, validation, and test purposes in proportion 8:1:1 respectively. One example how the BERT model can be imported and fed with `input_ids` and `attention_mask`:

```
import torch
import torch.nn as nn
from transformers import BertModel

class BertClassifier(nn.Module):
        def __init__(self):

                super(BertClassifier, self).__init__()
                D_in, D_out = 768, 7

                self.bert = BertModel.from_pretrained('bert-base-uncased')

                self.classifier = nn.Sequential(
                    nn.Linear(D_in, D_out),
                    nn.Sigmoid()
                )

        def forward(self, input_ids, attention_mask):

                outputs = self.bert(input_ids, attention_mask=attention_mask)

                # extract the mean of last hidden state for classification task
                pooled_output = torch.mean(outputs.last_hidden_state,1)

                # feed input to classifier
                logits = self.classifier(pooled_output)

                return logits
```

Listing 4.3: Bert Classifier model architecture example using *pytorch*

In *Listing 4.3* an additional classification layer is added on BERT model directly imported from [65]. In the example one dense layer followed by sigmoid as activation function is added to classify the input. Parameters of dense layer are the hidden size of BERT that is 768 for multilingual version, and the output dimension for n-dimensional classification. In forward function, dataset is passed as input to model. The mean value of the last hidden state is fed to classifier.

4.4 Fine-tuning BERT model

Fine-tuning over large pretrained language models (PLMs), which achieved remarkable performance over various benchmarks, has become the *de facto* paradigm for several current natural language processing (NLP) systems [69]. Compared to pre-training, fine-tuning is relatively inexpensive [64]. However, fine-tuning can be unstable in terms of significant variance in metrics, resulting in even worse-than-random failed models [69].

The layer on top can consist of one or more fully connected layers followed by an appropriate activation function. The weights of the pre-trained BERT can be frozen, ensuring that the valuable linguistic knowledge captured before is retained. On the other hand, by fine-tuning the model and updating weights of parameters, it can adjust its representation to adapt to tweet-specific elements and better understand special classification requirements. Since fine-tuning is often customized depending on the specific downstream tasks, it is often suggested to explore with approaches to choose the most convenient strategy.

During training, the input text is passed through the BERT model, which produces a sequence of hidden representations for each token in the input text. The mean of the last hidden state outputs is then used as input to a classification layer that predicts the output label. To find the best approach for this specific task and dataset, an experimental procedure was performed within this study evaluating the performance of two different architectures and hyperparameters to find the ones that give the finest results.

While assessing the results, certain evaluation metrics were supervised to weigh different architecture and parameters (see *Table 4.2*).

Granularity		Metric	
Micro	○ aggregates the contributions of all classes to compute the average metric ○ it gives equal weight to each instance in the dataset, regardless of the class ○ often used when the dataset is imbalanced or when the focus is on overall performance	Precision	measures true positives over the number of total positives predicted by the model
		Recall	measures true positive over the count of actual positive outcomes
Macro	○ computes the metric independently for each class and then take the average, treating all classes equally ○ treats each class equally ○ suitable when all classes are of equal importance or when the dataset is balanced	F1	harmonic mean between precision and recall ($\frac{2(P \cdot R)}{P+R}$)

Table 4.2: Evaluation metrics used and their granularities

There were 2 different strategies for building classifier layers. In the first case, a linear layer is added followed by sigmoid activation function. In the other approach, ReLU activation function together with dense layer is added before the two layers from the first case (see *Table 4.3*).

simple	complex

`nn.Linear (D_in, D_out),`
`nn.Sigmoid ()`
	`nn.Linear (D_in, H),`
	`nn.ReLU (),`
	`nn.Linear (H, D_out),`
	`nn.Sigmoid ()`

`D_in = 768`
`D_out = 7`
`H = 64`	

Table 4.3: Proposed architectures of the classifier while freezing pre-trained weights

ReLU is added usually to introduce nonlinearity to enable the network to learn complex patterns and relationships in the data. In both cases, binary cross-entropy loss function is used to minimize the errors during training our models. The goal was to check whether the architecture with additional ReLU layer can capture more intricate patterns or dependencies that are non-linear that simple architecture can.

The fine-tuning of the models is performed news media dataset as input since due to its higher number of valuable samples that have been manually classified by the experts in field.

4.4.1 Freezing pre-trained weights

Besides the fact that this approach is giving a lot of advantages in the terms of time (it does not require computing gradients or performing weight updates for BERT layers), it also mitigates the risk of overfitting, especially when the size of the dataset task is small. By keeping the weights fixed, it is ensured that the model focuses on learning specific patterns, rather than trying to fit the data too closely. Using this strategy, it is possible to leverage the knowledge from pre-trained model and use it as a feature extractor for a specific task is a transfer learning approach, that is particularly useful when labeled data are limited. The stability of the representations of natural language that pre-trained BERT has learn can be beneficial when working with noisy and unbalanced datasets.

The freezing of pretrained parameters is enabled (disabled) by changing the state of the following bool variable:

```
# freeze parameters
if freeze_bert:
    for param in self.bert.parameters():
        param.requires_grad = False
```

Listing 4.4: Freezing pre-trained parameters

The batch sizes suggested by BERT authors are 16 and 32 [64]. During the change of the batch size, the effective learning rate is affected too. For the learning rate, considering that the pre-trained weights are fixed, it has been decided to move on with a slightly higher learning rate than in the case of updating all weights. The options tested were 1e-3 and 1e-4. The number of epochs for training the model is chosen 5 as marginally higher value from the suggested ones, from the same reasoning as for the learning rates. The proposed architecture is obtaining the results reported in *Table 4.4*.

Freezing pre-trained weights	**BERT multilingual**			
Architecture	simple		complex	
Batch size	16	32	16	32
Learning rate	1e-4	1e-3	1e-4	1e-3
Epochs	5	5	5	5
Micro Precision	0.8453	0.8078	0.8100	0.8191
Micro Recall	0.4760	0.6323	0.5808	0.6529
Micro F1	0.6091	0.7093	0.6765	0.7266
Macro Precision	0.6069	0.7662	0.7234	0.8098
Macro Recall	0.3716	0.5333	0.4692	0.5601
Macro F1	0.4529	0.5879	0.5179	0.6148

Table 4.4: Results recap when freezing pre-trained parameters

When trying to put focus on the accuracy of positive predictions, and minimize false positives to decrease misclassification cases, precision is an appropriate metric to consider. It is equally important in this case not to miss important tweets of one category, as in that the true extent of the impact will remain unveiled. From these reasons, to target both problems, and make a balance between precision and recall, the advantage when choosing the best performed model is given to F1 score. The model shows the best results when using additional layers in classifier structure in case of freezing pre-trained weights.

It is also visible from the table that micro metrics obtained a higher score, which further indicates overall performance of BERT model in general. However, the poorer results for macro can demonstrate that there exists an imbalance between classes, and the model can be biased towards the majority classes.

In *Appendix – Accuracy of the model on categories of drought impact* available more information about performance of the model on individual categories.

4.4.2 Updating pre-trained weights

In cases when there is a large amount of labeled data for the specific task, or when the task requires more fine-grained language understanding, it is suggested updating parameters of the entire BERT model. The utilization of task-specific data is at a higher level as the model learns from both the general knowledge encoded during pre-training phase and the task-relatable information from labeled data. To prevent the model from forgetting previously learned knowledge when trained on a new task it is suggested to use this strategy. This approach though requires more computational resources and training time.

Unlike the first case, this time the learning rates chosen are more moderate, and commonly used in fine-tuning BERT [64]. In this context, when there are many parameters which already have meaningful representation, smaller learning rate is preferred to ensure adaptation to the model without drastically changing pre-trained representations. The number of epochs is chosen to be 4 as one of suggestions, for all the possible options.

Generally, when the pre-trained weights are unfrozen, there is a higher risk of overfitting, as the model has more flexibility to adjust the parameters to fit the task-specific data.

The outcome collected when unfreezing pre-trained parameters is presented in *Table 4.5.*

Unfreezing pre-trained weights	BERT multilingual			
Architecture	simple		complex	
Batch size	16	32	16	32
Learning rate	3e-5	5e-5	3e-5	5e-5
Epochs	4	4	4	4
Micro Precision	0.8723	0.8626	0.8853	0.8755
Micro Recall	0.8087	0.8082	0.7943	0.8033
Micro F1	0.8393	0.8345	0.8373	0.8378
Macro Precision	0.8372	0.8376	0.8714	0.8404
Macro Recall	0.7612	0.7640	0.7473	0.7377
Macro F1	0.7947	0.7969	0.8000	0.7817

Table 4.5: Results obtained when unfreezing pre-trained parameters

In this case, the differences in F1 score across different architecture as less significant which means that specific approach does not have significant impact on the performance of the model. When looking just at precision, the complex architecture yielded slightly better performed model than for simple architecture. The overall results show that the fine-tuned method with updating parameters acted better for this specific combination of the assignment and the dataset.

4.5 Multilabel tweet classification

In the previous step the model built was able to classify the tweets into 7 categoriesof drought impact with the respect to their content. The model has been trained on news media data to find the architecture that shows the best results. However, news media data can fail to capture the characteristics that tweet posts contain and primarily informal language of its users. As the focus of the study is social media content, and the assessment of the impacts recognized by Twitter population, the ultimate goal is to classify tweet datasets. To do that, it has been proceeded with two procedures:

- news-to-tweet transfer learning: evaluating the possibility of applying the model trained on news media data on tweet dataset.
  risk: the differences between sources of datasets may bring scarce end results.
- tweet training and evaluation: leverage of manually labelled tweet dataset to perform multilabel classification.
  risk: due to small number of samples in dataset the model may fail to learn specific patterns.

In this phase, the models performed better when using the technique of updating pre-trained weights of BERT model. Since there was no significant difference in results between classifiers' layers and hyperparameters, the architecture chosen was the simple one with lower batch and learing rate (see *Table 4.6*). was used to classify the content of the tweets in one or more categories, based on the impact that drought is having on the specific areas or spheres.

Updating pre-trained weights	BERT multilingual
Architecture	`nn.Linear (D_in, D_out),` `nn.Sigmoid ()`
Parameters	`D_in = 768` `D_out = 7`
Batch size	16
Learning rate	3e-5
Epochs	4

Table 4.6: BERT multilingual model used for multilabel classification

The model built was applied to downstream task of predicting the impacts for the tweets in Italian for 2020, 2021, 2022. The results are reported and discussed in *5.3 Model results*.

5. Results

A comprehensive analysis will be reported in order to assess the results of multilabel classifier model built on drought-related tweets in Italian. Additionally, the patterns and trends observed from model results and tweet datasets retrieved for three consecutive years, will be compared with the information obtained from traditional sources. In this way, it is possible to gain a more detailed insights into the strengths and limitations that this combined approach can bring. The goal is to provide valuable vision into the effectiveness of utilizing advanced data analysis techniques, such as machine learning models, in capturing and understanding drought-related information, as an additional source besides traditional one. The chapter can serve as a critical evaluation of using these methodologies for drought management, ultimately contributing to more effective decision-making process.

The chapter starts with a brief overview of characteristics and patterns captured from tweet datasets. This analysis is followed by classification results, that will be afterwards aligned with the observations made in *Chapter 2*.

5.1 Content peculiarities

To examine in more detail various characteristics, present within the dataset, the focus is set on exploring unique attributes and trends providing valuable insights into the specific drought relative aspects.

With the distribution of tweet length, it is possible to have some indications of possible patterns in user behaviour and preferences, such as whether is more likely for a user to engage with shorter or longer posts. Possible outliners are visualized in the distribution of the tweet length and word count using histogram and represented in *Figure 5.1.*

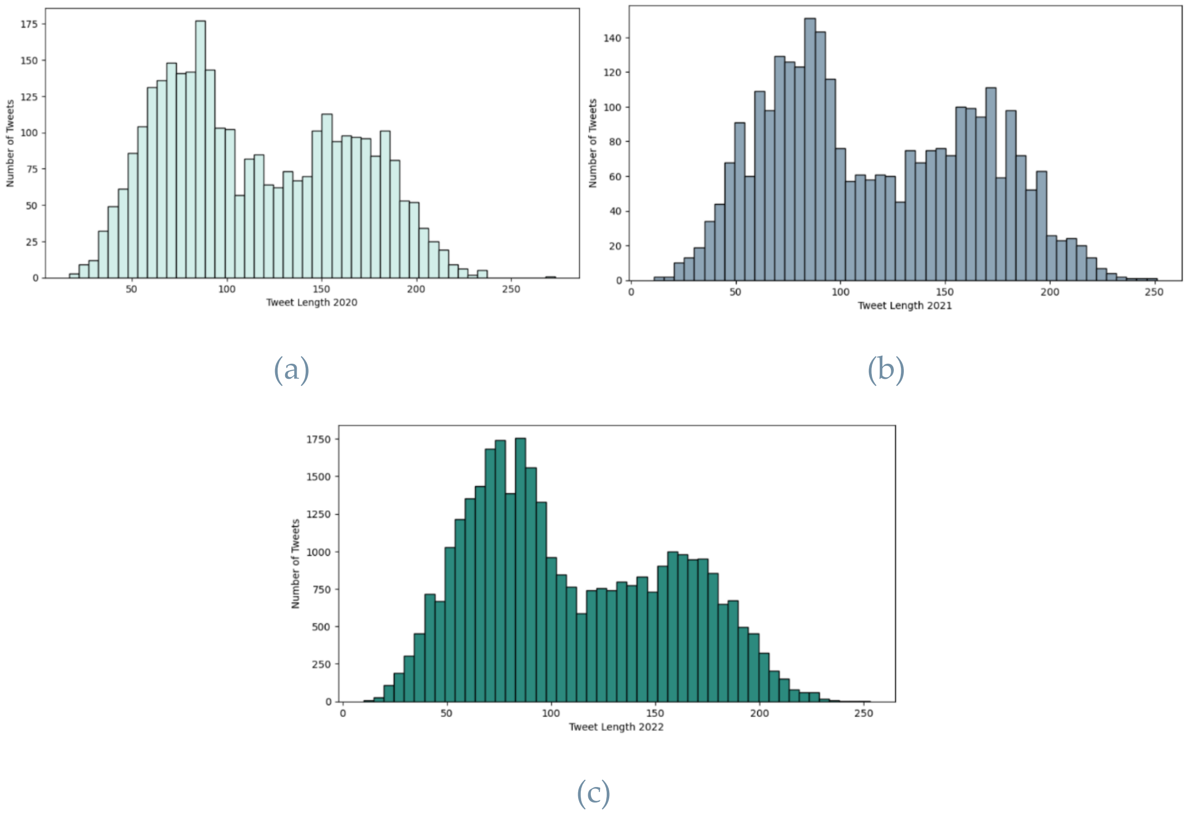
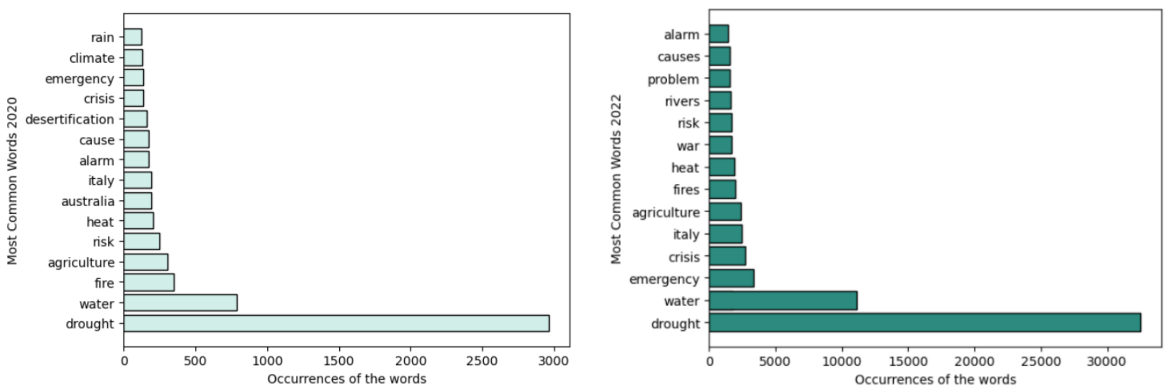


Figure 5.1: Distribution of text lengths of tweets for 2020 (a), 2021 (b) and 2022 (c)

The mean values for text length were variating from 110 to 116 characters per tweet. For all 3 datasets the tweet in average contained between 13 and 14 words. Approximately 13 words per tweet can indicate that the posts were marginally more elaborate than on average and tended to be more engaging in deeper conversation.

When examining the content of the tweet, in last 2020, 2021 and 2022, mainly the same topics were dominate in drought datasets. Note that the tweets were most of the time considered in their original form and translated since the process can be computationally demanding due to large amounts of samples and specific linguistic characteristics. *Figure 5.2* illustrates the most common words in 2020 and 2022 tweet dataset.

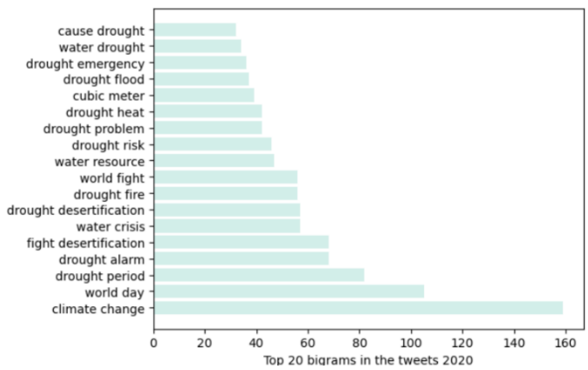


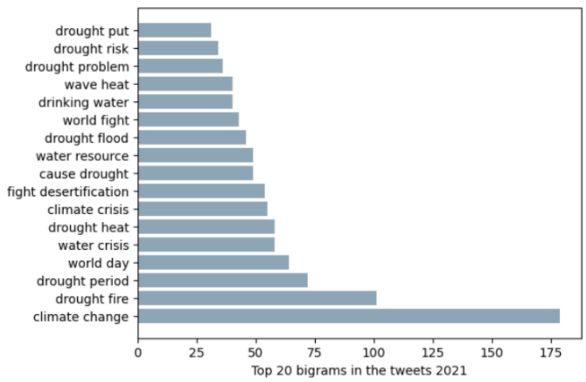
(a) (b)

Figure 5.2: Distribution of top 10 most common words in tweets 2020 (a) and 2022 (b)

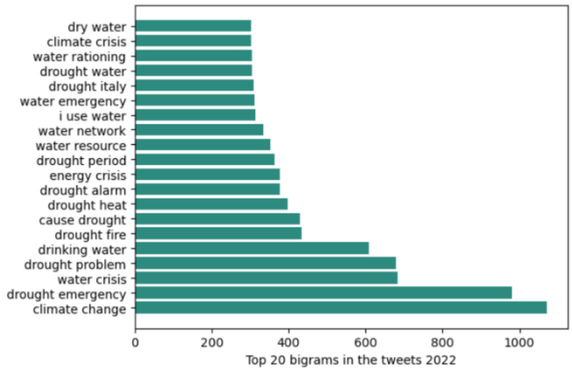
The drought term is left so that its dimension can be compared to the count of other words. The topics constantly mentioned during the years were *water*, *fire*, *agriculture*, *risk*, *heat*, *alarm*, *crisis and emergency*. Actually, *crisis* and *emergency* are certainly significant terms to consider, and it is marked the increase in importance through the years, since in 2022 they reached top 4 used words. This is important point for the monitoring of drought conditions. These words are reflecting droughts severity and urgency. Increasing the usage of this terms suggests drought situation worsened. Since drought can have significant socio-economic implications, as well as wars, the users often connect these two terms for the socio-political factors that have on society. The multiple appearances can have association with Russo-Ukrainian war started in February 2022.

To provide additional information from a higher point of view, the words are considered together with their neighbouring words. It was the way to try to make the entity recognition easier and capture more nuanced information. Commonly occurring pairs can provide intuition on the topics inside discussions. For these purposes, it is decided to analyse the pairs of 2 and 3 words together, to augment the knowledge gained (biagrams and trigrams are considered in Italian language and translated afterwards). The 2-words pairs can be seen on the Figure 5.3.





(a) (b)



(c)

Figure 5.3: The most popular biagrams in tweets for 2020 (a), 2021 (b) and 2022 (c)

The discussion for crisis and emergency can be indeed confirmed with these diagrams, where it can be further deducted on what these words really refer to.

From what reported above, there is no significant change in the tweet content that was broadcasted on Twitter by Italian users. However, the difference in the volume of data is evident. The drought topics became more ubiquitous in 2022 when comparing to previous two years.

5.1.1 Topic Modelling

As a statistical technique, topic modeling is often used to identify the underlying topics in a dataset or a document. It can discover the latent topics without any prior knowledge or labeled data. Simply put, each document in the corpus contains its own proportions of the topics discussed according to the words contained in them. The widely used algorithm for these purposes is Latent Dirichlet Allocation (LDA). It assumes that each dataset in the corpus is a mixture of various topics, and each topic is a distribution of words. LDA can be used to summarize, cluster, connect or process very large data because LDA generates a list of topics that are weighted for each document [70].

LDA is used for this study through *genism*[3], an open-source Python library for topic modeling, document similarity analysis, and other natural language processing (NLP) tasks.

After that the data has been pre-processed, document-term matrix is created whose columns represent words from tweets and the rows represent the tweets. Every single element is a frequency or weight of a particular word in the tweet. The LDA model is

---

[3] `https://radimrehurek.com/gensim/`

trained on this matrix using an iterative algorithm that assigns topics to each word and adjusts the assignments on statistical inference. Subsequentially, as a result, the probability distribution of topics is obtained for each tweet and distribution of the words for each topic, so it is possible to identify the most relevant pairs.

3 LDA models for 3 different dictionaries (tweets from 2020, 2021, 2022) are trained with number of topics equal to 5, in 10 passes.

For gain a better overview of a model trained, *pyLDAvis* visualization interface is used to have a closer look at topics and words associated with them (*Figure 5.4*).

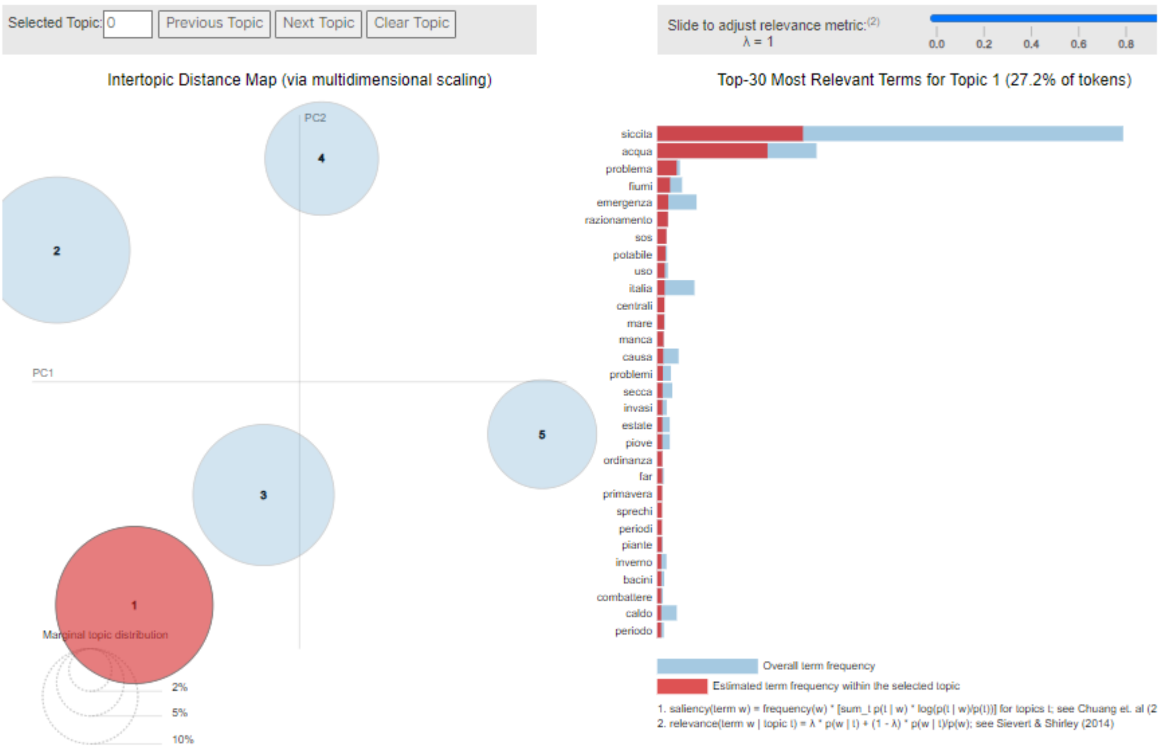


Figure 5.4: An example of words distribution from tweets extraction (2020) inside one topic

The size of the circles in left panel indicates the relative statistical weight of topics. This visualization of global topic view in space gives a sense of the statistical nearness or farness of topics from each other. A *relevance metric* slider scale at the top of the right panel controls how the words for a topic are sorted. More precisely, it combines two different ways of thinking about the degree to which a word is associated with a topic. In this case, lambda (λ) value in the slider is set to "1" which sorts words by their frequency in the topic (i.e., by the length of their red bars) [65].

The results obtained show 5 words for each topic with the highest estimated frequency within the selected topic (see *Table 5.1*).

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
2020	water risk alarm agriculture coldiretti[4]	water agriculture cause damage south	water fires climatic change children	australia fires water camels cause	water heat agriculture climate level
2021	fires water agriculture heat coldiretti	water risk agriculture hydro resource	risk water italy climatic donations	water fight desertification climatic italia	day world water crisis desertification
2022	water problem river emergency rationing	crisis war fires desertification change	water fires emergency hydro situation	agriculture production coldiretti risk grain	water lake cause level africa

Table 5.1: The words with highest distribution in each topic per year

There are some observations and patterns that are worth mentioning from the reported information. When taking a deeper look into the 2022 topics, it is possible to make a connection with topics considered in classifier. *Topic1* is indicating on water problems and possibly reasoning on how to implement measures to control and distribute limited water resources. *Topic2* can be an indication of what difficulties population may be facing in general, while *Topic3* can present combination of water and fire categories. *Topic4* topic can pinpoint to what it has been considered to belong to category of agriculture. *Topic5* can most likely express the effects of drought in water field. The role of water in different contexts is more understandable from the attached. Similar reasoning can be applied also for other years to gain more insights in the distribution of the topics. It should be mentioned though that there were some events that took place during these years such as Australia bushfires during 2020 and start of the war in 2022 that made people posting about it trying to address the possible connection they have with the drought.

[4] Coldiretti (Confederazione nazionale dei coltivatori diretti) agricultural organization at national level in Italy

However, there are posed some unique challenges when it comes to processing Twitter content. The fact that posts are short and use informal language with misspellings, acronyms and nonstandard abbreviations can lower the performance of LDA algorithm [71].

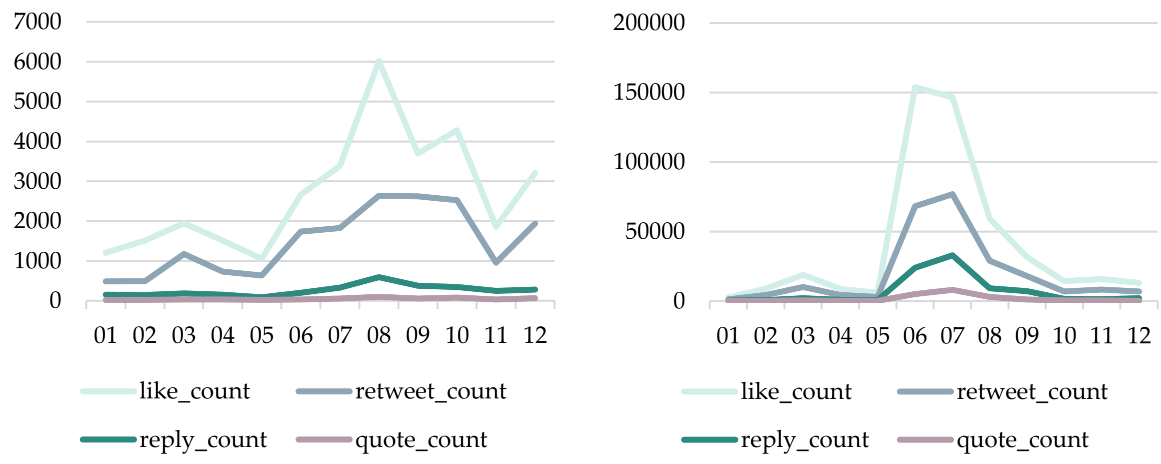
5.2 User Engagement Metrics

The number of retweets, likes, or replies are metrics that show the engagement and interaction of the users with tweets. They provide insights into the impact, popularity and reach and can help evaluate the effectiveness of the content and understand the level of audience interaction.

Retweeting or sharing someone else's tweet shows that the specific user found the content worth sharing and analyzing them can bring out the ones of the highest attention and reach among authors. The likes indicate that the user thinks the tweet is interesting, informative, or enjoyable. It can be used as a measure of popularity or overall positive sentiment towards the tweet. Replies are the responses or interaction with the tweets. Similarly, the replies can indicate the level of engagement, feedback, or opinion of an author on a particular topic.

The goal of this element of the study was to see how the users were engaging with the content on Twitter throughout the last 3 years, and to find the periods with the grow of attention to drought topics if any.

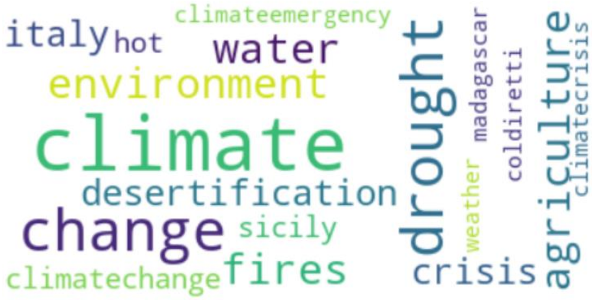
As drought depends mainly on precipitation and heat waves, usually those are the period of the year when it is talked about the most. This can be a valuable explanation why the tweets that are engaged the most are the ones posted during mentioned seasons, summer in first place (see *Figure 5.5*).



(a) 2021 (b) 2022

Figure 5.5: Evaluation of user engagement metrics for tweets from 2021 (a) and 2022 (b)

There are also other metrics that can be considered, besides the previously analyzed ones based on numbers and numerical statistics. Hashtags are keywords preceded by the '#' character with the goal of categorizing or grouping tweets related to a specific topic or event. Taking a deeper look at the hashtags can provide insight into trending topics or community discussions (see *Picture 5.1*).



Picture 5.1: A example of the most used hashtags in tweets about drought in 2021

From *Table 5.2* it is noticeable that just a small percentage of the overall tweets use hashtags or add some kind of attachments to it.

Year	Samples after cleaning	Hashtags	Attachments
2020	6949	2528	1443
2021	8313	2782	1581
2022	85478	36934	19295

Table 5.2: Number of tweets containing hashtags/attachments per year

(a)[5] (b)[6]

Picture 5.2: Example of attachments retrieved from tweet datasets

[5] https://pbs.twimg.com/media/EcwLo_eXgAEXhwt.jpg
[6] https://pbs.twimg.com/media/ESHjH4qWsAAUN7x.jpg

However, metrics can still help in entity detection and identification of emerging trends. Furthermore, analyzing all these parameters helped understand the effectiveness of content strategies, identifying audience responses and interaction.

5.3 Model results

Multilingual BERT was used to assess drought impacts discussed in Twitter posts. For that purpose, two models as trained, based on news media and tweet dataset. The results of multilabel classification of each of them are reported in following sections.

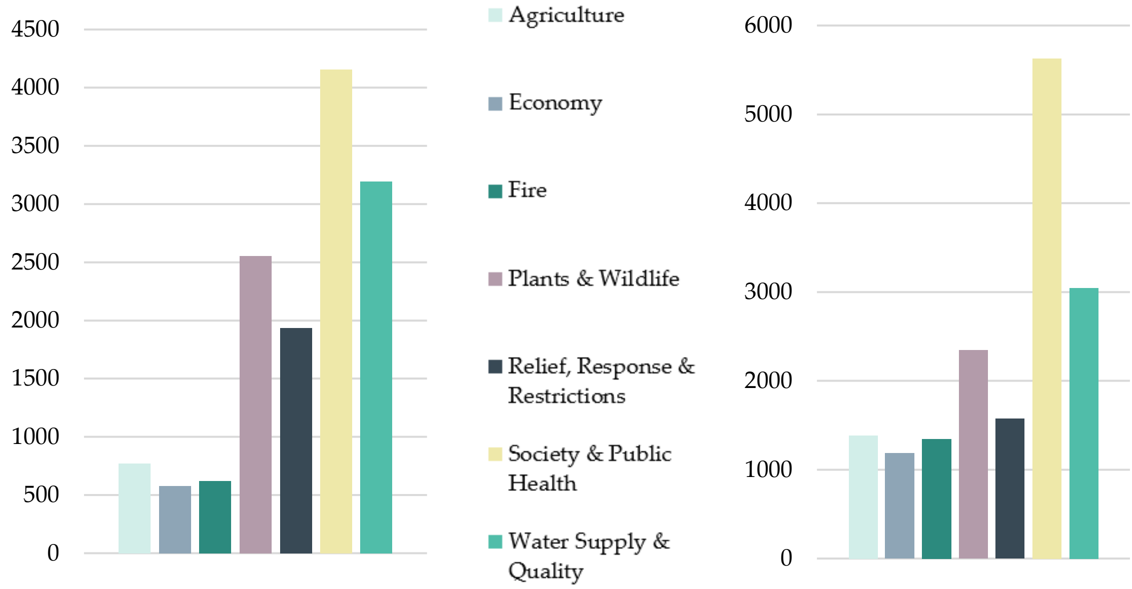
5.3.1 News-to-Tweets transfer learning results

This approach was meant to exploit the benefit from the pre-existing knowledge captured by the model trained on news media data and apply it to the task of tweet classification. Although the language usage in sources may be different, both datasets are from the same drought domain.

The tweets from 2022 signed as relevant have distribution for each category as showed in *Figure 5.6*.

Figure 5.6: Tweets belonging to categories of drought impact in 2022

The results of this approach indicate the category of *Society & Public Health* as the most impacted from drought. Since this category earned the most attention, it should be taken into consideration other category impacts that can be directly or indirectly connected to society. Although some associations are evident: *Agriculture* - food shortage; *Economy* - worsened living conditions; *Water Scarcity & Quality* – in household, other connections can be speculated. To try to capture dynamics in developing drought conditions through a three-year time frame, the model is applied to the previous two years.



(a) (b)

Figure 5.7: Tweets belonging to categories of drought impact in 2020 (a) and 2021 (b)

After further analysis for 2020 and 2021, there are not noticeable significant differences in the order of distribution for category impacts. When looking only at impact distribution, it appears that the drought had the same effects at the beginning as at the end. In real world scenario, this rarely happens since drought has slow onset effects that further progress differently in different stages. Additionally, it can be considered as surprising the scarce importance of *Agriculture* that the model results are showing. This category is one of the first ones to be hit by drought.

To verify the correctness of the predictions, the results will be compared to the ones obtained from official sources, after which more certain conclusions can be drawn on the model's performance and applying news-tweets transfer learning in general.

5.3.2 Tweet training and evaluation results

The model built had for goal to classify tweets containing Italian word for drought into categories of impacts. Multilingual BERT pre-trained model with added classification layer on top was fine-tuned on drought-related, manually labelled dataset. The model achieved the following results:

```
Micro Precision: 0.6807
Micro Recall: 0.6279
Mirco F1: 0.6532
Macro Precision: 0.6818
Macro Recall: 0.6284
Macro F1: 0.6258
```

Figure 5.8: The evaluation metrics of tweet based multilabel classifier

Final results of multilabel classifier are tweets assigned to zero or more categories based on its content. The proportion of relevant (at least one category) and irrelevant (not assigned to any of the categories) is reported below in *Figure 5.9.*

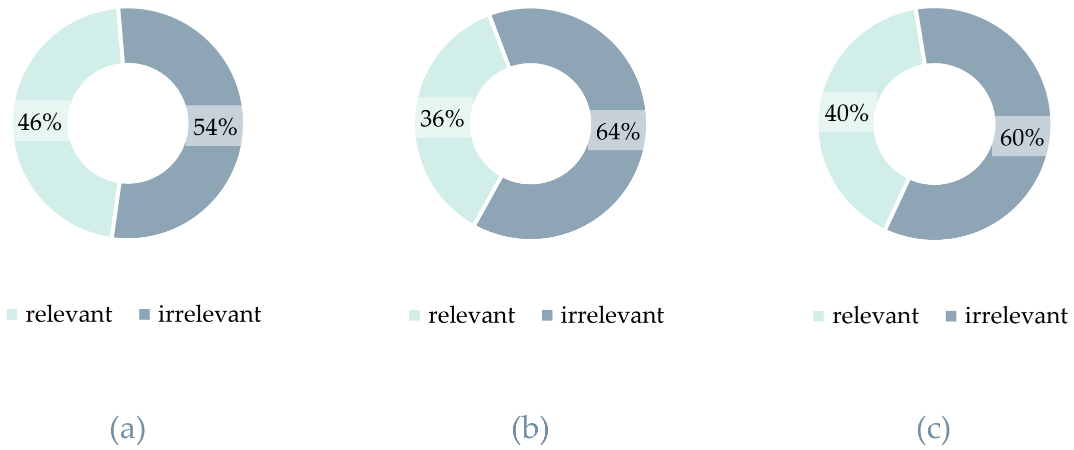


Figure 5.9: Summary of tweet relevance results from the model for 2020 (a), 2021 (b) and 2022 (c)

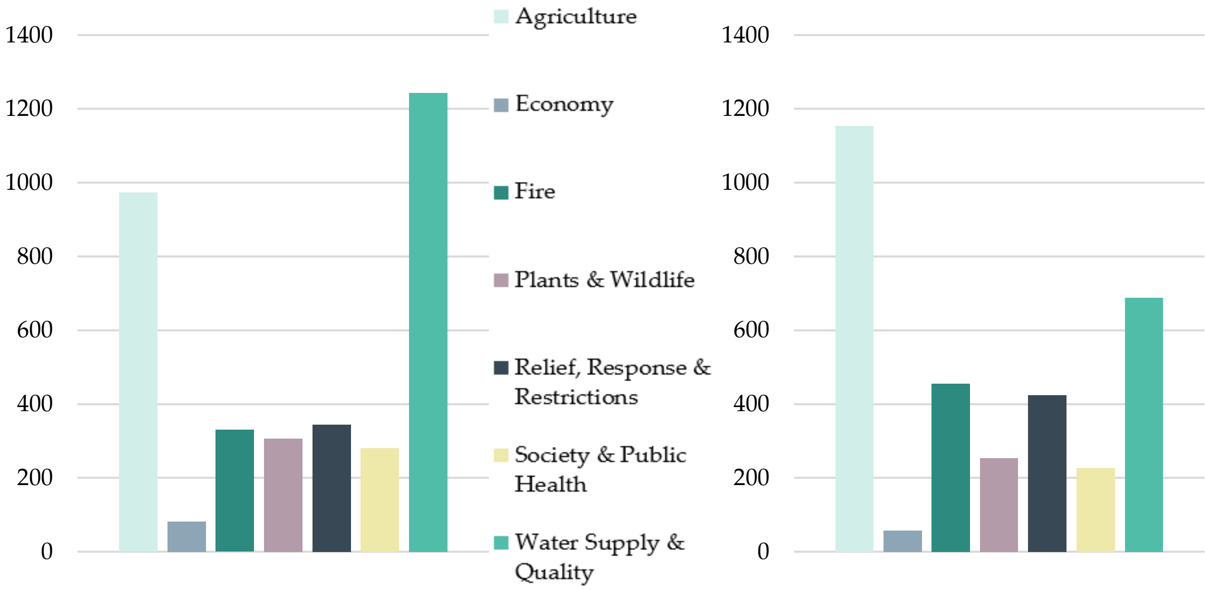
The tweets from 2022 signed as relevant have distribution for each category as showed in Figure 5.10.

Figure 5.10:Distribution of tweets over categories of impact in 2022

The model identified *Water Supply & Quality* category to have assigned the highest number of tweets. Water resources, quality, and conservation seem to be discussed the most on Twitter from Italian users. This suggests that water scarcity or deterioration of water resources is a significant concern during drought conditions. Of all tweets resulted relevant, 46% of them are assigned to water related issues, not exclusively (they may belong to other categories too). From the information obtained during analysis of most common words in tweet datasets, water was in all three datasets just after drought. Besides all mentioned direct impacts, water resources also affect ecosystems and wildlife habitats. There are different types of water resources that perceive the effects of drought, from surface water, groundwater, rainwater to wastewater and recycled water. From the attached, a comprehensive understanding of the water-related challenges and impacts can be gained, aiding in the development of appropriate strategies and interventions for drought management and mitigation.

Another significant topic in relation to drought is *Agriculture*. It takes second place on the most represented drought impacts captured by the model having 9141 tweets classified in this category. Whether the topics discussed were regarding the assessing the damage that the crops and soil suffered from or the reporting the effect it had on economy, this subject had high visibility. On the other side, water has several impacts on agriculture too. It is inevitable part in photosynthesis process as well as in crop irrigation. It is necessary to keep soil moisturised to prevent excessive drying out or waterlogging. Ultimately, it has impact on the quality and the yield of the crop. Thas is why water management can be considered as direct impact to water category and indirect in agriculture, as plays crucial role in ensuring sustainable and productive farming systems.

For the purpose of making some conclusions on how drought was developing from 2020, the predictions of the other two datasets are placed in *Figure 5.11*.



(a) (b)

Figure 5.11: Distribution of drought impacts contained in tweets from 2020 (a) and 2021 (b)

The results are showing presence of tweets containing response measures or suggestions, as well as communicating the constraints that the population or environment is facing with. The category that depicts these aspects is *Relief, Response & Restriction*. The Italian users on Twitter were showing the compassion and support to people that felt the effect of the drought in first person. As the drought conditions exceeded the capacity of affected individuals, and some systems failed to effectively respond and cope with its impacts, it escalated into a crisis event and it brought along various restrictions on water use, in the fields of agriculture, industry or public supply for essential needs. As visible, this category has various connections with other categories, and it is clearly visible in this example all the cumulative effects that drought impacts have.

Due to the interconnection of the categories, there were cases when the model identified in tweet content impacts from more than one category. The figure below is showing the number of categories that a tweet belongs to (see *Figure 5.12*).

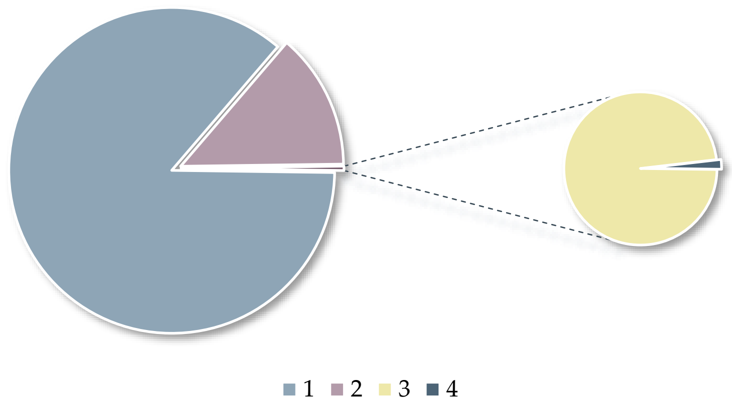
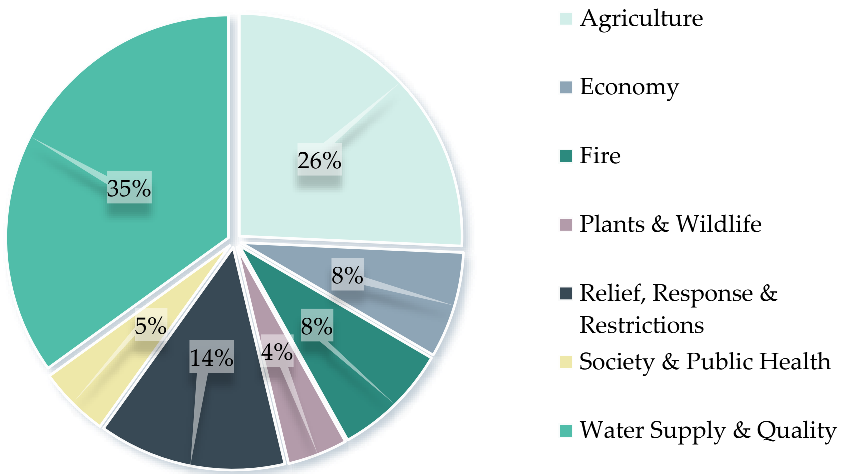


Figure 5.12: Tweets belonging to a certain number of categories in 2022

The most common case has been assigned to one or two categories of impacts. There were few cases that captured drought effects in more categories. The samples are represented below translated in English.

- *Orbetello: a cutting-edge company in the cultivation of plants and lawns that require very little water and maintenance. A policy opportunity to combat drought and desertification, promoting sustainability and civilization.*
  assigned categories: Agriculture, Plants & Wildlife, Relief, Response & Restrictions and Water Supply & Quality.
- *Drought: a prototype to irrigate fields with treated wastewater reduces costs in agriculture.*
  assigned categories: Agriculture, Response & Restrictions and Water Supply & Quality.
- *Marche, central Italy, affected by extraordinary drought, connected to a perceived high hydrogeological risk due to inadequate hydraulic infrastructure and increasing urbanization.*
  assigned categories: Economy, Response & Restrictions and Water Supply & Quality.
- *Driven out of their homes, extremist groups, hungry due to drought, thrown into despair, facing economic consequences, depleting resources due to COVID.*
  assigned categories: Agriculture, Economy, Society & Public Health.

A supplementary analysis is completed on subset of tweets from 2022 that received the highest attention from users (more than 50 likes). Once again, the model was estimated for these samples to verify the distribution of the classes of catgories.



Again *Water Supply & Quality* and *Agriculture* are dominating categories in the number of tweets that are assessing their drought imacts. It can be further dedicted that the number of likes, as one of the user engagement metrics, be considered as a valuable alernative help in research.

5.4 Comparing model results with traditional data sources

When it comes to assessing the results obtained, one of the best practices is to confront them with the ones coming from the official sources, taking into consideration their reliability and credibility. The correlation and combination of the two sources can help in defining the accuracy and validity of the findings. Differences in scope and granularity cannot be neglected either. Official sources may report detailed data on the impacts of natural disasters, while model's output is based on inferences learned from training dataset, that may does not show the entire range of impacts. However, since data coming from social media are by nature real-time, the classifier model can deliver the results serving them as early indicators, which may not be the case for official sources. Consequentially, a more comprehensive and reliable understanding of the situation is achieved by combining the advantages of both approaches. Therefore, it becomes essential to affirm the consistency of the outcome obtained from the model with domain expertise results.

An important conclusion drawn from the posts and also confirmed by other sources was that the drought situation is significantly worsening considering the three-year time frame. From the analysis performed on tweets, there is a momentous increase in tweets posted in 2022, 10 times more than 2021 and even 12 times more than 2020. The number of user engagement metrics reaches the momentous growth of attention this topic received from Italian Twitter users. The drought categorization using hashtags and attachments was expanded more in 2022 than in the other two years. Even though the drought has been detected according to the reports since the end of 2021 at the north of the country, some time needed to pass for users to start to discuss this topic since the peaks are visible in dry season months. All these metric values have contributed to raising awareness and understanding the level of audience interaction in 2022 more intensely than in the years before.

The official reports from EDO can be supported by the significant attention that was brought by people on Twitter, where it can be also seen a serious decrease in SPI measurements. As a matter of fact, one of the most considered topics on Twitter was water, in all forms, whether when talking about the water scarcity in rivers and households or water necessary for soil irrigation. It is anticipated also in the study [72] that the future will witness increased dynamics in hydro-meteorological variables around the world which will lead to frequent droughts whose impacts will be compounded by growing water demands. Even nowadays, water demand and restrictions in use are topics already discussed by Twitter users. *Water Scarcity & Quality* impact category certifies the substantial role that water plays in the drought effects, whether as a direct, or indirect impact on overall water cycle and ecosystem dynamics as well as its impact in other categories.

Due to all mentioned statistics, it is more likely that water resource is the category that has priority. Hence, it can be deducted that the model trained and evaluated on tweets has a higher possibility of depicting the real scenario on drought impacts.

Moreover, there are some other connections that we can make between Twitter analysis and traditional reports. The temperature deviations registered in Maximum Temperatures Anomaly can be justified by heat reaching one of the most common words discussed about for all three years. On the other side, the decrement in energy storage did not pass unnoticed on Twitter either. Energy crises as a term can be seen in the top 10 most common biagrams in 2022. The percentage of the population exposed to severe-extreme drought reached the highest value in July 2022 with the long-term overall percentage of 47%. The analysis of user engagement metrics in 2022 is indeed analogous to statistics since it shows significant spike in the months of June and July.

6. Conclusion and future development

Drought is characterized by a slow and difficult to define onset often long-lasting evolution. After flooding, it is the second natural disaster by the specter of effects it has on population, which puts crucial priority at effectively assessing its long-term impacts that is challenging [24]. Acquiring a better understanding of drought impacts becomes increasingly vital under a warming climate. Traditional drought indices describe mainly biophysical variables and not impacts on social, economic, and environmental systems [12]. On the qualitative side, society and emergency responders can glean new information from the content of tweets by the ones who are sharing personal experiences or experiences of the other people. The use of diversity statistics to filter surges of interest is an addition to methods used by the official sources [24], [49].

The Twitter analysis performed in this study can help to understand better its role and in general social media role in disaster management and the ways in which can be beneficial for the process. The secondary goal of deep tweet analysis was to detect the baseline rate of tweets in Italy over the last two years (2020-2021) and then use this information to detect whether the rate was within the boundaries for the third year (2022). This approach helped in assessing the quantitative metric for intensity of interest as comparing the results with official sources that were available for use. Indeed, the observed tweet volumes were analogous to the surges or spikes in attention that emergency managers use to identify events of interest [25].

The study was also complemented by the qualitative part about some most common experiences. From all the categories that may have been impacted by the drought conditions, the results showed that the most affected was water resources. Approximately 46% of tweets were related to this category, highlighting its central role during drought events, and even after leaving impacts for the future. Moreover, for other categories where water management plays a crucial role, the impacts are escalating, which puts the category of agriculture in the second place of the most impacted and damaged fields.

In one way, or another, they had their life affected. Direct impacts drought showed on water scarcity escalating with crisis. The plants and wildlife, as well as whole agriculture were also categories that were damaged from the first wave of precipitation deficiency. The data reported showed that it succeeds in recognizing the areas that suffered the most from drought and it could be further useful to allocate assistance in that direction. Furthermore, it can be considered as an additional metric to evaluate, as it captures drought from another, human aspect.

The awareness of the drought problem, raised by tweets analysis can positively impact the decision-making process in the impacted fields. Even basic content retrieved from Twitter by a single keyword search, like in this case, can provide insight on what type of impacts people are experiencing and eventually identify new types of categories to consider. If awareness is a precursor for the action, it is possible to infer that the innovation or advances are expected to take place in drought planning and policy.

By looking at the tweets extracted, it is noticed that they could also be grouped based on whether they are reporting information on the current situation and assessing the damage or providing possible suggestions for improvements instead. The division in this way could help distinguish pure drought assessment from assistance demand and provide relevant information to each field separately.

Another possible improvement can be exploring other search terms more specific to each category. Many tweets may relate to dry conditions even without explicitly using the word "*drought*".

There is a space for further advancement in geolocating drought tweets in some way. Even though the tweets extracted were in Italian, a lot of them were talking about drought situation outside Italy. With all possible obstacles that geolocation is bringing, it can become challenging to obtain some accurate results. Furthermore, focusing the study just on tweets written by the people confronted with drought in first person, can omit the tweets that are posted by others and therefore miss to bring attention to the others. It is necessary to find the balance between these aspects to be able to obtain the most suitable and effective results.

This topic is still under discussion, and it continues to attract the needed attention, as by considering moderate emission scenario (RCP4.5) drought events are calculated to become considerably more severe in the period 2071–2100 than 1981–2010 for Mediterranean region or even larger [73].

Bibliography

[1] M. Wiegmann, J. Kersten, H. Senaratne, M. Potthast, F. Klan, and B. Stein, "Opportunities and risks of disaster data from social media: A systematic review of incident information," *Natural Hazards and Earth System Sciences*, vol. 21, no. 5. Copernicus GmbH, pp. 1431–1444, May 06, 2021. doi: 10.5194/nhess-21-1431-2021.

[2] A. Crayton *et al.*, "Narratives and Needs: Analyzing Experiences of Cyclone Amphan Using Twitter Discourse," Sep. 2020, [Online]. Available: http://arxiv.org/abs/2009.05560

[3] H. Carrão, G. Naumann, and P. Barbosa, "Global projections of drought hazard in a warming climate: a prime for disaster risk management," *Clim Dyn*, vol. 50, no. 5–6, pp. 2137–2155, Mar. 2018, doi: 10.1007/s00382-017-3740-8.

[4] Nishadi. Eriyagama, Vladimir. Smakhtin, Nilantha. Gamage, and International Water Management Institute., *Mapping drought patterns and impacts : a global perspective*. [International Water Management Institute], 2009.

[5] J. Sodoge, C. Kuhlicke, and M. M. de Brito, "Automatized spatio-temporal detection of drought impacts from newspaper articles using natural language processing and machine learning," *Weather Clim Extrem*, vol. 41, p. 100574, Sep. 2023, doi: 10.1016/j.wace.2023.100574.

[6] S. Mukherjee, S. Wang, D. Hirschfeld, J. Lisonbee, and R. Gillies, "Feasibility of Adding Twitter Data to Aid Drought Depiction: Case Study in Colorado," *Water (Switzerland)*, vol. 14, no. 18, Sep. 2022, doi: 10.3390/w14182773.

[7] C. Havas *et al.*, "E2mC: Improving emergency management service practice through social media and crowdsourcing analysis in near real time," *Sensors (Switzerland)*, vol. 17, no. 12, Dec. 2017, doi: 10.3390/s17122766.

[8] Z. Zhang, N. He, D. Li, H. Gao, T. Gao, and C. Zhou, "Federated transfer learning for disaster classification in social computing networks," *Journal of Safety Science and Resilience*, vol. 3, no. 1, pp. 15–23, Mar. 2022, doi: 10.1016/j.jnlssr.2021.10.007.

[9] V. Linardos, M. Drakaki, and P. Tzionas, "A transformers-based approach on industrial disaster consequence identification from accident narratives," *Procedia Comput Sci*, vol. 217, pp. 1446–1451, 2023, doi: 10.1016/j.procs.2022.12.343.

[10] C. Wang, P. Nulty, and D. Lillis, "Congcong Wang et al. Transformer-based Multi-task Learning for Disaster Tweet Categorisation Transformer-based Multi-task Learning for Disaster Tweet Categorisation." [Online]. Available: https://github.com/wangcongcong123/crisis-mtl

[11] V. Balakrishnan *et al.*, "A Comprehensive Analysis of Transformer-Deep Neural Network Models in Twitter Disaster Detection," *Mathematics*, vol. 10, no. 24, Dec. 2022, doi: 10.3390/math10244664.

[12] B. Zhang, F. Schilder, K. H. Smith, M. J. Hayes, S. Harms, and T. Tadesse, "TweetDrought: A Deep-Learning Drought Impacts Recognizer based on Twitter Data."

[13] J. Spinoni, G. Naumann, J. Vogt, and P. Barbosa, *Meteorological droughts in Europe events and impacts : past trends and future projections*.

[14] Rodgers Alison Ince, "Understanding Droughts," *National Geographic Society*, 2022.

[15] A. K. Mishra and V. P. Singh, "A review of drought concepts," *Journal of Hydrology*, vol. 391, no. 1–2. pp. 202–216, Sep. 2010. doi: 10.1016/j.jhydrol.2010.07.012.

[16] I. R. Orimoloye, "Agricultural Drought and Its Potential Impacts: Enabling Decision-Support for Food Security in Vulnerable Regions," *Front Sustain Food Syst*, vol. 6, Feb. 2022, doi: 10.3389/fsufs.2022.838824.

[17] J. M. Ndayiragije and F. Li, "Effectiveness of Drought Indices in the Assessment of Different Types of Droughts, Managing and Mitigating Their Effects," *Climate*, vol. 10, no. 9. MDPI, Sep. 01, 2022. doi: 10.3390/cli10090125.

[18] K. H. Smith, A. J. Tyre, Z. Tang, M. J. Hayes, and F. Adnan Akyuz, "Calibrating human attention as indicator monitoring #drought in the twittersphere," *Bull Am Meteorol Soc*, vol. 101, no. 10, pp. E1801–E1819, Oct. 2020, doi: 10.1175/BAMS-D-19-0342.1.

[19] B. Zhang, F. K. A. Salem, M. J. Hayes, and T. Tadesse, "Quantitative Assessment of Drought Impacts Using XGBoost based on the Drought Impact Reporter," Nov. 2022, [Online]. Available: http://arxiv.org/abs/2211.02768

[20] A. Zargar, R. Sadiq, B. Naser, and F. I. Khan, "A review of drought indices," *Environmental Reviews*, vol. 19, no. 1. National Research Council of Canada, pp. 333–349, 2011. doi: 10.1139/a11-013.

[21] A. Singleton, "Forecasting Drought in Europe with the Standardized Precipitation Index An assessment of the performance of the European Centre for Medium Range Weather Forecasts Variable Resolution Ensemble Prediction System", doi: 10.2788/16522.

[22] "Copernicus European Drought Observatory (EDO) - SPI." [Online]. Available: https://edo.jrc.ec.europa.eu/

[23] W. Meteorological Organization and G. Water Partnership, "Integrated Drought Management Programme Handbook of Drought Indicators and Indices." [Online]. Available: www.droughtmanagement.info

[24] "Osservatorio siccità," *https://drought.climateservices.it/*.

[25] "Bollettino Drought December," *https://drought.climateservices.it/bollettino-italia/bollettino-dicembre-2022/*.

[26] "Copernicus European Drought Observatory (EDO) - FAPAR." [Online]. Available: http://edo.jrc.ec.europa.eu/

[27] B. Moccia, C. Mineo, E. Ridolfi, F. Russo, and F. Napolitano, "SPI-Based Drought Classification in Italy: Influence of Different Probability Distribution Functions," *Water (Switzerland)*, vol. 14, no. 22, Nov. 2022, doi: 10.3390/w14223668.

[28] H. Carrão, G. Naumann, and P. Barbosa, "Mapping global patterns of drought risk: An empirical framework based on sub-national estimates of hazard, exposure and vulnerability §", doi: 10.1016/j.gloenvcha.2016.04.012.

[29] K. Stahl *et al.*, "Impacts of European drought events Impacts of European drought events: insights from an international database of text-based reports Impacts of European drought events," *Nat. Hazards Earth Syst. Sci. Discuss*, vol. 3, pp. 5453–5492, 2015, doi: 10.5194/nhessd-3-5453-2015.

[30] National Integrated Drought Information System, "Sectors," *https://www.drought.gov/sectors*.

[31] D. García-León, G. Standardi, and A. Staccione, "An integrated approach for the estimation of agricultural drought costs," 2020, doi: 10.1016/j.landusepol.2020.104923.

[32] Fargione Riccardo, Modesto Lucrezia, and Vivani Dario, "Siccità, il mondo è a secco", Accessed: Jun. 17, 2023. [Online]. Available:

https://www.divulgastudi.it/wp-content/uploads/2022/07/Paper_5_Siccita%CC%80.pdf

[33] J. W. Lee, E. M. Hong, J. U. Kim, W. J. Jang, C. G. Jung, and S. J. Kim, "Evaluation of agricultural drought in South Korea using socio-economic drought information," *International Journal of Disaster Risk Reduction*, vol. 74, May 2022, doi: 10.1016/j.ijdrr.2022.102936.

[34] J. B. Houston *et al.*, "Social media and disasters: A functional framework for social media use in disaster planning, response, and research," *Disasters*, vol. 39, no. 1, pp. 1–22, Jan. 2015, doi: 10.1111/disa.12092.

[35] D. Feldman *et al.*, "Communicating flood risk: Looking back and forward at traditional and social media outlets," 2015, doi: 10.1016/j.ijdrr.2015.12.004.

[36] D. Wang, B. K. Szymanski, T. Abdelzaher, H. Ji, and L. Kaplan, "The Age of Social Sensing."

[37] V. Negri *et al.*, "Image-based Social Sensing: Combining AI and the Crowd to Mine Policy-Adherence Indicators from Twitter," Oct. 2020, [Online]. Available: http://arxiv.org/abs/2010.03021

[38] C. Bono *et al.*, "A Citizen Science Approach for Analyzing Social Media With Crowdsourcing," *IEEE Access*, vol. 11, pp. 15329–15347, 2023, doi: 10.1109/ACCESS.2023.3243791.

[39] P. T. Jaeger, B. Shneiderman, K. R. Fleischmann, J. Preece, Y. Qu, and P. Fei Wu, "Community response grids: E-government, social networks, and effective emergency management," *Telecomm Policy*, vol. 31, no. 10–11, pp. 592–604, Nov. 2007, doi: 10.1016/j.telpol.2007.07.008.

[40] C. Francalanci and B. Pernici, "Data integration and quality requirements in emergency services."

[41] N. Pourebrahim, S. Sultana, J. Edwards, A. Gochanour, and S. Mohanty, "Understanding communication dynamics on Twitter during natural disasters: A case study of Hurricane Sandy," *International Journal of Disaster Risk Reduction*, vol. 37, p. 101176, Jul. 2019, doi: 10.1016/J.IJDRR.2019.101176.

[42] S. E. Vieweg, "Qatar Computing Research Institute From the SelectedWorks of Sarah E. Vieweg 2012 Situational Awareness in Mass Emergency: A Behavioral and Linguistic Analysis of Microblogged Communications." [Online]. Available: http://works.bepress.com/vieweg/15/

[43] J. Coche, J. Kropczynski, A. Montarnal, A. Tapia, and F. Benaben, "Actionability in a Situational Awareness world: Implications for social media processing system design."

[44] F. Sufi, "A decision support system for extracting artificial intelligence-driven insights from live twitter feeds on natural disasters," *Decision Analytics Journal*, vol. 5, Dec. 2022, doi: 10.1016/j.dajour.2022.100130.

[45] J. Kersten, J. Bongard, and F. Klan, "Combining Supervised and Unsupervised Learning to Detect and Semantically Aggregate Crisis-Related Twitter Content." [Online]. Available: www.dlr.de/dw/en

[46] A. Vaswani *et al.*, "Attention Is All You Need," Jun. 2017, [Online]. Available: http://arxiv.org/abs/1706.03762

[47] J. Liu, T. Singhal, L. T. M. Blessing, K. L. Wood, and K. H. Lim, "CrisisBERT: a Robust Transformer for Crisis Classification and Contextual Crisis Embedding," May 2020, [Online]. Available: http://arxiv.org/abs/2005.06627

[48] "World Meteorological Organization." https://public.wmo.int/en (accessed Jun. 23, 2023).

[49] "European Drought Observatory," *https://edo.jrc.ec.europa.eu/edov2/php/index.php?id=1000*.

[50] "National Research Council (CNR)." https://www.cnr.it/en (accessed Jun. 10, 2023).

[51] "ISTAT," *https://www.istat.it/*.

[52] "EDO report March 2022." Accessed: Jun. 13, 2023. [Online]. Available: https://edo.jrc.ec.europa.eu/documents/news/GDO-EDODroughtNews202203_Northern_Italy.pdf

[53] "Heat and Cold Wave Index (HCWI)." [Online]. Available: http://edo.jrc.ec.europa.eu/

[54] N. Gobron *et al.*, "The state of vegetation in Europe following the 2003 drought," *Int J Remote Sens*, vol. 26, no. 9, pp. 2013–2020, May 2005, doi: 10.1080/01431160412331330293.

[55] Michele Bertelli, "'No water, no life': Drought threatens farmers and food in Italy," Apr. 14, 2023.

[56] "ISTAT archive," *https://www.istat.it/it/archivio/280257*.

[57] "ENTSOE," *https://transparency.entsoe.eu/*.

[58] "EDO report April 2022." Accessed: Jun. 13, 2023. [Online]. Available: https://edo.jrc.ec.europa.eu/documents/news/GDO-EDODroughtNews202204_Europe.pdf

[59] "EDO report July 2022." Accessed: Jun. 13, 2023. [Online]. Available: https://edo.jrc.ec.europa.eu/documents/news/GDO-EDODroughtNews202207_Europe.pdf

[60] "EDO report August 2022." Accessed: Jun. 13, 2023. [Online]. Available: https://edo.jrc.ec.europa.eu/documents/news/GDO-EDODroughtNews202208_Europe.pdf

[61] "Twitter Developer," *https://developer.twitter.com/en/docs/*.

[62] A. Kruspe, J. Kersten, and F. Klan, "Detection of informative tweets in crisis events", doi: 10.5194/nhess-2020-214.

[63] H. M. Zahera, I. Elgendy, R. Jalota, and M. A. Sherif, "Fine-tuned BERT Model for Multi-Label Tweets Classification." [Online]. Available: https://www.nltk.org/_modules/nltk/stem/wordnet.html

[64] J. Devlin, M.-W. Chang, K. Lee, K. T. Google, and A. I. Language, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." [Online]. Available: https://github.com/tensorflow/tensor2tensor

[65] "Hugging Face," *https://huggingface.co/*.

[66] T. Pires, E. Schlinger, and D. Garrette, "How multilingual is Multilingual BERT?," Jun. 2019, [Online]. Available: http://arxiv.org/abs/1906.01502

[67] A. Conneau *et al.*, "XNLI: Evaluating Cross-lingual Sentence Representations," Sep. 2018, [Online]. Available: http://arxiv.org/abs/1809.05053

[68] "BERT github." https://github.com/google-research/bert (accessed Jun. 01, 2023).

[69] C. Yang and X. Ma, "Improving Stability of Fine-Tuning Pretrained Language Models via Component-Wise Gradient Norm Clipping," Oct. 2022, [Online]. Available: http://arxiv.org/abs/2210.10325

[70] E. S. Negara, D. Triadi, and R. Andryani, "Topic Modelling Twitter Data with Latent Dirichlet Allocation Method."

[71] R. Mehrotra, S. Sanner, W. Buntine, and L. Xie, *Improving LDA Topic Models for Microblogs via Tweet Pooling and Automatic Labeling*. 2013. [Online]. Available: http://snap.stanford.edu/data/twitter7.html

[72] A. K. Mishra and V. P. Singh, "Drought modeling - A review," *Journal of Hydrology*, vol. 403, no. 1–2. pp. 157–175, Jun. 06, 2011. doi: 10.1016/j.jhydrol.2011.03.049.

[73] J. Spinoni, J. V. Vogt, G. Naumann, P. Barbosa, and A. Dosio, "Will drought events become more frequent and severe in Europe?," *International Journal of Climatology*, vol. 38, no. 4, pp. 1718–1736, Mar. 2018, doi: 10.1002/joc.5291.

A. Appendix – Anomaly indices information processing

The following code snippet (*Listing A.1*) is explaining a possible way how to represent data from *.nc* file type on the map. The specific file data contain measurements for maximum temperature anomaly recorded on a daily basis.

```
max_temp_data = xr.open_dataset('/kaggle/input/max-temp-anomaly/tpman_m_euu_2022010
1_20221231_d.nc')
max_temp_monthly = max_temp_data.resample(time='m').mean()
lat_temp = max_temp_data.variables['lat'][:]
lon_temp = max_temp_data.variables['lon'][:]
time_temp = max_temp_data.variables['time'][:]
max_temp = max_temp_data.variables['tpman']
max_temp_a = max_temp_monthly.variables['tpman']

m = Basemap(projection='merc', llcrnrlat=28.125, urcrnrlat=71.875, llcrnrlon=-31.87
5, urcrnrlon=49.875, resolution='l')
    lon2_temp, lat2_temp = np.meshgrid(lon_temp,lat_temp)
    x, y = m(lon2_temp, lat2_temp)
    fig = plt.figure(figsize=(35,15))
    m.drawcoastlines()
    parallels = np.arange(36,49, 12.)
    meridians = np.arange(6,21, 14.)
    m.drawparallels(parallels,labels=[1,0,0,0],fontsize=10)
    m.drawmeridians(meridians,labels=[0,0,0,1],fontsize=10)
    m.drawmapboundary(fill_color='white')
    cmap = plt.cm.get_cmap('jet')
    cs = m.pcolor(x,y,max_temp_a[11,:,:],cmap=cmap)
    cbar = m.colorbar(cs, location='bottom', pad="10%")
```

Listing A.1: Example of representing on the map maximum temperature anomaly measurements for one month

Below the code (*Listing A.2*) used for showing FAPAR measurements recorded and retrieved by EDO on the map. At the time of measurement extraction, the data was available for first eleven months of 2022.

```
fapar_data = xr.open_dataset('/kaggle/input/fapar-anomaly/fapan_m_euu_20220101_2022
1111_t.nc')

fapar_monthly = fapar_data.resample(time='m').mean()

lat_fapan = fapar_data.variables['lat'][:]
lon_fapan = fapar_data.variables['lon'][:]
time_fapan = fapar_data.variables['time'][:]
fapan = fapar_data.variables['fapan']
fapar_a = fapar_monthly.variables['fapan']

i = 0
while i < 11:

    m = Basemap(projection='merc', llcrnrlat=28.021, urcrnrlat=72.021, llcrnrlon=-3
2.021, urcrnrlon=49.979, resolution='l')
    lon2, lat2 = np.meshgrid(lon_fapan,lat_fapan)
    x, y = m(lon2, lat2)
    fig = plt.figure(figsize=(20,10))
    m.drawcoastlines()
    parallels = np.arange(36,49, 12.)
    meridians = np.arange(6,21, 14.)
    m.drawparallels(parallels,labels=[1,0,0,0],fontsize=10)
    m.drawmeridians(meridians,labels=[0,0,0,1],fontsize=10)
    m.drawmapboundary(fill_color='white')
    cmap = plt.cm.get_cmap('jet_r')
    cs = m.pcolor(x,y,fapar_a[i,:,:],cmap=cmap,vmin=-2, vmax=2)
    cbar = m.colorbar(cs, location='bottom', pad="10%")
    i = i + 1
```

Listing A.2: Processing measurements FAPAR anomaly obtained from EDO database

B. Appendix – Classification conventions

- Agriculture:
    - general information about agriculture
        - *Coldiretti: Italian agriculture is facing controversial phenomena, where exceptional waves of bad weather alternate with persistent drought in a few hours".*
    - soil and desertification
        - *Water is being plundered by private businesses as the country is facing a decade-long drought and desertification that is destroying agriculture and livestock*
    - specific culture
        - *Drought is causing a drop in olive production, called for a state of calamity*
    - farmers problems
        - *It's nice when in July, after months of drought, you go on the A21 and between Cremona and Piacenza you see endless kilometers of dry land and the farmers interviewed on Tg4 asking for damages.*
    - connections with other impacts
        - *Desertification on 70% of Sicily. 14 billion euro of damage to Italian agriculture*
- Economy:
    - impacts on the field of economy
        - *Venice under water. Whole islands disappeared within a few years. Droughts and floods cripple the economy*
    - assessment of damage
        - *Hydrogeological risk and drought cost 2.5 billion a year, 4,300 projects ready*

  - impact on energy field
    - *Heat and drought threaten the operation of nuclear power plants*
  - interconnection with other categories
    - *Potatoes 60% more expensive, due to drought and imports*
- Fire:
  - connection drought and fire
    - *Italy in the grip of drought: since the beginning of the year one fire a day, and -30% rainfall*
  - fire risks
    - *Abnormal heat and drought, Italy at risk of fires*
- Plants & Wildlife:
  - plants
    - *Droughts and other effects of global warming are making plants fragile and vulnerable. The #EPICON project is studying sorghum to understand how the plant reacts and defends itself.*
  - forests
    - *Urban forestation against droughts, heat waves, extreme winds, floods and landslides*
  - animals
    - *Thousands of fish died due to drought but not only...*
  - improvement of the plant resistance
    - *Towards the possibility of optimizing water-use efficiency for a better response of plants to drought*
- Relief, Response & Restrictions:
  - consequences
    - *For millions of human beings, the journey is something else: it is a desperate escape from the hell of wars, ethnic cleansing, absolute poverty, environmental disasters, hunger, drought.*
  - response and measures
    - *Floods, droughts, melting glaciers are the disastrous effects of climate change. The time has come to act against global warming, starting with schools and our cities. The future is in our hands*
  - suggested actions
    - *for climate change why not build reservoirs where rainwater is collected since there are then long periods of drought? This proposal was already made by other governments but never implemented*
  - limitations

        - *ARIF, the Puglia Region agency that provides irrigation services to farmers, is blocked by atavistic problems. In this period of prolonged drought I have asked to keep the wells efficient to save the crops, but adequate resources and men are urgently needed.*
    - restrictions
        - *Saving water shouldn't become a public problem only in times of drought, when #water is undeniably...*
    - assistance
        - *Porta acqua agli animali della savana per salvarli dalla siccità conseguente al riscaldamento globale: lo chiamano water man*
- Society&Public Health:
    - population affected
        - *Data shows that an increasing number of countries are exposed to extreme variations in climate: floods, droughts and storms are increasingly negatively impacting households.*
    - public health
        - *The Unesco heritage dried up by the most severe wave of drought in the last forty years. The economic and environmental damage is enormous. Seven million people risk dying of hunger.*
    - democratic migrations
        - *How many people are and will be put at risk from the climate crisis. How many people have died from drought or those who emigrated to escape drought to date?*
- Water Supply&Quality:
    - water resources
        - *Wastewater to beat drought. The University of Sassari and the European Union are betting on the project*
    - water crisis
        - *Italy risks running out of water: the WRI study on the effects of drought*
    - water conservation
        - *A future for our rivers. From instability to drought to the protection of freshwater ecosystems: strategies for adaptation to climate change*
    - water quality
        - *It has been raining for 4 days, after months of drought, unfortunately the maintenance and cleaning of the water drainage channels has not been performed*

C. Appendix – BERT pre-trained parameters

```
The BERT model has 199 different named parameters.

==== Embedding Layer ====

embeddings.word_embeddings.weight                          (119547, 768)
embeddings.position_embeddings.weight                        (512, 768)
embeddings.token_type_embeddings.weight                        (2, 768)
embeddings.LayerNorm.weight                                      (768,)
embeddings.LayerNorm.bias                                        (768,)

==== First Transformer ====

encoder.layer.0.attention.self.query.weight                  (768, 768)
encoder.layer.0.attention.self.query.bias                        (768,)
encoder.layer.0.attention.self.key.weight                    (768, 768)
encoder.layer.0.attention.self.key.bias                          (768,)
encoder.layer.0.attention.self.value.weight                  (768, 768)
encoder.layer.0.attention.self.value.bias                        (768,)
encoder.layer.0.attention.output.dense.weight                (768, 768)
encoder.layer.0.attention.output.dense.bias                      (768,)
encoder.layer.0.attention.output.LayerNorm.weight                (768,)
encoder.layer.0.attention.output.LayerNorm.bias                  (768,)
encoder.layer.0.intermediate.dense.weight                   (3072, 768)
encoder.layer.0.intermediate.dense.bias                         (3072,)
encoder.layer.0.output.dense.weight                         (768, 3072)
encoder.layer.0.output.dense.bias                                (768,)
encoder.layer.0.output.LayerNorm.weight                          (768,)
encoder.layer.0.output.LayerNorm.bias                            (768,)

==== Output Layer ====

encoder.layer.11.output.LayerNorm.weight                         (768,)
encoder.layer.11.output.LayerNorm.bias                           (768,)
pooler.dense.weight                                          (768, 768)

pooler.dense.bias                                                (768,)
```

Listing C.1: Pre-trained parameters that can be updated in BERT multilingual

D. Appendix – Accuracy of the model on categories of drought impact

To understand better if the model struggles to generalize well for minority classes, it has been also investigated the performance on individual classes to identify those with lower accuracy (see *Table D*.1*)*. Accuracy is calculated as the number of all true positive and true negative cases divided by the number of all cases.

Accuracy for individual classes	**BERT multilingual**			
Architecture	simple		complex	
Batch size	16	32	16	32
Learning rate	1e-4	1e-3	1e-4	1e-3
Epochs	5	5	5	5
Agriculture	0.8379	0.8471	0.8450	0.8760
Economy	0.8999	0.8999	0.8837	0.8971
Fire	0.8682	0.9197	0.9119	0.9281
Plants & Wildlife	0.7674	0.7851	0.7533	0.7865
Relief, Response & Restrictions	0.8436	0.8739	0.8527	0.8746
Society & Public Health	0.8584	0.8732	0.8696	0.8746
Water Supply & Quality	0.7829	0.8280	0.8175	0.8266

Table D.1: Accuracy for individual classes

From the attached, it can be inferred that *Plants & Wildlife* is showing scarce performance and may be lowering down the macro precision of the model. In general, the effects of this category can be relatively intertwined with the ones from *Agriculture* since its the common terms considered are flora and fauna, biodiversity, and habitat preservation that are intersecting to some some extent. Like for previous case, below reported the table with accuracy evaluated through all classes (*Table D*.2).

Unfreezing pre-trained weights	**BERT multilingual**			
Architecture	simple		complex	
Batch size	16	32	16	32
Learning rate	3e-5	5e-5	3e-5	5e-5
Epochs	4	4	4	4
Agriculture	0.9281	0.9274	0.9267	0.9366
Economy	0.9486	0.9345	0.9464	0.9408
Fire	0.9704	0.9676	0.9591	0.9634
Plants & Wildlife	0.8548	0.8640	0.8746	0.8689
Relief, Response & Restrictions	0.9168	0.9140	0.9049	0.9042
Society & Public Health	0.8971	0.8992	0.9126	0.9070
Water Supply & Quality	0.8999	0.8950	0.8922	0.8999

Table D.2: Accuracy results per category when unfreezing parameters

Similarly, as when freezing the parameters, also in this case the sparse performance is showing *Plants & Wildlife* along with *Water Supply & Quality* that is following it with slightly better results when it comes to evalute the accuracy. Despite some classess have shown high accuracy, examples for their classes were in minority and the model besides true (positive/negative) values predicted also a significant number of false (positive/negative) values. The explanation can lay in the fact that the samples for this classes were imbalanced in comparison to other classes. In fact, from *Table D.2* the two classes that contain the least number of samples are indeed the ones that show the poorest performance in terms of precision and recall (*Economy* and *Society & Public Health*).