



Constructing a Social Media-Based Index to Capture the Socio-Economic Impacts of Droughts

Jingxian Wang^{1,2}, Barbara Pernici², Matteo Giuliani², and Andrea Castelletti²

¹Scuola Universitaria Superiore IUSS Pavia, Pavia, Italy

²Politecnico di Milano, Milan, Italy

Droughts, unlike other natural disasters, have complex and multifaceted consequences that spread over vast regions and extended durations. Several indices, including the Standardised Precipitation Index (SPI) and Standardised Precipitation Evapotranspiration Index (SPEI), were designed to measure and quantify droughts. However, they predominantly focus on the meteorological and hydrological aspects of drought events, often overlooking the social and economic impacts. On the other hand, existing impact databases like the European Drought Impact Report Inventory (EDII) and the European Drought Impact Database (EDID) are usually constrained by temporal and spatial resolution due to the limitations of available data sources. Given this context, social media, bolstered by the rapid evolution of technology, offers a unique perspective. Users on social media can share their firsthand experiences and perceived impacts of droughts, providing a rich source of indirect socio-economic impact information that is often missed by traditional methods.

The objective of this study is to develop an ad-hoc drought index that reflects the socio-economic impacts of droughts using information gathered from social media, and compare the ad-hoc drought index to physical drought indices to evaluate the usefulness and accuracy of the state-of-the-art method. While current literature underscores the importance of integrating social media as a complementary data source to improve drought detection and response, most studies focus on classifying the impacts of droughts into different categories based on text mining. To the best of our knowledge, none have transformed these text-derived impacts from social media into a single numeric index to help decision-makers grasp the drought situation quickly and efficiently. Thus, we aim to address this gap and, ultimately, inform dynamic and adaptive drought management strategies.

While the goal is to capture the impacts of drought at a Pan-European scale, Italy was selected for preliminary studies, due to a significant drought event that occurred in 2022. This event drew attention from various sectors and offered a snapshot of socio-economic impacts on local communities. Notably, the number of tweets containing the keywords “siccità” or “siccita” (drought) increased more than tenfold in 2022 compared to 2020 and 2021. We conducted location extraction, topical modelling, and classification to filter out irrelevant tweets, identify the regions where information is shared, and categorise the sectors in which impacts are perceived by

local residents. Each tweet is scored based on the positivity or negativity of its narrative through sentiment analysis, which indicates the gravity of its impact. Subsequently, this score is combined with a manual evaluation of the intensity for constructing the ad-hoc drought index. Once established, the ad-hoc drought index is compared to physical drought indices. We do expect the ad-hoc drought index to reveal patterns that were not previously seen with physical drought indices, providing a broader and deeper understanding of the impacts of droughts on societies and economies.