

Article

Feasibility of Adding Twitter Data to Aid Drought Depiction: Case Study in Colorado

Sarbjit Mukherjee ¹, Simon Wang ^{1,*} , Daniella Hirschfeld ² , Joel Lisonbee ³  and Robert Gillies ¹¹ Utah Climate Center, Department of Plants, Soil & Climate, Utah State University, Logan, UT 84322, USA² Department of Landscape Architecture & Environmental Planning, Utah State University, Logan, UT 84322, USA³ NOAA/National Integrated Drought Information System, and Cooperative Institute for Research in the Environmental Sciences, University of Colorado Boulder, Boulder, CO 80305, USA

* Correspondence: simon.wang@usu.edu

Abstract: The use of social media, such as Twitter, has changed the information landscape for citizens' participation in crisis response and recovery activities. Given that drought progression is slow and also 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 how the changing usage facilitates drought detection. For this reason, contemporary analysis of how social media data, in conjunction with meteorological records, was conducted towards improvement in the detection of drought and its progression. The research utilized machine learning techniques applied over satellite-derived drought conditions in Colorado. Three different machine learning techniques were examined: the generalized linear model, support vector machines and deep learning, each applied to test the integration of Twitter data with meteorological records as a predictor of drought development. It is found that the integration of data resources is viable given that the Twitter-based model outperformed the control run which did not include social media input. Eight of the ten models tested showed quantifiable improvements in the performance over the control run model, suggesting that the Twitter-based model was superior in predicting drought severity. Future work lies in expanding this method to depict drought in the western U.S.

Keywords: drought; Twitter; machine learning; Colorado

Citation: Mukherjee, S.; Wang, S.; Hirschfeld, D.; Lisonbee, J.; Gillies, R. Feasibility of Adding Twitter Data to Aid Drought Depiction: Case Study in Colorado. *Water* **2022**, *14*, 2773. <https://doi.org/10.3390/w14182773>

Academic Editors: Zheng Duan and Babak Mohammadi

Received: 2 June 2022

Accepted: 24 August 2022

Published: 6 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Drought is increasingly impacting the American West, threatening major water supplies such as the Colorado River [1]. Due to drought's slow and elusive emergence and sometimes speedy intensification [2], the phenomenon presents a challenge for stakeholders and society as a whole, in timely responsiveness [3]. The drought research community has generally agreed that, due to its persistence, drought allows society to plan for mitigation strategies ahead of time as long as citizens are given actionable information about the developing state of affairs [4]. However, despite state-of-the-art drought monitoring and forecasting systems that are currently in place the drought-prone American West still suffered multi-billion-dollar losses from recent severe drought conditions [5]. Given the vulnerability of our society to future droughts, recent research has begun to examine not only the physical mechanism of drought, but also how society responds to drought [6]. Moreover, in the sparsely populated American West, the situation is further hindered by a lack of in situ meteorological and soil observations that underlies adequate drought depiction as well as its prediction.

The use of social media, such as Twitter, has changed the information landscape for citizens' participation in crisis response and recovery activities [7,8]. Social media has been used for information broadcasting during a variety of crisis events of both natural and human origins [9]. The majority of these social media systems work on the principle of

detecting variations from a baseline observation, such as the sudden increase in the use of certain predefined lexicons, as illustrated in [10–13]. Given that drought is a slow-moving process that is also spatially extensive [14], interesting questions arise on how Twitter usage may change during the progress of a major drought, i.e., one that is felt by a population at large, as well as how usage change might aid in the detection of drought. Past reports such as [15,16] have revealed heightened peoples' awareness, in recognizing the sudden threats posed by floods or fires, in Twitter posts; however, questions as to their awareness or information needs concerning droughts remain to be investigated. In the American West, we are also interested in first, learning the feasibility of human perception of drought through Twitter and second, can such observations can make up for data gaps in the ground station network?

Over the past few years, the computer science research community has made advancements in areas of machine learning and big data, which enables the use of technology towards exploring social responses to drought. The research of drought forecasting using machine learning models has gained prominence in recent years. Models have been developed using time series analysis [17,18], neural networks [19–24], fuzzy inference systems [25], support vector regression [26–28] and different ensemble techniques [29,30]. Although different models techniques have been designed in order to better understand and forecast drought, but various studies over the years have also suggested that a single indicator is not enough to explain the complexity and diversity of drought [31,32]. These models mainly use meteorological and hydrological observations as input without considering human-dimension data.

In the meantime, Twitter has made it easier to access past tweets through their Application Programming Interfaces (APIs) that enable two computer applications to access each other's data. The aforementioned means that social media data affords an investigator a huge source of unstructured big data. The use of deep learning techniques in a variety of applications (including natural language processing) also has further enabled researchers to analyze social media data at an expanded scale and with high accuracy [33,34]. Thus, the ability to obtain social media information coupled with the emergence of recent computer science techniques suggests a fresh tactic towards evaluating drought emergence.

Social media postings feature human emotion, and various researchers have analyzed social media data to extract sentiments of peoples' opinions in various contexts. Such approaches usually start by extracting text data from social media and then using sentiment analysis methods to capture user opinions and attitudes about a wide variety of topics including crisis management [11,35–40]. In the USA, "#drought" and related hashtags on Twitter have been found to increase correspondingly during high-impact drought events [8,41]. Researchers also have attempted to use data from Twitter to study climate change perceptions [42,43]. The aforesaid studies reflect the potential of using social media platforms like Twitter coupled with sentiment analysis methods for analyzing the progression of a drought. Nonetheless, there are challenges associated with such analysis: First, past studies such as [8,42] have found that peoples' concerns about climate-related matters were greatly influenced by media coverage, especially in the context of climate association with droughts and heat waves. Second, it is particularly difficult to automatically interpret humorous or sarcastic emotions in tweet content; this poses a considerable impediment in sentiment analysis.

It is feasible that a coupled analysis of social media data along with other meteorological sources, can enhance drought detection and capture the evolution of drought, especially in data-sparse regions like those that pervade the Western U.S. The analysis presented here features Twitter drought-related conversations during the most recent (2020–2021) drought in Colorado and how the addition of Twitter data affected drought monitoring.

2. Methodology and Data

The use of Twitter data to mine public opinion usually is structured as a pipeline that starts by collecting data regarding the event from Twitter, followed by processing and

cleaning the data, and then finishes by passing the data through a prediction model. We chose to collect Twitter data because it has grown to be popular in the USA, owing to its effective 140-character tweeting capability where people can simply use their smartphones to tweet about different topics. The prediction model is evaluated against the actual outcomes of the event based on the chosen evaluation metrics.

2.1. Data Collection

Researchers developed a number of drought indices based on meteorological or hydrological variables for drought monitoring and forecasting. Examples of drought indices include, but are not limited to, Standardized Precipitation Index (SPI), Standardized Precipitation Evapotranspiration Index (SPEI), Palmer Drought Severity Index (PDSI), Palmer Moisture Anomaly Index (Z-index), and China-Z index (CZI). Among them, the Palmer Drought Severity Index (PDSI, [44]) is one of the six more common drought indices for North America [45]. The US Drought monitor <https://droughtmonitor.unl.edu/> (accessed on 2 July 2021), uses the PDSI as one of the many indices subjected to weighting of subjective basis from individual authors. We obtained PDSI data from the gridMET dataset from the *Climatology Lab* (<http://www.climatologylab.org/gridmet.html>, accessed on 2 July 2021). The data is updated once every 5 days. Besides PDSI, we also analyzed the groundwater and soil moisture conditions derived from NASA's Gravity Recovery and Climate Experiment (GRACE) project given 94 their ability to measure the terrestrial water storage (i.e., variations in water stored at all 95 levels above and within the land surface). Through their "follow-on" satellites (GRACE-FO) and using data assimilation technique with the Catchment Land Surface Model, the fields of soil moisture and groundwater storage variations are derived from GRACE-FO's liquid water thickness observation [46]. The groundwater and soil moisture data used here as complementary indicators to PDSI are obtained from <https://nasagrace.unl.edu/> (accessed on 2 July 2021).

Twitter lets researchers access its data through two different tiered Application Programming Interface (API) services, Standard API and the Premium API. The Premium API is a subscription-based service, whereas the Standard API is free and primarily limited by a certain number of requests within a time frame. The premium tier of service provides access to the past 30 days of Twitter data or to the full history of Twitter data. We started by collecting tweets in real-time using the Standard Twitter API. For this, we searched for keywords which are closely related to drought, such as 'soil moisture', 'streamflow', 'drought', 'DroughtMonitor', 'drought20', 'Drought2020', 'drought21', 'less water', 'crops', 'farmer', 'dry', 'dried' etc. There could be many other terms or combination of terms for use. As shown in Table 1, we see the top used words in 2019 and 2020 depending upon their usage frequency in the tweets. We wrote a Python script for collecting tweets that had location information of Colorado. However, using the real-time streaming setup has a compromise in that most tweets did not have a geo-location attached to them.

Table 1. Sample of top used words in tweets describing drought in 2019 and 2020. Root words are presented here for a better understanding. For example: ['pray', 'praying', 'prayer', 'prayed'] becomes ['pray', 'pray', 'prayer', 'pray'].

Year	Top Words Used in Order of Their Corresponding Frequency of Use
2019	drought, colorado, water, plan, river, colorado river, climate, condition, contingency, monitor, change, dri, cowx, snow, droughtmonitor, continu, basin, need, help, news, increase, record, west, southwest, wildfire, summer
2020	drought, url, colorado, water, condit, year, dri, extrem, climat, week,rt ,fire, chang, cowx, wildfir, sever, area, across, droughtmonitor, time, record, rain, west, river, moderate , weather, expand, experienc, impact, high, warm, southern, dryness

Next, we used Twitter's Premium API to collect historical tweets along with the current ones. The premium API helped us add another search term to our query in the form of

'profile_region:colorado'. This helped identify the tweets regarding drought originating from users whose profile location is in Colorado. Based on this method, we collected close to 38,000 tweets originating from Colorado during 2019, 2020 & 2021 (For 2021 the data collection period was January–April). Concerning the role of drought, the period in which our models were trained coincided with a developing drought. Thus, the inclusion of data that would incorporate times during which drought was varying could lead the models to reflect better PDSI variations. However, there is a downside, in that tweets about drought would decrease under “peace or leisure times” conditions (i.e., no drought tweets), in such cases data would be fairly limited and there is not a clear way of fixing this issue.

2.2. Data Cleaning

One of the challenges of Twitter data mining was the data cleaning step. Twitter users refer to the term “drought” in a variety of contexts, from referring to traditional climate drought to “trophy drought” or “playoff drought” in the context of a sporting event, and so on. Along with the above, another challenge we faced in data cleaning concerns the localizing of the data. As previously mentioned, we collected tweets from users whose ‘profile_region’ was set to ‘Colorado’ on Twitter. As a result, we had a number of tweets which, although being generated from Colorado, described the drought status for regions outside of Colorado. There were also instances where the term ‘drought’ was used as a proper noun to refer to a very popular music label. In Table 2, we show some of the common terms we searched for to clean our data. Although in most cases we were able to completely remove a tweet if it contained one of the search terms, there were still some cases where a tweet containing an above search term could also have some reference to drought condition in Colorado. That made a labor-intensive manual work, as we had to be careful while removing tweets so that Colorado drought related tweets did not get removed by mistake. This level of work echoes the saying “designing a good Machine Learning system comes with ample Man Labor!”. We acknowledge that a complex linguistic technique could have been developed for data cleaning but the main goal of this project was not to design a perfect data cleaning algorithm. In Table 2, we present some of the search terms used to remove tweets not related to meteorological drought.

Table 2. Sample of the search terms for data cleaning and its corresponding category.

Categories	Search Terms
Sports	game, broncos, rams, sox, lakers, soccer, champion, super bowl, scoring, home run, league, touchdown, @Rockies, medal, rockies division, coach, playoff, hoops, jazz, galaxy, nuggets ...
Location	california, africa, australia, kansas, england, kingdom, somalia, sydney, pakistan, india, costa rica, vietnam, britain, china, ...
General	elephants, koala, music, beyonce, lil wayne, cancer, rapper, aussie, song, @GavinNewsom, ...

Tweets usually contain a lot of information apart from the text, like mentions, hashtags, urls, emojis or symbols. Normal language models cannot parse those data, so we needed to clean up the tweet and replace tokens that actually contains meaningful information for the model. The preprocessing steps we took are:

1. Lower Casing: Each text is converted to lowercase.
2. Replacing URLs: Links starting with ‘http’ or ‘https’ or ‘www’ are replaced by ‘<url>’.
3. Replacing Usernames: Replace @Usernames with word ‘<user>’. [e.g., ‘@DroughtTalker’ to ‘<user>’].
4. Replacing Consecutive letters: 3 or more consecutive letters are replaced by 2 letters. [e.g., ‘Heyyyy’ to ‘Heyy’].
5. Replacing Emojis: Replace emojis by using a regex expression. [e.g., ‘:)’ to ‘<smile>’]

6. Replacing Contractions: Replacing contractions with their meanings. [e.g., “ca not” to ‘can not’]
7. Removing Non-Alphabets: Replacing characters except Digits, Alphabets and pre-defined Symbols with a space. [e.g., \$heat@t> to heat t]

As much as the above preprocessing steps are important, the actual sequence in which they are performed is also important while cleaning up the tweets. For example, removing the punctuation before replacing the urls means the regex expression cannot find the urls. Same with mentions or hashtags. So in our data preprocessing step, we made sure that the logical sequence of cleaning was followed. The final count of the tweets from Colorado for 2019–2021 after the data cleaning was 25,597.

2.3. Sentiment Analysis

The goal of sentiment classification is to predict the general sentiment orientation conveyed by a user in a review, blog post or editorial. Such automated classification is generally conducted by two main approaches, a machine learning approach based on supervised learning of an annotated dataset, and a lexicon (or symbolic) approach which is based on lexicons and rules. Supervised machine learning techniques (such as KNN, Naive Bayes or SVM) use a manually annotated training dataset made up of samples which labelled as positive or negative with respect to the target event (i.e., the problem). Since these systems are trained on in-domain data, they do not scale well across different domains and are not easily generalized. For example, let us consider a supervised dataset consisting of labelled IMDB movie reviews, in which sentiments are labelled against the reviews which were written by expert people in movie industry using different words (dictionary) to explain their expert and lengthy opinions regarding movies. However in our case, the tweets are usually written by non-experts and generally short, usually 1–2 sentences (less descriptive) each. Thus, if we want to use the labelled movie dataset as our training data then it would not generalize well in case of tweets related to drought, which is not in the same domain as the opinions on movies. Another drawback of the above approach was that the training dataset needed to be sufficiently large and representative. To our knowledge, there is no drought related labelled dataset that could be used for our study. Towards that end, efforts have been made to develop techniques that rely less on domain knowledge. Such techniques include discourse analysis and lexicon analysis which takes into consideration several properties of natural language [47].

To associate sentiment orientation with the context of the words we use Opinion lexicons. The idea is that each word in a sentence is treated in a way such that it holds critical opinion information and therefore provide clues to document sentiment and subjectivity. For that purpose, we used SentiWordNet introduced in [47,48], which provides a readily available database of terms and semantic (synonyms, antonyms, preposition) relationships built to assist in the field of opinion mining. Its aim was to provide term level information on opinion polarity that had been built using a semi-automated process to derive the opinion information by using the WordNet database [49], and no prior training data is required. SentiWordNet is thus a lexical resource where every word on Wordnet is related to three numerical scores, namely Pos(s): a positivity score Neg(s): a negativity score Obj(s): an objectivity (neutrality) score. The scores are very precise, pertaining to the word itself along with its context. For each term in the WordNet database, a corresponding polarity score ranging from 0 to 1 is present in SentiWordNet. Each set of terms sharing the same meaning, a.k.a synsets, is associated with three numerical scores each ranging from 0 to 1, and the corresponding value indicates the synset’s objectiveness, positive and negative bias. In SentiWordNet it is possible for a term to have non-zero values for both positive and negative scores. A higher score carries a heavy opinion bias, or is highly subjective, and a lower score indicates a term is less subjective.

2.4. Machine Learning Methods

We chose 3 different machine learning techniques, generalized linear model, support vector machines and deep learning to analyze the data. Generalized linear models (GLM) are built on top of traditional linear models by maximizing the log-likelihood and also involves parameter regularization. GLMs are particularly useful when the models have a limited number of predictors with non-zero coefficients. The model fitting is comparatively faster than traditional linear models as the computations happen in parallel. A support vector machine (SVM) is defined as a technique that constructs a hyperplane or set of hyperplanes in a high or infinite dimensional space, which can be used for classification, regression, or other tasks. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the nearest training data points of any class (so-called functional margin), since in general the larger the margin the lower the generalization error. This SVM learning method can be used for both regression and classification and provides a fast algorithm and good results for many learning tasks. Parameters to suit the problem were selected by defining them in terms of a kernel function $K(x,y)$. In short, SVM finds an adequate function that partitions the solution space to separate the training data points according to the class labels being predicted, under the assumption that future prediction follows the same pattern.

Lastly, we adopted the deep learning (DL) model. The selected DL method is based on a multi-layer feed-forward artificial neural network and the training process is optimized using stochastic gradient descent using the back propagation step. The DL network can contain many hidden layers consisting of neurons with tanh, rectifier and maxout activation functions. The operator starts a 1-node local cluster and runs the algorithm on it. It uses the default number of threads for the system with one instance of the cluster as defined in RapidMiner. Details are referred to https://docs.rapidminer.com/latest/studio/operators/modeling/predictive/neural_nets/deep_learning.html (accessed on 2 July 2022). The selected Deep Learning operator is used to predict the “survived attribute” of the Twitter and GRACE datasets. Since the label is binominal, classification is first performed to check the quality of the model, while the Split Validation operator is used to generate the training and testing datasets. Our Deep Learning operator uses the default parameters in RapidMiner. This means that two hidden layers, each with 50 neurons was constructed. The operator then calculates the Accuracy metric for diagnostics.

2.4.1. Training Data and Control Run

To quantitatively evaluate the impact of social media data on the depiction of drought, we needed a baseline model to compare with. So we first created a regression model that only included meteorological variables gathered from the GRACE satellites as the independent variables, along with PDSI values as the dependent variable. We extracted shallow groundwater, root zone soil moisture and surface soil moisture from GRACE. There are two goals associated with this control run. First is to show what percentage of the observed variation in the PDSI values can be explained by the variation in the meteorological variables based on GRACE observations. Second, the obtained percentage would make the control run the benchmark for our social media-based models to compare against. We should emphasize that evaluating the correspondence between PDSI and GRACE measurements is not the focus of this paper.

We used the weekly GRACE data and PDSI data during the observation period in the training dataset, between 1 January 2019 to 31 December 2020, resulting in 104 data points. During the model training step, we separated out 40% of the data for testing purpose in order to calculate the performance of models, which means that we used 41 of the 63 data points and for testing. We also performed a 10-fold cross validation to remove bias in the models being trained. We applied additional “lags” of these variables from 1 to 3 weeks before the current week’s values in order to capture the progression of drought in terms of the weekly PDSI values. In total we had 12 independent variables as ‘features’ for building the baseline model with weekly PDSI as the dependent variable. The 12 independent variables

are as follows: groundwater, groundwater_1_week_before, groundwater_2_week_before, groundwater_3_week_before, root_zone_soil, root_zone_soil_1_week_before, root_zone_soil_2_week_before, surface_soil, root_zone_soil_3_week_before, surface_soil_1_week_before, surface_soil_2_week_before, surface_soil_3_week_before.

Model 1 from Table 3 shows the baseline control run model, which is the simplest model generated using only the meteorological variables. Figures 1–3 shows the results on test dataset for the generalized linear model, support vector machine and deep learning models, respectively. For each of those figures, the x-axis represents the true or the actual PDSI values and the y-axis represents the predicted PDSI values.

Table 3. Different combinations of Twitter data when added as individual features to the control run ‘D’. ‘P’ represents the count of positive tweets. ‘N’ represents the count of negative tweets. ‘xN’ & ‘xP’ represent the count of negative and positive tweets ‘x’ week before, respectively.

Model	Configuration
Model-1	Drought Indices (D)-control run
Model-2	D + N
Model-3	D + P + N
Model-4	D + P
Model-5	D + N + 1N
Model-6	D + P + 1P + N + 1N
Model-7	D + P + 1P
Model-8	D + N + 1N + 2N
Model-9	D + P + 1P + 2P
Model-10	D + P + 1P + 2P + N + 1N + 2N

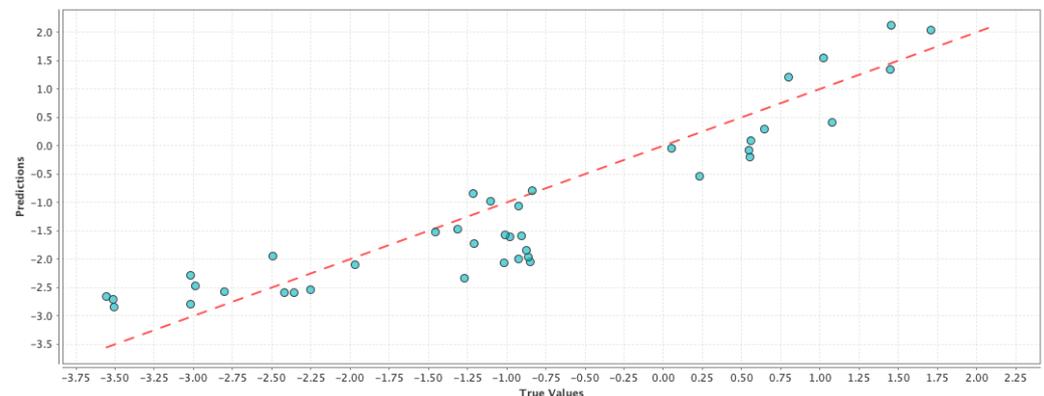


Figure 1. Result on the Test Set For Generalized Linear Model. The x-axis represents the true or the actual PDSI values and the y-axis represents the predicted PDSI values. Each small blue circle in the figures is represented by the tuple $(PDSI_{actual}, PDSI_{pred})$. The red dashed line is the reference line to test the model performance.

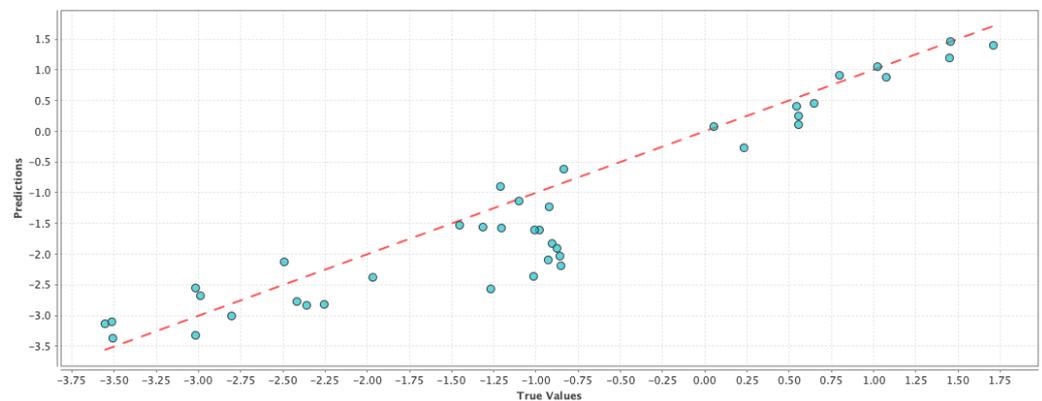


Figure 2. Same as Figure 1 but for Support Vector Machine model.

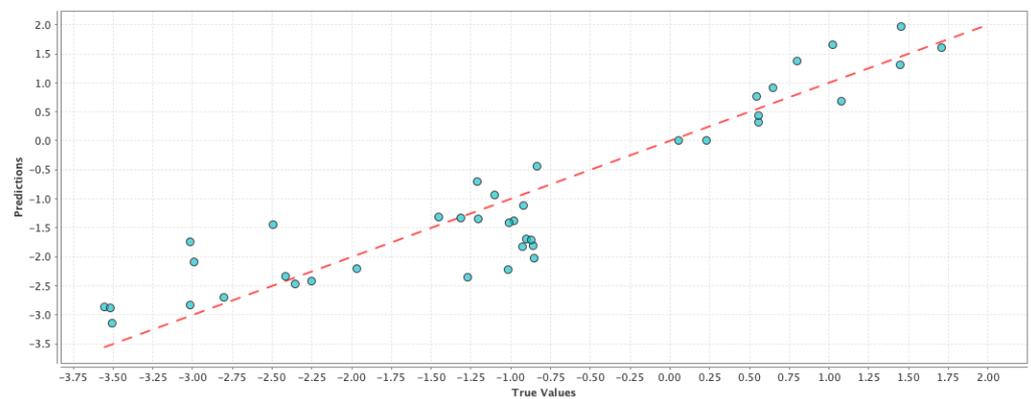


Figure 3. Same as Figure 1 but for Deep Learning model.

Each small blue circle in the figures is represented by the tuple $(PDSI_{actual}, PDSI_{pred})$, where $PDSI_{pred}$ are the PDSI values predicted by the individual machine learning models and $PDSI_{actual}$ are the actual PDSI values. The red dashed line is the reference line to test the model performance and each point on the line is represented as $(PDSI_{actual}, PDSI_{actual})$, thus the closer the blue circles are to the red line the smaller the error, and the better is the model performance. Root Mean Squared Error (RMSE) and correlation coefficient were chosen as the performance metrics and Table 4 shows the comparison of performance between the different models. From Table 4, we can see that the RMSE values are almost similar for all the cases and hence it is not possible to decide on a good model just based on RMSE values in this case. On the other hand, the correlation values in Table 4 indicate that there is a high correlation between predicted values and the actual PDSI values. From the above results, we can say that the simple model generated using only meteorological variables can be used as our baseline control run for our social media-based models to compare against. This ‘drought’ control run henceforth is denoted by ‘D’.

Table 4. Results on Test Set for Generalized Linear Model, Support Vector Machines and Deep Learning.

Models	RMSE	RMSE Std Dev	Correlation	Corr Std Dev
Generalized Linear Model	0.603	±0.116	0.936	±0.015
Support Vector Machines	0.584	±0.1	0.955	±0.006
Deep Learning	0.604	±0.107	0.922	±0.051

2.4.2. Twitter Models

Before diving deep into building social media-based models, we wanted to examine what words people commonly use in their tweets while referring to drought. Table 1 refers to the top used words in 2019 and 2020 depending upon their usage frequency in the tweets. Next, we proceeded with the sentiment analysis by using the 'sentlex' python library (Available at <https://github.com/bohana/sentlex>) to generate the polarity score of the relevant tweets. Upon initiation, the python library reads the SentiWordNet v3.0 language resource into memory and compiles word frequency data based on the frequency distribution of lexicon words in NLTK's Brown corpus. When we pass a sentence to the library, it first tokenizes or breaks the sentences and tags the relevant part of speech words (adjective, verb, noun and adverb) and assigns a tuple of numeric values (positive, negative) indicating word polarity known to the lexicon. It should be noted that when similar words in a sentence carry multiple meanings, then the opinion of each of those words is averaged out to obtain the output tuple of numeric values (positive, negative). In other words, the 'sentlex' library does not perform word sense disambiguation; rather, it just separates the words by part of speech. In the final step, we compare the positive and negative values of the tuple and then assign to a sentence, the label (positive or negative) which has the highest value in the tuple. Table 5 shows some sample tweets with the corresponding sentiment categories after using the 'sentlex' sentiment analysis library. We note that Twitter users may fall into a certain age distribution (e.g., gender, geography, geo-location, and education level), so collecting only Twitter data could bias the result with respect to user groups.

2.4.3. Twitter–Data Model

Next, we added Twitter data to our control run and examined the changes in explanatory power. Our goal was to see whether addition of social media data resulted in a performance improvement over the control run model. The first step was to classify each tweet as positive or negative by using the aforementioned sentiment analysis model. The following step was to generate the counts of positive and negative tweets related to drought per week.

Table 3 shows the different combinations of Twitter data when added as individual features to the control run 'D'. In Table 3, 'P' represents the count of positive tweets. 'N' represents the count of negative tweets. 'wN' and 'wP' represent the count of negative and positive tweets 'w' weeks before, respectively. For example, '1P' represents the count of positive tweets 1 week before and '2N' represents the count of negative tweets 2 weeks before. In the next two sections we will be analyzing the results from two of the Twitter-based models (Model 6 and Model 10 from Table 3).

We carried out our experiments about building a model able to predict PDSI values, using RapidMiner (<https://rapidminer.com/>, accessed on 2 July 2021) which is an integrated environment that enables efficient prototyping for machine learning applications that can be used across various domains. RapidMiner has a wide selection of machine learning models per the needs of the task in hand. Their software automatically optimizes and chooses the best weight values for individual models depending on the dataset.

Table 5. Sample positive and negative tweets referring to drought.

Original Tweet	Processed & Formatted Tweet	Category
With serious drought conditions likely to continue into 2021, the state of Colorado has activated the municipal portion of its emergency drought plan for the second time in history. https://t.co/GqD6l5CDOI (accessed on 9 September 2020)	with serious drought conditions likely to continue into 2021 the state of colorado has activated the municipal portion of its emergency drought plan for the second time in history <url>	Negative
Much of the Colorado River basin is enveloped in extreme or exceptional drought. The drought is making for dry soil, which will mean runoff in the spring is less effective as water seeps into the soil. https://t.co/CVDFXFdihn (accessed on 9 September 2020)	much of the colorado river basin is enveloped in extreme or exceptional drought the drought is making for dry soil which will mean runoff in the spring is less effective as water seeps into the soil <url>	Negative
#CORiver drought plans have helped, but key reservoirs are at historic lows, and more work is needed to protect the river, and keep water flowing to the millions that rely on its water, per a new report. https://t.co/PDS0clALoM (accessed on 9 September 2020)	coriver drought plans have helped but key reservoirs are at historic lows and more work is needed to protect the river and keep water flowing to the millions that rely on its water per a new report <url>	Negative
@JoshClarkDavis We are about to get hit 3 times with storms in the next 5 days. With the drought we have been in, it's very exciting. Getting our snow pack just to average would be big.	we are about to get hit 3 times with storms in the next 5 days with the drought we have been in it is very exciting getting our snow pack just to average would be big	Positive
All of Colorado is now in some level of drought, but fortunately Northglenn has sufficient water in storage for this time of year. Water conservation by every one of our residents will make a huge difference by helping to stretch our water reserves!	all of colorado is now in some level of drought but fortunately northglenn has sufficient water in storage for this time of year water conservation by every one of our residents will make a huge difference by helping to stretch our water reserves	Positive
The snowpack that feeds the Colorado River stands at 75% of the median for this time of year—and that line has dropped rapidly in just a few days. The soil in the watershed, baked dry over the past year, is starting to absorb melting snow like a sponge. https://t.co/GAaPK2wqii (accessed on 15 December 2020)	the snowpack that feeds the colorado river stands at 75 of the median for this time of year and that line has dropped rapidly in just a few days the soil in the watershed baked dry over the past year is starting to absorb melting snow like a sponge <url>	Positive

3. Results

We used the aforementioned machine learning techniques of generalized linear model, support vector machine and deep learning to train all of our Twitter-based models. The results are presented and discussed herein.

Model-6 D + P + N + 1P + 1N: The rationale behind this model was to capture the effect of people's opinion on Twitter from the current week in observation along with their opinion from the previous week, and how the combined change in user perception regarding drought was reflected on the change in PDSI values. In machine learning terms, our goal was to see if a high percentage of the observed variation in the PDSI values can be explained by the variation in the features. The results comparing the RMSE and correlation coefficient of the test results are shown in Table 6. If we compare Tables 4 and 6, we can see that the correlation values have improved for the generalized linear model and deep learning techniques, while the RMSE values have improved for the generalized linear model and support vector machine techniques. We can see a similar effect in Figures 4–6 where the blue circles are closer to the reference red line as compared to the ones in the case of the control run model. Although the improvement in performance is not significant, we can see that adding social media data as features gave more prediction power to the control run 'D'.

Table 6. Results on Test Set for Generalized Linear Model, Support Vector Machines and Deep Learning for Model 6: D + P + N + 1P + 1N.

Models	RMSE	RMSE Std Dev	Correlation	Corr Std Dev
Generalized Linear Model	0.578	±0.066	0.941	±0.043
Support Vector Machines	0.492	±0.123	0.948	±0.043
Deep Learning	0.624	±0.135	0.93	±0.026

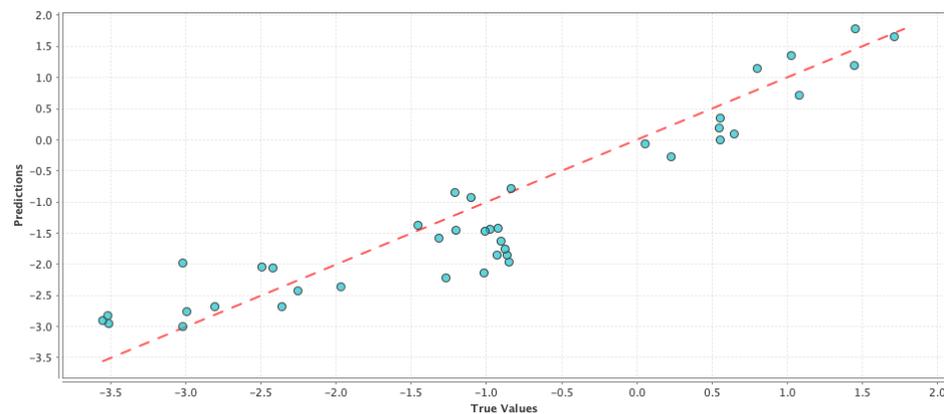


Figure 4. Result on Test Set For Generalized Linear Model using Model 6 from Table 3. The x-axis represents the true or the actual PDSI values and the y-axis represents the predicted PDSI values. Each small blue circle in the figures is represented by the tuple $(PDSI_{actual}, PDSI_{pred})$. The red dashed line is the reference line to test the model performance.

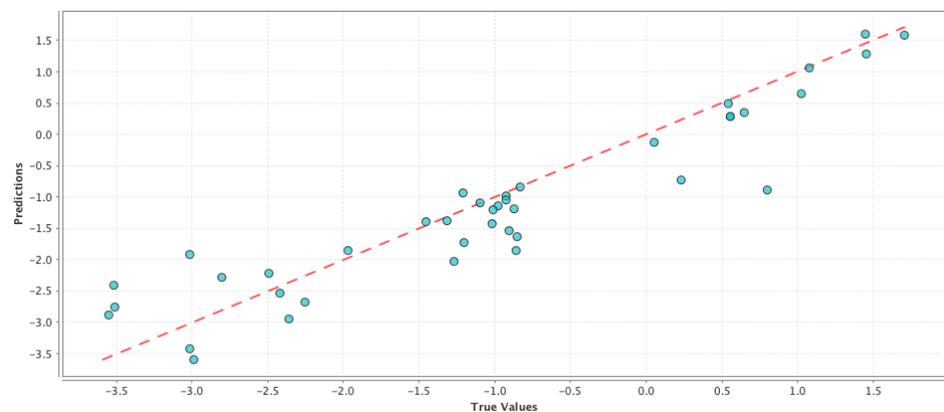


Figure 5. Result on Test Set For Support Vector Machine using Model 6 from Table 3. The x-axis represents the true or the actual PDSI values and the y-axis represents the predicted PDSI values. Each small blue circle in the figures is represented by the tuple $(PDSI_{actual}, PDSI_{pred})$. The red dashed line is the reference line to test the model performance.

Model-10 D + P + N + 1P + 1N + 2P + 2N: Similar to the previous model, the goal was to capture the change in the user perception regarding drought over a “two-week period”. Similar to the previous model, our goal was to see if a high percentage of the observed variation in the PDSI values can be explained by the variation in the features. The results comparing the RMSE and correlation coefficient of the test results are shown in Table 7. If we compare Tables 4 and 7, we can see that the correlation values with the inclusion of Twitter data have improved for all three cases and correspondingly the RMSE values have also improved for all three machine learning techniques. These effects are shown in Figures 7–9 where the blue circles are closer to the reference red line as compared

to the control run model. If we compare the results in Tables 6 and 7, we can see that there is a slight improvement in the performance due to the addition of an extra week of user perception from Twitter. Although the above performance improvements are not astonishing, we can still see that adding social media data as features resulted in a better prediction performance than simply using GRACE data alone.

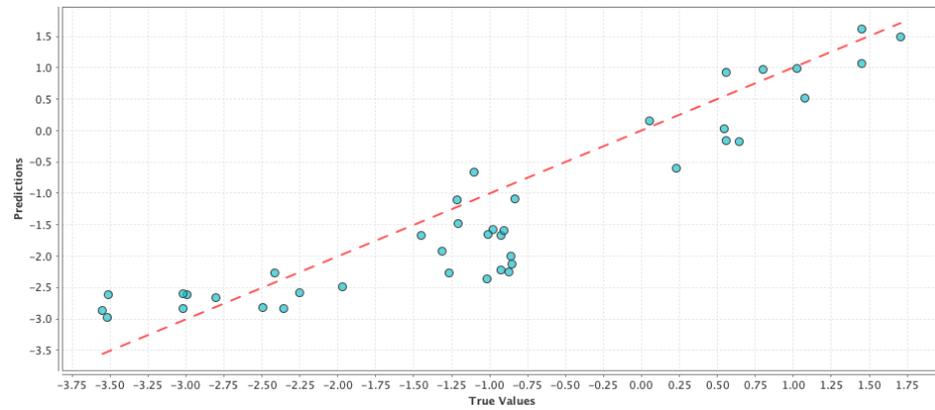


Figure 6. Result on Test Set For Deep Learning using Model 6 from Table 3. The x-axis represents the true or the actual PDSI values and the y-axis represents the predicted PDSI values. Each small blue circle in the figures is represented by the tuple $(PDSI_{actual}, PDSI_{pred})$. The red dashed line is the reference line to test the model performance.

Table 7. Results on Test Set for Generalized Linear Model, Support Vector Machines and Deep Learning for Model 10: D + P + 1P + 2P + N + 1N + 2N.

Models	RMSE	RMSE Std Dev	Correlation	Corr Std Dev
Generalized Linear Model	0.516	±0.088	0.96	±0.017
Support Vector Machines	0.497	±0.0112	0.956	±0.042
Deep Learning	0.554	±0.051	0.941	±0.045

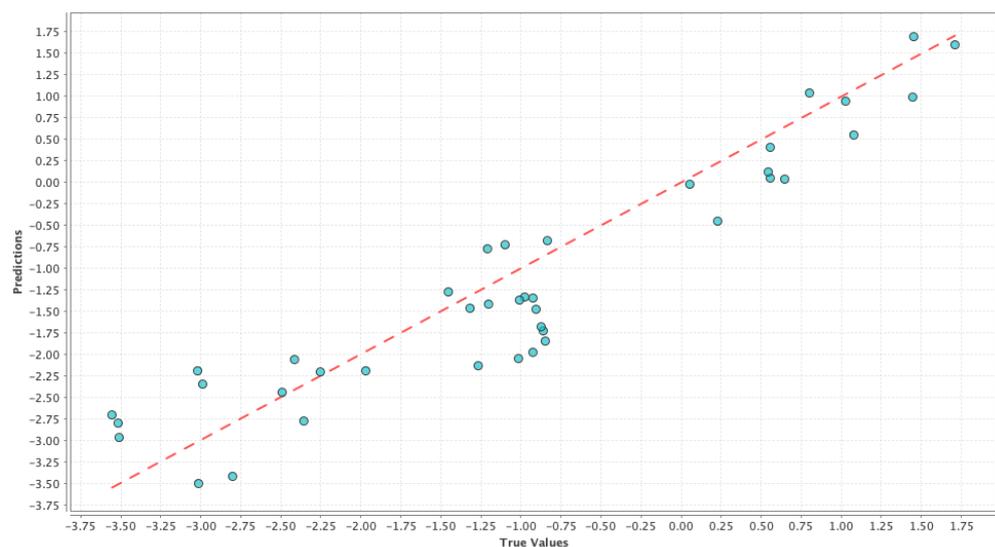


Figure 7. Result on Test Set For Generalized Linear Model using Model 10 from Table 3. The x-axis represents the true or the actual PDSI values and the y-axis represents the predicted PDSI values. Each small blue circle in the figures is represented by the tuple $(PDSI_{actual}, PDSI_{pred})$. The red dashed line is the reference line to test the model performance.

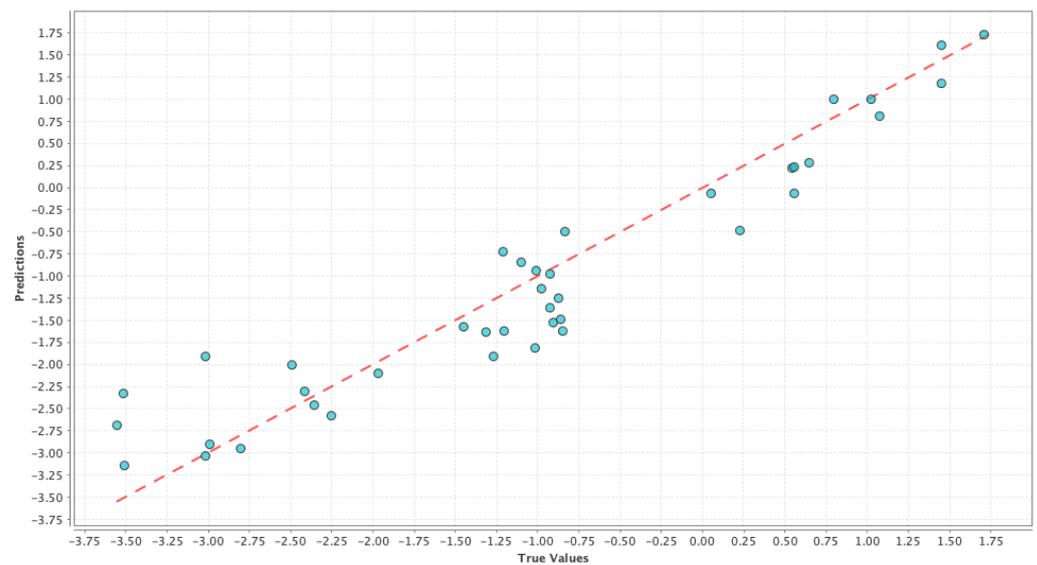


Figure 8. Result on Test Set For Support Vector Machine using Model 10 from Table 3. The x-axis represents the true or the actual PDSI values and the y-axis represents the predicted PDSI values. Each small blue circle in the figures is represented by the tuple $(PDSI_{actual}, PDSI_{pred})$. The red dashed line is the reference line to test the model performance.

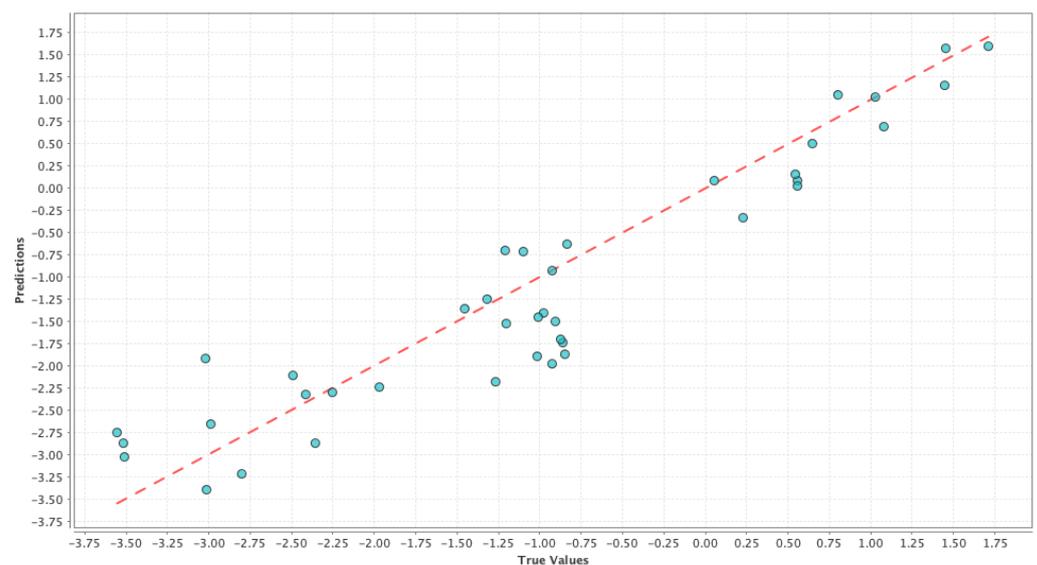


Figure 9. Result on Test Set For Deep Learning using Model 10 from Table 3. The x-axis represents the true or the actual PDSI values and the y-axis represents the predicted PDSI values. Each small blue circle in the figures is represented by the tuple $(PDSI_{actual}, PDSI_{pred})$. The red dashed line is the reference line to test the model performance.

Performance of Other Twitter Based Models: In this section, we will discuss the performance of all the models listed in Table 3. Figures 10–12 shows the performance comparison of individual Twitter-based models (Models 2–10 in Table 3) over the control run (Model 1 in Table 3). The x-axis in Figures 10–12 represent the difference between the RMSE values for the control run and the Twitter-based models. The y-axis represents the difference between the correlation coefficient values between the control run and the Twitter-based models. The coordinate position where the blue lines intersect is where the performance (in terms of RMSE and correlation coefficient) of the Twitter-based models and the control run are the same. In order to classify a model to be a better performer in comparison to the control run, its correlation value need to be higher and the corresponding

RMSE value needs to be lower. Thus, a better performing model would appear in the top left quadrant in Figures 10–12. If models fall in the top right quadrant, then the performance improvement only exists in terms of the correlation coefficient. If the models fall in the bottom left quadrant then the performance improvement is only for RMSE. Finally, if the models fall in the bottom right quadrant then there is no performance improvement. From Figures 10–12, we can see that, except for two cases (Model 2 and 5), the Twitter-based models have shown a better performance over the control run model.

We also constructed Figure 13, which shows the overall performance improvements of the two best-performing models, Model 6 and Model 10 in comparison to the baseline control run model. Each solid circle in Figure 13 represent either Model 6 or Model 10, which is a Twitter-based model (T_RMSE_{model} , $T_R^2_{model}$), where $model$ denotes either deep learning, support vector machine or generalized linear model. Each hollow circle in Figure 13 represents the control run model (C_RMSE_{model} , $C_R^2_{model}$) for either deep learning, support vector machine or generalized linear model. Arrows are drawn between the control run and Twitter-based models to visually delineate the performance improvements between the control run and the Twitter-based models (Model 6 and Model 10). A better performing model will have a lower RMSE value and a higher correlation coefficient (R^2) and the bottom right corner is the area which has the lowest RMSE and highest correlation values. Thus, any model for which the arrow points towards the bottom right means that the particular Twitter-based model has gained improvement over the control run. In Figure 13, we can see that the generalized linear model of both Model 6 and Model 10 has a marked improvement over Model 1. We can also see that the performance improvement is greater in case of Model 10, which contains user perceptions from the current time period as well as from one and two weeks before. In the case of support vector machine (SVM), we see that there is no significant performance improvement for both models. In terms of the deep learning model, a performance improvement is evident in Model 10 but not in Model 6, which has an increased RMSE value. From these results, we can say that Model 10 was the overall better performer and showed quantifiable improvement over the control run model.

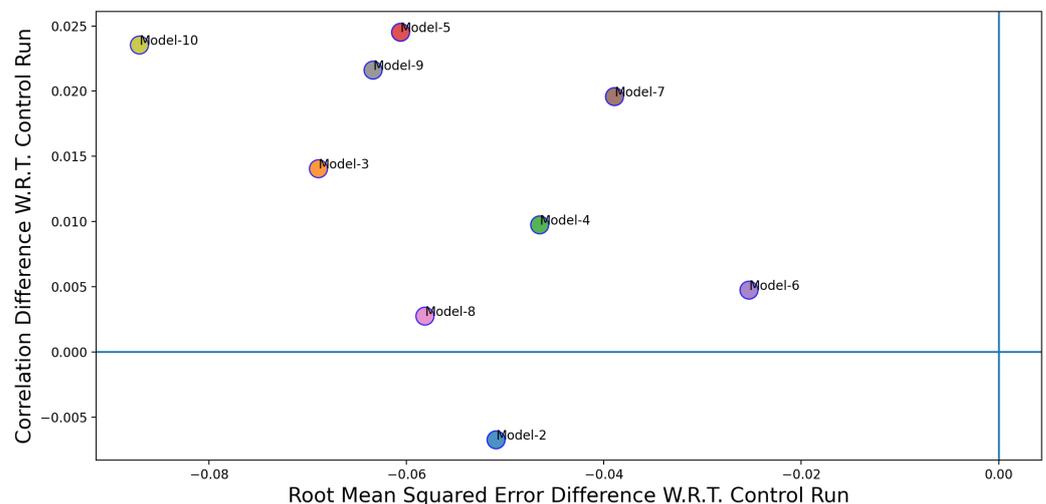


Figure 10. Improvement over control run for Generalized Linear Model. The x-axis represents the difference between the RMSE values for the control run and the Twitter-based models. The y-axis represents the difference between the correlation coefficient values between the control run and the Twitter based models.

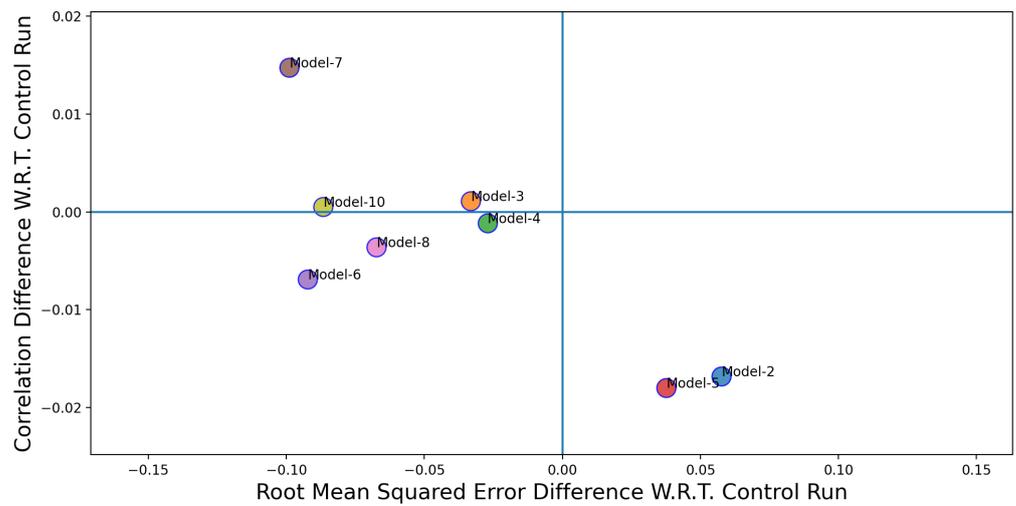


Figure 11. Improvement over control run for Support Vector Machine. The x-axis represents the difference between the RMSE values for the control run and the Twitter-based models. The y-axis represents the difference between the correlation coefficient values between the control run and the Twitter based models.

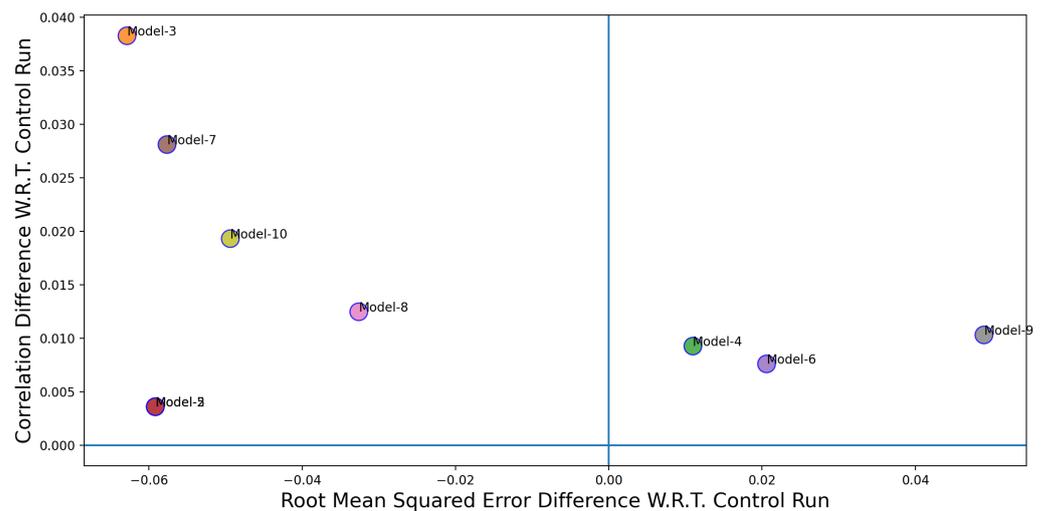


Figure 12. Improvement over control run for Deep Learning. The x-axis represents the difference between the RMSE values for the control run and the Twitter-based models. The y-axis represents the difference between the correlation coefficient values between the control run and the Twitter based models.

The following section summarizes the performance of our trained models on new and unseen data of 2021. We performed this analysis by first collecting Twitter data from Colorado during January–April 2021 (17 weeks of data). After cleaning the data and removing tweets based on the keywords and search terms in Table 2, we were able to retain 4960 tweets. In a similar way as described in the previous sections, we applied the sentiment analysis model to the tweets and gathered the count of the number of positive and negative tweets per week for evaluation time period. We also collected the necessary GRACE and PDSI data for the same time period, while we ran this analysis for all the models listed in Table 3, for brevity we will only discuss the performance result for the “best performing model”, i.e., Model 10 (D + P + N + 1P + 1N + 2P + 2N).

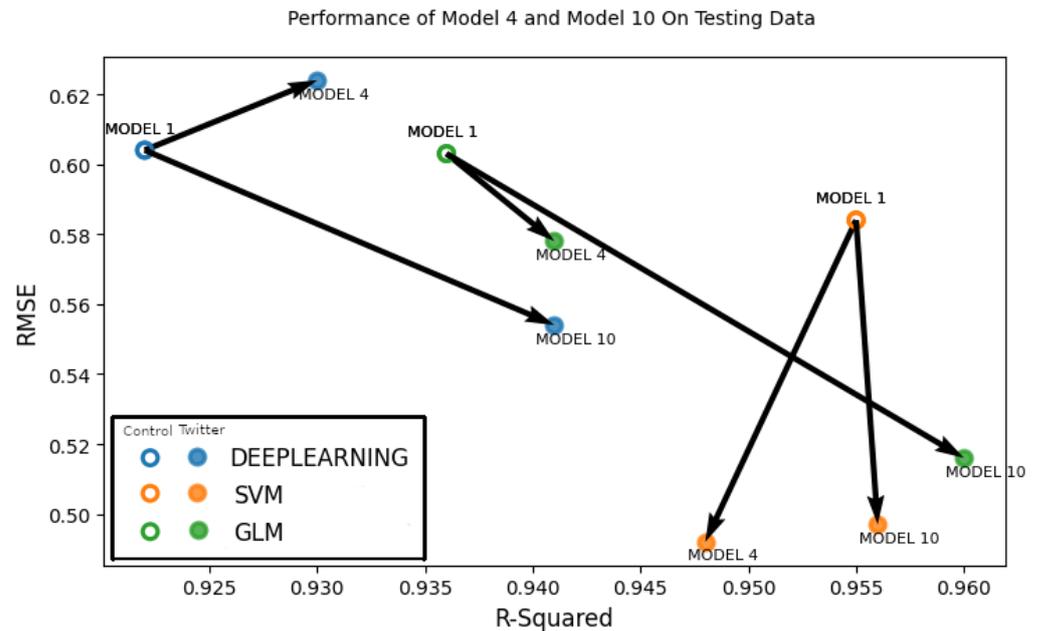


Figure 13. Performance comparison for Model 4 and Model 10 w.r.t. Model 1, which is the baseline control run, on the Test Dataset. The x-axis represents the correlation coefficient (R^2 , R-Squared) and the y-axis represents the RMSE values. The arrows indicate performance improvements between the control run and the Twitter-based models. Each solid circle represents a Twitter-based model. Each hollow circle represents the control run model.

Figures 14–16 show the prediction results for Model 10 when compared to the control run model. In each figure, x-axis represents the actual PDSI values and y-axis represents the predicted PDSI. The red circles are the predictions by Model 10 represented by tuple $(PDSI_{actual}, PDSI_{t_pred})$, where $PDSI_{t_pred}$ are the PDSI values predicted by Model 10 and $PDSI_{actual}$ are the actual PDSI values. Similarly, the green hollow circles in the figures are the predictions by the control run model and is represented by the tuple $(PDSI_{actual}, PDSI_{c_pred})$. The blue dashed line is the reference line to test the model performance and each point on the line is represented as $(PDSI_{actual}, PDSI_{actual})$, thus the closer the green and red circles are to the blue line, the smaller the error and the better is the model performance. From Figures 14–16, we can see that the red circles (Twitter-based model) are much closer to the blue line than the green circles (control run model). From these results we can conclude that the Twitter-based model consistently outperforms the control run in predicting the PDSI.

In Figure 17, the x-axis represents the R^2 values and the y-axis represents RMSE. Arrows are drawn between the circles representing control run and Twitter-based model for the deep learning, support vector machine and generalized linear model techniques, respectively. The arrows determine any performance improvement between the control run and the Twitter-based model (Model 10). The fact that all arrows are pointing to the bottom right, which indicates lower RMSE and higher R^2 , means that Model 10 indeed outperforms the control run model.

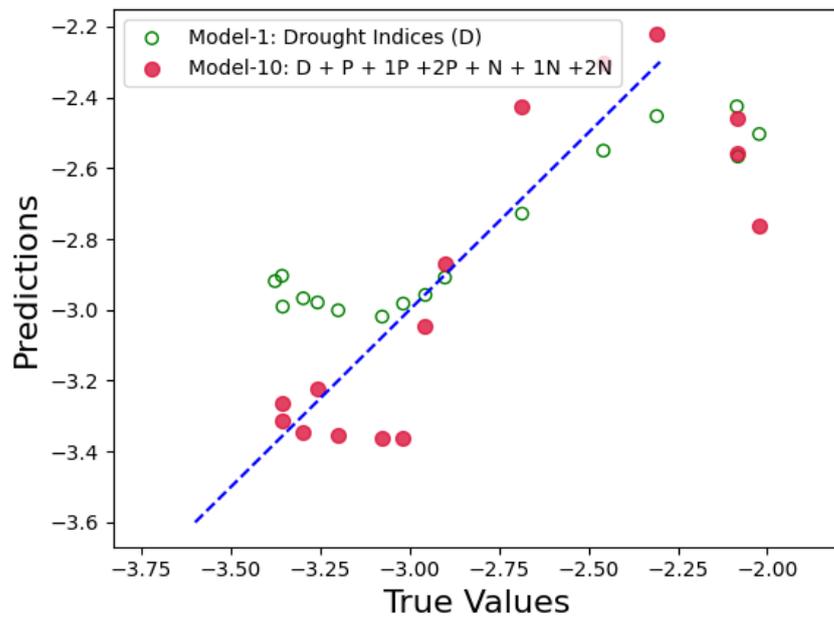


Figure 14. Result on unseen data during January–April 2021 using Generalized Linear Model and Model 10 from Table 3. The x-axis represents the true or the actual PDSI values and the y-axis represents the predicted PDSI values. The red circles are the predictions by Model 10 and the green hollow circles in the figures are the predictions by the control run model. The blue dashed line is the reference line to test the model performance.

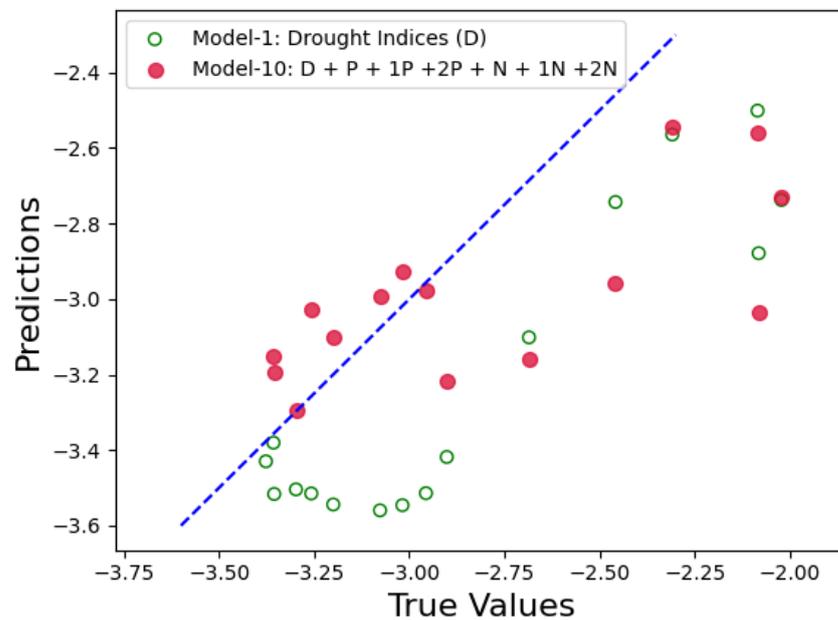


Figure 15. Result on unseen data during January–April 2021 using Support Vector Machine and Model 10 from Table 3. The x-axis represents the true or the actual PDSI values and the y-axis represents the predicted PDSI values. The red circles are the predictions by Model 10 and the green hollow circles in the figures are the predictions by the control run model. The blue dashed line is the reference line to test the model performance.

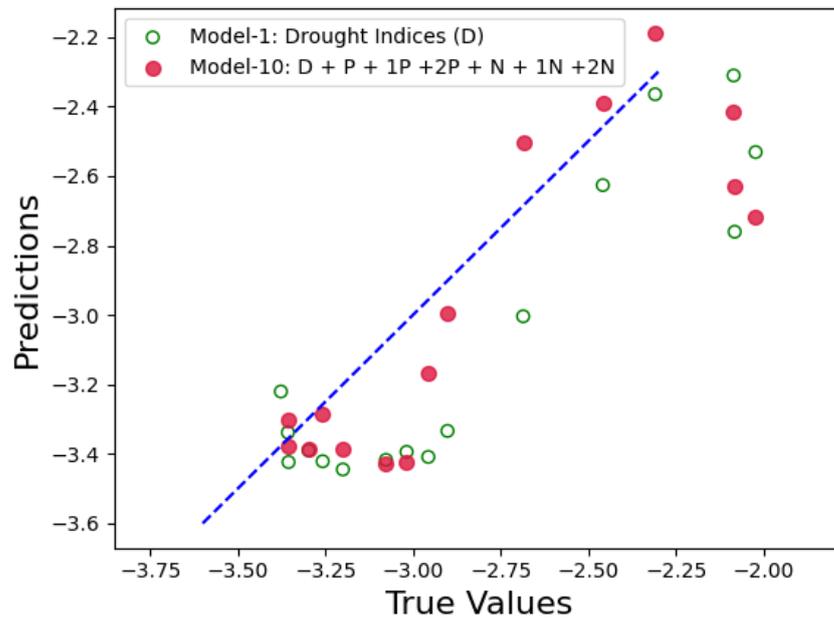


Figure 16. Result on unseen data during January–April 2021 using Deep Learning and Model 10 from Table 3. The x-axis represents the true or the actual PDSI values and the y-axis represents the predicted PDSI values. The red circles are the predictions by Model 10 and the green hollow circles in the figures are the predictions by the control run model. The blue dashed line is the reference line to test the model performance.

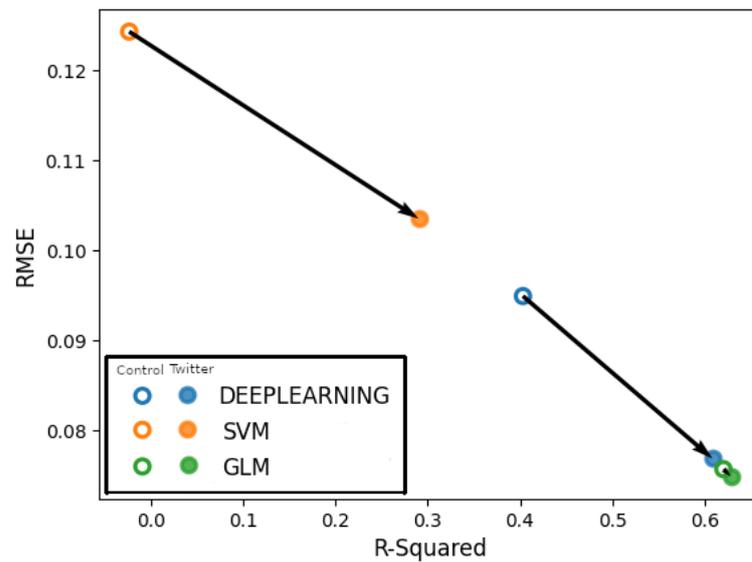


Figure 17. Overall performance comparison of Model 10 w.r.t control run model on data from January–April 2021.

Lastly, there could be a point of discussion about the possible effects of the release of the new drought.gov site released in January 2021. Although there is a possibility that recently initiated drought chatters might affect the tweet volume, it should be noted that we did not use the 2021 data to train our models. Furthermore, our models only tell us if it was a positive or negative tweet, so if people have retweeted the NIDIS tweets then our system was able to capture that engagement.

4. Conclusions

In this study, we tested the feasibility of developing social-media-based models, supplementary to meteorological predictors, to anticipate PDSI and drought in Colorado. The starting point was to build the control run model; this was trained using weekly hydrologic and PDSI data during the observation period (1 January 2019 to 31 December 2020). Shallow groundwater, root zone soil moisture and surface soil moisture was extracted from GRACE data and “lags” of 1–3 weeks for each of the variables were applied prior to the current week’s values; this in order to capture the progression of drought. As noted previously 12 variables were utilized to build a regression model where PDSI was the dependent variable. Subsequently, a control run model was constructed by using three different machine learning techniques, i.e., a generalized linear model, support vector machines, and deep learning. The meteorological control run model was treated as the baseline in the evaluation of the impact of including social media data in the machine learning models.

Next, by using Twitter as the social media-based platform, tweets were collected based on keywords which were closely related to drought. It was found that there were user discussions were varied regarding drought during the time period of study. Throughout the analysis of the frequency of different words used, we observed a change in the user perception of drought as it continued to worsen over the 2019–2020 period. Furthermore, noteworthy was that a considerable number of tweets in the period of the analysis included links to government and academic websites serving as sources of information about drought conditions; this observation supports the notion that people used Twitter not only to complain about degrading drought conditions, but also as a source of information with respect to drought. Next, by generating the polarity score of the tweets as either positive or negative, we added the different combinations of scores to the control run model. The different combinations delineated here served as Twitter-based models.

To conclude, the results indicate that Twitter-based machine learning model can improve the forecast of PDSI in times of a developing drought. Of the 10 models tested, eight showed quantifiable improvements in the performance over the control run model in terms of RMSE and correlation coefficient. This is supportive of the hypothesis that including social media data adds value to the PDSI depiction. Such an improvement is further reinforced by testing the control run and Twitter-based models on previously unseen data during January–April 2021. We found that the Twitter-based model consistently outperformed the control run in predicting the PDSI values as the drought worsened.

Machine learning (ML) models have gained prominence in the area of drought forecasting. Researchers have developed models using time series analysis [17,18], neural networks [19–24], fuzzy inference systems [25], support vector regression [26–28] and different ensemble techniques [29,30] to detect and forecast drought. Despite the new insights machine learning can provide, recent studies suggest that any single indicator alone is not enough to explain the complexity and diversity of drought [31,32].

ML models also thrive on good quality data. One of the important improvements lies in data collection. Thus, increasing the training dataset in terms of both Twitter and meteorological data can lead to an improvement of ML models. A further improvement would be in the sentiment analysis step, like how to reduce vulnerability to drought through an improved understanding of the interactions between society and physical processes, while there is a growing appreciation that the information landscape for citizens’ participation in crisis response and recovery activities is improving, it is not clear what the role of social media will be in slow developing threats as the case for drought. An improved language model could be deployed that could capture the word(s) embedded in tweets that depict how people tweet with respect to drought over different periods of time.

In this work, we focused on Colorado from where we were able to collect good quality Twitter data regarding drought. Since drought has been worsening over the American West, we believe that our approach could be used to study drought progression across

other states such as Utah, Arizona, California and Nevada. Thus, another future goal is to study how well the Twitter-based models perform across different states.

Author Contributions: Conceptualization, S.M. and S.W.; methodology, S.M. and S.W.; software, S.M.; validation, S.M., S.W. and J.L.; formal analysis, S.M.; investigation, S.M.; resources, S.W., R.G. and J.L.; data curation, S.M.; writing—original draft preparation, S.M.; writing—review and editing, S.M., S.W., D.H., J.L. and R.G.; visualization, S.M., S.W. and J.L.; supervision, S.W.; project administration, S.W. and J.L.; funding acquisition, S.W. All authors have read and agreed to the published version of the manuscript.

Funding: This project was funded by National Integrated Drought Information System (NIDIS) through NCAR/UCAR as a sub-award of NA18OAR4310253B. Simon Wang is partially supported by U.S. Department of Energy, Office of Biological and Environmental Research program under Award Number DE-SC0016605. This publication is also supported by the Utah Agricultural Experiment Station as paper #9607.

Institutional Review Board Statement: USU IRB Approval Code: 12855. Approval Date: 6 June 2022.

Informed Consent Statement: Written informed consent has been obtained from the patient(s) to publish this paper.

Data Availability Statement: Data are available upon request.

Acknowledgments: The authors appreciate the publication support by UCAR's CPAESS.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Fonseca, F.; Metz, S. Lake Mead Drops to a Record Low Amid Drought. 2021. Available online: <https://www.latimes.com/world-nation/story/2021-06-11/lake-mead-key-reservoir-colorado-river-record-low-drought#:~:text=Lake%20Mead%2C%20a%20key%20reservoir%20on%20the%20Colorado%20River%2C%20has,feet%20at%2011%20p.m.%20Wednesday> (accessed on 11 June 2021).
- Wang, S.Y.; Santanello, J.; Wang, H.; Barandiaran, D.; Pinker, R.; Schubert, S.; Gillies, R.; Oglesby, R.; Hilburn, K.; Kilic, A.; et al. An intensified seasonal transition in the Central U.S. that enhances summer drought. *J. Geophys. Res.* **2015**, *120*, 8804–8816. [[CrossRef](#)]
- Switzer, D.; Vedlitz, A. Investigating the Determinants and Effects of Local Drought Awareness. *Weather Clim. Soc.* **2017**, *9*, 641–657. [[CrossRef](#)]
- Shafiee-Jood, M.; Deryugina, T.; Cai, X. Modeling Users' Trust in Drought Forecasts. *Weather Clim. Soc.* **2021**, *13*, 649–664. [[CrossRef](#)]
- Smith, A.B. U.S. Billion-Dollar Weather and Climate Disasters, 1980—Present (NCEI Accession 0209268), 2020. Available online: <https://www.ncei.noaa.gov/access/billions> (accessed on 10 August 2021).
- Bolinger, B. How Drought Prone Is Your State? A Look at the Top States and Counties in Drought Over the Last Two Decades, 2019. Available online: <https://www.drought.gov/news/how-drought-prone-your-state-look-top-states-and-counties-drought-over-last-two-decades> (accessed on 10 August 2021).
- Sutton, J.; Palen, L.; Shklovski, I. Backchannels on the Front Lines: Emergent Uses of Social Media in the 2007 Southern California Wildfires. In Proceedings of the 5th International ISCRAM Conference, Washington, DC, USA, 4–7 May 2008.
- Smith, K.H.; Tyre, A.J.; Tang, Z.; Hayes, M.J.; Akyuz, F.A. Calibrating Human Attention as Indicator Monitoring #drought in the Twittersphere. *Bull. Am. Meteorol. Soc.* **2020**, *101*, E1801–E1819. [[CrossRef](#)]
- Hughes, A.; Palen, L. Twitter Adoption and Use in Mass Convergence and Emergency Events. *Int. J. Emerg. Manag.* **2009**, *6*, 248–260. [[CrossRef](#)]
- Abdelhaq, H.; Sengstock, C.; Gertz, M. EvenTweet: Online Localized Event Detection from Twitter. *Proc. VLDB Endow.* **2013**, *6*, 1326–1329. [[CrossRef](#)]
- Cameron, M.A.; Power, R.; Robinson, B.; Yin, J. Emergency Situation Awareness from Twitter for Crisis Management. In Proceedings of the 21st International Conference on World Wide Web, Lyon, France, 16–20 April 2012; WWW '12 Companion; Association for Computing Machinery: New York, NY, USA, 2012; pp. 695–698. [[CrossRef](#)]
- Mathioudakis, M.; Koudas, N. TwitterMonitor: Trend Detection over the Twitter Stream. In Proceedings of the 2010 ACM SIGMOD International Conference on Management of Data, Indianapolis, IN, USA, 6–10 June 2010; SIGMOD '10; Association for Computing Machinery: New York, NY, USA, 2010; pp. 1155–1158. [[CrossRef](#)]
- Sakaki, T.; Okazaki, M.; Matsuo, Y. Earthquake Shakes Twitter Users: Real-Time Event Detection by Social Sensors. In Proceedings of the 19th International Conference on World Wide Web, Raleigh, NC, USA, 26–30 April 2010; WWW '10; Association for Computing Machinery: New York, NY, USA, 2010; pp. 851–860. [[CrossRef](#)]

14. Svoboda, M.; LeComte, D.; Hayes, M.; Heim, R.; Gleason, K.; Angel, J.; Rippey, B.; Tinker, R.; Palecki, M.; Stooksbury, D.; et al. The drought monitor. *Bull. Am. Meteorol. Soc.* **2002**, *83*, 1181–1190. [CrossRef]
15. Glaser, M. California Wildfire Coverage by Local Media, Blogs, Twitter, Maps and More. 2007. Available online: <http://mediashift.org/2007/10/california-wildfire-coverage-by-local-media-blogs-twitter-maps-and-more298/> (accessed on 11 June 2021).
16. Stelter, B. How Social Media is Helping Houston Deal with Harvey Floods. 2017. Available online: <https://money.cnn.com/2017/08/28/media/harvey-rescues-social-media-facebook-twitter/index.html#:~:text=Hundreds%20of%20stranded%20Texas%20residents,high%20the%20flood%20waters%20were> (accessed on 11 June 2021).
17. Ömer Faruk, D. A hybrid neural network and ARIMA model for water quality time series prediction. *Eng. Appl. Artif. Intell.* **2010**, *23*, 586–594. [CrossRef]
18. Han, P.; Wang, P.; Zhang, S.; Zhu, D. Drought forecasting based on the remote sensing data using ARIMA models. *Math. Comput. Model.* **2010**, *51*, 1398–1403. [CrossRef]
19. Mishra, A.; Desai, V. Drought forecasting using feed-forward recursive neural network. *Ecol. Model.* **2006**, *198*, 127–138. [CrossRef]
20. Santos, J.; Portela, M.; Pulido-Calvo, I. Spring drought prediction based on winter NAO and global SST in Portugal. *Hydrol. Process.* **2014**, *28*, 1009–1024. [CrossRef]
21. Le, M.H.; Corzo, G.; Solomatine, D.; Nguyen, L.B. Meteorological Drought Forecasting Based on Climate Signals Using Artificial Neural Network—A Case Study in Khanhhoa Province Vietnam. *Procedia Eng.* **2016**, *154*, 1169–1175. [CrossRef]
22. Belayneh, A.; Adamowski, J.; Khalil, B.; Quilty, J. Coupling machine learning methods with wavelet transforms and the bootstrap and boosting ensemble approaches for drought prediction. *Atmos. Res.* **2016**, *172*, 37–47. [CrossRef]
23. Ali, Z.; Hussain, I.; Faisal, M.; Nazir, M.; Hussain, T.; Muhammad, Y.s.; Shoukry, A.; Gani, S. Forecasting Drought Using Multilayer Perceptron Artificial Neural Network Model. *Adv. Meteorol.* **2017**, *2017*, 5681308. Available online: <https://www.hindawi.com/journals/amete/2017/5681308/> (accessed on 11 June 2021). [CrossRef]
24. Mouatadid, S.; Raj, N.; Deo, R.; Adamowski, J. Input selection and data-driven model performance optimization to predict the Standardized Precipitation and Evaporation Index in a drought-prone region. *Atmos. Res.* **2018**, *212*, 130–149. [CrossRef]
25. Ali, M.; Deo, R.; Downs, N.; Maraseni, T. An ensemble-ANFIS based uncertainty assessment model for forecasting multi-scalar standardized precipitation index. *Atmos. Res.* **2018**, *207*, 155–180. [CrossRef]
26. Ganguli, P.; Janga Reddy, M. Ensemble prediction of regional droughts using climate inputs and the SVM–copula approach. *Hydrol. Process.* **2014**, *28*, 4889–5009. [CrossRef]
27. Deo, R.; Kisi, O.; Singh, V. Drought forecasting in eastern Australia using multivariate adaptive regression spline, least square support vector machine and M5Tree model. *Atmos. Res.* **2016**, *184*, 149–175. [CrossRef]
28. Xu, L.; Chen, N.; Xiang, Z.; Chen, Z. An evaluation of statistical, NMME and hybrid models for drought prediction in China. *J. Hydrol.* **2018**, *566*, 235–249. [CrossRef]
29. Khajehi, S.; Hamid, M. Towards an Improved Ensemble Precipitation Forecast: A Probabilistic Post-Processing Approach. *J. Hydrol.* **2017**, *546*, 476–489. [CrossRef]
30. Zhang, R.; Chen, Z.Y.; Xu, L.J.; Ou, C.Q. Meteorological drought forecasting based on a statistical model with machine learning techniques in Shaanxi province, China. *Sci. Total Environ.* **2019**, *665*, 338–346. [CrossRef] [PubMed]
31. Hayes, M.; Svoboda, M.; Comte, L.; Redmond, D.; Pasteris, P. Drought Monitoring: New Tools for the 21st Century. *Drought Water Cris. Sci. Technol. Manag. Issues* **2005**, *53*, 69. [CrossRef]
32. Wardlow, B.D.; Anderson, M.C.; Verdin, J.P. *Remote Sensing of Drought: Innovative Monitoring Approaches*; CRC Press: Boca Raton, FL, USA, 2012.
33. Tang, D.; Wei, F.; Yang, N.; Zhou, M.; Liu, T.; Qin, B. Learning Sentiment-Specific Word Embedding for Twitter Sentiment Classification. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, ACL 2014—Proceedings of the Conference, Baltimore, MD, USA, 22–27 June 2014; Volume 1, pp. 1555–1565. [CrossRef]
34. Severyn, A.; Moschitti, A. Twitter Sentiment Analysis with Deep Convolutional Neural Networks. In Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, Santiago, Chile, 9–13 August 2015; SIGIR '15; Association for Computing Machinery: New York, NY, USA, 2015; pp. 959–962. [CrossRef]
35. Agarwal, A.; Xie, B.; Vovsha, I.; Rambow, O.; Passonneau, R. Sentiment Analysis of Twitter Data. In Proceedings of the Proceedings of the Workshop on Languages in Social Media, Portland, OR, USA, 23 June 2011; LSM '11; Association for Computational Linguistics: Portland, OR, USA, 2011; pp. 30–38.
36. Liu, B. *Sentiment Analysis and Opinion Mining*; Morgan & Claypool Publishers: Williston, VT, USA, 2012.
37. Lachlan, K.A.; Spence, P.R.; Lin, X. Expressions of risk awareness and concern through Twitter: On the utility of using the medium as an indication of audience needs. *Comput. Hum. Behav.* **2014**, *35*, 554–559. [CrossRef]
38. Gruebner, O.; Lowe, S.; Sykora, M.; Shankardass, K.; Subramanian, S.; Galea, S. Spatio-Temporal Distribution of Negative Emotions in New York City After a Natural Disaster as Seen in Social Media. *Int. J. Environ. Res. Public Health* **2018**, *15*, 2275. [CrossRef]
39. Mittal, A.; Patidar, S. Sentiment Analysis on Twitter Data: A Survey. In Proceedings of the 2019 7th International Conference on Computer and Communications Management, Bangkok, Thailand, 27–29 July 2019; ICCCM 2019; Association for Computing Machinery: New York, NY, USA, 2019; pp. 91–95. [CrossRef]

40. He, Y.; Wen, L.; Zhu, T. Area Definition and Public Opinion Research of Natural Disaster Based on Micro-blog Data. *Procedia Comput. Sci.* **2019**, *162*, 614–622. [[CrossRef](#)]
41. Kam, J.; Stowers, K.; Kim, S. Monitoring of Drought Awareness from Google Trends: A Case Study of the 2011–2017 California Drought. *Weather Clim. Soc.* **2019**, *11*, 419–429. [[CrossRef](#)]
42. Kirilenko, A.; Molodtsova, T.; Stepchenkova, S. People as sensors: Mass media and local temperature influence climate change discussion on Twitter. *Glob. Environ. Chang.* **2015**, *30*, 92–100. [[CrossRef](#)]
43. Chen, X.; Zou, L.; Zhao, B. Detecting Climate Change Deniers on Twitter Using a Deep Neural Network. In Proceedings of the 2019 11th International Conference on Machine Learning and Computing, Zhuhai, China, 22–24 February 2019; ICMLC '19; Association for Computing Machinery: New York, NY, USA, 2019; pp. 204–210. [[CrossRef](#)]
44. Palmer, W. *Meteorological Drought*. *Weather Bureau Research Paper No. 45*; US Department of Commerce: Washington, DC, USA, 1965.
45. Tian, L.; Yuan, S.; Quiring, S.M. Evaluation of six indices for monitoring agricultural drought in the south-central United States. *Agric. For. Meteorol.* **2018**, *249*, 107–119. [[CrossRef](#)]
46. Houborg, R.; Rodell, M.; Li, B.; Reichle, R.; Zaitchik, B.F. Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations. *Water Resour. Res.* **2012**, *48*, W07525. [[CrossRef](#)]
47. Esuli, A.; Sebastiani, F. SENTIWORDNET: A Publicly Available Lexical Resource for Opinion Mining. In Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06), Genoa, Italy, 22–28 May 2006; European Language Resources Association (ELRA): Genoa, Italy, 2006.
48. Baccianella, S.; Esuli, A.; Sebastiani, F. SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. In Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10), Valletta, Malta, 19–21 May 2010; European Language Resources Association (ELRA): Valletta, Malta, 2010.
49. Miller, G.A.; Beckwith, R.; Fellbaum, C.; Gross, D.; Miller, K.J. Introduction to WordNet: An On-line Lexical Database. *Int. J. Lexicogr.* **1990**, *3*, 235–244. [[CrossRef](#)]