Feasibility Of Adding Twitter Data To Aid Drought Depiction: Case Study In Colorado

Sarbajit Mukherjee¹, Shih-Yu (Simon) Wang¹*, Daniella Hirschfeld², Joel Lisonbee³, Liping Deng⁴ and Robert Gillies¹

¹*Utah Climate Center, Dept. of Plants, Soil & Climate, Utah State University, Logan, 84322, UT, USA.
²Dept. of Landscape Architecture & Environmental Planning, Utah State University, Logan, 84322, UT, USA.
³NOAA/National Integrated Drought Information System (NIDIS), Organization, Boulder, 80305, CO, USA.
⁴College of Ocean and Meteorology, Guangdong Ocean University, Zhanjiang, China.

*Corresponding author(s): Shih-Yu Wang
E-mail(s): 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: How the usage of Twitter by a large population may change during the development of a major drought alongside how changing usage promotes drought detection? For this reason, by investigating contemporary procedures, this paper scrutinizes the potential to advance drought depiction. Hence, an analysis of how social media data, in conjunction with meteorological records, was conducted towards improvement in the detection of drought and furthermore, its progression. The research utilized machine learning techniques applied over satellite-derived drought conditions in Colorado. Specifically, 3 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 maintained that the data integration of resources is viable given that the Twitter-based model outperformed
the control run which did not include social media input. Furthermore
the twitter-based model was superior in predicting drought severity.

**Keywords:** Drought, Twitter, Machine Learning, Colorado

## 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 [https://www.drought.gov/drought-research], 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 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 the 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 remains 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. Meanwhile, 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 [17, 18]. 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, 19–24]. In the USA, “#drought” and related hashtags on Twitter have been found to increase correspondingly during high–impact drought events [8, 25]. Researchers also have attempted to use data from Twitter to study climate change perceptions [26, 27]. The aforesaid studies reflect the potential of using social media platforms like Twitter coupled to sentiment analysis methods for analyzing the progression of a drought. Nonetheless, there are challenges associated with such analysis: First, past studies such as [8, 26] 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 as 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. Among them, the Palmer Drought Severity Index (PDSI, [28]) is one of the six more common drought indices for North America [29]. The US Drought monitor https://droughtmonitor.unl.edu/ 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). 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 their ability to measure the terrestrial water storage (i.e. variations in water stored at all 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 [30]. The groundwater and soil moisture data used here as complementary indicators to PDSI are obtained from https://nasagrace.unl.edu/.

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 timeframe. 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 such, we searched for key words 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 a location information of Colorado. But using the real-time streaming setup has a compromise in that most tweets did not have a geo-location attached to them.

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).
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, drought-monitor, continu, basin, need, help, news, increas, record, west, southwest, wildfir, 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 |

Table 2: Sample of the search terms for data cleaning and it’s 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, ... |

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 context, 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-intense 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.

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>’. [eg: ‘@DroughtTalker’ to ‘<user>’].
4. Replacing Consecutive letters: 3 or more consecutive letters are replaced by 2 letters. [eg: ‘Heyyyy’ to ‘Heyy’].
5. Replacing Emojis: Replace emojis by using a regex expression. [eg: ‘:)' to ‘<smile>’]
6. Replacing Contractions: Replacing contractions with their meanings. [eg: “can’t” to ‘can not’]
7. Removing Non-Alphabets: Replacing characters except Digits, Alphabets and pre-defined Symbols with a space. [eg: $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
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 [31].

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 [31], [32], 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 [33], and no prior training data is required. 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, also known as 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.

3 Experiments

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. To keep the computational load reasonable, the mapping used by the SVM schemes are designed to ensure that dot products may be computed easily in terms of the variables in the original space, by defining them in terms of a kernel function $K(x,y)$.
selected to suit the problem. 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. Finally, the deep learning (DL) method is based on a multi-layer feed-forward artificial neural network and the training process is optimized using stochastic gradient descent during the back propagation step.

3.1 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 January 1, 2019 to December 31, 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 1–3 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, root\_zone\_soil\_3\_week\_before, surface\_soil, 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 1a, 1b, 1c shows the results on test dataset for 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. Each small blue circle in the figures is represented by the tuple \((PDSI_{\text{actual}}, PDSI_{\text{pred}})\), where \(PDSI_{\text{pred}}\) are the PDSI values predicted by the individual machine learning models and \(PDSI_{\text{actual}}\) are the actual PDSI values. The red
Fig. 1: Result on Test Set. 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 3: Different combinations of twitter data when added as individual features to the control run ‘D’. ‘P’ represent the count of positive tweets. ‘N’ represent 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                     |

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’.

3.2 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 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.

3.3 Adding Twitter Data To Control Run

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 used the machine learning techniques of generalized linear model, support vector machine and deep learning to train all of our twitter based models.

**Model6 – D+P+N+1P+1N:** The rational 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 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 the Table 4 and Table 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 generalized linear model and support vector machine techniques. We can see the similar effect in Figures 2a, 2b, 2c where the blue circles are closer to the reference red line as compared to the ones in case of the control
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/GqD6i5CDOI | 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 | 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 | 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’re about to get hit 3 times with storms in the next 5 days. With the drought we’ve 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 | 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 |
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      |

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      |

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’.

Model10 – 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 the Table 4 and Table 7, we can see that the correlation values with the inclusion of Twitter data have improved for all the three cases and correspondingly the RMSE values have also improved for all the three machine learning techniques. These effects are shown in Figures 3a, 3b, 3a where the blue circles are closer to the reference red line as compared to the control run model. If we compare the results in Table 6 and Table 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.

Performance of Other Twitter Based Models: In this section we will discuss the performance of all the models listed in Table 3. Figures 4a, 4b and 4c 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 4a, 4b and 4c represent the difference between the RMSE values for the control run and the Twitter based models. The y-axis represent 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.
Twitter Drought Study In Colorado

Fig. 2: Result on Test Set 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 $(P_{DSI_{actual}}, P_{DSI_{pred}})$. The red dashed line is the reference line to test the model performance.
Fig. 3: Result on Test Set 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 \((PDSI_{\text{actual}}, PDSI_{\text{pred}})\). The red dashed line is the reference line to test the model performance.
Fig. 4: The x-axis represent the difference between the RMSE values for the control run and the Twitter based models. The y-axis represent the difference between the correlation coefficient values between the control run and the Twitter based models.
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 the Figures 4a, 4b and 4c. If models fall in the top right quadrant, then the performance improvement only exists in terms of 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 not a performance improvement. From Figures 4a, 4b and 4c, 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.

**Fig. 5**: 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 represent a Twitter based model. Each hollow circle represents the control run model.

We also constructed Figure 5, 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 5 represent either Model 6 or Model 10, which is a Twitter based model ($T_{\text{RMSE}_{\text{model}}}$, $T_{R^2_{\text{model}}}$), where model denotes either deep learning, support vector machine or generalized linear model. Each hollow circle in Figure 5 represents the control run model ($C_{\text{RMSE}_{\text{model}}}$, $C_{R^2_{\text{model}}}$) 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 con-
trol run. In Figure 5, 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
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.

We carried out our experiments about building a model able to pre-
dict PDSI values, using RapidMiner (https://rapidminer.com/) which is an
integrated environment that enables rapid prototyping for building 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.

4 Testing On 2021 (Untrained) Data

This 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).

Figures 6a, 6b, 6c 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 pre-
dictions 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 \((PDSI_{\text{actual}}, PDSI_{\text{c,pred}})\). The blue dashed line is the reference line to test the model performance and each point on the line is represented as \((PDSI_{\text{actual}}, PDSI_{\text{actual}})\), 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 6a, 6b, 6c, 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 7, 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.

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 positive or negative tweet, so if people have retweeted the NIDIS tweets then our system was able to capture that engagement.

5 Summary and Conclusion

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 (January 1, 2019 to December 31, 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 upon keywords which were closely related to drought. It was found that there were user discussions were varied regarding drought during
Twitter Drought Study In Colorado

![Diagram of prediction results using different models.](image)

(a) Result on unseen data using Generalized Linear Model

(b) Result on unseen data using Support Vector Machine

(c) Result on unseen data using Deep Learning

Fig. 6: Prediction results on data from January-April 2021 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. 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.
Throughout the analysis of the frequency of different words used, we observed a change in the user perception about drought as it continued to worsen over the 2019 - 2020 period. Also, 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 model can improve the forecast of PDSI in times of a developing drought. In fact, eight of the 10 models tested 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.
6 Prospects

Machine learning (ML) models thrive on good quality data. Machine learning models also have gained prominence in the area of drought forecasting. Researchers have developed models using time series analysis [34, 35], neural networks [36–41], fuzzy inference systems [42], support vector regression [43–45] and different ensemble techniques [46, 47] 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 [48, 49].

There are a few additional steps that can improve the prediction models in the future. One of the important improvements lies in data collection. It is reasonable to suppose that increasing the training dataset in terms of both Twitter and meteorological data would lead to an improvement the ML models. 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 ample tweets about drought under certain conditions, e.g., “peace or leisure times”; in such cases data would be lost. A further improvement would be in the sentiment analysis step, i.e., 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.

Declarations

- Funding: This project was funded through NOAA’s National Integrated Drought Information System (NIDIS). Simon Wang is partially supported by U.S. Department of Energy, Office of Biological and Environmental Research program under Award Number DE-SC0016605. Liping Deng is supported by National Natural Science Foundation of China (Grant Number 41875071).
- Conflict of interest/Competing interests: Not applicable
- Ethics approval: Not applicable
- Consent to participate: Not applicable
- Consent for publication: All coauthors agreed to publish this work.
• Availability of data and materials: Data are available upon request.
• Code availability: Codes are available upon request.
• Authors’ contributions: SM and SW designed and wrote the paper. SM performed the data collection and analysis. JL and DH contributed to the discussion and writing. JD and RG edited the paper.

References

[1] Fonseca, F., Metz, S.: Lake Mead drops to a record low amid drought. Associated Press, LA Times (2021). https://www.latimes.com/world-nation/story/2021-06-11/lake-mead-key-reservoir-colorado-river-record-low-drought

[2] Wang, S.-Y., Santanello, J., Wang, H., Barandiaran, D., Pinker, R., Schubert, S., Gillies, R., Oglesby, R., Hilburn, K., Kilic, A., Houser, P.: An intensified seasonal transition in the central u.s. that enhances summer drought. Journal of Geophysical Research 120(17), 8804–8816 (2015). https://doi.org/10.1002/2014JD023013

[3] Switzer, D., Vedlitz, A.: Investigating the determinants and effects of local drought awareness. Weather, Climate, and Society 9(4), 641–657 (2017). https://doi.org/10.1175/WCAS-D-16-0052.1

[4] Shafiee-Jood, M., Deryugina, T., Cai, X.: Modeling users’ trust in drought forecasts. Weather, Climate, and Society 13(3), 649–664 (2021). https://doi.org/10.1175/WCAS-D-20-0081.1

[5] Smith, A.B.: U.S. Billion-dollar Weather and Climate Disasters, 1980 - present (NCEI Accession 0209268). NOAA National Centers for Environmental Information Dataset. Last Accessed on August 10, 2021 (2020). https://doi.org/10.25921/stkw-7w73

[6] Bolinger, B.: How Drought Prone Is Your State? A Look at the Top States and Counties in Drought Over the Last Two Decades. National Integrated Drought Information System. Last Accessed on August 10, 2021 (2019). https://www.drought.gov/news/how-drought-prone-your-state-look-top-states-and-counties-drought-over-last-two-decades

[7] Sutton, J., Palen, L., Shklovski, I.: Backchannels on the front lines: Emergent uses of social media in the 2007 southern california wildfires. Proceedings of the 5th International ISCRAM Conference (2008)

[8] Smith, K.H., Tyre, A.J., Tang, Z., Hayes, M.J., Akyuz, F.A.: Calibrating human attention as indicator monitoring #drought in the twittersphere. Bulletin of the American Meteorological Society 101(10), 1801–1819 (2020). https://doi.org/10.1175/BAMS-D-19-0342.1
Twitter Drought Study In Colorado

[9] Hughes, A., Palen, L.: Twitter adoption and use in mass convergence and emergency events. International Journal of Emergency Management 6, 248–260 (2009). https://doi.org/10.1504/IJEM.2009.031564

[10] Abdelhaq, H., Sengstock, C., Gertz, M.: Eventweet: Online localized event detection from twitter. Proc. VLDB Endow. 6(12), 1326–1329 (2013). https://doi.org/10.14778/2536274.2536307

[11] 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. WWW ’12 Companion, pp. 695–698. Association for Computing Machinery, New York, NY, USA (2012). https://doi.org/10.1145/2187980.2188183

[12] Mathioudakis, M., Koudas, N.: Twittermonitor: Trend detection over the twitter stream. In: Proceedings of the 2010 ACM SIGMOD International Conference on Management of Data. SIGMOD ’10, pp. 1155–1158. Association for Computing Machinery, New York, NY, USA (2010). https://doi.org/10.1145/1807167.1807306

[13] 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. WWW ’10, pp. 851–860. Association for Computing Machinery, New York, NY, USA (2010). https://doi.org/10.1145/1772690.1772777

[14] Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., Rippey, B., Tinker, R., Palecki, M., Stooksbury, D., Miskus, D., Stephens, S.: The drought monitor. Bulletin of the American Meteorological Society 83(8), 1181–1190 (2002). https://doi.org/10.1175/1520-0477-83.8.1181

[15] Glaser, M.: California wildfire coverage by local media, blogs, Twitter, maps and more. PBS MediaShift (2007)

[16] Stelter, B.: How social media is helping Houston deal with Harvey floods. CNN Money (2017)

[17] Tang, D., Wei, F., Yang, N., Zhou, M., Liu, T., Qin, B.: Learning sentiment–specific word embedding for twitter sentiment classification. In: 52nd Annual Meeting of the Association for Computational Linguistics, ACL 2014 - Proceedings of the Conference, vol. 1, pp. 1555–1565 (2014). https://doi.org/10.3115/v1/P14-1146

[18] 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. SIGIR ’15, pp. 959–962. Association for Computing Machinery, New York, NY, USA (2015). https://doi.org/10.1145/2766462.2767830. https://doi.org/10.1145/2766462.2767830

[19] Agarwal, A., Xie, B., Vovsha, I., Rambow, O., Passonneau, R.: Sentiment analysis of twitter data. In: Proceedings of the Workshop on Languages in Social Media. LSM ’11, pp. 30–38. Association for Computational Linguistics, USA (2011)

[20] Liu, B.: Sentiment Analysis and Opinion Mining. Morgan & Claypool Publishers, ?? (2012)

[21] 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. Computers in Human Behavior 35, 554–559 (2014). https://doi.org/10.1016/j.chb.2014.02.029

[22] 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. International Journal of Environmental Research and Public Health 15(10), 2275 (2018). https://doi.org/10.3390/ijerph15102275

[23] Mittal, A., Patidar, S.: Sentiment analysis on twitter data: A survey. In: Proceedings of the 2019 7th International Conference on Computer and Communications Management. ICCCM 2019, pp. 91–95. Association for Computing Machinery, New York, NY, USA (2019). https://doi.org/10.1145/3348445.3348466. https://doi.org/10.1145/3348445.3348466

[24] He, Y., Wen, L., Zhu, T.: Area definition and public opinion research of natural disaster based on micro-blog data. Procedia Computer Science 162, 614–622 (2019). https://doi.org/10.1016/j.procs.2019.12.030. 7th International Conference on Information Technology and Quantitative Management (ITQM 2019): Information technology and quantitative management based on Artificial Intelligence

[25] Kam, J., Stowers, K., Kim, S.: Monitoring of drought awareness from google trends: A case study of the 2011–17 california drought. Weather, Climate, and Society 11(2), 419–429 (2019). https://doi.org/10.1175/WCAS-D-18-0085.1

[26] Kirilenko, A., Molodtsova, T., Stepchenkova, S.: People as sensors: Mass media and local temperature influence climate change discussion on twitter. Global Environmental Change 30 (2015). https://doi.org/10.1016/j.gloenvcha.2014.11.003
[27] 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. ICMLC ’19, pp. 204–210. Association for Computing Machinery, New York, NY, USA (2019). https://doi.org/10.1145/3318299.3318382.

[28] Palmer, W.C.: Meteorological Drought. Weather Bureau Research Paper No. 45. Washington, DC: US Department of Commerce, (1965)

[29] Tian, L., Yuan, S., Quiring, S.M.: Evaluation of six indices for monitoring agricultural drought in the south-central united states. Agricultural and Forest Meteorology 249, 107–119 (2018). https://doi.org/10.1016/j.agrformet.2017.11.024

[30] 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 Resources Research 48(7) (2012) https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2011WR011291. https://doi.org/10.1029/2011WR011291

[31] 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). European Language Resources Association (ELRA), Genoa, Italy (2006). http://www.lrec-conf.org/proceedings/lrec2006/pdf/384.pdf.pdf

[32] 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). European Language Resources Association (ELRA), Valletta, Malta (2010). http://www.lrec-conf.org/proceedings/lrec2010/pdf/769_Paper.pdf

[33] Miller, G.A., Beckwith, R., Fellbaum, C., Gross, D., Miller, K.J.: Introduction to WordNet: An On-line Lexical Database*. International Journal of Lexicography 3(4), 235–244 (1990) https://academic.oup.com/ijl/article-pdf/3/4/235/9820417/235.pdf. https://doi.org/10.1093/ijl/3.4.235

[34] Ömer Faruk, D.: A hybrid neural network and arima model for water quality time series prediction. Engineering Applications of Artificial Intelligence 23(4), 586–594 (2010). https://doi.org/10.1016/j.engappai.2009.09.015

[35] Han, P., Wang, P., Zhang, S., Zhu, D.: Drought forecasting based on the remote sensing data using arima models. Mathematical and Computer
[36] Mishra, A., Desai, V.: Drought forecasting using feed-forward recursive neural network. Ecological Modelling 198, 127–138 (2006). https://doi.org/10.1016/j.ecolmodel.2006.04.017

[37] Santos, J., Portela, M., Pulido-Calvo, I.: Spring drought prediction based on winter nao and global sst in portugal. Hydrological Processes 28, 1009–1024 (2014). https://doi.org/10.1002/hyp.9641

[38] 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 Engineering 154, 1169–1175 (2016). https://doi.org/10.1016/j.proeng.2016.07.528

[39] 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. Atmospheric Research 172, 37–47 (2016). https://doi.org/10.1016/j.atmosres.2015.12.017

[40] 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. Advances in Meteorology 2017 (2017). https://doi.org/10.1155/2017/5681308

[41] 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. Atmospheric Research 212, 130–149 (2018). https://doi.org/10.1016/j.atmosres.2018.05.012

[42] Ali, M., Deo, R., Downs, N., Maraseni, T.: An ensemble-anfis based uncertainty assessment model for forecasting multi-scalar standardized precipitation index. Atmospheric Research 207, 155–180 (2018). https://doi.org/10.1016/j.atmosres.2018.02.024

[43] Ganguli, P., Janga Reddy, M.: Ensemble prediction of regional droughts using climate inputs and the svm–copula approach. Hydrological Processes 28, 4889–5009 (2014). https://doi.org/10.1002/hyp.9966

[44] Deo, R., Kisi, O., Singh, V.: Drought forecasting in eastern australia using multivariate adaptive regression spline, least square support vector machine and m5tree model. Atmospheric Research 184, 149–175 (2016). https://doi.org/10.1016/j.atmosres.2016.10.004

[45] Xu, L., Chen, N., Xiang, Z., Chen, Z.: An evaluation of statistical, nmm
Twitter Drought Study In Colorado

and hybrid models for drought prediction in China. Journal of Hydrology 566 (2018). https://doi.org/10.1016/j.jhydrol.2018.09.020

[46] Khajehei, S., Hamid, M.: Towards an improved ensemble precipitation forecast: A probabilistic post-processing approach. Journal of Hydrology 546, 476–489 (2017). https://doi.org/10.1016/j.jhydrol.2017.01.026

[47] 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. Science of The Total Environment 665 (2019). https://doi.org/10.1016/j.scitotenv.2019.01.431

[48] Hayes, M., Svoboda, M., Comte, L., Redmond, D., Pasteris, P.: Drought monitoring: New tools for the 21st century. Drought and Water Crises: Science, Technology, and Management Issues (2005). https://doi.org/10.1201/9781420028386.ch3

[49] Wardlow, B.D., Anderson, M.C., Verdin, J.P.: Remote Sensing of Drought: Innovative Monitoring Approaches. CRC Press, ??? (2012)