The Usefulness of Stable Water Isotopes in Improving Drought Prediction

by
Logan Adams

A THESIS

submitted to

Oregon State University

Honors College

in partial fulfillment of
the requirements for the
degree of

Honors Baccalaureate of Science in Ecological Engineering
(Honors Scholar)

Honors Baccalaureate of Arts in International Studies
(Honors Scholar)

Presented May 29, 2019
Commencement June 2019

AN ABSTRACT OF THE THESIS OF

Logan Adams for the degree of Honors Baccalaureate of Science in Ecological Engineering and Honors Baccalaureate of Arts in International Studies presented on May 29, 2019. Title: The Usefulness of Stable Water Isotopes in Improving Drought Prediction

Abstract approved: ____

Stephen Good

Drought represents an import part of natural hydrologic cycles that govern the rate and transportation of water on global and regional scales. (Trenberth, et al. 2015) Agriculturally, drought represents a threat to crop production and can endanger political and economic stability through the loss of resources and the threat of famine. (Sternberg, 2012) Ecosystem observations also indicate drought has a significant impact on biodiversity and productivity both of large-scale fauna and microbial communities. (Eisenhauer, et al., 2011) As the impacts of global climate change include an increased rate and severity of drought the need to gain a better understanding of drought mechanisms and patterns have also increased. This paper utilizes data from the National Center for Atmospheric Research's (NCAR) self-calibrated Palmer Drought Severity Index (scPDSI), the University of Utah's waterisotopes.org, and NASA's Tropospheric Emission Spectrometer (TES) to investigate the impact of adding stable water isotopes to drought prediction models. Model accuracy is reported as the root-mean-squared-error (RMSE), and the impact of stable water isotopes are reported as percent improvement. The analysis included linear, support-vector, ridge, and least absolute shrinkage and selection operator regression across multiple years in three distinct hydrologic regimes: atmospheric water vapor, precipitation, and surface water (lakes, rivers, and streams). Stable water isotopes were shown to be useful at improving drought prediction, especially at a time lag of +1 month, and when the isotopes were collected from surface water. The level of improvement varied greatly, with some models showing that it actually made it worse, but overall indicated that using stable water isotopes may assist in addressing the complexity of trying to predict drought. Future research needs to address the lack of geographically diverse stable water isotope data in precipitation and surface water and should seek to compare scPDSI values on a smaller scale that can better take into account the impact of internal dynamics within a region of interest.

Key Words: Drought, stable water isotopes, PDSI

Corresponding e-mail address: adamslo@oregonstate.edu

The Usefulness of Stable Water Isotopes in Improving Drought Prediction

by
Logan Adams

A THESIS

submitted to

Oregon State University

Honors College

in partial fulfillment of
the requirements for the
degree of

Honors Baccalaureate of Science in Ecological Engineering
(Honors Scholar)

Honors Baccalaureate of Arts in International Studies
(Honors Scholar)

Presented May 29, 2019
Commencement June 2019

Honors Baccalaureate of Science in Ecological Engineering and Honors Baccalaureate of Arts in International Studies project of Logan Adams presented on May 29, 2019.

APPROVED:

____
Stephen Good, Mentor, representing Biological and Ecological Engineering

____
David Noone, Committee Member, representing Water Resources Science and Engineering

____
Julia Jones, Committee Member, representing Geography, Environmental Sciences, and Marine Resource Management

____
Rebekah Lancelin, Committee Member, representing International Studies

____
Toni Doolen, Dean, Oregon State University Honors College

I understand that my project will become part of the permanent collection of Oregon State University, Honors College. My signature below authorizes release of my project to any reader upon request.

____
Logan Adams, Author

The Usefulness of Stable Water Isotopes in Improving Drought Prediction

Logan Adams

ABSTRACT

Drought represents an import part of natural hydrologic cycles that govern the rate and transportation of water on global and regional scales. (Trenberth, et al. 2015) Agriculturally, drought represents a threat to crop production and can endanger political and economic stability through the loss of resources and the threat of famine. (Sternberg, 2012) Ecosystem observations also indicate drought has a significant impact on biodiversity and productivity both of large-scale fauna and microbial communities. (Eisenhauer, et al., 2011) As the impacts of global climate change include an increased rate and severity of drought the need to gain a better understanding of drought mechanisms and patterns have also increased. This paper utilizes data from the National Center for Atmospheric Research's (NCAR) self-calibrated Palmer Drought Severity Index (scPDSI), the University of Utah's waterisotopes.org, and NASA's Tropospheric Emission Spectrometer (TES) to investigate the impact of adding stable water isotopes to drought prediction models. Model accuracy is reported as the root-mean-squared-error (RMSE), and the impact of stable water isotopes are reported as percent improvement. The analysis included linear, support-vector, ridge, and least absolute shrinkage and selection operator regression across multiple years in three distinct hydrologic regimes: atmospheric water vapor, precipitation, and surface water (lakes, rivers, and streams). Stable water isotopes were shown to be useful at improving drought prediction, especially at a time lag of +1 month, and when the isotopes were collected from surface water. The level of improvement varied greatly, with some models showing that it actually made it worse, but overall indicated that using stable water isotopes may assist in addressing the complexity of trying to predict drought. Future research needs to address the lack of geographically diverse stable water isotope data in precipitation and surface water and should seek to compare scPDSI values on a smaller scale that can better take into account the impact of internal dynamics within a region of interest.

1. INTRODUCTION

> "I talked with families who had lost their wheat crop, lost their corn crop, lost their livestock, lost the water in their well, lost their garden and come through to the end of the summer without one dollar of cash resources, facing a winter without feed or food—facing a planting season without seed to put in the ground.

> I shall never forget the fields of wheat so blasted by heat that they cannot be harvested. I shall never forget field after field of corn stunted, earless and stripped of leaves, for what the sun left the grasshoppers took. I saw brown pastures which would not keep a cow on fifty acres." (Roosevelt, 1936)

These were the words of President Franklin Delano Roosevelt on September 6, 1936, when addressing the nation about one of the greatest crises facing the United States at the time, the Dust Bowl. Though unsustainable agricultural practices played a part in the creation of the Dust Bowl, it was directly caused by an increased dryness within the region starting in 1930, and a series of severe droughts within the region that began in 1934 (Porter, 2012). The Dust Bowl also caused the displacement of nearly 3.5 million people between 1930 and 1940, made up not just of poorer farmers, but of doctors, lawyers, teachers, and other white-collar professions that could no longer be supported by broken communities. (Worster, 2012) In the long-term, the drought caused an economic decrease in the per-acre value of the land of 17-28% and an economic downturn that would still be felt in the 1950s. (Hornbeck, 2012). The Dust Bowl also had a major cultural impact on the United States, with stories of displacement told by John Steinbeck and documented by Dorothea Lange (Alexander & Nugent, 2018). An entire generation of people was impacted and the course of their lives and the economic course of the country was changed as a result. In the eighty-three years since President Roosevelt gave his speech, the United States has experienced a severe drought at least once every ten years, an occurrence that is reflective of patterns that exist everywhere in the world. (Cordova & Porter, 2015) Today, in the United States, when we discuss drought it is often more a conversation of resource usage and water allocation. (Stoutenborough & Vedlitz, 2014) The idea that someone would die as a direct result of drought is in many ways foreign to us, but for millions of people around the world that is still a reality. (UNICEF, 2019)

Drought represents an import part of natural hydrologic cycles that govern the rate and transportation of water on global and regional scales. (Trenberth, et al. 2015) Agriculturally, drought represents a threat to crop production and can endanger political and economic stability through the loss of resources and the threat of famine. (Sternberg, 2012) Ecosystem observations also indicate drought has a significant impact on biodiversity and productivity both of large-scale fauna and microbial communities. (Eisenhauer, et al., 2011)

One area of particular interest is that of South America. An extensive drought in 2005 was shown to have large impacts on the Amazon rainforest, making large swaths of the forest vulnerable to disease and death, and also impacting biomass production and carbon uptake within the forest. (Phillips et al., 2009; Markewitz et al., 2010) Authors also expressed concern that this impact and the expected increase in dryness in the region could point towards a positive feedback loop in which the rainforest's ability to uptake carbon would decrease, increasing carbon dioxide concentrations within the atmosphere and increasing the rate of climate change. Deforestation within the region is also expect to increase dryness as forestland shifts regimes into savannah, and as grassland begins to desertify, which will create a second positive feedback loop to increase the occurrence and severity of drought especially within the south-eastern region. (Staal et al., 2015) Drought within this region also has the ability to impact water security for the entire region. The Panatal wetland regulates water for large parts of Bolivia, Paraguay, and Brazil and safeguards biological diversity, both of which have become threatened by drought. (Bergier et al., 2017). Outside of the environmental impacts there is a very real human concern as over two million people in Bolivia alone face water insecurity, and the issue is only expected to increase in the coming century. (Bush et al., 2010) The impacts on South America represent a microcosm from which we can see the negative impact of increased drought on a larger scale. As the impacts of global climate change include an increased rate and severity of drought the need to gain a better understanding of drought mechanisms and patterns have also increased.

1.1 WHAT IS DROUGHT?

In their 1985 paper, “Understanding the Drought Phenomenon”, Wilhite and Glantz outlined four distinct types of drought, Meteorological, Agricultural, Hydrologic, and Socio-economic. (Wilhite & Glantz, 1985). Meteorological Drought is often the most commonly used and is defined as a site-specific degree of “dryness” for an area that is expected based off of historical trends within the region. Agricultural Drought focuses on the agricultural impacts of drought, such as soil moisture, and the impact on plant growth and thus crop yields. Hydrologic Drought is concerned with the impact of dry spells on hydrology, often at a watershed or river basin scale, and is used as a water balance mechanism. Finally, Socio-economic Drought often incorporates aspects of Meteorological, Agricultural, and Hydrologic Drought but within the context of the specific social implications that a drought may have, or the economic impact on an area that is more extensive than crop yield calculations. These four definitions of drought have been the most common since 1985, though, in a recent paper published in 2017, a new definition of drought was proposed, called Ecological Drought. Ecological Drought is defined as a water deficit that drives an ecosystem beyond a threshold of vulnerability, impacts ecosystem services, and triggers feedback in natural and/or human systems. (Crausbay et al., 2017) Although relatively new, this definition of drought takes into account the reaction of natural systems to drought outside of an anthropocentric view and represents an important shift in the perspective of many researchers. However, due to its recent development, more research using the Ecological Drought model is still needed.

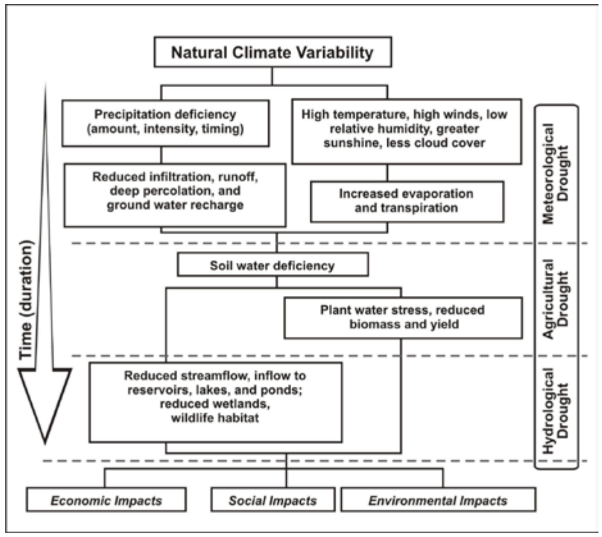

Fig. 1 Sequence of drought occurrence and impacts for commonly accepted drought types.(Source: NDMC)

Figure 1. From the National Drought Mitigation Center (NDMC) shows the connections between the four types of drought listed by Wilhite and Glantz and helps to illustrate the complex relationship between drought and society.

1.2 WHAT ARE STABLE WATER ISOTOPES?

The basic structure of water is that of two hydrogens and one oxygen (H_2O), but the specific atoms of Hydrogen and Oxygen making up a water molecule can differ. The basic and most abundant (99.98%) structure of a Hydrogen atom is one proton and one electron, known also as Hydrogen-1 or protium (δ^1H). However, Hydrogen can also exist as an isotope with either an additional neutron, which is known as Hydrogen-2 or deuterium (δ^2H) or two additional neutrons, which is known as Hydrogen-3 or tritium (δ^3H). (O'Leary, 2012) Tritium naturally occurs within the environment from interactions of cosmic rays with atmospheric gas and has also been released during nuclear weapons testing, however tritium has a half-life of 12.32 years (Miessler et al., 2014) which makes it detectable, but not practical for long-term monitoring or measurement. Protium and deuterium are both extremely stable, and are therefore very desirable for long-term monitoring. (Criss, 1999) Similar relationships exist within Oxygen isotopes. Oxygen-16 ($\delta^{16}O$) is the most common form of Oxygen (99.762%) and is made up of eight protons and eight neutrons. Oxygen-17 ($\delta^{17}O$), and Oxygen-18 ($\delta^{18}O$) also exist as stable isotopes, with Oxygen-18 being the most commonly used. (Cook & Lauer, 1968)

Deuterium and Oxygen-18 are useful because they are so rare compared to protium or Oxygen-16 that they can be used as a tracer within the environment, helping to track the movement of water through a system and acting as an isotopic fingerprint of a location. (Craig & Gordon, 1965) Furthermore, the additional mass within the heavier isotopes impacts their ability to evaporate and transpire, so that the lighter isotopes are slightly more likely to do so. (Dansgaard, 1964) Stable isotopes in water ($\delta^{18}O$, δ^2H) are important indicators for investigating hydrological processes across spatial and temporal scales and have been used effectively to track global weather patterns. (Good, et al., 2014) The relationship between the deuterium and oxygen-18 can also be defined as deuterium-excess (δ_{excess}) as a way to further understand the relationship between Hydrogen and Oxygen stable water isotopes and the specific isotopic signatures of a location. (Daansgaard, 1964)

1.3 DROUGHT PREDICTION

Several studies have concluded that drought is not the result of any single variable, but rather the result of complex interactions between global weather patterns, regional pressure systems, local topography, and internal dynamics of land surface processes. (NDMC, 2019). The number of variables that go into drought makes any attempt at prediction particularly difficult, and improvement of drought prediction methods represents an important step in our understanding of drought, and as a tool to address the issues that drought can cause.

1.4 RESEARCH GOALS

One of the key drivers of the movement of water is that of evapotranspiration (the direct evaporation of water from the surface of the earth, and the transpiration of water through vegetation). It is expected that drought will increase the rate of evapotranspiration that will occur as thermodynamic inputs increase the rate of cooling necessary for plant life. (Li et al., 2009) This offers the opportunity to use stable water isotopes to improve our understanding of the movement and occurrence of drought. Currently there are only three studies that investigate the relationship between drought and stable isotopes (Marchina et al, 2019;Wu et al., 2017; Vanplantinga et al., 2016) and they are limited to specific geographical areas and relatively small time-scales (one year), meaning that no metadata analysis of the relationship between drought and stable water isotopes has been performed. Recently the University of Utah’s Center for High-Performance Computing published one of the most comprehensive global datasets of water isotope data was published for public use with stable water isotope measurements from 1974 to the present for precipitation and surface waters across the world. Similarly, NASA’s Tropospheric Emissions Spectrometer, which measured atmospheric isotope ratios, has published data measured from 2005-2011. Using these two datasets it is possible to measure changes in stable water isotopes in three key hydrologic areas (precipitation, surface water, and the atmosphere), throughout the globe across several years. Furthermore, the National Center for Atmospheric Research has maintained a public dataset of global occurrences of drought using the Palmer Drought Severity Index (PDSI), a meteorological drought index. The goal of this research project is to determine how well the PDSI can be predicted using basic physical data (latitude, longitude, elevation, and time) and the impact of adding stable water isotope data to the drought prediction model.

2. DATA & METHODS

2.1 PALMER DROUGHT SEVERITY INDEX

The PDSI was first proposed in 1965 by meteorologist Wayne Palmer in his paper "Meteorological Drought", published by the U.S. Weather Bureau. (Palmer, 1965) The PDSI takes into account supply, measured as Precipitation (P) and demand measured as Potential Evapotranspiration (PE) to create an accounting system with which to calculate the departure of a region from normal dryness based off of a bucket-type model. The original PDSI utilized weather data from the central United States from 1947-1948 to create an algorithm in which historical data for a specific location was used to assume evapotranspiration and recharge rates. This was done to simplify the application of the PDSI model and allow estimates to be made without needing to take direct PE measurements, relying on historical data instead. Another simplifying assumption that was made was to calculate PE using the Thornthwaite equation which is based only on temperature, latitude, and month, instead of the more sophisticated Penman-Monteith equation, which can lead to errors in regions of limited energy. (Hobbins et al, 2008) This was done as a practical consideration as the data required to calculate the Penman-Monteith PE is not as readily available as that of the Thornthwaite equation. (Palmer, 1965)

The index value calculated by the PDSI refers to a standardized measure ranging from negative ten to positive ten. A negative value indicates increased dryness compared to the meteorological value, while a positive value indicates increased wetness. From Figure 2 you can see the original classification from Palmer which describes any deviation with an absolute value of greater than four as extreme, between three and four as severe, between two and three as moderate, between one and two as mild, between one-half and one as incipient, and between zero and one-half as normal.

TABLE 11.—*Classes for wet and dry periods*

X	Class
≥ 4.00	Extremely wet.
3.00 to 3.99	Very wet.
2.00 to 2.99	Moderately wet.
1.00 to 1.99	Slightly wet.
.50 to .99	Incipient wet spell.
.49 to −.49	Near normal.
−.50 to −.99	Incipient drought.
−1.00 to −1.99	Mild drought.
−2.00 to −2.99	Moderate drought.
−3.00 to −3.99	Severe drought.
≦ −4.00	Extreme drought.

Fig. 2 Original classification of the Palmer Drought Severity Index from Palmer, 1965. *X* indicates the specific index value calculated.

The original standardization set forth by Palmer in 1965 was based on data limited to fixed points within the central United States, so in 2004 Wells et al. proposed a way to improve the spatial comparability of the index using a self-calibrating PDSI (scPDSI) using local conditions, instead of using the fixed coefficients, which was found to perform better than the original (Wells, et al., 2004).

Both PDSI and scPDSI are designed to be strongly auto-correlated to account for long-term trends in drought. This practice of taking precedent conditions into account somewhat limits the usefulness of the index to analyzing longer-term trends in drought conditions, but in doing so it also makes it more useful than other statistically based drought indices by recognizing that drought is not wholly independent from itself over time. (Dai et al, 2011, van der Schrier et al., 2006) This longer-term analysis may also allow PDSI and scPDSI to better account for the impact of global warming, the effects of which have already started to be felt in the twentieth and twenty-first centuries. (Burke and Brown, 2008).

The scPDSI also makes full use of precipitation and surface temperature, which are two climate variables with long historical records, allowing for past values of scPDSI to be easily calculated. As such, the National Center for Atmospheric Research (NCAR) maintains a global scale database of monthly scPDSI values. These values are only for land and have a global spatial resolution of 2.5 degrees, which is extremely coarse but represents an opportunity to observe and analyze drought on a global scale. Years of observable data ranged from 1850 to 2014.

2.2 TROPOSPHERIC EMISSIONS SPECTROMETER

Launched in 2004 as part of NASA’s Aura spacecraft, the Tropospheric Emission Spectrometer (TES) uses infrared Fourier transform spectroscopy to provide atmospheric gas profiles inferred through a nonlinear optimization procedure that determines the vertical distribution most likely to produce the radiative spectra observed by the TES. (Worden et al., 2012; Clough et al., 2006; Good et al., 2015) Large uncertainties exist within the retrieved TES data because the instrument is not optimized to measure gases in the lower troposphere, (Sutanto et al., 2015) therefore the dataset used was cleaned and bias-corrected as best as possible by Good et al. Despite these uncertainties, the data has been used in past research to successfully track and analyze weather patterns in the El Nino Southern Oscillation (ENSO) and Superstorm Sandy. (Logan et al., 2008; Good et al., 2014), and represents a large global dataset of atmospheric deuterium values. Years of observable data ranged from 2004 to 2008

2.3 WATERISOTOPES.ORG

Operated by Dr. Gabriel Bowen from the University of Utah Department of Geology and Geophysics, and supported by the University of Utah Center for High-Performance Computing, waterisotopes.org is a database of water isotopes that seeks to compile isotope data published by government agencies and from data published in scientific articles to be used for non-commercial research and educational activities. (waterisotopes.org) As a result, the data resources provided by waterisotopes.org allows for access to global datasets of water isotopes across several areas, including groundwater, tap water, and precipitation. For the purpose of this research isotope data was broken into two categories, precipitation which ranged from 1974 to 2014, and surface water (lakes, rivers, streams) which ranged from 1977 to 2014. (Water Isotopes Database, 2018)

2.4 CLEANING

All of the datasets used were inspected and cleaned. Cleaning included the removal of recorded isotope values that represented an error within the system (values of -999) and the removal of precipitation and surface water data that didn't include elevation data. Each of these cleaned datasets was then projected into a 2.5 degree global map, and cross-referenced with the NCAR scPDSI dataset with time-lag values of -1 month, 0 lag, +1 month, +2 months, +3 months, and +6 months. After this, any data that did not have any recorded scPDSI values was removed because it could not be used in this analysis. The end result was 280,228 points of atmospheric data, 12,020 points of surface water data, and 10540 points of precipitation data. Each point was then projected into a global map of sampling sites which can be seen in Figure 3.

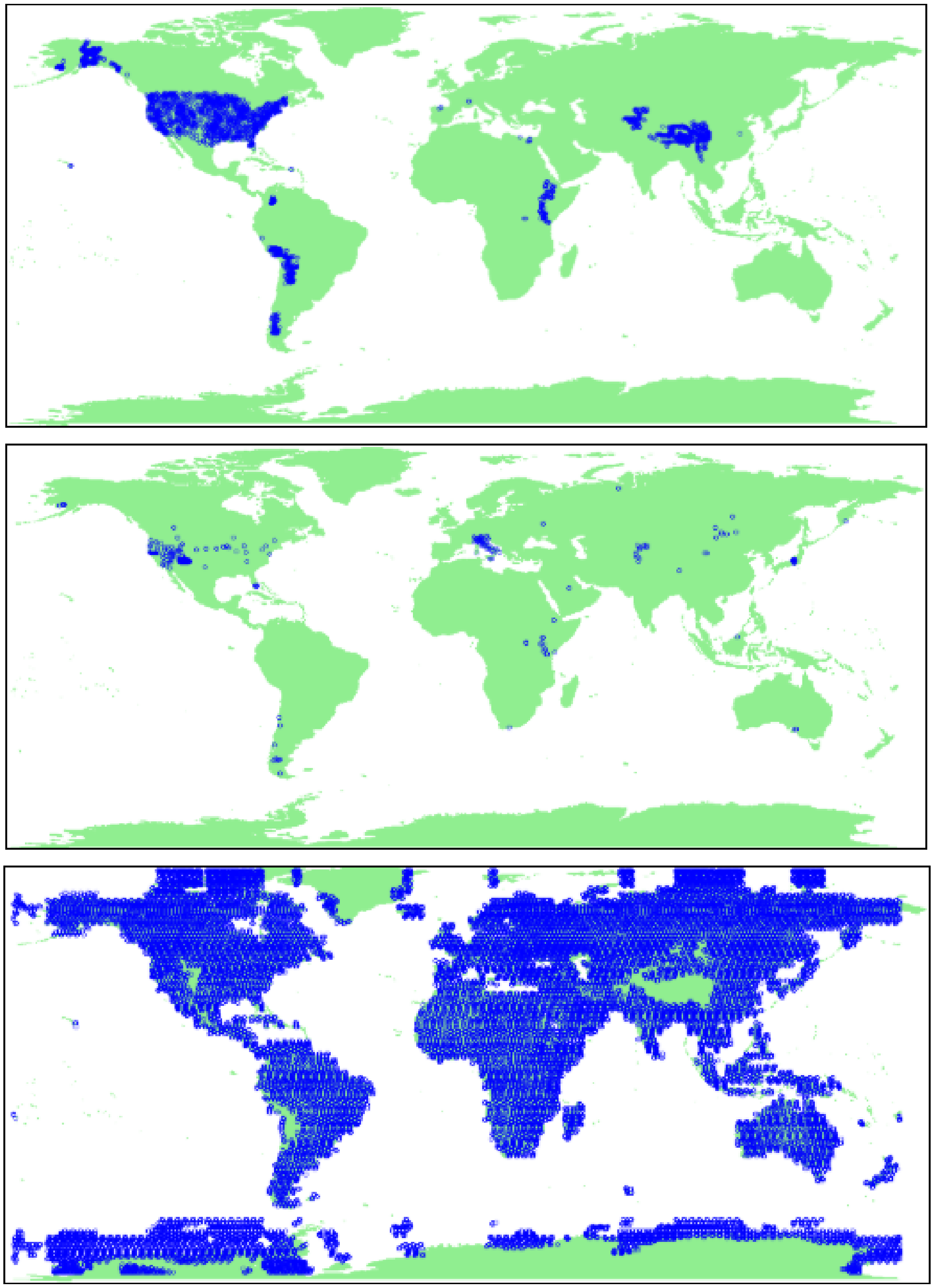

Fig. 3 Stable water isotope sampling locations. From top: Surface Water, Precipitation, TES Satelite. Each dot represents one sample taken from that location during the respective monitoring periods.

2.5 ANALYSIS

The analysis was completed by creating a null model which input readily available physical data and was used to predict scPDSI values, as shown in Equation 1, and comparing it to a full model which included both physical and stable isotope data as is shown in Equation 2.

$$f(x, y, z, t, t_{sine}, I_{-1}) = I_i$$

Eq. 1. Null model. x = Latitude, y = Longitude, z = elevation, t = DOY, t_{sine} = sDOY,
I_{-1}=scPDSI time lag of -1, I_i= calculated scPDSI value for a time lag of +0, +1,+2, +3, or +6

$$f(x, y, z, t, t_{sine}, I_{-1}, \delta^2H, \delta^{18}O, d_{excess}) = I_i$$

Eq. 2. Full model. x = Latitude, y = Longitude, z = elevation, t = DOY, t_{sine} = sDOY,
I_{-1}=scPDSI time lag of -1, δ^2H = Deuterium Isotope Data, δ^{18}O =Oxygen-18 Isotope Data,
d_{excess}=deuterium excess data,
I_i= calculated scPDSI value for a time lag of +0, +1,+2, +3, or +6

Physical data for the null model was defined as latitude, longitude, elevation (except for atmospheric data which did not report elevation), the day of the year (DOY) a sample was taken, the sine of the DOY (sDOY), and scPDSI with a time lag of -1 month. Latitude, longitude, and elevation were chosen because they are easily measurable and account for the specific location that a sample was taken. DOY and sDOY were used to account for the impact of the seasonality of a location, with the assumption that days that corresponded to the summer months of a region would have a greater likelihood of experiencing drought. scPDSI with a time lag of -1 month was used because as has already been stated, the PDSI model is highly autocorrelated, with the assumption that precedent drought conditions would inform current and future drought conditions.

Both null and full models were regressed using linear regression, Support-Vector Machine (SVM) regression, Ridge Regression, and Least Absolute Shrinkage and Selection Operator (LASSO) regression. Linear regression was chosen because it is one of the simplest and most widely-used statistical technique for predictive modeling. SVM, Ridge, and LASSO regression were chosen as machine-learning alternatives that were able to take into account non-linear regression, each with different strengths and weaknesses. (Freedman, 2012; Statnikov et al., 2006; Ng, 2004) While all three machines can be tuned to become more accurate, the computational requirements to tune an SVM are very large, and so results were taken from that machine without tuning it. In addition, due to the computation requirement of an untuned SVM, the full atmospheric dataset was unable to be analyzed as the time to analysis for SVM increases cubically with the number of data points given. Instead, a random selection of 5% of the atmospheric dataset was selected and run through the SVM. 5% was selected as a value because it resulted in an analysis of 14,011 data points which is similar in size to the precipitation and surface water datasets. For both Ridge and LASSO regression, the machines were cross-validated and trained using 75% of the available data, in order to test their predictive capabilities on the remaining 25% of the data.

Both models were analyzed using a root-mean-squared-error (RMSE) function. The RMSE function was chosen as the analysis metric for three reasons. First is simply that RMSE is widely used as a statistical metric for model performance in meteorology and climate research studies. (Chai and Draxler, 2014) Second is that the main goal of this thesis is to determine the usefulness of using stable water isotopes to improve drought prediction. Comparison of RMSE values allows us to determine the improvement in accuracy between null and full models without the confusion that can be introduced using a less definitive metric such as the coefficient of determination. Third, is that the scPDSI index varies significantly with relatively small changes in values, going from moderate to the extreme with only a difference of two. Being able to understand the mean error that occurs within the predictive models informs us of the usefulness of its application to drought prediction.

3.RESULTS

Results were reported in three ways. First, a comparison of the null and full model to the observed data. Second, comparison of the error to the latitude and the scPDSI value being predicted and the RMSE of models as they were projected into the future. Third, a comparison of the percent improvement to the latitude and the scPDSI value being predicted and the percent improvement in RMSE of models being projected into the future, with positive values signaling a decrease in RMSE values.

3.1 MODEL VS. OBSERVED

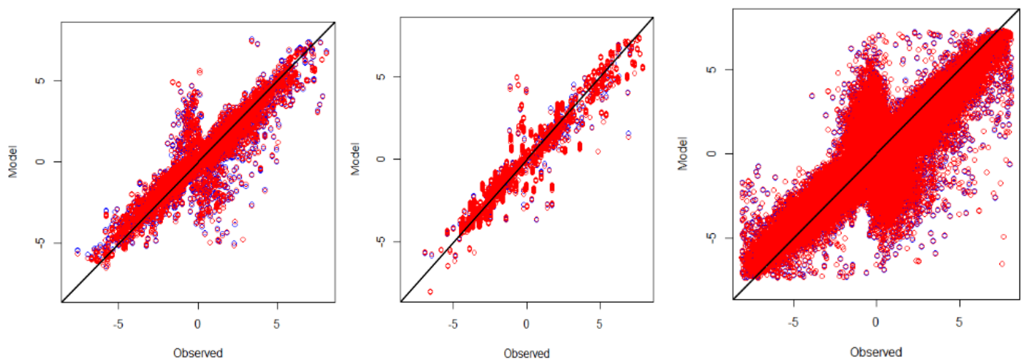

Fig. 4 Crossplot of model predictions of scPDSI vs. Observed values across surface water (left), precipitation (center), and atmospheric (right) data. Models represent linear regression at a time lag of zero, with the null model represented by blue dots, and the full model represented by red dots. Similar relationships were observed for other types of regression and other time-lags. (not shown)

3.2 ERROR

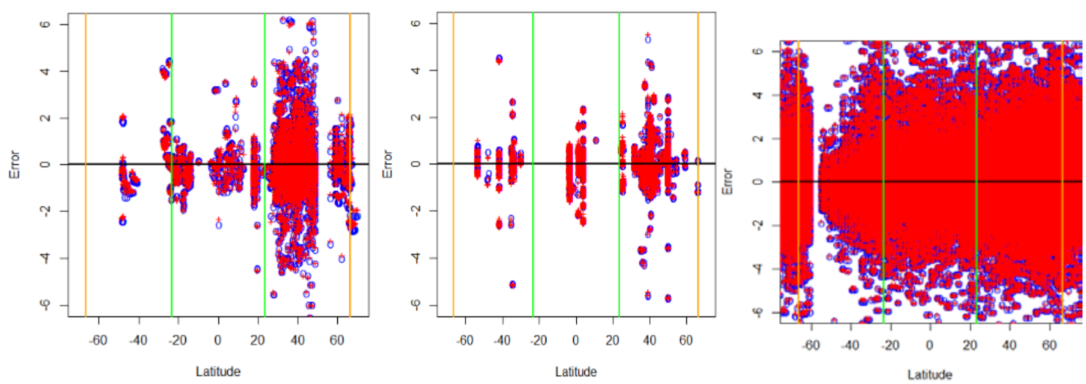

Fig. 5 Crossplot of individual prediction error vs. latitude values across surface water (left), precipitation (center), and atmospheric (right) data. Models represent linear regression at a time lag of zero with the null model represented by blue dots, and the full model represented by red dots. Green bars denote the Tropics of Capricorn and Cancer at +/- 23.5 degrees. Orange bars denote the Arctic and Antarctic Circles at +/- 66.5 degrees. Similar relationships were observed for other types of regression and other time-lags. (not shown)

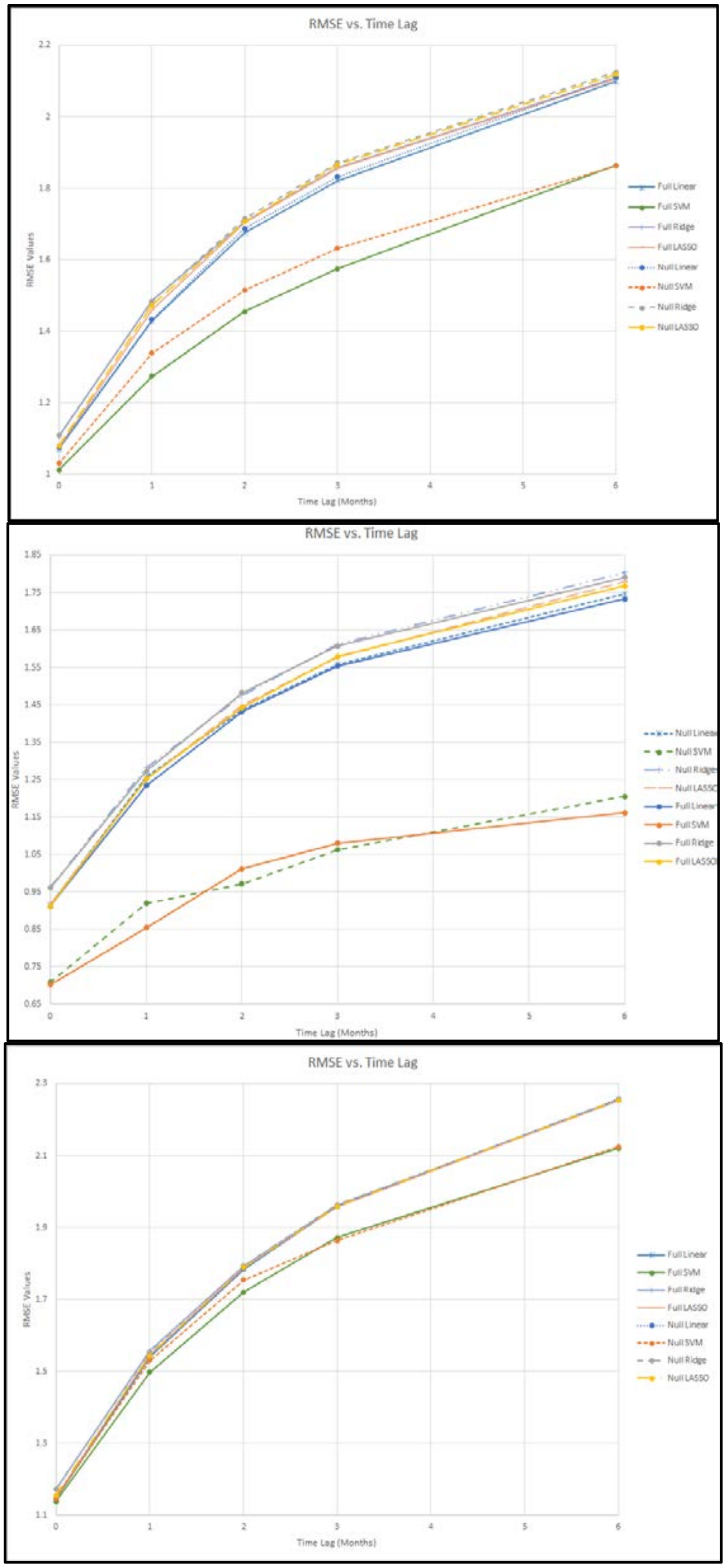

Fig. 6 Crossplot of RMSE Values vs. Time Lag in months showing both null and full models across surface water (top), precipitation (center), and atmospheric data (bottom)

3.3 IMPROVEMENT

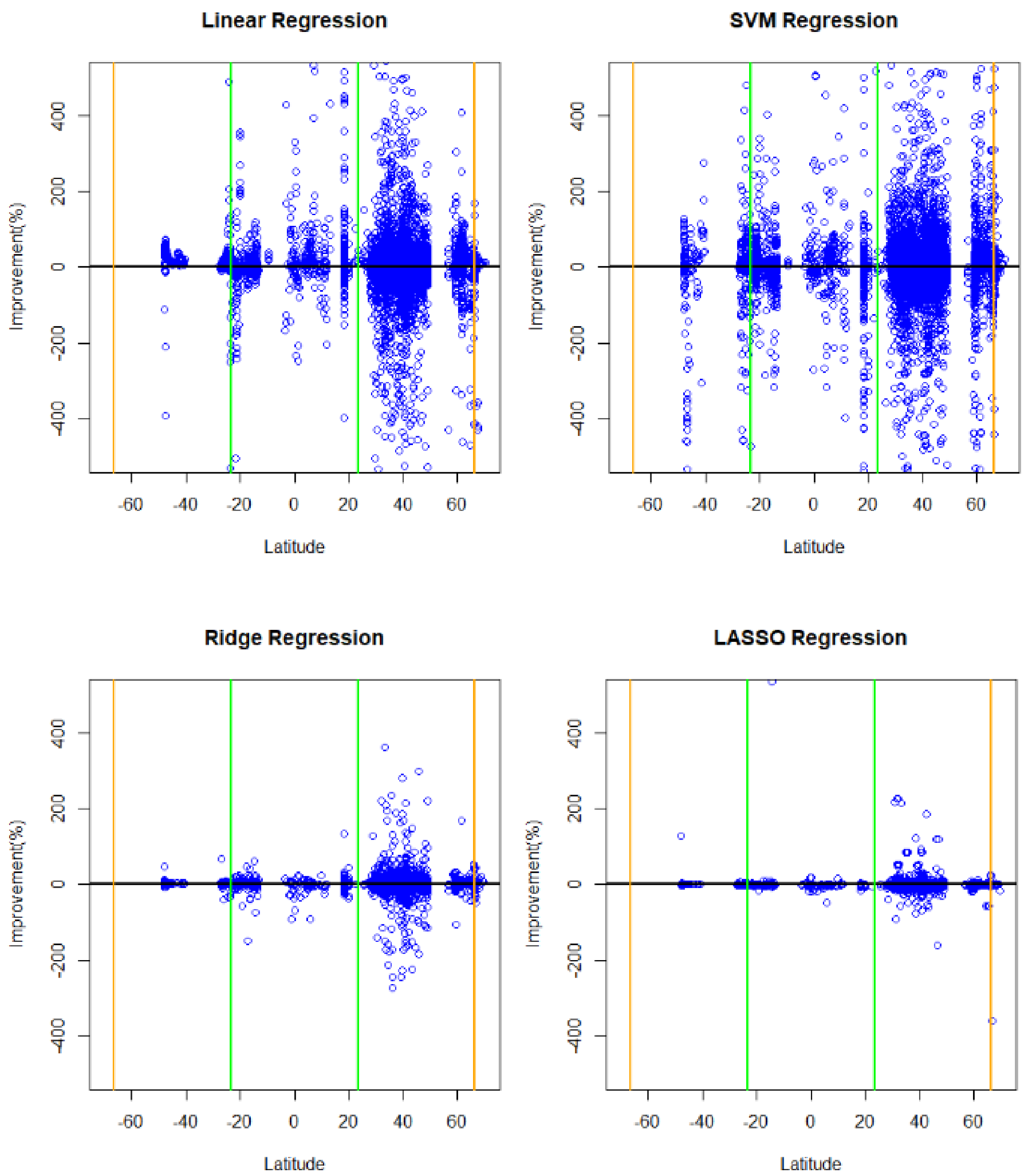

Figure 7. Crossplot of the percent improvement vs. Latitude for Surface Water Data across four types of regression at a time lag of zero. Green bars denote the Tropics of Capricorn and Cancer at +/- 23.5 degrees. Orange bars denote the Arctic and Antarctic Circles at +/- 66.5 degrees. Similar relationships were observed for other time-lags. (not shown)

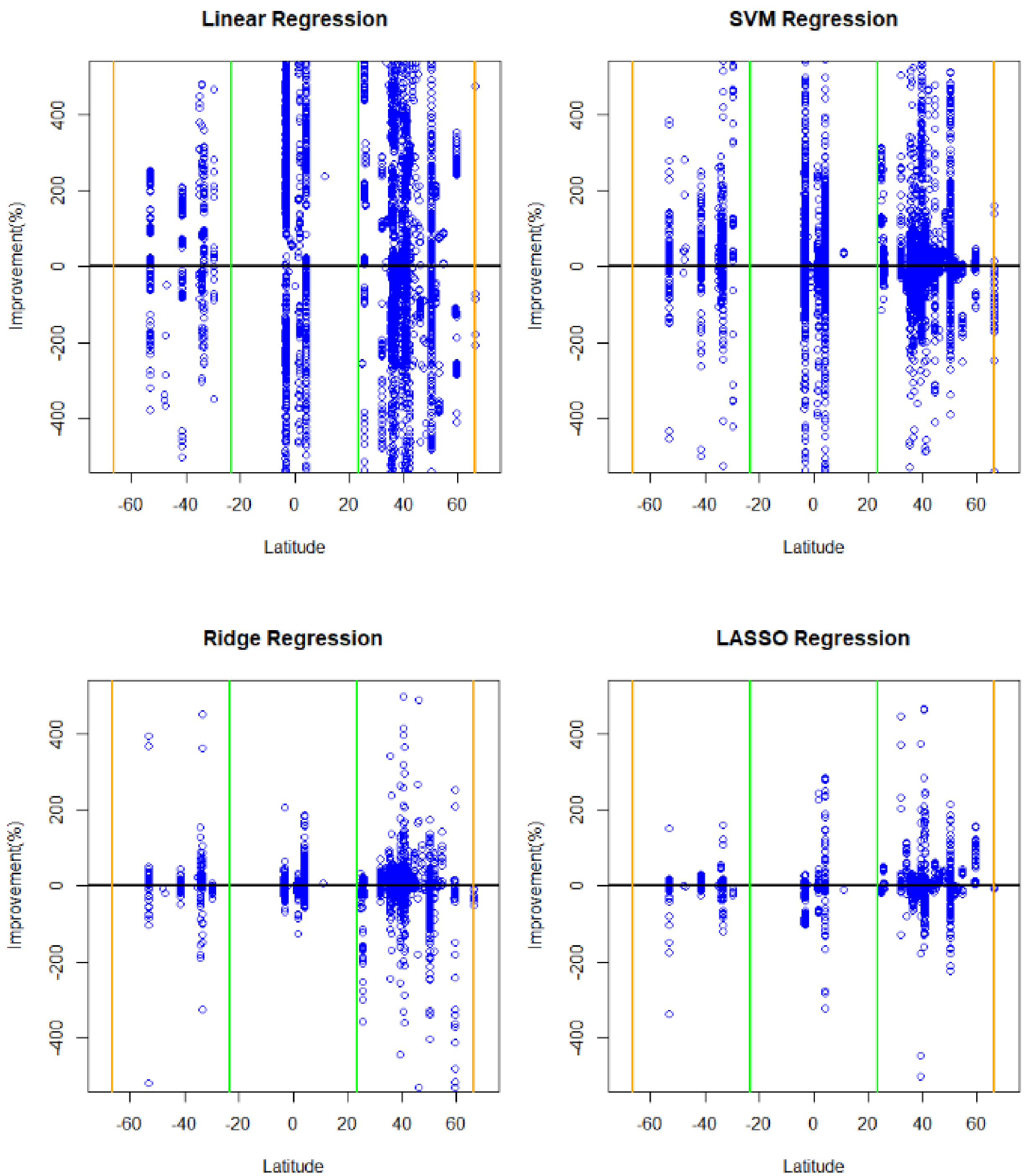

Figure 8. Crossplot of the percent improvement vs. Latitude for Precipitation Data across four types of regression at a time lag of zero. Green bars denote the Tropics of Capricorn and Cancer at +/- 23.5 degrees. Orange bars denote the Arctic and Antarctic Circles at +/- 66.5 degrees. Similar relationships were observed for other time-lags. (not shown)

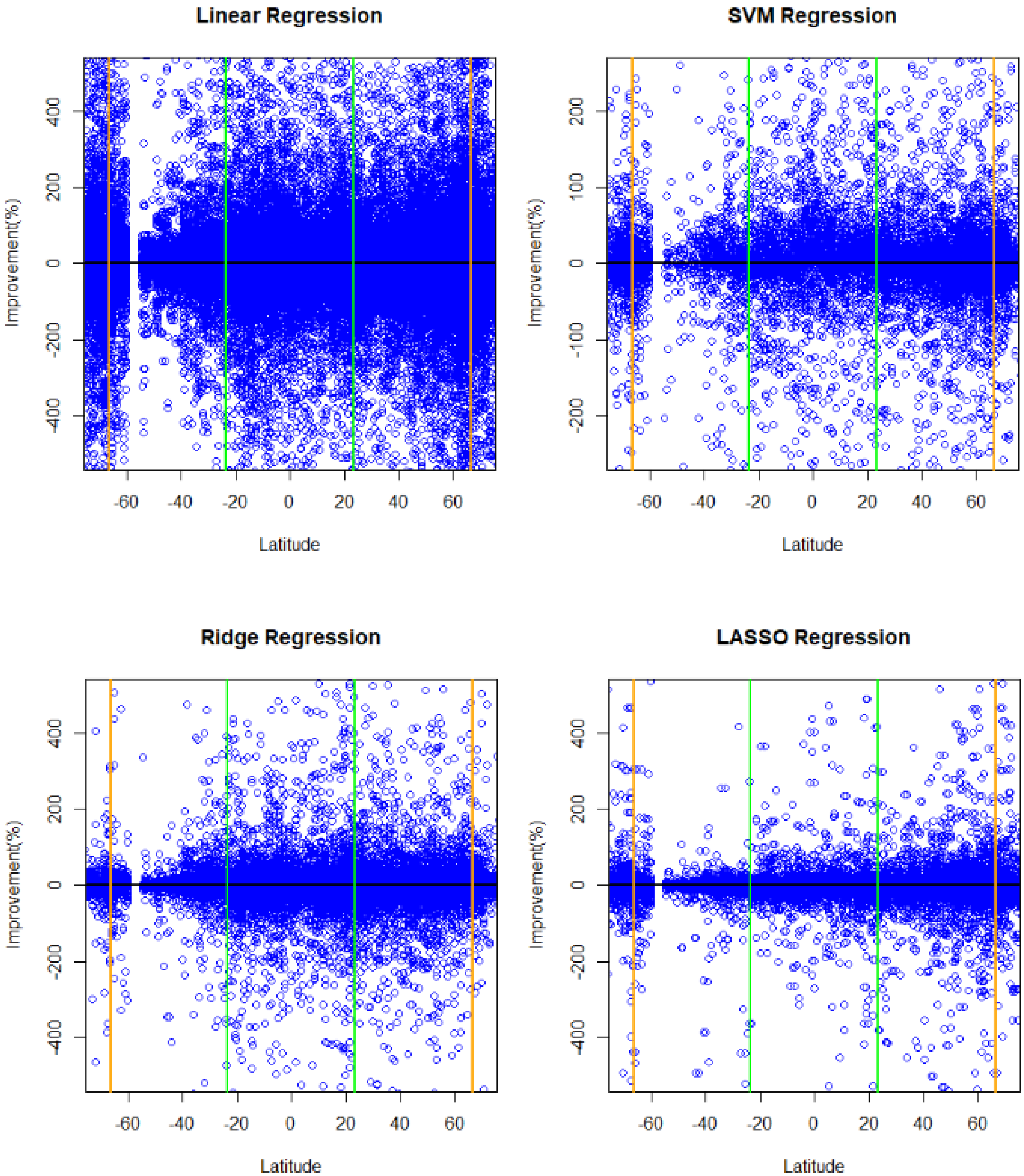

Figure 9. Crossplot of the percent improvement vs. Latitude for Atmospheric Data across four types of regression at a time lag of zero. Green bars denote the Tropics of Capricorn and Cancer at +/- 23.5 degrees. Orange bars denote the Arctic and Antarctic Circles at +/- 66.5 degrees. Similar relationships were observed for other time-lags. (not shown)

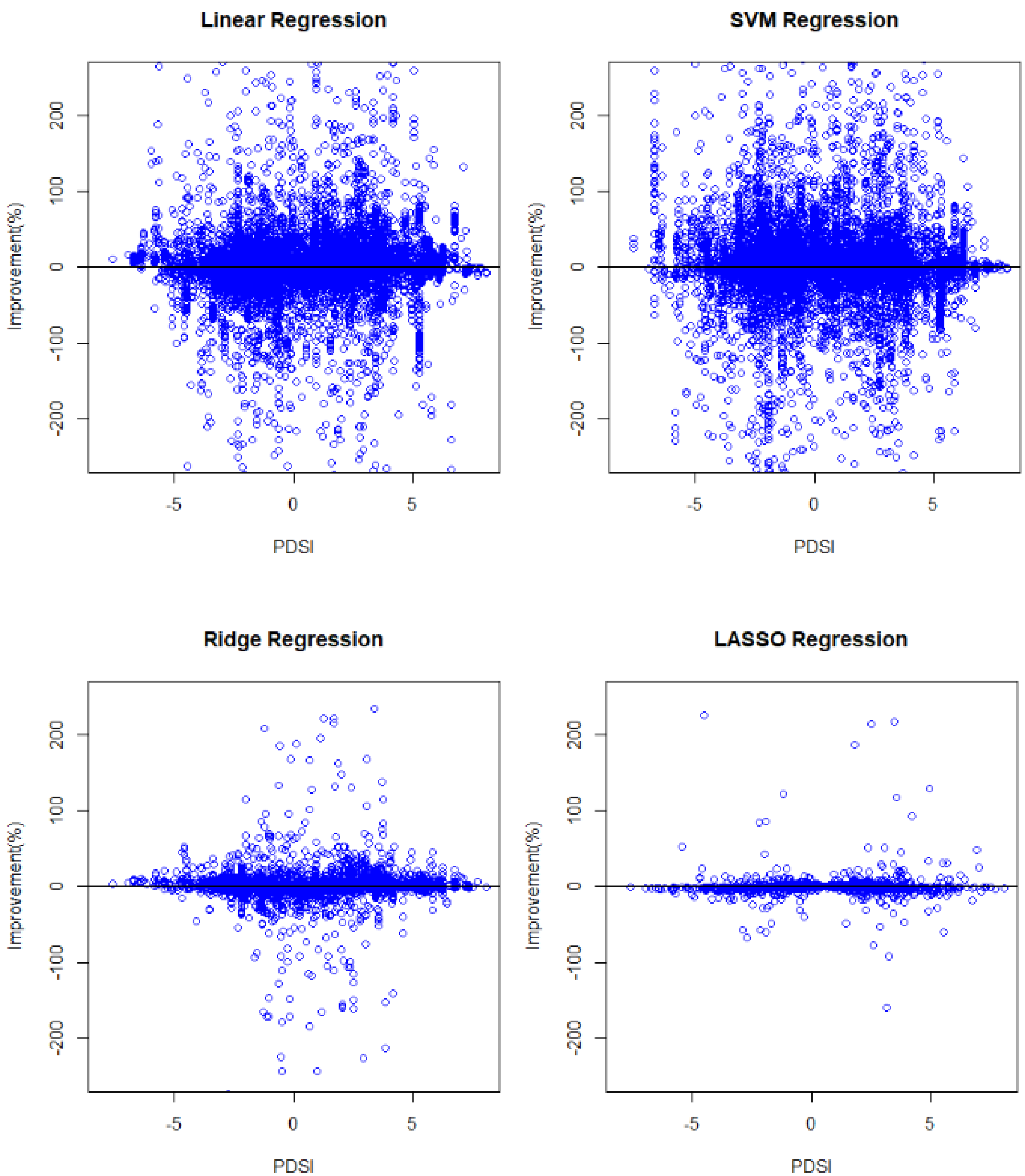

Figure 10. Crossplot of the percent improvement vs. PDSI for Surface Water Data across four types of regression at a time lag of zero. Similar relationships were observed for other time-lags. (not shown)

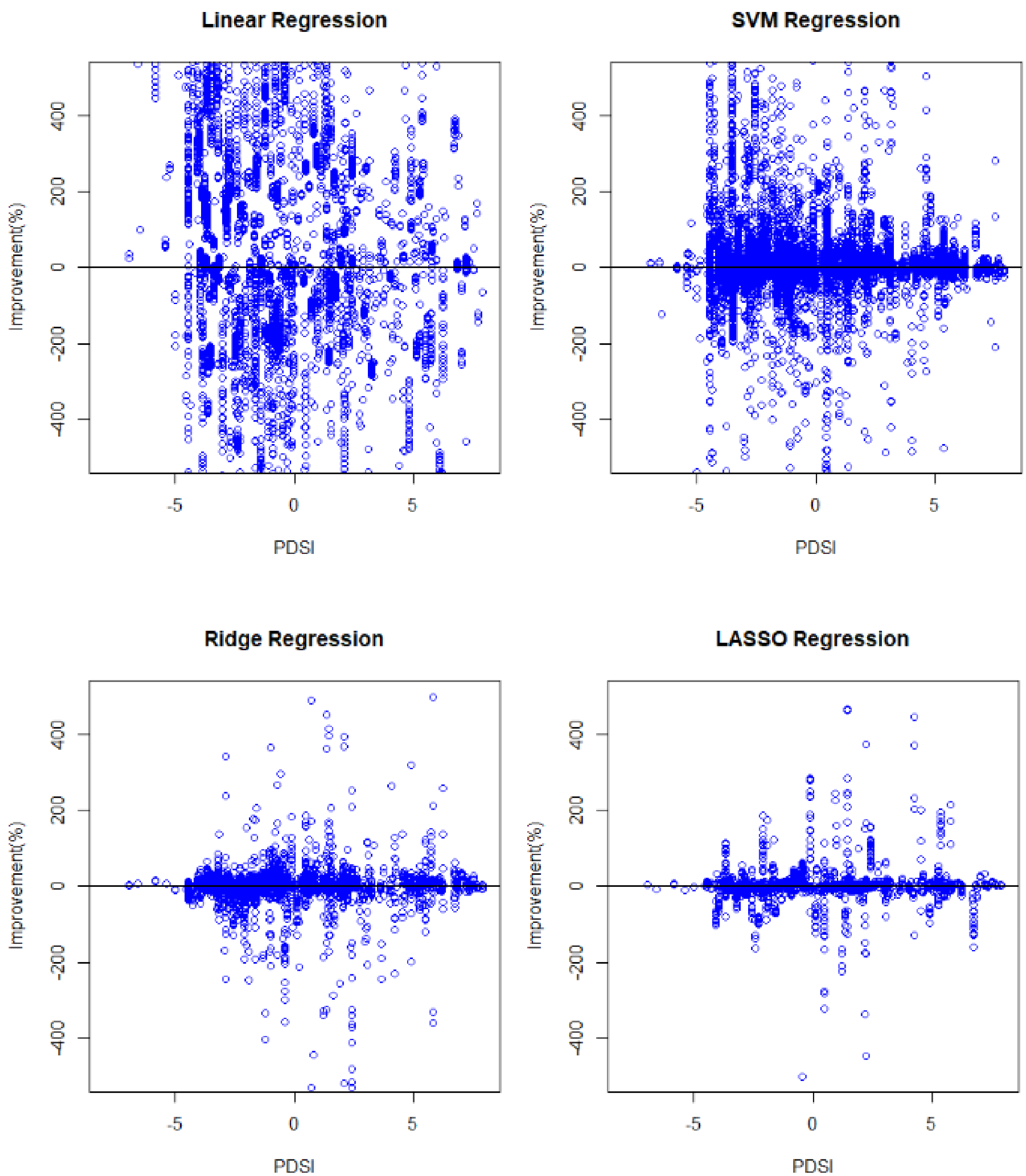

Figure 11. Crossplot of the percent improvement vs. PDSI for Precipitation Data across four types of regression at a time lag of zero. Similar relationships were observed for other time-lags. (not shown)

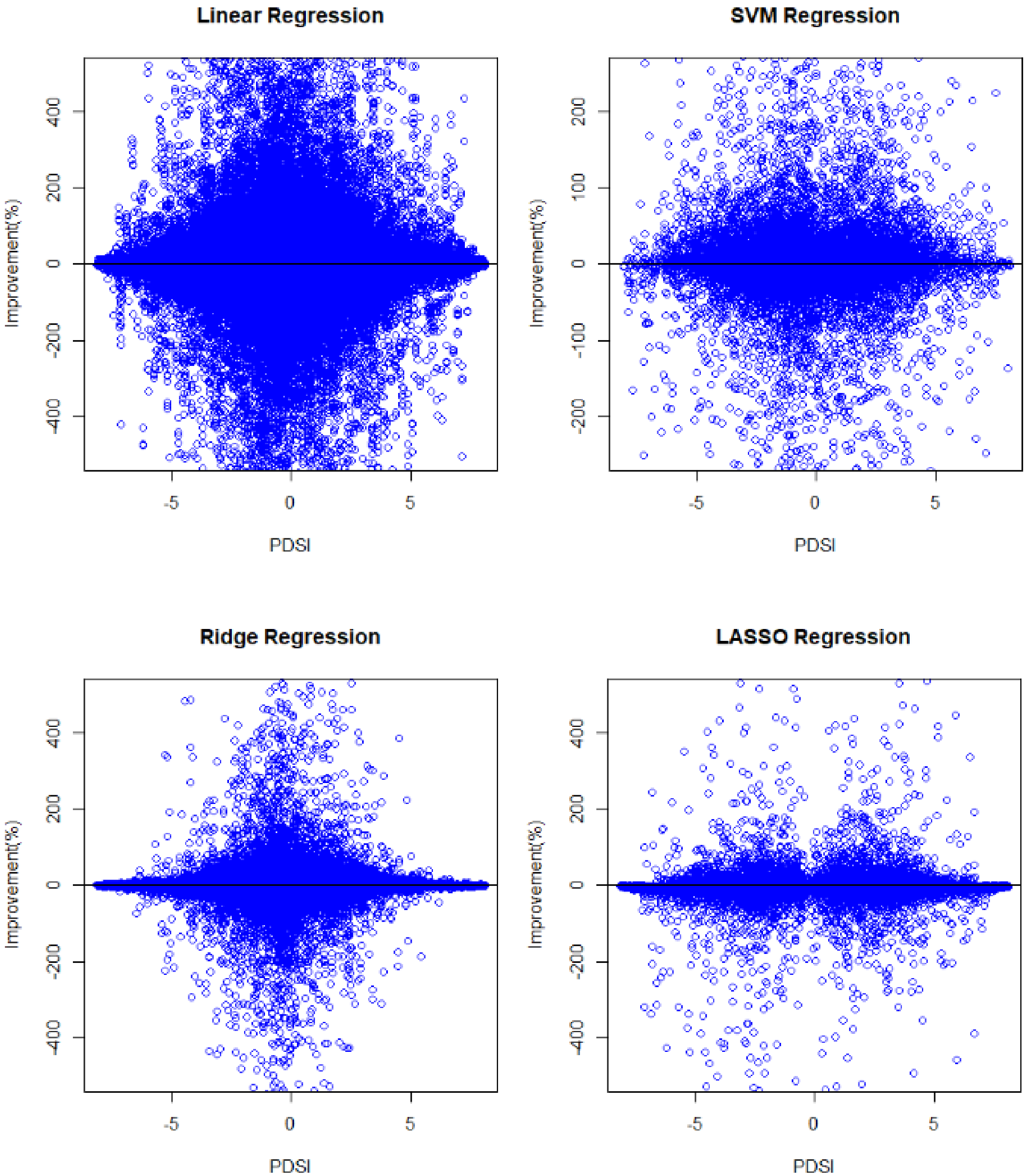

Figure 12. Crossplot of the percent improvement vs. PDSI for Atmospheric Data across four types of regression at a time lag of zero. Similar relationships were observed for other time-lags. (not shown)

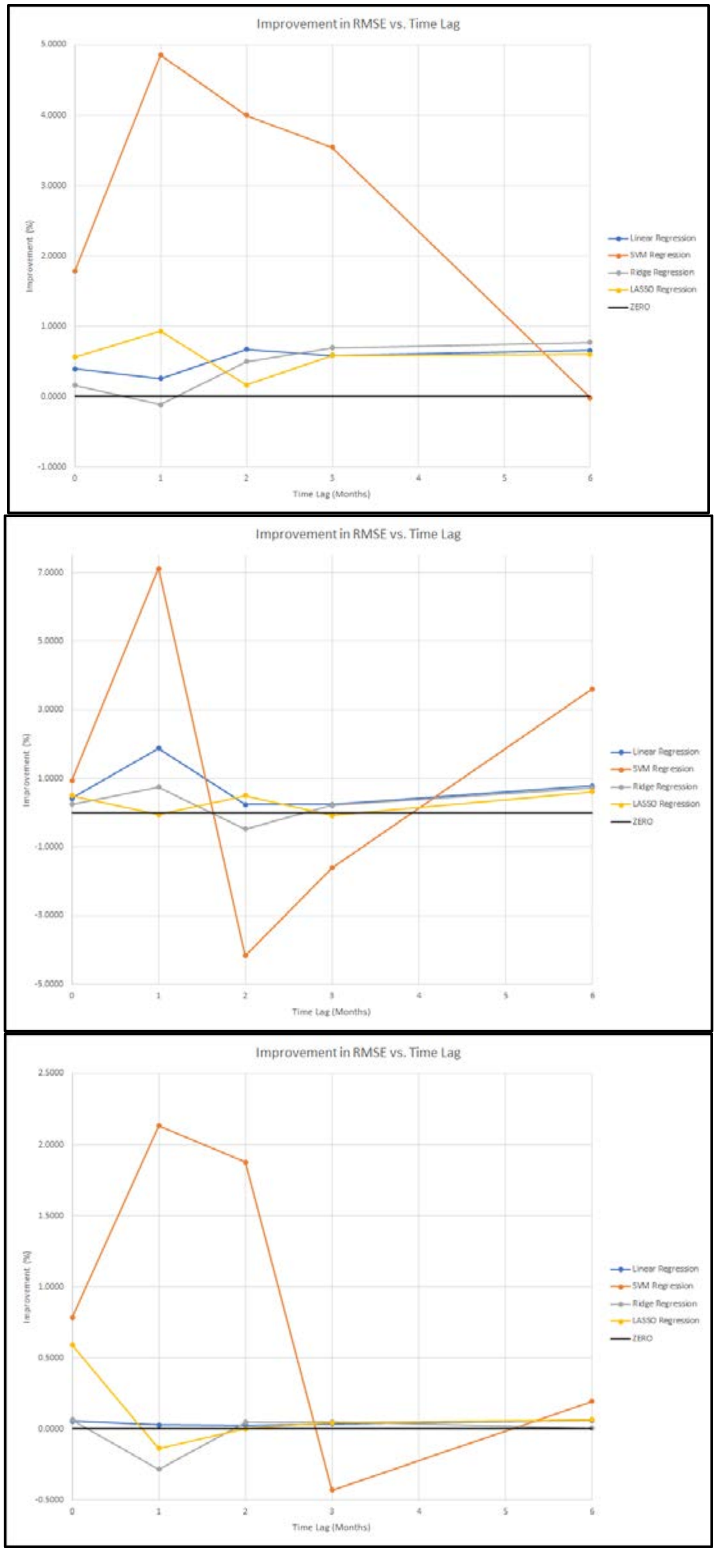

Fig. 13 Crossplot of percent improvement of RMSE Values vs. Time Lag in months across four types of linear regression across surface water (top), precipitation (center), and atmospheric data (bottom)

4.DISCUSSION

4.1 INTERPRETATION OF RESULTS

The Model vs. Observed crossplot within Figure 4 seems to indicate a generally good fit to the line across all three hydrologic regions, with deviation occurring in both the null and full model as scPDSI values approach zero. Figure 5 shows that error seems to act as a function of the number of observations, but does not seem to have any special relationship to the latitude of the sample site. Figure 6 confirms that the RMSE values of both the null and full models increase with an increase in the time-lag. This is what was expected as the farther into the future you go, the more unpredictable the variables become. Figure 6 also shows that RMSE values range from roughly 0.7 for the full SVM model of precipitation at a time lag of zero, up to 2.3 for the full linear regression model of atmospheric data at a time lag of plus six. This range indicates a wide variance that shows that this particular predictive modeling may be useful for drought management, but must be used with caution when applied using the scPDSI model due to the wide variance in drought conditions over a short span of index values.

Figures 7-9 show a similar trend as Figure 5 in that it seems that the percent improvement as a function of latitude seems to be more impacted by the number of samples taken than the latitude value. However, these figures also show that the percent improvement is greatest for linear regression, and decreases for SVM and Ridge regression, and is the smallest for LASSO regression. This is most likely because linear regression relies more heavily upon the input variables than any of the machine learning regression models, and it indicates that the LASSO model’s method of regularization is less impacted by the addition of further variables as it already seeks to perform variable selection within the null model.

Figures 10 and 12 show similar trends in the relation of improvement to the scPDSI value being predicted, with a diamond shape forming of minimum improvement of scPDSI values greater than +/- 5, and maximum improvement occurring at values near zero, except for LASSO regression, which again decreases in improvement as it approaches zero. This is interesting and indicates that the assumptions within the LASSO model are well-suited to predicting normal variation within the PDSI index. Figure 11 does not follow this pattern, and instead seems to show a greater tendency for improvement in values below zero (representing drought) which indicated that the relationship between stable water isotopes, evaporation, and precipitation described by Daansgaard in 1964 is strong and that stable water isotopes are only useful for precipitation in areas experiencing drought.

Figure 13 represents the answer to the key question of this analysis. Results seem to show that the addition of stable water isotopes can help to improve drought prediction, especially when using the SVM or linear regression model. This may indicate the the attempts at regularization found within the Ridge and LASSO models fails to take into account natural variation within meteorological systems, and so is not as well equipped as the SVM model to account for this complexity. Improvements are consistently largest at the +1 time lag, which also indicates that by the time an isotope sample is taken, it is more indicative of upcoming future values than it is of current values. Surface water also shows continuous improvement with the addition of stable water isotopes, whereas both precipitation and atmospheric water show considerable variability in whether or not the addition of stable water isotopes is useful. This may indicate that the collection processes that surface water undergo also act as a buffer, with the collection and mixing of isotopes which are taken at a single point representing the entire catchment basin above the collection point. These processes may also act as a time buffer, as it may take time for the precipitation and runoff collected in surface waters to reach the sample location. Because the scPDSI is at such a coarse resolution, and because it is meant to be autocorrelated over a longer time scale, these two things would allow for the stable water isotopes within surface water to be a better representative of scPDSI than the more instantaneous values taken for precipitation and atmospheric data.

4.2 SOURCES OF ERROR AND FUTURE RESEARCH

The two greatest weaknesses of this analysis are coarseness of the scPDSI dataset and the location bias within the surface and precipitation isotope data. At a resolution of 2.5 degrees (or roughly 173 miles), each cell within the NCAR scPDSI dataset essentially represents 29,929 square miles of land, with a single scPDSI value representing all of them, despite what we know about the impact of local topography and dynamics on drought. Additionally, having to project individual samples from specific locations into that 2.5-degree world map reduces some of the ability of that data to represent specific locations. Furthermore, as shown by Figure 3, there is a definitive sampling bias towards the United States and the Western Hemisphere. While this represents the first global analysis of both drought and stable water isotopes, it also highlights the need to improve isotopic collection in underrepresented areas if future analysis is to be made. Future research should try to investigate scPDSI on a smaller scale, which will allow for the impacts of the internal dynamics of a region to have a greater effect, and should investigate what regional differences exist in drought prediction and the use of stable water isotopes.

Of the regression models tested, SVM also appears to be the most useful in finding the smallest RMSE value and shows the greatest improvement with the addition of stable water isotope data. Future research should look into the impact of tuning this SVM model, and see whether or not it produces a significantly more accurate model.

4.3 APPLICATION TO SHARED GOVERNANCE AND INTERNATIONAL AID

While there is some variance within the improvement provided by the addition of stable water isotopes to a drought prediction model, the improvements that do exist indicate one example of how the use of stable water isotopes can be used to improve shared water governance and the administration of international aid and emergency management.

A key factor to any shared governance agreement is communication between stakeholders and agreeing upon burdens which can be shared equitably. (Hall et al., 2012) Research into drought propagation has also shown that monitoring is best performed at the basin-scale and has to take into account non-local hydrologic variables. (Oertel et al., 2018) Monitoring stable water isotopes can help with this in two ways. First, the maintenance of stable water isotope collection and analysis represents a regular, relatively low-cost management action which will ensure regular communication between stakeholders and allows each party to make a valuable contribution towards the management of water resources. Second, because the addition of stable water isotopes have been shown to be especially useful in improving drought prediction one month out, keeping a record of stable water isotopes can alert stakeholders of an impending change in scPDSI for a specific region with enough time for them to communicate and form a plan to address whatever the issue may be.

This also applies to applications for international aid, and emergency management and response.
Drought response and recovery typically takes place in stages of activity dictated by when a drought is confirmed to occur, and often take months to implement fully. (EPA, 2016) While more time is always useful, even one month represents time which can be put towards logistic planning that can help to build resiliency among the impacted communities and alleviate the impact of drought or flooding.

5. CONCLUSION

Stable water isotopes were shown to be useful at improving drought prediction, especially at a time lag of +1 month, and when the isotopes were collected from surface water. The level of improvement varied greatly, with some models showing that it actually made it worse, but overall indicated that using stable water isotopes may assist in addressing the complexity of trying to predict drought. Future research needs to address the lack of geographically diverse stable water isotope data in precipitation and surface water and should seek to compare scPDSI values on a smaller scale that can better take into account the impact of internal dynamics within a region of interest.

REFERENCES

Alexander, R., & Nugent, C. (2018). Cultural Responses to the Dust Bowl. *The Southwest Respiratory and Critical Care Chronicles,6*(22), 53. doi:10.12746/swrccc.v6i22.433

Bergier, I., Assine, M. L., Mcglue, M. M., Alho, C. J., Silva, A., Guerreiro, R. L., & Carvalho, J. C. (2018). Amazon rainforest modulation of water security in the Pantanal wetland. *Science of The Total Environment,619-620*, 1116-1125. doi:10.1016/j.scitotenv.2017.11.163

Burke, E. J., & Brown, S. J. (2008). Evaluating Uncertainties in the Projection of Future Drought. *Journal of Hydrometeorology,9*(2), 292-299. doi:10.1175/2007jhm929.1

Bush, M. B., Hanselman, J. A., & Gosling, W. D. (2010). Nonlinear climate change and Andean feedbacks: An imminent turning point? *Global Change Biology,16*(12), 3223-3232. doi:10.1111/j.1365-2486.2010.02203.x

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development,7*(3), 1247-1250. doi:10.5194/gmd-7-1247-2014

Clough, S., Shephard, M., Worden, J., Brown, P., Worden, H., Luo, M., . . . Mlawer, E. (2006). Forward model and Jacobians for Tropospheric Emission Spectrometer retrievals. *IEEE Transactions on Geoscience and Remote Sensing,44*(5), 1308-1323. doi:10.1109/tgrs.2005.860986

Cook, G. A., & Lauer, C. M. (1968). *Oxygen*(Clifford A. Hampel ed., Vol. The Encyclopedia of the Chemical Elements.). New York, NY: Reinhold Book Corporation.

Cordova, C., & Porter, J. C. (2015). The 1930s Dust Bowl: Geoarchaeological lessons from a 20th century environmental crisis. *The Holocene,25*(10), 1707-1720. doi:10.1177/0959683615594239

Craig, H., & Gordon, L. (1965). *Isotopic oceanography: Deuterium and oxygen 18 variations in the ocean and the marine atmosphere*. San Diego: Department of Earth Sciences, Scripps Institution of Oceanography, University of California.

Crausbay, S. D., Ramirez, A. R., Carter, S. L., Cross, M. S., Hall, K. R., Bathke, D. J., . . . Sanford, T. (2017). Defining Ecological Drought for the Twenty-First Century. *Bulletin of the American Meteorological Society,98*(12), 2543-2550. doi:10.1175/bams-d-16-0292.1

Criss, R. E. (1999). *Principles of stable isotope distribution*. New York, NY: Oxford Univ. Press.

Dai, A. (2011). Characteristics and trends in various forms of the Palmer Drought Severity Index during 1900–2008. *Journal of Geophysical Research,116*(D12). doi:10.1029/2010jd015541

Eisenhauer, N., Cesarz, S., Koller, R., Worm, K., & Reich, P. B. (2011). Global change belowground: Impacts of elevated CO2, nitrogen, and summer drought on soil food webs and biodiversity. *Global Change Biology,18*(2), 435-447. doi:10.1111/j.1365-2486.2011.02555.x

EPA, Water. (2016). *DROUGHT RESPONSE AND RECOVERY A Basic Guide for Water Utilities*.

Freedman, D. A. (2012). *Statistical models: Theory and practice*. New York: Cambridge University Press.

Good, S. P., Mallia, D. V., Lin, J. C., & Bowen, G. J. (2014). Stable Isotope Analysis of Precipitation Samples Obtained via Crowdsourcing Reveals the Spatiotemporal Evolution of Superstorm Sandy. *PLoS ONE,9*(3). doi:10.1371/journal.pone.0091117

Good, S. P., Noone, D., Kurita, N., Benetti, M., & Bowen, G. J. (2015). D/H isotope ratios in the global hydrologic cycle. *Geophysical Research Letters,42*(12), 5042-5050. doi:10.1002/2015gl064117

Hall, D. M., Gilbertz, S. J., Horton, C. C., & Peterson, T. R. (2012). Culture as a means to contextualize policy. *Journal of Environmental Studies and Sciences,2*(3), 222-233. doi:10.1007/s13412-012-0077-9

Hobbins, M. T., Dai, A., Roderick, M. L., & Farquhar, G. D. (2008). Revisiting the parameterization of potential evaporation as a driver of long-term water balance trends. *Geophysical Research Letters,35*(12). doi:10.1029/2008gl033840

Hornbeck, R. (2012). The Enduring Impact of the American Dust Bowl: Short- and Long-Run Adjustments to Environmental Catastrophe. *American Economic Review,102*(4), 1477-1507. doi:10.1257/aer.102.4.1477

Li, Y., Ye, W., Wang, M., & Yan, X. (2009). Climate change and drought: A risk assessment of crop-yield impacts. *Climate Research,39*, 31-46. doi:10.3354/cr00797

Logan, J. A., Megretskaia, I., Nassar, R., Murray, L. T., Zhang, L., Bowman, K. W., . . . Luo, M. (2008). Effects of the 2006 El Niño on tropospheric composition as revealed by data from the Tropospheric Emission Spectrometer (TES). *Geophysical Research Letters,35*(3). doi:10.1029/2007gl031698

Marchina, C., Natali, C., & Bianchini, G. (2019). The Po River Water Isotopes during the Drought Condition of the Year 2017. *Water,11*(1), 150. doi:10.3390/w11010150

Markewitz, D., Devine, S., Davidson, E. A., Brando, P., & Nepstad, D. C. (2010). Soil moisture depletion under simulated drought in the Amazon: Impacts on deep root uptake. *New Phytologist,187*(3), 592-607. doi:10.1111/j.1469-8137.2010.03391.x

Miessler, G. L., Fischer, P. J., & Tarr, D. A. (2014). *Inorganic chemistry*. Boston: Pearson.

National Drought Mitigation Center. Predicting Drought. (2017). Retrieved May 14, 2019, from https://drought.unl.edu/Education/DroughtIn-depth/Predicting.aspx

Ng, A. Y. (2004). Feature selection, L1 vs. L2 regularization, and rotational invariance. In *Proceedings of the 21st International Conference on Machine Learning*. Banff, Canada.

Oertel, M., Meza, F., Gironás, J., Scott, C. A., Rojas, F., & Pineda-Pablos, N. (2018). Drought Propagation in Semi-Arid River Basins in Latin America: Lessons from Mexico to the Southern Cone. *Water,10*(11), 1564. doi:10.3390/w10111564

Oleary, D. (2012). The deeds to deuterium. *Nature Chemistry,4*(3), 236-236. doi:10.1038/nchem.1273

Palmer, W. M. (1964). *Meteorological Drought*(USA, USDC, Weaher Bureau). Washington, D.C.: Office of Climatology, US Weather Bureau.

Phillips, O., Aragão, L., Lewis, S., Fisher, J., Lloyd, J., López-González, G., . . . Torres-Lezama, A. (2009). Drought Sensitivity of the Amazon Rainforest. *Science, 323*(5919), new series, 1344-1347. Retrieved from http://www.jstor.org/stable/25471642

Porter, J. (2012). Lessons from the Dust Bowl: Human-Environment Education on the Great Plains. *Journal of Geography,111*(4), 127-136. doi:10.1080/00221341.2011.624629

Roosevelt, F. D. (2017, April 26). September 6, 1936: Fireside Chat 8: On Farmers and Laborers. Retrieved May 25, 2019, from https://millercenter.org/the-presidency/presidential-speeches/september-6-1936-fireside-chat-8-farmers-and-laborers University of Virginia Miller Center Archives

Schrier, G. V., Briffa, K. R., Osborn, T. J., & Cook, E. R. (2006). Summer moisture availability across North America. *Journal of Geophysical Research,111*(D11). doi:10.1029/2005jd006745

Staal, A., Dekker, S. C., Hirota, M., & Nes, E. H. (2015). Synergistic effects of drought and deforestation on the resilience of the south-eastern Amazon rainforest. *Ecological Complexity,22*, 65-75. doi:10.1016/j.ecocom.2015.01.003

Statnikov, A., Hardin, D., & Aliferis, C. (2006). Using SVM Weight-Based Methods to Identify Causally Relevant and Non-Causally Relevant Variables. *Twentieth Conference on Neural Information Processing Systems*.

Sternberg, T. (2012). Chinese drought, bread and the Arab Spring. *Applied Geography,34*, 519-524. doi:10.1016/j.apgeog.2012.02.004

Stoutenborough, J. W., & Vedlitz, A. (2014). Public Attitudes Toward Water Management and Drought in the United States. *Water Resources Management,28*(3), 697-714. doi:10.1007/s11269-013-0509-7

Sutanto, S. J., Hoffmann, G., Scheepmaker, R. A., Worden, J., Houweling, S., Yoshimura, K., . . . Röckmann, T. (2015). Global-scale remote sensing of water isotopologues in the troposphere: Representation of first-order isotope effects. *Atmospheric Measurement Techniques,8*(3), 999-1019. doi:10.5194/amt-8-999-2015

Trenberth, K. E., Dai, A., Schrier, G. V., Jones, P. D., Barichivich, J., Briffa, K. R., &
Sheffield, J. (2014). Global warming and changes in drought. Nature Climate Change,4(1), 17-22. doi:10.1038/nclimate2067

UNICEF. (2019, May 01). Drought disasters. Retrieved from https://www.unicef.org/drought/drought-countries.htm

Vanplantinga, A. A., Grossman, E. L., & Roark, E. B. (2016). Chemical and Isotopic Tracer Evaluation of Water Mixing and Evaporation in a Dammed Texas River During Drought. *River Research and Applications,33*(3), 450-460. doi:10.1002/rra.3080

Wells, N., Goddard, S., & Hayes, M. J. (2004). A Self-Calibrating Palmer Drought Severity Index. *Journal of Climate,17*(12), 2335-2351. doi:10.1175/1520-0442(2004)0172.0.co;2

Wilhite, D. A., & Glantz, M. H. (1985). Understanding: The Drought Phenomenon: The Role of Definitions. *Water International,10*(3), 111-120. doi:10.1080/02508068508686328

Worden, J., Kulawik, S., Frankenberg, C., Payne, V., Bowman, K., Cady-Peirara, K., . . . Noone, D. (2012). Profiles of CH₄, HDO, H₂O, and N₂O with improved lower tropospheric vertical resolution from Aura TES radiances. *Atmospheric Measurement Techniques,5*(2), 397-411. doi:10.5194/amt-5-397-2012

Worster, D. (2012). *Dust Bowl: The southern Plains in the 1930s*. New York: Oxford University Press.

Wu, H., Li, J., Song, F., Zhang, Y., Zhang, H., Zhang, C., & He, B. (2017). Spatial and temporal patterns of stable water isotopes along the Yangtze River during two drought years. *Hydrological Processes*,*32*(1), 4-16. doi:10.1002/hyp.11382