JOURNAL OF APPLIED AND SERVICE CLIMATOLOGY

Published Online September 18, 2024

Exploring Relationships Between Drought Indices and Ecological Drought Impacts Using Machine Learning and Explainable AI

Annie Britton[1], Garrett Graham[2], Molly Woloszyn[3]

[1] *Johns Hopkins University, Baltimore, Maryland*
[2] *North Carolina Institute for Climate Studies, Asheville, North Carolina*
[3] *NOAA/National Integrated Drought Information System, and Cooperative Institute for Research in the Environmental Sciences (CIRES), University of Colorado Boulder, Boulder, Colorado*

Corresponding Author: Annie Britton, annie.eliz.britton@gmail.com, abritto4@jhu.edu

Rangeland ecosystems in the United States hold great ecological, economic, and cultural value. However, the increasing frequency and severity of droughts pose potential threats to these ecosystems. This study used remotely sensed data, machine learning, and explainable artificial intelligence (XAI) to explore relationships between drought indices and vegetation health in the Cheyenne River Basin, USA. The study employed XGBoost Regressor and Extra Trees Regressor models in conjunction with SHapley Additive exPlanations (SHAP) to identify the most influential drought indices and environmental variables for predicting Normalized Difference Vegetation Index (NDVI), thereby uncovering indicators of vegetation stress in the basin. The XGBoost regressor was moderately successful at predicting NDVI, making the model suitable for subsequent XAI analysis using SHAP. SHAP results revealed that the Palmer Drought Severity Index (PDSI), the 90-day Standardized Precipitation Index (SPI), and snow water equivalent (SWE) were the most influential predictors of NDVI, indicating their strong association with changes in vegetation health in the Cheyenne River Basin. This study demonstrates the feasibility and value of applying XAI methods to investigate both the strength and directionality of ecological drought indicators—an approach that has been underutilized in drought research. These insights can inform future drought research, improve monitoring efforts, and help anticipate ecological drought impacts in the region.

1. Introduction

The frequency of ecological drought is anticipated to increase as a result of human water usage, a warming climate, and resulting shifts in oceanic and atmospheric processes (IPCC, 2021). As these trends intensify, frequent ecological droughts may drive ecosystem transformations and cause the deterioration of vital ecosystem services (Crausbay et al., 2017). When various ecosystems are driven beyond key thresholds, which ecosystem services may be impacted, and how to properly measure and predict these variables, are all questions whose answers vary by impact, ecosystem, and scale. Further complicating these questions is the fact that links between traditional drought indices and ecological drought impacts, such as declines in vegetation health, are not well established (Bradford et al., 2020; Crausbay et al., 2020; Wiens & Bachelet, 2010).

The Normalized Difference Vegetation Index (NDVI) is a valuable tool for monitoring drought conditions and their impacts on vegetation health (Anyamba & Tucker, 2012; Tucker et al., 1986). Studies have found that declining NDVI values are indicative of reduced vegetation activity and productivity during drought events (Park et al., 2016; Pettorelli et al., 2005; Phillips et al., 2008). NDVI is therefore a valuable proxy for measuring the impacts of ecological drought on vegetation and has been successfully predicted in the past using ML methods (Li et al., 2021; Roy, 2021). However, NDVI lacks standardization based on historical data or baselines, making it less useful in indicating drought conditions compared to standardized drought indices like the Standardized Precipitation Index (SPI). Ecosystem-specific nuances further complicate the interpretation of NDVI, necessitating an understanding of the relationship between drought indicators and NDVI to accurately assess drought magnitude and its impact on an ecosystem. Identifying drought indices that correlate effectively with NDVI enables a more consistent and reliable as-

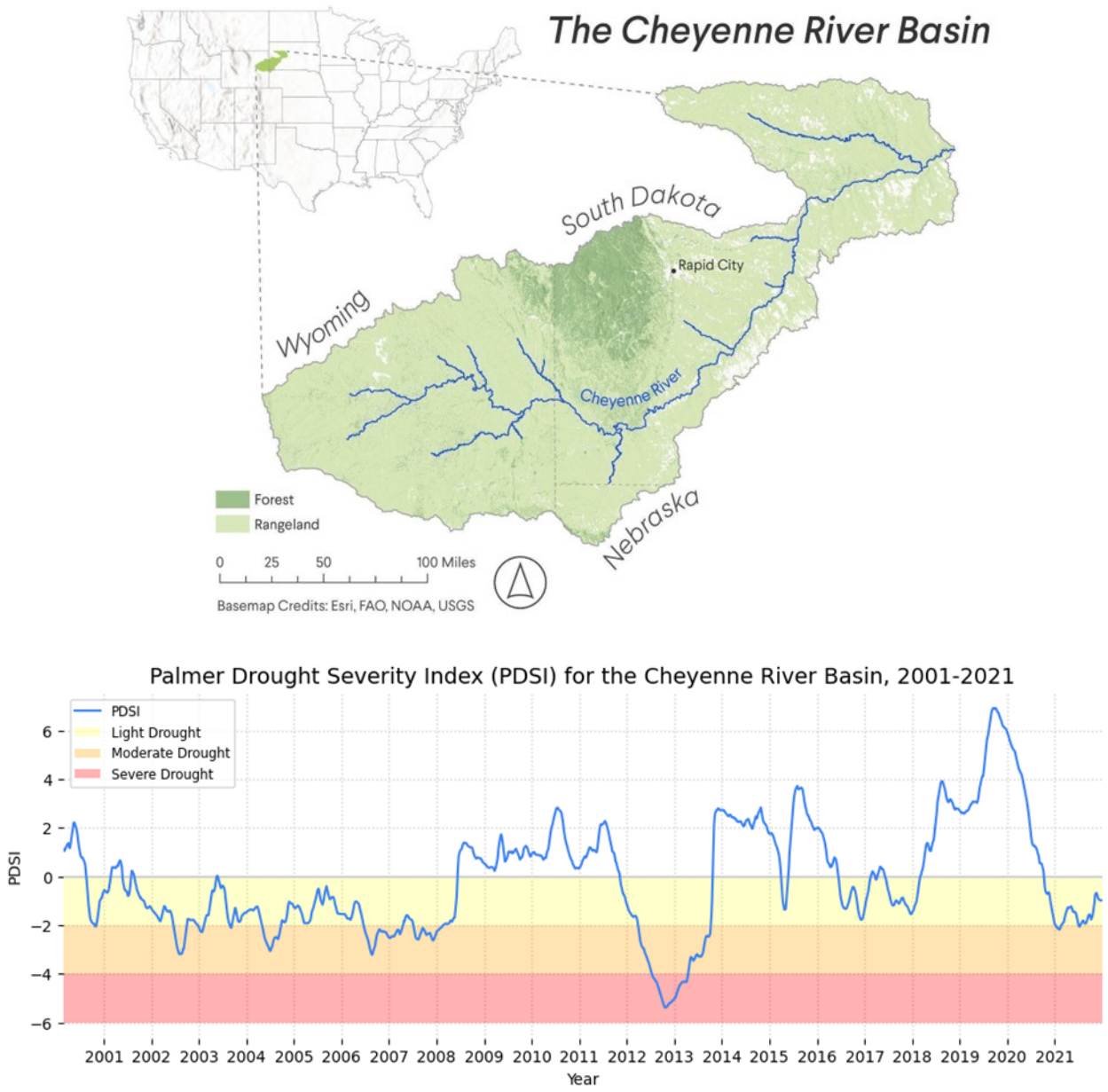

FIGURE 1: Top: The extent of the Cheyenne River Basin, including parts of Wyoming, Nebraska, and South Dakota. Bottom: A time series of Palmer Drought Severity Index values (-10: very dry to +10: very wet) between 2001-2021 for the basin, highlighting periods of light, moderate, and severe drought.

sessment of drought-related ecosystem stress. Additionally, the widespread availability of drought indices through open data makes them more practical for land managers compared to the calculation of standardized NDVI, contributing to effective and timely ecosystem monitoring.

To better understand links between drought indices and ecological drought, this research employed machine learning (ML) and SHapley Additive exPlanations (SHAP) to determine which drought indices and environmental variables contribute most towards predicting NDVI values in the Cheyenne River Basin, a subbasin in the Missouri River Basin (Lundberg & Lee, 2017). The Cheyenne River Basin was selected due to the presence of non-agricultural areas, which allows for a more consistent assessment of vegetation health, as measured by Terra MODIS-derived NDVI, and the occurrence of multiple drought events across the basin over the past two decades (Fig. 1).

ML techniques have increasingly been used to successfully model and predict meteorological and hydrological drought conditions (Belayneh & Adamowski, 2013; Dikshit et al., 2022; Park et al., 2016; Shamshirband et al., 2020; Sundararajan et al., 2021). However, as the number

of published studies on ecological drought increases, there remains a scarcity of studies that specifically investigate ecological drought using ML. Consequently, there is a significant opportunity to leverage ML methods that have proven successful for other forms of drought.

Despite the significant benefits of ML, studies have highlighted that ML approaches to drought monitoring often suffer from a lack of clarity in diagnosing the underlying logic of model decisions (Balti et al., 2020; Samek et al., 2017). To address this issue of explainability, this study used SHAP to interpret the relationships between features and results. SHAP was first introduced as an eXplainable AI (XAI) solution to determine the contributions of individual players in a cooperative game (Shapley, 1953). SHAP was then developed into a method for interpreting ML models by quantifying the directional contribution of model variables, called features, to the model's output (Lundberg & Lee, 2017; Shapley, 1953). SHAP measures how much a feature increases or decreases the predicted value by adding each feature individually to the model, assessing its directional contribution, and then averaging these contributions over all possible feature orderings (Lundberg & Lee, 2017). This measurement allows for a more nuanced understanding of how a model makes its predictions compared to methods such as feature importance scoring. As a result, previous efforts have demonstrated that using SHAP for drought prediction can significantly inform resulting decision-making practices (Dikshit & Pradhan, 2021a, 2021b).

Therefore, this study utilized SHAP to investigate connections between ecological drought impacts and drought indices in the Cheyenne River Basin. The study first sought to build a reliable ML regression model that incorporated both drought indices and environmental variables to predict vegetation health (NDVI), using R-squared to assess model performance. Model performance outcomes were solely intended to ascertain the adequacy of the model for facilitating XAI; a suboptimal model would render the XAI analysis ineffective. Using this model as a foundation, the research then used SHAP to explore associations between model features and NDVI.

2. Methods

2.1 Area of Study

The Cheyenne River Basin spans South Dakota, Wyoming, and Nebraska, and was chosen for this study for its extensive rangeland ecosystem and history of drought (Fig. 1). The study area is characterized by a range of geographical features, including rolling hills, plains, badlands, and plateaus (Culler et al., 1961; Ehlert, 2022). The Cheyenne River runs through the center of the region, originating in Wyoming and flowing eastward through South Dakota before eventually joining the Missouri River. The region's climate is characterized by semi-arid to arid conditions, with limited precipitation and high variability in temperature (Ehlert, 2022). The annual precipitation in the basin ranges from 10 to 20 inches, with the majority falling as rain during the spring and summer months (Culler et al., 1961; United States Bureau of Reclamation, 2019).

The basin is predominantly made up of rangelands, with forests occurring in the north-central portion of the study area, and a minimal amount of developed and agricultural land interspersed throughout (Fig. 1). Rangelands are a key ecosystem in this region and consist primarily of native grasses, forbs, and shrubs (Boden, 2023; Ehlert, 2023; "Introduction to Wyoming Rangelands," 2023). The definition of rangelands does not include a specific land use, underscoring that rangelands are defined by the ecosystems they sustain, rather than by how they are utilized. Thus, rangeland management strategies must be designed to be implemented across private, state, and federal lands, and to consider the relationship between climatic, environmental, and sociological factors.

2.2 Data Acquisition

The data acquisition process utilized the Google Earth Engine (GEE) Python API in a Jupyter Notebook hosted on Google Colaboratory. GEE provides open access to a diverse collection of satellite and remote sensing datasets, enabling reproducibility. This study used NDVI as a proxy for ecological drought and as the target data for the machine learner. NDVI is a widely used remote sensing index that quantifies the greenness of vegetation on the Earth's surface based on the reflectance of near-infrared and red light by plants (Pettorelli et al., 2005). The index generates values on a scale of -1 to +1, with negative values indicative of bare soil or little vegetative cover, low positive values indicative of unhealthy vegetative cover, and high positive values indicative of healthy vegetative cover.

Next, several drought indices and environmental variables were selected as features for the ML model to account for different aspects of drought dynamics. These additional indices and variables include short and long-term soil moisture status, atmospheric evaporative demand, precipitation anomalies, temperature, and snow cover. The selection of data involved considering relationships between drought and vegetation dynamics, as well as the inclusion of widely used and easily reproducible drought indices. It is important to note that, because the primary goal of this model is not to optimize NDVI prediction but to identify commonly used indicators that closely correlate with changes in NDVI, the presence of correlations between

TABLE 1: Earth observations acquired for the ML models.

Data Type	Data Platform	Variable(s)	Spatial & Temporal Resolution	Time Period
Drought Impact (Target)	MOD09GA v006: MODIS/Terra Surface Reflectance Daily L2G Global 1 km SIN Grid	Normalized Difference Vegetation Index (NDVI) - Derived	1 km, Daily	2000 – Present
Drought Indices (Features)	Gridded Surface Meteorological (gridMET) Dataset	Palmer Drought Severity Index (PDSI)	4 km, Pentads (5-day)	1979 – Present
		Palmer Z Score		
		Evaporative Demand Drought Index (EDDI) – 30, 90, 180-day		
		Standardized Precipitation Index (SPI) – 30, 90, 180-day		
		Standardized Precipitation Evapotranspiration Index (SPEI) – 30, 90, 180-day		
Environmental Variables (Features)	Gridded Surface Meteorological (gridMET) Dataset	Maximum & Minimum Temperature	4 km, Daily	1979 – Present
		Precipitation		
	Daymet	Snow Water Equivalent (SWE)	1 km, Daily	1980 – Present

drought indices and environmental variables in the model is acceptable due to how SHAP values are calculated (Lundberg & Lee, 2017; Janzing et al., 2020).

The following Earth observation data were acquired as image collections for the date range of February 24, 2000 (Terra MODIS's start date), and December 31, 2021: Terra Moderate Resolution Imaging Spectroradiometer (MODIS) surface reflectance data, Gridded Surface Meteorological (gridMET) Dataset, Palmer Drought Severity Index (PDSI), Evaporative Demand Drought Index (EDDI), Standardized Precipitation Index (SPI), Standardized Precipitation Evapotranspiration Index (SPEI), maximum and minimum temperature, precipitation data, and Daymet snow water equivalent (SWE) data (Table 1).

2.3 Data Preprocessing

To produce an NDVI band from the MODIS surface reflectance data, MODIS data were first quality controlled by masking cloudy and snowy pixels using bits from the QA band. An NDVI band was then computed by taking the normalized difference of the near-infrared and red bands. Spatially averaged values for each variable in Table 1 were calculated across the study area over time using the study area shapefile and GEE. To filter out seasonal variations, such as the annual spring vegetation green-up, standardized anomalies were computed by grouping the data for each variable by day of the year, subtracting the mean for each day from the original value, and dividing by the standard deviation:

All data were then resampled and smoothed to match the temporal frequency of the gridMET drought indices (pentads), using a 30-day rolling average for a five-day temporal resolution (Abatzoglou, 2013). By performing data standardization, emphasis was placed on identifying

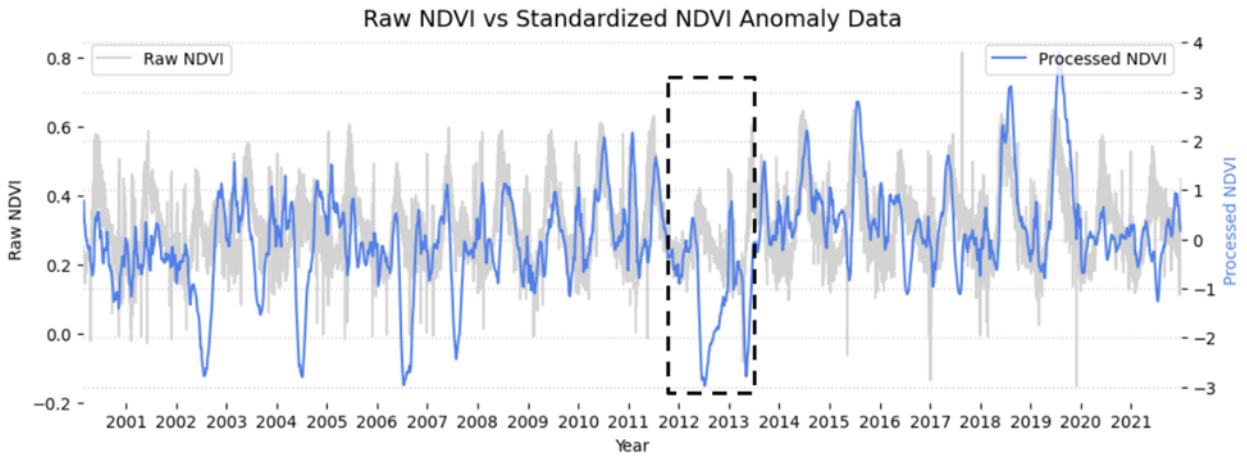

FIGURE 2: Comparison of raw NDVI and standardized NDVI anomaly data, illustrating the distinction between unprocessed and preprocessed data. The dashed box highlights a decrease in the processed NDVI data (blue line), indicating how data preprocessing accentuated the rapid decline in vegetation health during the 2012 drought. This event may have been challenging to discern in the raw data.

important deviations from expected drought conditions, such as drought-induced unseasonably low levels of NDVI (Fig. 2).

2.4 Model Selection

After preprocessing, NDVI was defined as the prediction target (y) and the remaining variables served as features (X). To choose an appropriate model, the Lazy Predict package was utilized to identify the regression models with the highest predictive performance for the given data (Pandala, 2022). Lazy Predict automates the process of testing over 40 types of untuned models, including a Linear Regressor, an XGBoost Regressor, and an Extra Trees Regressor, and provides statistics on model performance. Based on averaged cross-validated performance results, the Extra Trees Regressor and XGBoost Regressor frameworks were selected for further use. In testing, both frameworks had high R-squared and adjusted R-squared values, low Root Mean Square Error (RMSE) values, and moderate runtimes. Additionally, the Extra Trees and XGBoost Regressors are both tree-based models, which are generally considered to be more interpretable than other types of models (Molnar, 2023).

2.5 Model Tuning

After selecting the Extra Trees and XGBoost Regressor frameworks, the data were split into training and test sets, reserving the final 20% of data chronologically for testing. Hyperparameter optimization, a process used to tune model parameters for optimal performance, was conducted on each regressor in two steps: random search and, subsequently, grid search. This method of model tuning is an essential step in machine learning model development as it helps improve generalization to unseen data and reduce overfitting (Müller & Guido, 2016). Both search methods utilized k-fold cross-validation to ensure the generalization of results. Random search was repeated six times to narrow down the hyperparameter ranges based on the highest mean cross-validated R-squared values for each search. Grid search exhaustively searched arrays of possible combinations of hyperparameter values within the narrowed ranges discovered through random search. After three

TABLE 2: Selected hyperparameters for the Extra Trees Regressor and XGBoost Regressor models based on the results of random and grid search. These parameters, such as settings for bootstrap sampling, tree depth, and the number of estimators (trees) used, collectively shape the behavior of each model.

Model	Hyperparameters	
Extra Trees Regressor R-Squared: 0.48	bootstrap = False max_depth = 18 max_features = 1.0	min_samples_leaf = 1 min_samples_split = 2 n_estimators = 900
XGBoost Regressor R-Squared: 0.54	colsample_bytree = 0.75 max_depth = 7 learning_rate = 0.031	min_child_weight = 4 subsample = 0.62 n_estimators = 950

rounds of grid search, the best hyperparameters for each regressor were selected based on R-squared values (Table 2).

3. Analysis

3.1 Model Performance

Using the best hyperparameters discovered during the hyperparameter optimization searches, the ExtraTrees and XGBoost regressors were fit to the training data, producing two trained models. Predictions for each model were assessed using the model's predict() function on the test data. Mean Squared Error (MSE), RMSE, and R-squared scores were computed to evaluate model performance, and actual and predicted NDVI values were plotted.

3.2 Feature Importance and SHAP Values

Feature importance scores were calculated to provide a relative measure of how much each feature contributes to the model's predictive performance (Müller & Guido, 2016). Additionally, the SHAP library was utilized to explain the results of each model by providing insights into both the magnitude and direction of feature effects on predictions (Lundberg & Lee, 2017). A SHAP Explainer object was created to offer detailed explanations for model predictions. SHAP values were then computed for the test data in both models.

4. Results

4.1 Model Performance

The Extra Trees Regressor model returned an MSE of 0.54, an RMSE of 0.74, and an R-squared of 0.48 (Table 3). These metrics suggest that the Extra Trees Regressor model performed moderately well in predicting NDVI values based on the input features. However, the XGBoost model demonstrated better performance than the Extra Trees Regressor model, with a lower MSE of 0.48 and RMSE of 0.69 (Table 3). The model's R-squared value was higher at 0.54, indicating that the XGBoost model explained a larger proportion of the variance in the NDVI values, making it a more accurate predictor.

TABLE 3: Test set performance using the selected hyperparameters from Table 2.

Model	MSE	RMSE	R-Squared
Extra Trees Regressor	0.54	0.74	0.48
XGBoost Regressor	0.48	0.69	0.54

FIGURE 3: True and predicted NDVI values from the Extra Trees (top) and XGBoost (bottom) Regressors.

Model performance is visually represented in Figure 3 with plots of the true and predicted NDVI values from the Extra Trees and XGBoost Regressor models. Model predictions tended to be directionally correct, with much of the error coming from the underestimation of large NDVI spikes around indices 75 and 150, representing two intense vegetation green-ups, and the overall depression of NDVI values across the predicted set.

4.2 Feature Importance and SHAP Values

Based on the performance evaluation of both models, the XGBoost model was selected for further analysis due to the model's lower MSE and RMSE. Feature importance scores for the XGBoost model were ranked based on their contribution to the model's ability to predict NDVI values over time in the study area (Fig. 4, left). According to the feature importance scores, PDSI had the highest importance with a score of 0.26, followed by SPI 90-day with a score of 0.17, and SWE with a score of 0.11. Other important features included SPI 180-day with a score of 0.09, Palmer Z with a score of 0.09, and SPEI 180-day with a score of 0.06.

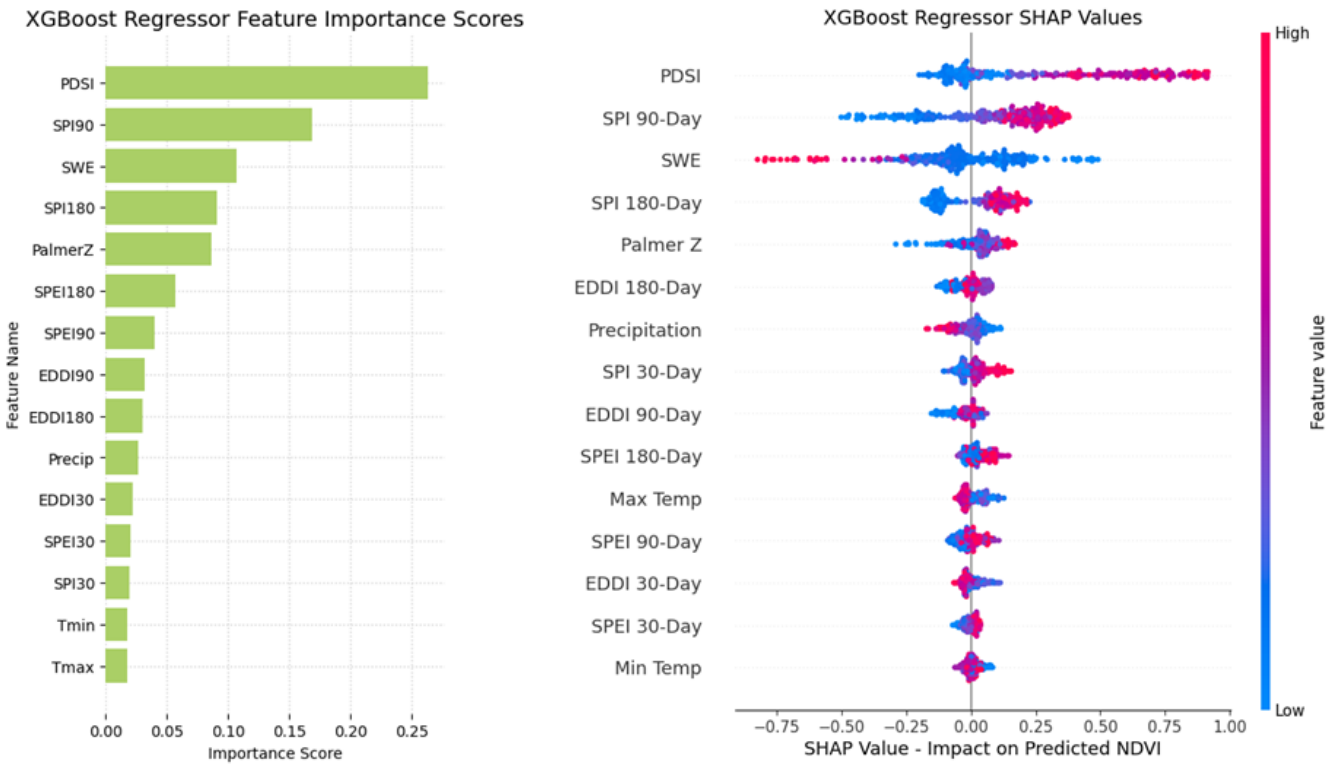

FIGURE 4: Feature importance scores (left) and SHAP values (right) from the XGBoost Regressor. On the right, shades of pink denote higher feature values while shades of blue indicate lower feature values. Negative impacts on NDVI are to the left of the vertical axis, while positive impacts are to the right of the vertical axis.

While feature importance scores provided a general indication of the relative importance of different features in the model, SHAP values offered a more detailed and directional explanation of the contribution of each feature for a specific predicted NDVI value. The right plot in Figure 4 depicts SHAP values for each feature as individual points grouped by feature name. Figure 4 shows that PDSI had the most significant overall contribution to NDVI prediction, with a direct relationship between high PDSI values and high SHAP values and vice versa. SHAP values for PDSI skewed positively, suggesting that PDSI was stronger at predicting high NDVI values, and weaker at predicting low NDVI values. The SPI 90-day index ranked second with a relatively balanced impact on positive and negative SHAP values compared to PDSI. SWE ranked third and had an inverse relationship with resulting SHAP values. Additionally, SHAP values for SWE had a wider distribution than SHAP values for SPI 90-day. Lastly, SPI 180-day and Palmer Z index had relatively high SHAP scores and mirrored the direct relationships observed for PDSI and SPI 90-day.

5. Discussion

5.1 Model Performance

Model performance results show that the XGBoost model was more effective than the Extra Trees Regressor at predicting NDVI. Therefore, the XGBoost model was chosen for analysis with SHAP. The moderate performance of the XGBoost model at predicting NDVI is essential as it establishes the credibility and reliability of SHAP results. Further, XGBoost has been the chosen method for several other drought and NDVI prediction studies, including Li et al.'s 2021 paper. In this study, the authors achieved an R-squared of 0.83, significantly higher than the R-squared value achieved here, using historical NDVI values along with six environmental variables as model features (Li et al., 2021). While Li et al. developed a model that better explains the variance in their NDVI data, it did not incorporate drought indices and therefore addressed a different goal than this research.

In comparison, this research aimed to assess the possibility of using a predictive model to inform the relationship of commonly used drought indices to vegetation health. This aim is an important distinction between past predictive

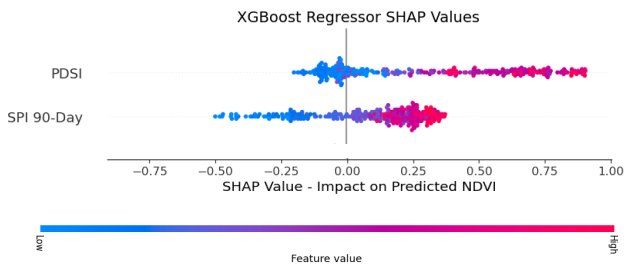

FIGURE 5: Distribution of SHAP values for PDSI and SPI 90-day. SPI 90-day is more evenly distributed across negative and positive SHAP values than PDSI, which skews positive.

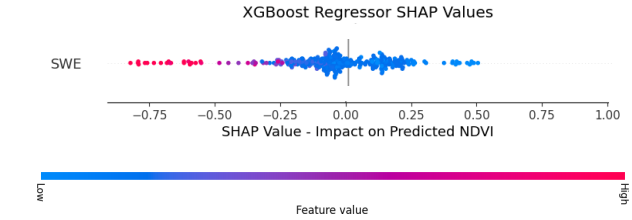

FIGURE 6: Distribution of SHAP values for SWE. High SWE values were associated with low NDVI values, while low SWE values were associated with both high and low NDVI values.

work and this study, as stakeholders throughout the wider Missouri River Basin have expressed the need to develop a more comprehensive understanding of drought indicators and their specific relationship to ecological drought (NOAA/NIDIS, 2020).

5.2 Feature Importance and SHAP Values

The feature importance scores and SHAP values from the XGBoost model demonstrate that drought indices (e.g., PDSI, SPI) were more important in predicting NDVI values than environmental variables (e.g., precipitation, temperature) in the Cheyenne River Basin. This assertion signifies that changes in the values of drought indices are more strongly linked to changes in the predicted NDVI values.

PDSI ranked highest in feature importance scores and had the highest absolute mean SHAP values. This result suggests that PDSI was the most important factor in predicting NDVI, demonstrating a direct relationship between PDSI values and NDVI. However, a key takeaway from PDSI's SHAP results is the importance of using SHAP to not only determine the strength of relationships but also the directionality of relationships between features and target variables. This takeaway is highlighted when noting the distribution of the SHAP values for PDSI and SPI 90-day, which performs second in terms of absolute mean SHAP values, across the x-axis (Fig. 5). SPI 90-day showed a more even distribution than PDSI across negative and positive SHAP values, while PDSI skewed positively overall. This positive skew implies that PDSI was stronger at predicting high NDVI values, and weaker at predicting low NDVI values. While PDSI has higher absolute SHAP values, the more even distribution of SPI 90-day across positive and negative SHAP values demonstrates that SPI 90-day may be more effective at predicting lower NDVI values, and, by extension, drought impacts on vegetation health, despite its lower absolute mean SHAP value.

In this study, SWE was the one environmental variable to rank highly in SHAP analyses. However, in contrast to other model features, high SWE values were associated with low NDVI values, and vice versa (Fig. 6). As SWE was incorporated into the model without a temporal lag, this inverse relationship is as expected due to snow accumulation in the winter months suppressing vegetation growth, resulting in higher SWE and lower NDVI values (Grippa et al., 2005; T. Wang et al., 2013; Y. Wang et al., 2022). As temperatures increase and SWE values lessen in the spring, snowmelt results in increased water availability for vegetation, leading to higher NDVI values (Matongera et al., 2021; Paudel & Andersen, 2013). This inverse relationship captures key information related to seasonal vegetation growth suppression and green-up that is otherwise unaccounted for in the model. This relationship likely explains the high SHAP values of SWE, as it captures seasonal vegetation dynamics that drought indices alone do not. However, in terms of monitoring ecological drought, SWE may be more useful as a lagged variable in conjunction with drought indices to provide information on water availability from snowmelt.

Overall, SHAP results suggest that vegetation health monitoring efforts in the Cheyenne River Basin should focus on using PDSI and SPI 90-day. However, the directionality of SHAP values for PDSI demonstrates that it may be less effective at predicting lower NDVI values than SPI 90-day. Therefore, when considering drought impacts, PDSI may be an even more powerful indicator when used in concert with SPI 90-Day.

Results also support the findings of the few previous studies that have used XAI techniques in the domain of drought research (Dikshit & Pradhan, 2021a). In particular, this study's successful use of SHAP aligns with findings in Dikshit and Pradhan's 2021 paper Explainable AI in Drought Forecasting, which suggested that SHAP is useful to understand the impact of variables within drought-related models. The utilization of SHAP in this study not only reinforces the findings presented by Dikshit and Pra-

dhan (2021) but also emphasizes the practicality of using SHAP to assess the influence of various variables within drought-related models. These collective findings contribute to the growing consensus on the effectiveness of XAI methodologies in enhancing the understanding of complex phenomena like drought and pave the way for further advancements in the field.

5.3 Limitations

There are four main limitations to this research. First, the MODIS-derived NDVI data have limitations that could affect their accurate representation of vegetation health. For instance, factors such as cloud cover, atmospheric conditions, and solar angle can affect the accuracy and consistency of NDVI. Additionally, NDVI may not capture changes in vegetation density or structure, which can also be a resultant impact of ecological drought conditions.

Second, the pentad drought indices are limited in their ability to describe daily drought conditions, as they represent only five-day intervals and may not capture spatial variability or lagged effects of drought on vegetation health. Daily data were resampled to this temporal resolution to avoid having to interpolate a significant portion of the data for many features used by the ML model.

Next, XGBoost, Extra Trees, and ML regression algorithms in general have limitations that can impact their ability to predict targets accurately. Most notably, the quality and representativeness of the training data play a critical role in the performance of a machine learner. Even with careful preprocessing, remotely sensed data contain some amount of error and uncertainty, potentially decreasing a model's accuracy or performance. For this research specifically, results are also limited by the moderate R-squared values of both models. Additionally, caution should be taken when interpreting and generalizing results from ML models, as they may not capture all complex relationships between features and targets and may be influenced by region-specific factors.

Lastly, while SHAP values can provide valuable insight into the importance of different features, they can be computationally expensive and may not provide accurate or meaningful insights if the model is poorly constructed or trained on biased data. SHAP values assume that input features are independent, which may not be true in all cases and may not fully capture joint interactions between features (Lundberg & Lee, 2017). Therefore, SHAP values should be used with caution alongside other explainability techniques to comprehensively understand the relationships between features and predictions.

6. Conclusions

This research stands apart from previous predictive NDVI research due to its focus on identifying the most influential drought indices and environmental variables for predicting NDVI, thereby uncovering indicators of vegetation stress in the study area. While relatively accurate NDVI prediction has been achieved in the past using machine learning methods (Li et al., 2021; Roy, 2021), the significance of this work lies in both its introduction of XAI methods into the ecological drought field and its identification of ecological drought indicators in the study area. By leveraging SHAP, this research not only provides insights into the ML model's predictions but also empowers human users to scrutinize their intuitions and validate them against model results. The use of XAI to interpret ML predictions represents a novel and valuable approach in the domain of ecological drought. Consequently, this research contributes to both the application of explainable predictive modeling techniques in the domain and the development of tools for drought monitoring and management.

Drought monitoring and management in the United States employ a diverse range of methods. The findings of this study not only inform this broader context of drought monitoring strategies but also offer tangible insights for using drought indices to guide on-the-ground actions, benefiting the Cheyenne River Basin region and serving as a model for other rangeland areas facing ecological drought challenges. While acknowledging the need for further studies to assess generalizability to different regions, the methods employed in this research pave the way for a more effective and nuanced approach to drought monitoring for a variety of stakeholders, from climatologists to natural resource managers, ultimately fostering more adaptive responses to mitigate the impacts of ecological drought.

References

Anyamba, A., & Tucker, C. J. (2012). Historical Perspectives on AVHRR NDVI and Vegetation Drought Monitoring. In B. D. Wardlow, M. C. Anderson, & J. P. Verdin (Eds.), Remote Sensing of Drought (0 ed., pp. 23–49). CRC Press. https://doi.org/10.1201/b11863-9

Balti, H., Ben Abbes, A., Mellouli, N., Farah, I. R., Sang, Y., & Lamolle, M. (2020). A review of drought monitoring with big data: Issues, methods, challenges and research directions. Ecological Informatics, 60, 101136. https://doi.org/10.1016/j.ecoinf.2020.101136

Belayneh, A., & Adamowski, J. (2013). Drought forecasting using new machine learning methods. Journal of Water and Land Development, 18(9), 3–12. https://doi.org/10.2478/jwld-2013-0001

Boden, D. (2023). Nebraska Rangelands: Rangelands. https://unl.libguides.com/c.php?g=51759&p=333380

Bradford, J. B., Schlaepfer, D. R., Lauenroth, W. K., & Palmquist, K. A. (2020). Robust ecological drought projections for drylands in the 21st century. Global Change Biology, 26(7), 3906–3919. https://doi.org/10.1111/gcb.15075

Clarivate Web of Science. (2023). "Ecological Drought" Analyze Results. Clarivate Web of Science. https://www-webofscience-com.proxy1.library.jhu.edu/wos/woscc/analyze-results/c64bd0be-48a8-48ee-842a-621365883054-831d4de7

Crausbay, S. D., Betancourt, J., Bradford, J., Cartwright, J., Dennison, W. C., Dunham, J., Enquist, C. A. F., Frazier, A. G., Hall, K. R., Littell, J. S., Luce, C. H., Palmer, R., Ramirez, A. R., Rangwala, I., Thompson, L., Walsh, B. M., & Carter, S. (2020). Unfamiliar Territory: Emerging Themes for Ecological Drought Research and Management. One Earth, 3(3), 337–353. https://doi.org/10.1016/j.oneear.2020.08.019

Crausbay, S. D., Ramirez, A. R., Carter, S. L., Cross, M. S., Hall, K. R., Bathke, D. J., Betancourt, J. L., Colt, S., Cravens, A. E., Dalton, M. S., Dunham, J. B., Hay, L. E., Hayes, M. J., McEvoy, J., McNutt, C. A., Moritz, M. A., Nislow, K. H., Raheem, N., & Sanford, T. (2017). Defining Ecological Drought for the Twenty-First Century. Bulletin of the American Meteorological Society, 98(12), 2543–2550. https://doi.org/10.1175/BAMS-D-16-0292.1

Culler, R. C., Hadley, R. F., & Schumm, S. A. (1961). Hydrology of the upper Cheyenne River basin: Part A. Hydrology of stock-water reservoirs in upper Cheyenne River basin; Part B. Sediment sources and drainage-basin characteristics in upper Cheyenne River basin. https://doi.org/10.3133/wsp1531

Dikshit, A., & Pradhan, B. (2021a). Explainable AI in drought forecasting. Machine Learning with Applications, 6, 100192. https://doi.org/10.1016/j.mlwa.2021.100192

Dikshit, A., & Pradhan, B. (2021b). Interpretable and explainable AI (XAI) model for spatial drought prediction. Science of The Total Environment, 801, 149797. https://doi.org/10.1016/j.scitotenv.2021.149797

Dikshit, A., Pradhan, B., & Santosh, M. (2022). Artificial neural networks in drought prediction in the 21st century–A scientometric analysis. Applied Soft Computing, 114, 108080. https://doi.org/10.1016/j.asoc.2021.108080

Ehlert, K. (2022). Landscape and Watershed Setting of the Northern Great Plains of Western South Dakota. South Dakota State University. https://extension.sdstate.edu/sites/default/files/2022-07/P-00235-03.pdf

Ehlert, K. (2023). Partner State Resources for South Dakota Rangelands. https://rangelandsgateway.org/states/south-dakota

Grippa, M., Kergoat, L., Le Toan, T., Mognard, N. M., Delbart, N., L'Hermitte, J., & Vicente-Serrano, S. M. (2005). The impact of snow depth and snowmelt on the vegetation variability over central Siberia. Geophysical Research Letters, 32(21), L21412. https://doi.org/10.1029/2005GL024286

Intergovernmental Panel on Climate Change (IPCC). (2021). Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, & B. Zhou (Eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. https://doi.org/10.1017/9781009157896

Introduction to Wyoming Rangelands. (2023). Wyoming Rangelands. https://uwyoextension.org/uwrange/

Janzing, D., Minorics, L. & Bloebaum, P. (2020). Feature relevance quantification in explainable AI: A causal problem. Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics, in Proceedings of Machine Learning Research 108:2907-2916. https://proceedings.mlr.press/v108/janzing20a.html

Li, X., Yuan, W., & Dong, W. (2021). A Machine Learning Method for Predicting Vegetation Indices in China. Remote Sensing, 13(6), Article 6. https://doi.org/10.3390/rs13061147

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Proceedings of the 31st International Conference on Neural Information Processing Systems, 4768–4777.

Matongera, T. N., Mutanga, O., Sibanda, M., & Odindi, J. (2021). Estimating and Monitoring Land Surface Phenology in Rangelands: A Review of Progress and Challenges. Remote Sensing, 13(11), Article 11. https://doi.org/10.3390/rs13112060

McEvoy, D. J., Hobbins, M., Brown, T. J., VanderMolen, K., Wall, T., Huntington, J. L., & Svoboda, M. (2019). Establishing Relationships between Drought Indices and Wildfire Danger Outputs: A Test Case for the California-Nevada Drought Early Warning System. Climate, 7(4), Article 4. https://doi.org/10.3390/cli7040052

Molnar, C. (2023). Interpretable Machine Learning. https://christophm.github.io/interpretable-ml-book/#

Müller, A. C., & Guido, S. (2016). Introduction to Machine Learning with Python: A Guide for Data Scientists. O'Reilly Media, Inc.

NOAA/NIDIS. (2020, December 14). 2021-2023 Missouri River Basin Drought Early Warning System Strategic Action Plan. Drought.Gov. https://www.drought.gov/documents/2021-2023-missouri-river-basin-drought-early-warning-system-strategic-action-plan

Park, S., Im, J., Jang, E., & Rhee, J. (2016). Drought assessment and monitoring through blending of multi-sensor indices using machine learning approaches for different climate regions. Agricultural and Forest Meteorology, 216, 157–169. https://doi.org/10.1016/j.agrformet.2015.10.011

Paudel, K. P., & Andersen, P. (2013). Response of rangeland vegetation to snow cover dynamics in Nepal Trans Himalaya. Climatic Change, 117(1), 149–162. https://doi.org/10.1007/s10584-012-0562-x

Pettorelli, N., Vik, J. O., Mysterud, A., Gaillard, J.-M., Tucker, C. J., & Stenseth, N. Chr. (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. Trends in Ecology & Evolution, 20(9), 503–510. https://doi.org/10.1016/j.tree.2005.05.011

Phillips, L. B., Hansen, A. J., & Flather, C. H. (2008). Evaluating the species energy relationship with the newest measures of ecosystem energy: NDVI versus MODIS primary production. Remote Sensing of Environment. 112: 4381-4392.https://doi.org/10.1016/j.rse.2008.04.012

Roy, B. (2021). Optimum machine learning algorithm selection for forecasting vegetation indices: MODIS NDVI & EVI. Remote Sensing Applications: Society and Environment, 23, 100582. https://doi.org/10.1016/j.rsase.2021.100582

Samek, W., Wiegand, T., & Müller, K-R. (2018). Explainable artificial intelligence: understanding, visualizing and interpreting deep learning models. ITU Journal: ICT Discoveries, 1(1), 39-48. http://handle.itu.int/11.1002/pub/8129fdff-en

Shamshirband, S., Hashemi, S., Salimi, H., Samadianfard, S., Asadi, E., Shadkani, S., Kargar, K., Mosavi, A., Nabipour, N., & Chau, K.-W. (2020). Predicting Standardized Streamflow index for hydrological drought using machine learning models. Engineering Applications of Computational Fluid Mechanics, 14 (1), 339–350. https://doi.org/10.1080/19942060.2020.1715844

Shapley, L. (1953). 17. A Value for n-Person Games. In H. Kuhn & A. Tucker (Ed.), Contributions to the Theory of Games (AM-28), Volume II (pp. 307-318). Princeton: Princeton University Press. https://doi.org/10.1515/9781400881970-018

Sundararajan, K., Garg, L., Srinivasan, K., Bashir, A., Kaliappan, J., Ganapathy, G., Selvaraj, S. K., & Thiruvadi, M. (2021). A Contemporary Review on Drought Modeling Using Machine Learning Approaches. Computer Modeling in Engineering and Sciences, 128, 447–487. https://doi.org/10.32604/cmes.2021.015528

Tucker, C. J., Justice, C. O., & Prince, S. D. (1986). Monitoring the grasslands of the Sahel 1984-1985. International Journal of Remote Sensing, 7(11), 1571–1581. https://doi.org/10.1080/01431168608948954

United States Bureau of Reclamation. (2019). Basin Precipitation—Monthly and Cumulative.

Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2010). A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index. Journal of Climate, 23(7), 1696–1718. https://doi.org/10.1175/2009JCLI2909.1

Wang, T., Peng, S., Lin, X., & Chang, J. (2013). Declining snow cover may affect spring phenological trend on the Tibetan Plateau. Proceedings of the National Academy of Sciences, 110(31), E2854–E2855. https://doi.org/10.1073/pnas.1306157110

Wang, Y., Chen, Y., Li, P., Zhan, Y., Zou, R., Yuan, B., & Zhou, X. (2022). Effect of Snow Cover on Detecting Spring Phenology from Satellite-Derived Vegetation Indices in Alpine Grasslands. Remote Sensing, 14(22), Article 22. https://doi.org/10.3390/rs14225725

Wiens, J., & Bachelet, D. (2010). Matching the Multiple Scales of Conservation with the Multiple Scales of Climate Change. Conservation Biology : The Journal of the Society for Conservation Biology, 24, 51–62. https://doi.org/10.1111/j.1523-1739.2009.01409.x

Data and Software

Abatzoglou, J. T. (2013). Development of gridded surface meteorological data for ecological applications and modelling [Data set]. International Journal of Climatology, 33: 121–131. https://doi.org/10.1002/joc.3413

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System (1.7.5) [Software]. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794. https://doi.org/10.1145/2939672.2939785

Google Colaboratory (2.84.0) [Software]. (n.d.). https://colab.research.google.com/

Google Earth Engine API (0.1.276) [Software]. (2021). Google. https://github.com/google/earthengine-api

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M.,

Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy (1.22.4) [Software]. Nature, 585(7825), Article 7825. https://doi.org/10.1038/s41586-020-2649-2

Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment (3.7.1) [Software]. Computing in Science & Engineering, 9(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions (0.41.0) [Software]. Proceedings of the 31st International Conference on Neural Information Processing Systems, 4768–4777.

McKinney, W., & others. (2010). Data structures for statistical computing in python (1.5.3) [Software]. Proceedings of the 9th Python in Science Conference (Vol. 445, pp. 51–56).

OpenAI. (2023). ChatGPT (Version GPT-3.5) [Software]. Accessed multiple times between January 2023 and April 2023 for coding assistance. https://chat.openai.com/chat

Pandala, S. R. (2022). Lazy Predict (0.2.12) [Software]. https://github.com/shankarpandala/lazypredict

Pedregosa et al. (2011). Scikit-learn: Machine Learning in Python (1.2.2) [Software]. JMLR 12, pp. 2825-2830, 2011. https://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html

Thornton, M.M., Shrestha, R., Wei, Y., Thornton, P.E., Kao, S-C., & Wilson, B.E. (2022). Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 4 R1 [Data set]. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/2129

Vermote, E., Wolfe, R. (2015). MOD09GA MODIS/Terra Surface Reflectance Daily L2G Global 1kmand 500m SIN Grid V006 [Data set]. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MODIS/MOD09GA.006