Comparative Evaluation of Weather Forecasting using Machine Learning Models

Md Saydur Rahman*, Farhana Akter Tumpa†, Md Shazid Islam*,
Abul Al Arabi‡, Md Sanzid Bin Hossain§, Md Saad Ul Haque¶
*University of California Riverside, USA,
†Ahsanullah University of Science and Technology, Bangladesh
‡CSE, Texas A&M University, USA
§University of Central Florida, USA, ¶ University of Florida, USA
Email: mrahm054@ucr.edu, Tumpafarhanaakter@gmail.com, misla048@ucr.edu,
abulalarabi@tamu.edu, md543636@ucf.edu, haque.m@ufl.edu,

Abstract—**Gaining a deeper understanding of weather and being able to predict its future conducts have always been considered important endeavors for the growth of our society. This research paper explores the advancements in understanding and predicting nature's behavior, particularly in the context of weather forecasting, through the application of machine learning algorithms. By leveraging the power of machine learning, data mining, and data analysis techniques, significant progress has been made in this field. This study focuses on analyzing the contributions of various machine learning algorithms in predicting precipitation and temperature patterns using a 20-year dataset from a single weather station in Dhaka city. Algorithms such as Gradient Boosting, AdaBoosting, Artificial Neural Network, Stacking Random Forest, Stacking Neural Network, and Stacking KNN are evaluated and compared based on their performance metrics, including Confusion matrix measurements. The findings highlight remarkable achievements and provide valuable insights into their performances and features correlation.**

Index Terms—**Machine Learning, Precipitation, Temperature, Boosting Algorithm, Weather Forecasting, Stacking Neural Network**

I. INTRODUCTION

The influence of weather on human survival and the development of civilization has long been recognized as pivotal. Accurate prediction of precipitation and temperature assumes paramount importance for the assessment of hydrological processes, effective management of water resources, anticipation of floods and droughts, and the evaluation of climate change impacts. Furthermore, the escalating rate of global warming and climate-related disasters has heightened the urgency for weather data prediction. While rain gauges provide localized data, radar-based models offer enhanced spatiotemporal resolution. Approaches to precipitation forecasting encompass dynamic methods based on physical models as well as empirical techniques reliant on data-driven methodologies. Notably, the empirical approach, leveraging artificial intelligence algorithms, has demonstrated promise in tackling meteorological variables' intricate and chaotic nature, thereby facilitating efficient and reliable predictions. Machine learning models, widely explored in the existing body of literature, have gained substantial recognition as effective tools for weather forecasting.

Machine learning algorithms have been very useful in processing multi-modal data [1]–[7] and optimization [8], [9]. These algorithms have been extensively researched and compared to assess their accuracy in predicting continually changing phenomena such as weather conditions. Key factors such as temperature, humidity, wind speed, surface pressure, and precipitation are vital in obtaining precise measurements for weather pattern predictions. Various popular machine learning algorithms, including Artificial Neural Networks (ANN), Support Vector Machines (SVM), Support Vector Regression (SVR), and Genetic Algorithms (GA), have been utilized with specific modifications or adaptations in these techniques. Recent advancements in Numerical Weather Prediction (NWP) models have enabled testing at finer resolutions, facilitating a more detailed representation of storm processes and hourly predictions. However, effectively placing and timing precipitation and temperature remains a challenge compared to models with coarser resolutions.

In order to tackle this challenge, ensemble methodologies utilizing storm-scale Numerical Weather Prediction (NWP) models have emerged, offering enhanced estimations of uncertainty by effectively sampling spatiotemporal errors linked to individual storms. The integration of ensemble predictions into accurate and actionable consensus guidance presents an area warranting further investigation. Boosting algorithms, such as AdaBoost, XGBoost, and Gradient Boost, have garnered considerable attention within the field of weather prediction, as they iteratively combine weak rules to establish a robust learner, thereby yielding more precise outcomes. Extensive analyses have been conducted to evaluate the performance of these algorithms in discerning weather conditions based on multiple contributing factors.

Bangladesh experiences significant fluctuations in rainfall and temperature, particularly in its capital city, Dhaka, thereby necessitating targeted investigations. The present study is dedicated to the analysis of an extensive collection of weather data encompassing a time span of around two decades (2003-2023) obtained from Dhaka City, Bangladesh. The dataset

comprises a diverse array of crucial variables, including wind speed, maximum and minimum temperatures, precipitation, wind direction, surface pressure, relative humidity, and specific humidity. In order to extract meaningful insights from the dataset, we have employed a suite of sophisticated algorithms, namely Gradient Boost, AdaBoost, Stacking KNN, and Stacking Random Forest. Our primary objective revolves around assessing the accuracy and performance of these algorithms in predicting long-term precipitation patterns and average temperature. Through the utilization of these advanced methodologies, we aim to enrich our comprehension of weather dynamics and bolster our forecasting capabilities in the Dhaka region.

II. Related Works

There has been a lot of work progressed in the context of weather pattern forecasting by machine learning. Recent research on IoT-based environment data collection (e.g.,[10]) is further fostering the weather prediction paradigm. In weather forecast research, authors involve the creation of a dataset by collecting data from previous months [11]. To simplify the analysis, the months are categorized into three seasons: summer, fall, and winter, with each season comprising four months. Subsequently, algorithms like linear regression, logistic regression, and Gaussian Naive Bayes are applied to the dataset for further analysis and prediction purposes.

In [3] authors aim to develop precipitation forecasting models for Northern Bangladesh with a lag time of up to 3 months. Two machine learning algorithms, M5P and SVR, were used, and a hybrid model combining both (M5P-SVR) was developed. The SVR parameters were optimized using the particle swarm optimization algorithm. This hybrid model is unique in the literature for predicting monthly precipitation in the monsoon climate of Northern Bangladesh. The study also includes a sensitivity analysis of the number of input parameters and lag time.

Moreover, the author proposes a data augmentation algorithm combining K-means clustering and SMOTE to enhance sample information for forecasting monthly grid precipitation in the Danjiangkou River Basin[12]. RF, XGB, RNN, and LSTM models are employed. The aim is to improve medium- and long-term precipitation forecasting accuracy. Results show that SMOTE-km-XGB performs best, and deep learning methods (RNN and LSTM) do not benefit significantly from the data augmentation. This study enhances precipitation forecast accuracy and provides insights for improving hydrological forecasting.

Here, authors analyze 20 years of weather data from Chittagong city in Bangladesh, including variables such as wind speed, temperature, precipitation, and surface pressure[13]. Various algorithms, including AdaBoost, XGBoost, Stacking KNN, and Stacking Random Forest, were applied to predict maximum temperature, minimum temperature, and average temperature over a long-term duration. The accuracy and performance of these algorithms were compared and analyzed in this study.

Utility companies facing a lack of access to a weather service that provides predicted hourly temperatures for short-term load forecasting is indeed a significant concern. To overcome this limitation, a temperature forecaster was developed, capable of forecasting hourly temperatures up to seven days in advance [14]. The mean absolute inaccuracy of the one-day predictions by this forecaster was found to be 1.48°F. The authors employed an Artificial Neural Network (ANN) for creating temperature predictions, allowing for adjustments in response to unusual meteorological events. However, it is important to note that training the ANN requires more computational resources compared to traditional forecasting methods.

In a study conducted in 2015, eight unique regression tree structures were employed for short-term wind speed prediction [15]. The author further compared the performance of the most effective regression tree approach against other artificial intelligence (AI) methods, including support vector regression (SVR), multi-layer perceptron (MLP), extreme learning machines, and multi-linear regression. The findings revealed that the best regression tree approach outperformed the other AI methods in wind speed prediction.

In [16], a lightweight data-driven weather forecasting model has been proposed using LSTM and TCN approaches. It compares the model's performance with classical machine learning and statistical forecasting approaches, as well as the NWP model. The LSTM and TCN models capture weather information from time-series data and are evaluated in multi-input multi-output and multi-input single-output regressions. The proposed deep learning networks show promise for accurate weather prediction using surface weather parameters.

In [5], authors examines a novel deep learning-based precipitation prediction approach. their model can adapt to modify in weather features such as temperature, wind flow, humidity, and so on. They aim to deal the limitations of traditional prediction models that only rely on target domain data by leveraging the information gained from various source domains. By adapting the prediction models to the target domain, they show upgrade the accuracy of rain precipitation prediction for multiple locations with various weather patterns.

III. Methodology

Data cleaning, or data mining, is a crucial process for extracting valuable insights from raw data and preparing it for input into mathematical models. We collected weather data from NASA Power and conducted an extensive analysis of 20 years' worth of Dhaka city weather data, spanning from January 1, 2003, to January 1, 2023 [17]. The location of the dataset is 23.8103° N and 90.4125° E. The dataset was carefully segmented, focusing on specific target values related to temperature and precipitation. With 16 extracted weather features, we aimed to predict average temperatures and precipitation accurately.

A. Dataset

Our findings from this dataset provide valuable insights into the relationships between different weather variables and

offer valuable information for weather forecasting and understanding weather patterns in the Dhaka region. The extracted features are given below:

TABLE I: Features Used in the Dataset

Features	Description
T2M	Average air temperature at 2 meters
T2MDEW	Dew point at 2 meters
T2MWET	Wet point at 2 meters
TS	Earth Skin Temperature
T2M RANGE	Merra-2 Temperature at 2 meters range
T2M MAX	Maximum temperature at 2 meters
T2M MIN	Minimum temperature at 2 meters
QV2M	Specific Humidity at 2 meters
RH2M	Relative Humidity at 2 meters
PRECTOT	Bias-corrected total precipitation at 2 meters
PS	Average surface pressure at surface
WS10M RANGE	Average wind speed 10 meters range
WS10M	Average wind speed 10 meters high
WD10M	Average wind direction 10 meters high
WS10M MAX	maximum wind speed 10 meters
WS10M MIN	Minimum wind speed 10 meters

B. Data Visualisation and Analysis

From the dataset, we have analyzed the correlation of data for precipitation and temperature. From fig-1 and fig-2 we have deciphered the density of the distribution for both of them. We have used MinMaxScaler to scale the dataset and standardize the distribution. From fig-3 heatmap, we have found the correlation between the features related to precipitation and temperature. Among the 16 features precipitation mostly correlated with dew point, specific humidity, and relative humidity whereas temperature mostly correlated with wet point and specific humidity. Figure 4 and 5 depicts the distribution of monthly precipitation and temperature. We can find that in the middle of the year, both precipitation and temperature are higher than at the end and start of the year.

IV. EXPERIMENT AND RESULTS

To assess the performance of the proposed weather data prediction model, we employed six different algorithms and evaluated their effectiveness on three target variables. The algorithms are Gradient Boosting, Ada Boost, Artificial Neural Network, Stacking Random Forest, Stacking Neural Network, and Stacking KNN. The dataset was divided into training and testing sets using a random split of 85% for training and 15% for testing.

In order to measure the performance of the models, we calculated various evaluation metrics, including accuracy, precision, recall, and F1-score, for each target variable. Higher
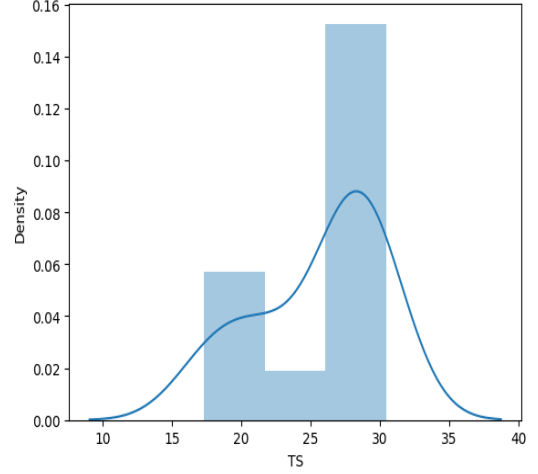
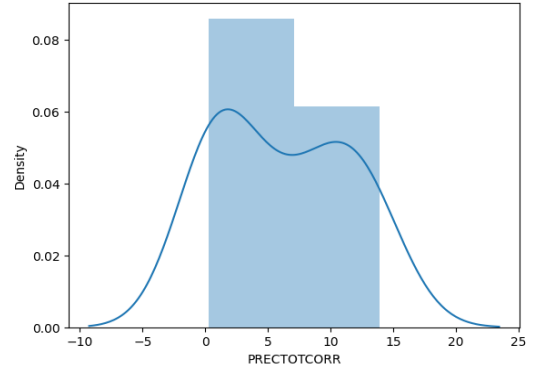

Fig. 1: Distribution plot for Temperature

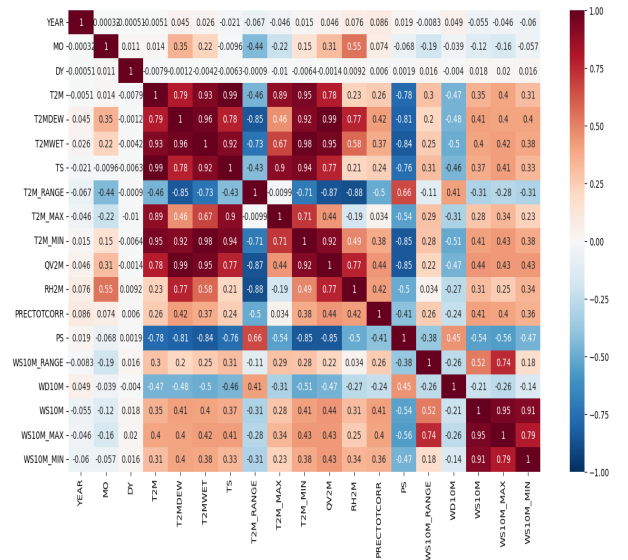Fig. 2: Distribution plot for Precipitation

Fig. 3: Correlation heatmap among the features of dataset
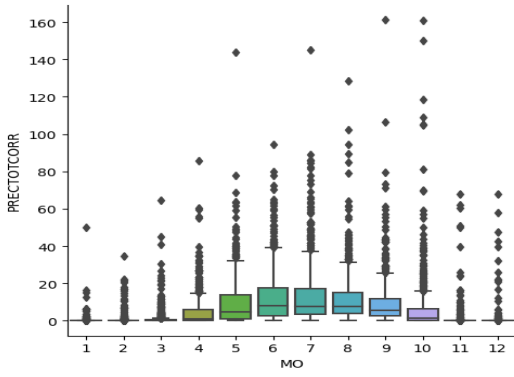
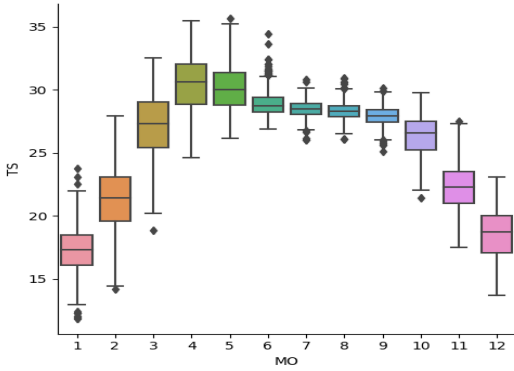Fig. 4: Monthly Precipitation Distribution from the dataset

Fig. 5: Monthly Temperature Distribution from the dataset

TABLE II: Algorithm Efficiency Overview of Precipitation

Algorithm	Accuracy	Precision	Recall	F1 score
Gradient Boost	92.11%	89%	93%	90%
Ada Boost	92.51%	89%	93%	90%
Artificial Neural Network	92.15%	88%	92%	89%
Stacking Random Forest	92.19%	89%	92%	90%
Stacking Neural Network	92.3%	89%	92%	90%
Stacking KNN	91%	88%	91%	89%
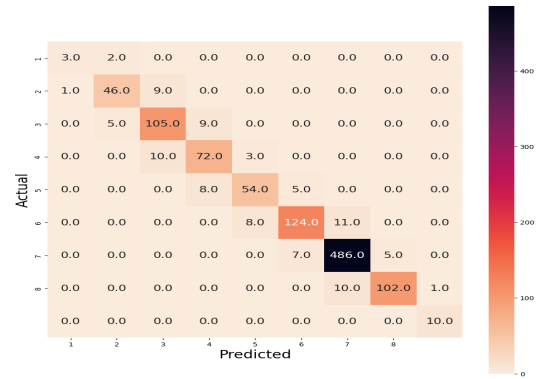
Fig. 6: Confusion Matrix for Ada Boost Algorithm (Precipitation)

values of these performance measures indicate better performance of the models. The results presented below demonstrate the performance of the algorithms on the selected target variables, providing valuable insights into their effectiveness in predicting weather data.

A. Precipitation

The performance measurements of different algorithms for precipitation extraction were analyzed. From table-1, the results showed that all the algorithms performed relatively similar in terms of accuracy, precision, recall, and F1 score. However, the algorithm that achieved the most accurate scores with a minimal difference was Ada Boost. Additionally, a confusion matrix was generated to visualize the performance of the best-performing algorithm in predicting precipitation. From the figure-6, in the confusion matrix, the correctly predicted classes are located along the diagonal. This matrix provides an overview of the algorithm's performance in terms of correctly and incorrectly classified instances.

B. Temperature

In the case of average temperature prediction, the same algorithms have shown close pattern results. On the other hand, the AdaBoost algorithm performed well in this particular task, compared to other algorithms (Table II). These results highlight the superior performance of certain algorithms in

accurately predicting average temperature. Moreover, from the figure-7, in the confusion matrix, the temperature's correctly predicted classes are situated along the diagonal. The provided matrix offers an overview of how the algorithm performed by displaying the count of instances accurately classified and those that were classified incorrectly.

V. FUTURE DIRECTION

Our forthcoming research endeavors will center on the integration of advanced models and techniques, with the primary objective of augmenting the precision and dependability of weather forecasting. This will encompass an exploration of ensemble learning methods, deep learning algorithms, and hybrid models that amalgamate multiple approaches. Additionally, considering the substantial data requirements of data-driven methodologies, a new research direction involves efficient feature selection techniques to mitigate computational costs. Furthermore, the incorporation of supplementary meteorological data sources and the exploration of feature engineering techniques to extract more informative features for prediction represent promising avenues of inquiry. The

TABLE III: Algorithm Efficiency Overview of Temperature

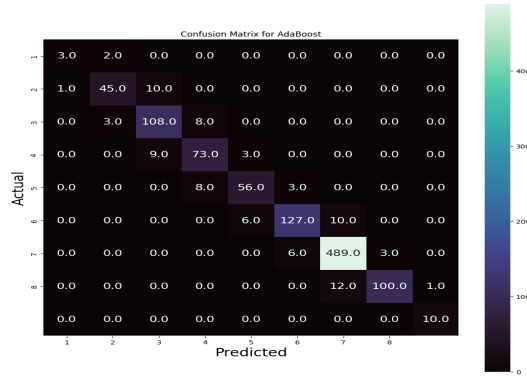| Algorithm	Accuracy	Precision	Recall	F1 score
Gradient Boost	91.05%	91%	91%	91%
Ada Boost	91.7%	92%	92%	92%
Artificial Neural Network	90.2%	90%	90%	90%
Stacking Random Forest	91.3%	91%	91%	91%
Stacking Neural Network	91.4%	92%	91%	91%
Stacking KNN	90.96%	91%	90.6%	90.4%

Fig. 7: Confusion Matrix for Ada Boost Algorithm (Temperature)

recent domain adaptation for unsupervised learning has a great potential in forecasting. Moreover, the application of big data analytics holds significant potential in effectively managing the vast volume, velocity, and variety of weather data. This entails leveraging distributed computing frameworks such as Apache Hadoop and Apache Spark for processing and storing extensive datasets. Furthermore, the utilization of Internet of Things (IoT) devices and sensor networks facilitate data collection from diverse sources, encompassing weather stations, satellites, drones, buoys, and ground-based sensors. The model's performance evaluation will be bolstered through validation using a wide range of challenging weather datasets.

VI. CONCLUSION

In conclusion, the algorithms subjected to testing in this study have exhibited exceptional performance, surpassing conventional weather prediction models. Moreover, the identification of significant weather variables and their corresponding scores has furnished valuable insights into meteorological patterns. Furthermore, our investigation delves into the exploration of influential aspects of weather events, thereby shedding light on critical factors including wind speed at specific altitudes and directions, absolute and relative humidity, precipitation, and surface pressure. Each of these variables is attributed significant scores, underscoring their profound influence on weather patterns and forecasting precision. Collectively, this research emphasizes the potential of the tested algorithms to enhance the accuracy of weather prediction and augment our comprehension of weather phenomena. Consequently, these findings pave the way for advancements in weather forecasting and the formulation of more robust prediction models.

REFERENCES

[1] M. S. Islam, M. S. Rahman, and M. A. Amin, "Beat based realistic dance video generation using deep learning," in *2019 IEEE International Conference on Robotics, Automation, Artificial-intelligence and Internet-of-Things (RAAICON)*, 2019.

[2] F. A. Tumpa, M. S. Rahman, and M. S. Islam, "Utilizing genetic evolution to enhance cellular automata for accurate image edge detection," EasyChair, Tech. Rep., 2023.

[3] F. Di Nunno, F. Granata, Q. B. Pham, and G. de Marinis, "Precipitation forecasting in northern bangladesh using a hybrid machine learning model," *Sustainability*, vol. 14, no. 5, p. 2663, 2022.

[4] M. S. Islam, A. S. M. J. Hasan, M. S. Rahman, J. Yousuf, M. S. I. Sajol, and F. A. Tumpa, "Location agnostic source-free domain adaptive learning to predict solar power generation," in *2023 IEEE International Conference on Energy Technologies for Future Grids, Wollongong, Australia*, 2023, in press.

[5] M. S. Islam, M. S. Rahman, M. S. U. Haque, F. A. Tumpa, M. S. Hossain, and A. A. Arabi, "Location agnostic adaptive rain precipitation prediction using deep learning," in *IEEE 9th International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE)*, 2023, in press.

[6] M. S. Hossain, M. S. Islam, M. S. U. Haque, and M. S. Rahman, "Gait phase classification from semg in multiple locomotion mode using deep learning," in *9th International Congress on Information and Communication Technology*, 2024, in press.

[7] S. Islam *et al.*, "Learning style of gestures to synthesize dance from music," 2020.

[8] A. S. M. Jahid Hasan, J. Yusuf, M. S. Rahman, and M. S. Islam, "Electricity cost optimization for large loads through energy storage and renewable energy," in *2023 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)*, 2023, pp. 46–50. DOI: 10.1109/ICICT4SD59951.2023.10303409.

[9] A. S. M Jahid Hasan, M. S. Rahman, M. S. Islam, and J. Yusuf, "Data driven energy theft localization in a distribution network," in *2023 International Conference on Information and Communication Technology for Sus-*

tainable Development (ICICT4SD), 2023, pp. 388–392. DOI: 10.1109/ICICT4SD59951.2023.10303520.

[10] A. H. Chowdhury, A. Al Arabi, and M. A. Amin, "In search of a low-cost iot system for real-time environment monitoring," in *2019 IEEE International Conference on Robotics, Automation, Artificial-intelligence and Internet-of-Things (RAAICON)*, 2019, pp. 87–92. DOI: 10.1109/RAAICON48939.2019.40.

[11] S. Jamal, T. Bappy, R. Pervin, and A. S. A. Rabby, "Weather status prediction of dhaka city using machine learning," in Jan. 2021, pp. 293–304, ISBN: 978-981-15-6875-6. DOI: 10.1007/978-981-15-6876-3_22.

[12] T. Tang, D. Jiao, T. Chen, and G. Gui, "Medium-and long-term precipitation forecasting method based on data augmentation and machine learning algorithms," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 1000–1011, 2022.

[13] M. S. Tahsin, M. Al Karim, M. U. Ahmed, Y. Rahman, F. Tafannum, and S. Abdullah, "Comparative analysis of weather prediction using ensemble learning models and neural network," in *2021 19th OITS International Conference on Information Technology (OCIT)*, IEEE, 2021, pp. 325–330.

[14] A. Khotanzad, M. H. Davis, A. Abaye, and D. J. Maratukulam, "An artificial neural network hourly temperature forecaster with applications in load forecasting," *IEEE Transactions on Power Systems*, vol. 11, no. 2, pp. 870–876, 1996.

[15] A. Troncoso, S. Salcedo-Sanz, C. Casanova-Mateo, J. Riquelme, and L. Prieto, "Local models-based regression trees for very short-term wind speed prediction," *Renewable Energy*, vol. 81, pp. 589–598, 2015.

[16] P. Hewage, M. Trovati, E. Pereira, and A. Behera, "Deep learning-based effective fine-grained weather forecasting model," *Pattern Analysis and Applications*, vol. 24, no. 1, pp. 343–366, 2021.

[17] P. Stackhouse. "Nasa power." Accessed: 2021-10-16. (), [Online]. Available: https://power.larc.nasa.gov/dataaccess-viewer/.