INTEGRATING EARLY CAREER STEM RESEARCH THROUGH MACHINE LEARNING BY THE NASA CLIMATE CHANGE RESEARCH INITIATIVE (CCRI)

Dr. Alicia Joseph[1], Gabriela Himmele[2], Matthew Pearce[3], Nicholas Pinder[4], Aditya Jain[5], Kari Espada[6], Joy Conway[7], Megan Mason[1], Dr. Carrie Vuyovich[1]

[1]NASA Goddard Space Flight Center, [2]West Virginia University, [3]NASA Goddard Institute of Space Studies, [4]University of South Florida, [5]The Harker School, [6]Virginia Space Grant Consortium, [7]Plum Point Middle School,

IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium | USGov | DOI: 10.1109/IGARSS53475.2024.10642493

ABSTRACT

Satellites are particularly well-suited to provide spatially distributed observations of global snow. For hydrological research and applications, Snow Water Equivalent (SWE) is the most important observation, but it is also our biggest gap in snow remote sensing. Currently, no satellite sensor has the ability to measure SWE globally at the accuracy, resolution and frequency needed, because of a number of factors that impact the signals such as forests, mountains, clouds and the snow characteristics themselves. The NASA Climate Change Research Initiative (CCRI) SnowEx team uses machine learning approaches to attempt to address the unanswered questions of snow science. The team has collaborated with other similar NASA wide programs and leveraged skills and resources which led to the formation of a community machine learning (ML) working group.

Index Terms— Machine Learning, Climate Change, STEM, SnowEx, Snow Water Equivalent (SWE).

1. INTRODUCTION

A good understanding of how snow water equivalent is changing due to climate change is needed to determine the impacts of these changes on things like vegetation, agriculture, Earth's albedo, and many other important hydrologic processes.

2. NASA CLIMATE CHANGE RESEARCH INITIATIVE (CCRI)

The Climate Change Research Initiative (CCRI) objective is to understand all facets of climate change, allowing diverse students, faculty, and STEM-centered individuals to contribute to understanding an ever-growing world crisis.

The NASA CCRI program in the fall and spring comprises of a STEM education specialist and a graduate student intern working with a project research mentor. During the fall session, the emphasis is on gaining the research skills necessary to take back to the classroom and to help develop a STEM curriculum. In the spring session, the team continue to work on the research project with more emphasis on the development of the STEM education curriculum in preparation for the arrival of the summer high school intern and the undergraduate intern. The team has the opportunity to impact NASA projects and missions at an early stage through their hands-on experience.

Fig. 1) SnowEx CCRI team at final roundtable presentation at GISS, NY.

3. INTEGRATING PROGRAMS

Connections and collaboration with NASA unique student engagement programs allow the building of relationships and thereby integrating internship programs within NASA. The early career programs are a good example of this integration. The SnowEx CCRI and soil moisture Student Airborne

IEEE IGARSS 2024 Paper #6010, Page 1 of 5
FOR REVIEW ONLY: CONTENTS MUST REMAIN CONFIDENTIAL

Research Program (SARP) interns work together and help each other on projects. One of the results of this interaction is the formation of the machine learning working group. Through the engagement with SMD, OSTEM, HQ program and Universities, we have recently added interagency partnerships to the machine learning working group. The outcome of this integration was the combined attendance at AGU for both CCRI and SARP interns. Traceable benefits from program integration are observed.

Fig. 2) SnowEx CCRI team and soil moisture SARP team at AGU, SAN Francisco (Dec. 2023).

4. SNOWEX

SnowEx is an intensive field campaign over several years to determine the best remote sensing technologies to observe snow through the growing season over different land covers. SnowEx uses a combination of remote sensing techniques, computer models, and ground observations for validation to

monitor changes in these parameters. There are different types of snow, and each snow type is important to the hydrology and climatology of the Earth.

Fig. 3) SnowEx CCRI team at SARP East on soil moisture ComRad truck at Virginia Commonwealth University, VA (July 2023).

5. MACHINE LEARNING (ML)

Machine learning is a subfield of Artificial intelligence (AI) that uses a model involving heavy mathematics and intense algorithms to learn from data and improve their performance on a specific task over time without being explicitly programmed. Overall, machine learning allows computers to identify patterns, make predictions, and adapt their behavior based on experience and input data.

This NASA SnowEx CCRI team started using machine learning to predict how radar backscatters varied against different kinds of snowpack. The current method of estimating radar backscatter includes using a Dense Media Radiative Transfer Multi-Layer Model (DMRT-ML), although the team had been looking to see how machine learning can outperform the model.

Fig. 4) SnowEx CCRI team at SARP East final presentation at William Mary University, VA (August 2023)

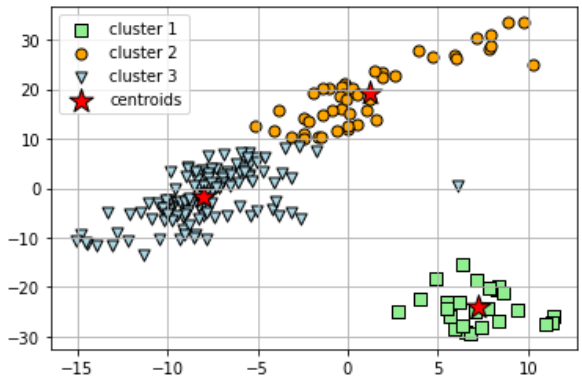
IEEE IGARSS 2024 Paper #6010, Page 2 of 5
FOR REVIEW ONLY: CONTENTS MUST REMAIN CONFIDENTIAL

6. ML METHODS

A research question of this team is as follows:
How can we accurately predict Snow Water Equivalent (SWE) measurements in areas that have forest cover spatially and temporally using the SWESARR instrument, with the goal of understanding how SWE will change with climate change?

Machine learning will be an integral part of this process, which will aid in quantifying nonlinear relationships between parameters which are required to predict SWE.

Fig 5) The above figure shows unsupervised clustering of SWESARR values, which will better inform the time series forecasting algorithm.



6.1. Unsupervised

Unsupervised machine learning is a type of machine learning algorithm that does not require human intervention/supervision and utilizes unlabeled data to group or structure the data in certain ways to find hidden patterns. This is seen in K-means clustering where the data is partitioned into a certain defined number of clusters (i.e., K), and then each data point is then paired with another that is nearest to it.

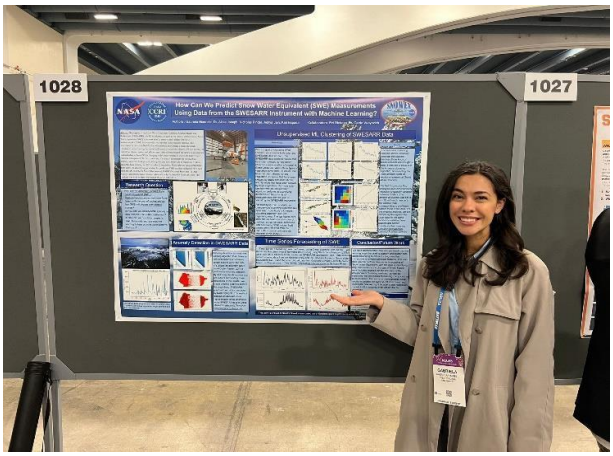


Fig. 6) SnowEx CCRI team graduate student, Gabriela Himmele, presenting at AGU (Dec. 2023)

6.2. Time Series

A sub-research question for NASA SnowEX is to understand the temporal variations of snow water equivalent (SWE). To address this question a Time Series Forecasting algorithm using deep learning was constructed to identify nonlinear patterns within the snow water equivalent (SWE).

6.3. Anomaly Detection

Anomaly detection is implemented using machine learning, and the algorithm chosen for this work was the Local Outlier Factor (LOF) algorithm. This project used an anomaly detection algorithm to build off the K-means clustering code to visualize spatially where the data is skewed by smaller foliage and tree canopy cover. LOF is an unsupervised learning anomaly detection algorithm that identifies data points that deviate from their "parent cluster." Scikit-Learn's LOF algorithm was manipulated to achieve the goal of determining where these parent clusters are located. Using more methods under this Scikit-Learn library allows the extraction of data about points that deviated from their local cluster, which can be used later.
After processing the data, feeding the files through the algorithm, and manipulating the dataset to be represented graphically, identification of the areas where the SWESARR instrument recorded measurements that have been skewed by foliage are made.

7. STEM INTEGRATING ML IN THE CURRICULUM

This team includes a high school teacher whose capstone project brings all the aspects of types of data, data collection categories, and data analysis together with a visual representation of the laboratory investigations and activities experienced. The capstone project explores machine learning models used to analyze the data that has been collected through binary examples, and data visualization recaps. Students will ultimately create a campaign to inform their peers and community. Students will learn that machine learning models can help to identify the results of data and determine next steps for further research.

The student experience aims to invoke a deeper understanding of climate change and climate research. Students' progress through the unit curriculum exploring topics in a hands-on manner to simulate the field expert's experience. The lessons consist of ecosystem experiences, revisiting the hydrologic cycle, rainfall and soil moisture data.

IEEE IGARSS 2024 Paper #6010, Page 3 of 5
FOR REVIEW ONLY: CONTENTS MUST REMAIN CONFIDENTIAL

This ties the unit together through the extension of exposing students to Jupyter notebooks and machine learning models using SnowEx datasets.

8. Conclusion

The CCRI team plan to add the physical characteristics of snow and snowflakes to the ML models to improve estimation of SWE. The team plans to expand the ML working group and continue to integrate programs within the agency. The overall hands-on experience and integration of programs create a unique NASA opportunity for student engagement. The retention rate of interns has increased. A pipeline to careers at NASA is fostered through the nurturing of this unique internship experience. Early demonstrations of the scientific processes and engagement with NASA projects and missions from the initial stages are created. The opportunity to add ML approaches to the curriculum through working with high school teachers gives the STEM field an advantage.

Fig. 7) CCRI members Lily Donaldson, Sangmin Pak, Gabriela Himmele, Dr. Alicia Joseph, and Dr. Ryan Miller at AI Expo at DC Convention Center May 7-8, 2024.

Fig. 8) Sangmin Pak, Lily Donaldson, Dr. Alicia Joseph, and Gabriella Himmele at AI Expo at DC Convention Center May 7-8, 2024.

IEEE IGARSS 2024 Paper #6010, Page 4 of 5
FOR REVIEW ONLY: CONTENTS MUST REMAIN CONFIDENTIAL

9. REFERENCES

[1] E. Kim et al., "NASA's snowex campaign: Observing seasonal snow in a forested environment," 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 2017, pp. 1388-1390, doi: 10.1109/IGARSS.2017.8127222.

[2] NASA SnowEx https://snow.nasa.gov/campaigns/snowex

[3] NASA GISS Climate Change Research Initiative https://www.giss.nasa.gov/edu/ccri/

[4] Yueh, S. H., Dinardo, S. J., Akgiray, A., West, R., Cline, D. W., & Elder, K. (2009). Airborne Ku-band polarimetric radar remote sensing of terrestrial snow cover. IEEE Transactions on Geoscience and Remote Sensing, 47(10), 3347-3364.

[5] National Aeronautics and Space Administration. NASA SnowEx Science Plan: Assessing Approaches for Measuring Water in Earth's Seasonal Snow (2020).

IEEE IGARSS 2024 Paper #6010, Page 5 of 5
FOR REVIEW ONLY: CONTENTS MUST REMAIN CONFIDENTIAL