AWS using utilizing machine learning to classify clouds from satellite data for process-oriented climate model assessment implementing for cloud computing

[1] K.Meghana, [2] R.Girija, [3] Nukala Saiteja, [4] Praneel Deva
[1,2,3]UG Scholar, Department of Computer Science and Engineering, St. Martin's Engineering College, Secunderabad, Telangana, India, 500100
[4]Assistant Professor, Department of Computer Science and Engineering, St. Martin's Engineering College, Secunderabad, Telangana, India, 500100
praneeldevacse@gmail.com

Abstract:

Road Clouds play a key role in regulating climate change but are difficult to simulate within Earth system models (ESMs). Improving the representation of clouds is one of the key tasks towards more robust climate change projections. This study introduces a new machine-learning based framework relying on satellite observations to improve understanding of the representation of clouds and their relevant processes in climate models. The proposed method is capable of assigning distributions of established cloud types to coarse data. It facilitates a more objective evaluation of clouds in ESMs and improves the consistency of cloud process analysis. The method is built on satellite data from the MODIS instrument labelled by deep neural networks with cloud types defined by the World Meteorological Organization (WMO), using cloud type labels from CloudSat as ground truth. The method is applicable to datasets with information about physical cloud variables comparable to MODIS satellite data and at sufficiently high temporal resolution.We recommend outputting crucial variables required by our method for future ESM data evaluation. This will enable the use of labelled satellite data for a more systematic evaluation of clouds in climate models. Keywords: Cloud , Cloud Computing , Machine Learning , Data Mining , Support Vector Machine (SVM) , Logistic Regression.

Keywords:Cloud,CloudComputing,MachineLearning,Data Mining , Support Vector Machine (SVM) , Logistic Regression.

1.INTRODUCTION

EARTH system models (ESMs, also referred to as climate models) are important tools not only to improve our understanding of present-day climate but also to project climate change under different plausible future scenarios. For this, climate models have been continuously improved and extended over the last decades from relatively simple atmosphere-only models to complex state-of-the-art ESMs including many processes such as the biogeochemical cycle, e.g. those participating in the latest (sixth) phase of the Coupled Model Intercomparison Project (CMIP6, [1]).

This increasing complexity of models is needed to represent key feedbacks that affect climate change, but also requires innovative and comprehensive model evaluation and analysis approaches to assess the performance of these models [2], given the increase in the number of tuneable parameters in the models. In particular, the simulation of clouds and their interactions with the climate system remain major challenges for ESMs [3]. As a consequence, cloud feedback mechanisms such as the shortwave radiative effect of low clouds, which are critical for long-term climate projections, have proven to be hard to quantify confidently [4]–[6].

Furthermore, the representation of clouds has been identified as one of the primary sources of intermodel spread in state-of-the-art ESMs [7]. An improved representation of cloud processes is therefore an essential component in addressing these issues [8]–[10].

2. LITERATURE SURVEY

A."Assessment of CMIP6 Cloud Fraction and Comparison with Satellite Observations," The seasonal and regional variations of cloud fractions are compared across two generations of global climate model ensembles, specifically, the Coupled Model Intercomparison Project-5 (CMIP5) and CMIP6, through the historical period in terms of skills and multimodel agreement. We find a wider spread of historical cloud fraction changes in the CMIP6 than was simulated by the CMIP5. The global mean cloud fractions of CMIP6 increased by about 4.5% from the CMIP5, which attributed to greater changes in the northern hemisphere than in the southern hemisphere. The CMIP6 cloud fractions in recent years are validated with the CALIPSO_CLOUSAT observations to understand the cloud fraction uncertainties in CMIP6 models. The CMIP6 ensemble mean of cloud fractions compares well with the observations with a mean difference of 0.5% in lower altitudes. The CMIP6 cloud fractions are higher than the observations at higher latitudes in both hemispheres in the upper troposphere, and the biases vary from one model to another. The spatial difference between the ensemble and observations is further revealed over the tropics: where the model displays a 3% higher bias. In addition, we observed a significant trend occuring in the northern hemisphere since the mid-20th century using calculations of cloud fraction trends based on the robust regression technique. Finally, we reduce the differences between the model and observations by applying a simple regression technique. The results exemplify that the model and modified observations compare well, with the root mean square value decreased by nearly 28%, and the correlation increased significantly."Observational constraints on low cloud feedback reduce uncertainty of climate sensitivity," Marine low clouds strongly cool the planet. How this cooling effect will respond to climate change is a leading source of uncertainty in climate sensitivity, the planetary warming resulting from CO2 doubling. Here, we observationally constrain this low cloud feedback at a near-global scale. Satellite observations are used to estimate the sensitivity of low clouds to interannual meteorological perturbations. Combined with model predictions of meteorological changes under greenhouse warming, this permits quantification of spatially resolved cloud feedbacks. We predict positive feedbacks from midlatitude low clouds and eastern ocean stratocumulus, nearly unchanged trade 6 cumulus and a near-global marine low cloud feedback of $0.19 \pm 0.12\,W\,m^{-2}\,K^{-1}$ (90% confidence). These constraints imply a moderate climate sensitivity (~3 K). Despite improved midlatitude cloud feedback simulation by several current-generation climate models, their erroneously positive trade cumulus feedbacks produce unrealistically high climate sensitivities. Conversely, models simulating erroneously weak low cloud feedbacks produce unrealistically low climate sensitivities."The Cloud Feedback Model Intercomparison Project (CFMIP) contribution to CMIP6," The primary objective of CFMIP is to inform future assessments of cloud feedbacks through improved understanding of cloud–climate feedback mechanisms and better evaluation of cloud processes and cloud feedbacks in climate models. However, the CFMIP approach is also increasingly being used to understand other aspects of climate change, and so a second objective has now been introduced, to improve understanding of circulation, regional-scale precipitation, and non-linear changes. CFMIP is supporting ongoing model inter-comparison activities by coordinating a hierarchy of targeted experiments for CMIP6, along with a set of cloud-related output diagnostics. CFMIP contributes primarily to addressing the CMIP6 questions "How does the Earth system respond to forcing?" and "What are the origins and consequences of systematic model biases?" and supports the activities of the WCRP Grand Challenge on Clouds, Circulation and Climate Sensitivity.

3. PROPOSED METHODOLOGY

We provide a novel method for evaluating ESMs that aims to alleviate some of the perceived drawbacks of using traditional observational data while also making process-oriented cloud assessment in climate models easier. We make advantage of preexisting information about the features of various cloud classes, which are derived from the World Meteorological Organization's (WMO) taxonomy of cloud types. Utilizing this earlier information, cloud operations may be emphasized for further analysis. Our strategy applies machine learningbased cloud categorization techniques recently developed for satellite data to climate models. Although machine learning-based cloud categorization is not a novel concept [e.g. 16], it has only recently been practical for largescale applications because of the rise in processing power that is now accessible and the fact that the various approaches have varied characteristics. The supervised vs unsupervised nature of categorization techniques is a key differentiator. Whereas the latter seeks to automatically discover unique new classes, the former depends on already given classes. While unsupervised approaches provide the user more flexibility over the composition of the classes, supervised classification makes the assumption that the classes allocated to them are appropriate for the task at hand. As a consequence, supervised approaches need a set of labeled data yet enable interpretation of the final findings without the need for extra analytic stages. Unsupervised approaches are better if finding as different classes as feasible is the aim, or if there are no accessible previously tagged data.

One of the WMO's key functions is to monitor and analyze global climate trends, working closely with organizations such as the Intergovernmental Panel on Climate Change (IPCC) to assess the impact of climate change. It also provides scientific data that supports international agreements like the Paris Agreement, which aims to limit global warming. In addition, the WMO plays a vital role in water resource management, aviation and maritime safety, and improving meteorological capabilities in developing countries. Through its reports and initiatives, the WMO continues to be at the forefront of addressing climate-related challenges and ensuring a safer, more resilient world.

Advantages: Machine learning models can learn complex patterns in satellite imagery, leading to more accurate cloud classification compared to traditional methods. Automated classification speeds up the process, allowing analysis of large volumes of satellite data in a relatively shorter time frame. Machine learning enables the analysis of high-resolution satellite imagery, capturing detailed cloud features that might be missed in manual analysis. Accurate classification provides detailed information about cloud types, their spatial distribution, and temporal variations, enriching the understanding of cloud behavior and its impact on climate dynamics.

4. EXPERIMENTAL ANALYSIS

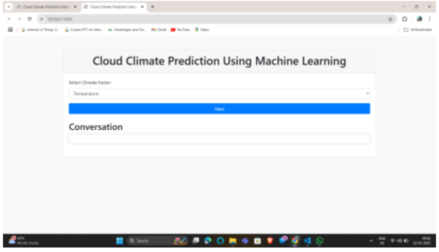

Fig: 6.1 Cloud Climate Prediction Using Machine Learning

The web-based application "Cloud Climate Prediction Using Machine Learning" is designed to predict cloud climate conditions using machine learning models. Running on a local server (127.0.0.1:5000), the interface is simple and user-friendly, allowing users to select different climate factors from a dropdown menu.This system could be valuable for meteorologists, environmental researchers, and individuals interested in weather forecasting. By analyzing historical data and real-time inputs, the model likely predicts future trends, helping users make informed decisions.

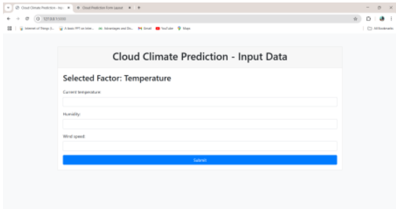

Fig: 6.2 Cloud Climate Prediction for Temperature

In this section, users are prompted to input key atmospheric parameters that influence precipitation. These include Current temperature, Humidity, and Wind Speed. Once the required values are entered into the respective fields, users can click the "Submit" button, which likely triggers the machine learning model to analyze the data and generate a Temperature prediction. This could be particularly useful for weather analysts, researchers, and individuals who need real-time climate predictions for planning and decisionmaking.

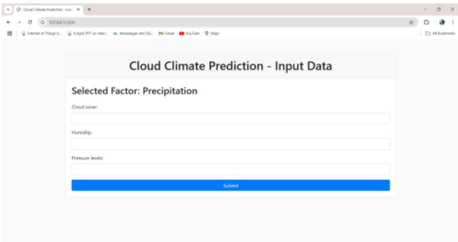

Fig: 6.3 Cloud Climate Prediction for Precipitation

The application shows to be designed for cloud climate prediction, specifically for inputting data related to precipitation.The webpage has a clean and simple layout with a centered form. The form consists of three labeled input fields: 1. Cloud cover – Likely meant to input data about the extent of cloud coverage. 2. Humidity – For entering the humidity level. 3. Pressure levels – For inputting atmospheric pressure data. At the bottom of the form, there is a blue "Submit" button for submitting the entered data.

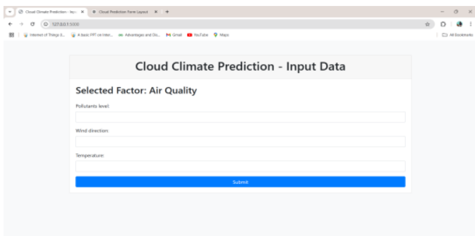

Fig: 6.4 Cloud Climate Prediction for Air Quality

The application is focused on cloud climate prediction, specifically for assessing air quality.At the top, the page has a heading that reads "Cloud Climate Prediction - Input Data". The form consists of three labeled input fields: 1. Pollutants level –To input the concentration of pollutants in the air. 2. Wind direction – For specifying the direction of the wind, which can affect air quality. 3. Temperature – For entering the current temperature, which influences atmospheric conditions. At the bottom of the form, there is a blue "Submit" button for submitting the entered data.

5. CONCLUSION

AWS's implementation of machine learning to classify clouds from satellite data presents a significant advancement in process-oriented climate model assessment. By leveraging cloud computing, this approach enhances scalability, efficiency, and accessibility for climate researchers and meteorologists. The integration of AWS services such as Amazon SageMaker, AWS Lambda, and Amazon S3—facilitates real-time data processing, model training, and inference at a global scale.This innovative solution enables more accurate cloud classification, improving climate models and supporting better predictions of weather patterns and climate change impacts. Moreover, the flexibility of AWS cloud computing ensures cost-effective and secure deployment of machine learning workflows, fostering collaboration among scientific communities. As the technology evolves, continued enhancements in AI models and cloud infrastructure will further refine climate assessments, contributing to more sustainable and data-driven decision-making in climate research.

REFERENCES

[1] P. P. Vignesh, J. H. Jiang, P. Kishore, et al., "Assessment of CMIP6 Cloud Fraction and Comparison with Satellite Observations," Earth and Space Science, vol. 7,no. 2, pp. 1–21, 2020.

[2] J.-L. Dufresne and S. Bony, "An assessment of the primary sources of spread of global warming estimates from coupled atmosphere–ocean models," Journal of Climate, vol. 21, no. 19, pp. 5135–5144, 2008.

[3] M. D. Zelinka, T. A. Myers, D. T. McCoy, et al., "Causes of Higher Climate Sensitivity in CMIP6 Models," Geophysical Research Letters, vol. 47, no. 1, pp. 1– 12, 2020.

[4] S. Bony, B. Stevens, D. M. Frierson, et al., "Clouds, circulation and climate sensitivity," Nature Geoscience, vol. 8, no. 4, pp. 261–268, 2015.

[5] T. Schneider, J. Teixeira, C. S. Bretherton, et al., "Climate goals and computing the future of clouds," Nature Publishing Group, vol. 7, no. 1, pp. 3–5, 2017.

[6] K. D. Williams and M. J. Webb, "A quantitative performance assessment of cloud regimes in climate models," Climate Dynamics, vol. 33, no. 1, pp. 141–157, 2009.

[7] A. Lauer, V. Eyring, M. Righi, et al., "Benchmarking CMIP5 models with a subset of ESA CCI Phase 2 data using the ESMValTool," Remote Sensing of Environment, vol. 203, pp. 9–39, Dec. 2017.

[8] H. Kawai, S. Yukimoto, T. Koshiro, et al., "Significant improvement of cloud representation in the global climate model mri-esm2," Geoscientific Model Development, vol. 12, no. 7, pp. 2875–2897, 2019.

[9] C. Jakob, "An improved strategy for the evaluation of cloud parameterizations in gcms," Bulletin of the American Meteorological Society, vol. 84, no. 10, pp. 1387 – 1402, 2003.

[10] S. Rasp, H. Schulz, B. B. Stevens, and S. Bony, "Combining crowd-sourcing and deep learning to understand meso-scale organization of shallow convection," in AGU Fall Meeting 2019, AGU, 2019.

[11]V. Zantedeschi, F. Falasca, A. Douglas, R. Strange, M. Kusner, and D. Watson-Parris, "Cumulo: A dataset for learning cloud classes," in Proceedings of the NeurIPS 2019 Workshop Tackling Climate Change with Machine Learning, Climate Change AI, 2019. [12] L. Denby, "Discovering the Importance of Mesoscale Cloud Organization 61 Through Unsupervised Classification," Geophysical Research Letters, vol. 47, no. 1, pp. 1–10, 2020.

[12]T. Kurihana, E. Moyer, R. Willett, D. Gilton, and I. Foster, "Data-Driven Cloud Clustering via a Rotationally Invariant Autoencoder," IEEE Transactions on Geoscience and Remote Sensing, pp. 1–25, 2021. arXiv: 2103.04885.

[13]C. Zhang, X. Zhuge, and F. Yu, "Development of a high spatiotemporal resolution cloud-type classification approach using Himawari-8 and CloudSat," International Journal of Remote Sensing, vol. 40, no. 16, pp. 6464–6481, 2019.

[14]W. J. Marais, R. E. Holz, J. S. Reid, and R. M. Willett, "Leveraging spatial textures, through machine learnThis article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2023.3237008 This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/ IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING 15 ing, to identify aerosols and distinct cloud types from multispectral observations," Atmospheric Measurement Techniques, vol. 13, no. 10, pp. 5459–5480, 2020.

[15]J. Lee, R. C. Weger, S. K. Sengupta, and R. M. Welch, "A Neural Network Approach to Cloud Classification," IEEE Transactions on Geoscience and Remote Sensing, vol. 28, no. 5, pp. 846–855, 1990.

[16]W. B. Rossow and R. A. Schiffer, "Advances in Understanding Clouds from ISCCP," Bulletin of the American Meteorological Society, vol. 80, no. 11, pp. 2261–2287, 1999.

[17]C. Jakob and G. Tselioudis, "Objective identification of cloud regimes in the Tropical Western Pacific," Geophysical Research Letters, vol. 30, no. 21, pp. 1–4, 2003.

[18]N. D. Gordon, J. R. Norris, C. P. Weaver, and S. A. Klein, "Cluster analysis of cloud regimes and characteristics dynamics of midlatitude synoptic systems in observations and a model," Journal of Geophysical Research D: Atmospheres, vol. 110, no. 15, pp. 1–13, 2005.

[20] R. Pincus, S. Platnick, S. A. Ackerman, R. S. Hemler, and R. J. Patrick Hofmann, "Reconciling simulated and observed views of clouds: MODIS, ISCCP, and the limits of instrument simulators," Journal of Climate, vol. 25, no. 13, pp. 4699–4720, 2012.