Evaluation and Implementation of Various Bayesian approaches to model predictions of future climate change

Saswata Lahiri
Depatment of Data Science and Business Systems,SRM Institute of Sceience and Technology
Chennai, India
sl9039@srmist.edu.in

Aagam Shah
Depatment of Data Science and Business Systems,SRM Institute of Sceience and Technology
Chennai, India
as4094@srmist.edu.in

Dr. T Veeramakali
Depatment of Data Science and Business Systems,SRM Institute of Sceience and Technology
Chennai, India
veeramat@srmist.edu.in

***Abstract*— Climate change is a major environmental issue that affects the entire planet. Accurate and reliable predictions of future climate change are essential to inform policy decisions and to develop effective mitigation and adaptation strategies. Bayesian modeling techniques have been shown to be effective in predicting future climate change by combining different sources of information, including historical data, climate models, and expert knowledge. This project aims to evaluate and implement various Bayesian approaches to model predictions of future climate change. The project will involve web scraping for data acquisition and preparation, model development using Bayesian methods, model evaluation, and implementation of the models. The project will use different sources of data, including historical climate data, climate models, and expert knowledge. The results of the project will help to improve the accuracy and reliability of climate change predictions and will inform policy decisions and the development of mitigation and adaptation strategies. The use of web scraping techniques will allow for the acquisition of a large and diverse set of data, enhancing the robustness and accuracy of the models.**

Keywords—***Bayesian Neural Network, Machine learning, Bayesian Approaches, climate change prediction***

I. Introduction

Research by various organizations has indicated that climate change will always have an adverse effect on the environment and our lives as long as we do not take proper measures to protect our enviornment.

In this line of thought, human activities are widely blamed for escalating effects of global climate change round the world (Hillel & Rosenzweig 2010). Only time will tell whether taming global climate change is feasible or impossible The impact of global climate change has been felt in every a part of the planet. consistent with United Nations Framework Convention on global climate change (UNFCCC), Asia, Africa and Latin America are among the regions of the planet, which have severely been suffering from the scourge. during a 2010 survey administered by global climate change Secretariat, Africa is under the pressure of global climate change and remains susceptible to these effects. Using BNNs in global climate change prediction involves training the network on historical climate data then using it to predict future climate conditions. The advantage of BNNs is that they will incorporate prior knowledge and expert opinions, which may improve the accuracy of the predictions. Additionally, the uncertainty estimation provided by BNNs allows decision-makers to form informed decisions supported the extent of confidence within the predictions.

II. Literature Survey

[1] during this study, they assessed the performance of three climate models by extracting and comparing their correlation pattern and RMSE(Root Mean Squared Value). CSIRO, HadCM3 and CCMA were ready to simulate both rainfall and temperature in Asia Pacific. Hindcasting was used as a way of testing the mathematical model. All estimated inputs of any past events have been entered into the model to ascertain how well the output matches the known results. This conceptual model was utilized in

evaluating the vulnerability of water fish habitats to the consequences of global climate change in Canada. an identical framework are often developed to assess the projected impacts of global climate change on aquatic environments within the Asia Pacific.

[2] They presented posterior inferences using four subsets of the Jones data which were corresponding to the different time periods mentioned in the paper. The clearly successively shorter periods of time have been suggested in the study since any trend or pattern in the global climate change are much more prominent and clear in the latter time period.

[3] The annual time-series of multi-model-weighted ensemble means and their 5–95% percentiles of probabilistic predictions of global mean SATs. The mean values of BF and EM are very close and similar to each other. They are larger than AEM where the difference increases with a maximum about 0.3 K. The 5% percentiles of the three predictions are very close to each other while the 95% percentiles of BF and EM are larger than that of AEM, which shows that the PDFs in the upper tail due to the Bayesian weighting.

[4] the mixture of analyses they presented here provides sufficient proof to question the notion of cooling climate being the first explanation for the population decline observed during the Late and Final Chulmun periods. If the cooling of the climate has been directly liable for effecting or causing the Chulmun population to decline then the population changepoint from a clearly positive to a negative growth rates would have occurred.

[5] In this paper they consider the spatiotemporal variability of the temperature of the surface of the sea in the North Atlantic Ocean. The largest spatial variation is in the latitudinal direction and while the most temporal variation is in the annual timescale and represents the seasonal cycle. The data is highly dispersed and the Spearman correlation coefficient is a modest 0.33. A plateau of correlation can be observed in a wide range of intermediate bin sizes. So as a conclusion the analysis suggests that the uncertainty information from Bayesian Deep Learning system can be used to estimate prediction error.

[6] the BMA method is limited because of the assumption that any simulated climate variables have Gaussian distributed. So there is a need to conduct a much more comprehensive and multivariate investigation. So it would be preferable to adopt some differnet non parametric statistical estimation approach or Multi Objective Optimization.

[7] They provided useful outputs: inferences and recommendations for decisions that are consistent with those inferences. It is based on a comprehensive and coherent framework rooted in normatively appealing assumptions. Compared to alternative paradigms, its concepts are familiar and, as the case studies show, its methods are often practical

[8] The evaluation of global climate change impact on hydrological cycle includes uncertainty. Their study aimed to understand the uncertainty of the impact of climate change on Zayandeh-Rud Reservoir during 2020–2049. The outputs of twenty-two GCM models are used under the three different emission scenarios .The Bayesian model averaging (BMA) was used here due to the uncertainty analysis for the output weights of twenty-two GCM models (precipitation and temperature) which supported the ability to simulate baseline 1980–2005 period. Then the statistic of equivalent precipitation and temperature were introduced to the hydrological model (i.e. IHACRES) and therefore the output of runoff was estimated under different global climate change scenarios. They used Downscaling GCM outputs and simulating runoff with Bayesian Model averaging.

The Root Mean Square Error (RMSE) is here applied to compare the abilities of different global climate models (GCMs) and the BMA method to generate the observed hydro-climatic variables' time series in the historical 1980-2005 period most of the GCMs showed considerable errors in projection of the observed precipitation in winter and spring months. But, the BCC-CSM1.1(m), FGOALS-g2, GFDL-CM3 and MIROC5 had relatively better abilities to estimate rainfall values in cold seasons. The PDF of annual temperature and precipitation during the baseline and future (after application of BMA) periods are indicated The average of annual temperature showed 0.5 to 1 C increase under different climate change scenarios for the future This increase in temperature can intensify the process of melting snows in the basin. Furthermore, the projected annual rainfall showed a reduction of 13 to 18 percent which will negatively alter the stream flow in the future In this study they concluded that, the outputs of 22 GCMs are the climate change impacts on the Zayandeh-Rud Reservoir Inflow for the future 2020-2049 period. To increase the reliability of hydro-climatic projections, the Bayesian Model Averaging was utilized for climate change impact assessment.

[9] The subjective Bayesian approach is now widely utilized for estimation of the equilibrium climate sensitivity. Development has been consistent, even though even the most recent estimates may have limitations. Even estimates based on the latest observational report are increasingly converging to a small price with a good estimate rarely far from 2 to 2. The actual weather is more complex than any model and the concept of an Equilibrium sensitivity might not be precisely definable within the actual global. Therefore, there ought to be a restriction on how accurately

this parameter may be meaningfully anticipated. Though, there may be no reason to presume we have but reached this limit, and it provides a valuable foundation for predicting the value of future climate change.

III. METHODOLOGY

Bayesian Neural Networks (BNNs) have gained significant attention in recent years due to their ability to perform deep learning tasks with low computational complexity and memory requirements. BNNs use binary values (-1 or 1) to represent the inputs, weights, and activations of the network. This paper provides an overview of BNNs, including their architecture, training, and applications. We discuss the advantages and limitations of BNNs, as well as the current state-of-the-art in BNN research.

Traditional neural networks are typically trained using deterministic optimization methods such as gradient descent. In contrast, BNNs use Bayesian inference to estimate the distribution of the network's parameters rather than finding a single set of weights that minimize a loss function. This approach allows for uncertainty quantification in the network's predictions and can also prevent overfitting by incorporating a prior distribution over the parameters.

The Bayesian approach to neural networks involves specifying a prior distribution over the weights of the network and updating this distribution using Bayes' theorem as new data is observed. This process involves computing the posterior distribution over the weights given the data and prior, which can be computationally challenging. One approach to addressing this challenge is to use Markov Chain Monte Carlo (MCMC) methods to sample from the posterior distribution. Another approach is to use variational inference to approximate the posterior distribution.

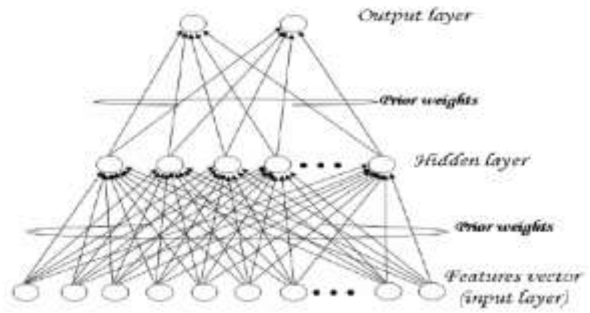

Figure 1 BNN Architecture

- Proposed Modules

(i) Data Acquisition and Preparation: This module will involve web scraping, downloading of existing datasets, and data cleaning to prepare the data for modeling. Algorithmic steps may include parsing and extraction of relevant data from websites, handling missing data, and merging data from different sources.

(ii) Bayesian Model Development: This module will involve the development of Bayesian models for climate change prediction. Various approaches such as Bayesian hierarchical models, Markov Chain Monte Carlo (MCMC), Bayesian networks, and Gaussian processes may be used. The algorithmic steps may involve model specification, prior selection, posterior estimation, model fitting, and model validation

(iii) Bayesian Neural Networks: This module will focus specifically on the use of Bayesian neural networks for climate change prediction. Algorithmic steps involves model architecture design, parameter initialization, model fitting, and model validation.

(iv) Model Evaluation: This module will evaluate the performance of the Bayesian models developed in the previous module. Algorithmic steps may include model comparison using statistical metrics such as mean square error, root mean square error, and R-squared, and cross-validation techniques to assess the generalization of the models.

(v) Model Implementation: This module will involve the implementation of the models for climate change prediction. The models may be integrated into web-based applications or decision support systems. Algorithmic steps may involve deploying the models in a production environment, monitoring their performance, and updating the models as new data becomes available.

IV. IMPLEMENTATION

We imported the required libraries including BeautifulSoup, csv, requests, time, pandas, urllib, re, and pickle. then sets a URL to scrape the data for January 2009. The page content of the URL is obtained, and BeautifulSoup is used to parse the page.

Afterward, using pandas' date range function, a list of dates is generated with a monthly frequency starting from January 1, 2009, until December 1, 2022. The dates list is looped over to scrape the data of each month. For each month, the URL is created by appending the year and month extracted from the dates list to the URL string, and the page content is obtained. The content is then parsed using BeautifulSoup, and the required data is extracted from the tables present on the page.

The data obtained from each month is appended to a list called df list. This list contains lists of data for each day of a month. The index of each record is also appended to a list called index, which is a combination of the year, month, and day.

We generated a unique index, the f index list is generated by looping over the index list and keeping only those records whose length is greater than 6. The unique index is then created by iterating over the f_index list and using the datetime module to convert the index's string to a datetime object, which is then converted to a string in the required format. Using the pandas DataFrame function, the df DataFrame is created from the data and the final_index list. The df DataFrame contains weather data for each day from January 2009 to December 2022

Then, a heatmap and a pairplot are generated using the seaborn library. The data frame df is then split into training and validation sets, and the input features in the training set are normalized.

Then we implemented a Bayesian neural network (BNN) in TensorFlow. A BNN is a neural network that uses Bayesian inference to make probabilistic predictions. It differs from a traditional neural network in that it represents the weights and biases of the network as probability distributions rather than fixed values.

The implementation consists of two main classes: Dense Flipout and Bayesian Neural Network.

(i) Dense Flipout is a custom layer that performs variational inference. It is a subclass of tf. keras.layers. Dense, which is a fully connected layer in a neural network. Variational inference is a method of approximating probability distributions by finding the distribution that is closest in KL-divergence to the true posterior. DenseFlipout is used to represent the weights and biases of the neural network as probability distributions. The call method of DenseFlipout calculates the output of the layer using the sampled weights.

(ii) Bayesian Neural Network is a subclass of tf .keras. Model, which is a high-level API for building neural networks in TensorFlow. It defines the architecture of the neural network, which consists of three fully connected layers (hidden_layer_1, hidden_layer_2, and output_layer) with ReLU activation functions in the hidden layers. The output layer has no activation function.

The log_prior and log_posterior methods of BayesianNeuralNetwork calculate the log probabilities of the prior and posterior distributions of the network weights and biases, respectively. The sample_elbo method calculates the evidence lower bound (ELBO), which is a lower bound on the log marginal likelihood of the data given the model. The ELBO is used as the loss function during training.

Then we did some pre processing, model compilation, training, and visualization of the training history. The input data is first normalized using the mean and standard deviation of the training data. The model is then compiled using the mean squared error (MSE) loss function and the Adam optimizer. The model is trained for a specified number of epochs and batch size. Finally, the training history is visualized using matplotlib.

V. RESULTS

Our paper identified several potential applications of Bayesian approaches for climate change prediction. One important application is in developing more accurate and reliable climate models. Bayesian approaches can be used to incorporate prior knowledge about the sources of uncertainty in climate models, such as future emissions scenarios and natural variability, leading to more accurate predictions. Another potential application is in developing strategies for mitigating and adapting to climate change. Bayesian approaches can be used to estimate uncertainty in predictions, which can inform decision-making and help identify effective strategies for reducing greenhouse gas emissions and building resilience to climate change impacts

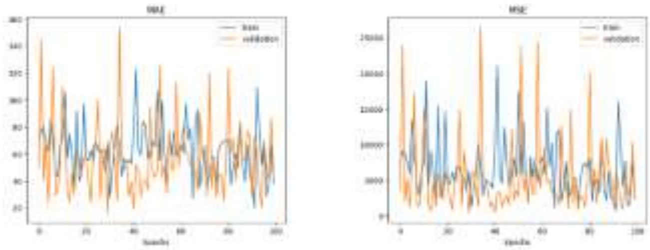

Figure 2 MAE V/s Epochs , MSE V/s Epochs

VI. CONCLUSION

The process of predicting climate change is difficult and unpredictable, but Bayesian methods provide a number of benefits that can increase the precision and reliability of forecasts. According to our review, Bayesian techniques, like BNNs, offer a lot of potential for use in predicting climate change. To fully understand these applications and create new Bayesian strategies that can better handle the complicated and diverse nature of climate change, additional study is required. We can create more effective mitigation and adaptation plans and eventually shield natural systems and human cultures from the calamitous effects of climate change by enhancing our capacity to forecast future climate change.

REFERENCES

[1] Patrick Gordan White, A Review of Climate Change Model Predictions and scenario selection for impacts on Asian Aquaculture, December 2013.

[2] L. Mark Berliner, Richard A. Levine and Dennis J. Shea, Bayesian climate change assessment, Manuscript received 20 September 1999, in final form 10 January 2000.

[3] Seung-Ki Min , Daniel Simonis and Andreas Hense; Probabilistic climate change predictions applying Bayesian model averaging; June 2007.

[4] Habeom kim, Gyoung-Ah Lee & Enrico R. Crema, Bayesian analyses question the role of climate in Chulmun demography, December 2019.

[5] Xihaier Luo, Balasubramanya T. Nadiga, Ji Hwan Park, Yihui Ren, Wei Xu, Shinjae Yoo; A Bayesian Deep Learning Approach to Near-Term Climate Prediction; September 2022

[6] Qingyun Duan, Thomas J. Phillips; Bayesian estimation of local signal and noise in multi model simulations of climate change; September 2010

[7] Benjamin F. Hobbs; Bayesian Methods for Analyzing Climate Change and Water Resource Uncertainties; May 2002.

[8] Ali Alinezhad, Alireza Gohari, Saeid Eslamian and Ramin Baghbani, Uncertainty Analysis in Climate Change Projection Using Bayesian Approach

[9] J. D. Annam, Recent Developments in Bayesian Estimation of Climate Sensitivity, November 2015.