CLIMATE CHANGE PREDICTION TECHNIQUES A STUDY

SUGASHINI K
I MSc Computer Science
Department of Computer Science
Bharathiar University
Coimbatore – 641046
sugashinikarthi0912@gmail.com

Dr.D.RAMYA CHITRA
Associate Professor
Department of Computer Science
Bharathiar University
Coimbatore – 641046
ramyachitra@buc.edu.in

B. KANNAKI DEVI
Ph.D Research Scholar
Department of Computer Science
Bharathiar University
Coimbatore – 641046
kavipalanichamy@gmail.com

Abstract-**Our main source of information on climate change is climate model data. Impact models are frequently used to quantify the effects of climate change. Global and regional climate models are generally biassed, and their resolution is frequently lower than needed, whereas impact models typically require high-resolution unbiased input data. As a result, many consumers of data from climate models downscale and apply bias correction. An essential presumption of bias correction is that the evaluated climate model generates knowledgeable input for a bias correction, including a convincing depiction of climate change. Existing bias correction techniques have limited downscaling potential and cannot realistically correct climate change trends. Cross-validation of marginal aspects must be supplemented with additional analyses in order to evaluate bias correction.**

I. INTRODUCTION

As a result of climate change, air temperatures are rising, and precipitation regimes, such as snowfall and the timing, amount, and inter-annual variability of rainfall, are altering (IPCC 2013). The enormous and long-lived ecosystems known as forests are intensively and extensively managed. Together with the communities and economies that depend on them, they may be sensitive to these longer- term climatic changes (Bernier and Schöne 2009).[1]

The statistical distribution of weather patterns across timescales ranging from decades to millions of years undergoes considerable and long-lasting change due to climate change. It could be a shift in the distribution of weather events around the average or in the typical weather conditions themselves (e.g., more or fewer extreme weather occurrences). The phrase is occasionally used to refer explicitly to anthropogenic climate change as opposed to changes in climate that might have occurred naturally as part of Earth's processes. Today, man-made global warming is synonymous with climate change.[2]

The development of a greenhouse and a climate forecast model go hand in hand. People can now enjoy tasty food both in and out of season thanks to the invention of the greenhouse. Modern agriculture is becoming more and more dependent on greenhouse cultivation. Crop growth in a greenhouse will be influenced by indoor temperature, indoor humidity, lighting, indoor carbon dioxide concentration, soil temperature, and soil humidity. An example of a thermophilic plant is the tomato, which can thrive at temperatures between 25 and 28 °C during the day and16 to 18 °C at night. Tomatoes require higher humidity.[3]

One of the most significant environmental problems facing the globe today is the climate change brought on by greenhouse gas emissions. It is now commonly acknowledged that it has the potential to significantly impact both the welfare of humans and the integrity of our ecosystems. Structures that are part of the infrastructure will have to endure a long period of changing climatic conditions (50–100 years). For this to be possible, building stocks, both present and future, must operate effectively in light of shifting climatic conditions.[4]

Updates to the climatic scenarios created for Finland within SILMU, the Finnish Research Project on Climate Change, between 1990 and 1995 are the primary goal (Carter et al. 1996). Based on the most current climate change simulations by a number of coupled atmosphere-ocean general circulation models, a new state-of- the-art set of climate scenarios for the 21st century will be given here (AOGCMs).According to the premise of FINSKEN, the scenarios are compatible with fourfictional storylines of the future world that were developed in the Special Report on Emissions Scenarios (SRES) by the Intergovernmental Panel on Climate Change (IPCC) (Nakienovi et al. 2000). (Also see Carter et al. (2004).[5]

II. TECHNIQUES

2.1 DATA ANALYSIS:

The process of modelling the linear relationship between a dependent variable and one or more independent variables is known as multiple linear regression (MLR). the reliant Sometimes a variable is referred to as a predictor, and predictors are indepent referred to as a predictor, and predictors are independent variables. A regression analysis was carried out using data from 1977 to 2006 (30 years) for three locationsand from 1993 to 1999 (7 years) for one location, with tea production (P) as the dependent variable and average maximum and minimum temperature difference (Tdiff), total rainfall (R), and average sunshine hours (S) as the independent variables from April to November. As a set of crucial factors related to yield was developed, non-contributing variables were eventually excluded one by one based on their p-values. Only three environmental variables—rainfall, temperature difference, or the difference between Tmin and Tmax, and sunshine duration—were shown to be the most important contributors to yield. These independent variables were used in several regression analyses to see how they impacted the production-dependent variable. The formulas created have the form

Y=0 + 1x1 + 2x2 + 3x3

and are multiple regression equations with three independent variables and one dependent variable, Y.[6]

2.2 CROSS VALIDATION:

With the provided training data, cross-validation is performed to determine the best value for prediction. Then, we select a number of subsets of size NCV from the training dataset (of size N).

(a) Our three subsets result in an NCV of about 6-7. Then, we go through a list of potential values and perform the subsequent actions for each one. Select from the list. As the validation data, choose one of the smaller datasets (size NCV).

(b) Reduce the remaining data (NNCV) to train the regression model .

(c) Utilize the validation dataset's inputs on the trained model to forecast results using and denote this by y.

Use an error metric such as root-mean-squared error (RMSE), accounting for all grid cells i = 1,..., p, and weighting by the grid-cell area, wi, to compare these predictions with the actual results of the validation dataset.

RMSEcv,λ=(∑ipwi∣ ∣ y∗ 2i−y2i∣)

For other subsets of validation data, repeat steps a through d (we use 3 in total). Determine the cross-validation score as the average RMSE for all three subsets for this value.

RMSEλ=∑cv3RMSEcv,λ \s

All of the list's values are subjected to the same procedure. In the final stage of training the model, where all training data is used, the value that yields the lowest RMSE is chosen as the parameter.[7]

2.3 GRADIENT DESCENT METHOD

Several machine learning methods are based on gradient descent, which is related to standard linear regression and tries to reduce the error (or "cost function," F)

$$U(n+1) = U(n) + \gamma \frac{\mathrm{d}F(U(n))}{\mathrm{d}U(n)}.$$

U(n) is a differentiable point (or vector), which ensures that the cost function decreases with each iteration: F(U(n)) >F(U(n + 1)), and n is the number of iterations. The gradient descent approach's step size, or "learning parameter," regulates how quickly data is gathered. Ineffective algorithms that either do not converge to the optimal solution or overcorrect the cost function result from poor optimization.[8]

2.4 DATA COLLECTION

The Nigerian Meteorological Agency's Oyo State office used Ibadan Synoptic Airport to gather the data for this study. The case data was collected over a 120-month span, from January 2000 to December 2009. At this point in the research, the following techniques were used: data cleaning, data selection, data transformation, and data mining.

2.5 CLEANING OF DATA

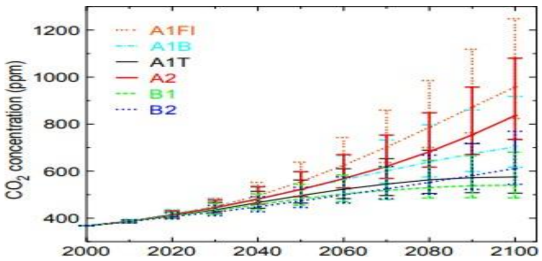

Fig.1. Atmospheric concentrations of CO2

This step involved creating a uniform format for the data model that dealt with missing data, locating duplicate data, and eliminating faulty data. The data was then converted into a data mining-friendly format.

2.6.DATACOLLECTION

At this point, the dataset's data that would be useful for the analysis was selected. The meteorological dataset contained ten (10) attributes. It presents their type and description, and presents an analysis of the numerical values. Both the sunlightdata and the cloud form data, which by their very nature have identical values, were not employed in the research due to their large percentages of missing values.

2.7 TRANFORMATION OF DATA

The term "data consolidation" also applies to this. It is the stage where the chosen data is converted into data mining-appropriate forms To lessen the impact of scaling on the data, the datasets were normalized, and the data file was saved in the Commas Separated Value (CVS) file format.

Data Mining Stage: Three phases made up the data mining stage. The meteorological datasets were analysed using all of the methods at each phase. Percentage splitting was used as the testing strategy in this study, which involved training on a portion of the dataset, cross-validating it, and testing on the remaining portion. After that, intriguing patterns that represented knowledge were found.[9]

Construction of climate scenarios from general circulation model output:With a horizontal resolution of a few degrees in longitude and latitude, experiments utilising AOGCMs generate time-dependent worldwide distributions of climate variables. We conducted an analysis of seasonal and annual mean surface air temperature change and precipitation change in fifteen global climate model experiments carried out with six AOGCMs using data fom the IPCC Data Distribution Center (Parry 2002) and the Climate Impacts LINK Project (Viner and Hulme 1997).

CO_2 concentrations in the atmosphere between the years 2000 and 2100 as a result of the six SRES emission scenarios based on the Bern-CC carbon cycle model (IPCC 2001: p. 221). The concentrations corresponding to the models' mid-range estimate are shown as curves. As a result of carbon cycle uncertainties, the error bars provide assessed ranges. In all calculations, a climate sensitivity of 2.5 °C for a doubling of CO_2 is assumed (for more information, see the IPCC).[10]

Extrapolating statistical method (degree- day method):For the purpose of estimating building energy consumption, the degree-day approach is essentially a single-measure steady-state method. The building industry frequently uses this method to link the energy consumption trends of buildings to the regional climate. The degree-day method's underlying premise is that there should be a base temperature (or balance point for external temperature) in order to maintain a particular inside temperature. A cooling load is necessary when the outside temperature is higher than the equilibrium point temperature. The demand for heatingarises when the outside temperature islower than the balance point temperature. Fig 1 depicts a link between buildingsenergy use and ambient temperature (Fig 2).

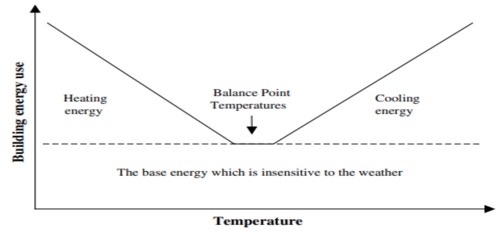

Fig. 1. The relationship between building energy use and the outdoor temperature.

The advantages of this imposed methods are:

1. The diurnal cycle's characteristics are preserved.
2. The IPCC or other pertinent national research groups frequently have information on the expected temperature rise caused by climate change.
3. For building simulation, the typical reference-year weather data for the current climate is also typically supplied.
4. It is possible to consistently compare the weather conditions of the present and the future. All comparisons must be made using the same base because impact studies are more interested in relative change.[11]

2.8.BIAS CORRESTNESS

The following shall refer to a simulated present-day model's (predictor) time series of length N and the matching observed(predictand) time series as $x_i\,p$and yp_i, respectively. Both may adhere to themarginaldistributions$xp\sim Dpiraw$ and $yp_i\sim Dp_{real}$ as discussed in "Definitions."

The uncorrected model's mean for a selected current period can be calculated as

$$\mu p\ \text{row}=xp\ \text{i}$$

(where the hat designates the estimator and the bar signifies time averaging), and the matching real mean canbe calculated as$\hat{\mu}p$ real $= \bar{y}p$ i. Then, an estimator of the model bias under the current circumstances is provided as

$$\widehat{Bias}(\mu^p) = \bar{x_i^p} - \bar{y_i^p}. \quad \text{---- (1)}$$

$$\widehat{Rel.\,Bias}(\mu^p) = \bar{x_i^p}/\bar{y_i^p}. \quad \text{---- (2)}$$

The value that is exceeded with a probability of 1 when sampling from the distribution is known as the quantile for a probability of a distribution D and is denoted as q D (α). Sorting the provided time series, let's say x I, and then taking into consideration the value at position N/100 will yield the equivalent empirical quantile ˆqD(α) (also called the rank of the data). Written as p D (q) =α, the cumulative distribution function's probabilities are indicated for a specific quantile, q D (α). Future simulations and their derived measurements will be identified by a ʃ.[12]

III. DATASET

3.1 QUANTILE MAPPING

More flexible bias correction methods also attempt to adjust the variance of the model distribution to better match the observed variance . A generalisation of all these approaches is quantile mapping, which employs a quantile-based transformation of distributions . In a widely used variant, a quantile of the present day simulated distribution is replaced by the same quantile of the present-day observed distribution(fig 3) . The variance of the model distribution is additionally adjusted through more adaptable bias correction techniques in an effort to better match the observed variance. Quantile mapping, which uses a quantile-based modification of distributions, is a generalisation of all three methods. In a common variation, the identical quantile of the current observed distribution is used to substitute a quantile of the current simulated distribution. Usually, climate models overestimate the number of rainy days (the so-called drizzle effect). Quantile mapping automatically modifies the number of wet days in this case (as the wet- day threshold is a quantile of the distribution). It has been suggested, for example, to randomly generate low precipitation levels if a climate model simulates too few wet days.implementation affects how the transfer function is actually formulated.

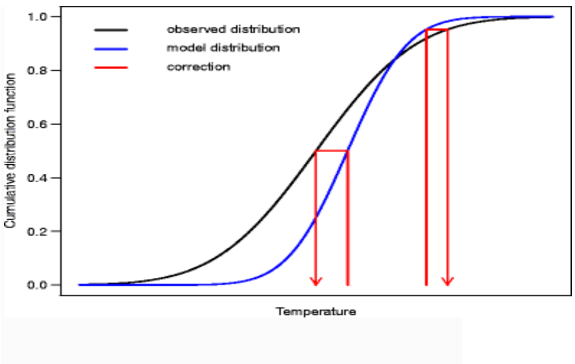

Fig. 2. Quantile mapping

The quantile of the observed distribution that corresponds to the same probability replaces a simulated value, or quantile, of the simulated distribution.[13]

3.2 UK SUMMER 2018 DROUGHT

Through the summer of 2018, the UK had a hot, dry year. Strong anomalies for the southern UK are seen in the ERA-Interim re-analyses (Dee et al. 2011) for the mean July and August temperature, rainfall, and soil moisture when compared to the mean conditions for these two months . Prior estimations of the meteorological conditions are provided by the ECMWF using a cutting-edge forecasting model, which is initialised from earlier forecasts. These estimates are then modulated by data assimilation from satellite observations, weather stations, and radiosondes. Such assimilation (or "4D-Var" optimization), according to (Dee et at(2011)). captures measurement uncertainty and prevents excessively altering models of atmospheric physics. Recent analyses contextualise the drought and emphasise that increasing global Rossby waves are likely to play a significant role in these climatic processes (Kornhuber et al. 2019). There is evidence that these wave characteristics are increasing in frequency, and they have the potential to lead to simultaneous extremes elsewhere (Kornhuber et al. 2019). However, the processes are still unclear, necessitating the identification and comprehension of potential parallel forcings, including prior soil drying, background warming rates, and warming-circulation change interactions, all of which ML may clarify. Planning for drought adaptation, such as when to sow crops, is made easier by any identified inter-seasonal links relating drought risk to springtime soil moisture or marine temperature factors. By combining the study of concurrent crop production and climate datasets, ML applications may benefit optimal crop timing itself

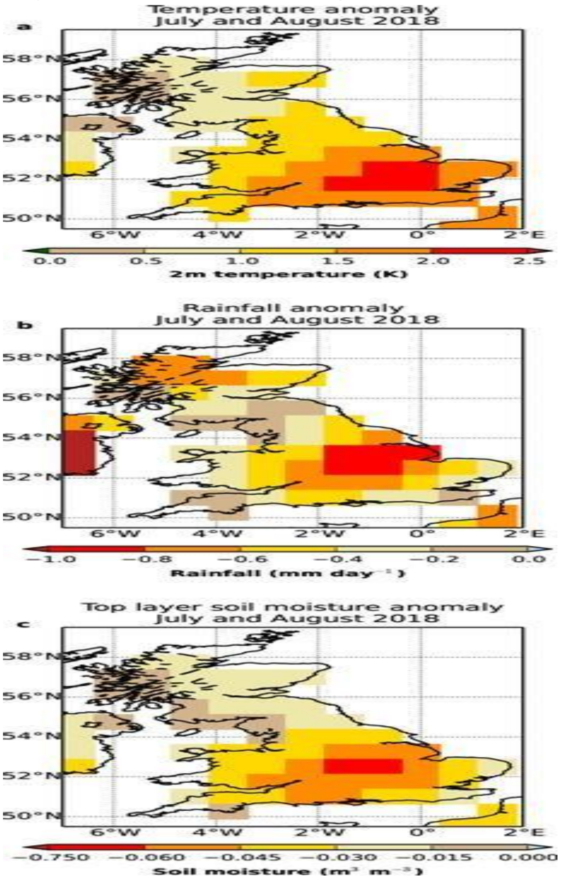

Fig. 3. UK July and August year 2018 temperature, rainfall and soil moisture anomalies.

UK Temperature, rainfall, and soil moisture anomalies in July and August of 2018. ECMWF reanalysis data shows the mean for the months of July and August for the year 2018 minus the mean for the same two months averaged over the years 1979– 2017.[14]

Tea Yield Prediction Model with Climatic Parameters: Following a successful MLR analysis, four equations representing the relationship between a region's unique yield and climatic characteristics were created for each of the four regions, resulting in four equations for tea production estimation. The resulting equations for the South Bank, Upper Assam, North Bank, and Cachar are provided as equations (1), (2), (3), and (4), respectively.

Table. 1. Coefficients of Correlation (R) and determination (R2) for all the four sites

Coefficients	South Bank	Upper Assam	North Bank	Cachar
R	0.8267	0.8000	0.7739	0.6555
R^2	0.6835	0.6400	0.5989	0.4296

It was now attempted to evaluate the values of the expected production for each region using these equations. The actual yield and expected yield were discovered to be at or very near one another.

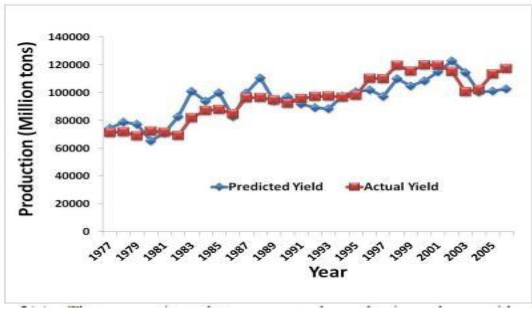

Fig. 4. Comparison of South Bank region production over a 30-year period between actual production and the equivalent expected value (1977-2006).

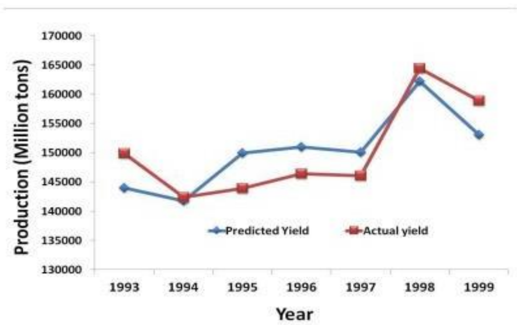

Fig. 5. The comparison of the Upper Assam region's actual production during a 30-year period with the equivalent expected value (1977-2006)

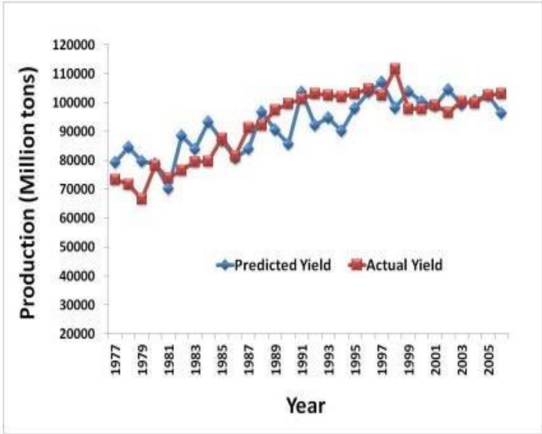

Fig. 6. The comparison of the North Bank region's actual production during a 30-year period with the equivalent expected value (1977-2006)

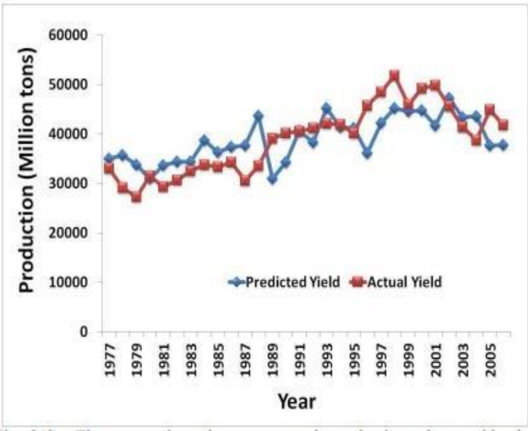

Fig. 7. Comparing actual production in the Cachar region over a 30-year period (1977–2006) with the equivalent expected value [15].

IV. CHALLENGES:

Since the early 1990s, climate issues have received a great deal of media attention, but the media has also played a crucial role in bringing doubt about climate change and contentious scientific theories to the public's attention. Science-averse people who find it hard to accept that a group of people can genuinely affect a large "planetary machine," as well as more organised lobbies with a stake in casting doubt on the possibility of human influence on climate, make up the majority of climate sceptics (e.g., Oreskes and Conway, 2010). The authors describe how groups of powerful advisors who had close ties to key political and business circles launched public relations operations to cast doubt on scientific understanding.

Skepticism about scientific matters raises the question of whether the scientific community is capable of explaining science in a way that is both understandable to a non-expert audience and convincing enough to refute the arguments made by a sceptic audience. In fact, the majority of researchers are educated in their specific fields of study and aren't always prepared to argue their points outside of the implied bounds of academic science. It can be challenging to explain the notion of climate change to a non-expert audience because many findings about present or future climatic conditions are based on temperature change relative to a certain baseline climate, such as the generally recognised temperature increase of 0.7°C since 1900 (e.g., IPCC, 2007).

Despite the care taken in preparing reports like those of the IPCC—including the widely read "Summary for Policy Makers"—communication of climate science remains a challenge that should not be taken lightly. There is still a significant gap between the messages that the scientific community wishes to convey and those that the end-recipients are willing or capable of receiving. It is becoming more and more important to explain properly the current state of our understanding of climate change and its effects, as well as the reasons why uncertainty occurs in complex systems and why decision-making is still possible in spite of these uncertainties.[16].

V. CONCLUSION

A compilation of future climate estimates shows that runoff from the Lule River Basin will generally rise, with peak spring floods happening around a month earlier than under the current climate. This suggests an increase in hydropower potential overall, on average, of about 34% more than the current situation. In evaluating hydrological change, the boundary conditions provided by the GCM, as opposed to the emissions scenario, are more important. When it comes to variations in runoff volumes, the scaling transfer and delta techniques produce results that are comparable, but they diverge when it comes to seasonal dynamics and exceptional river discharge. The extent to which the delta technique can benefit from improved RCM

resolution is constrained. Artificial Neural Networks don't require programming or the creation of intricate equations to model these relationships; instead, they can identify the links between the input variables and provide outputs based on the observed patterns inherent in the data. As a result, given enough data, ANNs are able to identify links between meteorological parameters and use these relationships to forecast future weather conditions.
Predictive ANN models for the forecasting of future values of wind speed, evaporation, radiation, minimum temperature, maximum temperature, and rainfall given the month and year were constructed using both TLFN neural networks and recurrent network architectures.

References

[1] Rodney J. Keenan , Climate change impacts and adaptation in forest management: a review, Annals of Forest Science 72 ,Published: 14 January 2015.

[2] Federal University Oye Ekiti, Nigeria, Application of Data Mining Techniques in Weather Prediction and Climate Change Studies, Article in International Journal of Information Engineering and Electronic Business, February 2012.

[3] Yuwen Liu1 | Dejuan Li2 | Shaohua Wan3 | Fan Wang1 | Wanchun Dou4 | Xiaolong Xu5 | Shancang Li6 | Rui Ma7 | Lianyong Qi1, A long short-term memory- based model for greenhouse climate prediction, 3 August 2021.

[4] Lisa Guan, Preparation of future weather data to study the impact of climate change on buildings, 20 May 2008,Building and Environment 44 (2009) 793– 800.

[5] Kirsti Jylhä, Heikki Tuomenvirta and Kimmo Ruosteenoja, Climate changeprojections for Finland during the 21st century,Volume 9, 7 May 2004, BOREAL ENVIRONMENT RESEARCH 9: 127– 152.

[6] Rupanjali D. Baruah, Use of Data Mining Technique for Prediction of Tea Yield in the Face of Climate Change of Assam, India, 2016 International Conference on Information Technology.

[7] L. A. Mansfield, P. J. Nowack, M. Kasoar, R. G. Everitt, W. J. Collins & A. Voulgarakis , Predicting global patterns of long-term climate change from short-term simulations using machine learning, npj Climate and AtmosphericScience volume 3, Article number: 44 (2020) .

[8] Chris Huntingford1, Elizabeth S Jeffers2, Michael B Bonsall2, Hannah M Christensen3, Thomas Lees4 and Hui Yang1,5 , Machine learning and artificial intelligence to aid climate change research and preparedness, Published 22 November 2019, Environmental Research Letters, Volume 14, Number 12.

[9] Federal University Oye Ekiti, Nigeria, Application of Data Mining Techniques in Weather Prediction and Climate Change Studies, Article in International Journal of Information Engineering and Electronic Business, February 2012.

[10] Kirsti Jylhä, Heikki Tuomenvirta and Kimmo Ruosteenoja, Climate change projections for Finland during the 21st century,Volume 9, 7 May 2004, BOREALENVIRONMENT RESEARCH 9: 127–152.

[11] Lisa Guan, Preparation of future weather data to study the impact of climatechange on buildings, 20 May 2008,Building and Environment 44 (2009) 793–800.

[12] Douglas Maraun, Bias Correcting Climate Change Simulations - a Critical Review, Current Climate Change Reports volume 2, 211–220 (2016).

[13] Douglas Maraun, Bias Correcting Climate Change Simulations - a Critical Review, Current Climate Change Reports volume 2, 211–220 (2016)

[14] Chris Huntingford1, Elizabeth SJeffers2, Michael B Bonsall2, Hannah MChristensen3, Thomas Lees4 and Hui Yang1,5 ,Machine learning and artificial intelligence to aid climate change research and preparedness, Published 22 November 2019, Environmental Research Letters, Volume 14, Number 12.

[15] Rupanjali D. Baruah, Use of Data Mining Technique for Prediction of Tea Yield in the Face of Climate Change of Assam, India, 2016 International Conference on Information Technology.

[16] Martin Beniston , Grand challenges in climate research, SPECIALTY GRAND CHALLENGE article, Volume 1 – 2013.