Time-Series Analysis for Forecasting Climate Parameters of Kashmir Valley Using ARIMA and Seasonal ARIMA Model.

Fahad Farooq J.[1], Khalid Hussain[1], Mir Salim N.[1], Sheikh Umar Farooq[2]

[1]Department of Computer Science and Engineering, North Campus, University of Kashmir
[2]Department of Computer Sciences, North Campus, University of Kashmir

Abstract
The valley of Kashmir being an extremely ecologically sensitive area deserves significant attention with regard to climate change. As the region is part of the Indian Himalayan Region (IHR) and home to multiple rivers and glaciers, it holds significant geopolitical, economic, as well as geographical importance. Climate change in the region can have ripple effects on the Indian subcontinent in general. As such effective prediction of the different climate variables of the region is of paramount importance and can help us to reverse or slow down the negative climate change process. In this paper, we investigate CRU TS4.04 time series data of 120 years (1901-2019) for three major climate variables - precipitation, temperature, and cloud cover observed in the region. The analysis revealed some key and concerning trends. Thereafter, forecasting of the climate parameters for the next 80 years (2020-2100) was performed using the ARIMA and S-ARIMA models after a thorough analysis of the data and proper hard tuning of the model hyper-parameters, separately for each of the three climate parameters. The resulting projections show drastic temperature changes, with a projected increase of about 2°C by the end of the century, slight changes in precipitation and cloud cover, and other alarming climatic conditions. This study also brings forth the inability of the ARIMA model to substantially forecast erratic changes in any climate variable. The insights gathered in this study may serve as a presage for the concerned government and stakeholders, and will pave way for the development of robust and efficient plans to tackle climate change in the area. This is the first kind of work that is using effective machine learning-based time-series forecasting models trained over such a large data set for all three major climate variables.

Keywords: ARIMA; climate; Kashmir valley; analysis; time-series forecasting; weather;

1. Introduction

Climate change is usually characterized by a long time change in various variables especially temperature and precipitation in any region. The region of interest can be a small local region or a region as large as a continent or even the entire planet. The studies focusing on understanding climate change usually analyze variations in various climate variables to ascertain the extent of change for the region of interest over a certain period. The climate time-series data can have trends (long-term changes in the data) and seasonality (changes and variations in the data that occur regularly at short intervals). Predicting and estimating the extent of change that will occur in the region helps in planning and managing natural calamities such as floods, droughts, and extreme heat/cold. The United Nations Organisation (UNO) calls climate change a "defining issue of our time" and the current times being a "defining moment" [1] i.e., immediate action needs to be taken to analyze, prevent and tackle climate change in as many regions as possible. A 2007 Intergovernmental Panel for Climate Change (IPCC) report by UNO highlights various damages that may result in an increase of 2°C in global temperatures [2]. However, a recent study in 2018 suggests that many such adverse impacts may occur if the global temperature will rise by 1.5°C [3]. Similarly, the last IPCC report from 2021 has called the current global climate scenario "*code red for humanity*" [4]. All these reports stress the need for immediate, significant, and effective efforts for analyzing, slowing, and/or reversing the climate change process. The fight for climate change needs to be global. However certain regions need more attention as compared to other regions as the adverse effects of climate change in those areas can have disastrous consequences. The valley of Kashmir being one such area holds great environmental, geographical, and economic importance. The region is located 1600 meters above sea level in the North-Western corner of the Himalayas, with an area ranging from (33.25°N,73.75°E) to (34.5°N,75.25°E), with some parts controlled by India and some by Pakistan as shown in Fig 1. The region is considered to have a sub-tropical climate, which is sometimes also classified as sub-Mediterranean due to the rainfall distribution pattern [5]. Mild summers and rigorous, severe winters are considered characteristic features of the climate in the region. Multiple rivers such as Jhelum, Bringhi, Lidder etc., originating from the glaciers located in the upper reaches of the valley, and then flowing through the valley, as well as towards other regions, such as Pakistan and Punjab, are acting as a source of irrigation. The high-altitude origins of the rivers allow for the generation of hydroelectricity, having an estimated potential of a whopping 20,000 Mega-Watt [6], and an estimated annual export revenue of $13 billion, although only about 20% of the total potential has been realized till date [7], [8]. Any drastic climate change can be disastrous to the hydroelectric potential of the region since it is directly dependent on the glaciers in the region. One industry that has already borne the brunt of climate change over the past few decades is the saffron industry, a sector with which the region has been involved for more than 2500 years, saffron has seen a steep decline in production, with the harvest in 2018 being only half of that in 1998 [9]. Multiple studies and reports have been published regarding the scenario of climate change in the region with some key findings; such as the 27-38% decrease in glacier sizes [10], a 19.44% decrease in the daily rainfall [11], a predicted 6.93°C increase in temperature by the year 2100 in RCP 8.5 Scenario [12].

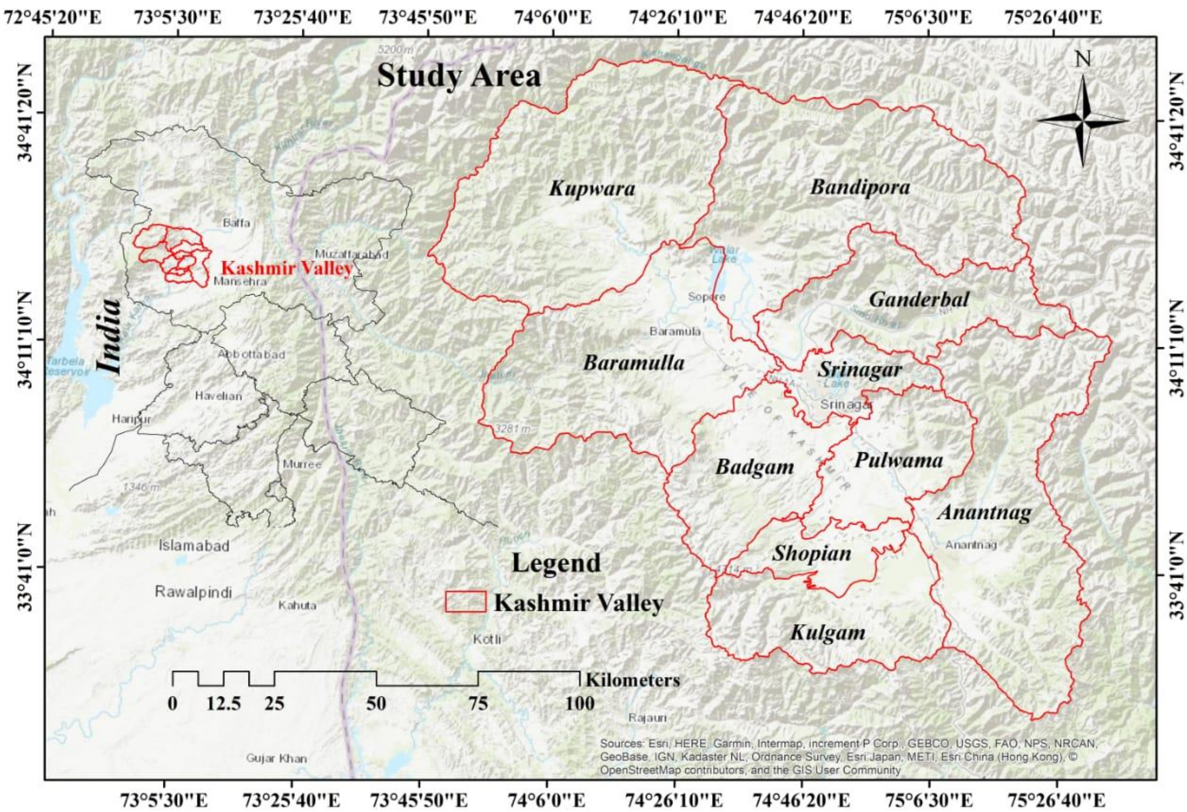


Fig 1 - Area of study (33° 15’ N, 73° 45’ E to 34° 30’N, 75° 15’E).

The major contributions of this study are as under:

i. Forecasting future trends:
One of the significant contributions of the study is forecasting the future climate trends of the valley for the next 80 years (2020-2100) based on a pre-processed large dataset ranging from 1901 to 2019. Despite being an ecologically sensitive and geographically important area, not enough efforts have been made to estimate the future trends of the climate variables such as temperature (mean, max, min) and precipitation using effective methods on the large high-quality dataset. The existing studies mostly use traditional mathematical approaches on limited-size datasets. Our study aims to improve upon those past results by using the proven ARIMA model for forecasting climate variables in the region. The validated forecasts will help in early planning and strategizing of methods to tackle the changes in the future, which can prove beneficial to the surrounding regions, as well as, the rest of the world.

ii. Identifying key trends.

Apart from forecasting climate change for the future, we also analyze the key trends in climate change that have occurred over the years (1901-2019) in the region of interest. In this study, effective analyses have been done on a much larger dataset, with some noteworthy findings that have been discussed, in detail, in section 4 of the study. The analysis may help us in better visualize how climate has changed over the past decades. The analysis can pave way for an immediate decision and policy-making regarding climate change in the region.

iii. Providing ready-to-use data.

A major hurdle in studying and analyzing climate change in Kashmir valley is the availability of ready-to-use, clean and readable data. The data used in the study is provided by the CRU (Climate Research Unit) and requires various pre-processing steps that may appear daunting to researchers of a non-technical background. The high-quality, processed, and ready-to-use CRU dataset has been made available publicly in a public repository (J&K Climate Data Github repository[1]). This repository can be used by further studies for replications or analyzing other variables such as potential evapotranspiration and vapor density etc.

The paper is further divided into multiple sections, with section 2 discussing the previously undertaken research work focusing on the subject matter. Section 3 follows by discussing the dataset used in the study, section 4 discusses inferences and notable observations made from the historical data. Section 5 discusses the methodology followed in the study. Section 6 discusses the forecasting results for each of the climate parameters; temperature, precipitation, and cloud cover. The conclusions and future scope of the study are discussed in section 7 and 8 of the study respectively.

2. Related Work

Many studies in the past have been conducted to forecast and study climate change in the Kashmir valley. The studies conducted differ in terms of variables used to study climate change, the area they have used in the study, the dataset, and/or the technique(s) used to carry out the forecasting. The studies conducted in the region of Kashmir present varying findings and results.

Linear regression analysis was used to examine the rate of change of climatic indices using data from 1961-2005 and an overall increasing trend was found in the seasonal and annual average temperatures. (Islam, Khan, Rao, 2013). Another study, including observational data from six stations within the

[1] https://github.com/fahad-farooq/climchange_jk

Kashmir valley, used Weather and Research Forecasting (WRF) model and ERA-Interim data. The main findings from the study were that the higher altitude stations exhibited a steep increase of 1.04°C to 1.13°C in the annual mean temperatures from 1980-2016 [13]. Investigation of future climate change trends for the 21st century was done under 3 emission schemes (AIB, RCP4.5, RCP8.5) with the baseline period of the data being 1961-1990. The study used GFDL CM2.1 model and the conclusion was that the annual mean temperature was projected to increase by 4.5°C, 3.98°C and 6.93°C respectively under the 3 emission schemes. The study also found the different climatic zones would experience significant changes [12]. Changes in the glacier sizes have also been studied, by comparing satellite imagery from 1980 to 2018, showing a decrease of 27-38% in the sizes of different glaciers, suggesting increasing temperatures and a decline in winter solid precipitation in the region, which if continues in the future, would adversely affect the economy in the region [10]. Similarly, glacier behavior estimation has been performed under climate change using a GRASS GIS module in Italy, with the results showing that the module can be used for a large number of glaciers to obtain spatial simulations to assess future scenarios [14]. Another study uses SDSM to make projections for maximum and minimum mean temperatures, showing an increase of 0.3°C to 2.3°C over the future decades, from 2030-2100 [15]. A study by [16] focusing on climatic extremes in the valley uses IMD data, ranging from 1980-2010, with future projections being generated for 2006-2100, showing clear upward trends in temperature extremes, and a general decreasing trend in precipitation.

The utilization of the ARIMA model has shown very promising results in various domains including climate change analysis. A study by Narayanan et al., has used the ARIMA model for analyzing trends and modelling pre-monsoon rainfall data with results indicating a significant rise in the pre-monsoon rainfall over the northwest part of the country [17].

Air pollution modelling and forecasting have also been done using ARIMA model, with findings showing that meteorological variables have a definite influence on the life cycle persistence of air pollutants [18]. Climate change has also been assessed and monthly rainfall forecasted over Khordha district in Odisha, India, with outstanding accuracy using the ARIMA model [19]. Precipitation and temperature changes in India's Bhagirathi River basin have been studied and forecasted using ARIMA, with the results showing an increasing trend for temperature in one station and a decreasing trend for another, while the precipitation is found to be over-predicted in case of extreme rainfall events [20].

Other domains where the ARIMA model has been successfully applied and has produced exceptionally accurate results include stocks, the economy and the COVID-19 Pandemic. ARIMA has been used for forecasting the GDP (Gross Domestic Product), which is a monetary measure of the market value of all the final goods and services produced in a specific time-period by countries, for Portugal and Germany until the year 2031 with results suggesting a steady growth in both the countries [21]. The Nigerian economy has also been studied using the ARIMA model, with a study showing that the living standards in Nigeria would most likely worsen over the next decade unless the economic policy stance is not reviewed [22].

For analyzing and forecasting the COVID-19 pandemic, in terms of the spread and infection rates in different regions, ARIMA has produced acceptable and proven results. A study focused on analyzing the top five affected countries, with the resulting forecasts being found to be within acceptable agreement with the observed data, and predictions of exponential curves in certain countries such as India, US and Brazil turning out to be true. (Sahai, Kumar et al., 2020). COVID-19 cases in India were also forecasted in another study using ARIMA, with data from the Ministry of Health and Family

Welfare (MoHFW), with results showing an increasing trend of COVID-19 cases with approximately 1500 cases per day [23].

None of the studies focusing on analyzing the climate for this specific region have taken advantage of machine learning approaches to forecasting the future climate trends and values for different variables such as temperature and precipitation, even though machine learning approaches can provide for better and more stable results than traditional mathematical models. Another avenue of improvement is the data used by the studies as almost all of them have taken only small datasets with records only available for a few decades [24]. Further, some studies have restricted their region of study to a smaller region within the valley excluding statistically and environmentally important regions [13], which limits the significance and utility of their findings and conclusions

3. Data Used

The timeseries data used in the study is sourced from the CRU TS4.04 dataset (CRU TS4.04, 2020[2]), which is in the public domain of the Tyndall Centre for Climate Change Research, School of Environmental Sciences, University of East Anglia in Norwich, UK. The main motive of the division is to provide data to the general public for research purposes. The data is an interpolated high-resolution gridded dataset (0.5° x 0.5° grids) consisting of global monthly mean temperature, precipitation, cloud cover, potential evapotranspiration, etc. The data ranges from 1901 to 2019. The 'Subsetter' tool from the web processing service (WPS) available on the CEDA web archive portal (The Subsetter Tool, Web Processing Service, 2021[3]) was used to extract the monthly precipitation, average cloud cover and mean temperature data for our region of study. While other datasets were available as well, such as the one from the Indian Meteorological Department (IMD) Pune available on the Open Governmental Data Platform and one from Berkeley Earth (Berkeley Earth Climate Data), these datasets have significantly lower number of entries than the CRU dataset and in some cases, the datasets provided are derivatives of the CRU timeseries dataset itself like the data available on the Indian Water Portal. The CRU data is also well documented, credible, accurate, and used extensively in research [20], [25]–[27]. Another reason for the selection of the CRU timeseries data in this study is that previous studies have used sparsely populated datasets, with data only available for the past 3-4 decades, and using an extensive dataset populated with datapoints from over a century as is the case with the CRU dataset, provides the potential for better understanding of the past trends as well as more accurate forecasting.

i. Pre-processing

The CRU data obtained in its original form required some pre-processing before being used for analysis and forecasting purposes. While the data is originally being provided as grid points, it needs to be averaged over the region of interest for each month, which is done by calculating the mean for cloud cover (%) and temperature (°C). To convert the gridded values to numeric values, the sum is taken for precipitation (mm/month).

ii. Detection and removal of outliers

Detection and removal of outliers allow machine learning models to better understand the underlying trends in the data and to model them more accurately. The 'z-score' of the distribution is calculated for this purpose and any values that had a z-score greater than 3 or less than -3, as is the norm with Gaussian (Normal) distributions, were replaced by the mean of the immediately neighboring values.

---

2 https://data.ceda.ac.uk/badc/cru/data/cru_ts/cru_ts_4.04

3 https://ceda-wps-ui.ceda.ac.uk/processes

$$z = \frac{x - \mu}{\sigma^2}$$

'μ' denoting the mean of the distribution and 'σ^2', the standard deviation.

iii. Training and testing values

The total collected data spanned time period of 1901 to 2019. The data (1901-2019) was split into two parts, with data from 1901-1999 being used for training the model and that from 2000-2019 used for validation of the forecasts made for the time period of 2000-2100.

4. Data Analysis and Notable Observations

A key part of our study was to analyze and visualize the extent of climate change that has occurred over the years up until now, especially since it has been a topic that has not received much attention when it comes to the valley of Kashmir [28]. The timeseries data was plotted using specialized software *Seaborn* and trend lines fitted upon the graphs to better visualize the direction and change in the temperatures, precipitation, and cloud cover over the period of study. Results are analyzed and discussed in the subsequent sections of the paper.

4.1 Drastic changes in temperature

The change in mean temperature over the period of 119 years in the region can be observed from the temperature (fig. 2-5), showing about 10-25% change in mean temperature in springs, summers, and autumns. The more worrying part is the gradual but extreme increase of about 200% in the mean temperatures over the past winters, as shown in Table 1 and Table 2. This is a worrying sign for the region, considering its ecologically sensitive nature. Such drastic and extreme increase in the mean temperature, especially for winters, can result in the accelerated melting of glaciers which may result in flooding and even affect the perennial nature of the rivers originating from them. The rise in mean temperature is also a problem for most of the fruit-bearing trees in the region such as apples, apricots, and walnuts, all of which have their own "Chilling hour" requirements, the insatiability of which can cause abnormalities in the yield [29]–[32].

4.2 Precipitation spikes

Upon analysis of the precipitation data, extreme spikes in precipitation were noticed in summers and autumns, causing flash floods in localized areas such as that in 2010 [33] and major floods like that of [34] and that of 2022 in Pakistan which resulted in a loss of more than 1600 lives and estimated damage of $30 billion [35]–[37]. The precipitation data when separately plotted for 1901-1990 and 1991-2019 shows a lower mean and deviation in the former time period and a higher mean and deviation, pointing to more erratic rainfalls. The precipitation keeps peaking in late summer and early autumn, as shown in Table 3 and Fig 6.

4.3 Potential for using temperature and cloud cover forecasts to predict heavy rainfall

Cloud cover is one of the most overlooked climate factors when it comes to analyses of the valley's climate. A trendline was fitted upon the data points to analyze the co-dependence of the variables. The trendline is a locally weighted polynomial regression line fit using weighted least squares giving more weight to points near the point whose response is being estimated and less weight to the points further away. Upon analysis of the data and the trendline, it was observed that the cloud cover usually sits around a small range of 41-47% and a relatively larger range of 23-34% in summers and autumns respectively, correlating to precipitations of 3000-8000 mm in summers and 1000-3000 mm in autumns. The deviances and spikes in precipitations can be correlated to cloud covers greater than 55% in summers and 40% in autumns, as shown in Fig 7 and Fig 8. A multi-variate approach, using the precipitation and cloud cover data, can thus be taken in early prediction of such spikes, which can thus aid in the prevention of catastrophic loss of property and life, such as the ones caused in 2010 and 2014.

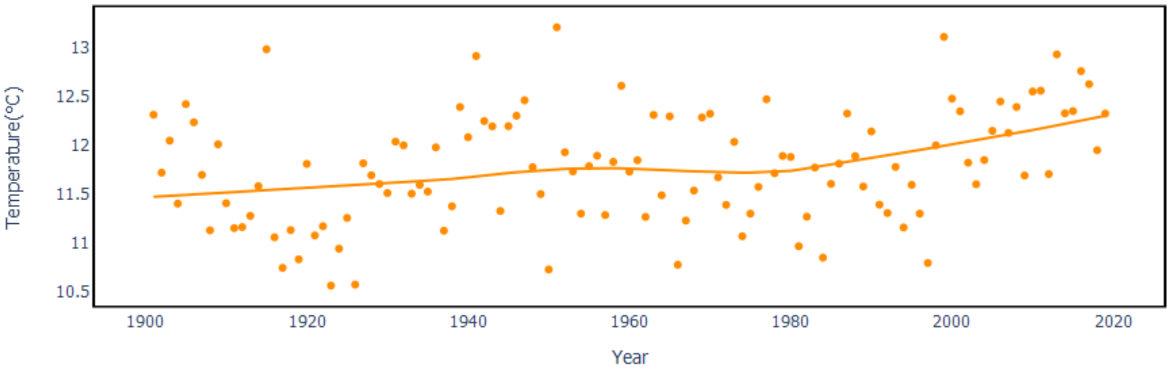


Fig 2 Changes in mean temperature over 1901-2019 in the autumn season (Sep, Oct, Nov).

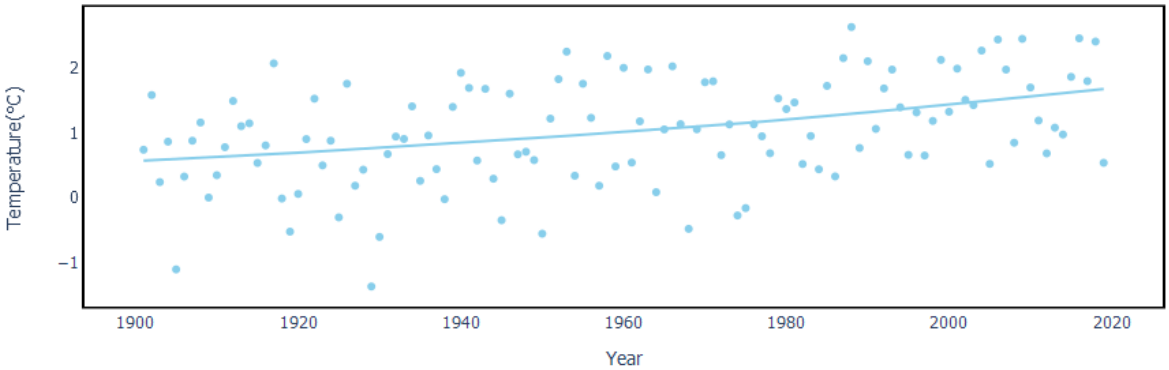


Fig 3 Changes in mean temperature over 1901-2019 in the winter season (Dec, Jan, Feb).

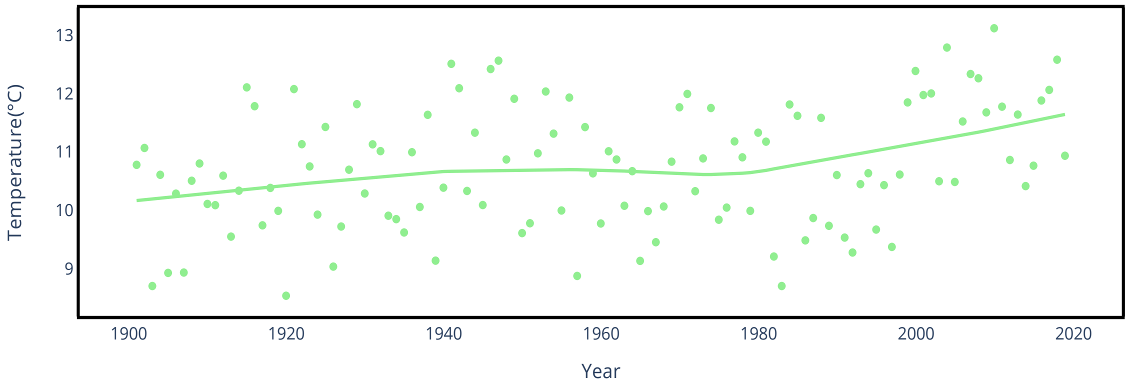


Fig 4. Changes in mean temperature over 1901-2019 in the spring season (Mar, Apr, May).

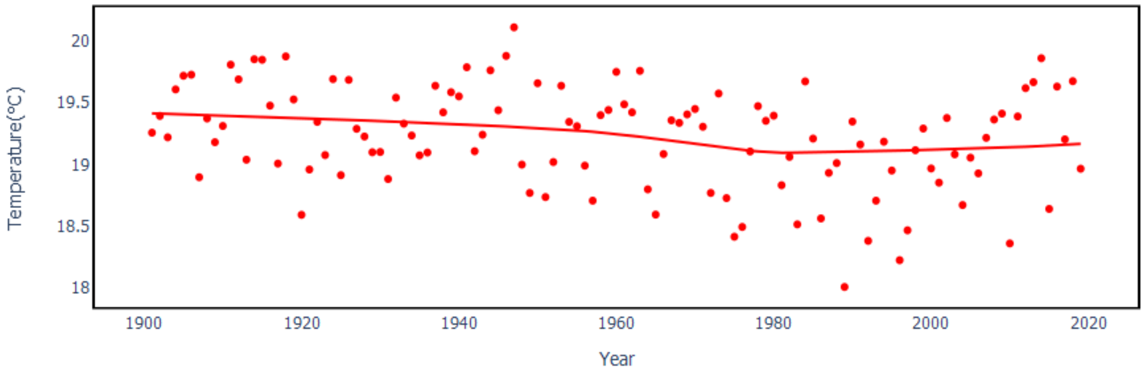


Fig 5. Changes in mean temperature over 1901-2019 in the summer season (Jun, Jul, Aug).

Years	Mean	Standard Deviation	Min	25%	50%	75%	Max
1901-1990	0.870	1.509	-3.275	-0.266	1.003	1.942	4.488
1991-2019	1.506	1.420	-1.206	0.359	1.669	2.709	5.000

Table 1. Changes in mean temperature for winter (1901-1990) and (1991-2019).

Years	Mean	Standard Deviation	Min	25%	50%	75%	Max
1901-1990	11.685	0.555	10.567	11.285	11.683	12.036	13.206
1991-2019	12.231	0.391	11.602	11.875	11.875	12.524	12.929

Table 2. Changes in mean temperature for autumn (1901-2000) and (1991-2019).

Years	Mean	Standard Deviation	Min	25%	50%	75%	Max
1901-1990	11960.7	2278.3	7490.7	10227	11651	13350	17319
1991-2019	13295.7	2427.7	10421	11490	12620	14761	18997

Table 3. Total yearly precipitation compared over the years, (1901-1990) vs. (1991-2019).

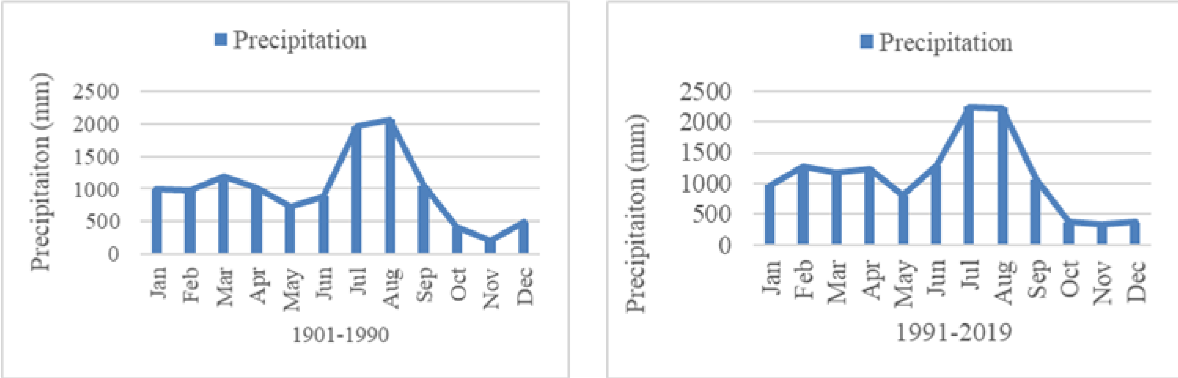


Fig 6. Comparison between mean monthly precipitation for 1901-1990 and that for 1991-2019.

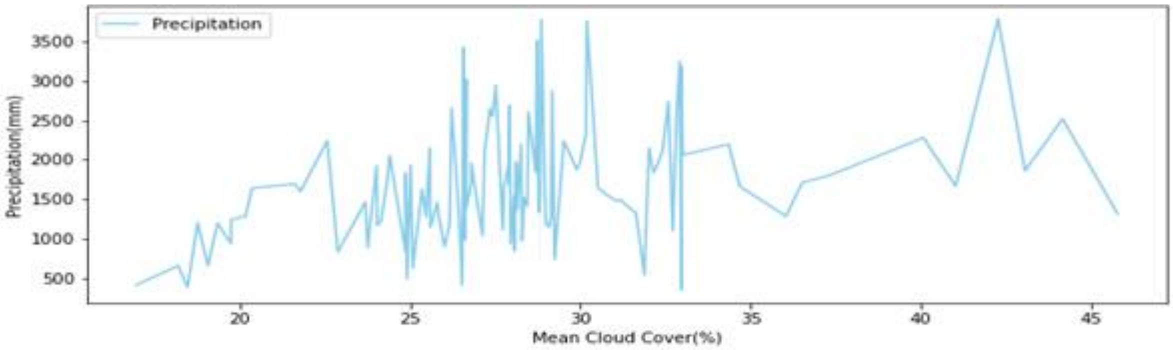


Fig 7. Correlation between cloud cover and precipitation spikes for the autumn season (Sep, Oct, Nov).

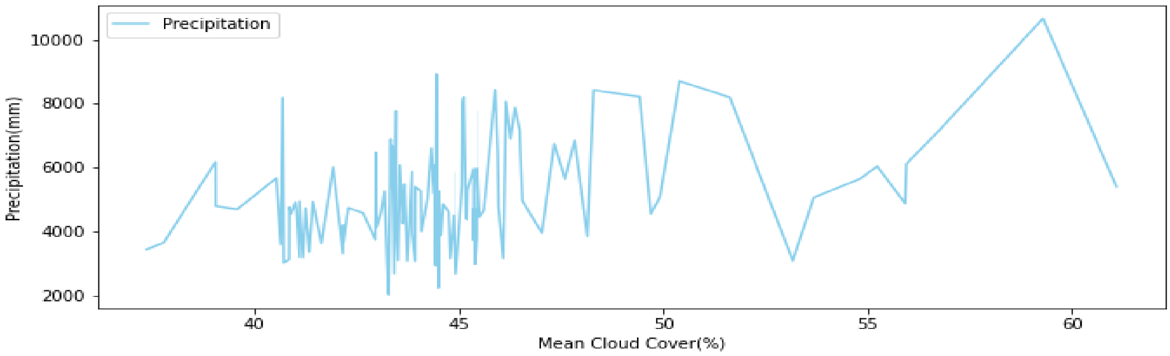


Fig 8. Correlation between cloud cover and precipitation spikes for the summer season (Jun, Jul, Aug).

5. Methodology

The data considered for the study is for the years 1901-2019 for temperature (*mean, maximum and minimum*), precipitation, and cloud cover. The data are provided in the gridded form and are prepared as monthly mean values, except for precipitation, which is taken as the sum total of precipitation observed in the entire region. Upon preparation of the data files, the models that produce the best results in terms of forecast need to be identified.

5.1 Model Description

Auto-Regressive Integrated Moving Average (ARIMA) model is used for time series analysis to better understand the underlying trends in the data and to produce forecasts based on the mean, deviation, and differences of the past values [21]–[23], [38]–[40]. The usage of the ARIMA model involves three main hyperparameters:

1. p – which denotes the order of the AR (auto-regressive) term i.e., the number of prior values the current value in a timeseries is regressed upon,
2. d – which denotes the I (integrated) term i.e., numbers of differencing required for the timeseries to make it stationary
3. q – which denotes the order of the MA (moving average) term i.e., the number of error values occurring at various time intervals in the past that the regression error is a linear combination of.

These p(AR), d(MA) and q(I) values need to be selected so that the model fits the data in the best way possible. If the timeseries to be used is non-stationary, its trend is removed by differencing to obtain a stationary timeseries. Non-seasonal ARIMA models are denoted as ARIMA (p, d, q) and seasonal ARIMA models as ARIMA (p, d, q) (P, D, Q) $_m$ where "m" denotes the number of periods in each season. ARIMA is considered to be one of the best models to use, when dealing with long and stable timeseries data, especially for approximating historical patterns for the future [41]–[43]. An ARIMA model can be mathematically described as:

$$\hat{y}_t = \mu + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q}$$

The equation follows the Box-Jenkins convention wherein the MA parameters (θ) are defined so that their signs are negative. Identifying the appropriate ARIMA model for some data begins by determining the order of differencing (d) required to make the timeseries stationary and remove any seasonality. Stationary series can still have autocorrelated errors, which is suggestive of the need for some number of AR terms ($p \geq 1$) and possibly some MA terms ($q \geq 1$) in the forecasting equation.

Separate ARIMA models have been fitted to forecast mean temperature, maximum temperature and minimum temperature, and cloud cover following the methodology presented in Figure 9. Precipitation achieved a relatively higher RMSE value, due to the inability of the ARIMA model to adapt to extreme changes in the data, with values ranging from 60mm/month, all the way up to around 19000 mm/month. identification of the best fit models was done using the metrics discussed in detail in section 5.2.

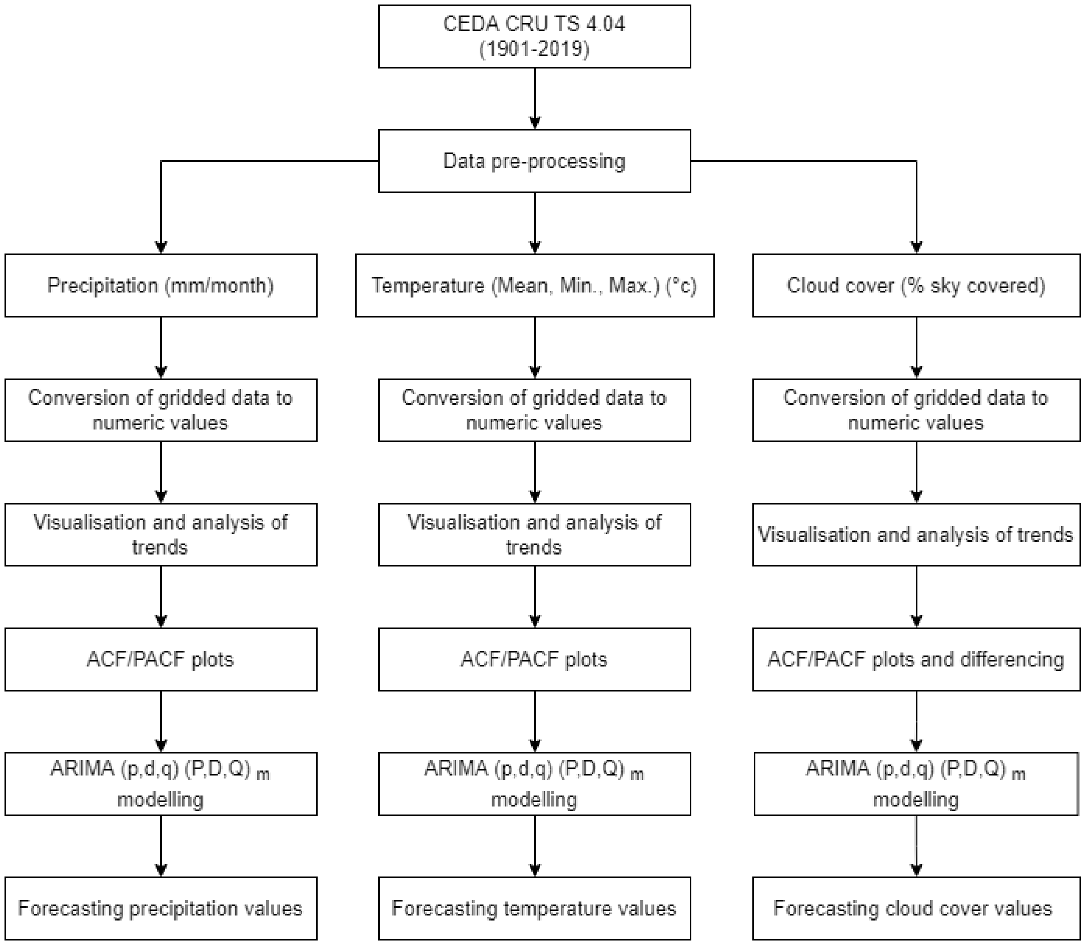


Fig 9. Methodology followed for forecasting the climate change.

5.2 Evaluation Metrics

The results are graded using multiple evaluation metrics, such as RMSE, AIC, BIC as well as distribution metrics such as skew, kurtosis, and stationarity R^2, all of which are inherent to the ARIMA model.

(i). Akaike Information Criterion (AIC)

AIC is an estimator of prediction error. When a statistical model is used to represent the process that generated the provided data, the representation is never exact, some information is always lost. AIC simply estimates this relative amount of information such as a part of the climate observations lost by our model. A lower AIC value implies a better model.

$$\mathrm{AIC} = 2 \times \mathrm{K} - 2 \times \ln(\mathrm{loglikelihood})$$

K being the number of parameters in the data.

(ii). Bayesian Information Criterion (BIC)

BIC tries to find the “True” model amongst a finite set of models. BIC introduces penalty terms into a model for each parameter that may increase the likelihood, thus preventing overfitting, i.e., preventing the model from simply repeating temperature or precipitation values instead of actually trying to understand the basis for the patterns. Similar to AIC, lower BIC values are preferred and considered to make for better models.

$$BIC = -2 \times loglikelihood + K \times \log(N)$$

K being the number of parameters in the data and N being the size of the dataset.

(iii). Root Mean Square Error (RMSE)

The RMSE denotes how close the values predicted by the trained model are to the actual observed values. The differences between the observed and predicted values, such as temperature or precipitation, known as residuals, help in the determination of a model's quality.

$$RMSE = \sqrt{mean(forecast - observed)^2}$$

(iv). R-squared

The R-squared value denotes how well the model has fit on the training data, thus providing a measure for the model's evaluation instead of the forecasts, i.e., how well the model has understood previous trends from the weather data, the closer the value to 1, the better the model is considered.

$$R^2 = 1 - \frac{unexplained\ variation}{total\ variation}$$

5.3 Stationarity Tests

The first step while dealing with any timeseries data is to determine whether it is stationary or not, which can be done using multiple tests. In this study, we used one such test known as the Augmented Dickey-Fuller Test (ADF). The p-value from the test is used to test the null hypothesis and any values higher than the significant threshold of 0.05 (5%) deem the rejection of the null hypothesis false. The conclusion from the results thus reached was that only the cloud cover timeseries data was non-stationary, while the rest of the data were stationary. To make the cloud cover data stationary, first-order differencing was performed. As auto-correlated errors could still exist in the differenced timeseries, AR ($p \geq 1$) and MA ($q \geq 1$) terms could be added to compensate for any mild under-differencing or over-differencing respectively. The ADF test results are shown in Table 4 and the differencing plots in figure 10.

	Mean Temperature	**Avg. Min. Temperature**	**Avg. Max. Temperature**	**Precipitation**	**Cloud Cover**
Test Statistics	-4.519	-4.186	-4.938	-5.494	-0.629
p-value	0.000	0.001	0.000	0.000	0.864
Lags	23	23	23	24	23
Observations	1404	1404	1404	1403	1404
Reject	True	True	True	True	False

Table 4 Augmented Dickey-Fuller Test results for the dataset.

5.4 Determination of AR(p) and MA(q) terms

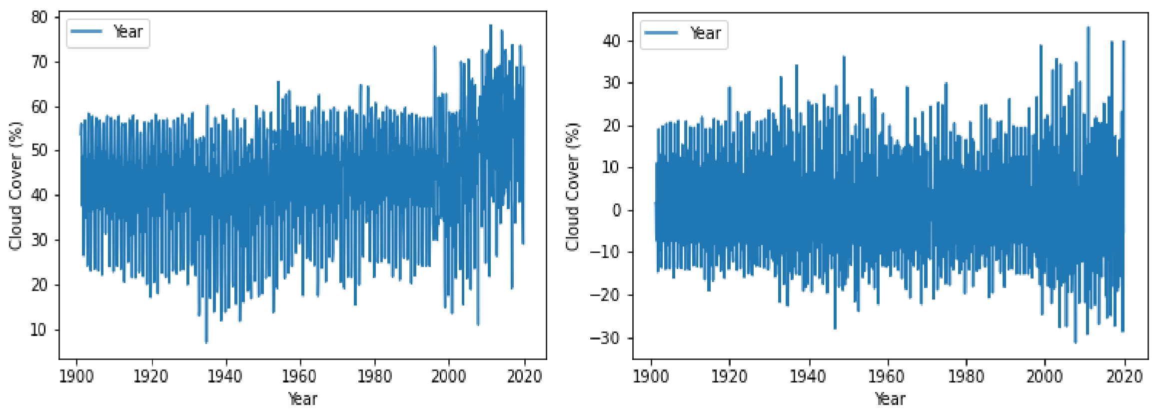
Auto-correlation function (ACF) and partial auto-correlation function (PACF) plots were used for determining the values for the AR (p) and MA (q) terms. The ACF and PACF plots showing large values decomposing very slowly over time shows the need for differencing. The plots also show a confidence level (95%), towards which, if the values show a sharp cut-off in ACF, denotes some over differencing in the series and the need for an MA(q) term. In case of PACF, a sharp cut-off represents under differencing and the need for an AR (p) term. The ACF-PACF plots are shown in fig. 11-15. It was found that in the case of mean temperature, maximum and minimum temperature, the ACF plot showed a sharp cut-off after the 3rd lag, thus providing us the value for the MA(q) term. The PACF plot showed a sharp cut-off after the 12th lag and a dip after the 4th lag, providing us with a value of either 12 or 4 for the AR(q) term, both of which were then used in the ARIMA model and the better performing of the two selected. So, the AR(p) term for mean temperature was 12 and that for minimum and maximum temperature was either 12 or 4, and the MA(q) term for all three of the temperature variables was 3, resulting in the selection of ARIMA (12,1,3) for mean temperature, ARIMA (12,1,3) for maximum temperature, ARIMA (4,1,3) for minimum temperature. Similar steps were followed to determine the 'p' and 'q' terms for both cloud cover and precipitation, with ACF plots for cloud cover showing a cut-off after the 12th lag and the PACF plot also showing a cut-off after the 12th lag, resulting in the selection of ARIMA (12, 1, 12). The ACF plot for precipitation shows a repeating pattern, indicative of one of the pattern cut-offs being the ideal value for the 'q' term, while its PACF plot shows a cut-off after the 11th lag, and ARIMA (11, 1, 12)

5.5 Determination of the 'm' term

The 'm' term is used to remove seasonality in the timeseries. Since the data is a monthly dataset, with each year having 12 months, the same value was chosen, i.e., m=12. Appropriate values for P, D, Q are then chosen for m=12 in a similar manner as the values for p, d and q.

5.6 Selection of the best ARIMA model

Once the appropriate values for the hyperparameters (p, d, q, P, D, and Q) are calculated, the best-fitting model is determined based on the residual values of the ARIMA models. Data from the years 1901-2000 is used for training the models and forecasting is performed from 2001-2100, with the data from 2001-

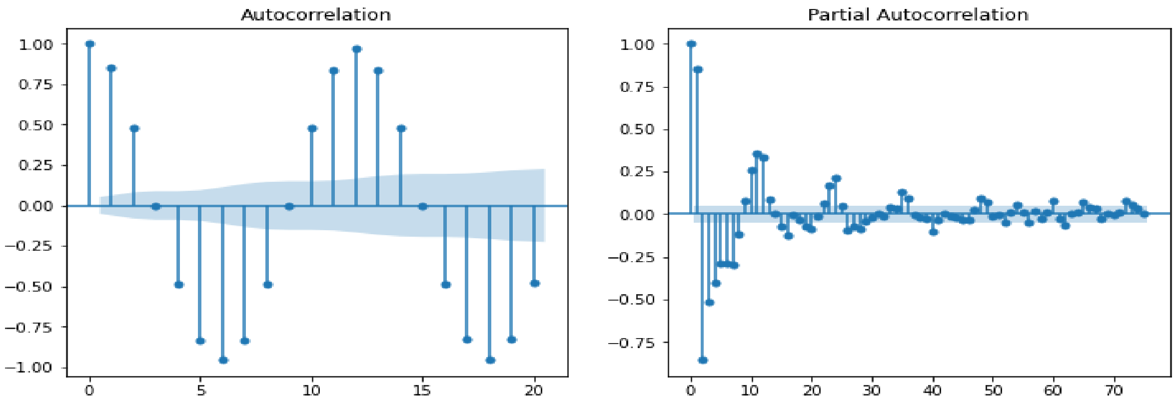


(a) **(b)**

Fig 10. Differencing plots for cloud cover: (a). Before Differencing and (b). After Differencing

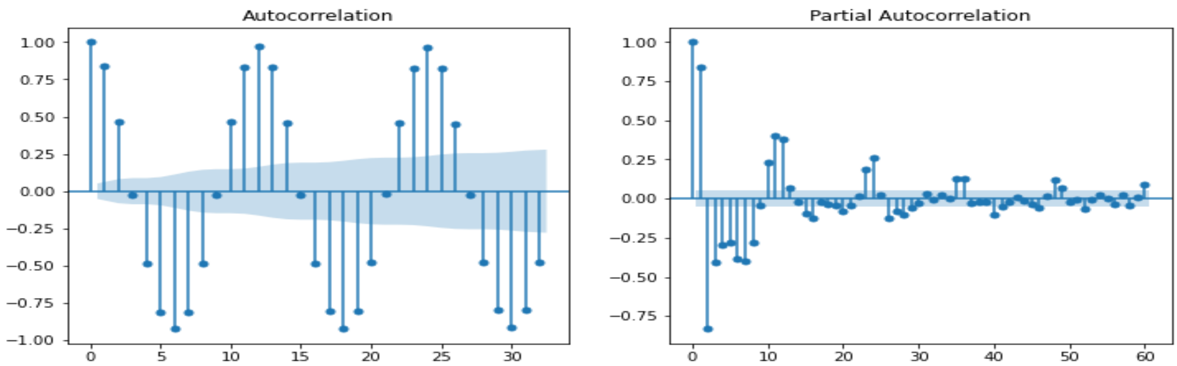
2019 acting as validation sets for the fitted models. The models are tested using the evaluation metrics described before in section 5.2.

The most appropriate model for mean temperature is found to be ARIMA (12,0,3) $(0,0,3)_{12}$, ARIMA (12,1,3) $(0,1,3)_{12}$ for maximum temperature and minimum temperature, ARIMA (4,1,3) $(0,1,3)_{12}$. The model selected for precipitation is ARIMA (11,1,12). In a similar manner, ARIMA (12, 1, 12) is selected for modelling the cloud cover data.



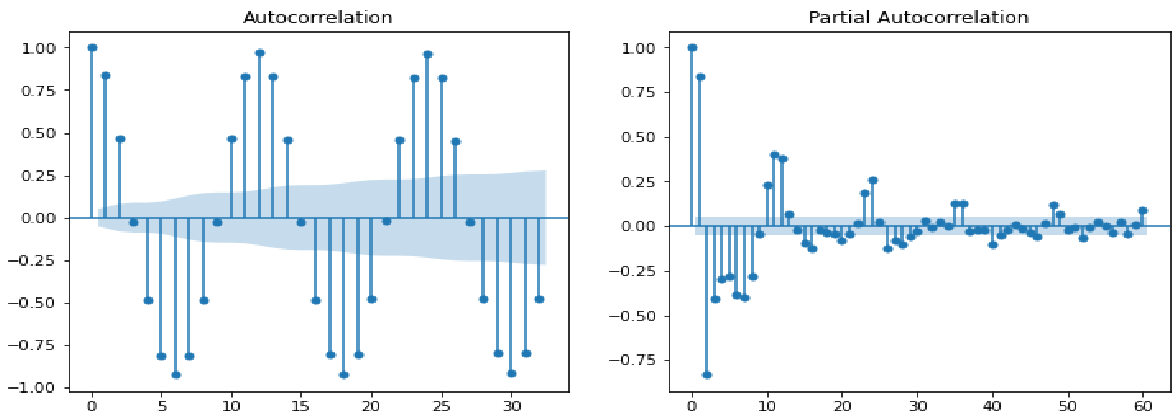
(a) **(b)**

Fig 11. ACF and PACF plots for Mean Temperature.



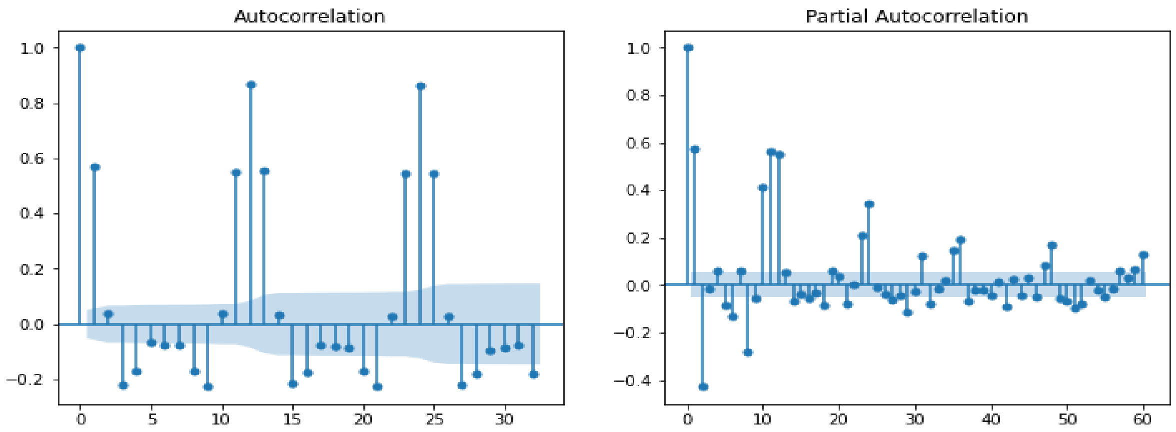
(a) **(b)**

Fig 12. ACF and PACF plots for Average Minimum Temperature.



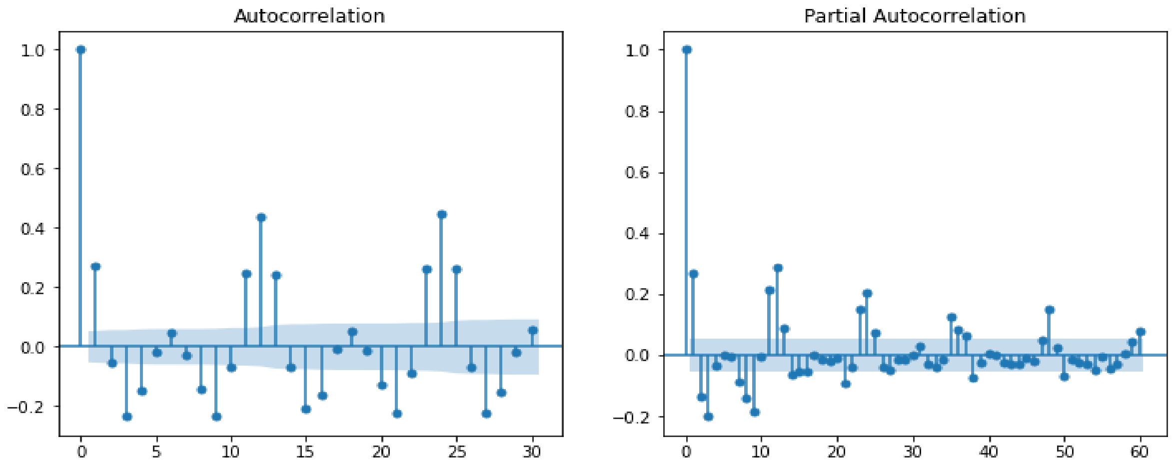
(a) **(b)**

Fig 13. ACF and PACF plots for Average Maximum Temperature.



(b) **(b)**

Fig 14. ACF and PACF plots for Cloud Cover.



(a) **(b)**

Fig 15. ACF and PACF plots for total precipitation.

6. Results

The time-series data for each of the variables *mean temperature, maximum temperature, minimum temperature, precipitation, and cloud cover* was separately trained upon & forecasted using different ARIMA models. The training data used was from the years 1901-1999 and forecasts were made for the next century i.e., 2000-2100 with data from 2000-2019 being used for validation of the forecasts. Although the ADF tests showed the data for precipitation, mean temperature, minimum and maximum temperature to be stationary, some seasonality was found using ACF/PACF plots in minimum and maximum temperature, which showed the timeseries having some seasonal properties such as sequentially recurring trends in the lags and slow decay over time.

A computer system with an AMD Ryzen 5 4500U APU, 16 GB RAM and AMD Vega 8 GPU was used for training and visualizing the data, as well as making forecasts. Microsoft Excel was mainly used to design and build the plots, for analysis and forecasting. The modeling and forecasting process was undertaken using Python 3.9 and Python libraries viz., as Matplotlib, Seaborn, Statsmodels, Pyplot, Plotly, NumPy, Pandas, SciKit Learn.

6.1 Temperature Forecasting Results

The results for the mean temperature, minimum temperature, and maximum temperature show a very low RMSE value, considering the fact that the average values of the parameters in the region lie around 18-23°C. The mean temperature shows the most prominent change, doubling by the end of the century for the month of January and showing an increase of about 50% on average for the other two months of winter (December and February) when compared to that in 2020. Springs are also predicted to get hotter, which further worsens the issue regarding the lack of the "chilling period" required by fruit trees in the region, as discussed earlier in the analysis phase. The summers and autumns are seen to be getting cooler when looking at the mean temperature for Jun-Nov, whilst also seeing an increase in the minimum and maximum temperatures, pointing to an increase in the erratic behavior seen in the said seasons over recent years.

The models fit for each of the temperature variables (mean, maximum and minimum) have been shown in detail in fig 16-18, and their evaluation results in table 5. The forecast results have been made yearly and shown as values in 20-year intervals between 2020 and 2100 in table 6-8, and to better visualize the change predicted in the temperature, plots showing the same are provided in fig 19-21.

Climate Variable	AIC	BIC	RMSE	Skew	Kurtosis	R^2
Mean Temperature	3442.07	3537.94	1.1296	-0.06	3.92	0.973
Minimum Temperature	3288.46	3343.84	1.1504	0.17	3.70	0.970
Maximum Temperature	3432.28	3527.93	1.1514	0.09	3.84	0.974

Table 5 - ARIMA model evaluation results for mean, minimum and maximum temperature.

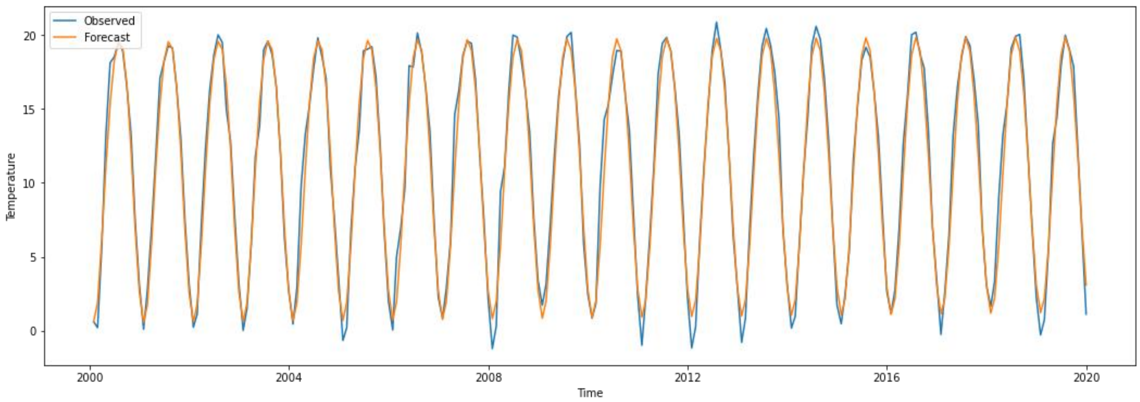


Fig 16 - ARIMA (12,0,3) $(0,0,3)_{12}$ model validation plot for mean temperature (°C).

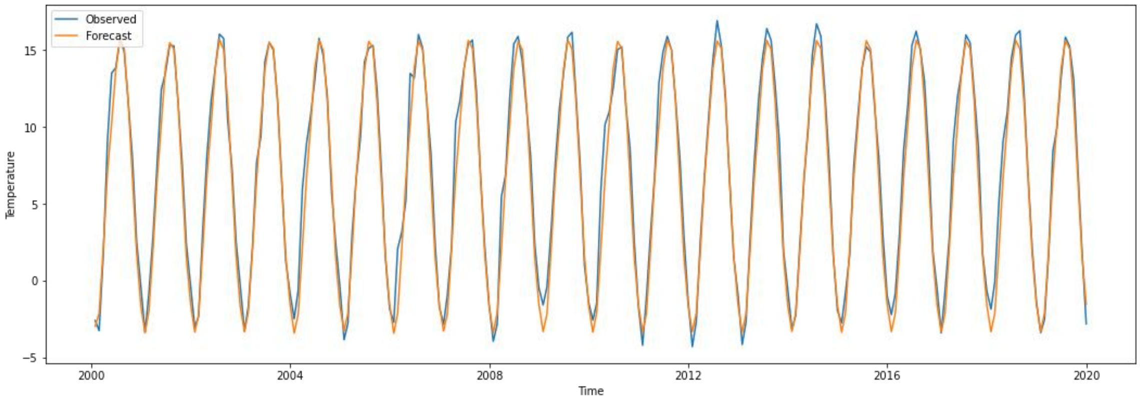


Fig 17. ARIMA (12,1,3) (0,1,3) $_{12}$ model validation plot for minimum temperature (°C).

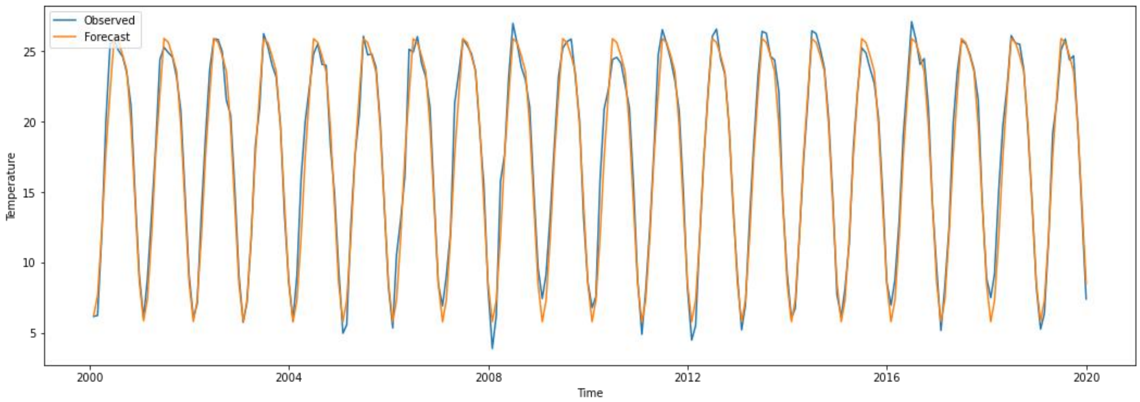


Fig 18. ARIMA (4,1,3) (0,1,3)$_{12}$ model validation plot for maximum temperature (°C).

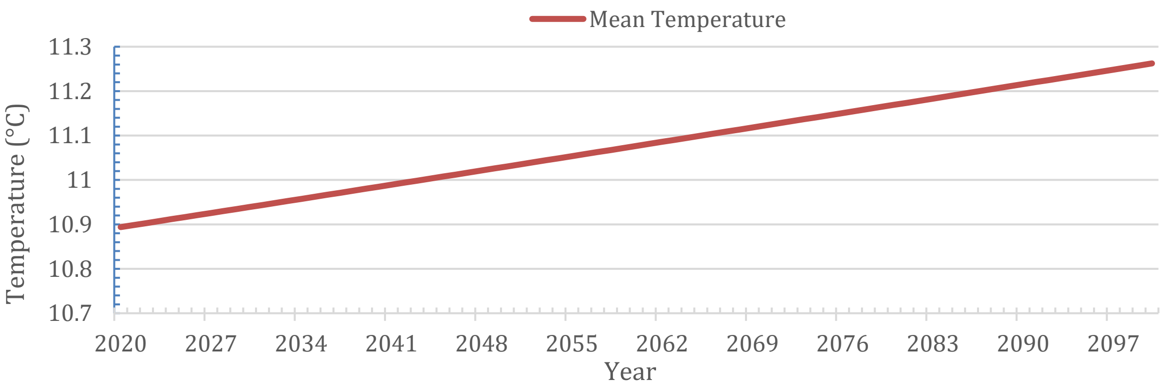


Fig 19. Yearly mean temperature (°C) changes over the years 2020-2100.

Month	Year				
	2020	**2040**	**2060**	**2080**	**2100**
JAN	**1.264**	1.837	2.246	2.574	**2.874**
FEB	**2.243**	2.723	3.165	3.544	**3.880**
MAR	5.773	5.951	6.251	6.565	6.856
APR	10.664	10.596	10.685	10.846	11.027
MAY	15.376	15.287	15.232	15.233	15.270
JUN	18.657	18.698	18.617	18.521	18.442
JUL	19.864	19.972	19.927	19.813	19.681
AUG	18.904	18.894	18.856	18.769	18.652
SEP	16.029	15.823	15.748	15.697	15.639
OCT	11.773	11.527	11.444	11.440	11.463
NOV	7.050	7.033	7.053	7.131	7.247
DEC	**3.132**	3.475	3.696	3.899	**4.114**

Table 6 Mean temperature (°C) forecasts for the years 2020-2100.

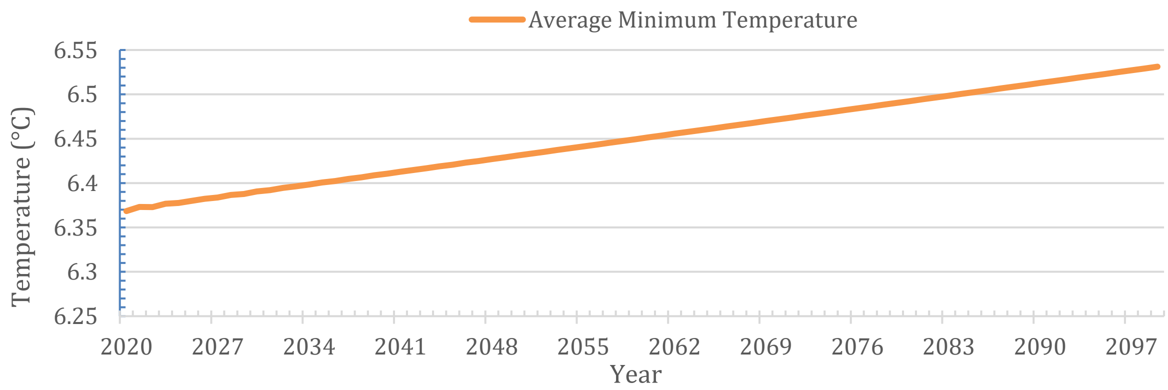


Fig 20. Yearly average minimum temperature (°C) changes over the years 2020-2100.

Month	Year				
	2020	**2040**	**2060**	**2080**	**2100**
JAN	-3.335	-3.289	-3.247	-3.206	**-3.165**
FEB	-2.127	-2.074	-2.034	-1.993	-1.952
MAR	2.133	2.163	2.203	2.244	2.284
APR	6.710	6.746	6.788	6.829	6.870
MAY	10.222	10.277	10.318	10.359	10.400
JUN	13.827	13.864	13.904	13.945	13.985
JUL	15.638	15.668	15.709	15.750	15.791
AUG	15.081	15.131	15.173	15.214	15.255
SEP	11.638	11.684	11.724	11.765	11.806
OCT	6.584	6.612	6.652	6.693	6.734
NOV	1.629	1.672	**1.714**	1.755	1.796
DEC	**-1.579**	-1.527	-1.486	-1.445	-1.404

Table 7 Average minimum temperature (°C) forecasts for the years 2020-2100.

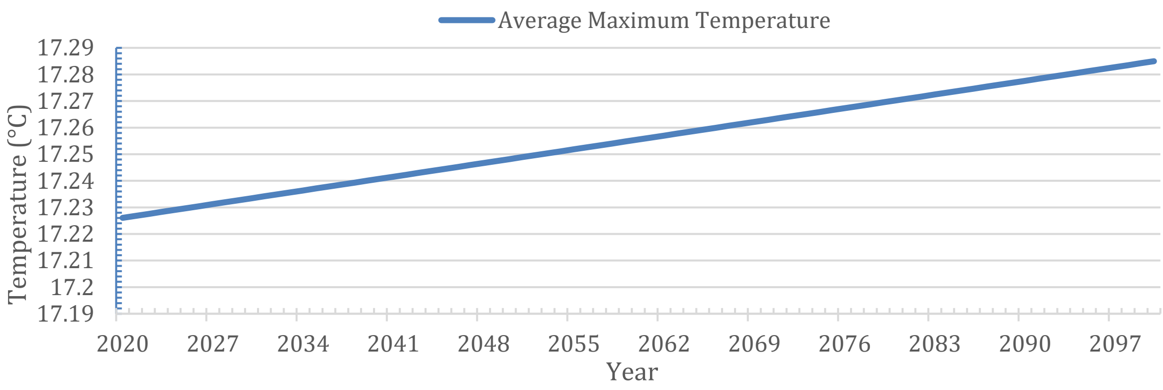


Fig 21. Yearly average maximum temperature (°C) changes over the years 2020-2100.

Month	**Year**				
	2020	**2040**	**2060**	**2080**	**2100**
JAN	5.799	5.814	5.829	5.844	5.858
FEB	7.324	7.339	7.354	7.369	7.383
MAR	11.890	11.905	11.920	11.934	11.949
APR	17.580	17.595	17.609	17.624	17.639
MAY	21.982	21.997	22.012	22.027	22.041
JUN	25.899	25.914	25.929	25.944	25.958
JUL	**25.623**	25.638	25.653	25.667	**25.682**
AUG	24.672	24.687	24.702	24.717	24.731
SEP	23.565	23.580	23.595	23.610	23.624
OCT	19.675	19.689	19.704	19.719	19.734
NOV	14.185	14.200	14.215	14.229	14.244
DEC	8.516	8.531	8.545	8.560	8.575

Table 8 Average maximum temperature (°C) forecasts for the years 2020-2100.

6.2 Cloud Cover Forecasting Results

The ARIMA model fitted for cloud cover showed an increase in the RMSE than that achieved for the model fitted for temperature values. The model thus tends to underestimate the peaks of the cloud and makes forecasts that are around 4-8% lower than what the actual values may be.

The evaluation metrics for the model are shown in table 9 and the model fit values are compared to actual observed values for validation in fig 22. The cloud cover values for different seasons have also been visualized in fig 23-26, with springs showing a very slight decrease, winters showing a relatively larger decline in values, summers and autumns showing slightly increasing values as the years go on.

The forecast results from the model are presented in table 10, in 20-year intervals.

	AIC	**BIC**	**RMSE**	**Skew**	**Kurtosis**	**R^2**
Cloud Cover	6262.51	6389.22	11.8918	-0.462	-0.967	0.2483

Table 9 ARIMA model evaluation results for cloud cover.

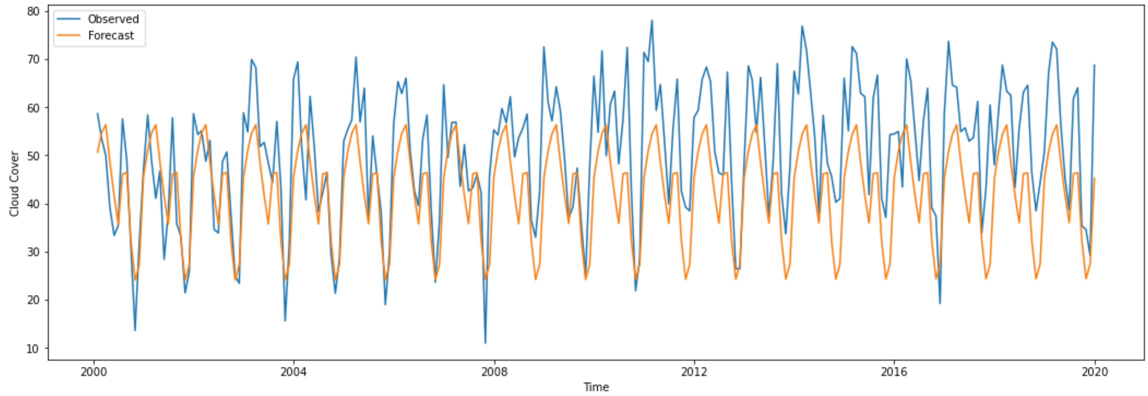


Fig 22. ARIMA (12,1,12) model validation plot for cloud cover (%).

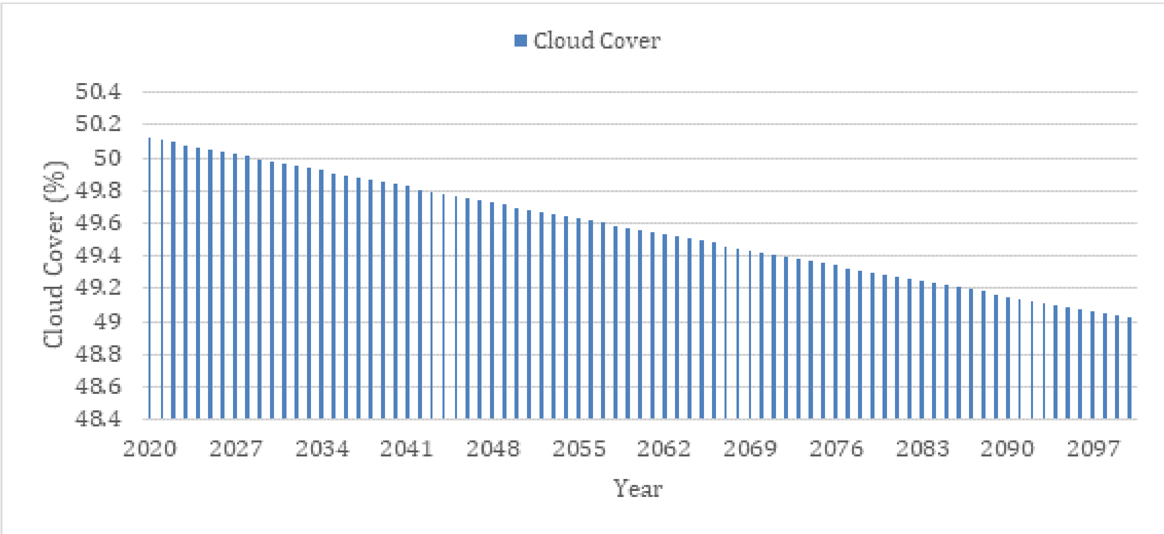


Fig 23. Cloud cover (%) forecast changes for winter (Dec, Jan, Feb) over the years 2020-2100.

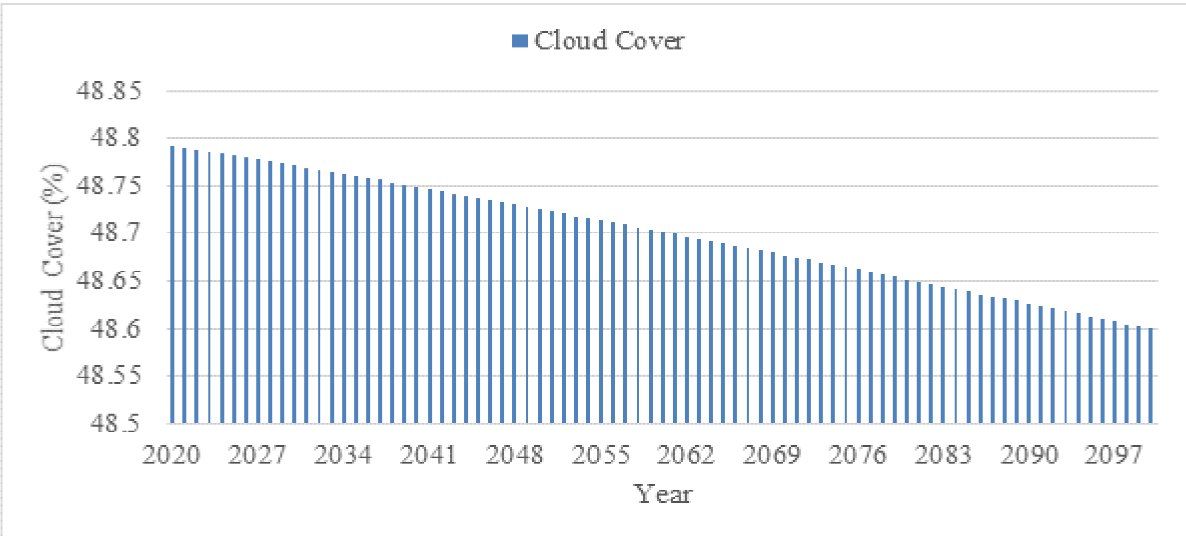


Fig 24. Cloud cover (%) forecast changes for spring (Mar, Apr, May) over the years 2020-2100.

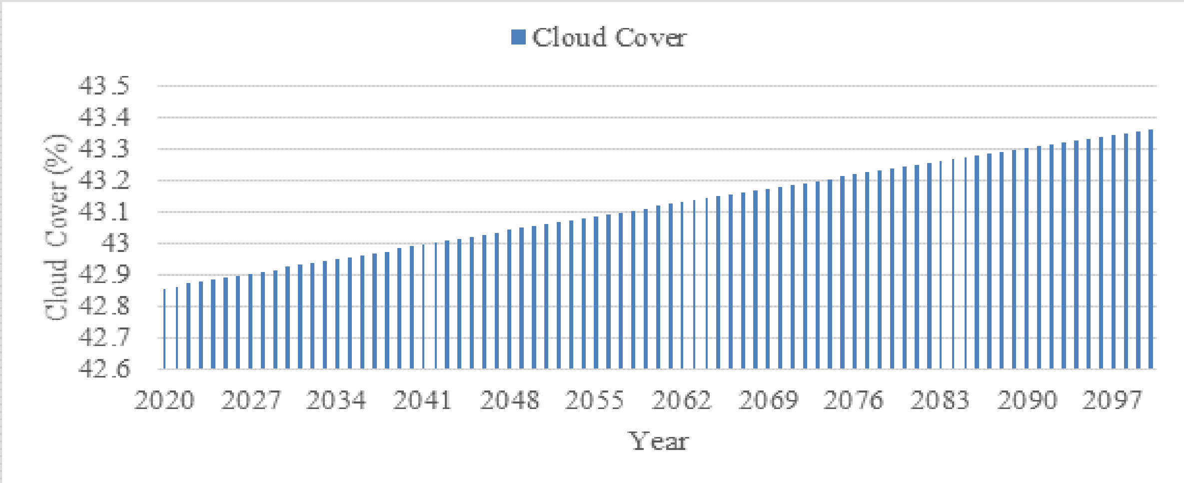


Fig 25. Cloud cover (%) forecast changes for summer (Jun, Jul, Aug) over the years 2020-2100.

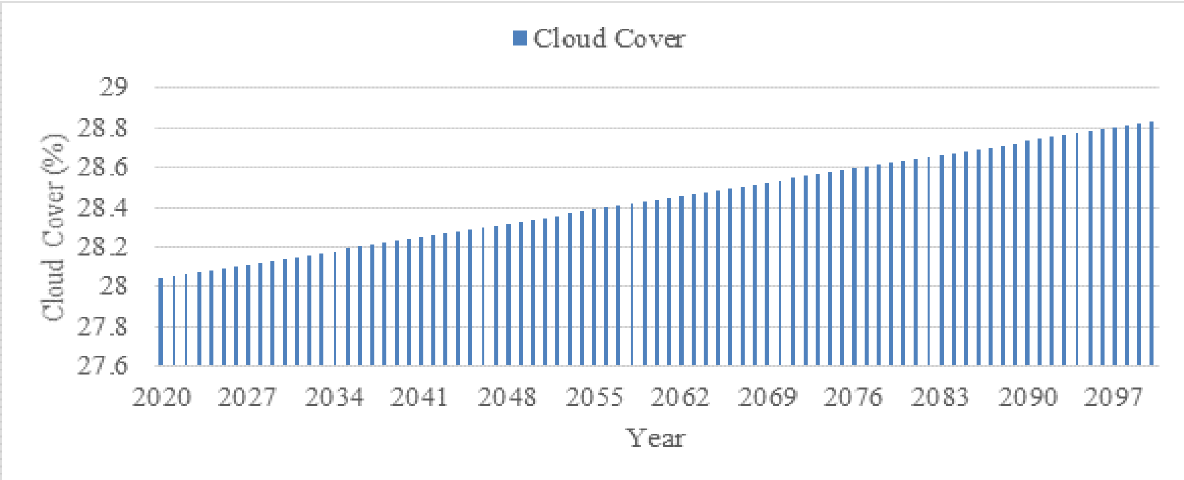


Fig 26. Cloud cover (%) forecast changes for autumn (Sep, Oct, Nov) over the years 2020-2100.

Month	Year				
	2020	2040	2060	2080	2100
JAN	50.845	50.598	50.356	50.122	49.894
FEB	54.339	54.054	53.778	53.509	53.247
MAR	56.310	56.225	56.127	56.020	55.904
APR	48.436	48.314	48.201	48.094	47.992
MAY	41.633	41.708	41.776	41.841	41.902
JUN	36.039	36.447	36.850	37.243	37.628
JUL	46.171	46.194	46.212	46.225	46.236
AUG	46.365	46.330	46.297	46.264	46.231
SEP	32.266	32.273	32.288	32.314	32.351
OCT	24.373	24.717	25.054	25.382	25.701
NOV	27.485	27.730	27.970	28.207	28.441
DEC	45.190	44.861	44.541	44.228	43.923

Table 10 Forecasts for average yearly cloud cover (%).

6.3 Precipitation Forecasting Results

The ARIMA model produces a high RMSE value for the precipitation dataset, mainly due to its inability to adapt to erratic changes in the precipitation which is essentially a high variance between the data points. The model is unable to replicate the peaks that are expected when comparing the forecasted values with the observed values and thus creates a curve with smaller peaks. The model evaluation results are discussed in table 11. The validation fit for the model has been shown as a plot in fig 27, and the values forecast using the model are presented in table 12, with visualizations for the same in fig 28, showing an overall decrease in the total precipitation in almost all seasons, especially the summers, pointing to drier climates.

	AIC	BIC	RMSE	Skew	Kurtosis	R^2
Precipitation	18005.49	18127.13	504.923	-0.073	-1.106	0.457

Table 11 ARIMA model evaluation results for precipitation.

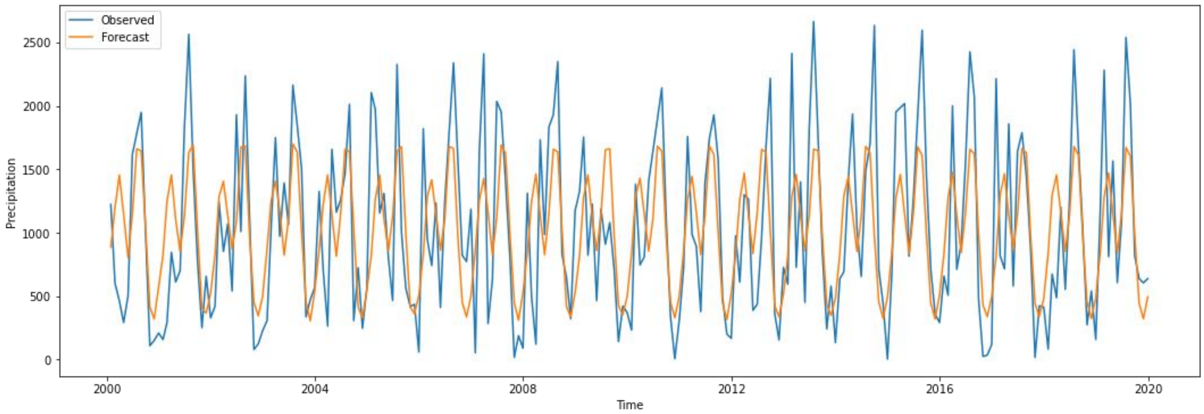


Fig 27. ARIMA (11,1,12) model validation plot for precipitation (mm/month).

Month	Year				
	2020	**2040**	**2060**	**2080**	**2100**
JAN	827.259	822.059	829.926	846.947	873.467
FEB	1292.45	1353.94	1408.87	1463.19	1515.35
MAR	1482.47	1513.28	1544.88	1571.42	1592.40
APR	1135.92	1144.92	1153.38	1163.77	1179.93
MAY	850.69	877.804	909.056	949.669	997.282
JUN	1163.02	1196.69	1236.43	1271.64	1298.92
JUL	1659.21	1655.83	1634.69	1600.16	1555.79
AUG	**1612.55**	**1547.77**	**1479.96**	**1415.04**	**1353.35**
SEP	983.912	945.346	921.974	904.800	891.905
OCT	434.783	458.288	471.460	478.244	478.894
NOV	336.650	317.251	290.615	258.438	220.468
DEC	483.213	427.805	381.267	**339.437**	**304.823**

Table 12 Forecasts for total precipitation (mm/month) for the years 2020-2100.

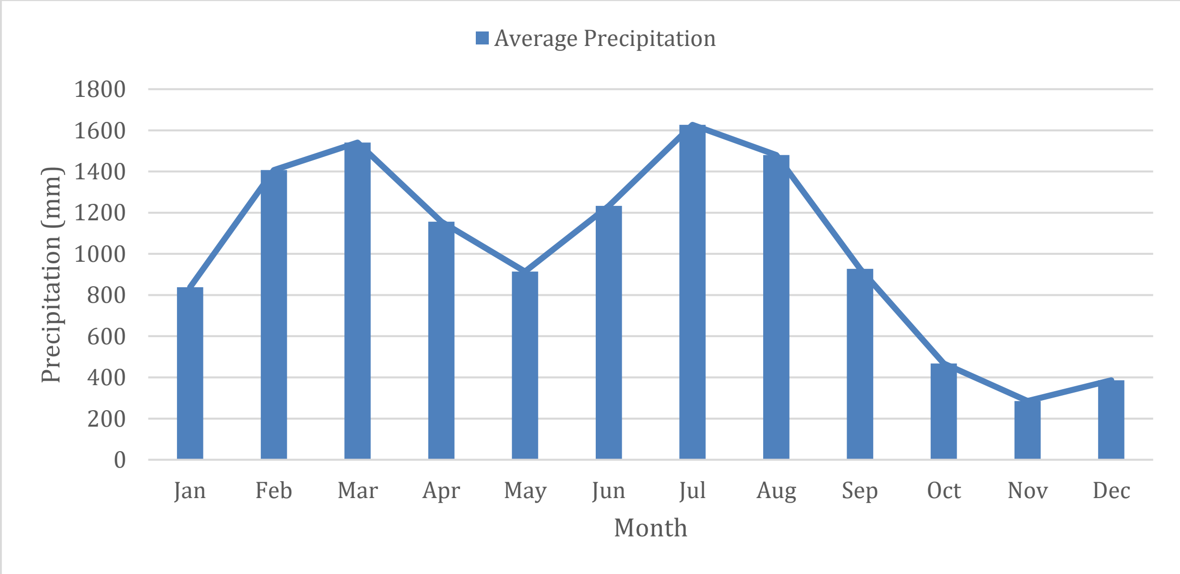


Fig 28. Forecasted average precipitation (mm/month) for the years 2020-2100.

7. Conclusion

The climate of any region is mainly dependent on the temperature and precipitation observed there. In the study, these factors, in addition to the cloud cover were studied using the ARIMA model, and forecasts were made for the valley of Kashmir.

The temperature variables i.e., mean, maximum and minimum temperature were predicted by the model with very low error values and it was found that the mean temperatures for the region, especially in winters are going to shift further from the natural range, towards an overall hotter climate, with a forecasted increase of nearly 2°C for winters by the end of the century when compared to that in 2019. Another concerning trend seen with the mean temperatures is such an increase that winter months with temperatures lower than 2°C are completely non-existent, starting from the year 2048, and for a region such as the Kashmir valley, such changes can prove to be detrimental to the ecosystem of the area, as well as to its economy.

The minimum and maximum temperatures also see a rise, although less pronounced than the mean temperature, with projections showing the summers and autumns in the region to get ever so slightly cooler, while the springs show a relatively higher increase in temperature.

The projected temperature shifts will deny important periods of low temperature during the springs and in a similar vein, periods of high temperature during the autumns, which is expected to adversely affect the fruit and crop yield of the valley, potentially proving to be dire to a large percentile of the population which is economically directly dependent on their crop yield.

Although the ARIMA model was found to be unable to predict extreme rainfall events with a reasonably small error value, the mean precipitation forecasts were found to be closer to actuality. The decrease in precipitation and increase in cloud cover during summers implies that an increasingly the humid climate is abound for the valley, which can further prove to be disastrous for the valley, considering the sensitive nature of the region.

Perhaps the most alarming finding in the study is that if the changes projected in all the climate parameters are taken in unison, they point to a calamitic outcome for the economic avenues of the region, such as hydroelectricity, tourism, and irrigation, which makes the issue even more deserving of attention. The temperature shift and increased cloud cover are bound to accelerate the melting of glaciers, which, in tandem with erratic precipitation, will cause an increase in the potential of regular flooding, not only in areas surrounding the glaciers but plains as well. It is high time that the government and stakeholders engage scientists and researchers to develop strategies and build roadmaps for tackling changes in the climate such as the ones projected in this paper, to prevent economic, environmental, and human losses in the region.

8. Future Scope

Although strongly accurate results with very low error values were achieved for mean, average minimum, and average maximum temperature variables, the results for cloud cover and precipitation did not reflect similar characteristics, owing to the inability of the ARIMA model to fit to, and replicate peaks and drastic changes. It can, thus, be said that with the use of different machine learning models and deep learning models, the results have the potential to be further improved upon, with a better prediction rate for extreme events.

Another area that can be improved upon, is the use of higher-resolution datasets, which can allow for finer, cluster-wise analyses and forecasts. Such datasets can allow for the separation of data for planes, and mountainous regions, providing a clearer picture of the impact on glaciers and snow-caps in the region.

Analysis and use of multiple different datasets such as the CRU dataset and satellite imagery datasets for glaciers, in a single study, can also aid in ironing out flaws that could be present in studies using one single dataset. Some room for improvement is also visible in collecting data for less studied variables in the region, such as CO_2 emissions, which is considered one of the most important factors in the health of the region's climate; Studies conducted on such datasets can paint a clearer picture of the extent to which the region's climate has deteriorated, especially given the move towards a higher level of industrialization in the region.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability

The dataset used in the study and the codebase is available at https://github.com/fahad-farooq/climchange_jk

Disclaimer

The authors have no financial or non-financial conflicts of interest to declare

References

[1] United Nations Organization, "Global Issues - Climate Change," 2021.

[2] A. Core Writing Team, Pachauri, R.K and Reisinger and (eds.)], "PCC, 2007: Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change," 2008.

[3] IPCC. Global Warming of 1.5°C, Intergovernmental Panel on Climate Change (IPCC) Special Report 2015, "No Title."

[4] "AR6 Climate Change 2021, Intergovernmental Panel on Climate Change (IPCC) Special Report 2021, IPCC."

[5] V. M. Meher-Homji, "On the mediterranean climatic regime of West Pakistan," *Archiv für Meteorologie, Geophysik und Bioklimatologie, Serie B*, vol. 19, no. 3, pp. 277–286, 1971.

[6] M. Ahmad, "Development of Hydropower Projects in Jammu and Kashmir," *International Journal of Trend in Research and Development*, vol. 6, no. 1, pp. 6–8, 2019.

[7] G. SMITh, "In Kashmir, water treaty means less power to the people." https://www.theglobeandmail.com/news/world/in-kashmir-water-treaty-means-less-power-to-the-people/article1389907/

[8] M. Ahmed, "Development of Hydropower Projects in Jammu and Kashmir," *International Journal of Trend in Research and Development*, vol. 6, no. 1, pp. 6–8, 2019.

[9] Parvaiz Bukhari, "Climate change ravages Kashmir's 'red gold' saffron crop," 2020.

[10] S. A. Romshoo, M. Fayaz, G. Meraj, and I. M. Bahuguna, "Satellite-observed glacier recession in the Kashmir Himalaya, India, from 1980 to 2018," *Environ Monit Assess*, vol. 192, no. 9, pp. 1–17, 2020.

[11] M. H. Wani, S. H. Baba, N. H. Bazaz, and H. Sehar, "Climate change in Kashmir valley: Is it initiating transformation of mountain agriculture?," *Indian J Econ Dev*, vol. 3, no. 2, pp. 142–154, 2015.

[12] S. A. Romshoo, J. Bashir, and I. Rashid, "Twenty-first century-end climate scenario of Jammu and Kashmir Himalaya, India, using ensemble climate models," *Clim Change*, vol. 162, no. 3, pp. 1473–1491, 2020.

[13] S. N. Zaz, S. A. Romshoo, R. T. Krishnamoorthy, and Y. Viswanadhapalli, "Analyses of temperature and precipitation in the Indian Jammu and Kashmir region for the 1980–2016 period: implications for remote influence and extreme events," *Atmos Chem Phys*, vol. 19, no. 1, pp. 15–37, 2019.

[14] D. Strigaro, M. Moretti, M. Mattavelli, I. Frigerio, M. De Amicis, and V. Maggi, "A GRASS GIS module to obtain an estimation of glacier behavior under climate change: A pilot study on Italian glacier," *Comput Geosci*, vol. 94, pp. 68–76, 2016.

[15] M. ul Shafiq, S. Ramzan, P. Ahmed, R. Mahmood, and A. P. Dimri, "Assessment of present and future climate change over Kashmir Himalayas, India," *Theor Appl Climatol*, vol. 137, no. 3, pp. 3183–3195, 2019.

[16] S. Ahsan, M. S. Bhat, A. Alam, H. Farooq, and H. A. Shiekh, "Evaluating the impact of climate change on extreme temperature and precipitation events over the Kashmir Himalaya," *Clim Dyn*, pp. 1–19, 2021.

[17] P. Narayanan, A. Basistha, S. Sarkar, and S. Kamna, "Trend analysis and ARIMA modelling of pre-monsoon rainfall data for western India," *Comptes Rendus Geoscience*, vol. 345, no. 1, pp. 22–27, 2013.

[18] F. Naseem *et al.*, "An integrated approach to air pollution modeling from climate change perspective using ARIMA forecasting," *Journal of Applied Agriculture and Biotechnology*, vol. 2, no. 2, pp. 37–44, 2018.

[19] S. Swain, S. Nandi, and P. Patel, "Development of an ARIMA model for monthly rainfall forecasting over Khordha district, Odisha, India," in *Recent Findings in Intelligent Computing Techniques*, Springer, 2018, pp. 325–331.

[20] T. Dimri, S. Ahmad, and M. Sharif, "Time series analysis of climate variables using seasonal ARIMA approach," *Journal of Earth System Science*, vol. 129, no. 1, pp. 1–16, 2020.

[21] F. Lhano, J. Gouveia, F. Perestrello, A. Melo, and M. Santa-Clara, "Using an ARIMA Model to Forecast GDP Until 2031 for Portugal and Germany," *Available at SSRN 3855082*, 2021.

[22] T. Nyoni, "Is Nigeria's economy progressing or backsliding? Implications from ARIMA models," 2019.

[23] F. M. Khan and R. Gupta, "ARIMA and NAR based prediction model for time series analysis of COVID-19 cases in India," *Journal of Safety Science and Resilience*, vol. 1, no. 1, pp. 12–18, 2020.

[24] I. Z. Ul and R. L. A. Khan, "Climate change scenario in Kashmir valley, India, based on seasonal and annual average temperature trends," *Disaster Advances*, vol. 6, no. 4, pp. 30–40, 2013.

[25] I. Harris, T. J. Osborn, P. Jones, and D. Lister, "Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset," *Sci Data*, vol. 7, no. 1, pp. 1–18, 2020.

[26] H. Shi, T. Li, and J. Wei, "Evaluation of the gridded CRU TS precipitation dataset with the point raingauge records over the Three-River Headwaters Region," *J Hydrol (Amst)*, vol. 548, pp. 322–332, 2017.

[27] A. R. Salvacion, D. B. Magcale-Macandog, P. C. S. Cruz, R. B. Saludes, I. B. Pangga, and C. J. R. Cumagun, "Evaluation and spatial downscaling of CRU TS precipitation data in the Philippines," *Model Earth Syst Environ*, vol. 4, no. 3, pp. 891–898, 2018.

[28] M. Ahmad, "It's High Time the Government Took Notice of Kashmir's Changing Climate," *WIRE*, 2018.

[29] E. T. Stafne, *Chilling-hour Requirement of Fruit Crops*. Mississippi State University Extension, 2017.

[30] N. R. Patel, A. Akarsh, A. Ponraj, and J. Singh, "Geospatial Technology for Climate Change Impact Assessment of Mountain Agriculture," in *Remote Sensing of Northwest Himalayan Ecosystems*, Springer, 2019, pp. 381–400.

[31] B. Das, H. Krishna, B. L. Attri, N. Ahmad, and J. K. Ranjan, "Harvest maturity standards and fruit quality of some apple cultivars under high altitudal conditions," *Indian Journal of Horticulture*, vol. 68, no. 2, pp. 170–179, 2011.

[32] A.-M. Salama *et al.*, "Temperate Fruit Trees under Climate Change: Challenges for Dormancy and Chilling Requirements in Warm Winter Regions," *Horticulturae*, vol. 7, no. 4, p. 86, 2021.

[33] "Kashmir flash floods leave dozens dead," *The Guardian*, Srinagar, Aug. 06, 2010.

[34] "The 2014 Kashmir Flood: The Extreme of the Extremes," 2015.

[35] S. Devi, "Pakistan floods: impact on food security and health systems," *The Lancet*, vol. 400, no. 10355, pp. 799–800, 2022.

[36] Raymond Zhong, "In a First Study of Pakistan's Floods, Scientists See Climate Change at Work," *https://www.nytimes.com/2022/09/15/climate/pakistan-floods-global-warming.html*, Sep. 15, 2022.

[37] UNICEF, "Devastating floods in Pakistan," *https://www.unicef.org/emergencies/devastating-floods-pakistan-2022*, Oct. 12, 2022.

[38] G. Ding, X. Li, Y. Shen, and J. Fan, "Brief Analysis of the ARIMA model on the COVID-19 in Italy," *medRxiv*, 2020.

[39] A. K. Sahai, N. Rath, V. Sood, and M. P. Singh, "ARIMA modelling & forecasting of COVID-19 in top five affected countries," *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, vol. 14, no. 5, pp. 1419–1427, 2020.

[40] J. Youssef, N. Ishker, and N. Fakhreddine, "GDP Forecast of the Biggest GCC Economies Using ARIMA," 2021.

[41] J. Jebeile, V. Lam, and T. Räz, "Understanding climate change with statistical downscaling and machine learning," *Synthese*, vol. 199, no. 1, pp. 1877–1897, 2021.

[42] D. Rolnick *et al.*, "Tackling climate change with machine learning," *arXiv preprint arXiv:1906.05433*, 2019.

[43] A. YoosefDoost, M. S. Sadeghian, M. NodeFarahani, and A. Rasekhi, "Comparison between performance of statistical and Low Cost ARIMA Model with GFDL, CM2. 1 and CGM 3 Atmosphere-Ocean General Circulation Models in assessment of the effects of climate change on temperature and precipitation in Taleghan Basin," *American Journal of Water Resources*, vol. 5, no. 4, pp. 92–99, 2017.