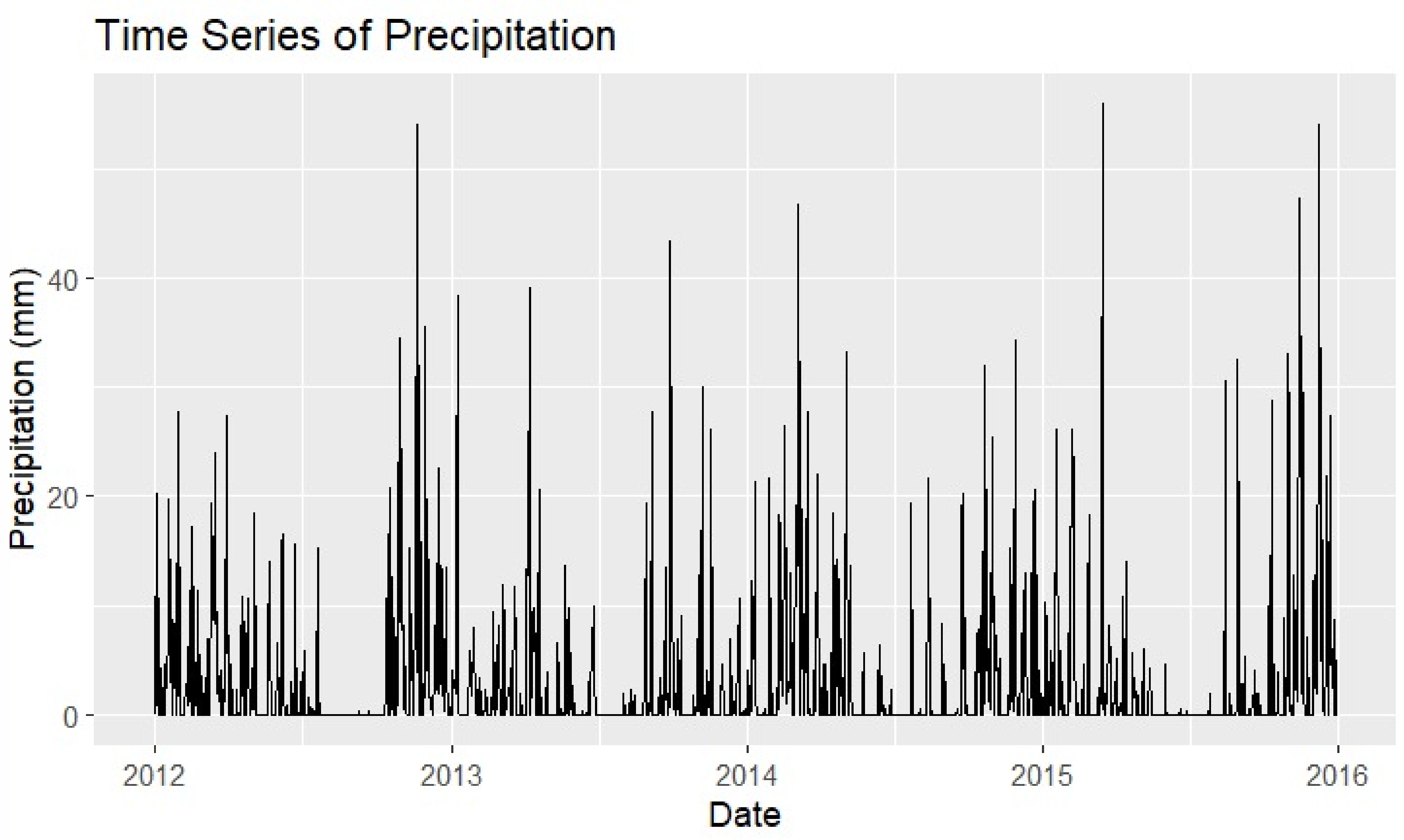
Analysis of Weather Data on R using Time Series

By: Tanmay D Salian
& Aditya Sharma

1) Plot the time series of precipitation, max temperature, and min temperature. What are the issues in each of them? Are each stationary? Why yes or why not?

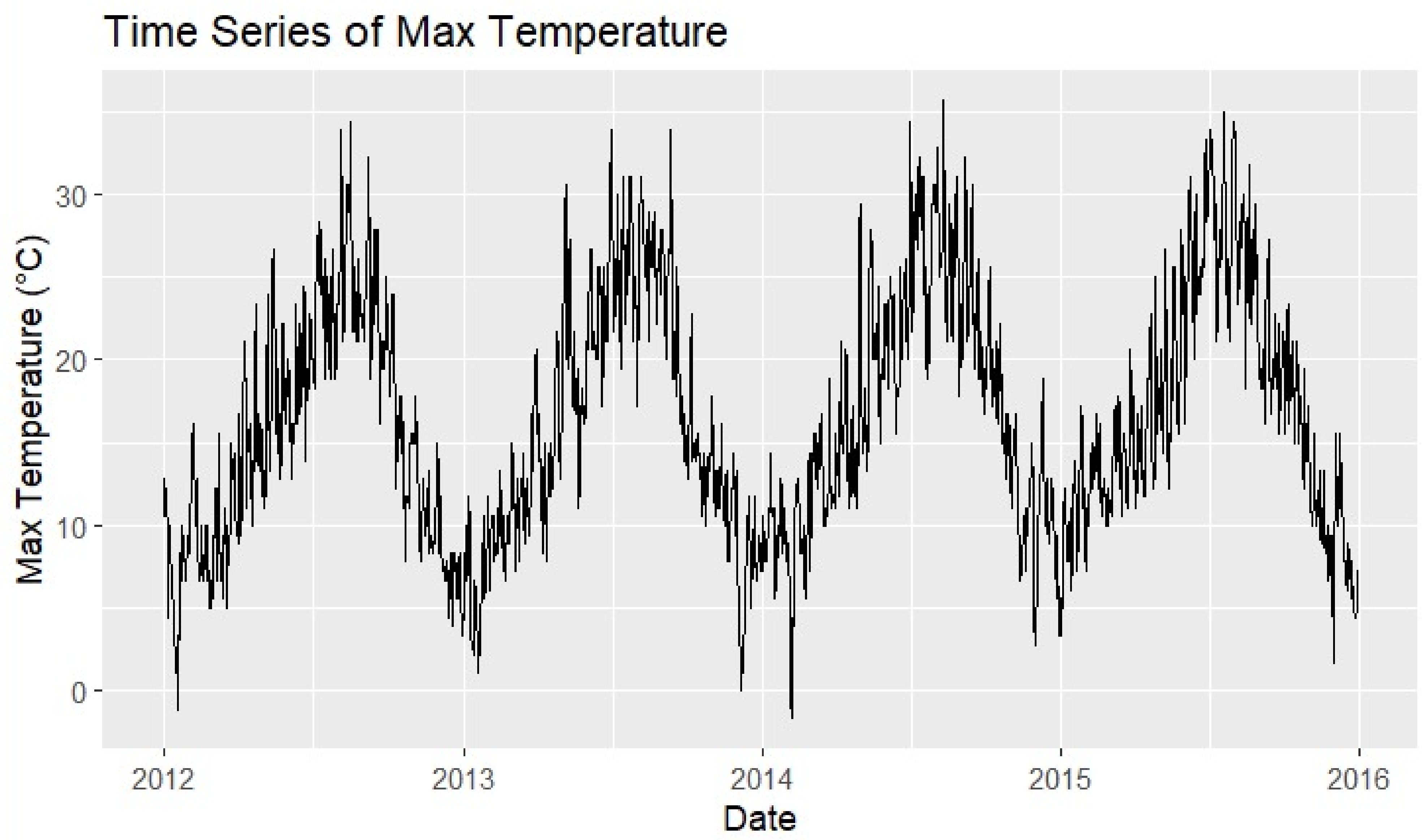
The time series plot of the precipitation data shows a clear seasonal pattern, with higher precipitation levels during the summer months and lower precipitation levels during the winter months. There is also some evidence of a trend, with precipitation levels appearing to increase over time.



In the case of the precipitation data, the Dickey-Fuller test statistic is -8.9531 and the p-value is 0.01. This means that we can reject the null hypothesis of non-stationarity at the 1% significance level. This means that there is strong evidence to suggest that the precipitation data is stationary.

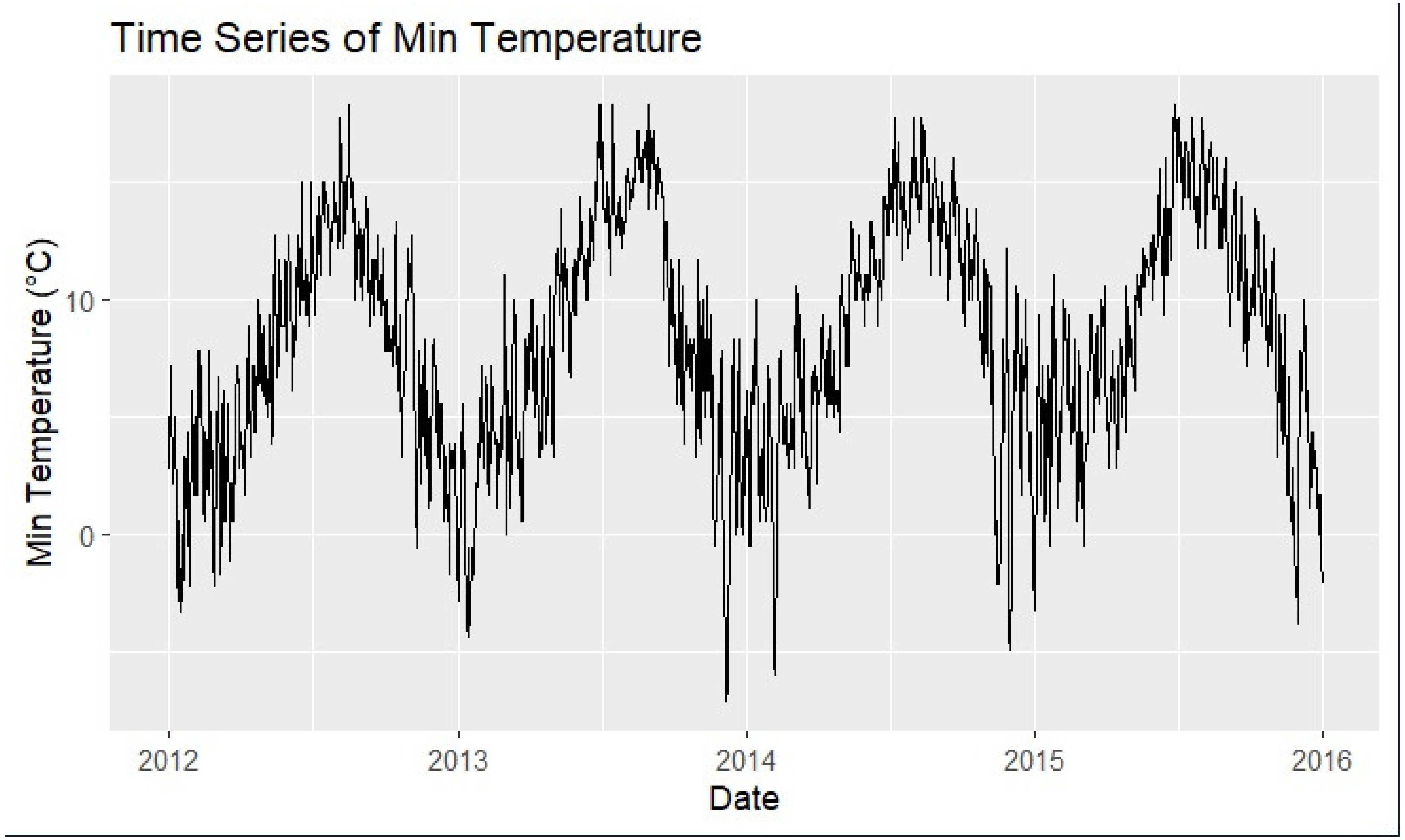
The Dickey-Fuller test is a statistical test that is used to determine whether a time series is stationary. A stationary time series is one in which the mean, variance, and autocorrelation structure do not change over time

The Dickey-Fuller test statistic for the temperature data is -2.4896 and the p-value is 0.3711. This means that we cannot reject the null hypothesis of non-stationarity at the 5% significance level. In other words, there is not enough evidence to suggest that the temperature data is stationary.



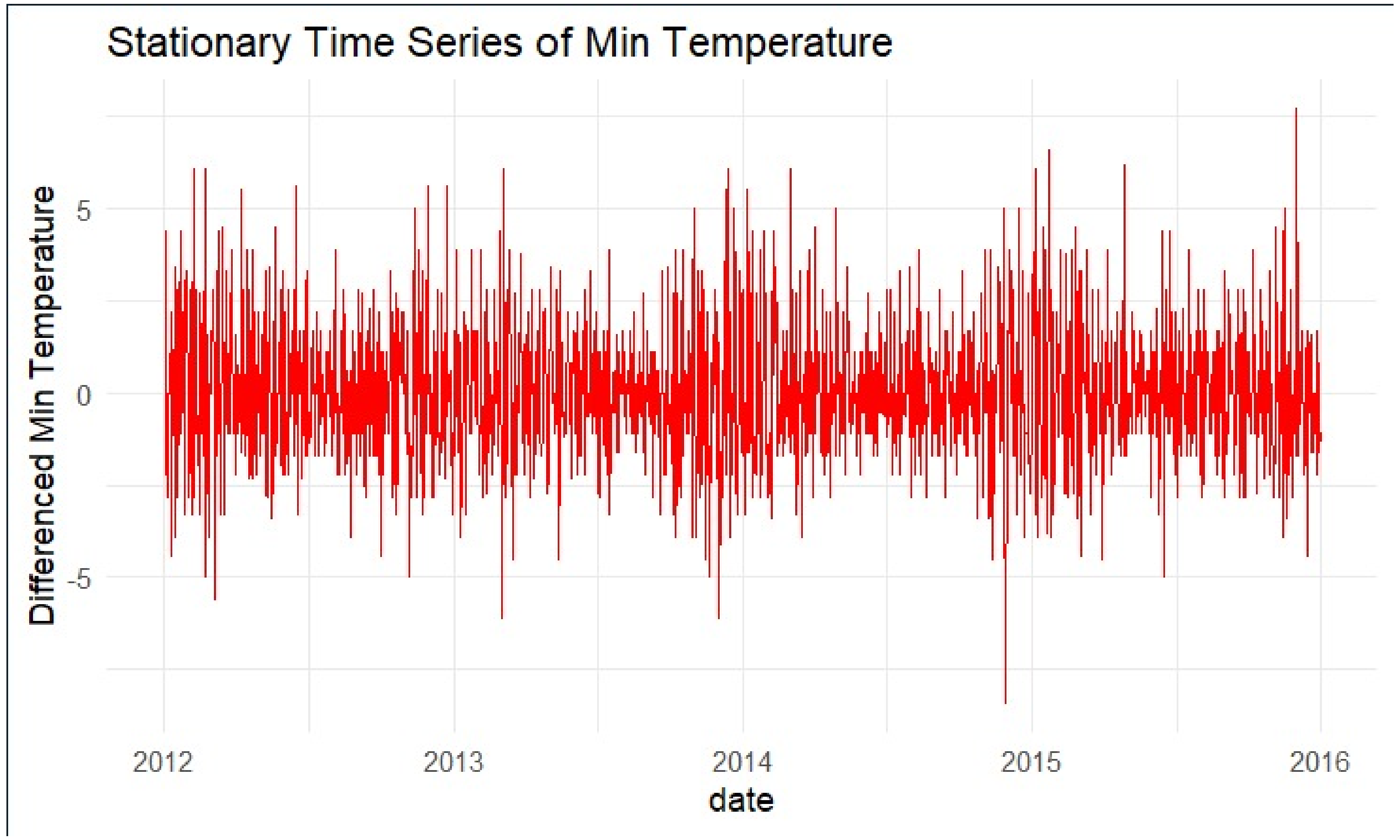
The time series plot of the temperature data shows a clear trend, with temperatures increasing over time. This trend is the reason why the temperature data is not stationary. the p-value of the test is greater than the significance level, so we cannot reject the null hypothesis of non-stationarity and conclude that the time series is not stationary.

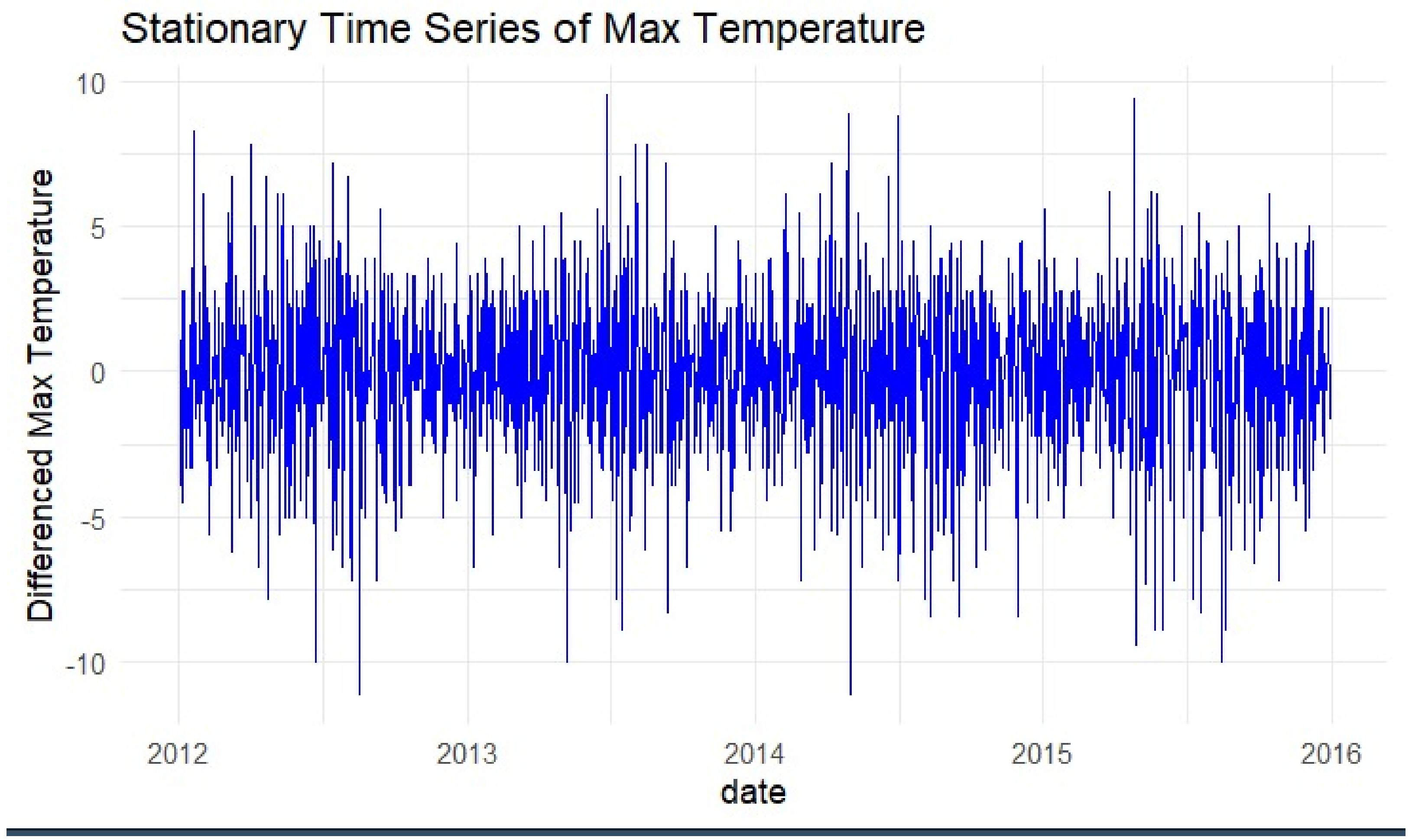
The Dickey-Fuller test statistic for the minimum temperature data is -2.7686 and the p-value is 0.2529. This means that we cannot reject the null hypothesis of non-stationarity at the 5% significance level. In other words, there is not enough evidence to suggest that the minimum temperature data is stationary.



The time series plot of the minimum temperature data shows a clear trend, with temperatures decreasing over time. This trend is the reason why the minimum temperature data is not stationary.

2) Making Min. and Max. temperature series stationary through variable transformations





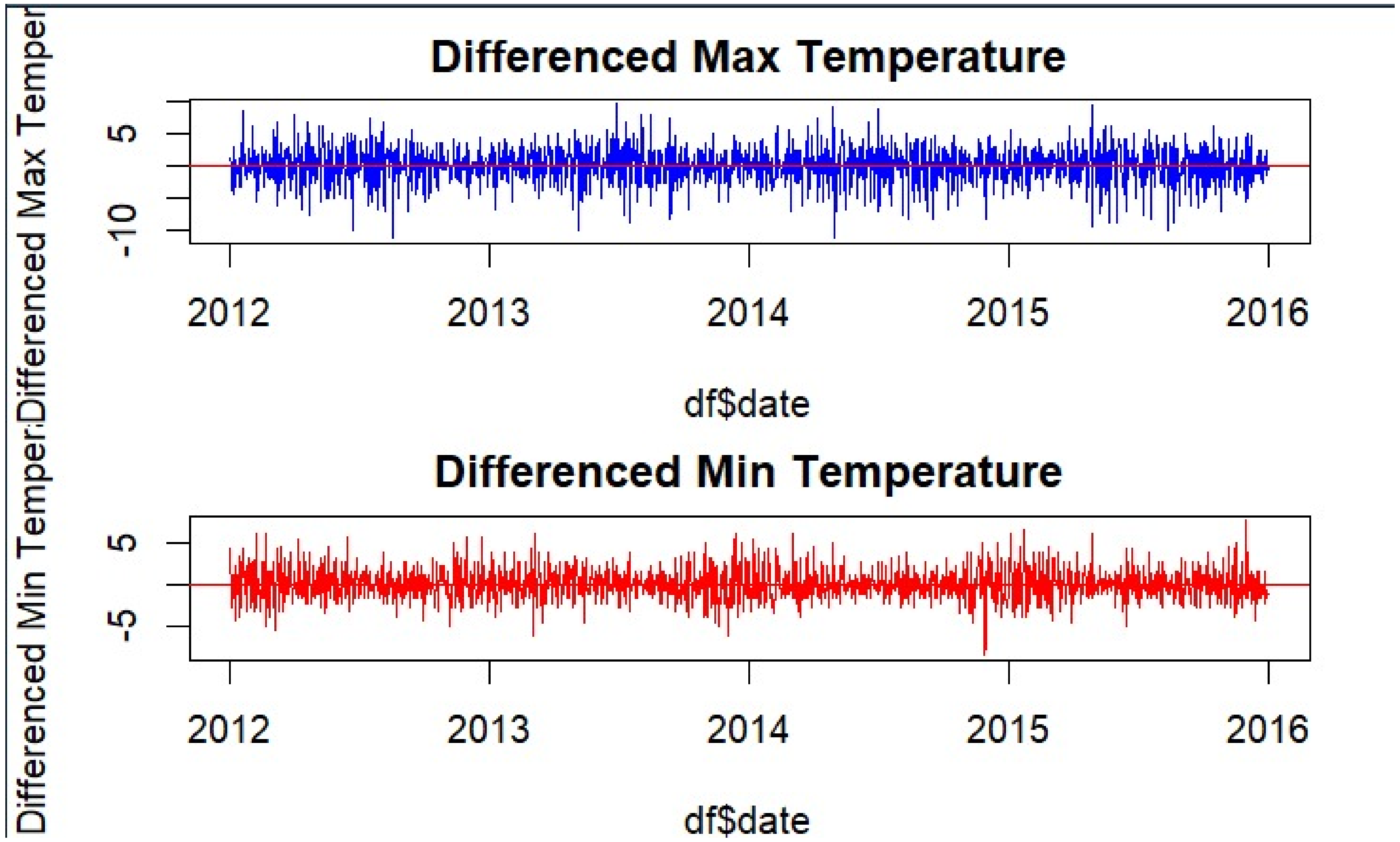
In time series analysis, variable transformations are techniques used to modify the characteristics of a time series data to make it more suitable for analysis and forecasting. Here, we are transforming the 2 non-stationary variables Max. and Min. Temperatures. Stationarity is important in time series analysis because it plays a crucial role in ensuring the effectiveness, interpretability, and accuracy of the models.

*****Variable transformation is a technique used to modify the characteristics of a variable to make it more suitable for analysis and modeling. It reduces skewness, improves linearity and removes outliers in the data*****

3) Checking if stationarity is achieved through tests and plots.

Yes, stationarity is achieved through tests and plots.

<u>Tests:</u>
The ADF test results for both df$temp_min_diff and df$temp_max_diff are significantly negative and have p-values less than 0.05. This suggests that we can reject the null hypothesis of non-stationarity and conclude that both of the differenced time series are stationary.



<u>Conclusion:</u>
Both the tests and the plots suggest that stationarity has been achieved through differencing. This is a positive outcome for time series analysis and modeling, as it means that we can proceed with building models to forecast future temperature values.

4) Fit an ARIMA model for PRECIPITATION data

Here is the output for the ARIMA model that was fit, which has later been used to forecast the next year's precipitation.

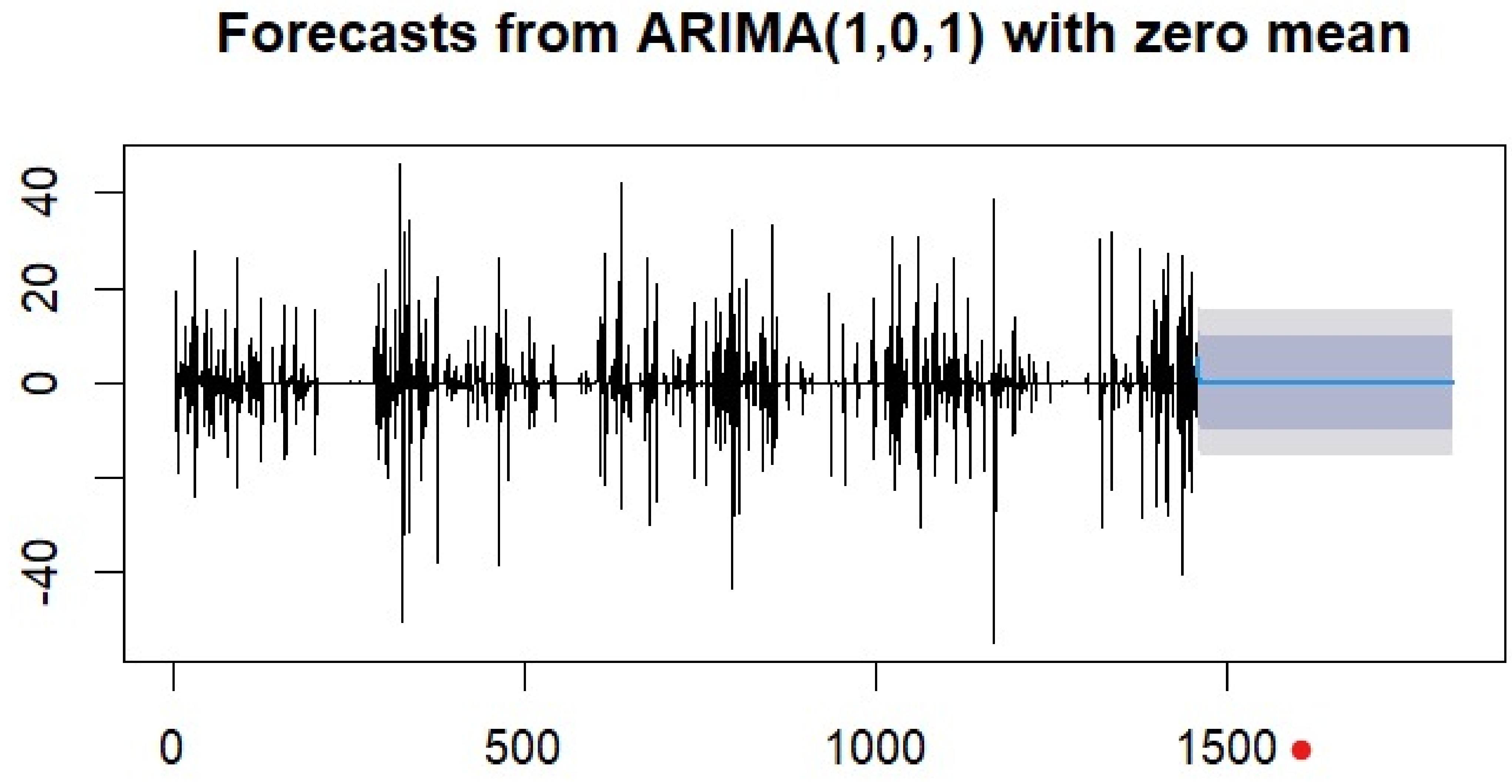
```
Series: df$precipitation_diff
ARIMA(1,0,1) with zero mean

Coefficients:
         ar1      ma1
      0.2450  -0.9634
s.e.  0.0271   0.0079

sigma^2 = 39.87:  log likelihood = -4758.9
AIC=9523.8   AICc=9523.81   BIC=9539.65

Training set error measures:
                     ME     RMSE      MAE MPE MAPE      MASE        ACF1
Training set 0.01481253 6.309623 3.598753 NaN  Inf 0.5394927 -0.01547171
```

5) Predicting the temperature/precipitation for the next 1 year using previous ARIMA model



The forecast graph of precipitation for the next 1 year with regards to the above ARIMA model shows that the precipitation is expected to remain relatively stable, with some minor fluctuations. The forecast shows that the precipitation is expected to be between 500 and 1500 millimeters for the next 12 months.

Thank You