Time series Project

Temperature Prediction of Austin

Akshay Kumar Varanasi

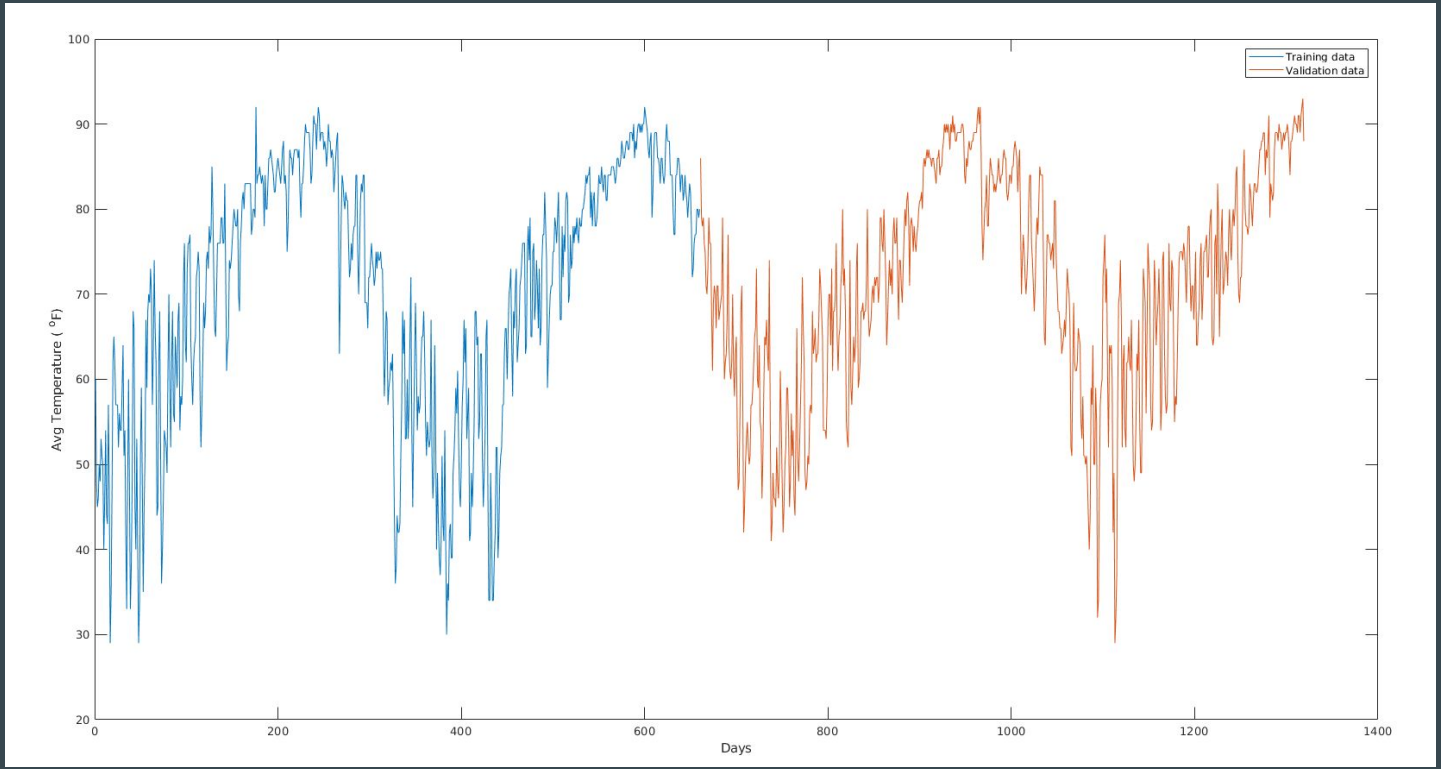
Outline

- About the data
- Stationary Model
- Stationary Model Forecast (1-step rolling)
- Non-Stationary Model
- Non-Stationary Model Forecast (1-step rolling)
- ARMAV
- ARMAV (1-step rolling)
- Conclusion

Overview of the data

- Accurate weather prediction is important for planning our day to day activities.
- In our case, we are predicting Temperature (whether it will be cool or hot tomorrow? Do we need to wear sweater or T-shirt?)
- Austin Weather Dataset from Kaggle was used which was obtained from WeatherUnderground.com, at the Austin KATT station. https://www.kaggle.com/grubenm/austin-weather
- This dataset contains data for every date from 2013-12-21 to 2017-07-31. (1319 data points)
- For training we used 50% data, which is around 660 data points and validated on the remaining data.

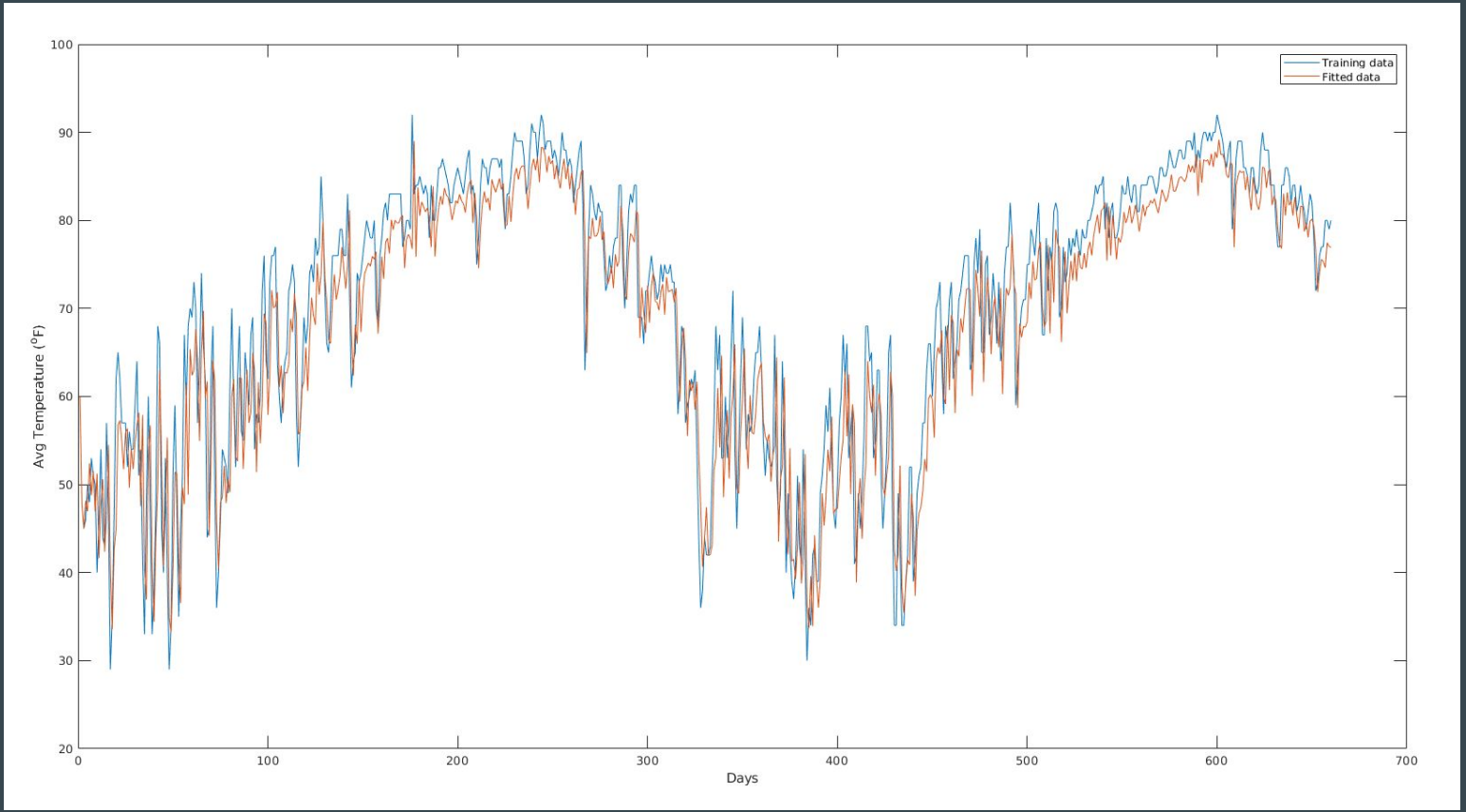
Temperature Data of Austin



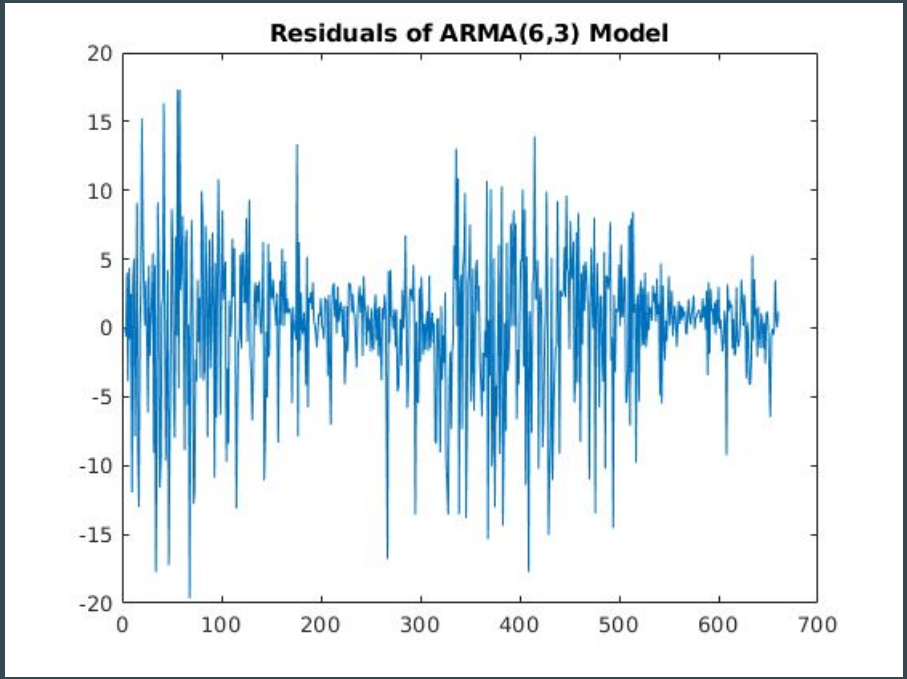
Stationary Model

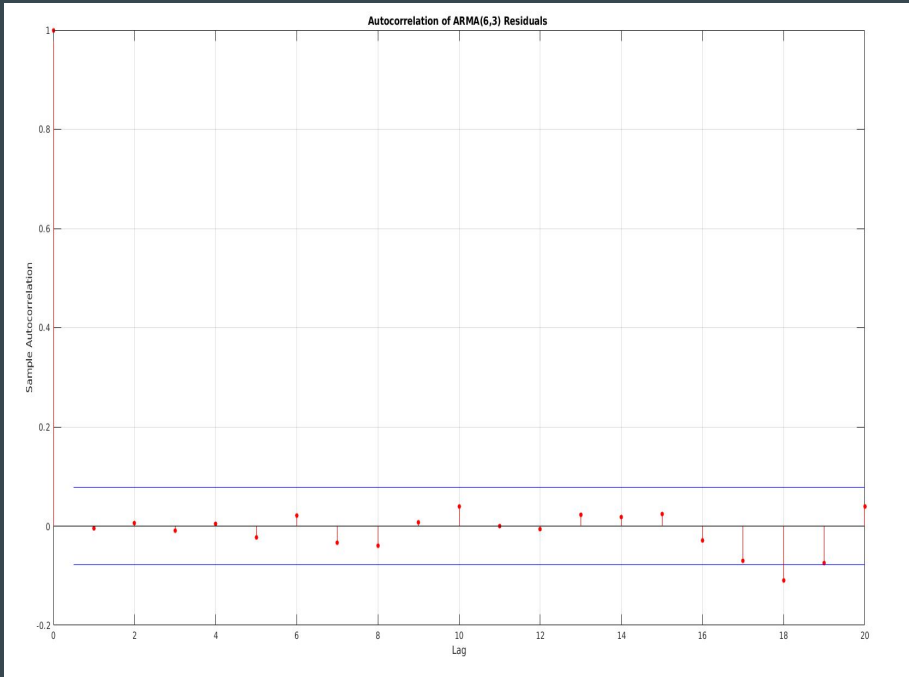
- ARMA(2n,2n-1) Technique, F-test was used
- ARMA(6,3) model was found adequate with RSS was 1.7312e+04

Stationary Model



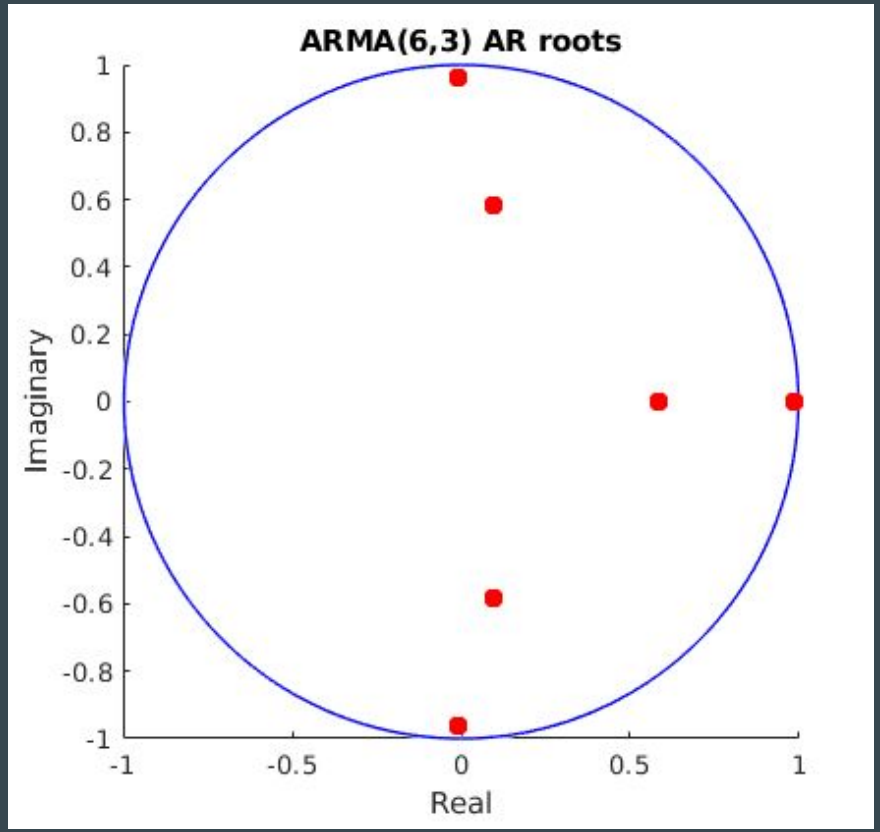
Residual plot





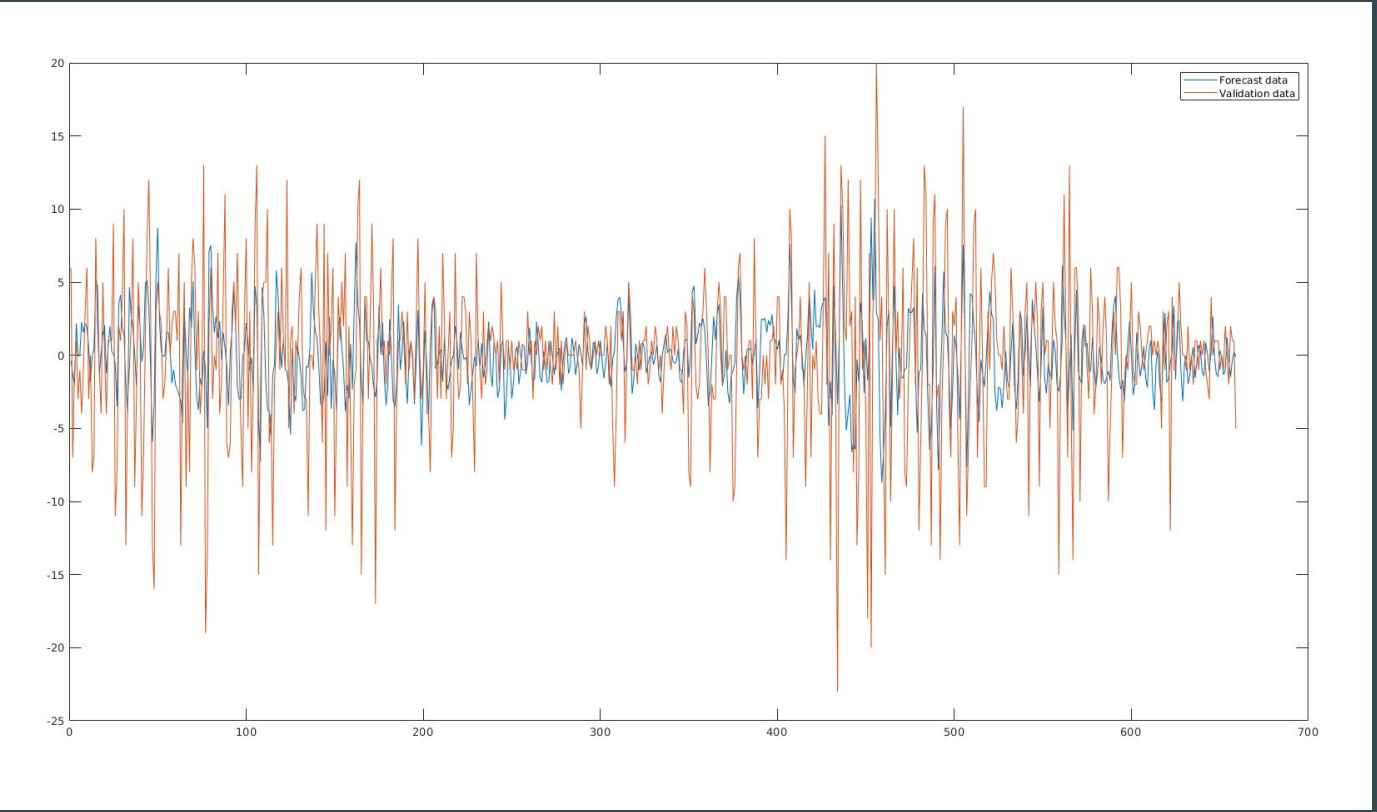
Auto-regressive roots

- Out of 6 roots, 2 roots are real and 4 are complex.
- 1 root is 0.9903 ≅ 1 so stochastic trend was checked it was found that it exits. RSS=1.7374e+04
- Complex roots give period 3.9748 ≅ 4 and 4.4614 ≅ 4.5. Seasonality was checked and it was found that it does not exists.



Forecast

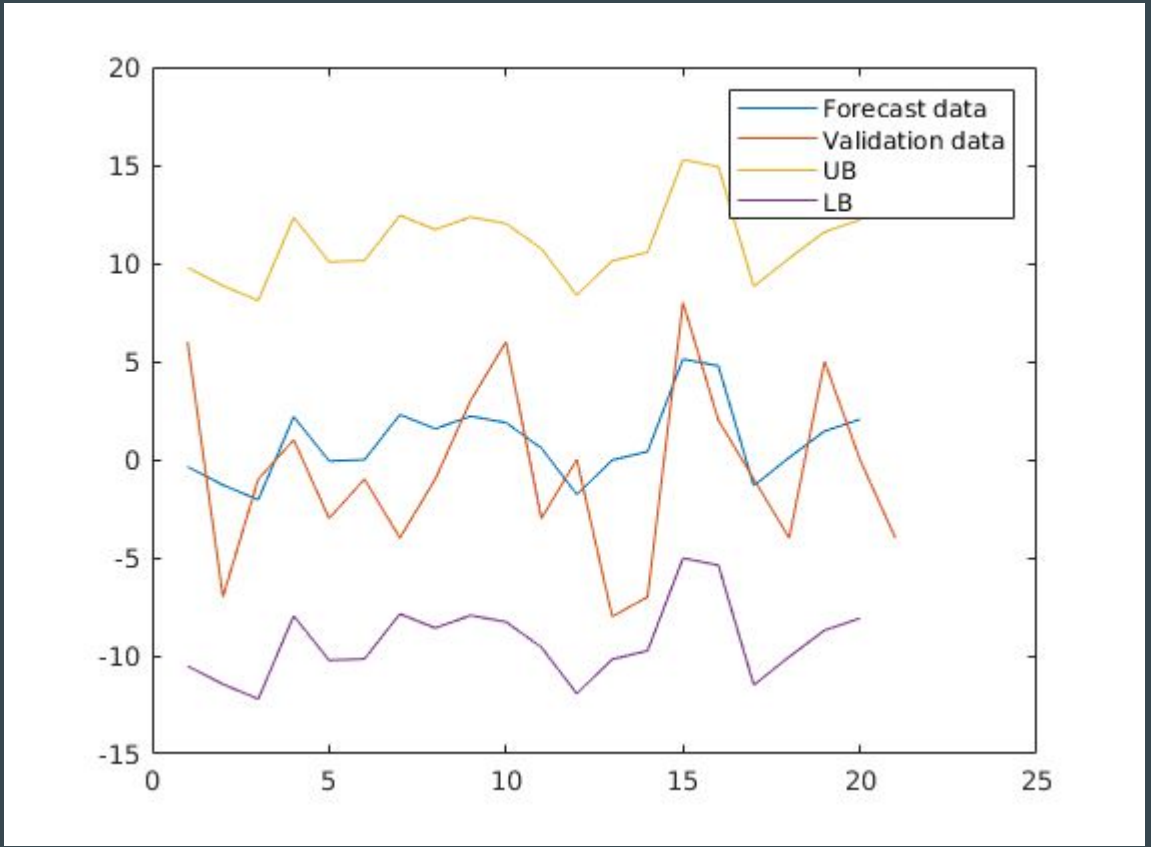
- The model used was that of parsimonious one i.e ARMA(5,3).



Forecast

MSE: 26.83 STD: 5.1798

RSS: 1.7235e+04



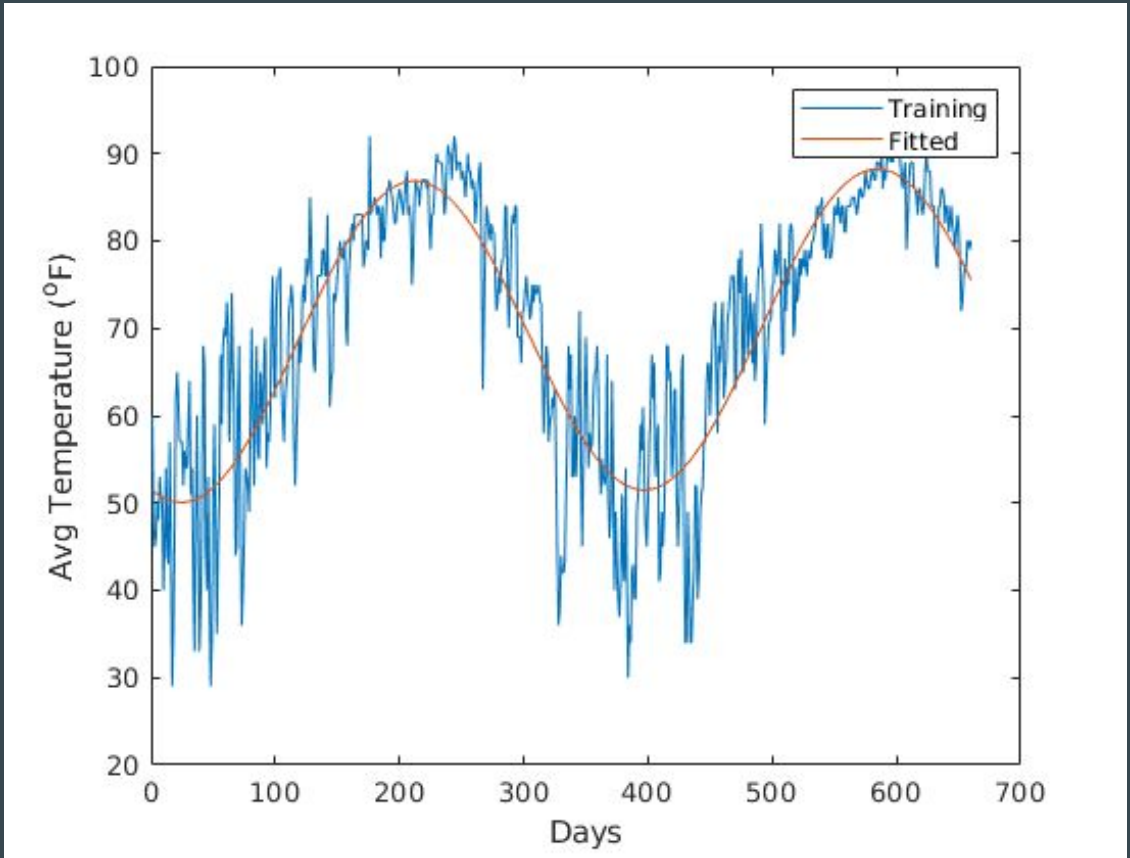
Non-Stationary Model

- Curve a1*sin(b1*x+c1) + a2*sin(b2*x+c2) was fitted and RMSE = 7.2470

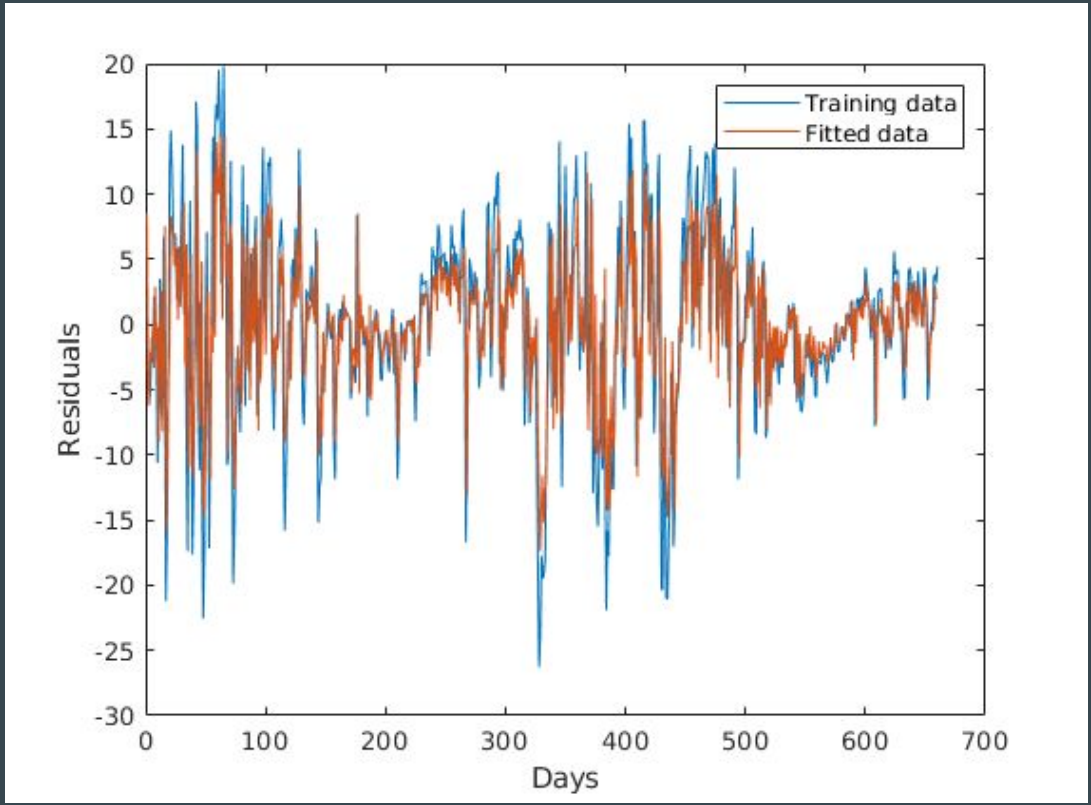
```
Coefficients (with 95% confidence bounds):
  a1 =       77.59  (-324, 479.2)
  b1 =    0.000105  (-0.002595, 0.002805)
  c1 =       1.069  (-8.322, 10.46)
  a2 =       18.04  (17.22, 18.85)
  b2 =     0.01686  (0.01656, 0.01717)
  c2 =      -1.992  (-2.111, -1.874)
```

- ARMA(2n,2n-1) Technique, F-test was used on the detrended data.
- AR(3) model is adequate and RSS is 1.6763e+04

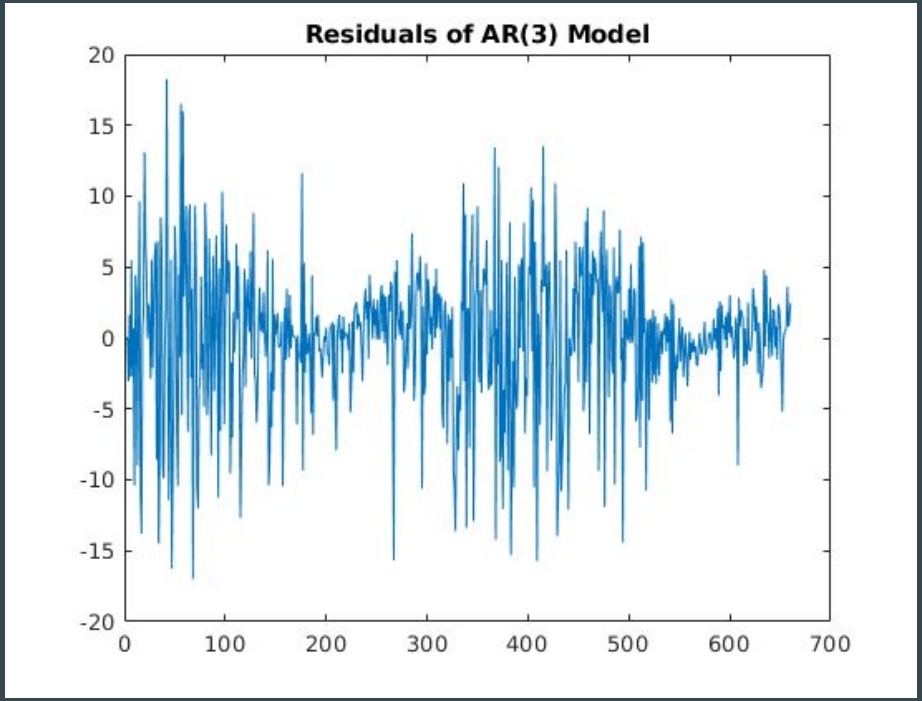
Deterministic trend

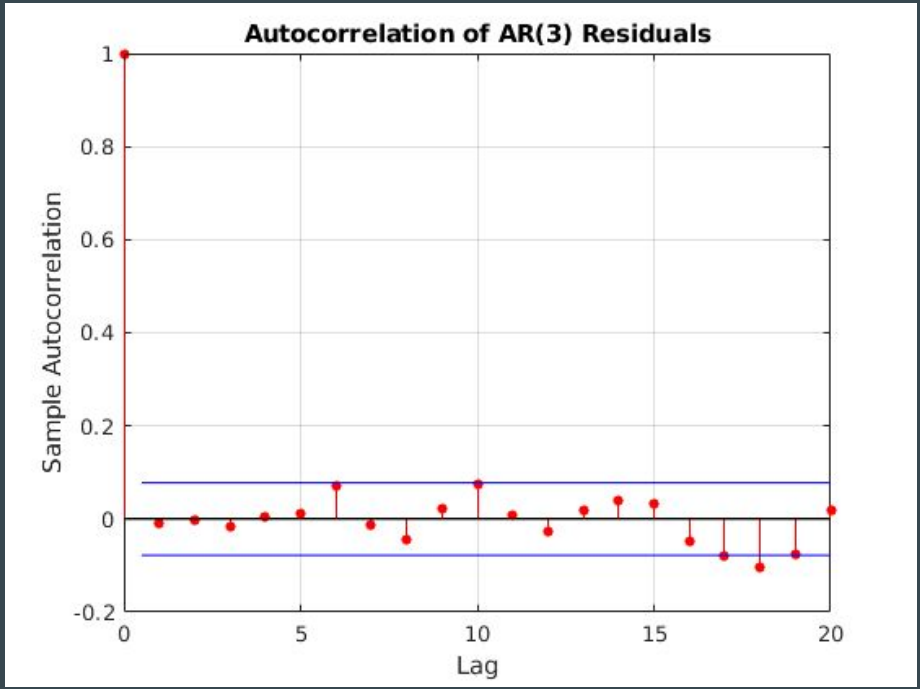


Detrended data



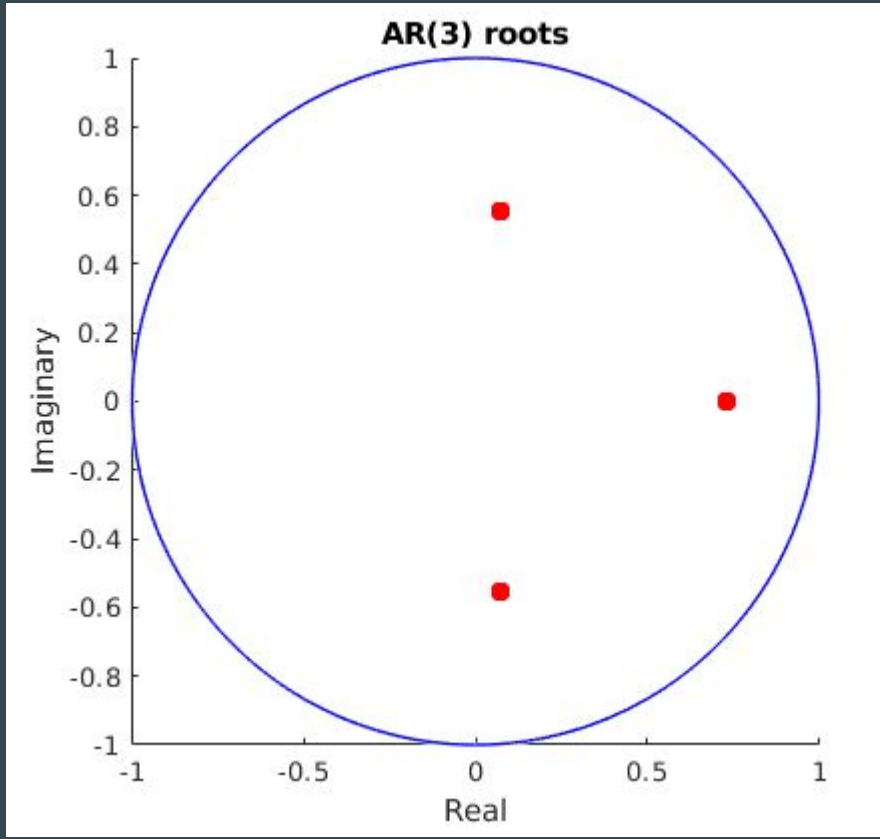
Residual Correlation





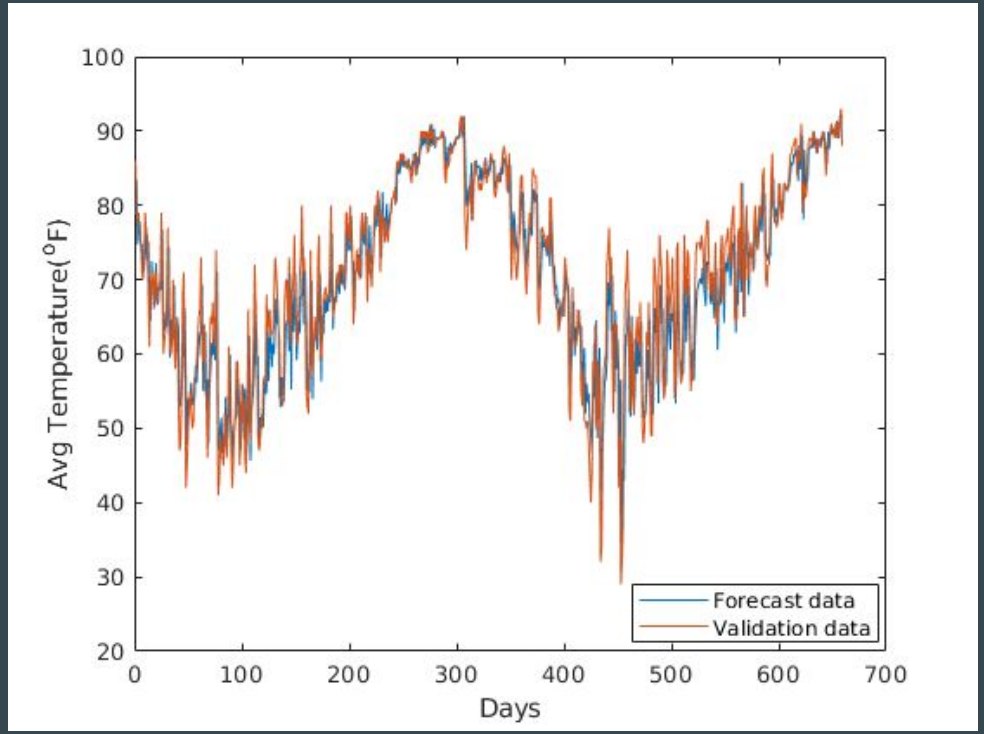
Auto-regressive roots

- Out of 3 roots, 1 roots are real and 2 are complex.
- 1 root is 0.7345 so stochastic trend was not checked
- Complex roots give period 4.3701 ≅ 4 or 4.5. Seasonality was checked and it was found that it does not exists.



Forecast

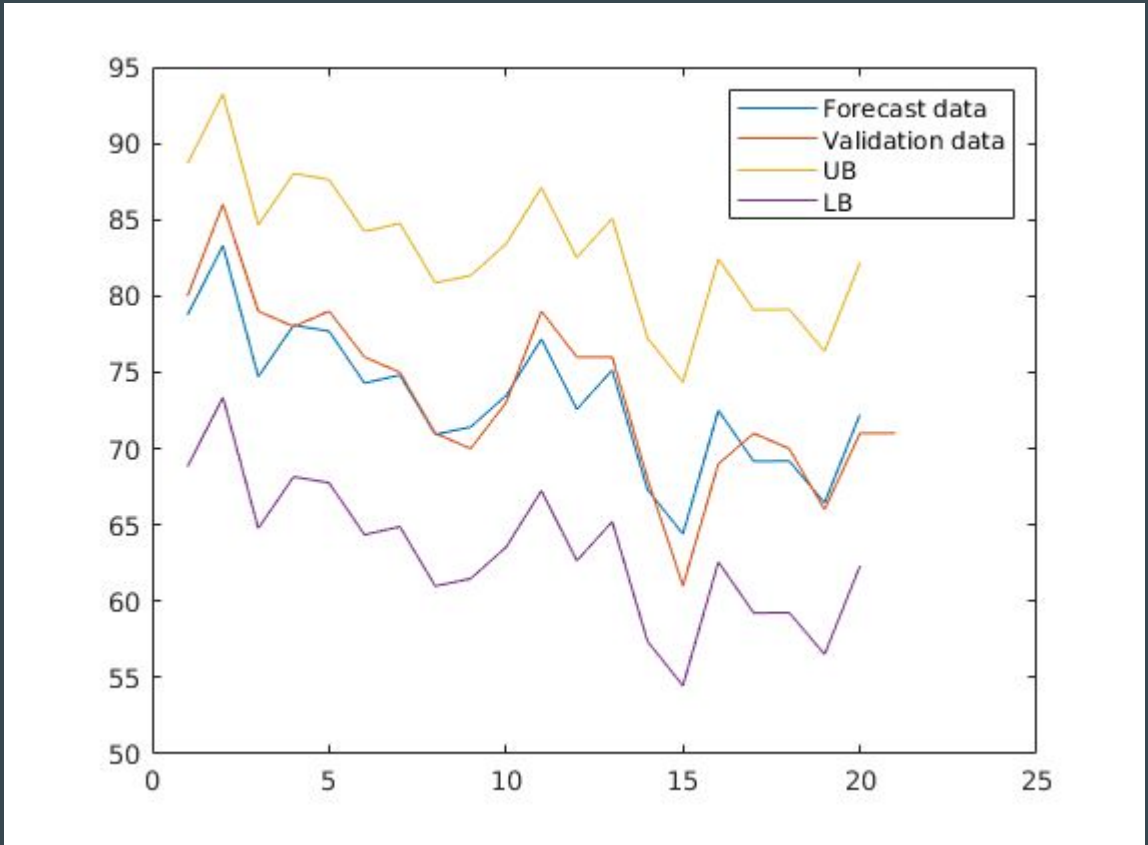
- AR(3) model along with the deterministic trend a1*sin(b1*x+c1) + a2*sin(b2*x+c2) was used.



Forecast

MSE: 25.69 STD: 5.0685

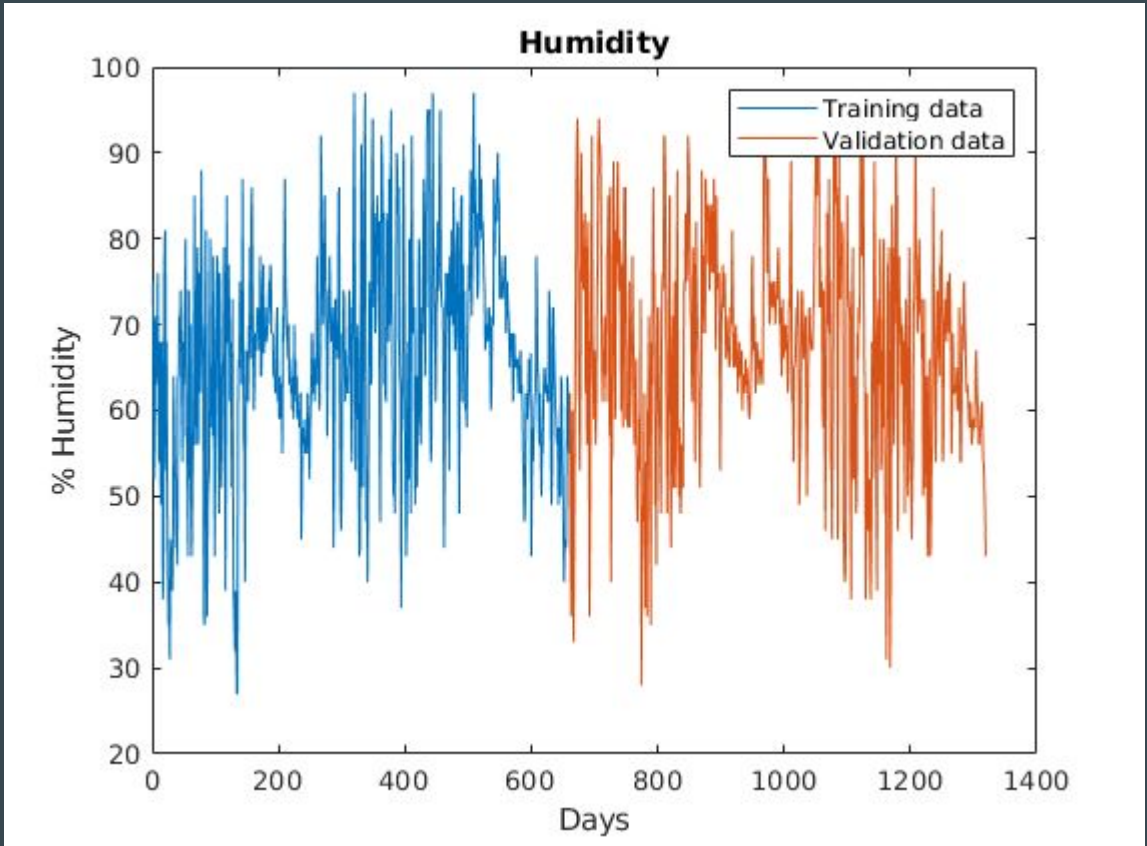
RSS: 1.6550e+04



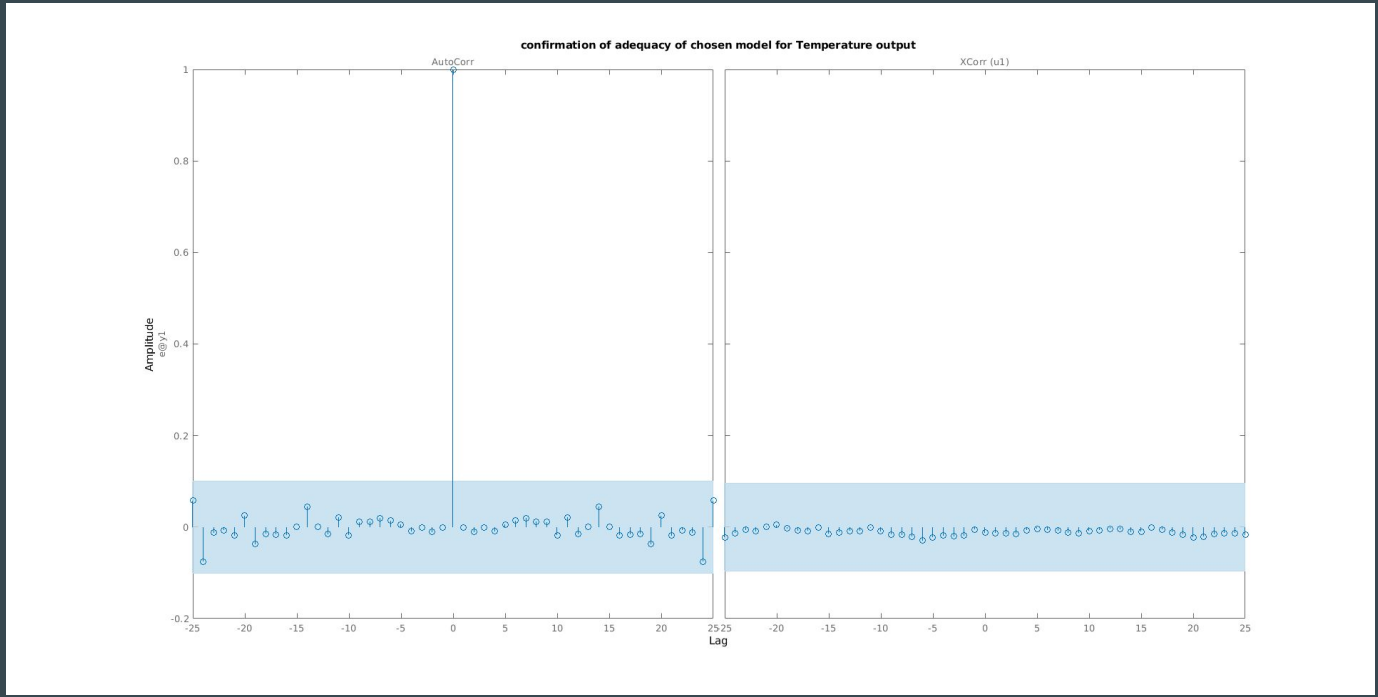
ARMAV

- Percentage humidity along with temperature data was used.
- Model was selected based on (n,n,n-1) modelling strategy with AIC criterion
- For temperature as the output, we get (25,25,24) as the optimum model with RSS = 1.1785e+04 and for humidity as the output, we get (23,23,22) as the optimum model with RSS = 3.9469e+04

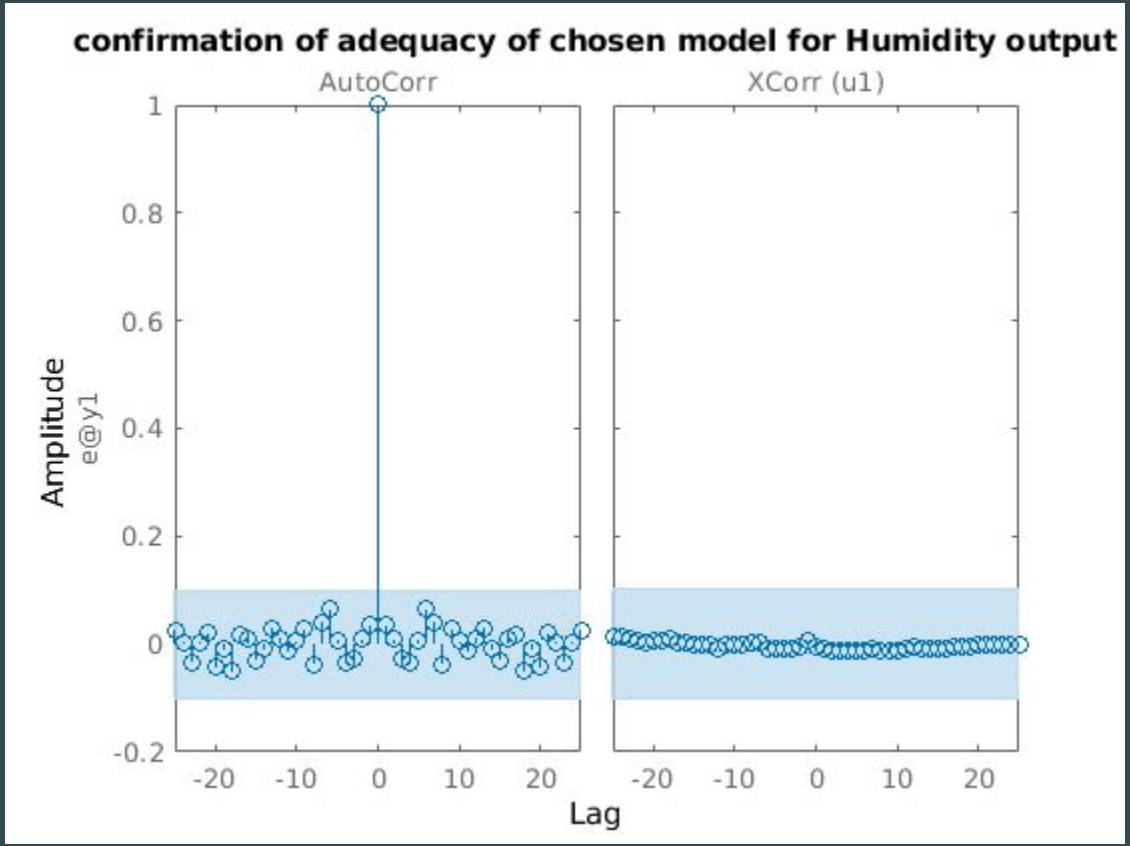
Humidity Data of Austin



Residuals


confirmation of adequacy of chosen model for Temperature output
AutoCorr
XCorr (u1)
Amplitude
e@y1
Lag
1
0.8
0.6
0.4
0.2
0
-0.2
-25
-20
-15
-10
-5
0
5
10
15
20
25

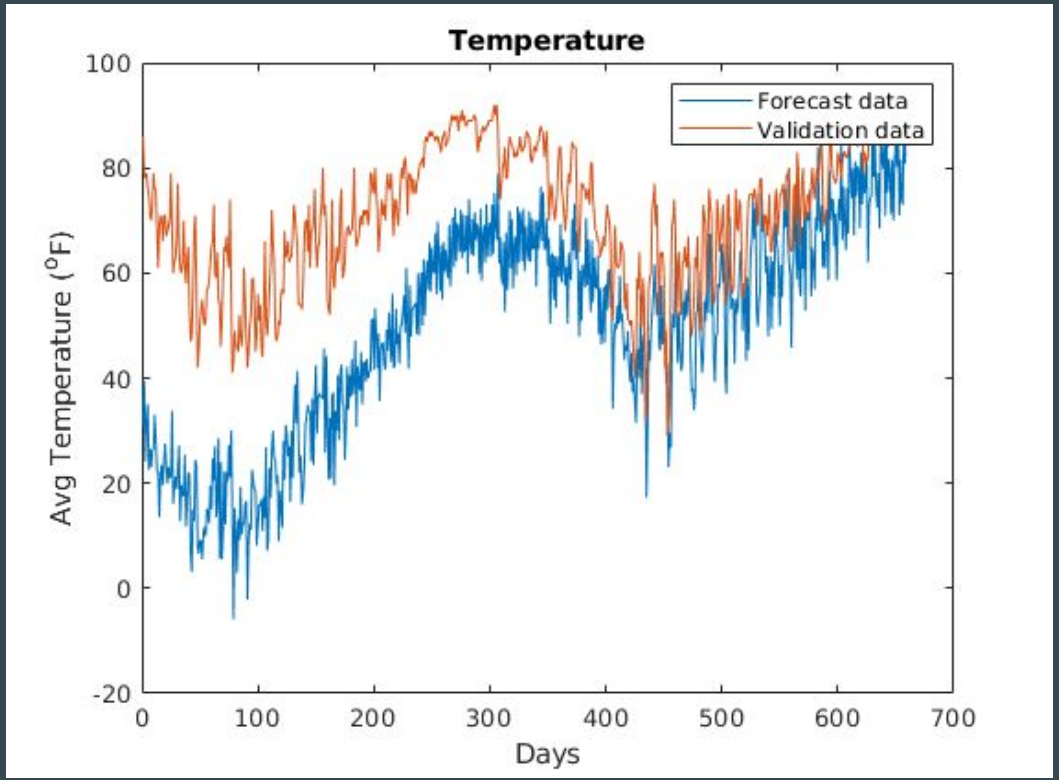
Residuals


confirmation of adequacy of chosen model for Humidity output
AutoCorr
XCorr (u1)
Amplitude
e@y1
1
0.8
0.6
0.4
0.2
0
-0.2
-20
-10
0
10
20
Lag

Temperature Forecast

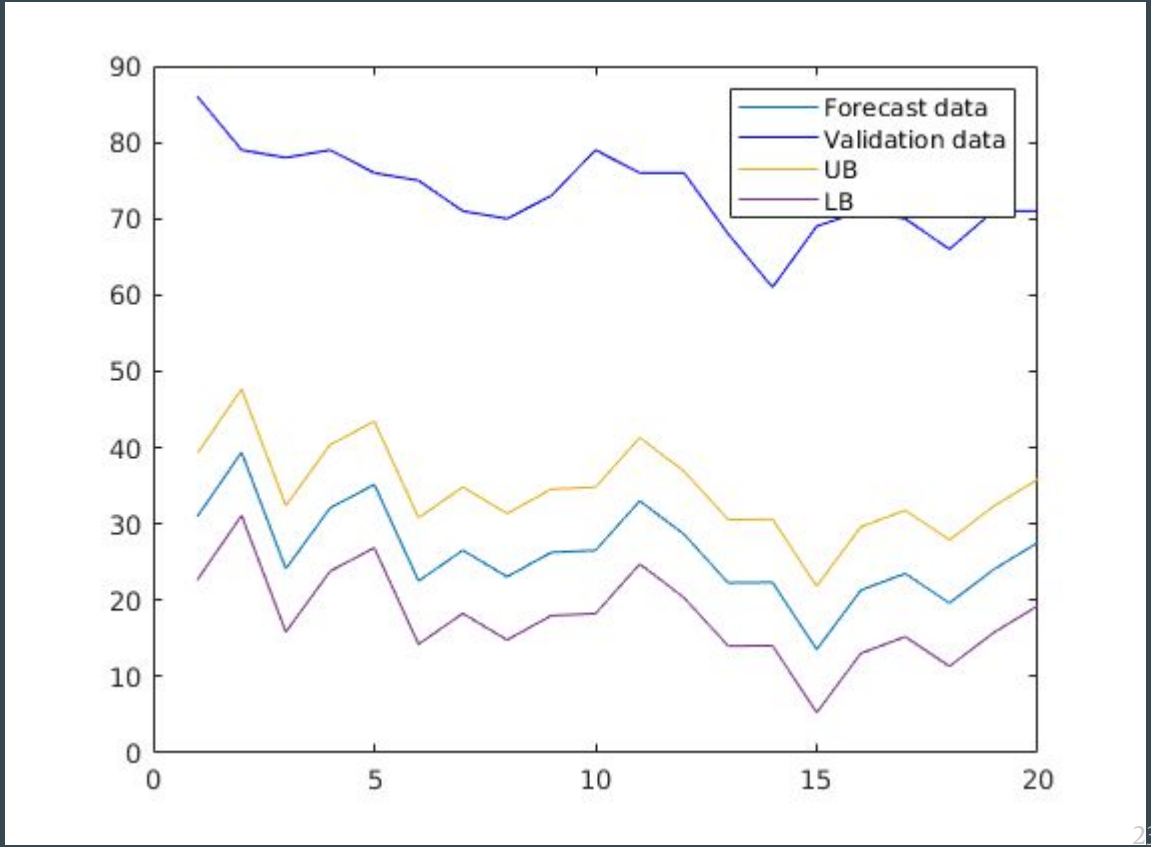
MSE: 17.86 , STD = 4.2261

RSS = 4.3771e+05



Temperature Forecast

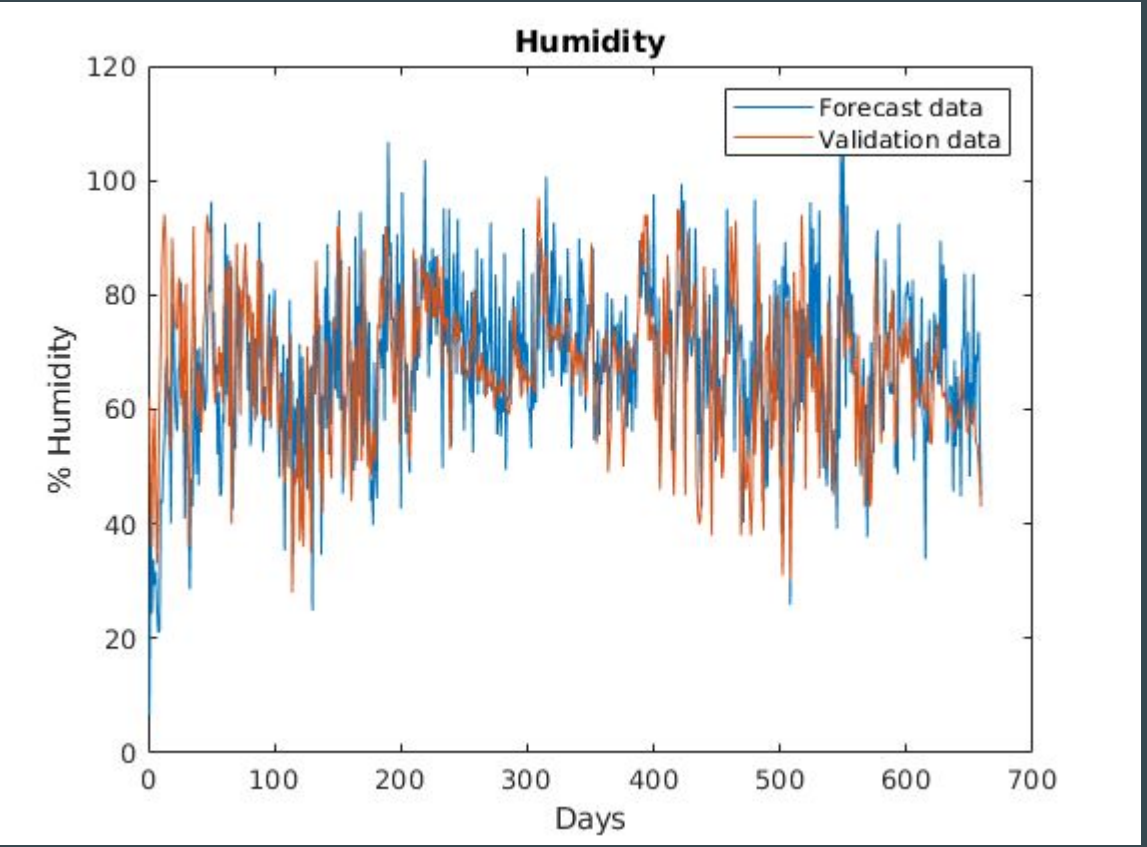
MSE: 17.86 , STD = 4.2261



Humidity Forecast

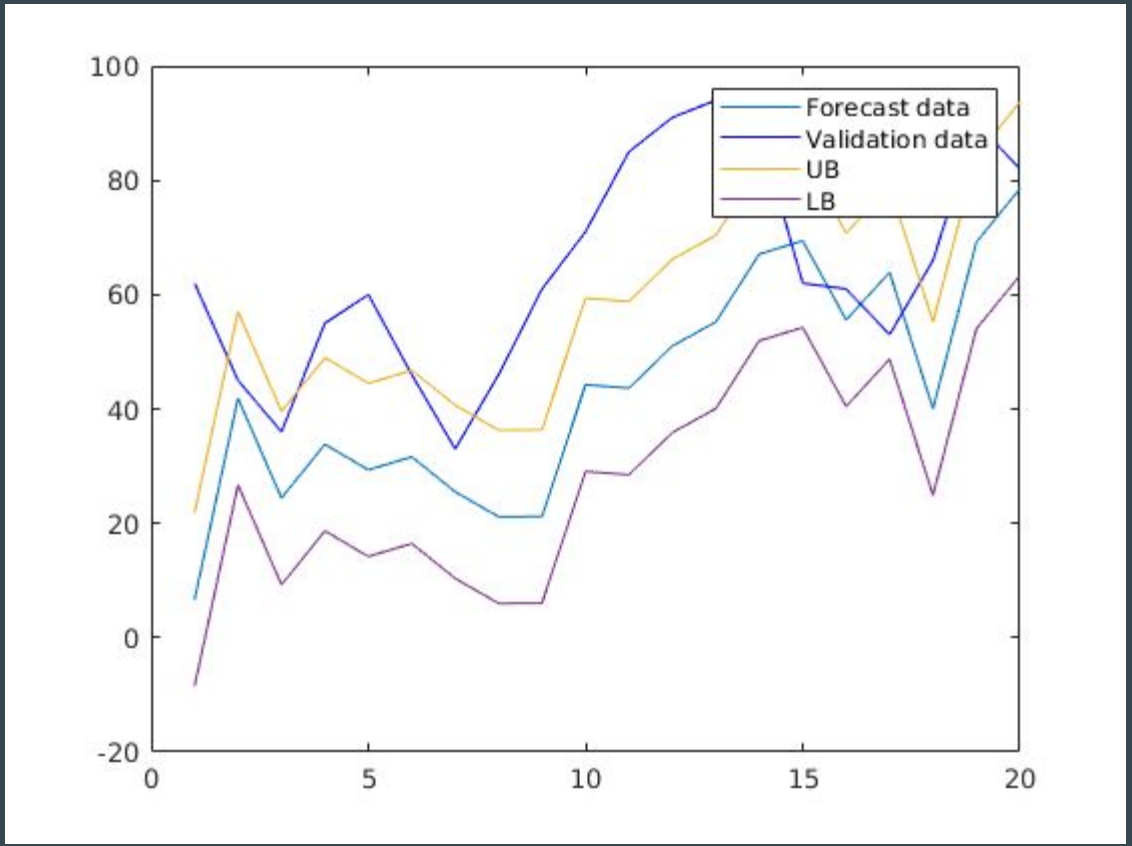
MSE: 59.8 , STD = 7.7330

RSS = 1.3996e+05



Humidity Forecast

MSE: 59.8 , STD = 7.7330



Conclusion

- Among the models, Non-stationary model seems to be the best model as it has the lowest forecast RSS of 1.6550e+04.
- ARMAV model under performs compared to other models which needs to be analysed.
- Should try n-step prediction like weekly prediction and compare with theory.