#### **Forecasting UK Covid-19 cases**

**INTRODUCTION**

This report contains a time series forecasting analysis of daily UK Covid-19 cases based on an ARIMA model. The data used is time series data from 1 January 2020 to 14 June 2020. The aim is to determine data stationarity, identify the process parameters in an ARIMA model, and forecast the next seven days from 15 June 2020 to 21 June 2020. In accordance with standard practice, all analyses were undertaken using IBM SPSS software.

**ASSESSING STATIONARITY**

In order to determine if the time series is stationary, it is appropriate to look at the ACF and PACF plots of the original series first.

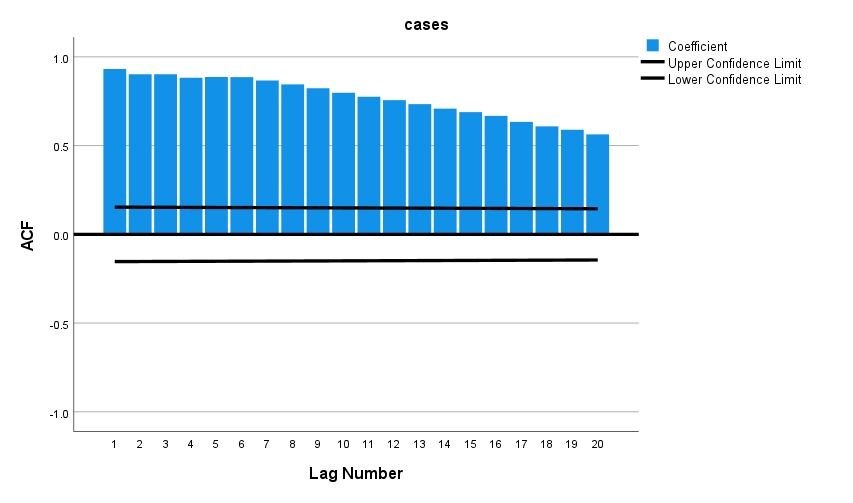


Figure 1 is the ACF plot and displays a slow exponential decay across 20 lags. The positive autocorrelation is strong, and the ACF indicates the original series is non-stationary.

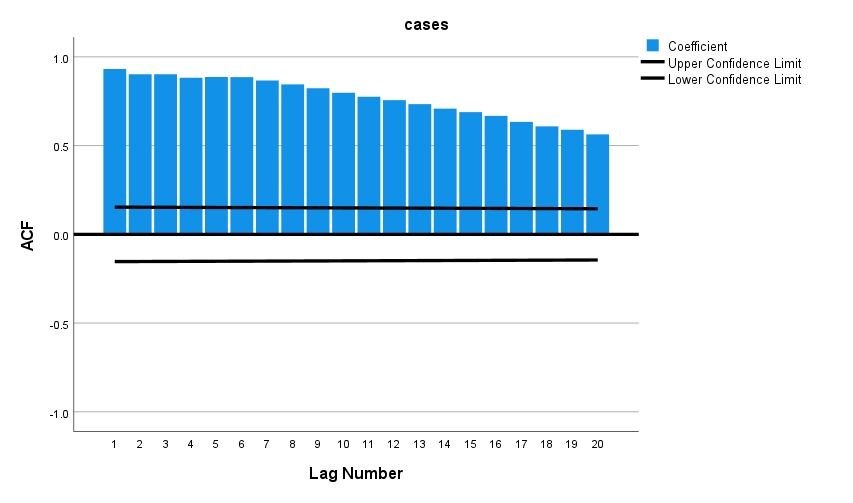


Figure 2 is the PACF plot and displays a large spike at lag 1 but then tapers off. The previous figure indicates this series appears to have autoregressive behaviour.

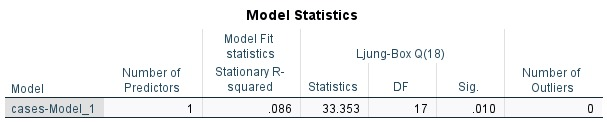
Given these figures the series was first differentiated once with the goal of stabilizing the mean of the series. After differencing the series, it appears to be stationary, as noted by the ACF and PACF, with the ACF cutting off quickly and the PACF indicating spikes that are within the limits of the confidence interval.

**MODEL IDENTIFICATION AND PARAMETER ESTIMATION**

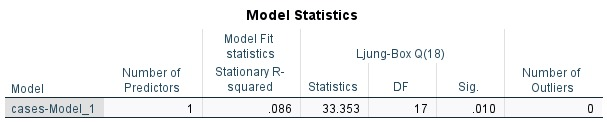
Given the structure of the differenced series and the autocorrelation, the PACF cuts off after lag 1, suggesting only 1 autoregressive term, while the ACF tapers off suggesting non-stationarity. The estimation indicates based on a differenced time rabson, an ARIMA(1,1,0) model.

A ARIMA(1,1,0) model has 1 autoregressive term, 1 difference, and 0 moving average terms. To identify the model the AIC was observed in SPSS and the significance of the parameters were observed for selection.

ARIMA (1,1,0)

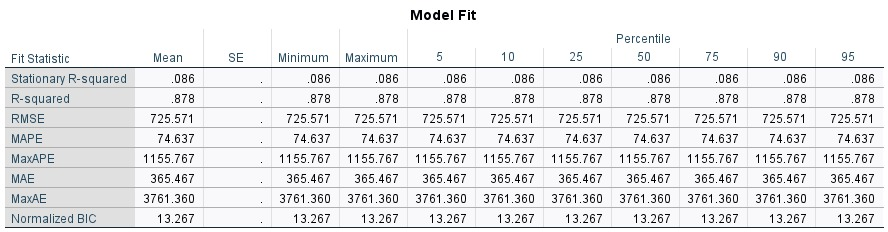


ARIMA ( 1,1,1)



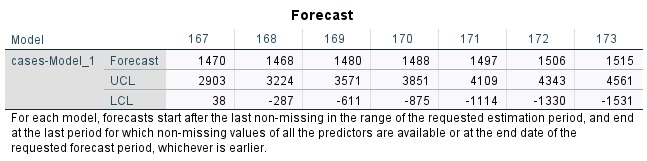
**FORECASTING AND EXPLORING MODEL EVALUATION**

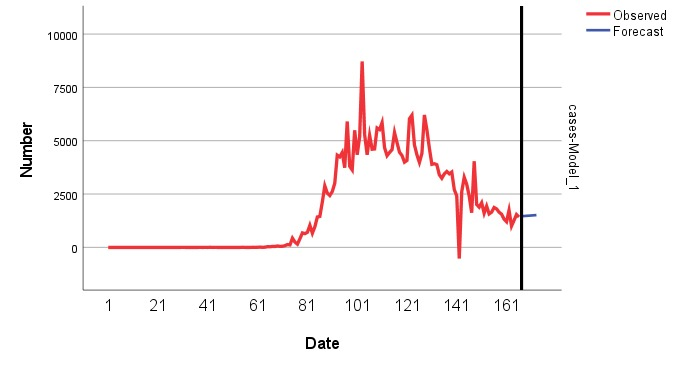
A 7-day forecast was prepared for the period from 15–21 June 2020 using the ARIMA (1,1,0) model. The forecast indicates that a downward trend, which started in very early June, is likely to continue.

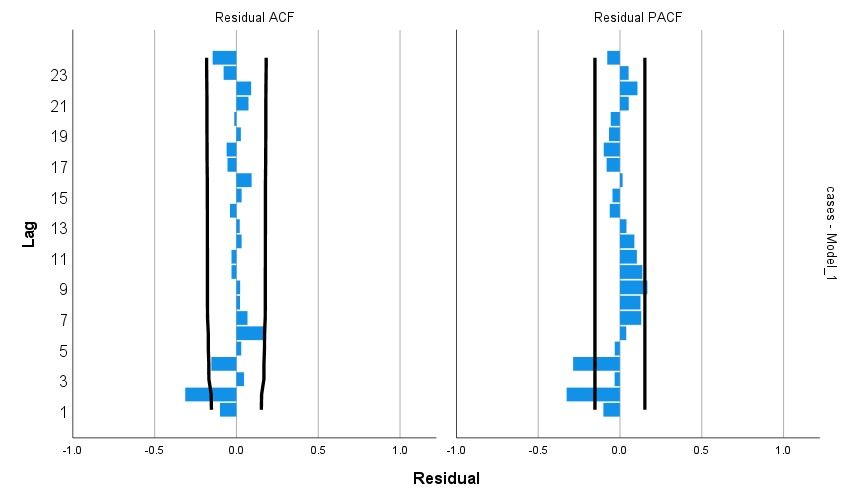
 The model evaluation metrics found were:

• Mean Absolute Error (MAE): 365.467

• Mean Squared Error (MSE): 725.571^2 = 526,452.5







**PREDICTION QUALITY AND PARSIMONY**

When compared with the possible fitness of simpler (ARIMA (0,1,0)) or richer (ARIMA (1,1,1)) models, the ARIMA (1,1,0) model is relatively simpler and less accurate. The previous models were either too poor of a fit or unnecessarily rich with insignificant improvements in prediction accuracy. Therefore, we feel that the ARIMA (1,1,0) model is parsimonious and statistically valid.

**CONCLUSION**

The report demonstrated that the Covid-19 time series was non-stationary and required a first-order differencing to create a stationary series. The chosen ARIMA (1,1,0) model showed statistically valid parameters, and reasonably forecasted periods of 7 days. Therefore, the ARIMA (1,1,0) model is recommended for short-term public health planning and future modelling, because it is simple and effective and provided an honest evaluation of its forecast.