

# Time Series forecasting of the USD/AUD exchange rate

Aiden Liu

## Background

The USD/AUD exchange rate represents the value of one US dollar (USD) in terms of Australian dollars (AUD). Interest rate differentials refer to the difference in interest rates between two countries. These differentials are a major determinant of exchange rate movements through mechanisms like the interest rate parity (IRP). There are some potential instances which affect the exchange rate via interest rates such as:

**Rising Australian Interest Rates Relative to the US:** The AUD tends to appreciate against the USD as investors move capital to take advantage of higher returns in Australian markets.

**Rising US Interest Rates Relative to Australia:** The USD appreciates against the AUD, reflecting increased demand for USD-denominated assets.

**Global Economic Uncertainty:** In times of market turbulence, the USD often strengthens as a safe-haven currency, even if interest rates in the US are lower than in Australia.

## Purpose

I have always had an interest in macroeconomics, particularly in international finance (as per my Economics major at university). The intention behind this project is to explore more about the behaviour of exchange rates and albeit practice time series modelling on real datasets. This document will be analysing the dataset of the monthly USD/AUD exchange rate from January 1971 to November 2024.

Source: <https://fred.stlouisfed.org/series/EXUSAL>

## Importing relevant libraries

```
#Importing necessary libraries  
library(tsibble)  
library(tidyverse)  
library(fable)  
library(feasts)  
library(tseries)
```

## Reading in the data

```
data = read.csv("EXUSAL.csv", header=TRUE)

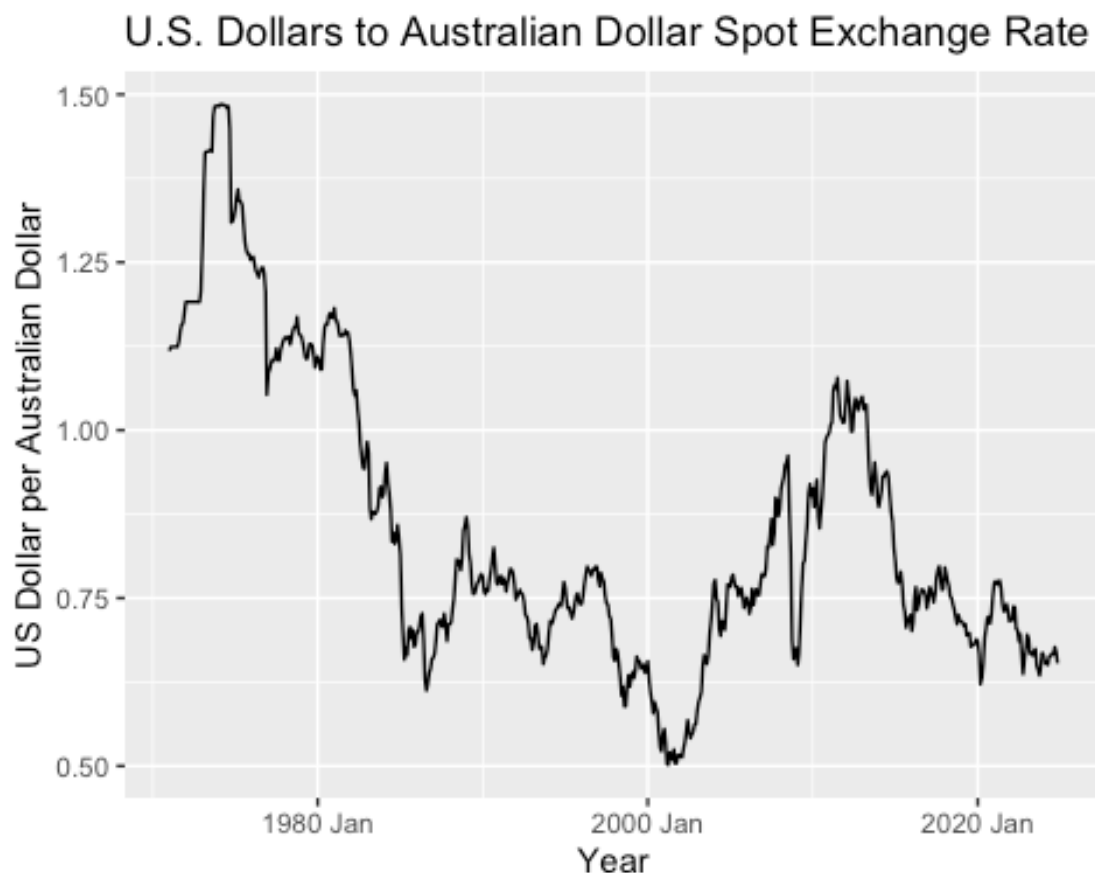
data$observation_date = yearmonth(data$observation_date) #Adjusting the time index to a year-month format

US.AUexchangerate = as_tsibble(data, index = observation_date) #Making a tsibble based on the data
```

## Time plots

Let's take a look at the time plot

```
US.AUexchangerate %>%
  autoplot(EXUSAL) +
  labs(y="US Dollar per Australian Dollar", x = "Year", title = "U.S. Dollars to Australian Dollar Spot Exchange Rate")
```

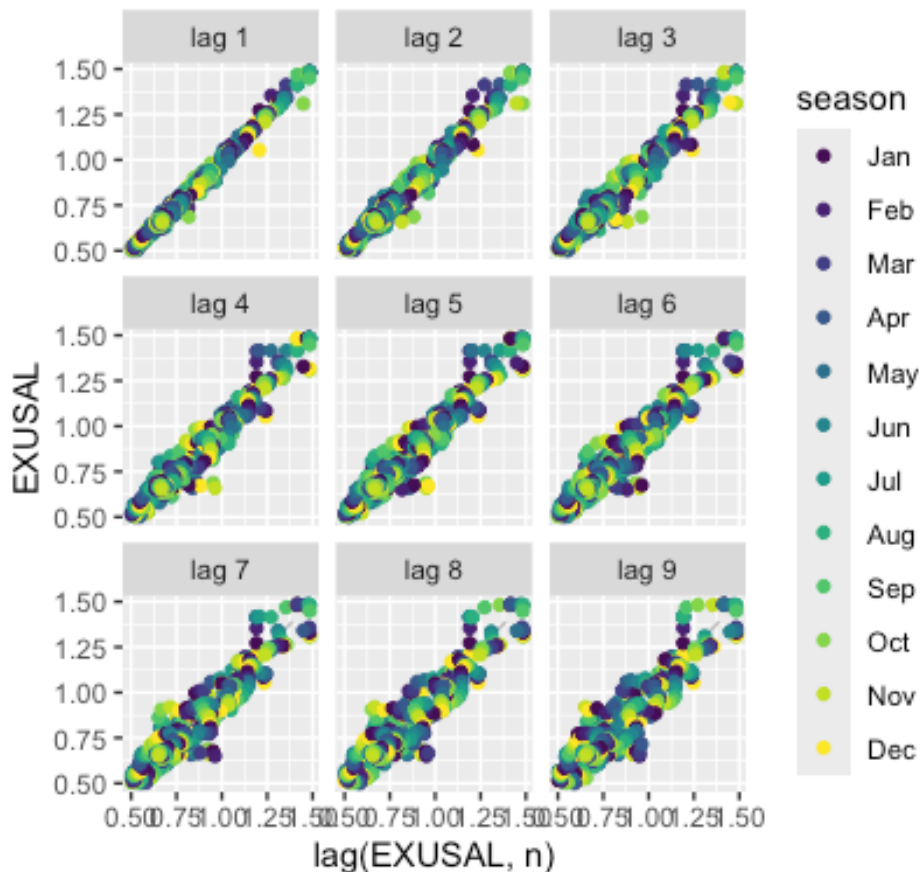


A decreasing trend is apparent. There seems to be some cyclicity in the U.S. Dollar to Australian Dollar spot exchange rate shown in this chart. Over time, you can observe

repeating patterns of rise and fall in the exchange rate, although the periodicity is not consistent.

Let's look at the relationship between the exchange rate and previous versions of the exchange rate at previous months

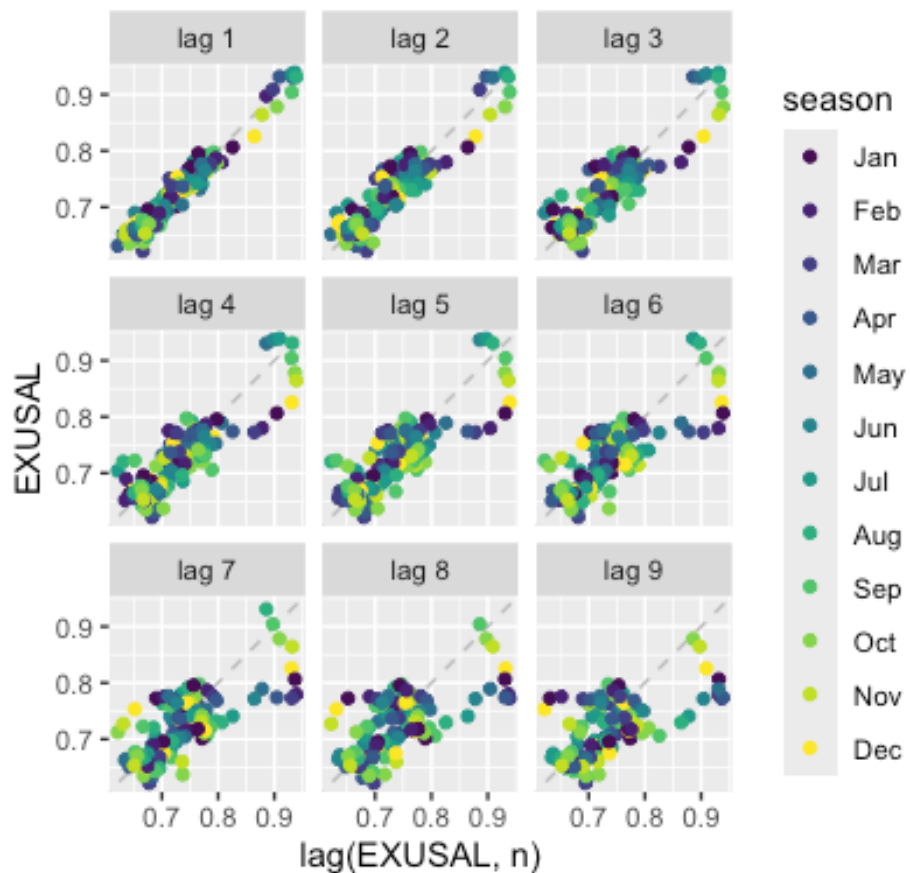
```
US.AUexchangerate %>%  
  filter(year(observation_date)>=1971) %>%  
  gg_lag(EXUSAL, geom = "point")
```



Overall, a strong linear relationship at all lags, meaning the exchange rate should have reverive behaviour.

What about from the last ten years and onwards?

```
US.AUexchangerate %>%  
  filter(year(observation_date)>=2014) %>%  
  gg_lag(EXUSAL, geom = "point")
```

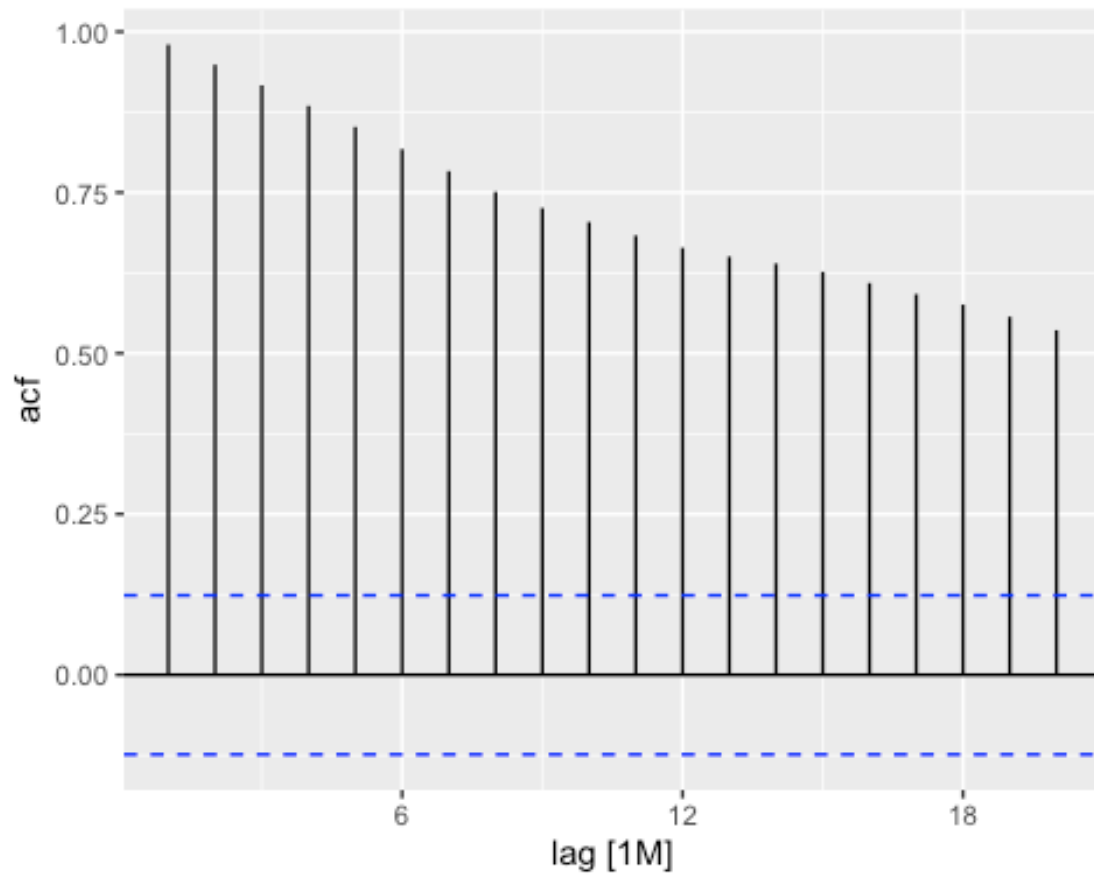


In the last 10 years, there is less of a positive relation with lagged versions.

Let's look at the autocorrelation function of the time series from the last 20 years

```
recent.exchange = US.AUexchange %>%
  filter(year(observation_date) >= 2004)

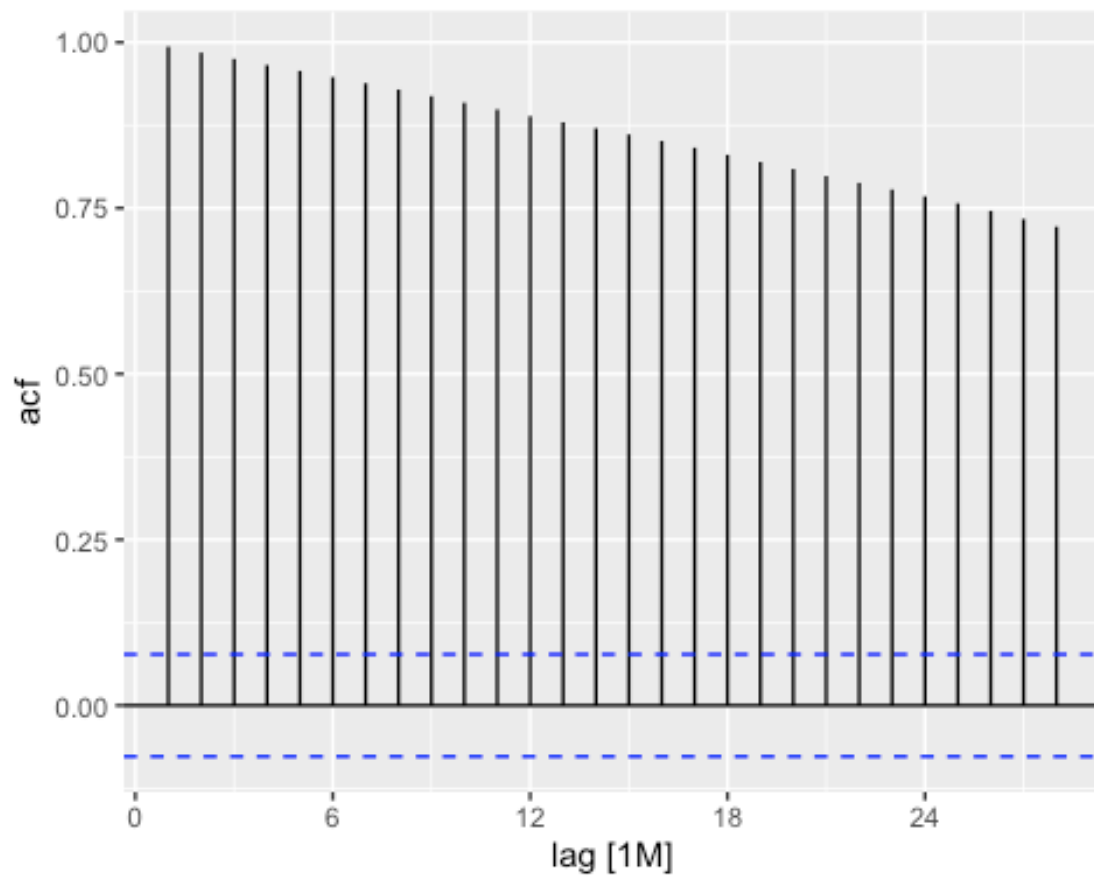
recent.exchange %>%
  ACF(EXUSAL, lag_max = 20) %>%
  autoplot()
```



As the lags increases, the ACF decays overtime due to the decreasing trend. Potentially no seasonality.

Without the lag-max constraint:

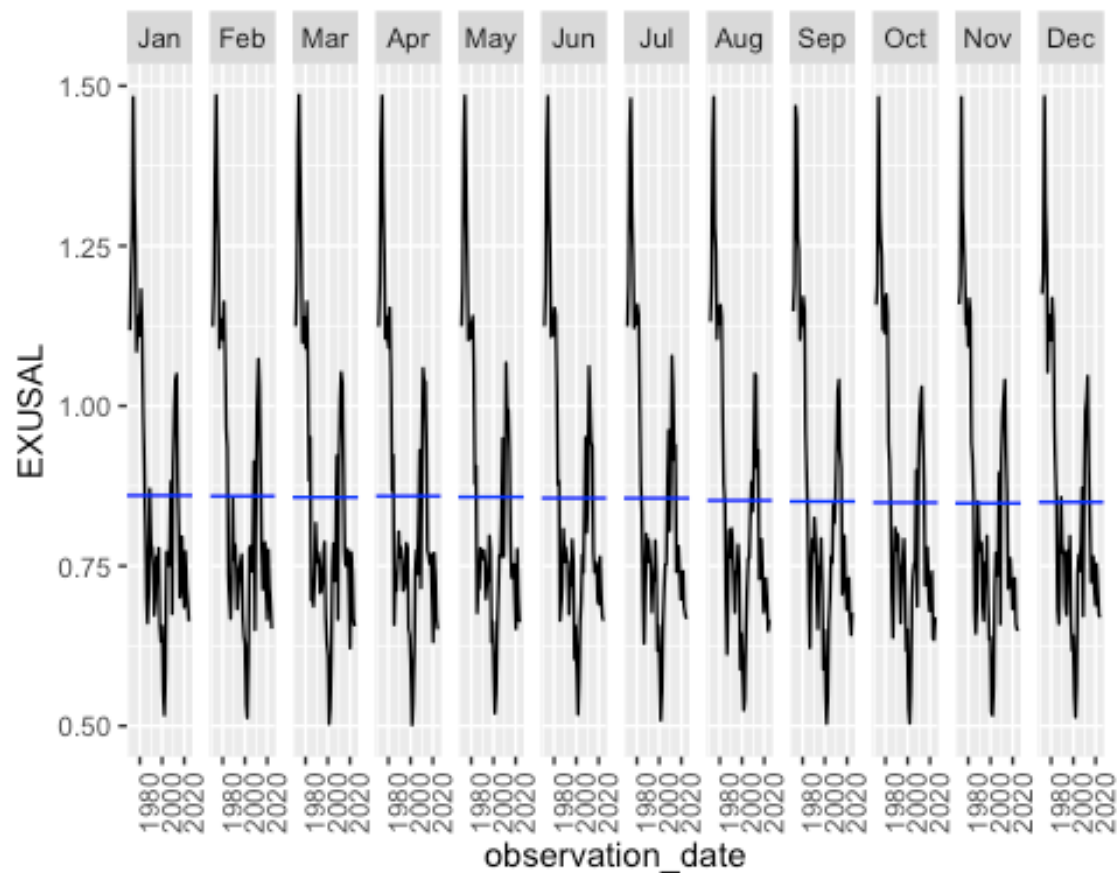
```
US.AUexchangerate %>%
  ACF(EXUSAL) %>%
  autoplot()
```



Same story as previously.

How about the subseries plot just to make sure there is no seasonal pattern.

```
US.AUexchangerate %>%  
  gg_subseries(EXUSAL)
```



Similar means across the board. As a result, the series has no or minimal seasonality. Therefore, the series is not white noise.

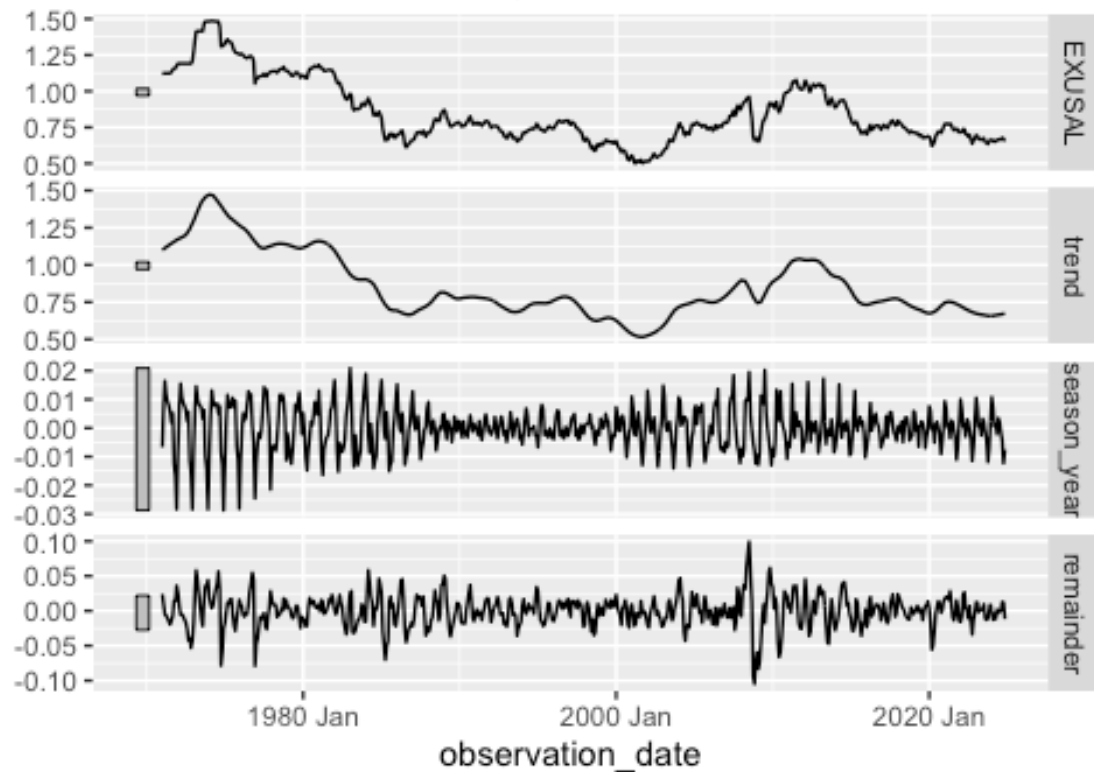
Performing a time-series decomposition

```
decomposition = US.AUexchangerate %>%
  model(stl = STL(EXUSAL))

components(decomposition) %>%
  autoplot()
```

## STL decomposition

EXUSAL = trend + season\_year + remainder



It is now certain that there is a decreasing trend, and, no seasonality. There are dips during pre 1980, 2008 and 2020 indicating that the 1970s crash, 2008 GFC and the COVID lockdowns have leaked into the remainder component.

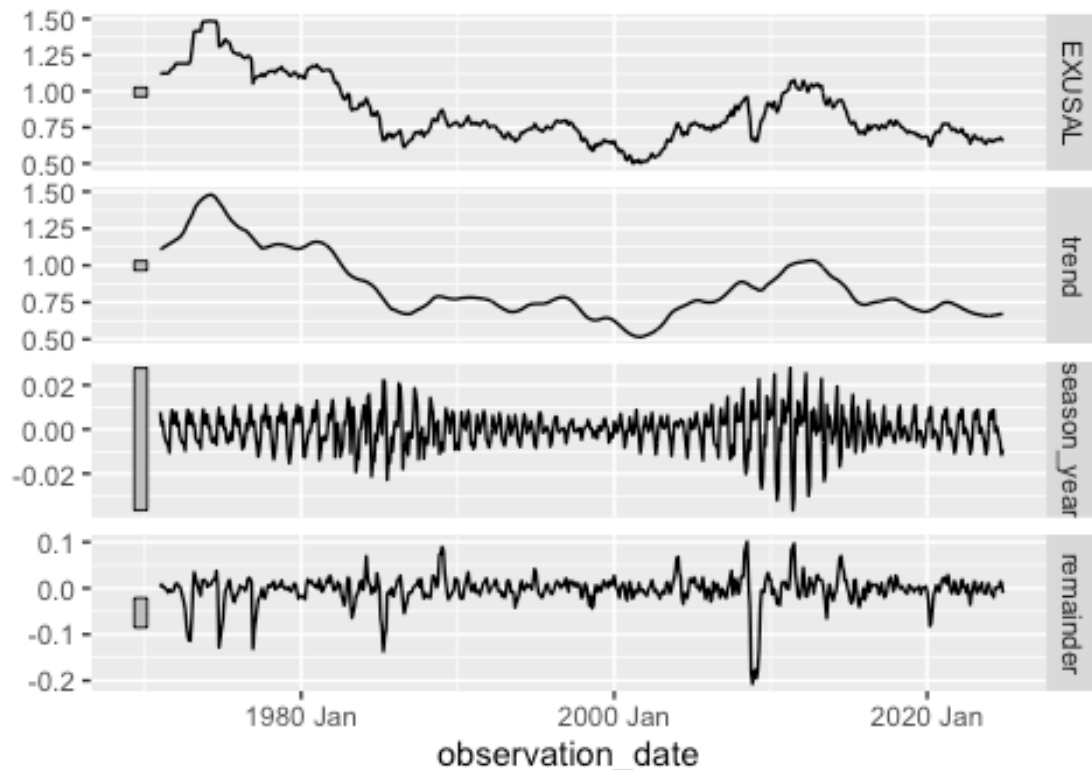
Adding `robust = TRUE` to mitigate the effects of the 1970s crash, 2008 GFC and 2020 COVID lockdowns.

```
decomposition = US.AUexchangerate %>%  
  model(stl = STL(EXUSAL, robust = TRUE))  
  
components(decomposition) %>%  
  autoplot()+  
  labs(title = "STL Decomposition with Robust Smoothing")
```



## STL Decomposition with Robust Smoothing

EXUSAL = trend + season\_year + remainder



Decomposition reinforces the idea of no seasonality. The trend is a bit wiggly, but appears to have an overall downwards trend. Notably, the remainder component still has significant deviations despite the robust smoothing, indicating the effect of the economic shocks. Therefore, we should consider including the outliers in the model as it could explain the nature of exchange rate in forecasts.

Due to cyclicity, let's fit an ARIMA model. First-order differencing since mean is not constant, but how many orders?

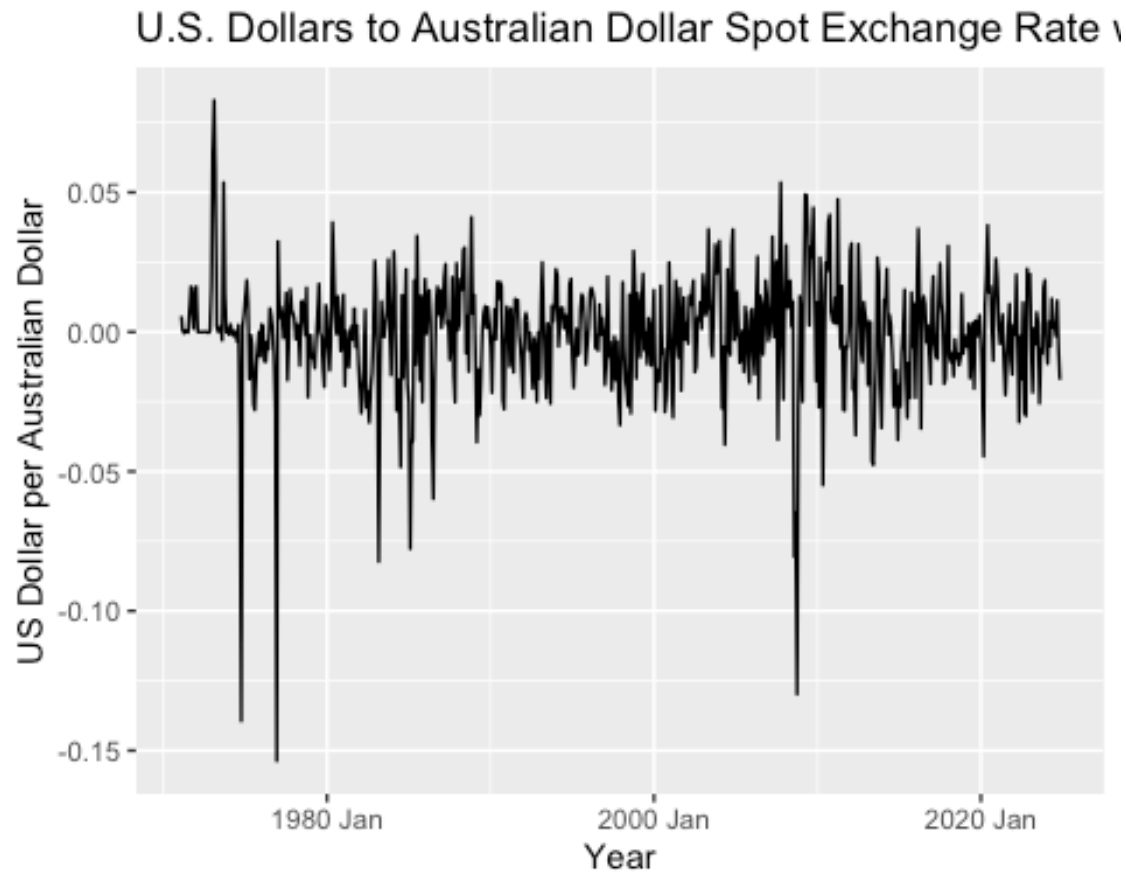
Performing a KPSS test

```
US.AUexchangerate %>%  
  features(EXUSAL, unitroot_ndiffs)  
  
## # A tibble: 1 × 1  
##   ndiffs  
##   <int>  
## 1     1
```

One order difference is sufficient.

Performing and plotting one order difference

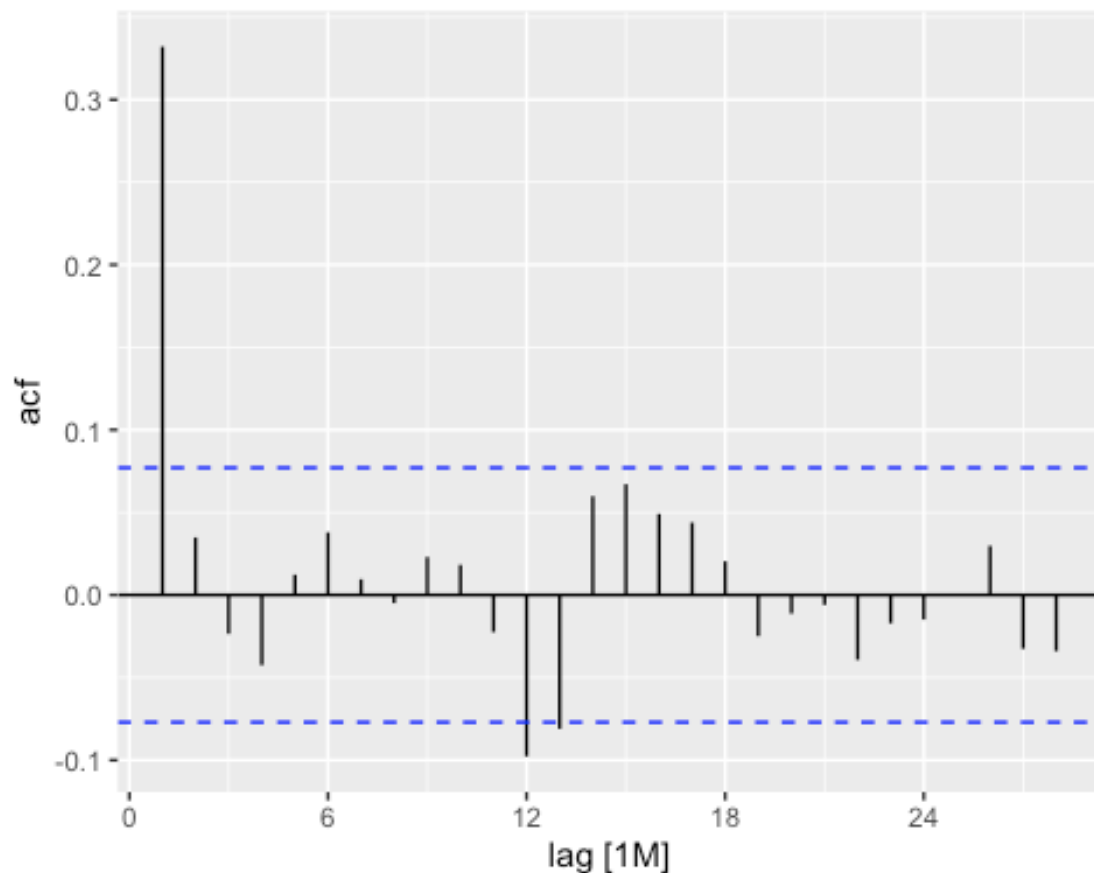
```
US.AUexchangerate %>%
  autoplot(difference(EXUSAL)) +
  labs(y="US Dollar per Australian Dollar", x = "Year", title = "U.S. Dollars
to Australian Dollar Spot Exchange Rate with first order differencing")
```



Extreme values due to economic shocks, but because of the sinusoidal nature of the business cycle we'll keep in the outliers.

ACF plot of differenced exchange rate

```
US.AUexchangerate %>%
  ACF(difference(EXUSAL)) %>%
  autoplot()
```



2 extreme points at lag 1 and 12. Let's perform a KPSS test to test for stationarity.

KPSS test comparison before and after differencing

```
US.AUexchangerate %>%
  features(EXUSAL, unitroot_kpss)
```

```
## # A tibble: 1 × 2
##   kpss_stat kpss_pvalue
##   <dbl>     <dbl>
## 1      3.92         0.01
```

P-value < 0.05 which reject the null hypothesis that the series is stationary and non-seasonal.

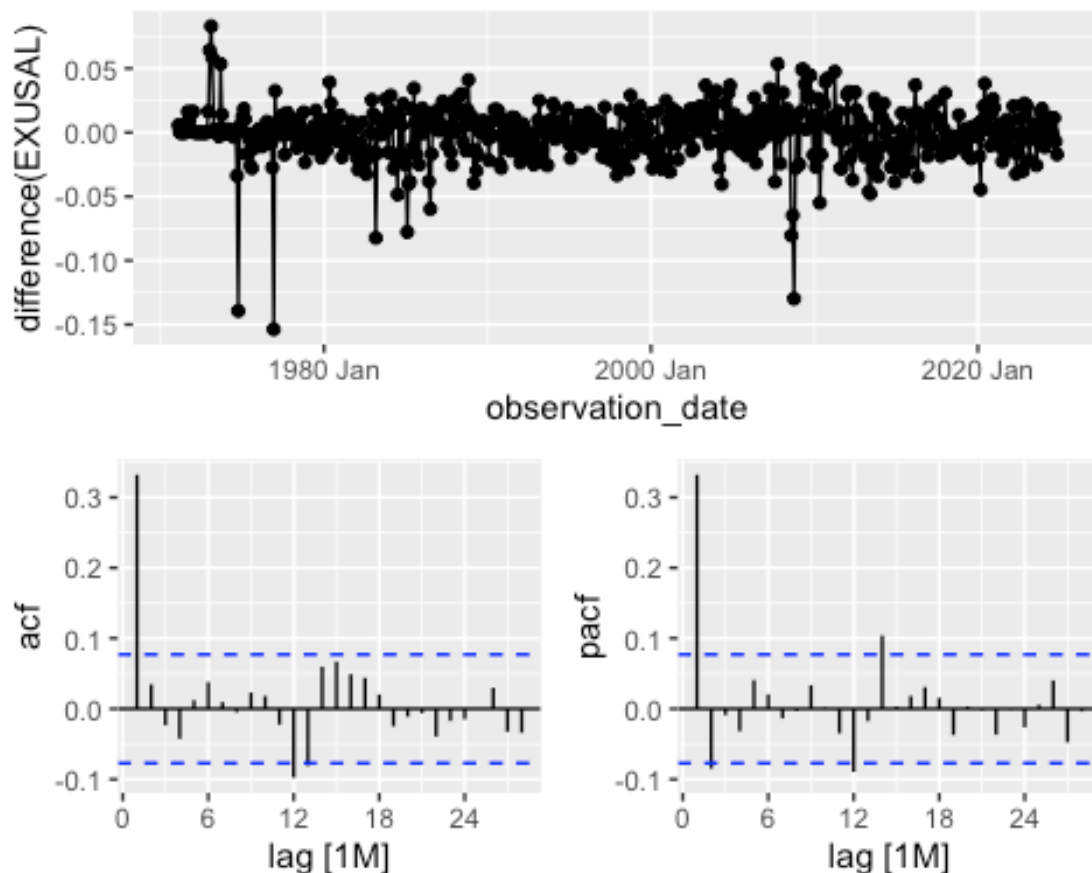
```
US.AUexchangerate %>%
  features(difference(EXUSAL), unitroot_kpss)
```

```
## # A tibble: 1 × 2
##   kpss_stat kpss_pvalue
##   <dbl>     <dbl>
## 1    0.0726         0.1
```

P-Value > 0.05 which fails to reject the null hypothesis that the series is stationary and non-seasonal

Since observations of the exchange rate are correlated, let's use the ACF and PACF to determine some candidate models.

```
US.AUexchangerate %>%  
  gg_tsdisplay(difference(EXUSAL), plot_type = "partial")
```



The ACF shows a significant spike at lag 1, and then a damped sinusoidal curve afterwards. Moreover, there is a significant spike at lag 1 in the PACF, but none beyond lag 2. I will not include a constant in all potential candidate models as first-order differencing is applied, and adding a non-zero constant would make long-term forecasts follow a straight line. Let's consider additional AR and MA terms in our candidate models incase it produces better forecasts.

## Model Building

Inspecting the candidate models

```
fits = US.AUexchangerate %>%  
  model(arima111 = ARIMA(EXUSAL~pdq(1,1,1)+ PDQ(0,0,0)), # From the
```

```

diagnostics
  arima211 = ARIMA(EXUSAL~pdq(2,1,1)+ PDQ(0,0,0)), # Lingering
significance at lag 2
  arima112 = ARIMA(EXUSAL~pdq(1,1,2)+ PDQ(0,0,0)), # ACF plot shows
significance beyond the first lag
  arima212 = ARIMA(EXUSAL~pdq(2,1,2)+ PDQ(0,0,0)), # Adding extra AR
and MA terms in case of complexity
  arima011 = ARIMA(EXUSAL~pdq(0,1,1)+ PDQ(0,0,0)), # Testing if AR
terms are significant
  arima110 = ARIMA(EXUSAL~pdq(1,1,0)+ PDQ(0,0,0)), # Testing if MA
terms are significant
  arimabaseline = ARIMA(EXUSAL~pdq(0,1,0)+ PDQ(0,0,0))) # Baseline
model with no AR pr MA terms

```

```

glance(fits) %>%
  arrange(AICc) %>%
  select(.model:BIC)

## # A tibble: 7 × 6
##   .model      sigma2 log_lik    AIC    AICc    BIC
##   <chr>      <dbl>   <dbl>  <dbl>  <dbl>  <dbl>
## 1 arima011    0.000385  1623. -3243. -3243. -3234.
## 2 arima111    0.000385  1624. -3242. -3242. -3229.
## 3 arima211    0.000385  1624. -3240. -3240. -3223.
## 4 arima112    0.000385  1624. -3240. -3240. -3222.
## 5 arima110    0.000387  1622. -3240. -3240. -3231.
## 6 arima212    0.000386  1624. -3239. -3238. -3216.
## 7 arimabaseline 0.000434  1584. -3166. -3166. -3161.

```

ARIMA(0,1,1) has the lowest BIC, AIC, and AICc. The difference between the Information Criterion for all the models are relatively small.

Inspecting the ARIMA(0,1,1) model

```

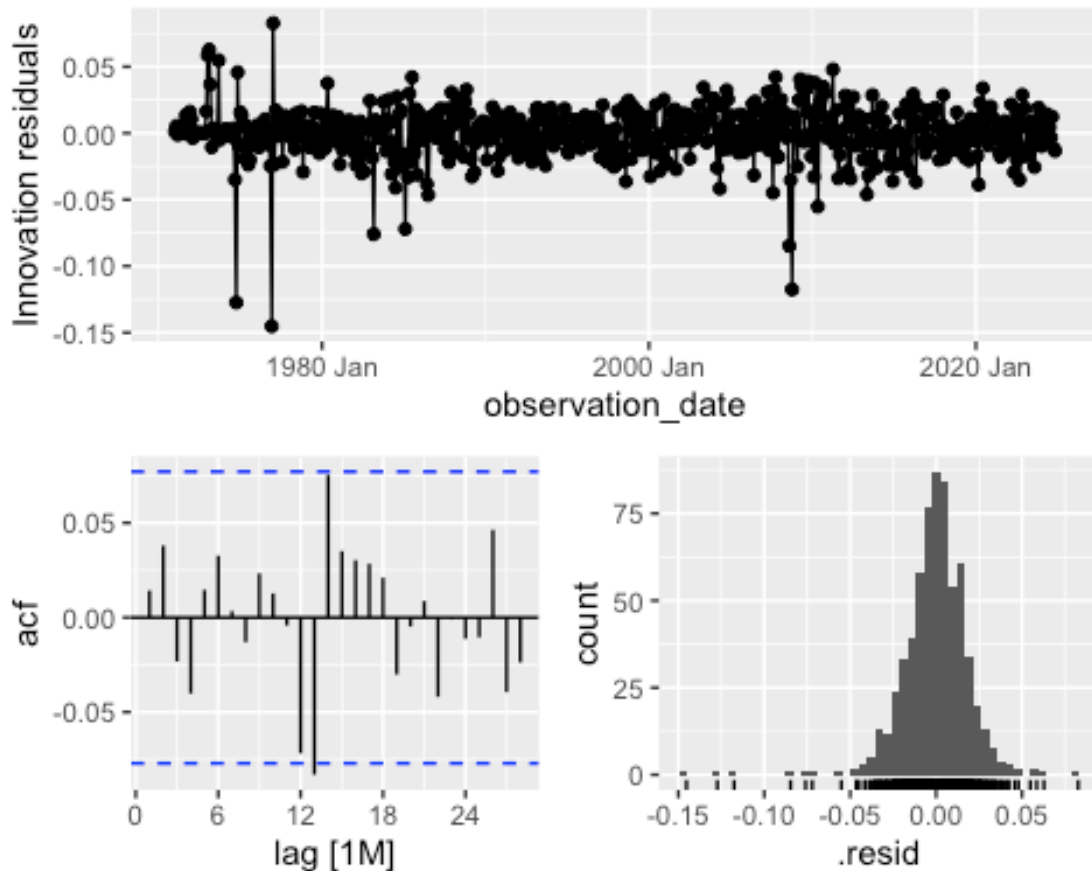
fits %>%
  select(arima011) %>%
  report()

## Series: EXUSAL
## Model: ARIMA(0,1,1)
##
## Coefficients:
##           ma1
##          0.3464
## s.e.    0.0357
##
## sigma^2 estimated as 0.0003848:  log likelihood=1623.48
## AIC=-3242.97  AICc=-3242.95  BIC=-3234.02

```

Looking at the innovation residuals

```
fits %>%
  select(arima011) %>%
  gg_tsresiduals()
```



Residuals seem to have a zero mean and are normally distributed.

Performing a Ljung-Box test to see if the chosen model is autocorrelated at multiple lags.

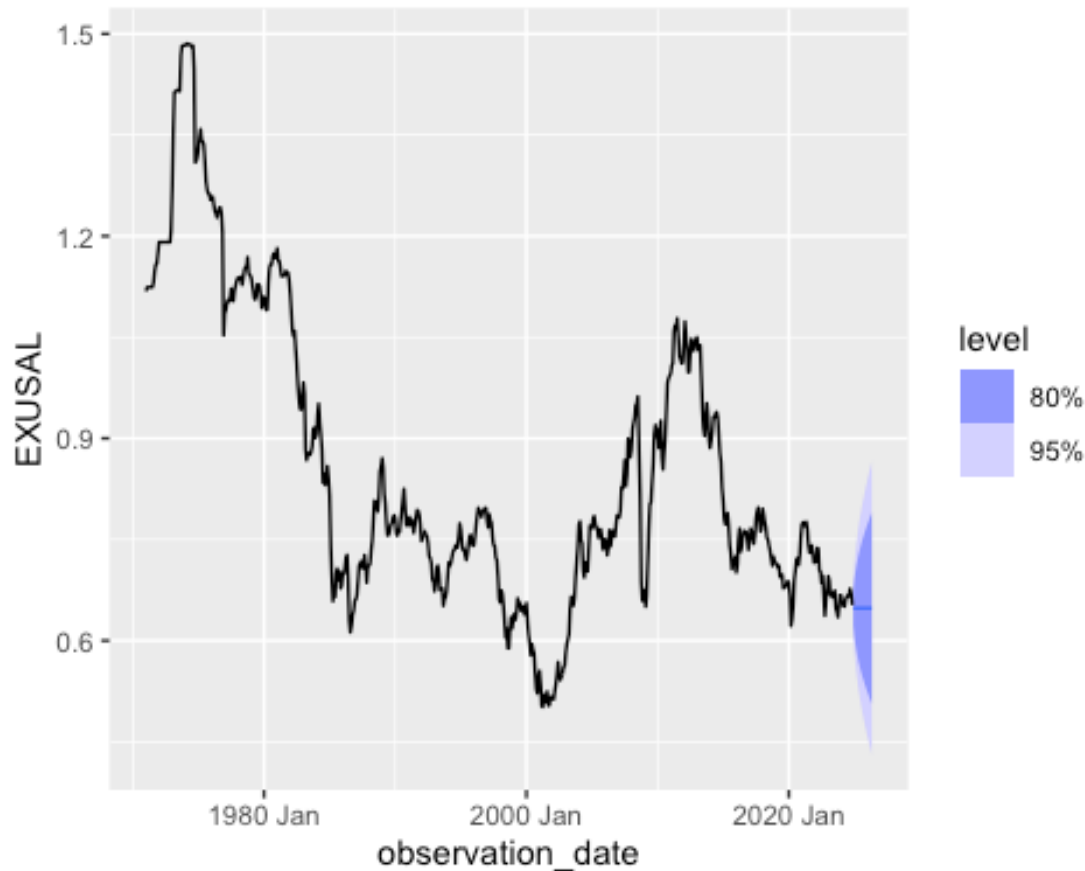
```
fits %>%
  select(arima011) %>%
  augment() %>%
  features(.resid, lbj_box, lag = 10)
```

```
## # A tibble: 1 × 3
##   .model  lb_stat lb_pvalue
##   <chr>    <dbl>    <dbl>
## 1 arima011    3.91    0.951
```

P-Value > 0.05 which fails to reject the null hypothesis that the the residuals are independently distributed (no autocorrelation). Thus, the series is likely white noise.

Forecasting using ARIMA(0,1,1) with 80% and 95% prediction intervals.

```
fits %>%
  select(arima011) %>%
  forecast(h = 18) %>% #18 months
  autoplot(US.AUexchangerate)
```



Wide intervals indicate greater uncertainty, which is common for long-term forecasts or highly volatile time series.

However, exchange rates can be influenced by past rates due to exchange rate expectation, interest rate expectation, and anticipation of the business cycle which influences confidence and expectations. Therefore, let's consider a model with at least 1 autoregressive term that captures the economic effects. Despite the BIC ranking from the candidate models, an AR term seems reasonable to include based on the PACF of the differenced series.

Let's consider at least 1 autoregressive term.

Picking the model with minimal errors

```
fits %>%
  accuracy() %>%
  arrange(MAE)
```

```
## # A tibble: 7 × 10
##   .model      .type      ME    RMSE    MAE      MPE    MAPE    MASE    RMSSE
ACF1
##   <chr>      <chr>      <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
<dbl>
## 1 arima011   Train... -5.38e-4 0.0196 0.0135 -0.0805  1.68 0.187 0.209
1.43e-2
## 2 arima212   Train... -5.33e-4 0.0196 0.0136 -0.0794  1.69 0.187 0.209
2.29e-4
## 3 arima211   Train... -5.32e-4 0.0196 0.0136 -0.0792  1.69 0.187 0.209
3.29e-4
## 4 arima111   Train... -5.15e-4 0.0196 0.0136 -0.0760  1.69 0.187 0.209
6.60e-4
## 5 arima112   Train... -5.15e-4 0.0196 0.0136 -0.0761  1.69 0.187 0.209 -
3.75e-5
## 6 arima110   Train... -4.87e-4 0.0196 0.0136 -0.0707  1.70 0.188 0.210
2.79e-2
## 7 arimabaseline Train... -7.17e-4 0.0208 0.0142 -0.114  1.76 0.196 0.222
3.32e-1
```

ARIMA(2,1,2) has the lowest MAE and RMSE after ARIMA(0,1,1) which are the most commonly used errors. The difference in AIC between the candidate are quite small. The log-Likelihood is also similar between the candidate models. Including the AR(2) terms enable us to better capture the effects of past exchange rates to produce better forecasts. Moreover, the MA(2) terms enable us to utilise past forecast errors to produce better forecasts as well. Since the USD/AUD exchange adjust via the business cycle and adjust through expectations, using AR and MA terms is appropriate as economies grow and shrink.

What about the ARIMA(2,1,2) model?

```
#Building the ARIMA(2,1,2) model
arima212 = US.AUexchangerate %>%
  model(ARIMA(EXUSAL~pdq(2,1,2)+PDQ(0,0,0)))
```

Is the difference in AIC significant? ARIMA(1,1,1) is a subset of ARIMA(2,1,2) and thus are nested models. Thus, we can perform a likelihood ratio test to determine if the difference in parameters make a difference.

Performing a Likelihood Ratio Test

```
# Extract the log-likelihood values
log_likelihood_null = fits %>%
  select(arima111) %>%
  glance() %>% #Null model is ARIMA(1,1,1)
  select(log_lik) %>%
  as.numeric()

log_likelihood_alt = glance(arima212) %>% #Alternative model is ARIMA(2,1,2)
  select(log_lik) %>%
  as.numeric()
```



```

# Compute the likelihood ratio test statistic
lr_statistic = -2 * (log_likelihood_null - log_likelihood_alt)

# Degrees of freedom: difference in number of parameters
df = length(coef(arima212)) - length(coef(fits %>%
  select(arima111)))

# Compute the p-value (Chi-squared distribution)
p_value = pchisq(lr_statistic, df, lower.tail = FALSE)

print(p_value)

## [1] 0

```

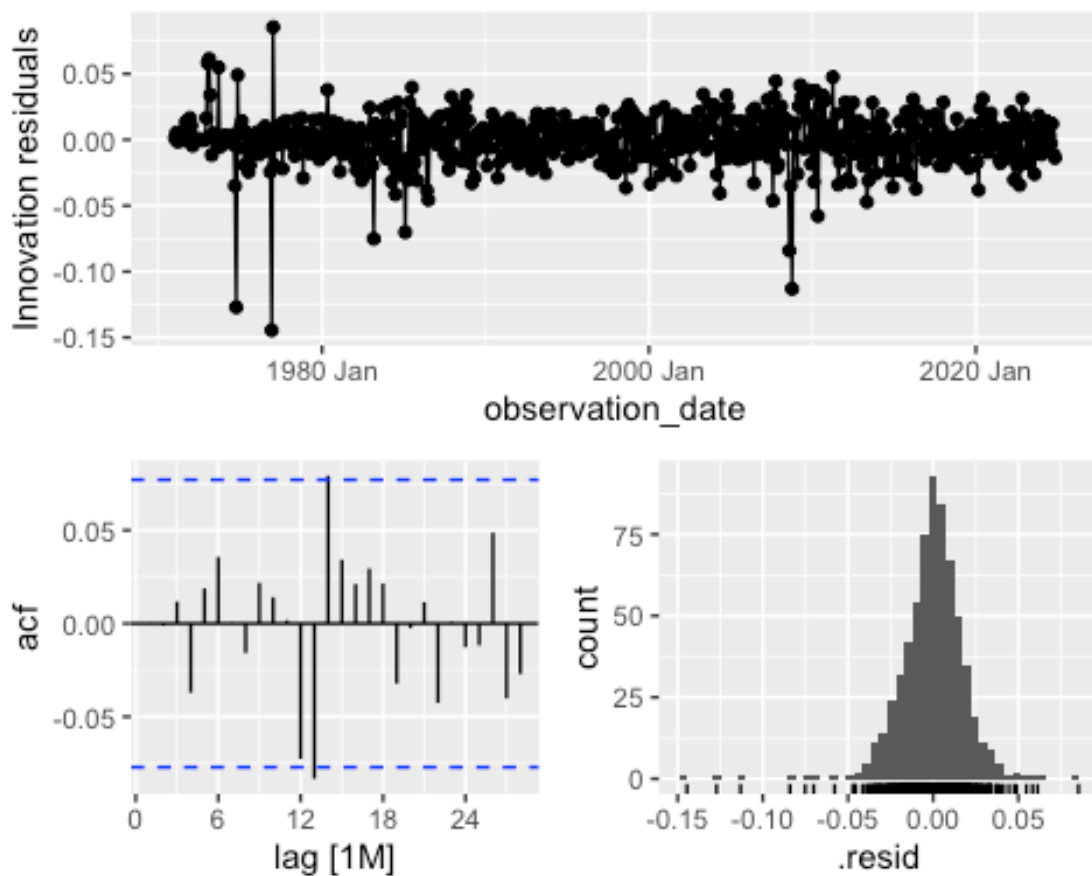
P-Value < 0.05 which rejects the null hypothesis and conclude that the alternative model (ARIMA(2,1,2)) provides a better fit

Looking at ARIMA(2,1,2) model's innovation residuals, ACF, and distribution of residuals.

```

arima212 %>%
  gg_tsresiduals()

```



Residuals are close to 0 and appear normally. However, there are some outliers (due to economic shocks) which will be included as part of understanding the business cycle or policy changes.

Performing a Ljung-Box test to see if the chosen model is autocorrelated at multiple lags.

```

arima212 %>%
  augment() %>%
  features(.resid, ljung_box, lag = 10)

## # A tibble: 1 × 3
##   .model                                lb_stat lb_pvalue
##   <chr>                                <dbl>    <dbl>
## 1 ARIMA(EXUSAL ~ pdq(2, 1, 2) + PDQ(0, 0, 0))    2.67    0.988

```

This is the probability of observing the test statistic under the Chi-squared distribution. P-Value > 0.05 which fails to reject the null hypothesis that the residuals are independently distributed (no autocorrelation). Thus, the series is likely white noise.

Inspecting the coefficients

```

arima212 %>%
  report()

## Series: EXUSAL
## Model: ARIMA(2,1,2)
##
## Coefficients:
##          ar1          ar2          ma1          ma2
##          0.5498   -0.1258   -0.1902   -0.0251
## s.e.    0.5650    0.1532    0.5653    0.1778
##
## sigma^2 estimated as 0.0003857:  log likelihood=1624.27
## AIC=-3238.54   AICc=-3238.45   BIC=-3216.19

```

## Final model

The general model is:  $(1 - \phi_1 B - \phi_2 B^2)(1 - B)y_t = (1 + \theta_1 B + \theta_2 B^2)\epsilon_t$

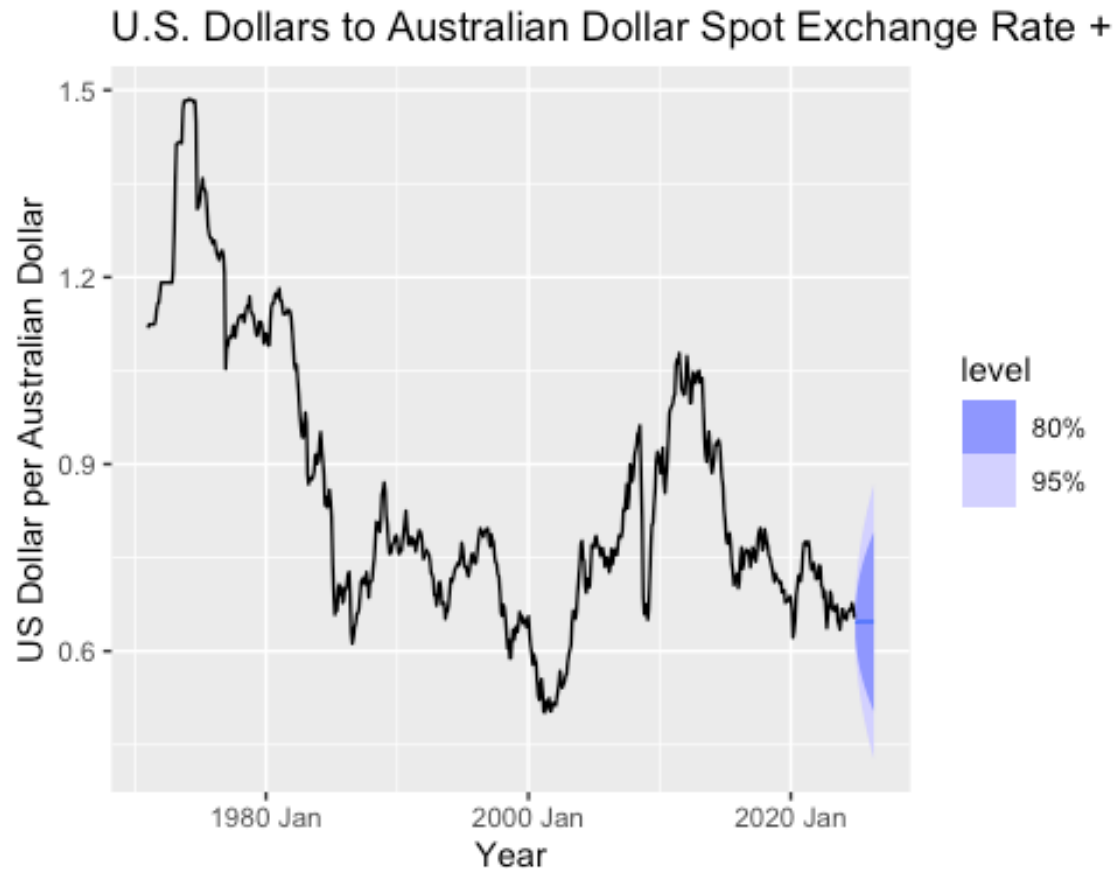
The model is:  $(1 - 0.55B + 0.13B^2)(1 - B)USD/AUD_t = (1 - 0.19B - 0.03B^2)\epsilon_t$

Forecasting using ARIMA(2,1,2) with 80% and 95% prediction intervals.

```

arima212 %>%
  forecast(h = 18) %>% #h = 18 indicates 18 months
  autoplot(US.AUexchangerate) +
  labs(y = "US Dollar per Australian Dollar", x = "Year", title = "U.S. Dollars
to Australian Dollar Spot Exchange Rate + Forecast")

```



Therefore, we know where the exchange rate is forecasted to be within the 18 months based on the ARIMA(2,1,2) model. At around the max point of the 95% predictive interval, the exchange rate is just around 0.88 which would indicate a depreciation of the USD against the AUD. At the lowest portion of the 95% predictive interval, the exchange rate is around 0.44 which would indicate a appreciation of the USD against the USD.

Decision making should be based on interest rate differential forecasts as well.