Time Series Models

Assingment I - Local Level Methods

Garret Frank (2738169)
Arvanitis Argyrios (2748465)
Beck DeYoung (2724909)
Max van der Meulen (2639613)

February 25, 2022

Supervisor:
prof. dr. SJ Koopman, Gabriele Mingoli, Karim Moussa, Ilka van de Werve

Contents

1 Introduction

For this assignment, we recreate the figures from the second chapter of the book *Time Series Analysis by State Space Methods: Second Edition* (Durbin & Koopman (2012)) using the Local Level Model (LLM). By recreating these figures, we become familiar with Kalman filtering, Kalman smoothing and forecasting. This is useful for a thorough understanding of the topics covered in the course Time Series Models as it relates to signal extraction.

After recreating these figures we have stated our findings for the figures. By analyzing these figures, we were able to observe differences in signal extraction using the Kalman filter and other methods. In addition, we are able to apply these methods in situations where data is missing to both interpolate and extrapolate data. We also, are able to forecast future data by treating it as missing.

Finally, we used the techniques that were used in the first part of this assignment to study inflation rates in the United States. In this case we used the commonly used *CPI* time series taken from the Federal Reserve Bank of St. Louis. We discuss how the data is utilized in the LLM following the methods set out in *Time Series Analysis by State Space Methods: Second Edition* (Durbin & Koopman (2012)) using Nile data.

2 Part (a): Nile Data Graphs from Chapter 2

The following section displays the graphs from 2.1 - 2.8 excluding 2.4 as outlined in *Time Series Analysis by State Space Methods: Second Edition* (Durbin & Koopman (2012)) using the same Nile Data. Further observations for each figure are found in the following section.

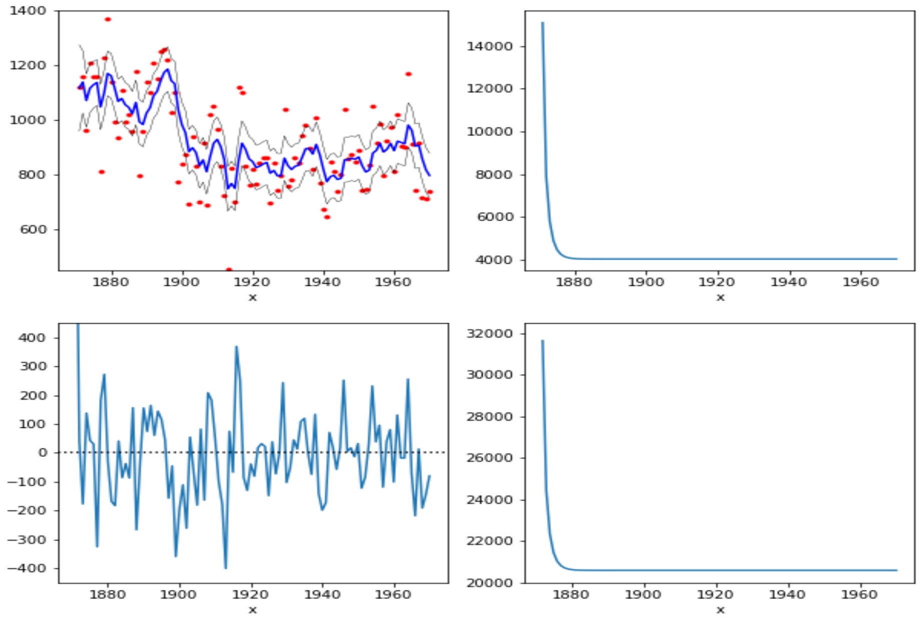

Figure 1: Nile data and output of Kalman filter: (i) data (dots), filtered state at (solid line) and its 90% confidence intervals (light solid lines); (ii) filtered state variance P_t; (iii) prediction errors vt; (iv) prediction variance F_t.

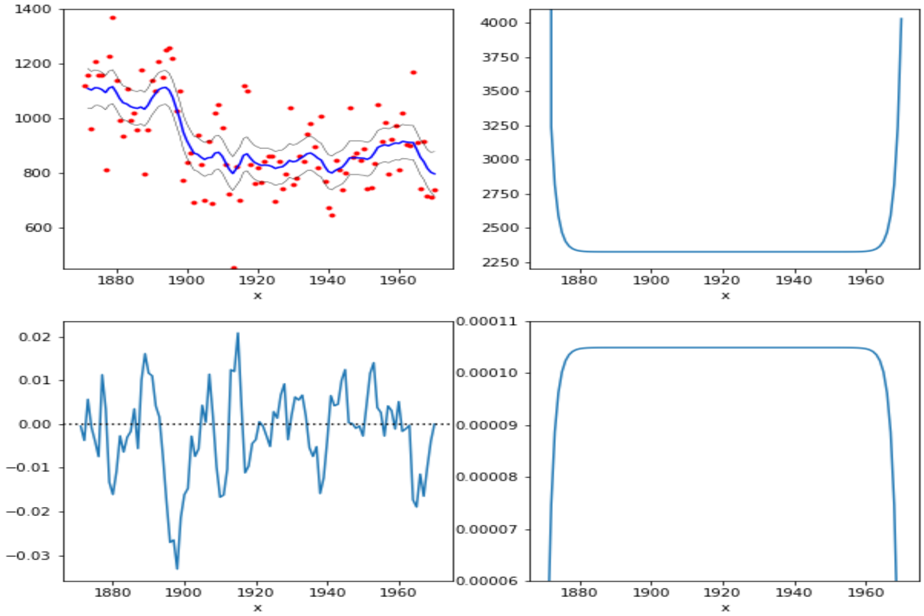

Figure 2: Nile data and output of state smoothing recursion: (i) data (dots), smoothed state α_t and its 90% confidence intervals; (ii) smoothed state variance V_t (iii) smoothing cumulant r_t; (iv) smoothing variance cumulant N_t.

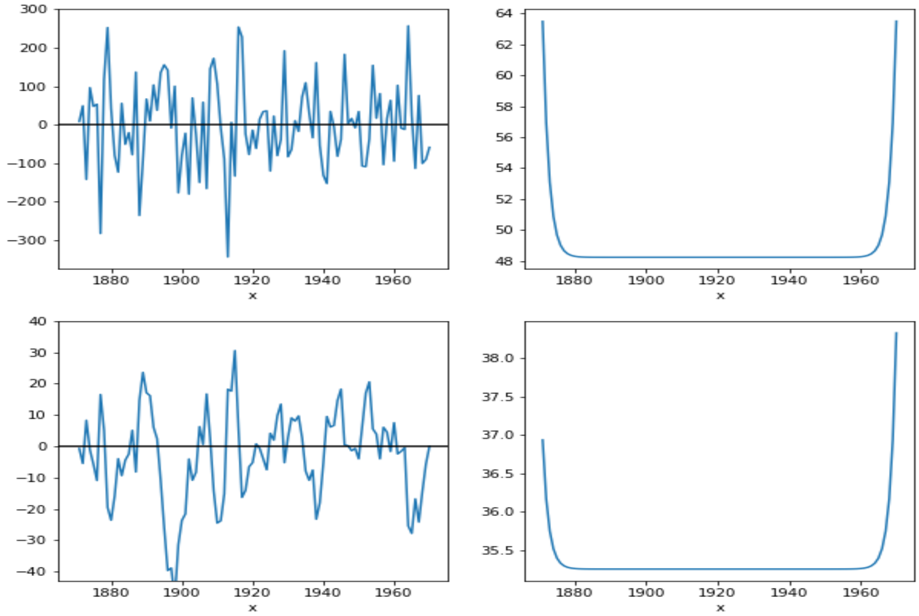

Figure 3: Output of disturbance smoothing recursion: (i) observation error $\hat{\varepsilon}_t$; (ii) observation error variance $Var(\varepsilon_t|Y_n)$; (iii) state error $\hat{\eta}_t$; (iv) state error variance $Var(\eta_t|Y_n)$.

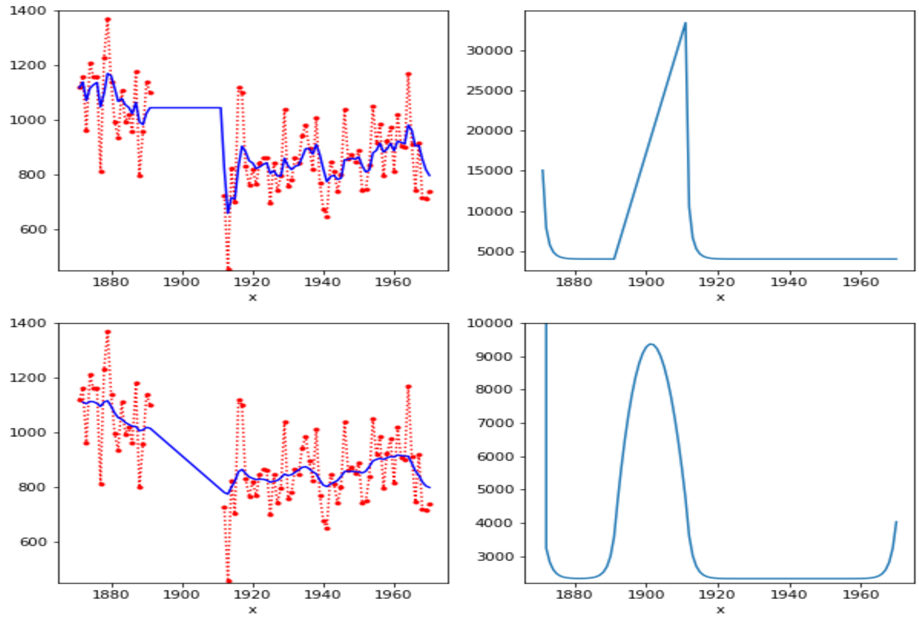

Figure 4: Filtering and smoothing output when observations are missing: (i) data and filtered state a_t (extrapolation); (ii) filtered state variance P_t; (iii) data and smoothed state $\hat{\alpha}_t$ (interpolation); (iv) smoothed state variance V_t.

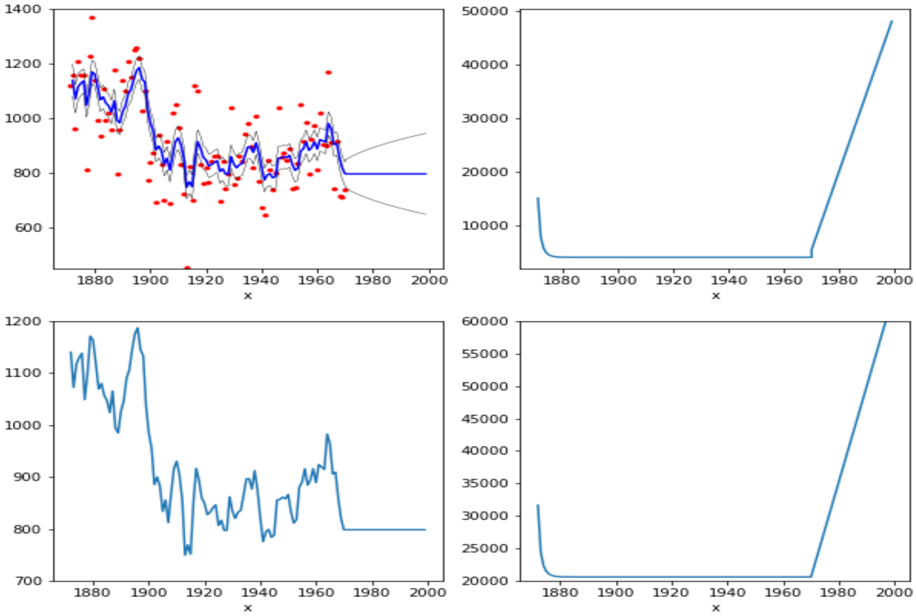

Figure 5: Nile data and output of forecasting: (i) data (dots), state forecast at and 50% confidence intervals; (ii) state variance P_t; (iii) observation forecast $E(y_t|Y_{t-1})$; (iv) observation forecast variance F_t.

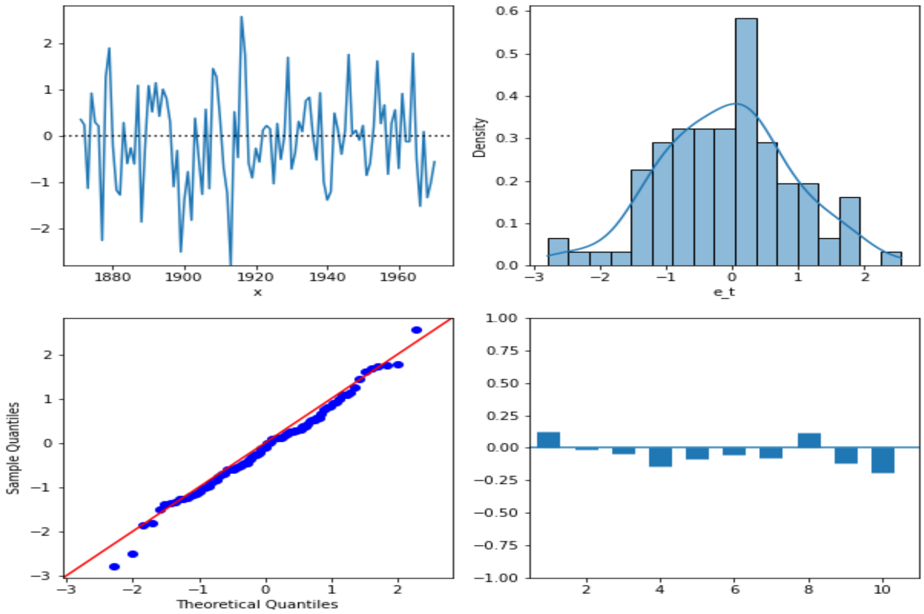

Figure 6: Diagnostic plots for standardised prediction errors: (i) standardised residual; (ii) histogram plus estimated density; (iii) ordered residuals; (iv) correlogram.

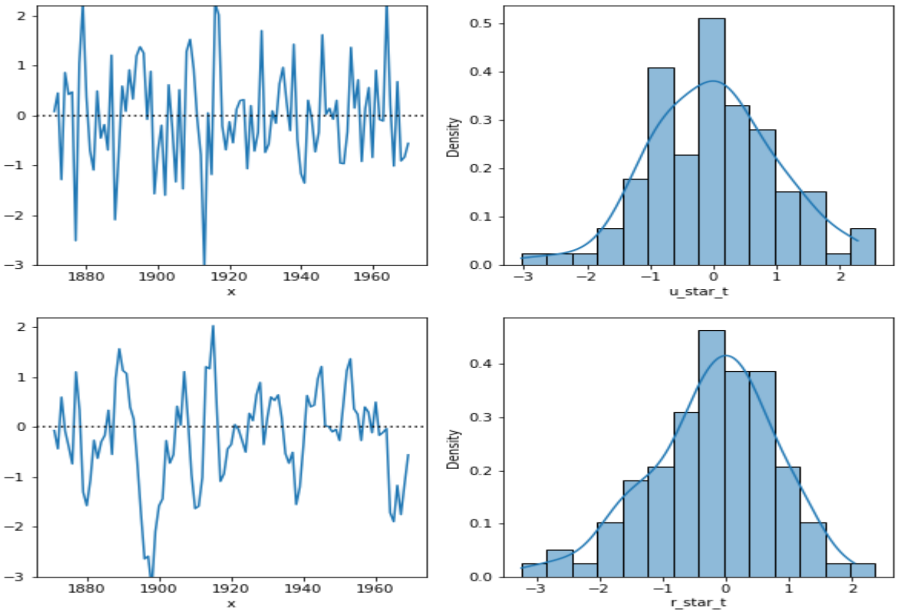

Figure 7: Diagnostic plots for auxiliary residuals: (i) observation residual u_t^* ; (ii) histogram and estimated density for u_t^*; (iii) state residual r_t^*; (iv) histogram and estimated density for r_t^*.

3 Part (b): Explanation Of The Nile Data Graphs

In this section we will explain the obtained figures and how they relate to each other.First, we start with Figure 1. In this figure we can see the Nile data and output of the Kalman filter. It is clear to see that both the filtered state variance P_t and the prediction variance F_t converge rapidly which confirms that the local level model has a steady state solution. Looking at the prediction errors, we can see that there is a lot of variation but that the mean is roughly zero.

In Figure 2 we can see the Nile data and the output of the smoothing recursion. The big difference between Figure 1 and Figure 2 is that, although the two estimated lines follow the same trend, the line in the first figure is much less smooth than the in the second figure. This is because Kalman smoothing uses future data along with past data and the data of today. This makes the extracted signal appear smooth.

Looking at Figure 3 we can see the output of the disturbance smoothing recursion. The main take away from this figure is that the errors of both the observation error variance $Var(\varepsilon_n)$ as the state error variance $Var(\eta_t|Y_t)$ move in the same way, but the scale is different and the state error variance is much smaller at the beginning of the sample. The scale is different since at the beginning and the end of the sample the observation error variance is almost twice as big as the state error variance. Which is directly translated to the plots of the observation error and the state error. There we can see that the minimum and maximum of the observation error are almost ten times larger than the minimum and maximum of the state error. Finally, it is trivial to see that the observation was already plotted in 1 and the state error variance was already plotted in 2.

Figure 4 shows the filtering and smoothing output when observations are missing. Since the smoothed state also uses the future data, the estimation for the missing observations is more reliable than the estimation of the filtered state. The filtered state only uses the past data and the present data to estimate, therefore the line is straight from the last observed point until it reaches the next observed point in (i). This is different with smoothing, since this algorithm also uses future data and therefore, the line is diagonal towards the next observation(ii). You can also observe a large spike in (ii) compared to 1 due to the flattening of data in (i). This is also observed in (iv) but in more of a parabolic or smooth manner.

In Figure 5 we can see the Nile data and the forecast output. The data set is extended by thirty missing values. When forecasting, only the Kalman filter is required since forecasting a future observation is equal to estimating a missing observation at time $t + n$. For the forecast the prediction error cannot be obtained, because we do not have the true observations for our prediction. Therefore, the prediction error variance will go to infinity and hence the Kalman gain will go to zero. The forecast is a straight line, since the Kalman filter uses only past observations and since the Kalman gain will go to zero, the forecast will not vary and thus stays straight. This takes place in both (i) and (iii) causing which inadvertenly caused the upward trends observed in (ii) and (iv)

Figure 6 shows the diagnostic plots for the standardised prediction errors. If we look at the QQ-plot of the ordered residuals we can assume that they are normally distributed. The majority of the residuals are close to the line that indicates the normal distribution. The tails contain more of the outliers however, they are not so big that we have to reject the assumption that these residuals are normally distributed. Next, the correlogram shows that the residuals are random. This concluded by the fact that the values in this correlogram are close to zero.

Finally, in Figure 7 we can see the diagnostic plots for the auxiliary residuals. Both have a distribution similar to the normal distribution, with a mean close to zero. However, the observations residuals are slightly positively skewed, whereas the state residuals are slightly negatively skewed. Looking at both figures, we do not observe heavy outliers which was also confirmed by the previous figure.

4 Part (c): LLM for CPI Data

In the final section, we analyze CPI data from the United States using quarterly data from 1960 to 2022. Using LLM, we replicate the figures in (Durbin & Koopman (2012)) for CPI data. The motivation behind using CPI data is that it displays an nonstationary behavior when plotted over time which can be seen in (i) of Figure 8.

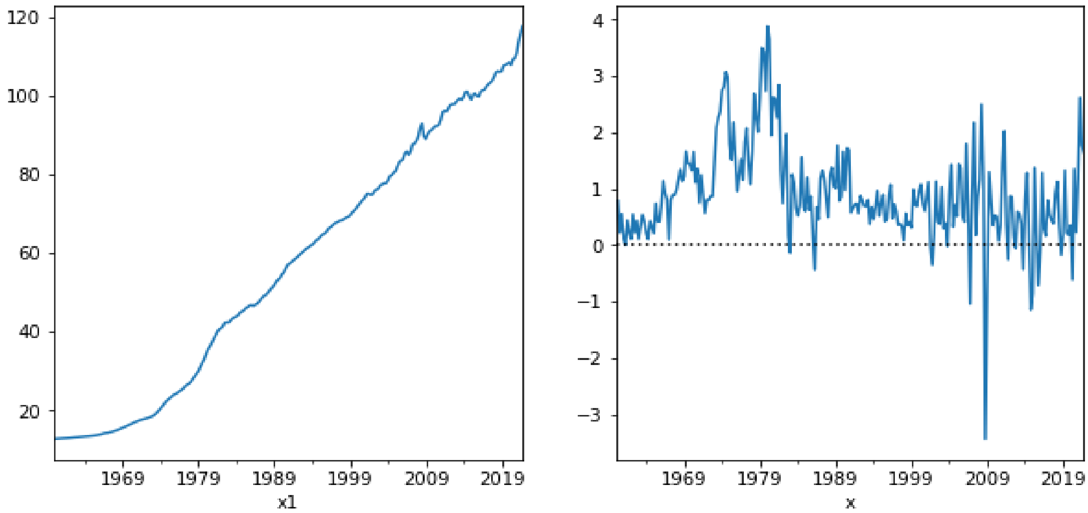

Figure 8: CPI Data: (i) unfiltered CPI data (ii) percent difference CPI data.

By taking the percent difference, we see that the data has noticeable outliers around 1979 and in 2009 and appears to have considerable noise which is typical in a LLM model. There is also no signs of seasonality in our data, which is fitting for the LLM because seasonality is not accounted for. By using an LLM model, we will be able to apply the Kalman filtering, Kalman smoothing, and forecasting to further interpret CPI data.

Using maximum likelihood, we are able to estimate $\sigma_\epsilon = 0.31$ and $\sigma_\eta = 0.049$. This implies that our signal-noise ratio is $q = 0.15$ meaning that the signal is stronger than the noise. These parameters are utilized in our replication of the figures used in previous sections to generate figures based on CPI data. Each figure is displayed below along with a discussion of results

First, we illustrate the output of the Kalman Filter using logarithm change in quarterly price of Consumer Price Index of the United States during the period from 1960 to 2020. In the graphs the filtered state is shown along with the 90 per cent confidence intervals and the prediction errors. The important information produced by these graphs is that the filtered state variance and the prediction variance converge rapidly to constant values which confirms that the local level model chosen for this data has a steady state solution.

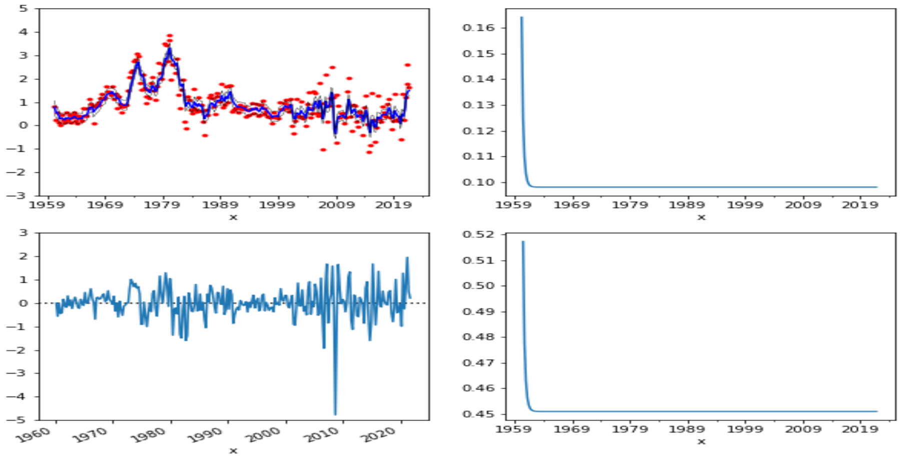

Figure 9: CPI data and output of Kalman filter: (i) data (dots), filtered state at (solid line) and its 90% confidence intervals (light solid lines); (ii) filtered state variance P_t; (iii) prediction errors vt; (iv) prediction variance F_t.

Figure 9 contains the results of state smoothing for our data, obtained by the backwards recursions of the local level model. Both the smoothed state and smoothed state variance are obtained by the state smoothing recursion which uses the smoothing cumulant and smoothing variance cumulant shown in graphs iii and iv. The smoothed state variance seems to be larger at the beginning and the end of the sample since the smoothing variant cumulant is zero for times after 2022 which drives the smoothed state variance higher.

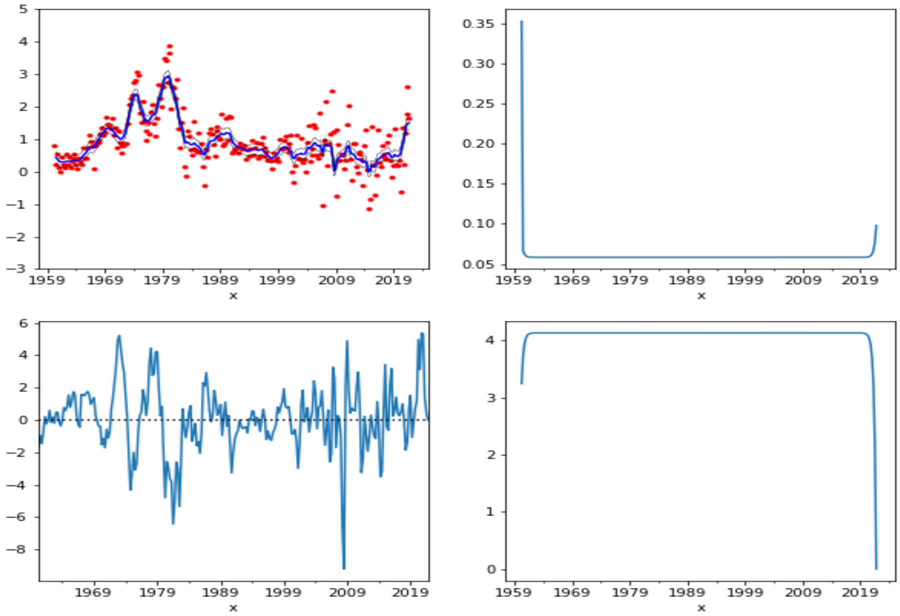

Figure 10: CPI data and output of state smoothing recursion: (i) data (dots), smoothed state α_t and its 90% confidence intervals; (ii) smoothed state variance V_t (iii) smoothing cumulant r_t; (iv) smoothing variance cumulant N_t.

Using the same form of recursion we calculate the smoothed disturbances and their variances and we illustrate them in figure 10. The key takeaways from these graphs are that the disturbance variances move in a different scale but in a similar way with the smoothed state variance for the same reasons mentioned in the previous graph and it is also clear that graph iii of the smoothed state error is the same with the smoothing cumulant r_t from the previous graph, since the state error is the smoothing cumulant r_t multiplied with the variance of the errors.

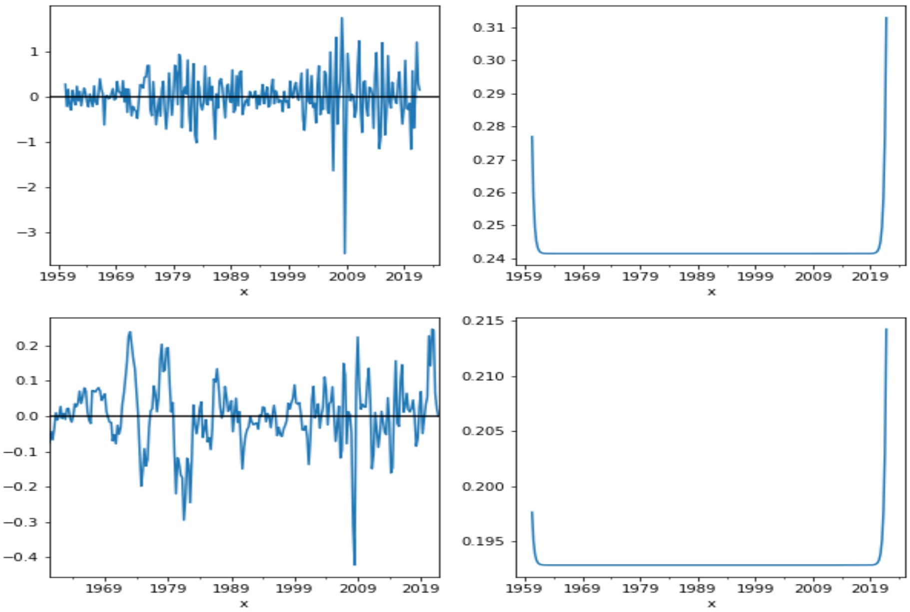

Figure 11: Output of disturbance smoothing recursion: (i) observation error $\hat{\varepsilon}_t$; (ii) observation error variance $Var(\varepsilon_t|Y_n)$; (iii) state error $\hat{\eta}_t$; (iv) state error variance $Var(\eta_t|Y_n)$.

Figure 11 shows the filtering and smoothing output when observations are missing. Graph i shows that the filtered state remains the same for the time periods of missing data since we consider the Kalman gain to be equal to zero for that period. The difference of the smoothed state is clearly visible in graph iii since the recursion used for its calculations takes into consideration both past and future observations for any point in time.

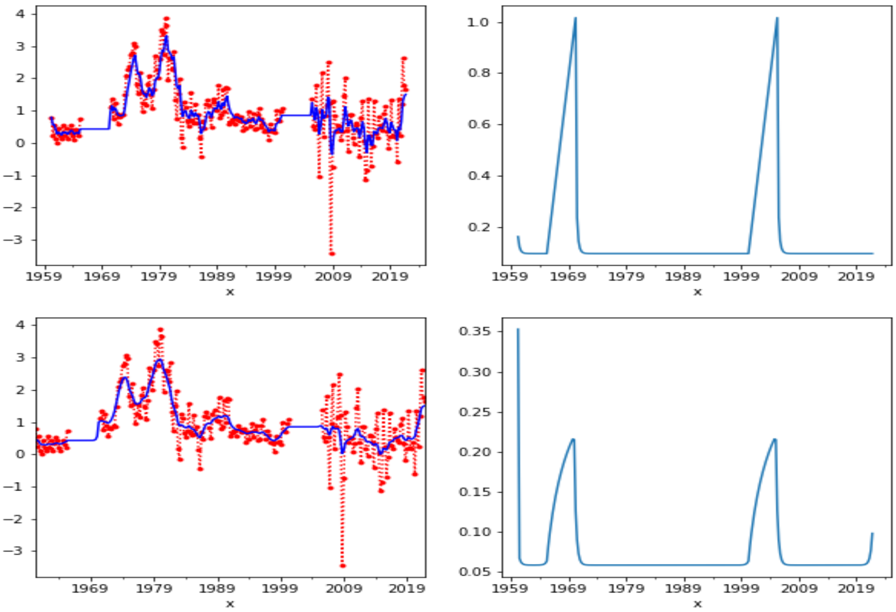

Figure 12: Filtering and smoothing output when observations are missing: (i) data and filtered state a_t (extrapolation); (ii) filtered state variance P_t; (iii) data and smoothed state $\hat{\alpha}_t$ (interpolation); (iv) smoothed state variance V_t.

Our data set is now extended by 40 observations. Treating these observations as missing like we did in the previous graphs allows for the Kalman Filter to be used for the forecasting of the time series. As previously, due to the lack of new observations, the state variance is going to grow to infinity and the kalman gain is going to remain to zero, leaving the state forecast as a straight line.

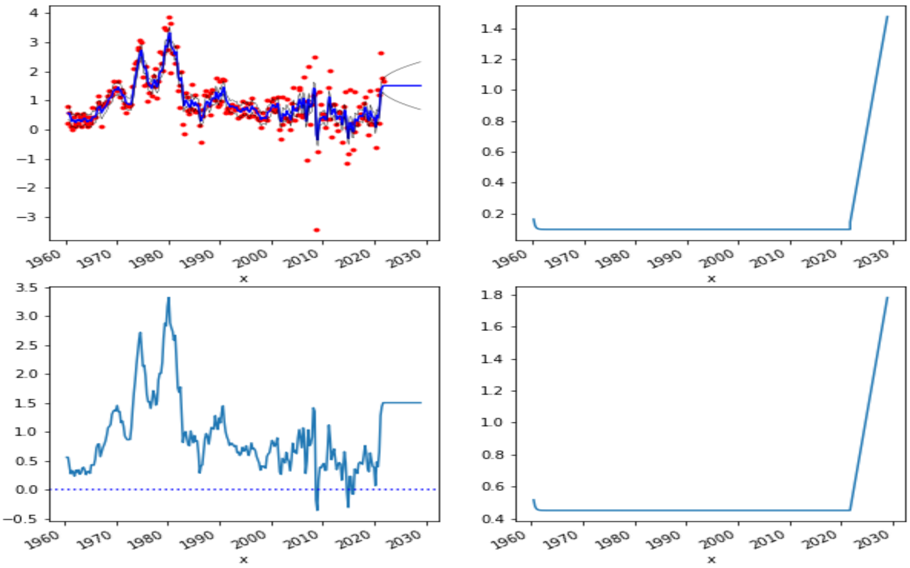

Figure 13: Nile data and output of forecasting: (i) data (dots), state forecast at and 50% confidence intervals; (ii) state variance P_t; (iii) observation forecast $E(y_t|Y_{t-1})$; (iv) observation forecast variance F_t.

One of the assumptions made for the local level model that we used for this data is that the disturbances t and t are normally distributed and serialy independent with constant variances. In figure 13 we show the standardized one-step ahead forecast errors et. Even though we notice the existence of some outliers we can claim that the errors are also normally distributed and serialy independent. The mean is zero with unit variance. From the QQ plot we can see that the sample quantiles do not have significant differences from the theoretical quantiles of a normal distribution. The correlogram states that there is no serial correlation between the errors and since the histogram appears to represent a normal distribution, we can conclude that the assumptions that we made about the local level model that we used are valid for our data.

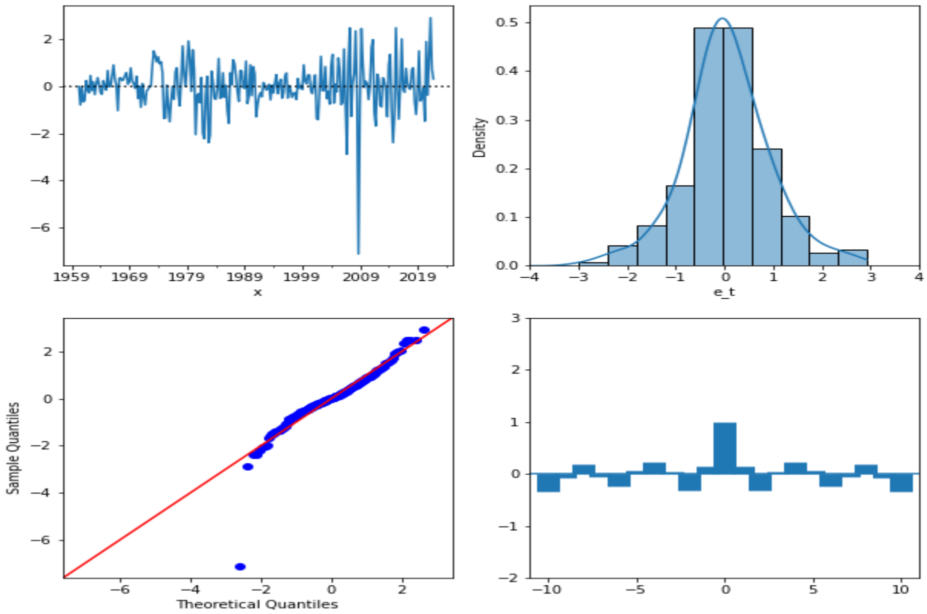

Figure 14: Diagnostic plots for standardised prediction errors: (i) standardised residual; (ii) histogram plus estimated density; (iii) ordered residuals; (iv) correlogram.

The last figure shows the diagnostic plots of the standardised residuals that we refer to as auxiliary residuals. From graph i, we are able to detect possible outliers in the data which in our case is true for the year 2008 and from graph iii we are able to detect possible level brakes which is the case for the years 1973, 1975, 1982. These observations are in line with the data shown in graph 2.1

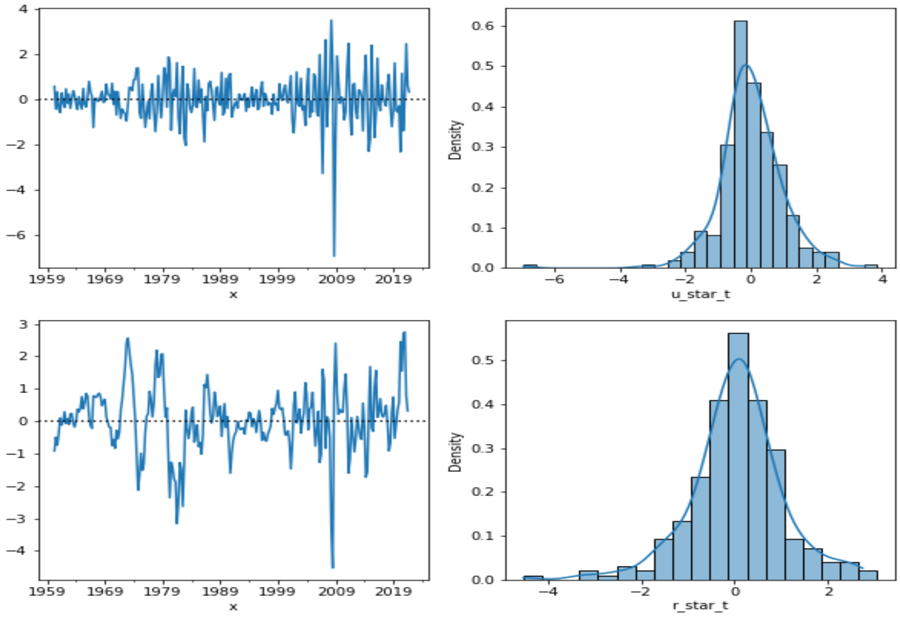

Figure 15: Diagnostic plots for auxiliary residuals: (i) observation residual u_t^* ; (ii) histogram and estimated density for u_t^*; (iii) state residual r_t^*; (iv) histogram and estimated density for r_t^*.

References

Durbin, J., & Koopman, S. J. (2012). *Time series analysis by State Space Methods*. Oxford University Press.