Signal extraction examples

Sonia Mazzi

13/12/2018

Signal extraction using a state-space model

Suppose that we have $y = \{y_t\}_{t=1}^n$ and $y_t = \mu_t + \epsilon_t$, where μ_t is a stochastic process (with linear dynamic) and ϵ_t's are iid with mean zero and constant variance. Note that y_t can be a vector.

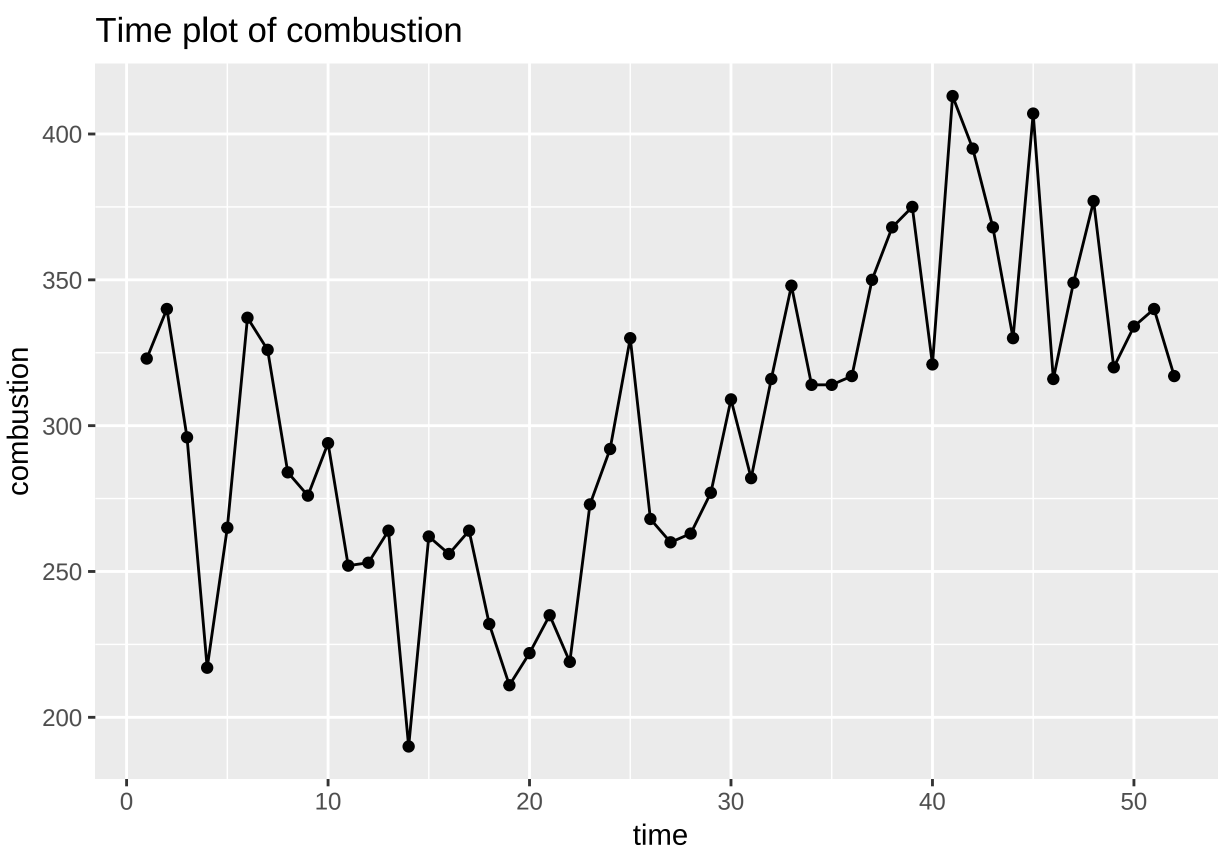
The stochastic specification of μ_t allows to model y in different ways.

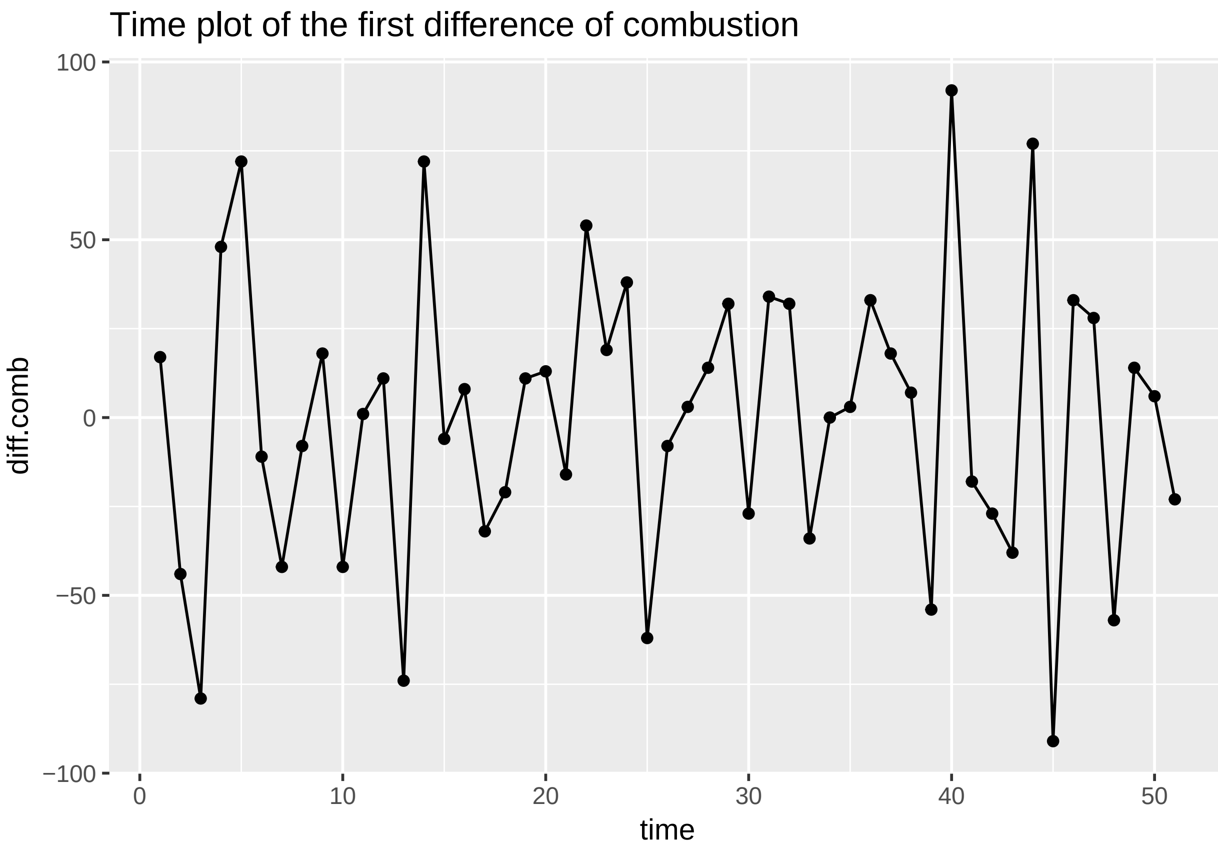
For more details on the state-space model and signal-extraction algorithms see

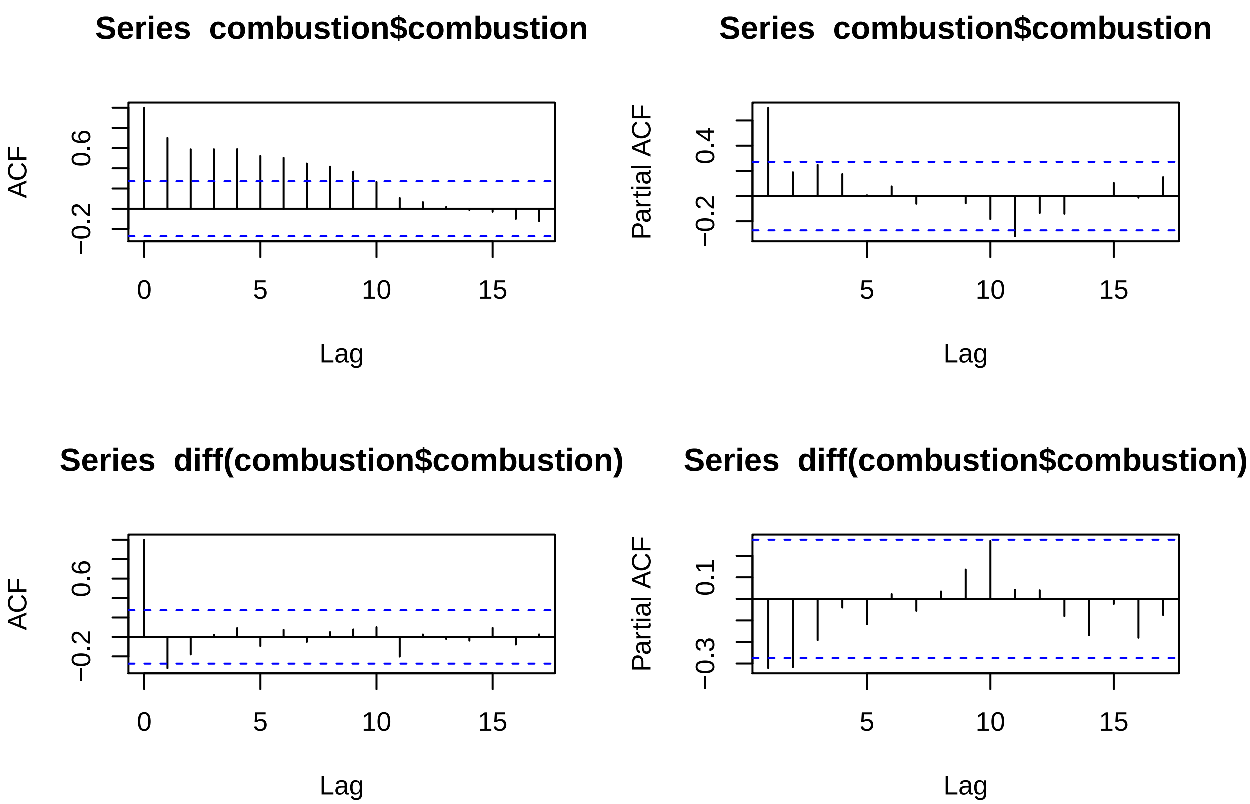
de Jong, P.(1991). The diffuse Kalman filter. *The Annals of Statistics*. Vol 19, No.2, pp 1073 - 1083.

I will illustrate the signal extraction prodedure with three sample quarterly time series: `combustion`, and `image`.

`combustion`







We can observe that the plots above indicate that `combustion` is a 'pure trend' series, with some extreme observations (possibly caused by other processes).

First I will fit a general locally linear model (LLM), which is essentially a random walk with a time varying drift.

In the LLM it is assumed

$y_t = \mu_t + \epsilon_t$, $\mu_{t+1} = \mu_t + d_t + \nu_t$, $d_{t+1} = \delta d_t + \eta_t$

where $\{\epsilon_t\}$, $\{\nu_t\}$, $\{\eta_t\}$ are mutually independent white noise processes with variances σ_ϵ^2, σ_ν^2 and σ_η^2.

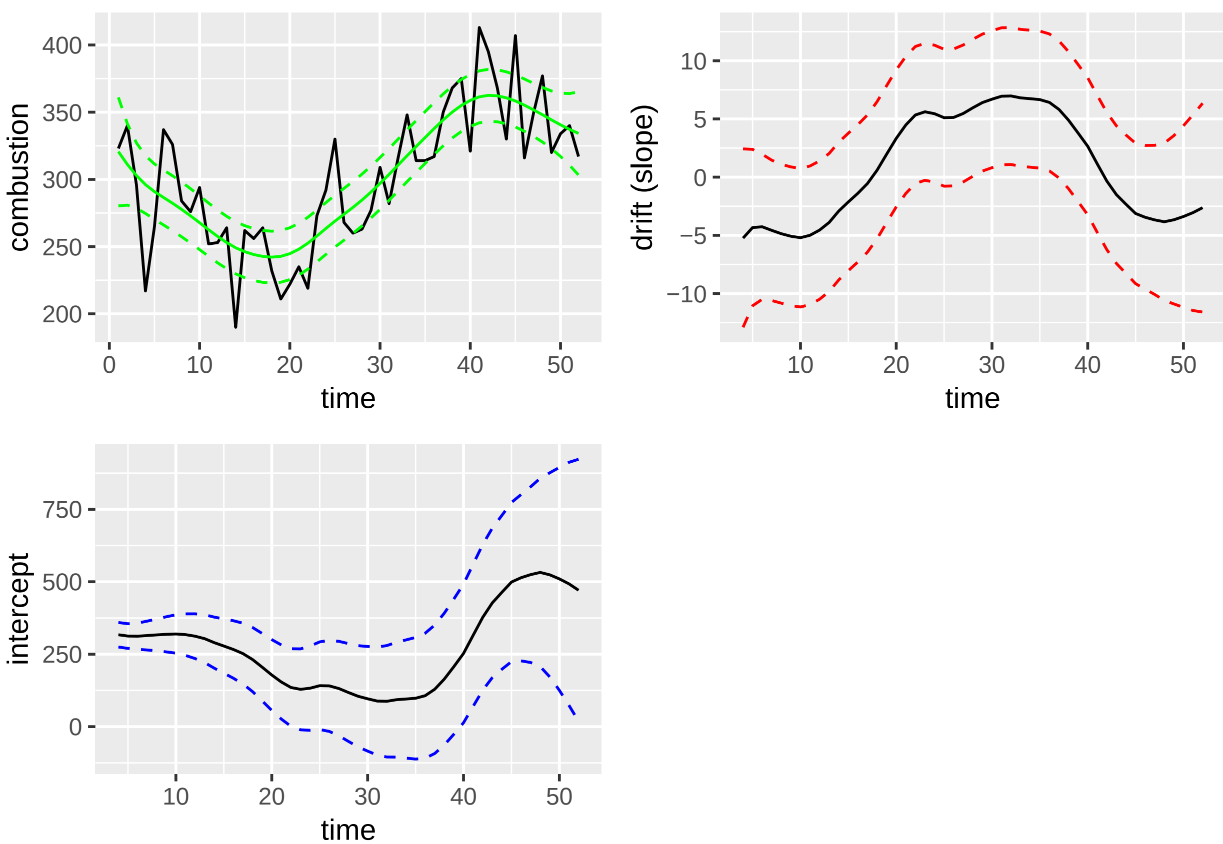
Fitting the LLM for `combustion`

Estimated parameters

Parameter	Estimate
σ_ϵ	29.7853952
σ_ν	0.0000019
σ_η	2.6454416
δ	0.8608259

Plots of predicted signal

```r
grid.arrange(p1, p2, p3, nrow = 2)
```

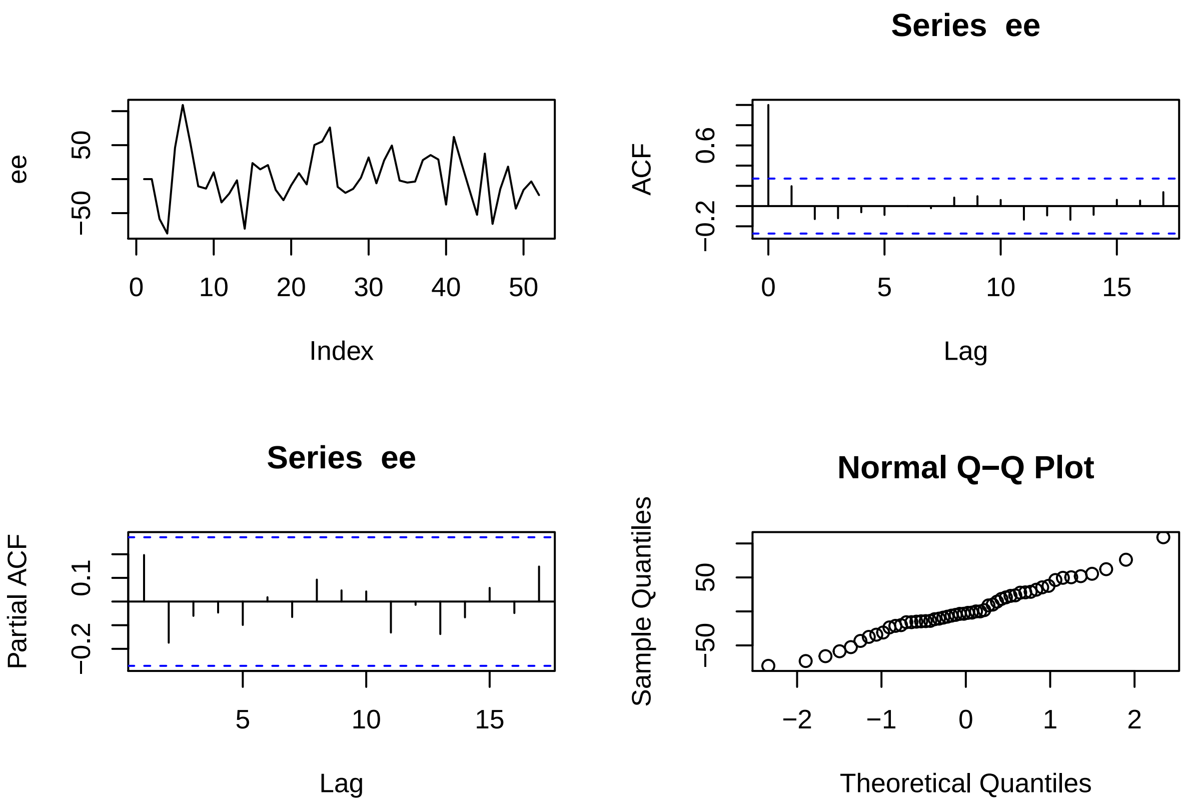


In the above plots, the green solid line has predicted μ_ts. All dotted lines are approximate 95% confidence bands (predicted value$\pm 2\sqrt{mse}$).

The "drift" is d_t and the "intercept" is $\mu_t - td_t$. The drift can be interpreted as the first derivative of the ideal smooth trend. So, when the drift is positive there is growth and when the drift is negative the trend is decreasing.

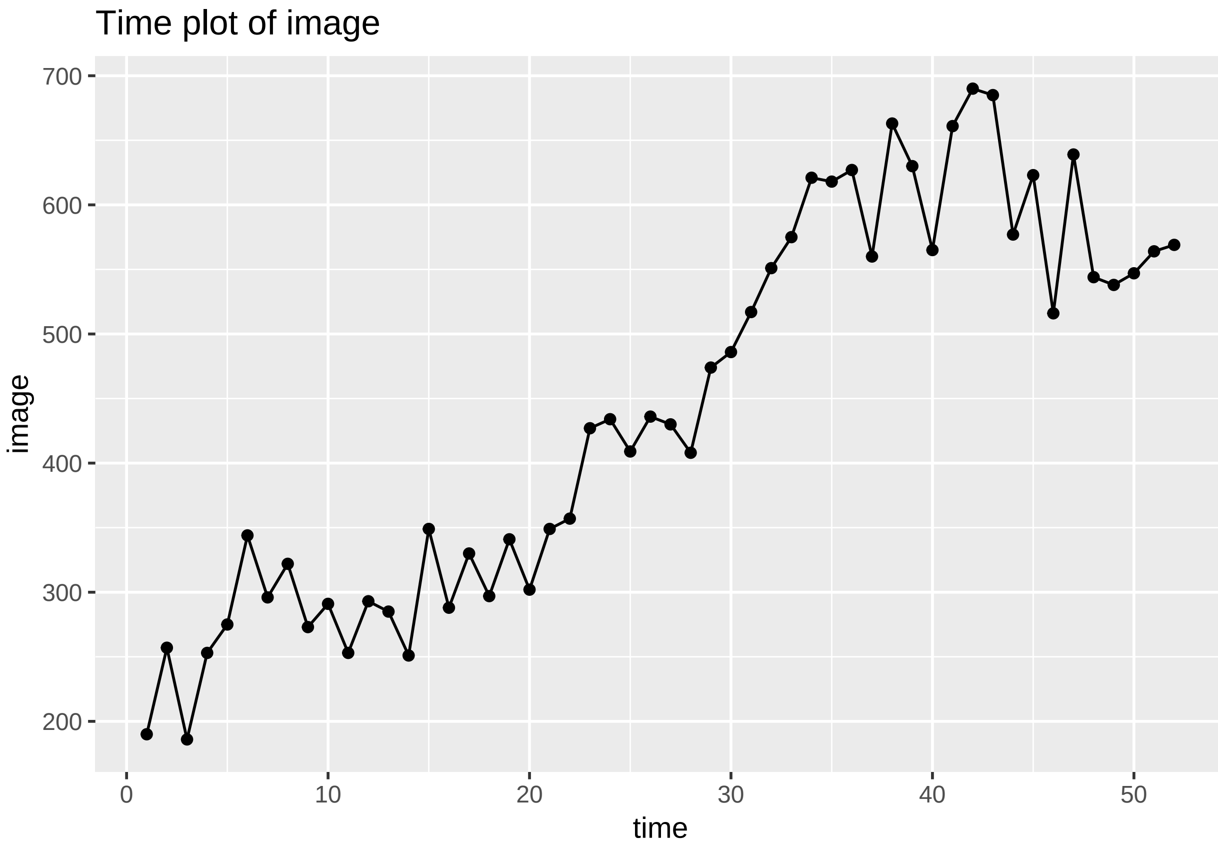
Note that for `combustion` the last few quarters, although the drift is negative, it has started to increase, possibly announcing that the decay in trend is decelarting and a period of growth in trend is coming once the drift crosses the zero line. This is re-inforced by an accompanying decrease in "intercept".

Residual Analysis

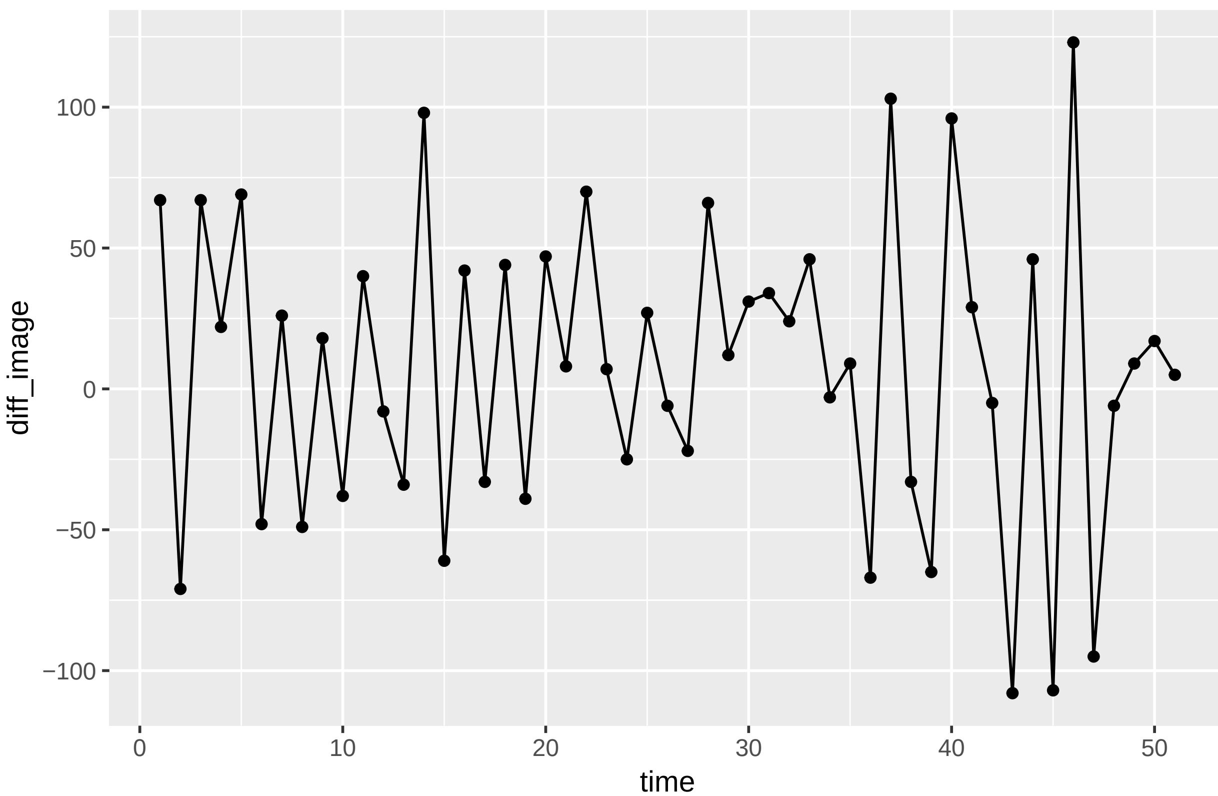


The plots above don't indicate that the model isn't fitting the data well.

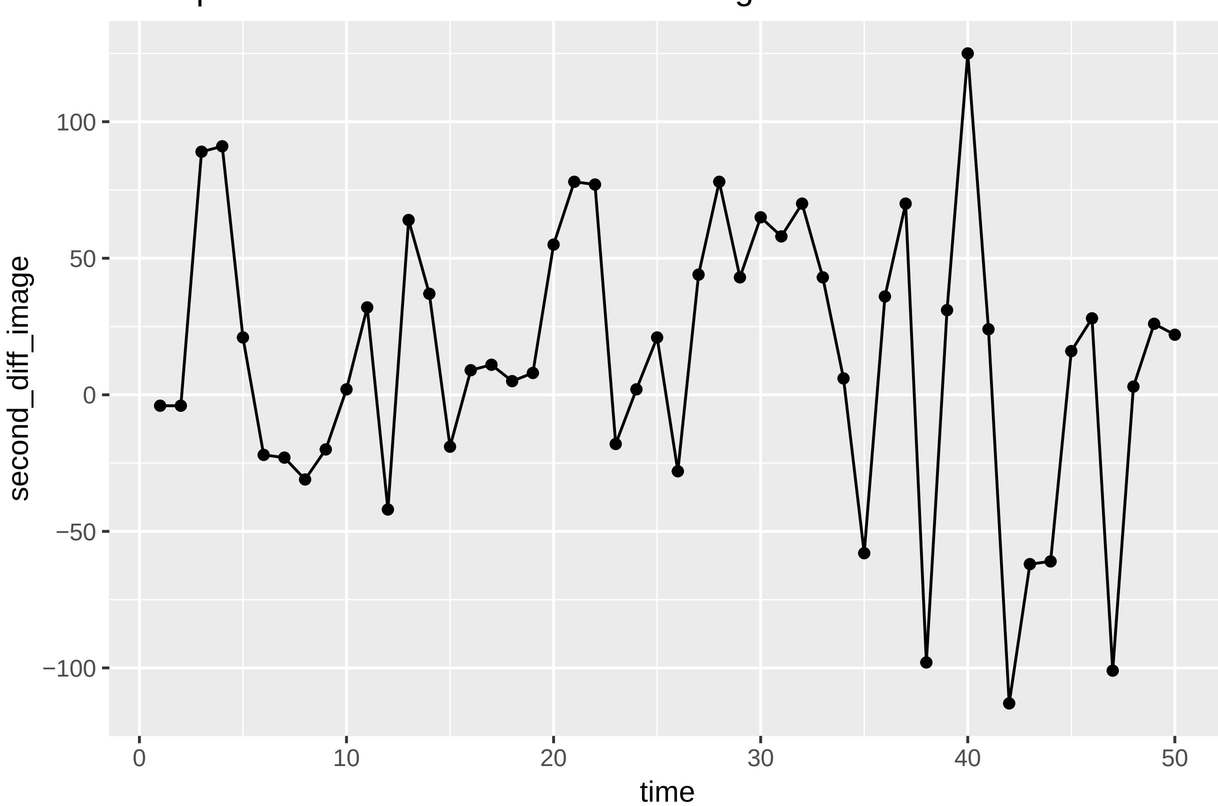
`image`

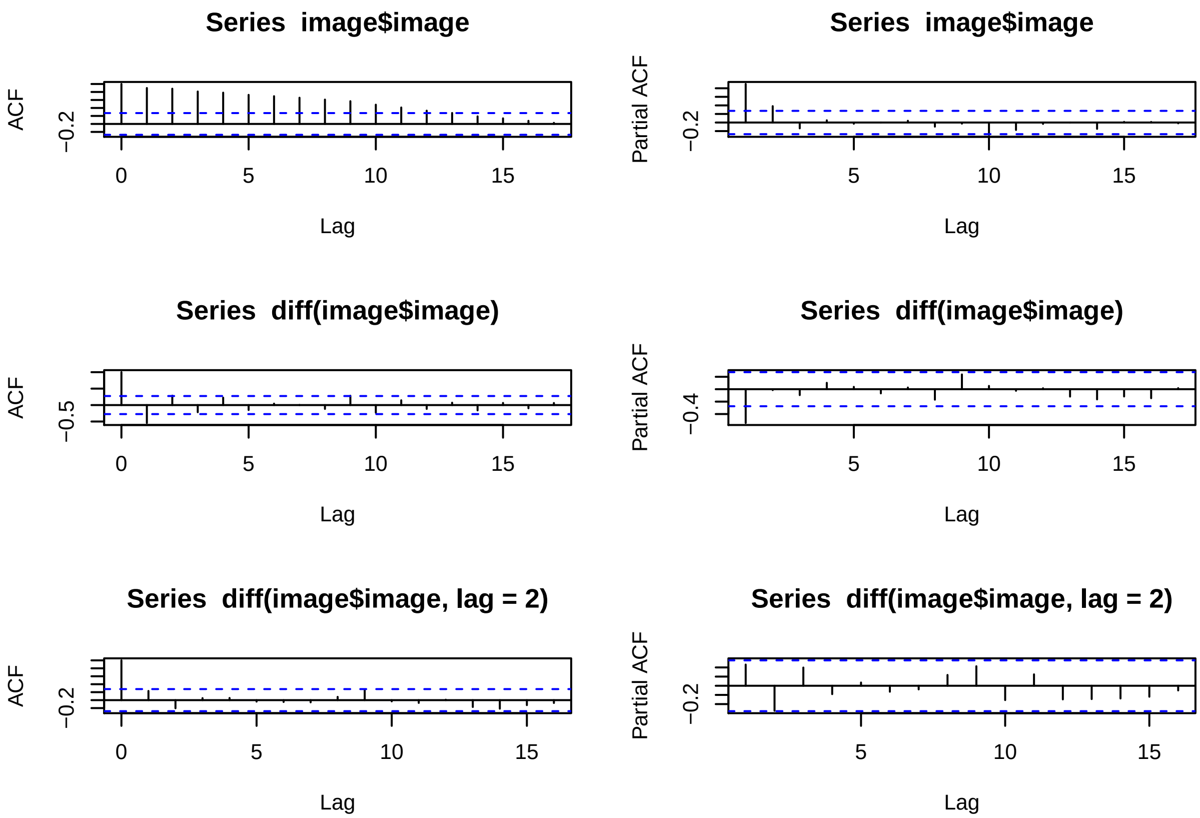

Time plot of image
700
600
500
400
300
200
image
0
10
20
30
40
50
time

Time plot of the first difference of image



Time plot of the second difference of image





We can observe that the plots above indicate that the autocorrelation function of the first difference is significant at lag one. Also, from the time plot of the first difference we clearly see that the series has a change in variance (it is heteroscedastic). So, a LLM would not be a appropriate. Instead a locally quadratic model (LQM) will be fitted.

In the LQM it is assumed

$y_t = \mu_t + \epsilon_t$, $\mu_{t+2} = 2\mu_{t+1} - \mu_t + d_t + \nu_t$, $d_{t+1} = \delta d_t + \eta_t$

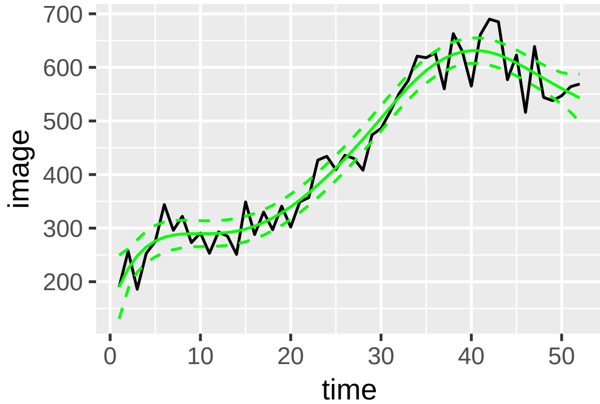
where $\{\epsilon_t\}$, $\{\nu_t\}$, $\{\eta_t\}$ are mutually independent white noise processes with variances σ^2_ϵ, σ^2_ν and σ^2_η.

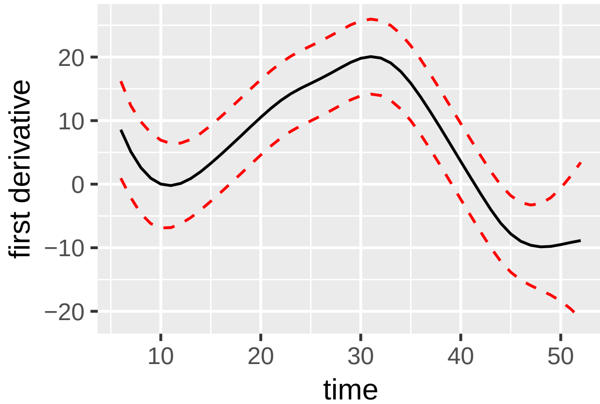
Fitting the LQM for `image`

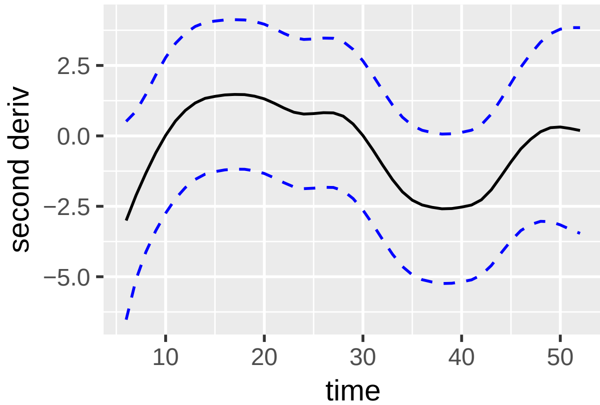
Estimated parameters

Parameter	Estimate
σ_ϵ	36.4279645
σ_ν	0.0000007
σ_η	1.2718623
δ	0.7311585

Plots of predicted signal





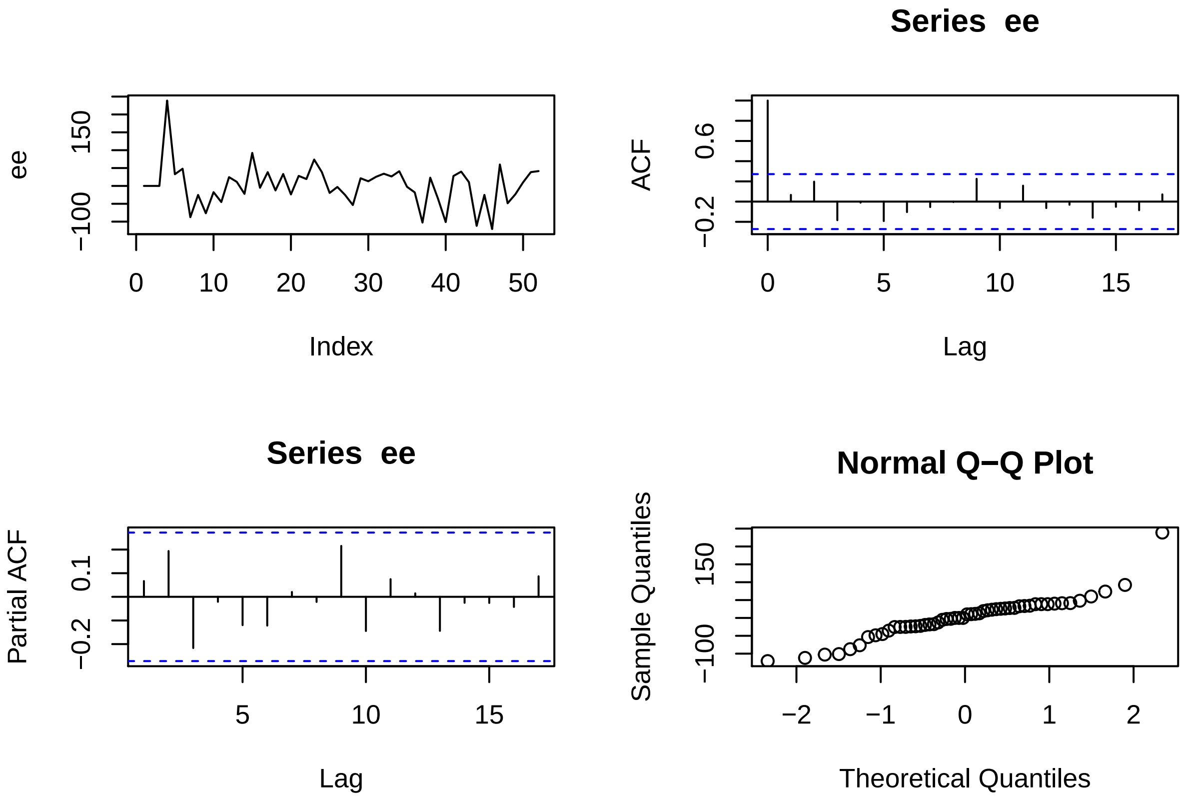


In the above plots, the green solid line has predicted μ_ts. All dotted lines are approximate 95% confidence bands (predicted value$\pm 2\sqrt{mse}$).

The "first derivative" is $\mu_t - \mu_{t-1} - 0.5 d_t$ and the "second derivative" is d_t.

Note that for `image` the last few quarters, although the drift is negative, it has started to increase, possibly announcing that the downfall trend is decelerating and may be a period of growth in trend is coming.

Residual Analysis



The plots above don't indicate that the model isn't fitting the data well.