

CCFs, Differencing, & AR(1) models

Lecture 08

Dr. Colin Rundel

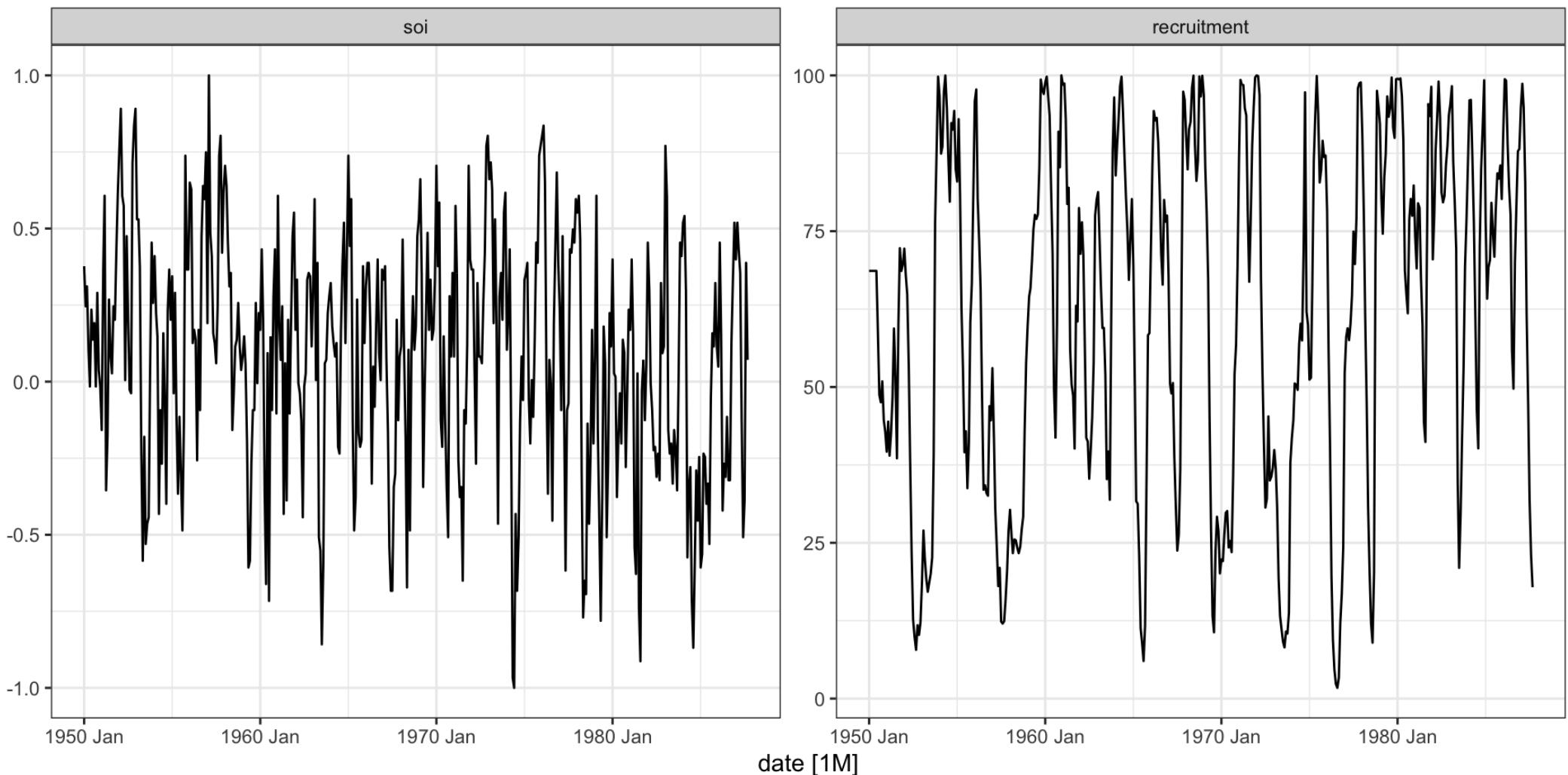
Lagged Predictors and CCFs

Southern Oscillation Index & Recruitment

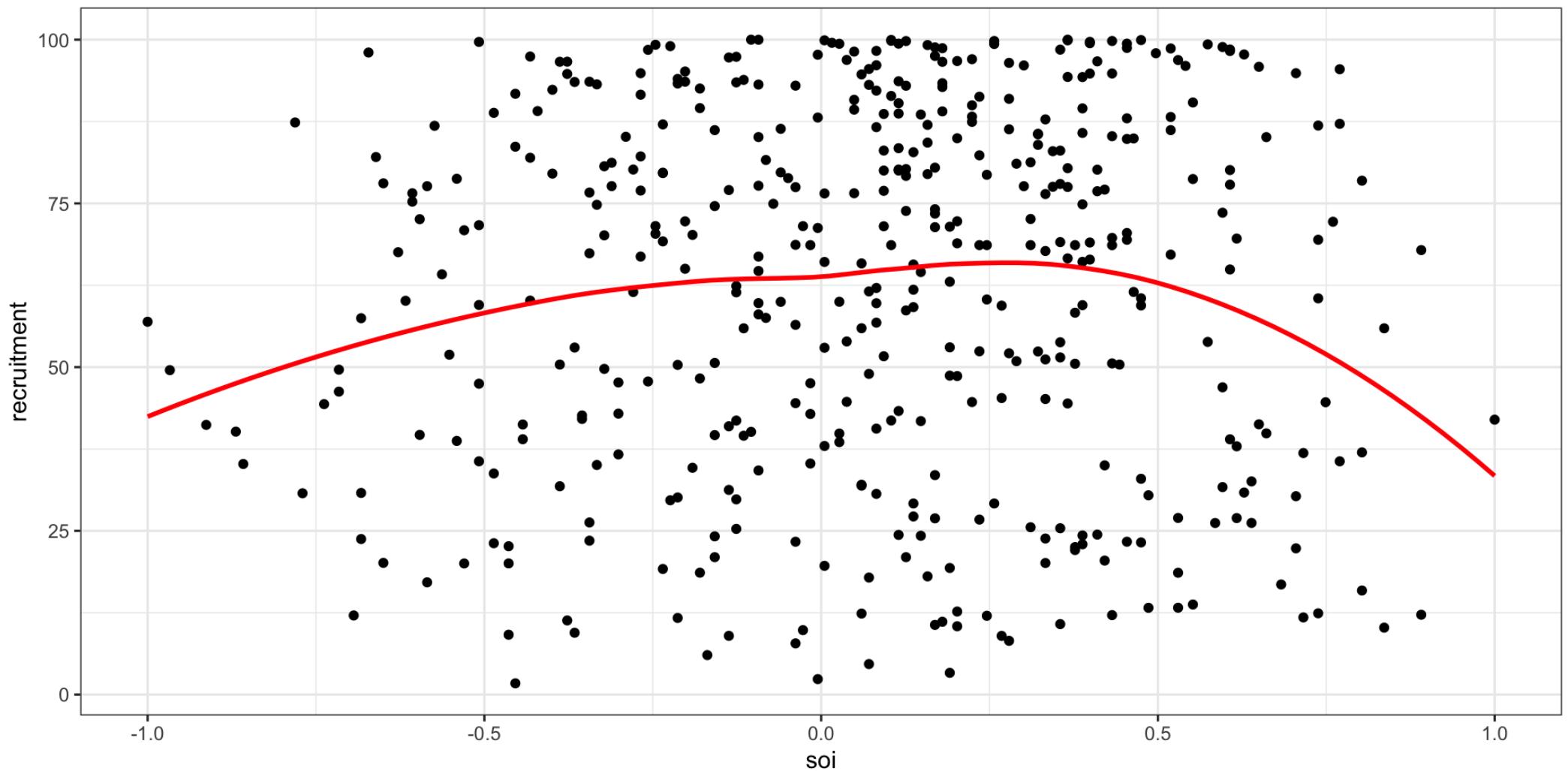
The Southern Oscillation Index (SOI) is an indicator of the development and intensity of El Niño (negative SOI) or La Niña (positive SOI) events in the Pacific Ocean. These data also included the estimate of “recruitment”, which indicate fish population sizes in the southern hemisphere.

```
# A tsibble: 453 x 3 [1M]
  date      soi recruitment
  <mth>    <dbl>     <dbl>
1 1950 Jan  0.377     68.6
2 1950 Feb  0.246     68.6
3 1950 Mar  0.311     68.6
4 1950 Apr  0.104     68.6
5 1950 May -0.016     68.6
6 1950 Jun  0.235     68.6
7 1950 Jul  0.137     59.2
8 1950 Aug  0.191     48.7
9 1950 Sep -0.016     47.5
10 1950 Oct  0.29      50.9
# i 443 more rows
```

Time series

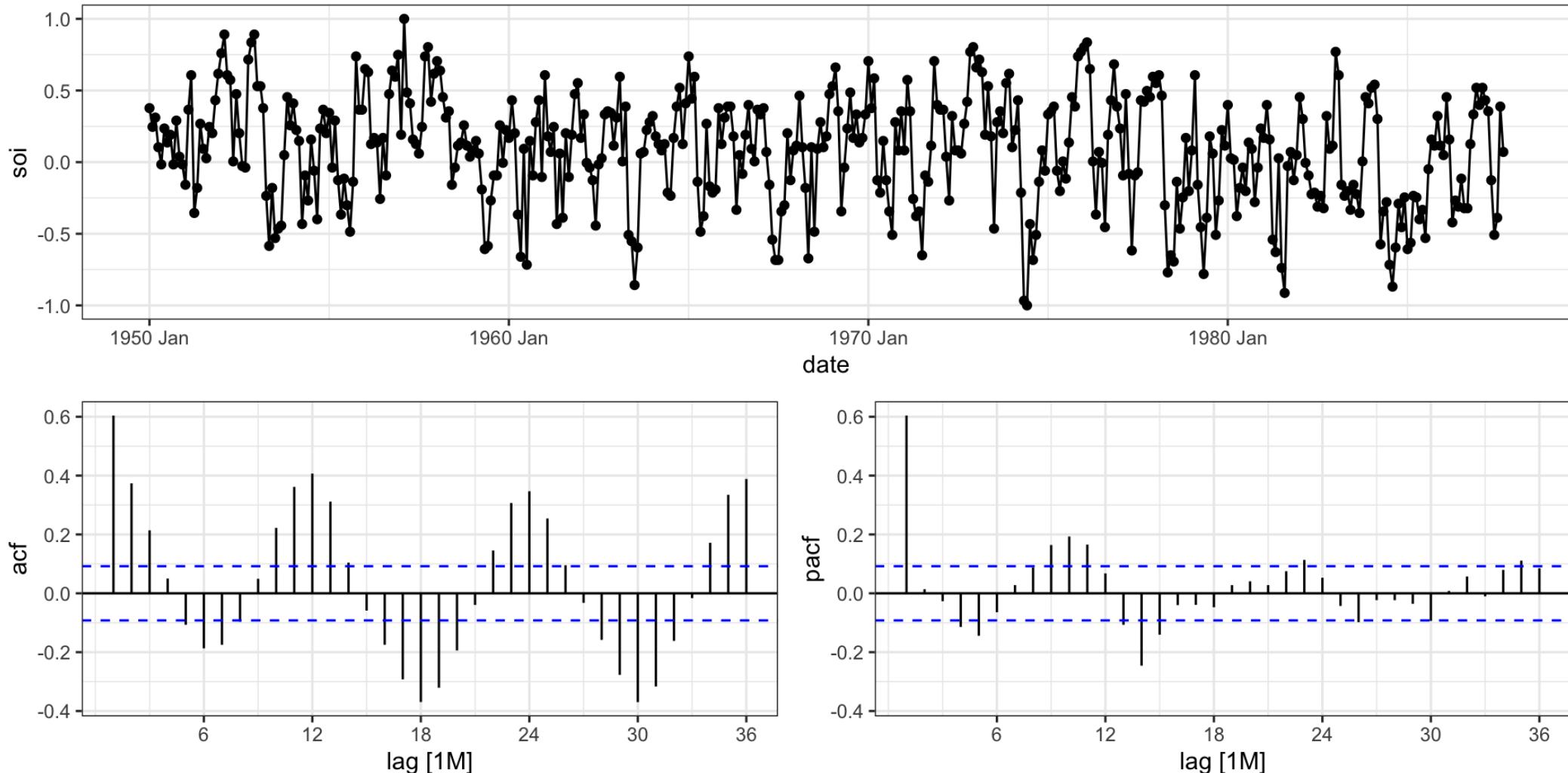


Relationship?



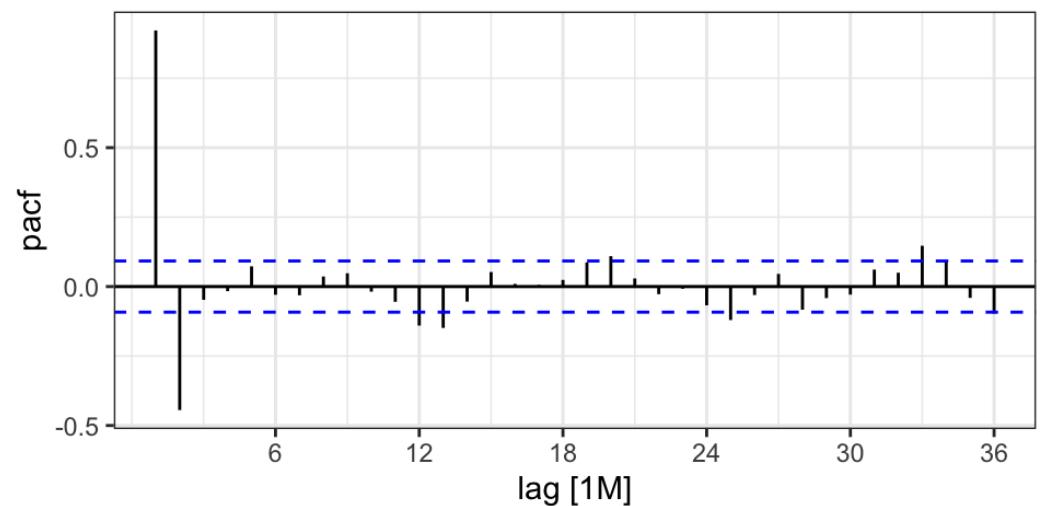
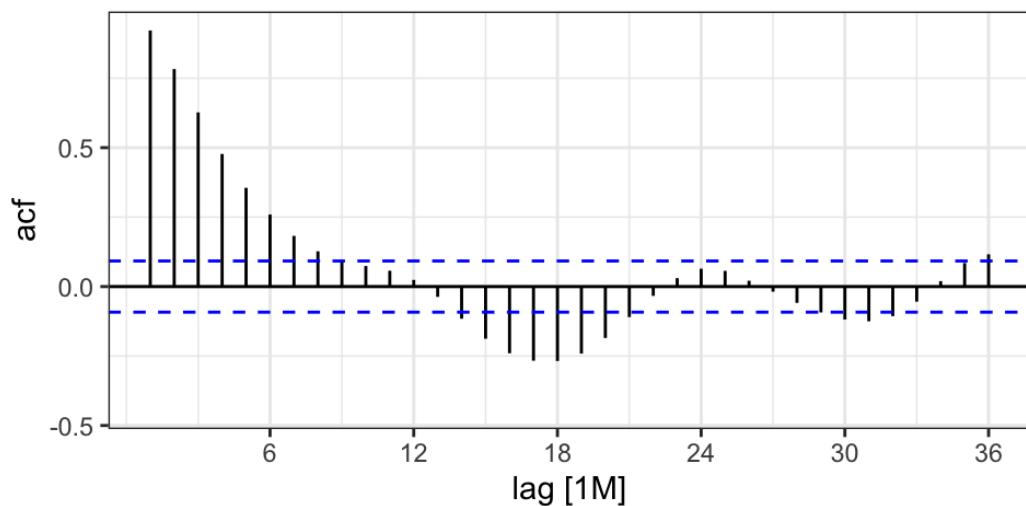
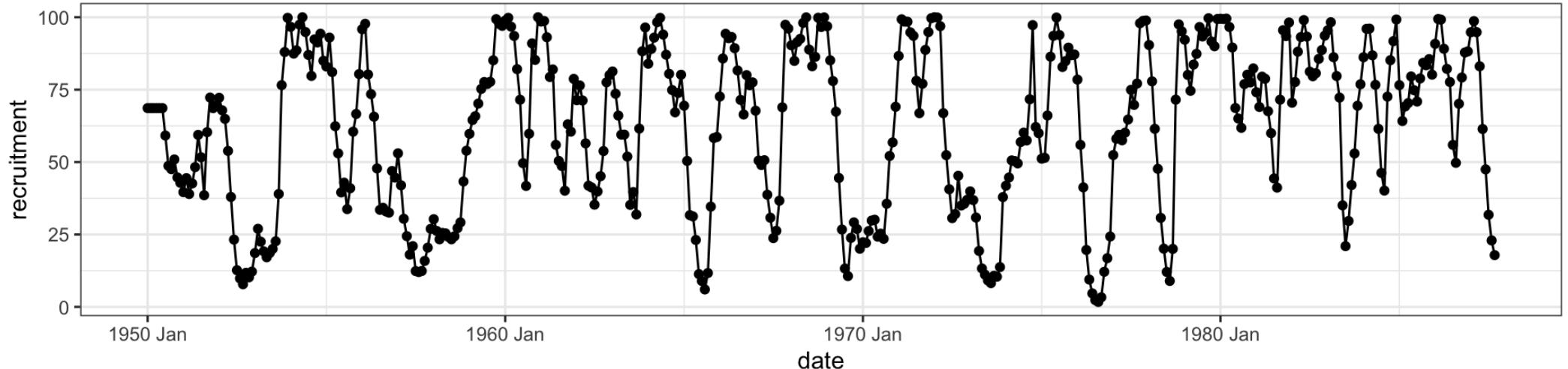
soi - ACF & PACF

```
1 feasts::gg_tsdisplay(fish, y=soi, lag_max=36, plot_type = "partial")
```



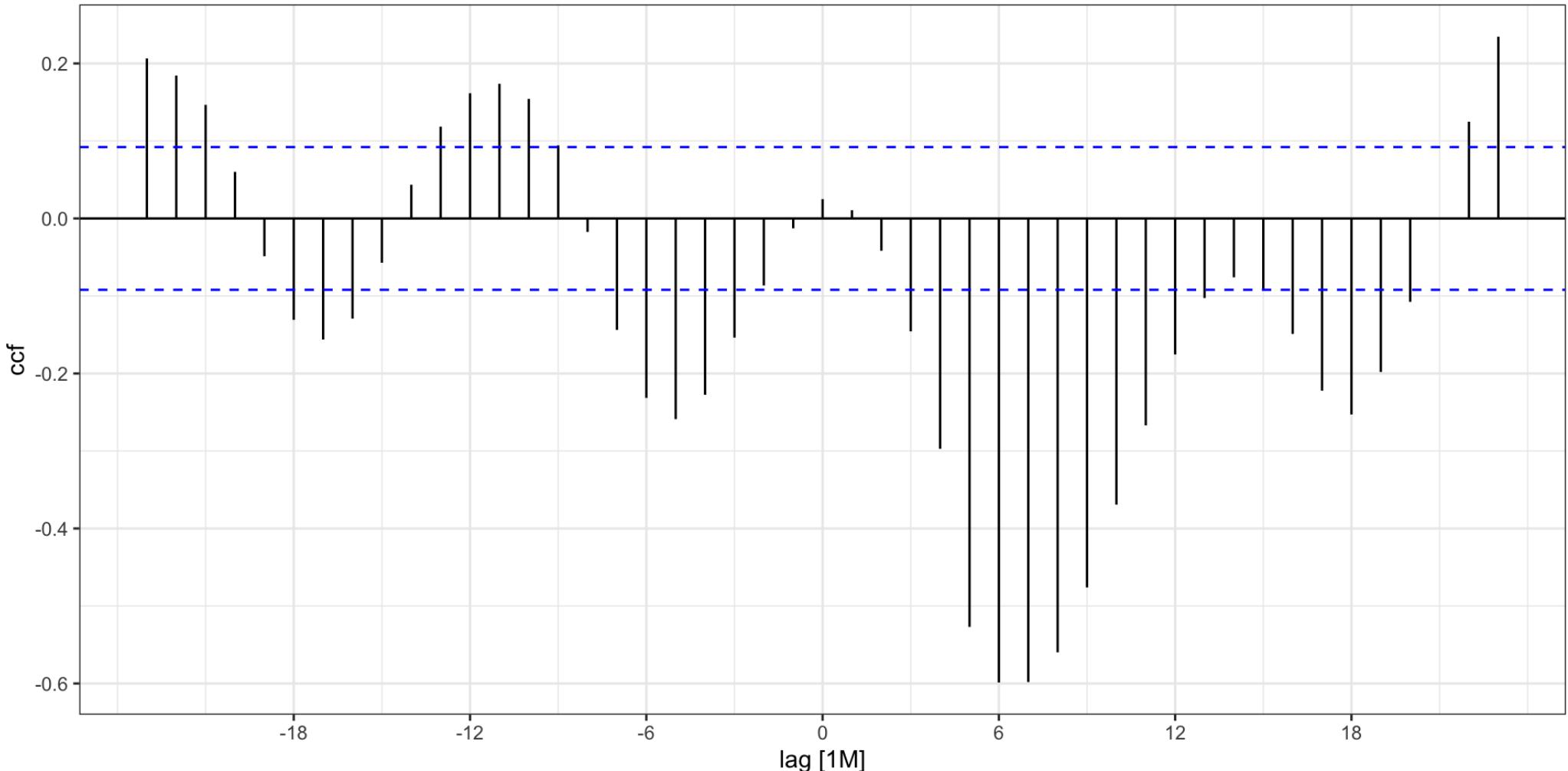
recruitment - ACF & PACF

```
1 feasts::gg_tsdisplay(fish, y=recruitment, lag_max=36, plot_type = "parti
```

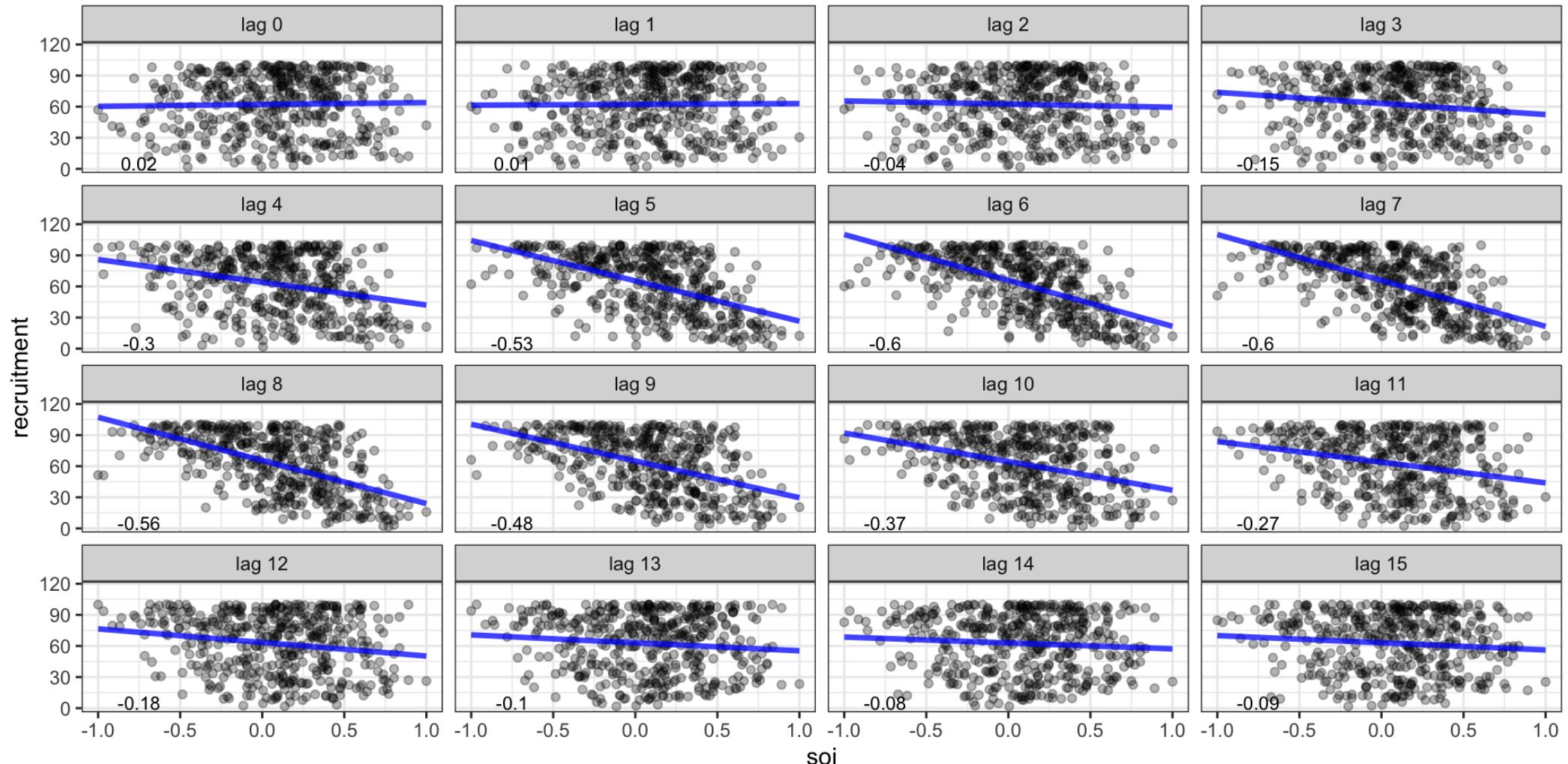


Cross correlation function

```
1 feasts:::CCF(fish, y = soi, x = recruitment) |>  
2 autoplot()
```



Cross correlation function - Scatter plots



The CCF gave us negative lags, why are we not considering them here?

Model

```
1 model1 = lm(recruitment~lag(soi,6), data=fish)
2 model2 = lm(recruitment~lag(soi,6)+lag(soi,7), data=fish)
3 model3 = lm(recruitment~lag(soi,5)+lag(soi,6)+lag(soi,7)+lag(soi,8), data=fish)
```

```
1 summary(model3)
```

Call:

```
lm(formula = recruitment ~ lag(soi, 5) + lag(soi, 6) + lag(soi,
7) + lag(soi, 8), data = fish)
```

Residuals:

Min	1Q	Median	3Q	Max
-72.409	-13.527	0.191	12.851	46.040

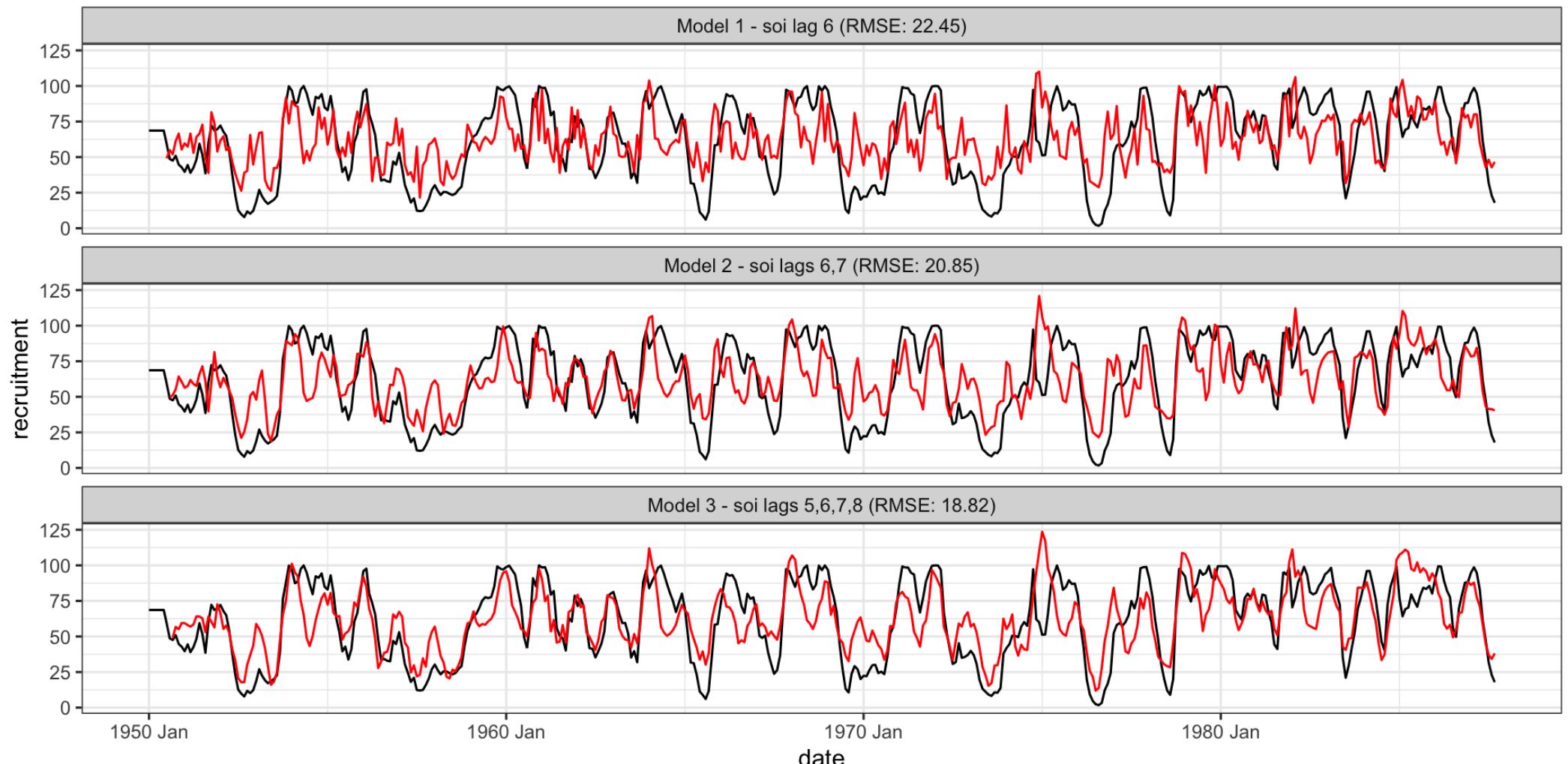
Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	67.9438	0.9306	73.007	< 2e-16 ***
lag(soi, 5)	-19.1502	2.9508	-6.490	2.32e-10 ***
lag(soi, 6)	-15.6894	3.4334	-4.570	6.36e-06 ***
lag(soi, 7)	-13.4041	3.4332	-3.904	0.000109 ***
lag(soi, 8)	-23.1480	2.9530	-7.839	3.46e-14 ***

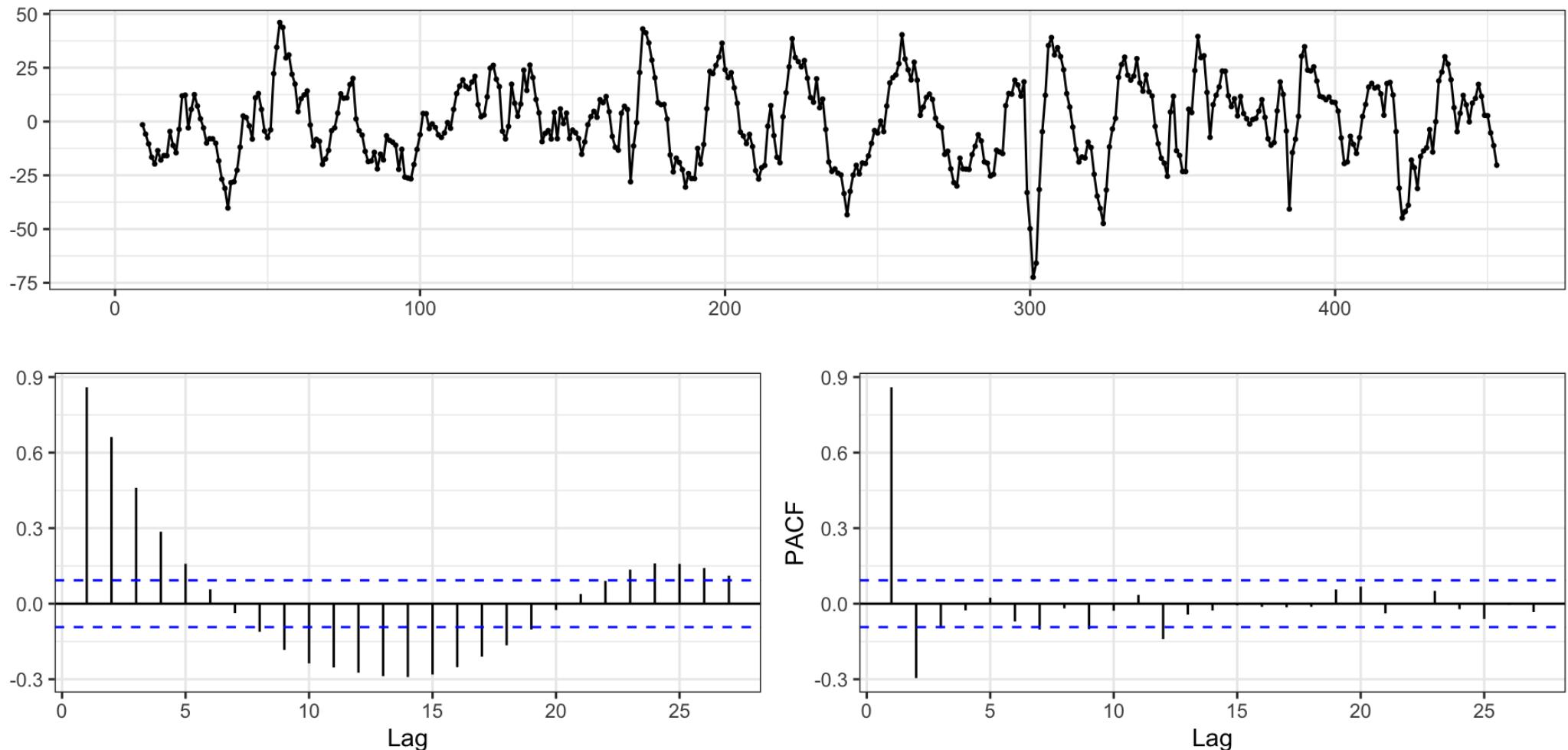
Signif. codes:	0 ****	0.001 **	0.01 *	0.05 .
	'	'	'	'

Residual standard error: 18.93 on 440 degrees of freedom
(8 observations deleted due to missingness)

Prediction



Residual ACF - Model 3



Autoregressive model 4

```
1 model4 = lm(  
2   recruitment~lag(recruitment,1) + lag(recruitment,2) +  
3           lag(soi,5)+lag(soi,6)+lag(soi,7)+lag(soi,8),  
4   data=fish  
5 )  
6 summary(model4)
```

Call:

```
lm(formula = recruitment ~ lag(recruitment, 1) + lag(recruitment,  
 2) + lag(soi, 5) + lag(soi, 6) + lag(soi, 7) + lag(soi, 8),  
 data = fish)
```

Residuals:

Min	1Q	Median	3Q	Max
-51.996	-2.892	0.103	3.117	28.579

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	10.25007	1.17081	8.755	< 2e-16 ***
lag(recruitment, 1)	1.25301	0.04312	29.061	< 2e-16 ***
lag(recruitment, 2)	-0.39961	0.03998	-9.995	< 2e-16 ***
lag(soi, 5)	-20.76309	1.09906	-18.892	< 2e-16 ***
lag(soi, 6)	9.71918	1.56265	6.220	1.16e-09 ***
lag(soi, 7)	-1.01131	1.31912	Sta 344/644 0 Fall 2023	0.4437

```
lag(soi, 8)      -2.29814     1.20730   -1.904    0.0576 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Autoregressive model 5

```
1 model5 = lm(  
2   recruitment~lag(recruitment,1) + lag(recruitment,2) +  
3           lag(soi,5) + lag(soi,6),  
4   data=fish  
5 )  
6 summary(model5)
```

Call:

```
lm(formula = recruitment ~ lag(recruitment, 1) + lag(recruitment,  
  2) + lag(soi, 5) + lag(soi, 6), data = fish)
```

Residuals:

Min	1Q	Median	3Q	Max
-53.786	-2.999	-0.035	3.031	27.669

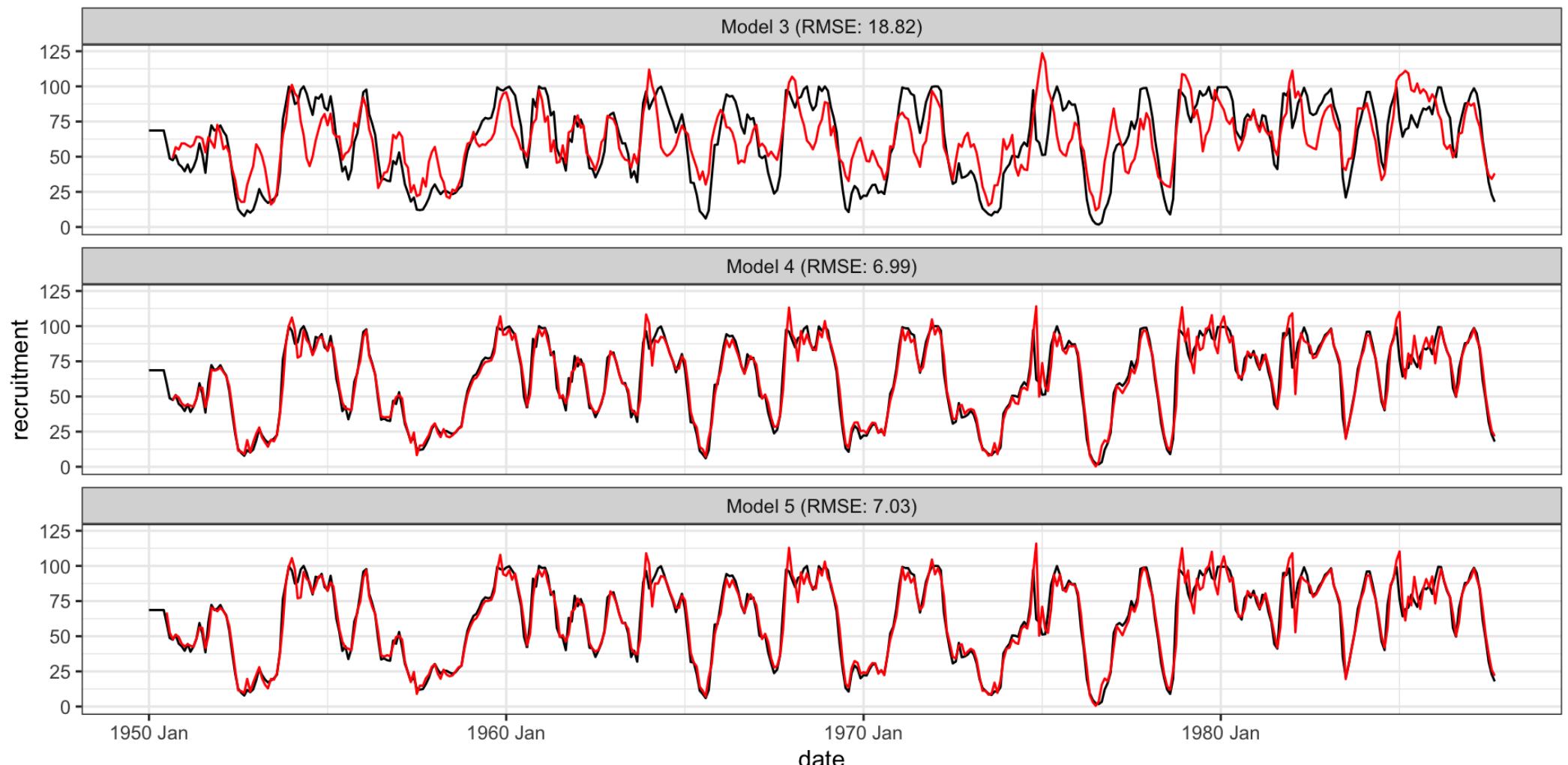
Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	8.78498	1.00171	8.770	< 2e-16 ***
lag(recruitment, 1)	1.24575	0.04314	28.879	< 2e-16 ***
lag(recruitment, 2)	-0.37193	0.03846	-9.670	< 2e-16 ***
lag(soi, 5)	-20.83776	1.10208	-18.908	< 2e-16 ***
lag(soi, 6)	8.55600	1.43146	5.977	4.68e-09 ***

Signif. codes:	0 '***'	0.001 '**'	0.01 '*'	Sta 94/05/44 - Fall 2023 1

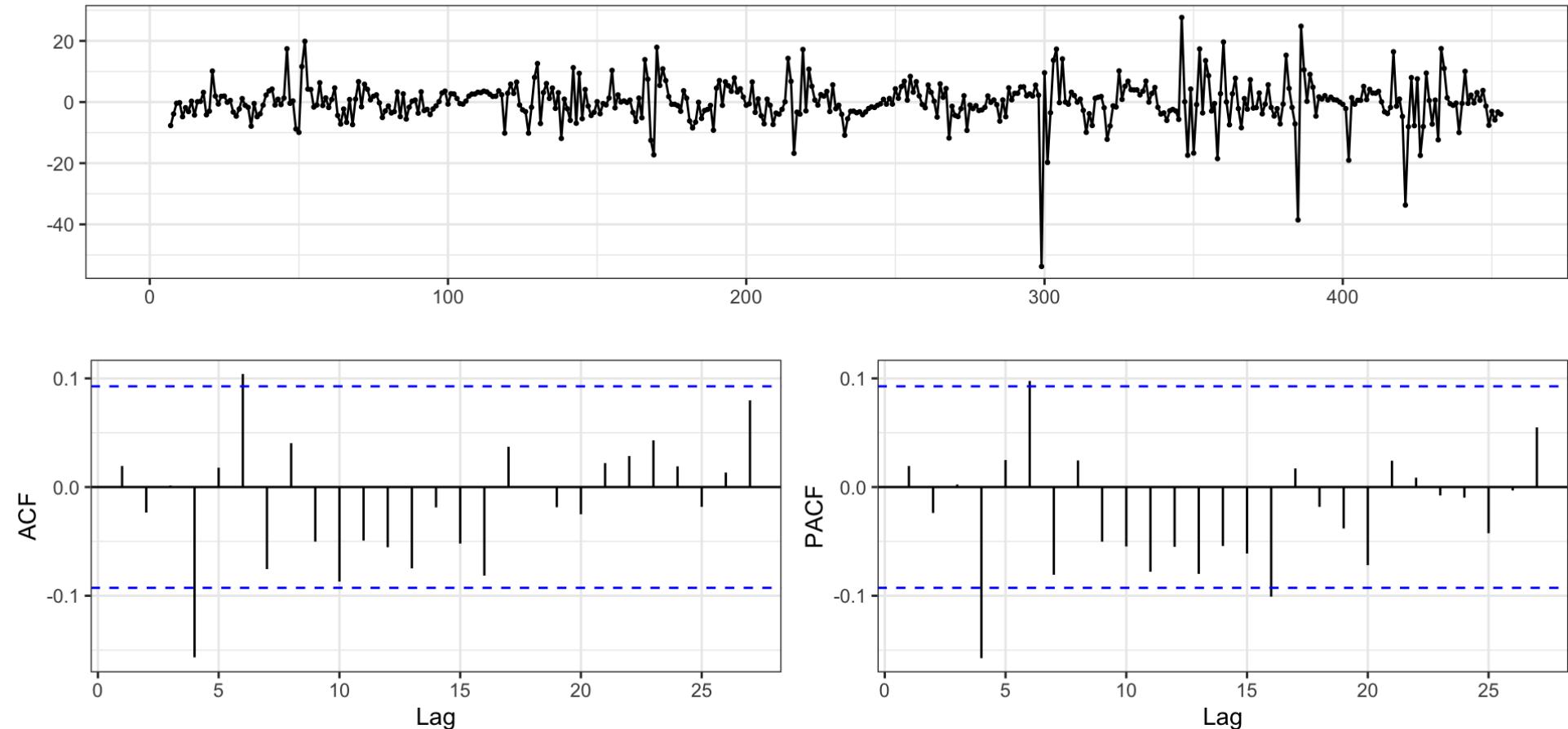
Residual standard error: 7.069 on 442 degrees of freedom
(6 observations deleted due to missinaness)

Prediction



Residual ACF - Model 5

```
1 broom:::augment(model5, newdata=fish) |>  
2 pull(.resid) |>  
3 forecast::ggtsdisplay()
```



Non-stationarity

Non-stationary models

All happy families are alike; each unhappy family is unhappy in its own way.

- Tolstoy, Anna Karenina

This applies to time series models as well, just replace happy family with stationary model.

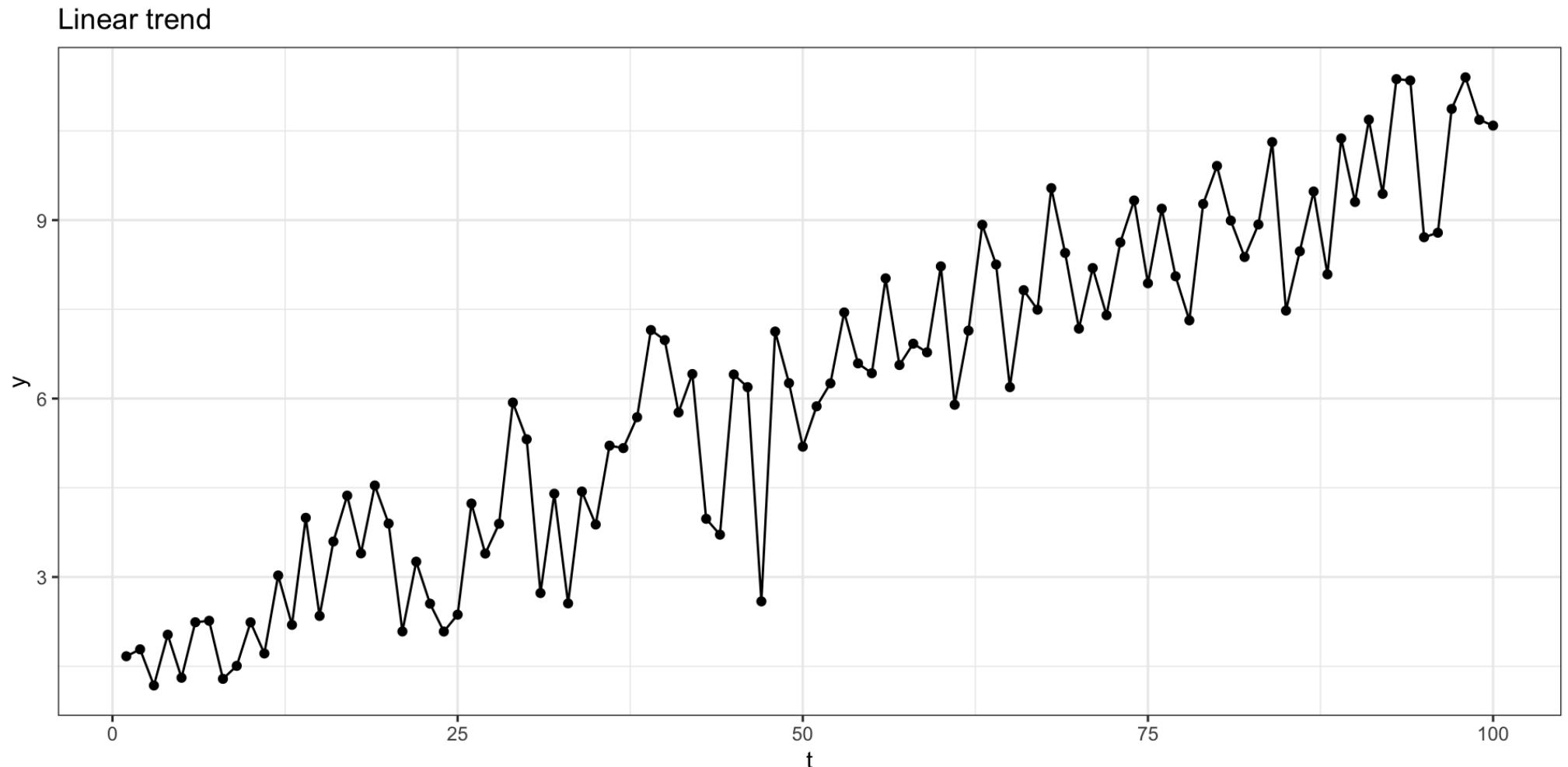
A simple example of a non-stationary time series is a trend stationary model

$$y_t = \mu(t) + w_t$$

where $\mu(t)$ denotes a time dependent trend and w_t is a white noise (stationary) process.

Linear trend model

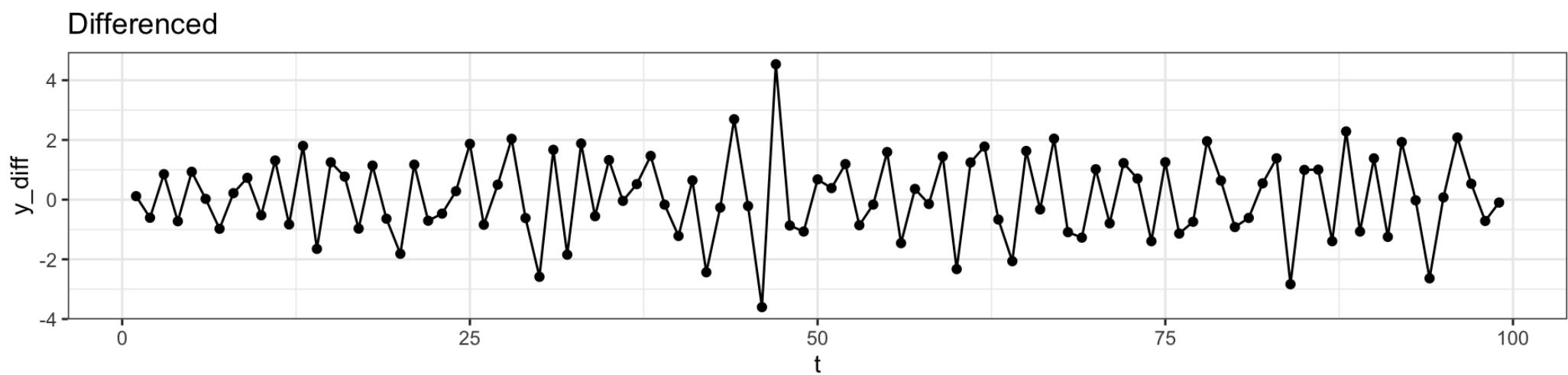
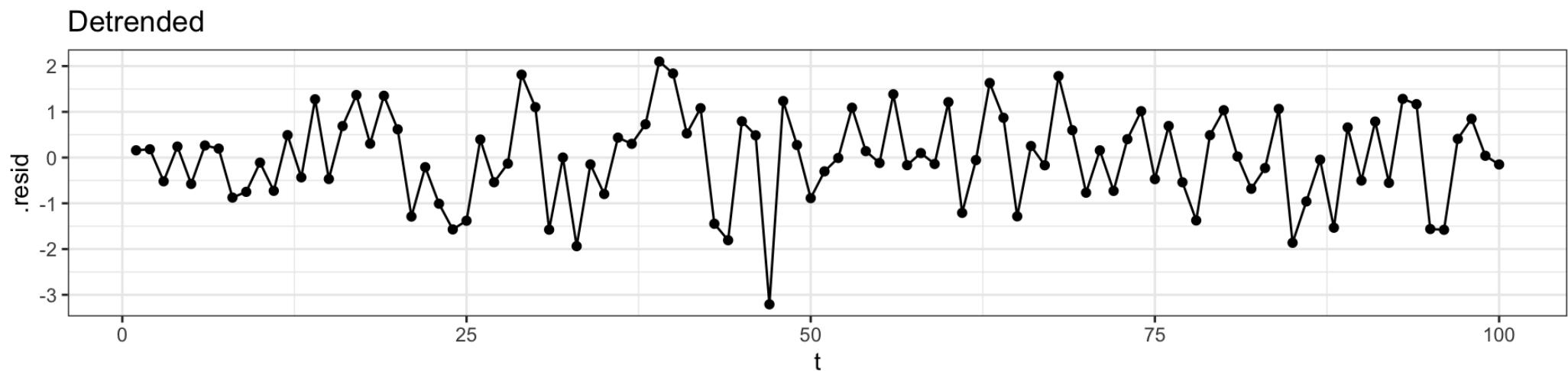
Lets imagine a simple model where $y_t = \delta + \beta t + x_t$ where δ and β are constants and x_t is a stationary process.



Differencing

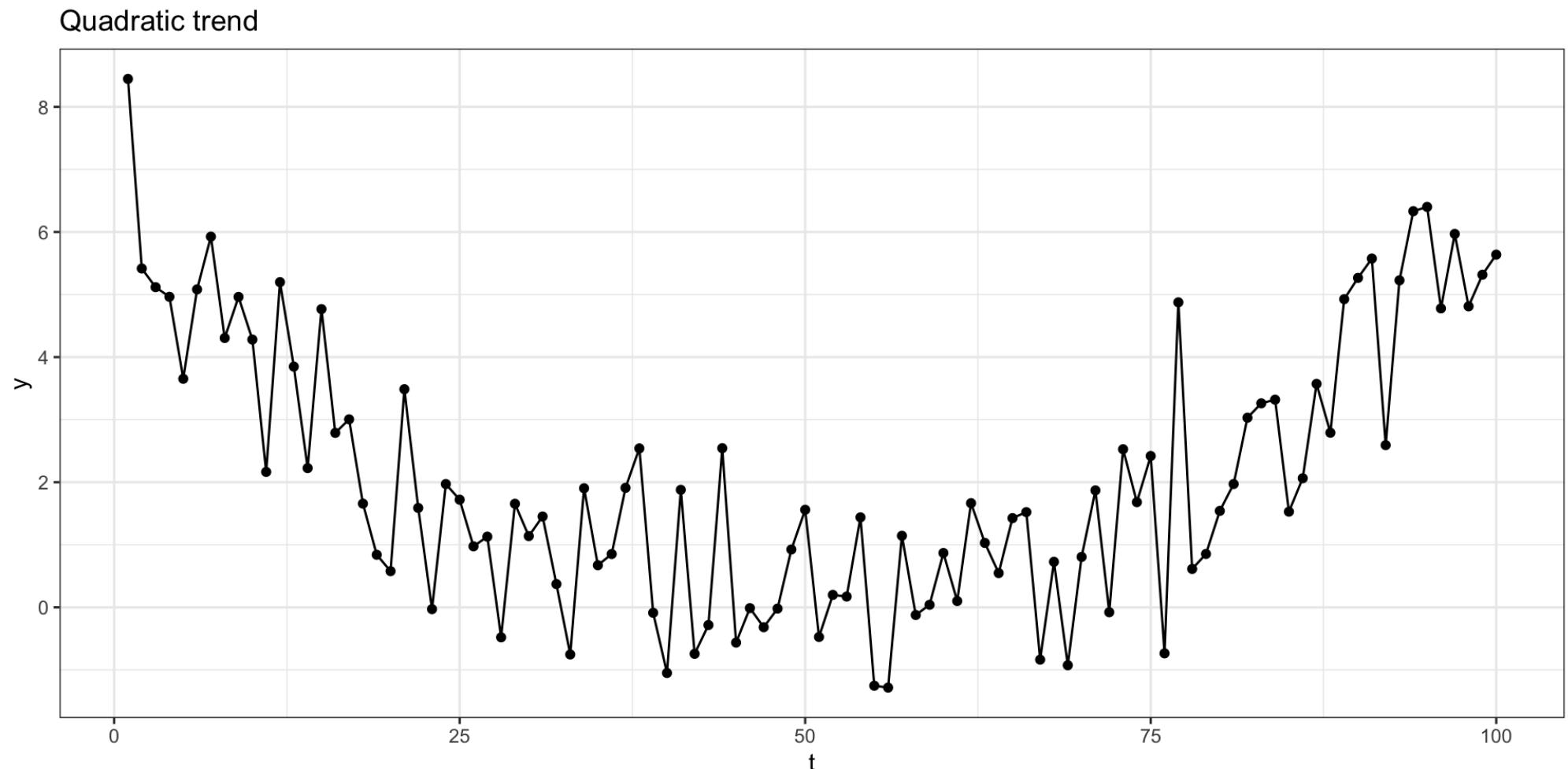
An simple approach to remove trend is to difference your response variable, specifically examine $d_t = y_t - y_{t-1}$ instead of y_t .

Detrending vs Differencing



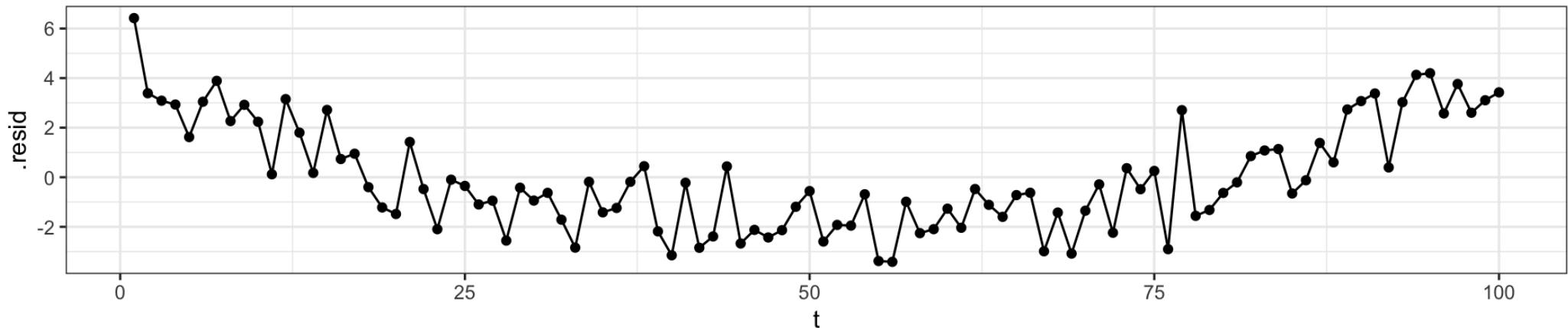
Quadratic trend model

Lets imagine another simple model where $y_t = \delta + \beta t + \gamma t^2 + x_t$ where δ , β , and γ are constants and x_t is a stationary process.

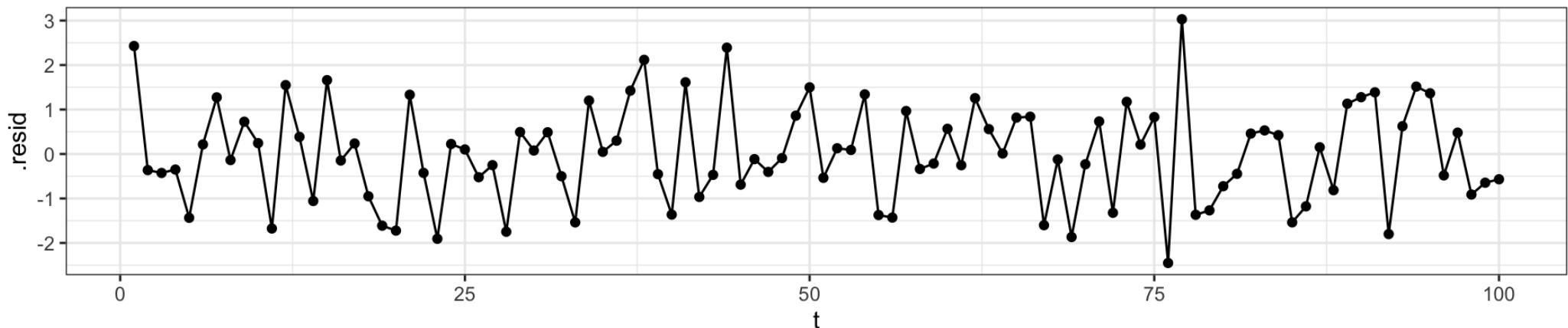


Detrending

Detrended - Linear



Detrended - Quadratic

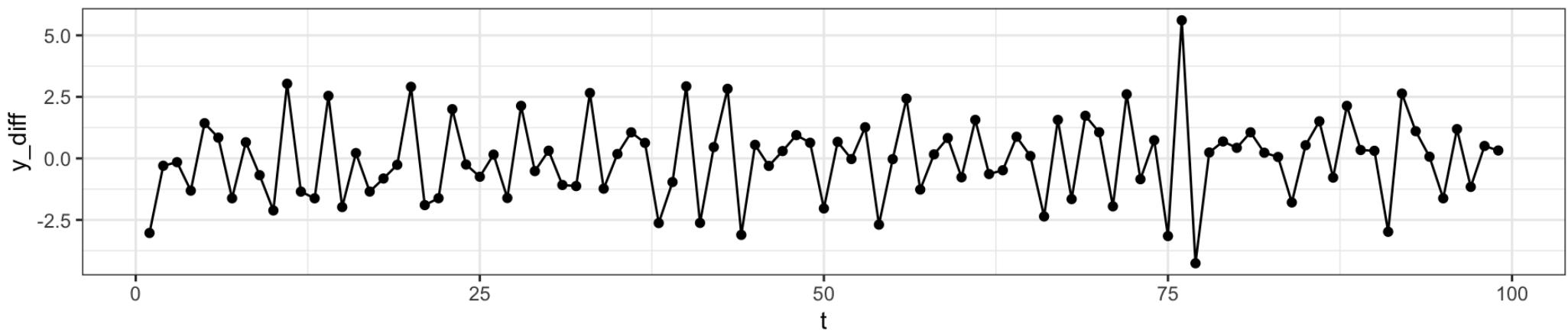


2nd order differencing

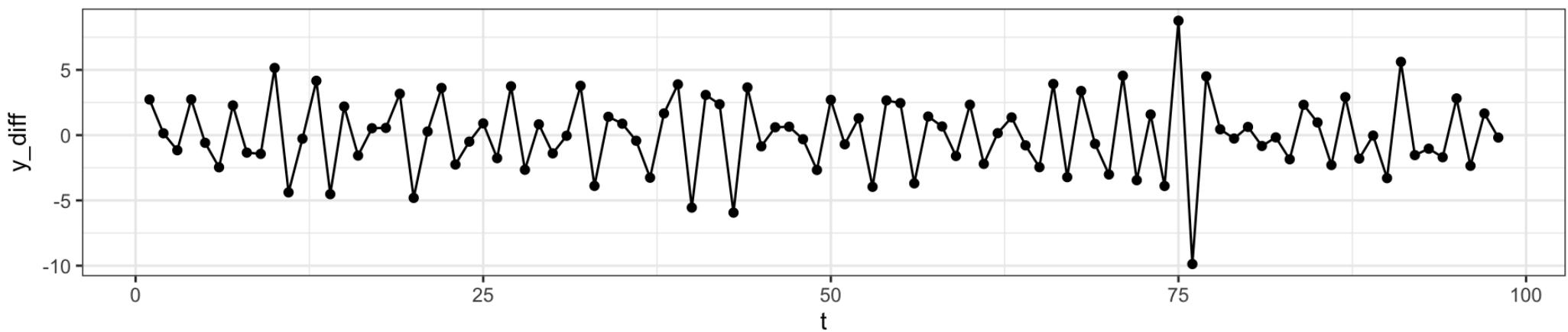
Let $d_t = y_t - y_{t-1}$ be a first order difference then $d_t - d_{t-1}$ is a 2nd order difference.

Differencing

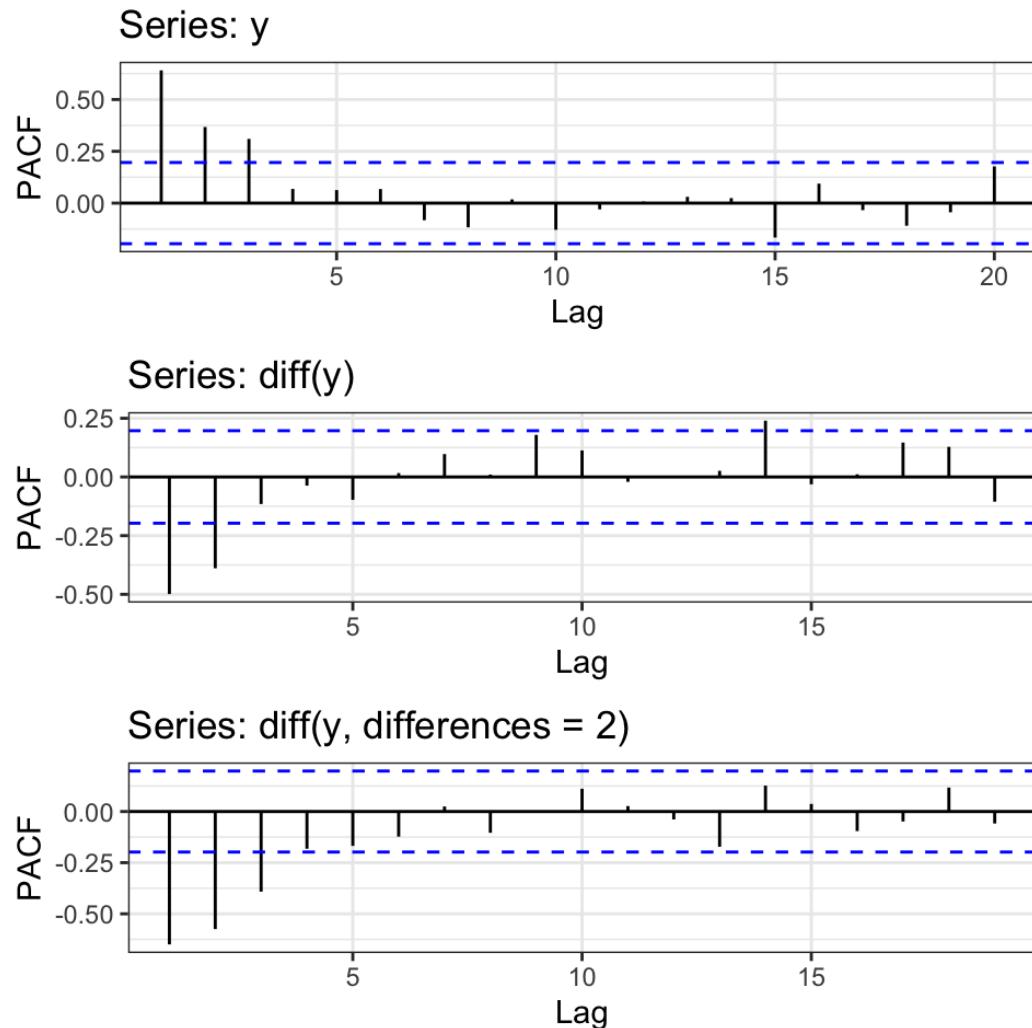
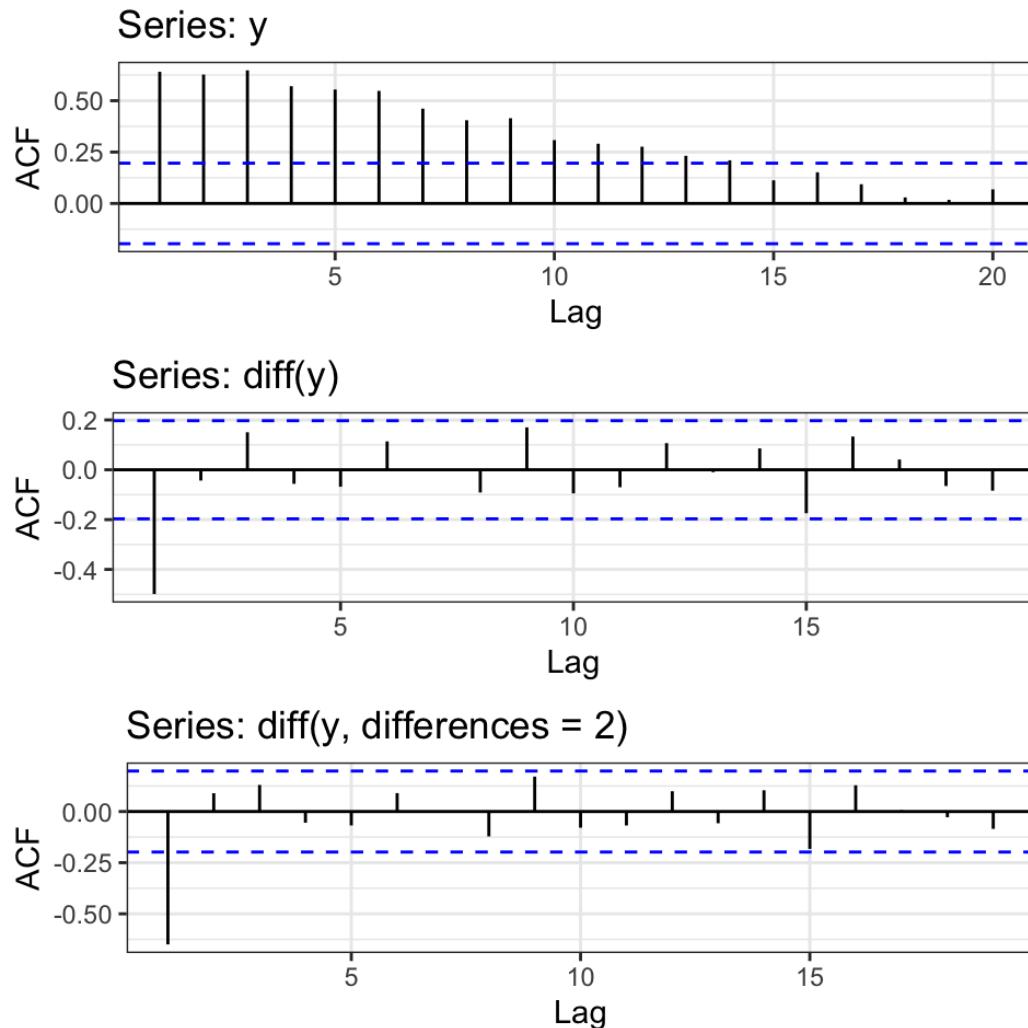
1st Difference



2nd Difference



Differencing - ACF



AR Models

AR(1)

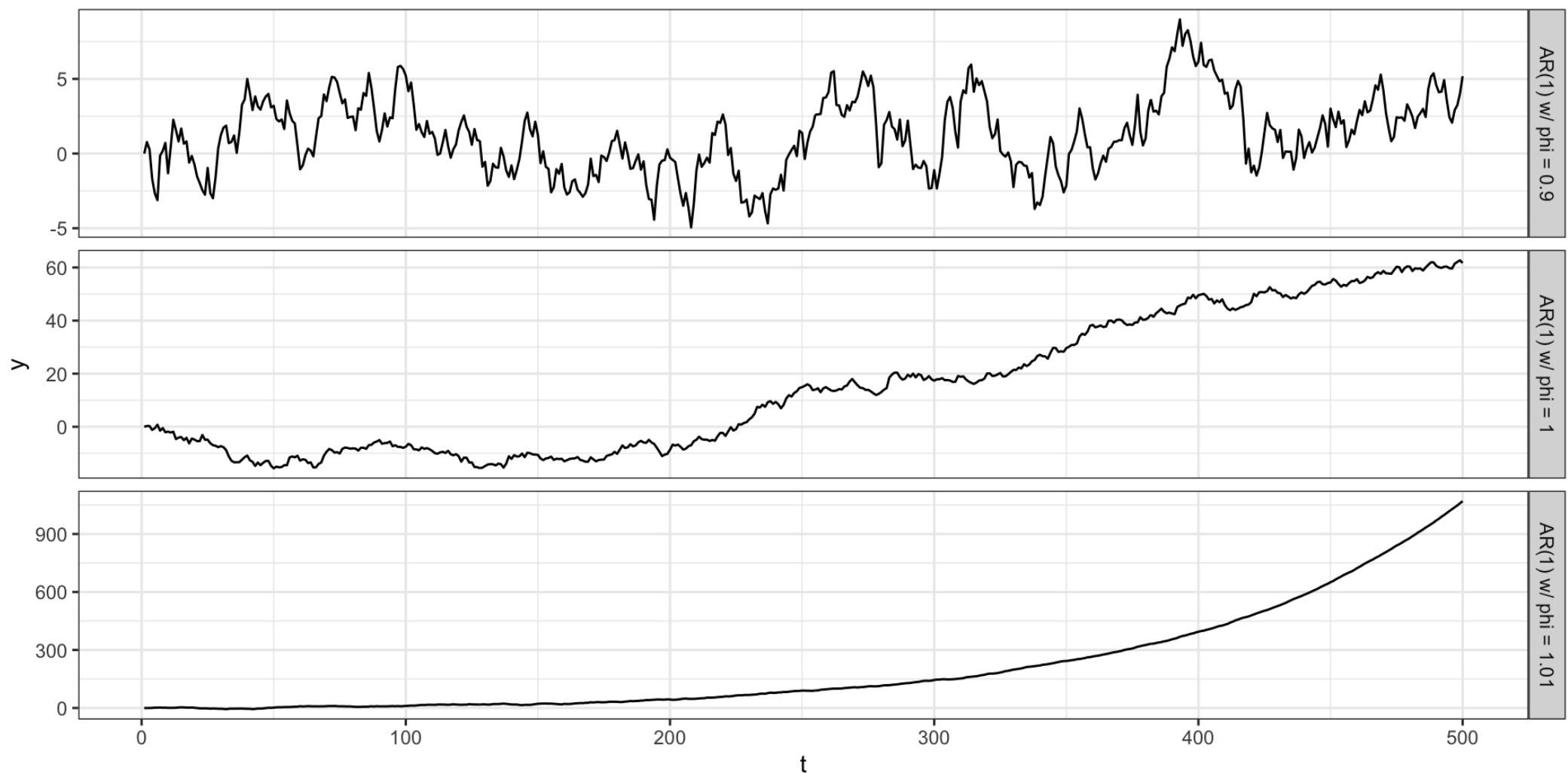
Last time we mentioned a random walk with trend process where
 $y_t = \delta + y_{t-1} + w_t$.

The AR(1) process is a generalization of this where we include a coefficient in front of the y_{t-1} term.

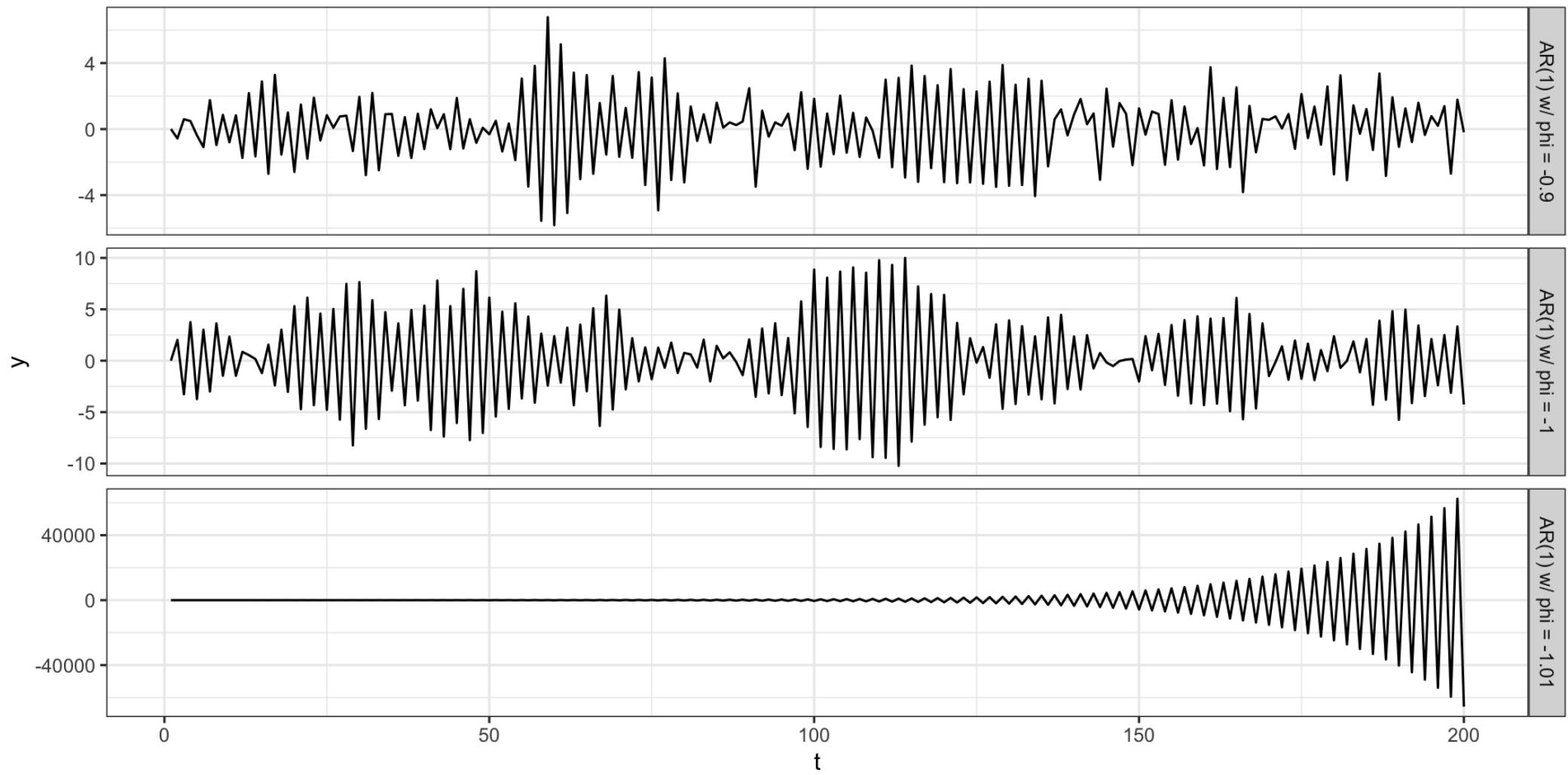
$$\text{AR}(1) : \quad y_t = \delta + \phi y_{t-1} + w_t$$

$$w_t \sim N(0, \sigma_w^2)$$

AR(1) - Positive ϕ



AR(1) - Negative ϕ



Stationarity of AR(1) processes

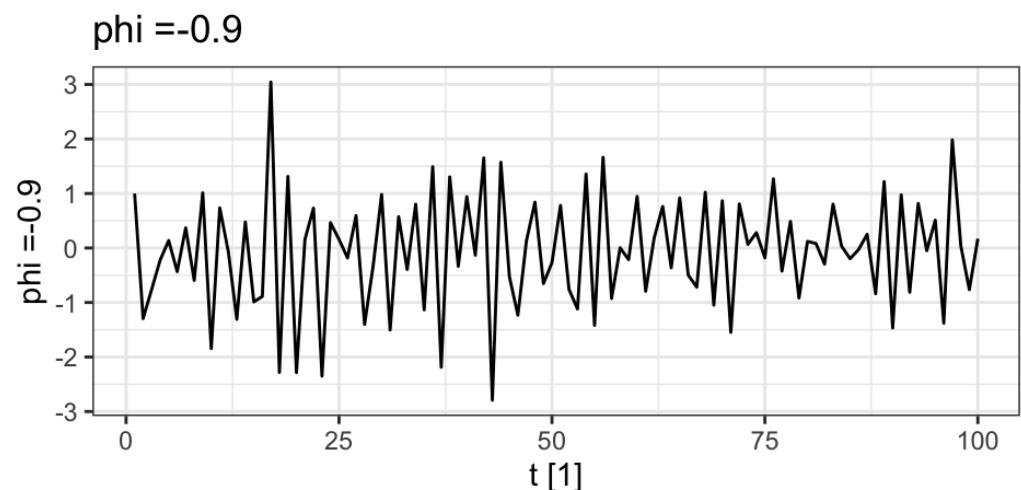
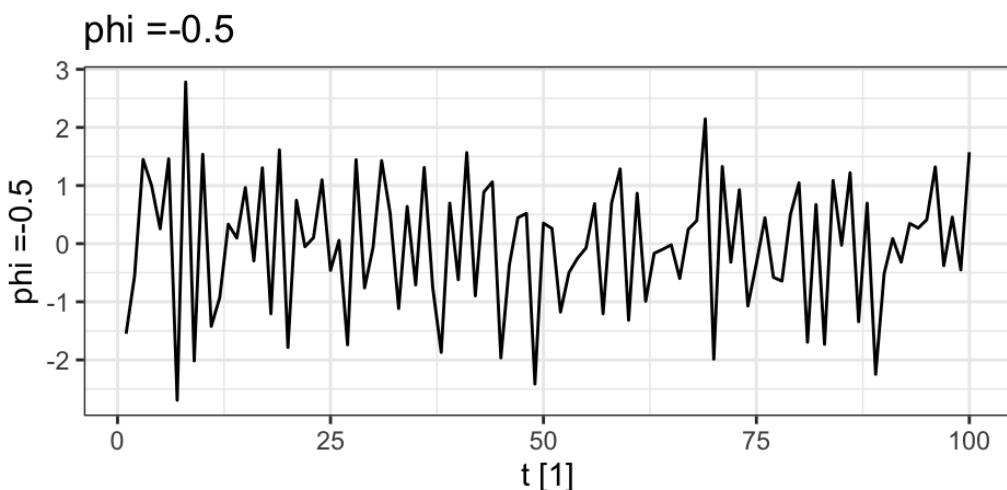
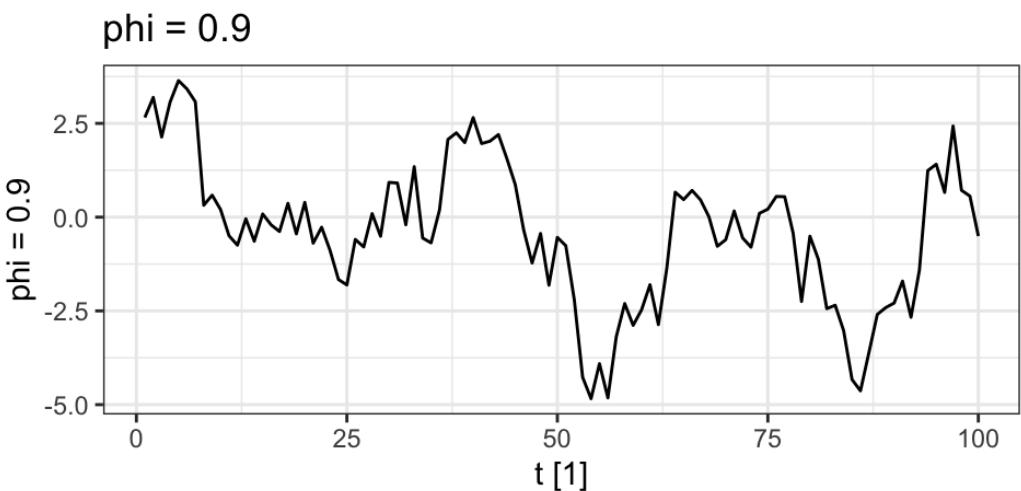
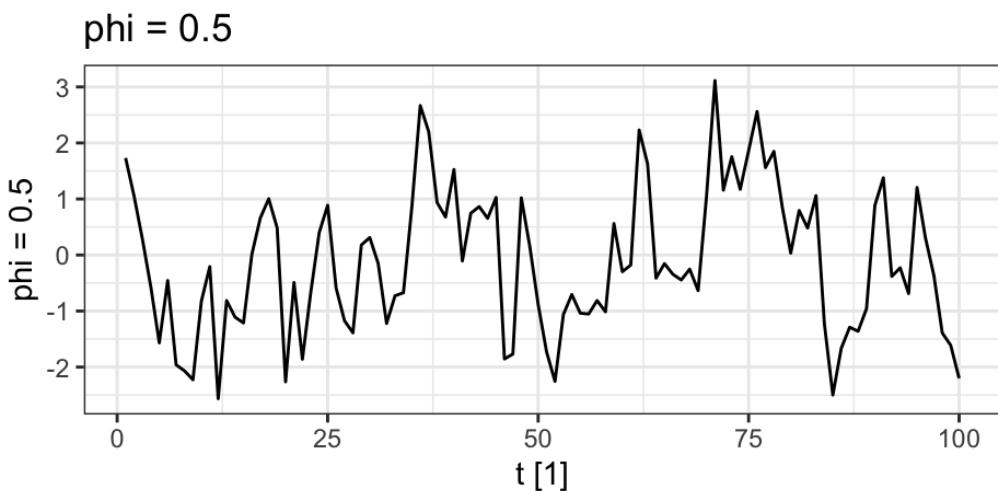
Lets rewrite the AR(1) without any autoregressive terms

Stationarity of AR(1) processes

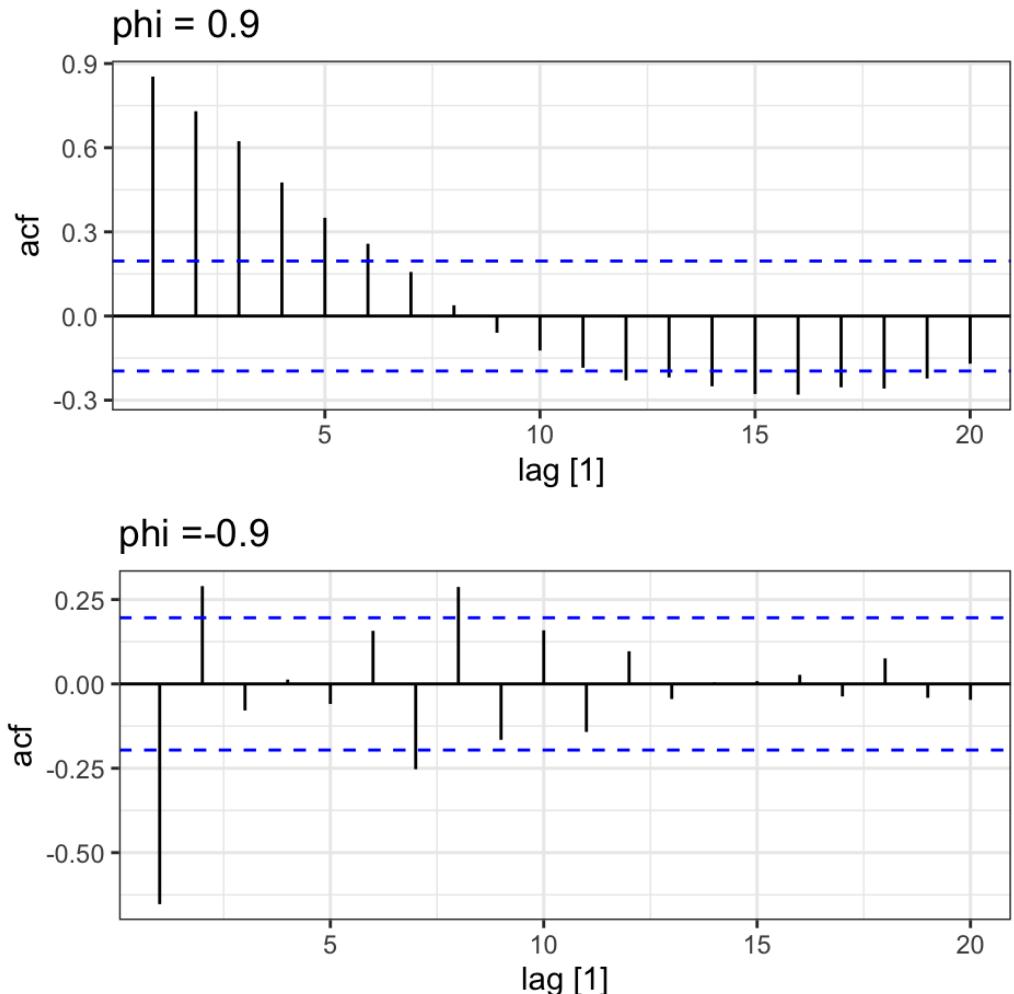
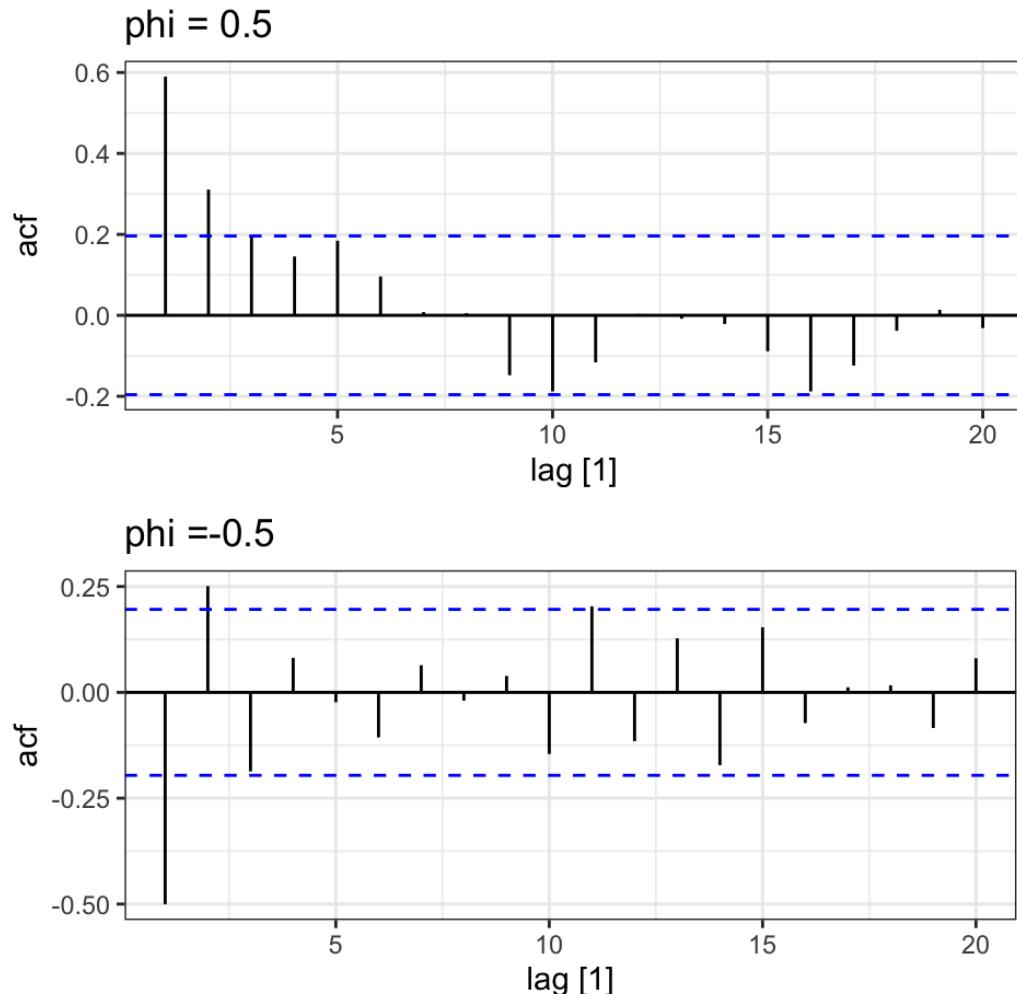
Under what conditions will an AR(1) process be stationary?

Properties of a stationary AR(1) process

Identifying AR(1) Processes



Identifying AR(1) Processes - ACFs



Identifying AR(1) Processes - PACFs

