Germany's Solar Power Generation (2015-2019)

Santoro Arnaldo mat:822274

05 settembre 2019

Power Generation Growth

- Photovoltaics is a fast growing market: The Compound Annual Growth Rate (CAGR) of PV installations was 24% between year 2010 to 2017.
- Concerning PV module production in 2017, China&Taiwan hold the lead with a share of 70%, followed by Rest of Asia-Pacific & Central Asia (ROAP/CA) with 14.8%. Europe contributed with a share of 3.1% (compared to 4% in 2016); USA/CAN contributed 3.7%.
- In 2017, Europe's contribution to the total cumulative PV installations amounted to 28% (compared to 33% in 2016). In contrast, installations in China accounted for 32% (compared to 26% in 2016).

Power Generation Growth

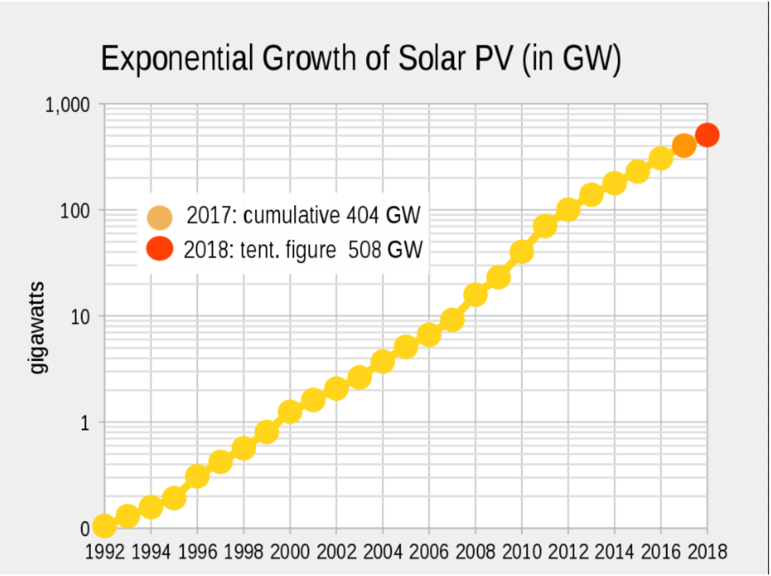


Figure 1: https://en.wikipedia.org/wiki/Solar_power

European horizontal irradiation

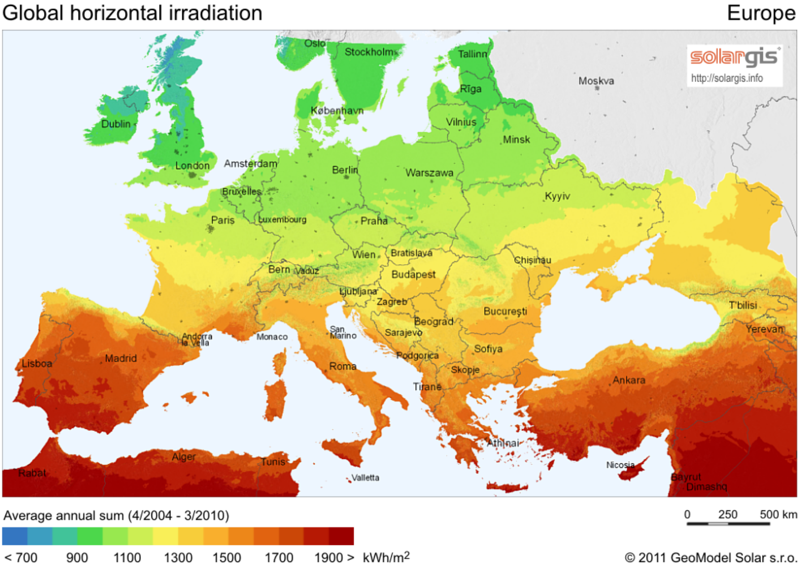


Figure 2: https://en.wikipedia.org/wiki/Solar_power

Solar Power in Germany

- Germany has about the same solar potential as Fairbanks, Alaska (Source: NREL, based on an average of 30 years of weather data).
- In 2018, Germany accounted for about 9% (45.4 GWp) of the cumulative PV capacity installed worldwide (515 GWp) with about 1.7 million PV systems installed in Germany. In 2018 the newly installed capacity in Germany was about 3.0 GWp; in 2017 it was 1.7 GWp.
- PV covered about 7% of Germany's electricity demand in 2017. Renewable sources delivered about 38% of the total net power consumption in 2017 in Germany.
- In 2017 about 19 Mio. t CO2 emissions have been avoided due to 38.4 TWh electrical energy generated by PV in Germany.
- PV system performance has strongly improved. Before 2000 the typical Performance Ratio was about 70%, while today it is in the range of 80% to 90%.

Solar Power in Germany

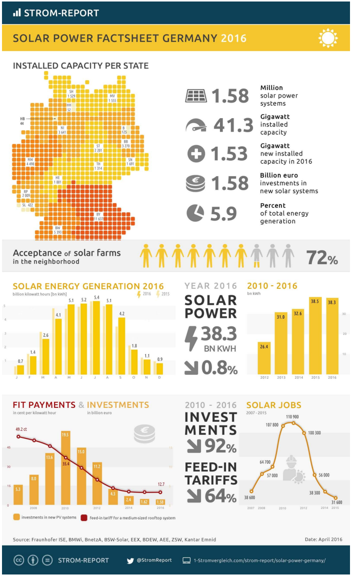


Figure 3: https://en.wikipedia.org/wiki/Solar_power_in_Germany

Solar Park in Meuro

Figure 4: Meuro and Schipkau, Germany 2012

Dataset Documentation

region	variable	attribute	Explanation
ISO 3166 area code and name or control area or bidding zone	load	actual_entsoe_power_statistics	Total load as published on ENTSO-E Data Portal
		actual_entsoe_transparency	Total load as published on ENTSO-E Data Portal/Power Statistics
		actual_tso	Total load as published by the TSO
		actual_net_consumption_tso	Total load exluding transmission losses as published onby the TSO
		actual_gross_generation_tso	Total power generation from national TSO
		day_ahead_forecast_entsoe_transparency	Day ahead total load forecast from ENTSO-E Transparency Platform
	solar / wind / wind_onshore / wind_offshore	generation_actual	Electricity produced by all power plants
		day_ahead_generation_forecast	Day-ahead generation forecast
		generation_actual_dso	Electricity produced by power plants connected to distribution grid
		generation_actual_tso	Electricity produced by power plants connected to transmission grid
		capacity	Aggregated installed capacity of power plants (actual availability not accounted for)
		profile	Share of installed capacity producing
	price	day_ahead	Day-ahead spot price

Figure 5: Variables of the Dataset

Dataset Documentation

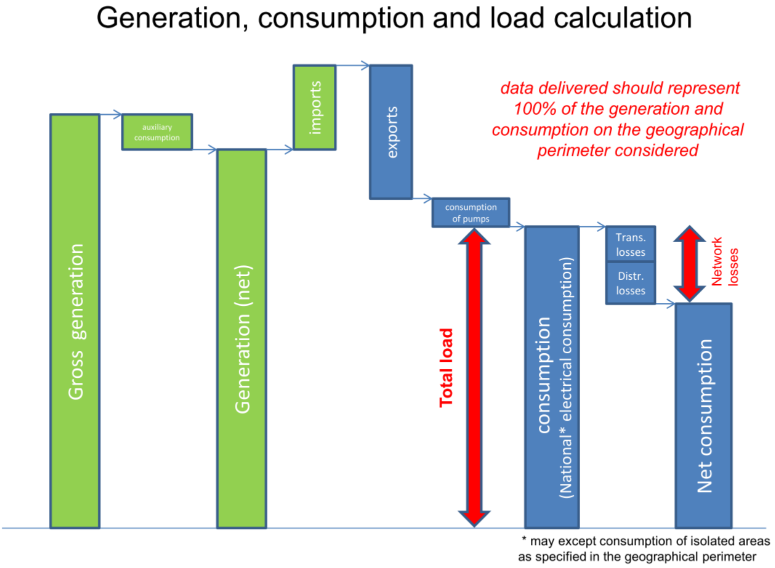


Figure 6: Variables of the Dataset

Dataset

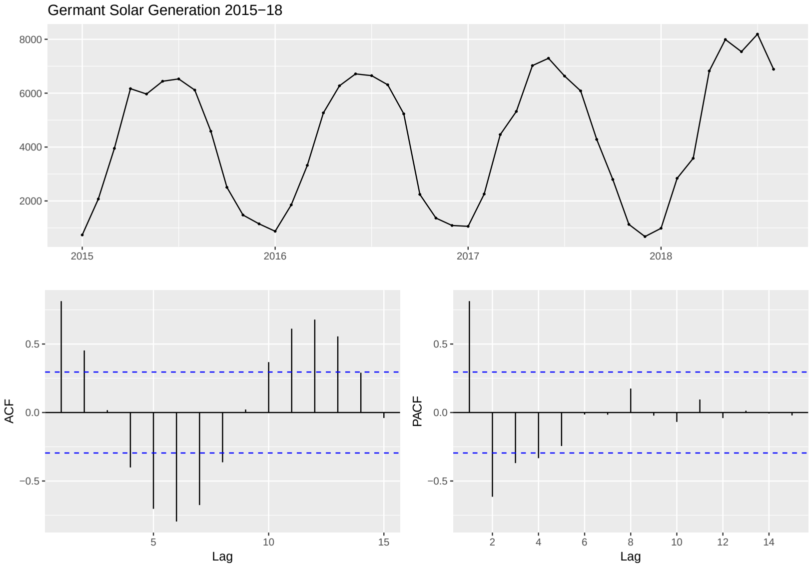
Problems with the dataset:

- Vast majority of solar data from 2005 to 2019 is NA.
- Data uses a different scale adjustment from 2005 till 2014.
- For these reasons the data used for this project goes from 2015 to 2019.

Data manipulation:

- Data was transformed from hourly to monthly. Last month of the test has only 9 observation avaliable, so its information is less reliable.
- The dataset was divided into train and test, using only the last 5 months for validation.
- Log transformation results very useful in energy production data, but Box-Cox transformations didn't result very useful, probably because of to few data, so readability was preferred.

Dataset



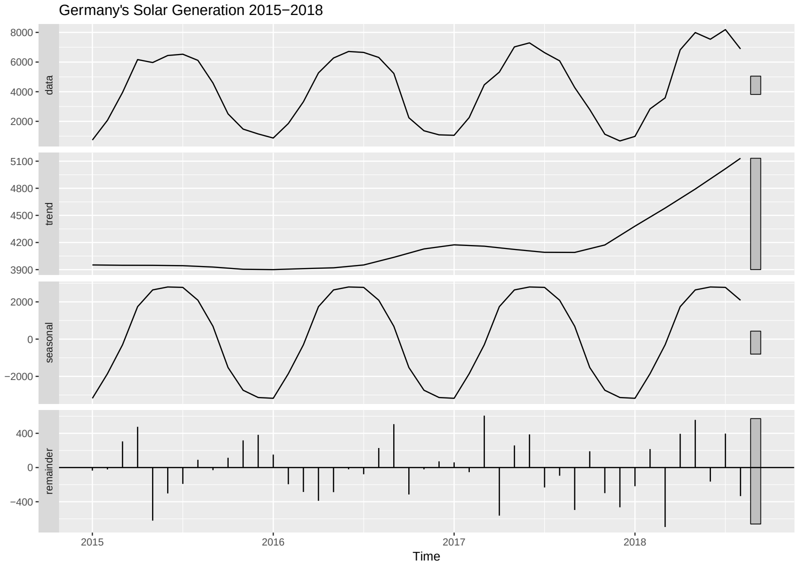
Germant Solar Generation 2015−18
8000
6000
4000
2000
2015
2016
2017
2018
ACF
0.5
0.0
−0.5
5
10
15
Lag
PACF
0.5
0.0
−0.5
2
4
6
8
10
12
14
Lag

STL Decomposition

Stands for Seasonal and Trend decomposition using loess

- loess is a kind of filter
- Seasonality found by loess smoothing on the seasonal sub-series (all Jan values, all Feb values,...): $loess(y_t) = \hat{s}_t$
- Seasonality is removed: $t_t \hat{+} e_t = y_t - \hat{s}_t$
- The Trend is found by smoothing the remainder: $loess(t_t \hat{+} e_t) = \hat{t}_t$
- The remainder component is the residuals from the seasonal plus trend fit: $\hat{e}_t = y_t - \hat{t}_t + \hat{s}_t$.
- This method provides values for $\hat{t}_t, \hat{s}_t, \hat{r}_t, t \in \{0, ..., \lfloor p/2 \rfloor - 1\}$ not provided by `decompose` function.

STL Decomposition



Auto Arima Algorithm

Hyndman-Khandakar algorithm for automatic ARIMA modelling

1. The number of differences $0 \le d \le 2$ is determined using repeated KPSS tests.

2. The values of p and q are then chosen by minimising the AICc after differencing the data d times. Rather than considering every possible combination of p and q, the algorithm uses a stepwise search to traverse the model space.

   a. Four initial models are fitted:
      - ARIMA$(0, d, 0)$,
      - ARIMA$(2, d, 2)$,
      - ARIMA$(1, d, 0)$,
      - ARIMA$(0, d, 1)$.

      A constant is included unless $d = 2$. If $d \le 1$, an additional model is also fitted:
      - ARIMA$(0, d, 0)$ without a constant.

   b. The best model (with the smallest AICc value) fitted in step (a) is set to be the "current model".

   c. Variations on the current model are considered:
      - vary p and/or q from the current model by ± 1;
      - include/exclude c from the current model.

      The best model considered so far (either the current model or one of these variations) becomes the new current model.

   d. Repeat Step 2(c) until no lower AICc can be found.

Figure 7:

Auto Arima Algorithm

- The Model Found has been recognized as a $ARIMA(0,1,1)_{12}(0,0,0)$ with drift.
- $\Theta = -0.5840$ with standard error 0.4488 (!)
- $\delta = 18.1580$ with standard error 6.2926
- $D = 1$.

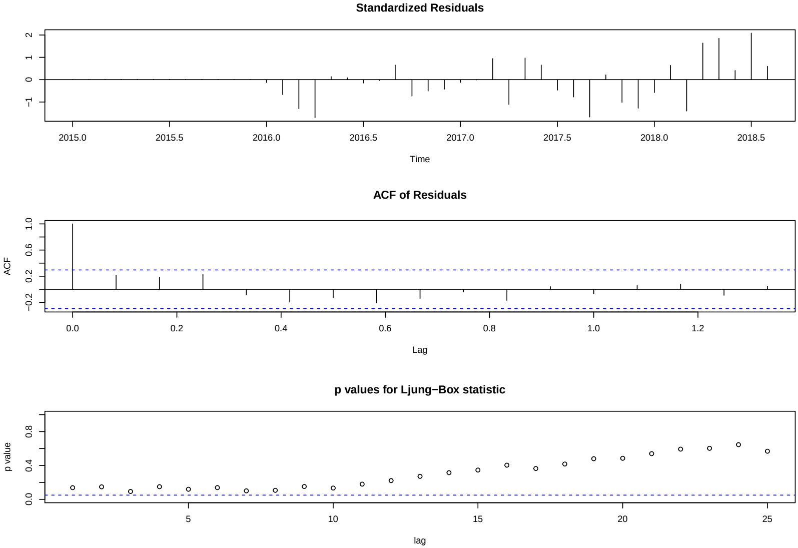
Other models were found, mainly to compare them with the automatically chosen one. They weren't necessairly chosen to be worse. Indeed their Parameters have lower standard errors, and their Information criteria are lower too.

Alternative Models

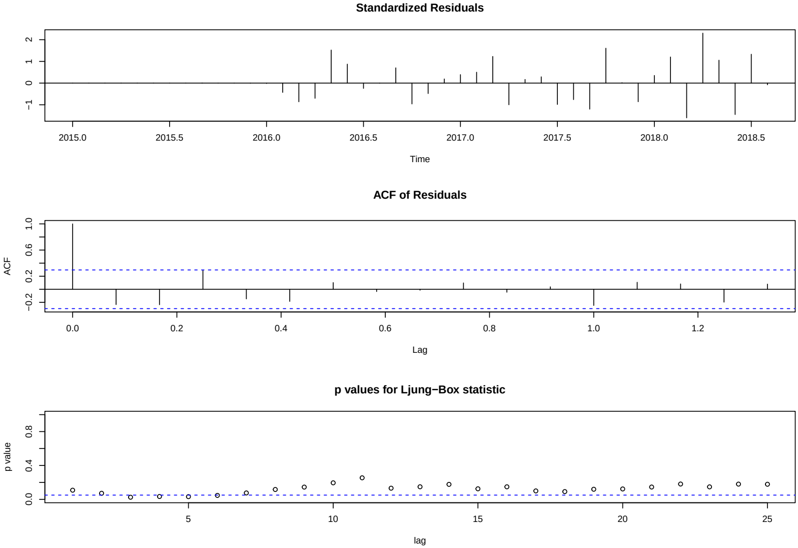
```
## Series: train
## ARIMA(0,1,1)(1,1,0)[12]
##
## Coefficients:
##           ma1     sar1
##       -0.6963  -0.4155
## s.e.   0.1548   0.2090
##
## sigma^2 estimated as 362283:  log likelihood=-242.83
## AIC=491.66   AICc=492.55   BIC=495.96
```


```
## Series: train
## ARIMA(0,1,1)(1,1,0)[12]
##
## Coefficients:
##           ma1     sar1
##       -0.6963  -0.4155
## s.e.   0.1548   0.2090
```

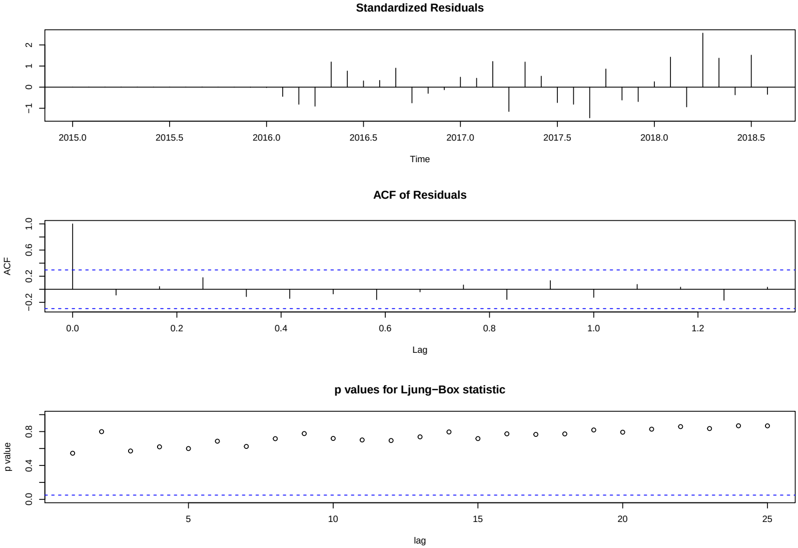
Portmanteau Test for Automated model



Pormanteau Test for Manual Model 1

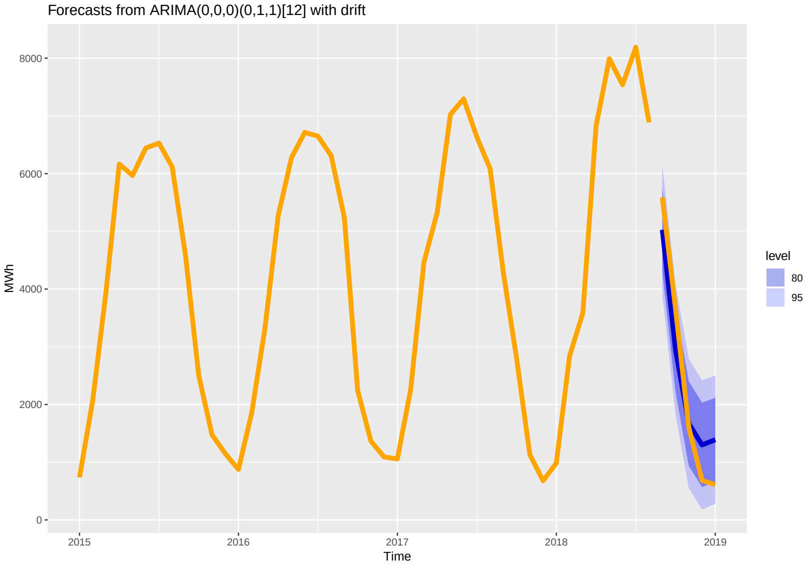


Pormanteau Test for Manual Model 2



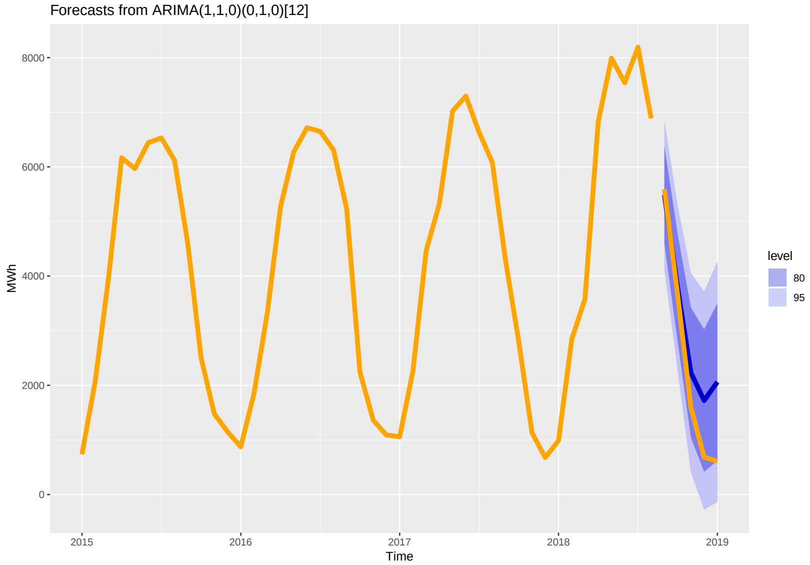
h-step ahead Forecasts

5-step ahead forecast for automatic model:



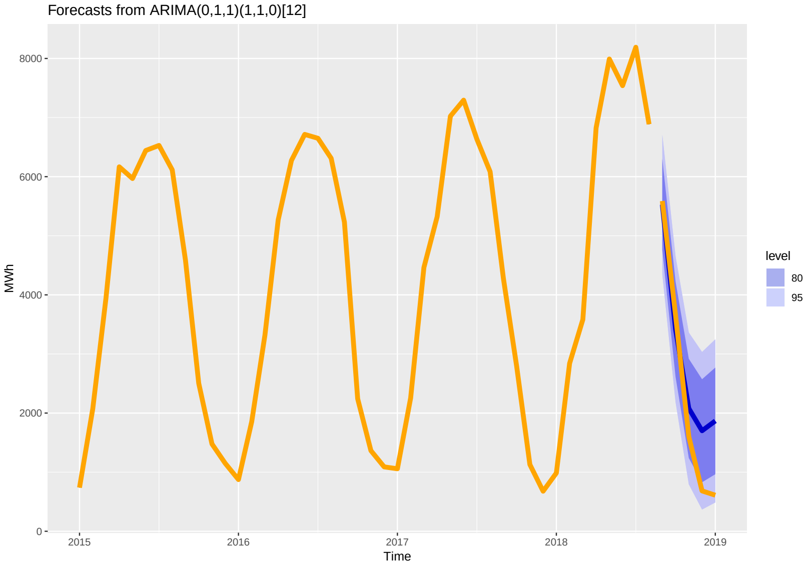
Forecasting

5-step ahead forecast for model 1:



Forecasting

5-step ahead forecast for model 2:



Accuracy of models

The accuracy measures of the models compared with themselves.

The third one is the result of `auto.arima` function.

Mean Absolute Scaled Error ($MASE$) is a scaled error measure.

It is based on MAE, the Mean Absolute Error.

Given a model m , $MASE(m) < 1$ means that the forecast is better than an average naive forecast, while $MASE(m) > 1$ means it's worse.

```
##                      ME     RMSE      MAE         MPE 
## Training set   33.49621 574.5916 401.8576  -0.8055412 12
## Test set     -630.94716 849.9963 671.7468 -86.2211959 86
##                    ACF1 Theil's U
## Training set -0.2342681        NA
## Test set      0.4300110   2.52961
```

```
##                      ME     RMSE      MAE         MPE 
## Training set   66.83796 488.6486 351.8273  -0.7150421 10
```

Cross Validation

RMSE for model 1, model 2, and the automated model.

```
## [1] 718.3601
```

```
## [1] 640.3908
```

```
## [1] 613.733
```

CV means A mean different than zero means that forecast is biased

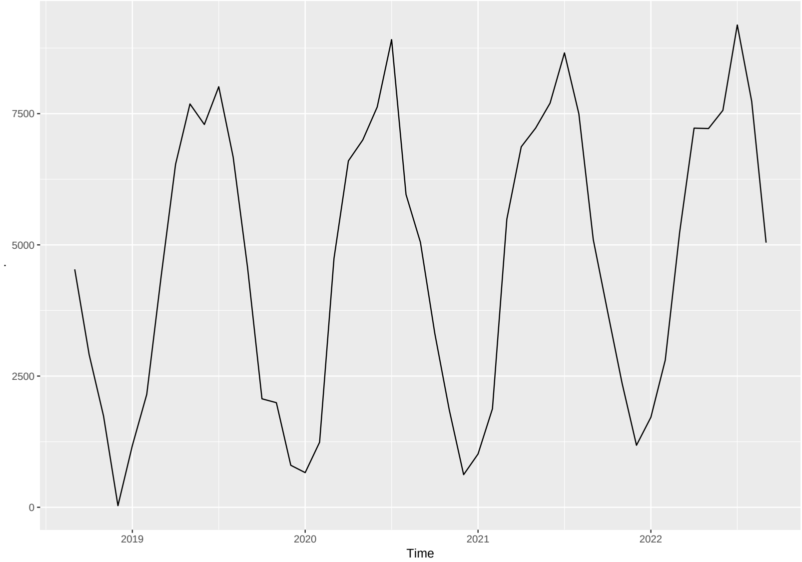
```
## [1] 15.73396
```

```
## [1] 69.18373
```

```
## [1] 59.51847
```

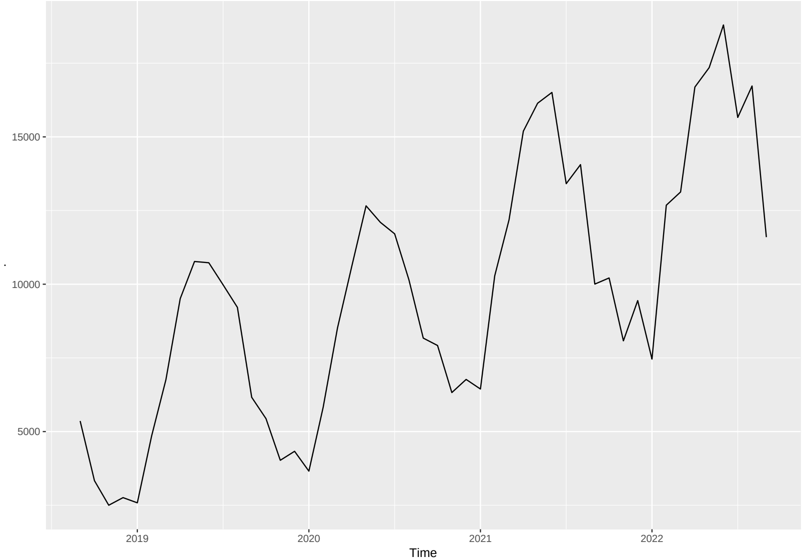
Simulation

Arima simulation with the `auto.arima` model:



Simulation

Arima simulation with model 1:



Simulation

Arima simulation with the model 2: