Table of Contents

1 AutoRegressive Moving Average models ¶

1.1 AutoRegressive Moving Average model (ARMA)

Let's consider the weighted moving average equation (**AutoRegression, AR**) as

$$\hat{y}_n = y_n + \alpha_1 y_{n-1} + \alpha_2 y_{n-2} + ... = y_n + \sum_{i=1}^{p} \alpha_i y_{n-i}.$$

Values of y_n can be also represented as its set of random-walk-like processes **Moving Average, MA**) as

$$y_n = \beta_0 \varepsilon_n + \beta_1 \varepsilon_{n-1} + \beta_2 \varepsilon_{n-2} + ... = \sum_{i=0}^{q} \beta_i \varepsilon_{n-i},$$

where $\beta_0 = 1$.

Thus we can join equations for describe the time series as both:

$$\hat{y}_n = \sum_{i=1}^{p} \alpha_i y_{n-i} + \sum_{i=0}^{q} \beta_i \varepsilon_{n-i}$$

where α_i and β_i are the model weight coefficients. This equation we will call

AutoRegressive Moving Average model (ARMA).

Pleas note:

1. In some sources you can find the following equation

$$\hat{y}_n - \sum_{i=1}^{p} \alpha_i y_{n-i} = \sum_{i=0}^{q} \beta_i \varepsilon_{n-i}$$

   which is different only by the sign of the coefficients α.

2. The strict ARMA equation derivation based on the harmonic analysis (or Laplacian and z-space analysis more precisely), where it is proved that every rational (physically real) spectrum of time series (or any time-depended system or procpess) can be represented as

$$Y^2(\omega) = \sigma^2 |B(\omega)|^2 / |A(\omega)|^2$$

   where $A(\omega)$ is the spectrum of **AR** process and $B(\omega)$ is the spectrum of **MA** process.

   The main consequence of this that ARMA process can be considered as y_n filtration with infinite impulse response filter. **MA** separately can be considered as noise filtration such that output have to be similar to y_n.

In the practice we can approximate ARMA by its first term that now we will call **AutoRegressive model(AR)** as

$$\hat{y}_n = \sum_{i=1}^{p} \alpha_i y_{n-i} + \varepsilon_n$$

or by second term that now we will call **Moving Average model (MA)** as

$$\hat{y}_n = \sum_{i=0}^{q} \beta_i \varepsilon_{n-i}$$

When we stay that model $\hat{y}_n$ depends on the $p+1$ coefficients α we say that model has AR of $p-th\ order$.

When we stay that model $\hat{y}_n$ depends on the q coefficients β we say that model has MA of $q-th\ order$.

The task of ARMA approximation is to find such weight coefficients α_i and β_i and its orders that approximate $\hat{y}_n$ the best.

As a rule this task can be solved by regression methods.

For this we have to select a part of series to approximation and check the result.

If we know ARMA coefficients we can predict the future values of y_n samples using ARMA equation.

1.2 Autoregressive integrated moving average (ARIMA)

The main drawback of the ARMA process in the previously described form is the intrinsic stationary requirement, which lay in the routine of coefficient searching.

> Indeed, if we require to obtain the best approximation (the smallest error) on the one part of the series - the series should be the similar (or have the similar behavior) for its future parts to ensure the same error.

Actually in Time series analysis for reducing (or eliminate) discussed requirement of stationary firstly we must reach the stationarity.

We can reduce task to stationary by trying some inversable transforms.

As a rule, this can be done by taking the numerical derivative.

This model is called **Autoregressive integrated moving average (ARIMA)**

Let's denote derivative as d, numerical derivative can be calculated as

$$d = 0: \ y'_n = y_n$$

$$d = 1: \ y'_n = y_n - y_{n-1}$$

end e.t.c.

> <u>In the case of considering y'_n instead of y_n we make assumption that our derivative (only change of values) have stationary behavior.</u>

Let us define an operator $L : S \to S$, a map which transform sequences in sequences. It is denoted as

$$L y_n = y_{n-1}$$

for all t. We should write $(Ly)_n = y_{n-1}$, with the meaning that for given sequence $y = (y_n)$ we introduce a new sequence Ly, that at time n is equal to the original sequence at time $n{-}1$, hence the notation $(Ly)_n = y_{n-1}$ (for shortness, we drop the brackets). The map L is called **time lag operator**, or **backward operator**.

In the same manner we can define

$$L^k y_n = y_{n-k},$$

and

$$L^{-k} y_n = y_{n+k}.$$

With this notation, we can define the **AR** model as

$$\varepsilon_n = \hat{y}_n - \sum_{i=1}^{p} \alpha_i y_{n-i} = (1 - \sum_{i=1}^{p} \alpha_i L^i) y_n$$

and **MA** model as

$$y_n = (1 + \sum_{i=1}^{q} \beta_i L^i) \varepsilon_n.$$

Thus **ARMA(p,q)** model canbe given in the lag form as

$$(1 - \sum_{i=1}^{p} \alpha_i L^i) y_n = (1 + \sum_{i=1}^{q} \beta_i L^i) \varepsilon_n$$

In the lag form we can define numerical derivative as

$$\Delta y_n = y_n - y_{n-1} = (1 - L) y_n$$

$$\Delta^d y_n = (1 - L)^d y_n$$

Using the derivative definition we can write the

ARIMA(p,d,q) model as

$$(1 - \sum_{i=1}^{p} \alpha_i L^i)(1 - L)^d y_n = (1 + \sum_{i=1}^{q} \beta_i L^i) \varepsilon_n$$

if $d = 1$, y_n has a linear trend; if $d = 2$, a quadratic trend and so on.

In some case you may come across the following forms of ARIMA

$$(1) : (1 - \sum_{i=1}^{p} \alpha_i L^i)\Delta^d y_n = (1 + \sum_{i=1}^{q} \beta_i L^i)\varepsilon_n$$

$$(2.1) : (1 - \sum_{i=1}^{p} \alpha_i L^i)\Delta^d y_n = c(n) + (1 + \sum_{i=1}^{q} \beta_i L^i)\varepsilon_n$$

$$(2.2) : (1 - \sum_{i=1}^{p} \alpha_i L^i)\Delta^d y_n = c + e \cdot n + (1 + \sum_{i=1}^{q} \beta_i L^i)\varepsilon_n$$

$$(3) : \phi(L)\Delta^d y_n = \theta(L)\varepsilon_n,$$

$$(4) : \phi(L)\Delta^d y_n = c + \theta(L)\varepsilon_n,$$

where c is some addition deterministic constant or slow change almost deterministic trend and $c(n) = c + e \cdot n$ is the deterministic linear trend;

$$\phi(L) \text{ is the AR operator: } \phi(L) = 1 - \sum_{i=1}^{p} \alpha_i L^i,$$

and

$$\theta(L) \text{ is the MA operator: } \theta(L) = 1 + \sum_{i=1}^{q} \beta_i L^i.$$

Pleas note if term c is constant - it can be eliminated by subtraction mean value.

*Notes about **ARIMA***:

1. ARIMA(0,1,0) random walk - the simplest model of the

$$\hat{y}_n = \hat{y}_{n-1} + \varepsilon_n.$$

2. ARIMA(0,1,1) simple exponential-like smoothing

$$\hat{y}_n = \hat{y}_{n-1} + \beta\varepsilon_{n-1}.$$

3. A model with no orders of differencing assumes that the original series is stationary (mean-reverting).
4. A model with one order of differencing assumes that the original series has a constant average trend (e.g. a random walk or simple moving average-type model).
5. A model with two orders of differencing assumes that the original series has a time-varying trend (e.g. a random trend or EMA-type model).

The first thing you need to do when work with ARIMA - is to determine derivation (to exclude trend and make data more stationary).

As a rule it is need to be 1-3 orders of derivative is applied, and rarely more.

For checking the stationarity it can be use a lot of statistical and visual techniques, among them:

- **Rolling Statistics**: Plot the rolling mean and rolling standard deviation. The time series is stationary if they remain constant with time (with the naked eye look to see if the lines are straight and parallel to the x-axis).

  This technique corresponds to the weak definition of stationary.

- **ACF analysis** - for the non-stationary process you will see slow decrease and other-wise for stationary.

  For stationary series autocorrelation function (ACF) plot decays fairly rapidly to zero or below.

- **Augmented Dickey-Fuller Test**: The time series is considered stationary by specific unit-root test, and the p-value is calculated also. If p-value is low (according to the null hypothesis) and the critical values at 1%, 5%, 10% confidence intervals are as close as possible to the ADF Statistics.

- **Kwiatkowski–Phillips–Schmidt–Shin (KPSS)** test which is also unit-root test, but differs from ADF in the case of deterministic trend with inflection points.

After integration order selection the AR and MA orders need to be selected.

It can be the following recommendations about it

ARMA order selection

ACF Pattern	PACF Pattern	Conclusion
Tapers to 0 in some fashion	Non-zero values at first p points; zero values elsewhere	$AR(p)$ Model
Non-zero values at first q points; zero values elsewhere	Tapers to 0 in some fashion	MA(q) model
Values that remain close to 1, no tapering off	Values that remain close to 1, no tapering off	Symptoms of a non-stationary series. Differencing is most likely needed.
No significant correlations	No significant correlations	Random Series.

1.3 Seasonal Autoregressive Integrated Moving Average (SARIMA).

If the model has the strong seasonality, then it may be compensated by the so-called seasonal derivative give as

$$\Delta_s y_n = y_n^s = y_{n+k_s} - y_n$$

In the operator form it can be presented as

$$\Delta_s y_n = (1 - L^{k_s}) y_n$$

Thus it can be introduced so-called

Seasonal Autoregressive integrated moving average (SARIMA).

In more general case the showed above introduce only AR seasonal derivative, in more general case the following seasonal derivative can be defined for each of the AR (**Seasonal AR, SAR**) AR integrated (**Seasonal AR integrated, SARI**)and for the MA (**Seasonal MA, SMA**) as

$$\textbf{AR}\ :\varepsilon_n = \phi(L)y_n = (1-\sum_{i=1}^{p}\alpha_i L^i)y_n$$

$$\textbf{SAR}\ :\varepsilon_n = \phi(L)\Phi(L^s) = (1-\sum_{i=1}^{p}\alpha_i L^i)\cdot(1-\sum_{i=1}^{P}A_i L^{is})y_n$$

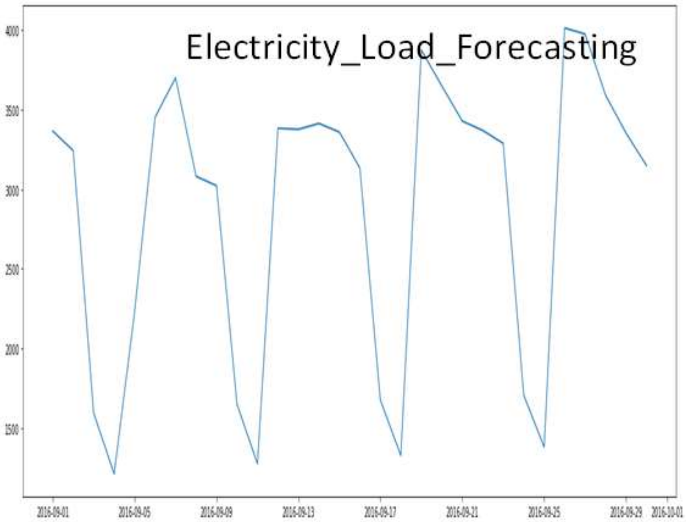
$$\textbf{SARI}\ :\varepsilon_n = \phi(L)\Phi(L^s)\Delta^d\Delta_s^D = (1-\sum_{i=1}^{p}\alpha_i L^i)\cdot(1-\sum_{i=1}^{P}A_i L^{is})\cdot(1-L)^d\cdot(1-L^s$$

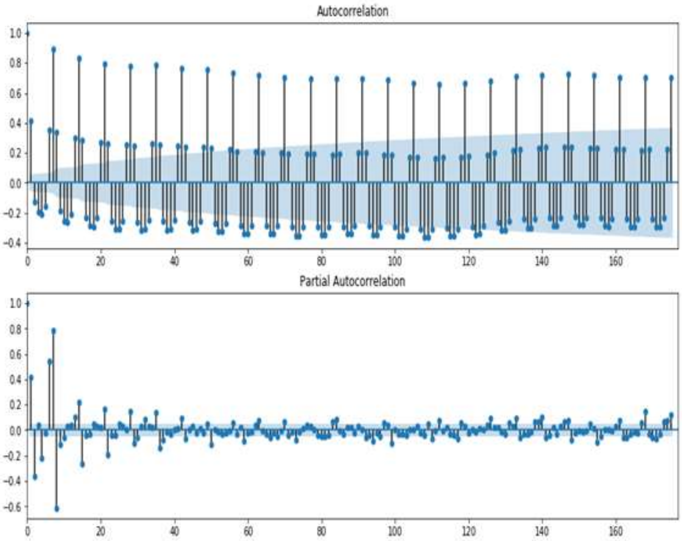
$$\textbf{MA}\ :y_n = \theta(L)\varepsilon_n = (1+\sum_{i=1}^{q}\beta_i L^i)\varepsilon_n$$

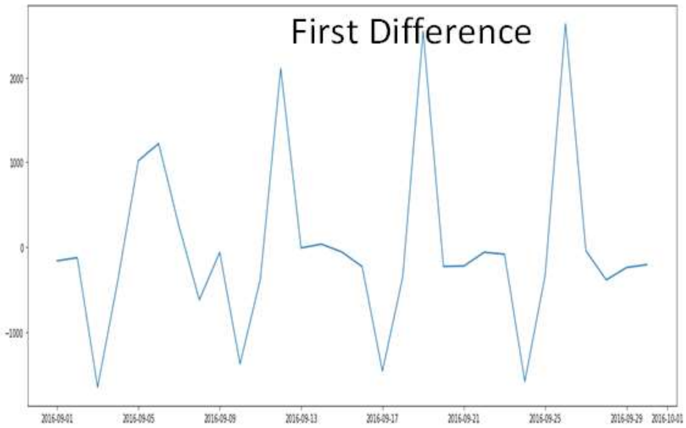
$$\textbf{SMA}\ :y_n = \theta(L)\Theta(L^s)\varepsilon_n = (1+\sum_{i=1}^{q}\beta_i L^i)\cdot(1+\sum_{i=1}^{Q}B_i L^{is})\varepsilon_n$$

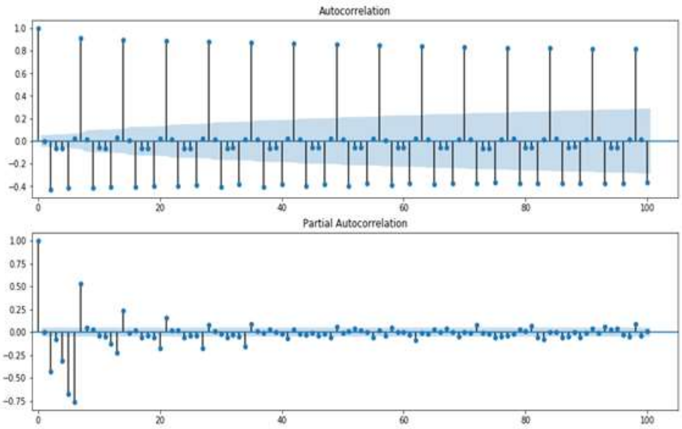
where A and B are the new coefficients for seasonal component compensation; P is the seasonal AR order and Q is the seasonal MA order and D is the order of seasonal derivative.

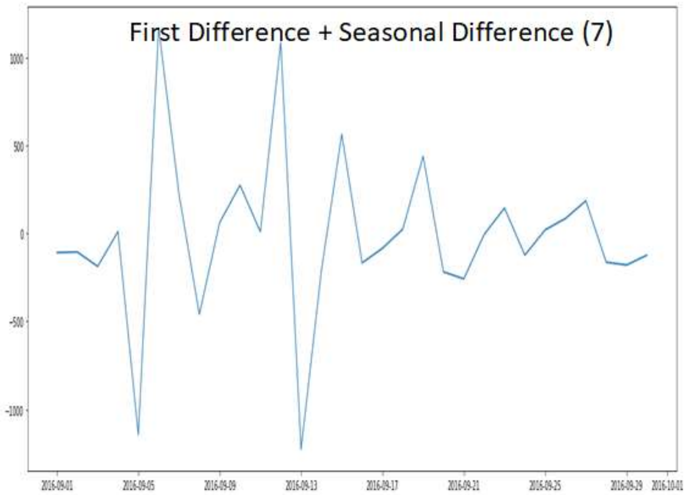
Pleas note that derivate are taken only of AR part.

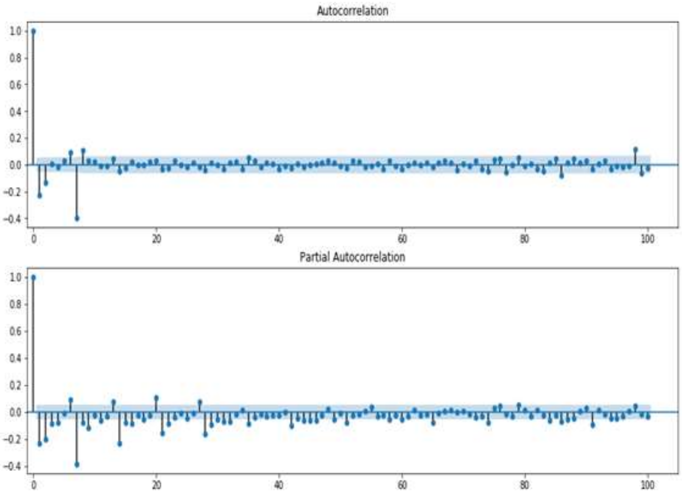

Electricity_Load_Forecasting

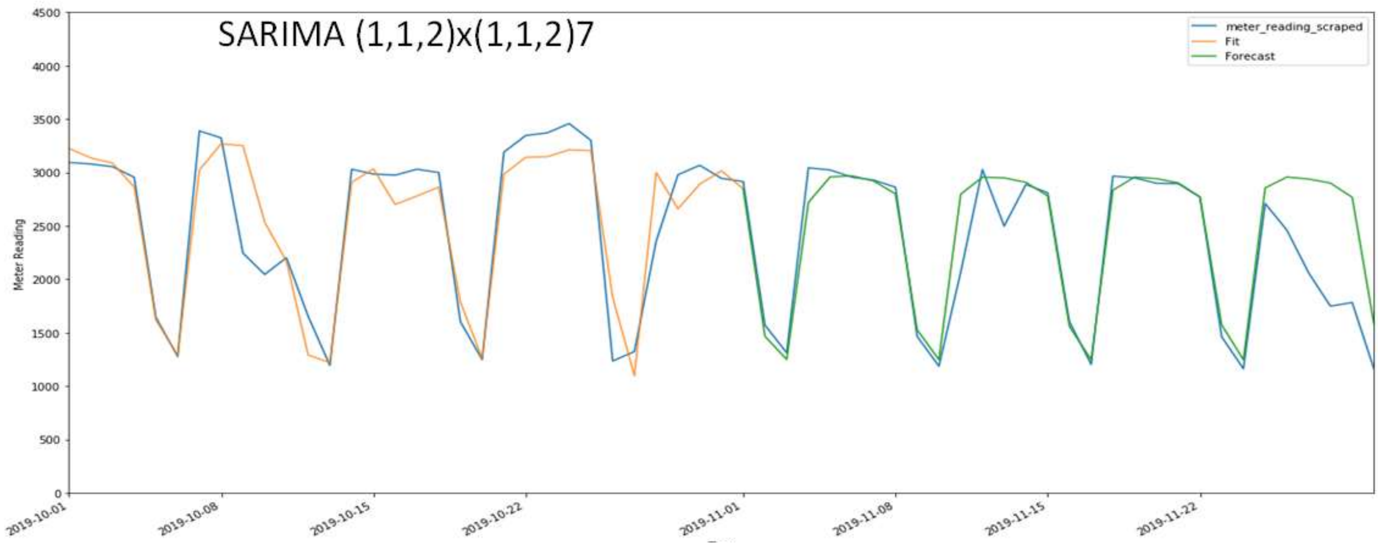

Autocorrelation
Partial Autocorrelation


First Difference


Autocorrelation
Partial Autocorrelation


First Difference + Seasonal Difference (7)


Autocorrelation
Partial Autocorrelation


SARIMA (1,1,2)x(1,1,2)7
meter_reading_scraped
Fit
Forecast
Meter Reading

Time

The **Full SARIMA Equation** can be given as

$$(1-\sum_{i=1}^{p}\alpha_i L^i)(1-\sum_{i=1}^{P}A_i L^{is})(1-L)^d(1-L^s)^D y_n =$$
$$= c(n) + (1+\sum_{i=1}^{q}\beta_i L^i)(1+\sum_{i=1}^{Q}B_i L^{is})\varepsilon_n$$

or in more compact form as

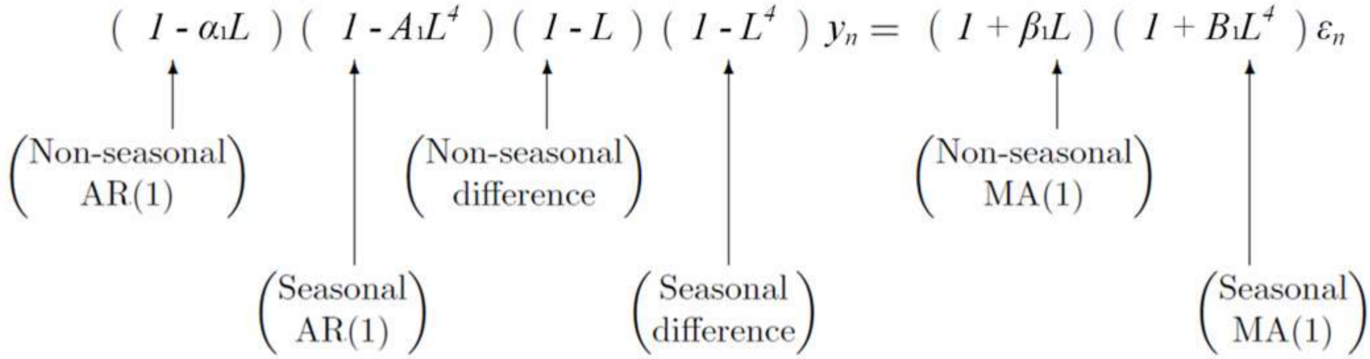
$$\phi(L)\Phi(L^s)\Delta^d\Delta_s^D y_n = c(n) + \theta(L)\Theta(L^s)\varepsilon_n$$

The model order fro SARIMA frequently written in the form

$$\text{SARIMA}(p,d,q)\times(P,D,Q)s \text{ or}$$
$$\text{SARIMA}(p,d,q)\times(P,D,Q,s)$$

where s is the season period.

Example for SARIMA $(1,1,1)\times(1,1,1)4$



Pleas note that here differentiation operator multiplication works as

$$(1-L)(1-L^4)y_n = (y_n - y_{n-1}) - (y_{n-4} - y_{n-5})$$

Example: Suppose $P = q = 1$ and $p = Q = 0$ with $s = 12$

$\rightarrow SARIMA(1, 0, 0) \times (1, 0, 0)12$. The model is then

$$y_n = Ay_{n-12} + \varepsilon_t + \beta\varepsilon_{t-1}$$

Example

Let's consider model SARIMA model $(0, 0, 1) \times (0, 0, 1)12$

$$y_n = (1 + B_1 L^1 2)(1 + \beta_1 L)\varepsilon_n = (1 + \beta_1 L + B_1 L^1 2 + \beta_1 B_1 L^1 3)\varepsilon_n = \\ = \varepsilon_n + \beta_1\varepsilon_{n-1} + B_1\varepsilon_{n-12} + \beta_1 B_1\varepsilon_{n-13}$$

The model include:

non-seasonal MA(1) polynomial $1 + \beta_1 L$

seasonal MA(1) polynomial $1 + B_1 L^{12}$.

Thus the model has MA terms at lags 1, 12, and 13. This leads to supposition that ACF for the model will have non-zero autocorrelations only at lags 1, 12, and 13. However, as you will see it is also will be non-zero lag 11 due to correlation properties. Here are ACF and PACF for discussed process with $\beta_1 = 0.7$ and $B_1 = 0.6$

The algorithm of fitting an SARIMA model.

1. Determination of the order of usual differencing needed to stationarize the series (to exclude the trend).

2. Checking for seasonal component influence and possibly eliminating seasonality.
3. Determination of AR and MA order.
4. Coefficients calculation.
5. Checking the results on validation tests.
6. If correct, then predicting, repeat otherwise.

In the explanation above we have notice that seasonal difference make data more stationary.

Then we need to choose the best initial guess about orders of SARIMA.

Rules of initial orders selection

1. **The correct order** d is the order of differencing that yields a time series with a noise like behavior

    i.e. fluctuates around a well-defined mean value and almost constant scatter, check for stationary by criterions named above.

2. Use seasonal derivative only in the case of strong seasonal pattern.
3. Then considering ACF and PACF for differenced series.

- The numbers of **AR terms** is determined as the last PACF lag before rapid decreasing from positive values to zero.
- The numbers of **MA terms** is determined as the last ACF lag before rapid increasing from negative values to zero.
- Add an **SAR term**, If ACF is periodically positive.

    Beside This the order be derived from PACF.

        Look at the number of significant lags, which are the multiples of the season period length.

            For example, if the period equals 24 and we see the 24-th and 48-th lags are significant in the PACF, that means the initial P should be 2.

- Add an **SMA term**, If ACF is periodically negative.

    Use the same rules of determine the number of lagas as for SAR

Note

- If your series is slightly **under-differenced**, add additional AR terms.
- If your series is slightly **over-differenced**,add additional MA terms.

- Try to avoid using more than one or two seasonal parameters (SAR + SMA) in the same model, as this is likely to lead to overfitting of the data and / or problems in estimation.
- if the series has **positive ACF values out to a high lag** .add differencing order.

Other Notes of initial order selection in SARIMA

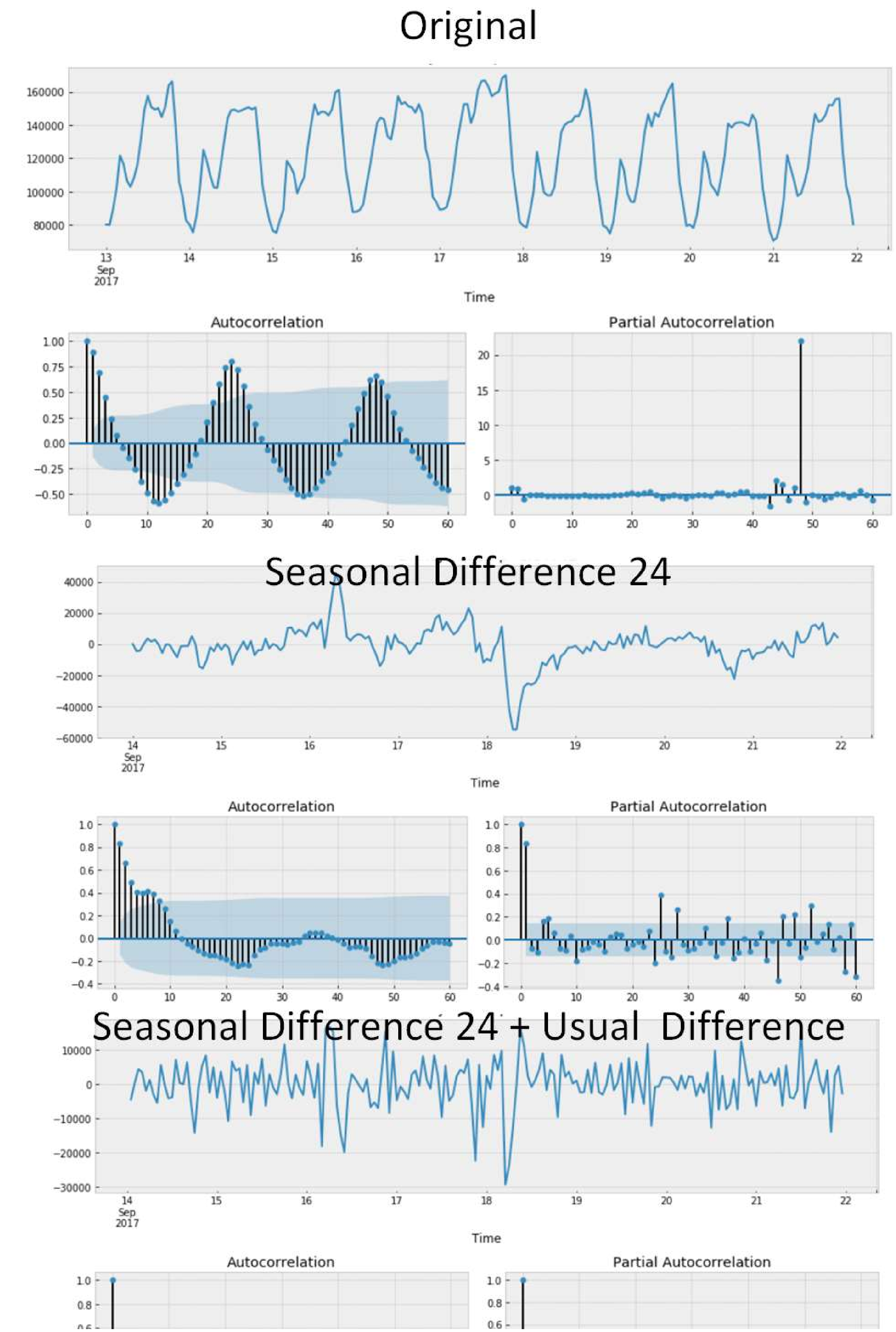
1. The optimal order of differencing is often the order of differencing at which the variance is lowest (you must try to avoid over-differencing).
2. If the PACF of the differenced series displays a sharp cutoff and/or the lag-1 autocorrelation is positive-i.e., if the series appears slightly "underdifferenced"--then consider adding an AR term to the model. The lag at which the PACF cuts off is the indicated number of AR terms.
3. If the ACF of the differenced series displays a sharp cutoff and/or the lag-1 autocorrelation is negative-i.e., if the series appears slightly "overdifferenced"-then consider adding an MA term to the model. The lag at which the ACF cuts off is the indicated number of MA terms.
4. If a full (AR and MA) model it is possible for an AR term and an MA term to cancel each other's effects. sIn

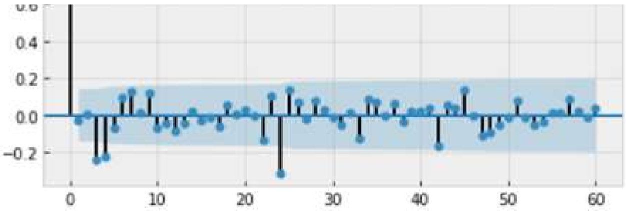
this case try a model with one fewer AR term and one fewer MA term-particularly if the parameter estimates in the original model require more than 10 iterations to converge.

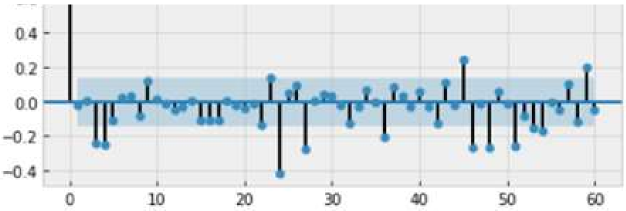
5. If the sum of the AR coefficients is almost exactly 1- you should reduce the number of AR terms by one and increase the order of differencing by one.
6. If the sum of the MA coefficients is almost exactly 1- you should reduce the number of MA terms by one and reduce the order of differencing by one.
7. If the series has a strong and consistent seasonal pattern, then you can use an order of seasonal differencing.
8. Use seasonal difference (D) for non-stationary seasonality.
9. If the series ACF has periodic positive bursts add seasonal order to SAR.
10. If the series ACF has periodic negative bursts add seasonal order to SMA.

Example 1:

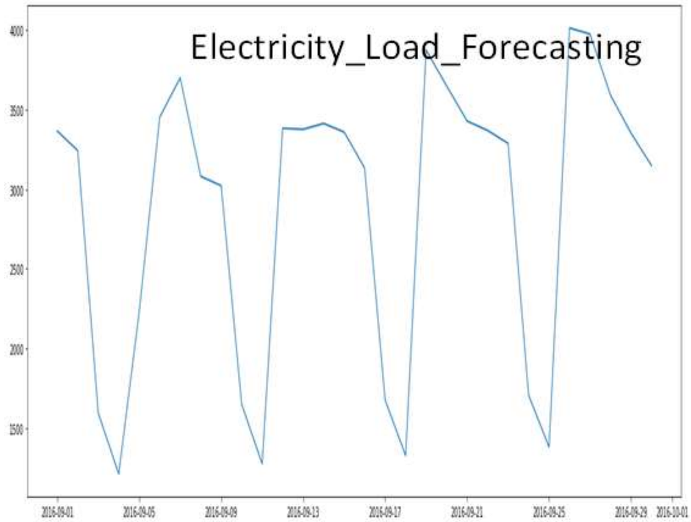
- p is most probably 4 since it is the last significant lag on the PACF, after which, most others are not significant.
- d equals 1 because we had first differences
- q should be somewhere around 4 as well as seen on the ACF
- P might be 2, since 24-th and 48-th lags are somewhat significant on the PACF
- D again equals 1 because we performed seasonal differentiation
- Q is probably 1. The 24-th lag on ACF is significant while the 48-th is not.

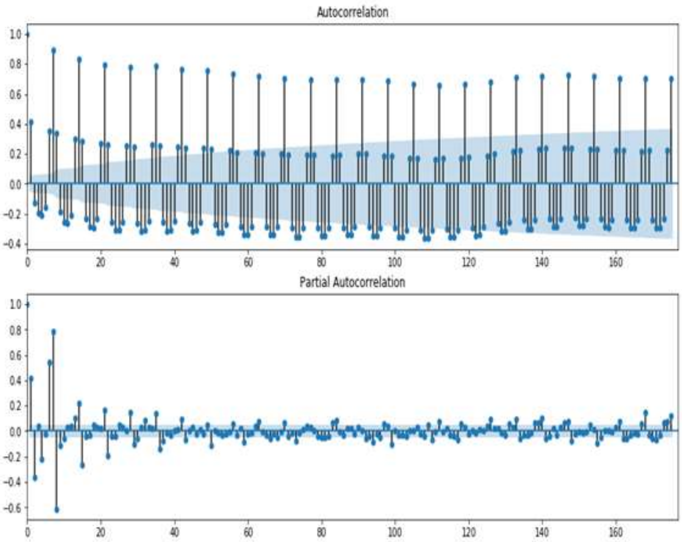

Original
Time
Autocorrelation
Partial Autocorrelation
Seasonal Difference 24
Time
Autocorrelation
Partial Autocorrelation
Seasonal Difference 24 + Usual Difference
Time
Autocorrelation
Partial Autocorrelation

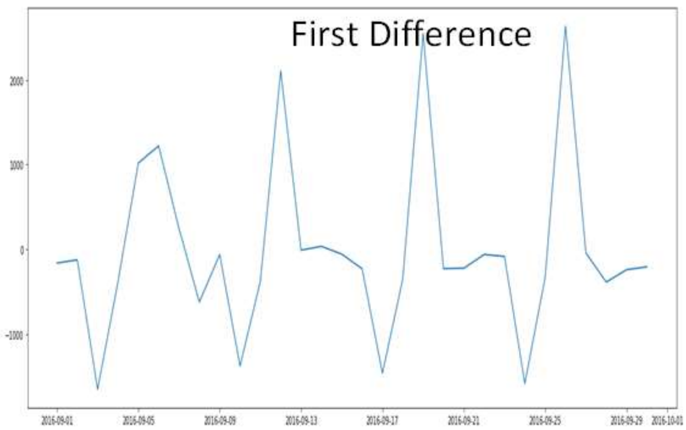


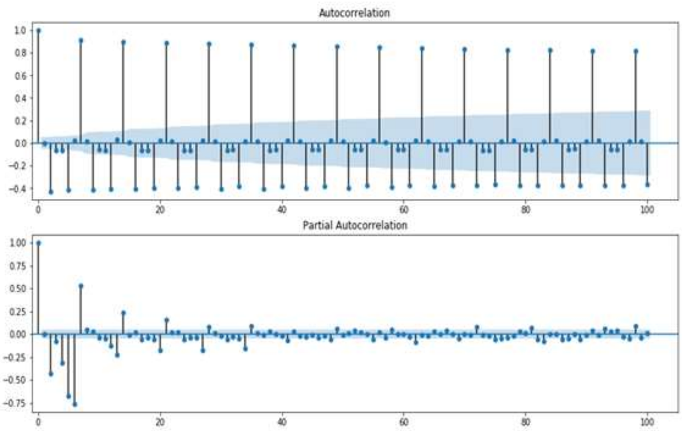


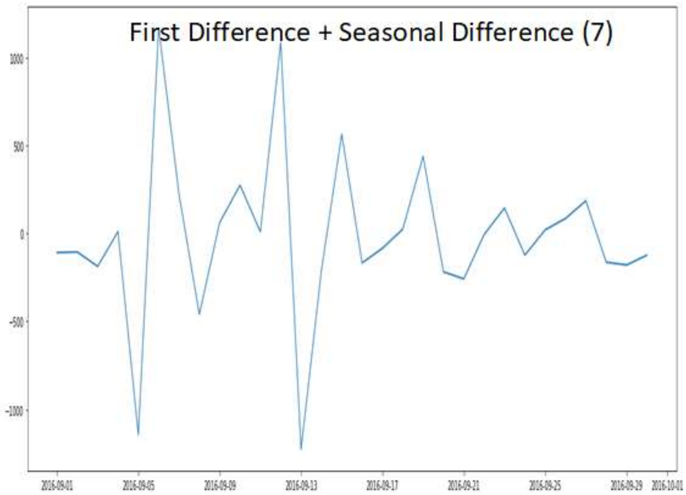
Example 2

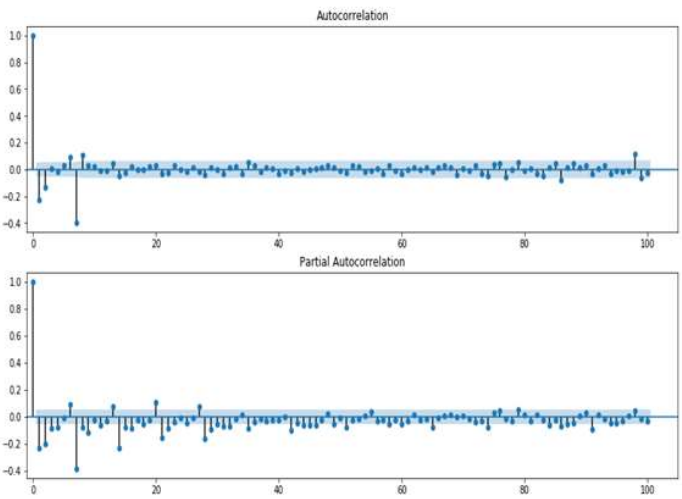

Electricity_Load_Forecasting

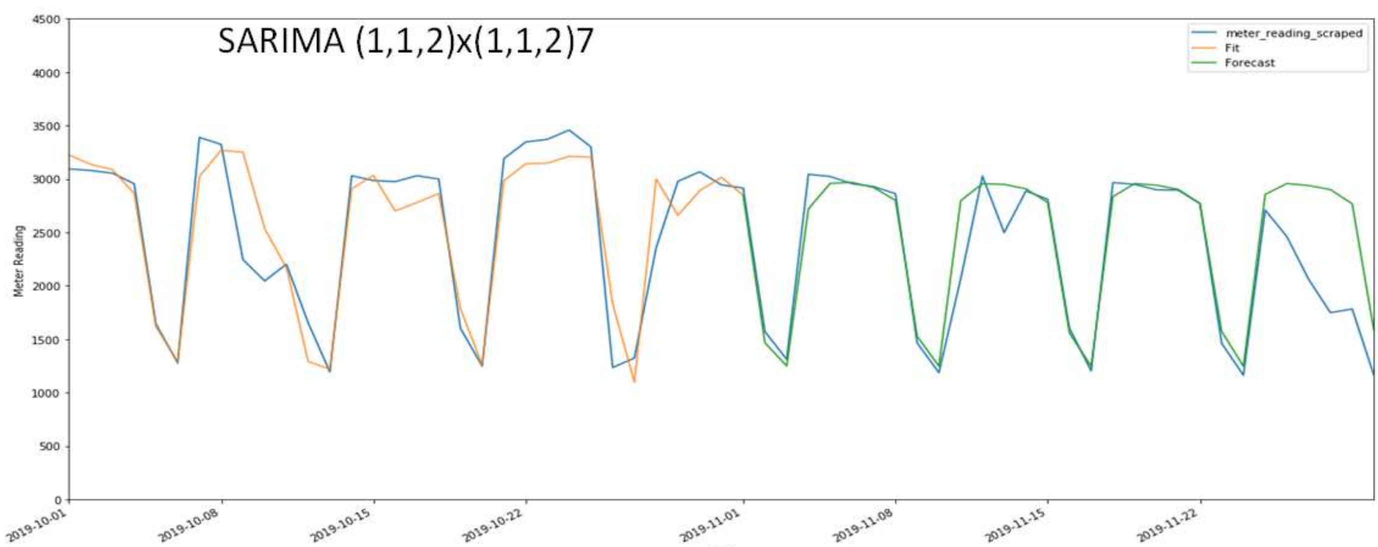

Autocorrelation
Partial Autocorrelation


First Difference


Autocorrelation
Partial Autocorrelation

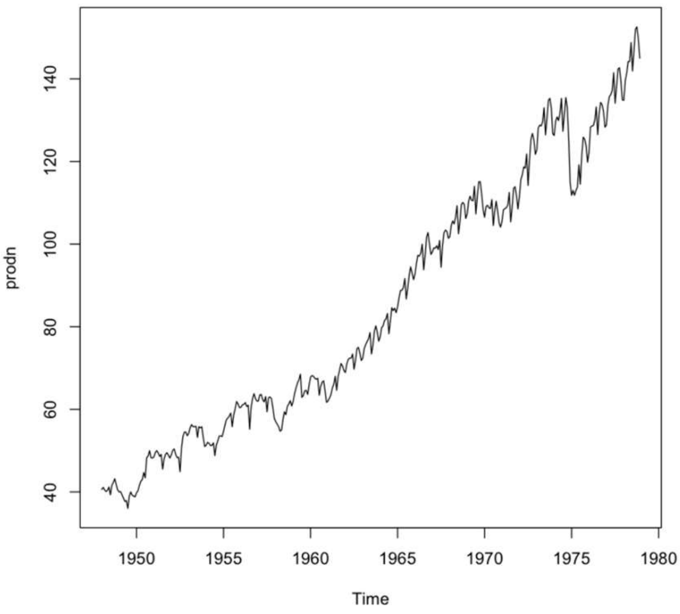

First Difference + Seasonal Difference (7)

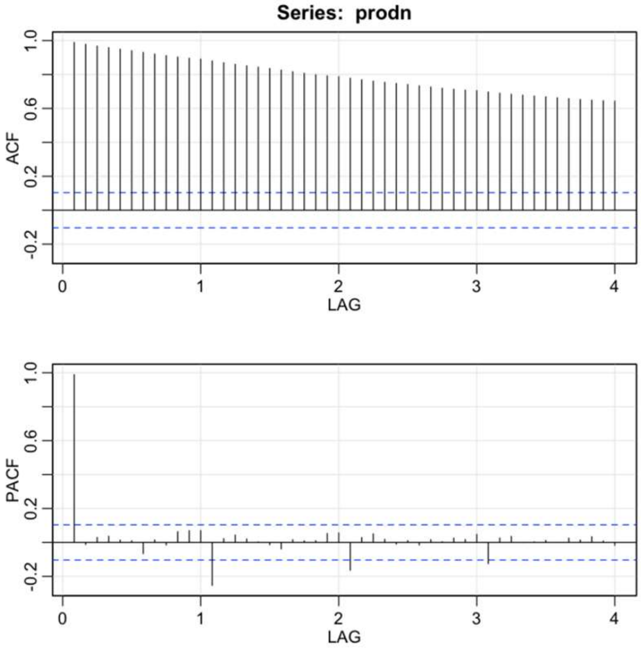

Autocorrelation
Partial Autocorrelation

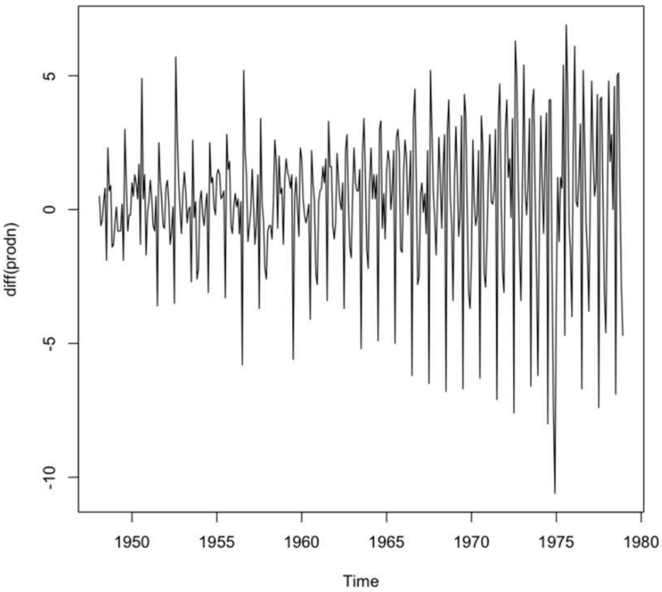

SARIMA (1,1,2)x(1,1,2)7
meter_reading_scraped
Fit
Forecast
Meter Reading

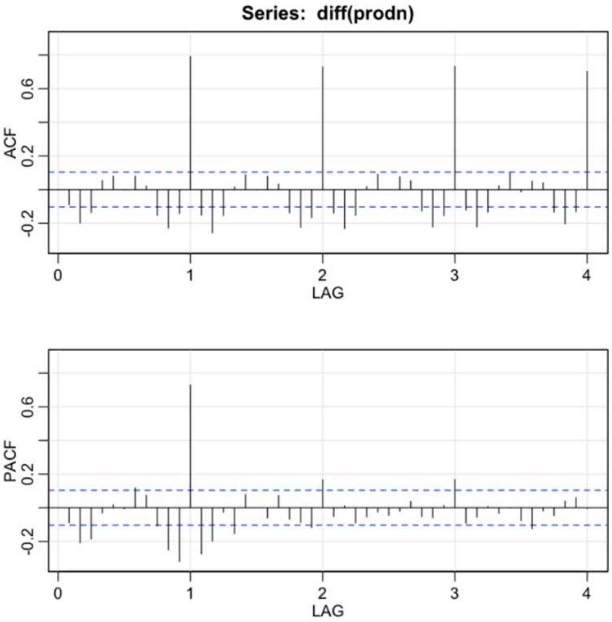
Time

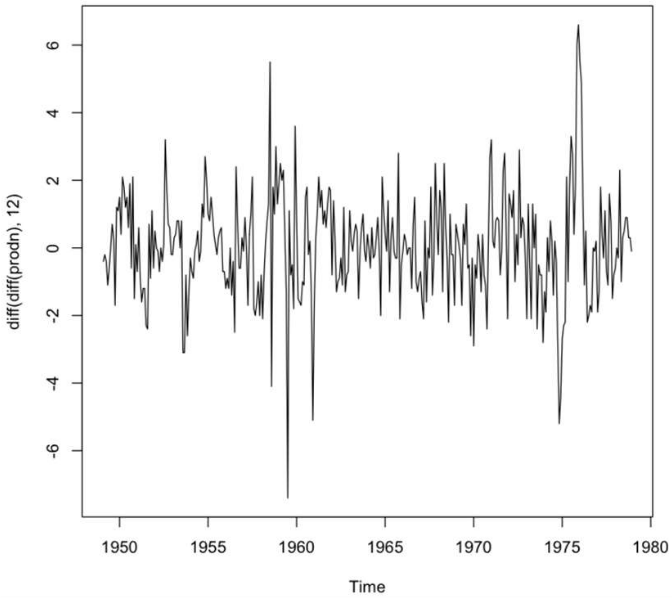
Example 3

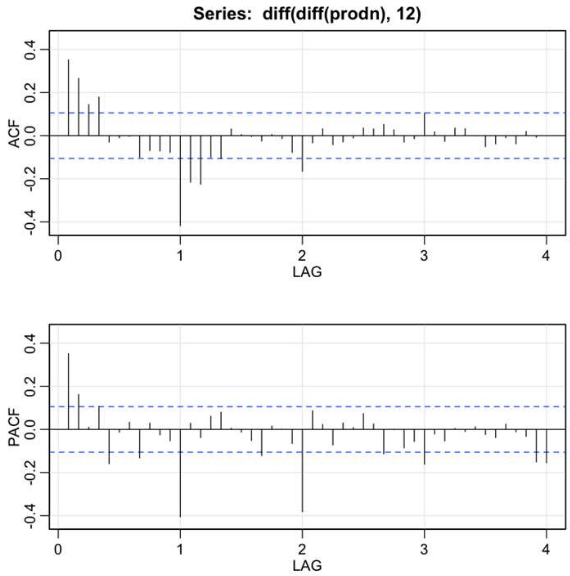

prodn
Time

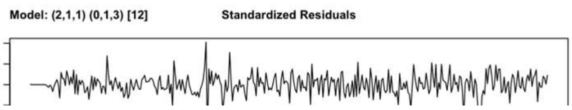

Series: prodn
ACF
PACF
LAG

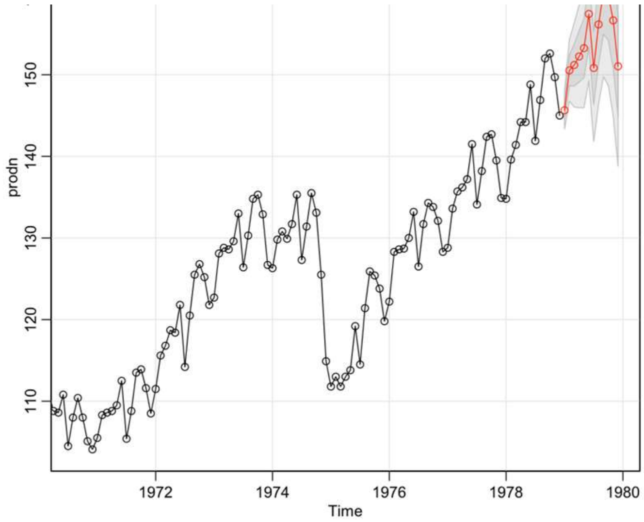

diff(prodn)
Time

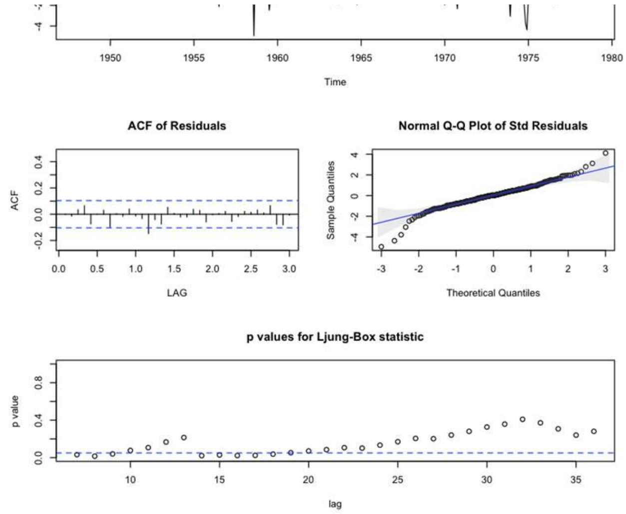

Series: diff(prodn)
ACF
PACF
LAG


diff(diff(prodn), 12)
Time

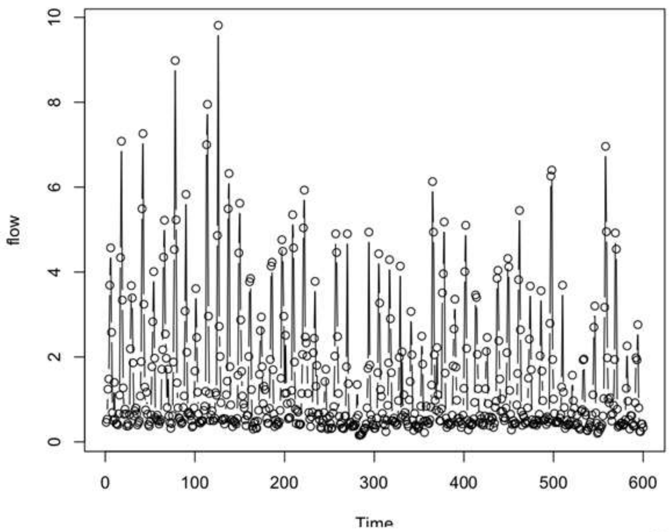

Series: diff(diff(prodn), 12)
ACF
PACF
LAG

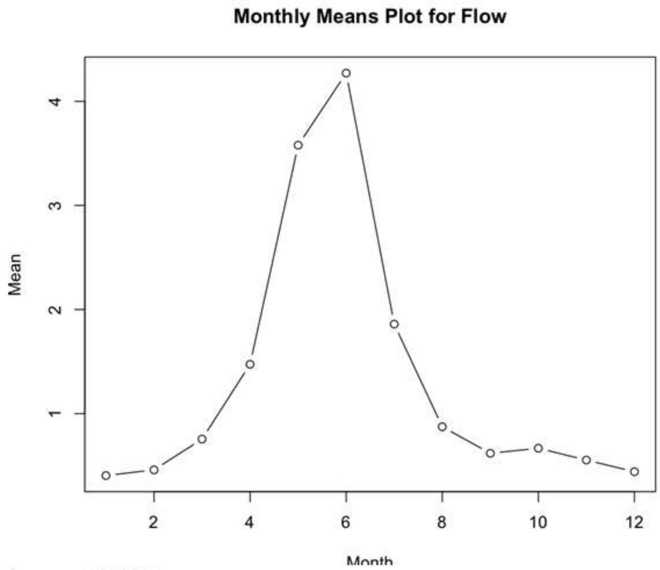

Model: (2,1,1) (0,1,3) [12]
Standardized Residuals

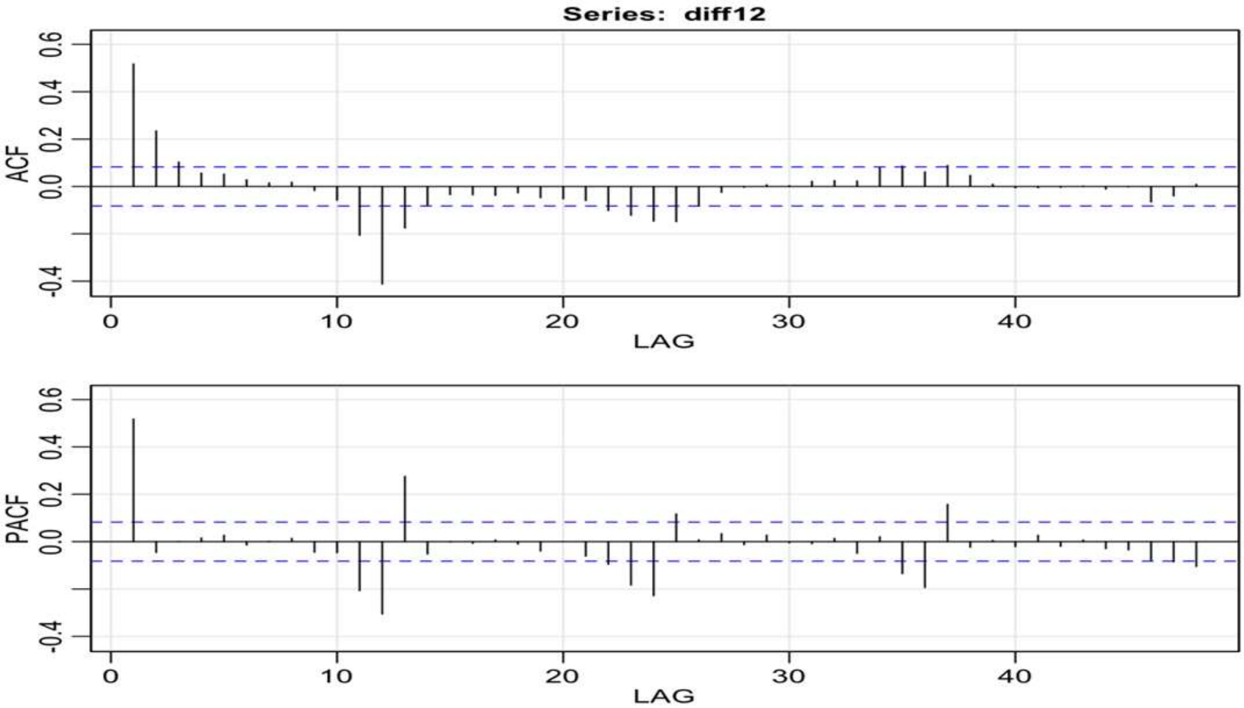




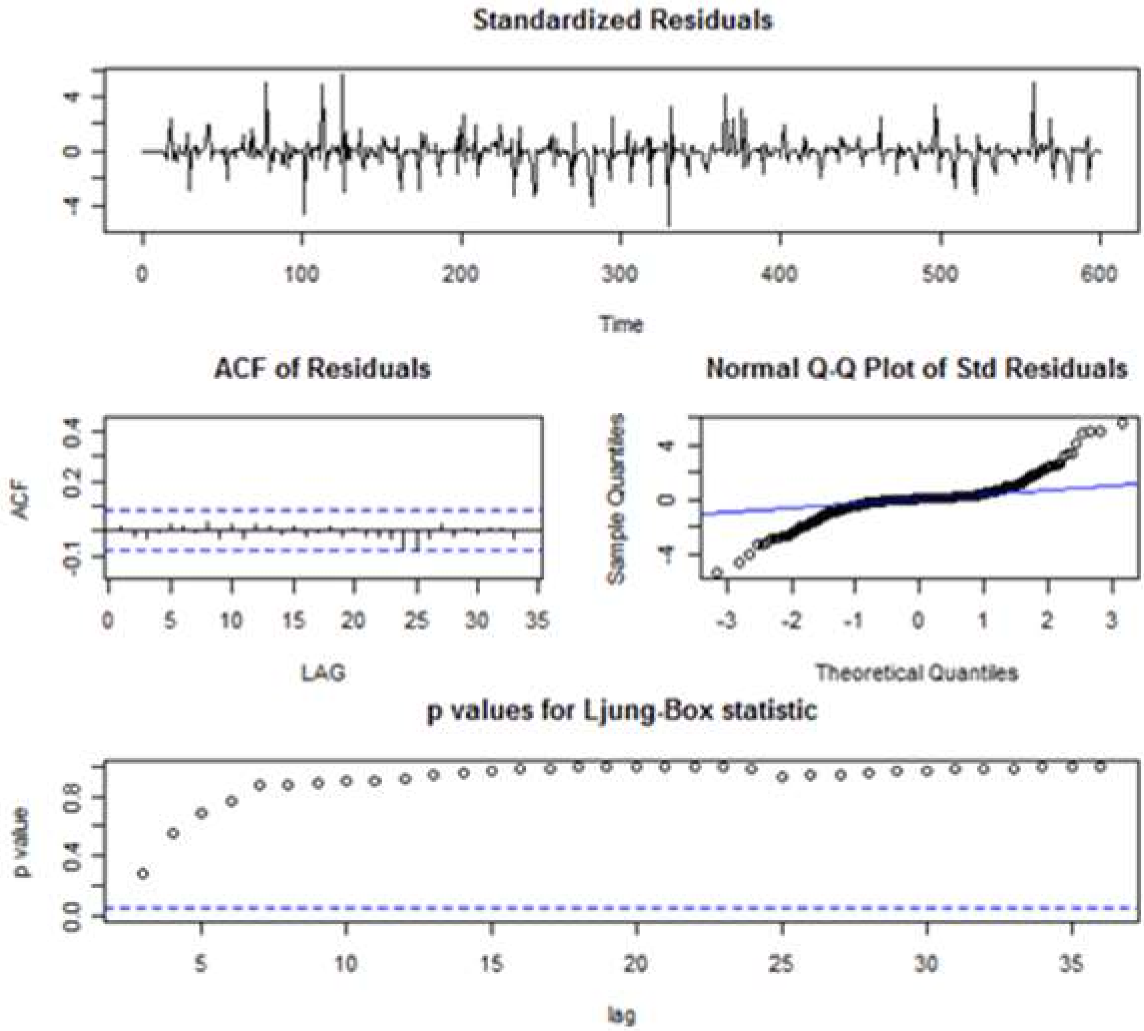
Example 4. The data series are a monthly series of a measure of the flow rate of the Colorado River, n = 600 consecutive months.

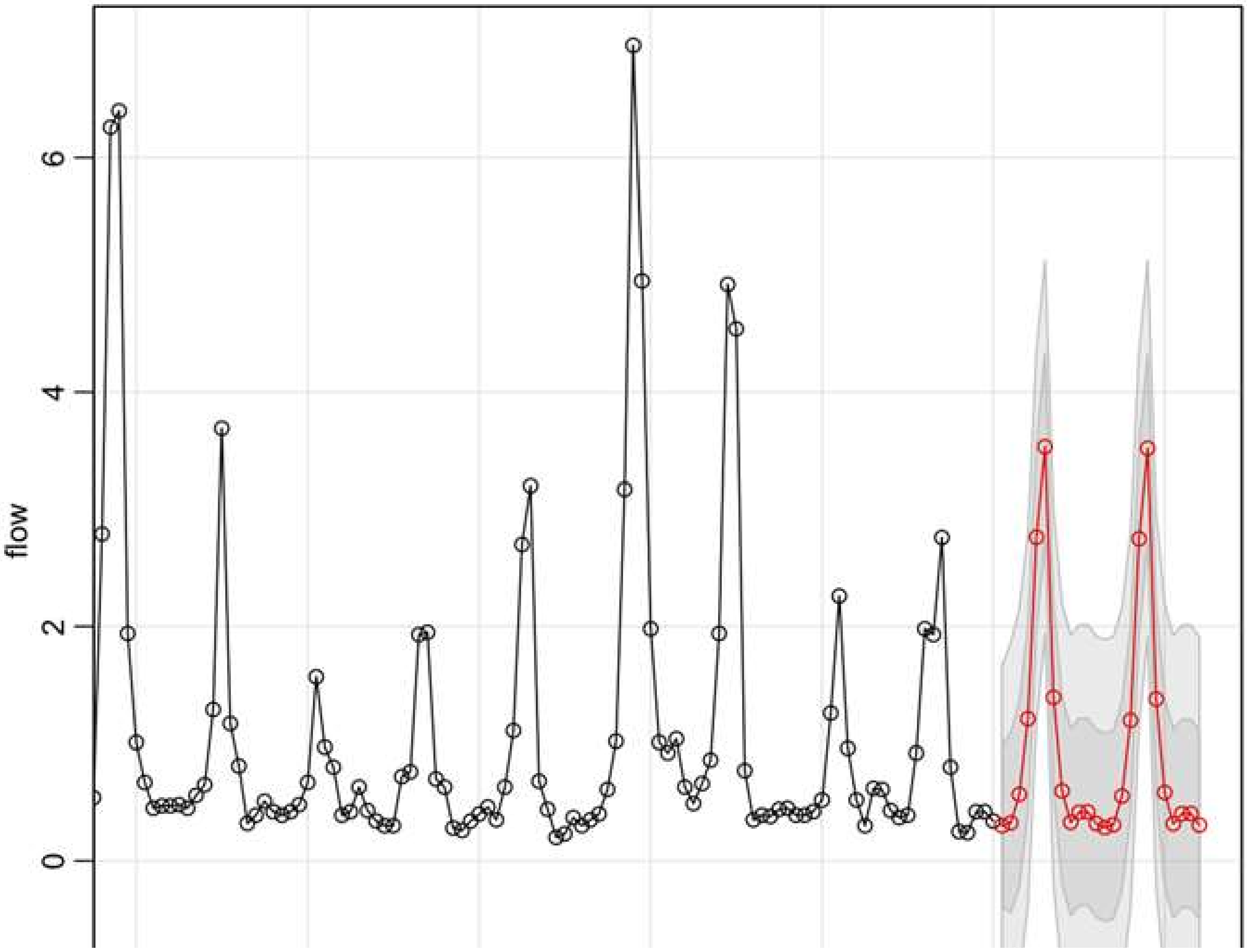


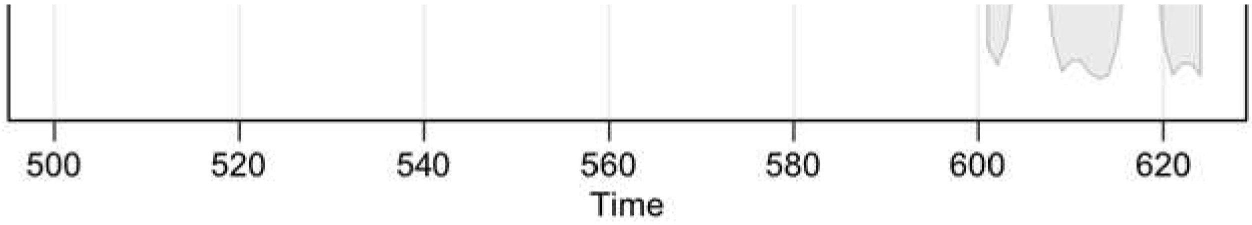




1. The PACF shows a clear spike at lag 1 and not much else until about lag 11. This is accompanied by a tapering pattern in the early lags of the ACF. A non-seasonal AR(1) may be a useful part of the model.
2. The PACF has significant lags at 12, 24, 36 and so on. This is seams that we need to set seasonal MA(1).


Standardized Residuals
Time
ACF of Residuals
ACF
LAG
Normal Q-Q Plot of Std Residuals
Sample Quantiles
Theoretical Quantiles
p values for Ljung-Box statistic
p value
lag


flow



Here are the matching table between ETS and some SARIMA models

ETS model	ARIMA model
ETS(A,N,N)	ARIMA(0,1,1)
ETS(A,A,N)	ARIMA(0,2,2)
ETS(A,A_d,N)	ARIMA(1,1,2)
ETS(A,N,A)	ARIMA(0,1,m)(0,1,0)$_m$
ETS(A,A,A)	ARIMA(0,1,$m+1$)(0,1,0)$_m$
ETS(A,A_d,A)	ARIMA(0,1,$m+1$)(0,1,0)$_m$

ARMA coefficient estimation

There are exist a several methods for parameters of AR and MA process estimation. **ARMA specific metrics**

> The ARMA parameters optimization can be performed with several specific metrics designed for this purpose. The most popular of them are the following.

- $AIC(k) = ln\hat{\sigma}^2 + 2\frac{k}{N}$; Akaike information criterion

- $BIC(k) = ln\hat{\sigma}^2 + lnN\frac{k}{N}$; Bayesian information criterion
- $AICc = AIC + \frac{2k(k+1)}{N-(k+1)}$, modified Akaike information criterion
- $HQC = N \cdot \ln\frac{RSS}{N} + 2 \times k \times \ln(\ln N)$, Hannan–Quinn information criterion

where

$$\hat{\sigma}^2 = RSS/(N-k)$$

- is the estimation of the model likelihood; k is the full number of parameters:

- for ARMA without constant $k = p + q$ in other case $k = p + q + 1$;
- for ARIMA $k = p + q + d + 1$;

  for SARIMA $k = p + q + d + P + Q + D + 1$.

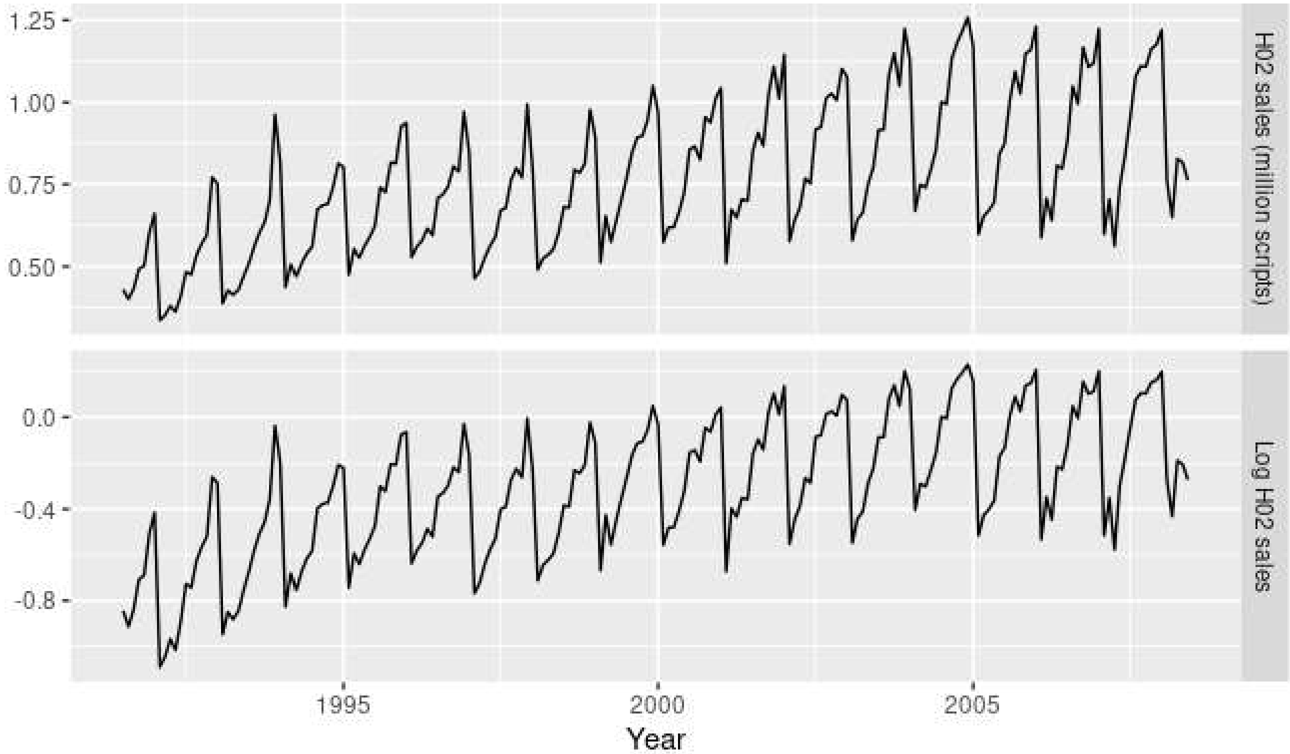
  The the specificity of these criteria lies in the joint optimization the number of parameters and RSS value.

  Between these criteria AIC is the most popular, but AICc is suppose to provide better result.

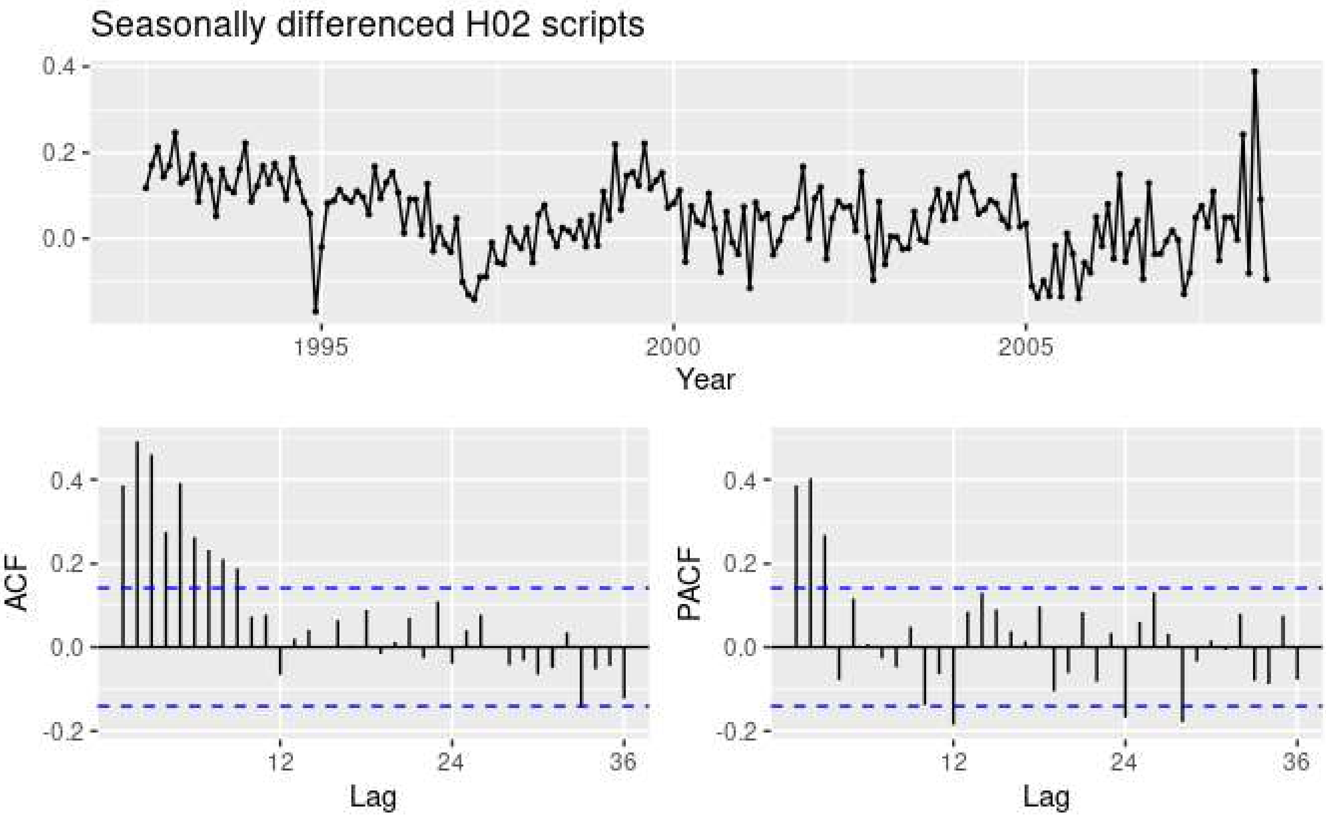
  The criteria could be applied of automatization parametric search routine.

Example: sales prediction. On the original plot there is a small increase in the variance with the level, so we take logarithms to stabilise the variance.

the example taken [here (https://otexts.com/fpp2/seasonal-arima.html)](https://otexts.com/fpp2/seasonal-arima.html).



The data are strongly seasonal and obviously non-stationary, so seasonal differencing will be used with seasonality 12.



It is not clear at this point whether we should do another difference or not.

We decide not to, but the choice is not obvious.

On the obtained plot there are spikes in the PACF at lags 12 and 24, but nothing at seasonal lags in the ACF. This may be suggestive of a SAR(2) term. In the non-seasonal lags, there are three significant spikes in the PACF, suggesting a possible AR(3) term. The pattern in the ACF is not indicative of any simple model.

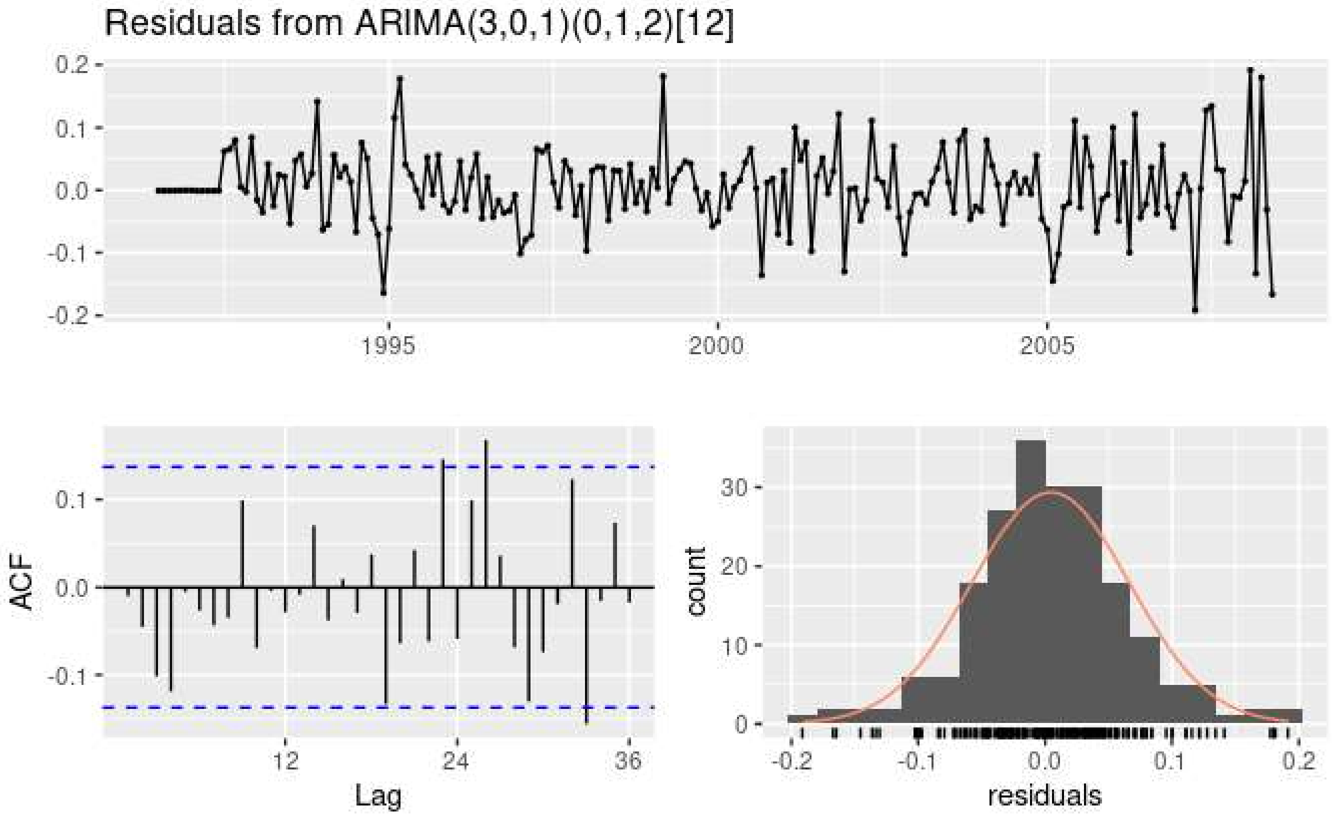
This initial analysis suggests that a possible model for these data is an

ARIMA(3,0,0)(2,1,0)12. It would be better to check this guess.

The variations of parameters decline the first guess - it is better to use ARIMA(3,0,1)(0,1,2)12 model.

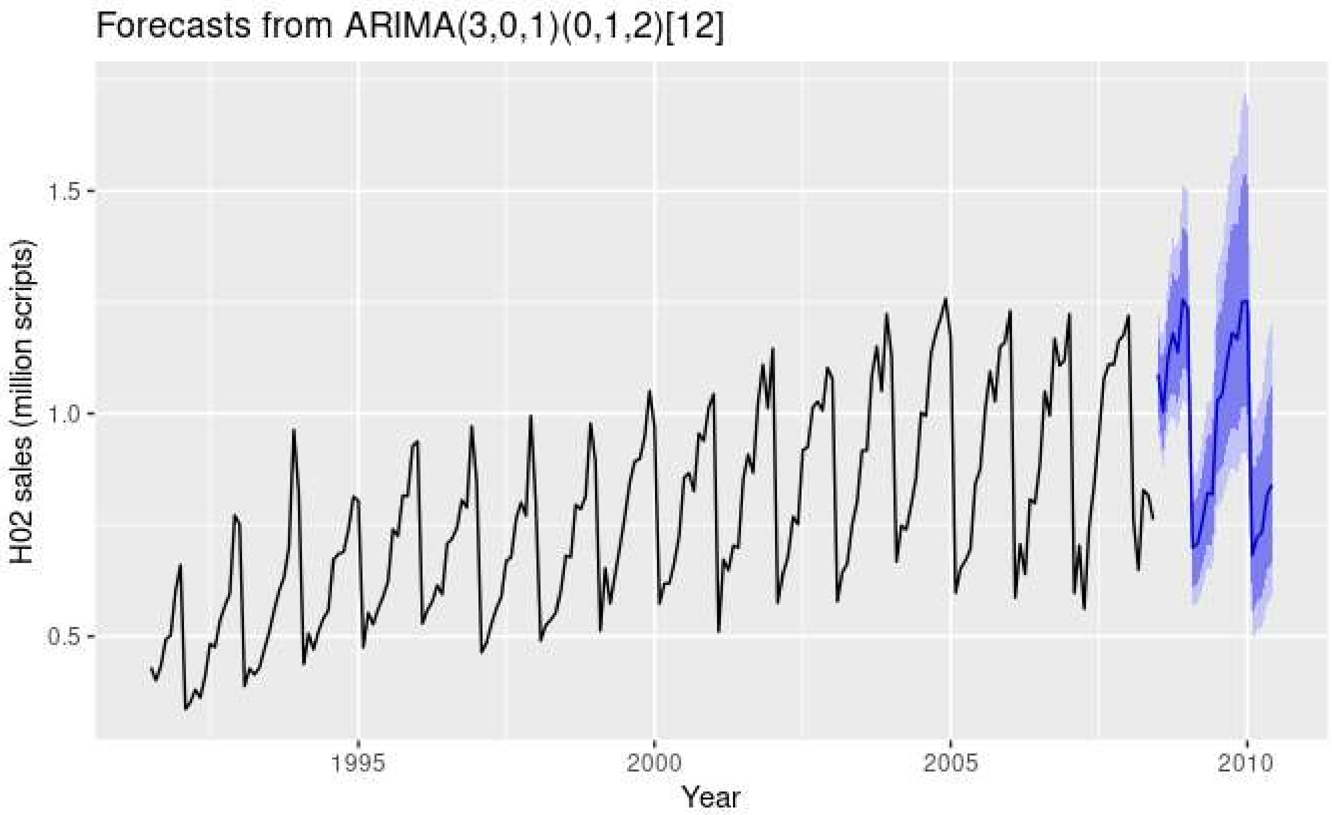
Model	AICc
ARIMA(3,0,1)(0,1,2)$_{12}$	-485.5
ARIMA(3,0,1)(1,1,1)$_{12}$	-484.2
ARIMA(3,0,1)(0,1,1)$_{12}$	-483.7
ARIMA(3,0,1)(2,1,0)$_{12}$	-476.3
ARIMA(3,0,0)(2,1,0)$_{12}$	-475.1
ARIMA(3,0,2)(2,1,0)$_{12}$	-474.9
ARIMA(3,0,1)(1,1,0)$_{12}$	-463.4

Then it is need to estimate the residuals



Then it is need to estimate the residuals There are a few significant spikes in the ACF thus the model fails the Ljung-Box test.

The model can still be used for forecasting, but the prediction intervals may not be accurate due to the correlated residuals.



1.4 Other SARIMA-Like Models

Then it is need to estimate the residuals Beside discussed above SARIMA model it usually considered a several other:

1. **SARIMAX** - SARIMA model with the exageneous addition factors.
2. **ARIFMA**- ARIMA model with uisng fractional derivatives (and fractional integration).

3. **VAR, VMA, VARMA** - multivariate models.
4. **NAR, NARX, NARMA** -Non-linear AR (NAR), Non-linear AR with exageneous factors (NARX), non-linear ARMA (NARMA) and e.t.c.

The **exogenous factor (exdogs, exogenous covariates)** x_n or a set of factors ($\sum_{j=1}^{r} x_j$) is the addition to aim-variable factors which are statistically independent from y_n but influence on it.

For example, if the previous week's average price of oil x_n, affects this week's exchange rate y_n, then x_n can be considered as exogenous time series. In this case ARIMAX can be use.

You can the SARIMAX model to check if a set of exogenous variables has an effect on a linear time series.

For **ARMAX** The factors can be taken into account as

$$\hat{y}_n = \sum_{i=1}^{p} \alpha_i y_{n-i} + \sum_{i=0}^{q} \beta_i \varepsilon_{n-i} + \sum_{i=1}^{r} \gamma_i x_n^{\{i\}},$$

where γ_i is the coefficient of exdogs influence; $x_n^{\{i\}}$ are the n-th sample of the i-th exogenous factor. In more generalized form

SARIMAX model can be give as

$$\phi(L)\Phi(L^s)\Delta^d \Delta_s^D y_n = c(n) + \theta(L)\Theta(L^s)\varepsilon_n + \Gamma x_n,$$

where Γ is the coefficients for $x_n^{\{i\}}$; x_n is the vector of $x_n^{\{i\}}$ for each i.

In opposite to exogenous factors the factors which are obviously correlate with the series can be considered as **endogenous(endogs)**.

The example of the endogenous factors is the Rainfall to plant growth - which are correlated and studied by economists since the amount of rainfall is important to commodity crops such as corn and wheat.

In [ ]:

```

```