

Forecasting av strømpriser

Data Science/Analytisk nettverksmøte

15. Mars

Valentyna Pysarieva, Analytical Consultant

&

Pia Skare Rønnevik, Customer Success Manager, FANS



Airline Data

A *time series* is an indexed set of *equally spaced* numbers.

Airline Passengers in 1000s U.S. Carriers

70000
60000
50000
40000
30000

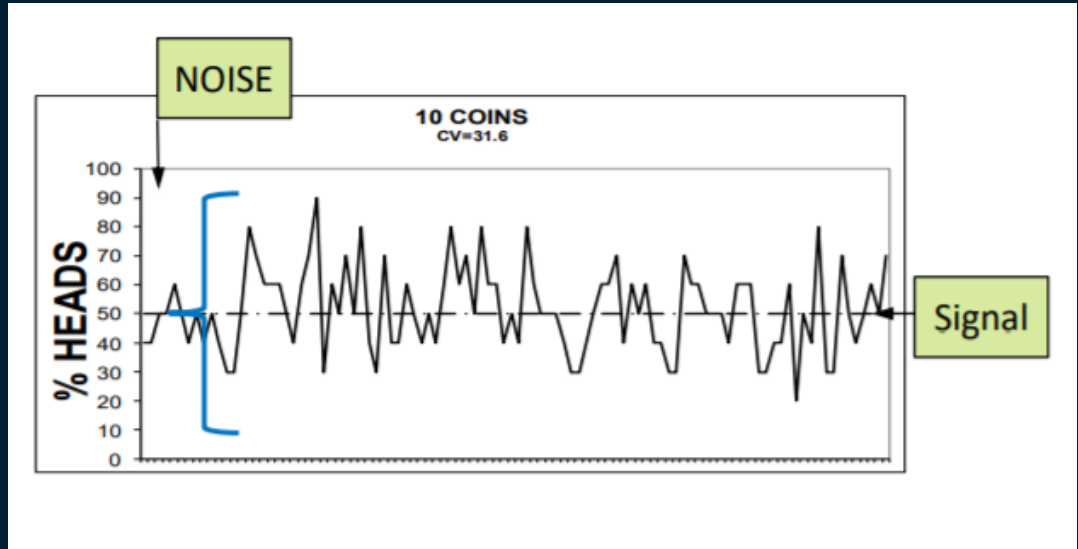
1990 1992 1994 1996 1998 2000

D

- A **time series** is an ordered sequence of values of a variable at equally spaced time intervals.
- **Time series analysis** accounts for the fact that data points taken over time may have an internal structure (such as autocorrelation, trend or seasonal variation) that should be accounted for.
- **Time series forecasting** is the use of a model to predict future values based on previously observed values.

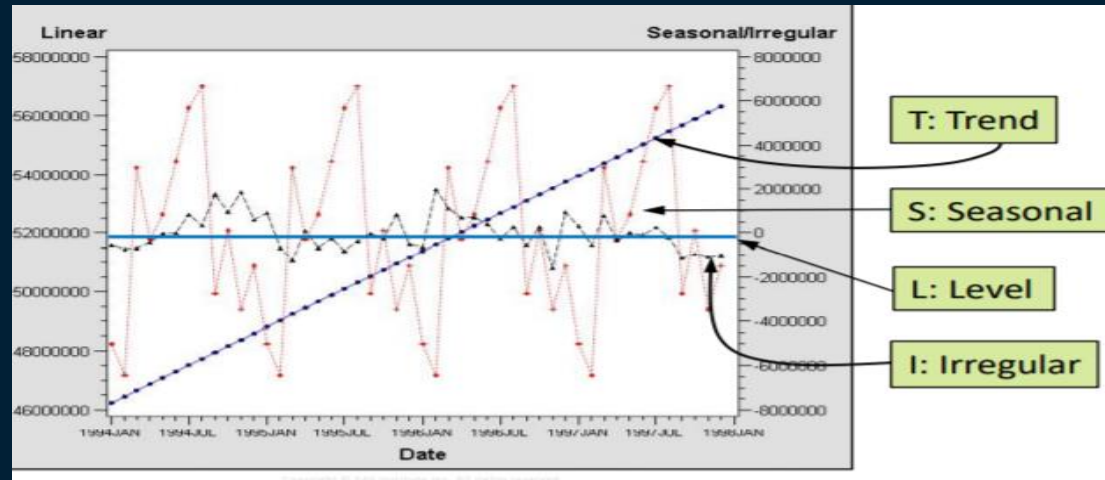
Variation in time series data

- Signal
- Noise



Signal Components

- **Level:** the average baseline for any time point
- **Trend:** the long-term pattern of the average of a serie.
- **Seasonality:** can be considered predictable deviations from the trend component. The pattern of the deviations has a regular period, for example a seasonal pattern in December each year.
- **Cycle:** is a repetitive cycle that does not have a regular period. A cycle can last for example 20 months/25 months.
- **Exogenous:** are not actually components of the series. However, they are sources of variation in the series that may be due to the values of another series. They are sometimes known as explanatory effects.
- **Irregular:** is what remains after the patterns for all other effects and components have been accounted for. That is the remaining variation.



Autocorrelation

- Autocorrelation simply means that current values in a time series are related with previous values.
- Autocorrelation is part of the signal in the irregular component.
- The irregular part of a series contains both noise and signal.

Exponential Smoothing Models (ESM)

Time series models

- Use weighted averages of past observations to forecast new values.
- Gives more weights to the recent values than older observations. Thus, as observations get older (in time), the importance of these values get exponentially smaller.
- Combines Error, Trend, and Seasonal components in a smoothing calculation (ETS). Each components can be combined either additively, multiplicatively, or be left out of the model.
- Work well when the time series shows a clear trend and/or seasonal behavior.

Time series models

AutoRegressive Integrated Moving Average with eXogenous variables Models (ARIMAX)

- AR: Autoregressive → Time series is a function of its own past.
- I: Integrated → Differenced values between successive time points can be modeled, and after modeling returned to the undifferenced metric.
- MA: Moving Average → Time series is a function of past shocks (deviations, innovations, errors and so on)
- X: Exogenous → Time series is influenced by external factors.

Unobserved Components Models (UCMs)

Time series models

- Also known as structural time series models.
- Can be considered as a multiple regression model with time-varying coefficients.
- Performs a time series decomposition into components:

$$Y_t = \text{Trend} + \text{Season} + \text{Cycle} + \text{Regressors}$$

- The components in the model have their own models and its own source of error and forecast.

Time series models

- Often used to compare your complex statistical models against:
 - Simple mean models
 - Simple random walk
 - Random walk drift
- The complex statistical models are only valuable if they perform better than Naive models.

Simple Regression Models

Time series models

- A good substitute if other forecasting models becomes problematic.
- Time can be used as an input variable.
- Seasonality can be modeled using dummy variables representing each season.
- Predefined trend components: linear, quadratic, cubic, log-linear, exponential, and so on.

Performance

- **Simple models:** have no performance issues.
- **ESM:** can be constructed quickly and easily, so they always have good performance.
- **ARIMAX:** require many more computer cycles than simple or exponential smoothing models, so some approximations and shortcuts are used to speed performance.
- **UCM:** are very computer-intensive and should be tried only on small data sets or individual time series.

Measurements of model fit

Mean Absolute Percent Error (MAPE)

- MAPE is one of the most common accuracy measures in business forecasting.
- As a selection criterion, choose the model with the smallest value of MAPE.
- MAPE is the average of all of the individual absolute percent errors.

Mean Absolute Error (MAE)

- MAE is not commonly used as an accuracy measure in business forecasting.
- As a selection criterion, choose the model with the smallest value of MAE.
- The average of all of the individual absolute errors.

Measurements of model fit

Root Mean Square Error (RMSE)

- RMSE is the square root of the average of all of the individual squared errors, adjusted for the number of estimated model parameters.
- RMSE is commonly used as an accuracy measure in industrial, economic, and scientific forecasting.
- As a selection criterion, choose the model with the smallest value of RMSE.

Simulating a Prospective Study in Model Studio

Divide the time series data into two segments:

- The fit sample is used to derive a forecast model.
- The holdout sample is used to evaluate forecast accuracy.

Full = Fit + Holdout data is used to fit a deployment model.

DEMO

Faktorer som påvirker strømprisene

Strømprisene påvirkes ikke kun av tilbud og etterspørsel, og her er noen av de andre faktorene:

- Kraftproduksjonen i Norge, samarbeidslandene og i land utenfor vårt kraftmarked
- Prisene på strøm utenfor vårt kraftmarked
- Forbrukernes strømforbruk
- Mengden strøm vi eksporterer og importerer til utlandet
- Dollarkursen / Eurokursen
- Prisen på CO₂-utslipp
- Vedlikeholdsarbeid
- Kabelkvaliteten på strømmettet
- Overføringsbegrensninger i strømmettet, også kalt flaskehals
- Vær og temperaturer som fører til høyere energiforbruk

Datagrunnlaget

- **Strømpriser fra Nord Pool**, oppgitt i øre/KWh inkludert MVA, <https://www.los.no/dagens-strompris/historiske-strompriser/>
 - NO1: Sørøst-Norge - Oslo
 - NO2: Sørvest-Norge - Kristiansand
 - NO3: Midt-Norge - Trondheim
 - NO4: Nord-Norge - Tromsø
 - NO5: Vest-Norge - Bergen
- **Elektrisitetsbalanse SSB**, Tabell 12824: Produksjon og forbruk av elektrisitet MWh, <https://www.ssb.no/statbank/table/12824>
 - 1 Total produksjon
 - 1.1 Vannkraft
 - 1.2 Varmekraft
 - 1.3 Vindkraft
 - 2 Import
 - 3 Eksport
 - 4 Bruttoforbruk
 - 5 Pumpekraftforbruk
 - 6 Tap i linjenettet
 - 7 Nettoforbruk
 - 8 ForbrukRåoljeNaturgass: Forbruk i utvinning av råolje og naturgass
 - 9 ForbrukIndustri: Forbruk i kraftintensiv industri
 - 9.1 ForbrukPapir: Forbruk i produksjon av papirmasse, papir og papp
 - 9.2 ForbrukKjemiskevarer: Forbruk i produksjon av kjemiske råvarer
 - 9.3 ForbrukMetall: Forbruk i produksjon av jern, stål og ferro
 - 9.4 ForbrukAndreMetaller: Forbruk i produksjon av aluminium og andre metaller
 - 10 ForbrukAlminnelig: Forbruk i alminnelig forsyning

Datagrunnlaget

- **Observasjoner og værstatistikk**, Norsk klimaservicesenter, <https://seklima.met.no/observations/>
 - **Maksimumstemperatur**: Høyeste registrerte temperatur per måned.
 - **Middelvind**: Høyeste middelvind per måned.
 - **Nedbør**: Antall dager med nedbør.
 - **Snødekke**: Månedsmiddel av snødekke, kode:
 - 0=ikke snø
 - 1=mest bar mark
 - 2= like mye snødekt som bar mark
 - 3=mest snødekt mark
 - 4=snø overalt
 - **Solskinnstid**: Hvor mange timer med sol siste måned.
- **Valutakurser**, Norges Bank, <https://www.norges-bank.no/tema/Statistikk/Valutakurser/?tab=currency&id=EUR>
 - Dollarkursen
 - Eurokursen

Våre mål med Forecasting

- Prøve å forecaste strømprisene i 2023 i både VA og Model studio, og sammenligne metodene.
- Prøve å forecaste strømprisene i 2022 for å se hvor godt vi treffer med de to ulike metodene.