KI Service Zentrum by Hasso-Plattner-Institut

Gefördert durch:
Bundesministerium für Forschung, Technologie und Raumfahrt

# Time Series Forecasting

## 1.8 Statistical Methods II

## Autoregressive models

Mario Tormo Romero

**Design IT.**
**Create Knowledge.**

# What we`ll cover in this video

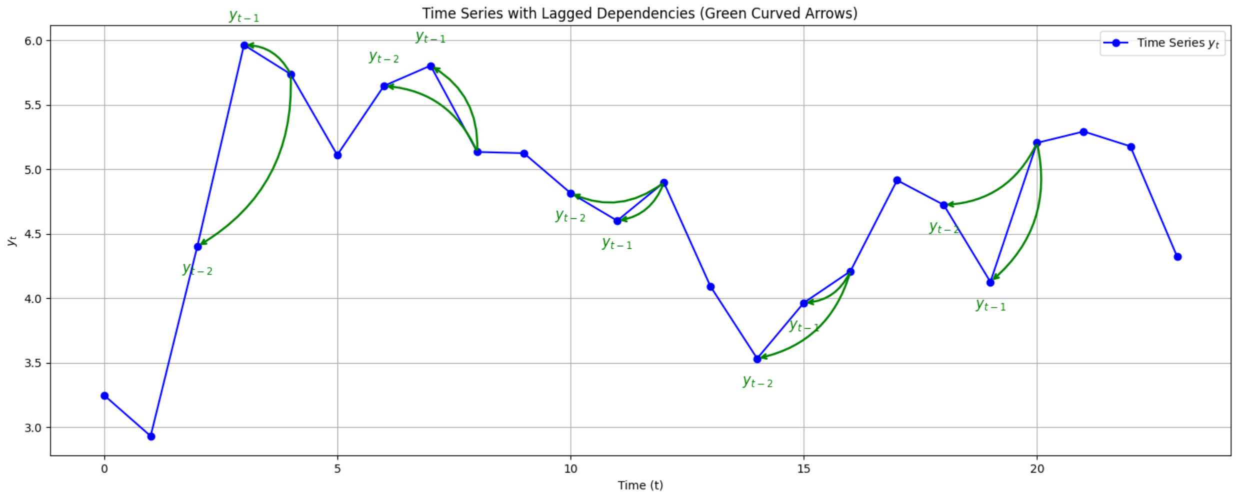
- Introduction to Autoregressive (AR) Models
- Moving Average (MA) Models
- ARMA and ARIMA Models — Handling Trends and Stationarity
- Seasonal Extensions: SARIMA, AutoARIMA, ARIMAX/SARIMAX

# Autoregressive (AR) Models

- Models that predicts future values using a linear combination of past values
- Work best when the time series is stationary
- The number of past values used is called the order $p$
- Formula:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t$$

  - $y_t$ is the value at time $t$
  - $c$ is a constant
  - $\phi_i$ are model coefficients
  - $\epsilon_t$ is the error term (noise)

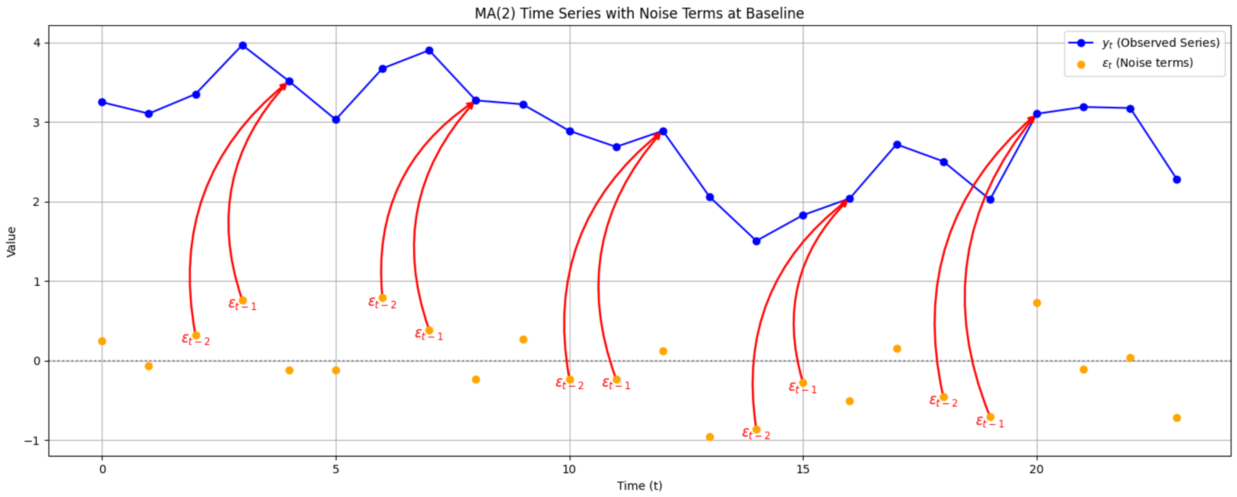



$$y_t = 3 + 0.7y_{t-1} - 0.3y_{t-2} + \varepsilon_t$$

# Moving Average (MA) Models

- Instead of using past values, MA models use past forecast errors to predict the current value
- The number of past errors considered is called the order q
- Formula:
- $y_t = \mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q}$

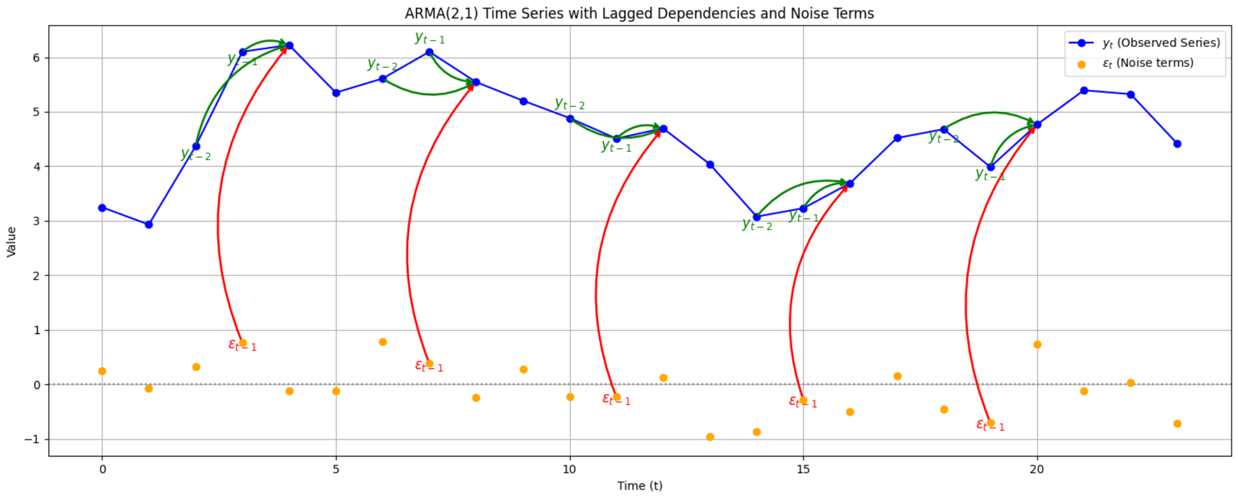
  - $y_t$ is the value at time $t$
  - $\mu$ is the mean of the series
  - $\epsilon_t$ is the current error term
  - $\theta_i$ are coefficients for past errors




$y_t = 3 + \varepsilon_t + 0.7\varepsilon_{t-1} + 0.3\varepsilon_{t-2}$

# ARMA Models (AutoRegressive Moving Average)

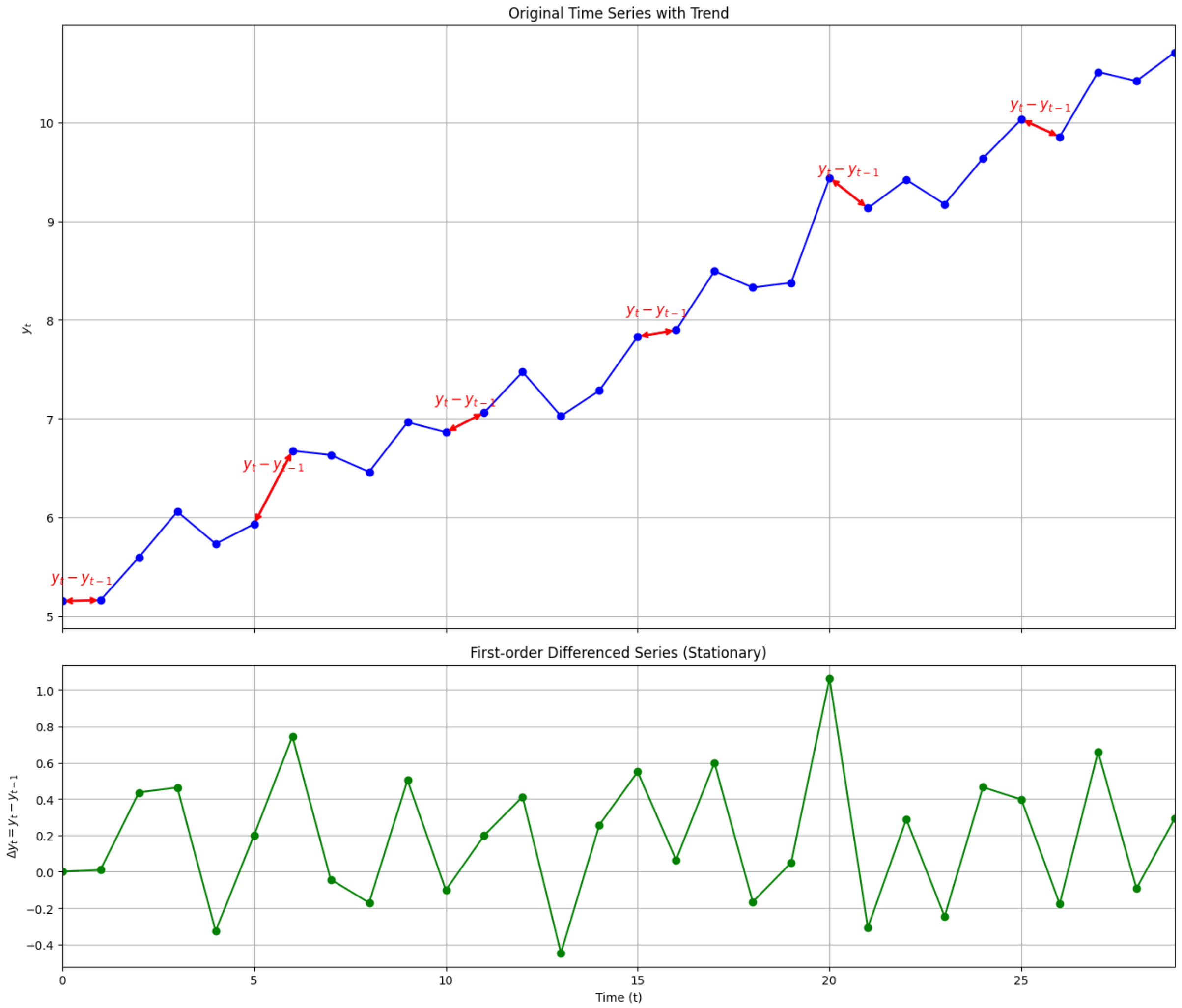
- Combines AR (AutoRegressive) and MA (Moving Average) models
- Uses both past values and past forecast errors to predict the current value
- Suitable for stationary time series
- Captures complex patterns by balancing memory of past data and adjustments for recent shocks




$$y_t = 3 + 0.7y_{t-1} - 0.3y_{t-2} + 0.5\varepsilon_{t-1} + \varepsilon_t$$

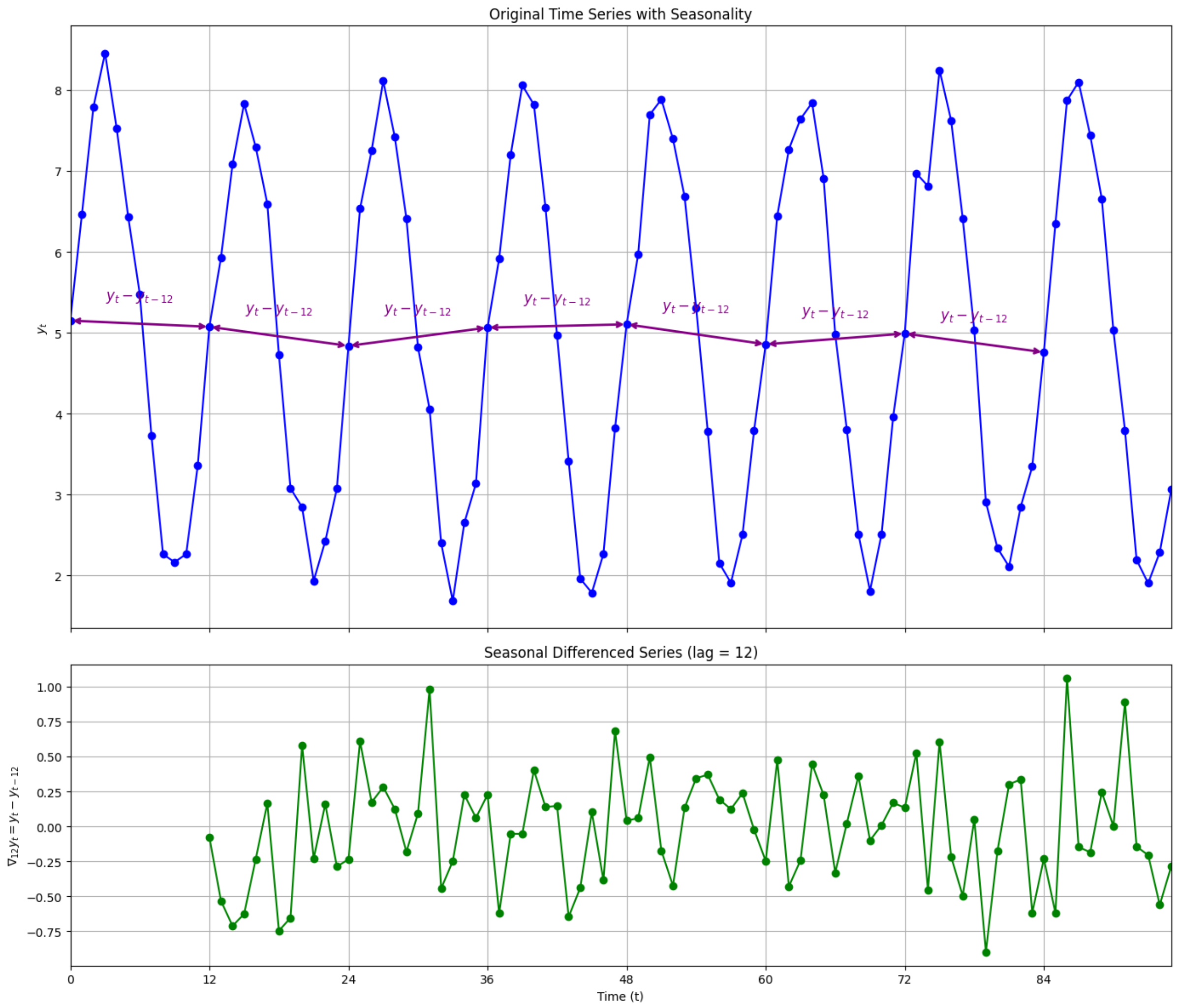
# ARIMA Models (AutoRegressive Integrated Moving Average)

- Extends ARMA to handle non-stationary time series. Adds the ‘Integrated’ (I) part, which means applying differencing to remove trends and stabilize the series
- Differencing means looking at the changes between consecutive points instead of raw values
- Very flexible and powerful for real-world data with trends or slow changes
- Parameters:
  - p = order of AR
  - d = number of differencing steps
  - q = order of MA



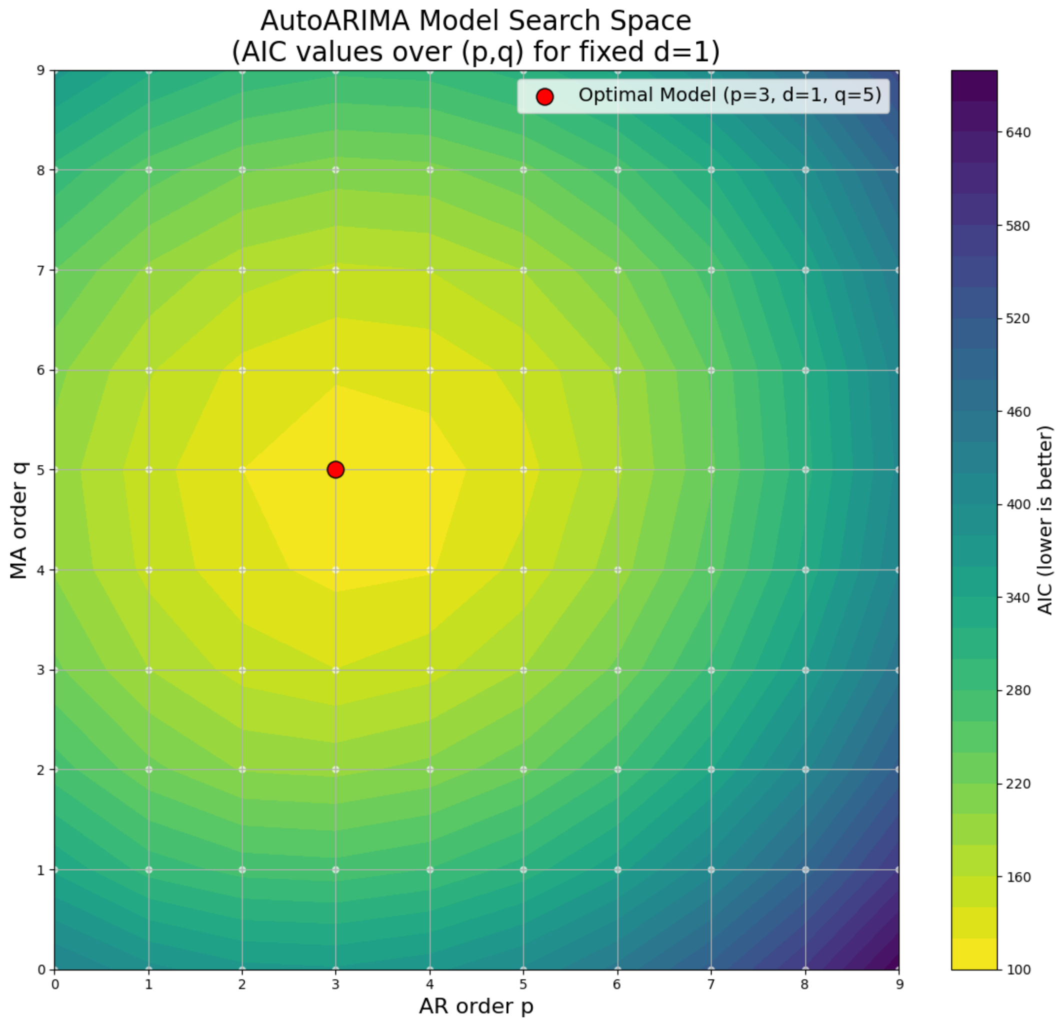
# SARIMA Models (Seasonal ARIMA)

- Extends ARIMA to handle seasonality
- Adds seasonal terms to the AR, I, and MA components
- Captures repeating seasonal
- It adds new parameters
  - P = seasonal order of AR
  - D = seasonal order of I
  - Q = seasonal order of MA
  - m = length of the seasonal cycle
- Useful for complex data with both trends and seasonality



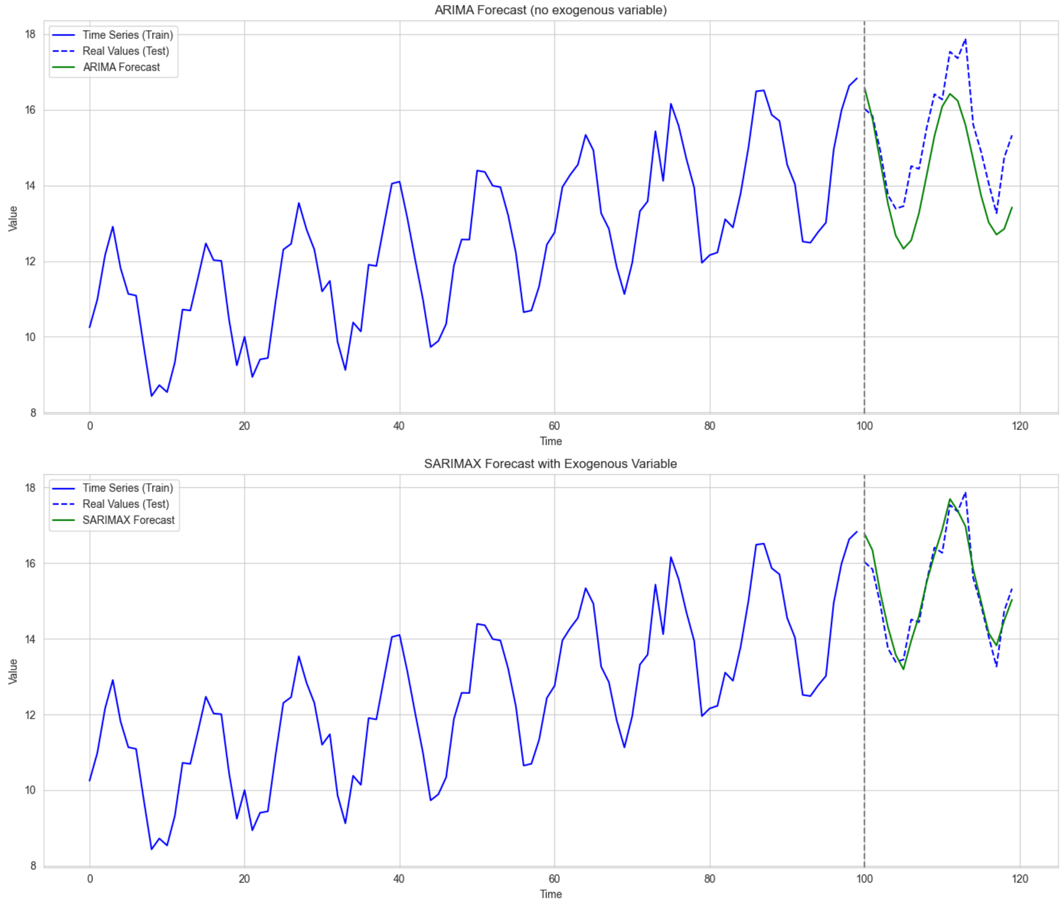
# AutoARIMA

- Automated approach to select the best ARIMA or SARIMA model parameters
- Searches over combinations of p, d, q, and seasonal P, D, Q, m to find the optimal model
- Uses statistical criteria like AIC (Akaike Information Criterion) or BIC (Bayesian Information Criterion) to evaluate models
- Saves time and reduces guesswork in model selection
- Ideal for beginners and practical when working with many datasets



# ARIMAX / SARIMAX Models

- Extensions of ARIMA and SARIMA that include exogenous variables
- Help improve forecasting by incorporating additional information beyond the time series itself
- Useful when other factors influence the target series, like promotions, weather, or economic indicators
- Parameters and seasonal components work like ARIMA/SARIMA but with extra inputs
- Flexible models for richer, more accurate forecasts



# What we‘ve learnt

- AR, MA, and ARMA models form the foundation for autoregressive modeling
- ARIMA adds differencing to handle trends and make data stationary
- SARIMA extends ARIMA to capture seasonal patterns
- AutoARIMA automates model selection, saving time and effort
- ARIMAX and SARIMAX incorporate external variables to improve forecasts
- Choosing the right model depends on your data’s characteristics and available information