Time Series Analysis and Forecasting

Chapter 0: Fundamentals

Daniel Traian PELE

Academia de Studii Economice din București
IDA Institute Digital Assets
Blockchain Research Center
AI4EFin Artificial Intelligence for Energy Finance
Academia Română, Institutul de Prognoză Economică
MSCA Digital Finance

Learning Objectives

By the end of this chapter, you will be able to:

1. **Define** time series and distinguish them from cross-sectional and panel data
2. **Decompose** time series into trend-cycle, seasonality, and residual components
3. **Apply** exponential smoothing methods (SES, Holt, Holt-Winters, ETS)
4. **Evaluate** forecasts using MAE, RMSE, MAPE, sMAPE
5. **Implement** train/validation/test splitting and cross-validation
6. **Model** seasonality using dummy variables or Fourier terms
7. **Remove** trend and seasonality through appropriate methods
8. **Distinguish** between deterministic and stochastic trends

Data Sources and Software Tools

Data Sources

- **Yahoo Finance**
  - Stock prices, cryptocurrencies, exchange rates
- **FRED** (Federal Reserve)
  - GDP, unemployment, interest rates
- **Eurostat / INS / BNR**
  - European and Romanian economic data
- **Classic datasets**
  - AirPassengers, Sunspots, CO_2

Python

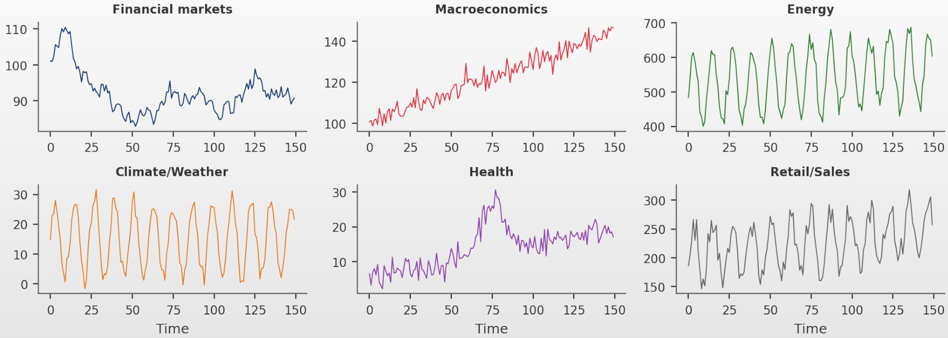
- `yfinance` — Yahoo Finance data
- `pandas_datareader` — FRED, Eurostat
- `statsmodels` — statistical models
- `pandas` — data manipulation
- `matplotlib` — visualization

R

- `quantmod` — financial data
- `tseries` — time series tests
- `forecast` — forecasting models
- `fredr` — FRED API access

Chapter Outline

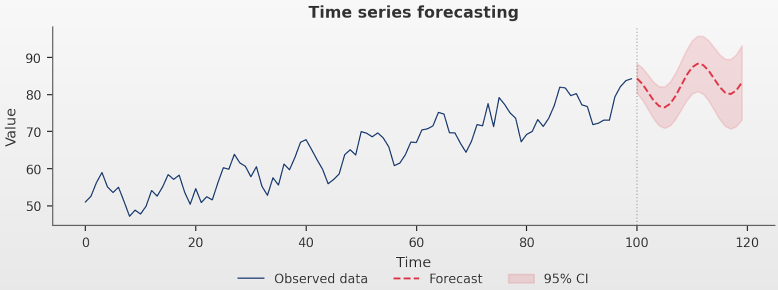
Time Series Are Everywhere



TSA_ch0_real_data

- **Finance**: Stock prices, exchange rates, volumes
- **Economics**: GDP, unemployment, inflation rates
- **Business**: Sales, website traffic, demand
- **Science**: Temperature, pollution, vital signs

Why Do We Study Time Series?

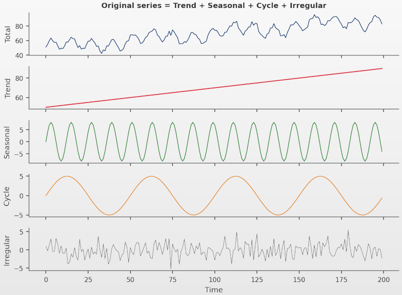


TSA_ch0_real_data

Main objective: forecasting

- We use historical patterns to predict future values $\succ$ essential for business planning, risk management, and policy decisions

Understanding Time Series Structure



TSA ch0 real data

Decomposition

- Any time series can be decomposed into: **trend-cycle** + **seasonality** + **noise**

Definition of a Time Series

Definition 1 (Time Series)

- **Time series**: a sequence of observations $\{X_t\}$ indexed by time: $\{X_t : t \in \mathcal{T}\}$ where $\mathcal{T}$ is a set of indices representing time points

Key Characteristics

- **Ordered**: natural temporal order
- **Dependent**: consecutive observations are correlated
- **Discrete/Continuous**: $t = 1, 2, 3, \ldots$

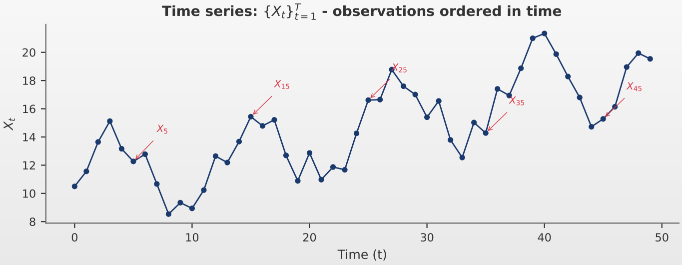
Notation

- X_t: observation at time t
- $\{X_t\}_{t=1}^{T}$: series with T observations

Time Series: Conceptual Illustration

Fundamental Elements

- **Formal notation**
  - X_t = value at time t
  - $t \in \{1, 2, \ldots, T\}$
- **Autocorrelation**
  - $\rho_k = \text{Corr}(X_t, X_{t-k})$
  - Measures temporal dependence

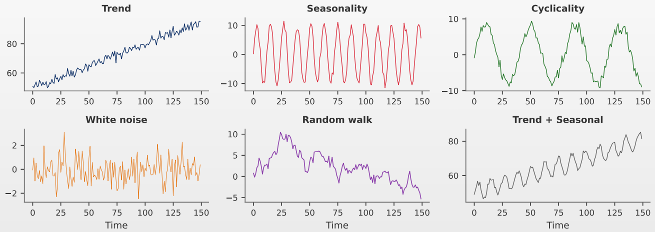


TSA_ch0_definition

Common Patterns in Time Series

Types of Patterns

- **Trend**
  - Long-term increase or decrease
- **Seasonal**
  - Regular periodic patterns
- **Cyclical**
  - Medium-term fluctuations (2–10 years)
- **Random**
  - Unpredictable fluctuations

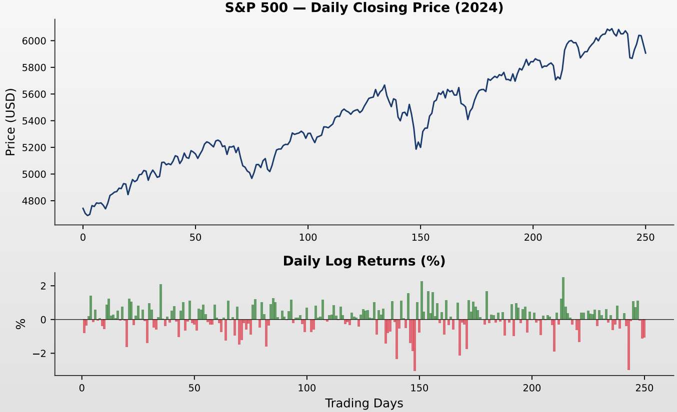


TSA_ch0_definition

Practical Example: Real Financial Data

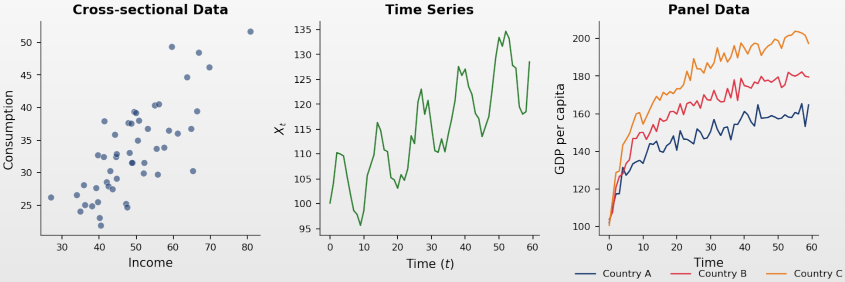
S&P 500 (2024)

- **Daily frequency**
  - $\approx$ 252 trading days/year
- **Observed characteristics**
  - Upward trend
  - Volatility clustering
  - Persistence (momentum)



TSA_ch0_definition

Data Types: Comparison



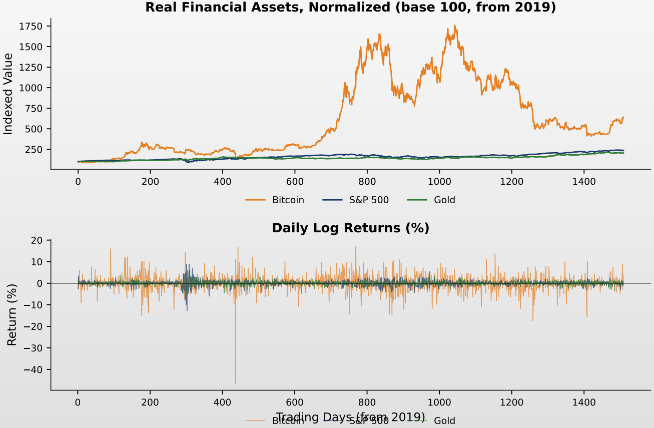
Data Type	Units (N)	Time (T)	Example
Cross-sectional	Many	1	Survey of 1000 households
Time series	1	Many	Daily S&P 500 prices
Panel	Many	Many	GDP of 50 countries, 20 years

TSA_ch0_definition

Examples of Time Series Data

Real financial data

- **Source**: Yahoo Finance (2019–2025)
  - Normalized to base 100
- **Bitcoin**: most volatile
- **Gold**: most stable



TSA_ch0_real_data

Why Do We Decompose a Time Series?

Objectives

- Understanding underlying patterns
- Removing seasonality for modeling
- Identifying trend direction
- Isolating irregular fluctuations
- Improving forecast accuracy

Components

- T_t: Trend-Cycle
  - Long-term movement
- S_t: Seasonal
  - Regular periodic pattern
- ε_t: Residual
  - Random noise

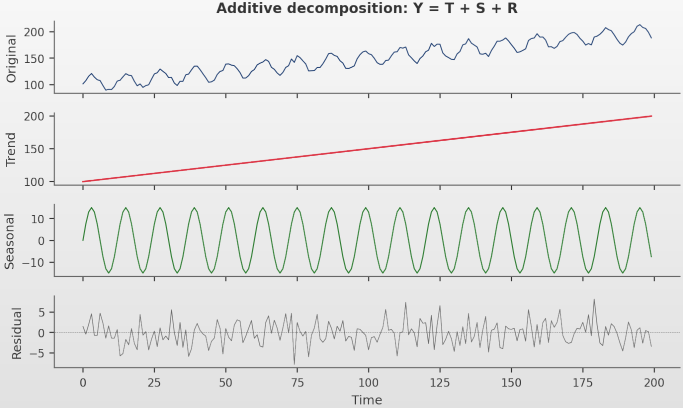
Classical decomposition models

- **Additive**: $X_t = T_t + S_t + \varepsilon_t$
  - Constant seasonal amplitude
- **Multiplicative**: $X_t = T_t \times S_t \times \varepsilon_t$
  - Seasonal amplitude grows with the level

Time Series Decomposition: Visual Example

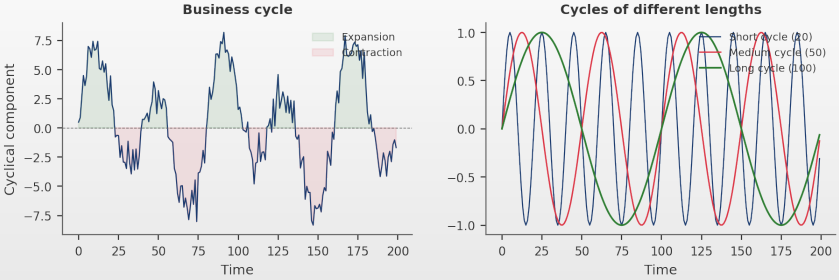
Components Explained

- **Original**
  - Observed series
- **Trend-Cycle**
  - Long-term movement
- **Seasonal**
  - Periodic pattern
- **Residual**
  - Random noise



TSA_ch0_decomposition

The Cyclical Component



TSA_ch0_decomposition

Characteristics

- **Duration**: medium-term fluctuations (2–10 years)
- **Aperiodic**: no fixed period (vs seasonality)
- **Origin**: reflects business cycles

In Practice

- **Combination**: cycle combined with trend
- **Difficulty**: hard to identify in short series
- **Solution**: usually absorbed into trend-cycle

The Additive Decomposition Model

Model

- **Equation**: $X_t = T_t + S_t + \varepsilon_t$
  - Components are added together to form the observed series

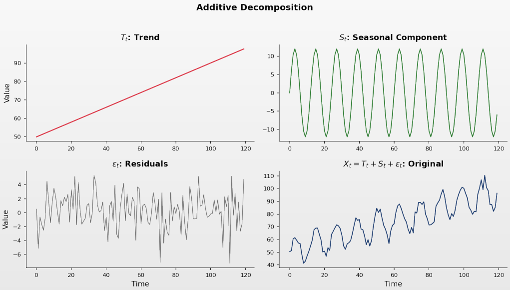
When to Use

- **Constant seasonal fluctuations**
  - Amplitude does not depend on the level
- **Stable series variance**
  - Measures dispersion around the mean
  - Estimator: $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$

Properties

- **Error**: $\mathbb{E}[\varepsilon_t] = 0$ (zero mean)
- **Seasonal**: $\sum_{j=1}^{s} S_j = 0$ (seasonal sum is zero)
- **Units**: S_t are the same as X_t

Additive Decomposition: Visualization



TSA_ch0_decomposition

Interpretation

- **Decomposition**: Original = Trend + Seasonal + Residual
- **Property**: constant seasonal amplitude, does not depend on the level

The Multiplicative Decomposition Model

Model

- **Equation**: $X_t = T_t \times S_t \times \varepsilon_t \succ$ components are multiplied

When to Use

- **Growing fluctuations**: seasonality increases with the level
- **Heteroscedasticity**: variance increases over time
- **Examples**: economic/financial data

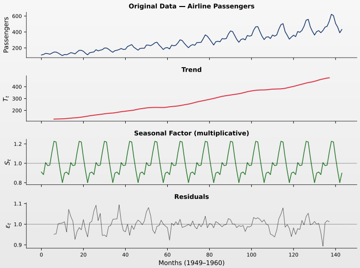
Properties

- **Error**: $\mathbb{E}[\varepsilon_t] = 1$ (centered at 1)
- **Seasonal**: $\frac{1}{s}\sum_{j=1}^{s} S_j = 1$ (mean is 1)
- **Units**: S_t is a dimensionless ratio

Tip

- **Log transformation**: multiplicative $\succ$ additive: $\log X_t = \log T_t + \log S_t + \log \varepsilon_t$

Multiplicative Decomposition: Real Data



Example

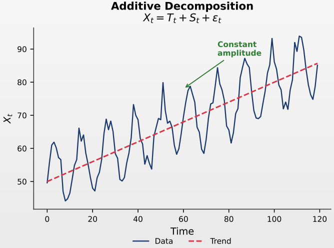
- Box-Jenkins data: monthly passengers (1949–1960). Seasonal amplitude increases with the level

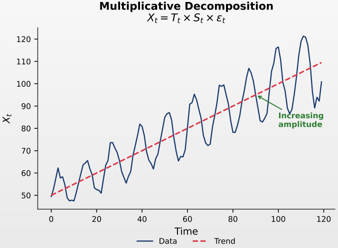
TSA_ch0_decomposition

Additive vs Multiplicative: Comparison

Key Difference

- **Multiplicative**
  - Seasonal component is a *ratio*
  - Centered at value 1
- **Additive**
  - Seasonal component in *absolute units*
  - Centered at value 0





TSA_ch0_decomposition

Trend Estimation: Moving Average

Definition 2 (Centered Moving Average)

- **Centered moving average** of order $2q+1$:

$$\hat{T}_t = \frac{1}{2q+1} \sum_{j=-q}^{q} X_{t+j} \tag{1}$$

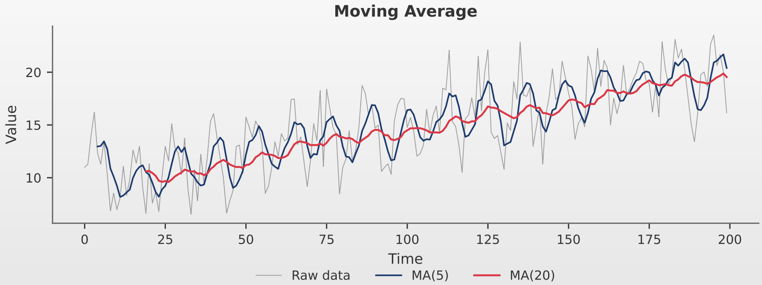
For Seasonal Data

- **Odd period** s
  - Use simple average
- **Even period** s
  - $2 \times s$ MA with half weights

Properties

- **Smoothing**: removes seasonal & random components
- **Large window** $\succ$ smoother estimate
- **Disadvantage**: data loss at endpoints

Centered Moving Average: Visual Illustration



Interpretation

- **Smoothing**: removes short-term fluctuations
- **Result**: reveals the underlying trend

TSA_ch0_definition

Classical Decomposition Algorithm

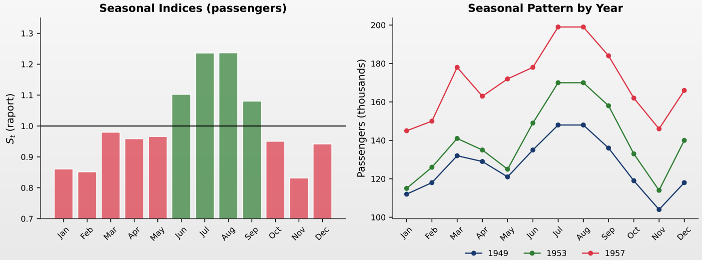
Steps for Multiplicative Decomposition

- **Step 1** $\succ$ **Estimate Trend**: $\hat{T}_t = MA_s(X_t)$
  - Centered moving average of order equal to the seasonal period
- **Step 2** $\succ$ **Detrend**: $D_t = X_t / \hat{T}_t$
- **Step 3** $\succ$ **Estimate Seasonal**: $\hat{S}_j = \text{mean}(D_t \text{ for season } j)$
- **Step 4** $\succ$ **Normalize**: scale so that $\frac{1}{s}\sum_{j=1}^{s} \hat{S}_j = 1$
- **Step 5** $\succ$ **Compute Residuals**: $\hat{\varepsilon}_t = X_t / (\hat{T}_t \times \hat{S}_t)$

Note

- **For additive decomposition**: operations change
  - Division $\succ$ subtraction
  - Multiplication $\succ$ addition

Seasonal Indices: Interpretation



Interpretation

- $S_t > 1$: above-average activity; $S_t < 1$: below average. Travel peak in July–August

TSA_ch0_seasonal

STL Decomposition: A Modern Approach

Definition 3 (STL - Seasonal-Trend Decomposition using LOESS)

- **STL**: uses locally weighted regression (LOESS): $X_t = T_t + S_t + R_t$

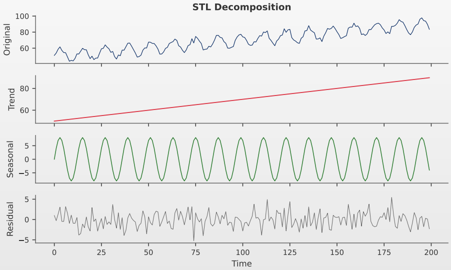
Advantages

- **Flexibility**: any seasonal period
- **Variability**: seasonality can evolve over time
- **Robustness**: resistant to outliers
- **Smoothing**: smooth trend estimates

Key Parameters

- `period`: seasonal period
  - E.g.: 12 for monthly data, 4 for quarterly
- `seasonal`: smoothing window
- `robust`: reduced weight for outliers

STL Decomposition: Visual Illustration



Key Idea

- STL (Seasonal-Trend-Loess): separates trend + seasonal + remainder using LOESS regression

TSA_ch0_decomposition

Exponential Smoothing: Overview

Definition

- **Exponential smoothing**: weighted averages of past observations
  - Weights decrease exponentially over time
  - Recent observations receive higher weights

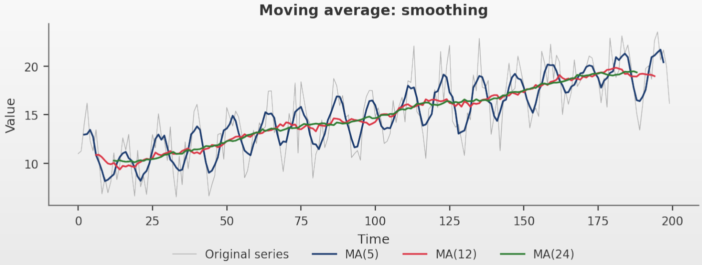
Why Exponential Smoothing?

- **Simple**: easy to implement and understand
  - A single smoothing parameter
- **Adaptive**: higher weights for recent data
- **Versatile**: handles trend and seasonality

Three Main Methods

- **SES** (Simple Exponential Smoothing): level only
  - The simplest exponential method
- **Holt**: level + trend
  - Captures the direction of evolution
- **Holt-Winters**: + seasonality
  - Complete model with all components

Moving Average Smoothing



Window Size Trade-off

- **Small window**: reactive but noisy
  - Captures rapid changes, but amplifies noise
- **Large window**: smooth but lagging
  - Removes noise, but reacts slowly

TSA_ch0_definition

Simple Exponential Smoothing (SES)

Model

- **Equation**: $\hat{X}_{t+1|t} = \alpha X_t + (1-\alpha)\hat{X}_{t|t-1}$
  - $\alpha \in (0,1)$ is the smoothing parameter

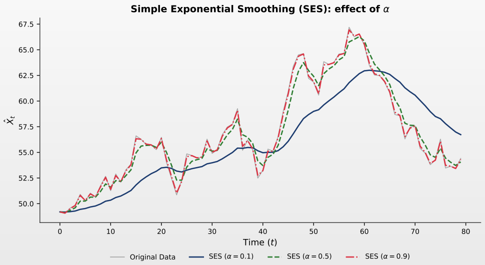
How It Works

- **Principle**: weights decrease exponentially
- **Large** α
  - Forecast reactive to changes
- **Small** α
  - Smoother, more stable forecast

Level Form

- **Equation**: $\ell_t = \alpha X_t + (1-\alpha)\ell_{t-1}$
  - ℓ_t = estimated level at time t
  - Forecast: $\hat{X}_{t+h|t} = \ell_t$ (constant)

Simple Exponential Smoothing: Effect of α



Trade-off

- **Small** $\alpha \succ$ smooth forecasts
  - More weight on distant history
- **Large** $\alpha \succ$ tracks the data
  - Fast reaction to recent changes

TSA_ch0_smoothing

SES: Step-by-Step Numerical Example

Data: Monthly Sales (thousands EUR)

- **Data**: $X_1 = 100,\ X_2 = 110,\ X_3 = 105,\ X_4 = 115,\ X_5 = 120 \quad (\alpha = 0.3,\ \hat{X}_{1|0} = 100)$

Iterative computation: $\hat{X}_{t+1|t} = \alpha X_t + (1-\alpha)\hat{X}_{t|t-1}$

t	X_t	$\hat{X}_{t\|t-1}$	e_t	Computation $\hat{X}_{t+1\|t}$
1	100	100.00	0.00	$0.3 \times 100 + 0.7 \times 100 = 100.00$
2	110	100.00	10.00	$0.3 \times 110 + 0.7 \times 100 = 103.00$
3	105	103.00	2.00	$0.3 \times 105 + 0.7 \times 103 = 103.60$
4	115	103.60	11.40	$0.3 \times 115 + 0.7 \times 103.6 = 107.02$
5	120	107.02	12.98	$0.3 \times 120 + 0.7 \times 107.02 = 110.91$

Forecast and Evaluation

$\hat{X}_{6|5} = 110.91$ MAE $= 7.28$ RMSE $= 8.97$

Holt's Linear Trend Method

Equations

- **Level**: $\ell_t = \alpha X_t + (1-\alpha)(\ell_{t-1} + b_{t-1})$
- **Trend**: $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1}$
- **Forecast**: $\hat{X}_{t+h|t} = \ell_t + h \cdot b_t$
  - Extrapolates the linear trend over h steps

Parameters

- α: level smoothing
  - Controls reactivity to level changes
- β^*: trend smoothing
  - Controls reactivity to slope changes

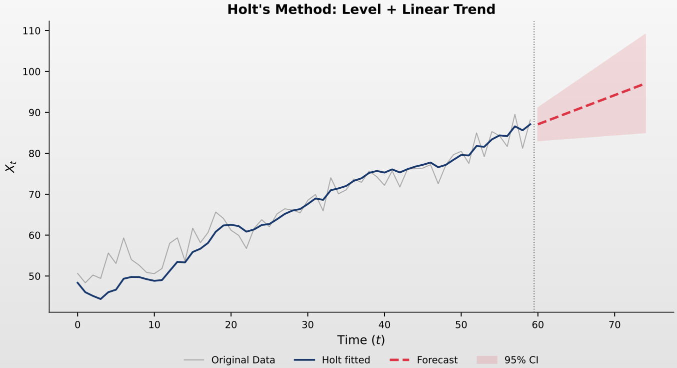
Components

- ℓ_t: estimated level
  - Local mean of the series
- b_t: estimated trend (slope)
  - Rate of increase/decrease

Holt's Method: Visualization

Interpretation

- **Holt's method**: captures level and trend
  - Projects them into the forecast horizon
- α: controls level changes
- β^*: controls trend changes



TSA_ch0_smoothing

Holt-Winters Seasonal Method

Equations (Additive Seasonality)

- **Level**: $\ell_t = \alpha(X_t - S_{t-s}) + (1-\alpha)(\ell_{t-1} + b_{t-1})$
- **Trend**: $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1}$
- **Seasonal**: $S_t = \gamma(X_t - \ell_t) + (1-\gamma)S_{t-s}$
- **Forecast**: $\hat{X}_{t+h|t} = \ell_t + h \cdot b_t + S_{t+h-s(k+1)}$
  - Where $k = \lfloor (h-1)/s \rfloor$

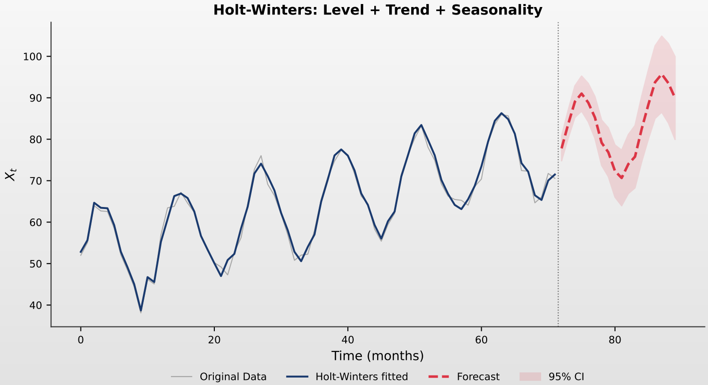
Parameters

- α — level
- β^* — trend
- γ — seasonal
- s — seasonal period
  - All in $(0,1)$; estimated by minimizing error

Holt-Winters: Capturing Seasonality

Key Feature

- **Complete decomposition**
  - Separates level, trend, and seasonal
- **Seasonal forecasts**
  - Includes both trend and periodic pattern



TSA_ch0_smoothing

The ETS Framework: Error-Trend-Seasonality

Definition 4 (ETS Models)

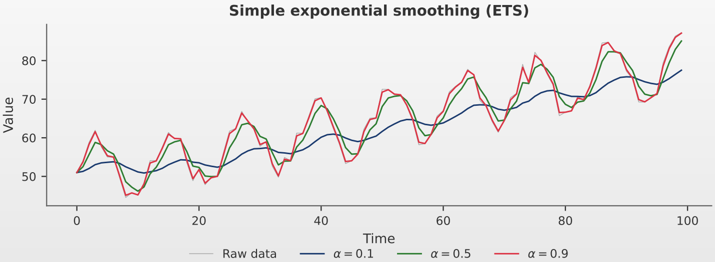
- **ETS framework**: generalizes exponential smoothing: ETS(E, T, S)

Component	**N**	**A**	**M**
Error (E)	–	Additive	Multiplicative
Trend (T)	None	Additive	Multiplicative
Seasonal (S)	None	Additive	Multiplicative

Examples

- **ETS(A,N,N)**: Simple Exponential Smoothing $\succ$ level only, no trend or seasonality
- **ETS(A,A,N)**: Holt's Linear Method $\succ$ level + additive trend
- **ETS(A,A,A)**: Additive Holt-Winters $\succ$ level + trend + additive seasonality

ETS: Exponential Smoothing Illustration



Interpretation

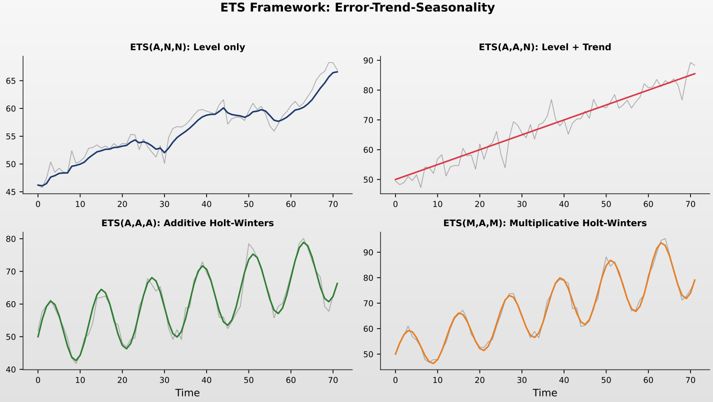
- Exponentially weighted observations: weights decrease with age; recent observations have greater importance

TSA_ch0_smoothing

ETS Model Selection

Automatic Selection

- **Information criteria**
  - AIC (Akaike Information Criterion)
  - BIC (Bayesian Information Criterion)
- **Optimal selection**
  - Balance between fit and complexity



TSA_ch0_smoothing

Damped Trend Methods

Damping Parameter

- **Parameter**: $\phi \in (0, 1)$
  - Prevents over-projection of the trend
  - Trend converges to a constant

Equations

- **Level**: $\ell_t = \alpha X_t + (1-\alpha)(\ell_{t-1} + \phi b_{t-1})$
- **Trend**: $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)\phi b_{t-1}$
- **Forecast**: $\hat{X}_{t+h|t} = \ell_t + \phi \frac{1-\phi^h}{1-\phi} b_t$

Key Idea

- **Asymptotic**: as $h \to \infty$, forecast $\to$ constant
  - Prevents unrealistic long-term extrapolation
- **Advantage**: often better for long horizons

Forecast Accuracy Metrics

Forecast Error

- **Definition**: $e_t = X_t - \hat{X}_t$ (actual minus predicted)
  - Positive $\Rightarrow$ underestimates; Negative $\Rightarrow$ overestimates

Scale-dependent

- **MAE**: $\frac{1}{n}\sum |e_t|$
- **MSE**: $\frac{1}{n}\sum e_t^2$
- **RMSE**: $\sqrt{\text{MSE}}$

Scale-independent

- **MAPE**: $\frac{100}{n}\sum \left|\frac{e_t}{X_t}\right|$
- **sMAPE**: $\frac{100}{n}\sum \frac{|e_t|}{(|X_t|+|\hat{X}_t|)/2}$

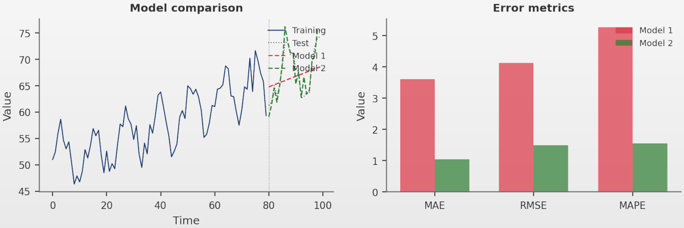
What to Use?

- **Same series**: RMSE, MAE $\succ$ compare models on the same data
- **Across different series**: MAPE, sMAPE $\succ$ percentage metrics, scale-independent

Forecast Evaluation: Visual Example

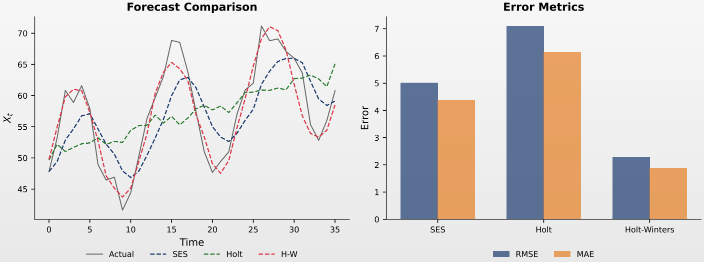
Observations

- **Top**: actual vs. forecast
  - Visual assessment of forecast quality
- **Bottom**: residuals
  - Zero mean
  - Constant variance
  - No pattern



TSA_ch0_forecast_eval

Comparing Forecast Methods



Interpretation

- **Left**: SES, Holt, Holt-Winters forecasts. **Right**: Error metrics. Visual and quantitative comparison

TSA_ch0_forecast_eval

Residual Diagnostics

Residual Properties

- **Zero mean**: $\mathbb{E}[e_t] = 0$
  - Forecast has no systematic bias
- **Uncorrelated**: $\text{Cov}(e_t, e_{t-k}) = 0$
  - No unexploited information remains
- **Constant variance**: $\text{Var}(e_t) = \sigma^2$
- **Normally distributed**: for confidence intervals

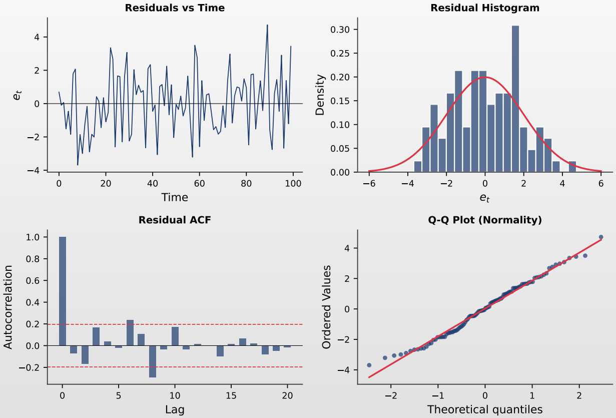
Diagnostic Tests

- **Ljung-Box test** (autocorrelation):
  - $Q = T(T+2)\sum_{k=1}^{h} \frac{\hat{\rho}_k^2}{T-k} \sim \chi_h^2$
- **Jarque-Bera test** (normality):
  - $JB = \frac{T}{6}\left(S^2 + \frac{(K-3)^2}{4}\right) \sim \chi_2^2$
  - S = skewness, K = kurtosis

Residual Diagnostics: Visualization

What to Check

- **Time plot**: no systematic patterns
- **Histogram**: normality check
- **ACF**: no significant autocorrelation
- **Q-Q plot**: normality confirmation



TSA_ch0_forecast_eval

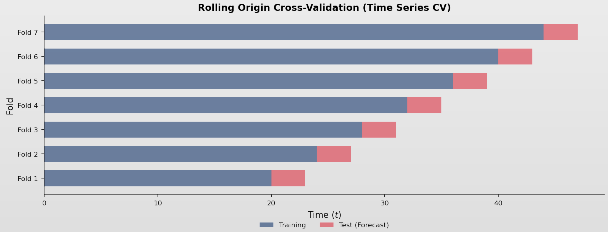
Cross-Validation for Time Series

Why Not Standard CV?

- **Temporal dependence**: observations are correlated
- **Order matters**: chronology must be respected
- **Standard k-fold** $\succ$ data leakage

CV with Rolling Origin

- **Step 1**: train on $\{X_1, \ldots, X_t\}$
- **Step 2**: forecast $\hat{X}_{t+h}$
- **Step 3**: increment t, repeat



TSA_ch0_forecast_eval

Train / Validation / Test Split

Training Set

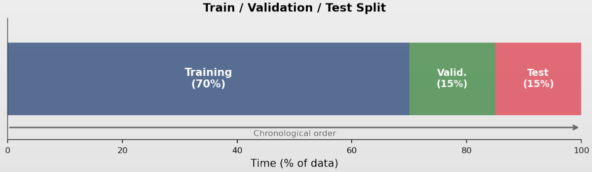
- Fitting model parameters
- Largest portion (60–80%)
- Used for estimation

Validation Set

- Hyperparameter tuning
- Comparing models
- Selecting the best approach

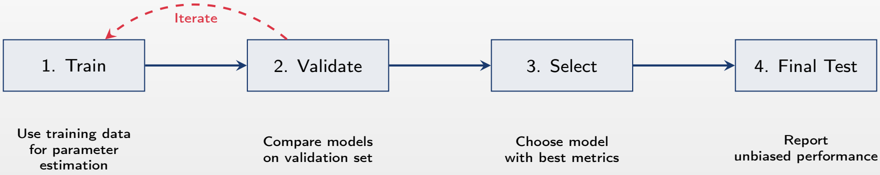
Test Set

- Final evaluation only
- Never used for tuning
- Unbiased performance



TSA_ch0_forecast_eval

Model Development Workflow



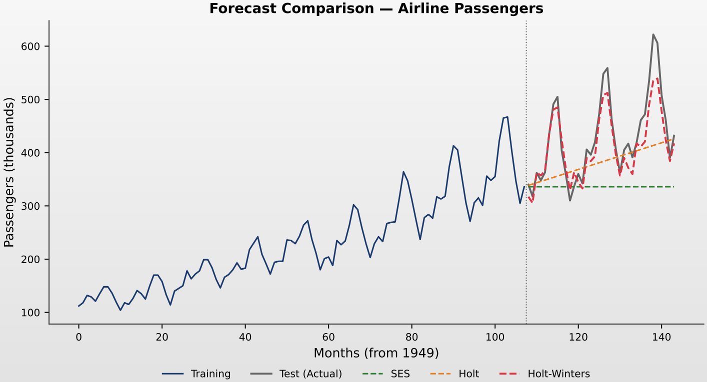
Critical Rule

- **Never use the test set for selection!**
  - Use it only for final evaluation
- **Avoid data leakage**
  - Overly optimistic performance estimates

Real Data: Comparing Forecasts

Interpretation

- **Data**: airline passengers
- **Best**: multiplicative Holt-Winters
  - Ideal for data with growing seasonality

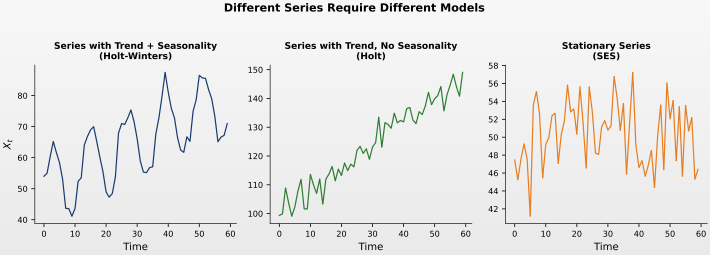


TSA_ch0_forecast_eval

Forecast Performance Across Different Datasets

Interpretation

- **Different series**
  - Require different models
- **Seasonal data**
  - Prefer seasonal methods
- **No universal model**
  - Test multiple approaches



TSA_ch0_forecast_eval

Modeling Seasonality: Two Approaches

1. Dummy Variables

- **Model**: $X_t = \mu + \sum_{j=1}^{s-1} \gamma_j D_{jt} + \varepsilon_t$
- $D_{jt} = 1$ if t in season j
- $s - 1$ parameters
- Any seasonal pattern

2. Fourier Terms

- **Model**:
  $X_t = \mu + \sum_{k=1}^{K} \left[\alpha_k \sin\left(\frac{2\pi kt}{s}\right) + \beta_k \cos\left(\frac{2\pi kt}{s}\right)\right]$
- Sinusoidal functions
- $2K$ parameters
- Smooth patterns

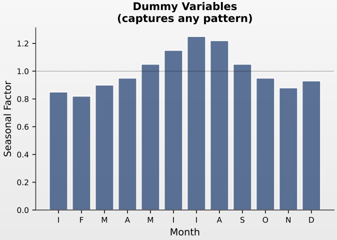
Trade-off

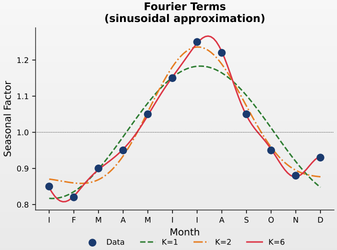
- **Dummy variables**
  - Any seasonal pattern, but more parameters
- **Fourier terms**
  - Smooth patterns, fewer parameters

Dummy Variables vs Fourier Terms

Comparison

- **Dummy variables**
  - Capture any shape
  - Require $s-1$ parameters
- **Fourier terms**
  - Only $2K$ parameters
  - Smooth, sinusoidal patterns





TSA_ch0_seasonal

Choosing Between Dummy and Fourier

Criterion	Dummy	Fourier
Parameters (monthly)	11	$2K$ (often 4–6)
Seasonal pattern	Any shape	Smooth/sinusoidal
Interpretation	Direct (monthly effects)	Frequency components
High-frequency seasons	Many parameters	Efficient
Multiple seasonality	Complex	Easy (add terms)

Recommendations

- **Use Dummy**
  - Irregular patterns, interpretable coefficients
- **Use Fourier**
  - Smooth patterns, high-frequency seasonality
  - Used in TBATS and Prophet

Why Do We Remove Trend and Seasonality?

Reasons for Detrending

- Stationarity requirement
- Focus on fluctuations
- Avoiding spurious regression
- Enabling valid inference

Reasons for Deseasonalizing

- Revealing the underlying trend
- Cross-season comparisons
- Simplifying modeling
- Focus on the irregular component

Important

- **We model the transformed series**
  - With trend and seasonality removed
- **We reverse the transformation**
  - Bring the forecast back to the original scale

Detrending Methods

Six Common Detrending Approaches

- **Differencing**: $\Delta X_t = X_t - X_{t-1}$
  - Most commonly used, removes stochastic trend
- **Linear regression**: $\hat{T}_t = \hat{\beta}_0 + \hat{\beta}_1 t$
- **Polynomial**: higher-order polynomial
- **HP filter**: balance between fit and smoothness
- **Moving average**: $\hat{T}_t = MA_q(X_t)$
- **LOESS**: local polynomial regression

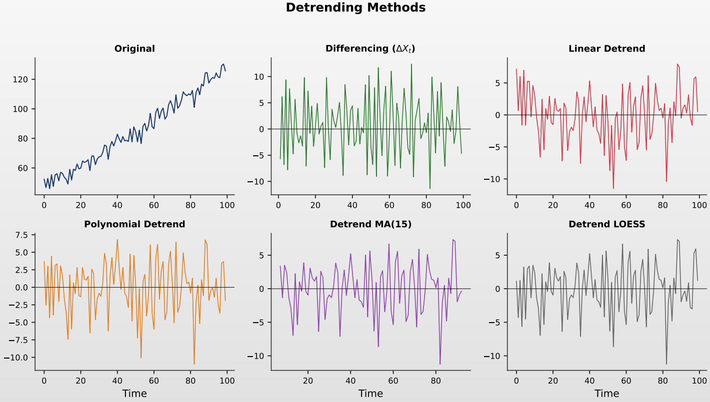
The Choice Depends on

- **Nature of the trend**
  - Deterministic vs stochastic
- **Purpose of the analysis**
  - Forecasting vs descriptive analysis

Detrending Methods: Comparison

Key Idea

- **Different methods**
  - Produce different residuals
- **Choose by trend type**
  - Consider the analysis objectives

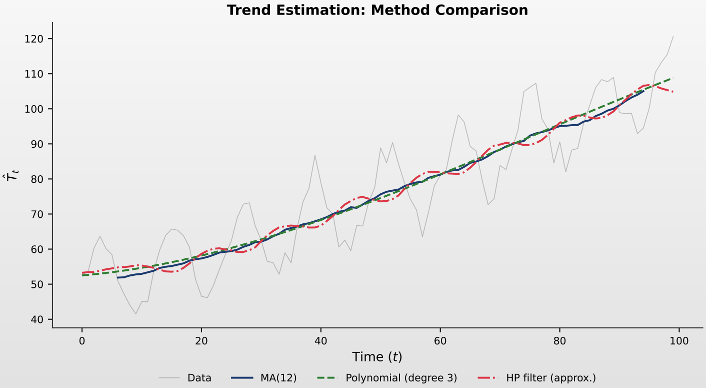


TSA_ch0_detrending

Trend Estimation: Multiple Approaches

Method Comparison

- **Moving average**
  - Simple but with lag
- **Polynomial regression**
  - Flexible, parametric
- **HP filter**
  - Macroeconomic standard



TSA_ch0_detrending

The Hodrick-Prescott (HP) Filter

Definition 5 (HP Filter)

- **HP filter**: decomposes X_t into trend τ_t and cycle c_t: $X_t = \tau_t + c_t$

$$\min_{\{\tau_t\}} \left\{ \sum_{t=1}^{T} (X_t - \tau_t)^2 + \lambda \sum_{t=2}^{T-1} [(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1})]^2 \right\}$$

Interpretation

- **First term**
  - Goodness of fit
- **Second term**
  - Smoothness penalty
- λ
  - Controls the balance between fidelity and smoothness

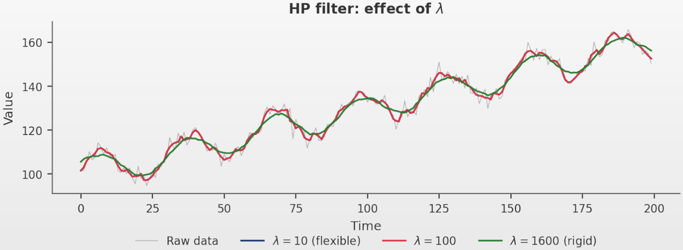
Standard λ Values (Ravn-Uhlig)

- **Annual**
  - $\lambda = 6.25$
- **Quarterly**
  - $\lambda = 1600$ (macroeconomic standard)
- **Monthly**
  - $\lambda = 129600$

HP Filter: Effect of λ

Trade-off

- **Small** λ: flexible trend
  - Follows the data closely
- **Large** λ: smooth trend
  - Approaches a linear trend

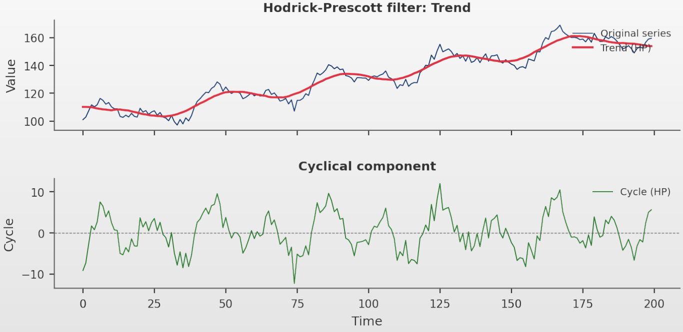


TSA_ch0_detrending

HP Filter: Business Cycle Extraction

Application

- **Macroeconomics**
  - Business cycle extraction
- **Common series**
  - GDP, unemployment, inflation



TSA_ch0_detrending

HP Filter: Limitations

Known Issues

- **Endpoint instability**
  - Trend estimates unreliable at the beginning and end
- **Spurious cycles**
  - Can create artificial dynamics
- **Choice of λ**
  - Results sensitive to the parameter

Alternatives

- **Band-pass filters**: Baxter-King, Christiano-Fitzgerald
  - Isolate specific frequencies
- **Hamilton filter**: regression-based
- **Unobserved components**: state-space models

Hamilton's Critique (2018)

- "Why You Should Never Use the Hodrick-Prescott Filter"
  - Suggests using regression on lagged values

Seasonal Adjustment Methods

Four Approaches for Seasonal Adjustment

- **Seasonal differencing**: $\Delta_s X_t = X_t - X_{t-s}$
  - Removes periodic pattern, simple to apply
- **Division** (multiplicative): $X_t^{adj} = X_t / \hat{S}_t$
- **Subtraction** (additive): $X_t^{adj} = X_t - \hat{S}_t$
- **X-13ARIMA-SEATS**: official US Census Bureau standard
  - Sophisticated method, used by statistical institutes

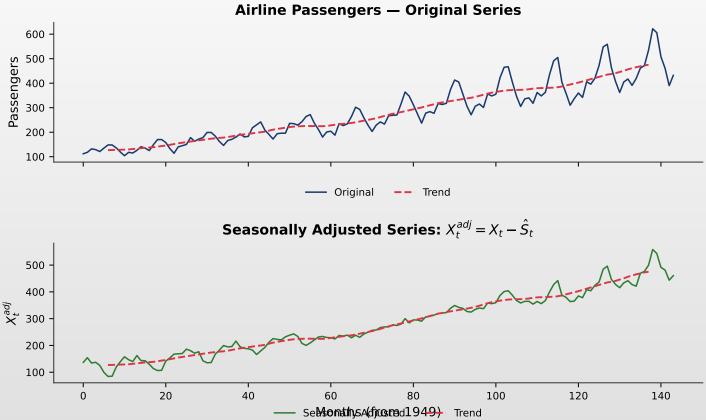
Seasonal Period s

- Monthly: $s = 12$ | Quarterly: $s = 4$

Seasonal Adjustment: Visualization

Result

- **Seasonally adjusted series**
  - Reveals the underlying trend
  - Removes periodic fluctuations



TSA_ch0_seasonal

Deterministic vs Stochastic Trend

Deterministic Trend

- **Model**: $X_t = \beta_0 + \beta_1 t + \varepsilon_t$
- **Characteristics**:
  - Trend is a function of time
  - ε_t is stationary
- **Method**: detrend by regression

Stochastic Trend

- **Model**: $X_t = X_{t-1} + \varepsilon_t$
- **Characteristics**:
  - Random walk component
  - ΔX_t is stationary
- **Method**: detrend by differencing

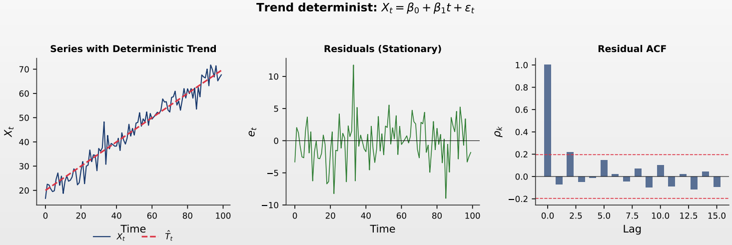
Wrong Method = Problems

- **Differencing a deterministic trend** $\succ$ over-differencing
  - Introduces artificial dependence in the series
- **Regression on a stochastic trend** $\succ$ spurious regression
  - Invalid statistical results

Example: Deterministic Trend

Key

- **Method**: regression
- **Result**: stationary residuals, ACF decays rapidly

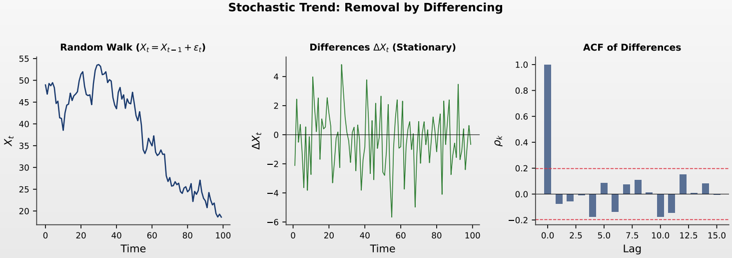


TSA_ch0_detrending

Example: Stochastic Trend (Random Walk)

Key

- **Method**: differencing
- **Result**: differences are stationary (white noise)

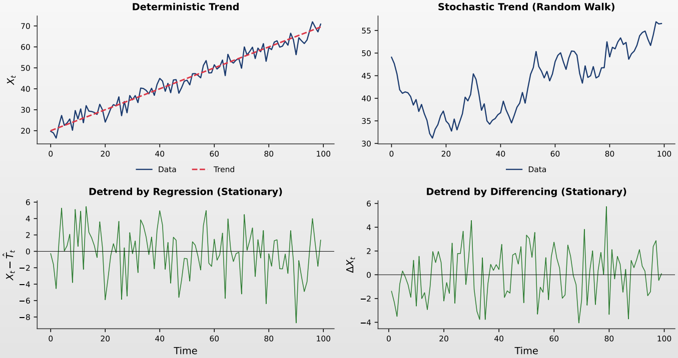


TSA_ch0_detrending

Side-by-Side Comparison

Remember

- **Deterministic trend**: use regression
  - Trend is a predictable function of time
- **Stochastic trend**: use differencing
  - Trend contains a random component



TSA_ch0_detrending

Experiment: ChatGPT vs Fundamentals

Prompt → Response

You: "I have monthly airline passenger data from 1949 to 1960. Analyze it and make a forecast."

ChatGPT: `seasonal_decompose(data, model='additive')`

"Trend extracted. Seasonal pattern identified. MAPE = 4.2%. *The model fits well.*"

Three errors a trained analyst catches immediately:

1. **Wrong decomposition type**: seasonal amplitude $\times 3$ from 1949 to 1960
   $\mathrm{Var}(S_t) \neq$ const $\succ$ additive violated $\succ$ use **multiplicative** or $\ln X_t$ first
2. **In-sample metric is meaningless**: MAPE $= 4.2\%$ is computed on **training data**
   Rolling-origin CV reveals true MAPE $= 8.7\%$ — the model is **twice as bad** as reported
3. **Residuals not checked**: ACF of residuals $\neq 0$ at lags 1–3
   Unexplained systematic pattern remains $\succ$ the decomposition is **misspecified**

Discussion: The code runs without errors. The output looks professional. *How do you know it's wrong?*

Summary

What We Learned in This Chapter

- Time Series Definition and Characteristics
  - Sequence of temporally ordered observations with dependence
- Decomposition (Additive vs Multiplicative)
  - Components: Trend-Cycle + Seasonal + Residual
- Exponential Smoothing Methods
  - SES (level), Holt (+ trend), Holt-Winters (+ seasonality), ETS
- Forecast Evaluation and Validation
  - Metrics: MAE, RMSE, MAPE; Cross-Validation with rolling origin

Key Idea

- **Understand Before Modeling**:
  - Visualize and decompose the data first
  - Choose additive vs multiplicative based on variance behavior

Quick Quiz

Test Your Knowledge

1. What is the difference between additive and multiplicative decomposition?
2. When should you use Holt-Winters instead of SES?
3. Why can't we use standard k-fold CV for time series?
4. What does $\alpha = 0.9$ mean in exponential smoothing?
5. How do you distinguish between a deterministic and stochastic trend?

Quiz Answers

Answers

1. **Additive vs multiplicative**: additive when seasonal amplitude is constant; multiplicative when it grows with the level
2. **Holt-Winters**: when data have trend AND seasonality; SES handles only the level
3. **CV**: standard k-fold ignores temporal order $\succ$ data leakage
4. $\alpha = 0.9$: high weight on recent observations, reacts quickly but more volatile
5. **Trend**: deterministic $\succ$ function of time (regression); stochastic $\succ$ random walk (differencing)

What's Next?

Chapter 1: Stochastic Processes and Stationarity

- **Stochastic Processes**: mathematical foundation, random variables indexed by time
- **Stationarity**: strict (invariant distribution) vs weak (invariant moments)
- **Fundamental Processes**: white noise and random walk $\succ$ building blocks for ARIMA
- **ACF and PACF**: tools for model identification

Questions?

Bibliography I

Time Series Fundamentals

- Wold, H. (1938). *A Study in the Analysis of Stationary Time Series*, Almqvist & Wiksell.
- Hamilton, J.D. (1994). *Time Series Analysis*, Princeton University Press.
- Brockwell, P.J., & Davis, R.A. (2016). *Introduction to Time Series and Forecasting*, 3rd ed., Springer.

Decomposition and Exploratory Analysis

- Cleveland, R.B., Cleveland, W.S., McRae, J.E., & Terpenning, I. (1990). STL: A Seasonal-Trend Decomposition Procedure Based on Loess, *Journal of Official Statistics*, 6(1), 3–33.
- Hyndman, R.J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice*, 3rd ed., OTexts.

Bibliography II

Exponential Smoothing and ETS Fundamentals

- Holt, C.C. (1957/2004). Forecasting Seasonals and Trends by Exponentially Weighted Moving Averages, *International Journal of Forecasting*, 20(1), 5–10.
- Winters, P.R. (1960). Forecasting Sales by Exponentially Weighted Moving Averages, *Management Science*, 6(3), 324–342.
- Hyndman, R.J., Koehler, A.B., Ord, J.K., & Snyder, R.D. (2008). *Forecasting with Exponential Smoothing: The State Space Approach*, Springer.

Online Resources and Code

- **Quantlet**: https://quantlet.com ≻ Code repository for statistics
- **Quantinar**: https://quantinar.com ≻ Learning platform for quantitative methods
- **GitHub TSA_ch0**: https://github.com/QuantLet/TSA/tree/main/TSA_ch0