HSLU Hochschule Luzern

Time Series Analysis

Foundations II

Dr. Ludovic Amruthalingam
ludovic.amruthalingam@hslu.ch

Informatik

Outline

- Stationarity
- Visual inspection
- Statistical tests
- Non-linear transforms
- Decomposition
- Differencing

Stationarity

Statistical properties constant over time

Strict stationarity: the joint distribution of any collection of the TS RVs is identical to that of the shifted RVs:

$$P\left[X_{t_1} = x_{t_1}, \dots X_{t_k} = x_{t_k}\right] = P\left[X_{t_1+h} = x_{t_1+h}, \dots X_{t_k+h} = x_{t_k+h}\right] \quad \forall\, k \geq 1, \forall\, t_1, \dots, t_k, \forall\, h \geq 0$$

(Weak) stationarity: both mean and variance are **constant** over time, while the auto-covariance depends **only on the lag** $h = |s - t|$ between observations :

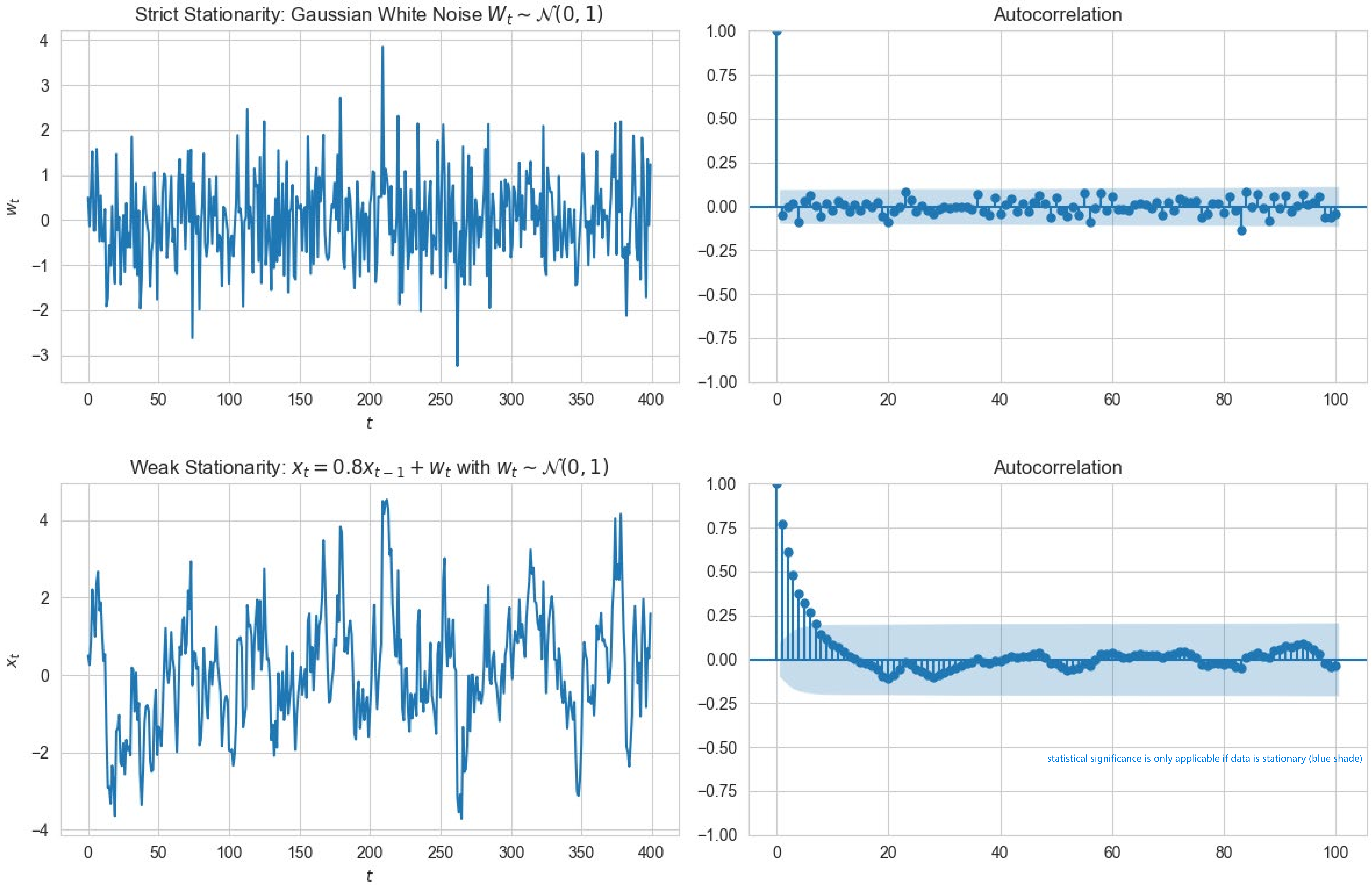
$$\gamma(t+h, t) = cov(X_{t+h}, X_t) = cov(X_h, X_0) = \gamma(h, 0) = \gamma(h) \quad \forall h, t$$

The **auto-correlation** only depends on the **shift** since

$$\rho(t+h, t) = \frac{\gamma(t+h, h)}{\sqrt{Var(X_{t+h})Var(X_t)}} = \frac{\gamma(h, 0)}{\sqrt{Var(X_h)Var(X_0)}} = \frac{\gamma(h)}{\gamma(0)} = \rho(h)$$

A **white noise process is stationary** since $W_t \sim WN(0, \sigma^2)$ and $\gamma_{W_t}(s, t) = 0\ \forall s \neq t$.

Strict stationarity implies stationarity, but the reverse is **not** true in general (Gaussian processes are an exception).


Strict Stationarity: Gaussian White Noise $W_t \sim \mathcal{N}(0, 1)$
w_t
t
Autocorrelation
Weak Stationarity: $x_t = 0.8x_{t-1} + w_t$ with $w_t \sim \mathcal{N}(0, 1)$
x_t
t
Autocorrelation
statistical significance is only applicable if data is stationary (blue shade)

Estimating the statistical properties of a time series

Typically, only a **single realization** of a TS is available. With non-stationary TS,

- Actual mean, variance and covariance (statistical properties) may **vary over time**.
- **Cannot** rely on past values to **estimate future** relationships.

With a stationary TS, its **statistical properties are constant over time** and can be estimated from past values.

Sample mean	Sample variance	Sample auto-covariance	Sample auto-correlation
$\hat{\mu} = \frac{1}{n}\sum_{i=0}^{n} x_i$	$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=0}^{n}(x_i - \hat{\mu})^2$	$\hat{\gamma}(h) = \frac{1}{n}\sum_{i=0}^{n-\lvert h\rvert}\left(x_{i+\lvert h\rvert} - \hat{\mu}\right)(x_i - \hat{\mu})$	$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}$

TS analysis typically starts with removing trends, seasonality, and heteroskedasticity to get **stationary residuals**, which can then be modeled.

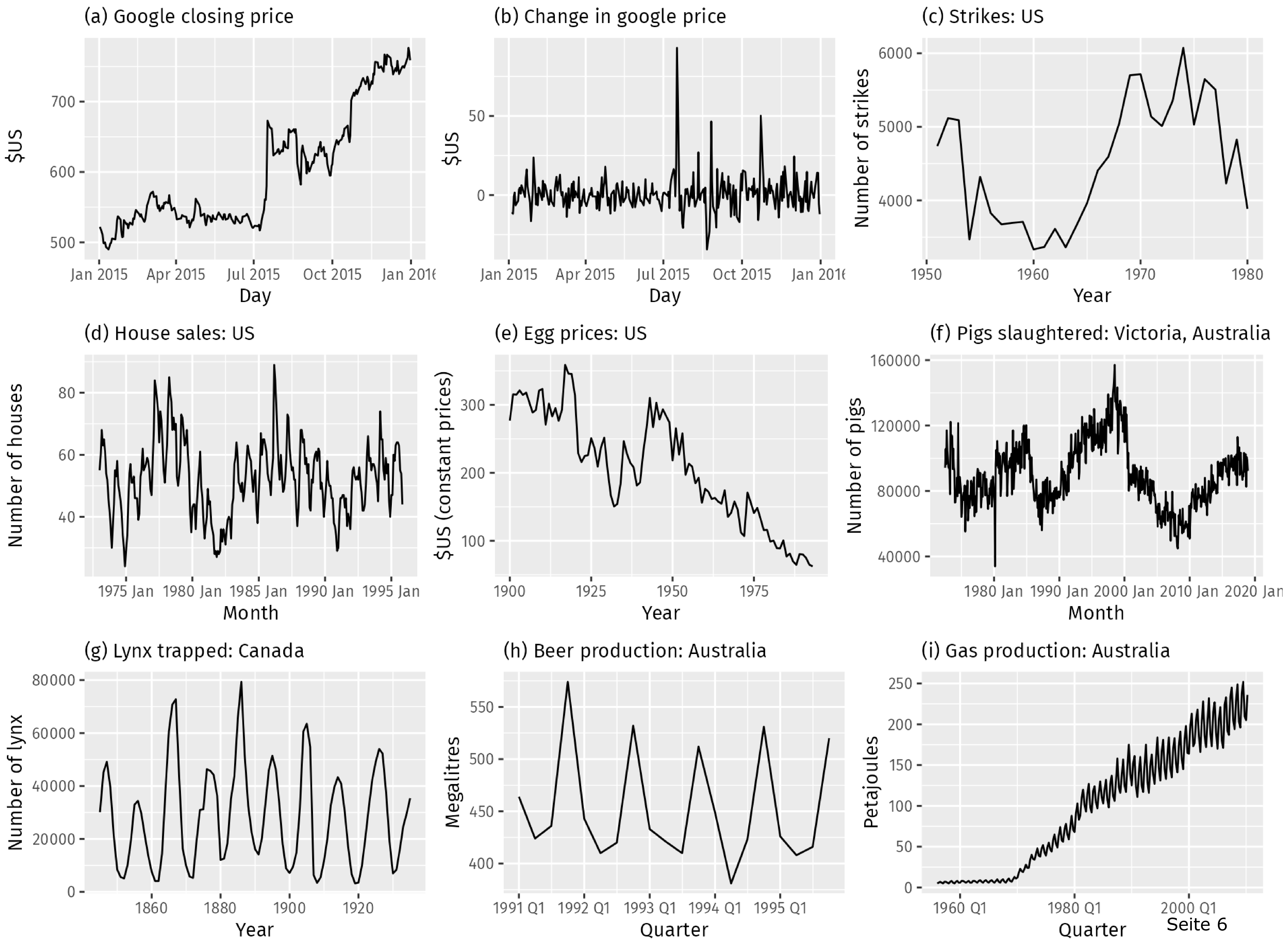
Checking stationarity – Visual inspection of time plots

Which of these TS are stationary?

Non-stationary patterns
- Trends
- Seasonality with trend
- Cycles with trend
- Changing variance
- Apparent breaks in the data

Seasonality

Cyclic patterns can be stationary: ups and downs are random but stable in distribution (time-invariant).

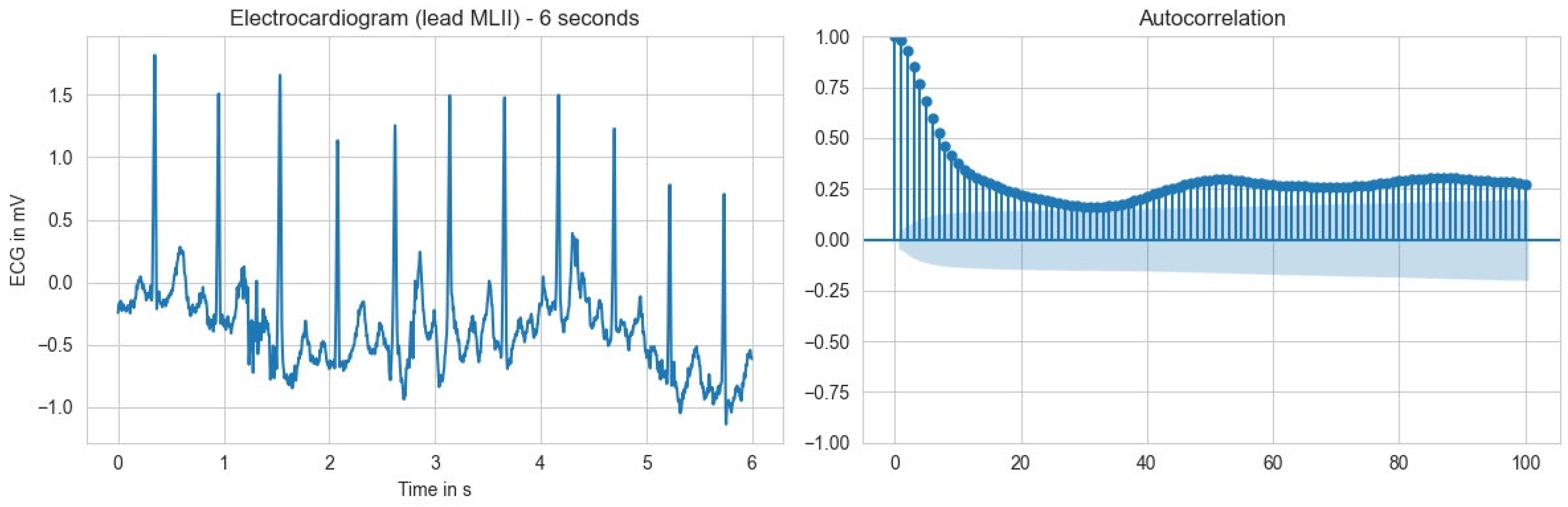


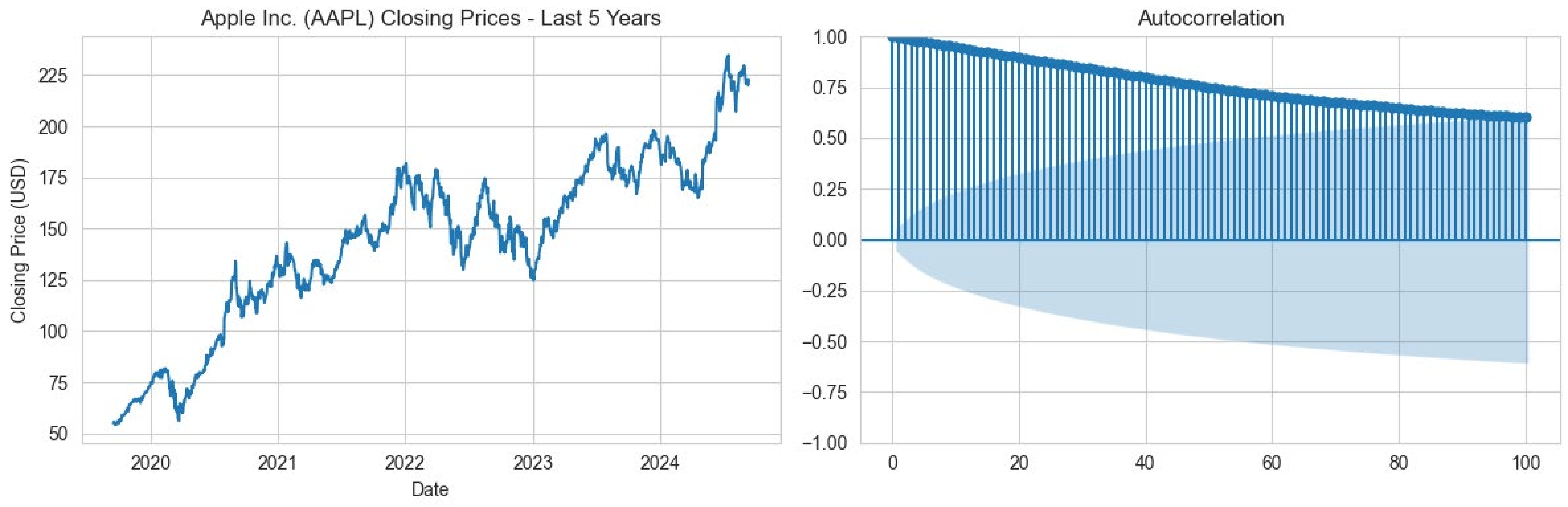
Forecasting: Principles and Practice (3rd ed)

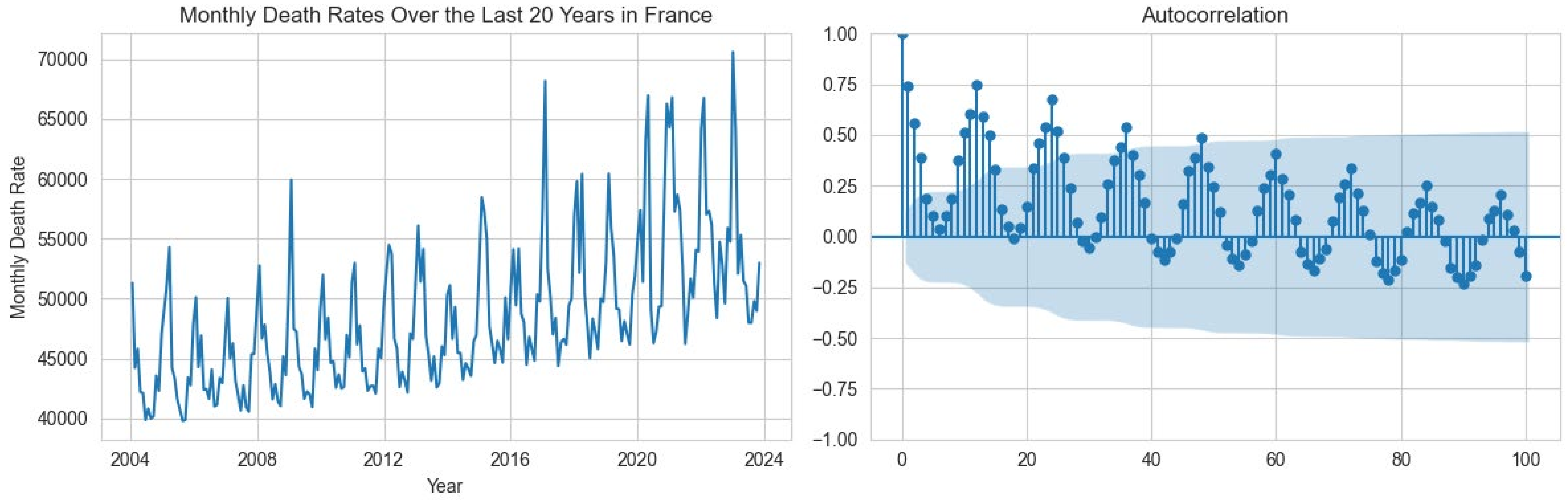
Checking stationarity – Visual inspection of correlograms

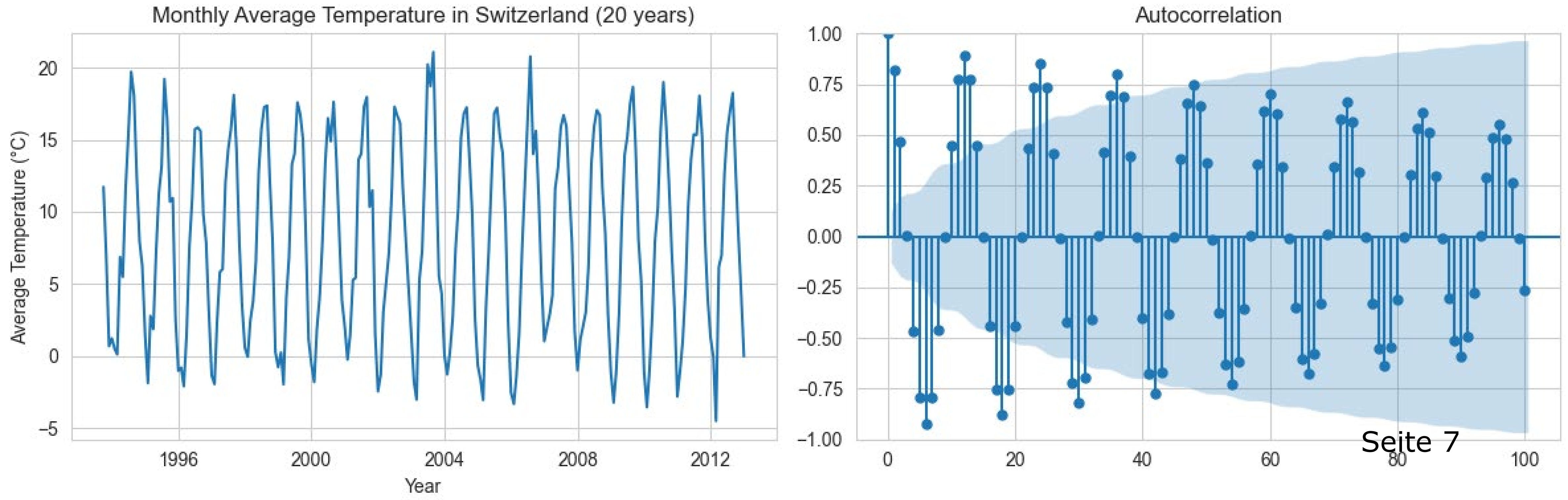
Correlogram: in stationary TS, the auto-correlations typically **drop off quickly** as the lag increases.
If the coefficients drop quickly but then show a pattern, it typically indicates **weak stationarity**.

ACF estimates are biased: shrunken towards zero for higher lags → consider only the **first ~ n/4 lags**.









Checking stationarity – Statistical tests

Null hypothesis (H_0): TS has a **unit root** e.g., random walk $x_t = \mathbf{1} \cdot x_{t-1} + w_t$

- TS value is strongly dependent on its previous value => time-dependent mean/variance, slow-decaying $\rho(h)$.
- Augmented Dickey-Fuller (ADF) test, Phillips-Perron (PP) test

H_0: TS is **trend stationary** (removing trend leaves a stationary TS)

- Detect whether TS is mean-reverting, has constant variance and auto-correlation over time.
- Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test.

Tests for structural breaks: Zivot-Andrews (ZA) test, Chow test, Bai-Perron test.

Statistical tests focus on specific stationarity aspects, but do **not** test stationarity in a broad sense.

They struggle to detect general forms of non-stationarity → **couple with visual analysis**.

Achieving stationarity

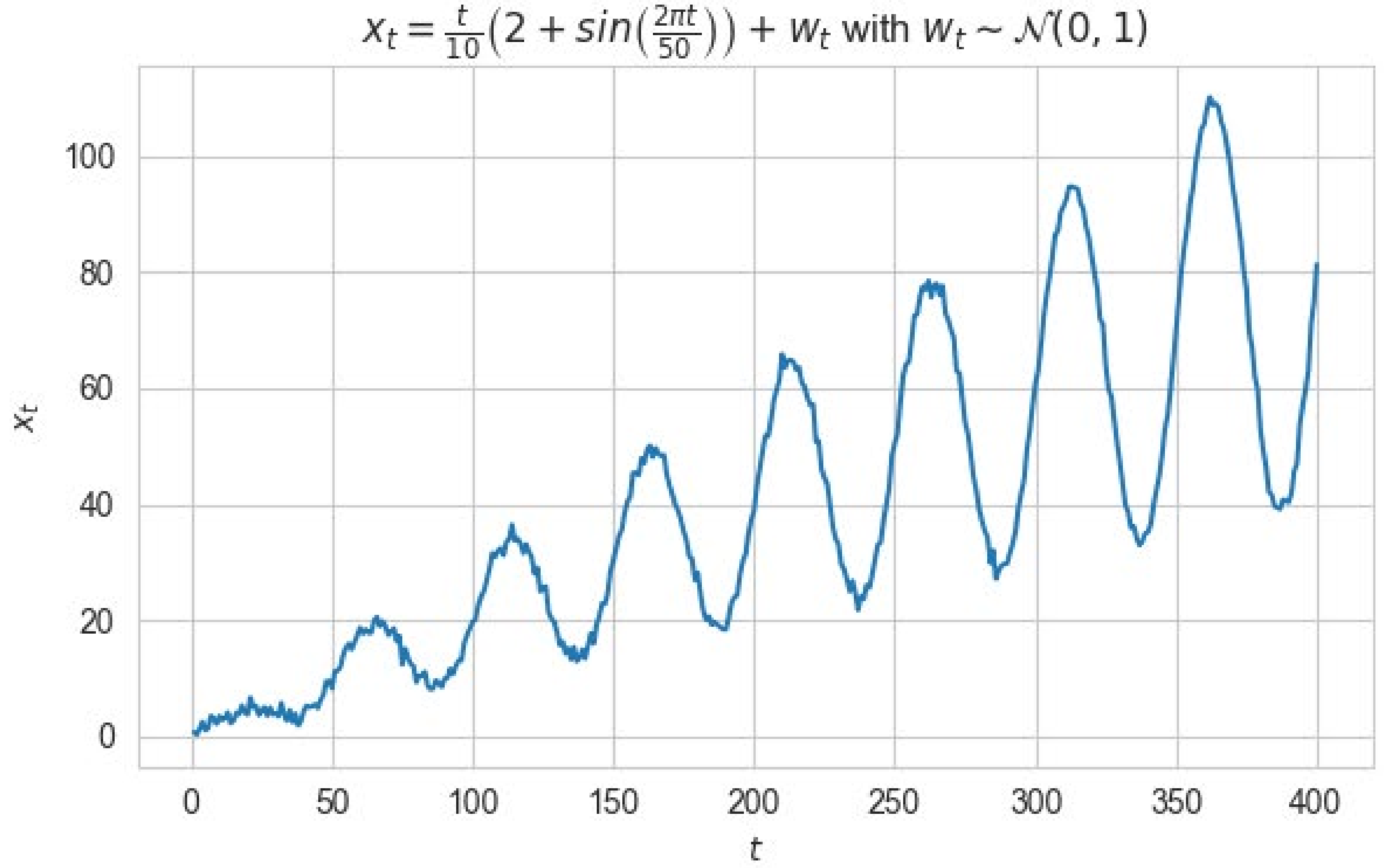
To make a TS stationary, apply either of the following steps (if necessary).

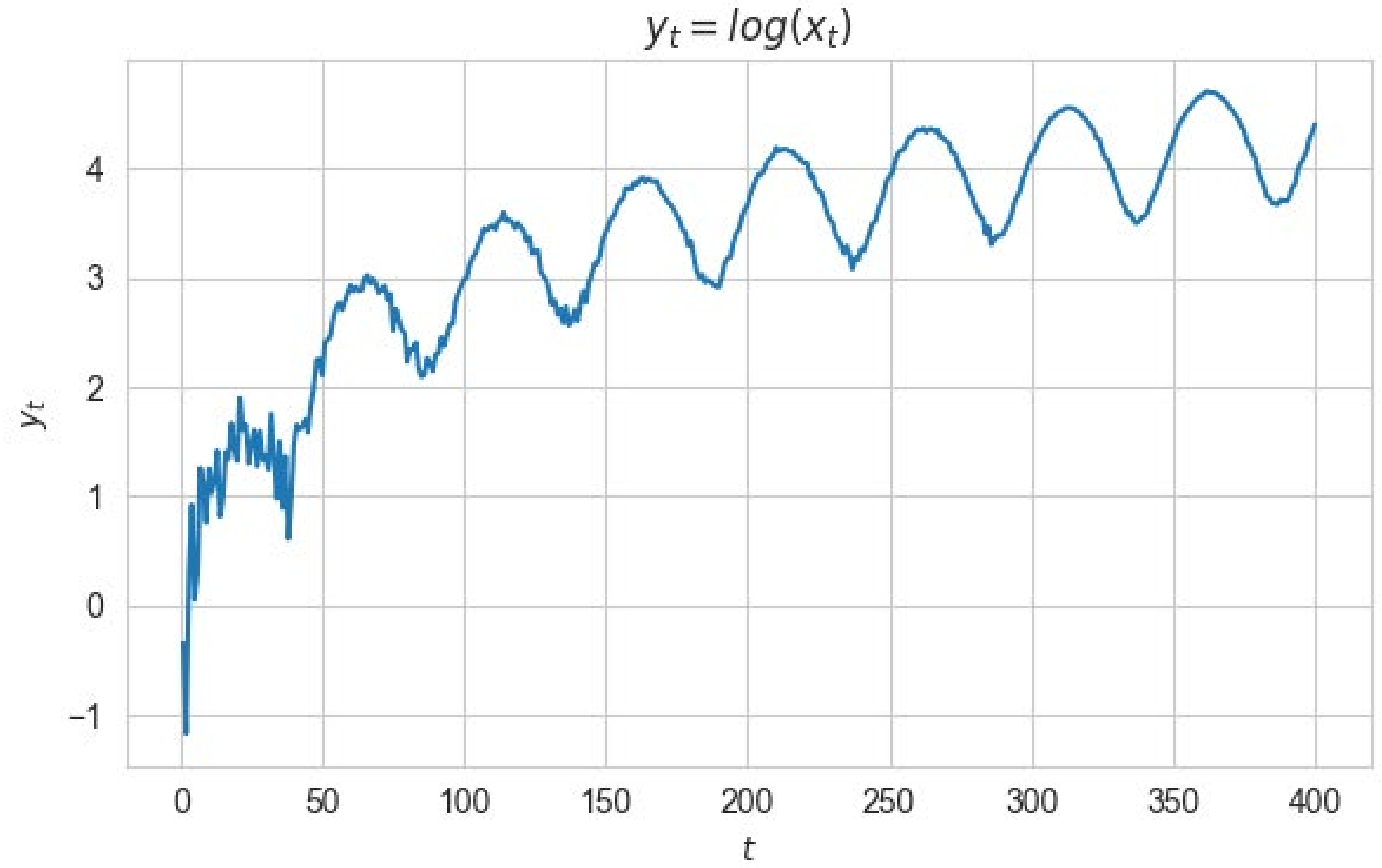
1. **Non-linear transformations** to stabilize variance
   - Log transform
   - Power transforms
   - Box-Cox transform
2. **Decomposition** to isolate trend and seasonal components
   - STL decomposition
3. **Differencing**
   - On the decomposition remainder if the residuals are non-stationary.
   - On the TS directly when decomposition is either difficult (e.g., weak patterns) or not needed.

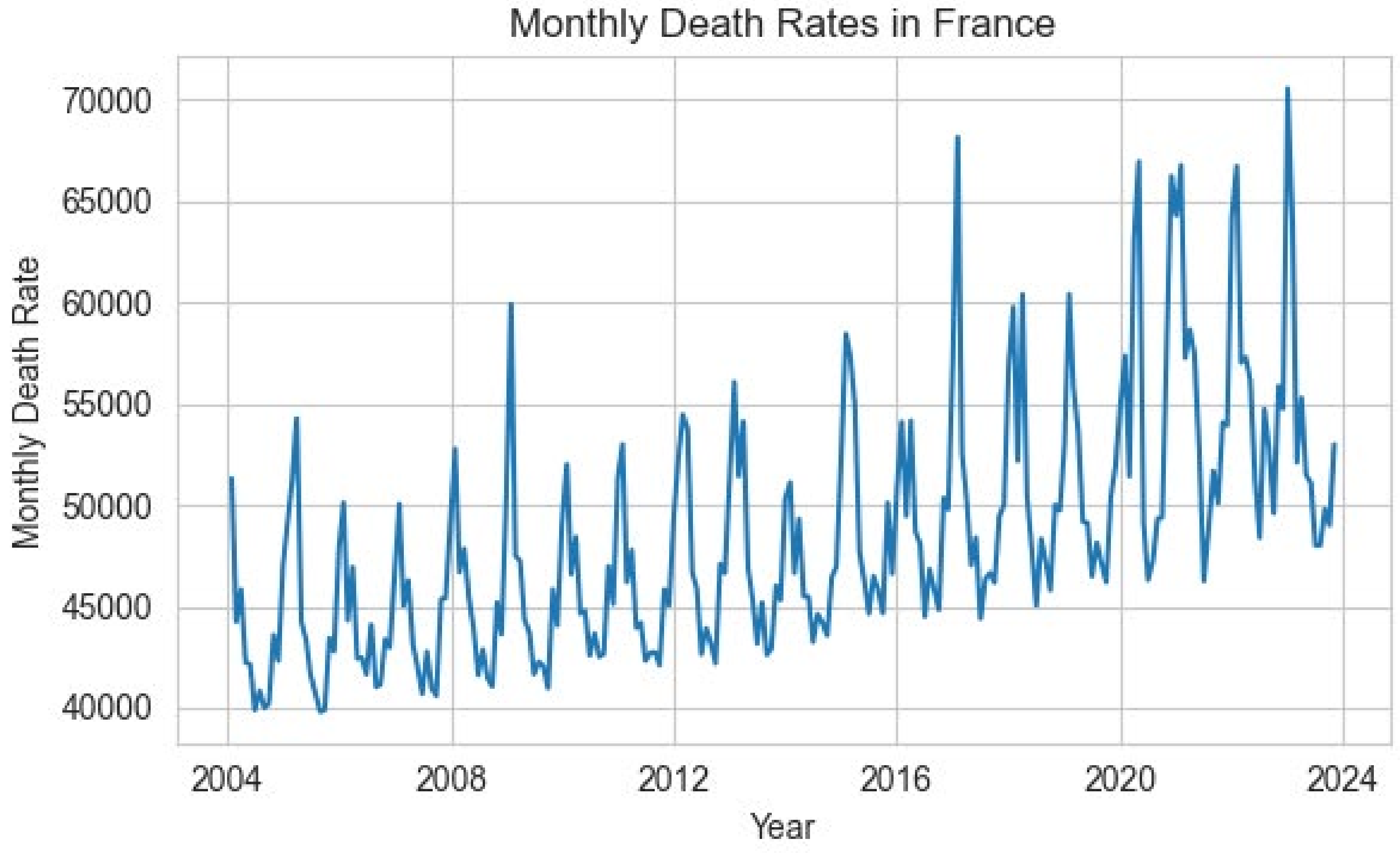
Non-linear transformations

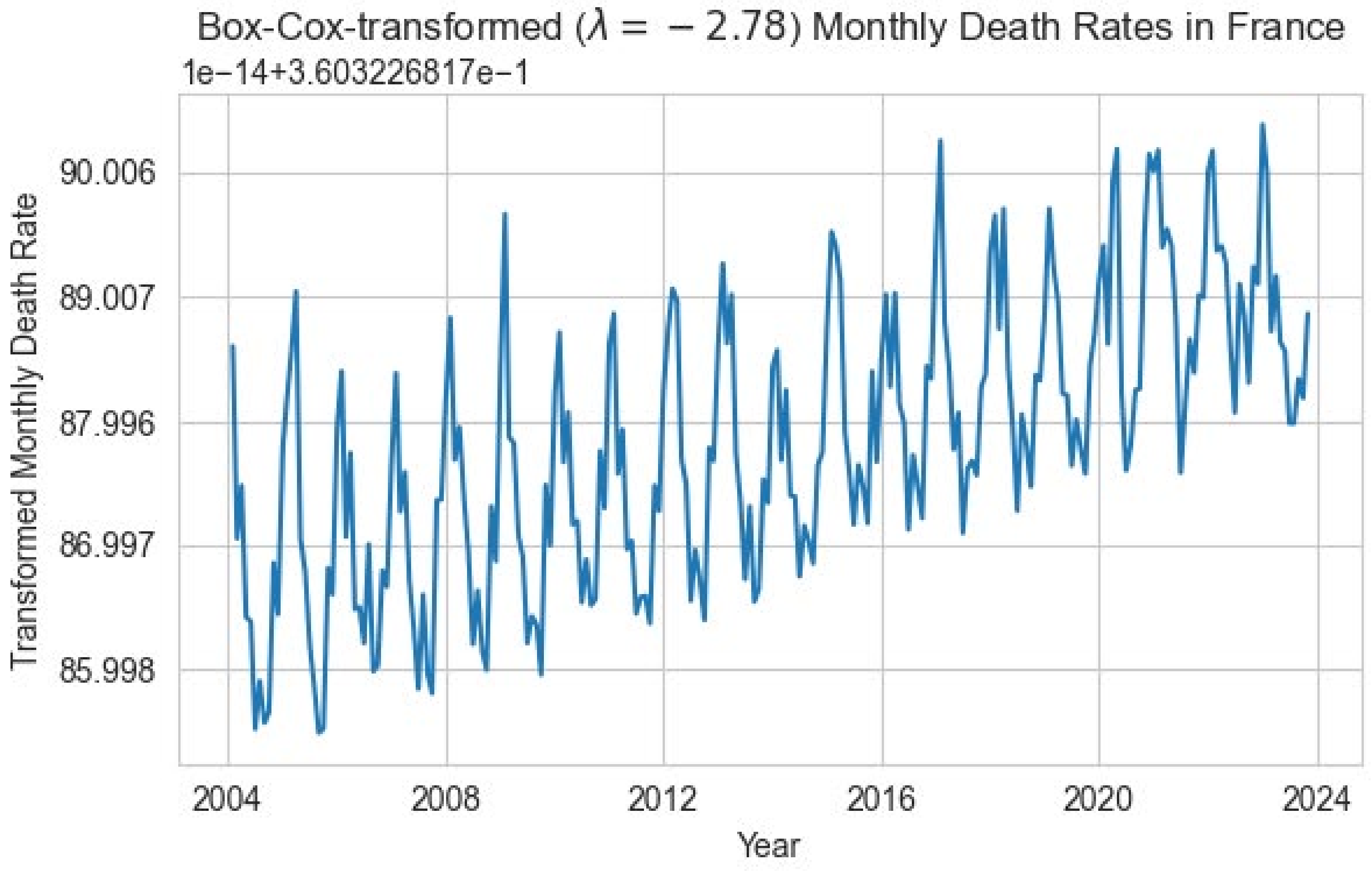
Goal: **stabilize** variance.

Log transform	Power transform	Box-Cox transform
$y_t = \log(x_t)$	$y_t = x_t^{\lambda}$	$y_t = \begin{cases} \log(x_t) & if\ \lambda = 0 \\ \dfrac{\left(sign(x_t)\lvert x_t\rvert^{\lambda} - 1\right)}{\lambda} & if\ \lambda \neq 0 \end{cases}$
• Interpretable • Use on skewed data with long right tail • Converts multiplicative relationships into additive	• Lack interpretability • Typically, $\lambda = \frac{1}{2}$ or $\frac{1}{3}$ • Use on count data (Poisson distribution) • Reduce impact of large values	• Flexible transformation • Determine optimal λ for each TS


$x_t = \frac{t}{10}\left(2 + sin\left(\frac{2\pi t}{50}\right)\right) + w_t$ with $w_t \sim \mathcal{N}(0, 1)$
x_t
t


$y_t = log(x_t)$
y_t
t


Monthly Death Rates in France
Monthly Death Rate
Year


Box-Cox-transformed ($\lambda = -2.78$) Monthly Death Rates in France
1e−14+3.603226817e−1
Transformed Monthly Death Rate
Year

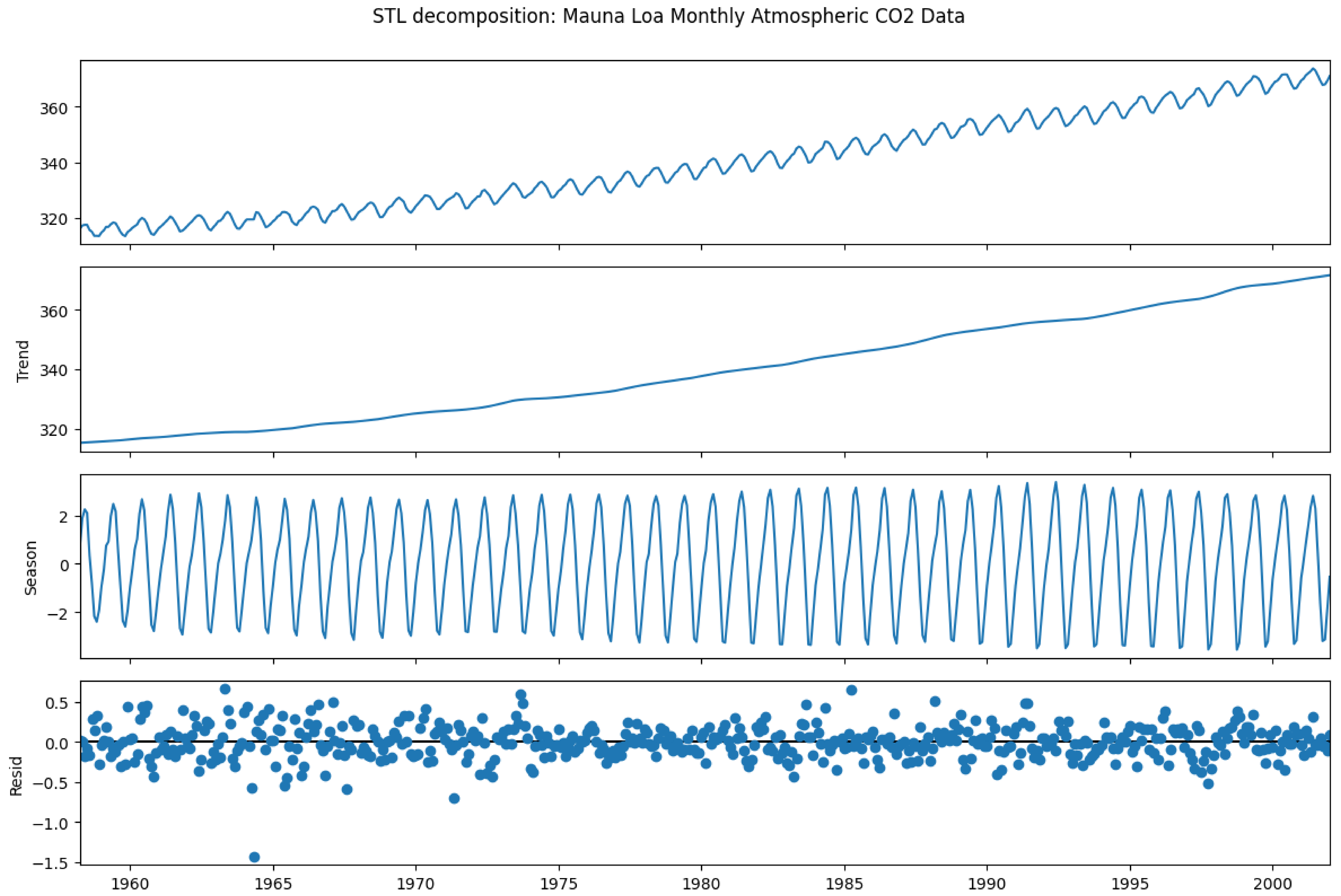
Recap – Decomposition

Separate TS into three components: $X_t = U_t + S_t + R_t$

- Trend/trend-cycle U_t: increasing/decreasing pattern in the data.
- Seasonality S_t: repeating pattern with roughly fixed period.
- Remainder $\mathrm{R_t}$: everything remaining (**residuals**).

Given a time series realization $\{x_1, x_2, \dots x_n\}$:

1. Estimate trend component $\hat{u}_t$
   – LOESS – locally estimated scatterplot smoothing (STL)
2. Detrend time series $\hat{d}_t \; = x_t - \hat{u}_t$
3. Estimate seasonality component $\hat{s}_t$
   – LOESS – locally estimated scatterplot smoothing (STL)
4. Deseasonalize TS to estimate the remainder $\hat{r}_t = \hat{d}_t - \hat{s}_t$



statsmodels.tsa.seasonal.STL

Keeling et al. "Atmospheric CO2 concentrations derived from flask air samples at sites in the SIO network." Trends: a compendium of data on Global Change (2004).

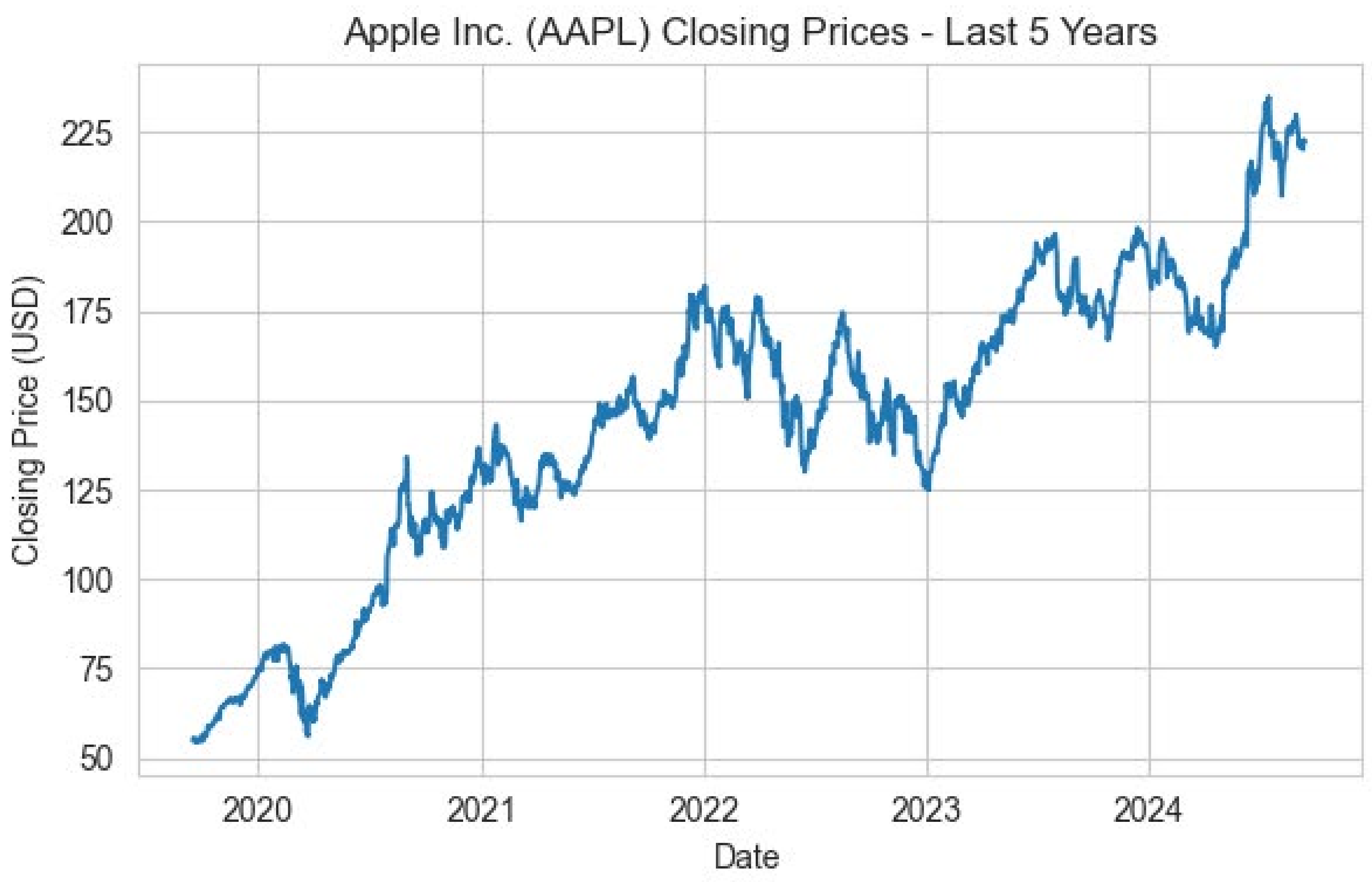
Differencing

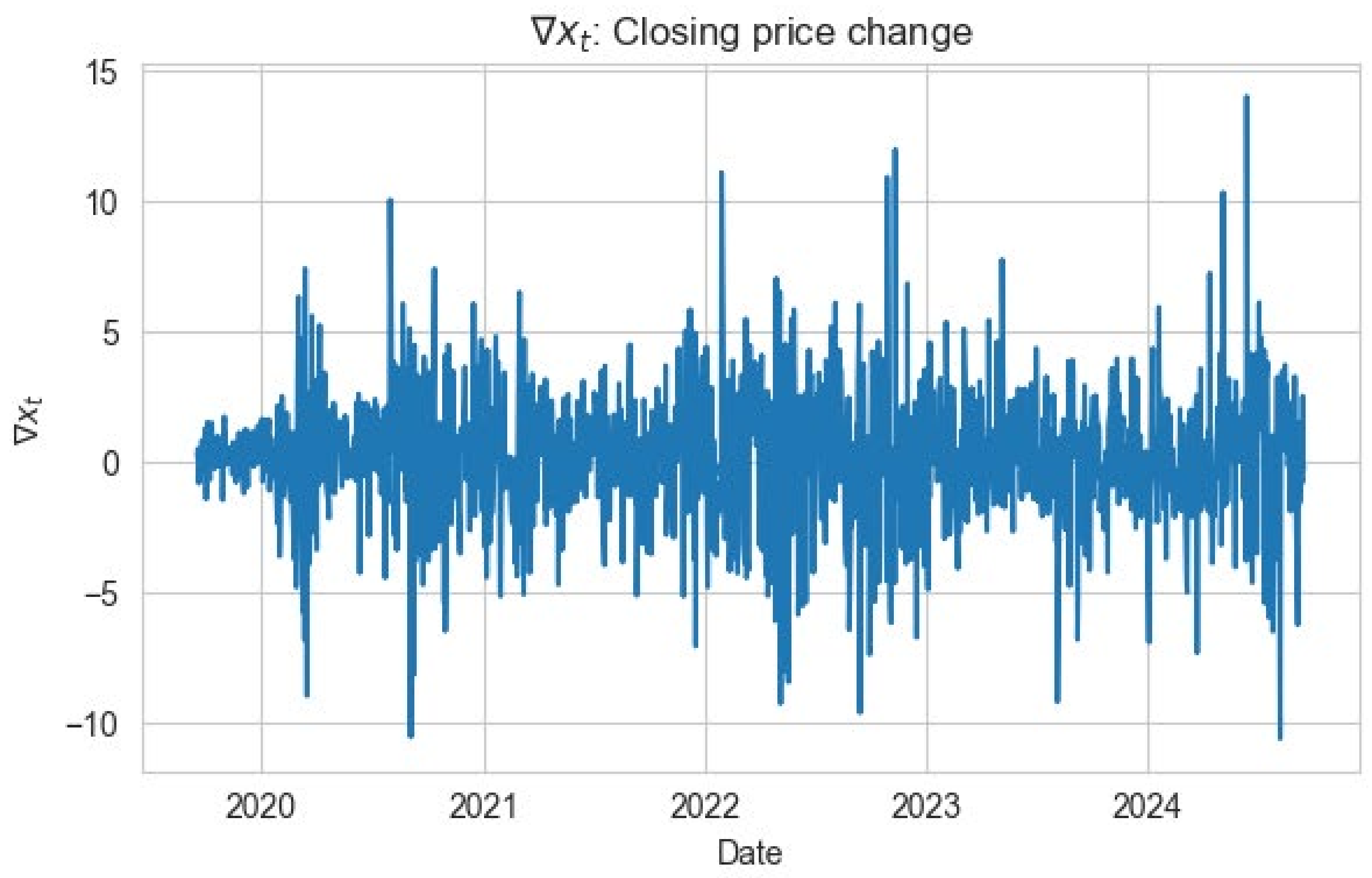
Studying a TS value **changes** rather than its values: $\nabla X_t = X_t - X_{t-1}$ (**first-order differencing**)
∇_{h} is called the **lag-h difference operator**: $\nabla_h X_t = X_t - X_{t-h}$ with ∇_1 denoted as ∇.
Differencing does **not** obtain explicit estimate of trend, season and remainder.

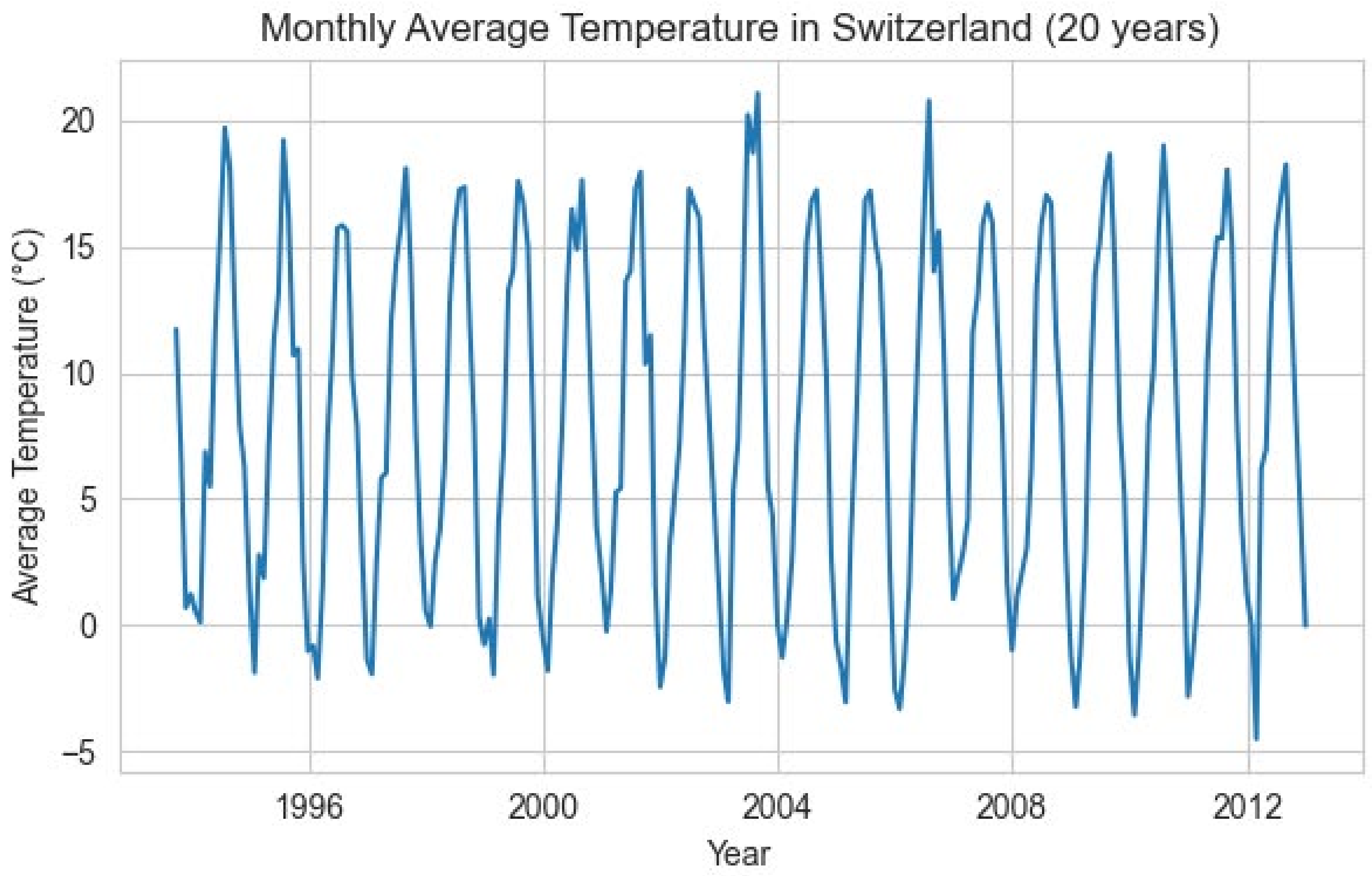
Differencing can **remove the trend** component: consider a TS with linear trend $X_t = \alpha + \beta t + W_t$
Then $\nabla X_t = \alpha + \beta t + W_t - (\alpha + \beta(t-1) + W_{t-1}) = \beta + W_t - W_{t-1}$

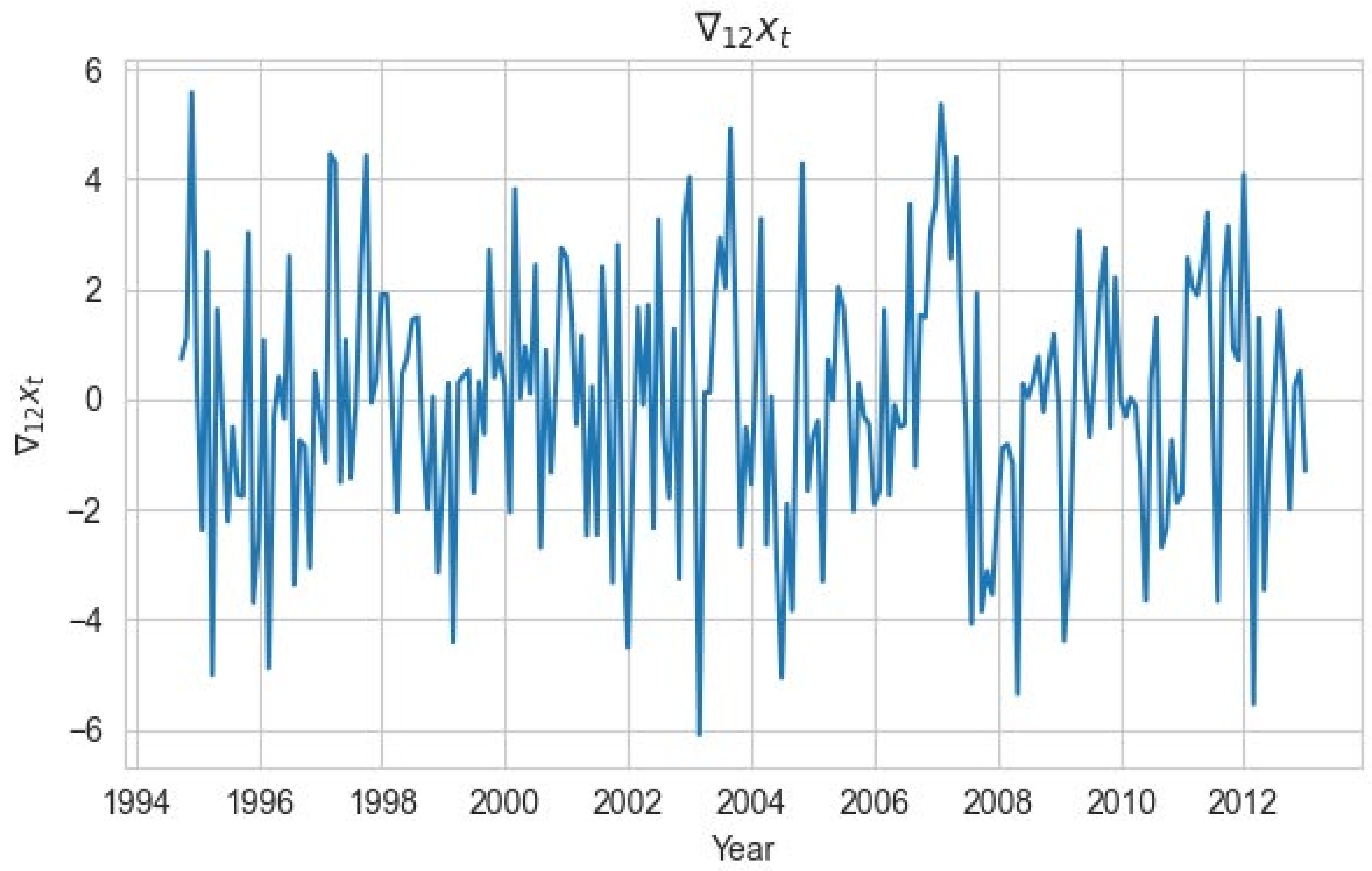
Differencing can remove the seasonal component: consider $X_t = A\sin\left(\frac{2\pi t}{P}\right) + W_t$
Then $\nabla_P X_t = X_t - X_{t-P} = A\sin\left(\frac{2\pi t}{P}\right) + W_t - \left(A\sin\left(\frac{2\pi(t-P)}{P}\right) + W_{t-P}\right) = W_t - W_{t-P}$ (**seasonal differencing**)

Differencing introduces **artificial dependencies**: $\mathrm{cov}(\nabla X_t, \nabla X_{t-1}) = cov(W_t - W_{t-1}, W_{t-1} - W_{t-2}) = -cov(W_{t-1}, W_{t-1}) \neq 0$
→ Keep differences **interpretable**: first-order, seasonal differencing.


Apple Inc. (AAPL) Closing Prices - Last 5 Years
Closing Price (USD)
50
75
100
125
150
175
200
225
2020
2021
2022
2023
2024
Date


∇x_t: Closing price change
∇x_t
−10
−5
0
5
10
15
2020
2021
2022
2023
2024
Date


Monthly Average Temperature in Switzerland (20 years)
Average Temperature (°C)
−5
0
5
10
15
20
1996
2000
2004
2008
2012
Year


$\nabla_{12} x_t$
$\nabla_{12} x_t$
−6
−4
−2
0
2
4
6
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
Year

Higher-order differencing

Iterative differencing: $\nabla^2 X_t = \nabla\nabla X_t = \nabla(X_t - X_{t-1}) = X_t - X_{t-1} - (X_{t-1} - X_{t-2}) = X_t - 2X_{t-1} + X_{t-2}$

B is called the **Backshift operator**: $B^h X_t = X_{t-h}$ with B^1 denoted as B and $BX_t = X_{t-1}$

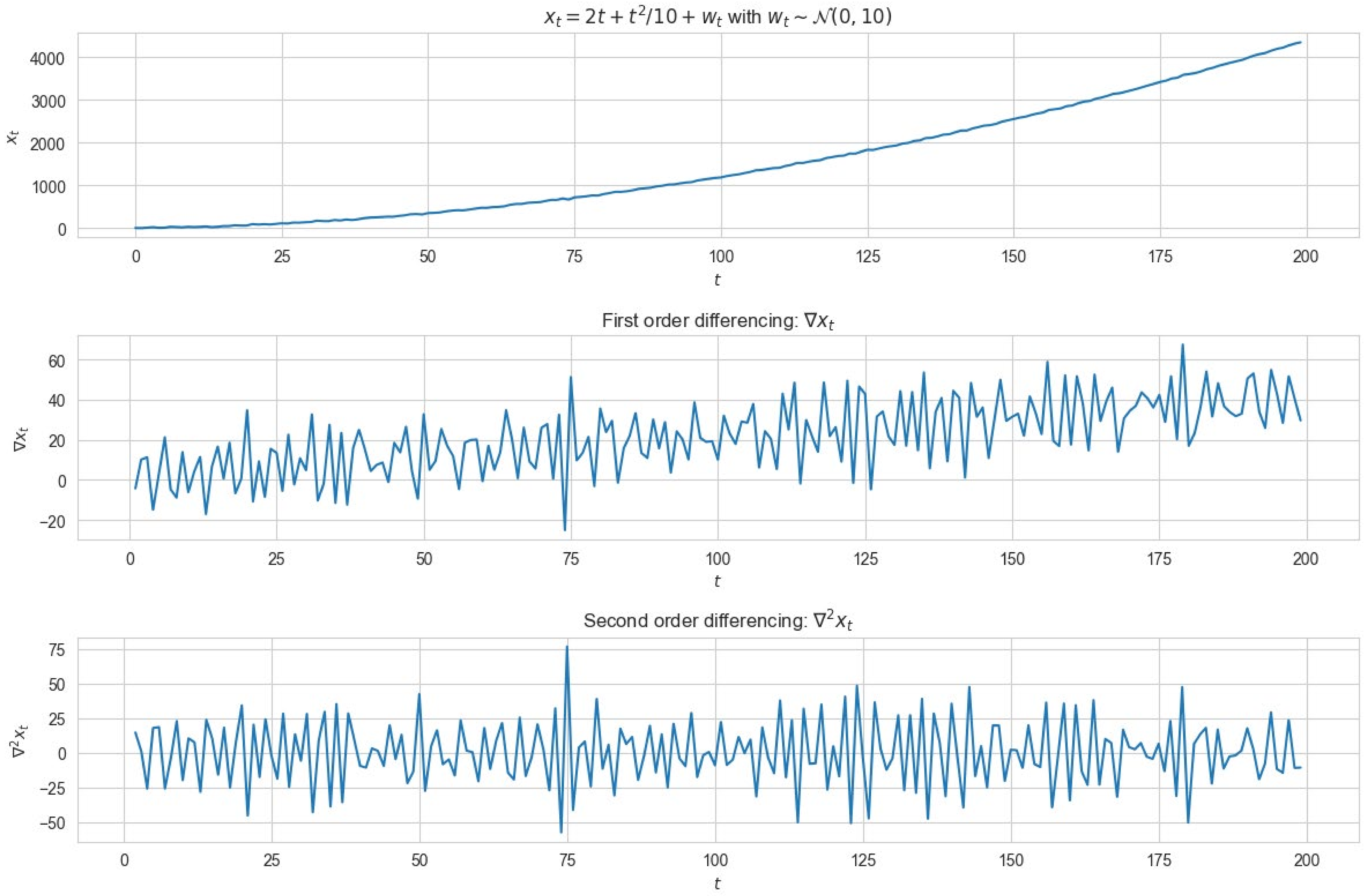
First-order differencing: $\nabla X_t = X_t - X_{t-1} = X_t - BX_t = (1-B)X_t$
Second-order differencing: $\nabla^2 X_t = \nabla\nabla X_t = \nabla(1-B)X_t = (1-B)^2 X_t$
k[th]-order differencing: $\nabla^k X_t = \nabla^{k-1}\nabla X_t = (1-B)^k X_t$

Note that $\nabla_h^k = \left(1 - B^h\right)^k$

Start with seasonal differencing, continue with differencing if some trend remain e.g., $\nabla^k \nabla_P X_t$

When trend is non-linear, apply non-linear transforms **before** any differencing.

Use higher-order differencing to remove **polynomial trend** e.g., second-order for quadratic trend.


$x_t = 2t + t^2/10 + w_t$ with $w_t \sim \mathcal{N}(0, 10)$
x_t
t
First order differencing: ∇x_t
∇x_t
t
Second order differencing: $\nabla^2 x_t$
$\nabla^2 x_t$
t

Exercise

Extend lecture 1 & 2 exercises with stationarity analysis: are the time series **stationary**?

- Visual inspection
- Statistical tests

Make the time series stationary

- Experiment with **non-linear transformations**.
- Experiment with **decomposition**.
- Experiment with **differencing**.

Explore **artificial auto-correlations** induced by differencing

- Start with a stationary time series.
- Apply differencing at various order.
- Compare **ACF**.