JOHNS HOPKINS
CAREY BUSINESS SCHOOL

BU.610.740: Forecasting Models for Business Intelligence
Chapter 3: ARIMA Models - Part II

Ali Eshragh

Spring II, 2024

Outline

1. **Integrated Models for Non-Stationary Data**
   - Exploratory Data Analysis
   - ARIMA Model
   - Box-Cox Transformation
   - SARIMA Model
   - Hybrid SARIMA-Regression Model

Non-Stationary Time Series

- All `ARMA` models are constructed based on the **stationarity** assumption:
  1. Constant **mean** over time
  2. **Autocorrelation** function depends only on the lag, not on the actual times

- Therefore, `ARMA` models and their derivatives, including `AR` and `MA` models, **cannot** be **directly** applied to non-stationary time series.

- In reality, many time series do **not** exhibit stationarity, often showing **trends** in mean.

Example

Example (Linear Trend in Mean)

Consider the **time series**

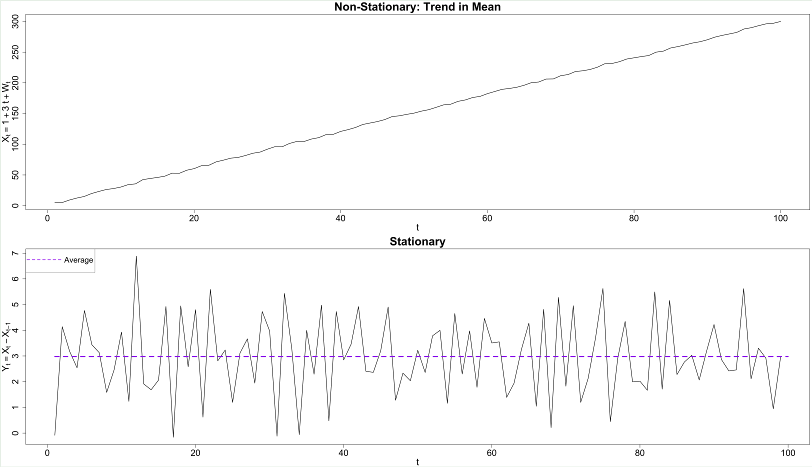
$$X_t = \beta_0 + \beta_1 t + W_t$$

where β_0 and β_1 are **constants** and W_t is a **Gaussian white noise** series with variance σ_W^2

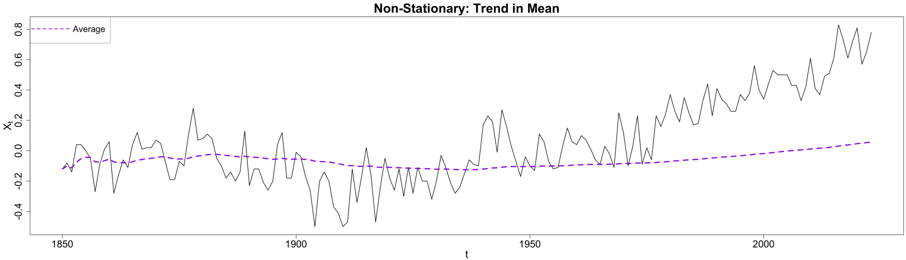
1. Determine whether X_t is **stationary**.
2. Show that the time series $Y_t := X_t - X_{t-1}$ is **stationary**.

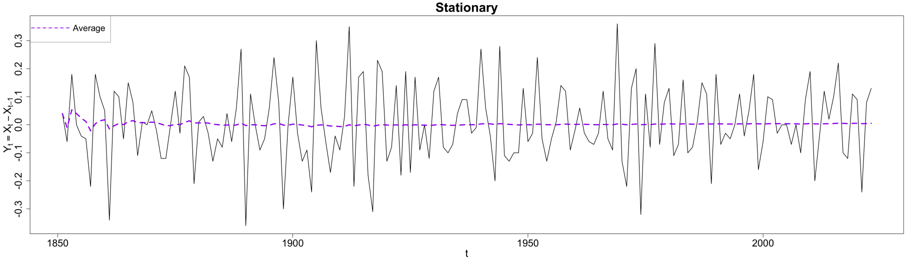
Example: Continued

Example



Case Study: Global Temperature vs. First Differencing





Differencing

Definition (Backshift Operator)

The **backshift operator** is defined by

$$\mathtt{B}x_t := x_{t-1}$$

and it is extended to **powers** such that

$$\mathtt{B}^k x_t = \mathtt{B}^{k-1}(\mathtt{B}x_t) = \mathtt{B}^{k-1}x_{t-1} = \cdots = x_{t-k}$$

Definition (Difference of Order d)

Difference of order d is defined by $(1-\mathtt{B})^d$, where the **operator** $(1-\mathtt{B})^d$ can be expanded **algebraically** to evaluate it for higher integer values of d.

Example

Example

- Difference of **order** 1 in the **Global Temperature** case study:

$$Y_t = X_t - X_{t-1} = X_t - \mathtt{B}X_t = (1 - \mathtt{B})X_t$$

- Difference of **order** 2:

$$(1 - \mathtt{B})^2 X_t = (1 - 2\mathtt{B} + \mathtt{B}^2)X_t = X_t - 2X_{t-1} + X_{t-2}$$

Assignment 5

Question 1

Consider the **time series**

$$X_t = \beta_0 + \beta_1 t + \beta_2 W_{t-1} + W_t$$

where β_0, β_1, and β_2 are **constants** and W_t is a **Gaussian white noise** series with variance σ_W^2

1. Determine whether X_t is **stationary**.
2. Show that the time series $Y_t := (1 - \mathtt{B})X_t$ is **stationary**.

Integrated Autoregressive Moving Average Model

Definition (ARIMA Model)

A **time series** $\{X_t : t = 0, \pm 1, \pm 2, \ldots\}$ is an $\texttt{ARIMA}(p, d, q)$ model if the time series Y_t, defined by the **difference of order** d, represented by

$$Y_t := (1 - \texttt{B})^d X_t$$

is **modeled** by an $\texttt{ARMA}(p, q)$ model.

Forecasting with ARMA Models for Non-Stationary Time Series

- Assuming that **differencing** a **non-stationary** time series X_t to order d results in a **stationary** time series $Y_t = (1 - \mathtt{B})^d X_t$
- Fit an $\mathtt{ARMA}(p, q)$ model to the stationary time series Y_t for **forecasting**, which is **equivalent** to fitting an $\mathtt{ARIMA}(p, d, q)$ model to X_t
- Use **forecasts** from Y_t with **linear transformation** to **predict** X_t values

Example

- Consider a **non-stationary** time series $X_1, \ldots, X_n$ which becomes **stationary** by differencing at one lag as $Y_t = X_t - X_{t-1}$ for $t = 2, 3, \ldots, n$
- **Forecast** future values of Y_t by fitting an ARMA model:

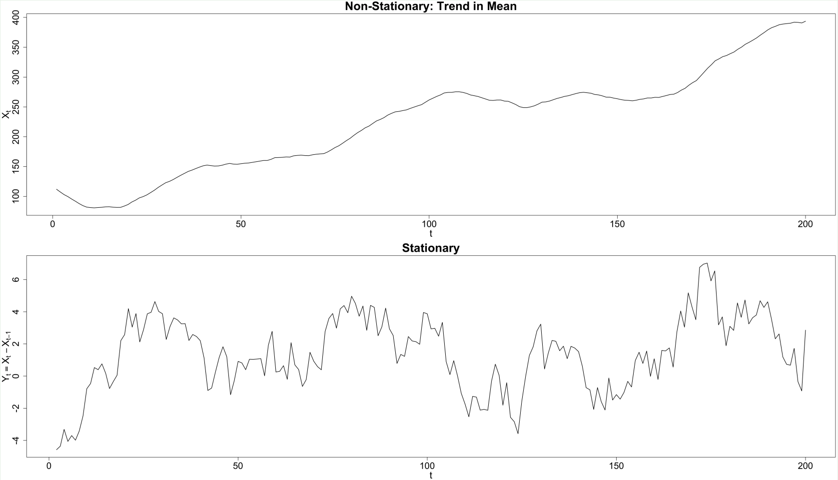
$$\widehat{Y}_{n+1}, \widehat{Y}_{n+2}, \ldots$$

- **Predict** future values of X_t as:

$$\widehat{X}_{n+1} = X_n + \widehat{Y}_{n+1}, \; \widehat{X}_{n+2} = \widehat{X}_{n+1} + \widehat{Y}_{n+2}, \; \ldots$$

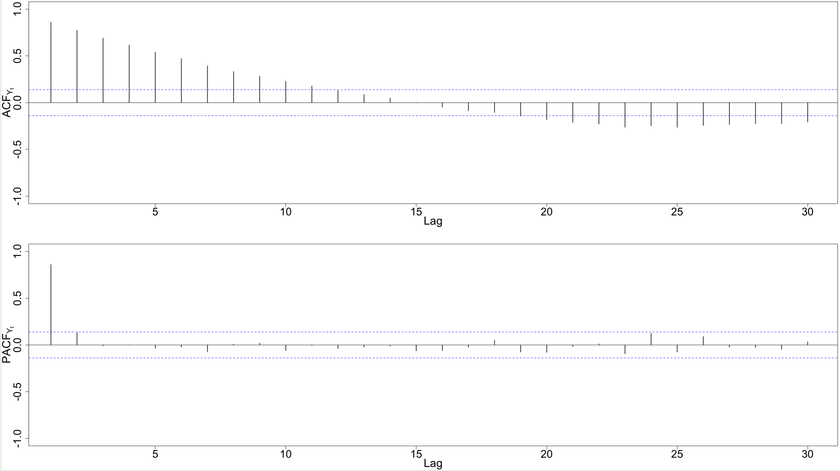
Example

Example (Non-Stationary Time Series Data with Trend in Mean)



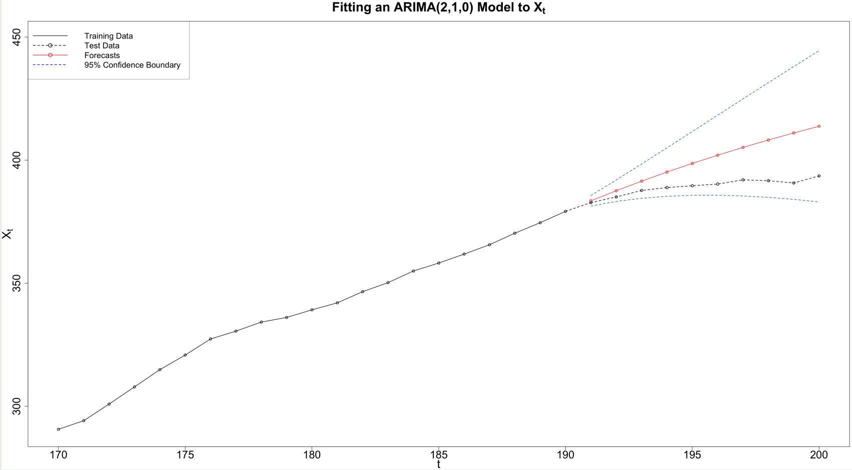
Example: Continued

Example (Non-Stationary Time Series Data with Trend in Mean)



Example: Continued

Example (Non-Stationary Time Series Data with Trend in Mean)



Lack of Fit Test: Ljung-Box Statistics

$$\begin{cases} H_0 : \text{Model does } \textbf{not} \text{ exhibit lack of fit} \\ H_a : \text{Model } \textbf{exhibits} \text{ lack of fit} \end{cases}$$

- The **Ljung-Box** test can be considered a **lack of fit** test for time series models
- It checks for **autocorrelation** in the **residuals** of a fitted `ARIMA` model
- If they exhibit **autocorrelation**, this suggests that the model has **not** fully captured the underlying **structure** of the time series data, indicating a **lack of fit**

Forecast Accuracy Metrics

- **Mean Absolute Error**: $\text{MAE} = \frac{1}{m}\sum_{t=1}^{m} |x_{n+t} - \hat{x}_{n+t}|$

  where x_{n+t} are the actual **test values**, $\hat{x}_{n+t}$ are the **predicted values**, n is the **number** of observations in the **training set**, and m is the **number** of observations in the **test set**

- **Mean Absolute Percentage Error**: $\text{MAPE} = \frac{1}{m}\sum_{t=1}^{m} \left|\frac{x_{n+t} - \hat{x}_{n+t}}{x_{n+t}}\right| \times 100\%$

- **Root Mean Square Error**: $\text{RMSE} = \sqrt{\frac{1}{m}\sum_{t=1}^{m} (x_{n+t} - \hat{x}_{n+t})^2}$

Further Reading

- D. Koutsandreas et al., On the selection of forecasting accuracy measures, *Journal of the Operational Research Society*, 73(5):937–954, 2022

Box-Cox Transformation: Stabilizing Variance and Normalizing Data

Definition (Box-Cox Transformation)

The **Box-Cox transformation** is a family of **power** transformations parameterized by the **parameter** λ to **stabilize variance** and **normalize** data distribution, and defined as:

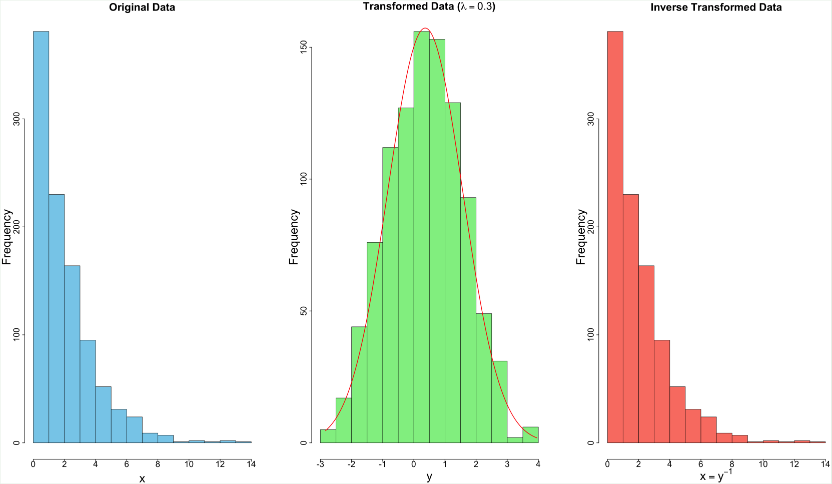
$$y := \begin{cases} \dfrac{x^{\lambda} - 1}{\lambda} & \text{if } \lambda \neq 0 \\ \log(x) & \text{if } \lambda = 0 \end{cases}$$

where x is the **original** data, and y is the **transformed** data.

- The parameter λ is **estimated** by finding the value that **maximizes** the **log-likelihood** for the transformed data, assuming a **normal** distribution.
- The Box-Cox transformation is **reversible**, ensuring the **original** data can be accurately **restored** from the transformed data without information loss.

Example

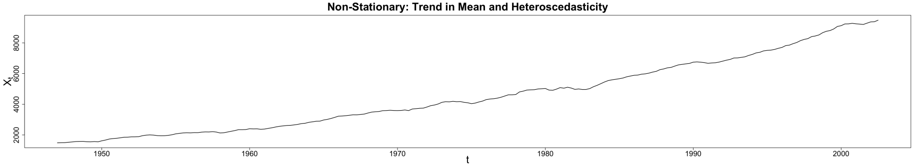
Example (Box-Cox Transformation: From Exponential to Normal Distribution)



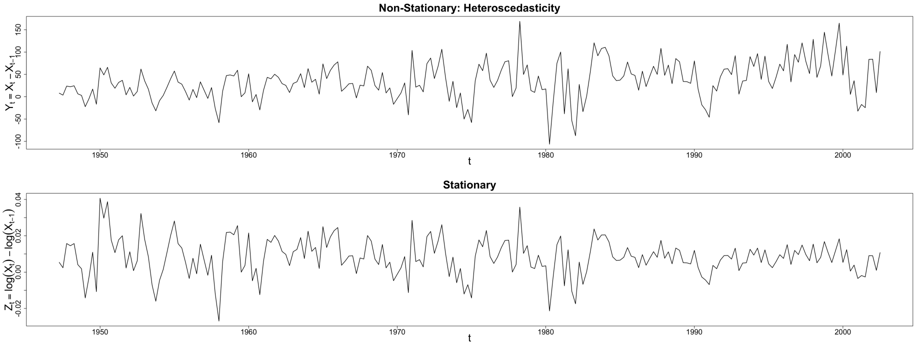
Box-Cox Transformation for Heteroscedastic Time Series Data

- In the context of **time series** analysis, the **Box-Cox** transformation reduces **heteroscedasticity** by adjusting the variance across the dataset.
- If $X_1, X_2, \ldots, X_n$ is a **heteroscedastic** time series, it can be **transformed** into a (hopefully) **homoscedastic** time series $Y_1, Y_2, \ldots, Y_n$ by choosing an **appropriate** value of λ
- Then, **fit** an appropriate `ARIMA` model to the **transformed** time series Y_t to make **predictions** $\widehat{Y}_{n+1}, \widehat{Y}_{n+2}, \ldots$
- The **inverse** transformation is then applied to these predictions to derive **forecasts** $\widehat{X}_{n+1}, \widehat{X}_{n+2}, \ldots$ on the **original** time series.

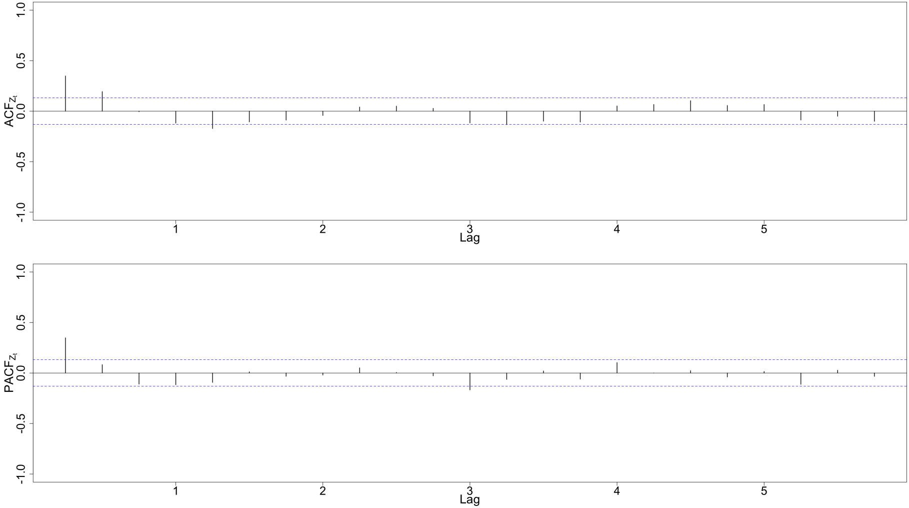
Case Study: Quarterly U.S. GNP from 1947 to 2001



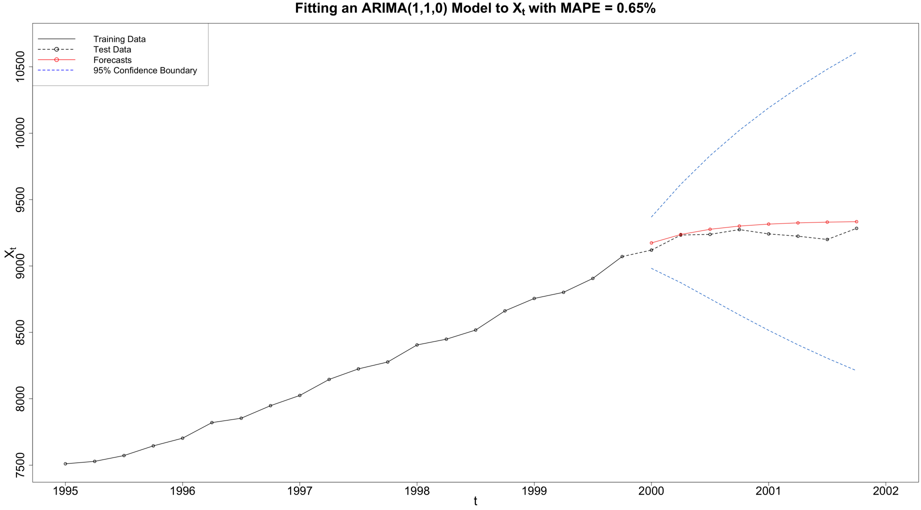
Case Study: Quarterly U.S. GNP... Continued



Case Study: Quarterly U.S. GNP. . . Continued



Case Study: Quarterly U.S. GNP... Continued



Further Reading

Review Study

- R.M. Sakia, The Box-Cox transformation technique: A review, *Journal of the Royal Statistical Society Series D: The Statistician*, 41(2):169-178, 1992

More Recent Review Study and Extensions

- A.C. Atkinson, M. Riani, and A. Corbellini, The Box–Cox transformation: Review and extensions, *Statistical Science*, 36(2)239-255, 2021

Challenges with Big Data Problems

- T. Zhang and B. Yang, Box–Cox transformation in big data, *Technometrics*, 59(2):189-201, 2017

Seasonal ARIMA Model

Definition (SARIMA Model)

A **time series** $\{X_t : t = 0, \pm 1, \pm 2, \ldots\}$ can be modeled by a **seasonal** ARIMA model, denoted by

$$\texttt{SARIMA}(p, d, q)(P, D, Q)_{\texttt{S}}$$

where the **non-seasonal** orders p, d, and q are as previously defined, and the **seasonal** orders include:

1. The **length** of each season, denoted by S
2. The seasonal AR order P, the seasonal MA order Q, and the seasonal **differencing** order D

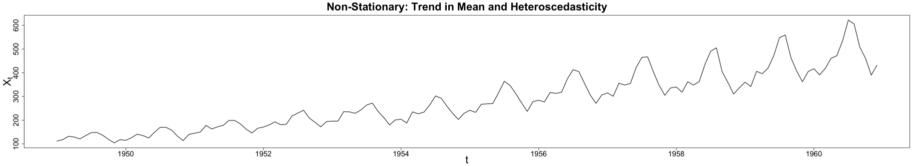
Estimating the Seasonal and Non-Seasonal Orders of SARIMA Models

- Estimating the **seasonal** orders P and Q

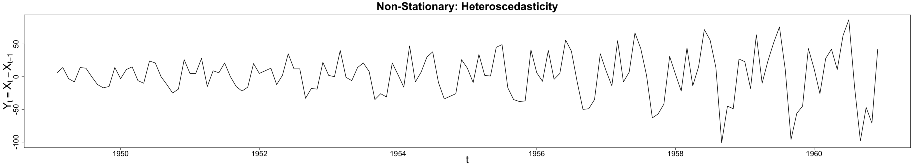
Model	ACF	PACF
Seasonal AR(P)	**Tails off** at lags $1S, 2S, \ldots$	**Cuts off** after lag PS
Seasonal MA(Q)	**Cuts off** after lag QS	**Tails off** at lags $1S, 2S, \ldots$
Seasonal ARMA(P, Q)	**Tails off** at lags $1S, 2S, \ldots$	**Tails off** at lags $1S, 2S, \ldots$

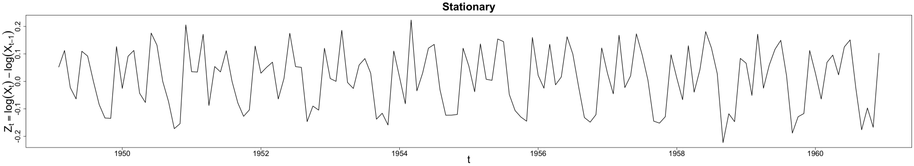
- After estimating the **seasonal** orders P and Q, the **non-seasonal** orders p and q are estimated **within** the lag interval between 1 and $1S$ as before

Case Study: Monthly Totals of International Airline Passengers

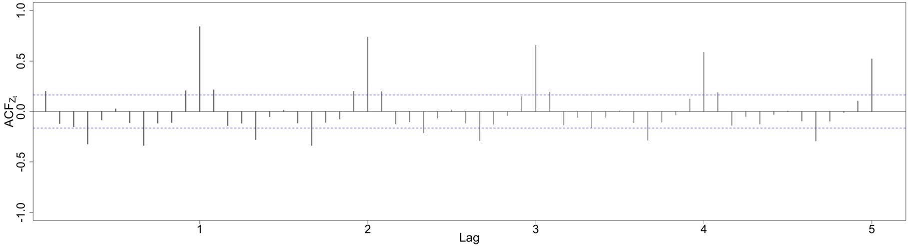


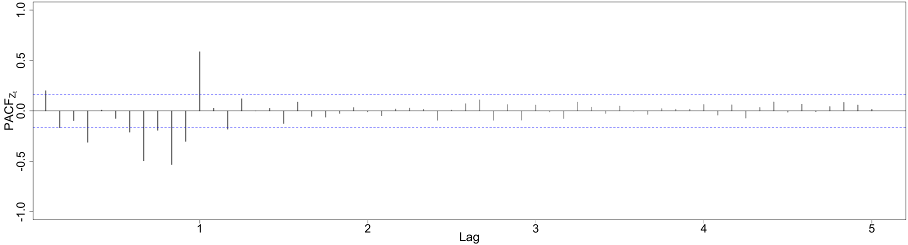
Case Study: International Airline Passengers. . . Continued



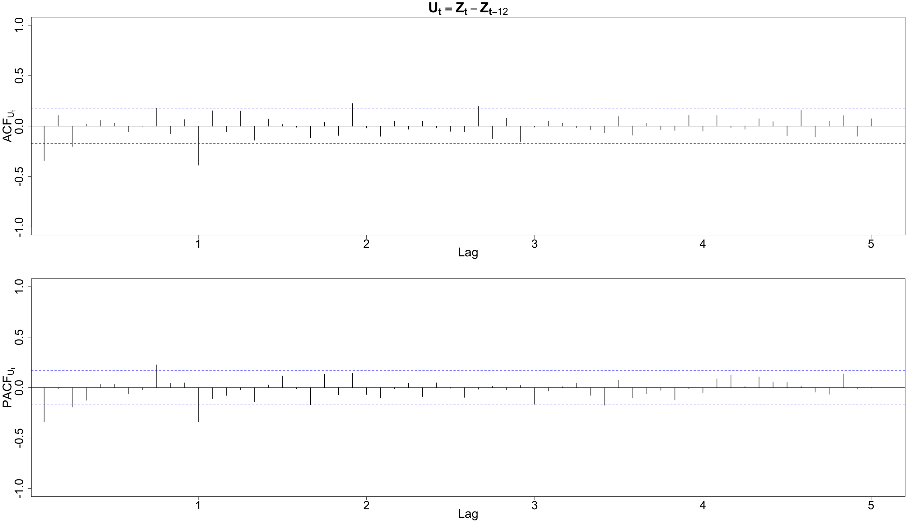


Case Study: International Airline Passengers... Continued

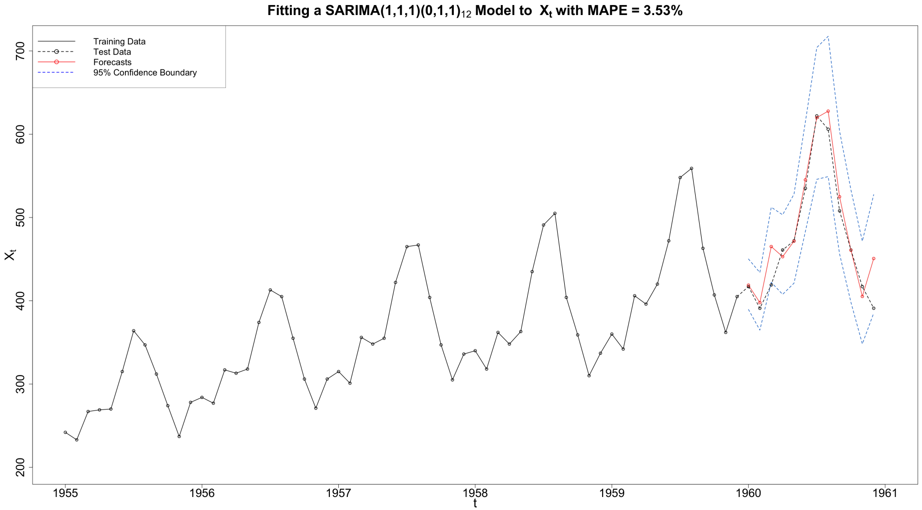




Case Study: International Airline Passengers. . . Continue



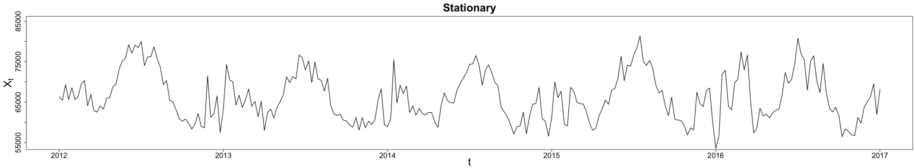
Case Study: International Airline Passengers. . . Continue



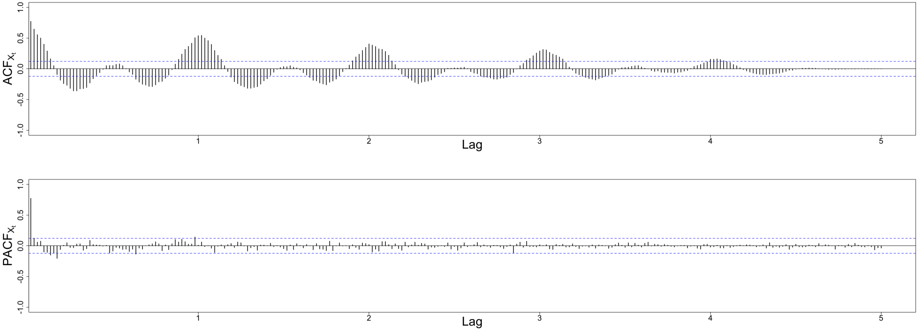
Hybrid SARIMA-Regression Models

- A **hybrid** model combines a SARIMA model with **linear regression** to leverage the strengths of both models.
- While the SARIMA model can model the **autocorrelation** and **seasonality** inherent in time series data, the **regression** model captures **exogenous** variables or trends not accounted for by time series analysis alone.
- The **hybrid** approach can **enhance** forecasting **accuracy** by capturing **both** seasonal autoregressive behaviors and linear relationships.

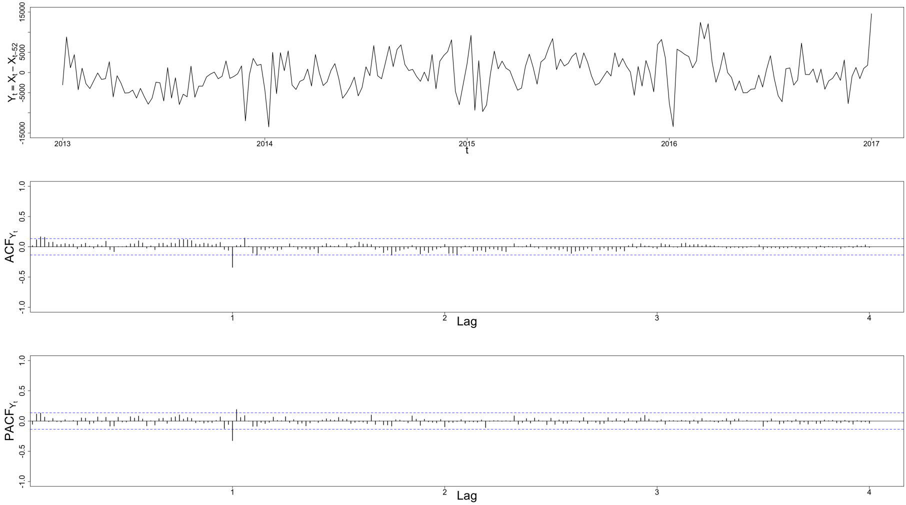
Case Study: Weekly Peak Australian Electricity Demand



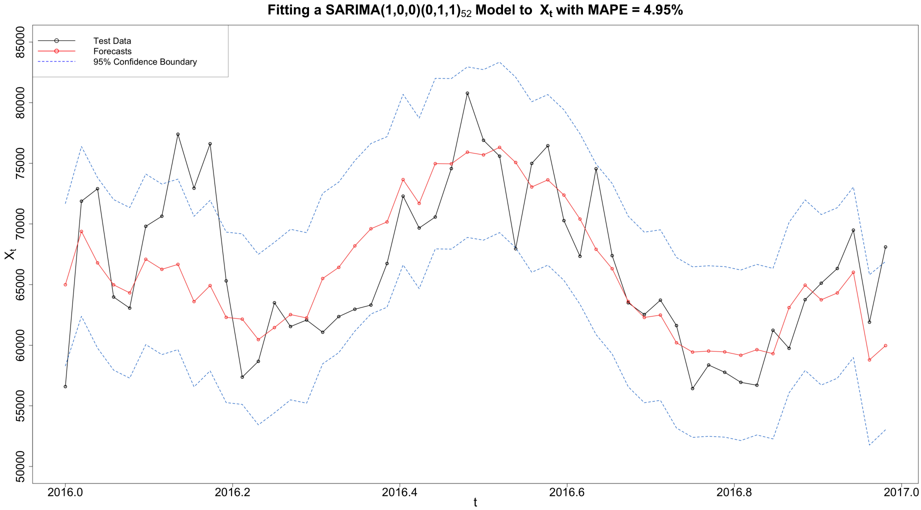
Case Study: Peak Australian Electricity Demand. . . Continue



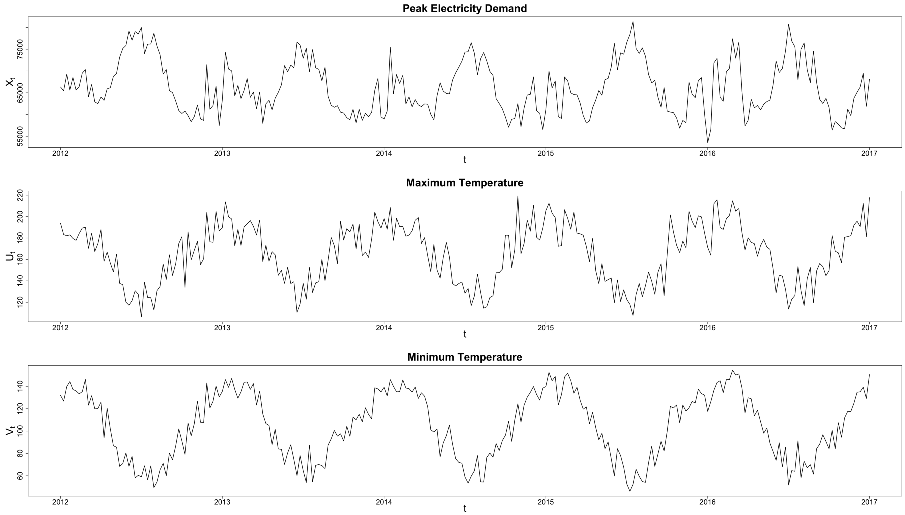
Case Study: Peak Australian Electricity Demand. . . Continue



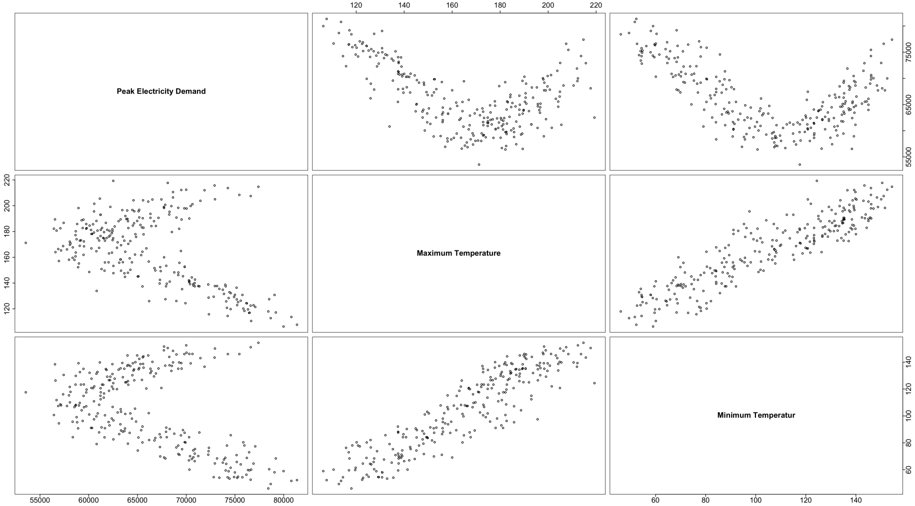
Case Study: Peak Australian Electricity Demand. . . Continue



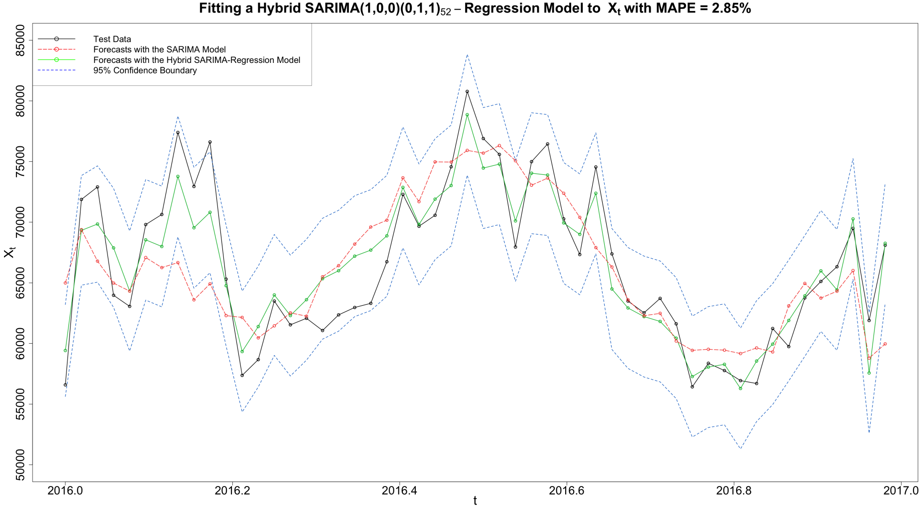
Case Study: Peak Australian Electricity Demand. . . Continue



Case Study: Peak Australian Electricity Demand. . . Continue



Case Study: Peak Australian Electricity Demand... Continue



Further Reading

Electricity Demand Forecasting

- A. Eshragh, B. Ganim, T. Perkins, and K. Bandara, The importance of environmental factors in forecasting Australian power demand, *Environmental Modeling & Assessment*, 27(1):1-11, 2022

Supply Chain Demand Forecasting

- M. Abolghasemi, J. Hurley, A. Eshragh, and B. Fahimnia, Demand forecasting in the presence of systematic events: Cases in capturing sales promotions, *International Journal of Production Economics*, 230:107892, 2020

Judgmental vs. Statistical Forecasting

- M.J. Lawrence, R.H. Edmundson, M. J. O'Connor, The accuracy of combining judgemental and statistical forecasts, *Management Science*, 32(12):1521-1532, 1986

References

1. R.H. Shumway and D.S. Stoffer, *Time Series Analysis and Its Applications With R Examples*, Springer, New York, 2017.