Appendix

Appendix E: Forecasting Model Overview

1 ARIMA: AutoRegressive Integrated Moving Average

The ARIMA model is a widely used approach in time series forecasting. It is particularly effective for data that exhibit autocorrelation and non-stationarity. The acronym stands for AutoRegressive Integrated Moving Average, and the model is defined by three components:

- **AR (AutoRegressive)**: The relationship between an observation and a number of lagged observations.
- **I (Integrated)**: The differencing of raw observations to make the time series stationary.
- **MA (Moving Average)**: The dependency between an observation and a residual error from a moving average model applied to lagged observations.

Mathematically, the ARIMA(p, d, q) model is expressed as:

$$Y_t = c + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t$$

Note: this equation is applied after the original time series has been differenced d times to ensure stationarity.

- Y_t is the observed value at time t (after differencing),
- c is a constant term,
- ϕ_i are the autoregressive coefficients,
- θ_j are the moving average coefficients,
- ε_t is white noise,
- d is the number of times the original series has been differenced to remove trends and achieve stationarity.

In this study, the `auto.arima()` function from the `forecast` package in R was used to automatically select the optimal model order by minimizing the Akaike Information Criterion (AIC).

2 Prophet: Decomposable Time Series Model

Prophet is a decomposable time series forecasting model developed by Facebook, designed to handle time series with strong trends and seasonality, even when data is missing or contains outliers. The model represents the time series as the sum of several components:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t$$

where:

- $g(t)$ represents the trend component (linear or logistic growth),
- $s(t)$ captures periodic seasonal effects using Fourier series,
- $h(t)$ represents the effects of holidays (if specified),
- ε_t is the error term.

For this study, the dataset consists of annual observations. As a result, the seasonality term $s(t)$ was disabled. The holiday component $h(t)$ was excluded as well, since holiday effects require higher-frequency data and cannot be meaningfully captured at the annual level. Despite the absence of these components, however, Prophet remains practically useful as a flexible, trend-focused forecasting model, well suited for capturing long-term patterns. Therefore, the use of the Prophet model remains appropriate for this study. The model was implemented using the `prophet` package in R, with forecasts generated through automatic trend fitting and uncertainty estimation. Prophet is widely used in applied forecasting for its flexibility, interpretability, and robust performance on real-world data.