

Forecasting: Quarterly International Arrivals from the UK

Tinozivashe N. Sibanda (10501541)

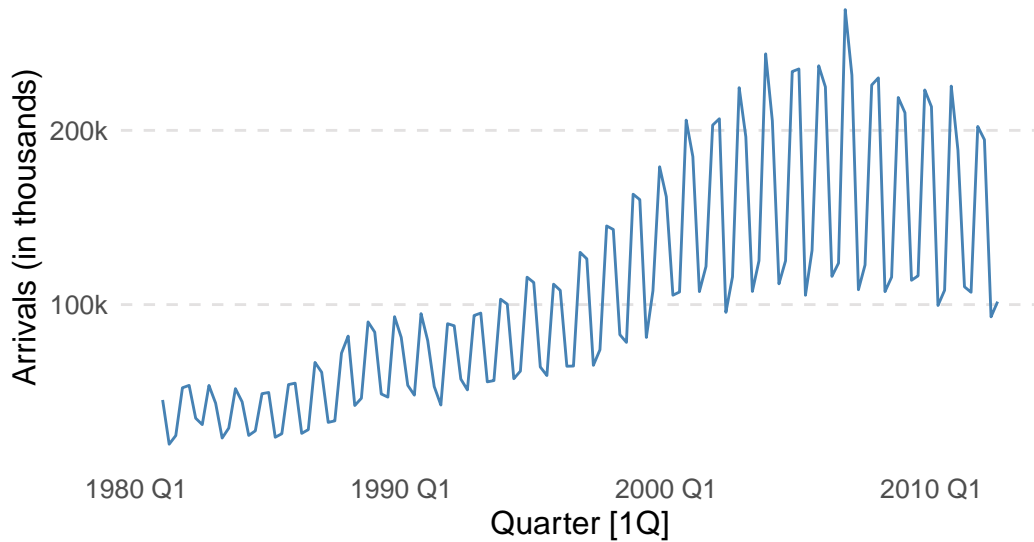
Table of contents

Introduction	2
Model Evaluation	3
Train-Test Split Approach	3
Time Series Cross-Validation Approach	5
Forecasts: Best Model	5
Discussion	6

Introduction

Figure 1 shows that international arrivals to Australia from the United Kingdom followed an upward trend from 1981 until around 2005, after which the series displays a gradual decline. It also indicates a clear seasonal pattern, with peaks occurring in Q4 (likely due to end-of-year travel and holidays) and troughs in Q2 and Q3. The variation of the seasonal pattern increases with time.

UK Arrivals to Australia



Source: Tourism Research Australia.

Figure 1: Quarterly international arrivals to Australia from the UK (1981 Q1 – 2012 Q3).

In this analysis, we evaluate two approaches for assessing forecasting model performance: **traditional train-test split** and **time series cross-validation (TSCV)**.

- The train-test split withholds the last two years of data from the model-fitting process. Specifically, the model is trained on data from *1981 Q1* to *2010 Q3*, and forecasts are generated on the holdout period from *2010 Q4* to *2012 Q3*. This enables straightforward evaluation by comparing the forecasts to actual observations over a fixed future window.
- Time series cross-validation (TSCV), on the other hand, provides a realistic and rigorous evaluation. Rather than relying on a single holdout period, TSCV repeatedly fits the model on expanding training windows and computes forecast accuracy over a rolling set of future observations. This better reflects operational forecasting workflows, where models are continually updated as new data arrives and are expected to perform consistently over time.

Using both methods allows us to evaluate how well each model fits historical data and how reliably it generalizes across different time periods.

Model Evaluation

We trained four forecasting methods, each incorporating a different strategy for capturing trend and seasonality in the data, namely:

- **Seasonal Naïve method (SNaive)**: A simple benchmark model that forecasts each quarter to be equal to the same quarter in the previous year. It captures seasonality but assumes no trend.
- **Holt-Winters Multiplicative method (HWM)**: Captures both a changing trend and seasonal patterns whose magnitude grows with the level of the series. This model is well-suited to the observed data and serves as our baseline exponential smoothing method.
- **Additive ETS method on Log-Transformed Data (ETS_LOG)**: Applies an exponential smoothing model with additive errors, damped trend, and additive seasonality to log-transformed arrivals. This approach stabilizes variance and can improve forecast reliability for series with exponential growth or volatility.
- **STL Decomposition with ETS (STL_ETS)**: Decomposes the log-transformed series using STL (Seasonal-Trend decomposition using Loess), then fits an ETS model to the seasonally adjusted component. This method provides flexibility in modeling complex seasonality and trend separately.

Train-Test Split Approach

Table 1 presents the forecasting accuracy of the four candidate models using RMSE and MAPE for both the training and test sets. We see that additive exponential smoothing method on log-transformed data (ETS_LOG) offered the most balanced performance, with low RMSE and MAPE on both training and test sets, making it the most reliable overall.

Model	RMSE		MAPE	
	Training	Test	Training	Test
SNaive	11639.74	14265.93	9.10	7.32
ETS_LOG	8103.41	8786.53	6.96	5.68
STL_ETS	7497.01	18979.61	5.99	12.19
HWM	8227.60	10955.29	6.97	7.26

Table 1: Forecast accuracy (RMSE and MAPE) of each model on training and test sets

Figure 2 shows that the innovation residuals are uncorrelated and have zero-mean, suggesting the forecasts generated by the additive exponential smoothing method on log-transformed data (ETS_LOG) are good.

Residual Diagnostics: ETS_LOG

Additive exponential smoothing method on log-transformed data

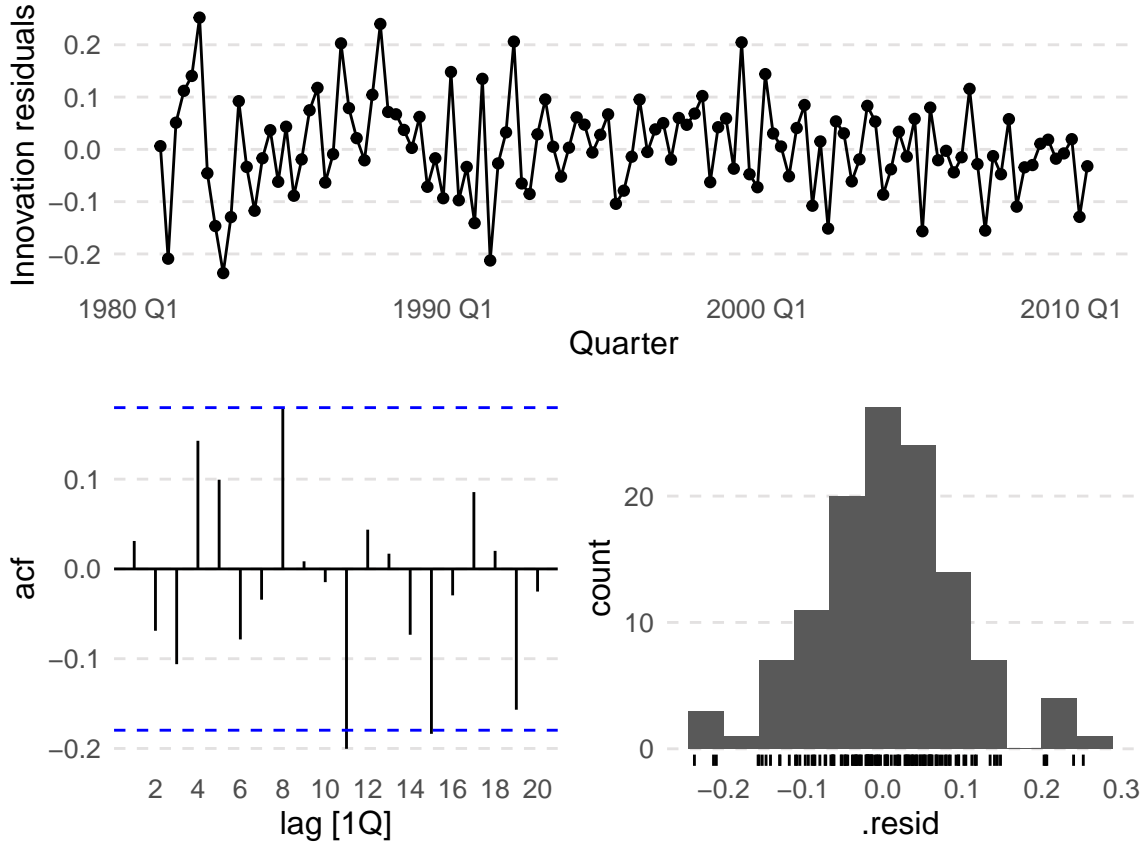


Figure 2: Residual diagnostics for the best fitted model to UK arrivals training data.

To make sure that the residuals are not distinguishable from a white noise series, we applied the Ljung-Box test. Table 2 shows the Ljung-Box test results ($\text{lag} = 10$) applied to the residuals of each model. The null hypothesis is that residuals are uncorrelated (i.e., resemble white noise). Specifically, for ETS_LOG, the p-value is 0.355, indicating no significant autocorrelation in the residuals — a good sign.

Model	LB_Statistic	P_Value
ETS_LOG	11.032	0.355
HWM	12.849	0.232
SNaive	51.949	0.000
STL_ETS	9.917	0.448

Table 2: Ljung–Box test (lag 10) on residuals from each forecasting model for UK arrivals.

We conclude that `ETS_LOG` is not only the most accurate model but also passes residual diagnostics — validating it as the best-performing model.

Time Series Cross-Validation Approach

We now apply time series cross-validation (TSCV) to compare the forecasting accuracy of the same four models evaluated earlier: `SNaive`, `ETS_LOG`, `STL_ETS`, and `HWM`. Table 3 reports the RMSE and MAPE for each forecasting model based on 8-step ahead rolling forecasts using time series cross-validation.

Warning: The future dataset is incomplete, incomplete out-of-sample data will be treated as missing. 7 observations are missing between 2012 Q4 and 2014 Q2

Model	RMSE	MAPE
ETS_LOG	13577.05	7.00
HWM	14089.91	7.42
SNaive	15681.63	6.85
STL_ETS	21495.83	11.14

Table 3: Forecast accuracy (RMSE and MAPE) of each model using time series cross-validation (8-step ahead rolling forecasts).

Again, we see that the `ETS_LOG` method demonstrates consistently strong performance, achieving the lowest RMSE and maintaining the best MAPE among the candidate models.

Forecasts: Best Model

The additive exponential smoothing method on log-transformed data (`ETS_LOG`) is suitable when the seasonal variation increases proportionally with the level of the series, as observed

in Figure 1 - this behavior—known as multiplicative seasonality — characterized by widening seasonal peaks and troughs over time.

Figure 3 shows that forecasts generated by the ETS_LOG method produce reasonable forecasts that are close to the actual arrival values in the testing set.

Forecasts: International arrivals to Australia from the UK

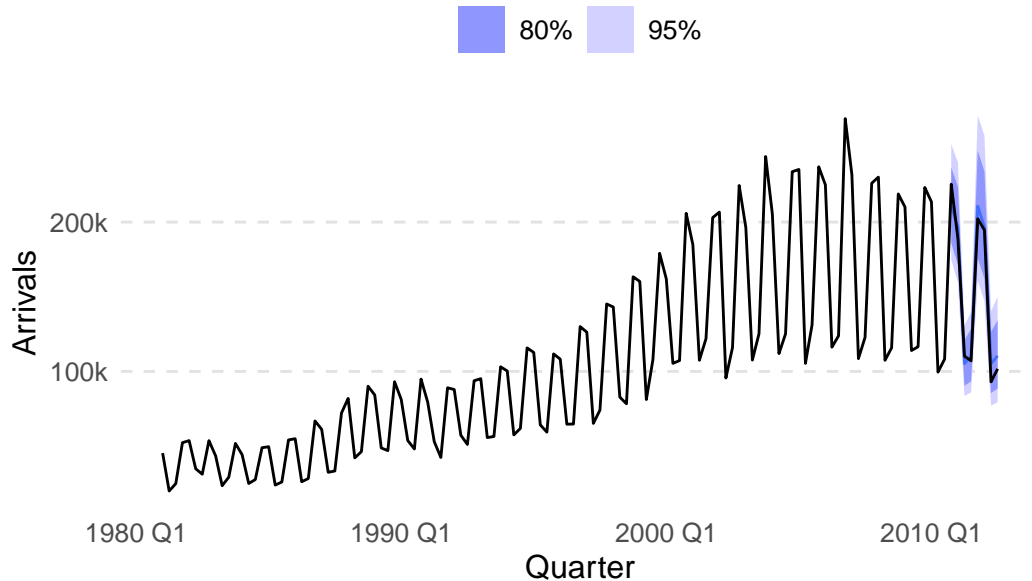


Figure 3: 2-Year forecast of international arrivals to Australia from the UK using best forecasting method.

Discussion

ETS_LOG provides the best RMSE and MAPE performance across both the train-test split and time series cross-validation approach. Its use of log transformation and damped trend appears to generalize well across time. These results reinforce ETS_LOG as the most effective model for forecasting UK arrivals.