0.0.1 Holt-Winters

The Holt-Winters method (Triple Exponential Smoothing for trend and seasonality) is an advanced form of exponential smoothing designed to model time series data exhibiting both trend and seasonal patterns. It is a univariate model, relying solely on the historical values of a single variable. To gain a deeper understanding of the Holt–Winters model, it is useful to first consider its simpler precursors—Simple Exponential Smoothing and Double Exponential Smoothing. While these methods are not discussed in detail in this project, they are thoroughly documented and readily available in the literature.

H.W. is a widely applied technique for forecasting business data that exhibits seasonality, evolving trends, and seasonal relationships [Ostertagova and Ostertag, 2012]. If you have a time series with a trend (upward or downward) and seasonal patterns, you can use Holt-Winters exponential smoothing to create short-term forecasts. This method estimates three parts at each time point: the current level (α), the trend slope (β), and the seasonal component (γ). The parameters α, β, and γ control how much weight is given to recent data when updating these estimates. All three range between 0 and 1 — when they are near 0, the model puts little emphasis on recent observations, relying more on past averages for forecasting [Coghlan, 2018].
Seasonality in time series data is generally categorized into two distinct types based on the nature of the seasonal pattern: additive and multiplicative. Additive seasonality implies that the seasonal effect remains constant in magnitude over time, regardless of the level of the time series (e.g., a consistent increase of 1 million dollars every December). In contrast, multiplicative seasonality indicates that the seasonal effect is proportional to the level of the series (e.g., a 40% increase in December), and thus varies in size depending on the series level. By plotting the time series, one can often distinguish between the two types of seasonality. If the seasonal fluctuations appear constant over time, the pattern is likely additive. If the fluctuations grow or shrink with the overall level of the series, the pattern is likely multiplicative [Kalekar, 2004].

Additive Seasonal Model
As previously discussed, additive seasonality involves adjusting forecasts by adding a constant amount, irrespective of the overall level of the time series.

$$\hat{x}_{n+h} = \text{level}_n + h \cdot \text{trend}_n + \text{seasonal}_{n+h-m}$$

The term level_n represents the estimated level at time n, h is the number of time steps ahead we want to forecast, and trend_n is the estimated slope (or change per period) at time n. The component seasonal_{n+h-m} provides the seasonal adjustment by aligning with the repeating seasonal pattern. Here, the parameter m refers to the length of the seasonal period, meaning the number of time steps in one complete seasonal cycle (for example, 12 for monthly data with yearly seasonality). The expression $n + h - m$ is used to correctly adjust the time index so that, when forecasting h steps ahead, the model applies the appropriate seasonal value from the past cycle [Thistleton and Sadigov, 2023]. We now examine the procedure for updating the estimates of the model parameters.

Smooth the Level (additive Model):
To compute the new level, we take a weighted average between the seasonally adjusted

observation (by removing the seasonal effect) and the non-seasonal forecast from the previous level and trend [Thistleton and Sadigov, 2023].

$$\text{level}_n = \alpha(x_n - \text{seasonal}_{n-m}) + (1-\alpha)(\text{level}_{n-1} + \text{trend}_{n-1})$$

- $\alpha(x_n - \text{seasonal}_{n-m})$: this part represents the seasonally adjusted level, where the observed value x_n is adjusted by removing the seasonal component from the same season in the previous cycle.

- $(\text{level}_{n-1} + \text{trend}_{n-1})$: this part represents the non-seasonal forecast, combining the previous level and trend to project forward without seasonal adjustment [Thistleton and Sadigov, 2023].

Smooth the Trend (additive Model):

$$\text{trend}_n = \beta \cdot (\text{level}_n - \text{level}_{n-1}) + (1-\beta) \cdot \text{trend}_{n-1}$$

Here, β is the smoothing parameter for the trend component. It determines how much weight is given to the most recent change in the level (i.e., $\text{level}_n - \text{level}_{n-1}$) compared to the previous trend estimate trend_{n-1}. A β value close to 1 gives more importance to recent changes, making the trend highly responsive, whereas a value close to 0 places more weight on the past trend, making the model less sensitive to recent fluctuations [Thistleton and Sadigov, 2023].

Smooth the Season (additive Model):

$$\text{seasonal}_n = \gamma \cdot (x_n - \text{level}_n) + (1-\gamma) \cdot \text{seasonal}_{n-m}$$

Here, γ is the smoothing parameter for the seasonal component. It controls how much weight is given to the most recent seasonally adjusted observation ($x_n - \text{level}_n$) compared to the previous seasonal estimate seasonal_{n-m}. A γ value close to 1 gives more importance to recent seasonal changes, making the seasonal component highly responsive, while a value close to 0 places more weight on past seasonal patterns, making it less sensitive to new fluctuations [Thistleton and Sadigov, 2023].
In the additive seasonal model, the seasonal factors are defined as deviations from the average level. To avoid introducing long-term bias into the forecast, it is necessary to normalize these seasonal components such that their total over a full seasonal cycle equals zero. This condition is expressed as:

$$\sum_{i=n-m+1}^{n} \text{seasonal}_i = 0$$

where m denotes the length of the seasonal cycle (e.g., $m = 12$ for monthly data with yearly seasonality). This ensures that the additive seasonal components, which represent either positive or negative deviations from the underlying level, balance out across the cycle and do not systematically shift the forecasted values upward or downward [Kalekar, 2004]. We now turn to the multiplicative seasonal model, in which the seasonal component varies proportionally with the series level.

Multiplicative Seasonal Model

Multiplicative seasonality means you adjust forecasts by multiplying by a constant factor, so the adjustment depends on the overall level [Thistleton and Sadigov, 2023].

$$\hat{x}_{n+h} = (\text{level}_n + h \cdot \text{trend}_n) \cdot \text{seasonal}_{n+h-m}$$

The structure of the multiplicative seasonal model follows the same component definitions as the additive case described previously; however, the seasonal adjustment is applied multiplicatively rather than additively. We next examine the procedure for updating the parameter estimates in the multiplicative seasonal model, beginning - as before - with the smoothing of the level, trend, and seasonal components.

Smooth the Level (multiplicative model):

$$\text{level}_n = \alpha \left(\frac{x_n}{\text{seasonal}_{n-m}} \right) + (1 - \alpha)(\text{level}_{n-1} + \text{trend}_{n-1})$$

- α: smoothing constant, with $0 < \alpha < 1$.
- $\frac{x_n}{\text{seasonal}_{n-m}}$: deseasonalized observation, where the seasonal component from the same season in the previous cycle is removed multiplicatively.
- $\text{level}_{n-1} + \text{trend}_{n-1}$: non-seasonal forecast, combining the previous level and trend to project forward without seasonal adjustment [Kalekar, 2004].

Smooth the Trend (multiplicative model):

The trend component in the multiplicative seasonal model is updated using the same structure as in the additive model, computing a weighted average between the most recent change in the level and the prior trend. This approach allows the model to adjust its sensitivity to recent fluctuations based on the value of β [Kalekar, 2004].

Smooth the Season (multiplicative model):

$$\text{seasonal}_n = \gamma \cdot \left(\frac{x_n}{\text{level}_n} \right) + (1 - \gamma) \cdot \text{seasonal}_{n-m}$$

- γ: smoothing parameter for the seasonal component, where $0 < \gamma < 1$.
- $\frac{x_n}{\text{level}_n}$: most recent observed seasonal factor, obtained by deseasonalizing the observation at time n.
- seasonal_{n-m}: previous estimate of the seasonal component from the same period in the last seasonal cycle.
- m: length of the seasonal cycle (e.g., $m = 12$ for monthly data with yearly seasonality) [Kalekar, 2004].

Since seasonal indices in multiplicative models represent proportions, it is recommended to normalize them after estimation to ensure that their average over a full seasonal cycle equals 1. This can be expressed as:

$$\sum_{i=n-m+1}^{n} \text{seasonal}_i = m \quad \text{[Kalekar, 2004]}$$

In the implementation phase, the model will be fitted in `R` using the `HoltWinters()` function. In this context, the smoothing parameters α, β, and γ are automatically estimated by the function, unless manually specified. These parameters are optimized to minimize the sum of squared errors on the training data, ensuring the best possible fit to the observed time series.

Ljung-Box test

After fitting the model to the in-sample data, it is important, as discussed in the context of the SARIMA model, to evaluate whether the residuals exhibit the properties of white noise. For this purpose, we apply the Ljung–Box test to assess whether the residuals exhibit the properties of white noise, that is, the absence of significant autocorrelation. A non-significant result would indicate that the model has adequately captured the underlying data structure [Hassani and Yeganegi, 2019].

Practical Implementation of Holt-Winters model

The complete implementation of the code can be accessed at the following repository: GitHub Repository (Holt-Winters). As previously discussed, the model is specifically designed to forecast time series that exhibit level, trend, and seasonal components. Consequently, the model was applied directly to price data, unlike the other models considered in this work, which were estimated on returns. Another key difference is that, in this case, only the training and test sets were used, omitting the validation set. This is because, unlike in regression models, the validation set cannot be utilised to fine-tune the TES model parameters. Finally, as shown in the implementation, applying Holt–Winters exponential smoothing to the entire dataset from 2010 to 2025 produces distorted results. The Ljung–Box test for autocorrelation, applied to the residuals of the in-sample fit, yielded a test statistic of $X^2 = 128.43$ with $df = 20$ and a p-value of 2.2×10^{-16}, indicating significant autocorrelation and suggesting that the model failed to fully capture the underlying dependence structure in the data. Therefore, the length of the training set was reduced to shorter windows (e.g., 1 year, 2 years, 3 years) in order to improve the quality and relevance of the forecasts. This adjustment is motivated by the observation that the model's structure is better suited to short-term forecasting. Using a window of three years, corresponding to 642 observations in the training set (85 percent), yielded satisfactory results [1]. The Ljung–Box test on the residuals indicated that the errors were consistent with white noise ($X^2 = 27.486$, $df = 20$, p-value $= 0.1221$). Further residual diagnostics revealed some evidence of heteroskedasticity, as the variance appeared

[1] The out-of-sample (test) period remains unchanged and corresponds to the last 15% of the full dataset (01.01.2010–31.12.2024). Only the in-sample (training) window was reduced to 642 observations (approximately three years). As a result, the usual 85%–15% partition between training and test sets does not strictly apply in this case. This adjustment was made to ensure that all models are evaluated on the same out-of-sample period, thereby preserving comparability of the performance metrics.

inconsistent, and the histogram of forecast errors displayed heavier tails than the normal distribution, suggesting deviations from normality. Nevertheless, the ACF and PACF plots confirmed the absence of significant autocorrelation, with lags close to zero and no notable partial autocorrelations. Overall, the residual diagnostics suggest that the TES model provides an adequate representation of the data, thereby allowing us to proceed with the out-of-sample forecasting. Applying the forecast to the out-of-sample data yields relatively low absolute error measures (MAE, RMSE, MAPE). However, the negative R^2 and a Directional Accuracy below 50% highlight limited explanatory capacity and poor predictive performance, indicating that the TES model fails to adequately capture the underlying dynamics of the test dataset. To address this limitation, we employ a rolling-window approach to generate one-step-ahead forecasts. In this framework, the model is iteratively re-estimated as the window advances through the series, incorporating each newly observed value at every step. Under this setting, the R^2 deteriorates further, but the Directional Accuracy improves, suggesting that while explanatory power decreases, the model gains marginal usefulness in predicting the direction of movements. Results are presented in the chapter on Empirical Evaluation.

Bibliography

Avril Coghlan. *A Little Book of R for Time Series*. 2018. Available at: https://a-little-book-of-r-for-time-series.readthedocs.io/en/latest/.

Hossein Hassani and Mohammad Reza Yeganegi. Sum of squared acf and the ljung–box statistics. *Physica A: Statistical Mechanics and its Applications*, 520:81–86, 2019. doi: 10.1016/j.physa.2019.01.063. URL https://doi.org/10.1016/j.physa.2018.12.028.

Prajakta S. Kalekar. Time series forecasting using holt-winters exponential smoothing. Technical report, Kanwal Rekhi School of Information Technology, 2004. Under the guidance of Prof. Bernard.

Eva Ostertagova and Oskar Ostertag. Forecasting using simple exponential smoothing method. *Acta Electrotechnica et Informatica*, 12(3):62–66, 2012. doi: 10.2478/v10198-012-0034-2.

William Thistleton and Elvin Sadigov. Introduction to forecasting: Holt-winters for trend and seasonality, 2023. Practical Time Series Analysis, Week 5.