Forecasting Monthly Power Generation of Photovoltaic Systems in Germany: A Comparative Study of ETS and ARIMA Models

Predictive Analytics (CDSCO1005E) - Home assignment (UC)

Copenhagen Business School
MSc. Business Administration and Data Science

Author: Ries, Alexander (Student ID: 158292)

Date of submission: 02.08.2023

Number of Pages: 9

Number of Characters: 22,197

Forecasting Monthly Power Generation of Photovoltaic Systems in Germany: A Comparative Study of ETS and ARIMA Models

Alexander Ries
Student - Copenhagen Business School
MSc. Business Administration and Data Science

August 2, 2023

Abstract

Due to its weather-dependency, forecasting the power generated by photovoltaic systems is essential for effective grid management, optimizing energy consumption, and promoting sustainable practices. This paper explores two forecasting methods, Exponential Smoothing (ETS) and Auto-regressive Integrated Moving Average (ARIMA), to identify the most suitable model for predicting the monthly power generated by photovoltaic systems in Germany over an 18-month horizon. The results indicate that both methods produce similar and well-performing models. Among them, the ETS (A,A,M) model exhibits the lowest forecast errors. Leveraging this model, the power generation from July 2023 to December 2024 is expected to follow the same seasonal developments observed in previous years. However, there is also the possibility of an all-time peak in solar power generation in Germany during summer 2024.

Keywords: *Predictive Analytics, Forecasting Methods, ARIMA, ETS, Solar Power Generation*

1 Introduction

The current developments in the climate crisis and geopolitical conflicts in Europe have shown that Germany's energy supply must become more climate-friendly and less dependent from imports. Therefore, the German government aims to place the energy supply on a broader basis and increase the energy produced by renewable energies. Solar photovoltaic (PV) systems play a significant role in this strategy (Presse- und Informationsamt der Bundesregierung 2023). However, since the power production from PV systems is weather-dependent and thus also depend on the seasons, accurate forecasting methods are needed to handle the fluctuating energy supply during the year.

This research paper aims to find a suitable forecasting method to forecast the monthly power generated by PV systems in Germany over a horizon of 18 months. These forecasts are significant for utility companies, policymakers, and industries enabling better grid management, informed energy consumption, and sustainable practices.

In the following course of this work, two forecasting methods are tested to identify the most reliable approach for predicting the power produced by PV systems. This begins with an exploratory data analysis to handle the dataset effectively. Next, ETS and ARIMA models are created, with the most promising model of each method selected for forecasting. Section 4 evaluates the forecasts produced by the models on a test set by visual inspection and forecast errors. Finally, the best model is applied to forecast solar power production over the next 18 months, providing valuable insights in the future energy supply from solar power in Germany.

The code for this paper, written in R, is available on a public Git Repository[1], along with the utilized dataset.

2 Exploratory Analysis

The underlying dataset for this paper is from SMARD - the information platform of the German Federal Network Agency [2]. It has 102 rows and 15 columns. Each row corresponds to one month starting from January 2015 until June 2023. The columns contain the power produced in megawatt hours (MWh) from different energy sources such as wind energy, biomass or solar energy. For the purpose of this paper, only the column with the values for power produced by PV systems is kept and renamed into a shorter name. Furthermore, the initial unit MWh is transformed to gigawatt hours (GWh) due to better readability.

Figure 1 shows the original time series. The x-axis shows the time in month and the y-axis presents the power produced by PV systems in Germany. The least amount of power was generated in Decmeber 2017 (495 GWh). The most power was produced recently in June 2023 (8736 GWh). On average, 3580 GWh were generated per month.

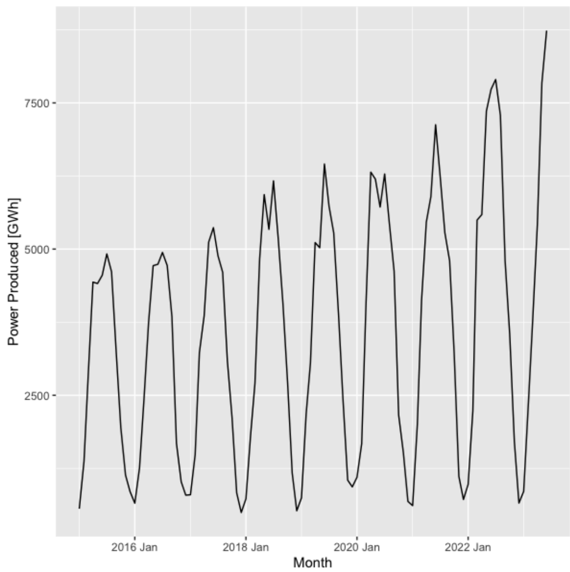


Figure 1: *Power produced by PV Systems in Germany - Original Time Series*

Looking only at Figure 1, a yearly seasonal pattern can be expected in the data where the power produced peaks in the summer months. This seems logical since the number of hours of sunshine in the summer are higher than in the winter. Furthermore, the variance of the data seem to increase over time. By focusing only at the local maxima, a trend could be identified but when looking at the local minima, this potential trend in the data is not confirmed.

In this section, an exploratory analysis of the data is conducted to gain insights into its characteristics by examining non-linearity, decomposing the time series into the components trend, season and remainders, and testing the data for unit roots and structural breaks.

2.1 Non-linearity

The yearly "up-and-down" pattern in the original time series (see Figure 1) and the increasing variance in the seasons is strong evidence against a

[1] https://github.com/Reese181/SolarPowerGenerationPrediction

[2] https://www.smard.de/home/downloadcenter/download-marktdaten/

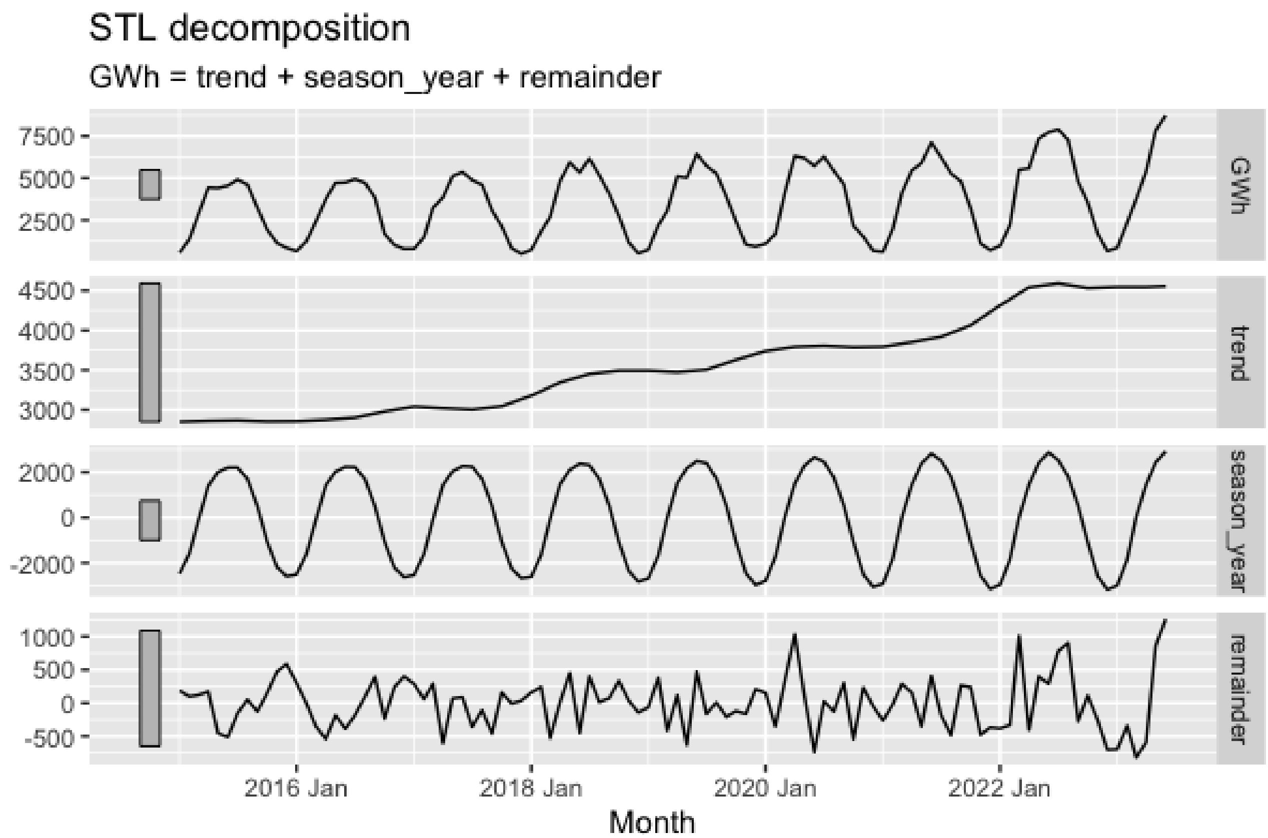


Figure 2: *STL Decomposition of Original Time Series*

linear relationship of the power produced by PV systems and the time. This gets even more clear by looking at Figure 7 in the Appendix, where the data is visualized as a scatter plot. Trying to fit a straight line to this points seems not to be possible. Furthermore the STL decomposition gives an indication that the original data is following a non-linear trend ruling out linearity too (see Figure 2).

2.2 Trend, Seasonality & Variance

To examine the existence of a trend or seasonality in the time series, the time series was decomposed by an STL model with default settings. Figure 2 shows the decomposed time series and confirms a slight increasing - but non-linear - trend in the data. The seasonality is the most significant component in the series with an increase in the variance over time. Both, the yearly seasonal pattern and the increasing variance are also confirmed in Figure 8 in the Appendix showing the course of the seasons over the same time period.

Similar to the seasonality, the variation of the remainders seem to increase over time. To address this variance we apply a logarithmic transformation to the data. The difference between the transformed time series to the original can be seen in Figure 9 in the Appendix. First of all, the scale of the time series has changed. In the original time series, we see absolute values of the power produced whereas a change in the new y-axis - the logarithmic scale - shows us the relative change to the original series. Moreover, we can see that the variation in the season peaks has been minimised. The STL decomposition of the transformed time series shows more clearly what has changed: The growing variance in the seasonality has decreased and the variance in the remainders is stabilised (see Figure 10 in the Appendix).

As an alternative for the logarithmic transformation, a box-cox transformation could have been applied. However, the calculated lambda, after utilizing the Guerrero feature in R, is really close to 0 (-0.06), which underlines the need of a transformation and confirms a logarithmic transformation as an appropriate measure. In any case, a transformation entails a loss of information and the time series has to be transformed back before forecasting.

2.3 Unit Roots

According to Hyndman and Athanasopoulos 2021, a time series is stationary, when its data points do not depend on the time at which they were observed. The assumption of a stationary time series is made by many statistical tests, such as for structural breaks (Quandt 1960), and by time series models like ARIMA. Furthermore, stationary time series might lead to more accurate forecasts. In this section, we will examine if the time series of power generated by PV systems is stationary. Time series with trends or seasonality are not stationary since these properties influence the value of the time series at different times. Therefore, methods like differencing or detrending are used making a time series stationary (Hyndman and Athanasopoulos 2021). The need for differencing can be determined by unit root tests such as augmented Dickey–Fuller test (ADF) and Kwiatkowski-Phillips-Schmidt-Shin test (KPSS). ADF has the null hypothesis that a unit root is present in the time series, thus the time series is non-stationary (Enders 2014), whereas KPSS makes the null hypothesis that the data is stationary (Kwiatkowski et al. 1992). This means a time series will be stationary if we can reject the null hypothesis of ADF and accept the null hypothesis of KPSS. All ADF and KPSS results for this paper can be seen in the Tables 5 and 6 in the Appendix.

Initially, the unit root tests have been applied to the logarithmic transformed time series. With ADF, the series was tested three times (with the test types "trend", "drift" and "none") to see if it is trend stationary, drift stationary or generally stationary. While the null hypothesis of trend and drift stationarity could have been rejected, the null hypothesis of non-stationary data without a constant and a trend cannot be rejected. The KPSS test was conducted with default settings and also with a specified number of lags because of the clearly seen seasonality in Section 2.2. The null hypothesis of the latter test could not be rejected confirming the results of the ADF results that the series is non-stationary and, moreover, indicated that the time series is not trend-stationary.

Therefore, a first difference of the logarithmic transformed time series was taken and the same testing procedure performed as before. The results of the ADF tests stated that the series is now stationary. However, the KPSS test type "tau" with the 12 lags used had a t-value (0.20) that is above the 5% critical value (0.18) so that we can reject the null hypothesis of stationary data. Furthermore, with a look at the plotted time series with a first lag difference (see Figure 11 in the Appendix), one can clearly see the still existing seasonality in the data indicating also that the data is not stationary.

Another possibility is seasonal differencing instead of taking a first lag difference. With lag-12 differences of the time series, the hypothesises of all ADF tests could be rejected and all hypothesis of both KPSS tests, default lags used and 12 lags used, could not be rejected. This indicates that the time series can be made stationary by seasonal differencing with 12 lags. Figure 11 in the Appendix illustrates the seasonal differenced time series displaying white noise characteristics which supports the test results that the time series is stationary. Since seasonal differencing introduced stationarity in the time series, differencing the series multiple times is not necessary.

2.4 Structural Breaks

A structural break is visually not identifiable when looking at Figure 1. Therefore, the

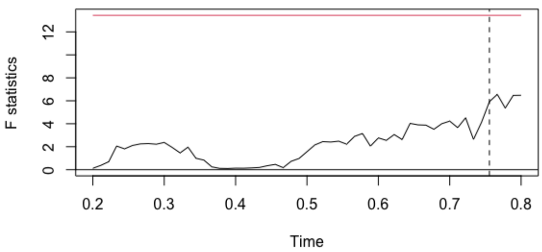


Figure 3: *QLR Test*

Quandt likelihood ratio (QLR) test, introduced by Quandt 1960, is applied to identify potential structural breaks that need to be handled before modeling. When applying this test, the p-value of the "supF" test statistic is higher than the significance level alpha (0.05) so that we do not reject the null hypothesis of no structural breaks (see also Figure 12 in the Appendix). Figure 3 visualises the FStatistics with its boundary. Since the graph does not exceed the red boundary, there is no evidence for a structural change (at level alpha = 0.05).

3 Methodology

This section applies two forecasting methods: Exponential Smoothing (ETS) and Autoregressive Integrated Moving Average (ARIMA). Both follow a different approach to identify patterns within data and make future predictions. The most promising model from each method is chosen by the Akaike Information Criterion (AIC), the Corrected Akaike Information Criterion (AICc), and the Bayesian Information Criterion (BIC).

Beforehand, a train-test split is necessary to test and evaluate the model performances on unseen data in the following Section 4. The train set starts from the beginning of the time series (January 2015) to December 2021. The time period from January 2022 to June 2023 is used as test set leading to split ratio of approximately 80:20 between train and test set.

3.1 ETS

ETS was introduced in the 50s and is a common forecasting technique to this day. It combines the components Error (E), Trend (T) and Seasonality (S) in a smoothing calculation. The components can be combined additively (A) or multiplicatively (M) or left out of the model completely (N) (Hyndman and Athanasopoulos 2021). This combination can be automatically derived from the ETS Function in R or one can manually instruct an ETS model how to combine. In this section, one potential combination will be estimated based on the exploratory analysis and compared to the ETS model automatically chosen.

In Section 2, we applied a logarithmic transformation to stabilize the variance of the remainders. Therefore, an additive error is used in our guessed ETS model. When looking at the decomposition of the logarithmic transformed time series (see Figure 10 in the Appendix), the trend shows no exponential behaviour so that an additive trend will be used. The seasonal component shows less variation in the magnitude after the logarithmic transformation. However, since there is still an increasing variation in the seasonal pattern the seasonal component will be combined multiplicatively in the guessed ETS model (Franco 2022) resulting in an estimated ETS(A,A,M) model. The automatic ETS model made the same choices in terms of an additive error and trend but has chosen an additive seasonal component instead (ETS(A,A,A)). The models are compared in Table 1 with their

	ETS(A,A,M)	ETS(A,A,A)
AIC	63.7	66.1
AICc	73.0	75.4
BIC	105.0	107.0

Table 1: *ETS Model Comparison*

information criteria AIC, AICc and BIC. The guessed ETS(A,A,M) has lower values in all three criteria indicating that it is expected to produce better forecasts than the automatic chosen ETS(A,A,A) (Hyndman and Athanasopoulos 2021). Therefore, we will proceed with the ETS(A,A,M) model in the following course of the paper. The smoothing parameters of the

α	β	γ
5.69×10^{-4}	2.46×10^{-4}	1.55×10^{-3}

Table 2: *Smoothing Parameters for ETS(A,A,M) Model*

ETS(A,A,M) model shown in Table 2 are all very close to zero. The small alpha value indicates that more weight is assigned to observations from the more distant past. A small beta means a slope that is not changing much over time indicating linear trend and a small gamma implies a roughly periodic time series (Hyndman and Athanasopoulos 2021).

The models suitability can also be evaluated by looking at its residuals. Therefore, the function gg_tsresiduals() is used which plots a visualized summary of the model's residuals (see also Figure 13 in the Appendix). In this summary, the ACF plot shows only one significant spike indicating little to no correlation between the residuals. The plotted residuals over time seem to occur randomly. However, there might be some pattern at the beginning of each year since the residuals at this point of time are higher than the others. The residuals follow also not a normal distribution. Moreover, we reject the null hypothesis of the Ljung-Box test meaning that there is still information left in the residuals which is not captured by the model (see also Figure 14 in the Appendix).

3.2 ARIMA

Another approach for time series forecasting involves using ARIMA models. While ETS models capture trend and seasonality, ARIMA models are designed to capture auto-correlation in the data. An ARIMA model is a combination of differencing, an auto-regressive model and moving-average model. ARIMA models can also model seasonal data. The structure of these models is described as ARIMA(p,d,q)(P,D,Q), where:

- p/P is the order of the autoregressive part
- d/D is the degree of differencing involved
- q/Q is the order of the moving average part
- uppercase letters represent seasonal parts
- lowercase letters represent non-seasonal parts

In this section, one ARIMA model is estimated manually and compared to a automatically cho-

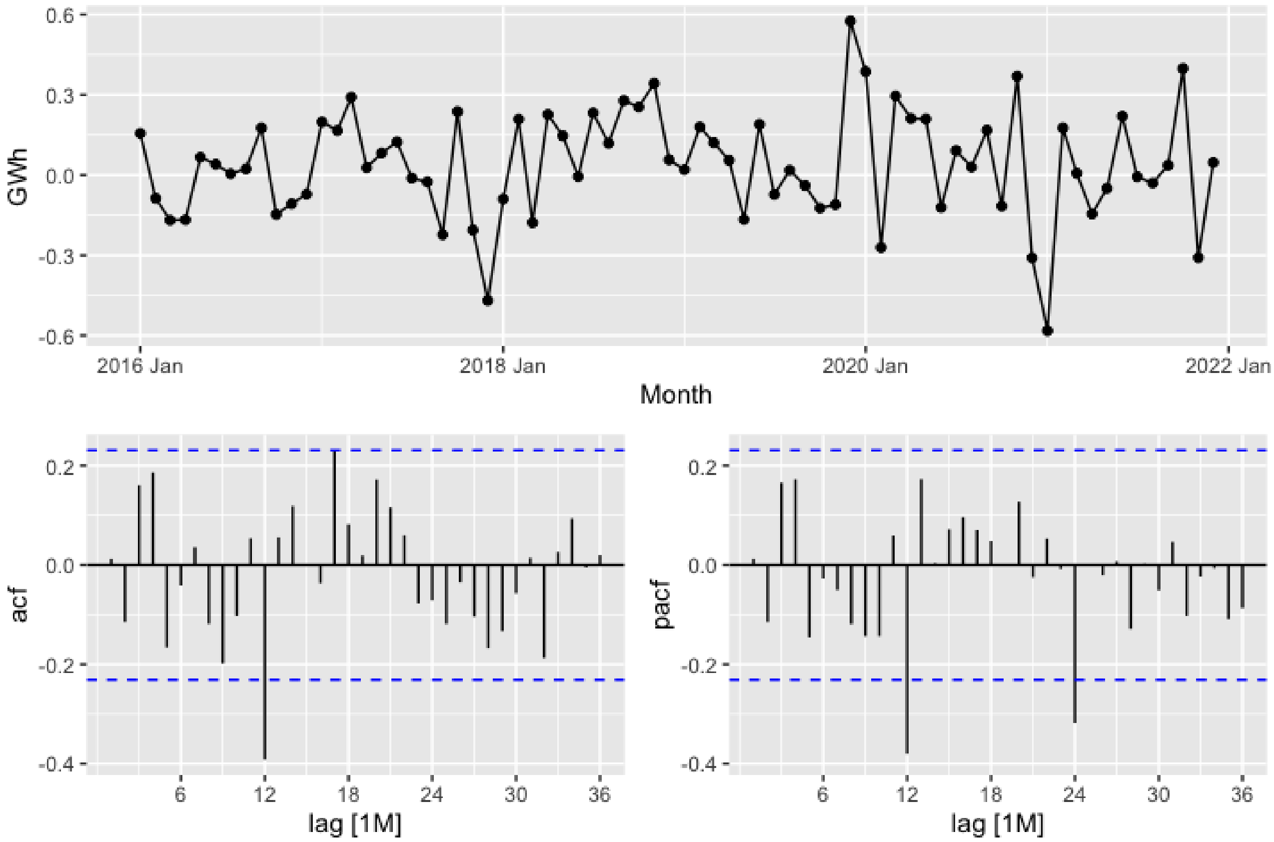


Figure 4: *ACF and PACF of Stationary Time Series*

sen ARIMA model. In order to determine the different parts of the ARIMA model, a stationary time series is needed. In Section 2.3, it was found that the time series can be made stationary by taking a seasonal difference which is expressed by D = 1. Looking at the stationary time series and its auto-correlation function (ACF) and partial auto-correlation function (PACF) in Figure 4, the further components of the ARIMA model can be determined. There is no clear pattern for a non-seasonal component in the model. However, the one significant spike at lag 12 in the ACF plot suggests a seasonal MA(1). Therefore, the guessed ARIMA model is an ARIMA(0,0,0)(0,1,1). On the contrary, the second ARIMA model which is automatically calculated has the structure ARIMA(0,0,0)(2,1,0). Both models state that there is no non-seasonal component. However, the guessed ARIMA shows a seasonal moving average model whereas the auto ARIMA shows an auto-regressive model (Hyndman and Athanasopoulos 2021). The two models are compared in Table 3 with the same information criteria used in Section 3.1. The auto-regressive model has in all three criteria lower values than the moving average model.

	ARIMA(0,0,0)(0,1,1)	ARIMA(0,0,0)(2,1,0)
AIC	-29.3	-42.1
AICc	-29.1	-41.5
BIC	-24.7	-32.9

Table 3: *ARIMA Model Comparison*

Since these results indicate that the ARIMA(0,0,0)(2,1,0) will produce better forecasts, the paper will continue with this ARIMA model. The residuals of the models, shown in Figure 15 in the Appendix, do not follow a clean normal distribution. However, the ACF plot exhibits only one significant spike and the residuals occur rather randomly as following a pattern and do not change their variance over time. Furthermore, the model passes the Ljung-Box test with 24 lags indicating that the model captures the relevant information of the data well (see also Figure 16 in the Appendix).

4 Results

After choosing the most promising ETS and ARIMA models in the Sections 3.1 and 3.2, the models will be used to forecast the time span of the test set (January 2022 - June 2023) in order to evaluate the models forecasting performance in two ways, visually and with forecasting errors.

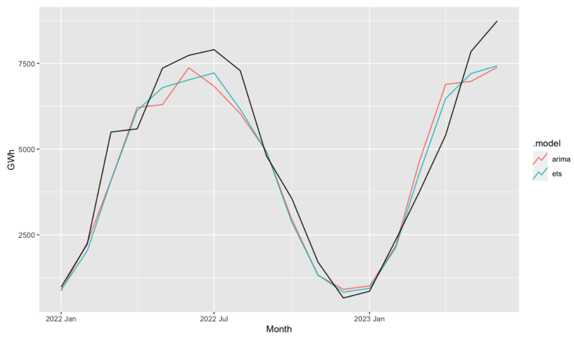


Figure 5: *Forecasted (ARIMA & ETS) and Actual Power Produced by PV Systems in Germany from January 2022 to June 2023*

The Figures 5 and 17 (in the Appendix) visualize the forecasts of the chosen models, the ARIMA(0,0,0)(2,1,0) and the ETS(A,A,M). Both models predict the upstream in spring phases, the downstream in autumn phases similar precise and capture the local minimum at the turn of the year pretty accurate. On the other side, both models seem to have difficulties predicting the peaks in the summer months correct since both predict in the summer times lower values as actually produced. The confidence intervals of the models appear to have similar widths and capture the actual data points within the 95% confidence interval most of the time. Moreover, the ARIMA model seem to get a bit less precise than the ETS the further the time goes in the forecast which might be an indicator that the ETS is more robust for longer forecasting periods. It cannot be said, which of the models produces the better forecasts by only looking at the visualizations. Therefore, the models' performances is evaluated by the forecast errors Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Mean Absolute Squared Error (MASE).

	RMSE	MAE	MAPE	MASE
ARIMA	837	680	15.8	1.57
ETS	717	589	13.4	1.36

Table 4: *Model Evaluation Metrics*

Since both models forecast the power produced on the same unit (GWh), two scale-dependent errors RMSE and MAE are used. Furthermore, the evaluation includes two scale-independent error measures with the MAPE and the MASE. Table 4 shows clearly that the ETS(A,A,M) outperforms the ARIMA(0,0,0)(2,1,0) in all of the four forecast errors. Therefore, the model ETS(A,A,M) is chosen to forecast the power generated by PV systems in Germany from July 2023 for the next 18 months.

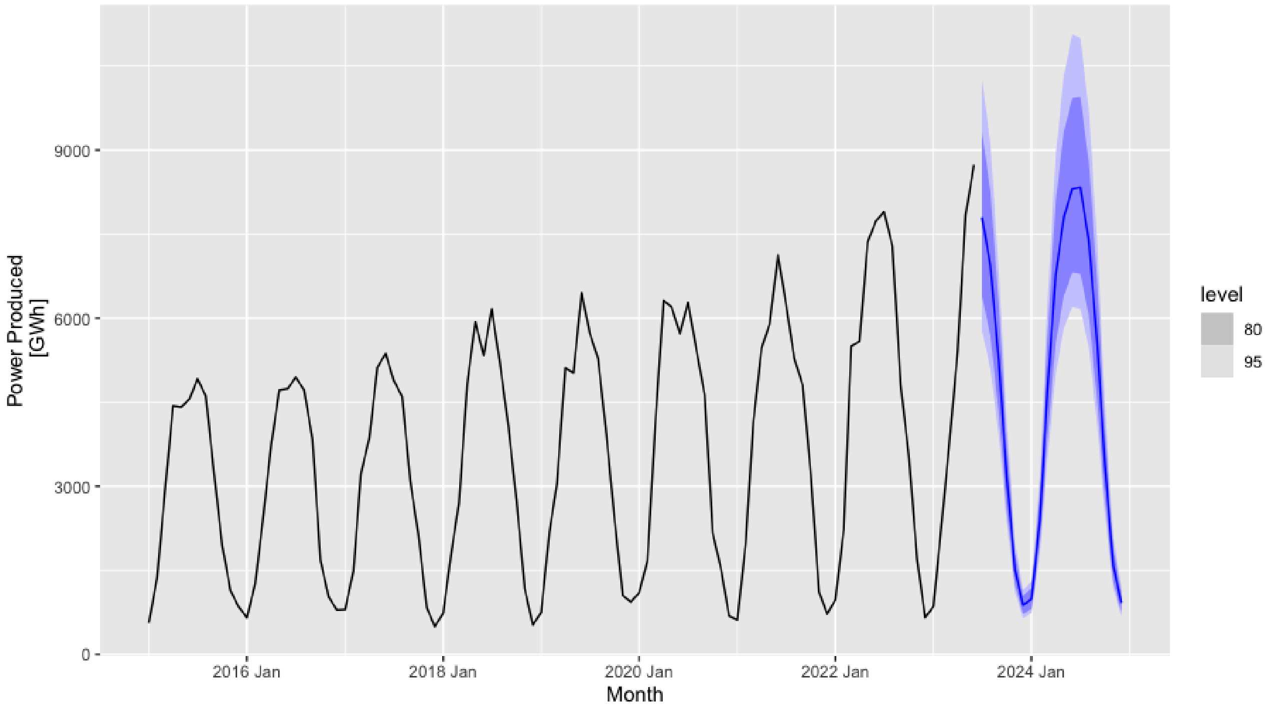


Figure 6: *Forecast of Monthly Power Produced by PV Systems in Germany from July 2023 to December 2024*

5 Conclusion

The goal of this paper was to find the model producing the best forecasts for the power produced by PV systems in Germany by applying two different forecasting methods.

In the beginning, an exploratory analysis of the data was performed. Non-linearity, trends, seasonality, and variance patterns were examined. The data was also tested for unit roots and structural breaks to anticipate potential measures ensuring a correct modeling process. In Section 3, both manual and automatic approaches were employed to derive ETS and ARIMA models, ensuring that a comprehensive range of potential models was explored. The selection of the best ETS and ARIMA model for forecasting was based on the criteria AIC, AICc, and BIC. The chosen models were used to forecast the test data in Section 4. A visual examination and a comparison of the forecast errors was conducted to evaluate the models and demonstrated that the models capture the seasonal behaviour of the data well but losing accuracy in the local maximums of each season. Furthermore, the evaluation indicated that the ETS(A,A,M) model is the most suitable to use for further forecasting.

In Figure 6, this model was leveraged to forecast the solar power generated in Germany from July 2023 to December 2024. Corresponding with the findings in the Results, the model seems confident in predicting the seasonal variations and the local minima. On the other hand, the model has wider confidence intervals in the summer months indicating that it is more difficult to predict the local maximum in a season. Overall, the forecast suggests a continuing seasonal behaviour and, according to the 95% confidence interval, a potential all-time high in the summer 2024 for the power produced by PV Systems in Germany.

References

Enders, W. (2014). *Applied econometric time series* (4th). Wiley.

Franco, D. (2022). Exponential smoothing methods for time series forecasting. Retrieved August 1, 2023, from https://www.encora.com/insights/exponential-smoothing-methods-for-time-series-forecasting

Hyndman, R. J., & Athanasopoulos, G. (2021). Forecasting: Principles and practice (3rd). *OTexts*. Retrieved July 30, 2023, from https://OTexts.com/fpp3

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, *54*(1-3), 159–178. https://doi.org/10.1016/0304-4076(92)90104-Y

Presse- und Informationsamt der Bundesregierung. (2023). Energieversorgung in deutschland - klimafreundlich und krisensicher. Retrieved July 30, 2023, from https://www.bundesregierung.de/breg-de/schwerpunkte/klimaschutz/energieversorgung-sicherheit-2040098

Quandt, R. E. (1960). Tests of the hypothesis that a linear regression system obeys two separate regimes. *Journal of the American Statistical Association*, *55*(290), 324–330. https://doi.org/10.1080/01621459.1960.10482067

Appendix

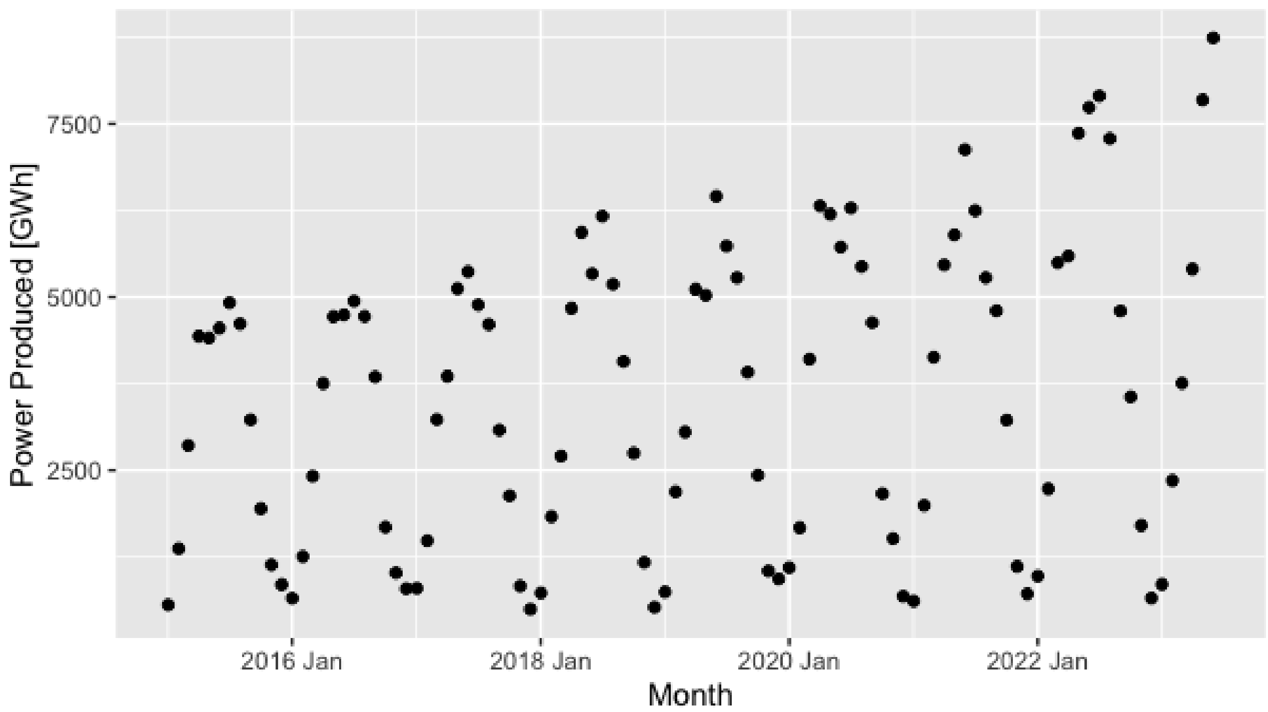


Figure 7: *Scatterplot of Original Time Series*

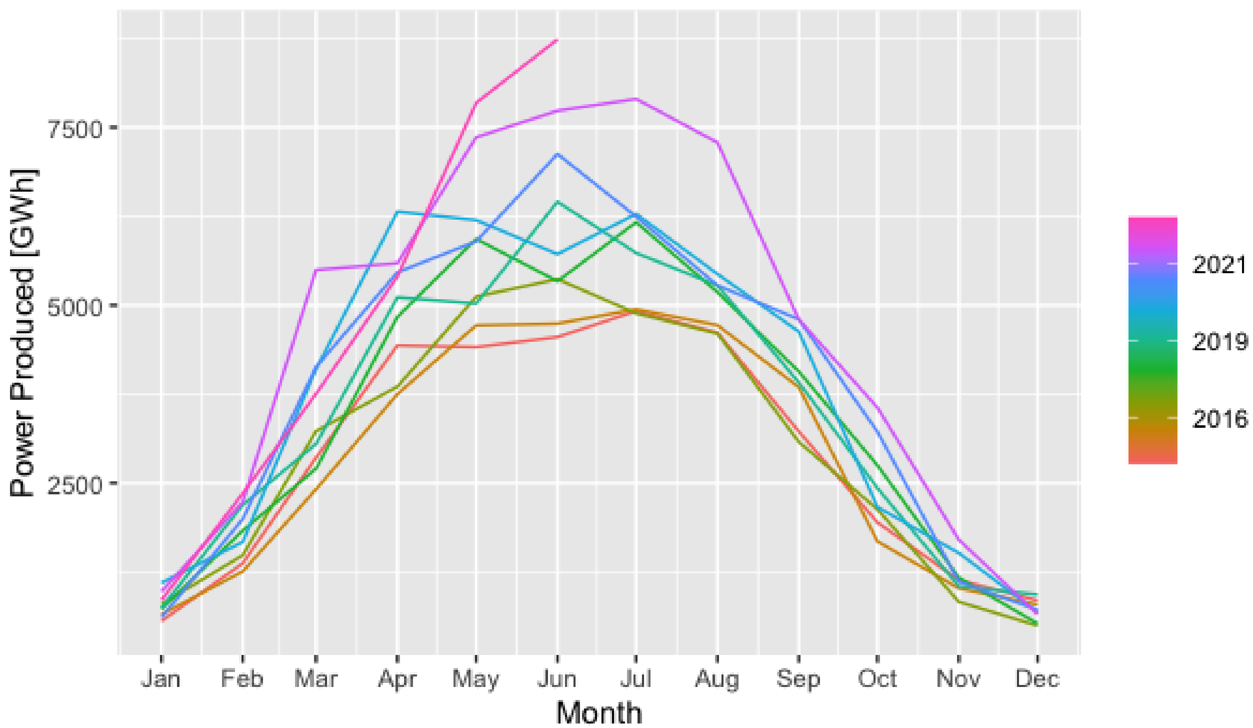
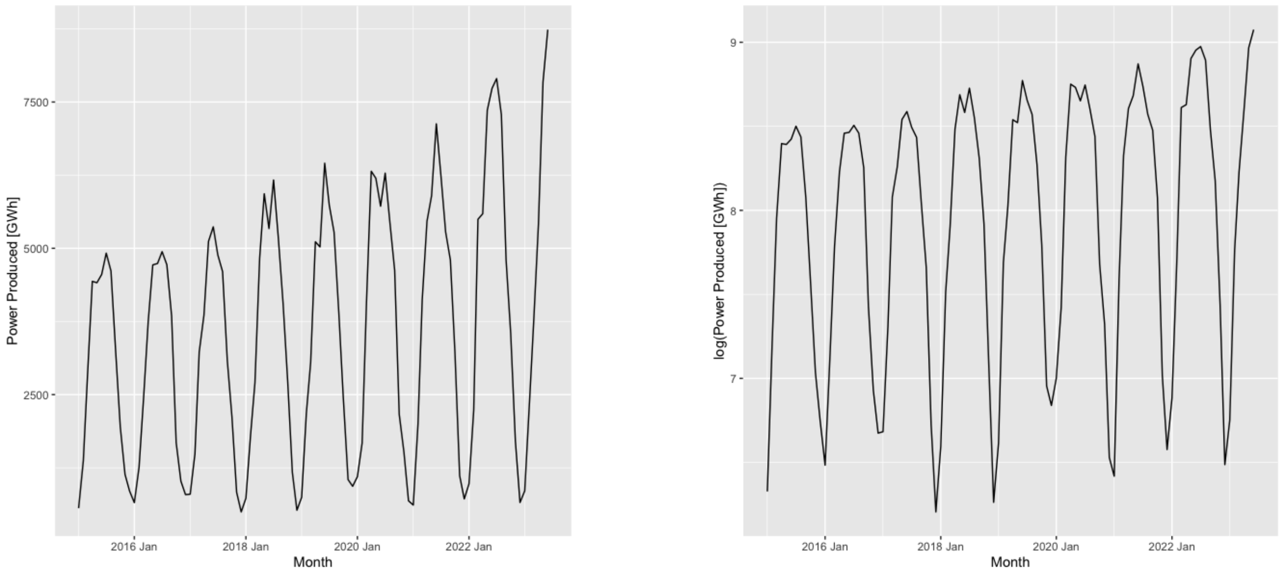


Figure 8: *Seasonal Plot of Original Time Series*



(a) Original Time Series

(b) Log. Transformed Time Series

Figure 9: *Comparison between Original and Transformed Time Series*

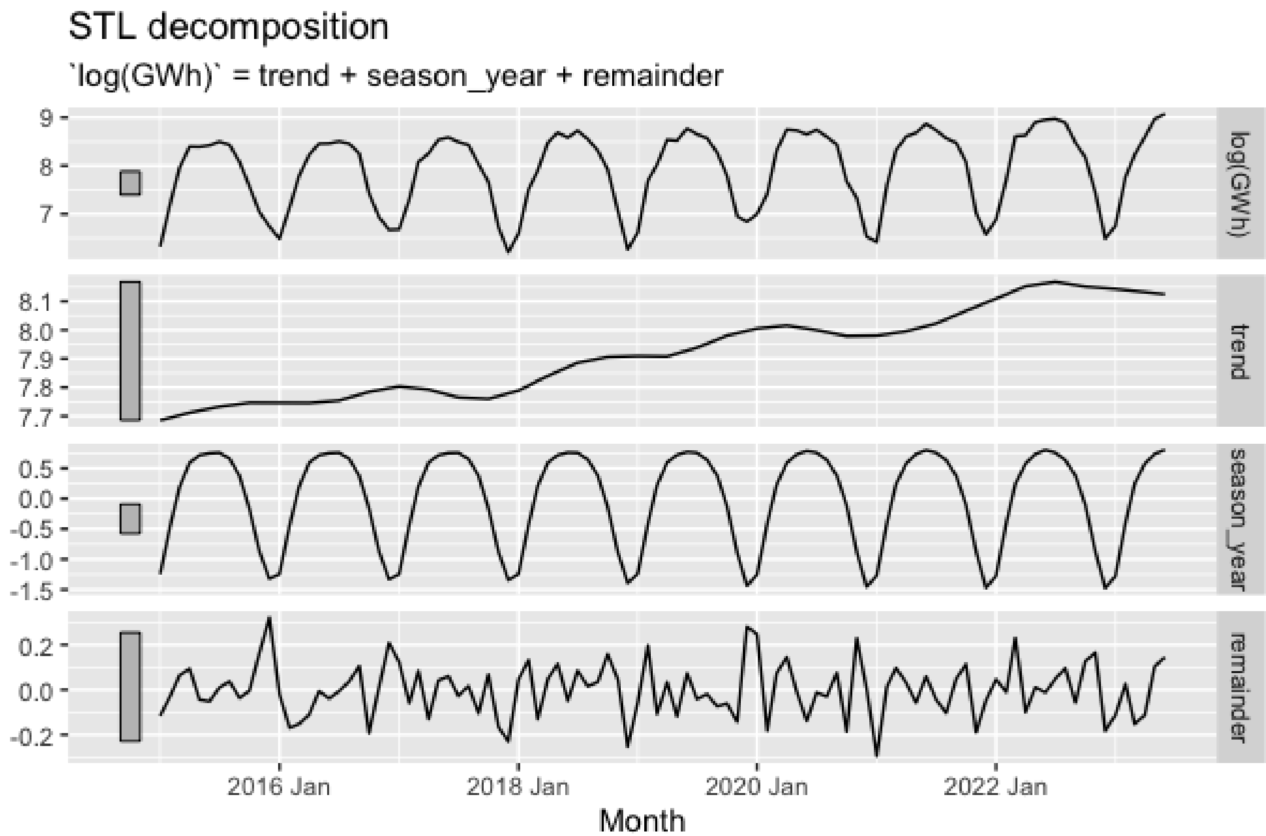


Figure 10: *STL Decomposition of Logarithmic Transformed Time Series*

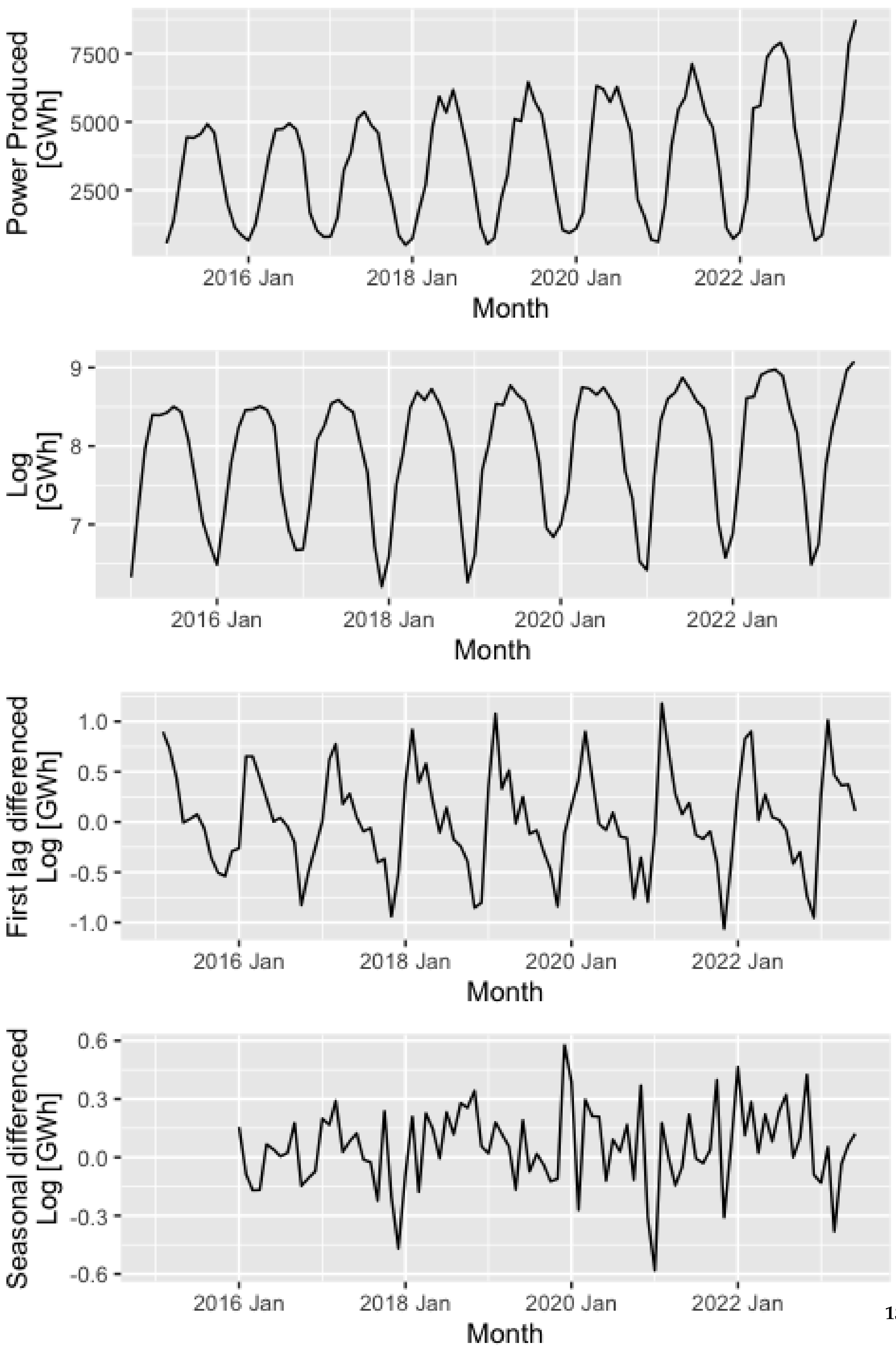


Figure 11: *Time Series in Original, Transformed and Differenced States*

Transformation	Type of test	t-value	$\tau3$ (1%)	$\tau3$ (5%)	$\tau3$ (10%)
log(GWh)	Trend	-8.06	-3.99	-3.43	-3.13
	Drift	-7.86	-3.46	-2.88	-2.57
	None	-0.50	-2.58	-1.95	-1.62
difference(log(GWh))	Trend	-6.37	-3.99	-3.43	-3.13
	Drift	-6.41	-3.46	-2.88	-2.57
	None	-6.43	-2.58	-1.95	-1.62
difference(log(GWh), lags = 12)	Trend	-7.02	-4.04	-3.45	-3.15
	Drift	-6.99	-3.51	-2.89	-2.58
	None	-6.41	-2.6	-1.95	-1.61

Table 5: *Augmented Dickey-Fuller Test Results*

Transformation	Type	KPSS (default lags)				KPSS (use.lag = 12)			
		t-val	1%	5%	10%	t-val	1%	5%	10%
log(GWh)	τ	.02	.22	.18	.12	.23	.22	.18	.12
	μ	.15	.74	.57	.35	.77	.74	.57	.35
diff(log(GWh))	τ	.03	.22	.18	.12	.20	.22	.18	.12
	μ	.03	.74	.57	.35	.20	.74	.57	.35
diff(log(GWh), lags = 12)	τ	.05	.22	.18	.12	.06	.22	.18	.12
	μ	.10	.74	.57	.35	.12	.74	.57	.35

Table 6: *KPSS Test Results*

```
> # QLR test
> qlr1 <- Fstats(Lag0 ~ Lag1 + Lag12, data = sb_data, from = 0.2)
> test1 <- sctest(qlr1, type = "supF")
> test1

        supF test

data:  qlr1
sup.F = 6.5551, p-value = 0.5219
```

Figure 12: *QLR Test*

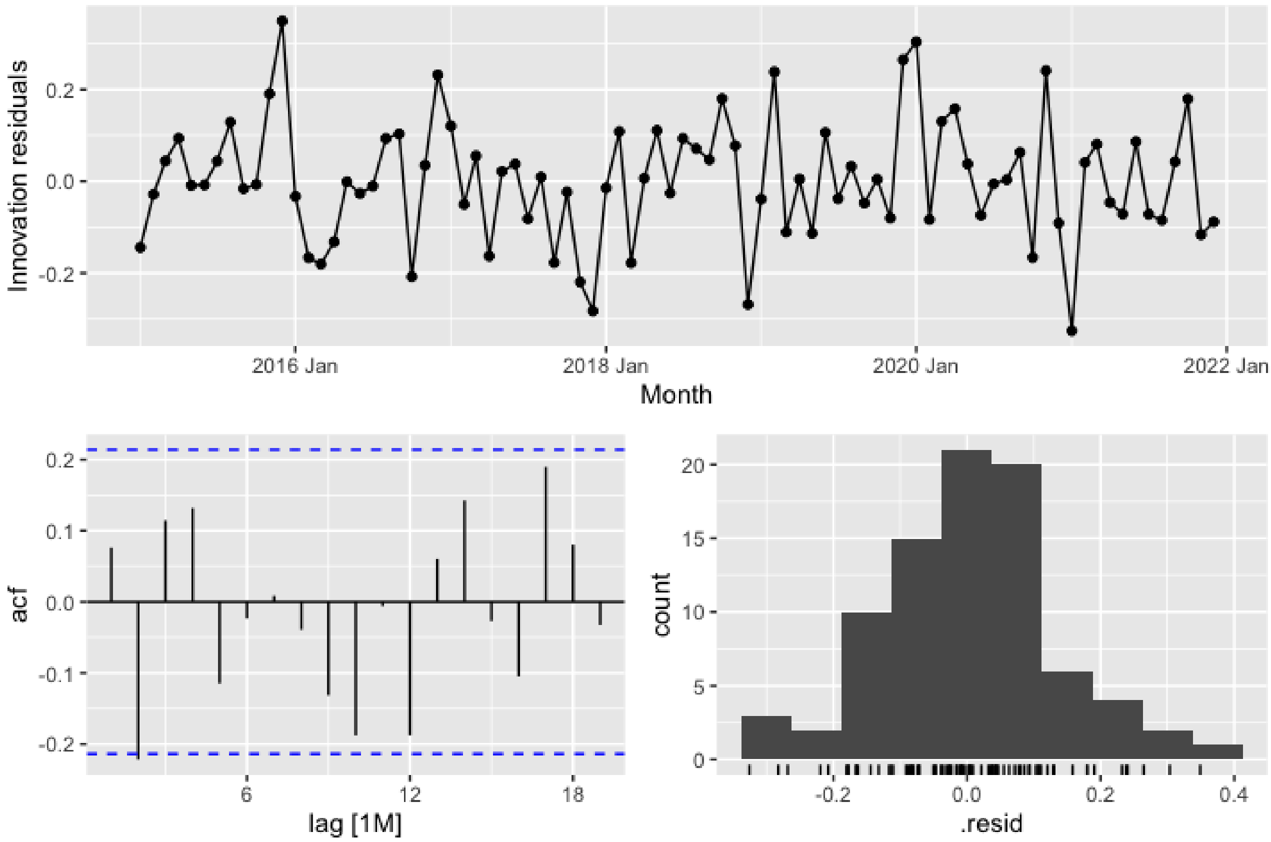


Figure 13: *Residuals of the ETS(A,A,M) model*

```
> # Ljung-Box Test
> augment(ets_fit[2]) %>%
+   features(.innov, ljung_box, lag = 24)
# A tibble: 1 x 3
  .model     lb_stat lb_pvalue
  <chr>        <dbl>     <dbl>
1 ets_guess     38.4    0.0318
```

Figure 14: *Ljung-Box Test for ETS(A,A,M)*

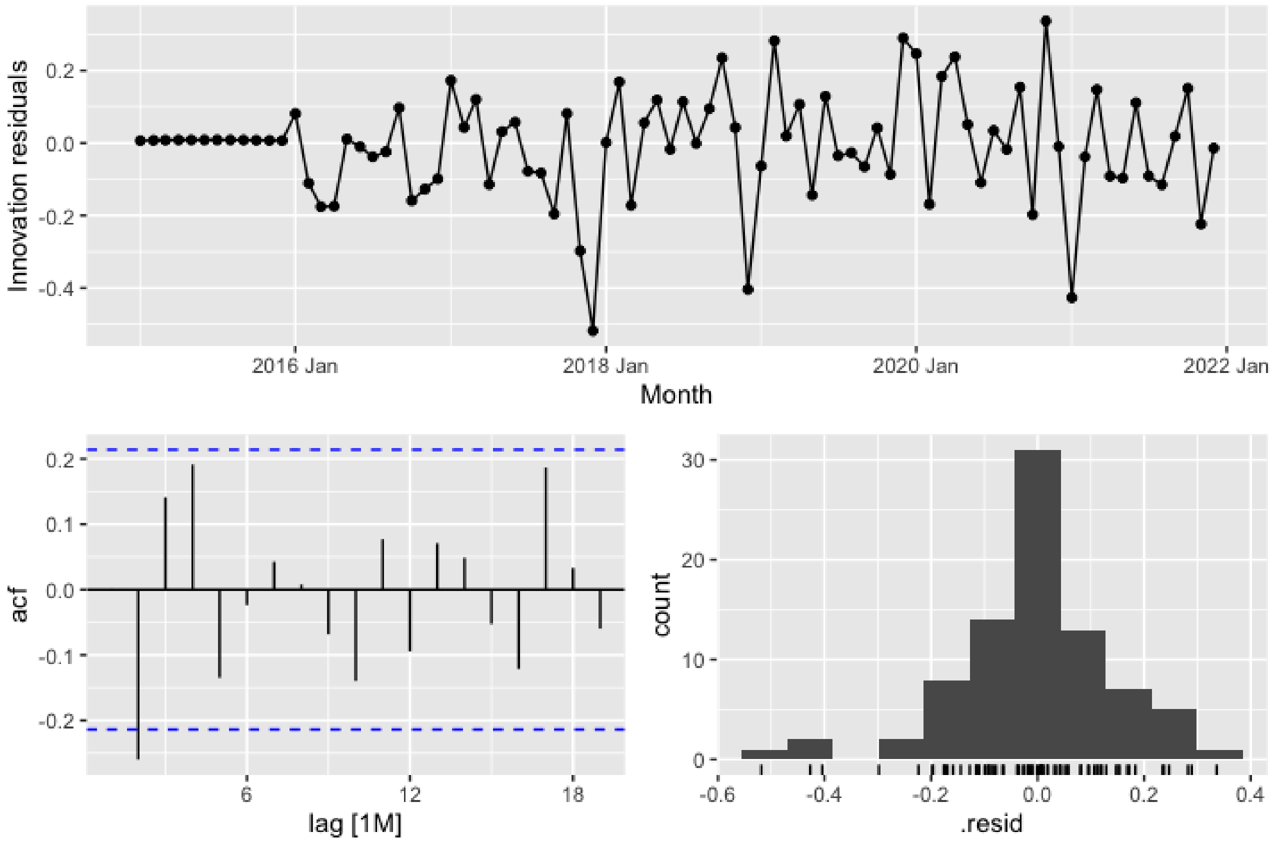


Figure 15: *Residuals of the ARIMA(0,0,0)(2,1,0) model*

```
> # Ljung-Box Test
> augment(arima_fit[2]) %>%
+   features(.innov, ljung_box, lag = 24)
# A tibble: 1 x 3
  .model      lb_stat lb_pvalue
  <chr>         <dbl>     <dbl>
1 arima_auto     30.4     0.171
```

Figure 16: *Ljung-Box Test for ARIMA(0,0,0)(2,1,0)*

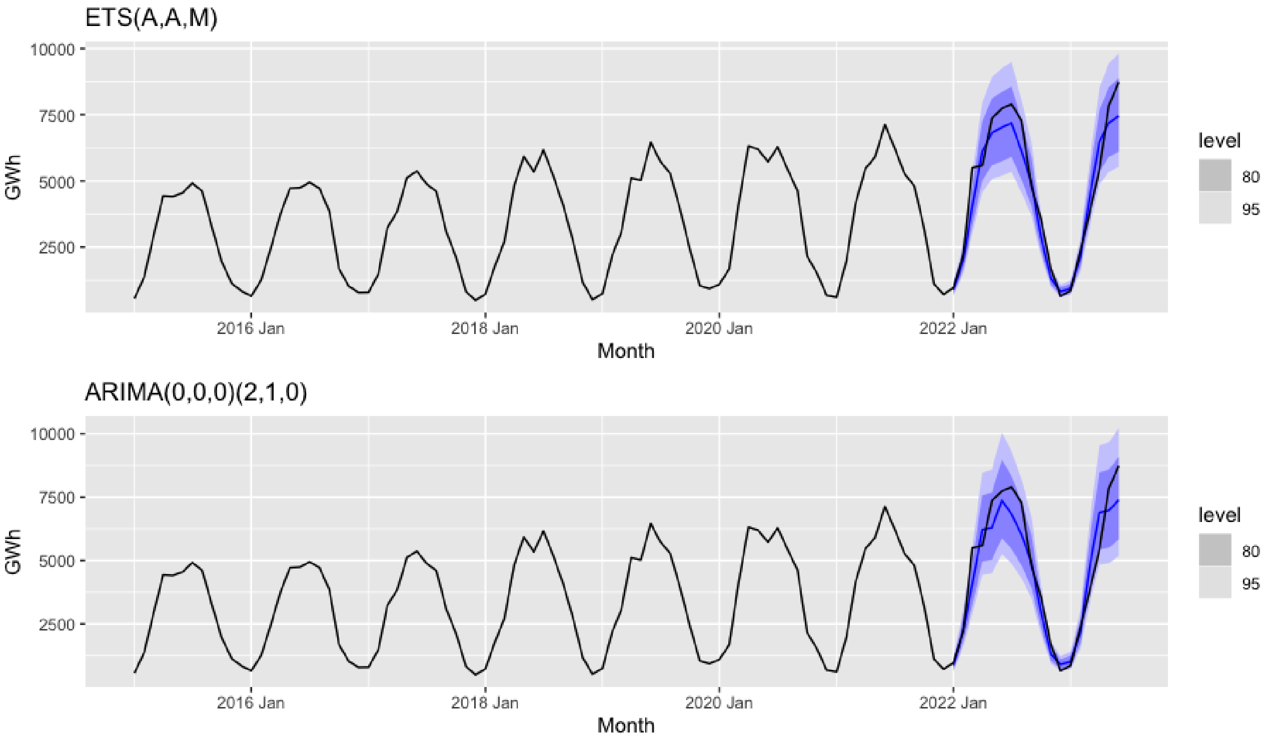


Figure 17: *Forecasted (ARIMA & ETS) and Actual Power Produced by PV Systems in Germany from January 2022 to June 2023 - Separated*