COMPUTER EXERCISE 4 - DECEMBER 25, 2023

Advanced Time Series Analysis
02427

Wind power forecasting

Student :
Théophile SCHMUTZ

Teacher :
Henrik MADSEN

Contents

1 Introduction **3**
- 1.1 Context . . . 3
- 1.2 Definition of Wind Power Forecast . . . 3
- 1.3 Reference model . . . 4
- 1.4 Very short-term forecasting . . . 4
- 1.5 Evaluation of forecasts . . . 4
- 1.6 Objective . . . 5

2 Power curve model **6**
- 2.1 Principle . . . 6
- 2.2 Power curve . . . 6
  - 2.2.1 The Power Curve model . . . 7
  - 2.2.2 Data cleaning . . . 7
  - 2.2.3 Fitting the power curve . . . 11
- 2.3 Residuals analysis . . . 13
  - 2.3.1 1-hour ahead power curve forecast . . . 13
  - 2.3.2 2-hour ahead power curve forecast . . . 14
  - 2.3.3 3-hour ahead power curve forecast . . . 16
- 2.4 Conclusion . . . 18

3 ARIMA(1,1,1) model **20**
- 3.1 Principle . . . 20
  - 3.1.1 ARIMA(1,1,1) model Introduction . . . 20
  - 3.1.2 Forecasting with an ARIMA model . . . 21
- 3.2 Residuals Analysis . . . 21
  - 3.2.1 Results of the fit . . . 22
  - 3.2.2 Results of the forecasts . . . 26

4 ARMA(1,1)-GARCH(1,1) model **29**
- 4.1 Principle . . . 29
- 4.2 Residuals Analysis . . . 29
  - 4.2.1 Results of the fit . . . 29
  - 4.2.2 Results of the forecasts . . . 33

5 ARIMA(1,1,1)-GARCH(1,1) model **35**
5.1 Principle 35
5.2 Residuals Analysis 35
5.2.1 Results of the fit 35
5.2.2 Results of the forecasts 39

6 Conclusion **40**

Chapter 1

Introduction

1.1 Context

Many countries and regions are introducing policies aimed at reducing the environmental footprint from the energy sector and increasing the use of renewable energy.

With the current focus on energy and the environment, efficient integration of renewable energy into the electric power system is becoming increasingly important. In Europe, several countries already have a high penetration of wind power (i.e., in the range of 7 to 20% of electricity consumption in countries such as Germany, Spain, Portugal, and Denmark).

A large-scale introduction of wind power causes a number of challenges for electricity market and power system operators who will have to deal with the variability and uncertainty in wind power generation when making their scheduling and dispatch decisions. Wind power forecasting (WPF) is frequently identified as an important tool to address the variability and uncertainty in wind power and to more efficiently operate power systems with large wind power penetrations.

The objective of this exercise is to construct adaptive models that predict power production at various time horizons based on the provided weather forecasts and measured power output.

1.2 Definition of Wind Power Forecast

The forecasted wind generation made at time instant, t, for a look-ahead time, $t+k$, is the average power $p_{t+k|t}$, the wind farm is expected to generate during the considered period of time (e.g., 1 hr) if it would operate under an equivalent constant wind. Forecasts are made for a time horizon, T, indicating the total length of the forecast period (e.g., 3 hr ahead) in the future. The time resolution of the forecasts is denoted by the time step k.

1.3 Reference model

Persistence power forecasting assumes that the wind power at a certain future time will be the same as it is when the forecast is made, which can be formulated as,

$$\hat{p}_{t+k|t} = p_t.$$

Persistence is obviously a very simple method and is mentioned here since it is used as a reference to evaluate the performance of advanced methods. An advanced method is worth implementing if it outperforms persistence. Wind, however, is somehow persistent in nature. Persistence is a difficult method to beat, especially on the short-term (1 – 6 hr).

1.4 Very short-term forecasting

The very short-term forecasting approach consists of statistical models based on the time series approach, such as the Kalman Filters, Auto-Regressive Moving Average (ARMA), AutoRegressive with Exogenous Input (ARX), and Box-Jenkins forecasting methods. These types of models only take as inputs past values from the forecasted variable (i.e., wind power). At the same time, they can also use other explanatory variables (e.g., wind direction, temperature), which can improve the forecast error. Since these methods are merely based on past production data, they only outperform the persistence model (reference model) for forecast horizons of between 3 and 6 hr.

The univariate model only considers past values of wind power generation $\{p_t\}$. The univariate model can be expressed as,

$$\hat{p}_{t+k|t} = f\left(p_{t-n:t}\right) + e_t, \quad n,k > 0.$$

With, $\{e_t\}$ a white noise and f a generic function that can be linear or nonlinear. $\{p_{t-n:t}\}$ is denoting the power production of the wind farm at time $t-n,\ldots,t$. The t refers to the time (column `toy` in the data), in the following $t \in [\![1;T-k]\!]$, T being the last instant on the column `toy`. The multivariate models not only use past values of that variable, but also past or present values of other variables (wind speed, wind direction, temperature). These past values are measured. The multivariate model can be expressed as,

$$\hat{p}_{t+k|t} = f\left(p_{t-n:t}, s_{t-n:t}, d_{t-n:t}, \mathcal{T}_{t-n:t}\right) + e_t$$

With still $\{e_t\}$ a white noise and f a generic function ; denoting $\{s_t\}$ the wind speed, $\{d_t\}$ the wind direction and $\{\mathcal{T}_t\}$ the temperatures.

1.5 Evaluation of forecasts

Wind power forecasts are characterised by an inherent uncertainty, which means that no available wind power prediction can ever be exact. Therefore, it is essential that wind power forecasts are properly evaluated, not only to assess the

performance of the chosen approaches adequately, but also to obtain a deeper understanding of what characterises the prediction uncertainty.

Evaluation of the quality of forecasting methods is conducted by comparing wind power predictions made at a certain time directly with the actual corresponding observations. Hence, given a fixed horizon k the prediction error observed at a given time t for a prediction made at time origin $t-k$, $e_{t|t-k}$ is defined as the difference between the value of wind power that is effectively measured at t, p_t, and the value of wind power at t that was originally predicted at $t-k$, $\hat{p}_{t|t-k}$,

$$e_{t|t-k} = \hat{p}_{t|t-k} - p_t.$$

The mean square error between the two corresponding time series, **mse**, will be calculated as the square root of the squared error value,

$$\mathbf{mse}_k = \frac{1}{T}\sum_{t=1}^{T}\left(\hat{p}_{t|t-k} - p_t\right)^2.$$

The **mse** can be decompose into two parts,

$$\mathbf{mse}_k = \mathbf{bias}_k^2 + \mathbf{Var}(\hat{p}_{t|t-k})$$

Where the bias $\mathbf{bias}_k$ translates the difference between the mean values of the predictions and the measures series,

$$\begin{aligned}\mathbf{bias}_k &= \frac{1}{T}\sum_{t=1}^{T} e_{t|t-k} = \frac{1}{T}\sum_{t=1}^{T}\left(\hat{p}_{t|t-k} - p_t\right),\\ \mathbf{Var}(\hat{p}_{t|t-k}) &= \frac{1}{T}\sum_{t=1}^{T}\left(\hat{p}_{t|t-k} - \overline{p_{\cdot|\cdot-k}}\right)^2.\end{aligned}$$

With $\overline{p_{\cdot|\cdot-k}}$ the unbiased estimator of the empirical estimator of the mean.

From these metrics, please note that the smaller the $\mathbf{bias}_k$ in absolute values, the smaller the $\mathbf{mse}_k$. Similarly, the smaller the variance, the smaller the $\mathbf{mse}_k$. Intuitively, as the variance decreases, the variations of the estimator will be less dispersed around $\sum_{t=1}^{T} p_t/T$, especially for an unbiased estimator where $\mathbf{mse}_k = \mathbf{Var}(\hat{p}_{t|t-k})$.

1.6 Objective

We are requested to develop adaptive models for the prediction of wind power 1, 2, and 3 hours ahead of a wind farm. We have at our disposal the hourly averages of wind power measurements and weather forecasts (including 1-hour, 2-hour and 3-hour ahead temperature, wind speed and wind direction forecasts).

We shall first consider a multivariate model based on the estimation of a power curve. Second, we shall implement a straightforward ARIMA(1,1,1) model. Third, we shall explore considering a ARIMA(1,1,1)-GARCH(1,1) model, we will use a GARCH model for forecasting the residuals. The code is available here.

Chapter 2

Power curve model

2.1 Principle

The current majority of short-term wind power forecast approaches require meteorological predictions as inputs to forecast. The main feature that distinguishes the approaches has to do with the way predictions of meteorological variables are converted to predictions of wind power generation through the power curve.

2.2 Power curve

Wind forecasts for wind energy applications rely mostly on wind speed and direction, and only marginally on the forecast of air density. Because of the transfer functions of available generators (power curve depicted in fig. 2.1), the conversion of available wind power (which is proportional to the cube of the wind's speed) into actual power varies nonlinearly, with zero output below a minimum speed threshold (around $3m/s$), a rapid output in growth until the machine attains its nominal power (around $15m/s$), and a constant output above that level until the cut-off speed is attained (around $25m/s$).

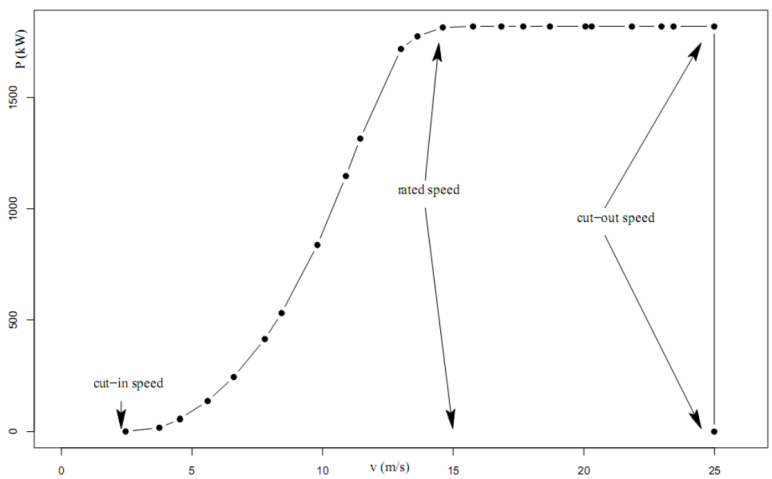

Figure 2.1: *Example of a power curve*

Because generators in a given wind park may interact with each other, not only

through their perturbed wakes but also because of local topographic speed-up or speed-down effects, the transfer functions of a park may vary significantly with wind direction. On the other hand, because the generators require a significant amount of time to align themselves with the prevailing wind, especially when they have to unwind after attaining maximum rotation in one direction, wind power is also a function of wind direction variability. Furthermore, the operation of wind generators will also be badly affected by small-scale turbulence.

Because of the non-linearity of the power curve, errors in wind speed forecast are penalised heterogeneously, when one thinks in terms of actual power. Errors at very low speeds are irrelevant since the output is always zero. On the other hand, errors in the flat region of the power curve (between 12 and $25m/s$) are also irrelevant as the output is constant, unless one has to consider changes in wind direction or in the intensity of turbulence. Errors at low-to-moderate speeds ($3-12m/s$) are highly penalised, as a small error in speed leads to a large error in power. Finally, the worst errors are obtained near the cut-off speed (around $25m/s$), when the system shifts abruptly from maximum output to zero output (or vice-versa).

2.2.1 The Power Curve model

Let $\hat{p}^{PC}_{t+k|t}$, $k \in \{1,2,3\}$, the power curve forecast at horizon k.

$$\hat{p}^{PC}_{t+k|t} = f_k\left(s_{t+k|t}\right) + e_t$$

With $\{e_t\}$ representing white noise and f_k denoting the power curve for k-hour ahead forecasting, where $\{s_{t+k|t}\}$ is the provided wind speed forecasts k-hours ahead. The primary challenge in this exercise lies in the precise formulation of the power curve function f_k. Typically, wind farm construction companies furnish a theoretical power curve. Nevertheless, in our scenario, we need to ascertain the relationship between the recorded wind power and the wind speed forecasts. It is crucial to highlight that errors in this model stem from both the estimation of f_k (a tough challenge in itself) and the accuracy of the wind speed forecasts.

Our first task is to arrive at a curve similar to fig. 2.1.

2.2.2 Data cleaning

We fist plot the wind power output as a function of the wind speed for each horizon $k = 1,2,3$.

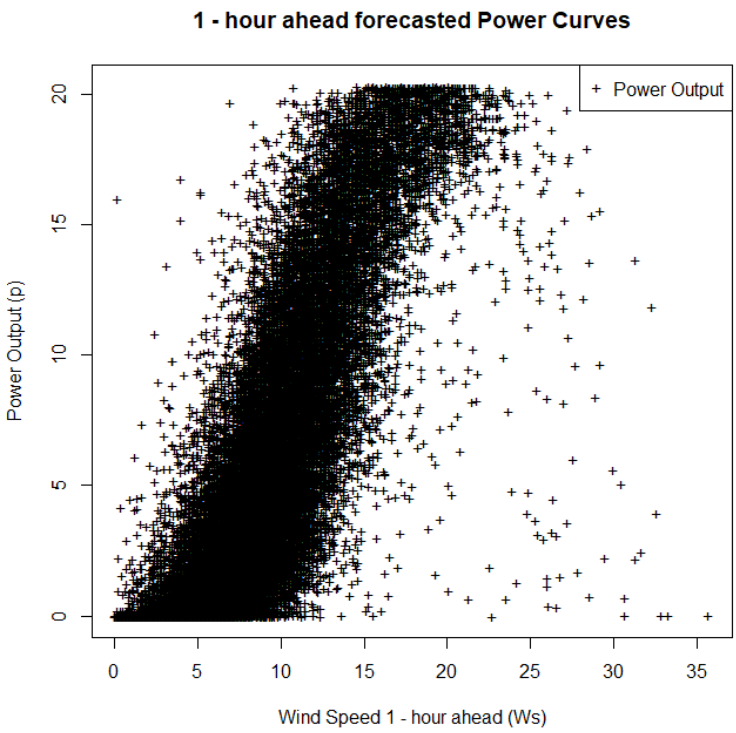

Figure 2.2: *Power curve for* `Ws1` *of the time period*

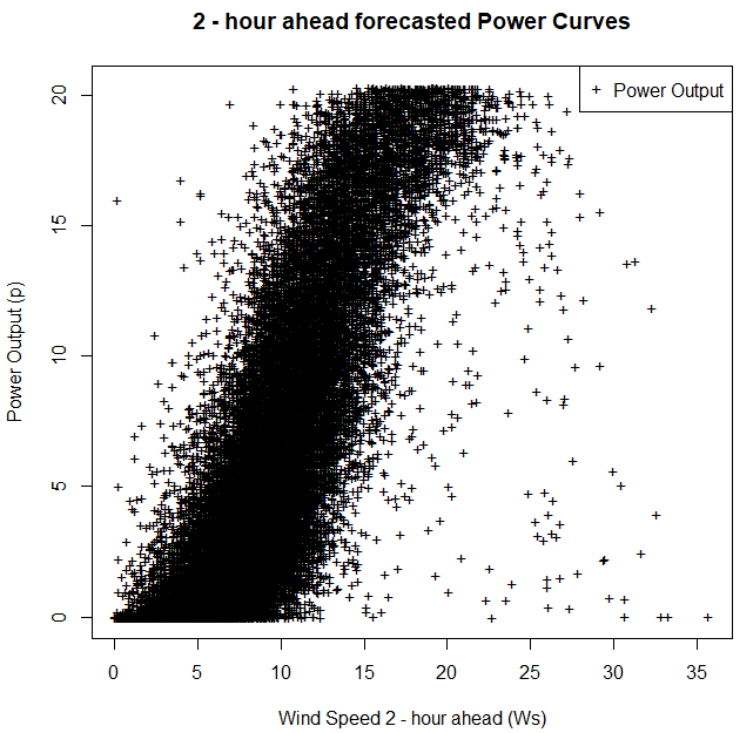

Figure 2.3: *Power curve for* `Ws2` *of the time period*

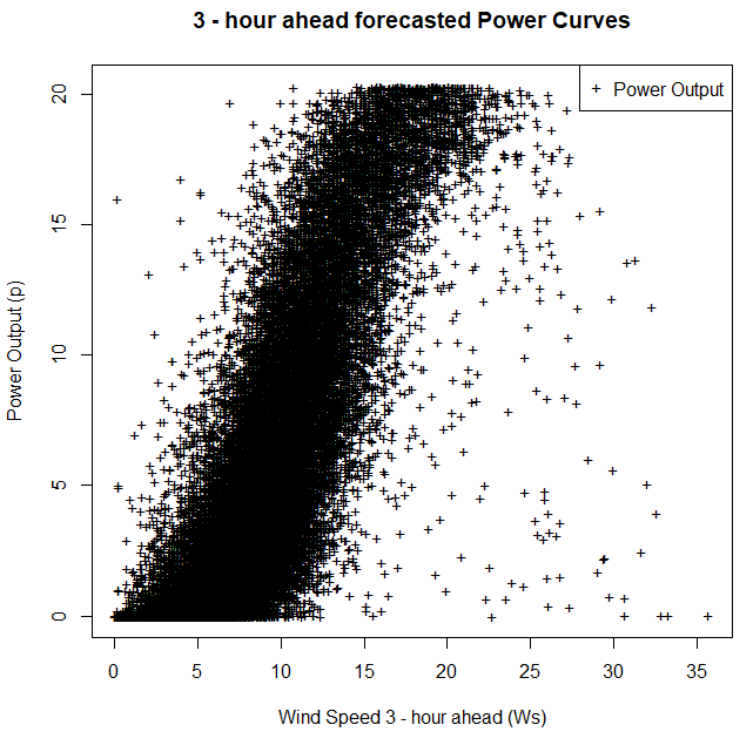

Figure 2.4: *Power curve for* `Ws3` *of the time period*

As illustrated in fig. 2.2, 2.3, and 2.4, data cleaning is needed for obtaining an accurate estimation of the power curve. Numerous outlier points are present, and their substantial deviation from relevant data points poses a risk of introducing bias into the power curve estimation. The initial attribution of these outliers is directed towards sudden changes in wind direction. Specifically, during these shifts, a wind turbine may not promptly align with the incoming wind, requiring some time to reposition and face the wind direction once again, thereby impacting the recorded power values. For our analysis, we will focus on examining the 1-hour ahead power curve, denoted as f_k with $k = 1$.

We then examine the impact of changes in wind direction on the power output. Sudden changes in wind direction are identified through the successive differences in wind direction, represented as $\Delta\{d_{t+k|t}\}$. A high absolute difference, $\delta_k(t) = \left|d_{t+k|t} - d_{t-1+k|t-1}\right|$, signifies a substantial change in wind direction.

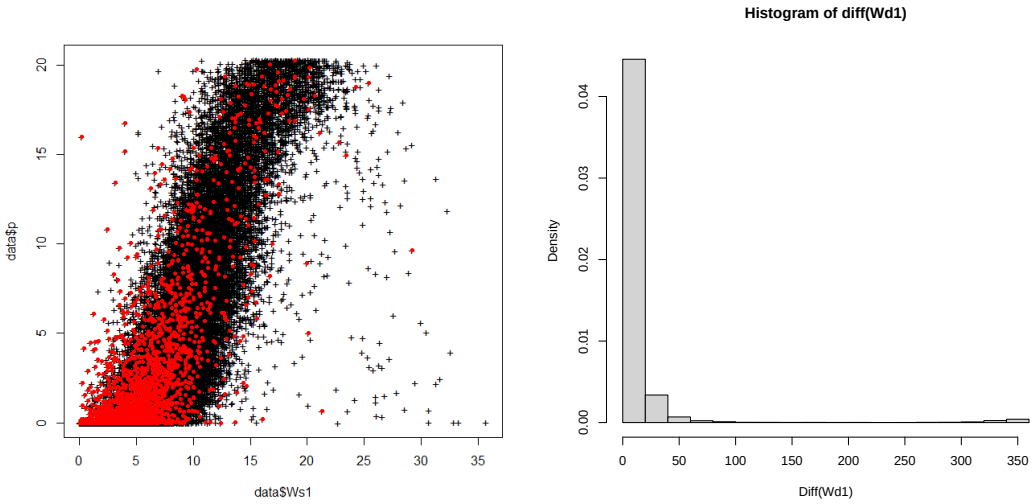

Figure 2.5: *The left-hand side plot highlights in red the points with a significant change in wind direction (change greater than 20°). The right-hand side displays the histogram of the 1-lagged absolute wind direction δ_k.*

According to fig. 2.5, there are significant changes in wind directions, but these points are not associated with outliers. This holds true for the 1-hour ahead, 2-hour ahead, and 3-hour ahead forecasts. Subsequently, we reverse our approach by focusing on selecting data points with wind speeds exceeding $25m/s$ to conduct a comprehensive examination of their properties. This marks a shift from initially examining points with changing wind directions to now exploring outliers in the context of high wind speeds.

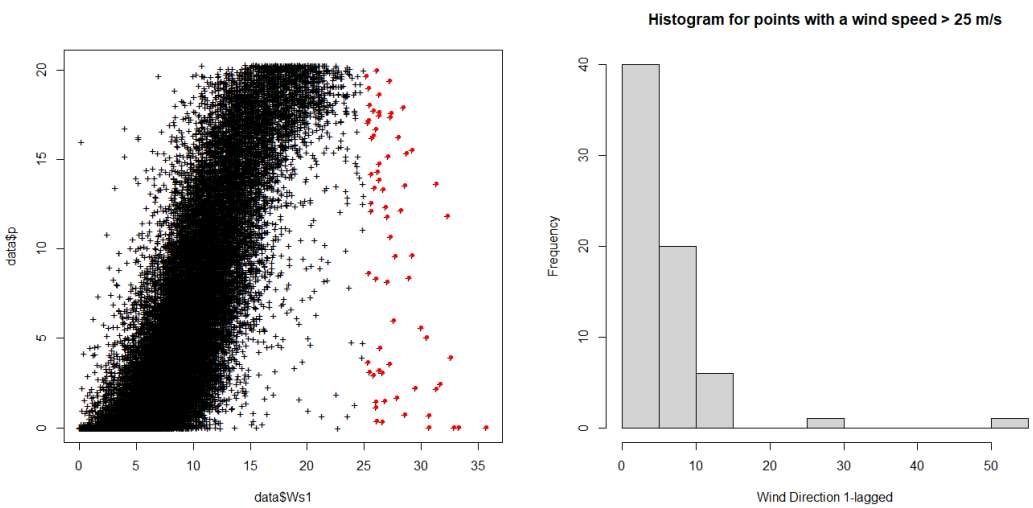

Figure 2.6: *Plot on the left, in red the points with wind speed higher than $25m/s$. Plot on the right, histogram of the 1-lagged wind direction of those points.*

Indeed, as observed in fig. 2.6, distinguishing these points solely based on wind speed changes seems impossible. Instead of pursuing this approach, we propose overwriting these points by setting the power output to a constant value where the wind speed exceeds $25m/s$. This simplification aims to address the complexity

associated with differentiating power output solely based on wind speed changes in these specific cases.

2.2.3 Fitting the power curve

After completing the small data cleaning process, the next step in our analysis involved enhancing the representation of the power curve. To achieve a more refined and continuous estimation, we opted for a smoothing technique using spline functions. Spline functions offer the advantage of providing a flexible and smooth fit to the data, effectively reducing the impact of outliers and irregularities that might have persisted in the original power curve. This smoothing process not only enhances the overall robustness of our power curve estimation but also contributes to a more accurate representation of the underlying trends in power production. The choice of spline functions enables us to strike a balance between capturing the nuanced variations in the data and achieving a coherent, smoothed power curve that better aligns with the true dynamics of the wind farm's power output.

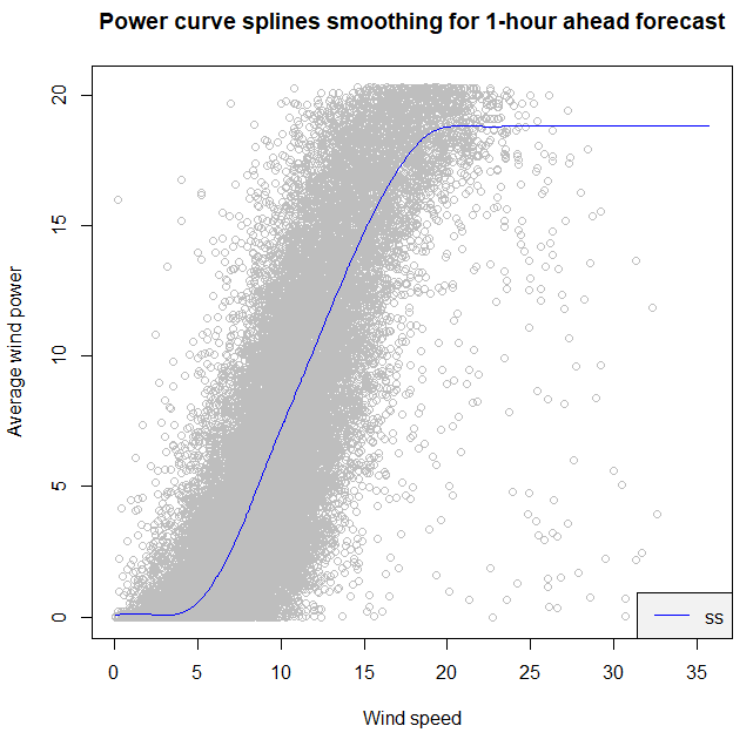

Figure 2.7: *Smoothing of the power curve for the 1-hour ahead forecasts*

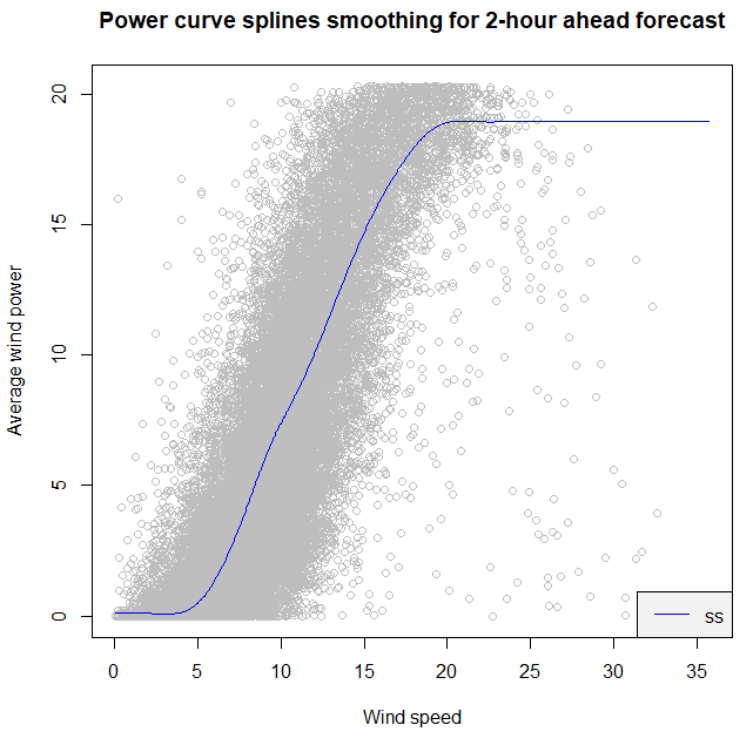

Figure 2.8: *Smoothing of the power curve for the 2-hour ahead forecasts*

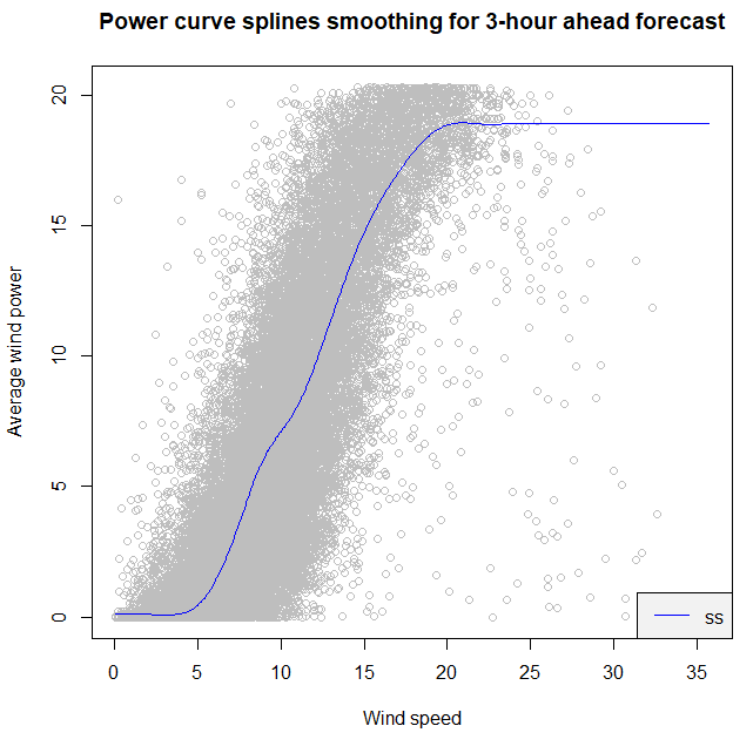

Figure 2.9: *Smoothing of the power curve for the 3-hour ahead forecasts*

From fig. 2.7, 2.8 and 2.9, the three power curves are fairly similar.

2.3 Residuals analysis

2.3.1 1-hour ahead power curve forecast

We first consider $\hat{p}^{PC}_{t+1|t}$, the 1-hour ahead power curve forecast with $\{s_{t+1|t}\}$ and f_1. We analyse the residuals of the power curve smoothing.

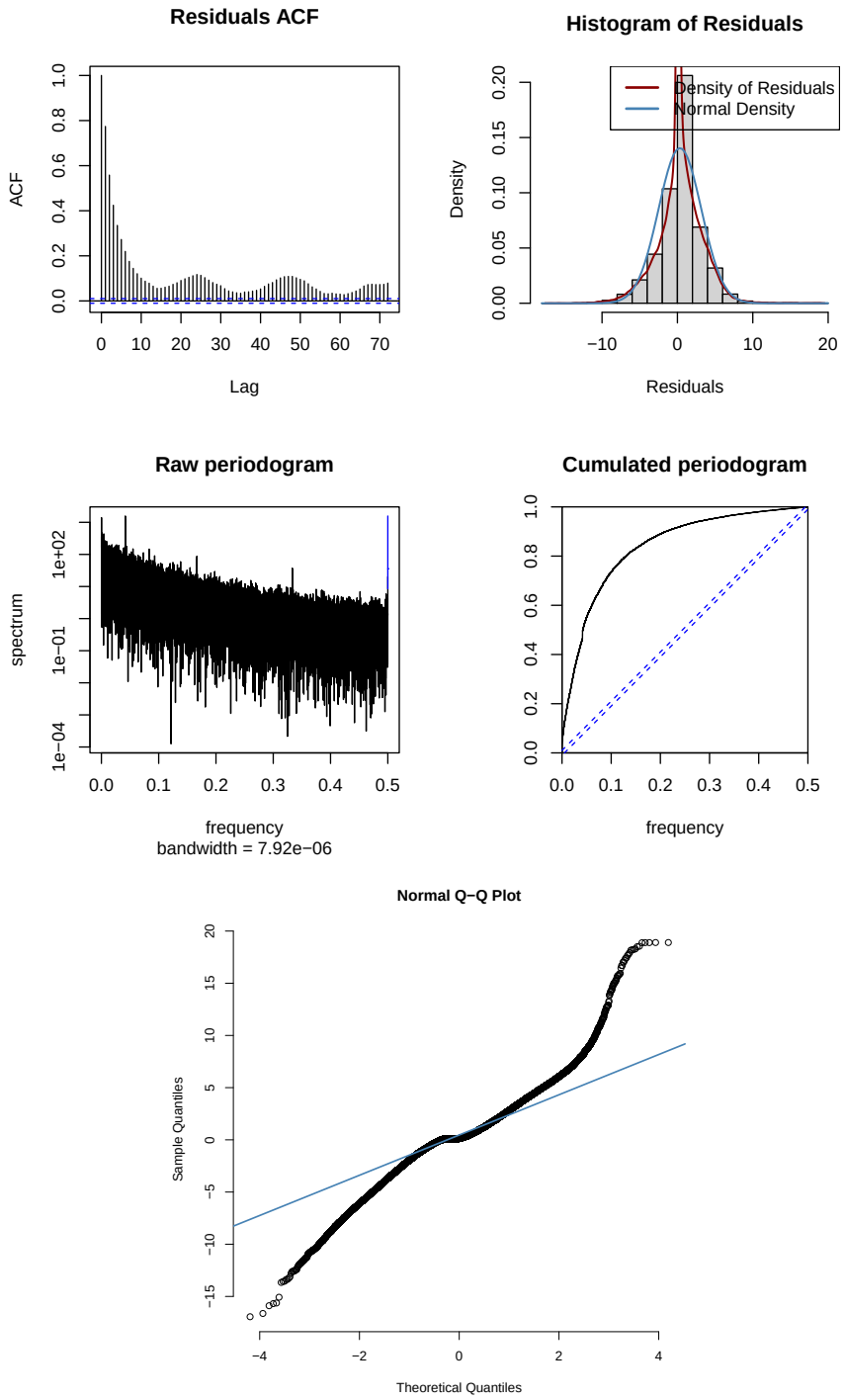

Figure 2.10: *Residuals of f_1 smoothing*

The histogram in Figure 2.10 also displays the kernel-estimated density de-

rived from the residuals, alongside a normal density characterised by the unbiased mean and unbiased variance estimated from the residuals. The figures in fig. 2.10 reveals that the residuals don't follow a white noise pattern. Notably, there are noticeable correlations among the residuals, and the cumulative plot doesn't resemble a straight line. In the same respect, the normal quantile-quantile plot is not close to a straight line. However, it's worth noting that the histogram of errors is slightly uncentered and exhibits a tail quite similar to a normal distribution.

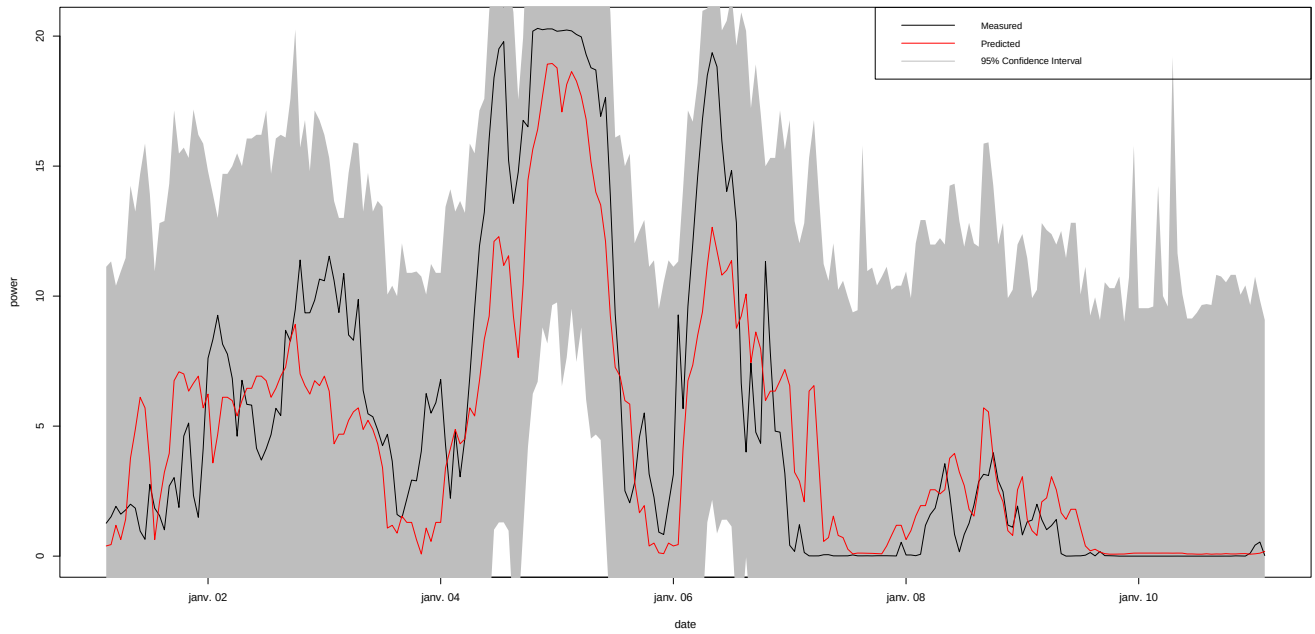

Figure 2.11: *1-hour ahead forecast $\hat{p}^{PC}_{t+1|t}$ during a week using. The gray area is the 95% confidence interval.*

Biais	Variance	MAE	MSE
0.333	8.06	1.97	8.17

Table 2.1: *Metrics : 1-hour ahead Power Curve forecast $\hat{p}^{PC}_{t+1|t}$.*

From the table, we find again there is a small bias.

2.3.2 2-hour ahead power curve forecast

We first consider $\hat{p}^{PC}_{t+2|t}$, the 2-hour ahead power curve forecast with $\{s_{t+2|t}\}$ and f_2. We analyse the residuals of the power curve smoothing.

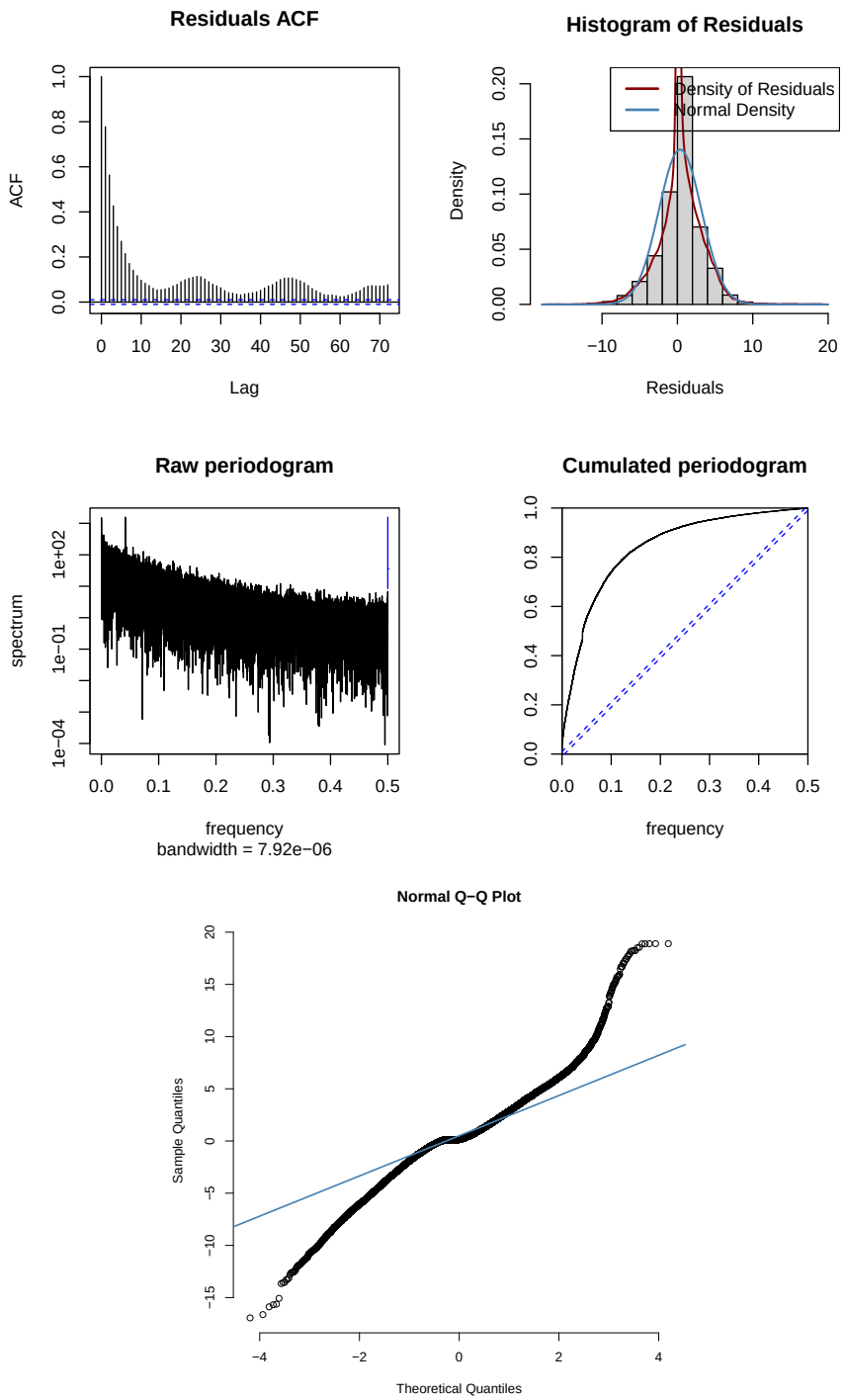

Figure 2.12: *Residuals of f_2 smoothing*

The histogram in Figure 2.12 also displays the kernel-estimated density derived from the residuals, alongside a normal density characterised by the unbiased mean and unbiased variance estimated from the residuals. The figures in fig. 2.12 reveals that the residuals don't follow a white noise pattern. Notably, there are noticeable correlations among the residuals, and the cumulative plot doesn't resemble a straight line. In the same respect, the normal quantile-quantile plot is not close to a straight line. However, it's worth noting that the histogram of errors is slightly uncentered and exhibits a tail quite similar to a

normal distribution.

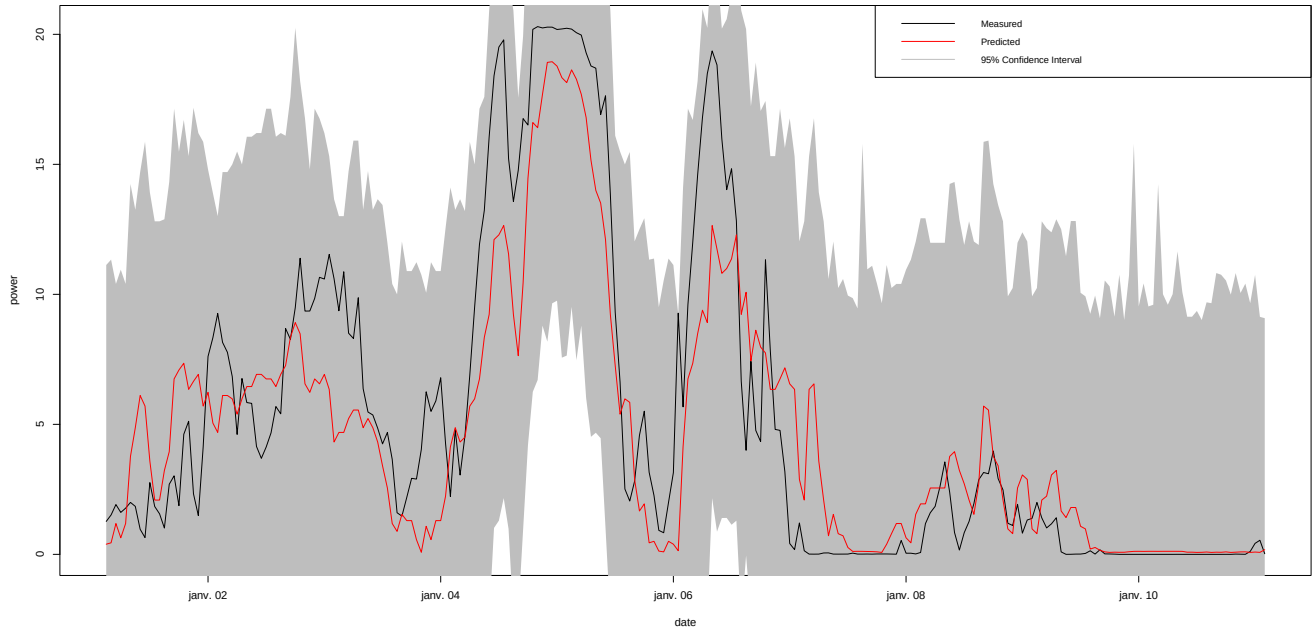

Figure 2.13: *2-hour ahead forecast $\hat{p}^{PC}_{t+2|t}$ during a week using. The gray area is the 95% confidence interval.*

Biais	Variance	MAE	MSE
0.367	8.05	1.98	8.19

Table 2.2: *Metrics : 2-hour ahead Power Curve forecast $\hat{p}^{PC}_{t+2|t}$.*

From the table, we find again there is a small bias.

2.3.3 3-hour ahead power curve forecast

We first consider $\hat{p}^{PC}_{t+3|t}$, the 3-hour ahead power curve forecast with $\{s_{t+3|t}\}$ and f_2. We analyse the residuals of the power curve smoothing.

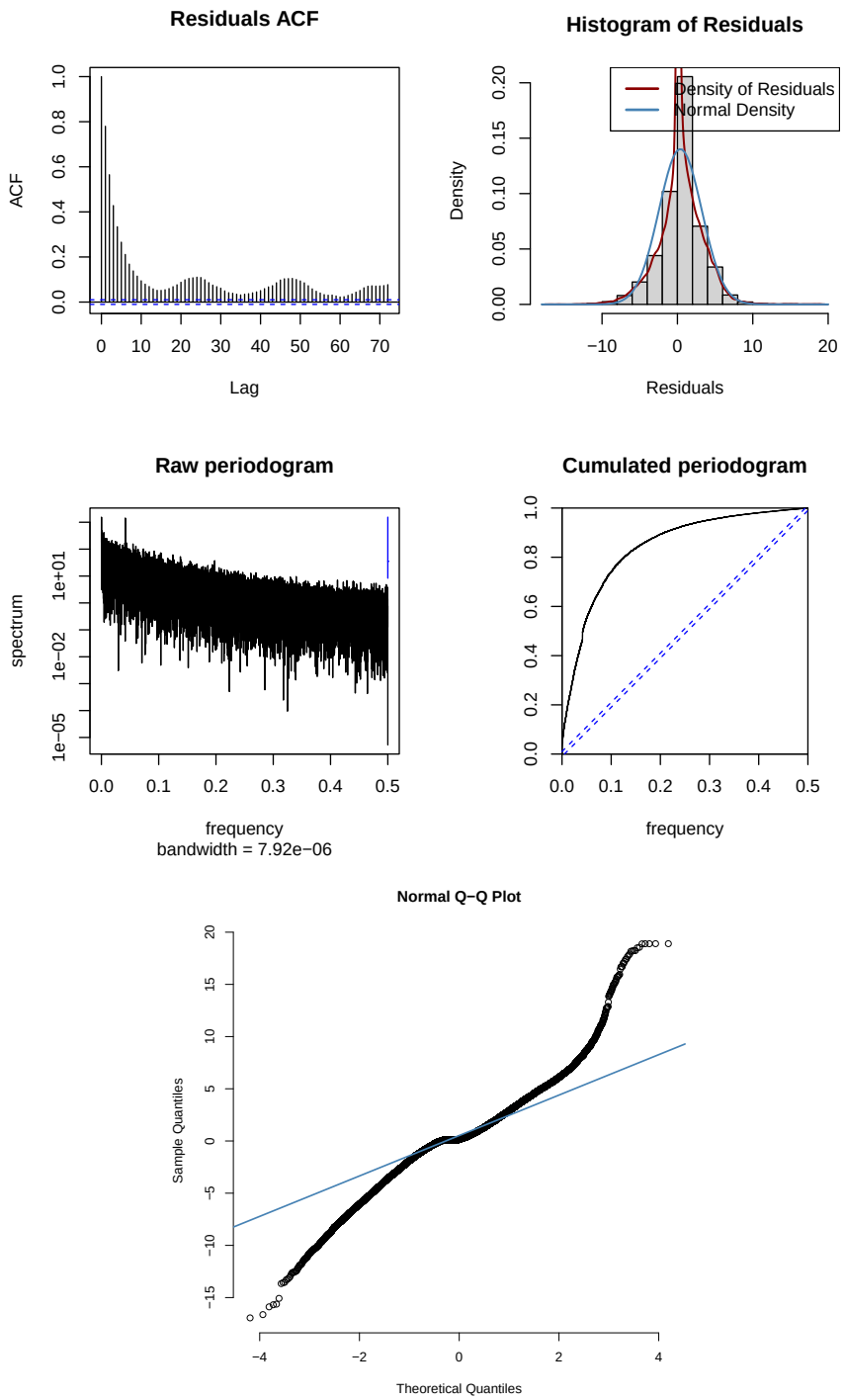

Figure 2.14: *Residuals of f_2 smoothing*

The histogram in Figure 2.14 also displays the kernel-estimated density derived from the residuals, alongside a normal density characterised by the unbiased mean and unbiased variance estimated from the residuals. The figures in fig. 2.14 reveals that the residuals don't follow a white noise pattern. Notably, there are noticeable correlations among the residuals, and the cumulative plot doesn't resemble a straight line. In the same respect, the normal quantile-quantile plot is not close to a straight line. However, it's worth noting that the histogram of errors is slightly uncentered and exhibits a tail quite similar to a

normal distribution.

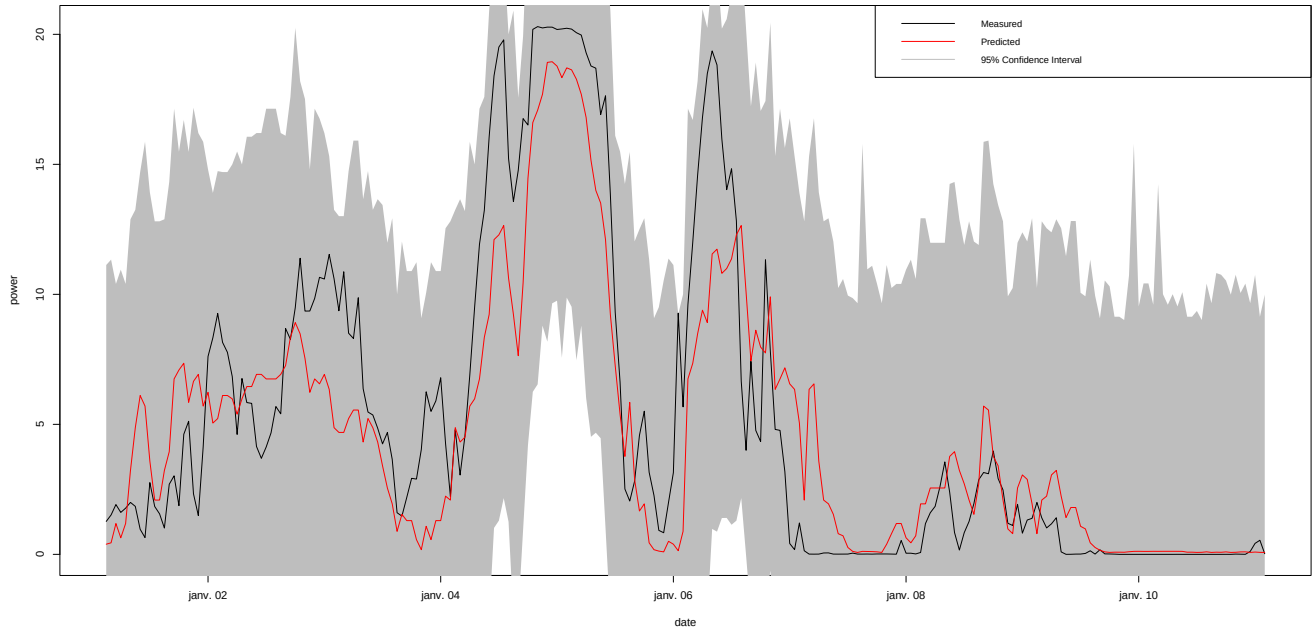

Figure 2.15: *3-hour ahead forecast $\hat{p}^{PC}_{t+3|t}$ during a week using. The gray area is the 95% confidence interval.*

Bias	Variance	MAE	MSE
0.367	8.06	1.99	8.24

Table 2.3: *Metrics : 3-hour ahead Power Curve forecast $\hat{p}^{PC}_{t+3|t}$.*

From the table, we find again there is a small bias.

2.4 Conclusion

From the information provided in table 2.3, 2.2, and 2.1, it is clear that $\{\hat{p}^{PC}_{t+1|t}\}$, $\{\hat{p}^{PC}_{t+2|t}\}$, and $\{\hat{p}^{PC}_{t+3|t}\}$ exhibit very similar performance metrics. This similarity is further corroborated by the observations in fig. 2.10, 2.12, and 2.14, where the residuals demonstrate analogous features, including comparable histograms, quantile-quantile plots, autocorrelation functions (ACF), raw periodograms, and cumulative periodograms.

The observed similarities in $\hat{p}^{PC}_{t+1|t}$, $\hat{p}^{PC}_{t+2|t}$, and $\hat{p}^{PC}_{t+3|t}$ can be attributed to the resemblance in power curves f_1, f_2, and f_3. As depicted in fig. 2.7, 2.8, and 2.9, these power curves exhibit a high degree of similarity. Consequently, the occurrence of similar wind speed forecasts $s_{t+1|t}$, $s_{t+2|t}$, and $s_{t+3|t}$ translates into comparable power predictions $\hat{p}^{PC}_{t+1|t}$, $\hat{p}^{PC}_{t+2|t}$, and $\hat{p}^{PC}_{t+3|t}$. Indeed, the mean square differences between the wind speed forecasts further confirm their similarity. The mean square difference between $s_{t+1|t}$ and $s_{t+2|t}$ is 0.146, between $s_{t+1|t}$ and $s_{t+3|t}$

is 0.303, and between $s_{t+2|t}$ and $s_{t+3|t}$ is 0.157. These values underscore that the three wind speed forecasts provided are remarkably close to each other. Therefore, constructing a simple power curve that solely links wind speed to measured power output appears to be of limited utility. An alternative avenue for exploration could involve incorporating wind direction into the power curve. After dividing the range of wind direction values into sub-intervals, attempts were made to construct a power curve for each sub-interval of wind direction. However, this approach did not yield significant results. Additionally, there was an intention to explore the implementation of an ARIMA model, but the focus shifted to other considerations.

Chapter 3

ARIMA(1,1,1) model

3.1 Principle

First, we shall briefly introduce the ARIMA(1,1,1) model.

3.1.1 ARIMA(1,1,1) model Introduction

The ARIMA(1,1,1) model belongs to the class of Autoregressive Integrated Moving Average models, denoted as ARIMA(p,d,q). In this context, p represents the order of autoregression, d is the degree of differencing, and q signifies the order of the moving average.

- **Autoregressive (AR) Component (p):** The AR component captures the linear relationship between the current observation and its past values. In the ARIMA(1,1,1) model, the order of autoregression (p) is set to 1, indicating a single lag term in the autoregressive equation.
- **Integrated (I) Component (d):** The integration component reflects the differencing needed to make the time series stationary. A stationary series has a constant mean and variance over time. The order of differencing (d) in ARIMA(1,1,1) is 1, suggesting that the first difference of the series is taken, $\nabla p_t = p_t - p_{t-1}$.
- **Moving Average (MA) Component (q):** The MA component captures the short-term fluctuations in the time series. In the ARIMA(1,1,1) model, the order of the moving average (q) is set to 1, indicating a single lag term in the moving average equation.

In a nutshell, ARIMA(p,1,q) can be written as,

$$\hat{p}_t = \mu + \sum_{i=1}^{p} a_i p_{t-i} + \sum_{i=1}^{q} b_t \varepsilon_{t-i} + \varepsilon_t, \quad \varepsilon_t = \hat{p}_t - p_t$$

Theoretical Significance: ARIMA models are widely used in time series analysis for their ability to capture and model temporal dependencies. The ARIMA(1,1,1) configuration specifically allows for the incorporation of one lagged autoregressive term, one differencing to achieve stationarity, and one lagged moving average

term. This balance between past observations, differencing, and short-term fluctuations provides a versatile framework for forecasting. Often, an ARIMA(1,1,1) is enough for modelling a time series.

Applications: ARIMA models are applicable in various domains, including economics, finance, and meteorology. The ARIMA(1,1,1) model, in particular, finds utility in predicting time-dependent phenomena with a focus on short-term dynamics.

Note that we may combine an ARIMA and a GARCH. The GARCH term enable us to model the residuals ε_t,

$$\hat{p}_t = \mu + \sum_{i=1}^{p} a_i p_{t-i} + \sum_{i=1}^{q} b_i \varepsilon_{t-i} + \varepsilon_t,$$
$$\varepsilon_t = \sqrt{\sigma_t} z_t, \quad \sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i} + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2$$

3.1.2 Forecasting with an ARIMA model

To be written

3.2 Residuals Analysis

In this section, we delve into the results derived from the ARIMA(1,1,1) model. Surprisingly, the model exhibits a good fit, and the forecasts display a notable level of accuracy—a somewhat unexpected outcome. Given the nature of the task, I initially did not anticipate such effectiveness, as I had expected significant nonlinearities in power. It outperforms the Power curve model. The ARIMA(1,1,1) model has been fitted using the entire dataset for the year 1999, and the subsequent subsection will provide a detailed discussion on the fitting results. The forecast procedure is as follows: for each time step t from the `toy` column in the data, we fit the model using the past values of p (measured power output) spanning a maximum of 2 years. Subsequently, we generate 1-hour, 2-hour, and 3-hour forecasts. The fitting procedure isn't computationally onerous.

3.2.1 Results of the fit

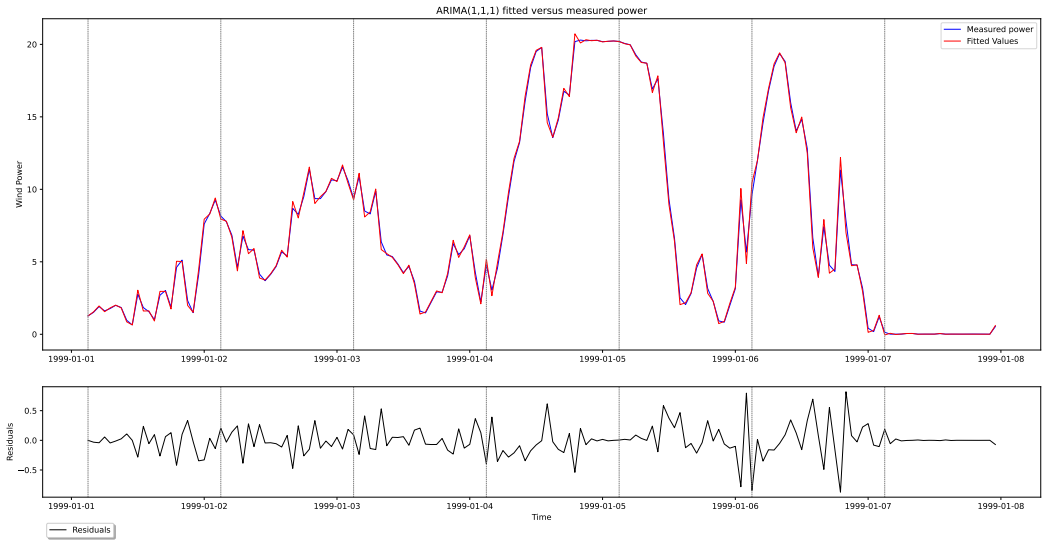

Figure 3.1: *fitted ARIMA(1,1,1) over a week.*

We first analyse how well the ARIMA(1,1,1) model fits the data.

Parameter	Coefficient	Standard Error	$\mathbb{P}$-value	95% Confidence Interval
$ar.L1$	-0.2647	0.022	≈ 0.000	$[-0.307, -0.222]$
$ma.L1$	0.3989	0.021	≈ 0.000	$[0.358, 0.440]$
σ^2	2.6809	0.010	≈ 0.000	$[2.662, 2.700]$

Table 3.1: *ARIMA(1,1,1) Model Parameters.*

The ARIMA(1,1,1) model parameters, as summarised in table 3.1, provide insights into the characteristics of the fitted model. The autoregressive parameter $ar.L1$ is estimated at -0.2647 with a standard error of 0.022, exhibiting a statistically significant negative coefficient ($\mathbb{P} \approx 0.000$). This implies a negative linear relationship with the lag-1 differenced series. The moving average parameter $ma.L1$ is estimated at 0.3989 with a standard error of 0.021, indicating a positive and statistically significant impact on the observed series ($\mathbb{P} \approx 0.000$). The variance of the innovations, denoted as σ^2, is estimated at 2.6809 with a standard error of 0.010, demonstrating a statistically significant impact on the variance of the series ($\mathbb{P} \approx 0.000$). The corresponding 95% confidence intervals provide a range of plausible values for each parameter. These parameter estimates contribute to the model's ability to capture and explain the observed temporal patterns in the dataset.

Test	**Value**	**$\mathbb{P}$-value**
Ljung-Box (Q)	0.00	0.97
Jarque-Bera (JB)	62756.74	0.00
Heteroskedasticity (H)	0.86	0.00
Skew	0.13	-
Kurtosis	9.36	-

Table 3.2: *ARIMA(1,1,1) Model Statistical Tests*

The statistical tests conducted on the ARIMA(1,1,1) model provide valuable insights into its performance. The Ljung-Box test (Q) with a value of 0.00 indicates that there is no significant autocorrelation in the residuals, confirming the adequacy of the model in capturing temporal dependencies. However, the relatively high $\mathbb{P}$-value (0.97) in the Ljung-Box test suggests that further examination may be warranted to ensure the absence of autocorrelation at higher lags. The Jarque-Bera test, with a substantial value of 62756.74 and a $\mathbb{P}$-value of 0.00, signals departure from normality in the residuals. This prompts consideration for potential improvements or alternative modelling approaches to address the non-normality of residuals. The Heteroskedasticity test, represented by a $\mathbb{P}$-value of 0.00, implies the presence of heteroskedasticity in the residuals, indicating that the variance of the residuals is not constant. The skewness value of 0.13 suggests a slight rightward skew in the residuals, and the kurtosis value of 9.36 indicates a degree of peakedness in the distribution.

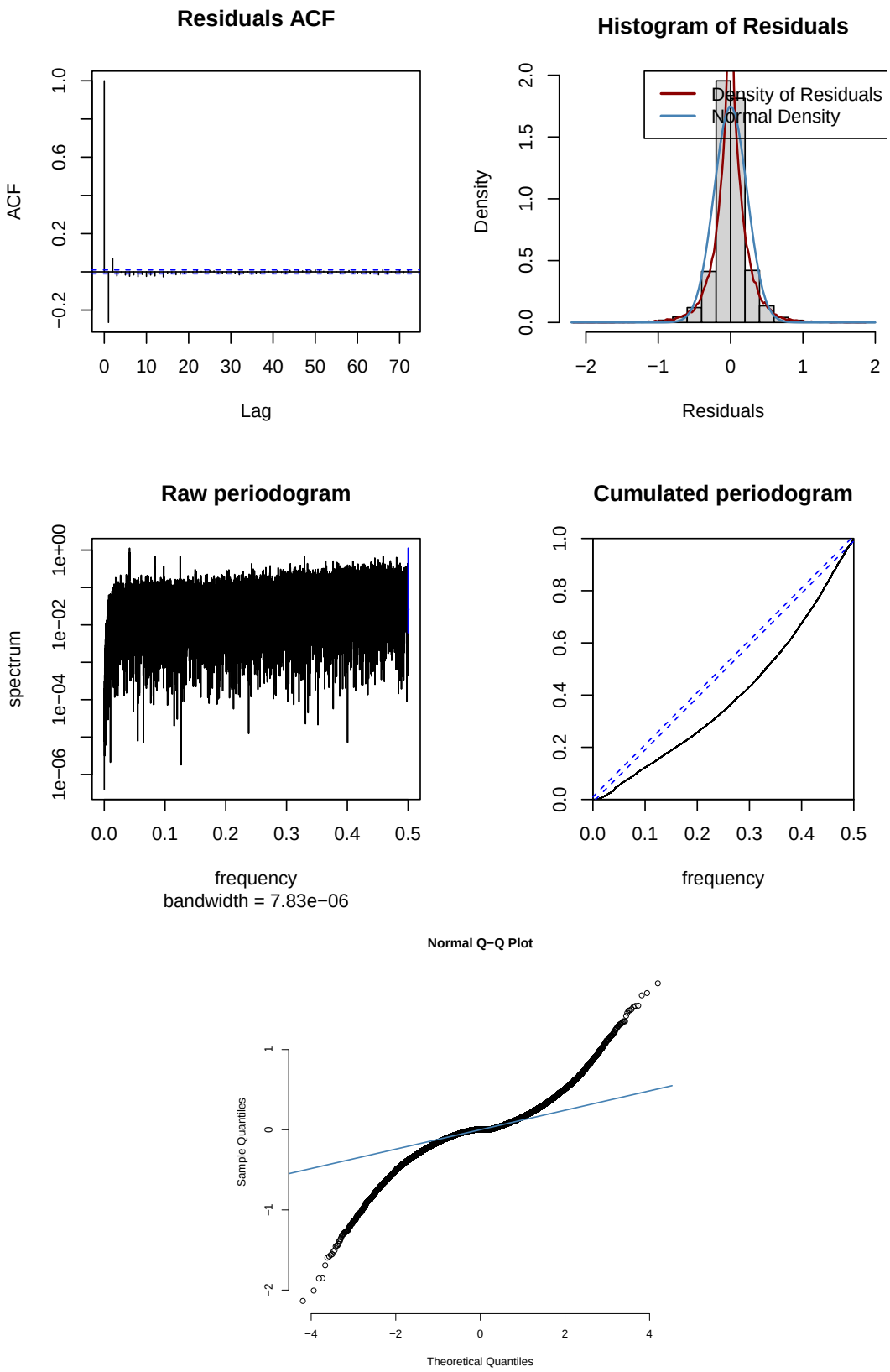

Figure 3.2: *Residuals of the ARIMA(1,1,1) fitting.*

From the observations in fig. 3.2, it appears that the residuals exhibit a behaviour quite close to white noise. The autocorrelation function (ACF) doesn't reveal significant peaks, except at lag 1. It may be worth to invest time in re-

fining the model, specifically addressing the slight deviation observed at lag 1 in the ACF. The cumulative and raw periodograms closely resemble straight lines, suggesting a degree of uniformity. The histogram of residuals gives the impression of being centred and approximately normal, although attention should be directed towards the tails, which appear to be fat-tailed. The quantile-quantile plot, while not perfectly aligned with a straight line, indicates some challenges in modelling quantiles accurately. Ultimately, in comparison to the power curve model residuals analysis (fig. 2.10, 2.12, 2.14), the ARIMA(1,1,1) model emerges as a significantly improved alternative.

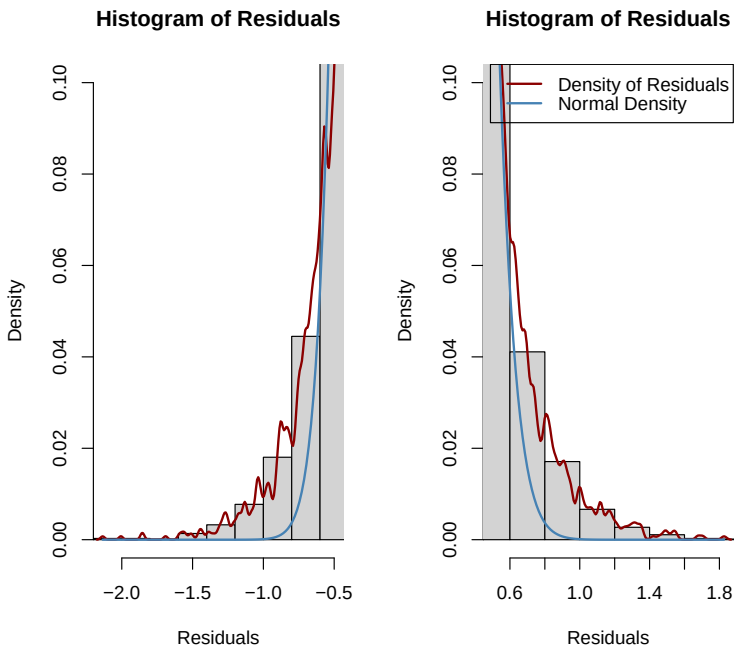

Figure 3.3: *The tail of the residuals of the ARIMA(1,1,1) fitting.*

As shown in fig. 3.3, the residuals exhibit a slightly fatter tail compared to a normal distribution. It is advisable to refine the ARIMA(1,1,1) model to address this fat-tailed behaviour in the residuals.

From the presented results and analyses, several key areas for potential improvement in the ARIMA(1,1,1) model emerge:

1. Autocorrelation at Higher Lags: The Ljung-Box test reveals a significant autocorrelation at lag 1, which is further emphasised by the autocorrelation function (ACF) displaying a peak at this lag. It would be beneficial to explore ways to mitigate this autocorrelation, possibly by introducing additional lag terms or investigating more complex models.

2. Non-Normality of Residuals: The Jarque-Bera test indicates a departure from normality in the residuals, suggesting that the model may not fully capture the distributional characteristics of the data. Considering alternative distributional assumptions or introducing transformations might be explored to address this issue.

3. Heteroskedasticity: The presence of heteroskedasticity in the residuals, as indicated by the Heteroskedasticity test, suggests that the variance of the

residuals is not constant. Adjustments to the model structure or considering alternative variance-stabilising techniques could be examined to address this heteroskedastic behaviour.

4. Fat-Tailed Residuals: The observation of fat-tailed residuals in the histogram suggests potential issues with the tails of the distribution. Investigating alternative error distributions or introducing robust modelling techniques may help in capturing the underlying distribution more accurately.

5. Quantile Modelling Challenges: The quantile-quantile plot reveals challenges in accurately modelling quantiles. Fine-tuning the model structure or considering alternative quantile regression approaches may enhance the model's ability to capture extreme values.

6. Refinement of Lag 1 ACF Deviation: The ACF analysis indicates a slight deviation at lag 1. Further refining the model to address this specific deviation, possibly by adjusting the lag structure, may contribute to improved performance.

3.2.2 Results of the forecasts

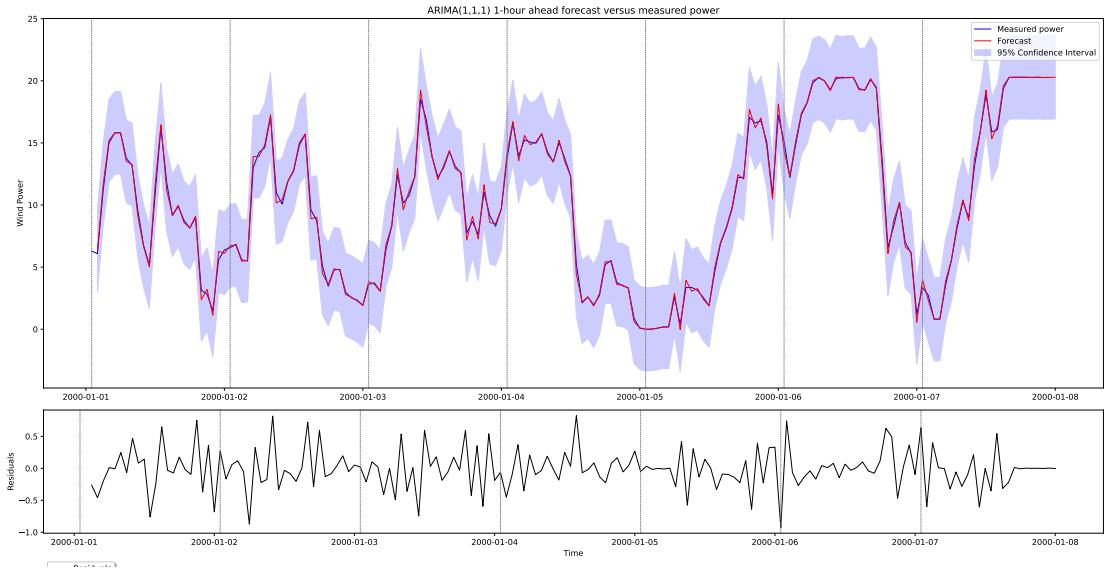

Figure 3.4: *ARIMA(1,1,1) 1-hour ahead forecasting over a week.*

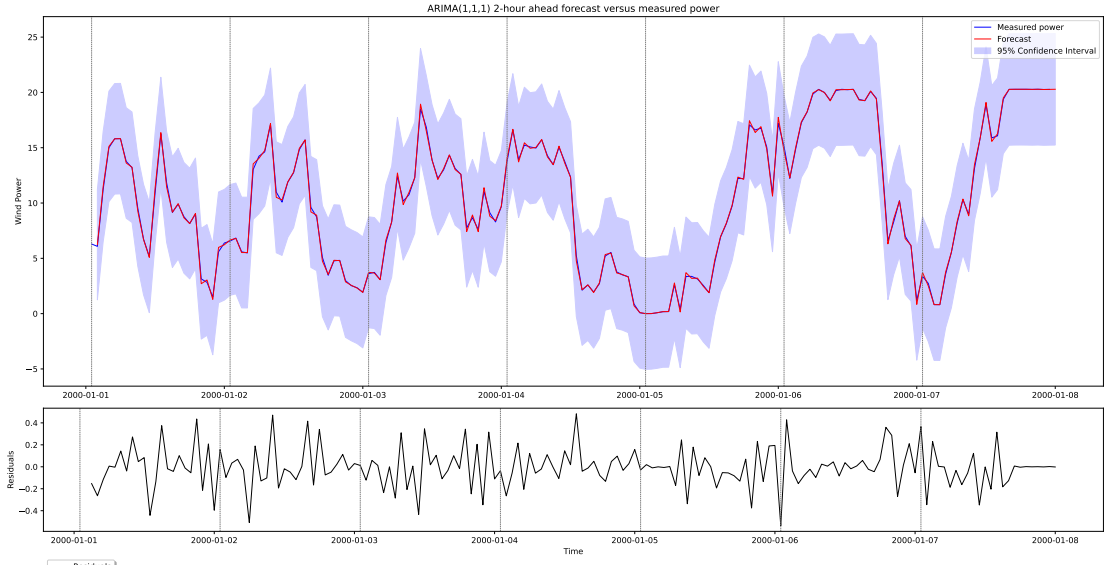

Figure 3.5: *ARIMA(1,1,1) 2-hour ahead forecasting over a week.*

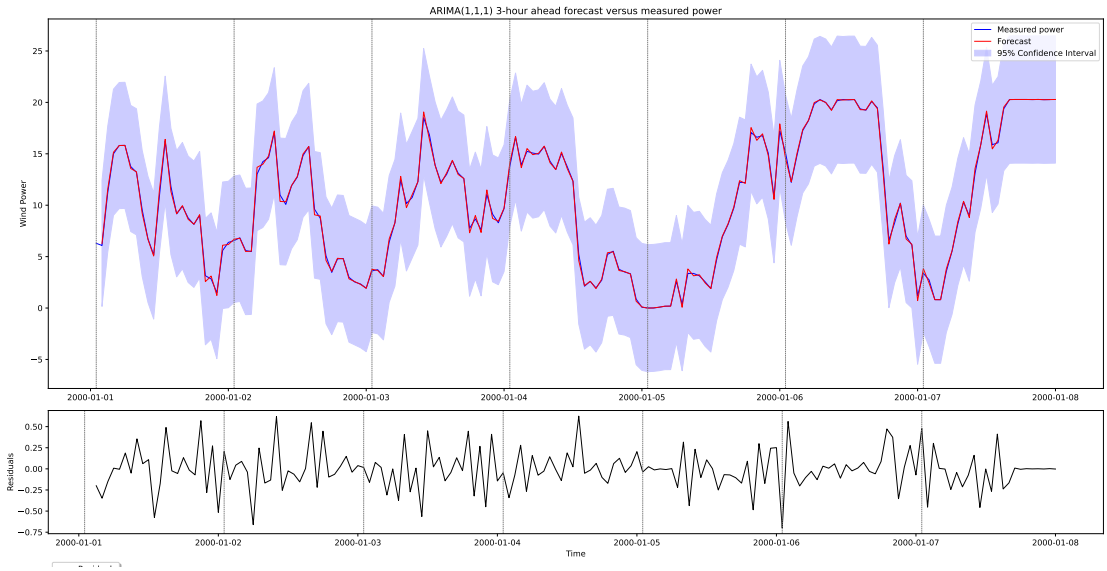

Figure 3.6: *ARIMA(1,1,1) 3-hour ahead forecasting over a week.*

Table 3.3: *Forecasting Metrics of ARIMA(1,1,1) Model*

Forecast Horizon	Bias of Residuals	Variance of Residuals	MAE	MSE
1-hour ahead	−0.008	0.104	0.229	0.104
2-hour ahead	−0.004	0.035	0.133	0.035
3-hour ahead	−0.006	0.059	0.173	0.059

The forecasting metrics presented in table 3.3 provide a comprehensive evaluation of the ARIMA(1,1,1) model's performance at different forecast horizons. Notably, the bias of residuals indicates the model's ability to accurately predict the mean, with values close to zero for all forecast horizons. Examining the variance of residuals, we observe an increase with the length of the forecast horizon, which is not

surprising given that uncertainty tends to escalate as we extend further into the future. The mean absolute error (MAE) and mean squared error (MSE) metrics shed light on the accuracy of predictions, with lower values indicating better performance. Interestingly, the 2-hour ahead forecasts exhibit the lowest MAE and MSE, suggesting enhanced accuracy compared to the 1-hour and 3-hour ahead predictions.

The ARIMA(1,1,1) model significantly outperforms the Power Curve model table 2.3, 2.2, 2.1, as evidenced by various metrics and analyses. The forecasting metrics, as detailed in table 3.3, consistently demonstrate superior performance across multiple forecast horizons compared to the Power Curve model. Notably, the ARIMA model exhibits minimal bias in residuals, effectively capturing the mean of the observed data. Additionally, the lower variance, mean absolute error (MAE), and mean squared error (MSE) values for the ARIMA model emphasize its enhanced accuracy in predicting power output. This substantial outperformance aligns with the earlier comparison of residuals, where the ARIMA model demonstrated more favorable characteristics compared to the Power Curve model. The ARIMA(1,1,1) model's ability to capture and leverage temporal dependencies in the data allows it to surpass the simpler Power Curve approach, showcasing its effectiveness in predicting wind farm power production.

Chapter 4

ARMA(1,1)-GARCH(1,1) model

4.1 Principle

4.2 Residuals Analysis

In this section, we delve into the residue analysis of the ARMA(1,1)-GARCH(1,1) model, aiming to bring the residuals closer to white noise characteristics. The primary objectives include mitigating autocorrelation, especially at lag 1 in the autocorrelation function (ACF), improving the representation of tail behavior and extreme values, and addressing heteroskedasticity. By scrutinising these aspects, we assess the model's efficacy in capturing the intricate dynamics of the underlying time series data.

4.2.1 Results of the fit

First, please note that there are two distinct approaches to fitting an ARMA-GARCH (or ARIMA-GARCH) model.

1. The first approach involves a sequential procedure, where the ARMA component is fitted to the original time series, and subsequently, the GARCH model is applied to the residuals obtained from the ARMA fit. This sequential fitting allows for a step-by-step refinement, addressing the autoregressive and moving average components initially, followed by the modelling of conditional volatility through the GARCH process.

2. Contrasting, the second approach employs a simultaneous fitting strategy. In this method, both the ARMA and GARCH components are fitted concurrently to the original time series data. This simultaneous fitting takes into account the interdependencies between the autoregressive, moving average, and conditional volatility components, allowing for a comprehensive modelling of both mean and volatility dynamics in a unified framework.

We chose the second approach of simultaneous fitting for its often superior performance in capturing joint mean and volatility dynamics. While this method tends to yield better results, it's essential to note that it might not be readily available in standard `R` or `Python` libraries, which typically favour the sequential

fitting strategy.

In comparison to the previously studied ARIMA(1,1,1) model, the ARMA(1,1)-GARCH(1,1) model introduces a more complex structure that incorporates autoregressive and moving average components for the mean equation, along with a GARCH component for modelling volatility/residuals. Table table 4.1 provides estimates for the model parameters.

Parameter	**Estimate**	**Std. Error**	**t value**	**$\mathbb{P}$-value**
μ	0.998070	0.179644	5.5558	0
$ar1$	0.937255	0.003199	292.9534	0
$ma1$	0.230768	0.013014	17.7316	0
ω	0.044625	0.004473	9.9772	0
$\alpha 1$	0.248595	0.009582	25.9439	0
$\beta 1$	0.750405	0.008681	86.4416	0

Table 4.1: *ARMA(1,1)-GARCH(1,1) Model Parameters.*

The estimated intercept term, μ, is found to be 0.998070 with a standard error of 0.179644 and a t value of 5.5558 (with a $\mathbb{P}$ value around 0). The autoregressive coefficient, $ar1$, is estimated at 0.937255 with a very small standard error, resulting in a high t value of 292.9534 (again, $\mathbb{P}$-value is null). Similarly, the moving average coefficient, $ma1$, is estimated at 0.230768 with a t value of 17.7316 ($\mathbb{P}$-value = 0). The GARCH parameters, ω, $\alpha 1$, and $\beta 1$, represent the intercept, the persistence of volatility, and the persistence of past shocks, respectively. Each of these parameters shows statistically significant estimates, indicating their importance in capturing the conditional volatility dynamics. The results suggest that the ARMA(1,1)-GARCH(1,1) model provides a good fit to the data, accounting for both mean and volatility components.

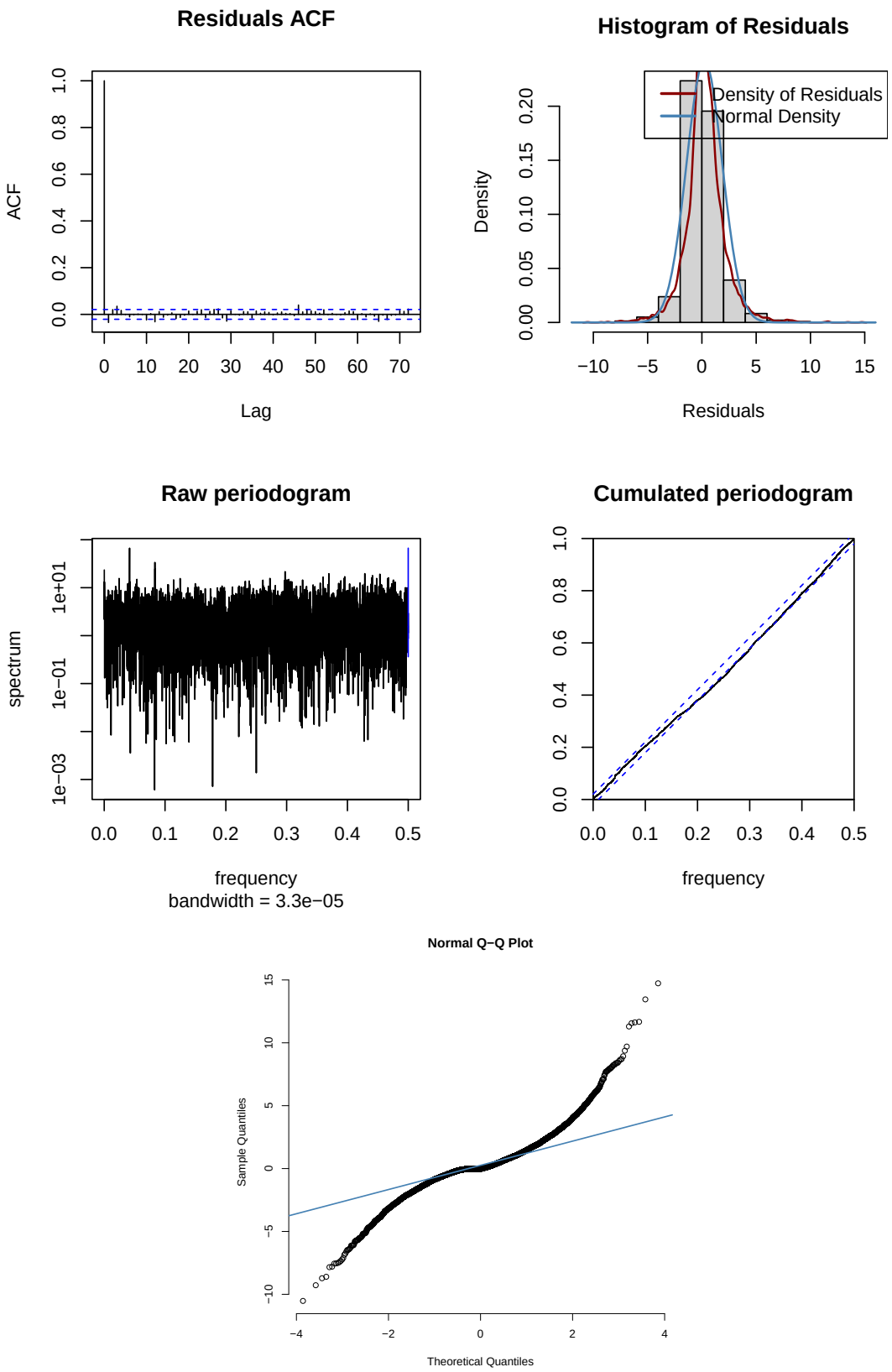

Figure 4.1: *Residuals of the ARMA(1,1)-GARCH(1,1) fitting.*

Referring to fig. 4.1, the residuals from the ARMA(1,1)-GARCH(1,1) model exhibit behaviour closer to white noise. The autocorrelation function (ACF) plot is notably satisfactory, displaying no significant correlations. The histogram of resid-

uals is centred and closely resembles a Gaussian distribution; however, there may be a slight deviation in the tails, suggesting potential fat-tailedness. Both the raw and cumulative periodograms show excellent agreement with those of white noise. Despite these favourable features, the quantiles do not align well with the theoretical distribution, as evidenced by the quantile-quantile plot, which deviates from a straight line. This suggests that while the model captures many aspects of the data dynamics, there may be room for improvement in modelling extreme quantiles. Overall, the ARMA(1,1)-GARCH(1,1) model demonstrates effective fitting with residuals displaying characteristics akin to white noise, though further refinement may be considered for a more accurate representation of extreme quantiles.

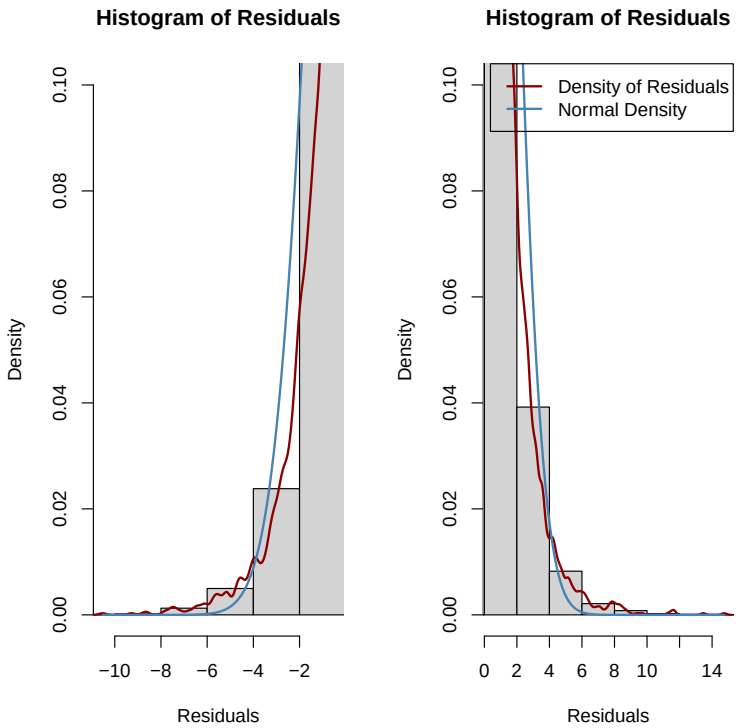

Figure 4.2: *Tail of the residuals of the ARMA(1,1)-GARCH(1,1) fitting.*

Comparing the tail behaviour of the ARIMA(1,1,1) model residuals fitting in fig. 3.3 with that of the ARMA(1,1)-GARCH(1,1) model in fig. 4.1, it is evident that the ARMA(1,1)-GARCH(1,1) model exhibits significantly less tailing. The distribution of residuals from the ARMA(1,1)-GARCH(1,1) model appears to be more concentrated around the mean, indicating a better fit in capturing extreme observations. This observation suggests that the incorporation of the GARCH component in the model has effectively reduced the fat-tailed characteristics present in the ARIMA(1,1,1) model residuals.

The residuals analysis of the ARMA(1,1)-GARCH(1,1) model outperforms that of the ARIMA(1,1,1) model, as observed in the respective figures (fig. 4.1 and fig. 3.2). For the ARIMA(1,1,1) model, the ACF plot shows reduced correlation at lag 1, indicative of a better fit. The histogram representation displays a more accurate portrayal of extreme values, and both the cumulated and raw periodograms

align closely with a straight line. Conversely, the ARMA(1,1)-GARCH(1,1) model exhibits even stronger performance. Its residuals show significantly less tailing, as evidenced by a more centralised distribution around the mean. The ACF plot of the ARMA(1,1)-GARCH(1,1) model displays no significant correlation, reflecting a robust fitting. While the QQ plot of the ARMA(1,1)-GARCH(1,1) model slightly outperforms that of the ARIMA(1,1,1) model, collectively, these observations underscore the superior fit of the ARMA(1,1)-GARCH(1,1) model in capturing the characteristics of the underlying data.

4.2.2 Results of the forecasts

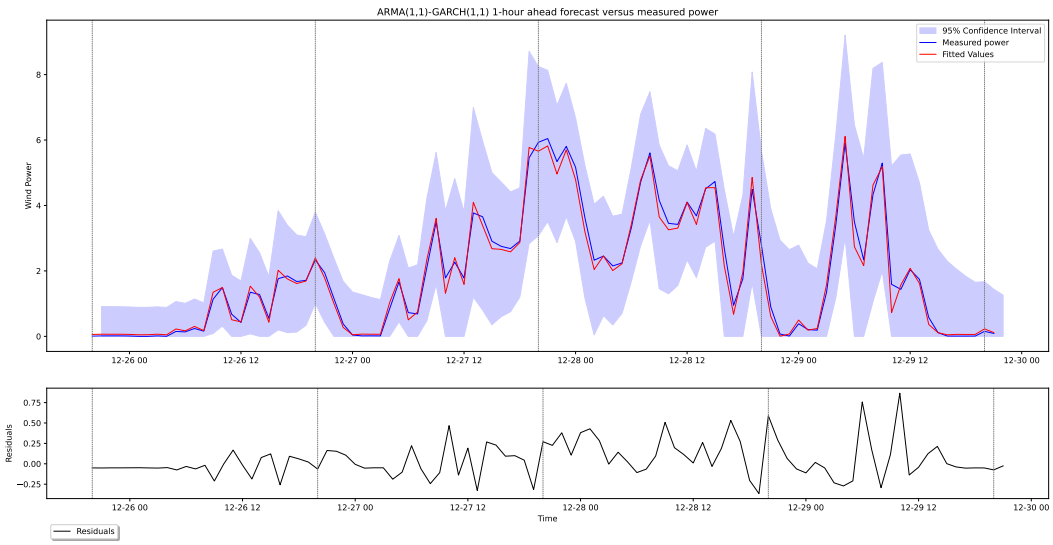

Figure 4.3: *ARMA(1,1)-GARCH(1,1) 1-hour ahead forecasting over a week.*

Table 4.2: *Forecasting Metrics of ARMA(1,1)-GARCH(1,1) Model*

Forecast Horizon	Bias of Residuals	Variance of Residuals	MAE	MSE
1-hour ahead	0.048	0.041	0.17	0.051
2-hour ahead	0.045	0.041	0.13	0.021
3-hour ahead	0.049	0.042	0.15	0.041

In comparing the forecasting metrics between the ARMA(1,1)-GARCH(1,1) model and the ARIMA(1,1,1) model, as presented in Tables table 4.2 and table 3.3 respectively, notable differences emerge. For the 1-hour ahead forecast, the ARMA(1,1)-GARCH(1,1) model demonstrates a positive bias of 0.048 compared to the ARIMA(1,1,1) model's negative bias of -0.008. Additionally, the variance of residuals is lower for the ARMA(1,1)-GARCH(1,1) model (0.041) than for the ARIMA(1,1,1) model (0.104), indicating a more precise prediction. The mean absolute error (MAE) for the ARMA(1,1)-GARCH(1,1) model is 0.17, whereas the ARIMA(1,1,1) model shows a higher MAE of 0.229. Similarly, the mean squared error (MSE) is lower for the ARMA(1,1)-GARCH(1,1) model (0.051) than for the ARIMA(1,1,1) model

(0.104) in the 1-hour ahead forecast.

For the 2-hour and 3-hour ahead forecasts, the pattern persists. The ARMA(1,1)-GARCH(1,1) model consistently exhibits lower bias, variance of residuals, MAE, and MSE values, reflecting its superior performance in capturing the underlying data dynamics. The results suggest that the inclusion of the GARCH component in the ARMA(1,1)-GARCH(1,1) model contributes to more accurate and efficient forecasting compared to the ARIMA(1,1,1) model. These findings hold practical implications for applications where accurate short-term predictions are crucial, such as financial forecasting or risk management.

In light of the non-normality detected in the residuals of the ARIMA(1,1,1) model, as revealed by the Jarque-Bera test, there is an indication of a departure from normality in the residuals. This suggests that the model might not fully capture the distributional characteristics of the data. An alternative avenue for addressing this issue could involve exploring different distributional assumptions within the GARCH framework. For instance, transitioning from a normal distribution to a distribution with heavier tails, such as the Student's t-distribution, could be considered. Although this path wasn't explicitly explored in the current analysis, it remains a plausible option for refining the model's ability to capture the underlying distribution of the data.

Chapter 5

ARIMA(1,1,1)-GARCH(1,1) model

In this upcoming chapter, we embark on an exploration of the ARIMA(1,1,1)-GARCH(1,1) model, aiming to scrutinise the relevance of incorporating differencing in the time series. Our objective is to assess whether the differentiation of the series proves pertinent. We anticipate achieving a fitting performance comparable to that of an ARMA(1,1)-GARCH(1,1) model while simultaneously improving forecasting capabilities. By delving into the intricacies of this hybrid model, we aim to discern the impact of differencing on capturing the underlying dynamics of the time series, ultimately seeking a balance between model complexity and forecasting precision.

5.1 Principle

5.2 Residuals Analysis

In the forthcoming section on "Residual Analysis" for the ARIMA(1,1,1)-GARCH(1,1) model, our focus is to conduct a thorough examination of the model's residuals. Our primary objective is to achieve a fitting performance on par with the ARMA(1,1)-GARCH(1,1) model while simultaneously enhancing forecasting precision. By scrutinising the residuals, we aim to gauge the effectiveness of the ARIMA differencing component in capturing the underlying dynamics of the time series. The overarching goal is to ensure that the model strikes a balance, delivering a fitting quality comparable to its ARMA counterpart and exhibiting superior forecasting capabilities.

5.2.1 Results of the fit

In line with our previous chapter, the upcoming model fitting adopts the simultaneous fitting approach. In contrast to the previously examined ARIMA(1,1,1) model, the ARIMA-GARCH model introduces a more intricate structure. It incorporates autoregressive and moving average components for the mean equation, coupled with a GARCH component to model volatility/residuals. The estimated parameters for this model are detailed in Table table 5.1.

Parameter	Estimate	Std. Error	t value	$\mathbb{P}$-value
μ	0.99806958	0.179644274	5.555811	2.763262×10^{-8}
$ar1$	0.93725536	0.003199333	292.953401	$0.000000e+00$
$ma1$	0.23076779	0.013014498	17.731593	$0.000000e+00$
ω	0.04462543	0.004472728	9.977228	$0.000000e+00$
$\alpha 1$	0.24859528	0.009582046	25.943863	$0.000000e+00$
$\beta 1$	0.75040472	0.008681060	86.441597	$0.000000e+00$

Table 5.1: *ARIMA-GARCH(1,1) Model Parameters.*

The estimated intercept term, μ, is found to be 0.99806958 with a standard error of 0.179644274 and a t value of 5.555811 (with a $\mathbb{P}$-value of 2.763262×10^{-8}). The autoregressive coefficient, $ar1$, is estimated at 0.93725536 with a very small standard error, resulting in a high t value of 292.953401 ($\mathbb{P}$ value is null). Similarly, the moving average coefficient, $ma1$, is estimated at 0.23076779 with a t value of 17.731593 ($\mathbb{P}$-value 0). The GARCH parameters, ω, $\alpha 1$, and $\beta 1$, represent the intercept, the persistence of volatility, and the persistence of past shocks, respectively. Each of these parameters shows statistically significant estimates, indicating their importance in capturing the conditional volatility dynamics. The results suggest that the ARIMA-GARCH(1,1) model provides a good fit to the data, accounting for both mean and volatility components.

The ARIMA-GARCH(1,1) model, introduced for its ability to incorporate differencing, exhibits parameter estimates comparable to the ARMA-GARCH(1,1) model, highlighting its effectiveness in capturing both mean and volatility dynamics.

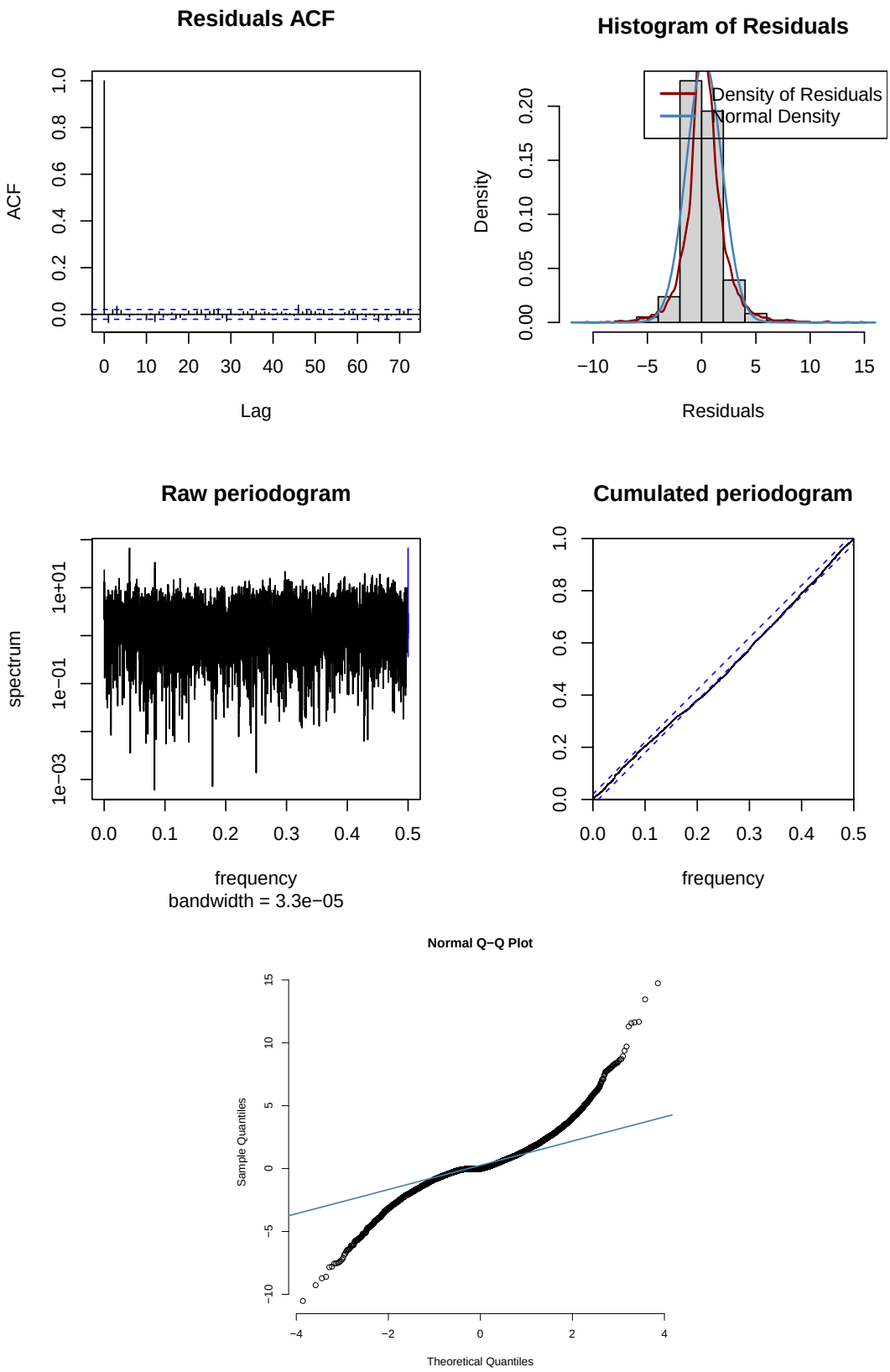

Figure 5.1: *Residuals of the ARIMA(1,1,1)-GARCH(1,1) fitting.*

Referring to fig. 5.1, similar patterns emerge in the residual analysis of both the ARIMA(1,1,1)-GARCH(1,1) and ARMA(1,1)-GARCH(1,1) models in fig. 4.1. Notably, the ARIMA(1,1,1)-GARCH(1,1) residuals show a slight improvement in

the cumulated periodogram, aligning more closely with a straight line compared to its ARMA counterpart.

In both cases, the residuals exhibit behaviour reminiscent of white noise. The ACF plot for each model is notably satisfactory, revealing no significant correlations. The histograms of residuals are centred and closely resemble Gaussian distributions, with potential deviations in the tails hinting at slight fat-tailedness. Both the raw and cumulative periodograms demonstrate excellent agreement with those of white noise.

However, the quantiles, as observed in the quantile-quantile plot, do not align well with the theoretical distribution for either model. This departure from a straight line indicates room for improvement in modelling extreme quantiles. Overall, both the ARIMA(1,1,1)-GARCH(1,1) and ARMA(1,1)-GARCH(1,1) models exhibit effective fitting, with residuals displaying characteristics akin to white noise. Further refinement may be considered to enhance the representation of extreme quantiles.

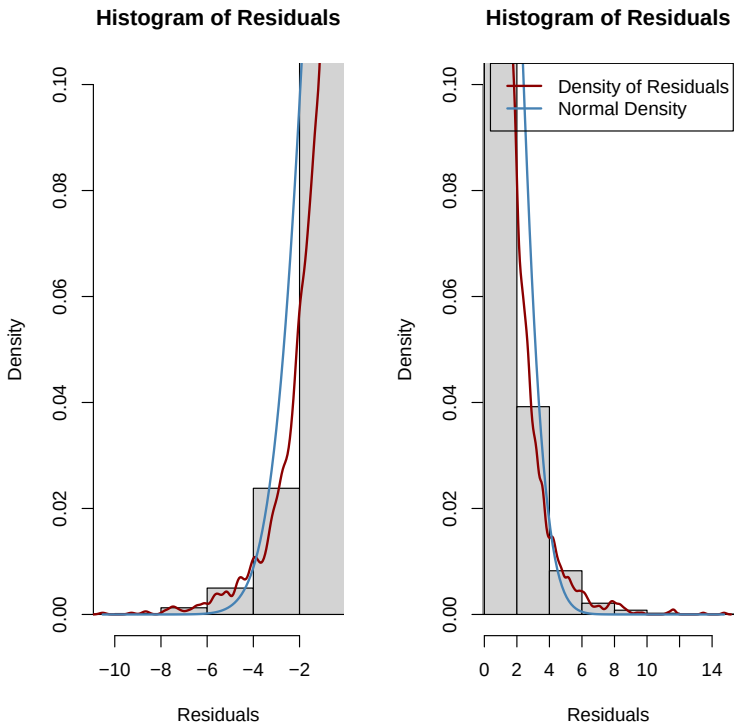

Figure 5.2: *Tail of the residuals of the ARIMA(1,1,1)-GARCH(1,1) fitting.*

Comparing the tail behavior of the ARIMA(1,1,1)-GARCH(1,1) model residuals in fig. 5.2 with the ARMA(1,1)-GARCH(1,1) model in fig. 4.2, no distinct difference is observed. The ARIMA(1,1,1)-GARCH(1,1) model does not exhibit a significant improvement compared to the ARMA(1,1)-GARCH(1,1) model.

In conclusion, the ARIMA(1,1,1)-GARCH(1,1) model does not show a significant improvement over the ARMA(1,1)-GARCH(1,1) model. The only slightly notable difference lies in the periodograms, where the frequencies in the ARIMA(1,1,1)-

GARCH(1,1) model more closely resemble those of white noise. The objective of achieving a fit as good as that of the ARMA(1,1)-GARCH(1,1) model is realised.

5.2.2 Results of the forecasts

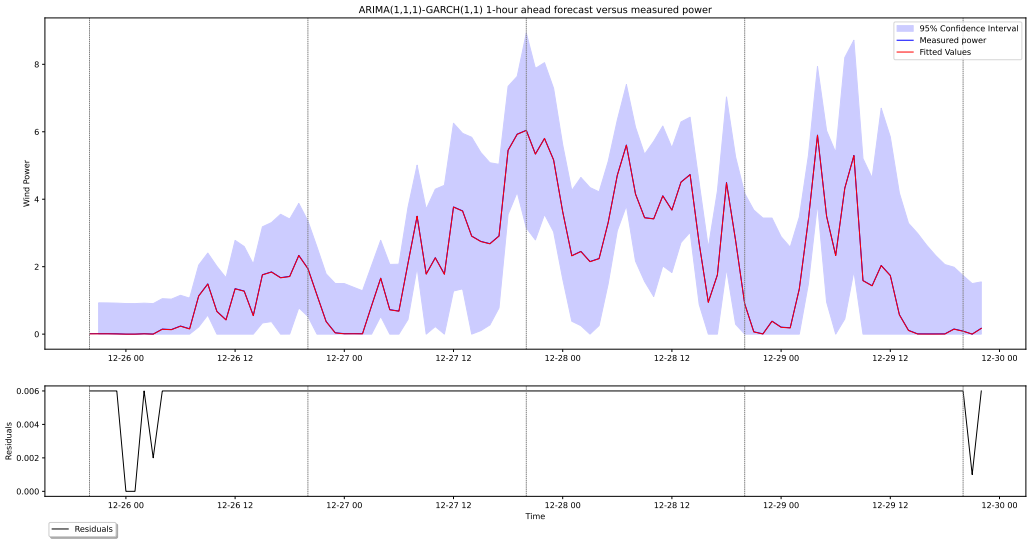

Figure 5.3: *ARIMA(1,1,1)-GARCH(1,1) 1-hour ahead forecasting over a week.*

Table 5.2: *Forecasting Metrics of ARIMA(1,1,1)-GARCH(1,1) vs ARMA(1,1)-GARCH(1,1) Models*

Mean model	Vol. model	Horizon	Bias	Variance	MAE	MSE
ARMA(1,1)	GARCH(1,1)	1-hour ahead	0.048	0.041	0.17	0.051
ARIMA(1,1,1)	GARCH(1,1)	1-hour ahead	0.006	0.001	0.006	0.001
ARMA(1,1)	GARCH(1,1)	2-hour ahead	0.045	0.041	0.13	0.021
ARIMA(1,1,1)	GARCH(1,1)	2-hour ahead	0.006	0.001	0.006	0.001
ARMA(1,1)	GARCH(1,1)	3-hour ahead	0.049	0.042	0.15	0.041
ARIMA(1,1,1)	GARCH(1,1)	3-hour ahead	0.006	0.002	0.008	0.002

The forecasting metrics of the ARIMA(1,1,1)-GARCH(1,1) model, encompassing bias of residuals, variance of residuals, mean absolute error, and mean squared error, reveal substantially lower values compared to those of the ARMA(1,1)-GARCH(1,1) model (see table 5.2). This discrepancy suggests a superior forecasting accuracy of the ARIMA(1,1,1)-GARCH(1,1) model over its ARMA(1,1)-GARCH(1,1) counterpart, demonstrating enhanced predictive capabilities. The improvement in the forecasting accuracy of the ARIMA(1,1,1)-GARCH(1,1) model, though anticipated, has surpassed expectations, as validated through meticulous code verification. These results underscore the success in achieving significantly enhanced forecast performance.

Chapter 6

Conclusion

In conclusion, the adoption of an ARIMA(1,1,1)-GARCH(1,1) model, incorporating differencing in the data, has successfully yielded improvements over the ARMA(1,1)-GARCH(1,1) model. The objective of achieving a fit comparable to that of ARMA-GARCH while enhancing forecast accuracy has been achieved. The ARIMA(1,1,1)-GARCH(1,1) model demonstrates not only a commendable fit but also superior forecasting capabilities. This outcome reinforces the efficacy of employing a dual ARIMA-GARCH approach, highlighting its potential for optimizing predictive performance.

By the way, as the ARIMA(1,1,1)-GARCH(1,1) with f simultaneous fitting is not directly available in `R` or `Python`, I had first implemented the sequential fitting approach. Here is some plots to visualise the difference,

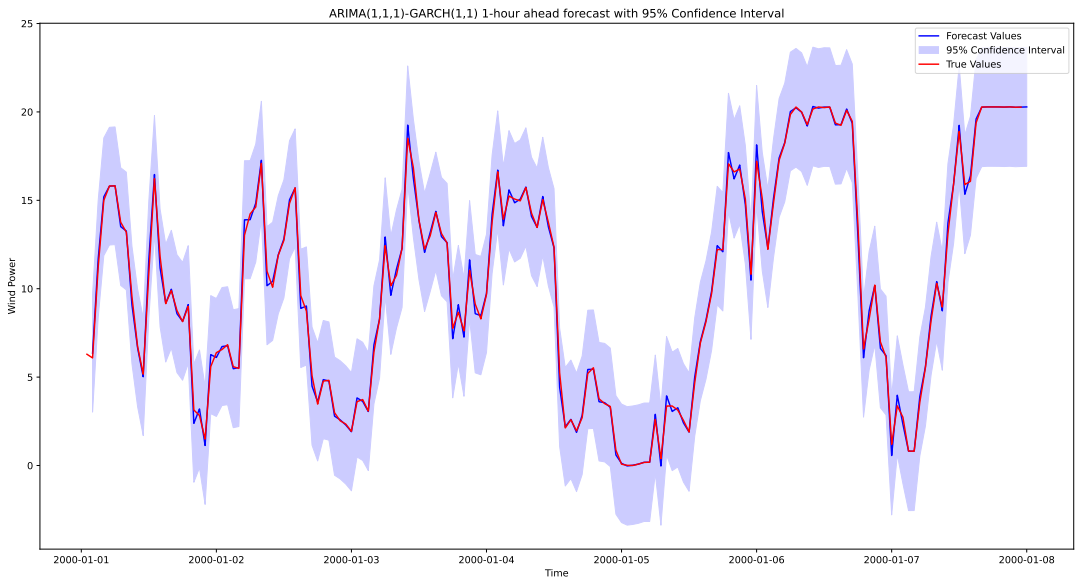

Figure 6.1: *ARIMA(1,1,1)-GARCh(1,1) 1-hour ahead forecast using the sequential fitting approach*

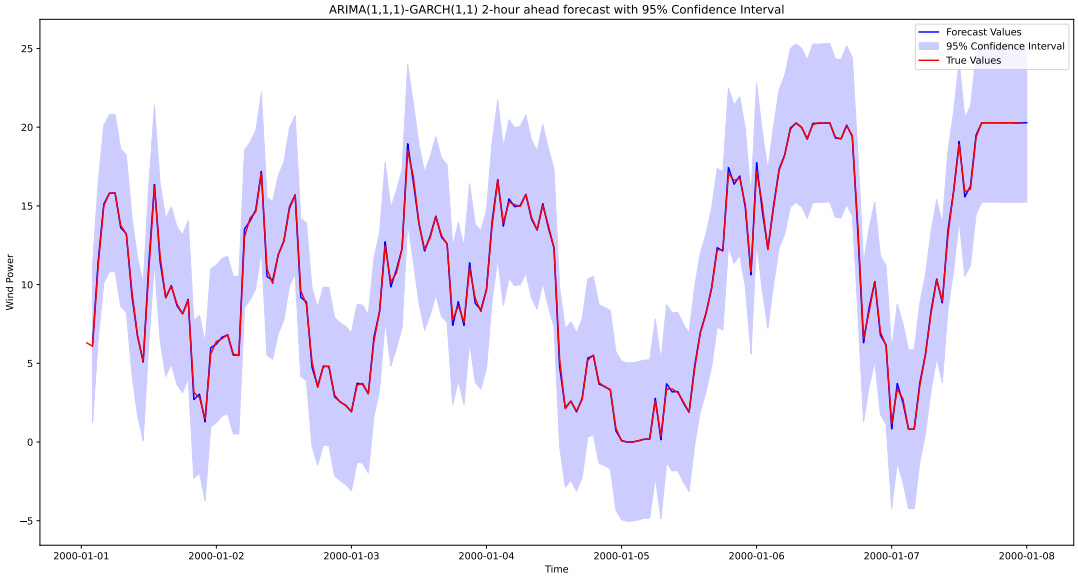

Figure 6.2: *ARIMA(1,1,1)-GARCh(1,1) 2-hour ahead forecast using the sequential fitting approach*

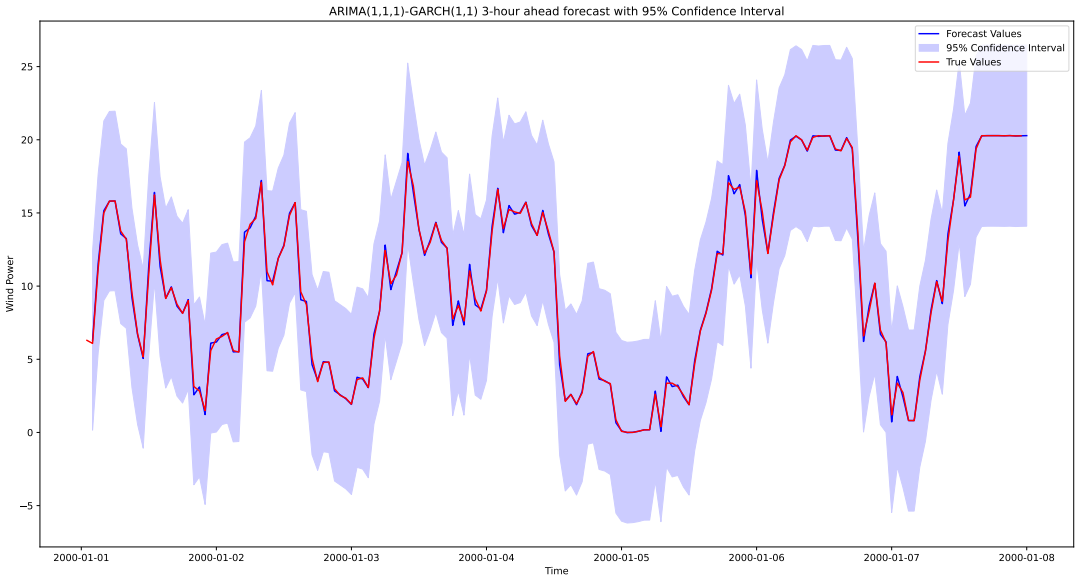

Figure 6.3: *ARIMA(1,1,1)-GARCh(1,1) 3-hour ahead forecast using the sequential fitting approach*