Application of Univariate Time Series Techniques to High Frequency Electricity Data

Andrew Boomer

June 30, 2020

Abstract

This paper is an extension of a previous project using the same dataset from Boomer and Kim (2020), where the model was not valid due to serial correlation of the residuals. This research will survey the existing time series techniques applied to high frequency data, with the goal of developing a valid specification for this previous work.

The dataset used in the previous model was compiled from the European Network of Transmission System Operators (ENSTO-e). The dataset included several relevant electricity market variables, and was aggregated into a panel of 20 regions, spanning 5 years of hourly data. This panel specification will be deconstructed, allowing analysis of each region as a univariate time series. This paper first provides an overview of the existing literature, detailing the history and progression of techniques. This summary then provides the basis for identification and understanding of the processes governing the data.

To finish, these techniques are applied using a suite of time series models, including both ARIMA and Structural. The results, as well as advantages and disadvantages of each approach are compared.

1 Introduction

1.1 Motivation

This research presents a time series treatment of several univariate high frequency datasets. The primary focus of the paper is a review of the existing modeling techniques and theories commonly used on this type of data, followed by an application of some of these models.

In the previous preliminary research on this subject Boomer and Kim (2020), the final model presented serial correlation in the residuals. This paper is motivated towards the techniques, processes, and models that can lead to a valid specification of this previous research. The general approach for accomplishing this will involve two methods. The first is to deconstruct the 20 region panel into a collection of univariate time series. This is motivated by an initial exploration of the data, which revealed heterogeneity of time series patterns for the dependent and independent variables. The heterogeneous phenomena found between regions included but was not limited to: non-parallel trends and differing numbers and frequencies of seasonal cycles. The second method will be to use a two step approach, taking the residuals of the previous model as the time series. This approach is explained in more detail in the Models Section.

This research is therefore not simply an extension of the initial research in Boomer and Kim (2020). While the first project was motivated by interest in the underlying economic hypothesis and its implications, this research will be driven by applying econometric techniques for high frequency time series analysis in a way that develops a tractable and valid model. For this goal, however, it will be necessary to develop a deeper understanding of the underlying behaviors and patterns present in both the dependent and independent variables. To this end of understanding the data and model building, there will first be a review of the current literature covering high frequency time series modeling.

1.2 Electricity Market Background

European electricity markets work to balance the demand and supply of power through the optimization of an auction market. The markets are cleared at the quantity demanded by electricty consumers, and the clearing price can be either uniform or pay-as-you-bid. From years in advance to minutes in advance, various types of consecutive electricity markets exist in order to correct for projection errors in the previous period and ultimately meet the flow time load. Balancing markets exist in order to provide a safety net for the projection error and extreme events[1]. The dependent variable in the initial model from Boomer and Kim (2020) was Activated Balancing Volume, which is the amount of balancing volume realized, as opposed to contracted or planned for. The residuals of the regression on this variable are used as the univariate time series in the current paper.

This research continues to use the dataset compiled by Boomer and Kim (2020). This data was gathered from the ENTSO-E's web-based platform dedicated to disclosure of European electricity market information. After data cleaning and null interpolation, we obtained hourly data spanning 5 years (43,800 records), from January 2015 through December 2019, per region. There were 20 regions with adequate data, and a subset of those regions will be modeled individually in this analysis. Further details regarding the aggregation and cleaning of this dataset are discussed in Boomer and Kim (2020).

[1] A more detailed explanation of these markets can be found in Appendix A

2 General Electricity Data Characteristics

There are several characteristics that make electricity data, specifically high frequency electricity data, unique in comparison to many other time series profiles [2]. One of these is the existence of more than one cycle of differing lengths, such as daily, weekly, monthly, and/or yearly. Sometimes, a smaller cycle can be thought of as a sort of subset of a larger cycle. For example, the roughly 12 hour cycle present in solar generation data could be thought of as contained within the larger 24 hour sun cycle. Framing in this way changes the appropriate approach as modeling these cycles separately does not necessarily make sense.

Additionally, each of these cycles can vary by translation within the time series, for example, the daily cycles present in electricity price and load data differ according to weekday/weekend/holiday/summer/winter Gould et al. (2008). Electricity data also exhibits strong mean reversion, and the class of autoregressive models can effectively reproduce this behavior [3] Cuaresma et al. (2004). This mean reversion can occur on a global or local level (i.e. reversion to a grouped mean).

Huisman et al. (2007) discusses a particularly interesting aspect of electricity data. They argue that not all electricity data fufills the assumption of the information set updating after each time period. For example, in the Day-Ahead electricity markets, prices for all 24 hours of the next day are simultaneously determined. That is, for Day Ahead electricity prices, the information set that determined the price at Hour 10 is the same that determined the price in the next period, Hour 11 of the same day. For this reason, Huisman et al. (2007) models Day Ahead electricity prices as a panel, where the regions are hours within the day, specifying cross sectional rather than serial correlation within a day. In the context of the present research, the dependent variable, however, does have an information set that updates after each period ENTSO-e (2018)

The literature I have found focuses solely on forecasting, a more amibitious task than mine. In the context of forecasting, extreme outliers are very important for metrics such as mean-squared error, but, if random, are less significant in contributing to serial correlation. For electricity price data, extreme outlying values are known as price spikes. These can be caused by a variety of market factors, including the non-storable nature of electricity, and the strict market clearing requirements Cuaresma et al. (2004). For the purposes of this paper, I will assume that these price spikes appear at random intervals. While these would be important to include in a forecasting model, the assumption of randomness implies no impact on serial correlation.

3 Review of Modeling Techniques

3.1 ARIMA Models

Given the goal of removing serial correlation in a time series dataset, a logic place to start is with the class of models known as ARIMA models (Autoregressive Integrated Moving Average). While all time series models involve using past data, ARIMA models are one of the first to formalize the estimation and evaluation process. From 1926 to 1938, Yule, Slutsky, and Wold put together an ARMA framework to model stationary time series processes Makridakis and

[2]Specifically financial data where much of the high frequency time series literature is focused. Cuaresma et al. (2004) notes "...time series of electricity spot prices exhibit more structure which can be used for forecasting compared to time series of financial securities, which are usually quite well described by Markovian processes."

[3]They however ignore seasonality and assumes the error structure is independent over time

Hibon (1997). In this framework, *AR* stands for autoregressive, a shifting of the independent variable expressed as:

$$L^h y_t = y_{t-h}$$

MA stands for moving average, a shifting of the error term express as:

$$L^h \epsilon_t = \epsilon_{t-h}$$

Therefore, the general ARMA regression model can be constructed using p AR terms p and q MA terms:

$$y_t = \alpha + \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + ... + \theta_q \epsilon_{t-q}$$

Both the AR and MA terms have reasonable economic interpretations. For AR terms, past values of the independent variable influence current and/or future values of the variable. For MA terms, past values of the error term (or the shock) influence current and/or past values of the variable.

For the context of the data in this project, each of these interpretations would be valid. For an AR term y_{t-24}, for example, it is reasonable to assume that the processes that influence Activated Volumes 24 hours ago would also be influence the current value, given the daily cyclical nature electricity data exhibits. Additionally, we can conjecture that recent shocks could persist for multiple periods. For example, a forced outage or line congestion that is difficult to immediately replace in the following electricity market could mean that ϵ_{t-1} influences current Activated Volumes.

Makridakis and Hibon (1997) describe four steps for modeling real world data with ARIMA processes. First is to make the data stationary in mean and variance [4]. Second is to specify the values of p and q for the AR and MA terms respectively. Next, estimate the parameters of the p and q lagged terms. Lastly, determine seasonality and modelling of this aspect if necessary.

ARIMA models can be further extended. For example Franses (2004)discuss periodic ARIMA models. These are models where the discrete components of the seasonal models are estimated separately. For example, in the context of 24 seasonal cycle, there could be 24 seperate parameter estimates for each hour of the day. Additionally, multiple seasonal cycles can be included in the general class of ARIMA models.

3.2 Structural/State Space Models

While the ARIMA modelling approach has long been used, and is commonly taught in universities, it is not the only approach to time series analysis. Placed in the broader context of time series models, ARIMA models are known as reduced form models. Another class of models, known as structural time series models (or state space models), are ones where the components of the time series process are explicitly modeled with independent error terms.

The simplest example of a structural time series model would be to model the process using the mean, or level of the series.

[4] In the Box-Jenkins methodology accomplished through differencing: the "I" term in ARIMA

$$y_t = \mu_0 + \epsilon_t$$

The unobserved component of the above equation would be μ_0 the level of the series. We could run a regression to obtain an estimate of this parameter. You could similarly add a linear trend term to this equation, with the model extending to:

$$y_t = \mu_0 + \beta t + \epsilon_t$$

From these static models, the next step that structural models take is to allow these parameters to vary over time, where each term follows a random process on its own. The level would become a local level, and the linear trend would become a local linear trend. Jalles (2009). For a local level model, we would have a set of two equations:

$$y_t = \mu_t + \epsilon_t$$
$$\mu_t = \mu_{t-1} + \eta_t$$

For a local linear trend model, we would get a system of three equations:

$$y_t = \mu_t + \epsilon_t$$
$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t$$
$$\beta_t = \beta_{t-1} + \zeta_t$$

Without going into further detail, from this baseline, we could further specify additional components, such as seasonality, in a way where those compenents are either static or allowed to vary stochastically (local). Harvey et al. (1998). These local parameters than can vary throughout the time series are often estimated using the Kalman filter, which is an equation that takes measurement error into account, smoothing current and past values. For a local level model, you could therefore re-estimate the level parameter at each period, and use the Kalman filter to determine whether the level of the model has in fact changed from period to period.

3.3 Comparing ARIMA and Structural Models

For the context of this research, it is important to understand the theoretical and practical differences between ARIMA and Structural models. While in simple time series processes, these two approaches are nearly one-in-the same, more complex models will highlight their differences. In ARIMA models, the process of estimation through forcing stationarity, removing seasonality, and inclusion of lagged terms to achieve independence of the residuals turns the model into a kind of black-box, according to Jalles (2009). In structural models, each component, such as the trend, level, and seasonality, are explicitly modeled, and these components may be of economic interest depending on the research context.

These two types of models can, however, theoretically be thought of as simply different represenations, with ARIMA being a reduced form version of a Structural Model as mentioned above. Harvey et al. (1998) explains how a local level structural model could be approximated by an ARIMA(0, 1, 1). For a model containing a stochastic trend, cycle, seasonal, and irregular components, it could be expressed as an ARMA(2, s + 4), after first and seasonal differencing.

Harvey et al. (1998) also mentions, however, with the quickly increasing complexity introduced with multiple cycles and levels of non-stationarity, that even if this type of ARIMA model could be identified, it would be near impossible to estimate. Given this, I would expect Structural Models to have the ability to more simply introduce complex processes than ARIMA models.

3.4 Periodic Models

One final type of model that will be tested in this research is called a Periodic Model. Periodic Models work by allowing the parameters to vary over time. Periodic modeling can be done using dummies, but for high frequency data, this can lead to a highly parameterized model. Another approach is to model the coefficients as a fourier series, summing sine and cosine functions over the seasonal cycles for each regressor. This limits the parameters, and allows coefficients to vary. Taylor (2006)

$$Y_t = \phi_0(t) + \phi_1(t)Y_{t-1} + \phi_2(t)Y_{t-24} + \phi_3(t)Y_{t-168} + ... + \epsilon_t$$

Where

$$\phi_p = \omega_0 + \Sigma_{i=1}^{4}\lambda_{pi}\sin(2\pi i\frac{d(t)}{24}) + \upsilon_{pi}\cos(2\pi i\frac{d(t)}{24}) + \kappa_{pi}\sin(2\pi i\frac{w(t)}{168}) + \eta_{pi}\cos(2\pi i\frac{w(t)}{168})$$

4 Descriptive Exploratory Analysis

4.1 Identifying Seasonality

The vital component that each of the models outlined above directly share is an identification of the cyclical seasonal behavior in the data. Each model requires understanding how many, and at what frequencies seasonal cycles occur. Identifying the seasonal cycles of the residuals of the initial model will therefore be a crucial part of this analysis. There are two methods that I will employ in this research to identify the seasonal cycles of the time series.

The first method is the Autocorrelation function. This function calculates the sample correlation between the time series and its lagged values. This function is useful for seeing the dependence of the data in the time domain, and can intuitively show how strong these relationships are. The formula is given by:

$$\widehat{\rho}(j) = \frac{\widehat{Cov}(y_t, y_{t-j})}{\widehat{Var}(y_t)}$$

The empirical covariance for lag j is given by

$$\widehat{Cov}(y_t, y_{t-j}) = \frac{1}{n-1}\Sigma_{t=j+1}^{n}(y_t - \bar{y})(y_{t-j} - \bar{y})$$

The empirical variance of y_t is given by

$$\widehat{Var}(y_t) = \frac{1}{n-1}\Sigma_{t=1}^{n}(y_t - \bar{y})^2$$

The second method uses the Fourier Transform to represent the time series in the frequency domain, by calculating how strongly the time series is represented by a cycle of a given frequency. The frequencies with the strongest power show the most prominent cycles present in the data. The periodogram is calculated by first representing the time series as the sum of sine and cosines functions across all possible frequencies:

$$y_t = \Sigma_{j=1}^{\frac{n}{2}}\beta_1(\frac{j}{n})cos(2\pi\frac{j}{n}t) + \beta_2(\frac{j}{n})sin(2\pi\frac{j}{n}t)$$

And then estimating the parameters β_1 and β_2 to get the power of the time series at each frequency:

$$P(\frac{j}{n}) = \widehat{\beta}_1^2(\frac{j}{n}) + \widehat{\beta}_2^2(\frac{j}{n})$$

Since the time series of the residuals will be messy, it will be helpful to understand how these functions work on a series that is clean and straightforward, like solar generation.

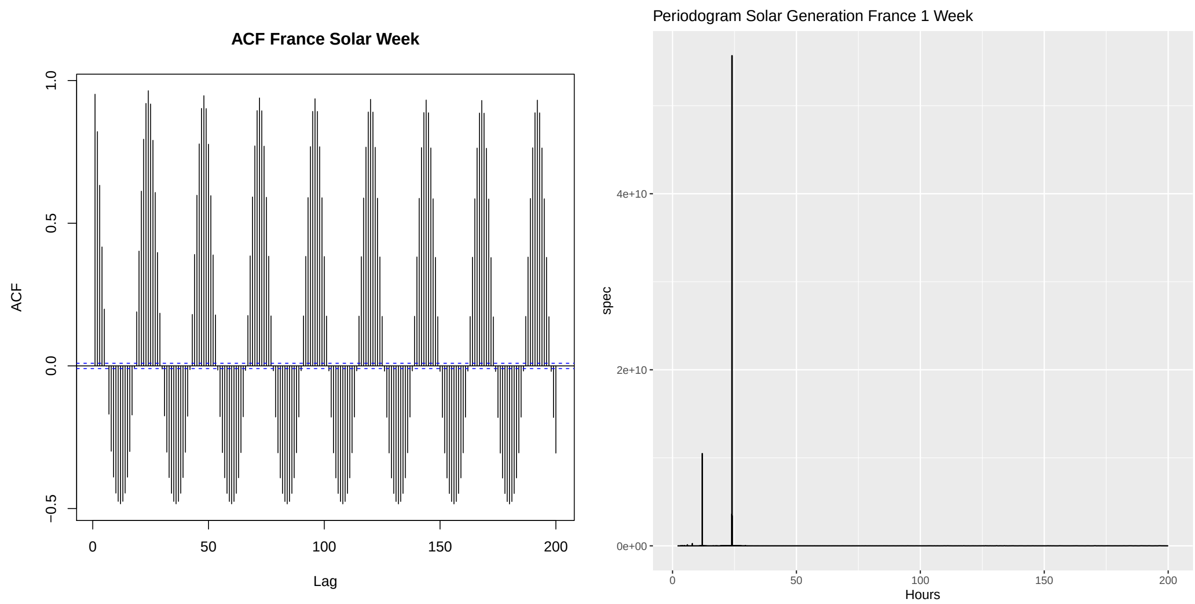

Figure 1: ACF and Periodogram 1 Week Hourly, France Solar Generation

Figure 1 shows the Autocorrelation Function and Periodogram function for solar generation in France for up to lags/frequencies of 200 hours. As expected with the diurnal shape of solar, both graphs clearly show the 2 seasonal cycles of 12 hours and 24 hours. In the ACF graph (time domain), these cycles are seen in the sinusoidal nature of the process, in the periodogram (frequency domain), these show up as high points in the graph.

With this intuition in mind, these two plots for the residuals of the initial model will be easier to interpret. The two graphs below show the ACF and the Periodogram for the residuals of the initial model in France.

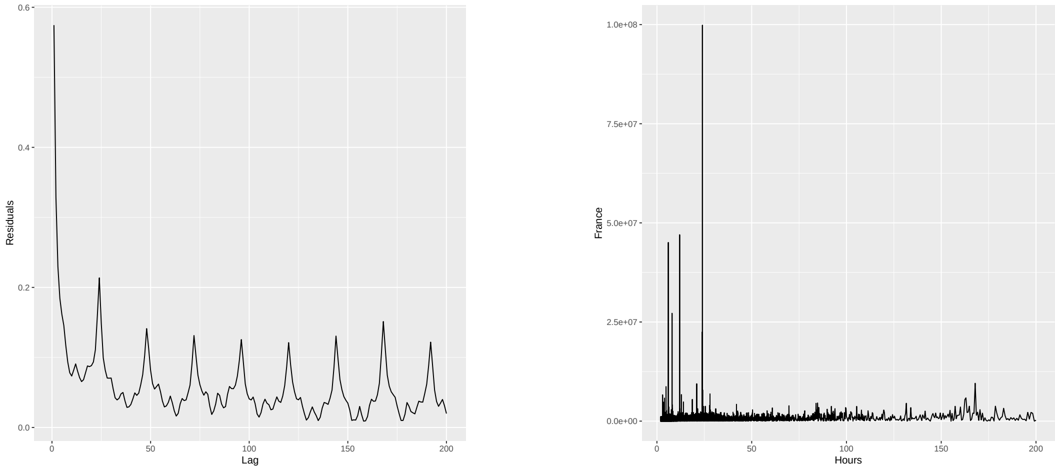

Figure 2: ACF and Periodogram 1 Week Hourly, France Residuals

There are two key takeaways from this comparison. One is just how much messier the graph is. There are several relatively strong cycles, rather than just the two clear ones present in solar generation. This draws directly from the chaotic nature of Activated Balancing Volume. Another key point to notice is the scale of the y-axis. This reinforces the point that not only is the graph messier, but the strength of each cycle is weaker (by construction since the strength is more spread out across multiple cycles compared to solar). We can see that the cycles with the greatest strength are 24 hours, 12 hours, 6 hours, and 8 hours (in that order). The rest in this truncated time span are much lower.

Expanding this analysis to the rest of the regions, Figure 3 shows the autocorrelation function of the residuals for every region, and the 3 strongest seasonal cycles within 200 hours for each region. The two black vertical lines in the right-hand graph delineate a cycle of 24 hours (daily) and 168 hours (weekly). From these two graphs, we can clearly see that there are remaining seasonaly cycles left over from the original model in each of the regions, and that these cycles are most prominently the daily and weekly ones.

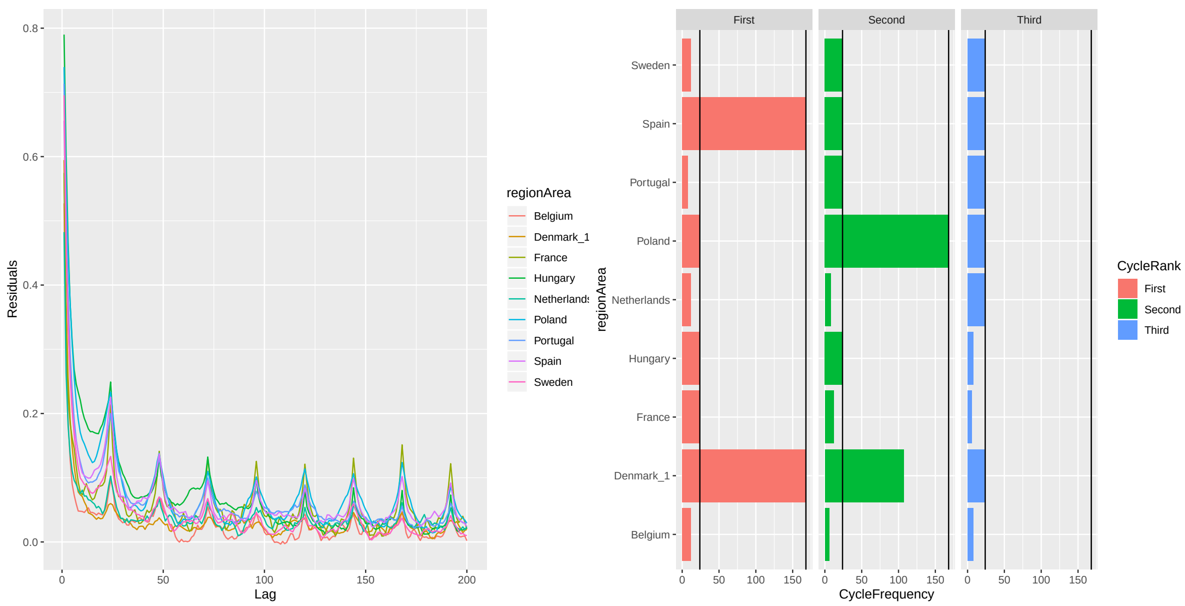

Figure 3: ACF and Periodogram 1 Week Hourly, All Regions' Residuals

4.2 Initial Effects of Modeling Seasonality

Having begun to identify some of the seasonal cycles in the residuals of the initial model, the next important step will be to gain some intuition on how modeling these cycles will affect the behavior of the data. Again, solar generation provides an excellent case study to understand how this may work in the later models.

Franses (2004) discusses a baseline techniques for modeling seasonal cycles. This is to simply run a regression of the time series on dummies for each component of the season. In the case of solar generation, this would mean 1 dummy for each hour of the day. This would translate into estimating the daily seasonality of solar by extracting the hourly mean generation value. We can take this seasonal component and repeat it so that we get an estimated seasonal curve. Figure 4 shows the results of the initial model, and the resulting behavior of the residuals.

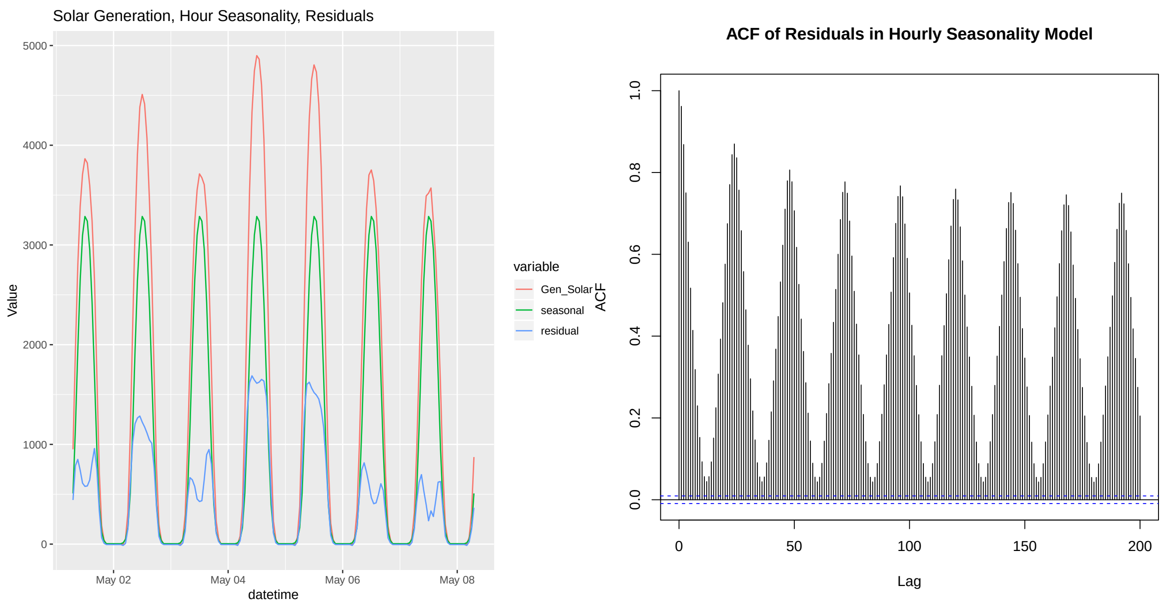

Figure 4: Initial Time Dummies Model of France Solar Generation

For a sample week in May, we see that the hourly seasonal estimation has systematically underestimated the series. This makes sense, as we only looked at cycles within a week time frame, and there was likely another, relatively weaker, cylce present on the monthly scale as solar is higher during the summer. As shown by the ACF graph in Figure 4, the daily cycle is still present in the residuals.

This is an important point to make. This was a specification meant to account for a seasonal cycle of 24 hours. It was an incomplete specification because it did not account for the possibility that the hourly cycle can shift its level throughout the year. When looking at the diagnostic information such as the ACF plot, however, we can see that the negative correlation at the 12 hour lag mostly vanishes, while the 24 hour lag remains. Understanding this behavior requires understanding the underlying process, which will be more complicated, but important for the other variables as well.

For solar, the 12 hour cycle is likely to shift less in magnitude throughout the year compared to the 24 hour cycle. The 12 hour cycle represents roughly day vs. night. North or south of the equator, the sun will shine for more or less hours during the day, but this will not be as impactful as the increase in luminosity at peak hours. Additionally, it would be reasonable to say that the 12 hour cycle is a consequence of or is secondary to the 24 hour cycle. That is, if the 24 hour cycle is correctly controlled for, the 12 hour cycle will be as well. While this may be counter intuitive at first, that a less granular cycle could control for a more granular cycle, it is a byproduct of understanding the processes. The same, however, would not work if we controlled for the monthly cycle, this would not take care of the daily cycle.

To summarize the intution gained from the basic seasonal modeling of solar generation in France, a higher frequency cycle may or may not control for a lower frequency cycle and vice versa. In the case of solar, it seems that controlling for the 24 hour frequency cycle does not account for the lower frequency monthly cycle, and in leaving this out, leaves the behavior of the 24 hour cycle we were controlling for in the residuals.

This intuition will be important when analyzing and comparing the final models. There is the possibility that an incomplete specification could leave traces of the cycles in the final residuals, even when they were controlled for in the model. Additionally, we may see that controlling for some seasonal cycles lessens or increases the relative strength of other cycles. Given these potentially confounding effects of modeling seasonality, from this initial analysis of the data, including seasonal cycles of 24 hours (one day) and 168 hours (one week), makes the most sense both economically, and according to the results from this initial analysis.

5 Models

5.1 Progression from Original Panel

The original 20 region panel approach will become, in this paper, a simple multivariate regression of a single region at a time. To fold in the exogenous regressors into the current time series analysis of the dependent variable, I will take a two step approach. I will first estimate the original model, and then apply the discussed time series identifications and approaches to these residuals. This allows for separating any of the serial correlation explained by the exogenous regressor variables. This approach is additionally motivated by the persistance of the same autocorrelation behavior present in the residuals of the intial model compared to the original time series.

In Figure 5, the left-hand graph displays the autocorrelation function of the original French time series, and the autocorrelation of the residuals of the French model. The right-hand side displays the difference between the autocorrelation values of the initial dataset and the residuals for each region. The remaining serial correlation is visible in the residuals of the original model. While the exogenous regressors reduce the power (unevenly across regions) of some of the residual autocorrelation, they do not fundamentally alter the seasonality present in the time series. Given the orthogonality of the residuals to the original time series, specifying a univariate time series model on the residuals is valid in the overall context.

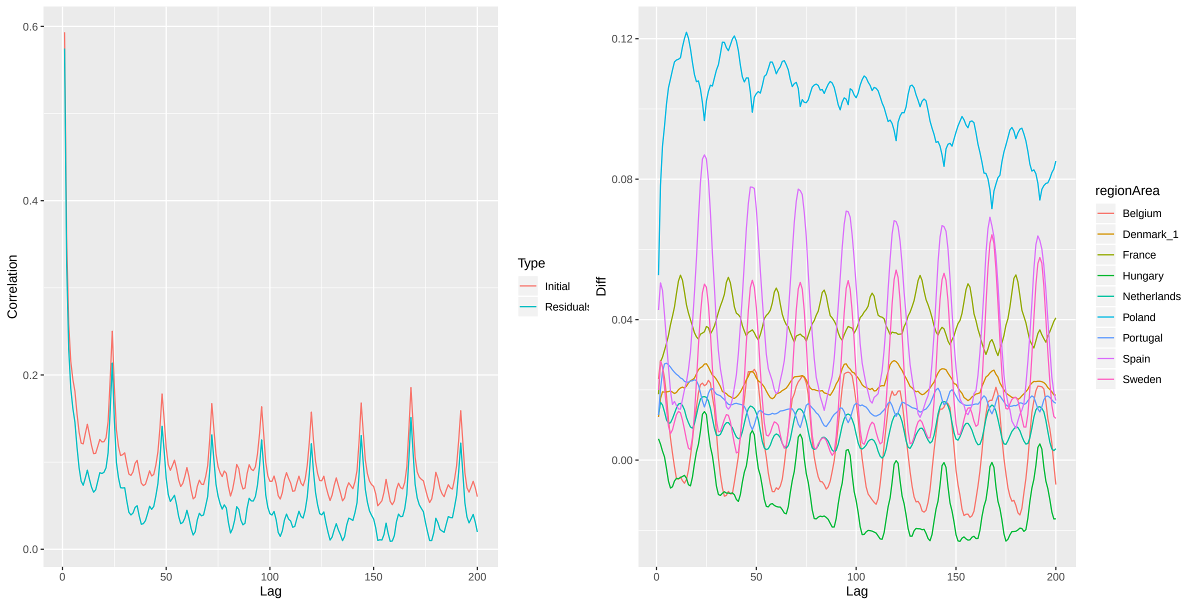

Figure 5: ACF Difference of France, All Regions, 1 Week Hourly

Examining the previous literature on this approach, Liu et al. (2006) forecasts hourly load as a two-step process as well. First fitting a non-parametric curve to the temperature-load relationship, then fitting the residuals to a linear parametric ARIMA model to account for serial correlation.

5.2 Current Models

All models will involve using the residuals of the initial regression as stated above [5].

Since this is the second regression of the two-step regression described previously, the dependent variable in these regression will be denoted as ϵ_t, while the error term (assumed to be white noise), will be denoted as η_t

5.2.1 Time Dummies Model

$$\epsilon_t = \Sigma_{i=0}^{22}\beta_i 1(hour(t) = i) + \Sigma_{j=0}^{6}\gamma_j 1(weekday(t) = j) + \Sigma_{k=0}^{11}\delta_k 1(month(t) = k) + \eta_t$$

This is a good first model to provide a baseline to compare the other models with. It defines a set of dummies for some of the major cycles seen in the data.

5.2.2 ARIMA(1, 0, 1) Model

$$\epsilon_t = \phi_1 \epsilon_{t-1} + \eta_t + \theta_1 \eta_{t-1}$$

This is another type of model to provide a baseline to compare the other models against. It doesn't directly have any seasonal components, just the past values of the AR and MA components.

[5]The Time Series Analysis (TSA) library in R is used for ARIMA models. The Bayesian Structural Time Series (BSTS) is used for the structural time series model

5.2.3 Seasonal ARIMA

$$\epsilon_t - \epsilon_{t-24} = \phi_1\epsilon_{t-1} + \phi_2\epsilon_{t-2} + \phi_3\epsilon_{t-24} + \phi_4\epsilon_{t-168} + \eta_t + \theta_1\eta_{t-1} + \theta_2\eta_{t-24} + \theta_3\eta_{t-168}$$

This is a double seasonal ARIMA model, which goes a step beyond the simple ARIMA model to directly model some of the prominent seasonal cycles seen in the data. The parameters are fixed at each of the lag values, so this can be nicely compared against periodic models.

5.2.4 Periodic AR

$$\epsilon_t = \phi_1(t)\epsilon_{t-1} + \phi_2(t)\epsilon_{t-2} + \phi_3(t)\epsilon_{t-24} + \phi_4(t)\epsilon_{t-168} + \eta_t$$

Where

$$\phi_p(t) = \omega_0 + \Sigma_{i=1}^{4}\lambda_{pi}\sin(2\pi i\frac{d(t)}{24}) + \upsilon_{pi}\cos(2\pi i\frac{d(t)}{24}) + \kappa_{pi}\sin(2\pi i\frac{w(t)}{168}) + \nu_{pi}\cos(2\pi i\frac{w(t)}{168})$$

This type of model will take a simple autoregressive model, and allow for the coefficients to vary in a cyclical manner over time. These cyclical parameters can be modeled with trigonometric functions, as above, or with dummies. This type of model relaxes some of the more rigid assumptions in the ARIMA class of models, and can ideally capture more of the variation and structure of the data.

5.2.5 Structural Model

$$\epsilon_t = \phi_1\epsilon_{t-1} + \mu_t + \Psi_{1,t} + \Psi_{2,t} + \eta_t$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \nu_t$$

$$\beta_t = \beta_{t-1} + \zeta_t$$

$$\begin{pmatrix} \Psi_{p,t} \\ \Psi^*_{p,t} \end{pmatrix} = \begin{pmatrix} cos(\lambda_p) & sin(\lambda_p) \\ -sin(\lambda_p) & cos(\lambda_p) \end{pmatrix} \begin{pmatrix} \Psi_{p,t-1} \\ \Psi^*_{p,t-1} \end{pmatrix} + \begin{pmatrix} \kappa_{p,t} \\ \kappa^*_{p,t} \end{pmatrix}$$

This is also known as a trend cycle model. $\Psi_{1,t}$ and $\Psi_{2,t}$ will correspond to evolving trigonometric cycle terms for the daily and weekly cycles respectively. λ_p is the value of the frequency for each of these cycles in the *sin* and *cos* terms Jalles (2009). μ_t and β_t represent the local level and trend components respectively. Separately defining error terms for each of the components will allow insight into the model components and structure.

6 Results and Conclusion

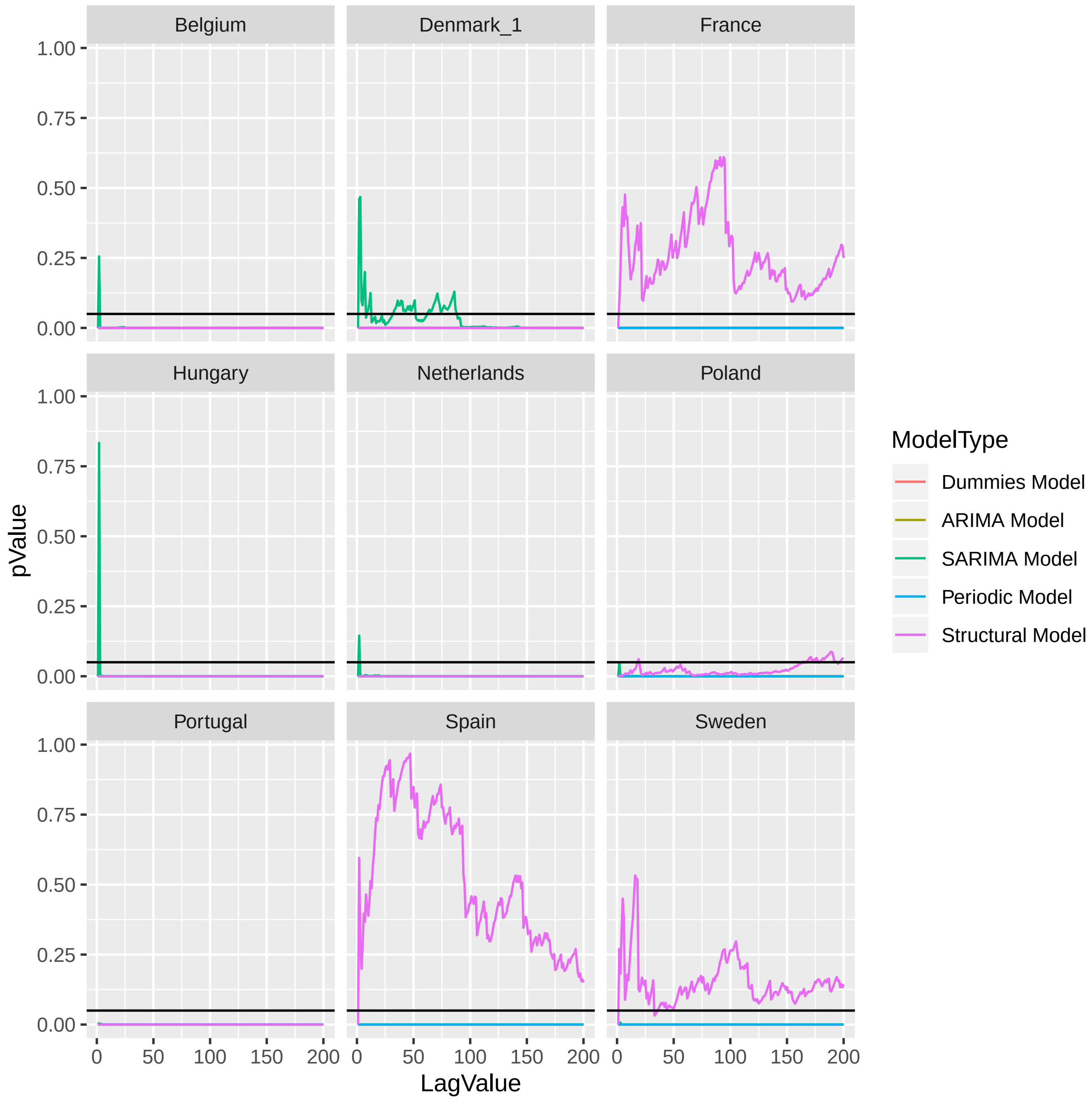

Figure 6: Test of Serial Correlation For All Models/Regions

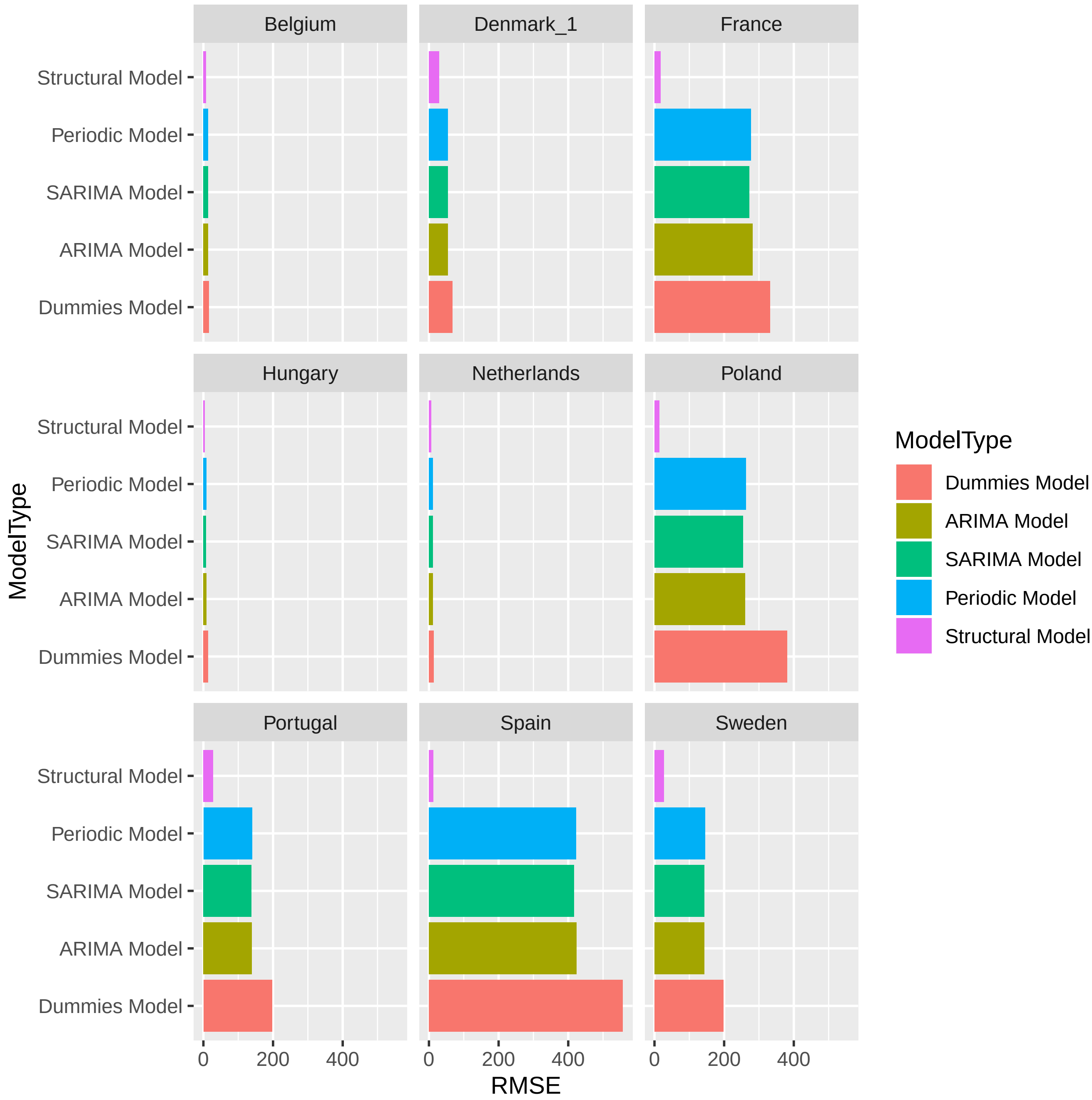

Figure 7: RMSE For All Model/Regions

Examining Figure 6 and Figure 7 showing the results of the models, we can see that the structual model peformed the best in minimizing root mean squared error. Additionally, it was the only model to trend towards no serial correlation [6] in any of the regions (4/9). For further interpretation of the efficacy of these models, it is also worth mentioning that serial correlation could theoretically arise from an overspecified model as well as an underspecified one. With applying the same models to each region in this research, this does become a possibility.

Diving in a little deeper, there does appear to be a relationship between the RMSE of the models and the serial correlation p-values. For example, in France, Poland, Spain, and Sweden, where the structural models trend towards passing the serial correlation test, the structural models have significantly lower rmse values compared to the other models. Figure 8 shows the

[6]The Box-Pierce test of serial correlation was used at each lag to obtain the p-value

ratio of the RMSE of the dummies model and the structural model on the x-axis, with the p-value of the structural model on the y-axis. While not a conclusive relationship, all ratio values to the right of Sweden (ratio of 7.3) have a significant p-value for the test of serial correlation at 200 lags.

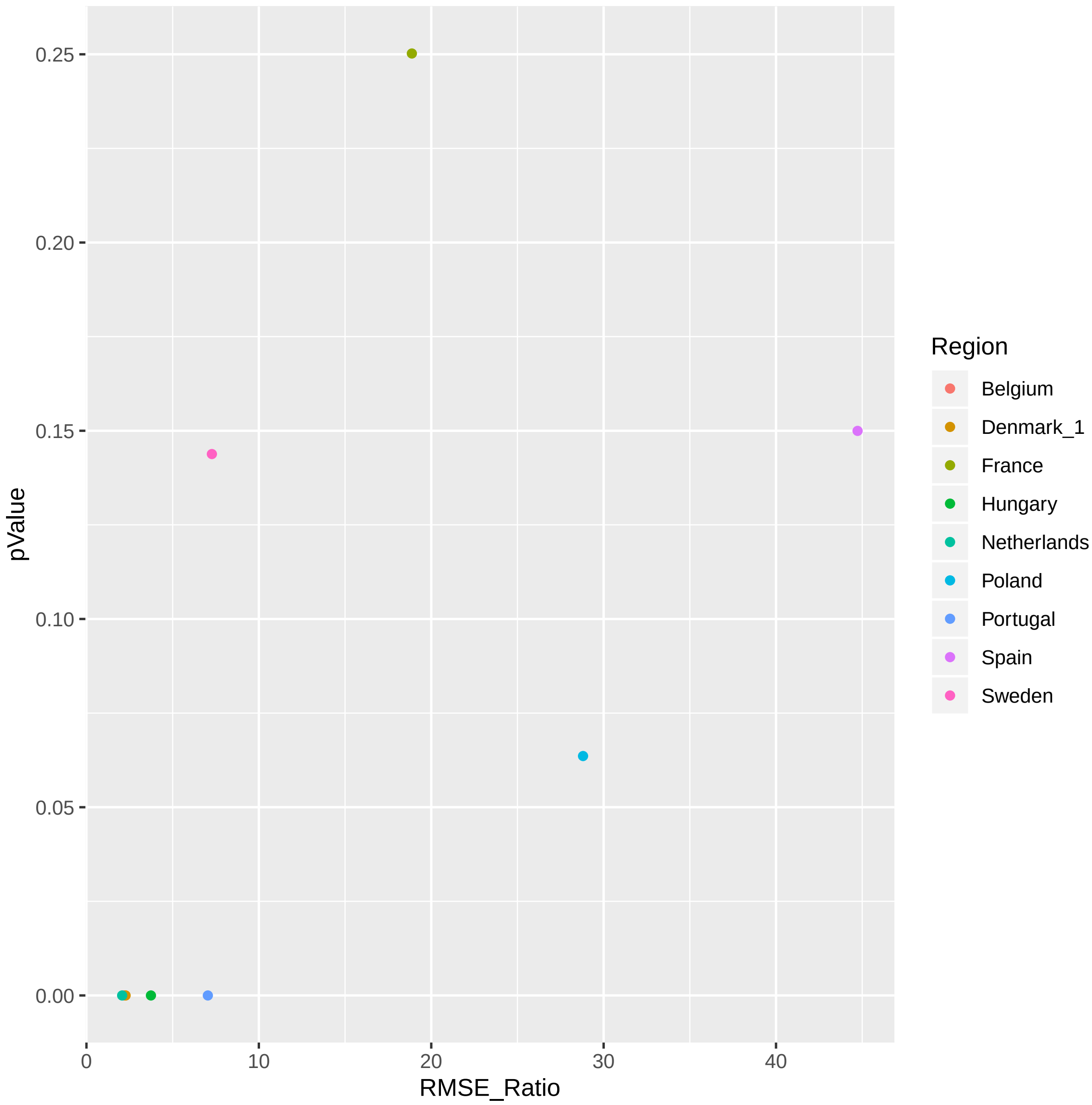

Figure 8: RMSE Ratio Compared to pValue at 200 Lags

On top of these numerical metrics of the performance of the models, it is worth including a graphical sanity check. Figure 9 shows the autocorrelation function out to 200 lags for each of the models and regions. The picture is not as clear cut here, with the SARIMA model appearing to perform as well as the structural model. In fact, in the case of Hungary, comparing the SARIMA and structural models, the SARIMA model performs better. Whether this is a situation of underspecification or overspecification on the part of the structural model requires further analysis. This graphical interpretation additionally may call into question the efficacy of the test used for determining serial correlation. Further research may yield other, more effective tests for this context.

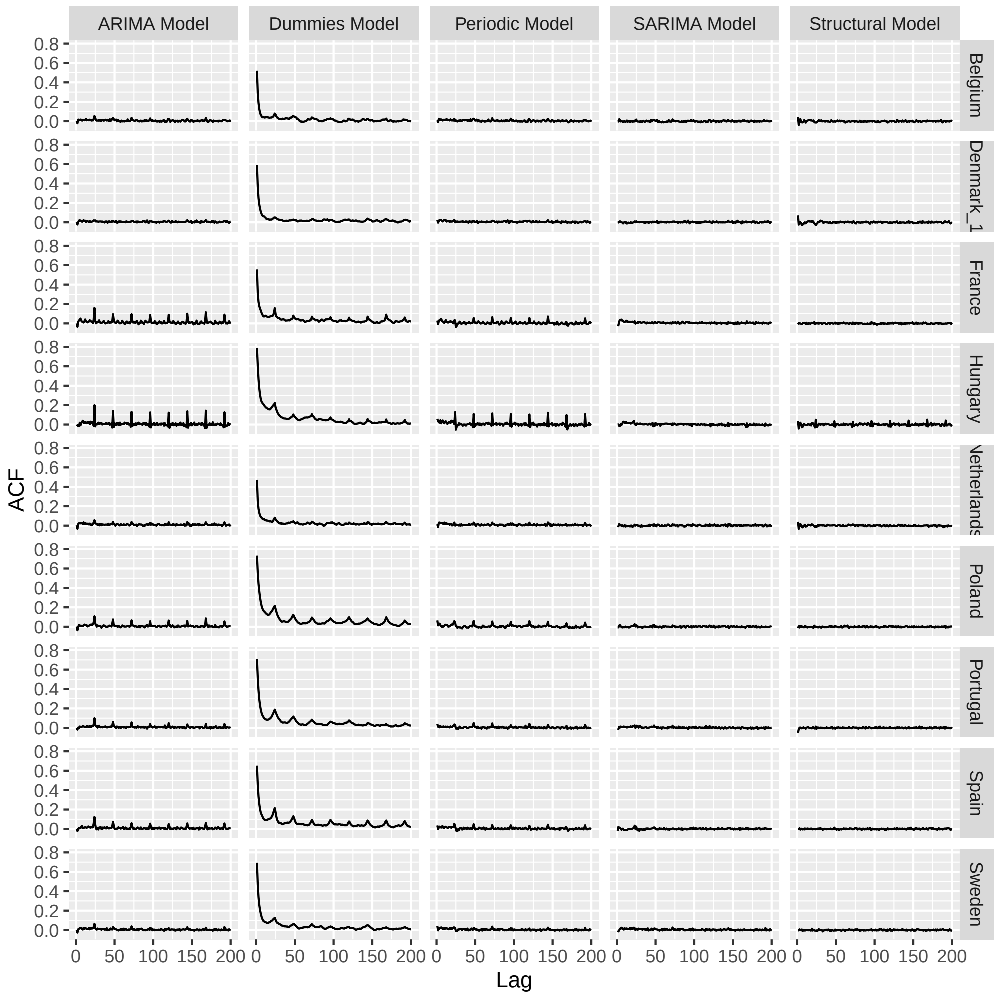

Figure 9: Autocorrelation Function per Model and Region

These results of the models relative to each other do make sense. In regards to the underlying specifications, the structural model included many of the components that the other models contained. The parameters were allowed to vary in the structural model, similarly to the periodic specification. The seasonal cycles were explicitly modeled using time varying parameters on trigonometrics terms, combining the benefits of both the periodic and seasonal ARIMA models. Additionally, unlike the other models, each of these components was modeled separately, allowing for a full understanding of the contribution of each of the components of the structural model.

While the structural model requires more background knowledge, and is more complex to estimate, this additional complexity comes with the benefit of a more robust interpretation and better performance compared to the next best model, the SARIMA model. On this note, future research could more systematically study the output of the structural model to understand the differing patterns in the data across regions. This could lend insight into both the time series aspect of the modeling as well as into potential omitted variables from the initial model by Boomer and Kim (2020).

In summary, due to the flexibility and capability for complexity given by structural time series models, as well as its overall performance towards the goal of eliminating serial correlation, this class of models seems to be the best choice for producing a valid specification to the original model initiated by Boomer and Kim (2020).

A Electricity Market Background

Balancing

1) Definition of Balancing

Balancing means all actions and processes through which Transmission System Operators (TSOs) continuously ensure grid stability. It can be achieved either by load or generation and also can take positive (Up regulation; generation ramp-up or negative demand response) or negative value (Down regulation; generation ramp-down or positive demand response).

Balancing volume is procured, or contracted in the balancing market, based on the extreme case scenario and on a minimum required amount, both of which are significantly higher than the activated volume in a normal case. [7].

2) Definition of Activated Balancing

Balancing Service Providers (which, once again, can either be load- or generation-providers) bid for and commit to being called upon a day ahead, and hence the market is called the Day-ahead Balancing Market. This commitment, however, may or may not be actually realized, depending on the discrepancy of the real-time market load and the actual load at flow time.

Electricity Market as an Auction Market & Merit Ordering

Though there exist decentralized over-the-counter markets directly between generator and Load Serving Entities, most of the electricity needs are satisfied through centralized markets that take the form of an auction market. The markets are cleared at the quantity demanded by Load Serving Entities (based on projections) and the clearing price can be either uniform or pay-as-you-bid.
With differing marginal costs, different types of generators form a stepwise supply stack where variable renewable electricity (VRE), low in production cost but also low in flexbility, is located at the lower left end, and traditional energy sources such as coal, high in production cost but also high in flexibility are located at the higher right end. Under this merit-order structure, VRE generators hold an advantage over generators of traditional energy sources, which creates the possibility of a crowding-out effect.

B Data Appendix

B.1 Data collection method

The raw data extracted from the platform is in XML files, which are then unpacked and transformed into CSV files by each generation type. We then merged the files into a single CSV file.

B.2 Variables

Activated Balancing Volume: The dependent variable

First we took the absolute value of Activated Balancing Volume data [8], since our motive is to see how much balancing is caused due to more renewable energy sources in the market,

[7]based on COMMISSION REGULATION (EU) 2017/1485 of 2 August 2017: a guideline on electricity transmission system operation

[8]As explained in Appendix C. Electricity Market Background, balancing can take either positive or negative value

and not at which direction. Then we aggregated the data into an hourly time-frame, in order to look at the total balancing volume without causing any information-loss with too much aggregation.

VRE Penetration: The main explanatory variable

$$\textit{VRE Penetration Rate} = \frac{\textit{Total generation by wind and solar}}{\textit{Total generation}}$$

VRE Forecast Error

$$\begin{gathered}\textit{VRE Forecast Error} = \\ |\textit{Day-ahead Generation Forecast for Wind and Solar} \\ -\textit{Actual Generation of Wind and Solar}|\end{gathered}$$

ENTSO-E provides data on the day-ahead forecast generation amount by production methods and the actual generation. The difference would indicate the error in prediction, hence we construct the variable as forecast error. We took the absolute value to maintain consistency across variables.

Load

Electricity load in the Day-ahead Market

Load Forecast Error

$$\textit{Load Forecast Error} = |\textit{Day-ahead Load Forecast} - \textit{Actual Load}|$$

ENTSO-E provides data on the day-ahead forecast load and actual load. The difference would, in this case as well, indicate the error in prediction. Again, we took the absolute value to maintain consistency across variables.

Percentage of Other Generation Methods

$$\textit{Percentage of Other Generation Methods} = \frac{\textit{Generation by all the other methods}}{\textit{(Total Generation - Generation by Renewable)}}^{*}$$

* In order to reduce down degree of mutli-collinearity, we subtract the generation of VRE from the total generation.

Other Generation Forecast Error

$$\begin{gathered}\textit{Other Generation Forecast Error} = \\ |\textit{Day-ahead Generation Forecast for All the Other Generation Methods} \\ -\textit{Actual Generation of All the Other Generation Methods}|\end{gathered}$$

Export Import

$$\textit{Export Import} = |\textit{Import Flow} - \textit{Export flow}|$$

C Glossary

Activated Balancing Volume
Realized balancing, which is a portion of procured balancing.

Balancing
All actions and processes through which TSOs continuously ensure the stability of the grid system.

ENTSO-E (European Network of Transmission System Operators for Electricity)
Consists of 42 TSOs from 35 European countries, ENTSO-E was established and given legal mandate by the EU's Third Legislative Package for the Internal Energy Market in 2009, which aims at further liberalising the gas and electricity markets in the EU.

Load
Load means the amount of electricity demanded by consumers. Load can be a forecast measure or the precise volume demanded at flow time.

TSO (Transmission System Operator)
Legal entity responsible for operating and maintaining, and if required, developing the transmission system in a given area

VRE (Variable Renewable Energy)
Refers to Wind and Solar

D Bibliography

Boomer, A. and M. Kim (2020). Maintaining the balance: Renewables and european electricity balancing markets.

Crone, S. F. and R. Dhawan (2007). Forecasting seasonal time series with neural networks: a sensitivity analysis of architecture parameters. In *2007 International Joint Conference on Neural Networks*, pp. 2099–2104. Ieee.

Cuaresma, J. C., J. Hlouskova, S. Kossmeier, and M. Obersteiner (2004). Forecasting electricity spot-prices using linear univariate time-series models. *Applied Energy 77*(1), 87–106.

Dilaver, Z. and L. C. Hunt (2011). Industrial electricity demand for turkey: a structural time series analysis. *Energy Economics 33*(3), 426–436.

ENTSO-e (2018). Electricity balancing in europe. `https://eepublicdownloads.blob.core.windows.net/public-cdn-container/clean-documents/Network%20codes%20documents/NC%20EB/entso-e_balancing_in%20_europe_report_Nov2018_web.pdf`.

Erzgräber, H., F. Strozzi, J.-M. Zaldívar, H. Touchette, E. Gutiérrez, and D. K. Arrowsmith (2008). Time series analysis and long range correlations of nordic spot electricity market data. *Physica A: Statistical Mechanics and its Applications 387*(26), 6567–6574.

Franses, P. H. (2004). Forecasting seasonal time series 1 . introduction.

Franses, P. H., P. H. B. Franses, et al. (1998). *Time series models for business and economic forecasting*. Cambridge university press.

Garcia, M. P. and D. S. Kirschen (2006). Forecasting system imbalance volumes in competitive electricity markets. *IEEE Transactions on Power Systems 21*(1), 240–248.

Gould, P. G., A. B. Koehler, J. K. Ord, R. D. Snyder, R. J. Hyndman, and F. Vahid-Araghi (2008). Forecasting time series with multiple seasonal patterns. *European Journal of Operational Research 191*(1), 207–222.

Harvey, A., S. J. Koopman, and J. Penzer (1998). Messy time series: a unified approach. *Advances in econometrics 13*, 103–144.

Huisman, R., C. Huurman, and R. Mahieu (2007). Hourly electricity prices in day-ahead markets. *Energy Economics 29*(2), 240–248.

Jalles, J. T. (2009). Structural time series models and the kalman filter: a concise review.

Liu, J. M., R. Chen, L.-M. Liu, and J. L. Harris (2006). A semi-parametric time series approach in modeling hourly electricity loads. *Journal of Forecasting 25*(8), 537–559.

Makridakis, S. and M. Hibon (1997). Arma models and the box–jenkins methodology. *Journal of Forecasting 16*(3), 147–163.

Sfetsos, A. (2000). A comparison of various forecasting techniques applied to mean hourly wind speed time series. *Renewable energy 21*(1), 23–35.

Taylor, J. W. (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of the Operational Research Society 54*(8), 799–805.

Taylor, J. W. (2006). Density forecasting for the efficient balancing of the generation and consumption of electricity. *International Journal of Forecasting 22*(4), 707–724.

Taylor, J. W., L. M. De Menezes, and P. E. McSharry (2006). A comparison of univariate methods for forecasting electricity demand up to a day ahead. *International journal of forecasting 22*(1), 1–16.

Taylor, J. W. and P. E. McSharry (2007). Short-term load forecasting methods: An evaluation based on european data. *IEEE Transactions on Power Systems 22*(4), 2213–2219.