Advanced Forecasting in Energy Markets

Summer Semester 2023/ 2024

Project: The Role Of Holiday In Probabilistic Load Forecasting: A Spanish Case Study

Authors: Öcal Kaptan & Ly Le Thi

Supervisors: Prof. Dr. Florian Ziel

1 Introduction

Accurate load forecasting is crucial for managing electrical power systems, balancing supply and demand, and optimizing operational costs. In this study, we tackled the challenge of probabilistic load forecasting for Spanish electricity demand for 168 hours ahead, with a particular focus on incorporating the holiday effects. Holidays cause significant deviations from normal consumption patterns due to changes in residential, commercial, and industrial activities. In Western countries like Spain, holidays like Christmas and Easter may decrease industrial demand, while other public holidays might boost residential consumption. To address these variations, we incorporate holiday effects into our forecasting models to improve accuracy and reliability.

Advanced techniques such as Distributional Neural Networks (DDNN), Generalized Additive Models for Location Scale and Shape (GAMLSS), and AutoRegressive with Exogenous Inputs (ARX) are utilized for modelling with the integration of holiday effects, weather conditions, and historical load data into the forecasts. The model performance will be evaluated using the Pinball Score to assess forecast accuracy across different quantiles and the Diebold-Mariano (DM) test for comparing forecasting methods.

The result shows that with the explore and combine different holidays and weather data into model, GAMLSS outperforms all other model, especially for fitting Skewed t distribution. DDNN models, though useful for capturing complexity patterns, did not achieve the same level of probabilistic accuracy as the GAMLSS models. ARX with the highest pinball score was less effective at capturing the probabilistic nature of load forecasts. In addition to this introduction, the report comprises sections on the dataset, methods, forecast performance evaluation, results, and limitations.

2 Data

The data for this project is provided by the study course. For load forecasting, three primary categories of features are crucial: calendar features (e.g., day of the week, hours of the day, season), weather features (e.g., temperature, wind speed), and socio-economic features (e.g., holidays, economic growth) (Monika & Ziel, 2024). These categories are recognized as the most significant and predictable factors affecting electricity demand. Therefore, our modeling approach will focus on these three key variables for load forecasting:

- Day-ahead Load of Spain. Measurement unit: Gigawatt (GW).
- Calendar data: days of the week and hours of the day.
- Holiday data: including 35 public and regional holidays in Spain,
- Weather data: forecast and actual temperature (Measurement unit: Degrees Celsius (°C))

Thus the main multivariate dataset used in this research contains a total of 38 variables with 81,838 data points equivalent to the hourly observations of 38 variables starting from 2018-01-01 at 00:00:00 and ending at 2024-06-01 at 21:00:00. For the data from this range we will utilize 5 years of data from 2018-01-01 00:00:00 to 2022-12-01 08:00:00 for training, one year for validation from 2022-12-01 09:00:00 to 2023-12-01 08:00:00 and the last 6 month from 2023-12-01 at 09:00:00 to 2024-06-01 at 08:00:00 for out of sample testing as described in Figure 1.

2.1 Load Day-ahead

Load day-ahead (DA) refers to the anticipated electricity demand for the next day. Figure 1 illustrates the stability of the DA Load in Spain from early 2018 to mid-2024. A noticeable decrease in power consumption occurred in March and April 2020, coinciding with the onset of the COVID-19 pandemic in Spain, which began on March 3, 2020.

Table 1 presents the summary statistics for Spain's Day-Ahead electricity load data, offering insights for further analysis and forecasting. The data exhibits significant volatility, as evidenced by a standard deviation of 4.53 GW, which reflects substantial fluctuations around the mean load of 27.51 GW. The median load value is 27.40 GW, which is quite close to the mean, suggesting that the distribution of load values is relatively symmetric. The range of the data spans from a minimum of 15.33 GW recorded on April 10, 2023, at 06:00:00, to a maximum of 41.78 GW observed on January 12, 2021, at 19:00:00. This broad range highlights the variability in electricity demand over the study period.

Additionally, the skewness of 0.136 and kurtosis of -0.744 provide insights into the shape

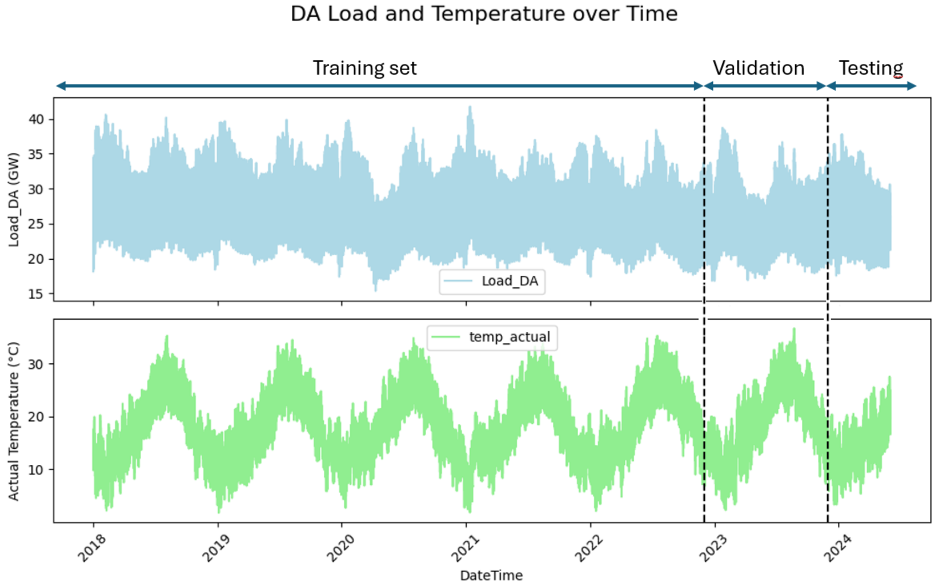

Figure 1: Day-ahead load over time

Table 1: Summary statistics of Spain's Load DA (Unit: GW)

Statistics	Mean	Median	Standard Deviation	Minimum	Max	Skewness	Kurtosis
Loads	27.51	27.40	4.52	15.33	41.78	0.136	-0.744

of the load distribution. A skewness value of 0.11 indicates that the distribution is almost symmetric, with a slight tendency for the load values to be slightly skewed to the right. Kurtosis, on the other hand, measures the "tailedness" of the distribution. A kurtosis value of -0.794 suggests that the distribution has lighter tails than a normal distribution and is less prone to extreme outliers. These characteristics suggest that Spain's Load DA is approximately normally distributed with a lighter tail than a normal distribution. Due to this reason and the prior fitting of distribution on the training set, this research chose the Normal distribution and Johnson SU (JSU) and Skewed Distribution for initial analysis. The normal distribution is chosen for its simplicity and the availability of established techniques, and the Johnson SU and Skewed distribution for its flexibility in modelling data with varying skewness and kurtosis.

2.2 Calendar data

For the calendar data, the research focuses on the daily, weekly and seasonal effects on the load forecasting model. For that reason, 3 dummy variables are selected: Hour of the week or Day of the week, hour of the day and hour of the year. The hour of the day plays a crucial role in shaping daily load profiles, which are characterized by peak consumption periods. As shown in Figure 2, there is a distinct pattern in the load profile: electricity consumption is comparatively lower during the early morning hours from 0:00 to 6:00 and in the late evening after 20:00. In contrast, the load experiences noticeable increases around 10:00 and 20:00.

The hour of the week or day of the week will capture the complexity of the electricity demand that varies between weekdays and weekends. Figure 2 reveals that the load is consistently lower during the weekend and it tends to be higher on normal weekdays due to industrial and commercial activities. There are some weekdays fall in the low load days are the holiday happens on weekdays. The hour of the year further expands this understanding by capturing seasonal and annual variations in electricity demand. Seasonal factors, such as temperature changes and daylight hours, significantly affect energy consumption patterns.

2.3 Holiday

Analysts often separate holiday data from regular data to understand holidays' unique effects on electricity demand. In the context of load forecasting for Spain, it is important to account for both public and regional holidays to accurately reflect variations in electricity demand patterns due to changes in consumption behavior. Public holidays are observed nationwide, affecting every county such as Christmas Day, New Year's Day, Labour Day, Good Friday, etc. In contrast, regional holidays are specific to certain states or regions, such as the Day of Madrid, which only impacts the Madrid region, Day of the Balearic Islands which only impacts the Balearic Islands or Maundy Thursday which happens in many counties in Spain such as Andalusia, Aragon, Balearic Islands,.... Both types of holidays need to be modelled differently due to their distinct characteristics and impacts on electricity demand.

2.3.1 Public holiday

When analyzing public holidays, we distinguish between fixed and flexible holidays. Fixed holidays are the type of holiday that occurs on the same date every year, such as Christmas Day (25/12), New Year's Day (01/01), and Labour Day (01/05). Flexible holidays, like Easter, occur on specific weekdays but vary each year. Each type requires a different approach to modelling. Flexible holidays, like Easter, Good Friday, and Maundy Thursday, vary annually as they depend on the lunar calendar. For instance, Easter fell on April 1st in 2018 and April 21st in 2019, with Good Friday and Maundy Thursday observed on the preceding Friday and Thursday, respectively.

Flexible holidays are somewhat more straightforward to model because they always fall on the same weekday each year. When a flexible public holiday occurs, it disrupts the usual weekly load structure, impacting not only the day of the holiday but also the hours leading up to and following it due to transitional effects. To capture these variations, we establish a 36-hour framework for each flexible holiday (denoted as D $\in$ Flex), encompassing 6 hours before the holiday, the 24 hours of the holiday itself, and 6 hours after the holiday (Ziel & Liu, 2016). The impact of flexible holidays will be approached as follows:

$$\mathcal{B}_k^D(t) = \begin{cases} 1, k \leq HoD \\ 0, Otherwise \end{cases} \tag{2.1}$$

where HoD_t is the hours of the day counting from 1, 2, . . . , 36 at time point t around the public holiday starting counting from $18:00$ the previous day to $6:00$ the following day of public holiday.

An alternative method for handling flexible holidays that we have tried is to set the 24 hours of that public holiday to 1 and all other days to 0. However, this method does not capture the load shape as effectively as the above approach, which involves setting 6 hours before the holiday, the 24 hours of the holiday itself, and 6 hours after the holiday to 1. The latter method provides a more comprehensive representation of the load variations associated with the holiday.

Fixed holidays present a more complex challenge because their impact varies depending on the day of the week they fall on. To address the complexity of modelling fixed holidays, we make use of an effective coefficient $C(t)$ (Ziel & Liu, 2016) for each hour of the week. $C(t)$ represents the coefficient for each hour of the week. This coefficient will help adjust the load patterns based on the day and time, considering the similarity to Sunday load patterns of fixed holidays. Using $C(t)$, we define the basis functions for fixed holidays F $\in$ Fix as follows:

$$\mathcal{B}_k^F(t) = \begin{cases} C(t), k \leq HoD \\ 0, Otherwise \end{cases} \tag{2.2}$$

In this model, HoD(t) represents the 36-hour period surrounding a public holiday, starting from 18:00 the day before the fixed holiday and extending through 6 hours after the holiday itself. To quantify the holiday's impact on electricity demand, we use coefficients $C(t)$ that are determined based on the specific weekday on which the holiday falls. - If the holiday occurs on a Sunday, the coefficient $C(t)$ is set to 0, reflecting that there is no additional impact beyond the usual Sunday load pattern. The Sunday load is also considered as low-level load targets and the low-level load targets are considered is 24-hour mean load values observed on Sunday of that week. - When a holiday falls on a core working day—such as Tuesday, Wednesday, or Thursday—the coefficient $C(t)$ is set to 1, indicating that the holiday's impact is at its maximum and we refer to this as high-level load targets. High-level load targets are considered is 24-hour mean load values observed on these three days of the considered week. - For holidays occurring on Monday, Friday, or Saturday, the impact is considered to be in between these above extremes, with $C(t)$ taking a value between 0 and 1 to reflect an intermediate level of impact. To determine these coefficients, we compare the actual load during the holiday period to the high-level and low-level load targets, using the formula:

$$C(t) = \max\left\{1 - \frac{\text{high level load target}(t) - \text{actual load target}(t)}{\text{high level load target}(t) - \text{low level load target}(t)}, 1\right\} \tag{2.3}$$

with the actual load target is considered as the means across all hours of the Load of the considered week.

2.3.2 Regional holiday

To assess the impact of regional holidays, we take the percentage of Spain's total population that resides in counties where the holiday is observed. We combine the percentage populations of all counties that celebrate the holiday and set a threshold of 10% of Spain's total population. If this combined population represents more than 10% of the national population, the 24 hours of that holiday is assigned a value of 1, indicating its significance, if the population falls below this threshold, the holiday is assigned a value of 0, indicating minimal regional impact. This approach helps in quantifying the influence of regional holidays based on their observed reach across the Spanish population.

2.3.3 Christmas holiday period

In Western countries, Christmas is often the most significant and extended holiday period for both educational institutions and the industrial sector. Given that many businesses and schools close for the entire duration of the Christmas holidays, there is a notable dip in electricity consumption patterns during this period which was observed by checking the load shape for this period. To leverage this opportunity, we analyzed the load data from previous years specifically during the Christmas holiday range. Our approach involved examining the variations in electricity load over the official Christmas holiday period in Spain (expatica, 2024), which typically starts on December 24th and lasts through January 6th, encompassing both Christmas and the extended holiday season. The days fall in this period will be set to 1 to show that this is the Chrismas days period and 0 otherwise

2.4 Weather

Temperature plays a crucial role in load forecasting due to its direct impact on electricity demand. The extreme temperature events, such as heatwaves or cold snaps, can cause significant spikes in electricity demand, which must be anticipated for effective grid management (Enric et al., 2001).

Figure 2 illustrates the close relationship between temperature and electricity load. The temperature-electricity load relationship reveals a clear "V" shaped curve, illustrating a non-linear trend. This curve shows that as temperatures deviate from a moderate range, electricity demand increases. Specifically, when temperatures fall below 12°C, the demand peaks due to higher heating needs, while temperatures above 27°C also drive up electricity use for cooling. Between 12°C and 26°C, the load is generally lower, with the lowest point around 21°C, where indoor comfort requires minimal heating or cooling.

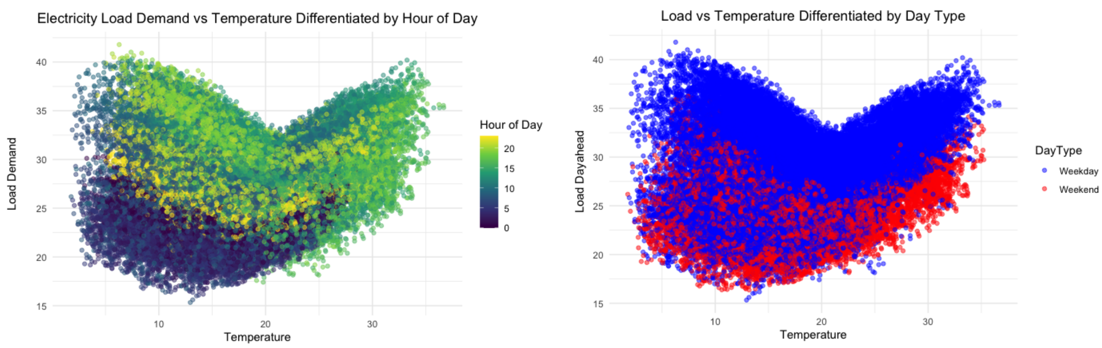

Figure 2: Daily, weekly and temperature affect to Load of Spain. (Unit: GW)

2.5 Features calculation

For Spain's Load Day ahead, by examining the autocorrelation function (ACF) and partial autocorrelation function (PACF) across all hours, lag values of 1, 2, 6 and 7 days are selected to capture daily fluctuations and weekly seasonality of historical Load DA. These lag values are introduced to enhance the predictive capability of the model for Spain's Load DA.

Additionally, due to the investigation of calendar and holiday data described in Section 2.2 and Section 2.3, the hour of the day, hour of the week, hour of the year, flexible holidays, fixed holidays, regional holidays and Christmas holiday period is also included in the model. After calculate the holiday as described in part 2.3 total of 21/35 holidays are chosen for modelling.

For model training, historical temperature data is utilized to develop and refine load forecasting models. For testing, we utilize 10-day forecast data averaged from ten weather stations in Spain's largest cities in population. This method ensures that models are evaluated under realistic future conditions, minimizes cumulative errors from predictions, and reflects broader regional weather trends. We consider temperature effects at various intervals—1 hour, 2 hours, 3 hours, 24 hours (1 day), and 48 hours (2 days)—to capture both immediate and delayed responses and smooth short-term fluctuations with a 3-hour mean temperature, and assess the minimum and maximum temperatures over the past 24 hours to gauge extremes, and analyze long-term temperature trends to identify seasonal patterns. These factors, through 9 selected temperature variables, are essential for understanding their combined impact on electricity demand and improving load forecast accuracy.

3 Methods

In this section, we provide an overview and description of the various methods used in this project, as well as the software utilized for their implementation.

3.1 Auto-Regressive with Exogenous Input

The Auto-Regressive with exogenous inputs (ARX) (Ziel & Weron, 2018) model is a type of time series model that incorporates both past values of the time series itself (auto-regressive components) and past values of external (exogenous) variables (regressors) to predict future values. The ARX model is commonly used in forecasting and can be particularly useful when the future behavior of the time series depends not only on its own past values but also on external factors. The ARX model can be represented by the equation 3.1 below:

$$L_t = \phi_0 + \phi_1 L_{t-24} + \phi_2 L_{t-48} + \phi_3 L_{t-144} + \phi_4 L_{t-168} + \phi_5 D_t + \varepsilon_t, \tag{3.1}$$

where L_t is the target variable (load data) at time t, L_{t-i} are the lagged electricity load for i = (24,48,144,168), D_t is the weekend dummy which is 0 for weekdays and 1 for weekends, ϕ_i are the coefficients for the auto-regressive and external terms, ε_t is the white noise error term. To generate forecasts, we fit a linear regression model using load lags and a weekend dummy variable. We extract the estimated coefficients and residuals and generated 10,000 random normal errors. We then loop over the forecast horizon to generate forecasts by adding these errors to the predicted values from the regression model.

3.2 Generalised Additive Model for Location, Scale and Shape

Generalised Additive Model for Location, Scale and Shape (GAMLSS) (Rigby & Stasinopoulos, 2005) is designed for fitting regression models. GAMLSS contains a wide range of highly skewed and kurtotic continuous and discrete distributions. GAMLSS model can estimate all parameters of the dependent variable's distribution such as location, scale, and shape—as linear, non-linear, or smooth functions of the explanatory variables. GAMLSS first proposed by Rigby and Stasinopolous (2005) as an extension to the Generalized Linear Models. One of the biggest advantage of using GAMLSS instead other classical approaches are that the dependent (response) variable is not restricted to follow exponential distribution family, and allows modelling scale and shape parameters.

Denoting Y as a response variable $(y_i \;\; i = 1, 2, ..., n)$ that are conditionally independent given a set of covariates, which follow a specific distribution $D(y_i; \theta_i)$, where θ_i is a set of parameters vector $\theta_i = (\mu_i, \sigma_i, \nu_i, \tau_i)$. A GAMLSS model can be written as

$$g_k(\theta_k) = X_k\beta_k + \sum_{j=1}^{J_k} s_{jk}(x_{jk}), \quad k = 1, \ldots, p, \tag{1}$$

where $g_k(.)$ is known as link function which is determined by the distribution parameter θ_k, X_k is design matrix, β_k is the parameters vector which is associated to X_k and s_{jk} is smooth functions (Eilers & Marx, 1996) of an explanatory variable x_{jk}. The parametric vectors β_k and the second term parameters (random effects γ_{jk}, for $j = 1, 2, \ldots, J_k$ and $k = 1, 2, 3, 4$), are estimated by maximizing the penalized log likelihood function ℓ_p given by

$$\ell_p = \sum_{i=1}^{n} \log f(y_i|\theta_i) - \frac{1}{2}\sum_{k=1}^{p}\sum_{j=1}^{J_k} \lambda_{jk}\gamma_{jk}^{\top}G_{jk}\gamma_{jk}$$

where G_{jk} is a symmetric matrix that depends on a vector of smoothing parameters λ_{jk}. The penalized log likelihood approach for estimating in the context of GAMLSS modifies the conventional back-fitting algorithm to achieve better estimates of the distribution parameters. This method is implemented in the R package gamlss.

3.3 Distributional Neural Networks

3.3.1 Modeling

The distributional neural networks (DDNN) will be utilized in this research with the aim of predicting the entire distribution of a future variable Y given input features X. The difference between this DDNN and deep neural network models lies only in the output layer. Instead of a point forecast, the output of the distributional neural network model will be the parameter layer $\Theta \in \mathbb{R}^{D \times S.P}$ consists of P distribution parameters for each of the S output features with $S = 168$ representing the 168 hourly load predictions for the next 7 days for Spain.

The optimization problem using DDNN aims at minimizing probabilistic losses. During training, the parameters Θ are adjusted to minimize the negative log-likelihood (NLL). A lower NLL indicates a better fit of the predicted distribution to the observed data. This process ensures that the model's predicted distributions are as close as possible to the actual data distributions. The optimization problem can be formulated as: $\hat{\Theta} = \arg\min_{\Theta} \left(-\sum_{i=1}^{n} \log f(L_i; \Theta, x_i)\right)$ where the $-\sum_{i=1}^{n} \log f(L_i; \Theta, x_i)$ is the negative log-likelihood of the observed data given the predicted distribution parameters. To mitigate the overfitting risk, regularization is applied for the loss function (Marcjasz et al., 2023):

$$
\begin{aligned}
\mathcal{L}_{\text{reg}}(L; F(\Theta; x)) = \mathcal{L}(L; F(\Theta; x)) &+ \sum_{i=0}^{I-1} \lambda_{1,i} \|H_i\|_k + \sum_{i=0}^{I-1} \lambda_{2,i} \|W_{i+1}\|_k + \sum_{i=0}^{I-1} \lambda_{3,i} \|b_{i+1}\|_k \\
&+ \sum_{p=0}^{P} (\lambda_{1,I,p} \|H_I\|_k + \lambda_{2,I,p} \|W_{I+1}\|_k + \lambda_{3,I,p} \|b_{I+1}\|_k)
\end{aligned}
\tag{3.2}
$$

In which: $\mathcal{L}_{\text{reg}}(L; F(\Theta; x))$ represents the regularized loss function of DDNN between the load and the predicted distribution. $\mathcal{L}(L; F(\Theta; x))$ is the original loss function of load and the predicted distribution. H_i denotes the output of the ith hidden layer. W_{i+1} and b_{i+1} represents the weights and bias of the $i + 1^{th}$ layer. $\lambda_{1,i}$, $\lambda_{2,i}$, $\lambda_{3,i}$ are the regularization parameters for the hidden layer output, weights, and biases, respectively.

3.3.2 Hyperparameter tuning

With the aim of minimizes the negative log-likelihood (NLL)the hyperparameter tunning also emplyed with the help of Optuna in Python. The key aspects of the model and the hyperparameters tuning and its results are as follows:

- Features: Load DA lag 1, 2, and 7, and Spanish holidays (holidays after selection and calculated in Section 2.3). The DDNN will not consider weather features. - Hidden Layers: The model consists of two hidden layers. Each layer's number of neurons ranges from 16 to 128 to balance complexity and performance. The tuning results reveal that the hidden layer for Normal distribution (ND) is $(71, 53)$ and JSU is $(75, 55)$ - Learning Rate: Learning rates are explored from 10^{-5} to 10^{-1} on a logarithmic scale to cover a wide range. After tuning, the learning rate for ND is 0.091 and for JSU is0.0026 - Number of Epochs: Fixed at 100 to balance the model's learning capacity and computation. - Batch Size: Set to 32, which is a common choice for effective training. - Dropout Rate: Explores rates between 0 and 1 to prevent overfitting. the JSU need the drop out layer with the rate of 0.0055, while Normal no need for drop out layers - Regularization Terms: Applies L1 regularization to activation and kernel weights with rates from 10^{-5} to 10 on a logarithmic scale to reduce overfitting by penalizing large weights. - Activation Functions: Tests various activation functions including ReLU, ELU, Sigmoid, Tanh, Softplus, and Softmax to identify the best fit for the forecasting task. For JSU, the ReLU is utilized and for ND the Softplus is used

3.4 Pinball score

To evaluate the accuracy of forecasts in probabilistic forecasting, a pinball score is employed. The pinball score is a metric used to assess the accuracy of a quantile forecast. It measures the average deviation of the predicted quantiles from the observed values. Let α be a target quantile, $L_{d,s}$ be a target value and $L_{d,s}^{\alpha}$ is the quantile forecast. The Pinball Score at quantile level α for a single forecast time point at hour s on day d is given by:

$$\text{Pinball}^{\alpha}_{(L_{d,s},\hat{L}_{d,s})} = \begin{cases} \alpha(L_{d,s} - \hat{L}_{d,s}) & \text{if } L_{d,s} \geq \hat{L}_{d,s} \\ (1-\alpha)(\hat{L}_{d,s} - L_{d,s}) & \text{if } L_{d,s} < \hat{L}_{d,s} \end{cases} \tag{3.3}$$

To provide a comprehensive measure of model performance, the pinball score is averaged across various quantiles. This averaging yields an aggregate pinball score that summarizes the model's performance across multiple quantiles and over different time periods. For this research, the pinball score is calculated daily by taking the mean of 7-day ahead forecasts across quantiles ranging from 0.05 to 0.95 with the step of 0.05 and considering all 24 hours within the target day. By using a sufficiently dense grid of quantiles, this average converges to the Continuous Ranked Probability Score (CRPS), which offers a robust measure that captures both the accuracy and the sharpness of probabilistic forecasts.

4 Results

4.1 Parameter Selection

In the context of GAMLSS, gamlss.lasso (Ziel Florian, 2022) implements Adaptive Lasso (Least Absolute Shrinkage and Selection Operator) and its variants to perform feature selection. The Adaptive Lasso method, in particular, improves upon the standard Lasso by applying different penalization weights to different coefficients based on initial estimates. The tuning parameters are chosen by minimizing the Bayesian information criterion (BIC). First, we included all external variables such as HoD, HoW, HoY, HolFix, HolFlex, trend, TempLag1, TempLag2, TempLag3, Temperature mean of the last 3 hours, maximum and minimum temperatures of the last 24 hours, lags of load and interactions between HoW, HoD, and HoY. We repeated this process every 30 days to fit the model. According to our parameter selection results, HolFix, HolFlex, XHol, HoW,HoD and Load Lag 24 and 168 were consistently chosen in each parameter tuning. Additionally, HoY, TempLag1, TempLag3, and TempMin and TempLag24 during winter months, and TempMax and TempMean during summer and spring months. Due to page limit constraints, we do not include the detailed parameter tuning results for each distribution parameters.

4.2 Pinball score

Table 2 : Average Pinball Loss Scores on Test Set

Model	GAMLSS.NO	GAMLSS.JSU	GAMLSS.ST	ARX	DDNN.NO	DDNN.JSU
PB score	0.5426	0.5392	0.5304	2.18	1.11	1.89

In this study, we evaluated the performance of various models for probabilistic load forecasting using the pinball score as a key metric. Among the models tested, the GAMLSS models, particularly the GAMLSS model to fit Skewed distribution (ST), demonstrated superior performance with the lowest pinball score of 0.5304 GW. This indicates that the GAMLSS.ST model provides the most accurate probabilistic forecasts, effectively capturing the skewness in the load distribution. The GAMLSS.NO model to fit normal distribution and GAMLSS.JSU model to fit JSU distribution also performed well, with pinball scores of 0.5426 GW and 0.5392 GW, respectively, highlighting their capability in handling different distribution assumptions for load forecasting.

On the other hand, the ARX model yielded a significantly higher pinball score of 2.18 GW, suggesting that it is less effective in capturing the probabilistic nature of load forecasts compared to the GAMLSS models. The DDNN models showed moderate performance, with pinball scores of 1.11 GW for the normal distribution and 1.89 GW for the JSU distribution. These results indicate that while DDNN can handle complex patterns, their probabilistic accuracy in this context is not as good as the GAMLSS models.

4.3 DM test

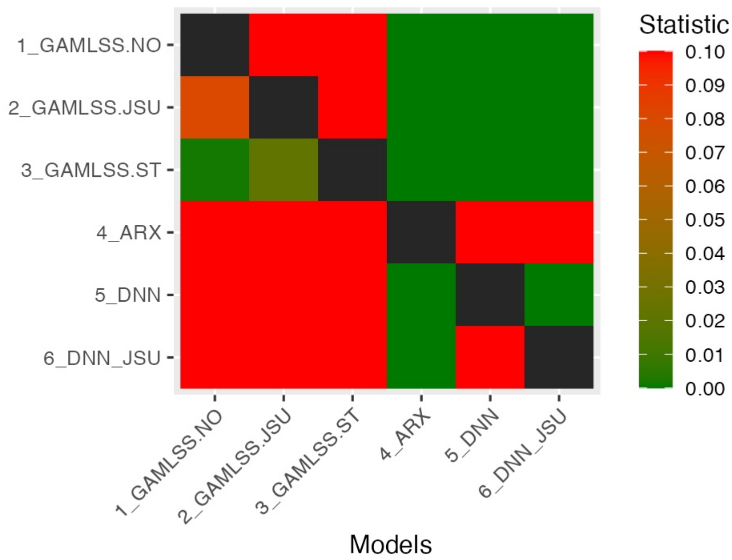

Figure 3: DM test results

In addition to the pinball score, we also conducted DM tests across various models to have more statistics prove which model outperforms others in probabilistic Spanish load forecasting. The outcome is visually depicted through a heat map in Figure 3. These maps illustrate the p-values associated with the null hypothesis of the forecasts. The low p-values (green colour) show the statistical significance that the model along the y-axis is significantly more accurate than the forecast along the x-axis. On the other hand, shades closer to red (p-values approaching 0.1) suggest weaker evidence against the null hypothesis. The results reveal that the GAMLSS models exhibit inferior performance compared to others especially for fitting Skewed distribution. DDNN for both Normal distribution and JSU distribution outperforms the ARX model.

5 Conclusion

In this study, we addressed the challenge of probabilistic load forecasting for Spain, focusing on the holiday effect and temperature variations. With the utilizing of advanced models like

DDNN and GAMLSS, we aimed to improve accuracy and account for key factors like holidays and temperature fluctuations that affect energy demand. In conclusion, our study demonstrates the effectiveness of the GAMLSS models for probabilistic load forecasting. By incorporating holiday effects, extra calendar data, and temperature variations, the GAMLSS model for fitting Skewed distribution showed superior performance, offering a robust approach for managing the complexities of energy demand forecasting during both regular and holiday periods. Our findings underscore the importance of these factors for enhancing forecast accuracy and informing energy management decisions.

For future research, we plan to incorporate additional weather data and optimize the DDNN model by including more holiday effects, and weather data, and exploring better hyperparameter settings to achieve improved results.

References

Eilers, P., & Marx, B. (1996). *Flexible smoothing with b-splines and penalties.* (Vol. 11).

Enric, V., Vicente, M., & Caselles, V. (2001). Daily air temperature and electricity load in spain. *Journal of applied Meteorology*, *40*, 1413–1421.

expatica. (2024). School holidays in spain: 2023 and 2024. https://www.expatica.com/es/education/children-education/school-holidays-in-spain-78291/#overview

Marcjasz, G., Narajewski, M., Weron, R., & Ziel, F. (2023). Distributional neural networks for electricity price forecasting. *Energy Economics*, (125).

Monika, Z., & Ziel, F. (2024). Efficient mid-term forecasting of hourly electricity load using generalized additive models. *arXiv*.

Rigby, R. A., & Stasinopoulos, D. M. (2005). *Generalized additive models for location, scale and shape.* (Vol. 51).

Ziel, F., & Weron, R. (2018). Day-ahead electricity price forecasting with high-dimensional structures: Univariate vs. multivariate modeling frameworks. *Energy Economics*, *70*, 396–420.

Ziel, F., & Liu, B. (2016). Lasso estimation for gefcom2014 probabilistic electric load forecasting. *International Journal of Forecasting*, *32*(3), 1029–1037.

Ziel Florian, S. M., Muniain Peru. (2022). Gamlss.lasso: Extra lasso-type additive terms for gamlss. https://CRAN.R-project.org/package=gamlss.lasso