Generation Deviation Penalties Forecast

Daniel Garcia Arana

***Abstract*—The electricty price in Spain is determined by auction markets. The Spanish electricity market operates through a structured system of day-ahead, intraday, and balancing markets, where energy producers commit to generation schedules and face financial penalties for deviations. These penalties vary largely, as they play a crucial role in ensuring grid stability, they are key to incentivizing accurate generation forecasting [1]. Some of the genaration plants that are specially susceptible to forecasting errors due to its dependence on variable meteorological conditions are solar power plants [2]. This report examines the deviation penalty framework in the Spanish electrical grid, with a focus on solar power generation. The deviation penalties are asymmetric and depend on the direction of the deviation relative to system needs: during shortages underproduction is penalized more severely than overproduction and viceversa. Deviations that support grid stability might even be financially rewarded in some cases. Forecasting accuracy is essential to minimize economic losses and make the solar power plant economically viable, optimizing market participation [3]. This study underscores the critical role of deviation penalties in shaping market behavior and emphasizes the need for investment in robust forecasting technologies to enhance profitability and reliability for renewable energy producers in Spain's competitive electricity market.**

***Index Terms*—IEEE, Forecast, Multivariate, Electricity market.**

I. Introduction

THIS paper focuses on forecasting the deviation penalties in the spanish electricity market. The Spanish electricity market operates through a sophisticated structure designed to balance supply and demand while integrating an increasing share of renewable energy. At its core, the system can be enumerated in 3 markets [4]:

1) Day-ahead market, which accounts for 78% of annual transactions via daily auctions
2) Intraday market, which accounted for 22% of day-ahead volume in 2023 through six auctions and continuous trading. In some cases the continuos market is accounted as the seventh market. The future plans for this market was to have it reduced to 3 markets.
3) Balancing market, which adresses real-time mismatches. It is managed by grid operator Red Eléctrica de España (REE).

For solar power generators, participation begins with day-ahead commitments submitted by 12:00 of the previous day, followed by intraday adjustments - a critical window where forecasting accuracy directly determines financial viability.

Deviation penalties in Spain are correlated to grid needs. During system shortages, underproduction incurs high penalties, whilst overproduction may even be compensated. This contrasts with scenarios where deviations align with grid requirements - for instance, reducing output during oversupply - which can generate financial rewards. The 2022 regulatory reforms intensified pressure through narrower tolerance bands, allowing just 5% overproduction before penalties apply, compared to 15% under previous rules [5]. One of the reasons being the high unestability of the electrical grid caused by the abundance of renewable energies, which have very little inertia, or none in case of photovoltaic plants. In fact this is believed to be the root of the blackout which occurred on April 28, 2025 [6].

Accurate solar power forecasting is highly challenging due to irradiance variability and cloud cover uncertainties. A 2024 case study revealed that forecasting errors exceeding 10% for photovoltaic plants triggered average penalty costs of €0.037/kWh [7]. These penalties compound with market risks: during the 2022 energy crisis, Spain extended windfall profit clawbacks to June 2022 for unhedged generation [8], incentivizing precise day-ahead commitments.

May 1, 2025

II. Approach/Methods/Materials

The data and code are available at https://github.com/dgarcia2025/Deviation_penalties. It was obtained via API from [9]. The software used to implement the code to complete the analysis has been Python version 3.10.5. The following libraries have been used: Scikit learn and Scipy.

The terminology used:

- 'Up penalization': The penalization for overproduction. In other words, penalization for under forecasting.
- 'Down penalization': Opposite of 'Up penalization'.
- 'Short state': The penalizations for under forecasting are lower than the ones for over forecast.
- 'Long state': Opposite of 'short state'.

Instead of focusing on each penalization separately (Up penalization and Down penalization), the models will be trained to forecast the difference between the Up and Down penalization. In other words, the subtraction of the Down penalization to the Up penalization.

The evaluation of the results may be difficult, since a model with high MAE may perform better than one with lower one when it comes to classifying if the state is short or long.

Department of Electrical and Computer Engineering, Florida Institute of Technology, Florida, FL, 32901 USA
Professor Anthony Smith

Hour	Long (%)	Short (%)	Neutral (%)
1	43.99	29.39	26.62
2	42.05	29.78	28.17
3	45.22	38.11	16.67
4	47.61	42.51	9.88
5	46.25	43.54	10.21
6	45.22	44.64	10.14
7	47.42	44.51	8.07
8	53.55	39.47	6.98
9	62.86	31.52	5.62
10	64.6	29.97	5.43
11	64.73	31.07	4.2
12	57.36	35.79	6.85
13	52.91	41.02	6.07
14	49.61	43.09	7.3
15	46.64	45.28	8.08
16	45.61	46.58	7.81
17	45.61	46.96	7.43
18	48.71	45.03	6.26
19	53.75	40.37	5.88
20	60.98	33.79	5.23
21	64.86	28.68	6.46
22	64.99	27.33	7.68
23	66.6	28.49	4.91
24	65.89	29.39	4.72

TABLE I: Average classification from 01/01/2021 to 03/28/2025

The following models have been trained to apply a multivariate analysis:

- Least Absolute Shrinkage and Selection Operator (LASSO) (Statistical Model)
- Bayesian Regression (Statistical Model)
- FeedForward Neural Network (Machine Learning Model)

It might be tempting to think that a statistical model such as SARIMAX could work, nonetheless such approax would give similar results to randomly trying to guess the state, as will be proven forward in the paper.

In addition to this, it is important to take into account that the forecasting of these values plays an important role during mid-day hours, when the generation coming from solar power is high and unpredictable, and knowing the penalization values plays a crucial role for companies to save thousands of euros. Then again, a model may perform very well overall, but won't neccesarily mean it's a good model, since it could perform poorly when it comes to forecasting and classifying correctly in the important hours of the day. It stands obviuos from table I that some hours are easier than others to classify.

It is also important to point out that some variables might hold a heavy role in the model and could be very relevant when it comes to forecasting yet they may not always be present. For example, the expected amount of energy generation from fuel is a heavy indicator of which class the difference might take, however, the amount of days in which the fuel generation differs from none are very few, thus the model would not be valid to use with good accuracy in many occasions.

The goal is to train the models to forecast the numeric value value of the difference between Up and Down (Up - Down), and then evaluate the accuracy in predicting short or long states.

As seen in the table I, it is not unusual for this difference to equal 0. This cases will not be taken into account to evaluate the accuracy in classification of states. The reason being when this happens commonly both are noticeably low, and generally do not end up causing a relevant economic impact. Be noted that the gravity of economic penalties electric market companies have to face is not uniform, two days of unlucky forecasting and falling in the wrong state can mean over 50% of the economic impact they have to face the whole month.

A. Lasso

Lasso (Least Absolute Shrinkage and Selection Operator) is a regression analysis method that combines variable selection and regularization to improve model accuracy and interpretability. One of the key characteristics is it applies an L1 penalty to the regression coefficients, which shrinks some coefficients to zero, effectively excluding irrelevant features [10]. Below is a detailed breakdown:

1) Formulation and Core Mechanism: Lasso enhances ordinary least squares (OLS) by incorporating a penalty term that depends on the sum of the absolute values of the model's coefficients, scaled by a factor λ. The aim is to minimize both the errors (residual sum of squares, or RSS) and the magnitude of the coefficients, which helps create a simpler model and reduces overfitting. The objective function is:

$$\text{Minimize: RSS} + \lambda \sum_{j=1}^{p} |\beta_j| \tag{1}$$

Here, RSS represents the model's errors, β_j are the coefficients, and λ determines the strength of the penalty on the coefficients.

2) Key Properties:

- Sparsity and feature selection: Lasso can decide which features are irrelevant and set their coefficients to 0, performing automatic feature selection.
- Handling correlated predictors: While Lasso may arbitrarily select one variable from a correlated group, extensions like elastic net address this limitation.
- Bias-variance tradeoff: Bias and variance can be calibrated.

3) Parameter Tuning and Interpretation:

- Cross-validation: Commonly used to select λ by balancing prediction accuracy and model simplicity.
- Bayesian interpretation: Lasso corresponds to using Laplace priors on coefficients, emphasizing sparsity.
- Effective degrees of freedom: Measures model complexity by counting non-zero coefficients.

4) Applications: Lasso is a common choice when dealing with high-dimensional series, specially when relevance of some datasets is not certain. For this reason, it is also used for mitigating overfitting in models with many correlated predictors [12].

B. Bayesian Regression

Bayesian regression is a model used to build linear models that use probabilities to include prior knowledge and measure how uncertainty of results. Unlike traditional methods that give a single "best guess" for the model's parameters, Bayesian regression treats these parameters as uncertain quantities. It updates their weihgts more data is taken into account, offering a range of possible values (posterior distributions) instead of just one answer. Here's a detailed explanation [13]:

1) Formulation and Core Mechanism: In Bayesian linear regression, the relationship between predictors X and response y is modeled as:

$$y \sim \mathcal{N}(X\beta, \sigma^2 I) \tag{2}$$

where β are the coefficients and σ^2 represents the noise. The posterior distribution $P(\beta, \sigma^2|X, y)$ combines likelihood and priors via Bayes' theorem:

$$P(\beta, \sigma^2|X, y) \propto P(y|X, \beta, \sigma^2) \cdot P(\beta) \cdot P(\sigma^2) \tag{3}$$

2) Key Properties: Bayesian regression interprets model parameters, such as coefficients β and noise variance σ^2, as random variables with prior distributions that reflect our initial beliefs or assumptions. Applying Bayes' theorem, these priors are updated with observed data through the likelihood function, resulting in posterior distributions. When priors are compatible, the posterior can be computed analytically, but for non-conjugate cases, methods like Markov Chain Monte Carlo (MCMC) need to be used. To make predictions the predictive posterior distribution is used, which compensates for uncertainty. Bayesian regression naturally regularizes through priors, provides uncertainty estimates via credible intervals, and performs well when data is limited or when strong prior knowledge is available [13].

3) Parameter Tuning and Interpretation: In Bayesian regression, uncertainty is measured in coefficient estimates using *credible intervals*, such as 95% highest posterior density (HPD) intervals, which tell us the range where the true parameter likely lies. Other estimates like *posterior mean*, *median* or *posterior variance* are also common estimates which provide feedback on point estimates or estimation confidence. For predictions, *predictive distributions* give us probabilistic forecasts, accounting for uncertainty in both the parameters and the data noise.

4) Applications: Bayesian regression shines when the size of datasets is small, hierarchical models, or problems requiring explicit uncertainty quantification. It is a flexible model which makes adaptable to many domain-specific constraints [14].

C. Deep Learning

Deep learning is a branch of machine learning that uses multi-layered artificial neural networks (ANNs) to model complex patterns in data. Below is a structured explanation of its core concepts, properties, and tuning processes.

1) Formulation and Core Mechanism: Deep learning relies on artificial neural networks (ANNs) inspired by the human brain's neuronal structure. These networks consist of [15]:

- Input layer: Receives raw data
- Hidden layers: Process data through weighted connections and non-linear activation functions (e.g., ReLU, sigmoid), extracting hierarchical features. Each layer refines representations, enabling recognition of intricate patterns
- Output layer: Produces predictions

In deep learning, forward propagation feeds data through neural network layers to predict outcomes, while backpropagation adjusts weights via gradient descent to reduce errors. Deep networks, with potentially hundreds of layers, model complex, non-linear relationships, outperforming shallower architectures.

2) Key Properties:

- **Handling raw data**: Specially useful when it comes to processing unprocessed, unlabeled data like text or images, without needing manual feature design.
- **Automatic feature discovery**: Learns key features on its own through hidden layers, such as detecting edges in images or understanding text relationships.
- **Scalable performance**: Gets better as datasets and network complexity increase.
- **Unsupervised learning**: Finds hidden patterns in unlabeled datas [16].

3) Parameter Tuning and Interpretation: Hyperparameters:

- Architecture: Number of layers/nodes, activation functions
- Training: Learning rate, batch size, epochs, regularization
- Optimization: Momentum, weight decay

Challenges:

- **High computational demand**: Training deep networks needs powerful GPUs and distributed systems for efficient processing.
- **Lack of transparency**: The inner workings of feature transformations are often unclear, giving deep learning a "black box" reputation, though methods like attention visualization and saliency maps help shed light on them.

Deep learning has led to major AI breakthroughs, like better image recognition and creative generative models. Still, figuring out how to keep it high-performing while making it easier to understand is a topic in research [17].

4) Applications: Deep learning is widely used for pattern recognition, prediction, and data generation across various domains, enabling machines to interpret complex data, automate decision-making, and learn from large datasets without explicit programming [18].

III. Experiment/Results

The data was split for all models was 70 % train (55% train + 15% validation in case of ANN) and 30 % test.

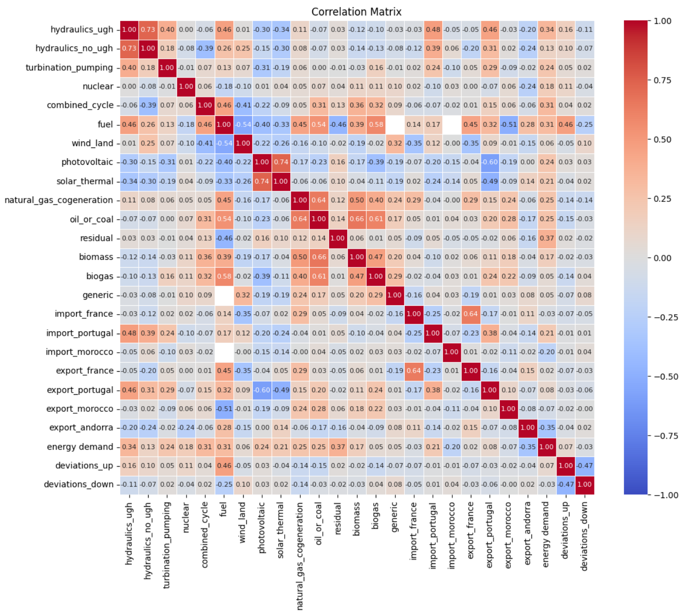

Fig. 1: Correlation Matrix

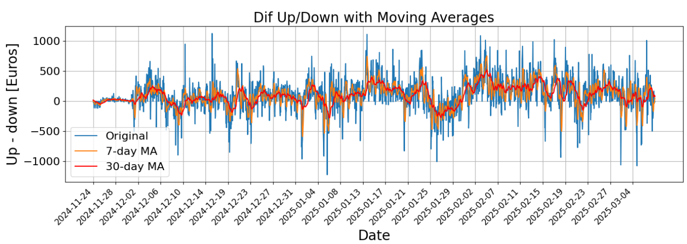

Fig. 2: Moving averages (Up - Down)

A. Exploratory Data Analysis

Figure 1 shows the correlation matrix of the data.

Figure 2 displays the moving averages for the values of the substraction of the penalization Down to the penalization Up.

Figure 3 shows a plot of the average values of the difference by hour.

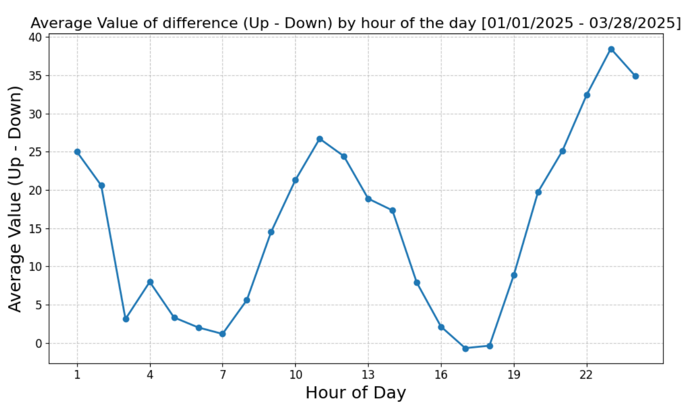

Fig. 3: Average values per hour (Up - Down)

The STL decomposition for the 2024 year data is shown in figure 4. The two seasonalities applied were the most relevant ones according to the fourier analysis, 12 hourly and daily (24 hours). In the discussion chapter we discuss why this figure demonstrates that a SARIMAX model would not correctly apply.

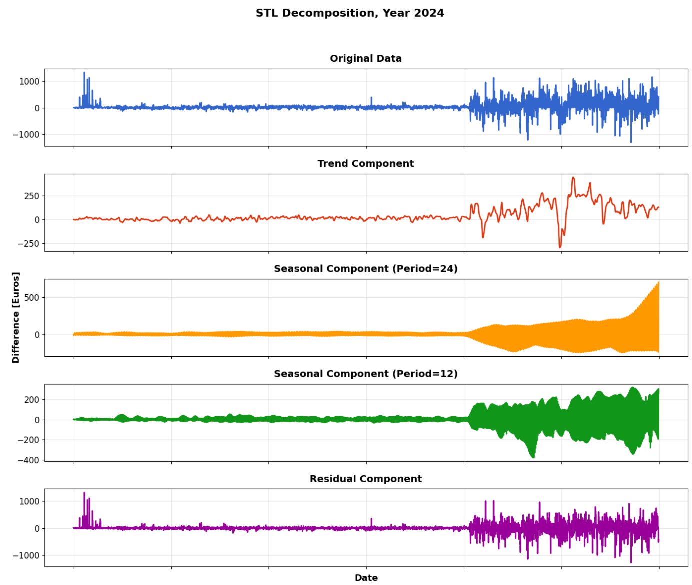

Fig. 4: STL decomposition

B. Lasso Model

The model took 1.9 seconds to train and test. The overall performance of the Lasso model is appreciated in figure 5.

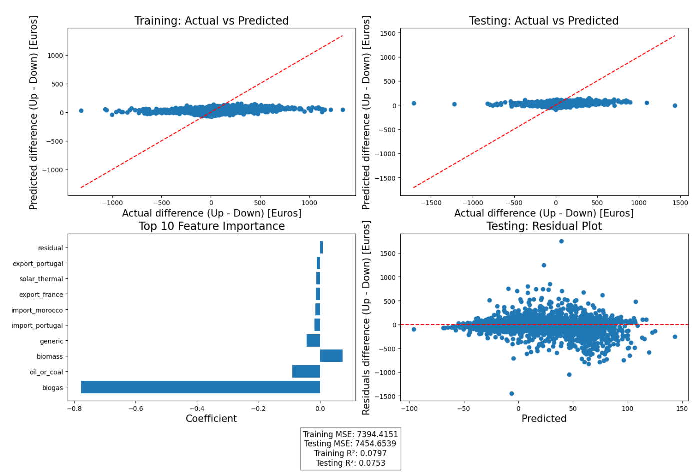

Fig. 5: Lasso Model results

The accuracy in classification by hour is shown in figure 6.

The best way to evaluate the model is by observing the table in figure 7.

C. BayesianRidge

The model needed 18.1 seconds to train and test. The metrics of the Bayesian Ridge model are shown in figure 8.

The performance in classification per hour as well as by per percentile range can be appreciated in the plots in figure 9.

Same way as in the Lasso model, the performance in forecasting classification by percentile can be observed through figure 10.

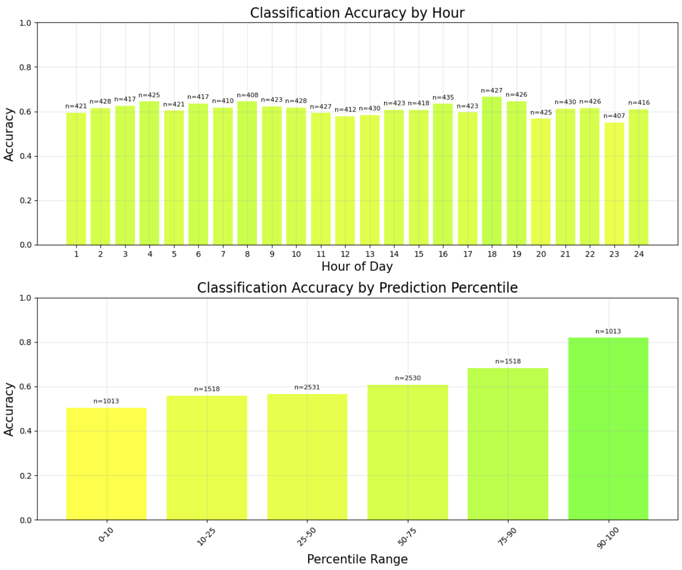

Fig. 6: Lasso Model: Accuracy by hour and percentile forecast - Testing set

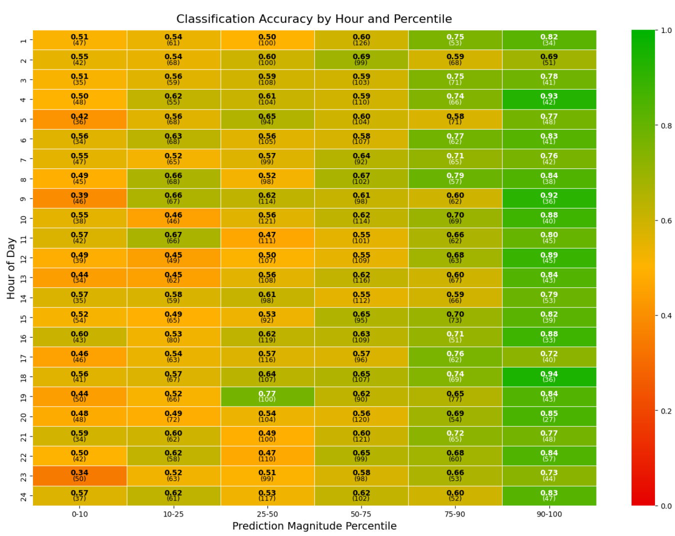

Fig. 7: Lasso Model performance by percentile forecast - Testing set

D. Deep Learning Model

The model architecture consisted of two hidden layers, with 64 and 32 neurons. A 20% dropout rate was applied to avoid overfitting. The activation function used was Rectified Linear Unit. Batch size was 32. The optimizer was 'Adam', and the monitor for early stopping was validation loss.

It took 2 minutes and 23.2 seconds to train and test the model. The evolution of the loss and MAE over epochs is shown in figure 11.

The performance of the deep learning model on the test set is displayed in figure 12.

The plot of residuals is shown in figure 13. Figure 14 shows the distribution of residuals applied to the test set.

The accuracy of this model by hour and by percentile range is shown in figure 15.

The performance in forecasting state classification by percentile can be observed through figure 16.

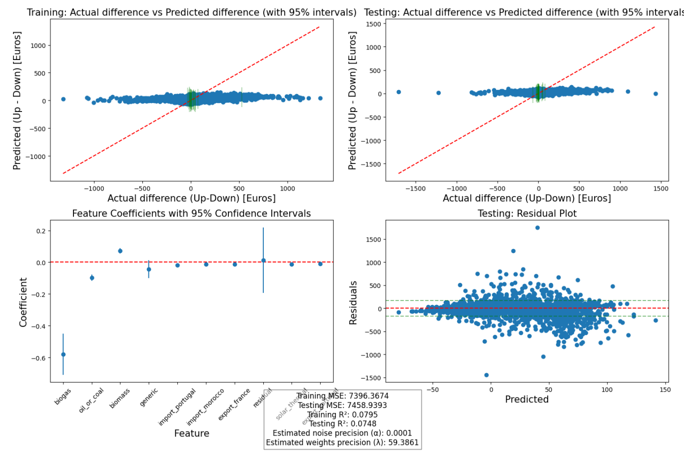

Fig. 8: BayesianRidge Model results

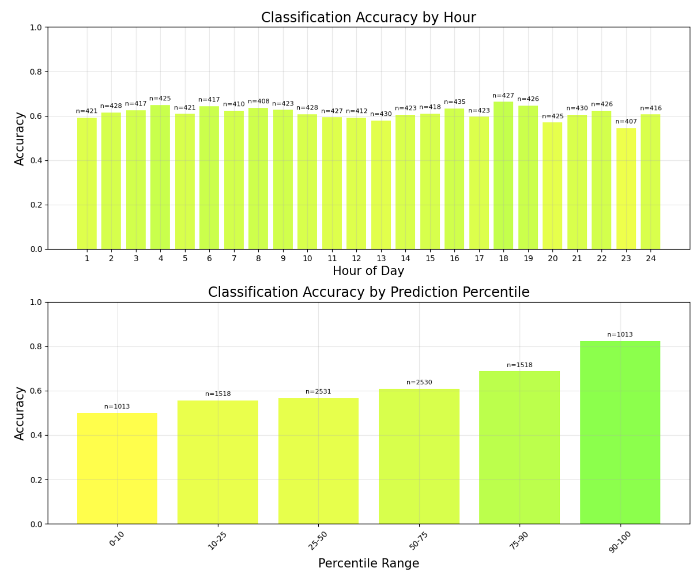

Fig. 9: BayesianRidge Model: Accuracy by hour and percentile forecast - Testing set

IV. Discussion

The classification of the state of the system is highly unpredictable and unstable, as we can see in figure 2. Specially during mid day hours (from 12 to 19), as they are influenced by the uncertainty of solar power generation. As solar power generation is reduced, the system becomes stable and the state is easier to predict, as we know from table I.

We can conclude from figure 4 that SARIMAX would not work succesfully for this time series data for two reasons:

1) The most relevant frequencies are 12 and 24 hours. The predictions need to be made more than 24 hours in advance, thus the two most relevant seasonalities cannot be taken iinto account, and without these, seasonality decomposition becomes an impossible challenge.

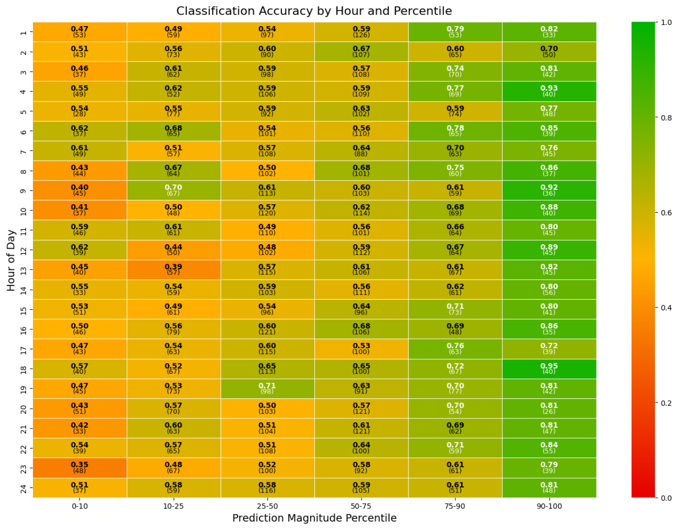

Fig. 10: BayesianRidge Model performance by percentile forecast - Testing set

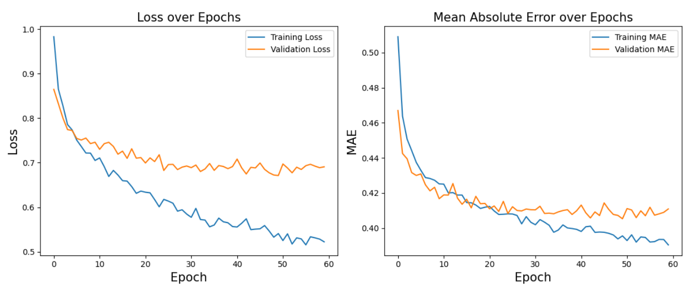

Fig. 11: Evolution of the training parameters - Deep Learning

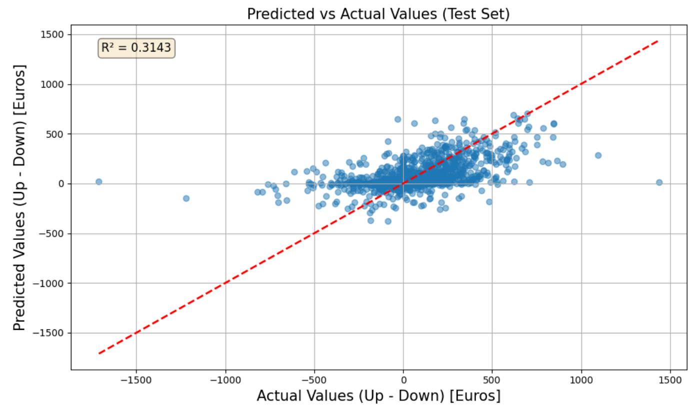

Fig. 12: Deep Learning Model correlation performance

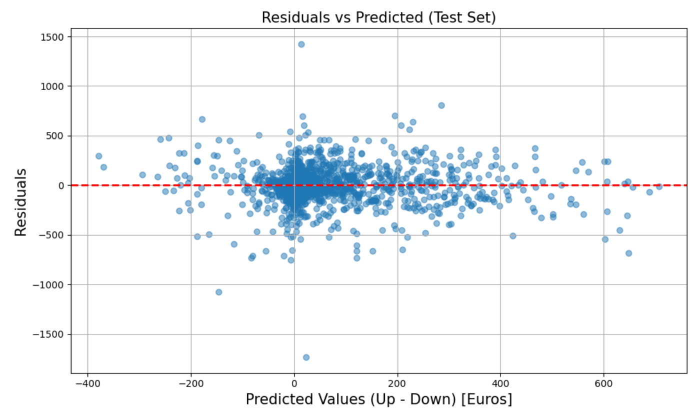

Fig. 13: Deep Learning model - Residuals plot

2) The seasonalities nor the trend are smooth or show consistency, proving a poor decomposition.
3) Residual values are very high relative to trend or seasonalities, suggesting too much variance.

A. Lasso Model

Figure 5 shows some very pesimistic results regarding the Lasso model. The testing R^2 has a very low value, just 0.0673, suggesting no linearity between the predictions and true data. Morover, we also learn from the figure that the two features with highest importance are biogas and residuals, which are not features you can count on to appear every day, meaning the model would likely underperform in days when these features are not present.

However, when it comes to classification accuracy, the model is consistent, and classifies state correctly over 60% of the times, as can be appreciated in figure 6. In addition to this, the accuracy in classification increases as the model forecasts values in high percentile, meaning the model correctly forecasts the state around 80% of the times the forecast is in the 90% percentile high. Thus, althought not viable for many of the days, the model can be useful when the forecast is very indicative. Figure 7 confirms that this performance is uniform accross all hours of the day.

B. Bayesian Ridge Model

Lasso model likewise, the first results are not very optimistic. Figure 8 makes it evident that the variance between the predicted and actual values is noticeably high. With a testing R^2 value of 0.067 it safe to say that the relation between the predicted and actual values is almost non-existent.

Nonetheless, figure 9 shows that the performance in classifying the state is not too bad. The accuracy seems to be consistent throughout the hours of the day, averaging around 60% accuracy. Although the performance is poor compared to simply taking the average percentage of times a state takes state Up or Down values, we see that most of the accuracy is lost for non mid-day hours, which are not the relevant ones, since the solar generation will equal 0. In addition to this, the figure also shows that as the percentile range increases the accuracy does too, also supported by figure 10

C. Deep Learning Model

It took about 50 epochs for the deep learning model to finish training (figure 11). The validation MAE is similar to the training MAE, a good sign that there is no overfitting.

We can appreciate from figure 12 that the correlation between actual values and the predicted values is relevant,

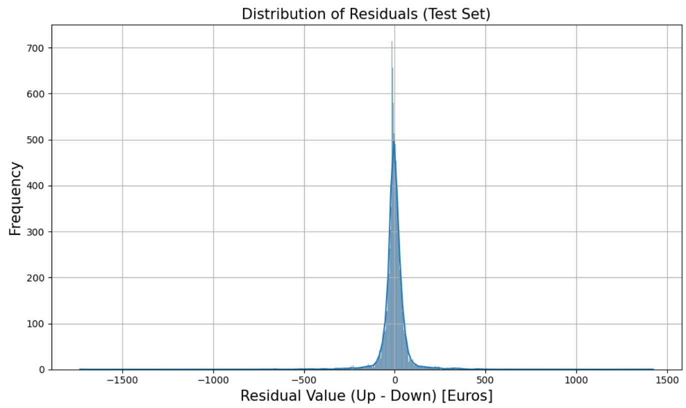

Fig. 14: Deep Learning model - Distribution of Residuals

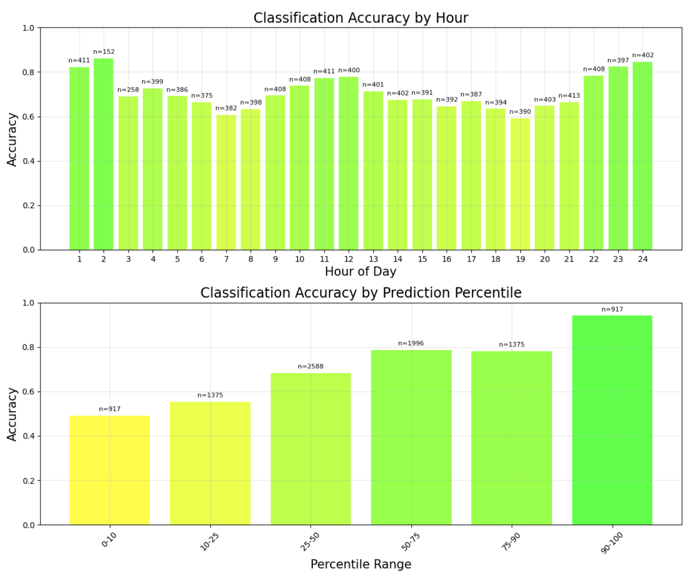

Fig. 15: Deep Learning Model: Accuracy by hour and percentile forecast - Test set

with an R^2 value of 0.3143. Despite existing some very high residuals which may be worrying, most of the residuals are close to 0 (figure 13). The distribution of residuals seems to follow a normal distribution (figure 14). A good sign, since it means that most of the residuals are close to zero and the model rarely has a forecast with a highly significant absolute error. If the plot did not have normal distribution shape it could be an indicative that a key variable is missing.

Figure 15 provides some very interesting information. First of all, we learn that the accuracy of state prediction is not constant through all the hours of the day. The model has very high accuracy in the hours 1, 2, 22, 23 and 24, correctly predicting the state over 80% of the times. Nonetheless, those hours are not important, and the hours that do count don't have an accuracy that high, with accuracy values slightly higher than 60% in some hours, like 18 and 19.

We also see in figure 15 a very high increase in accuracy as the percentile range augments, even surpassing 90% accuracy in the 90-100 % percentile. This being a very good indicator that the model is trustable.

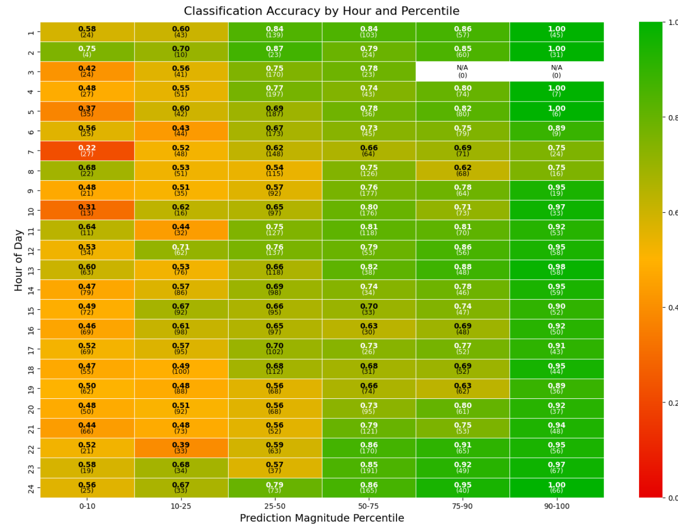

Fig. 16: Lasso Model performance by percentile forecast -Test set

Figure 16 shows that the statement that the model performs specially well during non mid-day hours holds true even with percentile ranges. Overall, the performance is great for the high percentile ranges, and commonly performs poorly when the forecast is in the low percentiles. This makes sense, since we see in the plot in figure 2 that the average value of the difference Up - Down is not zero. Thus, we can anticipate that the low percentile values are going to be close to 0, and as a consequence a slight forecast value error will cause an innacuracy in state classification.

D. Comparison of models

By comparing figures 5, 8 and 12 we immediately observe that the Deep Learning model is the one which captures best the relation between the variables and the output. As we would expect, the residual values are also lower in the deep learning model, yet the difference is not too noticeable in the plots (figure 13).

The performances of the Lasso and BayessianRidge models were almost identical. The BayesianRidge model shows slightly better results in classification accuracy, yet the difference is almost imperceptible.

Despite showing visible better results at first sight, the deep learning model does not perform that much better than the rest of models. The top plots of figures 6, 9 and 15 demonstrate that exceptuating hours 10-12, the three models have similar accuracy when classifying states in mid-day hours. Meaning a relevant fraction of the improvement of the model in comparison to the Lasso and BayesianRidge models is due to the fact that the model is able to capture the relationship between the state and the schedule.

Despite having some oddly bad performances in classifying states when the forecasted value falls into the low percentile range, like a 0.22 accuracy for the hour 7 when forecasting a 0-10 percentile in magnitude (figure 16), the deep learning exhibits as a more trustable model than Lasso and BayesianRidge models. With an outstanding accuracy in the high percentile ranges, over 90% accuracy for most hours whilst the Lasso and BayesianRidge average around 80%, the deep learning model proves to have the best classifying performance, not too much better than the Lasso and BayesianRidge models when the prediction magnitude percentile is low and the corresponding hours of the day are mid-day hours.

V. Conclusion

Based in the data presented, we can conclude that in electricity markets, the difference in economic penalties between under-generation and over-generation deviations fluctuates considerably depending on the time of day. This highlights the time-sensitive nature of imbalance costs, as system flexibility, demand patterns, and market prices vary throughout the day, making deviations more or less costly depending on when and how they occur.

The results indicate that there is a noticeable correlation between the values that were scheduled 48 hours in advance and the corresponding penalty values associated with any deviations from those scheduled figures. This suggests that earlier scheduled values are indicators that play a role in determining the extent of penalties incurred due to discrepancies or variations.

The presented results would suggest that the extent to which the deviation penalties depend on the time of day cannot be captured by linear models, which is why the Lasso and BayesianRidge models are unable to perform well in the hours 1,2 - 22, 23 and 24, yet have a consistent accuracy throughout the whole day.

In conlusion, the deep learning model is the best one, with a classification accuracy is notably higher, especially in high percentile ranges. The Deep Learning model's ability to leverage temporal patterns, particularly during non-mid-day hours, further enhances its effectiveness. Despite occasional underperformance in low percentile ranges, its overall robustness and trustworthiness make it the best model for accurate state classification in this context.

References

[1] Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. International Journal of Forecasting, 30(4), 1030-1081.

[2] Generation forecasting from variable RES. AVENSTON. Available at: https://avenston.com/en/articles/vre-generation-forecasting/. [Accessed: April 1, 2025].

[3] Red Eléctrica de España (REE), *Componentes precio energía cierre desglose*, [Online]. Available: https://www.ree.es/es/datos/mercados/componentes-precio-energia-cierre-desglose. [Accessed: April 1, 2025].

[4] Red Eléctrica de España (REE), ”Intraday Market,” Available online: https://www.sistemaelectrico-ree.es/en/spanish-electricity-system/markets/intraday-market, [Accessed: April 1, 2025].

[5] ”Renewable energy generation v. grid stability - A conflict of objectives?,” Energy, The Economic Times, July 18, 2024. Available online: https://energy.economictimes.indiatimes.com/news/renewable/renewable-energy-generation-v-grid-stability-a-conflict-of-objectives/113617437, [Accessed: April 1, 2025].

[6] The New York Times, “What We Know About the Power Outage in Spain and Portugal,” April 28, 2025, https://www.nytimes.com/2025/04/28/world/europe/spain-portugal-power-outage-what-we-know.html.

[7] ”Spanish company fined for selling solar power at night,” *pv magazine*, December 17, 2024. Available online: https://www.pv-magazine.com/2024/12/17/spanish-company-fined-for-selling-solar-power-at-night/, [Accessed: April 1, 2025].

[8] ”Important regulatory developments in the energy sector approved in Spain: Decree Law 6/2022,” Watson Farley & Williams, April 29, 2022. Available online: https://www.wfw.com/articles/important-regulatory-developments-in-the-energy-sector-approved-in-spain-decree-law- [Accessed: April 1, 2025].
”Improving accuracy of solar power forecasts,” *Solargis*, November 21, 2023. Available online: https://solargis.com/resources/blog/best-practices/improving-accuracy-of-solar-power-forecasts, [Accessed: April 1, 2025].
”Spain - Electricity security policy,” *IEA*, October 21, 2021. Available online: https://www.iea.org/articles/spain-electricity-security-policy, [Accessed: April 1, 2025].

[9] Red Eléctrica de España, https://www.ree.es/es, Accessed: April 13, 2025.

[10] Stats Stack Exchange, ”What is the lasso in regression analysis?”, https://stats.stackexchange.com/questions/17251/what-is-the-lasso-in-regression-analysis, Accessed: April 13, 2025.

[11] IBM, ”Lasso Regression”, https://www.ibm.com/think/topics/lasso-regression, Accessed: April 13, 2025.

[12] IBM, ”Lasso Regression”, https://www.ibm.com/think/topics/lasso-regression, Accessed: April 13, 2025.

[13] MathWorks, “Bayesian Lasso Regression,” https://www.mathworks.com/help/econ/bayesian-lasso-regression.html, [Online; accessed 19 April 2025].

[14] Bruna, W., “Bayesian Regression,” https://brunaw.com/phd/bayes-regression/report.pdf, [Online; accessed 19 April 2025].

[15] ISO, “Artificial Intelligence – Deep Learning,” https://www.iso.org/artificial-intelligence/deep-learning, [Online; accessed 19 April 2025].

[16] NVIDIA, “Deep Learning in a Nutshell: Core Concepts,” https://developer.nvidia.com/blog/deep-learning-nutshell-core-concepts/, [Online; accessed 19 April 2025].

[17] GeeksforGeeks, “Hyperparameter Tuning,” https://www.geeksforgeeks.org/hyperparameter-tuning/, [Online; accessed 19 April 2025].

[18] IBM, “What is Deep Learning?,” https://www.ibm.com/think/topics/deep-learning, [Online; accessed 19 April 2025].

Daniel García Arana