

Optimizing Energy Market Trading with Conformal Predictions and Conditional Value at Risk

Daniel Moore and Laila Saleh

Abstract—We optimized the market trading of a renewable energy generation operator with conditional value at risk based on probabilistic forecasts made with a conformalized Long Short-term Memory (LSTM) recurrent neural network. This work demonstrates an end-to-end workflow of how field data can be ingested, analyzed, and exploited to reduce risk exposure for the operator. This is financially beneficial to the individual operator, but taken to a large scale this methodology increases the incentive for renewable generation participation which should drive cost and emissions down. This work used only eight features with favorable results, so it is expected that further studies and more advanced models with the same architecture would provide better yield.

Index Terms—Julia, Flux, LSTM, Conformal Prediction, conditional value at risk, renewable energy, energy markets, optimization, revenue, risk management

I. INTRODUCTION

The [IEEE Hybrid Energy Forecasting and Trading Competition](#) challenges participants to make day-ahead, half-hourly probabilistic forecasts of solar and wind energy production for a solar farm and Hornsea-1 Wind Farm in the east of England with a combined 3.6 GW capacity and then maximize revenue through commitment in the day-ahead market. Any difference between the committed energy and actual energy is rewarded or punished according to the single settlement price (SSP). The implied task is to also forecast the market prices so that the operator can reduce their risk exposure from both the energy production and market prices.

A. Motivation

This is an appropriate capstone project for this course as it applies many topics covered ranging from unit commitment and energy market trading to advanced predictive and prescriptive analytics for complex and uncertain events. It is an interesting and practical opportunity to wrestle with the available resources to make the best decisions for the operator. Lastly, we find it a compelling problem because reducing the risk for renewable energy generation operators will encourage more participation and be of a net benefit to investors, consumers, and the environment - a rare triple-win.

B. Objectives

We will show an end-to-end workflow where we process data to train a forecasting model and conformalize it so that its point forecasts can be transformed into probabilistic forecasts. These forecasts enable us to make market-trading decisions that

consider the uncertainty in not only energy production but also in the market itself. Finally, we will demonstrate the financial benefit of leveraging the power of stochastic optimization to reduce the risk exposure of the operator.

C. Literature Review

What have other people done

II. DATA ANALYSIS

We have obtained datasets from two sources: the competition itself which provides the energy production data through the [Rebase API](#) and the [VisualCrossing API](#) which provides the weather data. The energy data details the solar and wind energy production and the DAP and SSP in half-hourly increments. The weather data is treated as historic for the period preceding a given forecast and as a weather forecast for the forecast horizon. If deployed, the model would need to operate only using forecasted weather data. This approach is acceptable for this study as 24-hour-ahead weather forecasts are typically very accurate and we are only incorporating basic weather features.

A. Data Insights

The tables below summarize the energy and weather data used throughout this report. We have the amount of energy produced from solar, wind, and combined total energy as well as the day-ahead and single settlement prices. Combining the solar and wind energy results in a more predictable value as we see the median total energy is greater than the sum of the median solar and wind energy and the mean value is closer to the median value. For the market prices we see similar mean and median values in the DAP and SSP but the domain of the SSP is much larger. Also it is notable that both have negative values indicating there are times of a surplus of energy which is penalized by the market.

Table I: Rebase Energy Data Summary

	variable	mean	min	median	max
	Symbol	Float32	Float32	Float64	Float32
1	Solar	276.497	0.0	7.85413	1853.73
2	Wind	543.425	0.0	698.674	826.254
3	TotalEnergy	819.922	0.0	778.34	2367.19
4	DAP	56.0618	-23.77	61.565	112.23
5	SSP	55.2193	-88.0	55.595	177.71

Examining the weather data, we see that everything but cloudcover is more or less normally distributed. Cloudcover, however, is heavily skewed towards being more cloudy as the mean is 71% and the median is 92%. Being England, this lives up to expectations.

Table II: Weather Data Summary

	variable	mean	min	median	max
	Symbol	Float32	Float32	Float64	Float32
1	temp	8.54791	-2.3	8.0	19.0
2	windspeed	16.9168	0.9	16.3	46.4
3	winddir	191.088	2.0	200.0	359.0
4	cloudcover	71.4654	0.0	91.6	100.0
5	visibility	14.4783	0.0	14.9	29.8

B. Data Visualizations

In the following plots we examine first the day-ahead and then the single settlement prices. The top plot shows the entire history of the data, the second provides a closer look at a single week. The bottom plots show the prices vs. temperature and total energy production with the right side showing the daily seasonality. The relative variability of the SSP compared to the DAP is highlighted by the scales being the same in the first and second plots.

For the DAP, we see a clear pattern with a few random dips in the price but eventually returning to the mean. The week plot provides a clearer look at the typical pattern. In the temperature plot we see that the price tends to show the same general fluctuations regardless of the price. There are two days around April 9th which seem to have low costs associated with the warmest days. It is unclear whether this is a coincidence. The bottom plot shows how the price changes with total energy production. There is not a clear pattern in the monthly plot but we do see that prices are at a constant high, but not peak, price when energy production is the greatest. This is a direct consequence of solar production and not a causation.

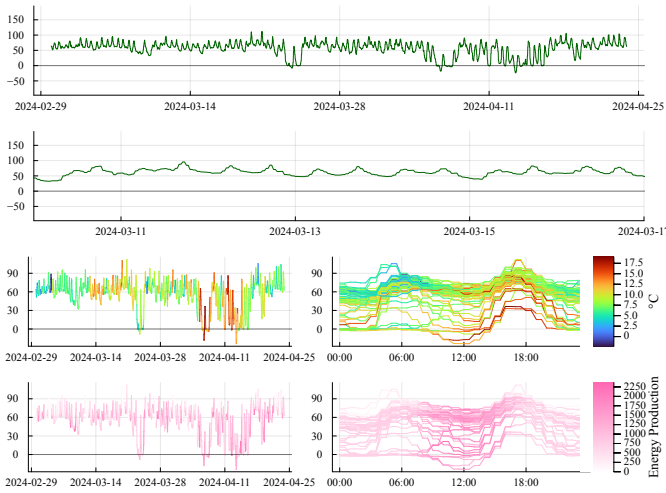


Fig. 1: Day Ahead Prices

The SSP is a different story as the time-series is hardly distinguishable from noise. We see the SSP fluctuates many times in a given day and there is no evident seasonality. There is also no clear pattern with temperature or total energy production. Finally, we note that the SSP covers a much larger domain than the DAP and frequently reaches near its extreme values.

The violin plots with overlaid boxplots provide context for the trend and seasonality for the market prices and energy production. First, we can see there is no clear trend for any

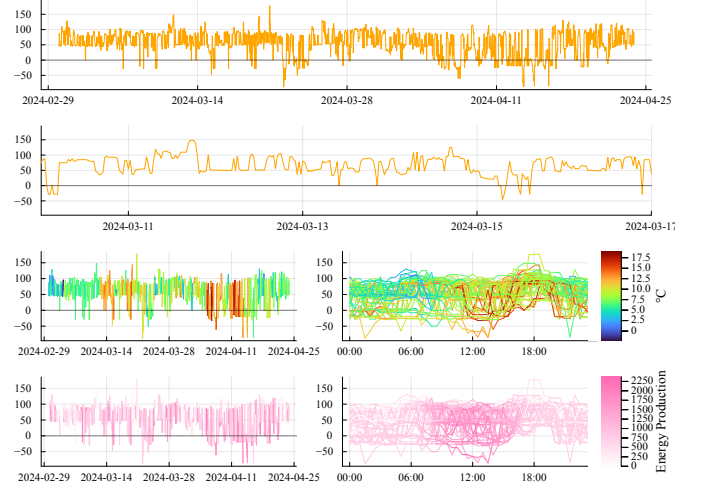


Fig. 2: Single Settlement Prices

across the entire time frame. Instead, we see cycles which hint at a dynamic relationship between the prices and energy production. Note how in April we see a relatively constant energy production associated with the most drastic decrease in prices. The day of the week plots do hint at some trends, but nothing strong enough to warrant a one-hot encoding. We do not have enough data to determine whether it is a coincidence that Saturday has lower prices or if there is actually some relationship that makes this a predictable event. Finally, in the hourly plots we can finally observe a clear trend. The DAP follows a close distribution at each hour of the day and the change throughout the day is predictable. We can see a correlation with the SSP, but still the SSP is more variable and not directly correlated. Finally, we see that the total energy production follows a clear pattern, but with upper and lower quantiles reaching the extreme values indicating there are often enough times with very little or very high production. Combining the wind and solar has resulted in a more stable distribution.

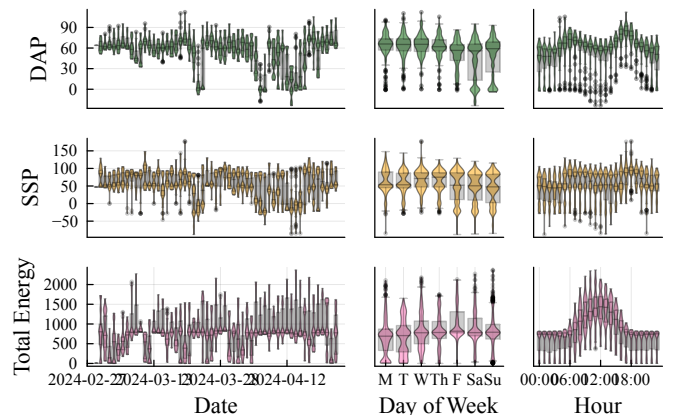


Fig. 3: Single Day, Weekly, and Daily Trends

III. LONG SHORT-TERM MEMORY MODEL

We used an LSTM Recurrent Neural Network (RNN) to predict the energy production and market prices for the

following day. RNNs are designed to handle time-series data and LSTMs are a special type of RNN that can learn long-term dependencies in the data. Combined with a dense neural network, this model can remember (and forget) time-dependent relationships and approximate the complex dynamics among the variables.

A. Training

We used Julia's deep learning library, Flux, to build and train the LSTM. We created a multi-target regressor to predict the energy production and market prices in single timestep increments. All features were normalized and time was encoded as a sine wave to capture the cyclical nature of time.

B. Performance

Qualitatively, we observe the LSTM predictions are sensible and follow the general characteristics of the target variables. It is able to predict the Total Energy very well, but the market prices are not as good.

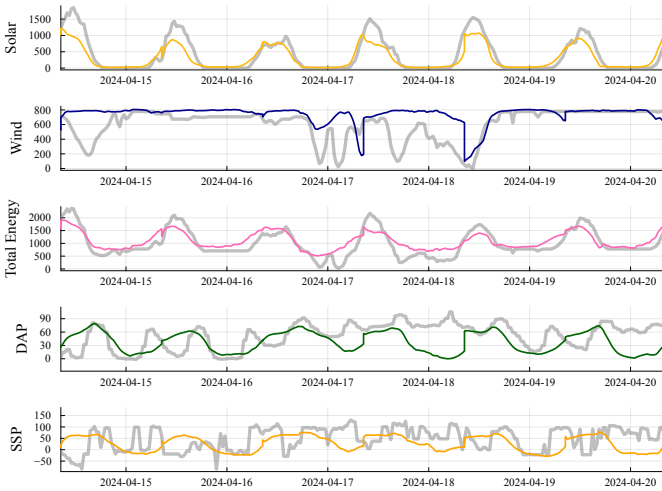


Fig. 4: Predictions

We check the Mean Absolute and Root Mean Squared Error for all three testing sets to ensure the model is not overfitting. The results are shown the table below.

Table III: Mean Absolute Error

	feature	Train	Calib	Test
	Symbol	Float32	Float32	Float32
1	Solar	102.419	131.789	175.321
2	Wind	166.083	112.212	174.395
3	TotalEnergy	259.734	203.678	280.884
4	DAP	18.8602	20.8735	29.6261
5	SSP	34.4089	37.4783	46.8583

Table IV: Root Mean Square Error

	feature	Train	Calib	Test
	Symbol	Float32	Float32	Float32
1	Solar	175.989	223.373	274.294
2	Wind	270.276	181.174	273.594
3	TotalEnergy	363.375	287.347	360.891
4	DAP	25.3194	27.2402	36.9338
5	SSP	42.9502	46.85	57.6788

We see that the residuals are normally distributed with a mean of 0, indicating we will be able to conformalize the model to obtain a probabilistic forecast. Solar has most values at zero because the model correctly predicts no solar energy production at night. We get around this by using the combined energy as we don't explicitly have to know the solar and the wind energy production.

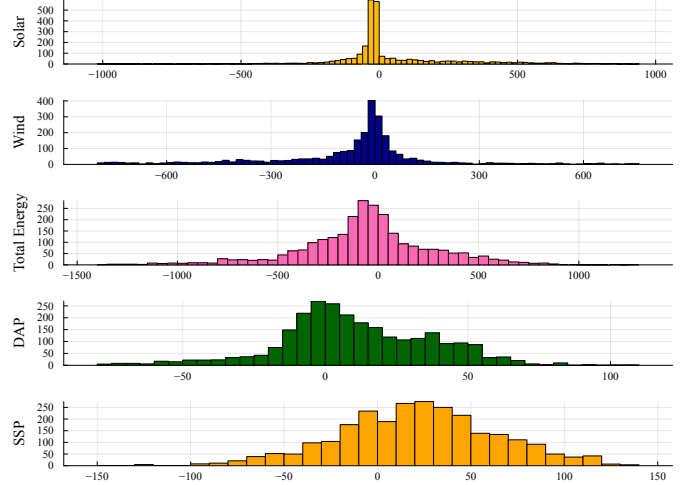


Fig. 5: Test Data Residuals

IV. CONFORMALIZING LSTM

The LSTM model has made predictions which capture the general behavior of the target variables in terms of the time-dependent dynamics. However, the model still has error that would make trading hazardous, especially because we do not have a sense of how confident the model is in its predictions.

A. Implementation

Conformal predictions are an elegant solution to quantifying the uncertainty of a point forecast model. It is empirically built from calibration data which the model has not seen. For this regression task, a simple residual nonconformity score is appropriate. We calculate the absolute error for each prediction in the calibration data and save it in a table for later reference. We can then retrieve the nonconformity score for any prediction and α level.

B. Performance

The nonconformity cumulative distribution functions are shown below. For both the energy production and market prices, the nonconformity scores grow exponentially once the quantile is past about 0.75. This indicates that $\alpha = 0.75$ is a good choice for the confidence levels because more than that and the range becomes too high while less than results in too little coverage.

The coverage of the values for the calibration and test data at

Finally, we confirm the coverage of the predictions for the test data at $\alpha = 0.75$ for the total energy and market prices for the training, calibration, and test data.

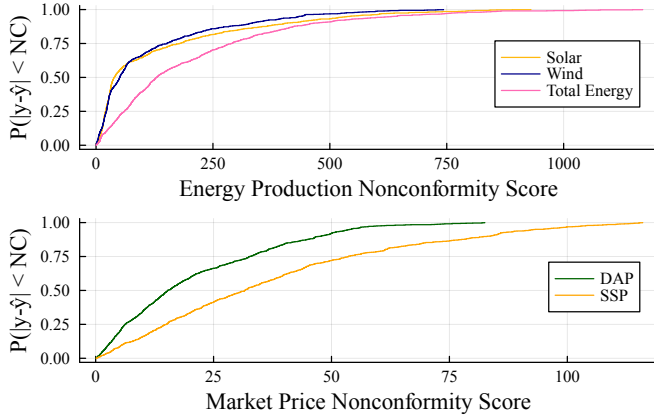


Fig. 6: Nonconformity Scores

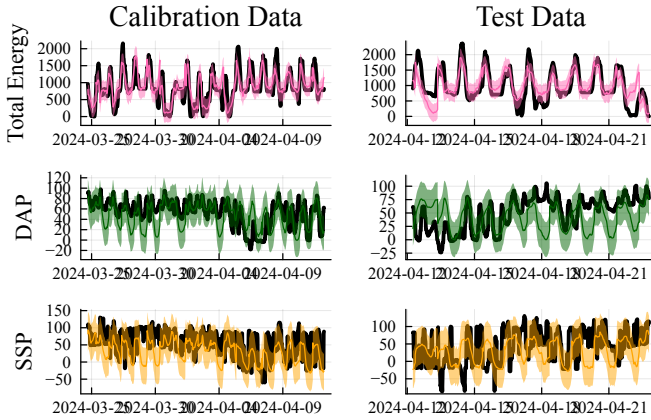


Fig. 7: Prediction Coverage

We observe that the values are consistent for the training and calibration data while the prediction intervals do not hold for the test data where the coverage drops to as low as 0.62 for the DAP. Still, the coverage for the Total Energy is at 0.72 which is acceptable. The degradation of the coverage in the test data is likely due to increased volatility in the markets which the model has not seen before.

V. MARKET TRADING

With prediction for the energy production and market prices we can make informed decisions about how much energy to commit to the day ahead market. We will explore three strategies for comparison. It must be assumed that the operator does not have the ability to store energy or reduce energy production because we have not been provided that information. If the operator could, they would opt to shut down production if prices are expected to be negative. We cannot consider this

outcome because the data does not provide information on the cost of shutting down production.

1. Commit the point prediction of the total energy
2. Commit the conformal prediction of the total energy at a given α level (as mentioned above we will use 0.75)
3. Commit the conformal prediction considering the joint probabilities of the total energy and market prices

Revenue is calculated for this competition according to this formula which approximates the impact of the difference between the energy prediction and actual production on the settlement price.

$$Revenue = Trade * DAP + \Delta_E(SSP - 0.07\Delta_E) \Delta_E = Actual - Trade$$

A. Revenue from Point Prediction

The predicted total energy is used as the traded amount in the day ahead market. Revenue is calculated above using the actual total energy and the actual market prices.

B. Revenue from Conformal Prediction

Here we determine the lower bound of the amount of energy to be produced by subtracting the nonconformity score of the 75th quantile from the point prediction. We then trade the greater of this amount and zero as negative values are not allowed. Revenue is calculated as before.

C. Revenue from Conditional Value at Risk

The shortcoming of the above methods are primarily that they do not consider the market prices and the joint probability of the errors. We must use joint probabilities to do so as the nonconformity scores of the total energy, DAP, and SSP are correlated.

We have two options for capturing the joint probabilities. First, we could assume these are multivariate normally distributed and we could sample them as such. While convenient, it introduces additional assumptions which may not be valid. Instead, we will sample the residuals from the calibration data. For each timestep, we modify the total energy and market predictions according to n samples from the calibration data. For all values of the possible traded energy, we evaluate the revenue for all scenarios and find the amount to trade which minimizes the expected shortfall. This process is depicted in the plot below.

The plot shows the revenue for each scenario as a scatter point colored on the amount of energy expected to be produced. We have overlaid the mean, median, 5th percentile value at risk, and the 5th percentile conditional value at risk. Also as scatter points are the trade and revenue amounts for a perfect energy production prediction, the point forecast, and the conformal forecast. Finally, we see the large green diamond indicating the revenue actually received for this time step at the trade amount set by the conditional value at risk. We note how the aggregation for the mean and value at risk smooths the curves to they have a convex shape.

Table V: Prediction Coverage

	feature	Train	Calib	Test
	Symbol	Float64	Float64	Float64
1	Solar	0.820447	0.749448	0.685328
2	Wind	0.667526	0.749448	0.666023
3	TotalEnergy	0.643471	0.749448	0.611969
4	DAP	0.763746	0.749448	0.586873
5	SSP	0.766323	0.749448	0.656371

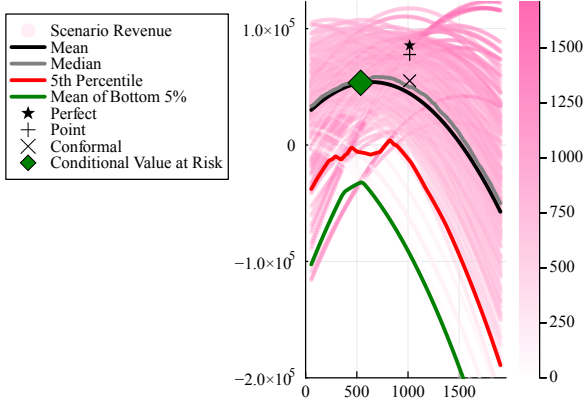


Fig. 8: Conditional Value at Risk

D. Performance

Below we examine the revenues for each trading strategy. The top plot shows that all strategies are typically making the same revenue but there is one moment where the predictions all suffer a catastrophic loss. The middle plot shows the cumulative revenue and we see over time they all recover but the dip is pronounced. Lastly at the bottom there is a plot which shows the relative cumulative value of each strategy compared to the perfect prediction. Eventually, they all recover and their values approach around 0.6 in the long run. Surprisingly the point forecast is the best followed by the CVaR and finally the conformal prediction.

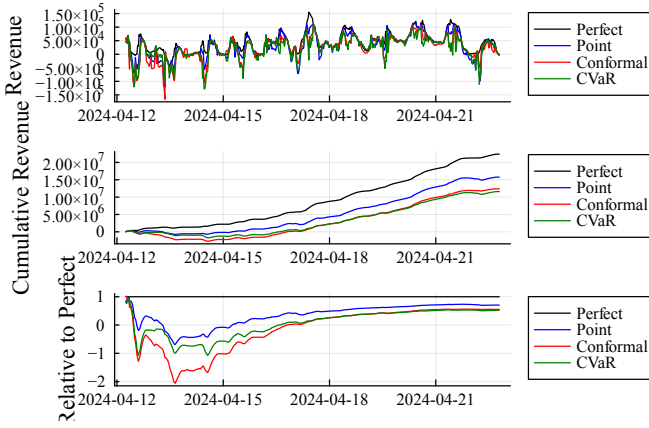


Fig. 9: Revenue vs. Time

VI. CONCLUSIONS

Suggestions for future work 1. Include more data, especially demand predictions 2. Include the cost of shutting down production 3. Predict a trade amount which will maximize the revenue with some α level directly