Conformalizing an LSTM to Optimize Revenue for a Renewable Energy Operator using Conditional Value at Risk

Energy Markets and Data Analytics, Rutgers, Spring 2024, Dr. Robert Mieth

Daniel Moore, Laila Saleh

2024-05-08

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 work flow of how field data can be ingested, analyzed, and exploited to reduce risk exposure for the operator. Obviously this is financially beneficial to the individual operator, but taken to a larget scale this methodology increases the incentive for renewable generation participation which should drive cost and emmissions down. This work carried out on a small number of features with great results, so it is expected that further studies and more advance models with the same architecture would provide better yield.

Table of contents

List of Figures

List of Tables

1 Introduction

The IEEE Hybrid Energy Forecasting and Trading Competition challenges participants to mak 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. The second task is to decide how much energy to commit to selling at the day-ahead price (DAP) to optimize reveneus. Any difference between the committed energy and actual energy is rewarded or punished according to the single settlement price (SSP). As discussed later in the data, beyond their volatility, the DAP and SSP can also be negative indicating a surplus of energy on the market. In rare events, underproduction could be rewarded due to a negative 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.

1.1 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.

1.2 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 which consider the uncertainty in not only the 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.

1.3 Literature Review

What have other people done

Table 1: Data from RebaseAPI

	DateTime	Solar	Wind	TotalEnergy	DAP	SSP
	DateTime	Float32	Float32	Float32	Float32	Float32
1	2024-03-05T03:00:00	0.0	587.3	587.3	54.33	44.0
2	2024-03-05T03:30:00	0.0	601.86	601.86	54.33	55.94
3	2024-03-05T04:00:00	0.0	595.92	595.92	61.81	52.82
4	2024-03-05T04:30:00	0.0	575.42	575.42	61.81	52.63
5	2024-03-05T05:00:00	0.0	563.54	563.54	71.01	57.49

Table 2: Ennergy Data Summary

	variable	mean	min	median	max
	Symbol	Float32	Float32	Float64	Float32
1	Solar	236.622	0.0	0.34961	1504.95
2	Wind	518.719	0.0	666.469	826.254
3	TotalEnergy	755.341	0.0	778.339	2271.71
4	DAP	59.4672	-18.2	62.015	112.23
5	SSP	61.5711	-88.0	61.565	177.71

2 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 being treated as historic fro time previoius to a given forecast and as a weather forecast for the given forecast horizon. This provides our model with more accurate information than it would have if it only had the historic weather data. However, we are only using the next 24-hours of weather forecast at a time and that short horizon will limit the unconservatism to an acceptable level for this study.

2.1 Data Summary

The tables below provide a sample of what the data from each source look like for a few time periods and summary statistics.

2.2 Data Visualizations

The first thing to notice when looking at the DAP and SSP time-series plots below is the volatility of the prices. The DAP clearly has some seasonality but the trend and cycles are

Table 3: Data from VisualCrossing

	DateTime	temp	windspeed	winddir	cloudcover
	DateTime	Float32	Float32	Float32	Float32
1	2024-03-05T03:00:00	6.9	14.8	132.0	100.0
2	2024-03-05T03:30:00	6.9	14.8	132.0	100.0
3	2024-03-05T04:00:00	6.3	13.9	140.0	94.0
4	2024-03-05T04:30:00	6.3	13.9	140.0	94.0
5	2024-03-05T05:00:00	6.2	12.5	150.0	100.0

Table 4: Weather Data Summary

	variable	mean	min	median	max
	Symbol	Union...	Any	Any	Any
1	DateTime		2024-02-29T23:00:00	2024-03-19T10:45:00	2024-04-06T23:30:00
2	temp	8.21836	-2.3	7.7	19.0
3	windspeed	16.0126	0.9	15.3	41.2
4	winddir	172.188	3.0	178.0	359.0
5	cloudcover	73.9165	0.0	93.7	100.0

not clear. The SSP is even more volatile as it traverses from the daily minimum to the daily maximum serveral times in a given day. This highlights how difficult it is for all participatnts to make accurate forecasts and fulfill their commitments.

The plots of the DAP and SSP for the first week of April below provide a closer look at the characteristics of these prices. While predicting the DAP appears feasible, the SSP is hardly distinuishable from noise.

Table @ref(monthly-energy-production-plot) shows the energy generation over March 2024 while Table @ref(weekly-energy-production-plot) shows the first week of April 2024. We see the Solar production is, as expected, seasonal while the wind production is very cycical. It apepars to go from nothing to its full capacity in a short amount of time and stay there for a random amount of time before dropping, usually back to nothing.

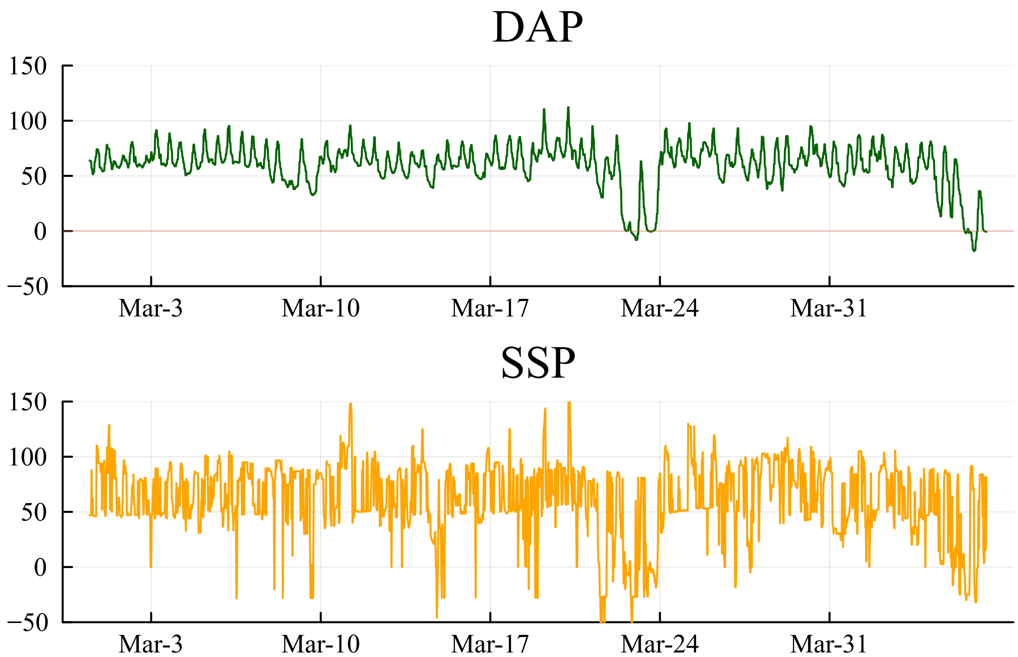


Figure 1: March 2024 Energy Prices

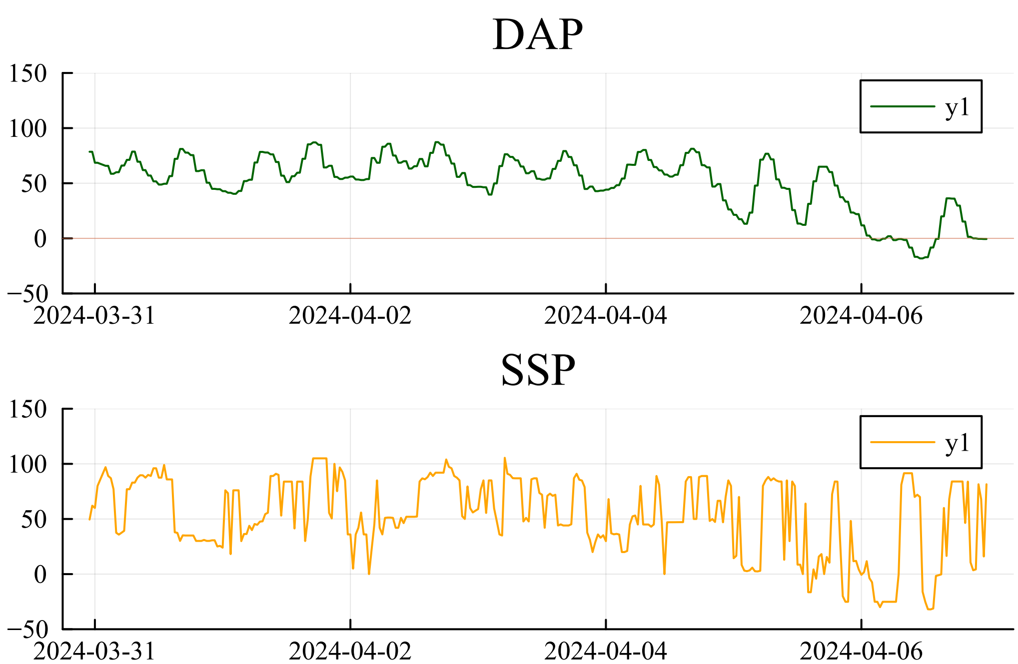


Figure 2: First Week of April 2024 Energy Prices

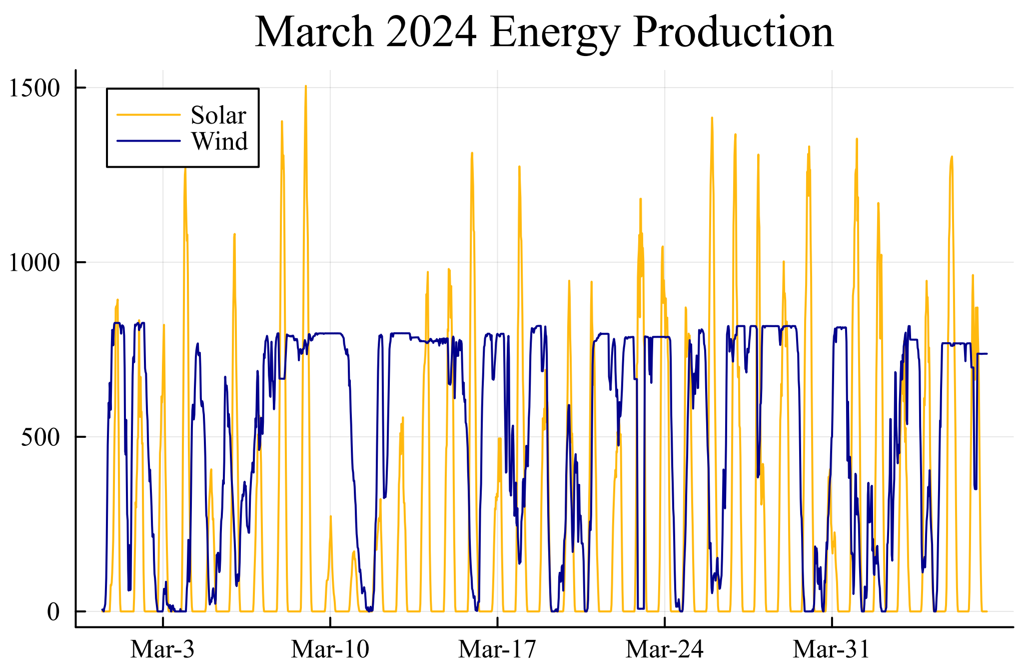


Figure 3: March 2024 Energy Production

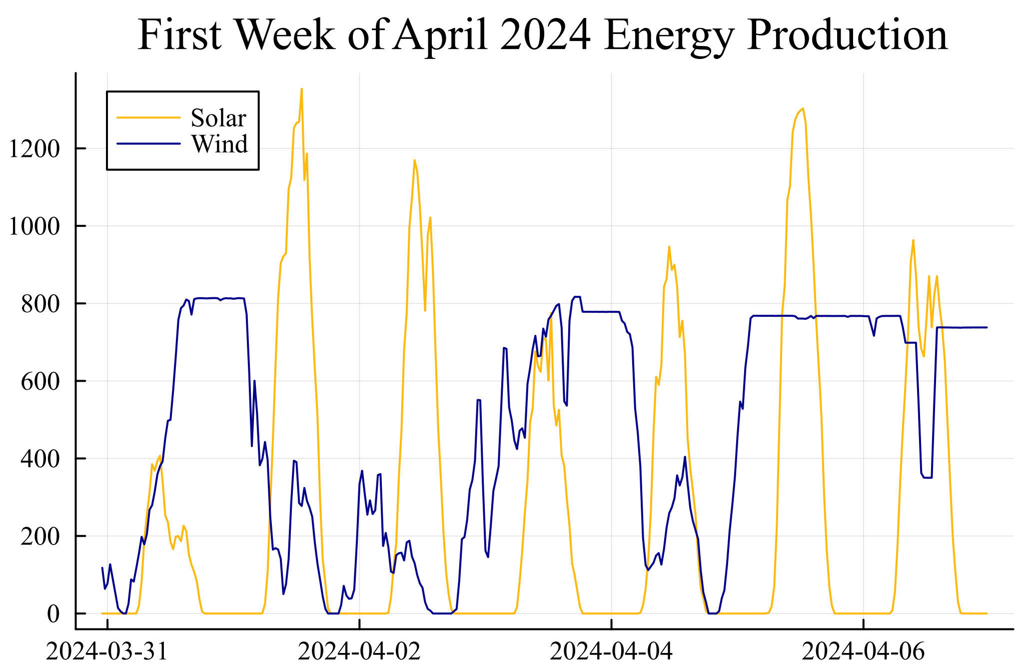


Figure 4: First Week of April 2024 Energy Prices

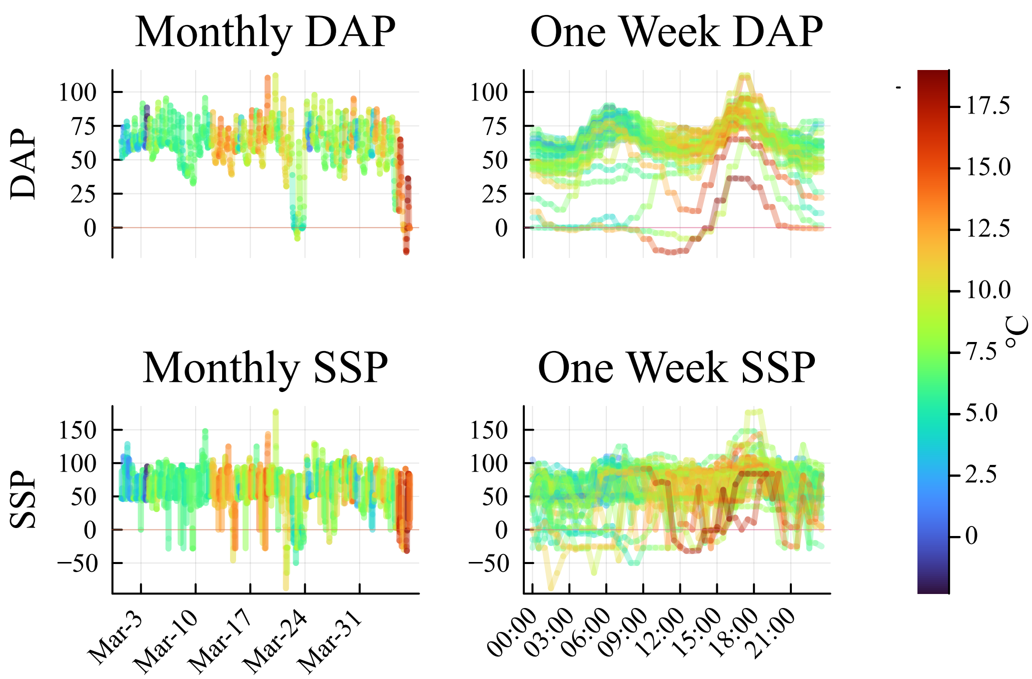


Figure 5: DAP, SSP vs Temperature

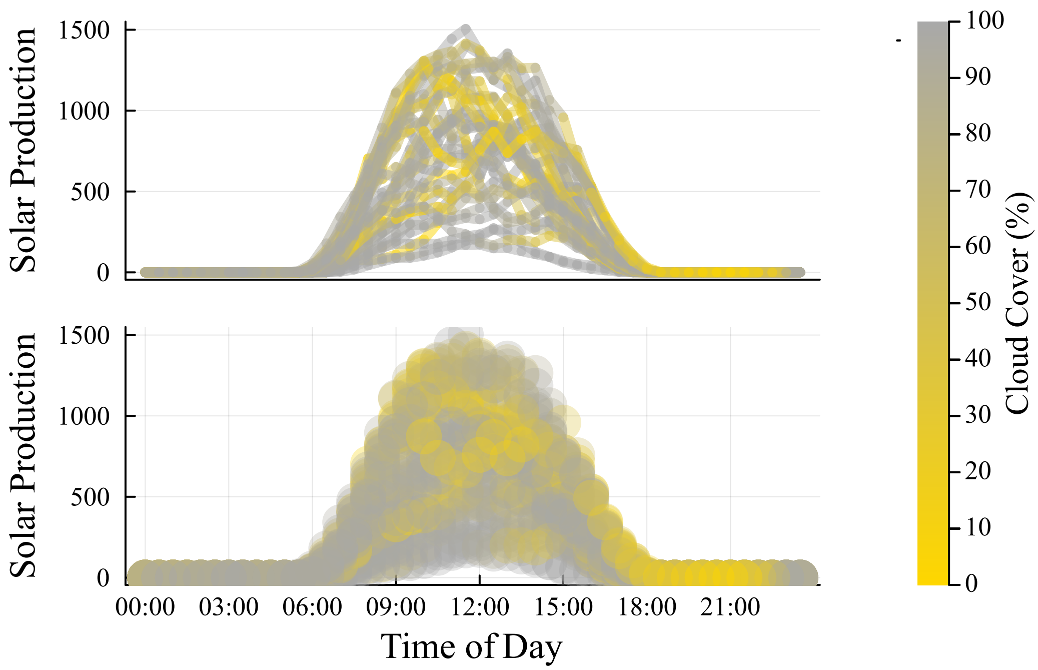


Figure 6: Daily Solar Production vs Cloud Cover

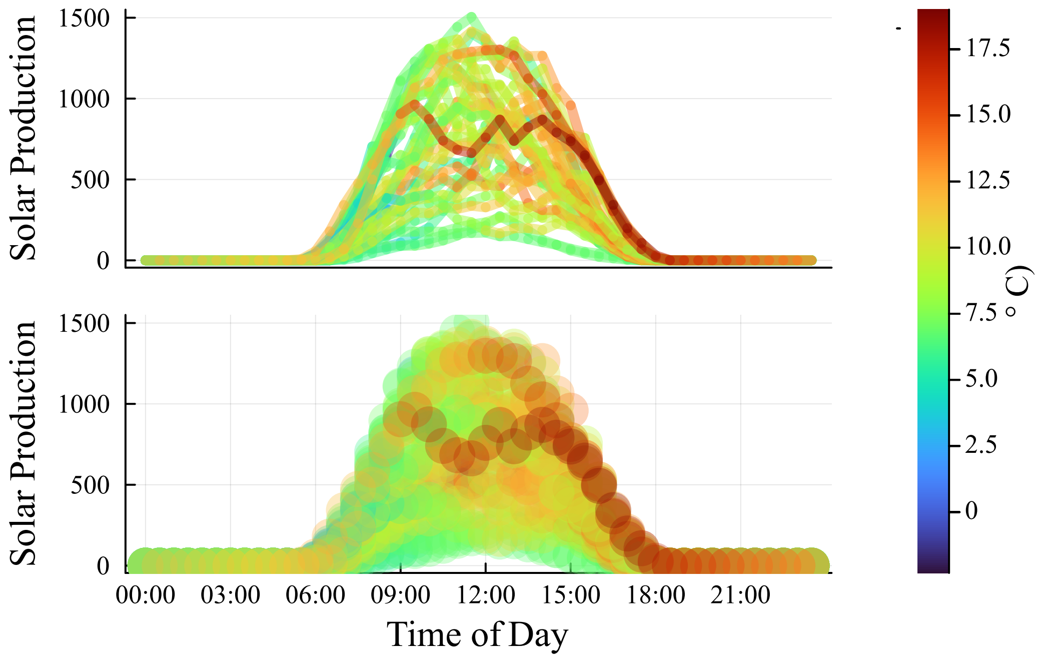
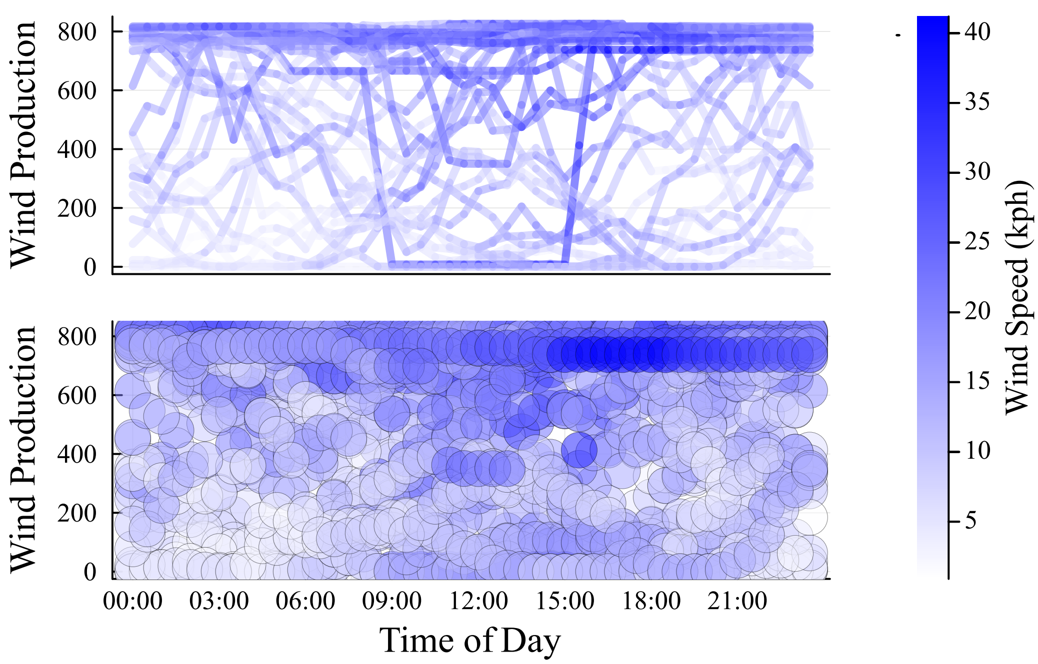
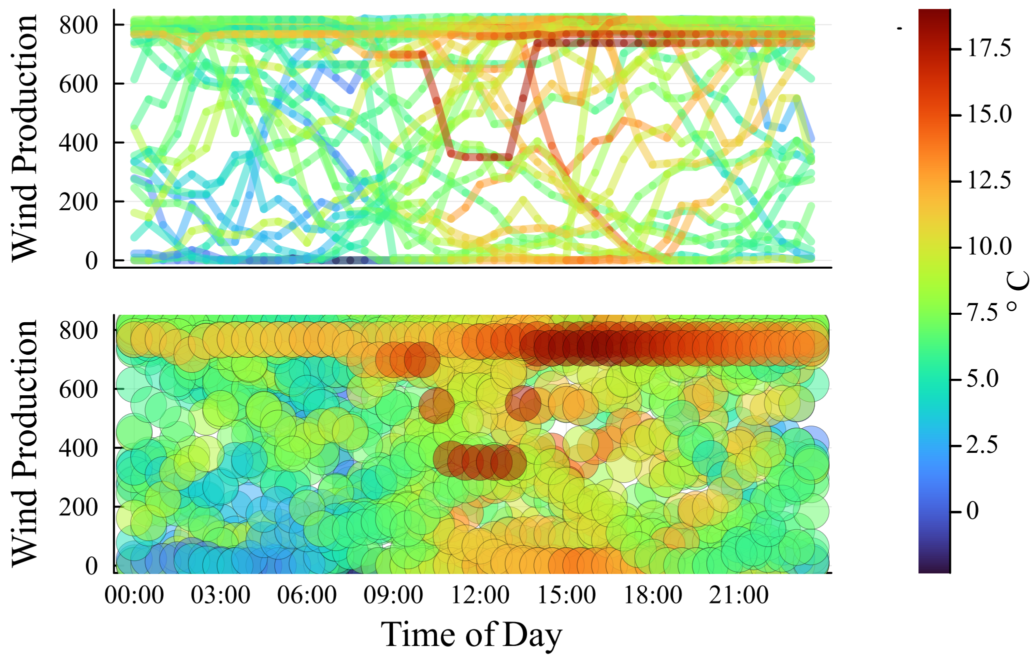


Figure 7: Daily Solar Production vs Temperature





3 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 which are able to learn long-term dependencies in the data. Combined with a dense neural network, this model can remember (and forget) time-depedent relationships and approximate the complex dynamics amongst the variables.

3.1 Training

How did we make one?

3.2 Performance

Is it any good?

4 Conformalizing LSTM

What are conforal predictions? Why are they good

4.1 How to?

How did we conformalize the LSTM?

4.2 Performance

Is it any good?

5 Conditional Value at Risk

What is CVAR

Why do we do it?

5.1 Implementation

How did we do it?

5.2 Performance

Was it any good?

6 Conclusions

Was any of this worth while?

What did we learn?

What could others do?