Predicting Cash Flow at Risk for a Natural Gas Power Plant

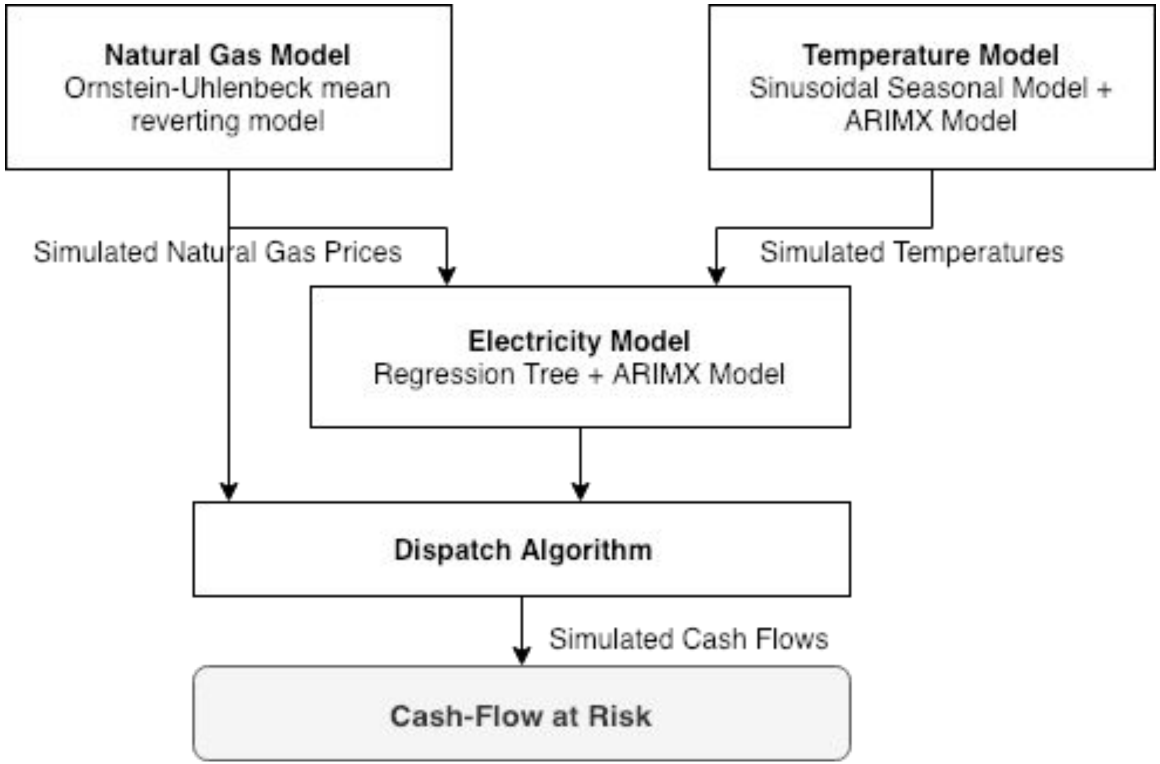
Data from New England

Natasha Ting, Peter Ogunmodede*, Rumman Waqar***

** Department of Economics, University of Alberta*
*** Department of Electrical Engineering, University of Alberta*
January 2019

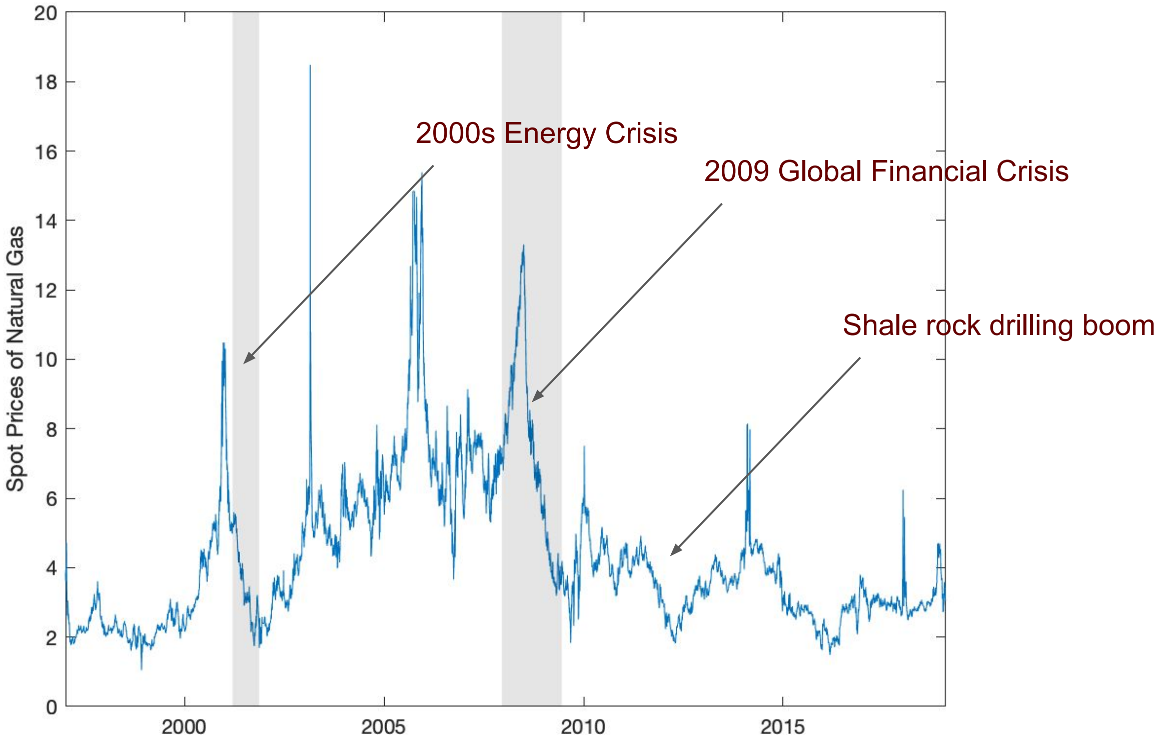
This work is produced for 2019 PRMIA Risk Challenge - Regional Round
We thank Professor Valentina Galvani for her mentorship, and Professor David P. Brown for his advice. All errors are our own.

The Models



Natural Gas Model
Ornstein-Uhlenbeck mean reverting model
Temperature Model
Sinusoidal Seasonal Model + ARIMX Model
Simulated Natural Gas Prices
Simulated Temperatures
Electricity Model
Regression Tree + ARIMX Model
Dispatch Algorithm
Simulated Cash Flows
Cash-Flow at Risk

Historical Natural Gas Prices



Ornstein-Uhlenbeck Mean Reversion Model

$$\Delta x_t = \alpha(\mu - x_t)\Delta t + \sigma dz_t, \text{ where } dz_t \sim N(0, \sqrt{\Delta_t})$$

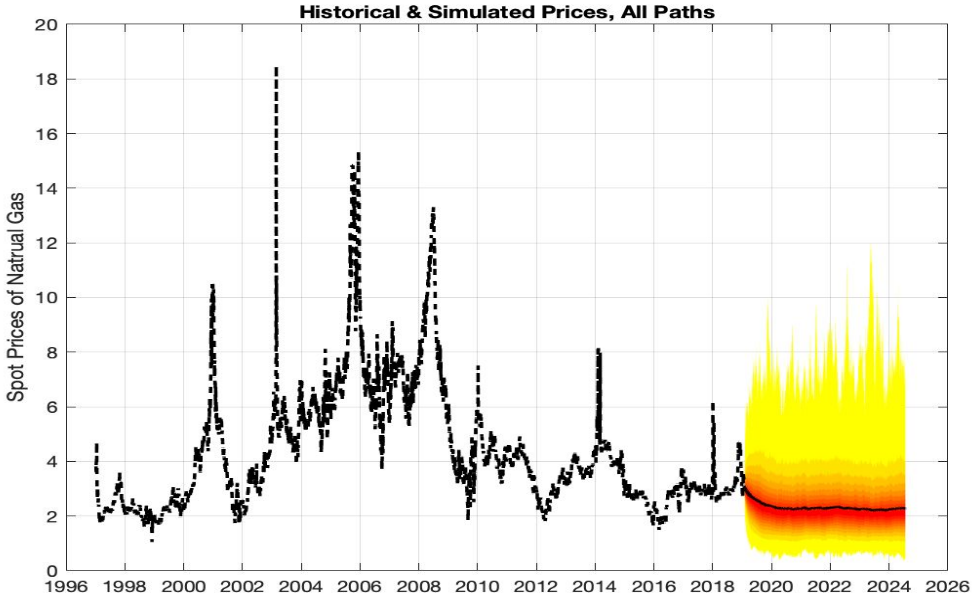
where x = Natural gas spot price, α = Mean Reversion rate, μ = Mean level of prices, and σ = Volatility

The model operates on the following assumptions:

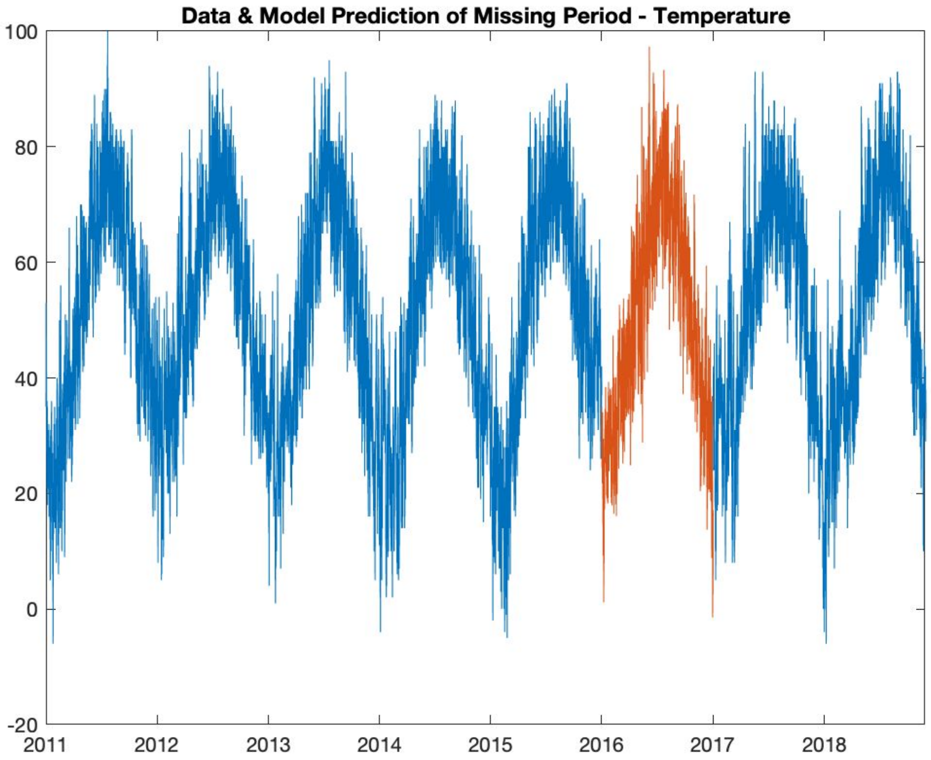
- Natural gas spot prices tend to move towards their arithmetic mean over time.
- The change in natural gas spot prices is dependent on a random normally distributed brownian movement.

Monte Carlo Simulation

The mean reversion model is used as the criteria to carry out our simulation. The simulation runs 1000 trials along with 2000 steps.



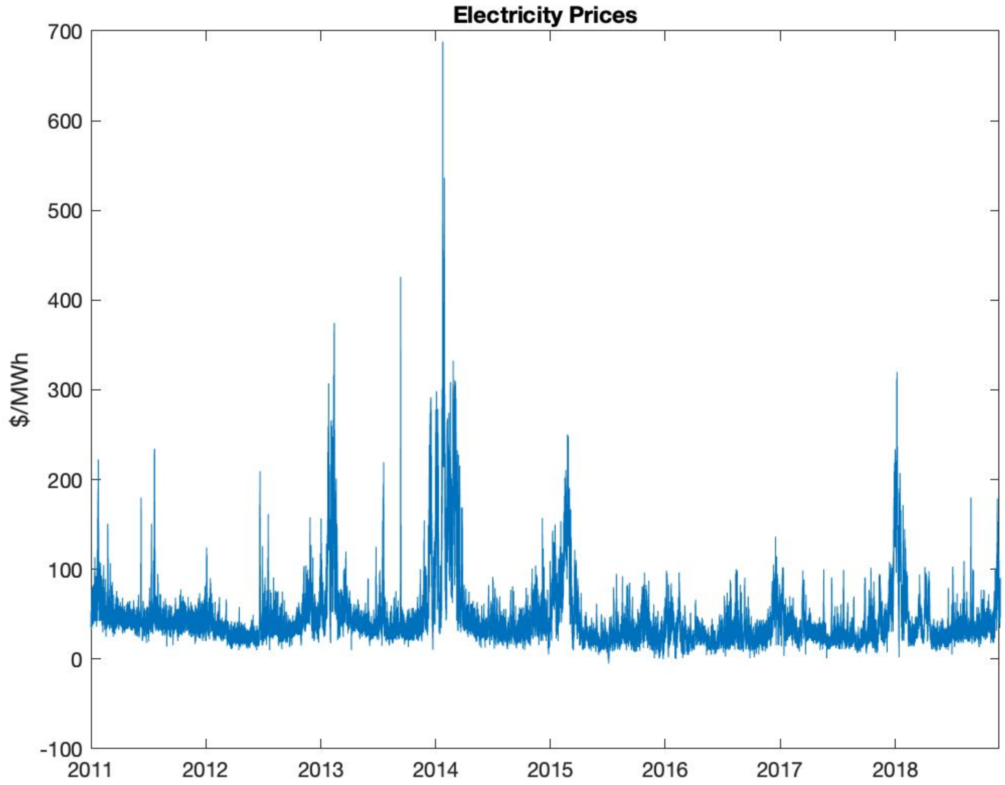
Simulated Temperature -- Aggregated



We filled in temperature data from 2016 to 2017 to validate the results of our simulations.

Temperature is simulated with a deterministic sinusoidal model + non-deterministic ARIMX model

Historical Electricity Prices - Day Ahead LMP

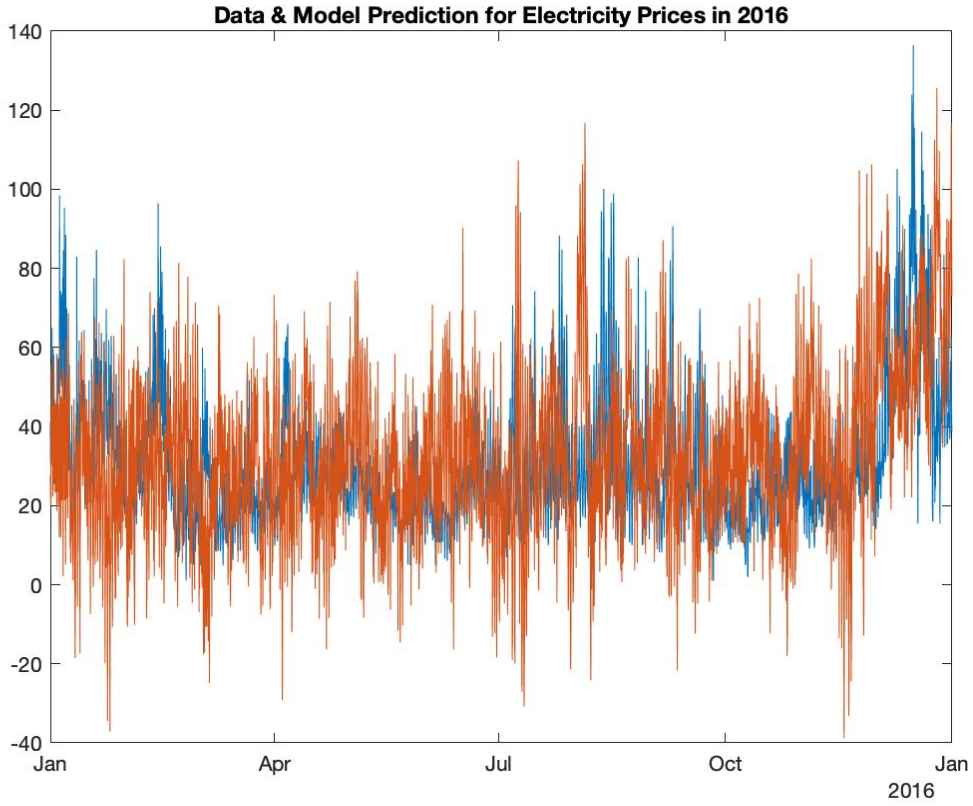


Prices shown are in USD$, wholesale hub day-ahead Locational Marginal Prices at the New England ISO.

Prices are influenced by:

- Supply:
    - Production efficiency at the plants
    - Storage Level
    - Bid prices
    - Temperature
    - etc.
- Demand:
    - Consumption (retail, industrial demand)
    - Temperature
    - Macroeconomic factors
    - etc.

Simulated Electricity Prices -- 2016



This graph shows predicted data (in orange) overlaid on realised value of DA prices (in blue)

The fitted values show variations from observed prices. This is because of reasons like seasonal correlation in temperature, rolling averages, other variables not accounted for, and strengths/weaknesses of statistical methods

Cash Flow At Risk

Definition:
Cash flow is equivalent to profits in our model

Method:
We ran the simulation process 30 times, applying the same shocks every time. The average of the output CFaR at different confidence levels is used here:

$3.61 Million

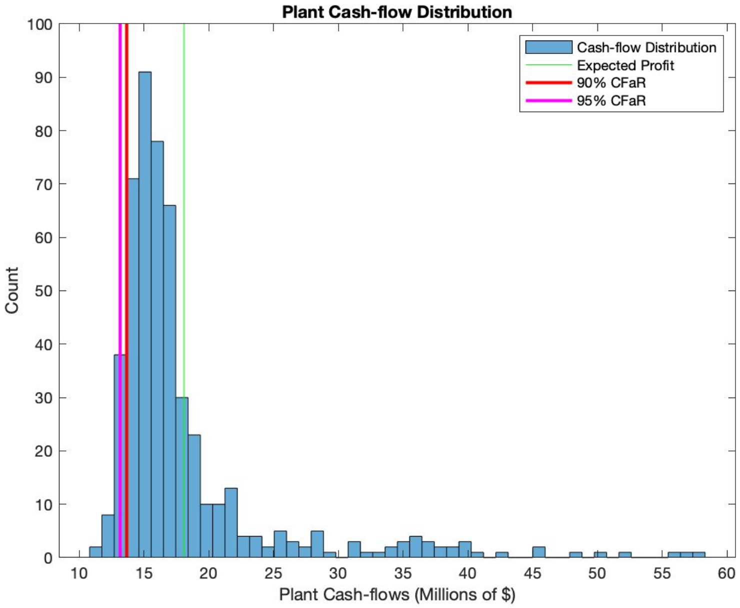
Cash-Flow-at-Risk at 90% confidence level

Std = 0.24

$4.16 Million

Cash-Flow-at-Risk at 95% confidence level

Std = 0.24



Model Assumptions

1. Dispatching

- Dispatch algorithm at ISO is taken as given

2. Market Position

- Not at a disadvantage (e.g. compared to wind plants)
- Faces a residual demand (not first mover), matters for HHI 440 NE Market
- Mid-range capacity
    - Dictates offer curve

3. Perfectly Inelastic Demand

- Consumers do not change consumption given price changes

4. DA LMP matters more than spot prices

- Predictions are done with regards to Day-Ahead prices

Model Assumptions

5. Shock is normally distributed

- Evidenced in the OU Model

6. Weather Station

- Temperature is taken from nearest weather station, not optimized
  - Nearest is not always the most appropriate, as congestion means engineers need to reconfigure the grid

Model Strengths

1. Minimises Mean Absolute Error

- The MAE is minimized at ~$8.48
- Stays constant overtime (the shocks introduced do not change the trend)

2. Optimized Variables

- Models are calibrated with variables that have the highest correlation coefficients
- In cases when not, we experimented:
    - Dry Bulb Temperature --> Dew Point = increased error to $9.8

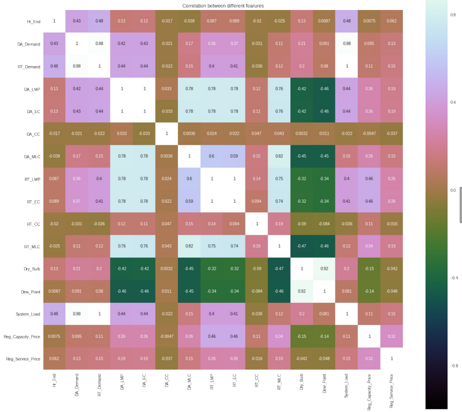
3. Completeness

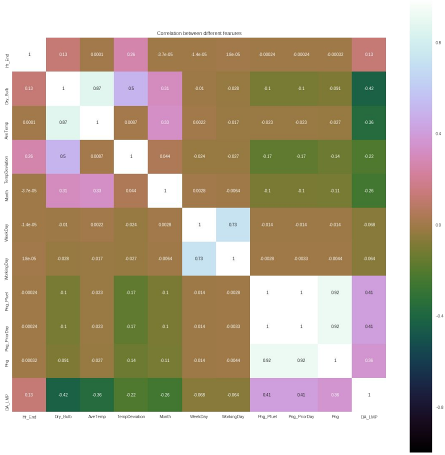
- Regression Tree
    - Using cross validation, we found an optimal level to prune regression tree to avoid overfitting
    - At the 116th level, a pruned tree leads to increased root mean squared error

4. Non-deterministic

- Models allow for multiple continuation at each possible step
    - This is a more realistic simulation for variables like weather
    - This adds randomness to the otherwise deterministic economic dispatch algorithm

Models are calibrated with variables that have the highest correlation coefficients:





Variable Used	Correlation with Day Ahead LMP
Dry Bulb Temperature	-0.42
Average Temperature	-0.36
Temperature Deviation	-0.22
Month	-0.26
Day is a week day	-0.068
Day is a working day	-0.064
Price of Natural Gas - fuel	0.41
Price of Natural Gas - prior day	0.41
Price of Natural Gas	0.36

Natasha Ting, Peter Ogunmodede, Rumman Waqar

Weaknesses / Areas for Improvement

1. New Shocks in ARIMX Model

- Since the output only depends on previous inputs, new fluctuations cannot be taken into account
    - e.g. economic shocks, recessions, policy revisions e.g. NE deregulations, are not accounted for
- *Suggestion for improvement*: carry out additional static analysis

2. Regression Tree

- Output is discrete, arising from the qualifying questions at each decision node
    - This raises root mean squared error
- *Suggestion for improvement*: change to a learning model with continuous outcome

3. Calibrate Specifications

- Use algorithm to find out optimal AR lags in temperature forecasting

4. Introduce Strategic Interaction

- Adding mathematical models to account for other firm's interaction before and after statistical forecasting will improve accuracy of actual cash flow

Weaknesses / Areas for Improvement

5. Explanatory Power of Model

- Adding more variables such as capacity, storage level, simulated forward market prices will add to increased R-squared
- Variables to include may be:
  - Renewable energy demand & supply (wind forecasts, solar radiation forecasts, rainfall, etc)
  - Generation unit outages
  - Import capacity
  - Offer curve

6. Testing other Statistical Methods

- Our model is good with average temperatures, but weak in predicting slumps and spikes
- Other time-series methods have strengths in predicting in the extremes that we can learn from

7. Validation

- Predictions should be tested against observed data to further improve the model