Bunched Random Forest in Monte Carlo Risk Simulation

Eina Ooka

May, 2017

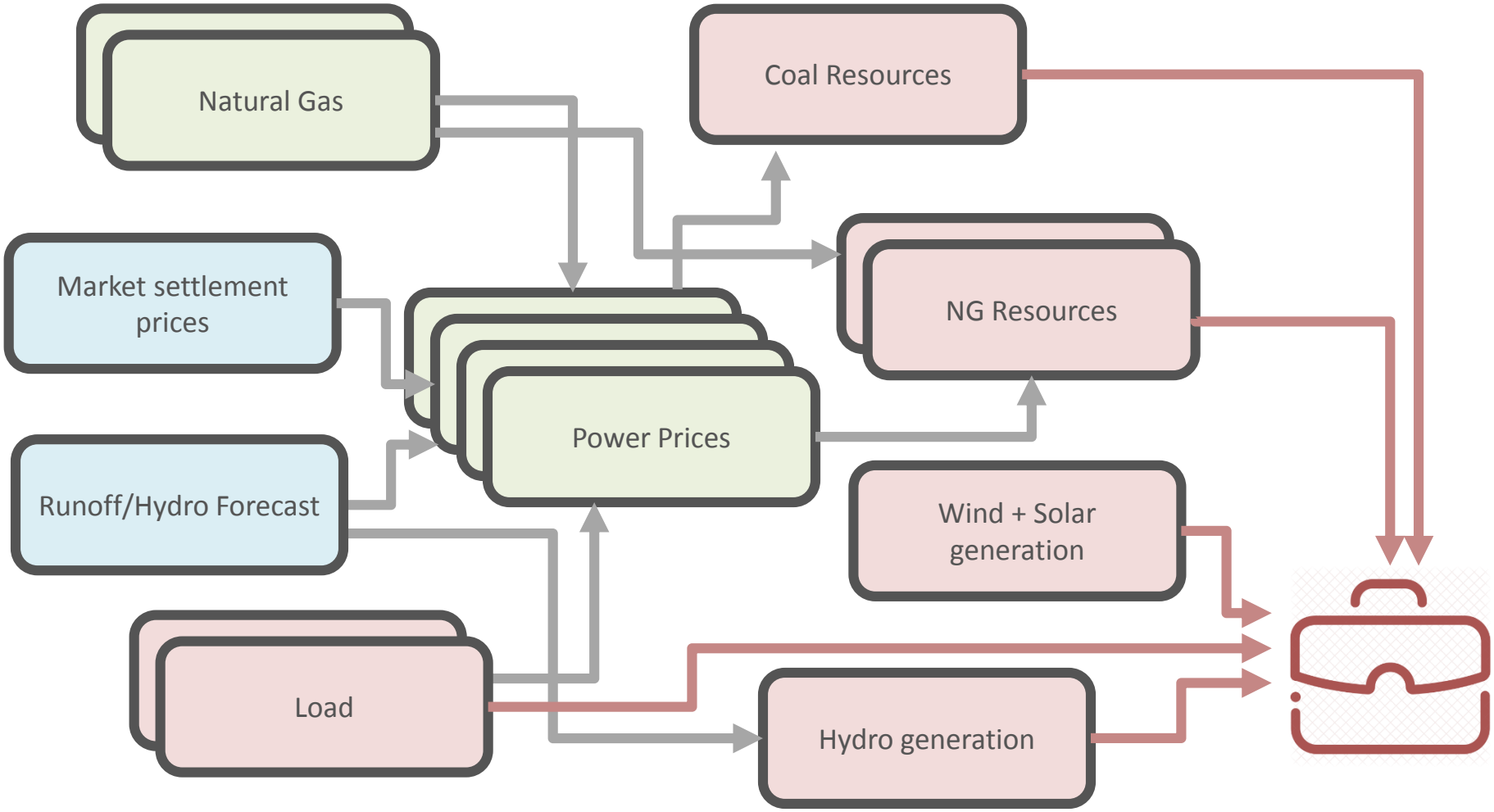
TEA THE EnergyAuthority

R/Finance 2017
In Chicago

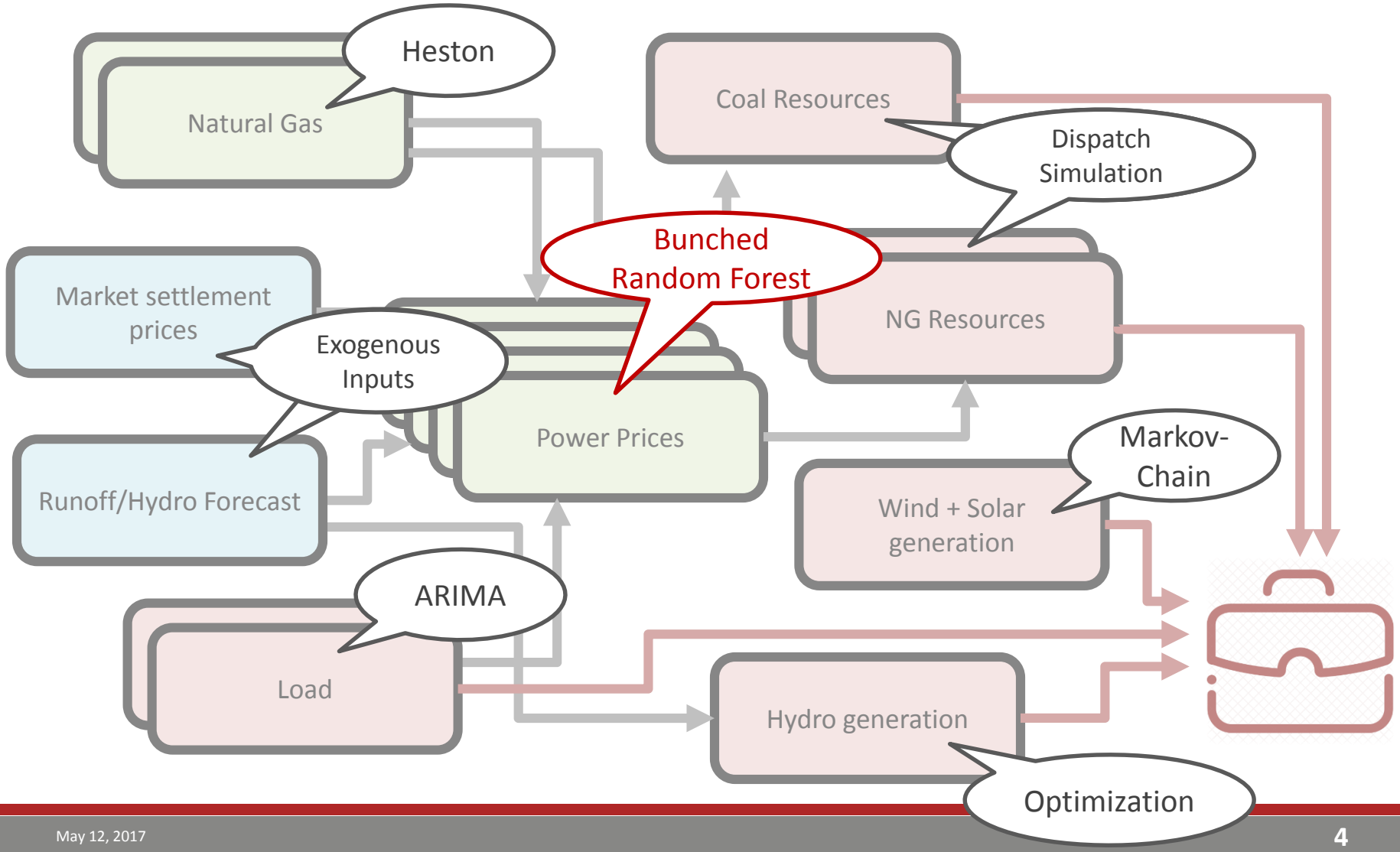
Power Utility Industry

- The Energy Authority serves public utilities nationwide for trading and analytics.
- Mid-term (1 month – 5 years) portfolio management.
- Stochastic simulation models for energy and gas market.

MC Simulation Approach

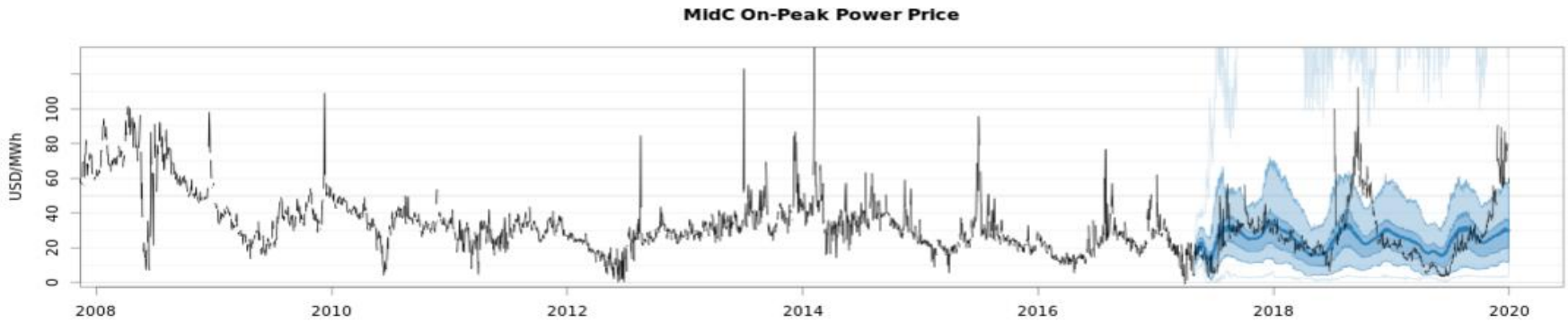


MC Simulation Approach



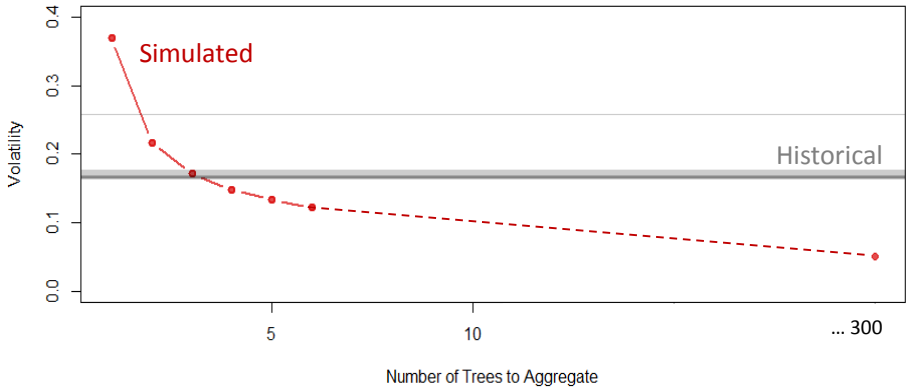
Power Price TS Characteristics

- Autocorrelation
- Seasonal and weekly shapes
- Volatility & Heteroscedasticity
  - Seasonal and weekly variability
- Multivariate cross-correlation and non-linear dependency
  - NG, load, regional hydro and other variable generation
- Non-normal distributions
  - fat tails
  - Extreme peaks and drops
- Negative prices
- Consistency with market expectations
- Consistency between monthly, daily and hourly data.

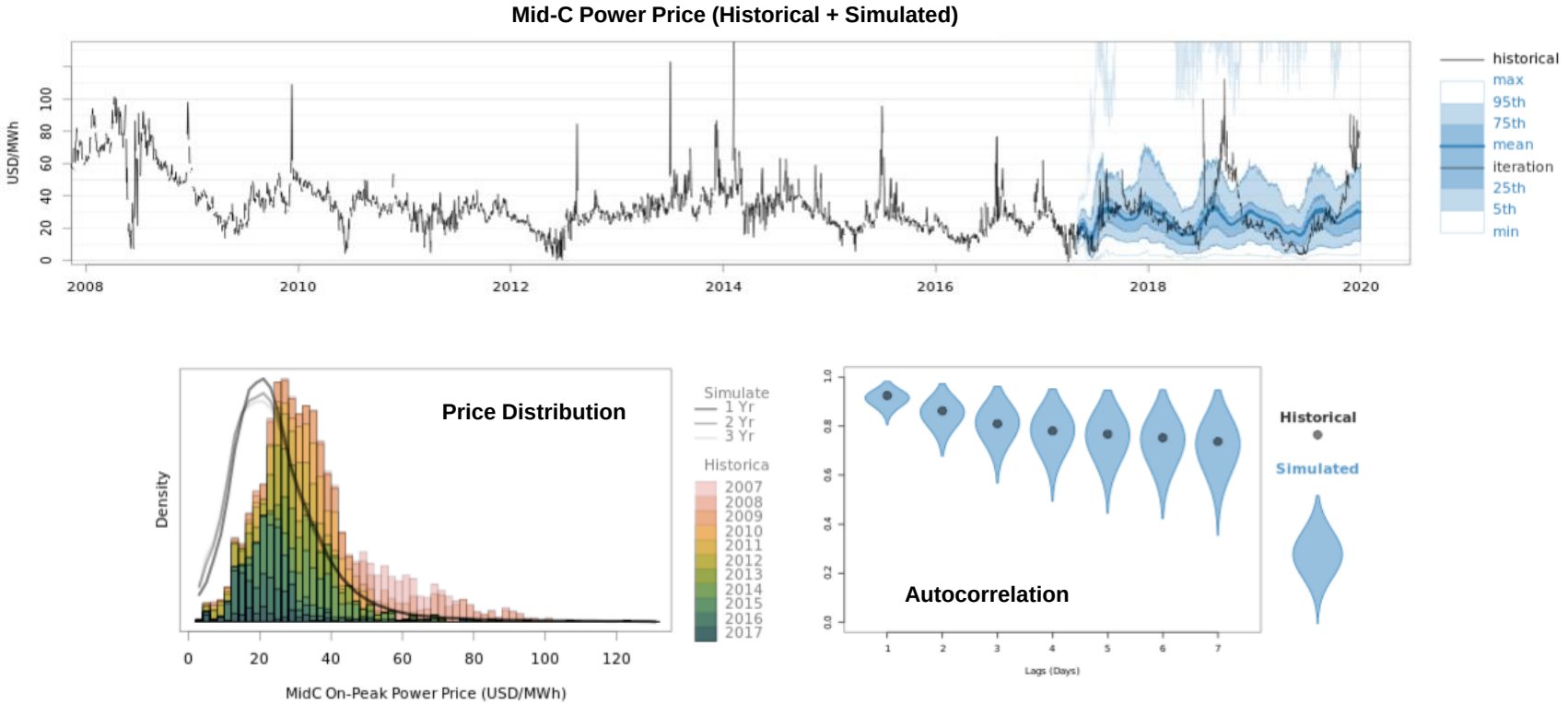


Bunched Random Forest

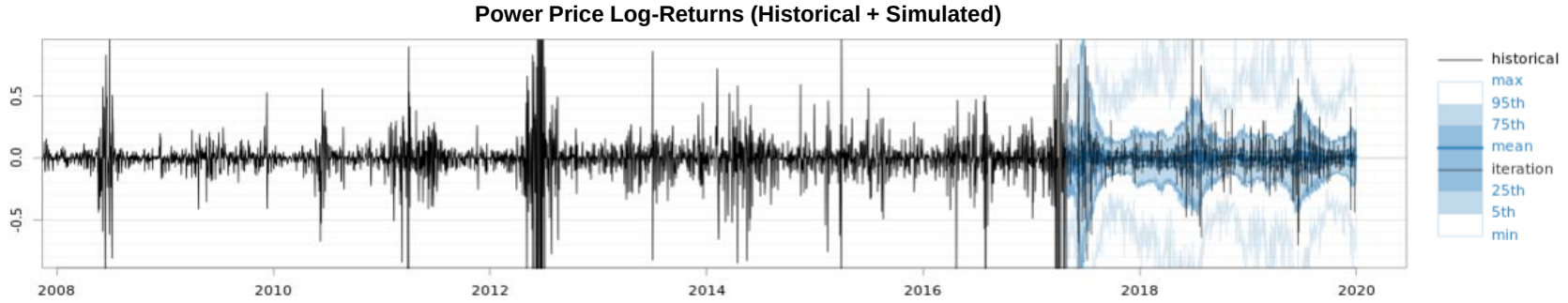
- Traditional Approach for MC
  - $Price = f(\ldots) + Error(\ldots)$
- Random Forest
  - Aggregation of a few hundred trees moderate values too much → **Low volatility**
  - a single tree encompasses too little predictability → **High volatility**
- Bunched RF in MC
  - Aggregating a selected number of trees for each Monte Carlo iteration
  - Achieve plausible volatility in each MC series, while maintaining the same predictability as RF in the whole MC simulation.
  - Implementation
    - randomForest::predict.randomForest() includes individual tree outputs.
    - After figuring out the bunching number based on historical volatility, sample and aggregate suitable number of individual tree outputs.

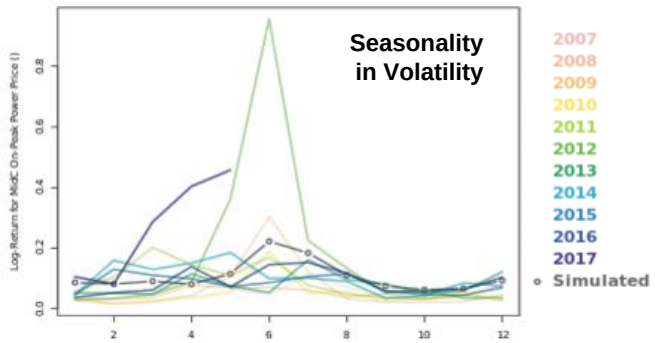


Results


Mid-C Power Price (Historical + Simulated)
USD/MWh
historical
max
95th
75th
mean
iteration
25th
5th
min
Price Distribution
Density
Simulate
1 Yr
2 Yr
3 Yr
Historica
MidC On-Peak Power Price (USD/MWh)
Autocorrelation
Historical
Simulated
Lags (Days)
MidC Power Heat Rate (mmBtu/MWh)
Non-linear dependency on regional hydro
Regional Hydro Generation (aMW)
Seasonality
Weekly Shape
Simulated

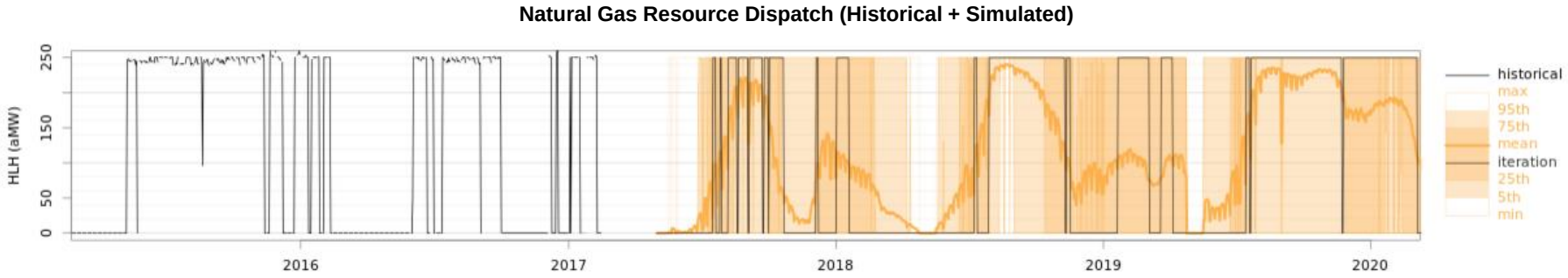
Results

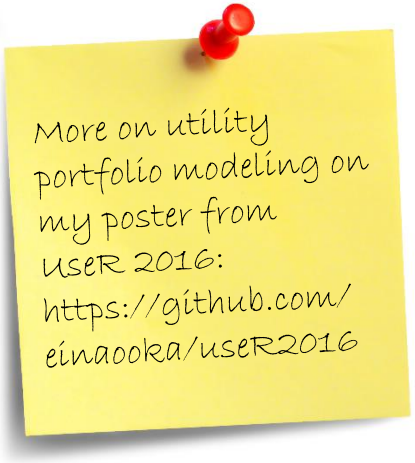




Scoring

- Pinball-loss scoring for stochastic forecasting
- Need to achieve all time series characteristics.
  - Volatility, seasonality, moments, …
- A visual inspection with heuristic reasoning is sometimes necessary.
- Backcast and feed the generated power prices to Natural Gas dispatch model, and look at pinball-loss scoring on generation outputs.





Thank you!

Contact:
eooka@teainc.org

! The Energy Authority is looking for an intern.