All-Inclusive but Practical **Multivariate Stochastic Forecasting** for Electric Utility Portfolio

AUTHOR: Eina Ooka
Sr. Structure and Pricing Analyst
The Energy Authority
eooka@teainc.org
https://github.com/einaooka/useR2016
(See this link for references)

Objective

- **Quantify market risk that utilities are exposed to, through stochastic simulation of loads, resources and market prices.**
- Portray appropriate tail risk with 1,000 or more stochastic iterations.
- Capture the following data characteristics:
  - i. Autocorrelation
  - ii. Heteroscedasticity
  - iii. Seasonal and weekly shape
  - iv. Mutivariate cross-correlation and non-linear dependency
  - v. Non-normal distributions
  - vi. Extreme peaks and drops
  - vii. Negative prices
  - viii. Max and min caps
  - ix. Reasonable long-term distributions
  - x. Consistency with market and utility expectations

Background Facts

- The wholesale power market is the most volatile market of all commodity markets.
- Consumers pay a relatively flat electricity rate to utilities. This means that utilities need to manage market risks.
- Some utilities set annual budgets probabilistically, meaning that the chance of insolvency given a specified cash reserve must be below a certain threshold.
- Many utilities hedge risk by trading forwards or options.

Scheduled Tasks: Stochastic Variable Generation; Resource Dispatch Optimization
On Demand: Hedge Analysis

DATA [objective]

Daily historical data
- On-peak **Load**
- Off-peak Load
- **Hydro** generations
- **Natural gas** prices
- **Power** prices

[i – viii]

External LT forecasts in **monthly** granularity.
- Econometric load forecasts
- Metrological hydro forecasts

[i – v, viii – x]

Market settlement data
- Forward prices

[x]

Wind data is independent and does not fit AR model. [iii, v, viii]

Hourly congestion price shapes at multiple (10 - 30) nodes. [i – viii]

Generation by thermal resources are determined economically based on prices and its capability. Need to take into account of future **outage schedules** set by utilities. [iv, x]

Ultimately, we want some **Risk Metrics** that guide us to avert risks.

STEPS

1. Data Scraping
2. Collinearity and Dependency Detection
3. Multiplicative Decomposition
   Outlier + Seasonality x Trend x Noise
4. Long-Term Stochastic Forecasting
5. Daily Volatility & Spike Model
6. Disaggregation and Combination
7. Judgmental Adjustment
8. Markov Chain Model
9. Predictive Modeling
10. Hourly-Shape Analysis
11. Resource Dispatch
12. Cash Flow Aggregation
13. Scenario & Hedge Analysis
14. Portfolio Optimization

METHODS

Hierarchical Clustering
Collinear variables are taken out of the main AR model and derived later using predictive models.

Grubb's Outlier Test
Applied twice before and after decomposition.

Fourier series

LOESS

Subset Regression
Weekly shapes and auto-regression factors are captured here after decomposition.

Multivariate Gaussian Model

Héston Model

ARIMA(1,2,0)
If an external LT forecast does not exist, it is derived from trend and seasonality components (step 3).

Seasonal Block Bootstrapping

Multiple Regression

Denton Disaggregation
Monthly LT forecasts (step 4) need to be disaggregated into daily time series, before combining with daily noise (step 5).

Temporal Hierarchical Reconciliation
Generated stochastic data is reconciled so that means match with market expectations.

Random Forest

Feature Selection
Collinear or dependent variables (step 2) are derived using predictive models. All trees of RF are utilized to capture variability.

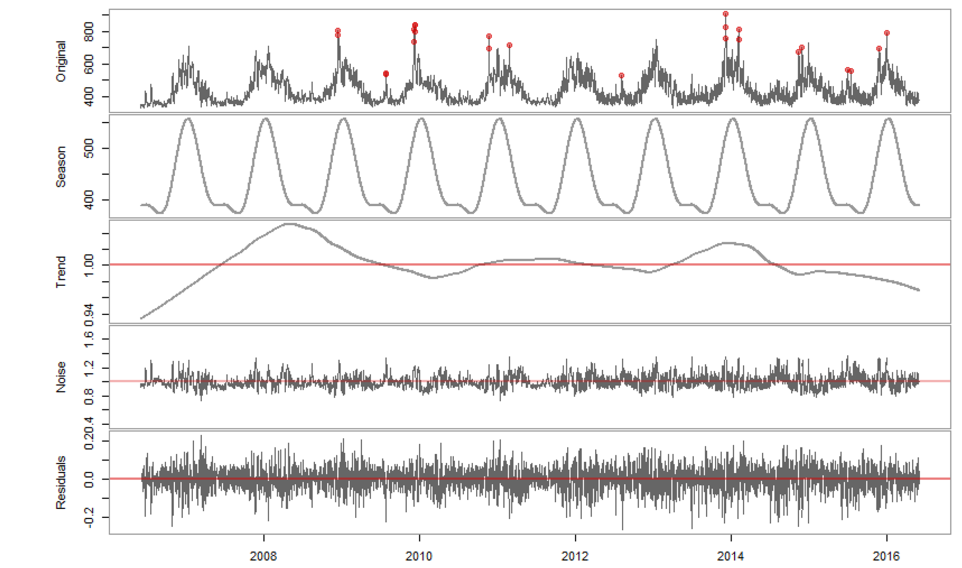
Logical Dispatch Calculation
Given resource unit specifications, determine the optimal economic dispatch schedule for each stochastic iteration. This achieves the perfect hindsight dispatch.

Logistic Regression/Random Forest
In reality, resources does not dispatch like the logical model, due to lack of knowledge in future prices. Predictive models simply derive dispatch patterns from the historical generation, instead of optimizing the cash flow.

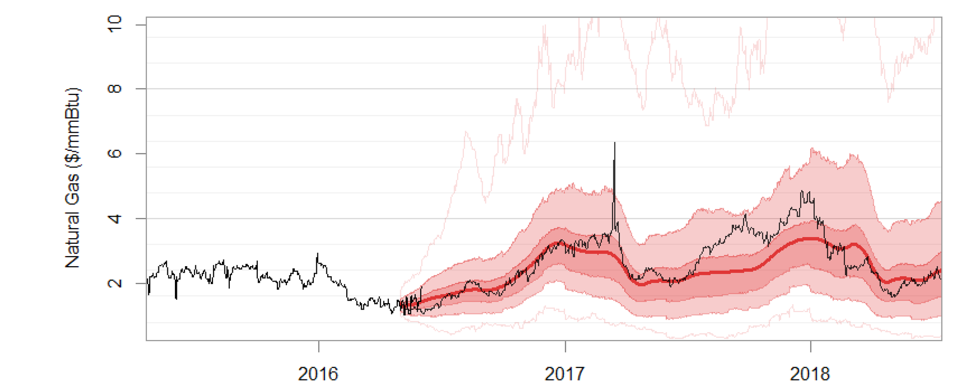
Portfolio Analysis through R-Shiny App
A user can interactively analyze stochastic scenarios and run hypothetical forward or option trades through shiny app.

Report through Rmarkdown
Hedge analysis can be made into a html report through the app.

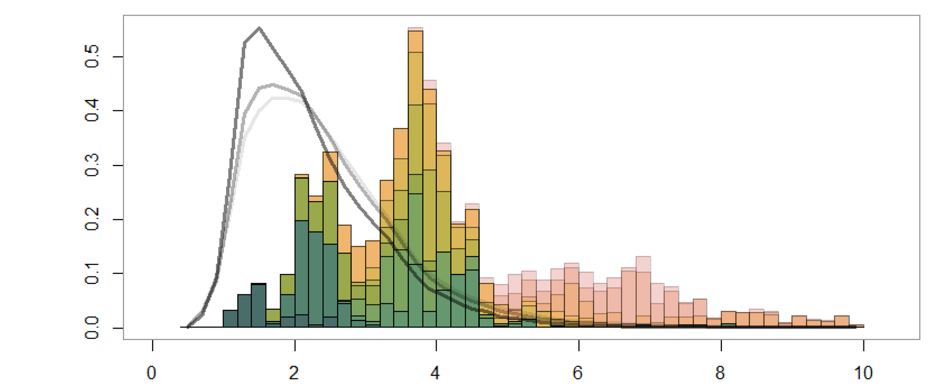
Time Series Decomposition



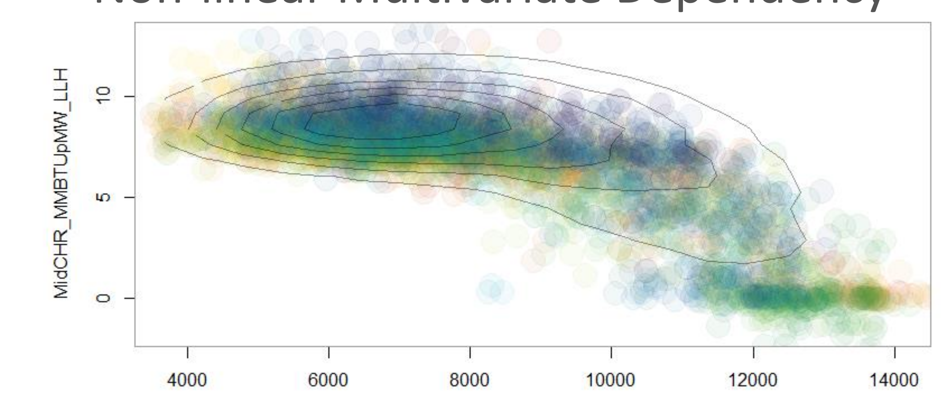
Simulated Time Series



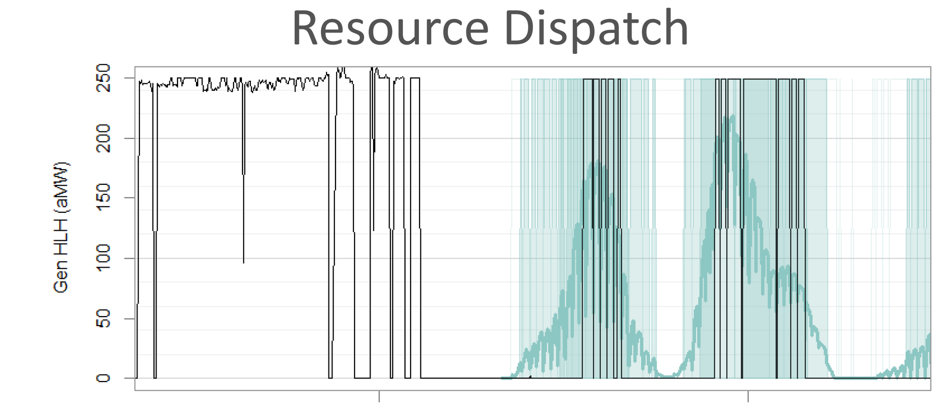
Historical vs. Simulated



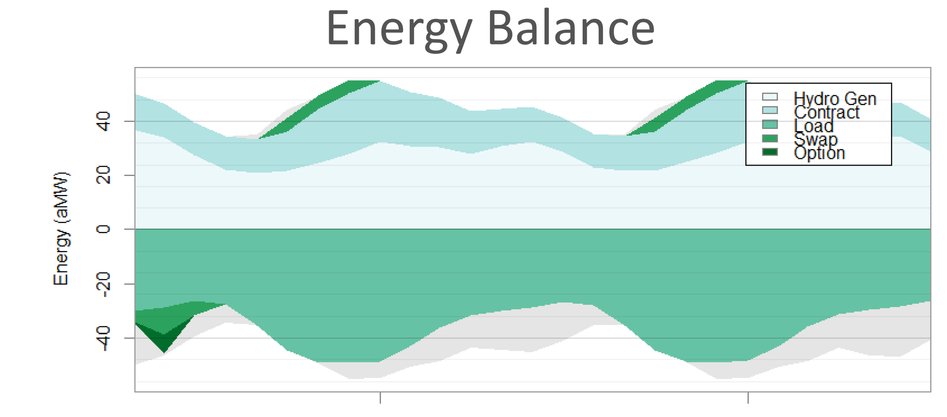
Non-linear Multivariate Dependency



Resource Dispatch



Energy Balance



Hedged vs. Unhedged Portfolio Cash Flow