

Project step 2: scenario-based stochastic optimisation

S. Drake Siard
DTU 31792, Spring 2021

21 Jan 2021

1 Problem definition

The day-ahead power market-clearing problem with transport constraints from [1], that was solved in step 1 of the project, was extended to include stochastic real-time market clearing as in [2].

The notation follows a similar convention as before, extended from a single market to the day-ahead and real-time markets. For a set of generators $\{g\}$ with production costs $\{C_g\}$, demands $\{d\}$ with utility $\{U_d\}$ and common load shedding disutility U_{shed} , buses $\{n\}$, and connections $\{t_{n,m}\}$, denote:

- day-ahead generation and consumption as $p_g^{G,DA}$ and $p_d^{D,DA}$
- the set of real-time scenarios as Ω , the individual scenarios as ω , and the corresponding probability π_ω
- real-time change in scenario ω from day-ahead commitment as $p_{\omega,g}^{G,RT}$ and $p_{\omega,d}^{D,RT}$
- maximum generation for g as \bar{P}_g^G , maximum demand for d as \bar{P}_d^D , and maximum flow through connection $t_{n,m}$ as $F_{n,m}$
- actual wind generation in scenario ω as $\bar{P}_{\omega,g}^{G,RT}$
- the difference in voltage angle between buses n and m as $\theta_n - \theta_m$ (in scenario ω : $\theta_{\omega,n} - \theta_{\omega,m}$), the susceptance of their connection $t_{n,m}$ as $B_{n,m}$, and the (rough) approximation of flow between them as $f_{n,m} = B_{n,m}(\theta_n - \theta_m)$
- the set of generators and demands in bus n as Ψ_n , and the set of buses connected to bus n via transmission lines as Ω_n
- θ_{n_0} as the reference voltage angle of some arbitrary bus n_0

The problem is still the maximisation of social welfare (the sum of the producer and consumer surplus), extended to the real-time market:

$$\max_{p_g^{G,DA}, p_d^{D,DA}, \theta_n} \left[\sum_d U_d p_d^{D,DA} - \sum_g C_g p_g^{G,DA} \right] + \sum_\omega \pi_\omega \left(\sum_d U_{shed} p_{\omega,d}^{D,RT} - \sum_g C_g p_{\omega,g}^{G,RT} \right) \quad (1.1)$$

In this case U_{shed} is equal to the inferred utility of the inelastic load, \$2000/MW. The objective is optimised subject to both the original day ahead constraints:

$$0 \leq p_d^{D,DA} \leq \bar{P}_d^D \quad \forall d \quad (1.2)$$

$$0 \leq p_g^{G,DA} \leq \bar{P}_g^G \quad \forall g \quad (1.3)$$

$$-F_{n,m} \leq f_{n,m} \leq F_{n,m} \quad \forall t_{n,m} \quad (1.4)$$

$$\theta_{n_0} = 0 \quad (1.5)$$

$$f_{n,m} = B_{n,m}(\theta_n - \theta_m) \quad \forall t_{n,m} \quad (1.6)$$

$$\sum_{d \in \Psi_n} p_d^{D,DA} + \sum_{m \in \Omega_n} f_{n,m} - \sum_{g \in \Psi_n} p_g^{G,DA} = 0 \quad \forall n \quad (1.7)$$

and to the real-time market constraints per scenario:

$$0 \leq p_d^{D,DA} + p_{\omega,d}^{D,RT} \leq \bar{P}_d^D \quad \forall \omega, d \quad (1.8)$$

$$p_{\omega,d}^{D,RT} \leq 0 \quad \forall \omega, d \quad (1.9)$$

$$0 \leq p_g^{G,DA} + p_{\omega,g}^{G,RT} \leq \bar{P}_g^G \quad \forall \omega, g \quad (1.10)$$

$$p_{\omega,g}^{G,RT} = 0 \quad \forall \omega, g \mid g \in \text{inflexible} \quad (1.11)$$

$$p_{\omega,g}^{G,RT} = \bar{P}_{\omega,g}^{G,RT} - p_g^{G,DA} - p_{\omega,g}^{spill} \quad \forall \omega, g \mid g \in \text{wind} \quad (1.12)$$

$$0 \leq p_{\omega,g}^{spill} \leq \bar{P}_{\omega,g}^{G,RT} \quad \forall \omega, g \mid g \in \text{wind} \quad (1.13)$$

$$-F_{n,m} \leq f_{\omega,n,m} \leq F_{n,m} \quad \forall \omega, t_{n,m} \quad (1.14)$$

$$\theta_{\omega,n_0} = 0 \quad \forall \omega \quad (1.15)$$

$$f_{\omega,n,m} = B_{n,m}(\theta_{\omega,n} - \theta_{\omega,m}) \quad \forall \omega, t_{n,m} \quad (1.16)$$

$$\sum_{d \in \Psi_n} (p_d^{D,DA} + p_{\omega,d}^{D,RT}) + \sum_{m \in \Omega_n} f_{\omega,n,m} - \sum_{g \in \Psi_n} (p_g^{G,DA} + p_{\omega,g}^{G,RT}) = 0 \quad \forall \omega, n \quad (1.17)$$

Once again, the flows are represented as additional variables with their own bounds (Eqs. 1.4 and 1.14) and explicitly linked to the voltage angles (Eqs. 1.6 and 1.16) for ease of retrieval in the implementation. Variable wind production is represented as a change in production which will either be used or spilled, with the net change represented by $p_{\omega,g}^{G,RT}$ (Eq. 1.12). No attempt was made to price flexibility in the objective: it was assumed that flexible real-time generation costs were symmetric and identical to day-ahead costs. This problem was implemented in MOSEK's Fusion API, as in step 1.

2 Scenario generation

As before, the day-ahead bid and offer data for 12:00-01:00 on October 7, 2020 from the MISO Market Data site [3] were used as a base, with some additional clean-up (such as excluding generators with the "Unit Available Flag" unset). Generators were divided into three categories: flexible, inflexible, and wind; generators with an offer of zero were assumed to be wind, and those with an offer price above the 75th percentile were assumed to be flexible. Flexible generators were assumed flexible over their entire generation range (from zero to the offered capacity).

For wind generation, MISO historical hourly wind generation data for that hour reported a total of 11937.76 MWh, which was approximately three times the offered capacity

attributed to wind generators. The default scenario was constructed by scaling the bid capacity of each (inferred) wind generator by a common factor so that the total wind capacity matched the historical total wind generation.

To create the scenario sets, reverse forecast error was then applied. [4] reports the first four moments of forecast error for the ERCOT interconnection in Texas for the year 2010. By sampling a uniform random variable mapped through the inverse CDF of a distribution¹ matching those moments, a distribution of wind levels that matches the distribution of the empirical forecast error can be derived. In order to produce correlated results, one sample per region was combined with one sample per generator to produce the per-generator per-scenario scalar:

$$\epsilon_{g,\omega} = \Phi_{\mu,\sigma,s,\kappa}^{-1}(\alpha X_{n,\omega} + (1 - \alpha)X_{g,\omega}) \quad \text{where} \quad X_{n,\omega}, X_{g,\omega} \sim U(0, 1), \alpha = 0.9 \quad (2.1)$$

The scalar ϵ_g was multiplied with the default scenario, and the whole process was repeated 1000 times. Fig. 1a shows the output distribution of two wind generators in the same region, and Fig. 1b shows the joint distribution for generators in different regions. The univariate distributions are visibly fat-tailed (leptokurtic) and negatively skewed, with a shape directly comparable to [4].

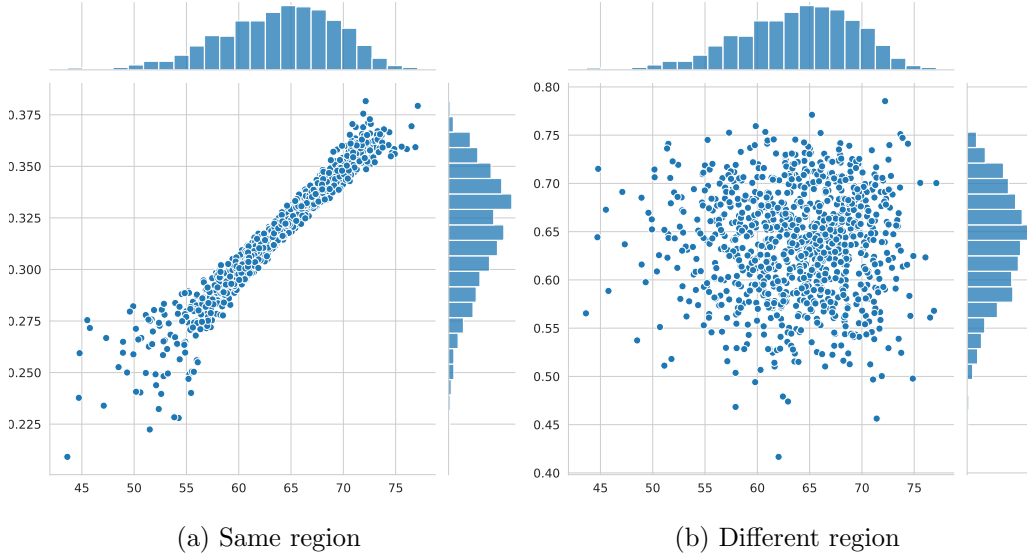


Figure 1: Original data correlations

3 Training and testing

Of the 1000 scenarios, the first 100 were used as a training set, to produce a day-ahead solution, and the remaining 900 as a test and validation set, using the fixed day-ahead solution from the training set². The training optimisation, with 346,336 constraints and 349,496 variables (before pre-solving), took 4.8 seconds; the test optimisation, with 3,086,107 constraints and 3,089,267 variables, took 9.1 seconds. In both cases, the actual optimisation time was dwarfed by the overhead in Python of problem setup, as it was in step 1.

¹See https://www.statsmodels.org/stable/generated/statsmodels.sandbox.distributions.extras.pdf_mvsk.html for more details.

²Implementation in `proj2-stochastic-training.ipynb` and `proj2-stochastic-test.ipynb`

For both the test and training sets, the per-scenario utility was calculated:

$$U_{\omega} = \left(\sum_d U_d p_d^{D,DA} - \sum_g C_g p_g^{G,DA} \right) + \left(\sum_d U_{shed} p_{\omega,d}^{D,RT} - \sum_g C_g p_{\omega,g}^{G,RT} \right) \quad (3.1)$$

Fig. 2 shows (normalised) histograms and density plots for the objective values of the test and training sets. Given that the scenarios are generated with a uniform seed distribution, any random sample of the scenario space is likely to be representative. In addition, the day-ahead solution from the training set came from a large (100) number of scenarios. It is therefore no surprise that the per-scenario objective values of the test and training data sets are a close visual match. More precisely, a (median-centered) Levene test fails to reject the null hypothesis of equal variance (p -value: 0.22), and a two-sample t -test fails to reject the null hypothesis of equal means (p -value: 0.21).

In summary, given forecast errors in line with the ones used above, the stochastic optimisation produces a something very close to an ideal day-ahead clearing solution. However, the extreme negative skew of the objective function values suggests that, depending on the system operator's goals, a more robust optimisation technique might be appropriate.

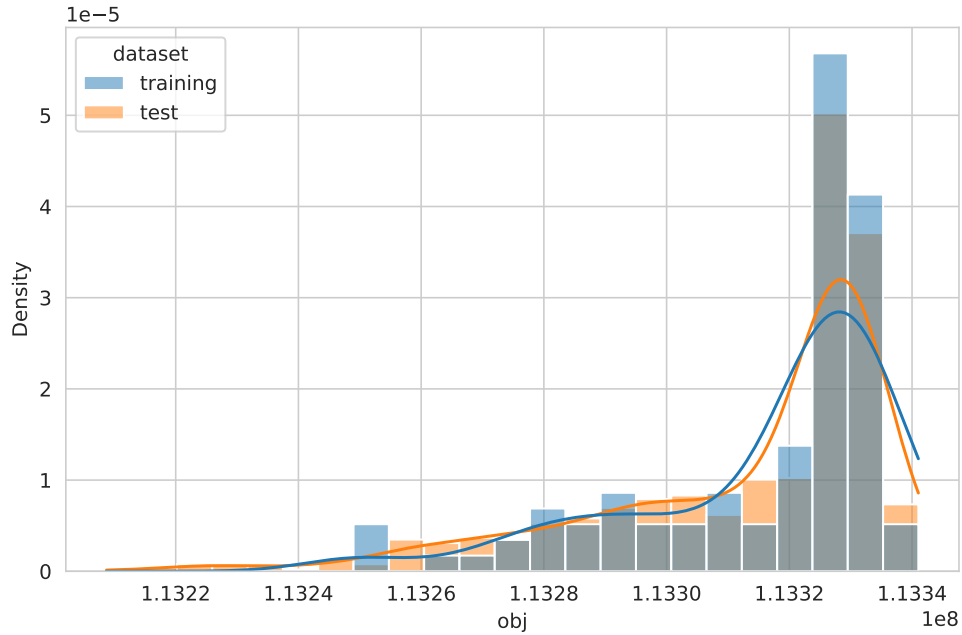


Figure 2: Objective function distribution

4 References

- [1] J. Kazempour, “Lecture 1: Market clearing as an optimization problem,” DTU 31792 – Advanced Optimization and Game Theory for Energy Systems, Jan. 2021. [Online]. Available: <https://cn.inside.dtu.dk/cnnet/filessharing/download/dcb15e39-cb18-4656-a13b-fec8ef3201bd>
- [2] —, “Lecture 5: Stochastic market clearing,” DTU 31792 – Advanced Optimization and Game Theory for Energy Systems, Jan. 2021. [Online]. Available: <https://cn.inside.dtu.dk/cnnet/filessharing/download/ab8abfcd-f87b-4087-b81b-f66bfeebfd4a>
- [3] “MISO Market Data.” [Online]. Available: <https://www.misoenergy.org/markets-and-operations/real-time--market-data/market-reports/#t=10&p=0&s=MarketReportPublished&sd=desc>
- [4] B. M. Hodge, D. Lew, M. Milligan, H. Holttinen, S. Sillanpaa, E. Gomez-Lazaro, R. Scharff, L. Soder, X. G. Larsen, G. Giebel, D. Flynn, and J. Dobschinski, “Wind Power Forecasting Error Distributions: An International Comparison; Preprint,” National Renewable Energy Lab. (NREL), Golden, CO (United States), Tech. Rep. NREL/CP-5500-56130, Sep. 2012. [Online]. Available: <https://www.osti.gov/biblio/1051129>