Value-oriented forecasting of net demand for electricity market clearing

Juan M. Morales[1]
juan.morales@uma.es

University of Malaga
OASYS group
oasys.uma.es

EURO
Espoo, July 4, 2022

[1]With the co-authorship of Miguel Á. Muñoz and Salvador Pineda

Introduction: Conditional Stochastic Problem

Solve

$$J^* := \inf_{x \in X} \mathbb{E}_{\mathbb{Q}} \left[f(x, y) \Big/ z = z_0 \right] = \mathbb{E}_{\mathbb{Q}_{\widetilde{\Xi}}} [f(x, y)]$$

with optimal solution x^*.

- The decision variable $x \in X \subseteq \mathbb{R}^n$.
- A random vector y with support set $\Xi_y \subseteq \mathbb{R}^{d_y}$, whereby we model the uncertainty affecting the value of the decision. Example: demand...
- Covariates (or features) modeled by a random vector z with support set $\Xi_z \subseteq \mathbb{R}^{d_z}$. Example: weather forecast, hashtag twitter...
- Side information: $\widetilde{\Xi} = \{(z, y) : z = z_0\}$

Introduction: Conditional Stochastic Problem

- At 10 am we have to decide how much ice cream to make (decision x)
- At 10 am we do not know the demand in the afternoon (uncertain parameter y)
- We can use some available information such as the temperature at 10 am (contextual data z)
- Obviously there is a relationship between the morning temperature (z) and the ice cream demand in the afternoon (y).
- We would like to use such a relation to make better decisions about ice cream quantity

Introduction: Conditional Stochastic Problem

Fundamental challenge

- The true joint distribution $\mathbb{Q}$ and the true conditional distribution $\mathbb{Q}_{\tilde{\Xi}}$ are unknown!
- All that we have is a data sample of size N, i.e., $(\widehat{z}_i, \widehat{y}_i)_{i=1}^N$.

Goal

We want to exploit the info on the features, z_0 (the context), in our favor to prescribe better decisions.

Some possible solution approaches

(Traditional) Forecasting

Compute $\widehat{y}(z_0) \approx \mathbb{E}[y|z = z_0]$ in the hope that

$$(arg) \inf_{x \in X} \mathbb{E}_{\mathbb{Q}} \left[f(x, y) \Big/ z = z_0 \right] \approx (arg) \inf_{x \in X} f(x, \widehat{y}(z_0))$$

Uncertainty, and thus its impact, are ignored.

Smart predict (value-oriented forecasting)

Find $\widehat{y}(\cdot) : \Xi_z \longrightarrow \Xi_y$ (very possibly different from $\mathbb{E}[y|z]$) such that

$$(arg) \inf_{x \in X} \mathbb{E}_{\mathbb{Q}} \left[f(x, y) \Big/ z = z_0 \right] \approx (arg) \inf_{x \in X} f(x, \widehat{y}(z_0))$$

In any case, we must *infer* from the data sample $(\widehat{z}_i, \widehat{y}_i)_{i=1}^N$

Value-oriented forecasting

- Predict, then optimize
  - learns the relation between y and z *ignoring* f and X
- *Smart* predict, then optimize (value-oriented forecasting)
  - learns the relation between y and z *taking into account* f and X

Contextual merit-order dispatch under uncertainty

A TWO-STAGE ELECTRICITY MARKET

Forward market

$$\min_{g_j,\, j\in G} \sum_{j\in G} c_j g_j$$

$$\text{s.t.} \quad \sum_{j\in G} g_j = \widehat{L}$$

$$0 \le g_j \le g_j^{\max}, \quad \forall j \in G$$

Real-time market

$$\min_{\Xi} \sum_{j\in G} (c_j^{\mathrm{U}} r_j^{\mathrm{U}} - c_j^{\mathrm{D}} r_j^{\mathrm{D}})$$

$$\text{s.t.} \quad 0 \le g_j^* + r_j^{\mathrm{U}} - r_j^{\mathrm{D}} \le g_j^{\max}, \quad \forall j \in G$$

$$0 \le r_j^{\mathrm{U}} \le R_j^{\mathrm{U}}, \quad \forall j \in G$$

$$0 \le r_j^{\mathrm{D}} \le R_j^{\mathrm{D}}, \quad \forall j \in G$$

$$\sum_{j\in G} (g_j^* + r_j^{\mathrm{U}} - r_j^{\mathrm{D}}) = \sum_{b\in\mathcal{B}} L_b$$

$$|\mathbf{m}_l(\sum_{j\in G_b} g_j^* + r_j^{\mathrm{U}} - r_j^{\mathrm{D}} - L_b)| \le \mathrm{Cap}_l, \forall l \in \mathcal{L}$$

- $\widehat{L}$ is a **point estimate**
- Units are sequentially dispatched from lowest to highest cost

- $\sum_{b\in\mathcal{B}} L_b$ is the *realized* net system demand
- $\Xi := \{r_j^{\mathrm{U}}, r_j^{\mathrm{D}} \in \mathbb{R}^+, j \in G\}$

Contextual merit-order dispatch under uncertainty

- $\widehat{L}$ is a **point estimate**.
- Available historical data: $(\widehat{f}_i, \widehat{\boldsymbol{L}}_i)_{i=1}^N$, where f is the classical point prediction and $\boldsymbol{L}$ is the vector of nodal net demands.
- Context: $z = f$ Uncertainties: $y = \boldsymbol{L}$

Forward market

$$\min_{g_j,\, j \in G} \sum_{j \in G} c_j g_j$$

$$\text{s.t.} \quad \sum_{j \in G} g_j = \widehat{L}$$

$$0 \leq g_j \leq g_j^{\max}, \ \forall j \in G$$

Option 1

$\widehat{L} := f$ (usual practice)

Option 2

$\widehat{L} := q_0 + q_1 f$ (forecast correction)

How do we estimate q_0 and q_1?

Estimation of q_0 and q_1: Model training

$$\min_{\mathbf{q},\Upsilon} \frac{1}{N}\sum_{i\in[N]}\sum_{j\in G}(c_j g_{ji} + c_j^{\mathsf{U}} r_{ji}^{\mathsf{U}} - c_j^{\mathsf{D}} r_{ji}^{\mathsf{D}})$$

$$\text{s.t.} \sum_{j\in G} g_{ji} = \widehat{L}_i := q_0 + q_1\widehat{f}_i, \ \forall i \in [N]$$

$$\text{s.t.} \ 0 \le g_{ji} + r_{ji}^{\mathsf{U}} - r_{ji}^{\mathsf{D}} \le g_j^{\max}, \ \forall j \in G, \ \forall i \in [N]$$

$$0 \le r_{ji}^{\mathsf{U}} \le R_j^{\mathsf{U}}, \ \forall j \in G, \ \forall i \in [N]$$

$$0 \le r_{ji}^{\mathsf{D}} \le R_j^{\mathsf{D}}, \ \forall j \in G, \ \forall i \in [N]$$

$$\sum_{j\in G}(g_{ji} + r_{ji}^{\mathsf{U}} - r_{ji}^{\mathsf{D}}) = \sum_{b\in\mathcal{B}} \widehat{L}_{bi}, \ \forall i \in [N]$$

$$|\mathbf{m}_l(\sum_{j\in G_b}(g_{ji} + r_{ji}^{\mathsf{U}} - r_{ji}^{\mathsf{D}}) - \widehat{L}_{bi})| \le \mathsf{Cap}_l, \ \forall l \in \mathcal{L}, \ \forall i \in [N]$$

$$u_{ji} g_j^{\max} \le g_{ji} \le u_{(j-1)i}\, g_j^{\max}, \ \forall j \in G : j > 1, \ \forall i \in [N]$$

$$u_{ji} g_j^{\max} \le g_{ji} \le g_j^{\max}, \ j = 1, \ \forall i \in [N]$$

$$u_{(j-1)i} \le u_{ji}, \ \forall j \in G : j > 1, \ \forall i \in [N]$$

$$u_{ji} \in \{0,1\}, \ \forall j \in G, \ \forall i \in [N]$$

$$\Upsilon := \{g_{ji}, r_{ji}^{\mathsf{U}}, r_{ji}^{\mathsf{D}}, u_{ji}\}_{\{i,j,l\}}$$

We train the *forecast correction rule* on the data set $(\widehat{f}_i, \widehat{\boldsymbol{L}}_i)_{i=1}^N$

Estimation of q_0 and q_1: Model training

$$\Upsilon := \{g_{ji}, r_{ji}^{\mathrm{U}}, r_{ji}^{\mathrm{D}}, u_{ji}\}_{\{i,j,l\}}$$

$$\min_{\mathbf{q},\Upsilon} \frac{1}{N} \sum_{i\in[N]} \sum_{j\in G} (c_j g_{ji} + c_j^{\mathrm{U}} r_{ji}^{\mathrm{U}} - c_j^{\mathrm{D}} r_{ji}^{\mathrm{D}})$$

$$\text{s.t.} \quad \sum_{j\in G} g_{ji} = \widehat{L}_i := q_0 + q_1 \widehat{f}_i, \quad \forall i \in [N]$$

$$\text{s.t.} \quad 0 \le g_{ji} + r_{ji}^{\mathrm{U}} - r_{ji}^{\mathrm{D}} \le g_j^{\max}, \quad \forall j \in G, \quad \forall i \in [N]$$

$$0 \le r_{ji}^{\mathrm{U}} \le R_j^{\mathrm{U}}, \quad \forall j \in G, \quad \forall i \in [N]$$

$$0 \le r_{ji}^{\mathrm{D}} \le R_j^{\mathrm{D}}, \quad \forall j \in G, \quad \forall i \in [N]$$

$$\sum_{j\in G} (g_{ji} + r_{ji}^{\mathrm{U}} - r_{ji}^{\mathrm{D}}) = \sum_{b\in\mathcal{B}} \widehat{L}_{bi}, \quad \forall i \in [N]$$

$$|\mathbf{m}_l (\sum_{j\in G_b} (g_{ji} + r_{ji}^{\mathrm{U}} - r_{ji}^{\mathrm{D}}) - \widehat{L}_{bi})| \le \mathrm{Cap}_l, \quad \forall l \in \mathcal{L}, \quad \forall i \in [N]$$

$$u_{ji} g_j^{\max} \le g_{ji} \le u_{(j-1)i}\, g_j^{\max}, \quad \forall j \in G : j > 1, \quad \forall i \in [N]$$

$$u_{ji} g_j^{\max} \le g_{ji} \le g_j^{\max}, \quad j = 1, \quad \forall i \in [N]$$

$$u_{(j-1)i} \le u_{ji}, \quad \forall j \in G : j > 1, \quad \forall i \in [N]$$

$$u_{ji} \in \{0, 1\}, \quad \forall j \in G, \quad \forall i \in [N]$$

Replicate real-time operation per data point $(\widehat{f}_i, \widehat{\boldsymbol{L}}_i)$

Estimation of q_0 and q_1: Model training

$$\Upsilon := \{g_{ji}, r_{ji}^{\mathsf{U}}, r_{ji}^{\mathsf{D}}, u_{ji}\}_{\{i,j,l\}}$$

$$\min_{\mathbf{q},\Upsilon} \frac{1}{N} \sum_{i\in[N]} \sum_{j\in G} (c_j g_{ji} + c_j^{\mathsf{U}} r_{ji}^{\mathsf{U}} - c_j^{\mathsf{D}} r_{ji}^{\mathsf{D}})$$

$$\text{s.t.} \sum_{j\in G} g_{ji} = \widehat{L}_i := q_0 + q_1 \widehat{f}_i, \ \forall i \in [N]$$

$$\text{s.t.}\ 0 \le g_{ji} + r_{ji}^{\mathsf{U}} - r_{ji}^{\mathsf{D}} \le g_j^{\max}, \ \forall j \in G, \ \forall i \in [N]$$

$$0 \le r_{ji}^{\mathsf{U}} \le R_j^{\mathsf{U}}, \ \forall j \in G, \ \forall i \in [N]$$

$$0 \le r_{ji}^{\mathsf{D}} \le R_j^{\mathsf{D}}, \ \forall j \in G, \ \forall i \in [N]$$

$$\sum_{j\in G} (g_{ji} + r_{ji}^{\mathsf{U}} - r_{ji}^{\mathsf{D}}) = \sum_{b\in\mathcal{B}} \widehat{L}_{bi}, \ \forall i \in [N]$$

$$|\mathbf{m}_l (\sum_{j\in G_b} (g_{ji} + r_{ji}^{\mathsf{U}} - r_{ji}^{\mathsf{D}}) - \widehat{L}_{bi})| \le \mathsf{Cap}_l, \ \forall l \in \mathcal{L}, \ \forall i \in [N]$$

$$u_{ji} g_j^{\max} \le g_{ji} \le u_{(j-1)i} g_j^{\max}, \ \forall j \in G : j > 1, \ \forall i \in [N]$$

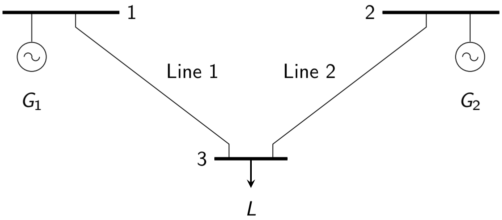
$$u_{ji} g_j^{\max} \le g_{ji} \le g_j^{\max}, \ j = 1, \ \forall i \in [N]$$

$$u_{(j-1)i} \le u_{ji}, \ \forall j \in G : j > 1, \ \forall i \in [N]$$

$$u_{ji} \in \{0, 1\}, \ \forall j \in G, \ \forall i \in [N]$$

Enforce merit-order dispatch per data point $(\widehat{f}_i, \widehat{\boldsymbol{L}}_i)$

Toy example



	c_j	c_j^{U}	c_j^{D}	$g_j^{\max}$
G_1	5	30	-20	60
G_2	15	15	10	150

- c_j, c_j^{U} and c_j^{D} in €/MWh; $g_j^{\max}$ in MW

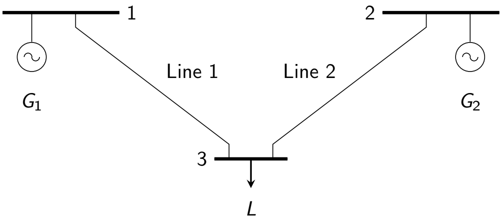
Joint distribution of (f, L)

Probability	f (MW)	L (MW)
0.5	25	U(20,30)
0.5	35	U(30,40)

- $R_j^{\mathrm{U}} = R_j^{\mathrm{D}} = g_j^{\max}$

- $Cap_1 = 30\text{MW}, \ Cap_2 = \infty$

Toy example

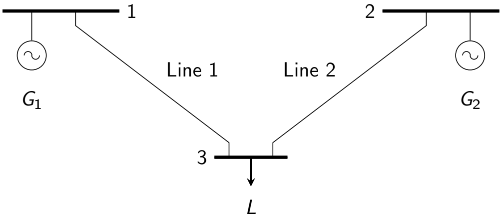


	c_j	c_j^{U}	c_j^{D}	$g_j^{\max}$
G_1	5	30	-20	60
G_2	15	15	10	150

- $Cap_1 = 30\text{MW}, \; Cap_2 = \infty$

Probability	f (MW)	L (MW)	$q_0 + q_1 f$ (MW)	Improvement(%)
0.5	25	U(20,30)	22.86	4.76
0.5	35	U(30,40)	30	33.82

Toy example



	c_j	c_j^{U}	c_j^{D}	$g_j^{\max}$
G_1	5	30	-20	60
G_2	15	15	10	150

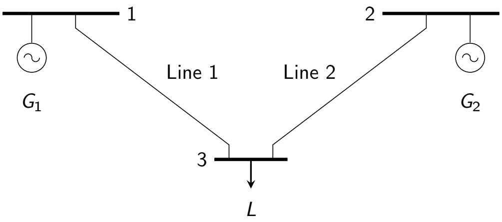
- $Cap_1 = 30\text{MW}, \; Cap_2 = \infty$

Probability	f (MW)	L (MW)	$q_0 + q_1 f$ (MW)	Improvement(%)
0.5	25	U(20,30)	22.86	4.76
0.5	35	U(30,40)	30	33.82

$$\mathbb{E}[Cost] = 5 \cdot 25 + \int_{20}^{25} 20 \frac{(25 - x)}{10} dx + \int_{25}^{30} 15 \frac{(x - 25)}{10} dx = €168.75$$

$$\mathbb{E}[Cost] = 5 \cdot 22.86 + \int_{20}^{22.86} 20 \frac{(22.86 - x)}{10} dx + \int_{22.86}^{30} 15 \frac{(x - 22.86)}{10} dx = €160.71$$

Toy example



	c_j	c_j^{U}	c_j^{D}	$g_j^{\max}$
G_1	5	30	-20	60
G_2	15	15	10	150

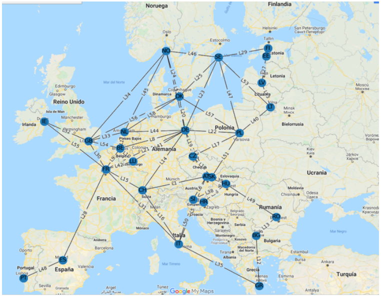
- $Cap_1 = 30\text{MW}, \; Cap_2 = \infty$

Probability	f (MW)	L (MW)	$q_0 + q_1 f$ (MW)	Improvement(%)
0.5	25	U(20,30)	22.86	4.76
0.5	35	U(30,40)	30	33.82

$$\mathbb{E}[Cost] = 5 \cdot 35 + 5 \cdot 20 + \int_{30}^{40} 15 \frac{(x-30)}{10} dx = €340$$

$$\mathbb{E}[Cost] = 5 \cdot 30 + \int_{30}^{40} 15 \frac{(x-30)}{10} dx = €225$$

A more realistic case study



- (Stylized) pipeline network model of the European electricity market
- 28 nodes, 1 per country

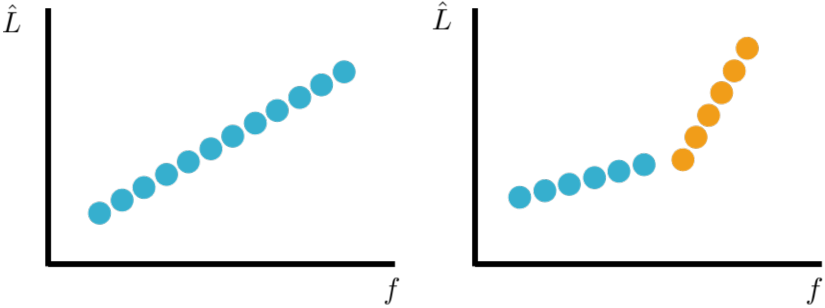
- Line and generation capacities from Nahmmacher et al. (2014)
- 1 base and 1 peak generator per node
  - **Base**: Cheap, but non-flexible (Nuclear, hard coal, oil ...)
  - **Peak**: Expensive, but flexible (Natural gas, Waste ...)
- Historical data of demand and renewables from ENTSO-e
- We derive the actual and predicted net demand of the whole system
- Training/test set 100/50 samples, 10 independent runs

Nahmmacher, P., Schmid, E., Knopf, B. (2014). Documentation of LIMES-EU – A long-term electricity system model for Europe. Potsdam Institute of Climate Impact Research (PIK)

Clustering technique to improve the *corrected* net demand forecast

We split the data set of forecast system net demands into K clusters (K-means), and estimate one affine model per cluster k:

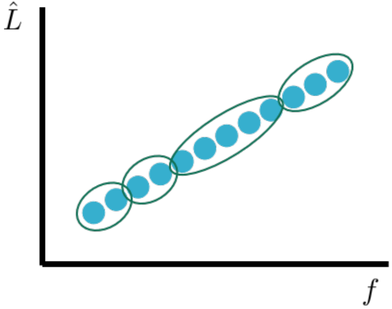
$$\widehat{L} = q_{0k} + q_{1k} f, \forall k \leq K$$

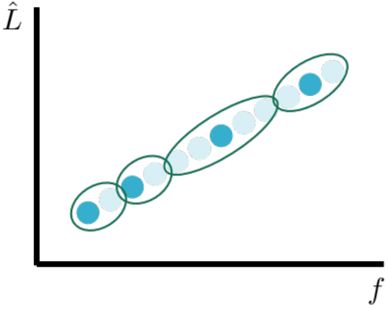


Increased flexibility!

Clustering technique to lighten the estimation of the correction rule

We only consider the most representative samples (*medoids*) from each cluster.





Reduced computational burden!

Some results

	K	1	2	5	7
	100%	2.83%	4.29%	4.74%	4.75%
r	50%	2.67%	4.23%	4.39%	4.06%
	20%	2.38%	4.12%	4.12%	3.97%

Table: Cost savings (out-of-sample). Percentage with respect to conventional approach

	K	1	2	5	7
	100%	2127.7	283.7	75.9	28.0
r	50%	180.0	27.2	7.4	5.5
	20%	8.3	3.2	1.1	1.4

Table: Computational time (s).

K: number of clusters into which the data set is divided
r: Sample reduction in each cluster

Further information in...

arXiv:2108.01003v2 [math.OC] 21 Mar 2022

Value-oriented forecasting of net demand for electricity market clearing

J. M. Morales*, M. A. Muñoz, S. Pineda

OASYS Group, Ada Byron Research Building, Arquitecto Francisco Peñalosa St., 18, 29010, University of Malaga, Malaga, Spain

Abstract

We consider a two-stage electricity market comprising a forward and a real-time settlement. The former pre-dispatches the power system following a least-cost merit order and facing an uncertain net demand, while the latter copes with the plausible deviations with respect to the forward schedule by making use of power regulation during the actual operation of the system. Standard industry practice deals with the uncertain net demand in the forward stage by replacing it with a good estimate of its conditional expectation (usually referred to as a *point forecast*), so as to minimize the need for power regulation in real time. However, it is well known that the cost structure of a power system is highly asymmetric and dependent on its operating point, with the result that minimizing the amount of power imbalances is not necessarily aligned with minimizing operating costs. In this paper, we propose a mixed-integer program to construct, from the available historical data, an alternative estimate of the net demand that accounts for the power system's cost asymmetry. Furthermore, to accommodate the strong dependence of

*Corresponding author

Email addresses: juan.morales@uma.es (J. M. Morales), miguelangeljmd@uma.es (M. A. Muñoz), spinedamorente@gmail.com (S. Pineda)

Concluding remarks

- Our approach prescribes how the *predicted* net demand should be *corrected* to minimize the expected total cost of a two-stage electricity market.
- Our simulations showcase the benefits of using several cluster techniques to provide prescriptions tailored to the operating point of the power system.
- Numerical experiments conducted on a stylized model of the European electricity market reveal that the cost savings implied by our approach are well above 2%.

Thanks for the attention!! Questions??

Juan Miguel Morales (juan.morales@uma.es)

OASYS
Optimization and Analytics for
Sustainable energY Systems

More info: oasys.uma.es

erc
European Research Council
Established by the European Commission

Supported by the project FlexAnalytics - Advanced Analytics to Empower the Small Flexible Consumers of Electricity.