PRICE STABILIZATION OF TWO-STAGE STOCHASTIC PROGRAMS WITH APPLICATION TO ENERGY GENERATION PROBLEMS

Clara Macedo Lage - ENGIE - Panthéon Sorbonne University - IMPA - (CIFRE)
Guillaume Erbs - ENGIE
Claudia Sagastizábal - Unicamp
Mikhail Solodov - IMPA

January 2020 - Paris

MOTIVATION - ENERGY PRICES

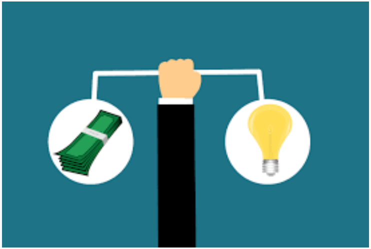
Modern energy markets involve a large number of units and different technologies to generate electricity.

MOTIVATION - ENERGY PRICES

In Brazil and Northern Europe, hydraulic generation is one of the main sources of energy. This technology has the following important characteristics:

- It is a renewable energy
- The low cost of hydro-energy generation if compared to others sources.
- It allows the system to store energy in the form of water in the reservoir.
- The difficulty in predicting the amount of rain or snow at any time scale makes the water inflows uncertain.

The price signal represents the opportunity cost, that is, consider the possibility of shortage of energy and the cost of other supply sources in future periods.



For example, if it rains less than expected, it can be necessary to activate different and more expensive power plants. This extra-cost is an important component in the price of hydro power plants.

MOTIVATION - ENERGY PRICES

The randomness that comes from the diverse inflow scenarios makes us consider many different possibilities in a future cost function.

We consider scenarios with large and small amount of inflow and their consequences for the system.

In long term planning problems, decisions are coupled in time. An example of the link between t and $t+1$ is the water balance equation.

Denoting the generation of the i-th unit by x_i, for one realization ξ of the uncertain Inflow, the 2-stage formulation for the energy generation problem is:

$$\begin{aligned} \min \quad & \langle \mathrm{Cost}_1, x_1 \rangle + \langle \mathrm{Cost}_2, x_2 \rangle \\ \text{s.t.} \quad & x_i \geq 0, i = 1, 2 \\ & Bx_i \leq b_i, i = 1, 2 \\ & Tx_1 + Wx_2 = \mathrm{Inflow}(\xi) \ \rightsquigarrow \ \bar{\pi}(\xi) \end{aligned}$$

The link between stages is represented by the matrices W and T.

The sub-vectors x_1 and x_2 represent, respectively, the parameters in the generation of the set of power plants, at time steps 1 and 2.

Variables x_2 are recourse variables that depend on the realization ξ.

Mathematically, the opportunity cost corresponds to the Lagrange Multiplier.

$$\begin{aligned} \min \quad & \langle \mathrm{Cost}_1, x_1 \rangle + \langle \mathrm{Cost}_2, x_2 \rangle \\ \text{s.t.} \quad & x_i \geq 0, i = 1, 2 \\ & Bx_i \leq b_i, i = 1, 2 \\ & Tx_1 + Wx_2 = \mathrm{Inflow}(\xi) \ \rightsquigarrow \ \bar{\pi}(\xi) \end{aligned}$$

Denoting $x = (x_1, x_2)$, the Lagrangian function is:

$$L(x, \pi, \mu_1, \mu_2) :=$$

$$\langle (\mathrm{Cost}_1, \mathrm{Cost}_2), x \rangle + \langle (B, B)x - b, \mu_1 \rangle + \langle -Ix, \mu_2 \rangle + \langle (T, W)x - \mathrm{inflow}(\xi), \pi(\xi) \rangle$$

if $L'(x, \bar{\pi}, \bar{\mu}_1, \bar{\mu}_2) = 0$, we call $(\bar{\pi}, \bar{\mu})$ Lagrange Multipliers.

$$-(\mathrm{Cost}_1, \mathrm{Cost}_2) = (B, B)^\top \bar{\mu}_1 - I\bar{\mu}_2 + (T, W)^\top \bar{\pi}(\xi)$$

The general multistage stochastic problem is:

$$\min_{A_1x_1=\xi_1,\ x_1\geq 0} c_1x_1+\mathbb{E}\Big[\min_{B_2x_1+A_2x_2=\xi_2,\ x_2\geq 0} c_2x_2+\mathbb{E}\Big[\ldots+\mathbb{E}\big[\min_{B_Tx_{T-1}+A_Tx_T=\xi_T,\ {}_T\geq 0}\big]\Big]\Big]$$

- x_t is called the decision variable.
- A_t and B_t are matrices.
- In the case of energy generation x_t is composed essentially by the level of the reservoirs of each hydro power plant, the generation of each power plant and the flow between them.

Difficulty 1 The set of Lagrange Multipliers $\{\pi\}$ is not commonly singleton. The price signal is one element in this set that depends on the way we model and the algorithm used to solve the problem. Is this price signal the best one for our application? What about the position of this price signal in this set?

Difficulty 2 Taking uncertainty into account means that the price will be a random vector $(\pi(\xi^1), ..., \pi(\xi^S))$, where $\xi^s \in \Omega, s \in \{1, ..., S\}$ are the scenarios. The distribution of the price signal depends on the scenarios we consider and on the probability P in this scenario set.

In a simulation, using a two stage model and $S = 80$ scenarios, we can see the difference of price signal distribution for two different samples P_1 and P_2 in $\Omega = \{\xi^1, ..., \xi^{80}\}$.

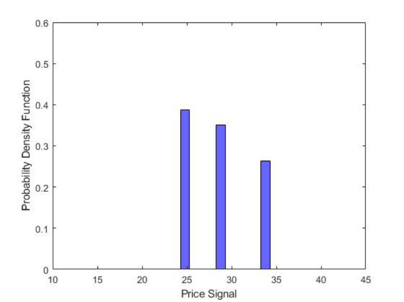


Figure: Price signal distribution for data distributed as P_1

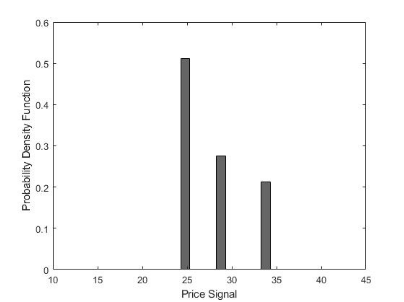


Figure: Price signal distribution for data distributed as P_2

How regularization can help us?

MATHEMATICAL MODEL

In the two stage model, for each realization ξ of the uncertainty, the price is given by the Lagrange Multiplier of the corresponding second stage problem.

First Stage Problem:

$$\begin{aligned} \min \quad & \langle \mathrm{Cost}_1, x_1 \rangle + E[Q(x_1, \xi)] \\ \text{s.t.} \quad & x_1 \geq 0 \\ & B_i x_i \leq b_i \end{aligned}$$

Second Stage Problem, fixed ξ_i:

$$Q(x_1, \xi_i) := \begin{cases} \min & \langle \mathrm{Cost}_2, x_2 \rangle \\ \text{s.t.} & W x_2 = \mathrm{inflow}(\xi_i) - T x_1 \\ & B_2 x_2 \leq b_2 \\ & x_2 \geq 0 \end{cases}$$

Second Stage Problem:

$$\begin{cases} \min & \langle \mathrm{Cost}_2, x_2 \rangle \\ \text{s.t.} & Wx_2 = \mathrm{inflow}(\xi_i) - Tx_1 \\ & B_2 x_2 \leq b_2 \\ & x_2 \geq 0 \end{cases}$$

Correspondent Lagrangian function:

$$L(x_2, \pi, \mu_1, \mu_2) :=$$

$$\langle \mathrm{Cost}_2, x_2 \rangle + \langle B_2 x_2 - b_2, \mu_1 \rangle - \langle I x_2, \mu_2 \rangle + \langle W x_2 - \mathrm{inflow}(\xi) - T x_1, \pi(\xi) \rangle$$

The problem can be rewritten as:

$$\min_{x\in\mathbb{R}^n}\left\{\sup_{(\pi,\mu)\in\mathbb{R}^m\times\mathbb{R}^k_+} L(x,\pi,\mu)\right\}$$

By definition the dual problem is:

$$\max_{(\pi,\mu)\in\mathbb{R}^m\times\mathbb{R}^k_+}\left\{\inf_{x\in\mathbb{R}^n} L(x,\pi,\mu)\right\}$$

Subject to:

$$\Delta = \left\{(\pi,\mu)\in\mathbb{R}^m\times\mathbb{R}^k : \inf_{x\in\mathbb{R}^n} L(x,\pi,\mu) > -\infty\right\}$$

.

For important class of optimization problems the dual problem can be rewritten as a classical optimization problem.

The second stage, also known as future cost function:

$$
Q(x_1, \xi) := \begin{cases} \min & \langle Cost_2, x_2 \rangle \\ \text{s.t.} & Wx_2 = inflow(\xi_i) - Tx_1 \\ & B_2 x_2 \leq b_2 \\ & x_2 \geq 0 \end{cases}
$$

has as dual:

$$
Q(x_1, \xi) := \begin{cases} \max & \langle \pi, inflow(\xi_i) - Tx_1 \rangle - \left\langle b_2, \pi^B \right\rangle \\ \text{s.t.} & W^\intercal \pi - B_2^\intercal \pi^B \leq Cost_2 \end{cases}
$$

Where π^B is the Lagrange Multiplier of the inequality constraint.

Consider the second stage problem.

Given a constant $\beta > 0$, the regularized second stage problem is:

$$Q^{\beta}(x_1, \xi) := \begin{cases} \max & \langle \pi, \text{inflow}(\xi) - Tx_1 \rangle - \left\langle b_2, \pi^B \right\rangle - \frac{\beta}{2}\|\pi\|^2 \\ \text{s.t.} & W^T\pi - B_2^T\pi^B \leq \text{Cost}_2 \end{cases}$$

or, computing its dual:

$$Q^{\beta}(x_1, \xi) = \begin{cases} \min & \langle \text{Cost}_2, x_2 \rangle + \frac{1}{2\beta}\|\text{inflow}(\xi) - Tx_1 - Wx_2\|^2 \\ \text{s.t.} & x_2 \geq 0 \\ & B_2 x_2 \leq b_2 \end{cases}$$

The Lagrange multipliers π can also be viewed as elements of the sub-gradient of Q and Q^β :

Defining:

$$\psi(x, \pi, \pi^B, \xi) = \langle \text{inflow}_\xi - Tx_1, \pi \rangle - \langle b_2, \pi^\beta \rangle - \beta \|\pi\|^2,$$

We have that ψ is convex, and:

$$Q^\beta(x_1, \xi) = \max_{W^T\pi - B_2^T \pi^B \leq \text{Cost}_2} \psi(x_1, \pi, \pi^B, \xi)$$

So by convex analysis theory:

$$\partial Q^\beta(x_1, \xi) = \text{conv}\{\psi'_{x_1}(x_1, \bar{\pi}, \lambda, \bar{\pi}^B, \xi) \mid \bar{\pi} \in \Pi(x_1)\} = \{-T\bar{\pi} \mid \bar{\pi} \in \Pi(x_1)\},$$

where $\Pi(x_1)$ is the set of optimal values of Q^β.

For $\Omega = \{\xi^1, ..., \xi^S\}$, with probabilities $\{p^1, ..., p^S\}$, the one level formulation of the regularized problem has the form:

$$\begin{cases} \min & \langle \mathrm{Cost}_1, x_1 \rangle + \sum_{s=1}^{S} p^s \left\{ \langle \mathrm{Cost}_2, x_2^s \rangle + \frac{1}{2\beta} \| \mathrm{inflow}(\xi^s) - Tx_1 - Wx_2^s \|^2 \right\} \\ \text{s.t.} & x_1 \geq 0, Bx_1 = b_1, x_2^s \geq 0, Bx_2 = b_2 \ \text{a.e } s = 1, ..., S. \end{cases}$$

THEORETICAL RESULTS

Difficulty 1 Often the set of Lagrange Multipliers $\{\pi\}$ is not a singleton. The price signal is a choice in this set that depends on the way we model and the algorithm used to solve the problem.
Is this price signal the best one to our application?
What about the position of this price signal in this set?

One level formultion:

$$\begin{cases} \textbf{min} & \langle \mathrm{Cost}_1, x_1 \rangle + \langle \mathrm{Cost}_2, x_2 \rangle + \frac{1}{2\beta} \| \mathrm{inflow}(\xi) - T x_1 - W x_2 \|^2 \\ \text{s.t.} & x_1 \geq 0 \,, x_2 \geq 0 \,, B_1 x_1 \leq b_1 \,, B_2 x_2 \leq b_2 \,. \end{cases}$$

From KKT equations, given a primal solution $\bar{x_1}^\beta$, $\bar{x_2}^\beta$ the regularized price signal will be:

$$\bar{\pi}^\beta = \frac{1}{\beta}(\mathrm{inflow}(\xi) - T\bar{x_1}^\beta - W\bar{x_2}^\beta).$$

Theorem If the one level original formulation has a unique solution $\bar{x} = (\bar{x}_1, \bar{x}_2)$, and the sequence $\beta_k \to 0$ is decreasing, then the solution $(x_1^k, x_2^k) = x^k \to \bar{x}$.

Our main interest is in the price π. Keeping ξ fixed, we know that:

$$\pi^k = \frac{1}{\beta_k}(\text{inflow}(\xi) - Tx_1^k - Wx_2^k).$$

The questions that arise naturally are:

1. Is π^k bounded?
2. Does π^k converge?

Theorem Let β_k be monotonically decreasing. Suppose that the original problem has a unique solution $\bar{x}$. Denote π^k the sequence of optimal regularized Lagrange multipliers. Under reasonble assuptions, there is a subsequence π^{k_j} of π^k that converges to $\hat{\pi}$, the minimum-norm optimal Lagrange multiplier.

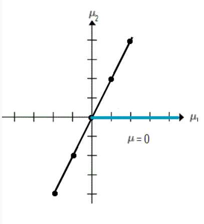
To explain the condition in the theorem, we remember that for one scenario we have $F = (\mathrm{Cost}_1, \mathrm{Cost}_2)$, $A = \begin{bmatrix} T \\ W \end{bmatrix}$. The regularized problem is:

$$\begin{cases} \min & F(x) + \frac{1}{2\beta}\|Ax - \xi\|^2 \\ \text{s.t.} & x \geq 0 \,. \end{cases}$$

The **necessary and sufficient condition** for boundedness of π^k is the following:

$$\mathrm{Im}\, A^T \cap \{\mu \in \mathbb{R}^n_+ \ : \ \mu_i = 0 \ \text{ if } \bar{x}_i > 0\} = \{0\}.$$

The condition is likely to be satisfied since there is no particular reason why some given lower-dimensional subspace intersects the axis $x_i = 0$, except at 0.



Dificulty 2 Taking uncertainty into account means that the price will be a random vector $(\pi(\xi^1), ..., \pi(\xi^S))$. The distribution of the price signal depends on the scenarios we consider and the probability P in this scenario set.

Let $\Omega = \{\xi^1, ..., \xi^S\}$, be the set of scenarios.

$$\mathbf{P} = \{P = (p^1, p^2, ..., p^S) : \sum_1^S p^i = 1\}$$

a perturbation of P is another probability

$$P_U = (p^1 + u^1, p^2 + u^2, ..., p^S + u^s).$$

The set of perturbations for probability P is:

$$\mathbf{U}_P = \{U = (u^1, u^2, ..., u^S), \sum u^s = 0, 0 \leq (p+u)^s \leq 1\}$$

and:

$$f_P^\beta(x_1, U) = \langle Cost_1, x_1 \rangle + \mathbb{E}_{P_U}\left[\mathbf{Q}^\beta(x_1, \xi)\right]$$

Define:

$$S_P^\beta(U) = \{\text{argmin}_{x_1 \in X} f_P^\beta(x_1, U)\},$$

Assume that $S_P^0(0)$ is a singleton and denote: $\bar{x}_1 = S_P^0(0)$.

$$X_1^{\beta,P}(u^1, u^2, ..., u^S) \mapsto \bar{x}_1^U = \text{argmin}\{\|x_1^U - \bar{x}_1\| \ : \ x_1^U \in S_P^\beta(U)\}.$$

We aim at understanding the properties of the function $X_1^{\beta,P}$.

Note that the price signal $\pi(X_1^{P,\beta})$ can be viewed as a function of $X_1^{P,\beta}$.

Theorems:

1. **Theorem 1** $S_P^{\beta}(U)$ is singleton.
2. **Theorem 2** $X_1^{P,\beta}$ is Lipschitz in $\mathbf{U}_P$ that is, there are $L_{X_1}, L_\pi > 0$, such that:
$$\|X_1^{U_1} - X_1^{U_2}\| \leq L_{X_1}\|U_1 - U_2\|$$
and
$$\|\pi^{U_1} - \pi^{U_2}\| \leq L_\pi\|U_1 - U_2\|$$
3. **Theorem 3** We can control $\mathrm{Var}(\pi(U))$ when $U \in \mathbf{U}_P$.

The constants L_{X_1}, L_π are proportional to $\frac{1}{\beta}$.

EXAMPLES

A simple example can illustrate the first theorem.

The first-stage cost $c \in \mathbb{R}^2$ while second-stage costs are deterministic $q^1 = q^2 = q \in \mathbb{R}^2$. The technology and recourse matrices are:

Non Regularized Problem:

$$\begin{cases} \min \ cx_1 + qx_2 \\ \text{s.t. } x_1 \geq 0 \\ Tx_2 + Wx_2^i = \xi^i, \ i \in \{1,2\}. \end{cases}, \quad T := \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad W := \begin{bmatrix} 2 & 0 \\ 1 & -1 \\ 1 & 2 \end{bmatrix},$$

The uncertain right-hand side terms are:

$$\xi^1 := (1,1,1) \text{ and } \xi^2 := (1,0,3).$$

We can rearrange terms to isolate first and second stage variables for both scenarios:

x is feasible if and only if, for some $y \geq 0$

$$x_1 := \left(y, \frac{3}{4}(1+y)\right)^T,\ x_2^1 := \left(\frac{1}{2}(1-y), \frac{1}{4}(1+y)\right)^T,\ x_2^2 := \left(\frac{1}{2}(1-y), \frac{1}{4}(5+y)\right)^T.$$

We arrive in the equivalent one level problem:

$$\min_{y \geq 0} \left(c + \frac{3}{4}c - \frac{1}{2}q_1 + \frac{1}{4}q_2\right)y + \frac{3}{4}c_2 + \frac{1}{2}q_1 + \frac{3}{4}q_2\,,$$

whose optimal solution is $\bar{y} = 0$, as long as

$$c \geq -\frac{3}{4}c + \frac{1}{2}q_1 - \frac{1}{4}q_2\,.$$

We can compute the primal solutions:

$$\bar{x}_1 := (0, \tfrac{3}{4})^T, \quad \bar{x}_2^1 := (\tfrac{1}{2}, \tfrac{1}{4})^T, \quad \bar{x}_2^2 := (\tfrac{1}{2}, \tfrac{5}{4})^T.$$

Remind the condition:

$$\text{Im } A^T \cap \{\mu \in \mathbb{R}^n_+ \ : \ \mu_i = 0 \ \text{ if } \bar{x}_i > 0\} = \{0\}.$$

If $\langle \bar{\mu}, \bar{x} \rangle = 0$. Therefore:

$$\bar{\mu} = \alpha e_1 \quad \text{ for some } \alpha \geq 0,$$

Suppose $\bar{\mu} = \alpha e_1 \in \text{Im}A^T$. Since $\text{Im}A^T$ and KerA are orthogonal:

$$\nu \in \text{Ker}A \Rightarrow \alpha \langle e_1, \nu \rangle = 0$$

Since KerA is a (unidimensional) subspace is generated by the vector $s := \Big(4, 3, -2, 1, -2, 1\Big)$, this means that :

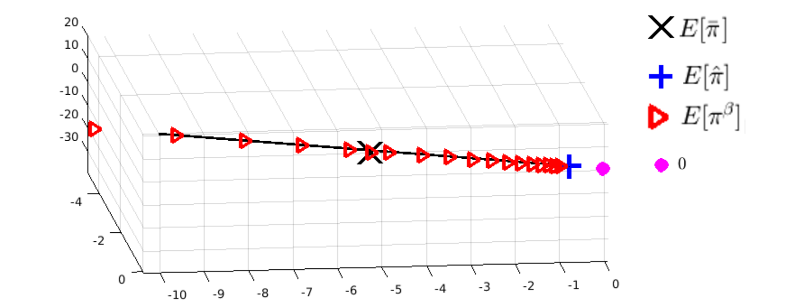
$$4\alpha = 0$$

which forces $\alpha = 0$, Showing that we satisfy the condition.

Table of primal and dual expected values for different values of β :

β	Primal Variable (First Stage)	Norm of Expected Price Signal
0	3.5	25.51
0.1	3.45	19.4
0.5	2.91	14.36
1	2.65	12.25

Graphically:



Consider the two stage stochastic problem in $\mathbb{R}$:

$$\begin{cases} \min & cx_1 + \mathbb{E}[Q(x_1,\xi)] \\ \text{s.t.} & x_1 \geq 0 \end{cases} \quad , \quad Q(x_1,\xi) := \begin{cases} \min & q^+y^+ + q^-y^- \\ \text{s.t.} & y^+ - y^- = x_1 - \xi \\ & y^+, y^- \geq 0 \end{cases}$$

So:

$$Q(x_1,\xi) = q^+(x_1-\xi)_+ + q^-(\xi - x_1)_+$$

$$\pi(\xi) = \begin{cases} q^+, & \text{if } x_1 > \xi \\ -q^-, & \text{if } x_1 < \xi. \end{cases}$$

If the distribution of ξ is symmetric and $q^+ = q^- = q$, we have that: $\mathbb{E}[\bar{\pi}] = 0$, and $\mathrm{Var}[\bar{\pi}] = \mathbb{E}[\bar{\pi}^2] = q^2$.

In the regularized case:

$$\begin{cases} \min & cx_1 + E[Q^\beta(x_1,\xi)] \\ \text{s.t.} & x_1 \geq 0 \end{cases},$$

$$Q^\beta(x_1,\xi) := \begin{cases} \min & q^+y^+ + q^-y^- + \frac{1}{2\beta}|y^+ - y^- - \xi + x_1|^2 \\ & y^+, y^- \geq 0, \end{cases}$$

And:

$$\pi^\beta(\xi) = \begin{cases} q^+, & \text{if } \xi \leq 2\beta q^+ + x_1 \\ -q^-, & \text{if } \xi \geq -2q^-\beta + x_1 \\ \frac{\xi - x_1}{2\beta}, & \text{if } 2q^-\beta + x_1 < \xi < 2\beta q^+ + x_1. \end{cases}$$

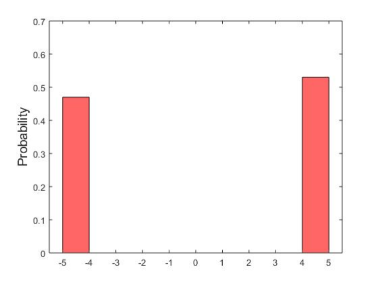
It is possible to estimate analytically that: $\mathrm{Var}[\pi^\beta] \leq \mathrm{Var}[\pi]$.

EXAMPLES

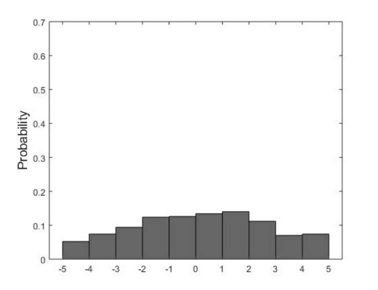
We simulate with:

- $\xi \sim N(0, 10)$
- $q = 5, \; c = 5$
- $\Xi = \{\xi^1, ..., \xi^N\}$ is a sample, and $N = 200$
- $\beta = 1$

Non Regularized Price Signal

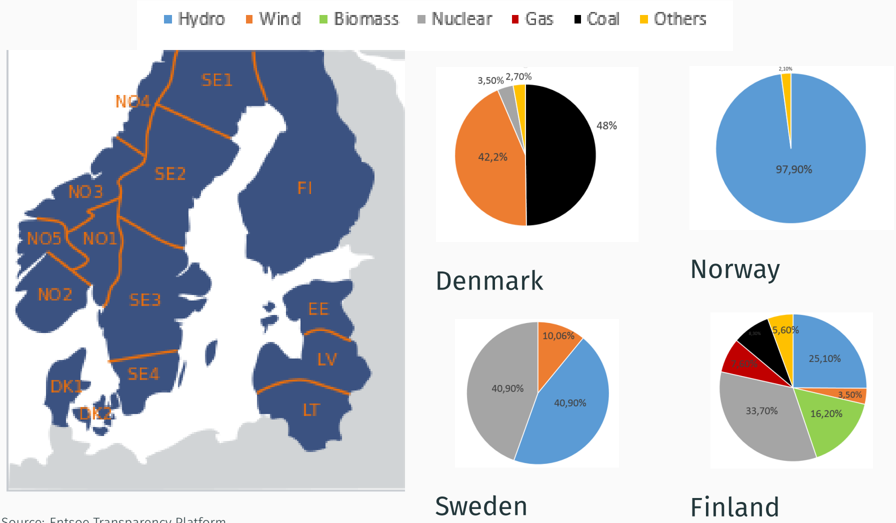


Regularized Price Signal



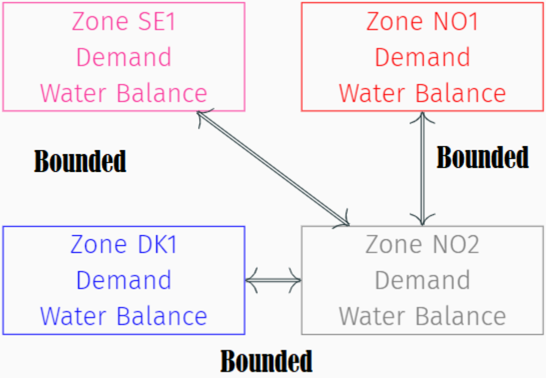
APPLICATION TO THE GENERATION PROBLEM

The Northern European generation system is composed by hydro, wind, thermal and solar power plants.



Source: Entsoe Transparency Platform

Zones are connected, but the flow between them are bounded. The price of one zone can deppend of the power plants of other zones



APPLICATION TO THE GENERATION PROBLEM

We modeled the generation of energy in the Northern Europe. Our model has the following characteristics:

- Multistage Model
- Uses real data from the Northern European energy system
- $t \in \{1, 2, ..., 365\}$, measures time in days, and $w \in \{1, .., 52\}$ weeks
- The model is deterministic for days inside each week and considers randomness for the first time of each week
- Inflow scenarios are generated from the historical mean and standard deviation, using a log-normal distribution
- Deterministic demand that varies with time
- The decision variable includes reservoir levels for hydro power plant, generation for each power plant, and spillage.

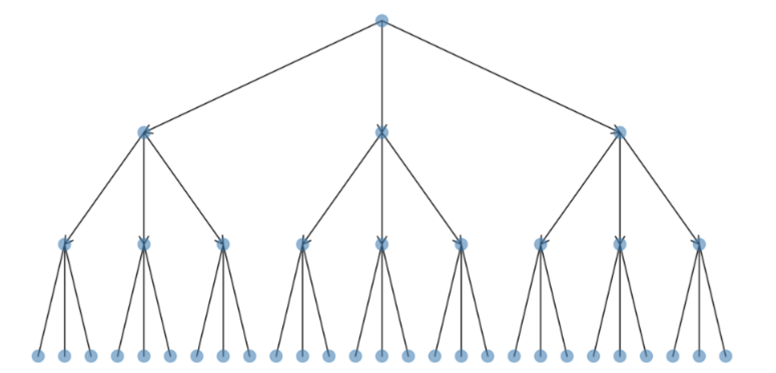
We use the rolling horizon algorithm keeping us in a two stage model

APPLICATION TO THE GENERATION PROBLEM

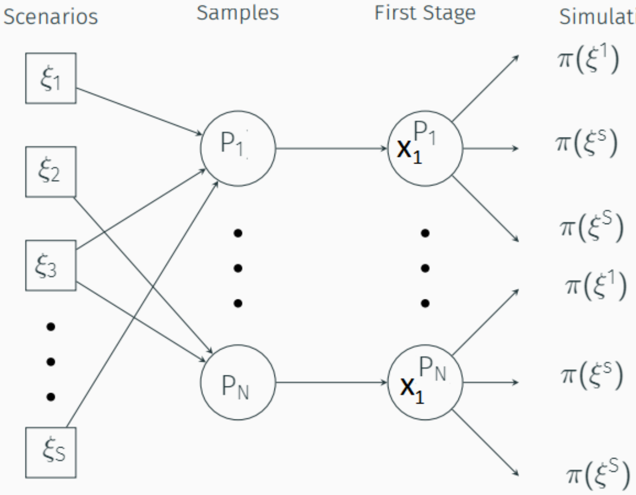
The general multistage stochastic problem is:

$$\min_{A_1 x_1 = \xi_1,\ x_1 \geq 0} c_1 x_1 + \mathbb{E}\Big[\min_{B_2 x_1 + A_2 x_2 = \xi_2,\ x_2 \geq 0} c_2 x_2 + \mathbb{E}\Big[\ldots + \mathbb{E}\Big[\min_{B_T x_{T-1} + A_T x_T = \xi_T,\ {}_T \geq 0} \Big]\Big]\Big]$$

Tree of scenarios:

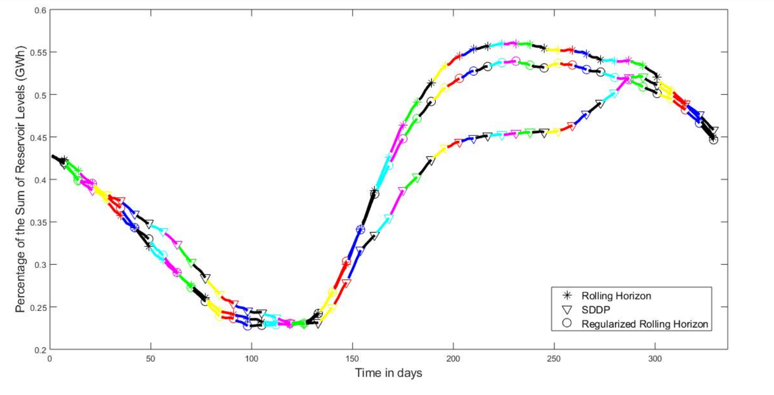


Model of the simulation test:



Data: 200 scenarios, 20 samples, 30 scenarios for optimization.

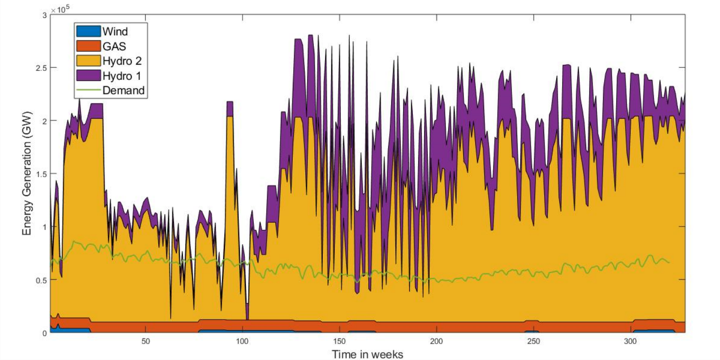
Reservoir management for SDDP, non regularized and regularized rolling horizon ($\beta = 30$) problems.



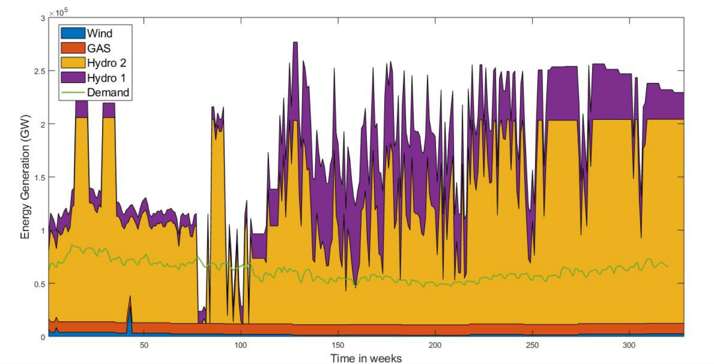
Since $Q(x_1, \xi^s) \geq Q^\beta(x_1, \xi^s), \quad \forall s \in \{1, ..., S\}$, we expect regularized decisions to be less conservatives.

Consequently, in the end of the period, we have more water and also constant generation levels (in the maximal generation level).

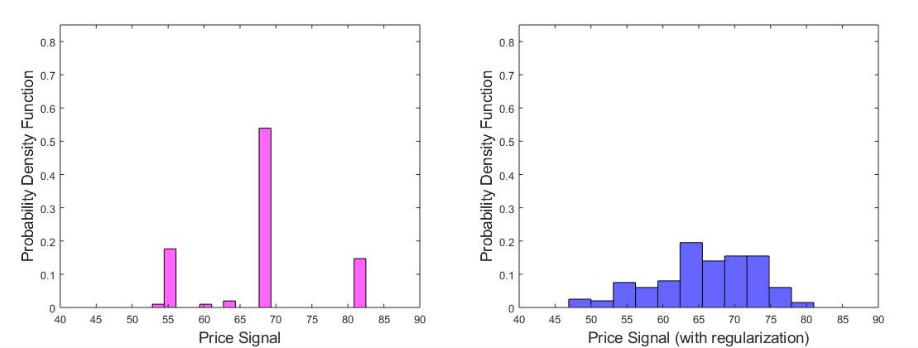
Non Regularized Generation Level for zone NO2 :



Regularized Generation Level for zone NO2 :

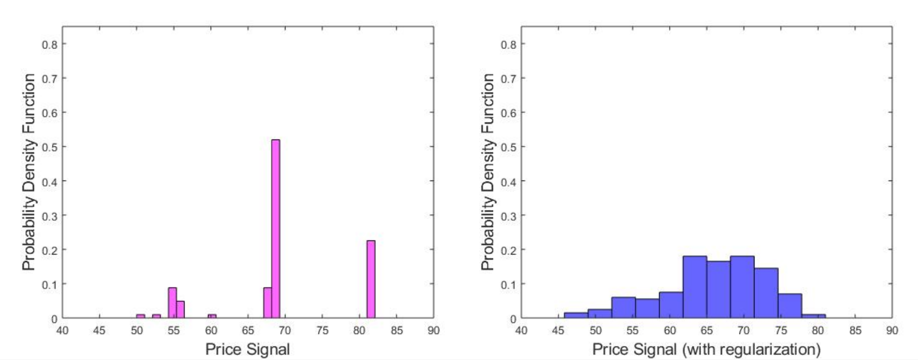


Histogram of the non regularized and regularized ($\beta = 30$) price signal for zone NO2, sample 1:



$E[\pi] = 64.8$ $E[\pi] = 64.2$

Histogram of the non regularized and regularized ($\beta = 30$) price signal for zone NO2, t = 169, sample 2:



$\mathsf{E}[\pi] = 65.6$ $\mathsf{E}[\pi] = 65.1$

Comparison of Wasserstein distance and the variance of expected value of histograms for 20 samples. 30 scenarios for optimization and 200 for simulation. Zone N02, $t = 169$:

	Mean	Variance of Mean	Wasserstein Distance
Regularized ($\beta = 30$)	64.31	0.5	33.20
Non-Regularized	64.56	8.32	146.61

Conclusion: Regularization helps to stabilize the price signal in respect to the distribution of inflows.

Thanks for your attention!