Robust Optimization in Energy and Environmental Problems

J.-P. Vial[3,4]

[3] Prof. Emeritus, University of Geneva
[4] Ordecsys, scientific consulting, Geneva, Switzerland

Optimization Days, May 2 - 4, 2011

ORDECSYS

Uncertainty in environment and energy modeling

Models for environmental and energy systems assessments are plagued with uncertainty.

- They are usually formulated as mid-term to long-term planning problems, crucially influenced by uncertain factors like climate change, economy growth, technological progress
- The increasing influence of markets makes prices volatile.
- Many parameters in the models are estimated, at best by statistical analysis, often by expert judgments.

Models for environmental and energy systems assessments are usually large, and can be solved only by the more advanced optimization technology. Standard tools like Stochastic Programming or Chance-Constrained Programming render the model numerically intractable.

Sources of Uncertainty

Many data in the large deterministic models for energy and environmental analyses are imperfectly known or subject to random fluctuations

- **Estimated parameters** Some parameters are estimated by statistical analysis.
  $\Longrightarrow$ Estimation errors
- **Parameters only revealed at posteriori** The true value of some parameters (e.g., demand, price, etc.) becomes known after the problem has been treated $\Longrightarrow$ Forecast errors
- **Ill-defined parameters** Some important parameters cannot be estimated with precision, neither a priori nor at posteriori (e.g., holding and shortage inventory costs) $\Longrightarrow$ Errors in subjective assessment
- **Implementation** If x is the “optimal” solution proposed the model, there is no guarantee that it will be implemented with full precision.
  $\Longrightarrow$ Implementation errors

Contents

1. Robust Optimization illustrated on a simple example
   - Deterministic formulation
   - Impact of uncertainty in pollutant diffusion
   - The robust optimization approach
2. Dynamic problem on the illustrative example
3. RO vs Chance-Constrained Programming
4. Energy Security in EU supply via TIAM
5. Concluding remarks

A simple environmental and energy planning model

- Model inspired by MARKAL-Geneva
- Simulate the evolution of the energy system under the assumption of an optimal use of resources (marginal cost efficiency) to maximize welfare.
- Local emissions of pollutants with diffusion in neighboring regions: cap on pollutant levels.
- Dynamic model with three periods.

F. BABONNEAU, J.-P. VIAL, AND R. APPARIGLIATO. Robust optimization for environmental and energy planning. In J. Filar and A. Haurie, editors, Handbook on *Uncertainty and Environmental Decision Making*, pages 79–126. Springer Verlag, 2010.

The energy system

Find the least cost energy production system

Production system must be expanded and operated so as to meet demand for energy at each period.

- Additional production capacity can be acquired (investment).
- The production units emit pollutant at levels which depend on the production technology
- Pollutant emitted in one region diffuses to neighboring regions

- Four regions $\mathcal{R} = \{r_1, r_2, r_3, r_4\}$
- Three production technologies $\mathcal{P} = \{p_1, p_2, p_3\}$
- Three time periods $\mathcal{T} = \{t_1, t_2, t_3\}$

The energy system

Cost	r_1	r_2	r_3	r_4
Maintenance (M_r)	1	1	1	1
Investment (I_r)	5	3	4	5
Operation (O_r)	2	2	3	2

Table: Maintenance, investment and operation costs for the three technologies.

	Technology p_1			Technology p_2			Technology p_3		
	t_1	t_2	t_3	t_1	t_2	t_3	t_1	t_2	t_3
r_1	1	1	-	0.5	-	-	-	-	-
r_2	-	-	-	-	-	-	-	-	-
r_3	-	-	-	-	-	-	-	-	-
r_4	-	-	-	-	-	-	1	1	-

Table: Residual capacities ($res_{p,t,r}$).

Emission and diffusion of pollutants

The technologies have different pollutant emission rates, given in Table 3, in each region.

	r_1	r_2	r_3	r_4
p_1	0.7	0.8	0.8	0.6
p_2	0.5	0.4	0.7	0.7
p_3	0.8	0.9	0.6	0.7

Table: Emission rates ($E_{p,r}$) by production technology and by region.

Emission and diffusion of pollutants

A pollution transport and dispersion process takes place. A source-receptor matrix specifies, for each source location, the proportion of the emitted pollutant that is deposited in the different receptor locations (see Table 4).

	r_1	r_2	r_3	r_4
r_1	0.5	0.1	0.1	0.05
r_2	0.1	0.4	0.04	0.1
r_3	0.09	0.05	0.5	0.1
r_4	0.05	0.1	0.1	0.6

Table: Source-receptor transfer matrix (G_{r_i,r_j}).

An environmental quality, $Q = 1.5$, is imposed on pollutant concentration for all regions and for all periods.

Production and demand constraints

$$x_{p,t,r} \leq z_{p,t,r} \quad \forall t \in \mathcal{T}, \forall p \in \mathcal{P}, \forall r \in \mathcal{R}$$

production $\leq$ installed capacity

$$z_{p,t,r} = res_{p,t,r} + \sum_{\tau \leq t} y_{p,\tau,r} \quad \forall t \in \mathcal{T}, \forall p \in \mathcal{P}, \forall r \in \mathcal{R}$$

installed capacity $=$ residual capacity $+$ investment

$$\sum_{p,r} x_{p,t,r} = d_t \quad \forall t \in \mathcal{T}$$

total production $=$ demand in each period

Pollutant emission and objective

$$\sum_p E_{p,r} x_{p,t,r} = e_{t,r} \quad \forall t \in \mathcal{T}, \forall r \in \mathcal{R}$$

unit emission $\times$ production $=$ local emission in region $\mathcal{R}$

$$\sum_{\rho \in \mathcal{R}} e_{t,\rho} G_{\rho,r} \leq Q \quad \forall t \in \mathcal{T}, \forall r \in \mathcal{R}$$

local emission $\times$ diffusion $=$ pollutant level $\leq$ cap on air quality

$$\min_{x \geq 0, y \geq 0, z, e} \left\{ \sum_r \left(O_r \sum_{p,t} x_{p,t,r} + I_r \sum_{p,t} y_{p,t,r} + M_r \sum_{p,t} z_{p,t,r} \right) \right\}$$

operating costs $+$ investment costs $+$ maintenance cost $\leq$ cap on air quality

Deterministic LP model

$$
\begin{aligned}
\min_{x\geq 0, y\geq 0, z, e} \quad & \sum_r \Big(O_r \sum_{p,t} x_{p,t,r} + I_r \sum_{p,t} y_{p,t,r} + M_r \sum_{p,t} z_{p,t,r}\Big) \\
& z_{p,t,r} = res_{p,t,r} + \sum_{\tau\leq t} y_{p,\tau,r}, \quad \forall t\in\mathcal{T}, \forall p\in\mathcal{P}, \forall r\in\mathcal{R} \\
& x_{p,t,r} \leq z_{p,t,r}, \quad \forall t\in\mathcal{T}, \forall p\in\mathcal{P}, \forall r\in\mathcal{R} \\
& \sum_{p,r} x_{p,t,r} = d_t, \quad \forall t\in\mathcal{T} \\
& \sum_{p} E_{p,r} x_{p,t,r} = e_{t,r}, \quad \forall t\in\mathcal{T}, \forall r\in\mathcal{R} \\
& \sum_{\rho\in\mathcal{R}} e_{t,\rho} G_{\rho,r} \leq Q, \quad \forall t\in\mathcal{T}, \forall r\in\mathcal{R}.
\end{aligned}
$$

The toy model has 108 variables and 99 constraints.

Optimal investment schedule (deterministic solution)

	t_1			t_2			t_3			
	p_1	p_2	p_3	p_1	p_2	p_3	p_1	p_2	p_3	Total
r_1	-	1.41	-	-	0.77	-	-	-	-	2.18
r_2	-	6.09	-	-	1.73	-	-	0.15	-	7.97
r_3	-	-	-	-	-	-	-	-	3.85	3.85
r_4	-	-	-	-	-	-	-	-	-	0
Sum	-	7.50	-	-	2.50	-	-	0.15	3.85	

Table: Optimal investment schedule.

Uncertainty

$$\min_{x\geq 0, y\geq 0, z, e} \quad \sum_r \left(O_r \sum_{p,t} x_{p,t,r} + I_r \sum_{p,t} y_{p,t,r} + M_r \sum_{p,t} z_{p,t,r}\right)$$

$$z_{p,t,r} = res_{p,t,r} + \sum_{\tau \leq t} y_{p,\tau,r}, \quad \forall t \in \mathcal{T}, \forall p \in \mathcal{P}, \forall r \in \mathcal{R}$$

$$x_{p,t,r} \leq z_{p,t,r}, \quad \forall t \in \mathcal{T}, \forall p \in \mathcal{P}, \forall r \in \mathcal{R}$$

$$\sum_{p,r} x_{p,t,r} = d_t, \quad \forall t \in \mathcal{T}$$

$$\sum_p E_{p,r} x_{p,t,r} = e_{t,r}, \quad \forall t \in \mathcal{T}, \forall r \in \mathcal{R}$$

$$\sum_{\rho \in \mathcal{R}} e_{t,\rho} G_{\rho,r} \leq Q, \quad \forall t \in \mathcal{T}, \forall r \in \mathcal{R}.$$

Model for the uncertainty on diffusion

The diffusion coefficients in G are uncertain; they cannot be measured with precision and they depend on random phenomena (wind, etc.). At best we can assign a range of possible values

$$\bar{G}_{ij} - \hat{G}_{ij} \leq G_{ij} \leq \bar{G}_{ij} + \hat{G}_{ij}$$

or

$$G_{ij}(\xi) = \bar{G}_{ij} + \hat{G}_{ij}\xi_{ij},\ \xi_{ij} \in [-1, 1].$$

Structural feature

The problem with uncertain diffusion coefficients has 3 periods. Nevertheless, we claim that it is static from the view point of uncertainty.

- Uncertainty is never fully revealed. One observes pollutant levels, but the inverse problem of finding the value of the diffusion coefficients is certainly tough. More importantly, knowing the value taken by the diffusion coefficients after one period says very little on their future value in the next period (e.g., the dominant winds may change).
- The observed air quality is not a usable information for the investment and production decisions. Those can be fixed in the first period.
- Investment and production are fixed in the initial stage. Their impact on the air quality is observed with no possible recourse to alleviate bad occurrences.

Simulation study

Underlying random factors ξ for the pollutant diffusion.

Assumption (Assumption for the simulation)

the underlying factors to be i.i.d. with a uniform distribution on the range $[-1, 1]$.

A scenario is a multidimensional vector of realizations of the random factors ξ. The proposed solution is tested on 1000 scenarios.

Impact of uncertainty on pollutant diffusion

Table 6 gives the simulation results with uncertainty on the pollutant diffusion coefficients. We observe that for only 46.5% of the simulations all the air quality constraints are satisfied. For 4.2%, 47.8%, and 1.5% of the cases, one, two and three air quality constraints are violated, respectively. The average and maximum relative violations are about 4.3% and 9.3%, respectively.

Average relative violation (%)	4.3
Maximum relative violation (%)	9.3
% of violation	53.5
% of one violation	4.2
% of two violations	47.8
% of three violations	1.5

Table: Simulation results with uncertain source-receptor matrix.

Hedging against uncertainty on diffusion

Could alternative choices in technologies and operating mode hedge against the uncertainty in the diffusion factors?

Constraint depending linearly on the uncertainty

Recall that the air quality constraint is linear in the uncertain diffusion coefficient. We develop the analysis under the base assumption of linearity.

$$\begin{cases} \sum_i a_i f_i(x) \leq b, \quad \text{(linear const. if } f_i(x) = x_i) \\ \text{with } a \text{ depending linearly on a base random factor } \xi \\ \\ a_i = a_i^0 + \sum_k a_i^k \xi_k \leq b \end{cases}$$

$$\underbrace{\sum_i a_i^0 f_i(x)}_{\text{certain}} + \underbrace{\sum_k (\sum_i a_i^k f_i(x))\xi_k}_{\text{uncertain}} \leq b. \tag{1}$$

From now on we focus on the uncertain term and on its possible large values.

Capturing the knowledge on uncertainty

In Robust Optimization we want to face the case that the decision-maker has limited knowledge on the uncertainty. Suppose the D-M can only provide a range of variation for each a_i:

$$\underline{a}_i \leq a_i \leq \bar{a}_i.$$

Define

$$\begin{aligned} a_i^0 &= \frac{\bar{a}_i + \underline{a}_i}{2} \\ a_i^k &= \frac{\bar{a}_i - \underline{a}_i}{2} \quad \text{for } k = 1,\ 0 \text{ otherwise} \\ \xi_k &\in [-1, 1] \end{aligned}$$

Worst possible situation according to the D-M

The worst value of the uncertain term is given by

$$\max\{\sum_i (a^i_i f_i(x))\xi_i : -1 \leq \xi_i \leq 1, \forall i\}.$$

This simple LP in ξ has dual

$$\min_u\{\sum_i u_i : -u_i \leq a^i_i f_i(x) \leq u_i\} = \sum_i |a^i_i f_i(x)|.$$

We recognize the norm 1 of the vector of component-wise products $a \bullet f(x)$

Remark

The analysis is performed in the ξ space, not in the x space.

Robust solution for the D-M

Replace the uncertain term by its worst case $\sum_i |a_i^i f_i(x)|$ and claim that the all x that are such that

$$\sum_i a_i^0 f_i(x) + \sum_i |a_i^i f_i(x)| \leq b$$

is a robust solution for the uncertain constraint

$$\sum_i a_i f_i(x) \leq b.$$

Remark

The analysis performed in the ξ space leads to a condition in the x space.

Summary of the Robust Optimization approach

1. RO applies to constraints in which the uncertain parameters appear in a linear form. (Extensions to nonlinear cases exist.)
2. The information on the uncertain parameters is captured into a set named the *uncertainty set*. In our first example, the uncertainty set is just a set of ranges for the parameters.
3. Using duality we majorized the uncertain term $\sum_k(\sum_i a_i^k f_i(x))\xi_k$ by its worst case value on the uncertainty set $\sum_i |a_i^j f_i(x)|$.
4. We plugged in this bound as a safety factor to be added to the nominal value $\sum_i a_i^0 f_i(x)$ and obtained a condition free of the uncertain factor ξ. This condition is named *equivalent robust counterpart*.

More sophisticated uncertainty set

In addition to the range information, the D-M believes that not all uncertain factors ξ_j can achieves simultaneously large absolute values.

1. Ellipsoidal uncertainty set

$$\Xi = \{\xi : (\sum_i \xi_i^2)^{\frac{1}{2}} \leq \sigma, \ -1 \leq \xi_i \leq 1\}$$

2. Polyhhedral uncertainty set

$$\Xi = \{\xi : \sum_i |\xi_i| \leq \sigma, \ -1 \leq \xi_i \leq 1\}$$

σ is an immunization factor. The larger σ, the larger the uncertainty set, and the larger the worst case value of the uncertain component of the constraint.

Computing the worst case on the uncertainty set

- Set $z_i = a_i^i f_i(x)$. The worst case value of the uncertainty term is given by the optimum of the LP problem

$$\max\{\sum_i z_i \xi_i : \sum_i |\xi_i| \leq \sigma, \ -1 \leq \xi_i \leq 1\},$$

  and also by the minimum of its dual. The dual is written compactly as

$$\min_u \left(||u||_1 + \sigma ||z - u||_\infty\right).$$

- The *robust counterpart* of the uncertain constraint is

$$\sum_i a_i^0 f_i(x) + \min_u \left(||u||_1 + \sigma ||z - u||_\infty\right) \leq b$$

  with $z_i = a_i^i f_i(x)$.
- If the constraint is embedded in an optimization problem, we can drop the min operator and let the overall optimization scheme manage the auxiliary variable u

$$\sum_i a_i^0 f_i(x) + ||u||_1 + \sigma ||z - u||_\infty \leq b.$$

Computing the worst case on the uncertainty set

With the ellipsoidal uncertainty set

$$\sum_i a_i^0 f_i(x) + ||u||_1 + \sigma ||z - u||_2 \leq b$$

with $z_i = a_i^i f_i(x)$.
If $f_i(x)$ is linear in x, the equivalent robust counterpart is

1. Conic quadratic if Ξ is ellipsoidal (as displayed above).
2. Linear if Ξ is polyhedral (the norm 2 above is replaced by the infinity norm).

The equivalent robust counterpart is numerically tractable by usual tools of convex optimization.

Application

Solve the planning problem with uncertain diffusion coefficients

$$G_{ij}(\xi) = \bar{G}_{ij} + \hat{G}_{ij}\xi_{ij},\ \xi_{ij} \in [-1, 1]$$

and the ellipsoidal uncertainty

$$\Xi = \{\xi : ||\xi||_2 \leq \sigma\}.$$

Replace the uncertain air quality constraint by its equivalent robust counterpart

$$\sum_{\rho \in \mathcal{R}} e_{p,\rho}\bar{G}_{r,\rho} + ||u||_1 + \sigma \, || \sum_{\rho \in \mathcal{R}} e_{p,\rho}\hat{G}_{r,\rho} - u||_2 \leq Q \quad \forall p \in \mathcal{P}, r \in \mathcal{R}.$$

Simulation study

	Deterministic ($\sigma = 0$)	Robust $\sigma = 0.8$	Robust $\sigma = 1$	Robust $\sigma = 1.2$
Cost performance	162.06	163.76	164.21	164.64
Cond. mean relative violation (%)	4.3	1.0	0.5	-
Maxi relative violation (%)	9.3	3.0	1.5	-
scenarios with violation (%)	53.5	25.4	6	0
scenarios with 1 violation (%)	4.2	17.2	5	-
scenarios with 2 violations (%)	47.8	6.4	1	-
scenarios with 3 violations (%)	1.5	1.8	-	-

Table: Variable transfer coefficients: behavior on the sample of 1000 scenarios of a robust solution with ellipsoidal uncertainty and different immunization factors.

Optimal investment schedule (RO vs deterministic)

	t_1			t_2			t_3			
	p_1	p_2	p_3	p_1	p_2	p_3	p_1	p_2	p_3	Total
r_1	-	1.41	-	-	0.77	-	-	-	-	2.18
	-	1.39	-	-	1.51	-	-	0.1	-	3
r_2	-	6.09	-	-	1.73	-	-	0.15	-	7.97
	-	5.94	-	-	0.91	-	-	0.16	-	7.01
r_3	-	-	-	-	-	-	-	-	3.85	3.85
	-	-	-	-	-	-	-	-	3.74	3.74
r_4	-	-	-	-	-	-	-	-	-	0
	0.17	-	-	-	0.08	-	-	-	-	0.25
Sum	-	7.50	-	-	2.50	-	-	0.15	3.85	
	0.17	7.33	-	-	2.50	-	-	0.15	3.85	

Table: Optimal investment schedule. The robust solution with immunization factor 0.8 is in red. The % of scenarios with violations drops from 53.5 to 12.4 and the cost increases form 162.06 to163.99.

Model for the uncertainty on demand

Our illustrative example with uncertain demands

$$
\begin{aligned}
d_1 &= \bar{d}_1 + \hat{d}_1 \eta_1 \\
d_2 &= \bar{d}_2 + \hat{d}_1 \eta_1 + \hat{d}_2 \eta_2 \\
d_3 &= \bar{d}_3 + \hat{d}_1 \eta_1 + \hat{d}_2 \eta_2 + \hat{d}_3 \eta_3
\end{aligned}
$$

where $\bar{d}_i$ and $\hat{d}_i$ are the average demands and the variability of the demands, respectively.

In the example $\bar{d}_1 = 10$, $\bar{d}_2 = 12$, $\bar{d}_3 = 14$ and $\hat{d}_i = 0.1\bar{d}_i$, $\forall i$.

The η are independent random perturbations. The range of variation of the demand around the average is $\pm 5\%$, $\pm 10\%$ and $\pm 15\%$, in periods 1, 2 and 3, respectively.

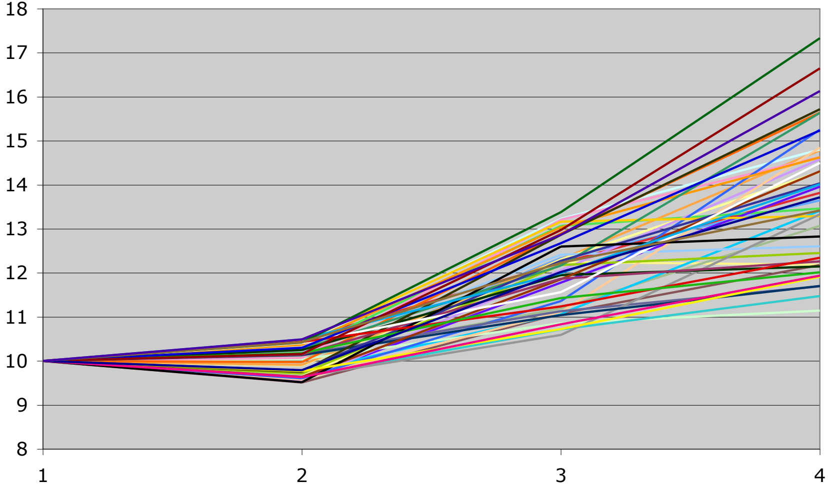
Simulation study

Underlying random factors η for the demand.

Assumption (Assumption for the simulation)

the underlying factors to be i.i.d. with a uniform distribution on the range $[-1, 1]$*.*

A scenario is a multidimensional vector of realizations of the random factors ξ. The proposed solution is tested on 1000 scenarios.



Examples of trajectories with η uniform on $[-1,1]$

Testing performance of the deterministic solution

Simulation scheme.

1. Solve the problem with the nominal demands only $(\bar{d}_1, \bar{d}_2, \bar{d}_3)$.
2. A triple (η_1, η_2, η_3) is selected at random and defines a scenario of demands. The investment and production used in the simulation are determined as follows
   - Investment variables take the value assigned by the optimization.
   - But production variables must be adapted:
     - if production is in excess of demand, scale it down to the meet the demand;
     - if production is insufficient, record a failure.
3. The deterministic solution is tested on 1000 scenarios

Result on simulations

Predicted cost performance	162.06
Observed cost performance	160.87
Scenarios with demand violation(s) in %	62.0
Conditional average relative violation in %	2.5
Average number of violations per scenario	2.0

Table: Simulation results with uncertain demands.

How come the observed cost is smaller than the one predicted by the deterministic model?

When demand is low, production is downscaled and the costs are decreased. When demand is high, production is unchanged, but shortfall occur. The bias: the cost of shorfall is not recorded.

Adjustable variable

To make simulations more meaningful, we downscaled production variables when possible to eliminate irrelevant over-production. This cuts operation costs whenever demand is less than nominal.

Suggestion

Make the adjustment part of the optimization modeling itself.

Adjustable variables should become an intrinsic feature in the modeling of uncertainty in dynamic problems. In a realistic optimization model, production should the track ups and downs of the demand.

Decision based on revealed information

In contrast with the static case, investment and production decisions should not be fixed once and for all.

- In each period, production should be set up so as to meet the manifested demand of the period (inasmuch as production capacities allow it).
- The observation of the demand in period p conveys useful information of the possible values of the demand in period $p+1$ (we have chosen an autoregressive model for the demand). So, investment decisions in period p should take advantage of this valuable information.

Production in periods $p = 1, 2, 3$ and investments in period $p = 2, 3$ are recourse variables

On the complexity of multistage stochastic programming

Shapiro and Nemirovski point out in 2005

> A multistage problem [...] is, generically, “severely computationally intractable”.

No hope!

Affine decision rules

We now define, for all pairs ($p \in \mathcal{P}, r \in \mathcal{R}$), the following linear decision rules for the production variables $x_{p,t,r}$ and the investment variables $y_{p,t,r}$.

$$
\begin{aligned}
x_{p,1,r} &= \alpha^0_{p,1,r} + \alpha^1_{p,1,r}\eta_1 \\
x_{p,2,r} &= \alpha^0_{p,2,r} + \alpha^1_{p,2,r}\eta_1 + \alpha^2_{p,2,r}\eta_2 \\
x_{p,3,r} &= \alpha^0_{p,3,r} + \alpha^1_{p,3,r}\eta_1 + \alpha^2_{p,3,r}\eta_2 + \alpha^3_{p,3,r}\eta_3 \\
y_{p,1,r} &= \beta^0_{p,1,r} \\
y_{p,2,r} &= \beta^0_{p,2,r} + \beta^1_{p,2,r}\eta_1 \\
y_{p,3,r} &= \beta^0_{p,3,r} + \beta^1_{p,3,r}\eta_1 + \beta^2_{p,3,r}\eta_2.
\end{aligned}
$$

We also relax the constraint production $=$ demand to production $\geq$ demand

Affine decision rules

The same authors (Shapiro and Nemirovski) add

> The only reason for restricting ourselves with affine decision rules stems from the desire to end up with a computationally tractable problem. We do not pretend that affine decision rules approximate well the optimal ones—whether it is so or not, it depends on the problem, and we usually have no possibility to understand how good in this respect is a particular problem we should solve. The rationale behind restricting to affine decision rules is the belief that in actual applications it is better to pose a modest and achievable goal rather than an ambitious goal which we do not know how to achieve.

Results (Robust solution with affine decision rule)

	Deterministic	LDR with robust		
		$k_{dem} = 0.4$	$k_{dem} = 0.6$	$k_{dem} = 0.7$
	Objective function			
Predicted cost performance	162.06	165.84	167.09	167.40
Observed cost performance	160.87	165.07	166.14	167.47
	Constraints on the demand			
% scen. with demand violation	62.0	20.4	9.9	0
Cond. average rel. violation in %	2.5	0.8	0.6	-
Av. # of viol. / scenario	2.0	1	1	-
	Constraints on the air quality			
Total # of viol. air quality constr.	0	53	39	82

Table: Variable demand: behavior on the sample of 1000 scenarios of a robust solution with LDR and different immunization factors k_{dem}.

1. Robust Optimization illustrated on a simple example
   - Deterministic formulation
   - Impact of uncertainty in pollutant diffusion
   - The robust optimization approach

2. Dynamic problem on the illustrative example

3. RO vs Chance-Constrained Programming

4. Energy Security in EU supply via TIAM

5. Concluding remarks

Robust Optimization

In Robust Optimization we wrote each constraint as an affine function of the uncertain factor ξ

$$z_0 + \sum_{i \geq 1} z_i \xi_i \leq 0$$

with $z_0 = \sum_i a_i^0 f_i(x) - b$ and $z_i = \sum_i a_k^i f_k(x)$.

The motto in RO is

1. The probability distribution of ξ is not known (or not usable).
2. Whatever is known about ξ is captured by an inclusion in a set $\xi \in \Xi$.
3. A robust solution is such that for all $\xi \in \Xi$, the constraint $z_0 + \sum_{i \geq 1} z_i \xi_i \leq 0$ remain feasible.
4. The set Ξ and the functions f_i must be such that the computation of a robust solution is numerically tractable.

Chance-Constrained Programming

What would be the formulation in CCP for a problem with typical constraint

$$\zeta = z_0 + \sum_{i \geq 1} z_i \xi_i \leq 0?$$

1. We know the probability distribution P of ξ.
2. We "pretend" we can compute

$$p(z) = \mathrm{Prob}_{\xi \sim P}(z_0 + \sum_{i \geq 1} z_i \xi_i \leq 0)$$

3. We assign a bound ϵ on the probability of constraint violation. A solution is CCP is

$$\boxed{p(z) \geq 1 - \epsilon}$$

for all constraints.

Drawbacks of CCP

1. Do we really know the probability distribution P of ξ?
2. Except for very few cases (e.g., the ξ_i are independent Gaussian variables), computing the probability $p(z)$ is numerically intractable.
3. Even if we can compute $p(z)$, the solution set in z of $p(z) \leq \epsilon$ (and ultimately in x through $z(x)$) is often non-convex if not disconnected.

Safe convex approximations of CCP

A functional representation of the probability of constraint violation:

$$\mathrm{Prob}(\zeta > 0) = E(I(\zeta))$$

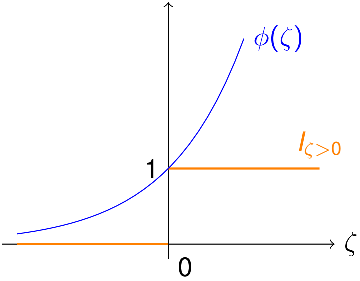
$$I(\zeta) = \begin{cases} 0 & \text{if } \zeta \leq 0 \\ 1 & \text{if } \zeta > 0. \end{cases}$$

How to build a convex approximation?

Replace $I(\zeta)$ by a dominating convex function $\phi(\zeta) \geq I(\zeta)$

$$E(I(\zeta)) \leq E(\phi(\zeta)) \leq \epsilon \quad \Rightarrow \quad \mathrm{Prob}(\zeta > 0) \leq \epsilon.$$

Safe convex approximations of CCP



$\phi(\zeta)$ convex implies $E(\phi(\zeta))$ is convex.

The constraint $E(\phi(\zeta)) \leq \epsilon$ is also convex

If we can compute the expectation efficiently, we are done.

Safe convex approx. as a robust constraint

General result

A safe convex approximation $E(\phi(\zeta)) \leq \epsilon$ of the chance constraint

$$\mathrm{Prob}(\zeta > 0) \leq \epsilon$$

is representable by a robust constraint

$$z_0 + \sum_{i \geq 1} z_i \xi_i \leq 0, \ \forall \xi \in \Xi$$

where Ξ is a convex set.

In other words, solving the constraint of the safe convex approximation is equivalent to solving the robust constraint with uncertainty set Ξ.

Tractability issue

In general, neither the safe convex approximation, nor the robust constraint formulation are numerically tractable.

- Computing the expectation $E(\phi(\zeta))$ is likely to be intractable.
- A formal of Ξ exists, but often not in a computationally usable format. And even when it is usable, the worst case computation

$$\max_{\xi}\{\sum_{i\geq 1} z_i\xi_i : \xi \in \Xi\}$$

may remain intractable.

A favorable case

Assumption

The ξ_i are independent.

Use the moment generating function $\phi(t) = \exp(t)$. By Assumption

$$E(\phi(t)) = \exp(z_0) \prod_i \exp(z_i \xi_i)$$

To compute the expectations on the right, we need to know the true distribution P_i of ξ_i. Even if we know the distribution, this may still turn out to be difficult.

Ambiguous chance constraint

Assumption

1. *The ξ_i are independent.*
2. *The ξ_i have common support $[-1, 1]$ and common expectation $E(\xi_i) = 0$.*

There are many distributions that are compatible with Assumption 2. They form a class $P_i \in \mathcal{P}$. But we may address the worst case in the class:

$$\max_{P_i \in \mathcal{P}, \forall i} \text{Prob}(z_0 + \sum_i z_i \xi_i > 0) \leq \epsilon.$$

Ambiguous chance constraint

We can identify the worst case distribution (*degenerate distribution with equal weights on* -1 *and* 1) and compute the expectation. After some more mathematical massaging, we obtain the sufficient condition

$$z_0 + \sqrt{2 \ln \frac{1}{\epsilon}} \sqrt{\sum_{i>0} z_i^2} \leq 0 \tag{2}$$

$$\Downarrow$$

$$\max_{P_i \in \mathcal{P}, \forall i} \text{Prob}(z_0 + \sum_{i>0} z_i \xi_i > 0) \leq \epsilon. \tag{3}$$

Because we assume $z_i(x)$ to be affine in x, the equivalent robust counterpart (2) is conic quadratic in x.

Other classes of ambiguous distributions

For many classes of ambiguous distributions, the general equivalent robust counterpart is of the form

$$z_0(x) + \sum_{i \geq 1} \max\{\mu_i^- z_i(x), \mu_i^+ z_i(x)\} + \sqrt{2 \ln \frac{1}{\epsilon}} \sqrt{\sum_i \sigma_i^2 z_i^2(x)} \leq 0.$$

The parameters $\mu_i^-, \mu_i^+, \sigma_i$ depend on the class of the ambiguous distribution.

For instance, let $\mathcal{P}$ be the class of distributions supported on $[-1, 1]$ and unimodular w.r.t. 0, then

$$\mu^+ = -\mu^- = \frac{1}{2},\ \sigma^2 = \frac{1}{12}.$$

Table from copied from the book

A. BEN-TAL, L. EL GHAOUI, AND A. NEMIROVSKI, *Robust Optimization*. Princeton University Press, 2009.

A priori information on P	P satisfies P.2 with parameters μ^-	μ^+	σ	Remark
$\mathrm{supp}(P) \subset [-1,1]$	-1	1	0	
$\mathrm{supp}(P) \subset [-1,1]$ P is unimodal w.r.t. 0	$-\frac{1}{2}$	$\frac{1}{2}$	$\sqrt{\frac{1}{12}}$	
$\mathrm{supp}(P) \subset [-1,1]$ P is unimodal w.r.t. 0 P is symmetric w.r.t. 0	0	0	$\sqrt{\frac{1}{3}}$	
$\mathrm{supp}(P) \subset [-1,1]$ $[-1 <]\ \mu^- \le \mathrm{Mean}[P] \le \mu^+\ [< 1]$	μ^-	μ^+	$\Sigma_{(1)}(\mu^-,\mu^+)$	(2.4.29)
$\mathrm{supp}(P) \subset [-1,1]$ $[-\nu \le]\ \mu^- \le \mathrm{Mean}[P] \le \mu^+\ [\le \nu]$ $\mathrm{Var}[P] \le \nu^2 \le 1$	μ^-	μ^+	$\Sigma_{(2)}(\mu^-,\mu^+,\nu)$	(2.4.32)
$\mathrm{supp}(P) \subset [-1,1]$ P is symmetric w.r.t. 0 $\mathrm{Var}[P] \le \nu^2 \le 1$	0	0	$\Sigma_{(3)}(\nu)$	(2.4.34)
$\mathrm{supp}(P) \subset [-1,1]$ P is symmetric w.r.t. 0 P is unimodal w.r.t. 0 $\mathrm{Var}[P] \le \nu^2 \le 1/3$	0	0	$\Sigma_{(4)}(\nu)$	(2.4.36)

A useful class

The ξ_i are Gaussian with partially known expectations

$$\mu_i^- \leq \mu_i \leq \mu_i^+$$

and a bound on the true variance ν_i^2

$$\nu_i^2 \leq \sigma_i^2.$$

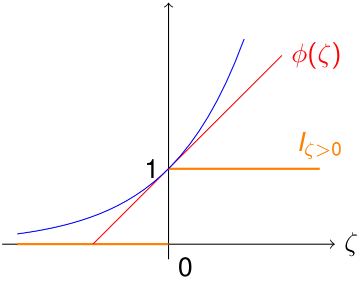
The best approximation is

$$z_0(x) + \sum_{i \geq 1} \max\{\mu_i^- z_i(x), \mu_i^+ z_i(x)\} + \mathrm{Errfcn}(\epsilon)\sqrt{\sum_i \nu_i^2 z_i^2(x)} \leq 0$$

where *Errfcn* is the error function (inverse of the standard normal distribution). For $\epsilon = 0.1$, the ratio $\mathrm{Errfcn}/\sqrt{2\ln\frac{1}{\epsilon}} \approx 0.6$, and tends to 1 quickly as $\epsilon \to 0$.

Least conservative safe convex approximation

Is there a least conservative safe convex approximation ? The idea is to start with the best upper approximation function ϕ



$$\phi(\zeta) = \max\{1+\zeta, 0\} = (1+\zeta)^+$$

What is $E((1+\zeta)^+)$?

Best convex approximation

$$\forall\alpha > 0, \quad E((1+\alpha\zeta)^+) \le \epsilon \Rightarrow \text{Prob}(\alpha\zeta > 0) \le \epsilon \Leftrightarrow \text{Prob}(\zeta > 0) \le \epsilon$$

$$E((1+\alpha\zeta)^+) \le \epsilon,\ \forall\alpha$$

$$\Updownarrow$$

$$\beta E((1+\tfrac{\zeta}{\beta})^+) - \beta\epsilon \le 0,\ \forall\beta = \alpha^{-1} > 0$$

$$\Updownarrow$$

$$\min_{\beta>0}\, [E((\beta+\zeta)^+) - \beta\epsilon] \le 0$$

$$\Updownarrow$$

$$\min_{\beta>0} \left[-\beta + \tfrac{1}{\epsilon} E((\beta+\zeta)^+)\right] \le 0$$

The last inequality displays on the left the CVaR at level ϵ of the random variable ζ.

Using CVaR to approximate chance constraint

CVaR was introduced as a measure of risk. See for instance R.T. ROCKAFELLAR AND S.URYASEV. Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26:14431471, 2002.

But

Main result

The CVaR constraint

$$\text{CVaR}(z) \leq 0$$

is the least conservative safe convex approximation of the chance constraint

$$p(z) = \text{Prob}(\zeta > 0) \leq \epsilon.$$

A caveat

We do not know how to compute the expectation in the CVaR formulation, unless the distribution is finite.

A frequent misunderstanding

Suppose $z(x)$ is robust with respect to Ξ. Clearly

$$\text{Prob}(z_0(x) + \sum z_\ell(x)\xi_\ell > 0) \leq \text{Prob}(\xi \notin \Xi). \tag{4}$$

People sometimes conclude that the robust solution requires that $\text{Prob}(\xi \in \Xi)$ is close to 1. Actually (4) may very loose: The left-hand side may be small (as guaranteed by the theorems on safe approximation) while the right-hand side is 1, that is, $\text{Prob}(\xi \in \Xi) = 0$!

A frequent misunderstanding

The degenerate distribution with equal weights on the vertices belongs to the class of distributions with zero mean. The theory says that an immunization factor $k = 1.4$ guarantees a probability of constraint satisfaction at least 0.62.

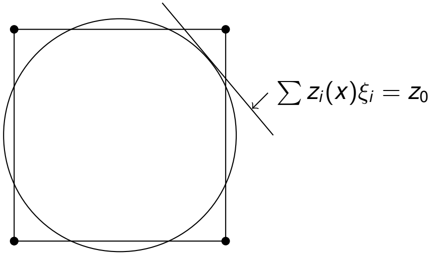


Figure: The uncertainty set Ξ does not contain any of the 4 vertices; it has thus a zero probability. But any support to Ξ separates at most one vertex from Ξ. The probability that the constraint is satisfied is at least $1 - \frac{1}{4}$.

1. Robust Optimization illustrated on a simple example
   - Deterministic formulation
   - Impact of uncertainty in pollutant diffusion
   - The robust optimization approach
2. Dynamic problem on the illustrative example
3. RO vs Chance-Constrained Programming
4. Energy Security in EU supply via TIAM
5. Concluding remarks

Introduction and Objective

- EU is strongly dependent on energy imports
- Some foreign energy sources are prone to interruptions, cost fluctuations, and other random events
- Increasing energy security might include:
  - Selection of less risky energy suppliers
  - Diversification of sources, for each energy form: oil, gas, uranium, biomass, electricity
  - Diversification of energy forms (e.g. smaller dependence on oil)
  - Reduction of energy imports
  - Reduction of total energy consumption?

Some modeling details for TIAM

- TIAM (Times Integrated Assessment Model) is a 16 region global energy model, with detailed representation of technologies, energy sources, energy trade, and demand sectors
- EU+ is one region, linked to the other 15 regions via 67 trade routes
- Each trade route is a technology, with an investment variable, an activity variable, and several technical and economic parameters

Trade route constraints in TIAM

Import via a single corridor

$$ACT_{k,t} \leq AF_{k,t} \times CAP_{k,t}$$

- k is an import corridor (from ROW to EU)
- CAP is the capacity of the corridor (decision variable)
- AF is the availability factor (random)

Total EU energy imports at period t

$$\sum_{k} (ACT_{k,t} - AF_{k,t} \times CAP_{k,t}) \leq 0$$

Uncertainty model

Uncertain availability factors

We assume that AF_k is random

$$AF_k = 1 - d_k \xi_k$$

- $0 \leq d_k \leq 1$ is a measure of the severity of the risk of corridor k
- ξ is the set of independent random variables with support $[0, 1]$ and mean μ
- $[1 - d_k, 1]$ is the range of uncertainty of the factor AF_k
- A small d_k means that the corridor has little variability, and conversely when $d_k = 1$, there is the possibility of a complete corridor shutdown

Uncertain total EU energy import constraint

$$\underbrace{\sum_k (ACT_k - CAP_k)}_{\text{certain}} + \underbrace{\sum_k d_k \cdot CAP_k \cdot \xi_k}_{\text{uncertain}} \leq 0. \tag{5}$$

Robustified transportation constraint

Let Ξ be the uncertainty set corresponding to our assumption.

$$\Xi = \{\xi : \xi_i = \frac{1+\eta_i}{2}, |\xi_i| \leq 1, ||\xi||_\infty \leq k\}$$

Proposition

Skipping all technical details, the robust constraint of

$$\sum_k (ACT_k - CAP_k) + \sum_k d_k \cdot CAP_k \cdot \xi_k \leq 0, \forall \xi \in \Xi$$

is given by deterministic system of inequalities

$$\sum_k (ACT_k - CAP_k) + d_k \mu_k CAP_k + \sum_k (1-\mu_k) u_k + \sqrt{\frac{K}{2} \ln \frac{1}{\epsilon}} \cdot v \leq 0 \tag{6a}$$

$$u_k + v - CAP_k \cdot d_k \geq 0, \; k = 1, \ldots, K \tag{6b}$$

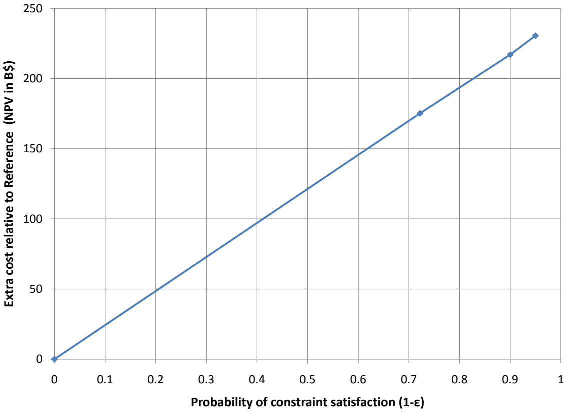
$$u_k \geq 0, \; v \geq 0, \; k = 1, \ldots, K \tag{6c}$$

It satisfies the energy constraint with probability at least $(1-\epsilon)$.

Application to EU via TIAM model

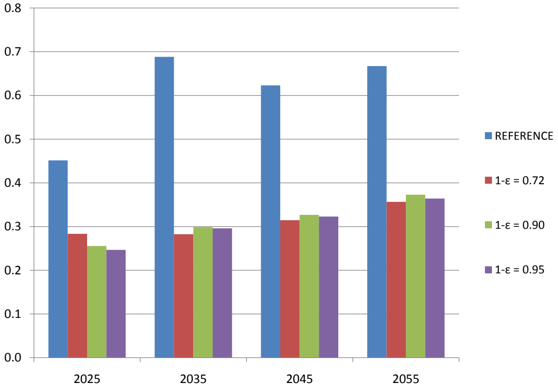
- In TIAM, there are 67 import corridors, for 4 energy forms (Oil, Gas, Coal, Electricity).
- We assumed that all corridors can be totally closed (they have the same domain of uncertainty with $d_k = 1$)
- We also assumed that all corridors have the same average availability factor $\bar{AF} = 0.6$ (i.e. $\mu_k = 0.4$) that is quite pessimistic
- Five robust constraints were created, for the 5 periods 2015, 2025, 2035, 2045, 2055 (ignoring the initial periods 2005, 2010)
- We tested three satisfaction probability levels 0.72, 0.90, and 0.95 and the reference scenario.

Cost-reliability trade-offs



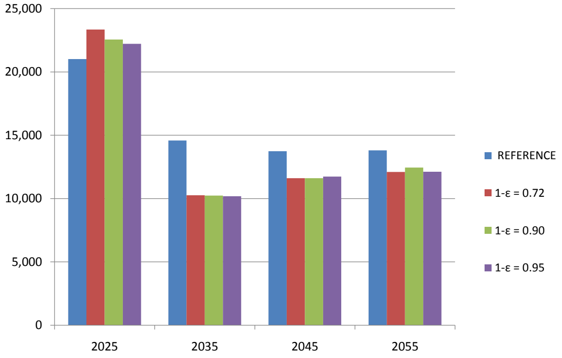
The extra costs for improving reliability range from 175 B$ to 230 B$, i.e. from 0.52% to 0.68% of total EU cost.

Average utilization of energy import channels



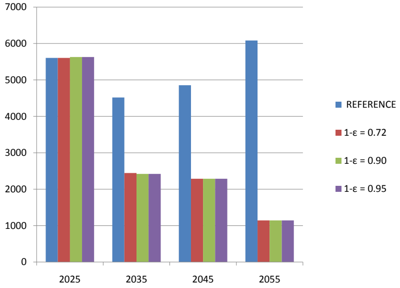
- The drop in utilization compared to Reference is caused by extra investments in channel capacities in order to insure reliability.
- Increasing reliability does not induce much extra investments in capacities.

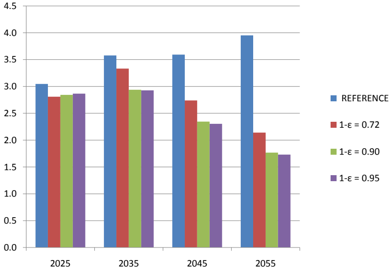
EU energy imports (PJ/yr)



- Energy imports slightly increase in 2025 and decrease significantly at later periods when reliability increases.
- We have observed that the increase in reliability is mainly achieved via shifts from imported to locally produced energy.

Maximum flow through channels and coefficient of variation of channel flows





- In 2025, not much impact is observed, due to the system's inertia
- Market dominance is reduced : Max flow is reduced, and flows tend to equalize

Comments

1. The supply of energy can be guaranteed with a known probability, under a very mild assumption.
2. Such reliability is achieved at a moderate extra cost (not exceeding 0.7% of the total EU energy cost).
3. The results show a significant reduction of the concentration of supply sources, a feature that is desirable in itself. RO favors a combination of several actions
   - Decrease imports selectively
   - Build extra corridor capacity (again in selective manner)
   - Equalize the market shares
4. The method is :
   - Easy to formulate
   - Easy to apply
   - Numerically tractable

Conclusions

- Robust optimization is a mature component of the OR toolbox to deal with uncertainty
  - Strong theoretical foundations (e.g., safe convex approx. of Ambiguous Chance Constrained Programming)
  - Practical, because it focuses on numerical tractability.
  - Supported by modeling languages, e.g., AIMMS.
- Environment assessment models and energy systems are ideal targets for RO applications.
  - Uncertainty is everywhere.
  - Models in use defy the use of classical tools, because of their intrinsic complexity.
  - The field is awaiting your contributions...

Conclusions

Extensions:

- Extensions to conic quadratic and semidefinite constraints.
- Non affine uncertain factors.
- Non affine decision rules.
- Globalized robust optimization.
- Distributionally robust optimization
- ...

ORDECSYS is developing a partnership with AIMMS to promote its powerful robust optimization add-on.

Optimization Software for Mathematical Programming and Solvers

AIMMS is an advanced development environment for building optimization based operations research applications and advanced planning systems. It is used by leading companies to support decision making in a wide range of industries. AIMMS is free for Academic use.

Paragon Decision Technology is part of the winning team of the INFORMS Edelman Award 2011

CHICAGO, April 11, 2011 – Midwest ISO, Alstom Grid, and Paragon Decision Technology have been awarded the Franz Edelman Award for outstanding Achievement in Operations Research and Management Sciences at a banquet sponsored by the Institute for Operations Research and Management Sciences (INFORMS) in Chicago tonight.

`http://www.aimms.com/`

Some references

The tutorial is available at `http://www.ordecsys.com`

R. APPARIGLIATO. Règles de décision pour la gestion du risque : Application à la gestion hebdomadaire de la production électrique. PhD thesis, Ecole Polytechnique, Paris, France, 2008.

F. BABONNEAU, A. KANUDIA, M. LABRIET, R. LOULOU, AND J.-P. VIAL. Energy security: a robust optimization approach to design a robust European energy supply via TIAM. Working paper (to appear in *Environmental Modeling and Assessment*), ORDECSYS, 2010.

F. BABONNEAU, J.-P. VIAL, AND R. APPARIGLIATO. Robust optimization for environmental and energy planning. In J. Filar and A. Haurie, editors, *Handbook on "Uncertainty and Environmental Decision Making",* pages 79–126. Springer Verlag, 2010.

F. BABONNEAU AND J.-P. VIAL. Modèle de prix fondamentaux à long terme du charbon avec investissements et incertitude sur la demande. Rapport interne pour EDF. 2010.

A. BEN-TAL, L. EL GHAOUI, AND A. NEMIROVSKI. *Robust Optimization*. Princeton University Press, 2009.