An introduction to Optimization under Uncertainty with special focus on Robust Optimization

Fabio D’Andreagiovanni

(email: d.andreagiovanni@hds.utc.fr)

École Polytechnique, Palaiseau, February 8th 2017

Presentation outline

- From Deterministic to Uncertain Optimization
- An overview of methodologies for Uncertain Optimization
- Fundaments of Robust Optimization
- A classic: the Bertsimas-Sim model

Why the special focus on Robust Optimization?

- I know most about this topic (theoretical + applied experience)
- Consulting experience in industry (optimization under worst case)
- (Reasonably) contained increase in problem complexity

DESIRABLE QUALITY FOR UNCERTAIN
HARD-TO-SOLVE REAL-WORLD PROBLEMS

It’s a stochastic world

Most of real-world optimization problems involve uncertain data

FINANCE

Stock value

AIRCRAFT SCHEDULING

Flight delays

TLC NETWORK DESIGN

Traffic flows

SURGERY SCHEDULING

Requests of operations

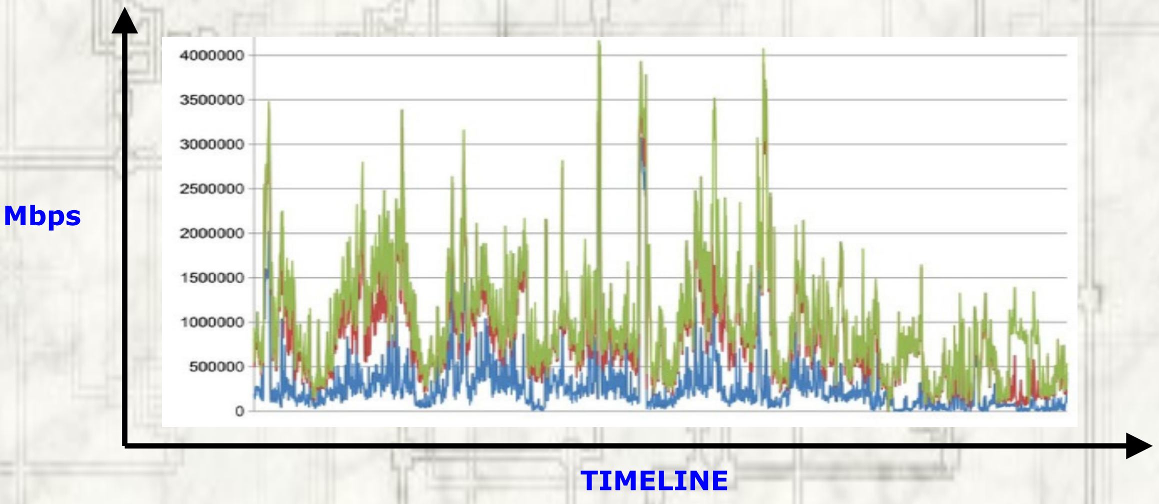
The topic of Uncertainty in Optimization was identified already by **George Dantzig**, the father of Linear Programming and an icon of Operations Research

(*Linear Programming under Uncertainty*, Management Science, 1955)

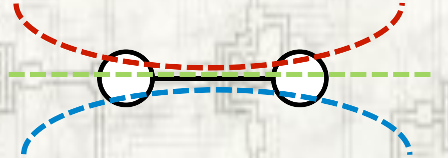
Given the presence of uncertainty in a problem, do we really need to take care of it?

- What if we neglect uncertainty? Do we risk to get meaningless solution?

An example: traffic uncertainty in Network Design



Traffic fluctuations of three O-D pairs
in the USA Abilene Network
(one-week observation)



- In every origin-destination pair, traffic volume heavily fluctuates over the week

- Overall fluctuation in a network link even more severe

- Solution of the professional: dimension network capacity by (greatly) overestimating demand

? CAN WE DEFINE A BETTER ROBUST SOLUTION THROUGH OPTIMIZATION ?

Data uncertainty in Optimization

CLASSIC OPTIMIZATION

$$\begin{aligned} \max \quad & c' \, x \\ & A \, x \leq b \\ & x \geq 0^n \end{aligned}$$

→ THE VALUE OF ALL COEFFICIENTS **IS KNOWN EXACTLY**

REASONABLE ASSUMPTION FOR ANY PROBLEM ? NO!

Neglecting data uncertainty may lead to bad surprises:

- nominal optimal solutions may result heavily suboptimal
- nominal feasible solutions may result infeasible

THEY OVERLOOKED DATA UNCERTAINTY…

To avoid such situations, we want to find **robust solutions:**

ROBUST SOLUTION = solution that remains feasible even when the input data vary (**PROTECTION AGAINST DATA DEVIATIONS**)

Relativity of Robustness

- The term Robustness is nowadays overused in Optimization

example: Robust Telecom Network Design (= robust against connection failures)

example: Robust Road Routing (= robust against non-rational decision makers)

- In this presentation:

ROBUST SOLUTION = solution protected against deviations of the input data

CRITICAL REMARKS

⚠

1) The question of how modeling the protection is open
2) Over the years, many protection models have been proposed
3) There is no evidence of the existence of a dominating model

⚠

ANYWAY

In my experience, Professionals like some models more than other models

(they can understand them and actively participate to their tuning! → better solutions)

It was not robust…

- A simple numerical example may clarify the effects of data deviations:

Suppose that we have computed an optimal solution x=1, y=1 for some problem with nominal constraint:

$$100\,x + 200\,y \leq 300$$

However, we have neglected that the coefficient of *x* may deviate up to 10%, so we could have

$$(100 + 10) + 200 > 300$$

OPTIMAL SOLUTION ACTUALLY INFEASIBLE!

- What if this was part of a problem to detect water contamination?

A simple example of uncertain problem

NEWSVENDOR PROBLEM

- Company producing **x units of a product** to meet a **demand d**
- Unitary production cost c
- Overproduction (x > d) → store left-over units (**unitary storage cost s**)
- Underproduction (x < d) → backorder missing units (**unitary order cost b: b > c**)
- **SCOPE:** establish the quantity to produce that **satisfies the demand** and **minimizes the total cost**

COST FUNCTION

$$\begin{aligned} f(x,d) &= c\,x + b\,\max\{0, d-x\} + s\,\max\{0, x-d\} \\ &= \max\{(c-b)\,x + b\,d,\ \ (c+s)\,x - s\,d\} \end{aligned}$$

→ PIECEWISE LINEAR FUNCTION WITH MINIMUM IN x* = d

OPTIMIZATION PROBLEM

$$\begin{aligned} \min \quad & f(x,d) = \max\{(c-b)\,x + b\,d,\ \ (c+s)\,x - s\,d\} \\ & x \geq 0 \end{aligned}$$

→ EQUIVALENT PROBLEM

$$\begin{aligned} \min \quad & u \\ & u \geq (c-b)\,x + b\,d \\ & u \geq (c+s)\,x - s\,d \\ & x \geq 0 \end{aligned}$$

If we know exactly the demand *d*, then we produce exactly *d* units of product

HOWEVER, future demand is generally unknown. How many units should we then produce?

Many ways of modeling data uncertainty (1)

Working hypothesis: the demand is a random variable D and we know its probability distribution

Naive way: solve the deterministic problem for the expected value of the demand

A more rational approach: minimize the **expected value of the objective cost function**

$$\begin{aligned} \min \quad & \mathbb{E}[f(x, D)] \\ & x \geq 0 \end{aligned}$$

with $$\mathbb{E}[f(x, D)] = b\,\mathbb{E}[D] + (c-b)\,x + (b+s)\int_0^x CDF_D(y)\,dy$$

and optimal solution $$x^* = CDF_D^{-1}\left(-\frac{c-b}{b+s}\right)$$

REMARKS:

- Closed form solution rarely available for real-world problems
- This solution can be very different from the one obtained for the expected demand value

Many ways of modeling data uncertainty (2)

Working hypothesis: we have characterized a number of **demand scenarios** $\mathbf{d_i}$ **, i = 1,…,I**

IF the number of scenarios is sufficiently large, THEN we could build an empirical distribution and operate as showed before

ALTERNATIVELY, we can consider a different expected value of the objective function:

$$\mathbb{E}[f(x, D)] = \sum_{i=1}^{I} p_i \; f(x, d_i)$$

PROBABILITY OF
REALIZATION
OF THE SCENARIO

$$\begin{aligned} \min \; & \sum_{i=1}^{I} p_i \; u_i \\ & u_i \geq (c - b)\, x + b\, d_i \qquad i = 1, \ldots, I \\ & u_i \geq (c + s)\, x - s\, d_i \qquad i = 1, \ldots, I \\ & x \geq 0 \end{aligned}$$

- generalization of the fixed-demand problem (= single scenario with p=1)
- decomposable structure

Many ways of modeling data uncertainty (3)

Working hypothesis: we have characterized a number of **demand scenarios d_i , $i = 1,\dots,I$**

Derive the **overall deviation range $[d^{low}, d^{up}]$** of the demand

WORST-CASE PROBLEM

$$\min_{x \geq 0} \boxed{\max_{d \in [d^{low},\ d^{up}]} f(x,d)} \longrightarrow \max\{f(x,d^{low}),\ f(x,d^{up})\}$$

$$\min_{x \in [d^{low},\ d^{up}]} \max\{f(x,d^{low}),\ f(x,d^{up})\}$$

with optimal solution $x^* = \dfrac{s\, d^{low} + b\, d^{up}}{s+b}$

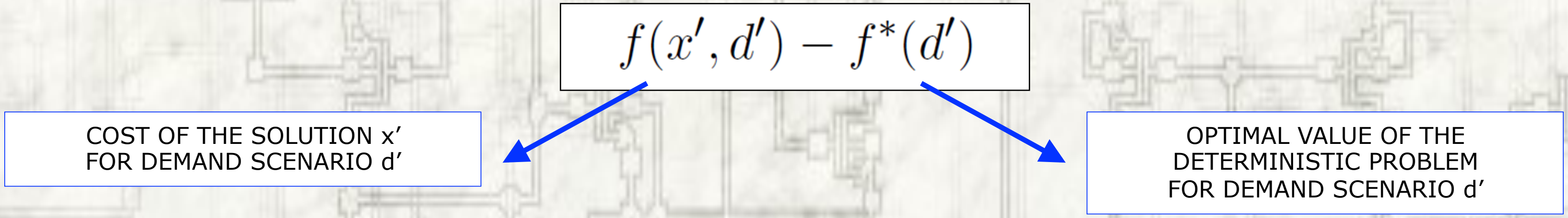
REMARKS:

- deterministically protected against all the specified deviations
- price of complete protection (Price of Robustness) = sensible increase in conservatism

Many ways of modeling data uncertainty (4)

Working hypothesis: we have characterized a number of **demand scenarios $\mathbf{d_i}$, i = 1,…,I**

Given a solution x’ for scenario d’, define its **regret** as the value:



$$f(x', d') - f^*(d')$$

COST OF THE SOLUTION x’
FOR DEMAND SCENARIO d’

OPTIMAL VALUE OF THE
DETERMINISTIC PROBLEM
FOR DEMAND SCENARIO d’

MIN-MAX
REGRET
PROBLEM

$$\min_{x \geq 0} \; \max_{d \in \mathcal{D}} \; [f(x,d) - f^*(d)]$$

Minimization of the maximum regret when considering all the possible scenarios

REMARKS:

- Takes into account all the relevant scenarios, not just the extreme deviations
- Reduced conservatism w.r.t. worst-case performance
- Remarkable increase in conmputational complexity

Many ways of managing data uncertainty (5)

A decision maker could choose to explicity control conservatism of produced solutions

$$f(x, D) \leq F$$

BUT, this could lead to problem infeasibility!

SOFTER STRATEGY: consider a probabilistic constraint

CHANCE-CONSTRAINED PROBLEM

$$\mathbb{P}\left[f(x, D) \leq F\right] \geq 1 - \alpha \qquad \alpha \in (0, 1)$$

REMARKS:

- the probabilistic constraint introduces non-convexities
- the problem becomes very hard to solve
- we need the probability distribution of D

Many ways of managing data uncertainty (6)

Other alternatives in brief:

- Robust Optimization
  - model uncertainty by additional hard constraints that cut off non-robust solutions
- Recoverable Robustness (Liebchen, Lübbecke, Möhring, Stiller, 2009)
  - solve the nominal problem
  - define (limited) reparation actions to adopt in case of bad deviations
- Light Robustness (Fischetti, Monaci, 2007)
  - a kind of Robust Optimization adding bound on the so-called Price of Robustness

Let’s take a first break

- **World is stochastic** and most of real-world optimization problems involve **uncertain data**, whose presence **cannot be neglected**

- Many models are available for representing uncertain data in optimization

- No model dominates the others from a theoretical point of view…
  …but **Robust Optimization** is emerging as the most effective way to **model and actually solve real-world problems**
  (and Professionals like it! - deterministic protection and accessibility)

Stochastic Programming – some more details

Let's say something more about Stochastic Programming (SP):

- oldest approach to Optimization under Uncertainty
- well-investigated topic
  (substantial literature – large community)

Anyway, I will limit the attention to fundaments of SP

In my experience, SP is still hard to adopt in real-world problems

- need for probability distributions of the uncertain data
- huge hard-to-solve problems
- not easily accessible to professionals

Two-stage Stochastic Programming (1)

We deal with an uncertain **min-cost problem** where:

- we must take a **decision in a first stage**
- after this first stage, the **uncertain data reveal** their actual values
- we may take a **second-stage decision** based on the observed data

Assuming to have a **set of uncertainty scenarios s = 1,…,S** we consider the problem:

$$\min \; c'\,x \;+$$

SOMETHING TO TAKE INTO ACCOUNT UNCERTAINTY AND SECOND-STAGE

FIRST-STAGE CONSTRAINTS

$$A\,x = b$$

$$x \geq 0^n$$

FIRST-STAGE - SECOND-STAGE LINKING CONSTRAINTS

$$E_s\,x + F\,y_s = g_s \qquad s = 1,\dots,S$$

$$y_s \geq 0^q \qquad s = 1,\dots,S$$

SECOND-STAGE VARIABLES
(DEPENDING UPON THE SCENARIO s)

Two-stage Stochastic Programming (2)

The second-stage variables represent the fact that we are not completely at the mercy of Nature

Given an uncertainty scenario s we can react smartly → RECOURSE ACTIONS

$$
\begin{aligned}
\min\ & c'\,x \quad + \ \text{RECOURSE COST} \\
& A\,x = b \\
& x \geq 0^{n} \\
& E_s\,x + F\,y_s = g_s \qquad s = 1, \ldots, S \\
& y_s \geq 0^{\,q} \qquad\qquad\qquad s = 1, \ldots, S
\end{aligned}
$$

The recourse is defined by:

- the second-stage decision variables y_s
- the recourse matrix F
- the cost of recourse (not yet characterized)

Two-stage Stochastic Programming (3)

We can reasonably model recourse cost by an additional vector **d** → $d' \, y_s$

… still, we have to consider scenario uncertainty → $\mathbb{E}_s \left[d' \, y_s \right]$

The resulting Stochastic Program:

$$
\begin{aligned}
\min \quad & c' \, x + \sum_{s=1}^{S} p_s \, d' \, y_s \\
& A \, x = b \\
& x \geq 0^n \\
& E_s \, x + F \, y_s = g_s && s = 1, \ldots, S \\
& y_s \geq 0^q && s = 1, \ldots, S
\end{aligned}
$$

PROBABILITY OF OCCURRENCE
OF SCENARIO s

- PRO: Linear Program
- CON: HUGE Linear Program

q recourse vars by *s* scenarios

m linking constraints by *s* scenarios

SP dimension may easily explode

As an example, consider a stochastic unit commitment problem

We are given:

- a set of power plants
- an estimate of the energy demand for each hour of the day and for each energy district

We want to choose the energy production level of each plant for each hour so that:

- the total production cost is minimized
- the estimated demand is satisfied

Two-stage stochastic perspective:

- first-stage cost is the (exactly known) energy production cost
- recourse cost = cost of balancing the network grid when not meeting district demand

Explosion of problem dimension even for coarse stochastic demand modeling:

3 demand estimates $\{d_z^{min}, d_z^{avg}, d_z^{max}\}$ for each district z

20 districts

→

3^20 scenarios
=
more than 3 billions
demand scenarios!

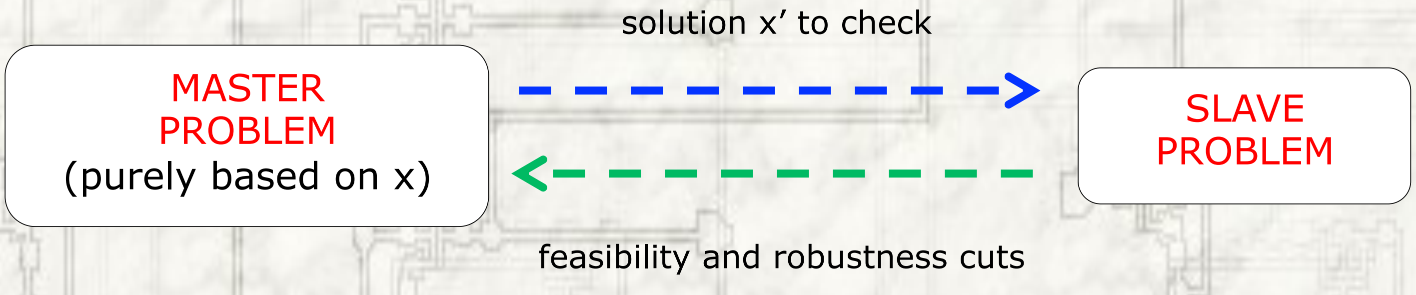
A decomposition approach

Let's visualize the problem in the following way:

$$\begin{array}{lcccccccccccl}
\min & c'\,x & + & p_1\,d'\,y_1 & + & p_2\,d'\,y_2 & + & \cdots & + & p_S\,d'\,y_S & & & \\
& A\,x & & & & & & & & & & = & b \\
& E_1\,x & + & F\,y_1 & & & & & & & & = & g_1 \\
& E_2\,x & & & + & F\,y_2 & & & & & & = & g_2 \\
& \vdots & & & & & & \ddots & & & & \vdots & \\
& E_S\,x & & & & & & & + & F\,y_S & & = & g_S \\
& x \geq 0^n, & & y_s \geq 0^q & & s = 1, \ldots, S & & & & & & &
\end{array}$$

Do you notice anything?

The problem is decomposable!



solution x' to check

MASTER
PROBLEM
(purely based on x)

SLAVE
PROBLEM

feasibility and robustness cuts

Another break before moving alone

- What you have seen about Stochastic Programming is just the tip of the iceberg!

(more than 50 years of research on the topic!)
uncountable SP modeling and solution methodologies

- I have tried to sketch essential features of the approach that will be useful to point out differences with respect to Robust Optimization

- For a more exhaustive introduction, I suggest the recent book:

A. Shapiro, D. Dentcheva, A. Ruszczynski,
“Lectures on Stochastic Programming: modeling and theory”
MPS-SIAM Series on Optimization, 2009

Uncertain problems – a remark

- ASSUMPTION: we have established that
  - our problem is uncertain
  - we must consider uncertainty

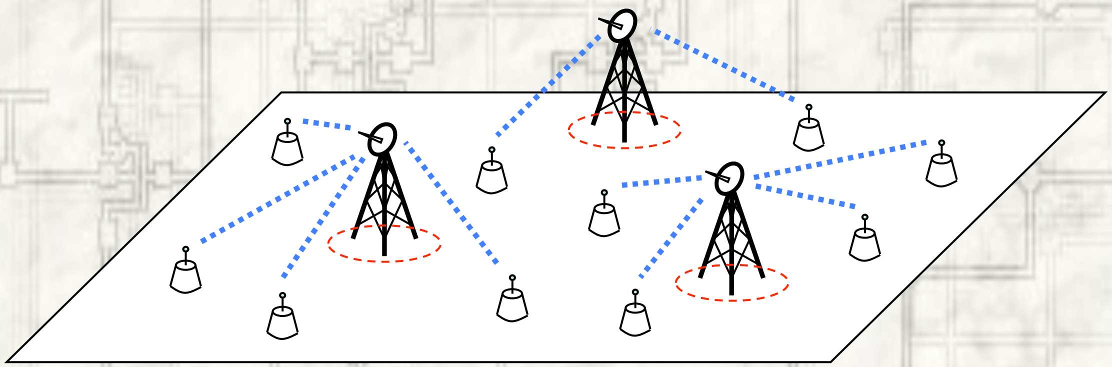
- We may tackle uncertainty by one of the methodogies sketched before

- If we know the distribution followed by the uncertainty, there could be the possibility to define a **(slightly) modified version** of the original problem

- I will illustrate this possibility by an application in Wireless Network Design

Wireless Networks

A Wireless Network can be essentially described as a **set of transmitters T** which provide for a telecommunication service to a **set of receivers R** located in a target area



- Every transmitter is characterized by a **set of parameters**

Positional (antenna height, geographical location)

Radio-electrical (e.g., power emission, frequency channel)

WIRELESS NETWORK DESIGN PROBLEM (WND)

→

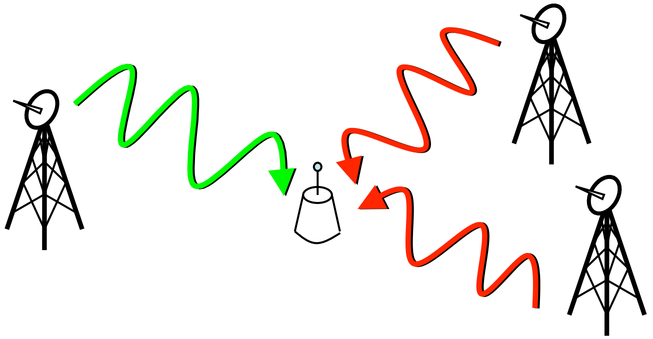
set the values of the parameters of each transmitter to maximize profit from service, while ensuring a minimum quality of service for each served receiver

Service coverage (1)

Every receiver r picks up signals from all the transmitters,

BUT:

- coverage is provided by a single transmitter, chosen as server of *r*
- all the other transmitters interfere the serving signal



If we introduce a continuous variable $0 \leq p_t \leq P^{\max}$ to represent power emission of transmitter t,

r is covered if the signal-to-interference ratio (SIR) is higher than a given threshold:

POWER RECEIVED FROM SERVER Tx

SUM OF POWER FROM INTERFERING Txs

$$\frac{a_{r\sigma} \cdot p_\sigma}{\sum_{t \in T \setminus \{\sigma\}} a_{rt} \cdot p_t} \geq \delta$$

COVERAGE THRESHOLD

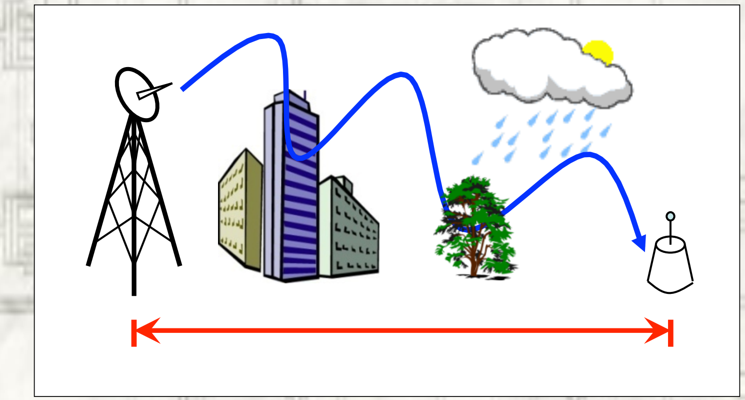
$$a_{r\sigma} \cdot p_\sigma - \delta \sum_{t \in T \setminus \{\sigma\}} a_{rt} \cdot p_t \geq 0$$

(SIR constraint)

Propagation and fading

A fading coefficient a_{rt} is usually computed through a propagation model and depends on several factors such as:

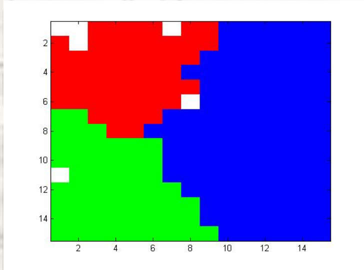
- the distance between t and r
- the presence of obstacles
- the weather

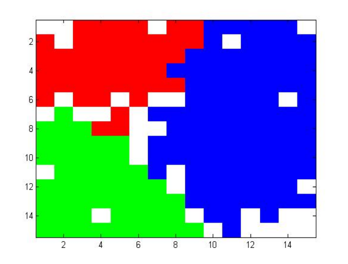


The fading coefficients are naturally subject to **uncertainty**

Neglecting uncertainty may lead to plans with **unexpected coverage holes**

EXPECTED COVERAGE





ACTUAL COVERAGE

Stochasticity of propagation (1)

What do network engineers actually consider to protect from signal uncertainty?

For each receiver to cover, they look at a probabilistic version of the SIR (signal-to-interference ratio):

$$SIR = \frac{U}{\sum_{k=1,\dots,K} I_k} \longrightarrow$$

Every signal S is a lognormal random variable

$$S \sim \text{Log-N}\,(m_S, \sigma)$$

However, a closed form for the summation of lognormal variables is not yet known.

so they must adopt one of the approximation proposed in literature

International Telecommunications Union (ITU) → k-LNM Method

- Sum Lognormal Vars = Lognormal var L
- $$m_L \;=\; \ln\left[\sum_{i=1}^{n} \exp\left(m_i + \frac{\sigma_i^2}{2}\right)\right] - \frac{\sigma_L^2}{2}$$
- $$\sigma_L \;=\; \ln\left\{1 - k \cdot \frac{\sum_{i=1}^{n}\left(1 - \exp \sigma_i^2\right)\cdot \exp\left(2m_i + \sigma_i^2\right)}{\left[\sum_{i=1}^{n} \exp\left(m_i + \frac{\sigma_i^2}{2}\right)\right]^2}\right\}$$

Stochasticity of propagation (2)

- Network engineers usually adopt a trial-and-error approach supported by simulation

COVERAGE PROBABILITY

$$P\left(\frac{U}{I} > \gamma\right) \geq p \quad \longrightarrow \quad \gamma \geq CDF^{-1}_{\frac{U}{I}}(1-p)$$

$$\frac{\ln \gamma - (m_U - m_I)}{\sqrt{2}\sqrt{\sigma_U^2 + \sigma_I^2}} \leq \operatorname{erf}^{-1}(1-2p)$$

- We can use the k-LNM method to get a constraint in the power variables of the transmitters, however the constraint is non-linear

- We can eliminate the non-linearity by making assumption on the deviation (strategy adopted in the design of the Italian DVB-T network in collaboration with AGCOM)

$$\sum_{i \in U} \tilde{a}_{ti} \cdot p_i - \sum_{i \in I} \tilde{a}_{ti} \cdot p_i \geq \delta'$$

ANYWAY, we have to check the validity of the solution and repair coverage errors if present

Intermediate remarks

- Most real-world optimization problems involve **uncertain data**, whose presence **cannot be neglected**
- **Many models** are available for **representing uncertain data** in optimization
- Until recent times, **Stochastic Optimization** has been the most used methodology for uncertain optimization
- **Robust Optimization** has emerged as a very competitive alternative to Stochastic Programming and is particularly appreciated by Professionals
- Sometimes there is the possibility that the uncertain problem may be "reformulated deterministically" exploiting problem-specific information about the uncertainty

Optimization under uncertainty

- Open question since the time of Dantzig [Management Science 1955]
- Many methodologies to deal with it proposed over the years

STOCHASTIC PROGRAMMING → probably the oldest and most studied approach to the question

Main features:

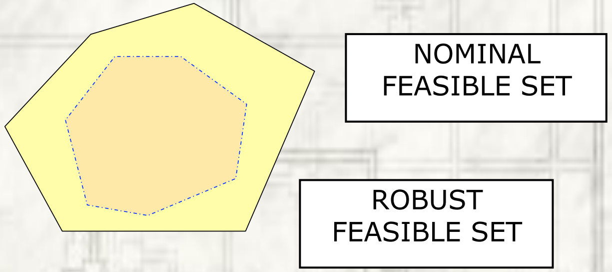
- find a solution that is feasible for (almost) all the realizations of the data
- optimize the expected value of the objective function

Drawbacks:

- precise knowledge of the uncertainty distribution is required
- (hard) very-large scale optimization problems
- solutions may still be infeasible

Robust Optimization

Uncertainty on coefficients is modeled as hard constraints that restrict the feasible set [Ben-Tal, Nemirovski 98, El-Ghaoui et. al. 97]



NOMINAL PROBLEM → Coefficients are uncertain!!! → ROBUST COUNTERPART

$$\max\ c'x$$
$$Ax \leq b$$
$$x \geq 0$$

$$a_{ij} = \bar{a}_{ij} + \delta_{ij}$$

ACTUAL VALUE

NOMINAL VALUE

DEVIATION

$$\max\ c'x$$
$$\tilde{A}x \leq b \quad \forall \tilde{A} \in \mathcal{A}$$
$$x \geq 0$$

- $\mathcal{A}$ is a subset of all the matrices allowed by deviations from nominal values
- “larger” $\mathcal{A}$ corresponds with higher risk adversion of the decision maker
- protection entails the so-called Price of Robustness

A breakthrough: the Bertsimas-Sim model

$$\max \sum_{j \in J} c_j \, x_j$$

$$\sum_{j \in J} a_{ij} \, x_j \leq b_i \qquad i \in I = \{1, \ldots, m\}$$

$$x_j \geq 0 \qquad j \in J = \{1, \ldots, n\}$$

Assumptions:

1) w.l.o.g. uncertainty just affects the coefficient matrix

2) the coefficients are independent random variables following an unknown **symmetric distribution over a symmetric range**

Deviation range: each coefficient a_{ij} assumes value in the symmetric range $a_{ij} \in [\bar{a}_{ij} - d_{ij}^{\max}, \bar{a}_{ij} + d_{ij}^{\max}]$

Row-wise uncertainty: for each constraint *i*, $\Gamma_i \in [0, n]$ specifies the max number of coefficients deviating from a_{ij}

ROBUST COUNTERPART
(NON-LINEAR)

$$\max \sum_{j \in J} c_j x_j$$

$$\sum_{j \in J} \bar{a}_{ij} x_j + DEV(x, \Gamma_i) \leq b_i \qquad \forall i \in I$$

$$x_j \geq 0 \qquad \forall j \in J$$

MAXIMUM DEVATION
OF CONSTRAINT i

The magic of duality

$$
\begin{aligned}
DEV(x,\Gamma_i) \;=\; \max\ & \sum_{j\in J} d_{ij}^{\max}\, x_j\ y_{ij} \\
& \sum_{j\in J} y_{ij} \le \Gamma_i \\
& y_{ij} \in \{0,1\} \qquad j \in J
\end{aligned}
$$

THE LINEAR RELAXATION HAS
THE SAME OPTIMAL VALUE

$$
\begin{aligned}
\max\ & \sum_{j\in J} d_{ij}^{\max}\, x_j\ y_{ij} \\
& \sum_{j\in J} y_{ij} \le \Gamma_i \\
& 0 \le y_{ij} \le 1 \qquad j \in J
\end{aligned}
$$

DEFINE THE
DUAL PROBLEM

$$
\begin{aligned}
\min\ & \Gamma_i\, w_i + \sum_{j\in J} z_{ij} \\
& w_i + z_{ij} \ge d_{ij}^{\max}\, x_j \qquad && j \in J \\
& w_i \ge 0 \\
& z_{ij} \ge 0 && j \in J
\end{aligned}
$$

$$
\begin{aligned}
\max\ & \sum_{j\in J} c_j\, x_j \\
& \sum_{j\in J} \bar{a}_{ij}\, x_j + \Gamma_i\, w_i + \sum_{j\in J} z_{ij} \le b_i \qquad && i \in I \\
& w_i + z_{ij} \ge d_{ij}^{\max}\, x_j && i \in I, j \in J \\
& w_i \ge 0 && i \in I \\
& z_{ij} \ge 0 && i \in I, j \in J \\
& x_j \ge 0 && j \in J
\end{aligned}
$$

ROBUST COUNTERPART [Bertsimas,Sim 04]
(LINEAR AND COMPACT)

Linearity requires:

- m n additional constraints
- m + m n additional variables

Other relevant results in Γ– Robustenss

0-1 COST UNCERTAIN LINEAR PROGRAM

$$\min\ c'\ x$$

$$x \in X \subseteq \{0,1\}^n$$

NON-NEGATIVE
UNCERTAIN COST VECTOR

$$c_j \in [\bar{c}_j, \bar{c}_j + d_j] \qquad j \in J$$

FEASIBLE SET
=
SUBSET OF ALL THE 0-1 VECTORS

- A robust optimal solution can be obtained by solving **n+1 nominal problems with modified cost vector c'**

PROBABILISTIC BOUND OF CONSTRAINT VIOLATION

- The robust optimal solution is **completely protected against at most** Γ_ι deviations occuring in constraint **i**
- This solution **may become infeasible** when more than Γ_i deviations occur
- Anyway, Bertsimas and Sim characterized a **bound on the probability** that the solution becomes infeasible

Robust Optimization VS Stochastic Programming

- **Soft probabilistic** VS **hard deterministic** constraints to represent uncertainty

  - need stochastic data
  - need to characterize the probability distribution
  - willingness to accept probabilistic guarantees

- **Rigid** budget of uncertainty (defined by the probability distribution)
  VS
  Flexible budget of uncertainty
  (shape the uncertainty set according to the decision maker's risk aversion)

- Computational tractability

KISS Bertsimas and Sim

- Mathematically elegant and accessible theory for dealing with uncertainty

- Starting point for many further theoretical developments
  (see the many subsequent papers mainly by Bertsimas and al. and Sim and al.)

- Notwithstanding the new developments, after ten years the model still remains a **central reference in applications**

Keep **I**t **S**imple
and **S**traightforward!

- Very plain and understandable uncertainty model
- Easily implementable
- Clear and direct control over robustness

Ideal robustness model for professionals and "more technical" research communities

It typically dissipates common questions like:

- Which is the sense of this model?
- How am I supposed to use this model?

Energy Offering for a Price-Taker (EnOff-PT)

PRICE-TAKER → producer that does not influence **market price**
(limited energy production)

The multi-unit offering problem can be decomposed into single-unit problems

For each generation unit of the producer :

Given:

- a planning horizon decomposed into a **set T of time periods**
- **the market price in each time period t**

We want to:

- choose the energy to offer in each time period in the market

So that:

- the total profit is maximized
- technical constraints of the units are satisfied (e.g., min up/down time, ramp limits)

A natural formulation for the EnOff-PT

RELEVANT FEATURES OF A GENERATION UNIT

$P^{\min}$	$P^{\max}$	(MIN and MAX ENERGY OUTPUT)
$R^{\nearrow}$	$R^{\searrow}$	(RAMP-UP and RAMP-DOWN RATE)
P^{SU}	P^{SD}	(MAX ENERGY OUTPUT AT START UP and BEFORE SHUT-DOWN)
U	D	(MIN UP and DOWN TIME)

DECISION VARIABLES

$p_t \geq 0$	$t \in T$	(ENERGY OUTPUT)
$u_t \in \{0,1\}$	$t \in T$	(STATUS ON/OFF)
$v_t \in \{0,1\}$	$t \in T$	(SWITCH ON)
$w_t \in \{0,1\}$	$t \in T$	(SWITCH OFF)

$$\max \sum_{t \in T} [\lambda_t p_t - c_t(p_t)]$$

PROFIT MAXIMIZATION
(REVENUE MINUS COSTS OF GENERATION AND START)

$$P^{\min} u_t \leq p_t \leq P^{\max} u_t \qquad t \in T$$

VARIABLE POWER BOUND

$$p_t \leq p_{t-1} + R^{\nearrow} u_t + \left(P^{SU} - R^{\nearrow}\right) v_t \qquad t \in T$$

$$p_t \geq p_{t-1} - R^{\searrow} u_{t-1} + \left(R^{\searrow} - P^{SD}\right) w_t \qquad t \in T$$

RAMP-UP AND –DOWN LIMITS

$$\sum_{\tau = t-U+1}^{t} v_\tau \leq u_t \qquad t \in \{U+1, \ldots, |T|\}$$

$$\sum_{\tau = t-D+1}^{t} w_\tau \leq 1 - u_t \qquad t \in \{D+1, \ldots, |T|\}$$

MIN UP AND DOWN TIME
(STRONG VERSION)

$$w_t = v_t + u_{t-1} - u_t \qquad t \in T$$

LINKING OF VARIABLES

Price uncertainty in the UC-PT

Major challenge for the price-taker → the hourly prices are not kwown exactly when the problem is solved (MARKET PRICE UNCERTAINTY)

The price-taker could solve its commitment problem using estimates of prices that he trusts…

…BUT he would risk a lot!

price estimates (sensibly) higher than the real market price → OVERPRODUCTION → LOSSES

price estimates (sensibly) lower than the real market price → UNDERPRODUCTION → REDUCED PROFIT

We cannot neglect price uncertainty and we must tackle it!

WHAT CAN WE DO?

Resuming the Bertsimas-Sim model (BS)

$$\max \sum_{j\in J} c_j\, x_j$$

$$\sum_{j\in J} a_{ij}\, x_j \le b_i \qquad i \in I = \{1,\dots,m\}$$

$$x_j \ge 0 \qquad\qquad j \in J = \{1,\dots,n\}$$

Assumptions:

1) w.l.o.g. uncertainty just affects the coefficient matrix

2) the coefficients are independent random variables following an unknown **symmetric distribution over a symmetric range**

Deviation range: each coefficient a_{ij} assumes value in the symmetric range $a_{ij} \in [\bar{a}_{ij} - d_{ij}^{\max}, \bar{a}_{ij} + d_{ij}^{\max}]$

Row-wise uncertainty: for each constraint *i*, $\Gamma_i \in [0, n]$ specifies the max number of coefficients deviating from a_{ij}

ROBUST COUNTERPART
(NON-LINEAR)

→

ROBUST COUNTERPART [Bertsimas, Sim 04]
(LINEAR AND COMPACT)

$$\max \sum_{j\in J} c_j x_j$$

$$\sum_{j\in J} \bar{a}_{ij} x_j \boxed{+ DEV(x, \Gamma_i)} \le b_i \qquad \forall i \in I$$

$$x_j \ge 0 \qquad \forall j \in J$$

$$\max \sum_{j\in J} c_j x_j$$

$$\sum_{j\in J} \bar{a}_{ij} x_j \boxed{+ \Gamma_i w_i + \sum_{j\in J} z_{ij}} \le b_i \qquad \forall i \in I$$

$$\boxed{\begin{array}{l} w_i + z_{ij} \ge d_i^{\max} j x_j \\ z_{ij} \ge 0 \\ w_i \ge 0 \end{array}} \qquad \begin{array}{l} \forall i \in I, j \in J \\ \forall i \in I, j \in J \\ \forall i \in I \end{array}$$

$$x_j \ge 0 \qquad \forall j \in J$$

Γ-Robustness for the price-uncertain UC-PT

Remarks about the UC-PT:

- data uncertainty only affects the objective function (uncertain price coefficients)

Γ-Robust Counterpart:

Given:

- the **nominal price** in each period λ_t^{NOM}
- the **worst deviation of price** w.r.t. the nominal price in each hour d_t
- the **number** $\Gamma > 0$ of price deviations for which protection is required

The robust counterpart is:

$$\max \sum_{t \in T} \left[\lambda_t^{\text{NOM}} p_t - c_t(p_t)\right] \boxed{- \Gamma z - \sum_{t \in T} q_t}$$

$$\left.\begin{array}{ll} z + q_t \geq d_t p_t & t \in T \\ z \geq 0 & \\ q_t \geq 0 & t \in T \end{array}\right\} \longrightarrow \boxed{\text{ADDITIONAL VARIABLES AND CONSTRAINTS FROM ROBUST DUALIZATION}}$$

$$\begin{array}{ll} p_t \in P_t & t \in T \end{array} \longrightarrow \boxed{\text{FEASIBLE ENERGY PRODUCTION SET}}$$

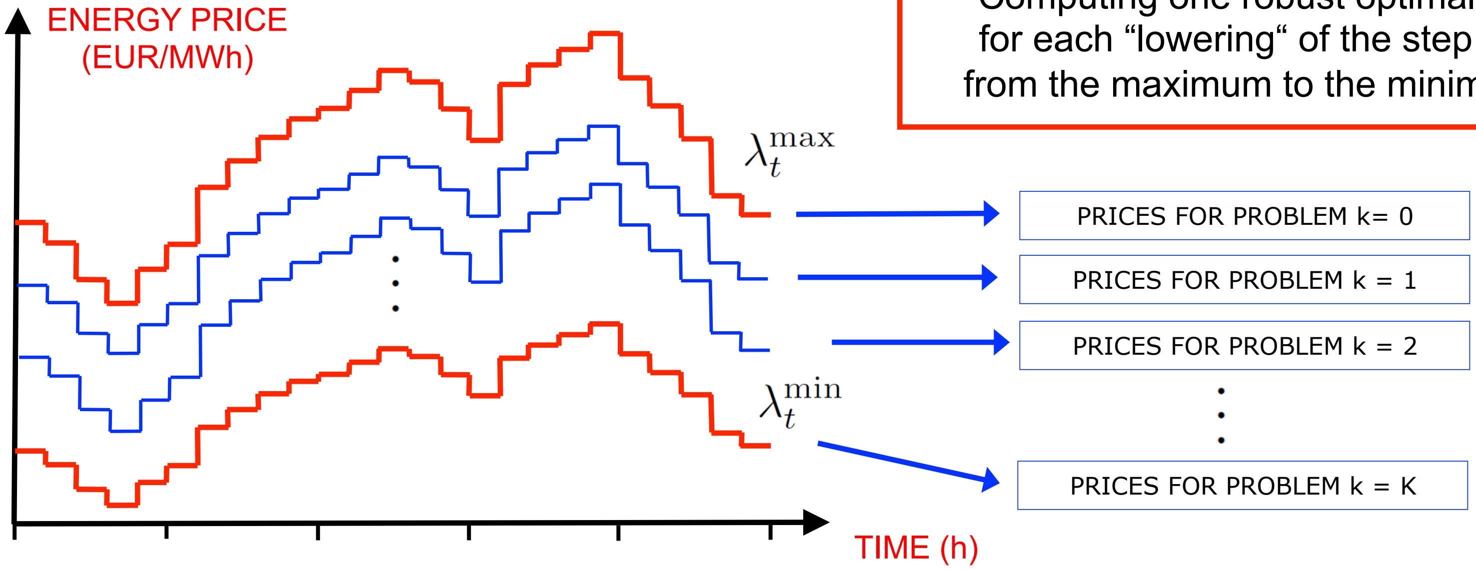
The Baringo-Conejo approach (1)

Highly cited work proposing a method for building energy offering curves for a price taker (2011)

Main steps:

- identify the **overall range of prices** in each period - maximum and miminum prices $\lambda_t^{\min}$ $\lambda_t^{\max}$
- define an **elementary price shortfall** $\delta \cdot (\lambda_t^{\max} - \lambda_t^{\min})$ with $0 < \delta < 1$
- Solve k = 0, 1, …, K Γ-Robust Counterpart where in each period
  - the **nominal price** is the **maximum price** of the range $\lambda_t^{\text{NOM}} = \lambda_t^{\max}$
  - the **worst deviation** is **k-times the elementary price shortfall** $d_t = k \cdot \delta \cdot (\lambda_t^{\max} - \lambda_t^{\min})$
  - **Γ = |T|** ⟶ FULL PROTECTION!

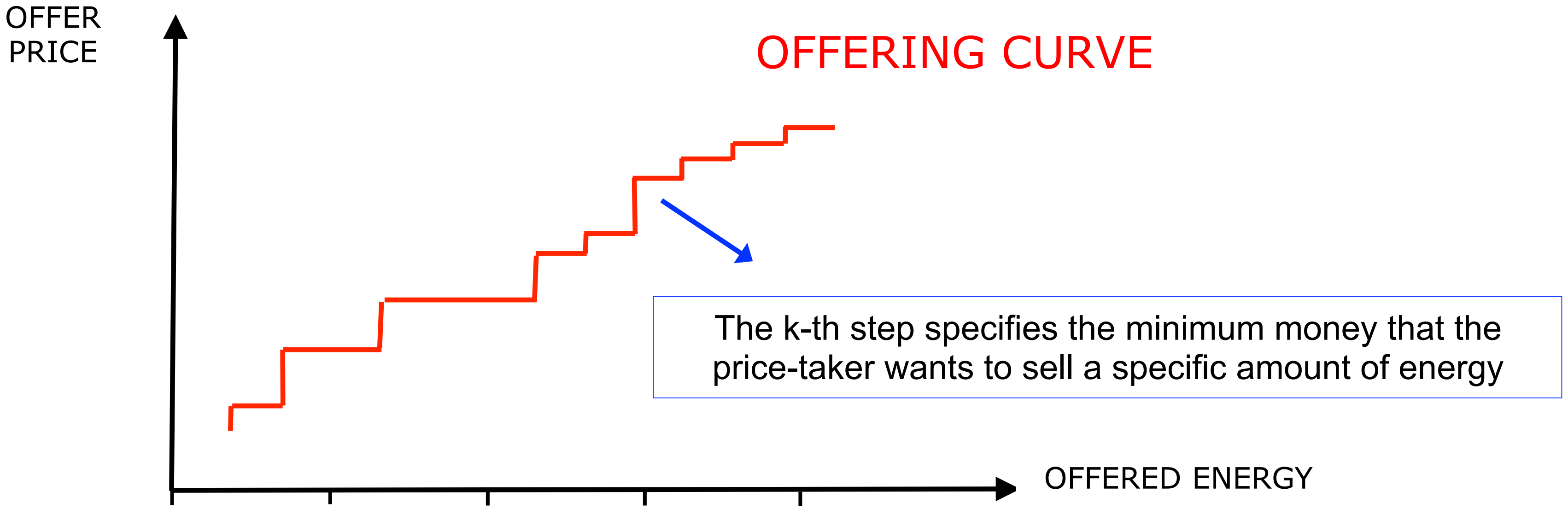
Computing one robust optimal solution
for each "lowering" of the step function
from the maximum to the minimum price



The Baringo-Conejo approach (2)

- For each step function k, we obtain a robust optimal solution
- The **robust optimal solutions are merged** to build one **energy offering curve** for each time period
- For each time period:

STEP FUNCTION k → MARKET PRICE k → OFFERED ENERGY k



The k-th step specifies the minimum money that the price-taker wants to sell a specific amount of energy

- The offering curve built for each time period are submitted to the Energy Exchange

The Baringo-Conejo approach – our critique

The approach presents **several issues** that have **NOT** been **pointed out until our work**

ISSUE 1: definition of offering curves that break market rules

An offering curve is built considering a high number of intermediate prices between the maximum and minimum prices (100 prices in experimentals tests)

→ **Violation of the limit on the number of steps of a curve** imposed by market rules

ISSUE 2: risk of non-acceptance

The offering curves **risk to be NOT accepted** in the market (minimum price asked for selling) → **BIG LOSSES**

ISSUE 3: compromised optimality and feasibility

The **offering curves defined merging distinct optimal robust solutions** obtained for different assumptions on the prices

→ optimality of energy production is compromised!

→ accepted portion of curves may result infeasible (e.g., violation of ramp constraints)

ISSUE 4: unnecessarily complex robust counterpart

The approach imposes full protection (worst price in each period)

→ it is not necessary to define the Γ-Robustness counterpart of increased dimension

Our revised approach based on Γ-Robustness (1)

OUR OBJECTIVES:

- (dramatically) increasing the chances that our energy offers are accepted
- defining robust solutions following the real spirit of Γ-Robustness (full protection is bad!)

BASIC FEATURES OF OUR STRATEGY:

- we do not compete on price and **all our selling offers are at zero price**

  → our offers are automatically accepted (≤ market price!)

- from historical market price data, we derive
  - the **nominal value** equals the **average price** over the past observations
  - the **worst deviation** is identified by **excluding the worst M observations** in a way that **better fits the practice of power system professionals**

- **we exclude extreme unlikely price shortfalls** and we show that partial protection grants (much) higher profits

Computational tests

- Tests on **45 realistic instances**:
  - **15 power plants** located in **3 distinct Italian price-zone**
  - **24 time periods (= hours in one day)**
  - **3 percentages of exclusions** of worst price observations **(0, 10, 20 %)**
- Experiments on a Windows machine with Intel 2 Duo-3.16 GHz processor and 8 GB of RAM
- Robust model coded in C/C++ interfaced through Concert Technology with CPLEX 12.5.1

Historical data and test period construction:

- For each hour:
  - we consider the **prices observed** in the price zone **in a time window of 4 weeks**
  - from these prices, we derive the nominal value and the max deviation of the uncertain price
- We compute the robust optimal solution **for each Γ=0 (=no protection), 1, 2, …, 24 (= full protection)**
- We test the performance of the computed robust optimal solution in the **week following the 4 weeks of the construction set**
- The 4-week time window is shifted through the entire year with steps of 1 week providing **24 evaluation periods**

Computational results

Unit ID	%Excluded	Γ Best w.r.t. $\Gamma = 0$		Γ Best w.r.t. $\Gamma = 24$	
		$\Delta\pi(EUR)$	$\Delta\pi\%$	$\Delta\pi(EUR)$	$\Delta\pi\%$
U1	0	+ 40399	+ 5.75	+ 213730	+ 40.45
	10	+ 44350	+ 6.32	+ 183161	+ 32.54
	20	+ 41921	+ 5.97	+ 152608	+ 25.82
U2	0	+ 23394	+ 5.00	+ 333543	+ 212.07
	10	+ 46063	+ 9.85	+ 234371	+ 83.96
	20	+ 42071	+ 9.00	+ 218607	+ 75.15
U3	0	- 1383	- 0.02	+ 1984511	+ 47.59
	10	+ 88980	+ 1.44	+ 1031465	+ 19.78
	20	+ 105253	+ 1.70	+ 627146	+ 11.13
U4	0	+ 43246	+ 6.27	+ 255124	+ 53.50
	10	+ 57386	+ 8.33	+ 181356	+ 32.11
	20	+ 51614	+ 7.49	+ 148634	+ 25.12
U5	0	+ 15454	+ 3.57	+ 340567	+ 319.00
	10	+ 45327	+ 10.49	+ 240506	+ 101.61
	20	+ 45331	+ 10.49	+ 199406	+ 71.78
U6	0	+ 14273	+ 5.30	+ 2030185	+ 44.87
	10	+ 91766	+ 10.58	+ 1117143	+ 20.25
	20	+ 152707	+ 11.77	+ 675172	+ 11.22
U7	0	+ 307690	+ 5.73	+ 1312795	+ 30.13
	10	+ 268508	+ 5.00	+ 909989	+ 19.28
	20	+ 195207	+ 3.64	+ 792081	+ 16.62

Generation units of increasing capacity

DIFFERENCE OF TOTAL PROFIT
(IN EUR, SUM OF 24 TEST PERIODS)
best protection - no protection **best protection - full protection**

Unit ID	%Excluded	Γ Best w.r.t. $\Gamma = 0$		Γ Best w.r.t. $\Gamma = 24$	
		$\Delta\pi(EUR)$	$\Delta\pi\%$	$\Delta\pi(EUR)$	$\Delta\pi\%$
U8	0	+ 465184	+ 12.62	+ 1295788	+ 45.41
	10	+ 492568	+ 13.37	+ 1052152	+ 33.68
	20	+ 387575	+ 10.52	+ 888433	+ 27.91
U9	0	- 179096	- 0.54	+ 8606255	+ 35.47
	10	+ 249549	+ 0.75	+ 4586552	+ 15.97
	20	+ 253871	+ 0.76	+ 3008976	+ 9.93
U10	0	+ 579613	+ 11.70	+ 1711145	+ 44.78
	10	+ 662118	+ 13.36	+ 870430	+ 18.34
	20	+ 502539	+ 10.14	+ 594417	+ 12.22
U11	0	+ 612087	+ 19.62	+ 2471071	+ 196.02
	10	+ 465014	+ 14.90	+ 1270815	+ 54.92
	20	+ 534279	+ 17.12	+ 788373	+ 27.51
U12	0	+ 23935	+ 19.62	+ 12397479	+ 49.27
	10	+ 409219	+ 14.90	+ 4751121	+ 14.31
	20	+ 438452	+ 17.12	+ 3506377	+ 10.17
U13	0	+ 111221	+ 0.34	+ 8244775	+ 33.56
	10	+ 421916	+ 1.29	+ 8244775	+ 16.04
	20	+ 479952	+ 1.46	+ 1809060	+ 5.76
U14	0	+ 231532	+ 6.14	+ 2103904	+ 111.02
	10	+ 391966	+ 10.40	+ 1184916	+ 39.83
	20	+ 485693	+ 12.89	+ 494986	+ 13.17
U15	0	- 76685	- 0.21	+ 11622838	+ 49.54
	10	+ 524423	+ 1.49	+ 5459972	+ 18.06
	20	+ 442969	+ 1.25	+ 3175311	+ 9.79

In almost all cases we can:

- **greatly increase the profit** w.r.t. a practice that we observed among professionals (average price)
- **dramatically increase the profit** w.r.t. full protection

Some concluding remarks

- **World is stochastic** and most of real-world optimization problems involve **uncertain data**
- Uncertain optimization problems can be really tricky
- Many models are available for representing uncertain data in optimization problems
- No model dominates the others from a theoretical point of view…
- …but **Robust Optimization** is itself as way to **model and actually solve real-world problems** (and Professionals like it! - deterministic protection and accessibility)
- the **Bertsimas-Sim model for Robust Optimization** is still a central reference and is used in many (practical) studies also outside the Mathematical Programming community

Thanks for your attention!

For additional discussions and references I am at your disposal

Some useful References

- A. Shapiro, D. Dentchevea, A. Ruszczynski,
  Lectures on Stochastic Programming: modeling and theory
  Book freely available at:
  http://www2.isye.gatech.edu/people/faculty/Alex_Shapiro/SPbook.pdf

- D. Bertsimas, D. Brown, C. Caramanis,
  Theory and Applications of Robust Optimization
  Survey freely available at:
  http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.259.8234&rep=rep1&type=pdf

- F. D‘Andreagiovanni, G. Felici,
  Revisiting the use of Robust Optimization for optimal energy offering under price uncertainty
  Paper freely available at: https://arxiv.org/pdf/1601.01728.pdf