Noname manuscript No.
(will be inserted by the editor)

Sergio V. Bruno · Claudia Sagastizábal

Optimization of Real Asset Portfolio using a Coherent Risk Measure: Application to Oil and Energy Industries

Version from 22nd July 2010

Abstract We consider the problem of optimally determining an investment portfolio for an energy company owning a network of gas pipelines, and in charge of purchasing, selling and distributing gas. We propose a two stage stochastic investment model which hedges risk by means of Conditional Value at Risk constraints. The model, solved by a decomposition method, is assessed on a real-life case, of a Brazilian integrated company that operates on the oil, gas and energy sectors.

Keywords Portfolio Optimization · Real Assets · CVaR · Decomposition methods

1 Introduction

In many industries, investment is part of the most important planning activities. In the past decades, mathematical programming models have been widely used in capacity planning and facility location to support investment decisions. Such initial techniques evolved to the use of enterprise portfolio management, very common in the energy industry. Nowadays, risk management is one of the top priorities on the planning processes of companies. As a consequence, state of the art planning models are stochastic, and usually consider some kind of risk measure as well as financial instruments to hedge the investment portfolio.

We consider a portfolio management problem posed by Petrobras, an energy company, owning a network of gas pipelines in Brazil. Petrobras not only is re-

Sergio V. Bruno
Petrobras, Rio de Janeiro, Brazil
E-mail: svbbruno@gmail.com

Claudia Sagastizábal
CEPEL, Electric Energy Research Center, Eletrobrás Group
On leave from INRIA Rocquencourt, France.
Research partially supported by CNPq Grant No. 300345/2008-9, FaPERJ and PRONEX-Optimization E-mail: sagastiz@impa.br

sponsible for the purchase, sale, and transportation of gas nationwide, but also is the major agent of the energy industry in Brazil. Its assets include oil and gas fields, refineries, thermo-electrical power plants, natural gas processing plants, and several oil pipelines, gas pipelines, re-gasification terminals[1]. The natural gas distribution, in particular, has very specific characteristics, shown in Figure 1, containing a schematic representation of the gas network operated by Petrobras.

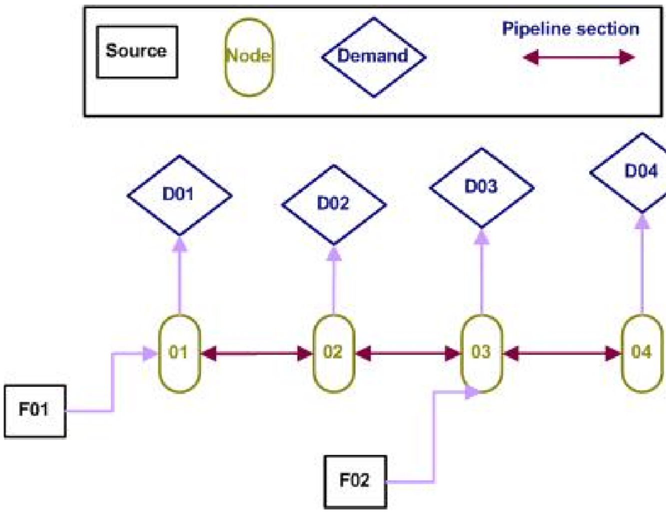

Fig. 1 Schematic representation of Petrobras gas network

Gas supply in Brazil is made up mostly by gas produced nationally by Petrobras and its partners (64%) and, in a lesser proportion (36%), by gas imported by Petrobras from other countries. This gas is currently imported from Bolivia, through the Brazil-Bolivia pipeline, but new re-gasification terminals in Rio de Janeiro and the Northeast have been recently built. These terminals allow the purchase of imported gas in a more flexible manner, through the acquisition of Liquid Natural Gas (LNG) from Africa, for example.

The demand for gas is composed of three major groups. The demand of the Petrobras refineries for fuel and inputs in their units (30%), the demand of local distribution companies, which serve to industrial, residential, and vehicular natural gas (51%), and thermoelectric power plants (19%), which use gas as fuel to generate electricity[2]. Since the beginning of the 90's, various national policies encouraged the use of natural gas, making the natural gas participation in the Brazilian energy matrix increase from 5.4% in 2000 to 9,3% in 2005, according to [5]. As shown in Figure 2, the gas pipeline followed this increase with various expansions of the gas network capacity.

[1] For more details, see http://www.petrobras.com

[2] Source: Balanço Energético Nacional, Ministério de Minas e Energia, Brazil

Fig. 2 Current Gas network and pipes under construction in Brazil

In view of the above, one of the main long-term planning decisions of the company is to determine its investment portfolio. To meet the long-term demand, the company is expected to invest in the expansion of its pipelines network, in order to increase the gas transportation capacity, as described in Figure 3.

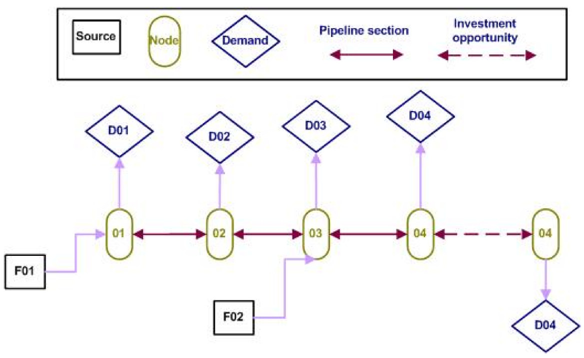

Fig. 3 Schematic representation of the investment opportunity

However, determining new investments is difficult, due to the complexity of the problem, in which there are several gas suppliers and customers sharing the same resources and also because the inclusion of an investment in the portfolio directly affects the economic viability of another project. For this reason, investment models in the company are usually deterministic. Notwithstanding, since the long-term demand is highly uncertain, decision makers often are forced to resort to heuristic methods, such as sensitivity studies, to somehow foresee the impact of deviations in the forecasted (deterministic) demand.

The objective of this work is to integrate uncertainty in the investment problem for real assets, that is, investment in physical assets that have some degree of irreversibility. The assets that compose the investment portfolio are those typically considered by the company, such as pipelines, processing units, or LNG terminals. We develop a model to support the choice of investment projects, given certain optimality criteria and operational and financial constraints, dealing with uncertainty by means of Conditional Value at Risk (CVaR) constraints.

It will be shown in this paper that there are significant potential gains in using stochastic models in the resolution of this problem. Besides, a stochastic framework offers the opportunity to manage various financial risks inherent to the investment decision.

This work is organized as follows. Section 2 describes the importance of modeling uncertainty in certain decision frameworks, in particular when modeling mathematical programs for real asset optimal allocation. We explain how oversimplified models can fall under the so-called "flaw of averages" and we propose the use of a risk measure to avoid undesired outcomes in unfavorable scenarios. Sections 3 and 4 present, respectively, the risk-averse planning model, and some decomposition methods for the solution of the model in reasonable CPU time. Section 5 contains some numerical experience assessing the proposed methodology. Finally, some concluding remarks are given in Section 6.

2 The importance of properly modeling uncertainty

As mentioned, traditional investment portfolio models in Petrobras are deterministic and use the average demand of the various scenarios to represent future demand. This simplification makes the resulting optimization problem deterministic and easy to solve, since it is just a medium size linear program. But there is a negative impact on the quality of the decision obtained with this deterministic-like model.

2.1 The flaw of averages

Forcing a deterministic framework may lead to the so-called *flaw of averages*. Namely, when taking the average as a proxy for the probabilistic distribution of demand, the demand variability disappears from the modeling. When scenarios (representing the demand distribution) have a large variability, neglecting such deviations can severely degrade the validity of the model. An illustration of this issue, highlighted by Sam Savage in [8], is called the "drunk on the highway". On this example, Savage considers the state of a drunk, wandering back and forth on a busy highway. His average position is the centerline of the highway. Clearly, the state of the drunk at his average position is “alive”, while the average state of the drunk is “dead”.

In a more formal setting, the difference between the two, opposed, drunk states above is explained by the fact of wrongly interchanging the expected value function $E[\cdot]$ with a certain function $f(\cdot)$. Suppose that $f(\cdot)$ denotes the state of the drunk (dead or alive), given as a function of a random variable z, his position in the highway. Then the wrong approach, which amounts to take $f(E[z])$, would declare the drunk alive. The right approach, instead, would consider the average of the outcome, $E[f(z)]$, and declare the drunk dead.

Petrobras previous models used average demand to decide on investments in the network. More precisely, assume uncertainty is represented by a finite set of scenarios $w = 1, \ldots, N_w$, each one having probability of occurrence p_w. Suppose, in addition, that decision variables are split into deterministic and uncertain ones, denoted by x and y_w, respectively (x represents the investment decision on certain asset, while y_w is related to operational decisions, for a given realization w of the uncertainty). Then the average model used by Petrobras corresponds to replacing the stochastic linear program with uncertain right hand side:

$$\begin{cases} \min\limits_{\{x,(y_w)_1^{N_w}\}} & c^\top x + \sum\limits_{w=1}^{N_w} p_w q^\top y_w \\ \text{s.t.} & Ax = b \\ & Tx + W y_w = h_w\,, \quad w = 1, \ldots, N_w \\ & x \geq 0\,, y_w \geq 0\,, \quad\ \ w = 1, \ldots, N_w \end{cases} \tag{1}$$

by the deterministic linear program

$$\begin{cases} \min\limits_{\{x,\bar{y}\}} c^\top x + q^\top \bar{y} \\ \text{s.t.}\ \ Ax = b \\ \qquad Tx + W\bar{y} = E_w[h] \\ \qquad x \geq 0\,, \bar{y} \geq 0\,. \end{cases} \tag{2}$$

2.2 Stochastic model

To replace (2), in this work we adopt a *two stage stochastic programming* point of view, with first stage deterministic investment variables x and uncertain second stage operational variables $y_w\,, w = 1, \ldots, N_w$.

As the investment decision in real assets is of long term, a two stage approach is preferred by the decision maker. A multi-stage, “wait-and-see”, strategy, postponing the investment decision until uncertainty is revealed, is not interesting, since it would correspond to start building a new pipeline only after the gas demand becomes known (demand would then be left unsatisfied for all the years it takes to build the pipeline).

Since we consider a long-term problem (over 5 years or more), uncertainty appears at various levels. There are political and economic uncertainties, such as purchase and selling prices, or taxes. There are also technical uncertainties, related to the operational nature of the problem, such as the quality of gas received from a source, the volume available for supply and the volume of demand. For the case of interest, thermo-electric demand is the most relevant source of uncertainty and volatility, so only this uncertainty is considered in the stochastic model.

The thermo-electric demand is provided as information given by a program, which typically produces data with $N_w \in [80, 200]$ future scenarios. These scenarios are organized in a tree with the "fan" structure from Figure 4. The fan format represents the fact that there is a first decision stage, in which the investment portfolio is determined. In Figure 4, the rectangle represents the investment decision

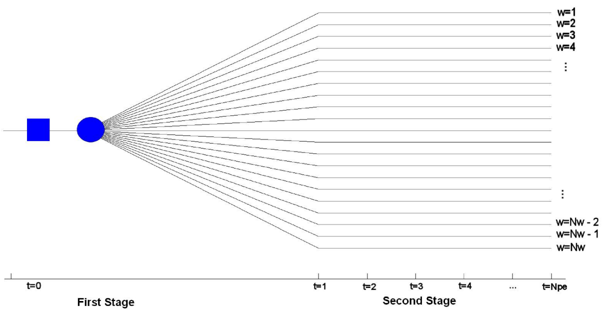

Fig. 4 Scenarios in a "fan" structure.

and the circle represents the subsequent realization of a demand scenario. There is also a second stage, consisting of several (N_{pe}) subsequent periods of time. During the second stage, previously decided investments, x, start to operate and the best operational decisions, y_w, are taken considering the realization of scenario w (among N_w possible).

Uncertainty is handled in (1) in two different manners. The expected function, $c^\top x + E_w[q^\top y_w]$, is used in the objective. In addition, to hedge against volatility, we introduce a coherent risk measure, the Conditional Value at Risk [6] and [7], to define acceptable levels of risk of the resulting portfolios.

Specifically, given a loss distribution function $f(x,w)$, for a portfolio x subject to uncertainties w, and given a confidence level $\beta \in [0,1]$, the CVaR_β, Conditional Value at Risk of level β, also known as *Expected Shortfall*, is the average of the quantiles $Q_{\psi(x)}(p)$ of the cumulative distribution function associated to $f(x,w)$:

$$CVaR_\beta(x) := \frac{1}{1-\beta} \int_0^{1-\beta} Q_{\psi(x)}(s) ds \,. \tag{3}$$

Artzner et al [1] proposed a class of risk measures based in axiomatic properties, know as coherent risk measures. These measures have the properties of Sub-Additivity, Sensibility, Positive Homogeneity and Translation Invariance. The $CVaR_\beta$ is a coherent risk measure with great popularity. One of the most desired properties of these measures is the Sub-Additivity: $CVaR_\beta(x+x') \leq CVaR_\beta(x) + CVaR_\beta(x')$, representing the fact that portfolio diversification mitigates risk.

Another advantage of this risk measure is in its implementation, not only $CVaR_\beta$ generates a convex feasible region, thus granting global optimal solutions, but also can be represented by using linear relations, yielding a linear programming formulation. Linear programming theory is well understood and there are efficient available solvers, both free and commercial. The linear programming representation can be used if underlying distributions are discrete or if we resort to some sampling technique. For this reason, the $CVaR_\beta$ measure is suitable for stochastic optimization.

The linear formulation proposed by [4] is useful if the probability distribution is discrete or is sampled in a finite number of N_w scenarios, like in our application. In these cases, the right hand side in (3) is well approximated by the function

$$\tilde{F_\beta}(x,a) = a + \frac{1}{1-\beta} \sum_{w=1}^{N_w} p_w [f(x,w) - a]^+ , \tag{4}$$

where $[\cdot]^+ = \max(\cdot, 0)$ denotes the positive-part function.

If, in addition, the loss function $f(x,w)$ is linear with respect to x, then the function $\tilde{F_\beta}(x,a)$ is convex and piecewise linear. Such is the case in (1), taking as loss function $f(x,w) = c^\top x + q^\top y_w$. In this case, having a parameter ρ_β, determining the minimum value for the average return of the worst β scenarios, adding a constraint of the form $F_\beta(x,a) \leq \rho_\beta$ in (1) yields the following piecewiselinear CVaR-model:

$$\left\{ \begin{array}{lll} \min\limits_{\{x,a,(y_w)_1^{N_w}\}} & c^\top x + \sum\limits_{w=1}^{N_w} p_w q^\top y_w & \\ \text{s.t.} & Ax = b\,, & \\ & Tx + Wy_w = h_w\,, & w = 1,\ldots,N_w \\ & a + \frac{1}{1-\beta} \sum\limits_{w=1}^{N_w} p_w [c^\top x + q^\top y_w - a]^+ \leq \rho_\beta\,, & \\ & x \geq 0\,, y_w \geq 0\,, & w = 1,\ldots,N_w\,. \end{array} \right. \tag{5}$$

An explicit representation of (5) as a two-stage linear program, using auxiliary variables z_w to express the positive-part function, that is,

$$z_w = [c^\top x + q^\top y_w - a]^+ \iff z_w \geq c^\top x + q^\top y_w - a \text{ and } z_w \geq 0\,,$$

is considered in the next section.

3 The optimal portfolio problem

The objective of the proposed model is to maximize the expected value of the company portfolio, considering gas commercialization over an horizon of at least 5 years. The main constraints are related to gas offer and demand and network capacity. For convenience, we will represent the model as a minimization problem, as in (5):

$$
\begin{cases}
\min\limits_{\{x,a,(y_w,z_w)_1^{N_w}\}} & c^\top x + \sum\limits_{w=1}^{N_w} p_w q^\top y_w \\
\text{s.t.} & Ax = b\,, & \text{(a)} \\
 & B_w y_w = d_w\,, w = 1,\dots,N_w\,, & \text{(b)} \\
 & Tx + Wy_w = h_w\,, w = 1,\dots,N_w\,, & \text{(c)} \\
 & a + \dfrac{1}{1-\beta} \sum\limits_{w=1}^{N_w} p_w z_w \leq \rho_\beta\,, & \text{(d)} \\
 & z_w \geq c^\top x + q^\top y_w - a\,, w = 1,\dots,N_w & \text{(e)} \\
 & x \geq 0\,, y_w \geq 0\,, z_w \geq 0\,, w = 1,\dots,N_w\,.
\end{cases}
\tag{6}
$$

where:

- x represents the set of investment variables. There are m assets and $N_{pe}+1$ periods on the time horizon. Each variable $x_{i,j}$ represents the proportion of invested capital on the i-th asset, on the j-th period, so $0 \leq x_{i,j} \leq 1$, for all $i = 1,\dots,m$ and $j = 0,\dots,N_{pe}$.
- c is the investment cost vector, of dimension $m(N_{pe}+1)$.
- y_w contains several variables representing gas flow, amount of gas acquired by a source, amount of gas delivered to a client and other operational variables referring to the scenario w.
- $q^\top y_w$ represents net operation costs on scenario w. Cash flows are adjusted to their present values.
- The investment constraints are represented by (6) (a).
- h_w, d_w is, respectively, the demand and supply data for scenario w.
- $p_w \in [0,1]$ is the probability of occurrence of scenario w ($\sum_{w=1}^{N_w} p_w = 1$).
- ρ_β is the required minimum value for the average return of the worst β scenarios, in monetary units.
- a is an auxiliary, unconstrained, scalar variable, used to define the risk constraints. The approach guarantees that, under reasonable assumptions, the optimal value of variable a equals the Value at Risk of level β of the portfolio.
- Relations (6) (b) and (c) describe the second stage constraints of the problem. These constraints are related to physical, financial, and economic business rules.
- Constraints (6) (d) and (e) represent the $CVaR_\beta$ constraint, noting that (6) (d) is set over the sum of the net operational costs and investment costs.

4 Decomposition methods

Following [2], we will refer to the direct solution of our investment problem (6) as *solving the equivalent deterministic* problem. However, for a large number of scenarios, the equivalent deterministic problem becomes untractable. In this case, decomposition approaches must be applied, taking into account the structure of the constraints (the objective function is separable on the variables). We see that some constraints couple deterministic variables, $X := (x,a)$, and stochastic ones, $Y_w := (y_w, z_w)$ for $w = 1,\dots,N_w$:

- Stagewise coupling: (6) (c) couples the first and second stage variables. This coupling can be treated using Benders decomposition [3, Ch. 11.1.3].
- Scenario-wise coupling: in (6) (d), the first stage variable a is coupled with all the scenarios. This coupling needs to separate individual scenarios, by price decomposition [3, Ch. 11.1.1].

We begin by writing (6) as a bilevel optimization problem, taking as first stage variables the deterministic ones, $X = (x,a)$, letting the stochastic variables $Y_w := (y_w, z_w)\,, w = 1,\dots,N_w$ for the second stage.

$$\begin{cases} \min\limits_{\{x,a\}} \; c^\top x + V(x,a) \\ \text{ s.t. } \; Ax = b\,, \\ \qquad 0 \le x \le 1\,. \end{cases} \tag{7}$$

The function $V(x,a)$ is the value function of the second level problem:

$$V(x,a) := \begin{cases} \min\limits_{\{(y_w,z_w)_1^{N_w}\}} & \sum\limits_{w=1}^{N_w} p_w q^\top y_w & \\ \text{s.t.} & B_w y_w = d_w\,, & w = 1,\dots,N_w \\ & W y_w = h_w \; - Tx\,, & w = 1,\dots,N_w \\ & \frac{1}{1-\beta} \sum\limits_{w=1}^{N_w} p_w z_w \le \rho_\beta \; - a\,, & \text{(d)} \\ & z_w - q^\top y_w \ge c^\top x - a\,, & w = 1,\dots,N_w \\ & y_w \ge 0\,, z_w \ge 0\,, & w = 1,\dots,N_w\,. \end{cases} \tag{8}$$

In order to better emphasize the dependence of the constraints on the first level variables x and a, in the feasible set of (8) we gathered on the left hand side those terms depending on the second stage variables only. Accordingly, first stage variables appear always as the rightmost terms in the corresponding constraints.

4.1 Benders-like decomposition

Like (6)(d), the $CVaR_\beta$ constraint (d) in (8), couples all scenarios. This is the reason why we consider a different splitting between first and second stage variables, including z_w in the first stage of decision: $X := (x, a, z_w, w = 1, \dots N_w)$. We obtain, instead of (7), a first level problem of the form

$$\begin{cases} \min\limits_{\{x,a,(z_w)_1^{N_w}\}} & c^\top x + \sum\limits_{w=1}^{N_w} p_w V_w(x,a,z_w) \\ \text{s.t.} & Ax = b\,, \\ & a + \frac{1}{1-\beta} \sum\limits_{w=1}^{N_w} p_w z_w \le \rho_\beta\,, \\ & 0 \le x \le 1\,, z_w \ge 0\,, \qquad w = 1,\dots,N_w\,. \end{cases} \tag{9}$$

The fact of considering variables z_w at the first stage can be interpreted as a target the first stage sets for the difference between the loss function of the investment $f(x,w)$ and its Value at Risk a, for each scenario w. The advantage of this approach is that the second level function is now separable into individual value functions for each scenario $w = 1, \ldots, N_w$:

$$V_w(x,a,z_w) = \begin{cases} \min\limits_{\{y_w \geq 0\}} & q_w^\top y_w \\ \text{s.t.} & B_w y_w = d_w\,, \\ & W y_w = h_w \;- T x\,, \\ & -q^\top y_w \geq c^\top x - a - z_w\,. \end{cases} \tag{10}$$

Moreover, being the optimal value of a linear programming problem, for which variables (a,x,z_w) are right hand side terms of the feasible region, each function $V_w(x,a,z_w)$ is is convex and piecewise polyhedral. As such, it can be iteratively approximated by a cutting-planes model $\check{V}_w^k(x,a,z_w)$, defined below.

Suppose first that, for all scenario $w = 1, \ldots, N_w$, the feasible sets in (10) are nonempty for given $(x^i, a^i, z_w^i, w = 1, \ldots, N_w)$, so that the corresponding linear programs have both primal and dual solutions. Then, when solving (10) written with (x,a,z_w) therein replaced by (x^i, a^i, z_w^i), not only the value $V_w(x^i, a^i, z_w^i)$, but also optimal multipliers (λ_w^i, μ_w^i), with $\mu_w^i \geq 0$, corresponding to the two last constraints in (10), are available. As a result, recalling that a, z_w, and μ_w are scalars, the subgradient

$$g_w^i = \begin{pmatrix} -T^\top \lambda_w^i + \mu_w^i c \\ -\mu_w^i \\ -\mu_w^i \end{pmatrix} \in \partial V_w(x^i, a^i, z_w^i) \tag{11}$$

can be computed, too. It is then possible to define an affine hyperplane supporting from below each function (10):

$$V_w(x^i, a^i, z_w^i) + g_w^{i\;\top}(x - x^i, a - a^i, z_w - z_w^i) \leq V_w(x,a,z_w)\,.$$

Hence, the maximum of the hyperplanes obtained for different points (x^i, a^i, z_w^i), for $i = 1, \ldots, k$, gives a cutting-planes model $\check{V}_w^k$.

If for some scenario w, the parameters (x^i, a^i, z_w^i) make the feasible set in (10) empty, instead of "optimality cuts" as the hyperplanes above, one should generate "feasibility cuts". Basically, this amounts to solve a linear program like (10), adding slack variables s and replacing the objective function by a penalization of the constraint violation, to obtain the feasibility value function

$$U_w(x,a,z_w) = \begin{cases} \min\limits_{\{y_w \geq 0, s := (s_1,s_2,s_3) \geq 0\}} & |s|_1 \\ \text{s.t.} & B_w y_w = d_w\,, \\ & W y_w + s_1 - s_2 = h_w \;- T x\,, \\ & -q^\top y_w - s_3 \geq c^\top x - a - z_w\,. \end{cases} \tag{12}$$

This linear program has always primal and dual solutions, with its optimal value being null whenever problem (10) is feasible. In particular, denoting once more (λ_w^i, μ_w^i) with $\mu_w^i \geq 0$ the optimal multipliers corresponding to the two last constraints in (12), the vector g_w^i from (11) gives a subgradient for the function U_w at the point (x^i, a^i, z_w^i). Since the feasibility function is always nonnegative, whenever $U_w(\cdot) \leq 0$ the feasible set in (10) is nonempty. This is why the relation

$$U_w(x^i, a^i, z_w^i) + g_w^{i\;\top}(x - x^i, a - a^i, z_w - z_w^i) \leq 0$$

defines a "feasibility cut" for (10). We refer to [3, p. 170] for more details; see also [2].

While the weighted sum of the optimality cuts give a cutting-planes model for the second-level function in (9), feasibility cuts define a polyhedral region for which (10) is well defined. At each iteration i, either an optimality cut or a feasibility cut is generated. More precisely, suppose that after k iterations, past iterations in the index set O_k generated optimality cuts, while iterations in the set F_k generated feasibility cuts. We have then that $O_k \cup F_k = \{1, \ldots, k-1\}$ and the approximate first level problem is a linear program of the form:

$$\begin{cases} \min\limits_{\{x,a,r,(z_w)_1^{N_w}\}} & c^\top x + r \\ \text{s.t.} & Ax = b\,, \\ & a + \frac{1}{1-\beta} \sum\limits_{w=1}^{N_w} p_w z_w \leq \rho_\beta\,, \\ & r \geq \sum\limits_{w=1}^{N_w} p_w \Big(V_w(x^i, a^i, z_w^i) + {g_w^i}^\top (x - x^i, a - a^i, z_w - z_w^i) \Big) \quad i \in O_k \qquad (a) \\ & 0 \geq U_w(x^i, a^i, z_w^i) + {g_w^i}^\top (x - x^i, a - a^i, z_w - z_w^i)) \qquad i \in F_k \\ & 0 \leq x \leq 1\,, z_w \geq 0\,, \qquad w = 1, \ldots, N_w\,. \end{cases} \tag{13}$$

A primal solution $\{x^k, a^k, r^k, (z_w^k)_1^{N_w}\}$ gives the next point to call the second level problems (10) or (12). In the problem above, the additional variable r represents the cutting-planes model, and its optimal value satisfies $r^k = \sum_{w=1}^{N_w} p_w \check{V}_w^k(x^k, a^k, z_w^k)$. Since by construction $\check{V}_w^k \leq V_w$ for all scenario w, and no cut is deleted along the process, for some subsequence of iterates the nonnegative difference

$$\sum_{w=1}^{N_w} p_w V_w^{k_j}(x^{k_j}, a^{k_j}, z_w^{k_j}) - r^{k_j} \to 0 \quad \text{as } j \to \infty\,. \tag{14}$$

Thus, the iterative procedure can stop when such difference becomes smaller than a given tolerance with a good estimate for a solution to (6).

4.2 Price decomposition

We now present an alternative decomposition by prices, based on Lagrangian Relaxation. In a classical approach, the variables that are not indexed by scenarios are placed on the first stage and the remaining variables on the second stage, obtaining the bilevel problem (7)-(8).

Lagrangian Relaxation is specially interesting in problems that present complicating constraints that increase the computational cost for the problem solution. In (6), constraint (d), coupling scenarios, appears as a natural candidate for relaxation. The dualization of this constraint using a scalar multiplier $\eta \geq 0$ gives the dual problem

$$\max_{\eta \geq 0} \theta(\eta)$$

with the (concave) dual function defined by

$$\theta(\eta) := \begin{cases} \min\limits_{x,a,(y_w,z_w)_1^{N_w}} & c^\top x + \sum\limits_{w=1}^{N_w} p_w q^\top y_w \; + \eta\,(a + \frac{1}{1-\beta}\sum\limits_{w=1}^{N_w} p_w z_w - \rho_\beta) \\ \text{s.t.} & Ax = b\,, \\ & B_w y_w = d_w\,,\ w = 1,\ldots,N_w \\ & Tx + W y_w = h_w\,,\ w = 1,\ldots,N_w \\ & z_w \geq c^\top x + q^\top y_w - a\,,\ w = 1,\ldots,N_w \\ & 0 \leq x \leq 1\,, y_w \geq 0\,, z_w \geq 0\,,\ w = 1,\ldots,N_w\,. \end{cases} \tag{15}$$

The evaluation of the dual function at a point η^i requires to solve the problem above. As before, along with the optimal value $\theta(\eta^i)$ a supergradient for the dual function is available. More precisely, if $\{x^i, a^i, (y_w^i, z_w^i)_1^{N_w}\}$ denotes a primal solution to (15), then the scalar

$$a^i + \frac{1}{1-\beta}\sum_{w=1}^{N_w} p_w z_w^i - \rho_\beta \text{ is a supergradient for } \theta \text{ at } \eta^i\,. \tag{16}$$

Since (15) is a problem on all the variables $\{x, a, (y_w, z_w)_1^{N_w}\}$, when there is a large number of scenarios, the exact evaluation of the dual function and a supergradient can be too difficult. For this reason, similarly to [9], we obtained approximate values by inducing separability among scenarios, as before. Splitting the decision variables into first level and second level ones, (x,a) and $(y_w, z_w)_1^{N_w}$, respectively, we rewrite $\theta(\eta) = -\eta\rho_\beta + \varphi(\eta)$, with

$$\varphi(\eta) := \begin{cases} \min\limits_{\{x,a\}} \; c^\top x + \eta a + \sum\limits_{w=1}^{N_w} p_w W_{w,\eta}(x,a) \\ \text{s.t.} \qquad Ax = b\,, \\ \qquad\qquad 0 \leq x \leq 1\,. \end{cases} \tag{17}$$

In the problem above, each individual value function has the form

$$W_{w,\eta}(x,a) := \begin{cases} \min\limits_{\{y_w \geq 0, z_w \geq 0\}} & q^\top y_w + \dfrac{\eta}{1-\beta} z_w \\ \text{s.t.} & B_w y_w = d_w\,, \\ & W y_w = h_w \; - Tx\,, \\ & z_w - q^\top y_w \geq c^\top x - a\,, \end{cases} \tag{18}$$

similar to (10) (but not the same). As a result, letting $\{y_w^i, z_w^i\}$ be a primal solution to (18) written with $(x,a) = (x^i, a^i)$, with λ_w^i and $\mu_w^i \geq 0$ denoting optimal multipliers for the last two constraints in (18), the affine hyperplane

$$q^\top y_w^i + \frac{\eta}{1-\beta} z_w^i + (-T^\top \lambda_w^i + \mu_w^i c)^\top (x - x^i) - \mu_w^i (a - a^i)$$

supports from below $W_{w,\eta}(\cdot)$. Repeating the procedure for $i = 1,\ldots,k$ and making the weighted sum over all scenarios gives a cutting-planes model $\check{W}_\eta^k$ for the second level function in (17). In turn, this gives the approximation

$$\varphi^k(\eta) := \begin{cases} \min\limits_{\{x,a\}} c^\top x + \eta a + W^k_\eta(x,a) \\ \text{s.t.} \qquad\qquad Ax = b\,, \\ \qquad\qquad\quad 0 \le x \le 1\,. \end{cases}$$

that can be used to define a cutting-planes model in (15), and obtain an approximation for the supergradient in (16).

When compared to Benders decomposition, we see that Lagrangian relaxation eliminated the scenario-wise coupling at the expense of introducing an additional level of optimization. Since the dual variable η is scalar, this may not be a handicap, especially when the number of scenarios is large.

5 Numerical Results

In order to assess the validity of the approach, we consider a case study based on real data for the Brazilian network of gas pipelines.

For small instances, the equivalent deterministic model can be solved with commercial packages, like CPLEX 11. To check the corresponding computational cost, we initially created a problem as in (6) with only one scenario ($N_w = 1$). This model and the deterministic version of the problem are solved in about 0.5 seconds. The number of scenarios was then gradually increased, until the computational resources of the available computer were exhausted. As the number of scenarios increases, the CPU time needed to solve (6) directly increases significantly. Making a smart choice of the solver's parameters, it is possible to obtain the solution of equivalent deterministic models with up to $N_w = 63$ scenarios. The corresponding linear program (6) has about 1,845,000 constraints and 1,860,000 variables. The computer was unable to allocate more memory beyond this scenario cardinality.

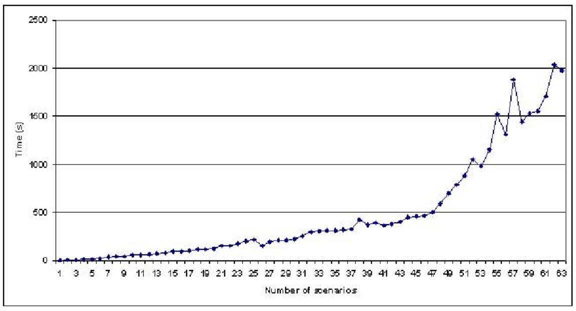

Fig. 5 CPU time of equivalent deterministic problem versus number of scenarios.

Figure 5 plots the CPU time needed to solve (6) directly, as a function of the number of scenarios. It can be seen that, as expected, there is an exponential growth in time for solving the investment problem. For this reason, since the typical cases of interest use $N_w \in [80, 200]$ scenarios, instead of looking for a more powerful computer, we used decomposition methods.

Employing the previously presented decomposition methods, we solved some stochastic instances with 80 scenarios. The reference data comes from the deterministic model (2), using average data for the demand, giving an investment portfolio with Net Present Value (NPV) equal to 5151 (10^6\$) and CVaR-90% equal to 3434 (10^6\$).

A first instance, dropping the $CVaR_\beta$ constraints (d) and (e) in (6), represents a stochastic risk neutral model, as in (1). The attained portfolio has mean NPV of 5491 (10^6\$) and CVaR-90% of 3938 (10^6\$). This result already represents an improvement over the reference, of 6.61% in the NPV and 14.7% in the CVaR. This improvement is a direct consequence of the stochastic approach, which avoids the phenomenon of the flaw of averages.

In Figure 6 we see, for each scenario, the portfolio NPVs obtained with both the stochastic risk neutral model (1) and its deterministic replacement (2).

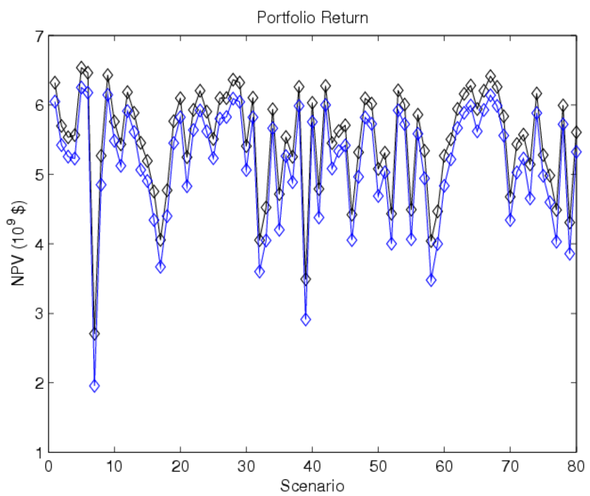

Fig. 6 Asset allocation with the deterministic and stochastic models.

Clearly, the portfolio determined by the stochastic model is different from the one found by the deterministic model. As expected, we observe that the return obtained by the stochastic model dominates the return of the deterministic model: the stochastic model found a portfolio with higher NPV than the deterministic model, for every scenario.

In Table 1 we report the optimal allocation levels of both models, for every available asset.

It appears that, even in a risk neutral variant, the stochastic model diversifies more investment, making smaller allocations in some assets and investing in assets

Asset	Stochastic allocation	Deterministic allocation
1	100,00%	100,00%
4	62,60%	76,90%
7	100,00%	0,00%
8	100,00%	100,00%
9	25,40%	25,40%
0A	100,00%	100,00%
0B	100,00%	100,00%
0C	17,61%	35,71%
0D	0,00%	0,00%
0E	29,46%	0,00%
0F	0,00%	0,00%
0G	100,00%	100,00%
0H	16,79%	18,55%
0I	35,36%	35,36%
0J	0,00%	0,00%

Table 1 Assets chosen by the deterministic and stochastic models.

that had not been chosen by the deterministic model. It is still possible to enable the restriction of CVaR, obtaining a solution with CVaR of 3941 (10^6\$) and expected NPV of 5489 (10^6\$). This result represents an improvement of 6.57% in the NPV and 14.76% in CVaR in relation to the deterministic model.

We observed in the studied case that the variability of CVaR is much smaller than the corresponding one of returns. This is a consequence of the absence of proposed projects with risk mitigating profiles. This is an indicative that the decision maker must propose new investments in order to improve the return for the worst scenarios, if he wants to have greater protection against the risk.

Typically, a run with Benders' model took less than 30 minutes, but could reach 9h for some cases when the CVaR constraint was active, as shown in Table 3. In order to reduce the computational requirements of the approaches, some additional tests were performed.

In the first set of tests, inactive cuts were deleted at each iteration. For example, in (13), instead of having that $O_k \cup F_k = \{1, \dots, k\}$ for all iterations, the next index set of optimality cuts, O_{k+1}, contains the optimality cut generated at iteration k and those indices in O_k for which constraint (13) (a) held with equality at the solution. Somewhat similarly to the compression mechanism in bundle methods, by suppressing inactive optimality cuts we expected to reduce the size of the linear programs (13). Note however that, unlike bundle methods, there is no longer a theoretical guarantee for convergence: a result like (14) can only be ensured if the information generated at all iterations is kept in problems (13).

We tested two instances of the problem, one where the CVaR constraints were active at the optimal solution, and another one, without CVaR constraints. Results are summarized on Table 2.

It is clear from Table 2 that the CVaR constrained instances require more iterations to reach optimality. In the table, those fields with a * correspond instances for which Benders model did not converge before 1000 iterations . We may conclude, at least for Benders approach and with with our data, that inactive cuts may become active in further iterations, thus systematically cleaning inactive cuts has a negative impact in the overall process, due to the increase in the number of it-

Model	CVaR constraint	Feasibility Cuts	Optimality Cuts
All cuts preserved	No	4	96
All cuts preserved	Yes	916	916
Only active cuts	No	10	813
Only active cuts	Yes	*	*

Table 2 Number of cuts of each model instance

erations. This feature indicates that there can be potential gains in using a bundle method [3], instead of a cutting-planes algorithm, as considered in this paper.

We also noticed that there was systematically a large number of iterations generating cuts which did not contribute significantly in improving the algorithmic process, making convergence speed stall. This is due to the existence of several degenerate optimal solutions, resulting in cycling. In order to mitigate this effect, we explored alternative formulations of the problem. In particular, we increased the optimality tolerance when solving subproblems (10), but this approach did not reduce the overall computational time of the algorithm.

In order to assess the performance of the proposed decomposition schemes, we tested several instances with different problem sizes. The summary of the tests results are presented in Table 3. The fields with the * correspond to infeasible instances, resulting from setting CVaR requirement that was unattainable with the corresponding data.

# Scenarios	Problem	Min CVAR	Time (s)	Iterations	CVAR
10	Benders inactive CVAR	-	649	106	2.705
20	Benders inactive CVAR	-	896	85	3.380
25	Benders inactive CVAR	-	1167	87	3.837
30	Benders inactive CVAR	-	1446	93	3.837
40	Benders inactive CVAR	-	1554	69	3.574
50	Benders inactive CVAR	-	2062	76	3.744
10	Benders active CVAR	2.706	8817	2214	2.706
20	Benders active CVAR	3.381	5897	2139	3.381
25	Benders active CVAR	3.838	*		
30	Benders active CVAR	3.838	10888	3385	3.838
40	Benders active CVAR	3.574	32623	4588	3.574
50	Benders active CVAR	3745	14612	4252	3.745
10	Lagrange inactive CVAR	-	623	104	2.705
20	Lagrange inactive CVAR	-	1975	89	3.380
25	Lagrange inactive CVAR	-	1055	76	3.837
30	Lagrange inactive CVAR	-	2016	110	3.837
40	Lagrange inactive CVAR	-	1983	89	3.574
50	Lagrange inactive CVAR	-	2235	80	3.744
10	Lagrange active CVAR	2.706	4439	506	2.706
20	Lagrange active CVAR	3.381	1251	106	3.381
25	Lagrange active CVAR	3.838	*		
30	Lagrange active CVAR	3.838	1688	98	3.838
40	Lagrange active CVAR	3.574	6508	283	3.574
50	Lagrange active CVAR	3745	8401	413	3.745

Table 3 Performance of the decomposition methods over several instances

In instances where the CVaR constraint was inactive, both algorithms had similar results. The main differences may be noticed in the cases were the CVaR constraint was active. Recall that the Benders decomposition approach needs to solve the feasibility cuts subproblem at any iteration with first stage iterate infeasible for the second level. In all the runs of our study, whenever the algorithm rendered an infeasible solution, several feasibility cuts had to be generated before finding a new feasible first stage iterate. This behavior greatly increased the iteration count and solution time of this approach, especially for those instances with active optimal CVaR constraints. By contrast, the price decomposition approach rapidly enters the feasible region, usually converging within 1 or 2 iterations of its first level problem. Due to this feature, the Lagrange model was superior to Benders', both in solution time and iteration count, as shows in Figure 7. Furthermore, in instances whose data made the CVaR constraint infeasible, the Lagrangian approach reported final points that would still provide valuable information for the user, while the Benders approach would only return an infeasible solution status. More precisely, since our portfolio problem (6) is convex, barring problems with infeasible data, the Lagrangian approach always solves the initial problem. If the instance was infeasible due to the CVaR constraint, this fact was easily detected by the method, that would exit with an output corresponding to an optimal solution for a a similar instance, with relaxed CVaR constraint.

Figure 7 shows a comparison between Benders and the price decomposition approaches, for several types of problems. Both models are comparable for unconstrained (risk neutral) instances, while the price decomposition approach is superior for CVaR constrained instances.

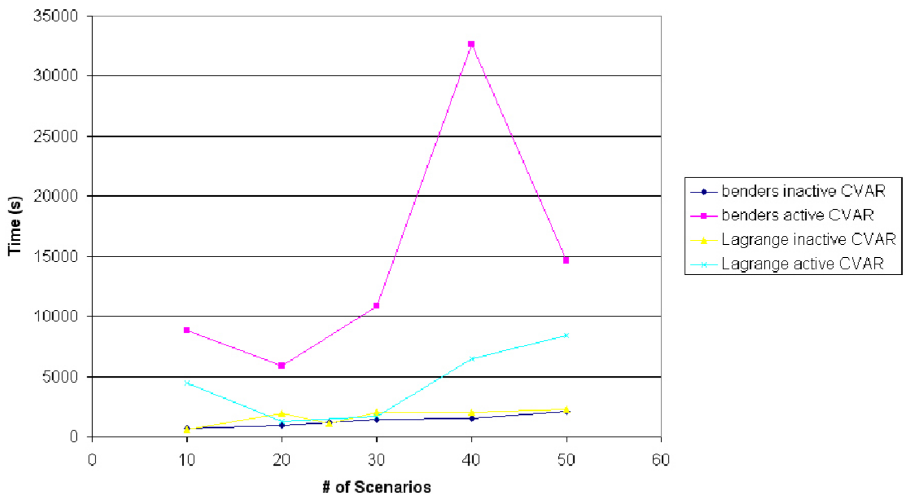

Fig. 7 Comparison of time results of several problem instances.

Considering the overall performance and the above mentioned side benefits for the infeasible instances, for our case study we conclude that the Price (Lagrange) decomposition approach has more practical appeal than the Benders' approach.

6 Concluding remarks

This work developed a model to support the choice of a real assets investment portfolio, given certain operational and financial constraints. The stochastic model includes CVaR constraints, allowing risk management of the portfolio.

In the case study we observed that the proposed stochastic model was able to use the stochastic demand information more efficiently than the deterministic model, providing allocations with higher expected profit and lower risk profile.

The inclusion of CVaR constraints is an interesting approach to manage risk in a sound and computationally tractable manner. However, for the CVaR approach to be meaningful, it is essential for the decision maker to provide enough investments in assets that can potentially mitigate risks for the portfolio, allowing the model to find natural hedges. A good example for the considered application would be to incorporate alternatives for flexible supply and demand of gas, such as investments in Liquefied Natural Gas terminals and flexible contracts with consumers. Without the presence of alternative investment profiles, problem (6) may be unfeasible and the model may be unable to allocate resources in a manner that satisfies the required risk constraint.

The proposed decomposition techniques successfully solved the stochastic model in reasonable time. The Price decomposition presented some advantages over the Benders decomposition scheme. Both methods also offer the possibility to be parallelized, to further reduce CPU time. The incorporation of these advances in the complex planning of the petroleum industry is just starting. From our preliminary numerical experience, we believe that those oil companies that are expanding their activities to become integrated energy companies can obtain significant gains by using the presented techniques in their corporate planning.

References

1. P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Math. Finance*, 9:203–228, 1999.
2. John R. Birge and François Louveaux. *Introduction to stochastic programming*. Springer-Verlag, New York, 1997.
3. J. Frédéric Bonnans, J. Ch. Gilbert, C. Lemaréchal, and C. A. Sagastizábal. *Numerical Optimization: Theoretical and Practical Aspects (Universitext)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2 edition, 2006.
4. P. Krokhmal, J. Palmquist, and S. Uryasev. Portfolio optimization with conditional value-at-risk objective and constraints. *The Journal of Risk*, 4(2):11–27, 2002.
5. C. Prates, E. Pierobon, R. Costa, and V. Figueiredo. Evolução da oferta e da demanda de gás natural no Brasil. *BNDES Setorial*, 2006/09.
6. R. Rockafellar. Coherent approaches to risk in optimization under uncertainty. *INFORMS tutorial*, November 2007.
7. R. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *The Journal of Risk*, 2(3):21–41, 2000.
8. S. Savage. The flaw of averages. *Harvard Business Review*, November:20–21, 2002.
9. Golbon Zakeri, Andrew B. Philpott, and David M. Ryan. Inexact cuts in Benders decomposition. *SIAM J. Optim.*, 10(3):643–657 (electronic), 2000.