Optimization Techniques for the Brazilian Natural Gas Network Planning Problem

Leonardo A. M. Moraes[a,*], Sergio V. B. Bruno[a], Welington de Oliveira[b]

[a] *PETROBRAS, Operations Research Department, Rio de Janeiro, 20020-100, Brazil*
[b] *IMPA, Instituto Nacional de Matemática Pura e Aplicada, Rio de Janeiro, 22460-320, Brazil*

Abstract

This work reports on modeling and numerical experience in solving the long-term design and operation planning problem of the Brazilian natural gas network. A stochastic approach to address uncertainties related to the gas demand is considered. Representing uncertainties by finitely many scenarios increases the size of the resulting optimization problem, and therefore the difficulty to solve it. We calculate the value of applying a stochastic programming framework. The numerical tractability of the problem is obtained by applying state of the art techniques of optimal scenario reduction and decomposition by bundle methods. The benefits of each approach are considered. Numerical experiments in a real-life case are assessed.

Keywords: Stochastic programming, Long-term planning, Natural gas, Large scale optimization, OR in energy

1. Introduction

Expansion and operation planning of existent supply chain is a key activity in the natural gas market. By expansion, one can consider physical expansion of pipelines, construction of new ones, and negotiation of prices and volumes for supply and demand contracts – new or existent ones. There is also a need for an integration between expansion and operation plans, due to the quest for a higher profitability. Thus, existent and new infrastructure must be operated in an optimal way, considering the following options for natural gas use:

- ∘ *acquisition* – production or purchase of imported gas or liquefied natural gas (LNG) loads;
- ∘ *consumption* – local distribution companies demand, internal consumption, and thermoelectric power plants fueling.

In Figure 1 an example of a natural gas supply chain is shown. This chain starts with exploration and production activities that, together with importation activity, correspond to gas *acquisition*. Gas is shipped to delivery nodes that represent *consumption*. Activities covered by the model presented in this work are marked with a rectangle.

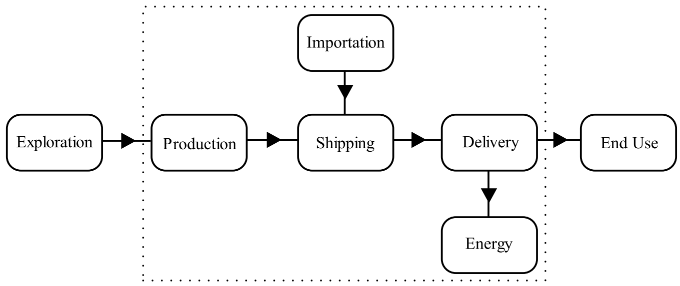

Figure 1: Activities of natural gas supply chain

Investments in new natural gas sources usually take up to ten years until production is initiated reliably and safely and involve large sums of money. Thereby, long-term planning of a natural gas supply chain, taking into account infrastructure development and long-term operation of the system, is of great importance for the profitability of the natural gas industry, Li et al. (1).

In Brazil, natural gas is expected to have a more relevant role in the energy market over the coming decades. In its majority, Brazilian natural gas is associated to oil production and this portion should be expanded as soon as more pre-salt oil fields come into operation. In 2012 natural gas average supply was about 2.65 BCF/day, of which 1.39 BCF/day is related to domestic production, 0.95 BCF/day were imported from Bolivia and the remaining 0.31 BCF/day refers to LNG loads, specially from Nigeria and Trinidad and Tobago. These numbers correspond to a 22% increase, if compared to 2011 numbers.

As of 2019, Brazilian natural gas consumption should be about 6 BCF/day, of which 30% must be related to industrial sector and about 9.5% must be used for energy generation. In that same year, Brazil/Bolivia 20-year supply contract must be renegotiated.

Natural gas has several uses: in residences, in industries, in oil fields for re-injection purposes, as fuel for refineries and rigs, for vehicles, and for energy generation. Some of these demands can be deterministically forecasted due to its low variability, such as residences, industries, and vehicles consumption. However, gas demand for energy generation purposes can be highly uncertain, as explained below. This characteristic turns the natural gas network planning into a stochastic problem.

The Brazilian generating system is hydro dominated and characterized by large reservoirs with multi-year regulation ca-

*Corresponding author. Tel.: +55 21 3229 0697; Fax: +55 21 3229 0444.
Email addresses: `leonardo.moraes@petrobras.com.br` (Leonardo A. M. Moraes), `sbruno@petrobras.com.br` (Sergio V. B. Bruno), `wlo@impa.br` (Welington de Oliveira)

pability, arranged in complex cascades over several river basins. Hydroelectric plants are responsible for about 85% of the total annual energy generation. Due to this hydro predominance, operation is driven by the rainfall events – occurrence and forecasting. As an example, if hydroelectric reservoirs are at low levels and a dry period (in terms of rainfall) is expected, then thermoelectric dispatch is high, in order to save the water in hydroelectric reservoirs. On the other hand, if reservoirs are at regular or high levels, and no dry period is being expected in the coming months, so thermoelectric dispatch tends to be small because there is no need to save water.

Therefore, thermoelectric dispatch depends on the rainfall, which is uncertain. As a consequence, thermoelectric dispatch is also uncertain. This dispatch is centralized by an independent system operator (ONS) supported by a computational model, described in Maceira et al. (2). This model is based on the Stochastic Dynamic Dual Programming technique, introduced by Pereira and Pinto (3). One of the outputs of this model is a set of scenarios of thermoelectric dispatch that can be used to represent the stochasticity of the gas demand for energy generation purposes.

A natural way to address gas demand uncertainties in the Brazilian natural gas network planning problem consists in modeling the problem into the framework of stochastic programming with recourse, as defined in Birge and Louveaux (4). By using this framework, a set of possible scenarios (representing uncertain data realizations) approximates the *real* stochastic distribution of these uncertainties.

More precisely, in this work we consider a case based on a real-life natural gas industry problem from PETROBRAS, the Brazilian oil and gas company. Such problem, which is of large-scale and difficult to solve, was modeled as a two-stage stochastic linear problem, meaning that some decisions – investments in infrastructure and new contracts – are of *here-and-now* type, i.e., they do not depend directly on the uncertain data. In each scenario the decision maker can *wait-and-see* the uncertainty revealed, and take recourse decisions such as natural gas volumes, pipelines use and thermoelectric plants operation.

Some authors have already studied the use of stochastic programming methods in one or more planning levels – strategic, tactical or operational. In Sagastizábal (5) one can find an extensive list of examples of decomposition models applied to energy systems optimization. Santoso et al. (6) present a stochastic model for supply chain network design under uncertainty. The goal of the latter work was to route the flow of products from a supplier to customers, and to define which processing centers should be built. The modeling results in a mixed-integer linear problem, which is solved via sample average approximation, Kleywegt et al. (7). Schütz et al. (8) also studied supply chain design under uncertainty, for the Norwegian meat industry. The authors apply sample average approximation and Lagrangian relaxation techniques to the resultant two-stage stochastic problem. Cutting-plane Kelley (9) and bundle methods Hiriart-Urruty and Lemaréchal (10) are used to solve the Lagrangian dual problem. Another example of integrated design and operation problem is presented by Li et al. (11), who propose a large-scale mixed-integer problem solved by the so called Nonconvex Generalized Benders Decomposition Algorithm, Li et al. (12).

Due to the size and complexity of the Brazilian natural gas network planning problem, we have chosen to model it in a continuous framework, instead of using a mixed-integer approach. We study a compromise between uncertainty representation and solvability for the considered application: considering a large number of gas demand scenarios the resulting problem is solved by applying a decomposition technique akin L-Shaped method Slyke and Wets (13), but using a bundle method instead. An additional approach consists in selecting a much smaller but *representative* amount of scenarios to represent the gas demand uncertainties. The representative scenarios are selected by applying the optimal scenario reduction technique developed in Heitsch and Römisch (14).

The main contributions of this work are the empirical comparison of the benefits of decomposition techniques and optimal scenario reduction, as well as their effect over a natural gas network planning problem - the Brazilian case. The paper is organized as follows. In Section 2 some characteristics of the considered problem are explained and the mathematical modeling is presented. In order to validate the stochastic approach for the problem, stochastic measures are reviewed and compared in Section 3. A decomposition technique and a bundle method are considered in Section 4. Section 5 address the optimal scenario reduction technique. Some concluding remarks are reported in Section 6. Numerical experiments for a 20-year horizon planning problem are given throughout Sections 3, 4 and 5.

2. The long-term planning problem

2.1. Problem statement

The Brazilian natural gas network planning problem is composed of:

– natural gas supply nodes – representing domestic gas production, imported gas, and LNG loads;
– demand nodes – representing local distribution companies and internal consumption (e.g. fertilizers factories and refineries);
– pipelines for the gas transportation;
– compressors to help with the gas transportation through pipelines;
– natural gas processing units – NGPU; and
– thermoelectric plants.

As we are dealing with a long-term (20-years horizon) planning problem, operational constraints are modeled with far less details than they would be in a short-term model, but some of them are still present in the model, such as pipelines maximum allowed flows and the use of a heat rate to convert natural gas into energy at thermoelectric plants.

The main goal of the model is to determine an optimal investment plan, including pipelines capacities expansion (or pipeline construction), volumes for new demand contracts (with, e.g., local distribution companies) and optimal operation of the whole network, i.e, gas flows through pipelines, gas volumes sent to demand nodes and absorbed in supply nodes, and thermoelectric plants operation.

In the modeled problem, a company buys natural gas at some input nodes (supply nodes) and must carry this gas to demand nodes or thermoelectric plants through its own pipelines network. For the sake of simplicity, we consider that no gas loss occurs during transportation. Each demand node expects to receive a certain amount of gas, varying from a minimal to a maximal value. If the optimal decision is to deliver less than the minimal natural gas demand at a node, company must pay for this unsatisfied amount of gas. Thermoelectric plants demands are related to their dispatch levels, i.e, energy dispatch is converted into gas demand taking into account generator machines' efficiency.

Investments may be made in certain periods within the planning horizon and are offered in projects. A project may consist of any combination of investment in: new pipelines or expansion of existent ones, supplying contracts minimal and maximal volumes, and demand contracts minimal and maximal volumes. As an example, a project may consist of a new pipeline to carry gas from a developed field to a new demand node, also included in the project, with minimal and maximal demand values.

The expansion and operation planning problem consists of deciding (i) which investment projects must be made, when and at which level, and (ii) given an optimal investment policy, volumes of gas to be bought at supply nodes, delivered at demand nodes (including thermoelectric plants), and volumes carried by pipelines.

In Figure 2 a simple representation of the problem is given. Supply nodes, where natural gas can be bought by the company, are represented by boxes with capital letter S, while demand nodes, where natural gas is delivered, are represented by boxes with capital letter D. Gas can also be delivered to thermoelectric plants. Company internal network is composed of pipelines, compressors, and natural gas processing units (NGPU), represented by capital letters C e U, respectively.

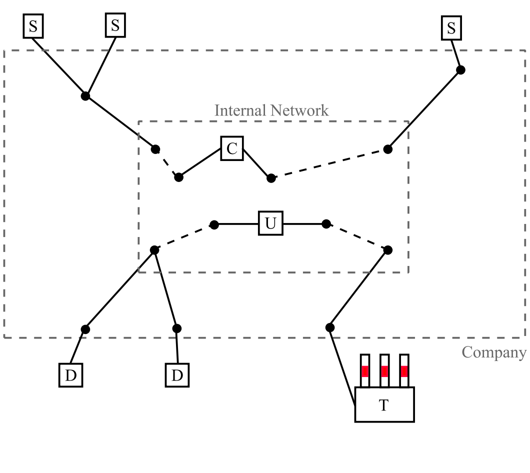

Figure 2: Natural gas network representation

2.2. *Mathematical Model*

Let us first define two main sets: $\mathcal{N}$ and $\mathcal{P}$ stand for the set of nodes and pipelines, respectively. Each node $n \in \mathcal{N}$ can represent a gas input (resp. output) point, representing a supplier (resp. demand) or a concentration point, at which two or more pipelines segments converge to or diverge from. Two nodes can be connected by a pipeline segment, a compressor or a processing unit. Compressors and processing units sets are represented by sets $\mathcal{C}$ and $\mathcal{U}$, respectively. Each thermoelectric plant $\theta \in \Theta$ consists of a set of engines $\mathcal{E}_\theta$ with different heat rates[1]. Supply and demand nodes are represented by sets $\mathcal{S} \subset \mathcal{N}$ and $\mathcal{D} \subset \mathcal{N}$. It turns out that the set of thermoelectric is contained in the set of demand note, i.e., $\Theta \subset \mathcal{D}$. Finally, let $\mathcal{I}$ denote the set of investment projects and $\mathcal{T}$ the set of periods. The decision variables of the considered problem are:

– $x_i^t \in [0,1]$, the percentage of investment i to be made at period t;
– $w_{ij}^t \geq 0$, the flow of gas from node i to node j at period t;
– $e_\theta^t \geq 0$, representing the amount of energy generated in plant θ, at period t; and
– $u_j^t \geq 0$, natural gas deficit in node j, period t.

The deterministic version of the considered natural gas network planning problem can now be stated and symbolized by

$$\min_{x,w,e,u} \sum_{t\in\mathcal{T}} \left\{ \sum_{i\in\mathcal{I}} ci_i \cdot \left(x_i^t - x_i^{t-1}\right) + \sum_{j\in\mathcal{S}} \left(cg_j^t \cdot \sum_{k\in\mathcal{N}} w_{jk}^t \right) + \sum_{j\in\mathcal{S}\cup D\backslash\Theta} cs_j \cdot u_j^t - \sum_{j\in\mathcal{D}\backslash\Theta} \left(rg_j^t \cdot \sum_{k\in\mathcal{N}} w_{kj}^t \right) - \sum_{\theta\in\Theta} re_\theta^t \cdot e_\theta^t \right\} \tag{1a}$$

$$\text{s.t.} \sum_{k\in\mathcal{N}} w_{jk}^t + u_j^t \geq v_j^t + \sum_{i\in\mathcal{I}} \left(x_i^t \cdot vi_{ij}\right), \quad \forall j \in \mathcal{S}, \forall t \in \mathcal{T} \tag{1b}$$

$$\sum_{k\in\mathcal{N}} w_{jk}^t \leq \upsilon_j^t + \sum_{i\in\mathcal{I}} \left(x_i^t \cdot \upsilon i_{ij}\right), \quad \forall j \in \mathcal{S}, \forall t \in \mathcal{T} \tag{1c}$$

$$\sum_{k\in\mathcal{N}} w_{kj}^t + u_j^t \geq v_j^t + \sum_{i\in\mathcal{I}} \left(x_i^t \cdot vi_{ij}\right), \quad \forall j \in \mathcal{D}\backslash\Theta, \forall t \in \mathcal{T} \tag{1d}$$

$$\sum_{k\in\mathcal{N}} w_{kj}^t \leq \upsilon_j^t + \sum_{i\in\mathcal{I}} \left(x_i^t \cdot \upsilon i_{ij}\right), \quad \forall j \in \mathcal{D}\backslash\Theta, \forall t \in \mathcal{T} \tag{1e}$$

$$w_{ij}^t \leq \phi_{ij}^t + \sum_{i\in\mathcal{I}} \left(x_i^t \cdot \phi i_{ij}\right) \tag{1f}$$

$$hr_\theta \cdot e_\theta^t = \sum_{k\in\mathcal{N}} w_{k\theta}^t, \quad \forall \theta \in \Theta, \forall t \in \mathcal{T}, \tag{1g}$$

$$e_\theta^t \geq ed_\theta^t, \quad \forall \theta \in \Theta, t \in \mathcal{T}, \tag{1h}$$

$$x_i^t - x_i^{t-1} \geq 0, \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}, \tag{1i}$$

$$w \in \mathcal{F}, e \in \Sigma, x \in \mathcal{X}. \tag{1j}$$

In this model, ci_i represents the total investment cost for project i and it is applied to the difference $\left(x_i^t - x_i^{t-1}\right)$, which is equivalent to the amount of investment made at period t. Parameters

[1]Heat rate is an expression of the conversion efficiency of power generating engines in terms of their consumption – in the considered case, natural gas consumption.

cg_j and rg_j denote unitary acquisition cost and sale price for natural gas at node j, respectively. Additional cost term cs_j is included to penalize gas shortfall. Energy remuneration is represented by parameter re_θ, for thermoelectric plant θ. Thus, objective function (1a) consists of minimizing investment, gas acquisition and penalty costs minus demand and thermoelectric remuneration. In other words, solving the above problem is equivalent to maximize the total profit of the company subject to operational and investment constraints.

Constraints (1b) and (1c) require that minimum and maximum values of natural gas acquisition be satisfied. Parameters v_j^t and vi_{ij} represent minimum amounts of natural gas to be bought – the former one represents original minimal value, and the later one represents the minimal values due to investments. A similar analysis can be done to parameters υ_j^t and υi_{ij} for maximal values and constraints (1d) and (1e) represent the same requirements for demand nodes, *mutatis mutandis*. Pipelines capacities are represented in constraint (1f).

Energy generation is represented by constraints (1g) and (1h) – the former one shows the relation between natural gas consumption, $\sum w_{k\theta}^t$, and energy generation, e_θ^t at thermoelectric plant θ, and the latter one ensures a minimal generation, representing energy dispatch. In constraint (1i), investments are forced to be non-decreasing, i.e., disinvestment is not allowed. Sets $\mathcal{F}$, Σ, and $\mathcal{X}$, in constraint (1j) represent other (linear) restrictions, such as maximum energy generation of a plant θ.

2.3. Compact formulation

For convenience, model (1a)-(1j) is represented with the more compact notation

$$\left\{\begin{array}{ll} \min_{x,y} & c^\top x + q^\top y \\ \text{s.t.} & Tx + Wy = h \\ & x \in X,\, y \ge 0\,, \end{array}\right. \tag{2}$$

where vector x keeps representing investment decisions, and decision variable $y = (w \;\; e \;\; u \;\; s)^\top$, being s a slack variable with appropriate dimension. Thus, cost vector $c = ci$, while q is defined by

$$q = ((cg - rg) \;\; -re \;\; cs \;\; 0)^\top.$$

The polyhedral set X represents (1i) and the operational constraints $x \in \mathcal{X}$. Notation T represents a matrix composed of parameters vi, υi, and ϕi. Vector h is composed of parameters v, υ, ϕ, and ed, and also parameters of constraints Σ and $\mathcal{F}$ in (1j). Matrix W, among others, incorporates coefficients of constraints Σ and $\mathcal{F}$, and coefficients of the slack variables.

For practical interests, problem (2) (and thus (1)) might be simplistic. In fact, some parameters involved in problem (2) may not be considered deterministic, such as the energy dispatch represented by parameter ed_θ^t whose reasons were presented in Section 1. Thereby, the deterministic model symbolized by problem (1) needs to be examined in a stochastic context.

3. Incorporating stochasticity into the problem

The question we address in this section is the following one: Is there a significant gain by introducing stochasticity into the model represented by problem (1)?

In order to answer this question we use energy dispatch (gas demand) scenarios represented by the random vector $\xi = \left\{ed_\theta^t\right\}$ in appropriated probability space $(\Xi, \sigma(\Xi), P)$. The resulting formulation is as follows

$$\left\{\begin{array}{ll} \min_{x,y} & \mathbb{E}[c^\top x + q^\top y(\xi)] \\ \text{s.t.} & Tx + Wy(\xi) = h(\xi) \\ & x \in X,\, y(\xi) \ge 0 \text{ a.s.}\,, \end{array}\right.$$

where $\mathbb{E}$ stands for the expectation operator with respect to probability measure P (*a.s.* means almost surely).

It is assumed throughout this work that the random variable ξ is approximated by finitely many N scenarios $\xi_1, \xi_2, \ldots, \xi_N$, with associated probability $p_i > 0$ for all $i = 1, \ldots, N$. For the considered problem, a set of 2,000 gas demand scenarios ξ_i is available by a computational program responsible, among others, for dispatching thermal energy generation to meet a power load in the Brazilian electric energy system; see Maceira et al. (2) for more details.

Since the number N of scenarios is finite, the above problem can be written in the *deterministic equivalent* form

$$z_{\texttt{RP}} := \left\{\begin{array}{ll} \min_{x,y_i} & c^\top x + \sum_{i=1}^N p_i[q^\top y_i] \\ \text{s.t.} & Tx + Wy_i = h(\xi_i) \\ & x \in X,\, y_i \ge 0 \text{ for all } i = 1, \ldots, N. \end{array}\right. \tag{3}$$

In the quest of answering the stated question, we now review and apply some useful measures in stochastic programming.

3.1. Stochastic Measures

In this section we describe two important measures in stochastic programming: the expected value of perfect information (EVPI) and the value of stochastic solution (VSS). A more detailed discussion of these measures can be found in Birge and Louveaux (4).

3.1.1. The Value of Stochastic Solution

In practice, it is not unusual to find situations in which the deterministic equivalent problem (3) is not solvable. This may occur when this problem is large enough to become computationally intractable or when solving time is too excessive. Instead of applying more sophisticated techniques to solve the problem one may be attempted to solve a much simpler problem, replacing all uncertain parameters by their expected values: $\bar{\xi} := \sum_{i=1}^N p_i \xi_i$. In this case, the measure VSS is useful to evaluate the cost of ignoring uncertainty in making a decision.

With purpose of defining the measure VSS, one has to solve the following problem, called the expected value problem (EV)

$$z_{\texttt{EV}} := \left\{\begin{array}{ll} \min_{x,y} & c^\top x + q^\top y \\ \text{s.t.} & Tx + Wy = h(\bar{\xi}) \\ & x \in X,\, y \ge 0\,. \end{array}\right. \tag{4}$$

We denote $(x_{\mathtt{EV}}, y_{\mathtt{EV}})$ its optimal solution and $z_{\mathtt{EV}}$ its correspondent objective value. In order to measure how good decision $x_{\mathtt{EV}}$ is, one may test this solution by solving the following problem and by evaluating the expected cost of using the EV solution:

$$z_{\mathtt{EEV}} := c^\top x_{\mathtt{EV}} + \sum_{i=1}^{N} p_i \left\{ \min_{y_i \geq 0} q^\top y_i : \; W y_i = h(\xi_i) - T x_{\mathtt{EV}} \right\}. \quad (5)$$

Since problem (5) is decomposable for each scenario, its optimization is not a too much time-consuming task, in general.

As defined in Birge and Louveaux (4), one may calculate the value of the stochastic solution VSS as

$$\mathtt{VSS} = z_{\mathtt{EEV}} - z_{\mathtt{RP}} \,,$$

where $z_{\mathtt{RP}}$ is the optimal value of problem (3), considering N scenarios. As previously said, VSS represents the cost of ignoring uncertainty in the problem. Therefore, a small value for VSS means that there is no (enough) gain in solving the much more difficult problem (3), instead of solving (4).

3.1.2. The Expected Value of Perfect Information

Another important measure in stochastic programming is the *expected value of perfect information* – EVPI. This value represents obtained gain if complete and accurate information on future is available at the moment decision is made. Therefore, it is a useful measure to evaluate the scenarios used to define problem (3).

The EVPI is defined as the difference between the optimal value $z_{\mathtt{RP}}$ of the deterministic equivalent problem (3) and the wait-and-see optimal value, given by

$$z_{\mathtt{WS}} := \sum_{i=1}^{N} p_i \left\{ \min_{x_i \in X, y_i \geq 0} c^\top x_i + q^\top y_i : \; T x_i + W y_i = h(\xi_i) \right\}, \quad (6)$$

i.e.,

$$\mathtt{EVPI} := z_{\mathtt{RP}} - z_{\mathtt{WS}} \,.$$

The bigger EVPI is the more important is the role played by the uncertainties. Therefore, more effort should be made in order to obtain a representative set of scenarios. It is shown in Birge and Louveaux (4) that the optimal values of the given formulations satisfy

$$z_{\mathtt{WS}} \leq z_{\mathtt{RP}} \leq z_{\mathtt{EEV}} \,.$$

3.2. VSS and EVPI measures for the considered problem

We now report some numerical results obtained by applying the stochastic measures described in the previous subsection to the Brazilian natural gas network planning problem.

3.2.1. Test problem

The test problem is based on a realistic Brazilian network planning problem, composed of 65 supply nodes, 305 demand nodes, 33 thermoelectric plants and 68 pipeline segments. The planning horizon is 20 years, split into 76 planning periods – each time period can represent either a month or a whole year, depending on the proximity of this period.

Nine investments projects are offered: four of them are related to the expansion of pipelines capacities, three represent new supplying contracts opportunities and the remaining two represent new demand contracts that need to be evaluated. It is important to the company to decide if it is worth to invest in these contracts. As stated above, only thermoelectric dispatch was considered to be uncertain in this work. With purpose of representing *real* uncertainty, we used a set of $N = 200$ scenarios randomly selected from a sample of 2,000 thermoelectric dispatch scenarios given by the official Brazilian energy planning program described in Maceira et al. (2).

All computations were carried out on a Intel Xeon X5650 2.67GHz, with 2 processors, 48 Gbyte of RAM Memory, running 64 bit Windows Server 2008. All implementations were performed using AIMMS 3.13 and Gurobi 5.1 to solve resultant linear problems.

3.2.2. Numerical assessment

Table 1 reports on the difference between wait-and-see problem optimal value and the optimal values of the deterministic problem equivalent (3) and the expected cost problem (5). Values are given in thousand US$. We chose to inform the differences between the optimal values so that the magnitudes of these differences are more easily understood. Problem (3) was solved according to the method to be described in Section 4.

RP	EEV	WS
2 323.31	39 928.40	0

Table 1: Difference between optimal values

We give in Table 2 the stochastic measures, also in thousand US$.

VSS	EVPI
37 605.09	2 323.31

Table 2: Stochastic measures

Note that VSS is about US$ 37.6 million, while EVPI is about US$ 2.3 million. The moderate value of EVPI means that the set of $N = 200$ scenarios represents the problem's uncertainties in a satisfactory way. On the other hand, a big VSS means that there is a huge loss of profitability if one chooses to solve the EV problem (4) instead of solving deterministic equivalent problem (3). A solution to the stochastic version of the considered problem provides an increase in profit of around US$ 37.6 million, when compared to the deterministic variant. Hence, results presented in Table 2 indicate that a stochastic approach should be used in the considered problem.

4. Decomposition

We have concluded in the previous section that considering uncertainties about the demand gas is an important matter for the considered natural gas network planning problem. Therefore, in order to make a better planning decision, problem (3) must be solved considering a given set of gas demand scenarios ξ_i, $i = 1, \dots, N$. Notice that problem (3) is a linear programming problem, which is expected to be easily solved by commercial LP solvers. However, for the considered application and $N = 80$ gas demand scenarios, the problem has approximately 20, 883, 165 variables and 6, 752, 380 constraints; being too hard to be solved by traditional LP solvers even using a powerful computer like the one described in § 3.2.1. In this manner, for solving the problem with $N \geq 80$ scenarios one may consider using decomposition techniques.

4.1. Two-stage stochastic linear programming formulation

Following the lead of Shapiro et al. (15), linear problem (3) can be decomposed as

$$\min_x \; f(x) \;\text{ s.t. }\; x \in X \quad \text{with} \quad f(x) := c^\top x + \sum_{i=1}^{N} p_i Q(x, \xi_i), \qquad (7)$$

where $\sum_{i=1}^{N} p_i Q(x, \xi_i)$ is the expectation of the second stage costs given by

$$Q(x,\xi) := \begin{cases} \min_y & q^\top y \\ \text{s.t.} & Wy = h(\xi) - Tx \\ & y \geq 0\,. \end{cases} \qquad (8)$$

Given the characteristics of problem (1) and slack variables inclusion in problem (2), problem (8) has a solution for every given point $x \in X$, and for all scenarios ξ_i, $i = 1, \dots, N$. This property is known in the stochastic programming literature as *relatively complete recourse*. Given these assumptions, it is well known that problem (7)-(8) is convex, finite valued, but nonsmooth; see (Shapiro et al. (15), § 2.1) for further information.

Most optimization techniques for solving nonsmooth problems rely only on first-order information provided by an oracle (black-box). For the considered two-stage stochastic linear problem (7)-(8), an oracle is a decomposable optimization procedure that for a given feasible point x, it computes the value of the objective function and a subgradient at this point. Such procedure is described below:

Oracle 4.1. *(Oracle for two-stage stochastic linear problems).*

▷ *Step 0*

Input: $x_k \in X$

▷ *Step 1*

for $i = 1, 2, \dots, N$ ***do***
  Solve problem (8) *for* $x = x_k$ *and* $\xi = \xi_i$
  Get the optimal value $Q(x_k, \xi_i)$
  Obtain a Lagrange multiplier u_i *for* (8)
end for

▷ *Step 2*

Compute the value of the function:
$f(x_k) \leftarrow c^\top x_k + \sum_{i=1}^{N} p_i Q(x_k, \xi_i)$
Compute a subgradient:
$g_k \leftarrow c - \sum_{i=1}^{N} p_i T^\top u_i$

Output: $(f(x_k), g_k)$.

Proposition 2.2 in Shapiro et al. (15) ensures that g_k computed above is indeed a subgradient of f at point x_k, i.e.,

$$f(x) \geq f(x_k) + g_k^\top (x - x_k) \;\text{ for all }\; x \in \mathbb{R}^n\,.$$

In the following section we present the optimization tool employed in this work for solving problem (7)-(8), making using of Oracle 4.1.

4.2. Proximal bundle method

Bundle methods are designed to solve nonsmooth convex optimization problems by making use of only first-order information, see Hiriart-Urruty and Lemaréchal (10) and Bonnans et al. (16). Such methods are well known by their robustness and by having an efficient stopping test. Moreover, differently from cutting-plane methods like L-Shaped Slyke and Wets (13), most bundle methods have limited memory: the bundle of oracle information can be kept bounded saving computational memory without impairing convergence. This is particularly interesting for large-scale optimization problems, as the considered one.

This section restricts itself to the presentation of a proximal bundle algorithm suitable for solving problem (7)-(8). We refer to Ruszczyński (17), Fábián (18), Fábián and Szőke (19), and Oliveira and Sagastizábal (20) for more bundle method variants for solving two-stage stochastic linear programming problems.

4.2.1. Description of the method

The method generates a sequence of feasible iterates $\{x_k\} \subset X$. For each point x_k, Oracle 4.1 is called to compute $f(x_k)$ and a subgradient g_k. With such information, the method creates the linearization

$$\bar{f}_k(x) := f(x_k) + g_k^\top (x - x_k) \quad (\leq f(x))\,.$$

At iteration k a polyhedral *cutting-plane* model of f is available:

$$\check{f}_k(x) := \max_{j \in \mathcal{B}_k} \bar{f}_j(x) \quad \text{with} \quad \mathcal{B}_k \subset \{1, \dots, k\}\,. \qquad (9)$$

The set $\{x_j,\; f(x_j), g_j\}_{j \in \mathcal{B}_k}$ is called information bundle. For instance, the L-Shaped method Slyke and Wets (13) takes $\mathcal{B}_k = \{1, \dots, k\}$ for all k. In contrast, bundle methods can keep $\mathcal{B}_k$ with only two well-chosen linearizations, as shown in Remark 4.2 below.

Given a parameter $\mathtt{t}_k > 0$ and denoting $\hat{x}_k$ a stability center (the best past iterate) at iteration k, the next iterate x_{k+1} is the unique solution to the quadratic program

$$\min_x \; \check{f}_k(x) + \frac{1}{2\mathtt{t}_k}|x - \hat{x}_k|^2 \;\text{ s.t. }\; x \in X\,, \qquad (10)$$

which is equivalent to

$$\begin{cases} \min_{x,r} & r + \frac{1}{2\mathtt{t}_k}|x - \hat{x}_k|^2 \\ \text{s.t.} & \bar{f}_j(x) \leq r, \;\forall\, j \in \mathcal{B}_k \\ & x \in X,\; r \in \mathbb{R}\,. \end{cases} \qquad (11)$$

Denoting by i_X the indicator function of the set X, the optimality conditions for (10) give

$$x_{k+1} = \hat{x}_k - \mathtt{t}_k \hat{g}_k, \qquad (12)$$

with

$$\hat{g}_k = p_f^k + p_X^k \quad \text{and} \quad \begin{cases} p_f^k := \sum_{j \in \mathcal{B}_k} \alpha_j^k g_j & \in \partial \check{f}_k(x_{k+1}) \\ p_X^k := -\dfrac{x_{k+1} - x_k}{\mathtt{t}_k} - p_f^k & \in \partial i_X(x_{k+1}) \,. \end{cases} \tag{13}$$

The simplicial multiplier α^k satisfies the following relations for all $j \in \mathcal{B}_k$

$$\sum_{j \in \mathcal{B}_k} \alpha_j^k = 1 \,, \quad \alpha_j^k \geq 0 \,,$$
$$\alpha_j^k [\check{f}_k(x_{k+1}) - \bar{f}_j(x_{k+1})] = 0,$$

and can be used to save storage without impairing convergence. More precisely, "inactive" indices, corresponding $\alpha_j^k = 0$, can be dropped: $\mathcal{B}_{k+1} \supset \{j \in \mathcal{B}_k : \alpha_j^k \neq 0\}$.

Now comes an important convergence parameter:

$$\phi_k := \mathfrak{f}(\hat{x}_k) - \check{f}_k(x_{k+1}) - \hat{g}_k^\top x_{k+1} \,.$$

It is shown in Oliveira et al. (21), § 3 that a proximal bundle algorithm can stop with a satisfactory solution $\hat{x}_k$ when both ϕ_k and $|\hat{g}_k|$ are small. In fact, Oliveira et al. (21), Theorem 3.2 ensures that convergence amounts to obtaining the following property:

a subsequence $\{(\phi_k, \hat{g}_k)\}_{k \in K}$ converges to $(\phi, 0)$ with $\phi \leq 0$.

A *rule* decides whether to move the center ($\hat{x}_{k+1} = x_{k+1}$, descent-step) or to keep it ($\hat{x}_{k+1} = \hat{x}_k$, null-step). This rule views $f(\hat{x}_k)$ as a threshold, which each iteration strives to improve: a descent-step is performed if it improves the threshold by a definite amount; say

$$f(x_{k+1}) \leq \mathfrak{f}(\hat{x}_k) - \kappa \mathtt{v}_k, \tag{14}$$

with $\mathtt{v}_k := f(\hat{x}_k) - \check{f}_k(x_{k+1})$ and some fixed $\kappa \in (0, 1)$.

Algorithm 1 outlines the considered proximal bundle method variant. If (x_{k+1}, r_{k+1}) is a solution to problem (11), then $r_{k+1} = \check{f}(x_{k+1})$. Hence, the updating rule of $\mathtt{v}_k$ on line 9 of the algorithm coincides with the definition in (14).

Remark 4.2 (Bundle compression)**.** *It is worth mentioning that the set $\mathcal{B}_k$ gathering bundle information can be kept with at most $M_{\max}$ indices, for some chosen integer $M_{\max} \geq 2$. In fact, if at Step 4 of Algorithm 1 one has $|\{j \in \mathcal{B}_k : \alpha_j^k \neq 0\}| = M_{\max}$, one can choose any two indices $j, i \in \mathcal{B}_k$ and replace the two triples $(x_j, f(x_j), g_j)$ and $(x_i, f(x_i), g_i)$ by the artificial one $(x_{k+1}, \check{f}_k(x_{k+1}), \hat{g}_k)$. In this manner, one of the indices j or i will be related to the triple $(x_{k+1}, \check{f}_k(x_{k+1}), \hat{g}_k)$, while the other one can be eliminated from the set $\mathcal{B}_k$. Since the bundle updating incorporates the new index $k+1$ into $\mathcal{B}_{k+1}$, the bundle size will remain $M_{\max}$. This strategy is called* bundle compression*, and it is an efficient manner to keep the auxiliary problem* (11) *easy to solve, without impairing convergence of the algorithm.*

Another proximal bundle method variant designed for two-stage stochastic linear programming is the so called *regularized decomposition*, proposed by Ruszczyński (17). Algorithm 1 is more general than the regularized decomposition in the sense

Algorithm 1 Proximal bundle method

▷ Step 0: initialization
1: Select $\kappa \in (0, 1)$ and $\mathtt{t}_1 \geq \mathtt{t}_{\min} > 0$
2: Choose $x_1 \in X$ and stopping tolerances, tol_ϕ, $\mathrm{tol}_g > 0$
3: Call Oracle 4.1 to compute $(f(x_1), g_1)$
4: Set $\hat{x}_1 \leftarrow x_1$ and $\mathcal{B}_1 \leftarrow \{1\}$,
5: **for** $k = 1, 2, \ldots$ **do**
▷ Step 1: Next iterate
6: Obtain (x_{k+1}, r_{k+1}) and α^k by solving (11)
7: $\hat{g}_k \leftarrow (\hat{x}_k - x_{k+1})/\mathtt{t}_k$
8: $\mathtt{v}_k \leftarrow f(\hat{x}_k) - r_{k+1}$
9: $\phi_k \leftarrow \mathtt{v}_k - \hat{g}_k^\top x_{k+1}$
▷ Step 2: Stopping test
10: **if** $\phi_k \leq \mathrm{tol}_\phi$ and $|\hat{g}_k| \leq \mathrm{tol}_g$) **then**
11: **return** $\hat{x}_k$ and $f(\hat{x}_k)$
12: **end if**
▷ Step 3: Oracle call
13: Call oracle 4.1 to compute $(f(x_{k+1}), g_{k+1})$
14: Define $\mathtt{t}_{\mathtt{aux}} = 2\mathtt{t}_k[1 + (f(\hat{x}_k) - f(x_{k+1}))/\mathtt{v}_k]$
15: **if** $f(x_{k+1}) \leqslant f(\hat{x}_k) - \kappa \mathtt{v}_k$ **then**
16: $\hat{x}_{k+1} \leftarrow x_{k+1}$
17: $\mathtt{t}_{k+1} = \min\{\mathtt{t}_{\mathtt{aux}}, 10\mathtt{t}_k\}$
18: **else**
19: $\hat{x}_{k+1} \leftarrow \hat{x}_k$
20: $\mathtt{t}_{k+1} = \min\{\mathtt{t}_k, \max\{\mathtt{t}_{\mathtt{aux}}, \mathtt{t}_{\min}\}\}$
21: **end if**
▷ Step 4: Bundle management
22: Choose $\mathcal{B}_{k+1} \supset \{j \in \mathcal{B}_k : \alpha_j^k \neq 0\} \cup \{k+1\}$
23: **end for**

that Algorithm 1: (i) does not assume that the problem has a solution; (ii) its stopping test is more sophisticate (see Oliveira et al. (21) for more details); and (iii) its uses the bundle compression mechanism. Convergence analysis of Algorithm 1 can be obtained in Hiriart-Urruty and Lemaréchal (10), and in a more general manner in Oliveira et al. (21).

4.3. Numerical assessment

As mentioned, for the considered planing problem and computer whose description was given in § 3.2.1, formulation (3) becomes computationally intractable for 80 gas demand scenarios or more. To solve the deterministic equivalent problem using $N = 200$ scenarios we then used Algorithm 1 and Oracle 4.1.

The corresponding problem has more than 52 million variables and almost 17 million constraints, and it was solved in 79 iterations for tolerances $\mathrm{tol}_\phi = 10^{-3}\sqrt{n}$ and $\mathrm{tol}_g = 10^{-3}\sqrt{n}$, with $\sqrt{n} = 684$.

Figure 3 shows the evolution of $\{-f(x_k)\}$ along iterations 10 to 79. Circular markers indicate iterations at which a descent-step was found. We recall that for the considered application, minimizing the objective function in problem (3) corresponds to maximizing the total profit of the company: $-f(x)$ is the profit provided by decision x. Algorithm 1 using Oracle 4.1 was able to solve the deterministic equivalent problem, but it took 45 hours and 48 minutes in order to satisfy the stopping

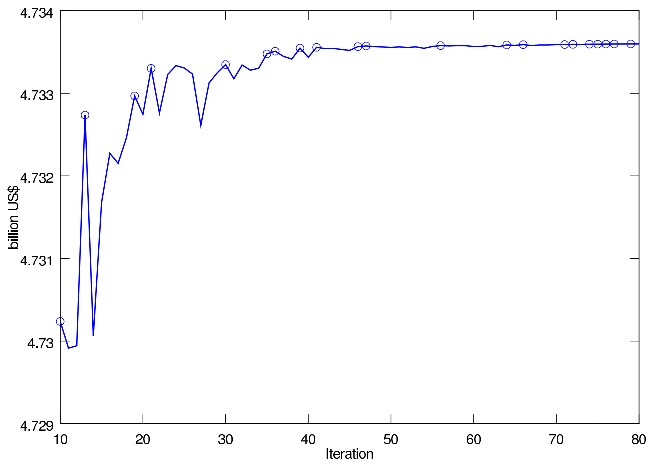

Figure 3: Evolution of $-f(x^k)$, $k = 10, \cdots, 79$.

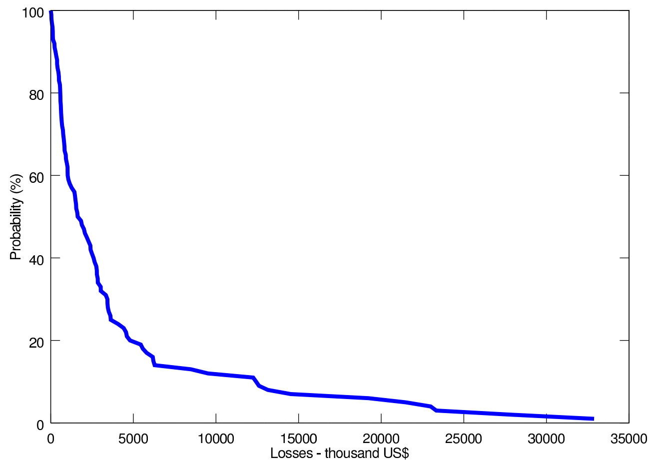

Figure 4: Probability of profitability losses

criteria. It turns out that each call to Oracle 4.1 takes around 34 minutes. Roughly speaking, allocating computational memory and solving problem (8) for a given x and ξ takes 10 seconds of CPU time. Therefore, supposing problem (3) with $N = 1,000$ scenarios (a good representation of the gas demand uncertainties) and estimating 80 iterations for Algorithm 1 to stop, the total CPU time for solving the resulting problem would be more than 9 days. In strategic level of planning, as it is considered in this work, this amount of CPU time may be acceptable. However, other strategies should be tested as an attempt to reduce solving time while still considering uncertainties.

As a first attempt one could suggest to solve a smaller problem by randomly choosing a smaller number of scenarios. This approach was considered as follows. We solved the original problem 100 times considering 10 scenarios randomly chosen from the same reference sample with 200 scenarios used previously. After this, it is possible to estimate the probability of having profit losses by using this kind of scenario selection approach.

Figure 4 shows on the x-axis the loss ℓ and on the y-axis the probability of having a loss greater or equal to ℓ, i.e, the ordinate component of the curve, y, represents

$$y = P(z_{\mathrm{RP}} - f_{200}(\bar{x}_{10}) \geq \ell)$$

with $\bar{x}_{10}$ denoting a solution to problem (3) with 10 scenarios randomly chosen, and f_{200} denoting the objective function in problem (3) considering the reference sample with $N = 200$ scenarios. Therefore, we can see that there is a significant probability – about 20%, of having, for instance, US$ 5 million or more of profitability loss.

In Table 3 some statistics on these losses are presented, such as the average, standard deviation, minimum and maximum values. One can conclude that choosing randomly a 10 scenarios set can lead to losses of up to approximately US$ 33 million. On average, the losses are around US$ 4 million. Thereby, we conclude that a smaller set of scenarios should not be chosen randomly but, as showed below, by using a more sophisticated procedure.

Statistics	(10^3 US$)
Average	3 938.26
Std Deviation	6 168.83
Min	0
Max	32 881.94

Table 3: Statistics on objective function values

5. Optimal scenario reduction

In this section we show how to select efficiently, among all scenarios $\{\xi_1, \xi_2, \ldots, \xi_N\}$ with probability p_i for $i = 1, \ldots, N$, a reduced subset with fewer representative scenarios. Namely, reducing the number of scenarios entails redistributing the initial set of probabilities $P := \{p_i : i = 1, \ldots, N\}$ by taking an index set $I_{\mathtt{rep}} \subset \{1, 2, \ldots, N\}$ such that:

$$\tilde{P} := \{\tilde{p}_i : i = 1, \ldots, N\} \quad \text{with} \quad \begin{cases} \tilde{p}_i = 0 \text{ for } i \notin I_{\mathtt{rep}} \\ \sum_{i \in I_{\mathtt{rep}}} \tilde{p}_i = 1\,. \end{cases}$$

With such a redistribution $\tilde{P}$, the resulting (smaller size) deterministic equivalent problem is

$$\min_x \; \tilde{f}(x) \;\text{ s.t. }\; x \in X \quad \text{with} \quad \tilde{f}(x) = c^\top x + \sum_{i \in I_{\mathtt{rep}}} \tilde{p}_i Q(x, \xi_i)\,. \quad (15)$$

Let $\mathtt{VAL}(P)$ and $\mathtt{VAL}(\tilde{P})$ be the optimal values for the problems (7) and (15), respectively. Accordingly, $S(P) \subset X$ and $S(\tilde{P}) \subset X$ are the respective solution sets (it is assumed that both $S(P)$ and $S(\tilde{P})$ are nonempty sets). Let $N_{\mathtt{rep}} = |I_{\mathtt{rep}}|$ (the number of representative scenarios) be given. Naturally, $N_{\mathtt{rep}} < N$. What one wishes is to determine $\tilde{P}$ such that the distance between the optimal values $|\mathtt{VAL}(P) - \mathtt{VAL}(\tilde{P})|$ and the error $\epsilon \geq 0$ such that $S(\tilde{P}) \subset S(P) + \epsilon B(0, 1)$ are as small as possible (here $B(0, 1)$ stands for the unit ball in $\mathbb{R}^n$). In order to do so, Heitsch and Römisch (14) propose to use a probabilistic metric to estimate the distance between the optimal values, which allows at same time for stability on the solution sets, i.e., a small $\epsilon \geq 0$.

Given $N_{\mathtt{rep}}$, the best redistribution $\tilde{P}$ of P is determined by solving a combinatorial optimization problem, as shown below.

The result in (Dupacová et al. (22), Thm. 2) shows that given a norm or pseudonorm $d : \Xi \times \Xi \to \mathbb{R}_+$ (for instance, $d(\xi_i, \xi_j) = |\xi_i - \xi_j|_2$ can be the Euclidean norm) and a set $I \subseteq \{1, \ldots, N\}$, the particular case of the Monge-Kantorovich functional

$$\mathtt{MK}(I) := \sum_{j \notin I} p_j \min_{i \in I} d(\xi_j, \xi_i)$$

is (except for a multiplicative constant) an upper bound for the distance between the given probability P and the new one $\tilde{P}$, defined by

$$\text{for each } i \in I \quad \tilde{p}_i = p_i + \sum_{j \in J_i} p_j, \qquad (16)$$

$$\text{where} \quad J_i := \{j \notin I : i \in \arg\min_{l \in I} d(\xi_j, \xi_l)\}\,.$$

Providing that the set feasible X is compact, it follows by (22, Eq. (4) and Thm. 2) that

$$\left| \sum_{i=1}^{N} p_i Q(x, \xi_i) - \sum_{i \in I} \tilde{p}_i Q(x, \xi_i) \right| \leq L \cdot \mathtt{MK}(I) \quad \text{for each } x \in X\,,$$

where $L > 0$ is a constant that depends on the problem and on the chosen pseudonorm d. Given the rule (16) for the new probability $\tilde{P}$, a natural criterion for proximity between solutions of problems (7) and (15) should strive to minimize the functional $\mathtt{MK}(I)$, by choosing the best set $I = I_{\mathtt{rep}} \subset \{1, \ldots, N\}$ with $N_{\mathtt{rep}}$ indices. This is a combinatorial optimization problem, and thus difficult to solve. Following Heitsch and Römisch (14), the representative scenarios index set $I_{\mathtt{rep}}$ is chosen by a heuristic method, called *Fast Forward Selection*. More precisely, this work applies Algorithm 2.4 proposed in Heitsch and Römisch (14) to select representative scenarios. The main idea of this algorithm is to iteratively solve problems of the form

$$\min_{I \subset \{1,2,\ldots,N\}} \mathtt{MK}(I) \;\; \text{s.t.} \;\; |I| = N - i\,,$$

for $i = N, N-1, \ldots, N-N_{\mathtt{rep}}$. For more information on scenario reduction for two-stage stochastic programming, see Dupacová et al. (22) and Heitsch and Römisch (14). See also Küchler (23), Heitsch and Römisch (24), Oliveira et al. (25) and Pflug and Pichler (26) for the multistage setting.

5.1. Numerical assessment

By using scenario optimal reduction technique, a subset $\Xi_{N_{\mathtt{rep}}} = \left\{\xi_{j_1}, \xi_{j_2}, \cdots, \xi_{j_{N_{\mathtt{rep}}}}\right\}$ is chosen from the set of $N = 200$ scenarios used in § 4. Then, the reduced deterministic equivalent problem (15) is solved, using only $N_{\mathtt{rep}}$ selected scenarios $\xi \in \Xi_{N_{\mathtt{rep}}}$ with new probability $\tilde{P}$. As usually $N_{\mathtt{rep}} \ll N$, it takes less time to solve (15) than to solve problem (7).

We have considered 7 different instances of problem (15), corresponding to take

$$N_{\mathtt{rep}} \in \{10,\ 20,\ 30,\ 40,\ 50,\ 60,\ 70\}\,.$$

Let $\bar{x}_{N_{\mathtt{rep}}}$ be an optimal investment decision to problem (15) and $\tilde{f}(\bar{x}_{N_{\mathtt{rep}}}) = c^\top \bar{x}_{N_{\mathtt{rep}}} + \sum_{\in I_{\mathtt{rep}}} \tilde{p}_i Q(\bar{x}_{N_{\mathtt{rep}}}, \xi_i)$ be its optimal value.

In Table 4 we compare first stage (investment) cost, $c^\top \bar{x}_{N_{\mathtt{rep}}}$, *expected* second stage costs

$$\sum_{i \in I_{\mathtt{rep}}} \tilde{p}_i Q(\bar{x}_{N_{\mathtt{rep}}}, \xi_i) \;\; \text{and} \;\; \sum_{i=1}^{N} p_i Q(\bar{x}_{N_{\mathtt{rep}}}, \xi_i)\,, \qquad (17)$$

number of variables and constraints, and CPU time for each instance $N_{\mathtt{rep}}$. As expected, the bigger $N_{\mathtt{rep}}$, the better quality solution $\bar{x}_{N_{\mathtt{rep}}}$: compare the values

$$\sum_{i=1}^{200} p_i Q(\bar{x}_{N_{\mathtt{rep}}}, \xi_i) \;\; \text{and} \;\; \sum_{i=1}^{200} p_i Q(\bar{x}_{200}, \xi_i)\,.$$

Moreover, as reported in Table 4 the values in (17) become closer to each other as $N_{\mathtt{rep}}$ increases, showing effectiveness of the scenario reduction technique.

5.2. Optimal investments comparison

In many practical applications, decision makers are concerned not only with optimal values deviation, but also with optimal decision variables deviation. Ultimately, the decision maker is interested in the *optimal policy* - that is, the best first stage decisions. In our application, deviation on optimal solutions might result in very different company's actions: for instance, expanding a pipeline rather than importing gas. Figure 5 shows different decisions for a specific investment project when solving the problem with different number of scenarios.

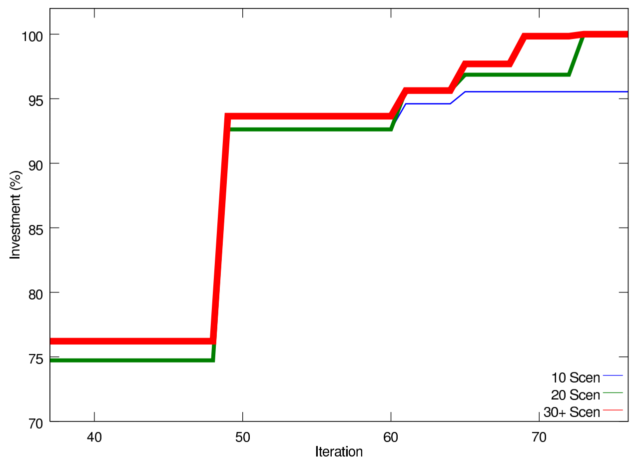

Figure 5: Different investment decisions

The bigger the number of scenarios $N_{\mathtt{rep}}$ is, the closer results are to the optimal decisions (obtained with $N = 200$ scenarios). By using $N_{\mathtt{rep}} = 30$ scenarios or more, optimal decisions are achieved for this investment project.

6. Concluding remarks

In this work we have analyzed the Brazilian natural gas network planning problem. We have provided numerical results

$N_{\texttt{rep}}$	Obj. Function (10^6 US\$)			Problem Size		CPU (s)
	$c^\top \bar{x}_{N_{\texttt{rep}}}$	$\sum_{i \in I_{\texttt{rep}}} \tilde{p}_i Q(\bar{x}_{N_{\texttt{rep}}}, \xi_i)$	$\sum_{i=1}^{200} p_i Q(\bar{x}_{N_{\texttt{rep}}}, \xi_i)$	Variables	Constraints	
10	-8.69	4 367.07	4 733.38	2 610 995	845 080	151
20	-8.70	4 473.30	4 733.54	5 221 305	1 688 980	260
30	-8.72	4 515.01	4 733.60	7 831 615	2 532 880	471
40	-8.72	4 566.87	4 733.60	10 441 925	3 376 780	560
50	-8.72	4 597.55	4 733.60	13 052 235	4 220 680	843
60	-8.72	4 645.87	4 733.60	15 662 545	5 064 580	1 025
70	-8.72	4 692.17	4 733.60	18 272 855	5 908 480	1 207
200	-8.72	4 733.60	4 733.60	52 206 885	16 879 180	164 879

Table 4: Problem comparison for $N_{\texttt{rep}} \in \{10, 20, 30, 40, 50, 60, 70\}$

indicating that a stochastic setting to the problem should be considered. The gain obtained by using a stochastic approach to the problem was estimated to be about US\$ 38 million.

Solving the problem with a small sample of gas demand scenarios ($N = 10$) randomly chosen provides unstable optimal values and investment decisions. On the other hand, solving the problem for bigger samples, say 80 scenarios or more, is only possible via decomposition technique and specialized nonsmooth optimization methods. For $N = 200$ gas demand scenarios the CPU time needed to solve the problem was almost 46 hours, and the estimate of solving time for the case with $N = 1,000$ scenarios is of 9 days, using the computer described in § 3.2.1. The tool used to solve large instances of the decomposed problem was the (state of the art) proximal bundle algorithm, presented with details in § 4.2.

Since the CPU time required to solve the problem might not be affordable for the company's studies, smaller instances of the problem were considered and obtained by applying optimal scenario reduction, described in § 5. We have demonstrated that the use of this strategy can lead to stable results (indicating the same optimal investments) in a considerable smaller solving time.

Finally, we mention that a straightforward extension of this work is the use of the inexact bundle method as in (Oliveira et al. (27)), aiming at reducing solving time for larger instances of the problem.

References

[1] Li X, Armagan E, Tomasgard A, Barton PI. Long-term planning of natural gas production systems via a stochastic pooling problem. In: American Control Conference (ACC), 2010. 2010, p. 429 –35.

[2] Maceira MEP, Duarte VS, Penna DDJ, Moraes LAM, Melo ACG. Ten years of application of stochastic dual dynamic programming in official and agent studies in Brazil – description of the NEWAVE program. In: Power Systems Computation Conference, 2008. 2008, p. 429 –35.

[3] Pereira MVF, Pinto LMVG. Multi-stage stochastic optimization applied to energy planning. Mathematical Programming 1991;52:359–75.

[4] Birge JR, Louveaux FV. Introduction to Stochastic Programming. Springer Series in Operations Research Series; Springer London, Limited; 1997. ISBN 9780387982175.

[5] Sagastizábal C. Divide to conquer: decomposition methods for energy optimization. Mathematical Programming 2012;134:187–222.

[6] Santoso T, Ahmed S, Goetschalckx M, Shapiro A. A stochastic programming approach for supply chain network design under uncertainty. European Journal of Operational Research 2005;167(1):96 – 115.

[7] Kleywegt AJ, Shapiro A, Homem-de Mello T. The sample average approximation method for stochastic discrete optimization. SIAM Journal on Optimization 2002;12(2):479–502.

[8] Schütz P, Tomasgard A, Ahmed S. Supply chain design under uncertainty using sample average approximation and dual decomposition. European Journal of Operational Research 2009;199(2):409 –19.

[9] Kelley Jr. J. The cutting-plane method for solving convex programs. Journal of the Society for Industrial and Applied Mathematics 1960;8(4):703–12.

[10] Hiriart-Urruty JB, Lemaréchal C. Convex Analysis and Minimization Algorithms. No. 305-306 in Grund. der math. Wiss; Springer-Verlag; 1993. (two volumes).

[11] Li X, Chen Y, Barton PI. Nonconvex generalized benders decomposition with piecewise convex relaxations for global optimization of integrated process design and operation problems. Industrial & Engineering Chemistry Research 2012;51(21):7287–99.

[12] Li X, Tomasgard A, Barton P. Nonconvex generalized benders decomposition for stochastic separable mixed-integer nonlinear programs. Journal of Optimization Theory and Applications 2011;151:425–54.

[13] Slyke RV, Wets RB. L-shaped linear programs with applications to optimal control and stochastic programming. SIAM Journal of Applied Mathematics 1969;17:638–63.

[14] Heitsch H, Römisch W. Scenario reduction algorithms in stochastic programming. Computational Optimization and Applications 2003;24:187–206.

[15] Shapiro A, Dentcheva D, Ruszczyński A. Lectures on Stochastic Programming. Modeling and Theory; vol. 9 of *MPS-SIAM series on optimization*. SIAM and MPS, Philadelphia; 2009.

[16] Bonnans J, Gilbert J, Lemaréchal C, Sagastizábal C. Numerical Optimization: Theoretical and Practical Aspects. Springer-Verlag; 2nd ed.; 2006.

[17] Ruszczyński A. A regularized decomposition method for minimizing a sum of polyhedral functions. Math Program 1986;35:309–33.

[18] Fábián CI. Bundle-type methods for inexact data. In: Proceedings of the XXIV Hungarian Operations Researc Conference (Veszprém, 1999); vol. 8 (special issue, T. Csendes and T. Rapcsk, eds.). 2000, p. 35–55.

[19] Fábián C, Szőke Z. Solving two-stage stochastic programming problems with level decomposition. Computational Management Science 2007;4:313–53.

[20] Oliveira W, Sagastizábal C. Level bundle methods for oracles with on demand accuracy. Available at http://wwwoptimization-onlineorg/DB_HTML/2012/03/3390html 2012;.

[21] Oliveira W, Sagastizábal C, Lemaréchal C. Bundle methods in depth: a unified analysis for inexact oracles. Available at http://wwwoptimization-onlineorg/DB_HTML/2013/02/3792html 2013;.

[22] Dupacová J, Gröwe-Kuska N, Römisch W. Scenario reduction in stochastic programming: An approach using probability metrics. Mathematical Programming 2003;95:493–511.

[23] Küchler C. On stability of multistage stochastic programs. SIAM J on

Optimization 2008;19(2):952–68.
[24] Heitsch H, Römisch W. Scenario tree reduction for multistage stochastic programs. Computational Management Science 2009;6:117–33.
[25] Oliveira WL, Sagastizábal C, Penna DDJ, Maceira MEPn, Damázio JM. Optimal scenario tree reduction for stochastic streamflows in power generation planning problems. Optimization Methods and Software 2010;25(6):917–36.
[26] Pflug G, Pichler A. A distance for multistage stochastic optimization models. SIAM Journal on Optimization 2012;22(1):1–23.
[27] Oliveira W, Sagastizábal C, Scheimberg S. Inexact bundle methods for two-stage stochastic programming. SIAM Journal on Optimization 2011;21(2):517–44.