Procedia Computer Science
Volume 80, 2016, Pages 1875–1886
ICCS 2016. The International Conference on Computational Science

A Stochastic Approach to Solving Bilevel Natural Gas Cash-Out Problems

Vyacheslav Kalashnikov[1,2,3], Vassil Alexandrov[4*] , and Nataliya Kalashnykova[5]
[1]*Tecnológico de Monterrey (ITESM), Campus Monterrey, Mexico*
[2]*Central Economics and Mathematics Institute (CEMI), Moscow, Russia*
[3]*Sumy State University, Sumy, Ukraine*
[4]*ICREA-BSC, C/Jordi Girona 29, Barcelona, Spain*
[5]*Universidad Autónoma de Nuevo León (UANL), Monterrey, Mexico*
kalash@itesm.mx, vassil.alexandrov@bsc.es, nkalash2009@gmail.com

Abstract
We study a special bilevel programming problem that arises in transactions between a Natural Gas Shipping Company and a Pipeline Operator. Because of the business relationships between these two actors, the timing, and objectives of their decision-making process are different. In order to model that, bilevel programming was traditionally used. Apart from the theoretical studies of the problem to facilitate its solution a linear reformulation is required, as well as heuristic approaches, and branch-and-bound techniques may be applied. We present a linear programming reformulation of the latest version of the model, which is easier and faster to solve numerically. This reformulation makes it easier to theoretically analyze the problem, allowing us to draw some conclusions about the nature of the solution.

Since elements of uncertainty are definitely present in the bilevel natural gas cash-out problem, its stochastic formulation is developed in the form of a bilevel multi-stage stochastic programming model with recourse. After reducing the original formulation to a bilevel linear problem, a stochastic scenario tree is defined by its node events, and time series forecasting is used to produce stochastic values for data of natural gas price and demand. Numerical experiments were run to compare the stochastic solution with the perfect information solution and the expected value solutions.

Keywords: Operations Research, Big Data, natural gas cash-out problems; stochastic algorithms

[*] Distinguished Visiting Professor, Tecnológico de Monterrey (ITESM), Campus Monterrey

Selection and peer-review under responsibility of the Scientific Programme Committee of ICCS 2016

doi:10.1016/j.procs.2016.05.483

1 Introduction

Since 1992 in the United States (EIA, 1992[a], 1992[b], 2005), (Soto, 2008), and since 1995 in the European Union (IHS, 2007), the society regulating bodies issued and endorsed a concatenation of special rules aiming to separate the cardinal operations inherent for the natural gas supply chain (*see* the details in Kalashnikov *et al*., 2010[b]). The related merchandises gravely required separation of the transportation and sale processes. As a consequence of such an epitomic change and the due transformations of the gas markets, many new attributes and subtle details appeared that urgently required a thorough study.

Among the most pertinent conundrums the natural gas supply chain stumbled about was that of equilibrating the gas volumes supplied through a pipeline network. Such an equilibrating design is of great importance for the pipeline operating body (POB), due to the fact that under a well-marshalled gas transfer via the pipeline a strict control of its amount is indispensable. Moreover, each (natural gas) shipping company (SC) is also concerned about the hauled volumes being as insignificantly imbalanced as possible. A natural gas shipping company's business is to distribute gas by transferring it through pipelines to customers: it is obliged to operate in line with signed contracts first, and after that it can sell the exceeding gas volumes to earn a maximum gain. Aiming at that, the SC has to arrange the gas supply at every selling point (called *pipeline meters*) having in mind the balance, the selling prices, and the net profit. The comprehensive mathematical framework of this problem can be found, e.g., in (Kalashnikov *et al*., 2005).

Although natural gas pipeline networks have been thoroughly studied, most of the well-known models concentrate mainly on the material movement facets of the natural gas supply chain and less on the SC–POB cooperation issues: *cf.*, e.g., the optimization of network operations (Borraz-Sánchez and Ríos-Mercado, 2005), (Chebouba *et al*., 2009), or stationing appliances (Kabirian and Hemmati, 2007). In some other publications, the natural gas supply chain is implemented as a multilevel structure where both SC and POB are present and make decisions as partners; *see*, e.g., (Gabriel *et al*., 2005) and (Egging *et al*., 2008). In these works, the authors span the whole supply chain with more attention to the traders (financial problems of the natural gas producers), so that imbalances in the system (resulting from the interaction between the SCs and POBs) aren't investigated.

A lot of authors admit (*cf*., Arano and Blair, 2008; Hawdon, 2003) the existence of some new difficulties in the SC–POB system generated by the principal changes, yet it is possible to find only a small number of references that provide reliable instruments to deal with the above-mentioned problems. The paper (Esnault, 2003), for example, shows that the SC requires a depot if the gas volume maintaining rules are too restrictive, either due to some business habits or owing to some technical blemishes. Nevertheless, the crucial part of the modern natural gas supply chain management is to keep a balance as strictly as possible when carrying out contracts. However, up to date, no policy has been yet admitted as comprehensive about the mode how the imbalances generated by the SC, are materially and economically managed. Essential devices assisting the POB in the restoration of the balance are the arbitrage punishment approach allowing that the POB rearrange the imbalances inside the system and levies some fees to the shipper.

In (Kalashnikov and Ríos-Mercado, 2001), a model is developed to fit better the punishment part of the imbalance problem. There, the penalization deals only with the cash-out taking place between the SC and POB bearing no reference to real market conditions that are very important to the shipper (SC). The paper provides a solution algorithm for a somewhat modified problem, together with the examination of how this modification affects the objective function and the obtained solutions. In (Kalashnikov and Río-Mercado, 2006), the authors collate two procedures solving the specified problem. In (Dempe *et al*., 2005), the imbalance cash-out problem was split into several generalized transportation problems, which made it easier to calculate the optimal solution.

In (Dempe *et al*., 2015) and (Kalashnikov *et al*., 2010[a]), an extended version of the gas imbalance cash-out problem was studied, in which the upper- level objective function involves extra terms reflecting the leader's (the shipping company's) expected net profit. Nonetheless, that scheme presupposes the perfect information about the tendencies of gas price along time, which is hardly possible in the real life systems. Even more, this assumption isn't very useful, because the resulting function doesn't clearly explain the logic behind the shipper's activity. Hence, in the framework of our previous paper (Kalashnikov *et al*., 2010[b]), we describe here a stochastic formulation of the cash-out problem where the shipper is able to predict (to some degree of accuracy) the natural gas demand within a series of intervals of time, which helps plan the volumes of gas to extract from the pipeline meters. The model in question is a stochastic variation of the original mixed-integer bilevel optimization problem, and two algorithms to solve it are presented and discussed.

To our best knowledge, there are few literature sources explicitly dealing with the shipper-pipeline subsystem by formulating a bilevel optimization problem involving the operations to resume the balance. For stochastic optimization methods, it is especially important that the number of the upper-level variables be not too high, in order to save on the number of branches of scenario trees. Recent results obtained in this direction are also mentioned in this paper.

The remaining part the manuscript has the following structure. After having specified the examined model in Section 2, a mathematical procedure of solution is developed in Section 3. The paper ends with the conclusions and the aims for the future research, followed by an acknowledgment and the list of references.

2 Model Specification

Proceeding along the lines previously elaborated in (Kalashnikov *et al*., 2010[b]), we examine a bilevel programming model, in which the upper level decision maker (the leader), namely, the natural gas shipping company (SC) buys the gas, supplies it into an (interstate) pipeline at its starting meter station and extracts certain amounts of gas from the pipeline meters in several pool zones embracing inhabited territories (consumers). At the same time, the pipeline management (called the pipeline operating body, or POB) behaves as a follower. In more detail, the POB permits the SC to extract quantities of natural gas that may not match exactly with the originally claimed amounts, which often leads to imbalances of the different sign (positive or negative).

Such a policy is indispensable for the fuel market flexibility and the operational dynamics within the natural gas supply chain. Albeit, these imbalances use to conduct to (unexpected) excess expenditures suffered by both the SC and POB. Therefore, the shipping company considers doing that only if the future market conditions forecasts hint that the aggregate gross revenues may prevail over the financial losses triggered by the POB's penalization procedures towards the shipper.

Hence, the pair SC–POB operates according to the following mode:

1) The SC produces a solid foretelling of the gas demand for next time period and estimates probabilities for different scenarios and trends.

2) The SC books a certain amount of natural gas at each pool zone for every time slot of the planned horizon.

3) With respect to every consecutive time slot, the shipper sets what volume of gas to excerpt and trade. Any aberrations of these volumes from those stipulated in the agreement with the POB use to bring about (positive and/or negative, daily and/or final) imbalances at the meters in question.

4) The pipeline (POB) traces the emerging daily and terminate imbalances and is in its right of revamping them according to the appropriate business rules.

5) The POB charges the shipper with a penalty for the final (reshuffled) imbalances, which may also occur negative, i.e., the pipeline returns its debt to the SC.

6) The shipper (SC) reckons up the net profit as its aggregate revenue minus the penalty.

The above-outlined paradigm is referred to as a bilevel multi-stage stochastic optimization problem (Kall and Wallace, 1994). Here, the shipper is the upper level decision maker (the leader) maximizing its net profit as the difference between the gross revenue from the sales of its gas in the pipeline and the possible penalty charged by the POB. On the other hand, the pipeline (POB) plays at the lower level (being the follower) by trying to minimize the absolute figure of the penalty cash-out flow, either positive or negative. The first stage of the stochastic process consists in booking the supply by the shipper; these reserved supply fixed during the whole time horizon. At all the subsequent stages, the decision variables are the daily excerpt volumes, unsatisfied demand, and the penalty cash-outs charged by the pipeline.

.

Remark 1. Even though the pipeline (POB) seems to play a more important role in the above-described system, still the shipper (SC) is the leader in the bilevel (Stackelberg duopoly) model. In order to conclude which of the parties enjoys the more or less important influence, one has to consider and analyze many aspects of different processes in a system. In our model, the timing of the decision making is pivotal for the classification of who is the principal and who accepts the follower's role, because the latter determines whose actions are dependent on whose resolution.

A. Notation

Throughout this paper, the following notation is used:

Sets

P The quantity of time periods at each node; $P \in Z_{++}$;

N The number of pool zones; $N \in Z_{++}$;

L The number of event nodes in the process; $L \in Z_{++}$;

B The number of stages in the process; $B \in Z_{++}$;

$\mathbf{T}$ Set of time periods at any given node; $\mathbf{T} = \{1, 2, \ldots, P\}$;

$\mathbf{J}$ Set of pool zones; $\mathbf{J} = \{1, 2, \ldots, N\}$;

$\mathbf{L}$ Set of event nodes; $\mathbf{L} = \{1, 2, \ldots, L\}$;

$\mathbf{L}^{\beta}$ Set of nodes at stage β; $\beta = 1, 2, \ldots, B$.

Upper Level Parameters

$x_{\ell tj}^{Lower}, x_{\ell tj}^{Upper}$ Lower and upper bounds for the imbalances on day t at node ℓ, in pool zone j; $j \in \mathbf{J}$; $t \in \mathbf{T}$; $\ell \in \mathbf{L}$;

$x_{\ell t}^{Lower}, x_{\ell t}^{Upper}$ Lower and upper bounds for the total imbalances on day t at node ℓ; $t \in \mathbf{T}$; $\ell \in \mathbf{L}$;

$s_{\ell tj}^{Lower}, s_{\ell tj}^{Upper}$ Lower and upper bounds on imbalance swings on day t at node ℓ, in pool zone j; $j \in \mathbf{J}$; $t \in \mathbf{T}$; $\ell \in \mathbf{L}$;

x_{0j} Initial imbalance at the beginning of day 1 at node 1, in pool zone j; $j \in \mathbf{J}$;

$ED_{\ell tj}$ Expected demand on day t, at node ℓ, in pool zone j; $j \in \mathbf{J}$; $t \in \mathbf{T}$; $\ell \in \mathbf{L}$;

$\pi_{\ell tj}$ Unit price for the gas extracted/sold in line with the contract on day t at node ℓ in zone j; $j \in \mathbf{J}$; $t \in \mathbf{T}$; $\ell \in \mathbf{L}$;

$RC_{\ell tj}, BC_{\ell tj}$ Recourse and booking capacity costs per gas unit on day t at node ℓ, in pool zone j; $j \in \mathbf{J}$; $t \in \mathbf{T}$; $\ell \in \mathbf{L}$;

p_ℓ The probability of node ℓ to appear in any scenario; $\ell \in \mathbf{L}$.

Lower Level Parameters

f_{ij} Fraction of gas used as the fuel being hauled from pool zone i to pool zone j; $i, j \in \mathbf{J}$; $i < j$;

v_{ij} Forward haul unit cost for moving gas from pool zone i to pool zone j; $i, j \in \mathbf{J}$; $i < j$;

n_{ij} Backward credit for "returning" a unit of gas from pool zone j to pool zone i; $i, j \in \mathbf{J}$; $i < j$;

m_j Cash-out penalty coefficient in pool zone j; $j \in \mathbf{J}$.

Upper-Level Decision Variables

$x_{\ell tj}$ Imbalance on day t at node ℓ in pool zone j; $j \in \mathbf{J}$; $t \in \mathbf{T}$; $\ell \in \mathbf{L}$;

$s_{\ell tj}$ Imbalance swing on day t at node ℓ in pool zone j; $j \in \mathbf{J}$; $t \in \mathbf{T}$; $\ell \in \mathbf{K}$;

$AE_{\ell tj}$ Amount of gas actually extracted on day t at node ℓ in pool zone j; $j \in \mathbf{J}$; $t \in \mathbf{T}$; $\ell \in \mathbf{L}$;

APE_{kti} Amount of gas planned to be extracted (i.e., the booked pipeline capacity) on day t at node k in pool zone i; $i \in \mathbf{J}$; $t \in \mathbf{T}$; $k \in \mathbf{K}$;

$AS_{\ell tj}$ Amount of gas actually extracted and sold on day t at node ℓ in pool zone j; $j \in \mathbf{J}$; $t \in \mathbf{T}$; $\ell \in \mathbf{L}$;

$UD_{\ell tj}$ The part of demand $ED_{\ell tj}$ unmet on day t at node ℓ in pool zone j; $j \in \mathbf{J}$; $t \in \mathbf{T}$; $\ell \in \mathbf{L}$.

Lower Level Decision Variables

z_j Final imbalance in pool zone j; $j \in \mathbf{J}$;

u_{ij} Amount of gas moved from pool zone i to pool zone j; $i, j \in \mathbf{J}$; $i < j$;

b_{ij} Amount of gas decided to "return" from pool zone j to pool zone i; $i, j \in \mathbf{J}$; $i < j$;

y Total cash-out charge imposed on the shipper (SC) by the pipeline (POB).

Auxiliary Variables

q A binary variable equal to 1(0) if the final imbalances z_j are all non-negative (non-positive). In the special case when $z_j = 0$, $\forall j \in \mathbf{J}$, we set $q = 1$.

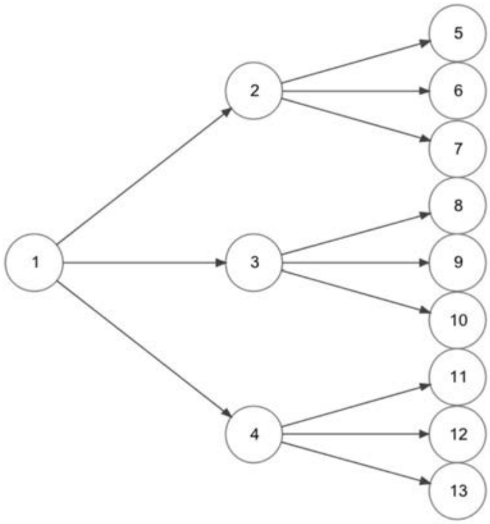

Fig. 1. A 3-staged, ternary scenario tree.

Figure 1 illustrates a three-staged scenario tree generating nine scenarios. Scenario 1 "passes through" nodes 1, 2, and 5; scenario 2 through nodes 1,2, and 6, and so on. Every node is endowed with the probability of occurring to be equal to the sum of probabilities of all the scenarios "passing" through it; the leaf (termination) nodes occurs with the same probabilities that the corresponding scenarios finishing at them.

Since the probabilities through all the scenarios sum up to 1, one has $\sum_{\ell \in \mathbf{L}^{\beta}} p_{\ell} = 1$, where $\ell \in \mathbf{L}^{\beta}$ enumerates the set of all leaf nodes (all them related to the last stage *B*.) The root evidently has the probability of 1. Therefore, for each consecutive stage, the sum of probabilities of all nodes at that stage also gives 1. The probabilities p_{ℓ} depend on the scenario tree forecast scheme.

The function $\alpha : \mathbf{L} \to \mathbf{L}$ determines the predecessor of any node ℓ (except the root, whose predecessor doesn't exist); that is, $\alpha(\ell) = \ell'$ if ℓ and ℓ' are the extremes of an arc in the scenario tree, and $\ell \in \mathbf{L}^{\beta} \Rightarrow \ell' \in \mathbf{L}^{\beta-1}$, $\beta = 2,3,\ldots,B$. In the particular case of ℓ = 1, we set $x_{\alpha(1),P,j} := x_{0j}, j \in \mathbf{J}$.

Because the costs RC $,BC$ aren't random variables, they have the same values at every node for a fixed stage β, that is, $RC_{\ell tj} = RC_{\ell' tj}, BC_{\ell tj} = BC_{\ell' tj}$, for all $\ell, \ell' \in \mathbf{L}^{\beta}, \beta = 2,3\ldots,B; j \in \mathbf{J}$. The same holds true for the variables *APE*, as explained below in (1i).

Lastly, the node imbalance matrix x_{ℓ} is defined by fixing one node ℓ and arranging the imbalance matrix elements in the following form: $x_{\ell} = \left(x_{\ell tj}\right)_{t=1,j=1}^{P \quad N}$, while the last day imbalance vector $x_{\ell P}$ is gathered in a classic way: $\left(x_{\ell Pj}\right)_{j \in \mathbf{J}}$.

3 Problem Statement

Following the lines of the model described in (Kalashnikov and Ríos-Mercado, 2001), we propose a multi-stage stochastic bilevel optimization problem defined below with (1a)–(1j) and (4a)–(4l). After that, we will reduce the model to a bilevel linear program, the lower level of which may be recognized as a generalized transportation problem, or a quadratic assignment problem. It is worthy to mention that bilevel programs defined over networks (see, e.g., Chiou, 2005; and Cruz et al., 1999) often arise when studying transportation problems; cf., also (Yang and Bell, 2001), (Ben-Ayed et al., 1988).

A. Upper Level Model

Relationships (1b)–(1j) present the upper level of the bilevel program. The upper level problem is stochastic and reflects a ternary scenario tree similar to that depicted in Figure 1.

$$Min\ W_1\left(x,s,AE,APE,UD;z,u,b,y,a,q\right)=$$

$$=\sum_{\ell\in\mathbf{L}} p_\ell\left[\sum_{t\in\mathbf{T}}\sum_{j\in\mathbf{J}}\left(BC_{\ell tj}UD_{\ell tj}-\pi_{\ell tj}min\left\{AE_{\ell tj},UD_{\ell tj}\right\}\right)+RC_{\ell tj}APE_{\ell tj}\right]+ \quad (1a)$$

$$+\sum_{\ell\in\mathbf{L}^\beta} p_\ell w\left(x_{\ell P};z,u,b,z,a,q\right), \quad (1b)$$

subject to

$$x_{\ell tj}^{Lower}\le x_{\ell tj}\le x_{\ell tj}^{Upper},\ \ \ell\in\mathbf{L},t\in\mathbf{T},j\in\mathbf{J}; \quad (1c)$$

$$s_{\ell tj}^{Lower}\le s_{\ell tj}\le s_{\ell tj}^{Upper},\ \ \ell\in\mathbf{L},t\in\mathbf{T},j\in\mathbf{J}; \quad (1d)$$

$$x_{\ell t}^{Lower}\le\sum_{j\in\mathbf{J}}x_{\ell tj}\le x_{\ell t}^{Upper},\ \ \ell\in\mathbf{L},t\in\mathbf{T}; \quad (1e)$$

$$x_{\ell tj}=\begin{cases}x_{\alpha(\ell),P,j}+s_{\ell tj}, & \text{if } t=1;\\ x_{\ell,t-1,j}+s_{\ell tj}, & \text{otherwise;}\end{cases}\quad \ell\in\mathbf{L},t\in\mathbf{T},j\in\mathbf{J}; \quad (1f)$$

$$x_{\ell tj}=APE_{\ell tj}-AE_{\ell tj},\ \ell\in\mathbf{L},t\in\mathbf{T},j\in\mathbf{J}; \quad (1g)$$

$$UD_{\ell tj}=max\left\{0,ED_{\ell tj}-AE_{\ell tj}\right\},\ell\in\mathbf{L},t\in\mathbf{T},j\in\mathbf{J}; \quad (1h)$$

$$APE_{\ell tj}=APE_{\ell' tj},\ \ell,\ell'\in\mathbf{L}^\beta,\beta=1,\dots,B;t\in\mathbf{T},j\in\mathbf{J}; \quad (1i)$$

$$AE_{\ell tj}\ge 0,\ \ell\in\mathbf{L},t\in\mathbf{T},j\in\mathbf{J}. \quad (1j)$$

In (1b), the terms $w\left(x_{\ell P};z,u,b,y,a,q\right)=y$ combine the variables from both levels and have the meaning of the optimal solutions of the corresponding lower level problems (4a) – (4l). In other words, they represent the penalty cash-out flows that the shipper (SC) has to pay to the pipeline (POB). The product $\pi_{\ell tj}\ min\left\{AE_{\ell tj},ED_{\ell tj}\right\}$ is the revenue earned with the gas excerpt out from the wells and sold in pool zone *j* on day *t* at node ℓ. Further, $BC_{\ell tj}UD_{\ell tj}$ is the full value of the unsatisfied demand, which could be interpreted either as the redemption of the consumers, or the cost of a stock purchased from a third party in the amount just sufficient to meet the demand completely.

The term $RC_{\ell tj}APE_{\ell tj}$ means the expenses suffered by the SC when booking capacity in the pipeline on day *t* at node ℓ in pool zone *j*. The latter term is especially important to guarantee the non-triviality of solutions. Indeed, if the capacity booking were free for the shipper, the latter would simply increase the values of variables *APE* and trade the available gas at the price *π*.

It is noteworthy that if there are multiple stages, then $\sum_{\ell \in \mathbf{L}} p_\ell > 1$. This inequality holds because the first term of the objective function is an expected value not *over the nodes* but rather *over the scenarios*. Here, the non-anticipativity constraints (*cf.*, Kall and Wallace, 1994) have been implicitly exposed by applying the node formulation instead of a scenario one.

Constraints (1c)–(1e) impose the technological bounds on the generated imbalances at each node, as well as on their daily totals, and the daily imbalance swings. Equalities (1f) describe the structure of the everyday imbalance swings. At every node, time period and pool zone, the imbalance on one day isn't allowed to vary too much day after day; this is outlined in different ways depending upon whether one is at an in-node (i.e., when $t = 2,\dots,P$) or at a cross-node (when $t = 1$).

If we are at an in-node, then the swing of the imbalance from $x_{\ell,t-1,j}$ to $x_{\ell tj}$ is determined with the value of the swing variable $s_{\ell tj}$. The latter is bounded from above and below with the values of parameters $s_{\ell tj}^{Upper}$ and $s_{\ell tj}^{Lower}$, respectively. It also holds for the cross-nodes but the day before the first day at such a node is the last day at the predecessor node, hence, one has to use the update formula $x_{\ell tj} = x_{\alpha(\ell),P,j} + s_{\ell tj}$.

We also remark the following property of the imbalance swing variables *s*: their values are completely determined by the imbalance variables *x*, as combined with the predefined initial imbalances x_0. So we introduce the swing variables merely to make the imbalance movements more transparent, but in fact, they are not necessary for the solution process. When solving the modified upper level problem as explained later with equations (2b)–(2j), we simply drop these variables in favour of equivalent bounds on imbalances *x*.

Equations (1g) illustrate the interrelations among the imbalance, the booked capacity, and the excerption from each pool zone within every time slot, while (1h) calculates the unmet demand values. Then, constraint (1i) characterizes the one-stage feature of the variables *APE*, in the sense that each node on a given stage must boast the same value of those variables.

Even though program (1a)–(1j) is an appealing model of the SC–POB subsystem, it is nonlinear due to the presence of the operators *max* and *min*, which makes it much harder to solve as compared to, say, a linear programming problem.

Nevertheless, we can diminish the considered program's computational difficulty by introducing certain auxiliary variables in such a way that the objective function isn't affected. The equivalent problem gets rid of *max* or *min* operators, yet it remains nonlinear because the upper level objective function contains the variable *y* controlled by the lower level as an optimal response to the upper level actions. Now we are going to eliminate *y* from the upper level objective function in order to reduce our problem to a completely linear model. Moreover, once the lower level can be also reformulated as a linear program, the resulting bilevel model is called a *bilevel linear programming* problem.

One can qualify the optimization problem (2a)–(2j) as an "almost" linear program equivalent to the original upper level formulation (as is explained in detail in Kalashnikov *et al.* 2010[a], 2010[b]).

$$Min\ W_2\left(x,AE,AS;z,u,b,y,a,q\right)=$$

$$= \sum_{\ell \in \mathbf{L}} p_\ell \left[\sum_{t \in \mathbf{T}} \sum_{j \in \mathbf{J}} BC_{\ell tj}\left(UD_{\ell tj} - AS_{\ell tj}\right) - \pi_{\ell tj} AS_{\ell tj} + RC_{\ell tj}\left(x_{\ell tj} + AE_{\ell tj}\right) \right] + \tag{2a}$$

$$+ \sum_{\ell \in \mathbf{L}^\beta} p_\ell w\left(x_{\ell P}; z,u,b,y,a,q\right), \tag{2b}$$

subject to

$$x_{\ell tj}^{Lower} \le x_{\ell tj} \le x_{\ell tj}^{Upper}, \quad \ell \in \mathbf{L}, t \in \mathbf{T}, j \in \mathbf{J}; \tag{2c}$$

$$x_{\ell t}^{Lower} \le \sum_{j\in\mathbf{J}} x_{\ell tj} \le x_{\ell t}^{Upper}, \quad \ell \in \mathbf{L}, t \in \mathbf{T}; \tag{2d}$$

$$s_{\ell 1j}^{Lower} \le x_{\ell 1j} - x_{\alpha(\ell),P,j} \le s_{\ell 1j}^{Upper}, \quad \ell \in \mathbf{L}, j \in \mathbf{J}; \tag{2e}$$

$$s_{\ell tj}^{Lower} \le x_{\ell tj} - x_{\ell,t-1,j} \le s_{\ell tj}^{Upper}, \ t = 2,\dots,P; \ \ell \in \mathbf{L}, j \in \mathbf{J}; \tag{2f}$$

$$AS_{\ell tj} \le ED_{\ell tj}, \ \ell \in \mathbf{L}, t \in \mathbf{T}, j \in \mathbf{J}; \tag{2g}$$

$$AS_{\ell tj} \le AE_{\ell tj}, \ \ell \in \mathbf{L}, t \in \mathbf{T}, j \in \mathbf{J}; \tag{2h}$$

$$x_{\ell tj} + AE_{\ell tj} = x_{\ell' tj} + AE_{\ell' tj}, \ \ell, \ell' \in \mathbf{L}^{\beta}, \beta = 1,\dots,B; t \in \mathbf{T}, j \in \mathbf{J}; \tag{2i}$$

$$AE_{\ell tj} \ge 0, \ x_{\ell tj} \ge 0, \ell \in \mathbf{L}, t \in \mathbf{T}, j \in \mathbf{J}. \tag{2j}$$

Here $w(x_{\ell P}; z,u,b,y,a,q) = y$ is the lower level optimal response to the upper level move, as specified below in (3a)–(3l).

B. Lower Level Model

The lower level problem exactly coincides with the one examined in (Kalashnikov and Ríos-Mercado, 2006); it uses the linear objective function a to substitute equivalently the absolute-value lower level objective function with the help of inequalities:

$$Min \ w_1(x_{\ell P}; z,u,b,y,a,q) = a \tag{3a}$$

subject to

$$z_j = x_{\ell Pj} + \sum_{i\in\mathbf{J}: i<j} \left[(1 - f_{ij}) u_{ij} - b_{ij} \right] + \sum_{k\in\mathbf{J}: k>j} (b_{jk} - u_{jk}), \ j \in \mathbf{J}; \tag{3b}$$

$$\sum_{j\in\mathbf{J}: j>i} u_{ij} + \sum_{k\in\mathbf{J}: k<i} b_{ki} \le max\{0, x_{\ell Pi}\}, \ \ell \in \mathbf{L}, i \in \mathbf{J}; \tag{3c}$$

$$u_{ij} \le \begin{cases} x_{\ell Pi}, & \text{if } x_{\ell Pi} > 0 \text{ and } x_{\ell Pj} < 0; \\ 0, & \text{otherwise;} \end{cases} \quad \ell \in \mathbf{L}; i, j \in \mathbf{J}, i < j; \tag{3d}$$

$$b_{ij} \le \begin{cases} x_{\ell Pi}, & \text{if } x_{\ell Pj} < 0 \text{ and } x_{\ell Pi} > 0; \\ 0, & \text{otherwise;} \end{cases} \quad \ell \in \mathbf{L}; i, j \in \mathbf{J}, i < j; \tag{3e}$$

$$min\{0, x_{\ell Pi}\} \le z_i \le max\{0, x_{\ell Pi}\}, \ \ell \in \mathbf{L}; \ i \in \mathbf{J}; \tag{3f}$$

$$-\mathbf{M}_1(1-q) \le z_i \le \mathbf{M}_1 q, \ \ i \in \mathbf{J}; \tag{3g}$$

$$y = -\sum_{i\in\mathbf{J}} m_i z_i - \sum_{(i,j): i<j} n_{ij} b_{ij} + \sum_{(i,j): i<j} v_{ij} (1 - f_{ij}) u_{ij}; \tag{3h}$$

$$-a \le z \le a; \tag{3i}$$

$$y, z_i \in R; \ i \in \mathbf{J}; \tag{3j}$$

$$u_{ij} \ge 0, b_{ij} \ge 0, \ i, j \in \mathbf{J}, i < j; \tag{3k}$$

$$q \in \{0,1\}. \tag{3l}$$

Here, $\mathbf{M}_1 > 0$ is a large fixed scalar parameter (an analog of the Big-*M* concept in linear programming).

4 Conclusions

In this paper, we present a bilevel multi-stage stochastic optimization model, which is developed to deal with a certain subsystem of the Natural Gas Supply Chain. While former models were focused on the arbitrage policies in a deterministic setting, here we have expanded the problem to include such elements as gas sales and booking costs and added a stochastic framework to model the uncertainty in demand and prices faced by the upper level decision maker (the leader).

The developed model was implemented numerically and compared to the Perfect Information Solution (PIS) and the Expected Value Solutions (EVS). Experimental findings show that 19 of the 21 instances deliver implementation values of over half of the PIS, whereas only one of the EVS implementation values has a relative error below 0.75. The Stochastic Solution Implementation values are better than those of the EVS values in all but one case – which corresponds to the simplest instance tested, – which testifies in favor of our approach. The performed linear reformulation also proved advantageous, as solving the original model with nonlinear levels takes considerably longer time and does not provide better solutions after up to 10 hours of running time in 20 of the 21 experiments.

Future work includes assessing the convenience of using heuristic approaches for solving the lower level (as opposed to using a specialized linear solver,) and reformulating the linear lower level in the form of its duality conditions, adding these to the upper level to solve a single-level problem instead of a bilevel one. We also intend to study these models under different time series not showing seasonality is also planned, as it is the implementation of a rolling horizon approach to remedy the lack of accuracy over long-period problems. In addition, we present here our first steps in the direction of development of techniques allowing one to reduce the quantity of the upper level variables thus decreasing essentially the number of branches of the scenario trees.

Another prospective direction of future research is to study some approaches helping one to reduce the dimension of the upper level problem in a bilevel optimization model. This question is important because sometimes the dimension of the upper level is a critical parameter for applications of stochastic optimization algorithms.

Indeed, if a stochastic process relies on generation of scenario trees, the quantity of tree branches/nodes grows exponentially along with the increasing number of upper level variables and possible outcomes. If the latter is high enough, then even when only three possible outcomes are taken into account, the scenario trees expand so fast that after 5-6 stages the examined problems become numerically intractable.

To diminish the quantity of variables controlled at the upper level, one could think of introducing an artificial ("dummy") follower added to the initially existing follower in the original problem. The newly inserted follower is endowed with the objective function that jibes with that of the leader, whereas part of the originally governed variables of the upper level, are controlled by the dummy follower at the lower level. In this way, the lower level problem is also transformed and becomes a Nash equilibrium problem for the initial and dummy followers. One can try to find the conditions that provide that both the transformed and original bilevel programing models share (at least one) optimal solution.

5 Acknowledgements

The research activity of the authors was financially supported by the Research Department of the Tecnológico de Monterrey (ITESM), Campus Monterrey, and by the SEP-CONACYT project CB-2013-01-221676, Mexico. The third author was also supported by the PAICYT project No. CE250-09 and by the SEP-CONACYT project CB-2009-01-127691.

References

Arano, K.G., and Blair, B.F. (2008). An ex-post welfare analysis of natural gas regulation in the industrial sector. *Energy Economics*, vol. 30, pp. 789-806.

Ben-Ayed, O., Boyce, D.E., and Blair, C.A. (1988). A general bilevel linear programming formulation of the network design problem. *Journal of Transportations Research*, vol. 22, pp. 311-318.

Borraz-Sánchez, C., and Ríos-Mercado, R.Z. (2005) A hybrid meta-heuristic approach for natural gas pipeline network optimization. In: *Hybrid Metaheuristics*, Lecture Notes in Computer Science, Vol. 36, M. J. Blesa et al. (Eds.) New York: Springer, pp. 54-65.

Brockwell, P.J., and Davis, R.A. (2002). *Introduction to Time Series and Forecasting*. New York: Springer.

Chebouba, A., Yalaoui, F., Smati, A., Amodeo, L., Younsi, K., and Tairi, A.(2009). Optimization of natural gas pipeline transportation using ant colony optimization. *Computers and Operations Research*, vol. 36, pp. 1916-1923.

Chiou, S.W. (2005). Bilevel programming for the continuous transport network design problem. *Transportations Research Part B*, vol. 39, pp. 361-383.

Cruz, F.R.B., Smith, J.M.J., and Mateus, G.R. (1999). Algorithms for a multilevel network optimization problem. *European Journal of Operational Research*, vol. 118, pp. 164-180.

Dempe, S., Kalashnikov, V.V., Pérez-Valdés, G.A., and Kalashnykova, N.I. (2015). *Bilevel Programming Problems: Theory, Algorithms and Applications to Energy Networks*. Berlin-Heidelberg: Springer.

Dempe, S., Kalashnikov, V.V., and Ríos-Mercado, R.Z. (2005). Discrete bilevel programming: Application to a natural gas cash-out problem. *European Journal of Operational Research*, vol. 166, pp. 469-488, 2005.

Egging, R., Gabriel, S.A., Holz, F., and Zhuang, J. (2008). A complementarity model for the European natural gas market. *Energy Policy*, vol. 36, pp. 2385-2414.

EIA (1992[a]). *Energy Information Administration, FERC policy on system ownership since 1992*. Retrieved from http://www.eia.doe.gov/oil gas/natural gas/analysis publications/ngma-jorleg/fercpolicy.html

EIA (2005). *Energy Information Administration, Natural gas summary*. Retrieved from http://tonto.eia.doe.gov/dnav/ng/ng sum lsum dcu nusm.html

EIA (1992[b]). *Energy Information Administration, FERC order 636: The restructuring rule*. Retrieved from http://www.eia.doe.gov/oil gas/natural gas/analysis publications/ngma jorleg/ferc636.html

Esnault, B. (2003). The need for regulation of gas storage: the case of France. *Energy Policy*, vol. 31, pp. 167-174.

Gabriel, S.A., Zhuang, J., and Kiet, S. (2005). A large-scale linear complementarity model of the North American natural gas market. *Energy Economics*, vol. 27, pp. 639-665.

Hawdon, D. (2003). Efficiency, performance, and regulation of the international gas industry: A bootstrap DEA approach. *Energy Policy*, vol. 31, pp. 1167-1178.

IHS (2007). *IHS Engineering: EC proposes new legislation for European energy policy*. Retrieved from http://engineers.ihs.com/news/eu-enenergy-policy-9-07.html

Kabirian, A., and Hemmati, M.R. (2007). A strategic planning model for natural gas transmission networks. *Energy Policy*, vol. 35, pp. 5656-5670.

Kalashnikov, V.V., Pérez-Valdés, G.A., and Kalashnykova, N.I. (2010[a]). A linearization approach to solve the natural gas cash-out bilevel problem. *Annals of Operations Research*, vol. 181, pp. 423-442.

Kalashnikov, V.V., G.A. Pérez-Valdés, G.A., Tomasgard, A., and Kalashnykova, N.I. (2010[b]). Natural gas cash-out problem: Bilevel stochastic optimization approach. *European Journal of Operational Research*, vol. 206, pp. 18-33.

Kalashnikov, V.V., and Ríos-Mercado, R.Z. (2001). A penalty-function approach to a mixed-integer bilevel programming problem. In: *Proceedings of the 3rd International Meeting on Computer Science*, C. Zozaya et al. (Eds.), Vol. 2, pp. 1045-1054.

Kalashnikov, V.V., and Ríos-Mercado, R.Z. (2006). A natural gas cash-out problem: A bilevel programming framework and a penalty function method. *Optimization and Engineering*, vol. 7, pp. 403-420.

Kall, P., and Wallace, S.W. (1994). *Stochastic Programming*. Chichester: John Wiley & Sons.

Keyaerts, N., Meeus, L., and D'Haeseleer, W. (2008). Analysis of balancing-system design and contracting behavior in the natural gas markets. In: *Proceedings of the European Doctoral Seminar on Natural Gas*, Delft, The Netherlands, 26 pp.

Kinderlehrer, D., and Stampacchia, G. (1980). *An Introduction to Variational Inequalities and Their Applications*. New York: Academic Press.

Midthun, K.T. (2007). *Optimization Models for Liberalized Natural Gas Markets*. Ph.D. Thesis, Norwegian University of Science and Technology (NTNU): Trondheim. – 205 p.

Ríos-Mercado, R.Z., Secomandi, N., and Buraparate, V. (1999). *A mixed-integer bilevel programming model for the problem of minimizing cash-out penalty costs of a natural gas shipping company*. Technical report, Working paper, Postgraduate Program in System Engineering, UANL, San Nicolás de los Garza, Mexico.

Rockafellar, R.T. (1970). *Convex Analysis. Princeton*: Princeton University Press.

Soto, A. (2008). *FERC order 636 & 637*. Retrieved from http://www.aga.org/Legislative/issuesummaries/FERCOrder636637.html

Yang, H., and Bell, M.G.H.(2001). Transport bilevel programming problems: recent methodological advances. *Transportation Research Part B*, Vol. 35, pp. 1-4.