

Discrete Bilevel Programming: Application to a Natural Gas Cash-Out Problem

Stephan Dempe* Vyacheslav Kalashnikov[†] Roger Z. Ríos-Mercado[‡]

Abstract

In this paper, we present a mathematical framework for the problem of minimizing the cash-out penalties of a natural gas shipper. The problem is modeled as a mixed-integer bilevel programming problem having one Boolean variable in the lower level problem. Such problems are difficult to solve. To obtain a more tractable problem we move the Boolean variable from the lower to the upper level problem. The implications of this change of the problem are investigated thoroughly. The resulting lower level problem is a generalized transportation problem. The formulation of conditions guaranteeing the existence of an optimal solution for this problem is also in the scope of this paper. The corresponding results are then used to find a bound on the optimal function value of our initial problem.

Key words: Nonlinear programming, bilevel programming, Stackelberg games, gas cash-out problem, generalized transportation problem.

1 Introduction

In many decision processes there is a hierarchy of decision makers, and decisions are made at different levels in this hierarchy. One way to handle such hierarchies is to focus on one level and include the other levels' behaviors as constraints. In terms of modeling, the feasible set associated with a bilevel programming problem is implicitly determined by a series of two optimization problems which must be solved in a predetermined sequence: we have one leader (associated with the upper level) who selects his decision first and the follower (associated with the lower one) replies on this decision. With other words, the variables

*Freiberg University of Mining and Technology, Freiberg, Germany; e-mail: dempe@math.tu-freiberg.de, corresponding author

[†]ITESM, Monterrey, Nuevo León, Mexico; on leave from the Central Economics and Mathematics Institute (CEMI), Russian Academy of Sciences, Nakhimovsky prospekt 47, Moscow 117418, Russian Federation; e-mail: slavkamx@mail.ru

[‡]Graduate Program in Systems Engineering, Universidad Autónoma de Nuevo León, México; e-mail: roger@uanl.mx

of this problem are partitioned into two parts (x, y) . $y \in \mathbb{R}^m$ is the leader's variable and $x \in \mathbb{R}^n$ is that of the follower. Using y as a parameter, the follower solves a parametric optimization problem thus computing an optimal solution $x^* = x^*(y)$. This optimal solution $x^*(y)$ is the reaction (or the reply) of the follower on the leader's selection. The bilevel programming problem is the leader's problem. Knowing the reactions of the follower on his decisions he has now to minimize a certain function $f(x^*(y), y)$ with respect to his decision variables. Often this problem is interpreted such that the decision variables y of the leader compose "environmental data" for the follower's problem and that the leader is interested in optimizing a certain welfare function depending on those environmental data together with the optimal reaction of the follower on these data. In [1, 3, 11], one can find both the essential fundamentals of bilevel optimization and its applications to the solution of real systems.

A particular case of the bilevel programming problem is presented by the following mixed-integer model arising from the problem of minimizing the cash-out penalty costs of a natural gas shipping company. In many countries the natural gas industry has been going through a deregulation process since the mid-1980s leading to significant market changes. Now the decision making procedure of gas buying, selling, storing, transporting, etc., is immersed into a very complex world in which producers, pipelines (transporters), and brokers, all play quite important roles in the chain. This problem becomes even more complex if we take into account the whole network of pipelines transporting gas and oil throughout the continents. In this paper we investigate one particular problem in this complex system. This problem reflects the shipper's difficulties to deliver the correct amount of gas. It arises when a shipper draws a contract with a pipeline company to deliver a certain amount of gas at several points. What is actually delivered may be more or less of the amount that had been originally agreed on (this phenomenon is called *an imbalance*). When an imbalance occurs, the pipeline penalizes the shipper by imposing a cash-out penalty policy. As this penalty is a function of the operating daily imbalances, an important problem for the shippers is how to carry out their daily imbalances so as to minimize their incurred penalty.

Initial attempts and theoretical investigations of the considered problem can be found in [10]. There, penalty function approaches known from solution algorithms for variational inequalities [5, 9] are applied and first results of numerical experiments have been reported.

In [10], a mathematical framework for the above described problem is presented. The problem is modeled as a mixed-integer bilevel linear programming problem (BLP) where the shipper is the leader (upper level) and the pipeline represents the follower (lower level). Even the simplest version of a multilevel optimization problem, a linear problem with two levels, is known to be \mathcal{NP} -hard in the strong sense [4, 8]. This even implies that there cannot be a fully polynomial approximation scheme solving this problem. Mixed-integer BLP possess an even higher degree of difficulty as the typical concepts for fathoming in traditional branch-and-bound algorithms for mixed-integer programming cannot be directly applied to mixed-integer BLP [12]. There are only a small number of attempts to solve (mixed-) discrete bilevel programming problems (see [1, 3] and the references therein). One possible reason

for this can be the difficulties to treat parametric (mixed-) discrete optimization problems.

The position of constraints in the upper respectively in the lower level problem is critical. A shift of some constraint from the lower to the upper level problems can drastically change the optimal solution up to the situation that one of both problems has an optimal solution and the other one has no one. For (mixed-) discrete bilevel programming problems the position of the integrality conditions is of special importance [13]. Originally, in our problem, the Boolean variable belongs to the lower level problem. Knowing the difficulties in solving discrete bilevel programming problems we feel that it is easier to solve them if the discreteness demand is in the upper rather than in the lower levels. This forced us to move the integrality condition from the lower to the upper level problem. But surely, this implies a change of the problem. This is illustrated in the figures 4 to 6. The investigation of the differences and relations between the two problems is one of the topics in this paper.

After we have moved the discreteness demand into the upper level problem we solve an auxiliary problem which in our situation reduces to the solution of two linear bilevel programming problems. These can be solved by any of the known algorithms; we suggest to use the exact penalty function approach from [14]. Preliminary computational results show that the move of the Boolean variable to the upper level implies a much “smoother” solution process and produces not too large errors with respect to the optimal objective function value.

The resulting lower level problems are generalized transportation problems for which we derive necessary and sufficient conditions guaranteeing the existence of optimal solutions. These are then used to find bounds on the optimal objective function value of the original problem.

The paper is organized as follows. The problem is specified in Section 2, whereas the penalty function method is described in Section 3. The algorithm convergence results are also presented in Section 3.

2 Problem Specification

First, we describe the problem in terms of [10]. Assume that a shipper has entered into a contract (with other customers) to deliver a given amount of natural gas from a receipt to a delivery meter in a given time frame. (From now on, we treat “natural gas” and “gas” as synonyms). The shipper must stipulate title transfer agreements with the meter operators and a transportation agreement with the pipeline. Under such agreements, the shipper nominates a daily amount of gas to be injected by the receipt meter operator into the pipeline and to be withdrawn by the delivery meter operator from the pipeline. The pipeline transports the gas from the receipt meter to the delivery meter.

Due to the nature of the natural gas industry, what is actually transported is inevitably different from what is nominated. Such a difference constitutes an imbalance. There exist operational and transportation imbalances. The first type of imbalance refers to differences

between nominated and actual flows, while the latter involves differences between net receipts (receipts minus fuel) and deliveries. While pipelines allow for small imbalances, they issue penalties for higher (both operational and transportation) imbalances to the other parties. In the following, we assume that the shippers are held responsible for imbalance penalties, and we analyze the cash-out penalties associated with operational imbalances.

On the shipper's side, an operational imbalance can either be positive or negative. A positive [negative] imbalance arises when the shipper leaves [takes] gas in [from] the pipeline. Alternatively, a positive [negative] imbalance means that the actual flow is smaller [greater] than the nominated amount of gas. A positive [negative] end-of-the-month imbalance implies a cash transaction from the pipeline [shipper] to the shipper [pipeline]. Cash-out prices are set in a way that whenever a shipper sells [buys] gas to [from] the pipeline, he does that at a very low [high] price. The relation between cash-out price and imbalance position depends non-linearly on the average, maximum and minimum gas spot price for the past month.

Shippers daily nominate gas flows taking into account the constraints deriving from their buy/sell activity, their contractual constraints, and future market opportunities. The gas price is one of the major factors affecting their decisions. In the absence of cash-out provisions, historically shippers would take out high cost gas in the winter from the pipeline (causing negative imbalances), and pay the transporter back with low cost gas in the summer. This corresponds to a speculative behavior by the shippers, whereby imbalances are created and managed as pseudo-storage in order to take advantage of movements in the gas price. Cash-out penalties were designed in order to avoid such pricing arbitrages. In the framework below, shippers are concerned with minimizing the cash-out penalties.

In addition to the penalties stated in [10], we also try to prevent the shipper making large pseudo-storage by introducing supplementary fines for positive final imbalances.

2.1 Notation

As stated in the previous subsection, the decision making process for the shipper (leader) is to determine how to carry out its daily imbalances so as to minimize the penalty that will be imposed by the pipeline (follower). The following notations are used to describe the model.

Indices and Sets

- i, j, k zone pool indices; $i, j, k \in J = \{1, 2, \dots, P\}$;
- t time index; $t \in T = \{1, 2, \dots, N\}$.

Parameters

- x_{ti}^L, x_{ti}^U bounds on daily imbalances at the end of the day t in zone i ; $t \in T, i \in J$;
- x_t^L, x_t^U bounds on total daily imbalances at the end of the day t ; $t \in T$;
- s_{ti}^L, s_{ti}^U bounds on balance swings during day t in zone i ; $t \in T, i \in J$;
- $e_{ij} \in (0, 1)$ percentage of fuel retained for moving one dekatherm of gas from zone i to j ; $i, j \in J$;

- f_{ij} transportation charge for moving one dekatherm of gas from zone i to j ; $i, j \in J$, $i < j$;
- b_{ij} backward haul credit for moving one dt of gas from zone j to i ; $i, j \in J$, $i < j$;
- x_{0j} initial imbalance (start of day 1) in zone j , $j \in J$;
- r_j fine for the negative final imbalance in zone j , $j \in J$;
- δ_j fine for the positive final imbalance in zone j , $j \in J$.

Decision Variables

- x_{ti} imbalance at the end of the day t in zone i ; $t \in T$, $i \in J$;
- s_{ti} imbalance swing during day t in zone i ; $t \in T$, $i \in J$;
- y_i final imbalance in zone i ; $i \in J$;
- u_{ij} forward haul volume moved from zone i to j ; $i, j \in J$, $i < j$;
- v_{ij} backward haul volume moved from zone j to i ; $i, j \in J$, $i < j$;
- z total cash-out revenue for shipper.

Auxiliary Variables

- q binary variable equal to 1 (0) if final imbalances are nonnegative (nonpositive). If all final imbalances are zero, it is convenient for us to set $q = 0$.

2.2 Mathematical Model

Here we provide the set of constraints involved in both the upper and lower levels of the problem.

Upper level problem (Shipper's problem):

Objective: Maximize the shipper's revenue with respect to all variables (x, s, y, u, v, z) :

$$\max h_1(x, s, y, u, v, z) = z, \quad (2.1)$$

subject to:

$$x_{ti}^L \leq x_{ti} \leq x_{ti}^U, \quad t \in T, i \in J; \quad (2.2)$$

$$s_{ti}^L \leq s_{ti} \leq s_{ti}^U, \quad t \in T, i \in J; \quad (2.3)$$

$$x_t^L \leq \sum_{i \in J} x_{ti} \leq x_t^U, \quad t \in T; \quad (2.4)$$

$$x_{ti} = x_{t-1,i} + s_{ti}, \quad t \in T, i \in J. \quad (2.5)$$

and (y, u, v, z) being an optimal solution of the lower level problem

$$\text{for fixed } (x, s). \quad (2.6)$$

In case of non-unique optimal solution in the lower level problem, we will use the optimistic approach [3]. This means that that optimal solution of the lower level problem is taken which is the best one from the leader's point of view. This is reflected by the minimization with respect to all variables. But even under this assumption the above problem need not to have an optimal solution since the feasible set of bilevel programming problems with lower level discreteness assumptions is in general neither open nor closed [13].

Lower level problem (Problem of the pipeline):

Objective: The penalty is determined by minimizing the amount of cash transactions. In many cases, both shipper and pipeline agree in a policy that represents a compromise between them two, so rather than minimizing the revenue for the shipper, it is agreed to minimize the deviations from zero. Hence, given (x, s) , the following objective is to be minimized with respect to (y, u, v, z) :

$$\min h_2(x, s, y, u, v, z) = |z|, \quad (2.7)$$

subject to the constraints below.

Balance constraints: This constraint identifies the relationship between the imbalance at the day $N = |T|$, forward and backward haul volumes, retained fuel, and final imbalance at zone j :

$$y_j = x_{N,j} + \sum_{i:i<j} (1 - e_{ij})u_{ij} + \sum_{k:k>j} v_{jk} - \sum_{k:k>j} u_{jk} - \sum_{i:i<j} v_{ij}, \quad j \in J; \quad (2.8)$$

Gas conservation: This constraint ensures no gas loss occurs. This constraint follows directly from (2.8) after summation with respect to all $j \in J$.

$$\sum_{i \in J} y_i + \sum_{(i,j):i<j} e_{ij}u_{ij} = \sum_{i \in J} x_{N,i}. \quad (2.9)$$

Note that $\sum_{(i,j)} e_{ij}u_{ij} \geq 0$, hence $\sum_i y_i \leq \sum_i x_{N,i}$.

Zone upper bounds: This constraint prevents cyclic movements of gas. It simply states that, from any given zone, we cannot move more than any initial positive imbalance.

$$\sum_{j:j>i} u_{ij} + \sum_{k:k<i} v_{ki} \leq \max\{0, x_{N,i}\}, \quad i \in J. \quad (2.10)$$

Forward haul upper bounds: These bounds prevent positive-to-positive and negative forward movement of imbalances.

$$u_{ij} \leq \begin{cases} x_{N,i} & \text{if } x_{N,i} > 0 \text{ and } x_{N,j} < 0; \\ 0 & \text{otherwise,} \end{cases} \quad i, j \in J. \quad (2.11)$$

Backward haul upper bounds: These bounds prevent positive-to-positive and negative backward movement of imbalances.

$$v_{ij} \leq \begin{cases} x_{N,j} & \text{if } x_{N,j} > 0 \text{ and } x_{N,i} < 0; \\ 0 & \text{otherwise,} \end{cases} \quad i, j \in J. \quad (2.12)$$

Bounds on final imbalances: These bounds ensure that all final imbalances have the “right” sign, i.e. an imbalance must not change sign.

$$\min\{0, x_{N,i}\} \leq y_i \leq \max\{0, x_{N,i}\}, \quad i \in J. \quad (2.13)$$

Sign of final imbalances: This is a business rule that states that final imbalances for all zones must have the same “sign” (i.e. all nonpositive or nonnegative); that means that the imbalances must not change sign from zone to zone:

$$-M(1 - q) \leq y_i \leq Mq, \quad i \in J, \quad (2.14)$$

where M is a large number and q is a binary variable.

Shipper’s revenue: This equation represents the revenue from the shipper’s point of view:

$$z = \sum_{i \in J} [r_i y_i - \delta_i (y_i)_+] + \sum_{(i,j): i < j} b_{ij} v_{ij} - \sum_{(i,j): i < j} f_{ij} (1 - e_{ij}) u_{ij}; \quad (2.15)$$

where $y_+ = \max\{0, y\}$.

Variable types:

$$y_i, \quad i \in J, \quad z \text{ free}; \quad (2.16)$$

$$u_{ij}, v_{ij} \geq 0 \quad i, j \in J; \quad (2.17)$$

$$q \in \{0, 1\}. \quad (2.18)$$

We used this definition of the objective function in (2.15) to allow for different fines for positive and negative imbalances. If the imbalance is negative the fine per unit is equal to r_i , and if it is positive the per unit fine is $r_i - \delta_i$.

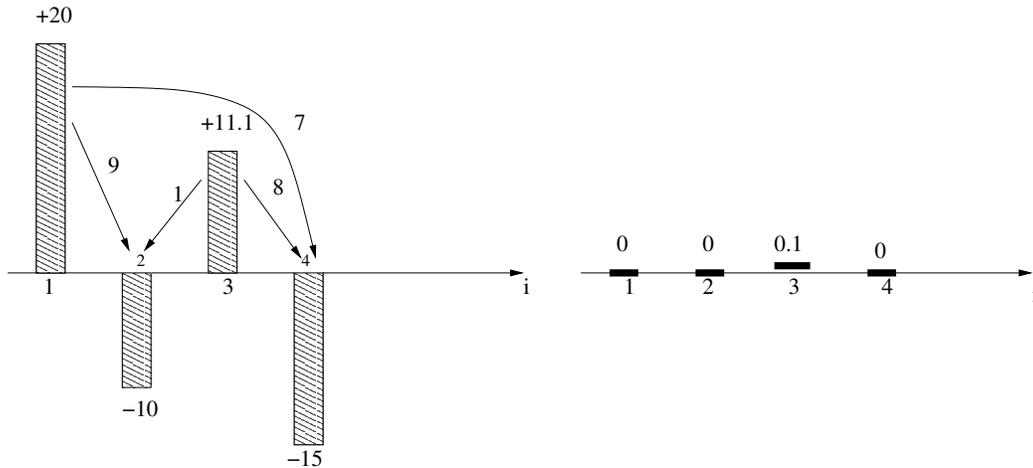


Figure 1: Final imbalances for $e_{12} = 0.1, e_{14} = 0.3, e_{34} = 0.2$ and $u_{12} = 10, v_{23} = 1, u_{14} = 10, u_{34} = 10$

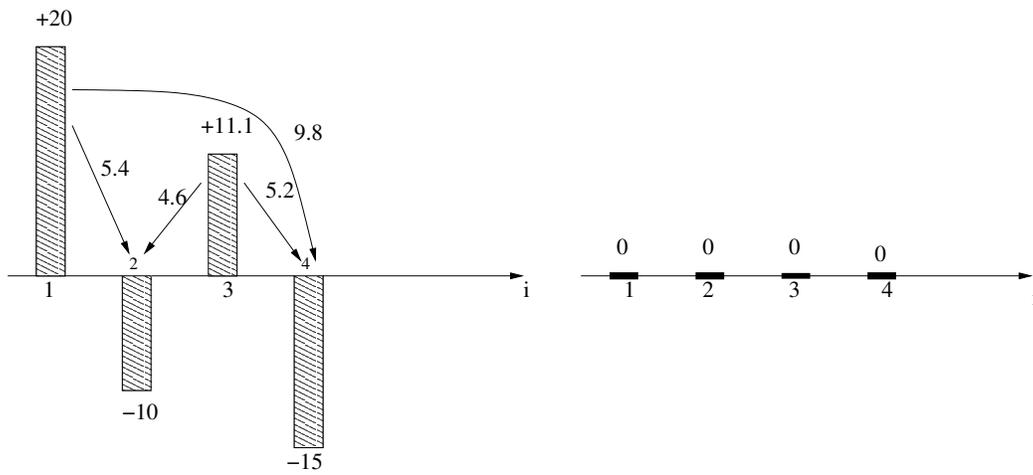


Figure 2: Final imbalances for $e_{12} = 0.1, e_{14} = 0.3, e_{34} = 0.2$, and $u_{12} = 6, v_{23} = 4.6, u_{14} = 14, u_{34} = 6.5$

Figures 1 to 3 show feasible solutions of the lower level problem. Here data $x_{N,i}$ (if positive then depicted at the bar above zero, if negative on the one below zero: $x_{N,1} = 20, x_{N,2} = -10, x_{N,3} = 11.1, x_{N,4} = -15$) are given together with the values for $e_{ij}, i, j \in J$. The numbers at the arrows indicate values of components of some feasible solution $u_{ij}, v_{ij}, i, j \in J$. All not given values are equal to zero. The final imbalances $y_i, i = 1, \dots, 4$, are shown in the right-hand pictures.

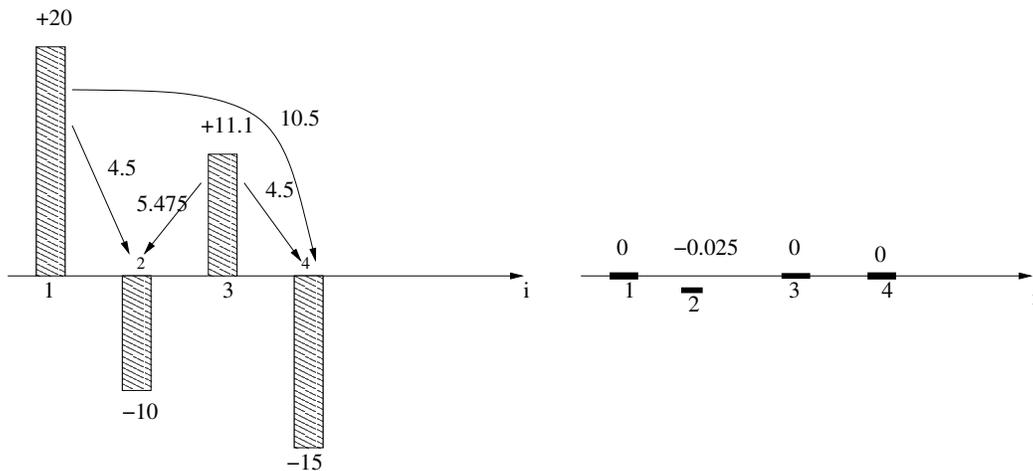


Figure 3: Final imbalances for $e_{12} = 0.1, e_{14} = 0.3, e_{34} = 0.2$, and $u_{12} = 5, v_{23} = 5.475, u_{14} = 15, u_{34} = 5.625$

3 The Modified Problem and its Properties

The mathematical model (2.1)–(2.18) described above is a mixed-discrete *bilevel program*. Implied by non-unique optimal solutions of the follower for some values of (x, s) , the leader has to maximize a generally discontinuous function in solving the problem. To explain this and the change of the problem induced by a move of the Boolean variable from the lower to the upper level problems, consider first two right-hand side perturbed linear optimization problems. Assume for simplicity that they have linear optimal value functions (in reality these functions are piecewise linear) depending on the parameter x and that each of them corresponds to one value of the Boolean variable q . The functions are given as $f_0(x)$ and $f_1(x)$ in Fig. 4.

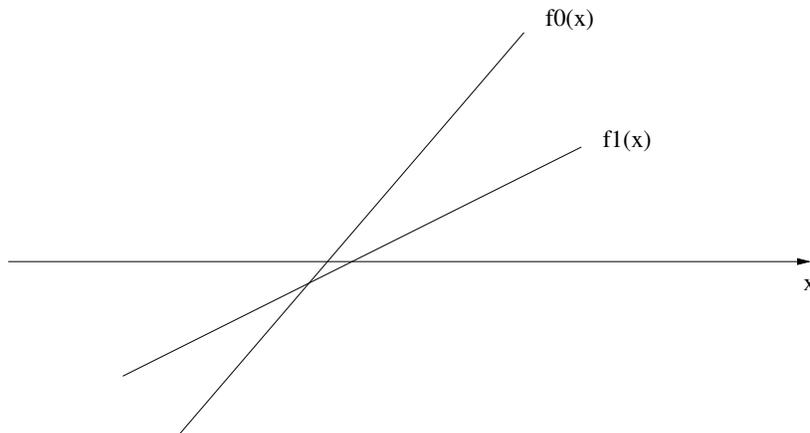


Figure 4: Example optimal response functions $z|_q$ of the pipeline (follower) for $q = 1$ and $q = 0$, i.e. functions to be maximized in the auxiliary problems with upper level integrality condition

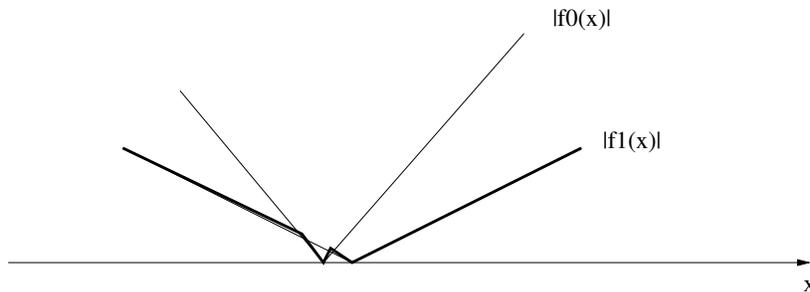


Figure 5: Absolute value function of the optimal response of the pipeline (follower) in the initial problem, i.e. the problem with lower level integrality condition

Then, if the objective function of both problems is changed to their absolute value, the optimal value functions are no longer linear but piecewise linear and nonnegative. Now, if the

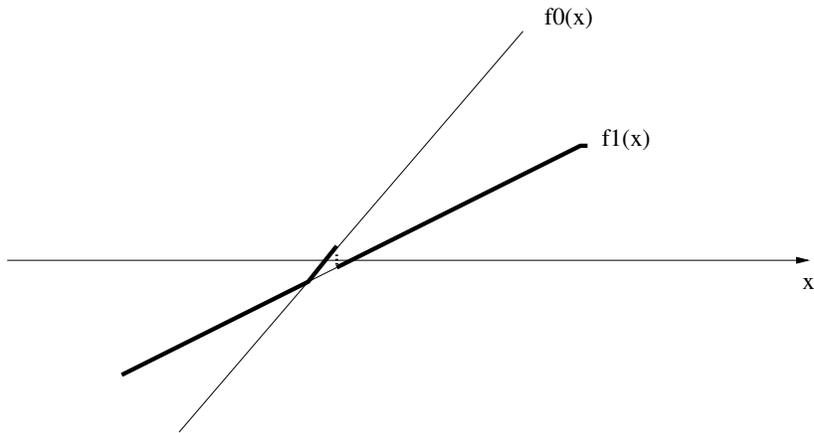


Figure 6: Resulting function to be maximized in the initial problem, i.e. the problem with lower level integrality condition. It has a break point $x = -8/3$, so only local maximum may be guaranteed by the solutions of the auxiliary problems

Boolean variable appears in the lower level problem, the follower has to take the smaller of the two absolute values as the optimal one (cf. the thick-lined function in Fig. 5). The leader does not consider the absolute value but he has to bear in mind the follower's reaction on his choice of x . This leads to the discontinuous thick-lined function in Fig. 6 to be minimized by the leader. Note that the jumps in the objective function of the leader can appear at any value of x depending on the follower's problem.

To make the problem more tractable, we move the integrality condition $q \in \{0, 1\}$ from the lower to the upper level. This leads to the problem

$$\begin{aligned}
 & \max && h_1(x, s, y, u, v, z) = z, && (3.1) \\
 & \text{subject to} && (2.2) - (2.5), q \in \{0, 1\} \text{ and to} && (3.2) \\
 & \min && h_2(x, s, y, u, v, z) = |z|, && (3.3) \\
 & \text{subject to} && (2.8) - (2.17). && (3.4)
 \end{aligned}$$

As first benefit from this change we get now a problem which has an optimal solution if the optimistic approach is used [3, 13]. Also, since there is only one Boolean variable q we simply solve the problem (3.1)–(3.4) for both values of q and compare the optimal objective function values of both auxiliary problems. The follower's problem is a parametric linear one having a piecewise-linear continuous optimal value function. The leader has now to maximize the pointwise maximum of the two functions in Fig. 4. The maximal function value of this maximum function can be different from that in the original problem. As an approximation of the objective function value of the original problem we take that optimal objective function value for both relaxations which is closer to zero. In the illustration in Fig. 4 this means occasionally that we take the function $f1(x)$ in each case. Clearly, our modifications change the problem and we have to investigate the properties of the resulting

solution.

3.1 Generalized Transportation Problem

Before we present the exactness result, we examine some properties of the lower level feasible sets W^β for $\beta \in \{0, 1\}$, where

$$W^\beta(x, s) = \{(y, u, v) \text{ defined by (2.8)-(2.14), (2.16)-(2.17), } q = \beta\}, \quad (3.5)$$

Lemma 3.1 *The subset W^β is convex and compact for each $\beta \in \{0, 1\}$.*

Proof: It is easy to see that the constraints (2.2)–(2.3) imply that the values of the variables (x, s) are bounded, whereas the values of y belong to the cube $[-M, M]^P$ (cf. (2.14)). The values of variables u and v are bounded since they satisfy (2.11), (2.12) and (2.17), and the values of $x_{N,j}$ are also bounded. Therefore, the subset W^β is bounded, and being closed it is compact.

To prove the convexity of W^β notice that all the constraints that define W^β are linear, except for constraints (2.11), (2.12) and (2.13). But it is evident that combined with (2.17), both (2.11) and (2.12) obtain the convex subsets. Finally, although (2.13) define a non-convex subset, they deliver convex subsets combined with (2.14). Thus, the subset W^β is convex which completes the proof. q.e.d.

We now show the equivalence of problem (3.3)–(3.4) to a generalized transportation problem, in which each pool zone with a positive last day imbalance ($x_{N,i} > 0$) is a point of supply, whereas the pool zones with negative last day imbalances ($x_{N,i} < 0$) are the points of demand. The transportation problem is a generalized one because of the coefficients $(1 - e_{ij})$ being different from 1 in equation (2.8).

More precisely, introduce the following notation:

$$\begin{aligned} J_+ &= \{i \in J \mid x_{N,i} > 0\}, & J_- &= \{j \in J \mid x_{N,j} < 0\}, \\ a_i &= x_{N,i}, \quad i \in J_+, & b_j &= -x_{N,j}, \quad j \in J_-, \\ p_{ij} &= \begin{cases} v_{ji}, & \text{if } i \in J_+, j \in J_-, i > j, \\ u_{ij}, & \text{if } i \in J_+, j \in J_-, i < j. \end{cases} \\ \lambda_{ij} &= \begin{cases} 1, & \text{if } i \in J_+, j \in J_-, i > j, \\ 1 - e_{ij}, & \text{if } i \in J_+, j \in J_-, i < j. \end{cases} \end{aligned}$$

These settings are consistent with (2.11), (2.12) which imply that in (2.8) either both negative sums (if $x_{N,i} > 0$) or both positive sums appear (if $x_{N,i} < 0$). Using these notations and $y_i \leq 0$ if $q = 0$ by (2.14) we derive that we have equations in (2.8) for $x_{N,i} > 0$ and inequalities for $x_{N,i} < 0$. Hence, the system (2.8) is equivalent to

$$\sum_{j \in J_-} p_{ij} = a_i, \quad i \in J_+; \quad \sum_{i \in J_+} \lambda_{ij} p_{ij} \leq b_j, \quad j \in J_-; \quad p_{ij} \geq 0, \quad i \in J_+, j \in J_- \quad (3.6)$$

for $q = 0$. In an analogous way, if $q = 1$ this system is equivalent to

$$\sum_{j \in J_-} p_{ij} \leq a_i, \quad i \in J_+; \quad \sum_{i \in J_+} \lambda_{ij} p_{ij} = b_j, \quad j \in J_-; \quad p_{ij} \geq 0, \quad i \in J_+, j \in J_-. \quad (3.7)$$

The upper bounds in the constraints (2.11) and (2.12) are corollaries of (3.6) and (3.7).

Remark 3.2 *A deeper look at the construction of the above systems of equalities and inequalities shows that, depending on the values of $x_{N,i}$ at least one of the systems (3.6) or (3.7) arises. Namely, if $\sum_{i \in J} x_{N,i} \leq 0$, then (2.9) and (2.14) imply that $y_i \leq 0, i \in J$ and, hence, (3.6) arises. In this case, system (3.7) has no solution at all. On the other hand, if $\sum_{i \in J} x_{N,i} > \left(\max_{(i,j): i < j} e_{ij}\right) \sum_{i \in J_+} x_{N,i}$ then, again by (2.9) and (2.14) we come to $y_i \geq 0, i \in J$ and hence system (3.7) appears. Now, system (3.6) is of no value. For other values of $x_{N,i}$ both systems are to be investigated.*

Let $J_+ = \{1, \dots, m\}, J_- = \{1, \dots, n\}$. To investigate nonemptiness of the sets $W^\beta(x, s)$ for $\beta \in \{0, 1\}$ it is sufficient to investigate solvability of generalized transportation problems of the following two kinds:

$$\begin{aligned} \sum_{i=1}^m \sum_{j=1}^n c_{ij} p_{ij} &\rightarrow \min_p \\ \sum_{j=1}^n p_{ij} &\leq a_i, \quad i = 1, \dots, m \\ \sum_{i=1}^m \lambda_{ij} p_{ij} &= b_j, \quad j = 1, \dots, n \\ p_{ij} &\geq 0, \quad i = 1, \dots, m, \quad j = 1, \dots, n \end{aligned} \quad (3.8)$$

and

$$\begin{aligned} \sum_{i=1}^m \sum_{j=1}^n c_{ij} p_{ij} &\rightarrow \min_p \\ \sum_{j=1}^n p_{ij} &= a_i, \quad i = 1, \dots, m \\ \sum_{i=1}^m \lambda_{ij} p_{ij} &\leq b_j, \quad j = 1, \dots, n \\ p_{ij} &\geq 0, \quad i = 1, \dots, m, \quad j = 1, \dots, n \end{aligned} \quad (3.9)$$

where $\lambda_{ij} \in (0, 1)$ for all $i = 1, \dots, m, j = 1, \dots, n$.

Problems of such type (where also upper bounds on the variables p_{ij} are allowed, but only with equality constraints) have been investigated in [6]. Our investigations in the following are in some sense parallel to the ideas there.

We start with the investigation of problem (3.8). We need some notation. Each (generalized) transportation problem corresponds to a bipartite graph $G = (E, V)$ with node set

$E = \{1, \dots, m\} \cup \{1, \dots, n\}$ and edge set $V = \{1, \dots, m\} \times \{1, \dots, n\}$. Take a subset $U \subset V$ in this graph constituting a forest, i.e. a finite collection of trees $U = \bigcup_{k=1}^t U^k$. Denote

$$I(V) = \{i \mid \exists j, (i, j) \in V\}; \quad J(V) = \{j \mid \exists i, (i, j) \in V\}.$$

The set U is called exhausted if $\bigcup_{k=1}^t I(U^k) = \{1, \dots, m\}$ and $\bigcup_{k=1}^t J(U^k) = \{1, \dots, n\}$. Then, by the properties of trees, $|I(U^k)| + |J(U^k)| - 1 = |U^k|$ which implies that the system of equations

$$\mu_i - \lambda_{ij} \nu_j = 0, (i, j) \in U^k; \quad \mu_{i_1} = 1 \text{ for some } i_1 \in I(U^k) \quad (3.10)$$

has a unique solution $\bar{\mu}_i > 0, i \in I(U^k), \bar{\nu}_j > 0, j \in J(U^k)$.

Let U be an exhausted forest and let $\bar{\mu}, \bar{\nu}$ denote the solution of the systems (3.10). Using this solution we can transform the generalized transportation problem into an equivalent one: Multiply the inequalities with $\bar{\mu}_i$ and the equations with $\bar{\nu}_j$ and substitute the variables p_{ij} with $\bar{\mu}_i p_{ij}$:

$$\begin{aligned} \sum_{i=1}^m \sum_{j=1}^n c_{ij} p_{ij} &\rightarrow \min_p \\ \sum_{j=1}^n \bar{\mu}_i p_{ij} &\leq \bar{\mu}_i a_i, \quad i = 1, \dots, m \\ \sum_{i=1}^m \lambda_{ij} \frac{\bar{\nu}_j}{\bar{\mu}_i} \bar{\mu}_i p_{ij} &= \bar{\nu}_j b_j, \quad j = 1, \dots, n \\ p_{ij} &\geq 0, \quad i = 1, \dots, m, \quad j = 1, \dots, n. \end{aligned}$$

Note that the result of this reformulation is now that the coefficient λ_{ij} is transformed into 1 for all $(i, j) \in U$. Call this problem canonical with respect to U .

The condition in the following theorem means that using the above transformation we derive a classical transportation problem under one assumption.

Theorem 3.3 ([7]) *The generalized transportation problems (3.8) as well as (3.9) can be reduced to the classical one if, and only if rank $\Lambda=1$, where $\Lambda = \|\lambda_{ij}\|_{m,n}$ is the matrix of the coefficients λ_{ij} .*

Lemma 3.4 *Problem (3.8) has a feasible solution if and only if there is an exhausted forest $U = \bigcup_{k=1}^t U^k$ such that for each k there exists $i_k \in I(U^k)$ with: The following system has a nonnegative solution:*

$$\sum_{j \in J(U^k)} \bar{\mu}_i p_{ij} = \bar{\mu}_i a_i \quad i \in I(U^k) \setminus \{i_k\} \quad (3.11)$$

$$\sum_{i \in I(U^k)} \bar{\mu}_i p_{ij} = \bar{\nu}_j b_j \quad j \in J(U^k) \quad (3.12)$$

and we have

$$\sum_{i \in I(U^k)} \bar{\mu}_i a_i \geq \sum_{j \in J(U^k)} \bar{\nu}_j b_j. \quad (3.13)$$

Proof: First, let problem (3.8) have a feasible solution \bar{p} . Extend the problem by adding slack variables in the system of inequalities. Then, this problem has also a feasible solution. Hence, there is also a basic feasible solution containing $m + n$ basic variables. Take U as the set of all indices of the basic variables without the slack variables. Consider the bipartite graph corresponding to the generalized transportation problem. Then, U corresponds to a subset of the edges and this subset does not contain cycles. Hence, $U = \bigcup_{k=1}^t U^k$ is an exhausted forest. Note that to each tree U^k there exists a unique slack variable $p_{i_k, n+1}$ being in the basis in the extended problem.

Now we can perform the above transformation of the generalized transportation problem which does not change the values of \bar{p} . Inserting the solution \bar{p} and considering only the basic variables we obtain for each k :

$$\begin{aligned} \sum_{j \in J(U^k)} \bar{\mu}_i \bar{p}_{ij} &= \bar{\mu}_i a_i \quad i \in I(U^k) \setminus \{i_k\} \\ \sum_{i \in I(U^k)} \bar{\mu}_i \bar{p}_{ij} &= \bar{v}_j b_j \quad j \in J(U^k) \end{aligned}$$

and, for $i = i_k$

$$\sum_{j \in J(U^k)} \bar{\mu}_i \bar{p}_{ij} + \bar{\mu}_i p_{i_k, n+1} = \bar{\mu}_i a_i.$$

Moreover, solvability of the problem implies that

$$\sum_{i \in I(U^k)} \bar{\mu}_i a_i = \sum_{i \in I(U^k)} \sum_{j \in J(U^k)} \bar{\mu}_i \bar{p}_{ij} + \bar{\mu}_i p_{i_k, n+1} = \sum_{j \in J(U^k)} \bar{v}_j b_j + \bar{\mu}_i p_{i_k, n+1}$$

Now, $\bar{\mu}_i p_{i_k, n+1} \geq 0$ gives the desired inequality (3.13).

Now let there be an exhausted forest with the given properties. We want to construct a feasible solution. Since the trees in the forest are not connected by edges it is sufficient to describe the construction for just one tree. Take a solution of (3.11) and (3.12). Then, the already fixed variables $p_{i_k, j}$ with $(i_k, j) \in U^k$ have the value

$$\bar{\mu}_{i_k} p_{i_k, j} = \bar{v}_j b_j - \sum_{i \in I(U^k) \setminus \{i_k\}} \bar{\mu}_i p_{ij}.$$

Note that there is one more variable $p_{i_k, j}$ in the i_k th row which has not been fixed yet (namely the slack variable in the extended problem) and we get

$$\begin{aligned} \bar{\mu}_{i_k} \sum_{j \in J(U^k)} p_{i_k, j} &= \sum_{j \in J(U^k)} \bar{v}_j b_j - \sum_{i \in I(U^k) \setminus \{i_k\}} \sum_{j \in J(U^k)} \bar{\mu}_i p_{ij} \\ &= \sum_{j \in J(U^k)} \bar{v}_j b_j - \sum_{i \in I(U^k) \setminus \{i_k\}} \bar{\mu}_i a_i \leq \bar{\mu}_{i_k} a_{i_k} \end{aligned}$$

by (3.13). Hence, we can set

$$\bar{\mu}_{i_k} p_{i_k, n+1} = \bar{\mu}_{i_k} a_{i_k} - \bar{\mu}_{i_k} \sum_{j \in J(U^k)} p_{i_k, j} \geq 0.$$

This gives the desired feasible solution.

q.e.d.

Consider the tree U^k . By the properties of basic feasible solutions, the system (3.11), (3.12) has a unique solution provided that (3.13) is satisfied. To compute this solution we can use the following simple algorithm written down for the original problem for simplicity.

Redraw the tree as being rooted in i_k . Each node of the tree corresponds to exactly one of the equations (3.11) or (3.12). If such a node is a leaf, summation is done over only one element whose value can now be fixed. Since a tree has at least one leaf, fix the value of the unique edge incident with this leaf. Then delete this leaf together with this edge and update the value for the other vertex incident with the deleted edge. More precisely:

Step 1 Let $V := U^k$ be drawn such that i_k is its root.

Step 2 Let l be the number of a leaf in the tree U^k . If $l \in J(U^k)$ then goto Step 3, otherwise to Step 4.

Step 3 Let (s, l) be the unique edge incident with l in V . Put $p_{sl} = b_l/\lambda_{sl}$, set $a_s := a_s - p_{sl}$, delete node l together with edge (s, l) , denote the resulting graph again by V . If V contains the unique node i_k stop, else goto Step 2.

Step 4 Now, $l \in I(U^k)$. Let the unique edge incident with node l be denoted by (l, s) . Put $p_{sl} = a_l$, $b_s := b_s - \lambda_{ls}p_{ls}$, delete node l together with edge (l, s) , denote the resulting graph again by V and goto Step 2.

Correctness of the algorithm is easy to see since it only writes down the solution of a solved system of equations. Using this algorithm now enables us to pose assumptions guaranteeing the existence of feasible solutions for the problem (3.8).

Theorem 3.5 *Problem (3.8) has a feasible solution if and only if there is an exhausted forest $U = \bigcup_{k=1}^t U^k$ such that for each k there exists $i_k \in I(U^k)$ with:*

$$\sum_{i \in I(U^k)} \bar{\mu}_i a_i \geq \sum_{j \in J(U^k)} \bar{\nu}_j b_j \quad (3.14)$$

and for each subtree U_t^k in U^k we have either

$$0 \leq \sum_{j \in J(U_t^k)} \bar{\nu}_j b_j - \sum_{i \in I(U_t^k) \setminus \{t\}} \bar{\mu}_i a_i \leq \bar{\mu}_t a_t \quad (3.15)$$

if node $t \in I(U^k)$ is the root of the subtree or

$$0 \leq \sum_{i \in I(U_t^k)} \bar{\mu}_i a_i - \sum_{j \in J(U_t^k) \setminus \{t\}} \bar{\nu}_j b_j \leq \bar{\nu}_t b_t \quad (3.16)$$

if $t \in J(U^k)$ is the root of the subtree U_t^k .

Proof: To show the validity of the theorem, it is sufficient to verify that the above algorithm works if the assumptions of the theorem are satisfied.

If the problem (3.8) has a feasible solution, then by Lemma 3.4 it has a basic solution with the properties in Lemma 3.4 and, given the set of basic variables, this solution can be computed applying the above algorithm to all trees in the exhaustive forest U . In each iteration of the algorithm one variable is set such that either equation (3.11) or (3.12) is satisfied. Now, rounding this variable up or down we arrive at the inequalities in (3.15) or (3.16).

Now, let there be given an exhausted forest decomposing into the trees U^k and let besides (3.14) for all subtrees U_t^k of U^k the corresponding two inequalities in either (3.15) or (3.16) be satisfied. Consider the above algorithm. Then, if the algorithm reaches node t we have to compute the right value for one variable p_{tj} if $t \in I(U^k)$ or p_{it} in the other case $t \in J(U^k)$. In this moment, all the variables belonging to nodes in the subtree U_t^k have got their values and the equations (3.11) and (3.12) are satisfied. Now, a simple calculation as in the proof of Lemma 3.4 shows that we have enough place to determine p_{tj} or p_{it} if the assumption (3.15) respectively (3.16) is valid. This shows that the theorem is correct. q.e.d.

In a way which is absolutely similar to Theorem 3.5 we get a condition for the existence of feasible solutions for problem (3.9):

Corollary 3.6 *Problem (3.9) has a feasible solution if and only if there is an exhausted forest $U = \bigcup_{k=1}^t U^k$ such that for each k there exists $i_k \in I(U^k)$ with:*

$$\sum_{i \in I(U^k)} \bar{\mu}_i a_i \leq \sum_{j \in J(U^k)} \bar{\nu}_j b_j$$

and for each subtree U_t^k in U^k we have either

$$0 \leq \sum_{j \in J(U_t^k)} \bar{\nu}_j b_j - \sum_{i \in I(U_t^k) \setminus \{t\}} \bar{\mu}_i a_i \leq \bar{\mu}_t a_t$$

if node $t \in I(U^k)$ is the root of the subtree or

$$0 \leq \sum_{i \in I(U_t^k)} \bar{\mu}_i a_i - \sum_{j \in J(U_t^k) \setminus \{t\}} \bar{\nu}_j b_j \leq \bar{\nu}_t b_t$$

if $t \in J(U^k)$ is the root of the subtree U_t^k . Here again, $(\bar{\mu}, \bar{\nu})$ is the solution to (3.10) corresponding to the exhausted forest U .

The following algorithm which implements a modified North-West Corner Rule can be used to verify that, for each values a, b at least one of the problems (3.8) and (3.9) has a feasible solution (and is therefore solvable by boundedness of the feasible set).

Lemma 3.7 *For any natural numbers n, m and positive vectors (a_1, \dots, a_m) , (b_1, \dots, b_n) , at least one of the problems (3.8) and (3.9) has a feasible solution.*

Proof: To prove our assertion, we use the modified North-West Corner Rule in the following form.

Step 0 Set $k = 0$, $i_k = 1$, $j_k = 1$.

Step 1 Put

$$p_{i_k, j_k} = \begin{cases} a_{i_k}, & \text{if } \lambda_{i_k, j_k} a_{i_k} \leq b_{j_k}; \\ b_{j_k} / \lambda_{i_k, j_k}, & \text{otherwise.} \end{cases}$$

Step 2 If $i_k < m$ then set

$$i_{k+1} = \begin{cases} i_k + 1, & \text{if } \lambda_{i_k, j_k} a_{i_k} \leq b_{j_k}, \\ i_k, & \text{otherwise.} \end{cases}$$

If $i_k = m$ and $\lambda_{i_k, j_k} a_{i_k} \leq b_{j_k}$, then set $K = k$ and stop. Goto Step 3.

Symmetrically, if $j_k < n$, put

$$j_{k+1} = \begin{cases} j_k + 1, & \text{if } \lambda_{i_k, j_k} a_{i_k} \geq b_{j_k}, \\ j_k, & \text{otherwise.} \end{cases}$$

At last, if $j_k = n$ and $\lambda_{i_k, j_k} a_{i_k} \geq b_{j_k}$, then set $K = k$ and stop.

Step 3 Now we update the parameters:

$$a_{i_k} := a_{i_k} - p_{i_k, j_k}; \quad b_{j_k} := b_{j_k} - \lambda_{i_k, j_k} p_{i_k, j_k}.$$

If $a_i = 0$ for all $i = 1, \dots, m$, or $b_j = 0$ for each $j = 1, \dots, n$, then set $K = k$ and stop. Otherwise, set $k := k + 1$ and return to Step 1.

Let the algorithm stop in Step 3. Then by construction, either $a = 0$ which means that we have $\sum_{j=1}^n p_{ij} = a_i$, $i = 1, \dots, m$ or $b = 0$ implying $\sum_{i=1}^m \lambda_{ij} p_{ij} = b_j$, $j = 1, \dots, n$. Since the other system of inequalities is also satisfied, we obtain non-emptiness of the feasible set of either problem (3.9) or (3.8). If the algorithm stops in Step 2 then either $i_K = m$ and $\sum_{j=1}^n p_{ij} = a_i$, $i = 1, \dots, m$ or $j_K = n$ and $\sum_{i=1}^m \lambda_{ij} p_{ij} = b_j$, $j = 1, \dots, n$. In both cases, the other system of inequalities is also satisfied by construction. This again implies non-emptiness of the feasible set of either problem (3.9) or (3.8). The lemma is proved completely. q.e.d.

3.2 A Solution Method

The solution of problem (3.1)–(3.4) essentially reduces to the solution of two linear bilevel programming problems and to use that optimal solution of one of these problems which has the larger objective function value as an optimal solution of (3.1)–(3.4) or that one with optimal function value closer to zero as an approximation of an optimal solution of the original problem (2.1)–(2.18). Many algorithms solving linear bilevel programming problems

can be found in the monograph [1]. Most efficient (at least for small problems) are branch-and-bound algorithms being based on an implicit enumeration of the faces of the underlying convex polyhedron. But knowing the nonpolynomial behavior of branch-and-bound algorithms we decided to apply the exact penalty function approach of [14] since this produces at least an approximation of the optimal function value if it runs out of time. The approach is shortly outlined as follows.

Fix q to either of its values and reformulate the resulting linear bilevel programming problem in (3.1)–(3.4) as

$$c_1^\top x + d_1^\top y \rightarrow \max_{x,y} \quad (3.17)$$

$$A_1 x \leq b_1, x \geq 0 \quad (3.18)$$

$$\text{where } y \text{ solves} \quad (3.19)$$

$$d_2^\top y \rightarrow \max_y \quad (3.20)$$

$$A_2 x + B_2 y = b_2 \quad (3.21)$$

$$y \geq 0. \quad (3.22)$$

To do it, it suffices to introduce a new variable x instead of the variables (x, s) of the upper level, denote the lower level variables (y, u, v) by a new variable y , and finally introduce some auxiliary non-negative variables (included into the new vectors x and y) to get rid of the absolute value $|\cdot|$ and the non-negative part $(\cdot)_+$ in the objective functions of the auxiliary problems. The latter is accomplished by introducing new two-sided linear inequalities involving the new variables and the initial objective function of the auxiliary problem.

For a fixed x the vector y solves the lower level problem (3.20)–(3.22) if and only if there exists a vector λ satisfying

$$A_2 x + B_2 y = b_2, y \geq 0 \quad (3.23)$$

$$B_2^\top \lambda \geq d_2 \quad (3.24)$$

$$y^\top (B_2^\top \lambda - d_2) = 0 \quad (3.25)$$

We intend to use a penalty function approach for solving the problem (3.17)–(3.22). Following this aim we penalize the complementarity constraint (3.25). This leads to the following problem:

$$c_1^\top x + d_1^\top y - K y^\top (B_2^\top \lambda - d_2) \rightarrow \max_{x,y,\lambda} \quad (3.26)$$

$$A_1 x \leq b_1, x \geq 0 \quad (3.27)$$

$$A_2 x + B_2 y = b_2, y \geq 0 \quad (3.28)$$

$$B_2^\top \lambda \geq d_2 \quad (3.29)$$

We will use the following assumption:

Assumption 1: The set $\{\lambda : B_2^\top \lambda \geq d_2\} \neq \emptyset$.

If this assumption is satisfied, then the lower level problem (3.20)–(3.22) has an optimal solution if and only if $\{y \geq 0 : A_2x + B_2y = b_2\} \neq \emptyset$. Using the results in [14] the following theorem can easily be verified to be true.

Theorem 3.8 *Assume that problem (3.26)–(3.29) has an optimal solution for some positive K_0 . Then, the problem (3.26)–(3.29) describes an exact penalty function approach for problem (3.17)–(3.22), i.e. there is a number K^* such that the optimal values of the problems (3.17)–(3.22) and (3.26)–(3.29) for all $K \geq K^*$ coincide.*

We have used this idea to attack problem (3.1)–(3.4) and have compared the corresponding solution process with one for solving the original problem (2.1)–(2.17). The main difference between both processes was it that the solution process for (2.1)–(2.17) very often jumped between solutions of the lower level problems for $q = 0$ and $q = 1$ even for problems of very small dimension. Almost surely, this discontinuous effect will intensify for problems with larger dimension. In contrast, the problems (3.1)–(3.4) have been solved very smoothly.

The results of our preliminary test calculations are as follows. The test problems used for checking the performance of the exact penalty function algorithm have various numbers of pools (4 to 9) and time periods (2 to 30). Exact specification of even the minimum size (4 pools and 2 days) problems is too large, so we mention only that we fixed the initial imbalance vector (like the vector $x_0 = (-10, -4, 3, 6)$ in the minimum size problem) and the daily imbalance upper and lower bounds, while varying the daily swing s_{ij} upper and lower bounds, the penalty parameter K and of course, the integer parameter $q \in \{0, 1\}$.

To solve the standard mathematical programming problem (3.1)–(3.4), we composed a program in the GAMS [2] language (version 2.50A) and conducted experiments in a Sun Ultra 10 workstation running Solaris 7. The test problems were processed very fast, at the total (compilation, generation and execution) time of 0.020–0.050 sec. The exact penalty approach performance has proved to be more efficient than the direct algorithm (which models the problem directly by calculating the optimal response of the follower to the leader’s choice of the last day imbalance; here, the follower’s optimal response includes also the optimal choice of the value of parameter q) at least in one aspect. Namely, the exact penalty algorithm has always reached the optimal solution with only one attempt, in contrast to the direct one which has repeatedly needed 3 to 5 restarts after having “jammed” far from the optimum solution.

4 Interpretation of the Obtained Solution

Recall that in our simple problem with only one Boolean variable we solve the problem (3.1)–(3.4) for q being fixed to 0 and for q being fixed to 1 and compare both optimal function values. In doing so we obtain lower and upper bounds for the optimal value of the initial problem (3.1)–(3.4). In the two distinguished cases in Remark 3.2, it is evident that the solution of the corresponding auxiliary problem (that with nonempty feasible set

W^β) coincides with the solution of the initial one. In the remaining cases, the situation may be complex enough, as it is illustrated in Figures 4 to 6. The investigation of this more interesting case is the topic of this section. We will see that, under certain assumptions, we are able to determine a locally optimal solution respectively to find bounds on the optimal function value of the original problem. For this we use the variables of problem (2.1)–(2.18).

Theorem 4.1 *Let $(x^0, s^0, y^0, u^0, v^0, z^0)$ be a locally optimal solution of problem (3.1)–(3.4) for $q = 0$ and let (y^1, u^1, v^1, z^1) be an optimal solution of the lower level problem (3.3), (3.4) for fixed $x = x^0, s = s^0, q = 1$. If $0 < z^0 < z^1$, then $(x^0, s^0, y^0, u^0, v^0, z^0)$ is a locally optimal solution of the problem (2.1)–(2.18).*

Proof: Since $0 < z^0 < z^1$, the point (y^0, u^0, v^0, z^0) is an optimal solution of the lower level problem (2.7)–(2.18) and, hence, the point $(x^0, s^0, y^0, u^0, v^0, z^0)$ is feasible for the bilevel programming problem (2.1)–(2.18). Now, let (x, s) be sufficiently close to (x^0, s^0) and feasible for (2.2)–(2.5). Then, by continuity of the optimal value function of parametric linear programming problems, the optimal function values of the problems (3.3), (3.4) for $q = 0$ resp. $q = 1$ are close to z^0 and z^1 , resp. Hence, the optimal solution of the problem (2.1)–(2.18) is obtained for $q = 0$. Then, local optimality of $(x^0, s^0, y^0, u^0, v^0, z^0)$ is implied by local optimality of this point for (3.1)–(3.4) with $q = 0$. q.e.d.

Remark 4.2 *There are three other cases which can be treated analogously:*

1. $(x^0, s^0, y^0, u^0, v^0, z^0)$ is locally optimal for (3.1)–(3.4) with $q = 0$ and $0 > z^0 > z^1$,
2. $(x^1, s^1, y^1, u^1, v^1, z^1)$ is locally optimal for (3.1)–(3.4) for $q = 1$ and $0 > z^1 > z^0$, and
3. $(x^1, s^1, y^1, u^1, v^1, z^1)$ is locally optimal for (3.1)–(3.4) with $q = 1$ and $0 < z^1 < z^0$.

Remark 4.3 *To compute points satisfying the assumptions of Theorem 4.1 or Remark 4.2 we can solve two problems:*

$$\max_{x,s} \max_{q \in \{0,1\}} z \text{ subject to (2.2)–(2.17)}$$

and

$$\max_{x,s} \min_{q \in \{0,1\}} z \text{ subject to (2.2)–(2.17)}.$$

Then, Theorem 4.1 can be applied if either the optimal value of the first problem is less than zero or the optimal value of the second problem is larger than zero. To solve the second problem we can adapt a descent algorithm combined with a check after each iteration if we remain on the correct branch resp. a jump on the other branch.

To derive some more relations between the feasible sets of the initial problem (2.1)–(2.18) and (3.1)–(3.4) introduce the following notation. Let

$$H = \{(x, s) \text{ satisfying (2.2) – (2.5)}\}.$$

Denote

$$t(\bar{y}, \bar{u}, \bar{v}) = \sum_{j \in J} [r_j \bar{y}_j - \delta_j (\bar{y}_j)_+] + \sum_{i < j} [b_{ij} \bar{v}_{ij} - f_{ij} (1 - e_{ij}) \bar{u}_{ij}]$$

and

$$\begin{aligned} V(x, s) = & \{(x, s, \bar{y}, \bar{u}, \bar{v}, \bar{z}) : (\bar{y}, \bar{u}, \bar{v}) \in W^0(x, s) \cup W^1(x, s) \\ & \text{and } |\bar{z}| = \min\{|z^0|, |z^1|\}, \text{ where } \bar{z} = t(\bar{y}, \bar{u}, \bar{v}) \text{ and} \\ & |z^\beta| = \min\{|t(y, u, v)| : (y, u, v) \in W^\beta(x, s)\}, \beta \in \{0, 1\}\}, \end{aligned}$$

where $W^\beta(x, s)$ is defined by (3.5). Moreover, put

$$\bar{q}(x, s) = \begin{cases} 0, & \text{if } \bar{z} = z^0 \\ 1, & \text{if } \bar{z} = z^1 \end{cases}$$

and $\bar{z} = \bar{z}(x, s)$. In the case $|z^0| = |z^1|$ we set $\bar{q}(x, s) = 0$. Analogously, we define

$$\begin{aligned} V^\beta(x, s) = & \{(x, s, \tilde{y}, \tilde{u}, \tilde{v}, \tilde{z}) : (\tilde{y}, \tilde{u}, \tilde{v}) \in W^\beta(x, s), |\tilde{z}| = |t(\tilde{y}, \tilde{u}, \tilde{v})| = |z^\beta|, \\ & \text{where } |z^\beta| = \min\{|t(y, u, v)| : (y, u, v) \in W^\beta(x, s)\}\} \end{aligned}$$

Denote $z^\beta = z^\beta(x, s)$ with the convention that $z^\beta(x, s) = \infty$ if $W^\beta(x, s) = \emptyset$ (and thus $V^\beta(x, s) = \emptyset$, too).

Lemma 4.4 *For each $(x, s) \in H$, the following relationships are valid:*

$$V(x, s) = \begin{cases} V^0(x, s) & \text{if } |z^0(x, s)| < |z^1(x, s)| \\ V^1(x, s) & \text{if } |z^0(x, s)| > |z^1(x, s)| \\ V^0(x, s) \cup V^1(x, s) & \text{if } |z^0(x, s)| = |z^1(x, s)| \end{cases} \quad (4.1)$$

In particular, if we introduce $V^\beta = \bigcup_{(x,s) \in H} V^\beta(x, s)$ then (4.1) implies

$$V = \bigcup_{(x,s) \in H} V(x, s) \subseteq V^0 \cup V^1. \quad (4.2)$$

Proof: Let $(x, s, \bar{y}, \bar{u}, \bar{v}, \bar{z}) \in V(x, s)$. If $\bar{q}(x, s) = 0$ and $|z^0(x, s)| < |z^1(x, s)|$, then clearly $\bar{z} = z^0$. Hence, $(x, s, \bar{y}, \bar{u}, \bar{v}, \bar{z}) \in V^0(x, s)$. Similarly, if $\bar{q}(x, s) = 1$ then $(x, s, \bar{y}, \bar{u}, \bar{v}, \bar{z}) \in V^1(x, s)$. Vice versa, $(x, s, \bar{y}, \bar{u}, \bar{v}, \bar{z}) \in V^0(x, s)$ and $|z^0(x, s)| < |z^1(x, s)|$ imply that $(x, s, \bar{y}, \bar{u}, \bar{v}, \bar{z}) \in V(x, s)$ with $\bar{q}(x, s) = 0$, whereas $(x, s, \bar{y}, \bar{u}, \bar{v}, \bar{z}) \in V^1(x, s)$ and $|z^0(x, s)| > |z^1(x, s)|$ means that $(x, s, \bar{y}, \bar{u}, \bar{v}, \bar{z}) \in V(x, s)$ with $\bar{q}(x, s) = 1$. The degenerated case $|z^0(x, s)| = |z^1(x, s)|$ is analyzed in the similar manner which establishes (4.1). Inclusion (4.2) follows immediately from (4.1). q.e.d.

Remark 4.5 *Lemma 4.4, in particular, (4.2) implies that the optimal value $z^* = \max_{(x,s) \in H} \bar{z}(x, s)$ of the initial bilevel problem (2.1)–(2.18) can be estimated as follows:*

$$\min_{\beta \in \{0,1\}} z^{*\beta} \leq z^* \leq \max_{\beta \in \{0,1\}} z^{*\beta} \quad (4.3)$$

where $z^{*\beta} = \max_{(x,s) \in H} z^\beta(x, s)$, $\beta \in \{0, 1\}$ are the optimal values of the auxiliary problems (3.1)–(3.4) for $q = 0$ and $q = 1$, respectively.

Unfortunately, as the Figures 1 to 3 show, there may exist vectors $(x, s) \in H$ such that both $V^0(x, s)$ and $V^1(x, s)$ are non-empty but $|z^0(x, s)| \neq |z^1(x, s)|$. Namely, Figures 1 – 3 show feasible solutions $u_{ij}, v_{ij}, i, j \in J$, for the lower level problem (2.7)–(2.18) for different values of q . Note that the final imbalances are restricted in sign according to the value of q .

The latter means that in general, $V \subset V^0 \cup V^1$ without $V = V^0 \cup V^1$ which forbids the equality in estimate (4.3). However, due to Remark 3.2 this “bad” case can take place only if $0 < \sum_{i \in J} x_{N,i} \leq (\max_{(i,j):i < j} e_{ij}) \sum_{i \in J_+} x_{N,i}$, or equivalently, if

$$-\sum_{i \in J_-} x_{N,i} < \sum_{i \in J_+} x_{N,i} < -(1 - \max_{(i,j):i < j} e_{ij})^{-1} \sum_{i \in J_-} x_{N,i}. \quad (4.4)$$

If the factors e_{ij} are close to 0, then (4.4) means that the final imbalance $\sum_{j \in J} y_j$ is also very close to zero. In such a case, the pipeline may ignore this imbalance and not to fine the shipper at all, which sets $z = 0$.

5 Conclusions

The paper considers a mixed-integer bilevel programming problem arising in applications. The problem is interesting and difficult in that even under assumptions of linearity, the follower’s optimal response may be a discontinuous function. To avoid this difficulty, the authors propose to shift the integrality constraints from the lower level to the upper level of the original problem. Of course, the latter changes the problem, but the paper proposes ways to estimate the proximity of solutions of the modified and original problems.

The above-described modification of the original problem allows one to apply an exact penalization approach to reduce the modified problem to a standard mathematical program which is almost linear one: the only nonlinear terms are the complementarity products arising in the K-K-T optimality conditions. If these products are multiplied by a penalty parameter and subtracted from the objective function, we come to the mathematical program with linear constraints and nonlinear objective function. The latter makes it easier to solve the modified problem by the modern mathematical programming algorithms. The test numerical experiments confirm the smoothness of the solution process.

Then it has been noticed that in the application of these problem, the lower level part is a generalized transportation problem, and the classical existence results are applicable to it. The latter helps to examine the structure of the solution set.

At last, the numerical experiments reveal that the original problem needs some improvements to prevent the leader from practicing an arbitrage resulting in an artificial accumulation of positive imbalances at the “far” pools and negative ones at the “nearby” pools. These improvements will be presented in a forthcoming paper.

Acknowledgments: The authors acknowledge financial support from various institutions: Deutscher Akademischer Austauschdienst (first and second author); Mexican National Council for Science and Technology (third author under CONACyT award J33187–A); Univer-

sidad Autónoma de Nuevo León (second author under UANL-PAICYT award CA562-01; third one under UANL-PAICYT awards CA555-01 and CA763-02). We are also grateful to Nicola Secomandi (Carnegie Mellon University, USA) and Viroj Buraparate (PROS Revenue Management, USA), who had a major role on working and developing the first version of the mathematical model for this application. Their assistance, suggestions, and discussion is highly appreciated. Last but not least, we also would like to thank the very helpful remarks made by two anonymous referees.

References

- [1] Bard, J.F., *Practical Bilevel Programming: Algorithms and Applications*, Kluwer, Dordrecht, 1998.
- [2] Brooke, A., Kendrick, D., and Meeraus, A., *GAMS: A User's Guide*, The Scientific Press, San Francisco, 1992.
- [3] Dempe, S., *Foundations of Bilevel Programming*, Kluwer, Dordrecht, 2002.
- [4] Deng, X., “Complexity issues in bilevel linear programming,” in: A. Migdalas, P. Pardalos, and P. Värbrand (eds.), *Multilevel Optimization: Algorithms and Applications*, Kluwer, Dordrecht, 1998.
- [5] Fukushima, M., “Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems,” *Mathematical Programming*, 53 (1992) 99-110.
- [6] Gabasov, R.F., and Kirillova, F.M., *Linear Programming Methods. Part 2. Transportation Problems*, Izdatel'stvo Belorusskogo Universiteta, Minsk, 1978 (in Russian). Russian title: Metody linejnogo programmirovaniya. Chast' 2. Transportnye zadachi.
- [7] Gol'shtein, E.G., and Yudin, D.B., *Linear Programming Problems of the Transportation Type*, Nauka, Moscow, 1969 (in Russian).
- [8] Hansen, P., Jaumard, B., and Savard, G., “New branch-and-bound rules for linear bilevel programming,” *SIAM Journal on Scientific and Statistical Computing*, 13 (1992) 1194–121.
- [9] Kalashnikov, V.V., and Kalashnikova, N.I., “Solving two-level variational inequalities,” *Journal of Global Optimization*, 8 (1996) 289–294.
- [10] Kalashnikov, V.V., and Ríos-Mercado, R.Z., “Solving a natural gas cash-out bilevel program by a penalty function method,” Technical Report PISIS-2002-03, Graduate Program in Systems Engineering, Universidad Autónoma de Nuevo León, San Nicolás de los Garza, México, 2002.

- [11] Migdalas, A., Pardalos, P., and Värbrand, P. (eds.), *Multilevel Optimization: Algorithms and Applications*, Kluwer, Dordrecht, 1998.
- [12] Moore, J., and Bard, J.F., “The mixed integer linear bilevel programming problem,” *Operations Research*, 38 (1990) 911–921.
- [13] Vicente, L.N., Savard, G., and Judice, J.J., “The discrete linear bilevel programming problem,” *Journal of Optimization Theory and Applications*, 89 (1996) 597–614.
- [14] White, D.J., and Anandalingam, G., “A penalty function approach for solving bilevel linear programs,” *Journal of Global Optimization*, 3 (1993) 397–419.