

USING COLUMN GENERATION TO SOLVE A COAL BLENDING PROBLEM

STÉPHANE AURAY¹, DANIEL DE WOLF² AND YVES SMEERS³

Abstract. In this paper, we formulate and solve a real life coal blending problem using a Column Generation Approach. The objective of the model is to prescribe optimal mixes of coal to produce coke. The problem is formulated as a mixed integer program. It involves various types of constraints arising from technical considerations of the blending process. The model also incorporates nonlinear constraints. It results in a large-scale problem that cannot be solved by classical operations research methods. Defining three heuristic methods based on column generation techniques, this paper proposes reasonable solutions for the industry.

Keywords. Column generation, coal blending.

Mathematics Subject Classification. 90B30, 49M27.

1. INTRODUCTION

In this paper, we consider a real life problem: a coal blending problem for coke production. The model includes not only quality constraints for the coke but also all the constraints resulting from the different transportation modes and the inventory constraints.

Different types of constraints should be considered during the production. Firstly, the coke is subject to a collection of *quality constraints* which, in some

Received August 26, 2013. Accepted May 21, 2014.

¹ CREST-Ensai, Rue Blaise Pascal, 35172 Bruz, France; Equipe, Université du Côte d'Opale and CIRPEE. stephane.auray@ensai.fr

² TVES, Université du Littoral Côte d'Opale 189B avenue Maurice Schumann, BP 5526, 59379 Dunkerque Cedex 1, France daniel.dewolf@univ-littoral.fr

³ Center for Operations Research and Econometrics, UCL, 34 voie du Roman Pays, 1348 Louvain-La-Neuve, Belgium. yves.smeers@uclouvain.be

cases, may depend on the clients. These constraints are all linear. But there are also, for technical reasons, a *maximum number of coal types in the mix*. This number is dependant on the number of gates of the plant. A *mix* consists simply of using particular ratios of different coal types during a certain period of time.

There is also, for technical reasons, a *minimal proportion for each coal present* in the mix. As we shall see in Section 3, taking these conditions into account requires the introduction of *integer variables* and *nonlinear constraints*. This may be linearized under a restrictive condition (see Williams [10]).

The coke delivered to the customers can be obtained using different mixes of coals. However, the total number of mixes used by a plant in a given time period has an upper bound: there is a significant fixed cost involved in changing the mix and only two different mixes can be used each month. Finally, we must take into account *capacity constraints* as well as *minimum production levels*.

The industrial problem tackled in this paper is more difficult than the classical coal blending problem of the Operations Research literature considered by Sarker and Gunn [6] or Greenberg [4] due to the *presence of binary variables*. These binary variables are required to model the minimum and/or maximum level of use of each coal. Sarker and Gunn [6] solve the tactical planning/coal blending problem. This leads to a nonlinear nonconvex problem for which solutions are computed using simple Successive Linear Programming. They used the Lasdon SLP implementation based on the XLP code of Marsten (see Lasdon [5]). Greenberg [4] looks at the same pooling problem that arises when blending materials, but considers a different application, namely the blending of crude or refined petroleum. He presents a new method based upon computational geometry which provides exact answers to questions of sensitivity analysis for this nonlinear nonconvex problem.

We solve a more complicated problem through the following column generation technique (see Vanderbeck and Wolsey [7] for an exact algorithm for IP column generation, Vanderbeck [8] for the branch-and-price algorithm or Desaulniers, Desrosiers and Solomon [2] for a didactic introduction to the use of column generation technique in integer programming). For each time period, and each plant, by fixing the coal delivery prices at the entry of the plant, *i.e.* a price that represents the unit purchasing and transportation costs, we determine the *optimal coal mix for one unit of production*. This results in *subproblems*. Then, the *Master problem* determines the level of use of these mixes. The Master problem also determines the coal delivery plan to the plants and the coke delivery plan to the clients. As we shall see in Section 4, the solution of the Master problem leads to reviewing the delivery prices at the entry of the plant for each coal. New mixes based on these new prices are generated by solving the subproblems. Consequently, new columns are generated for the Master problem. Note that there are also integer variables at the upper level, since only a limited number of different mixes can be considered at each time period.

We would like to mention that other types of constraints have been considered in coal blending problems for coke production. For instance, Vasko *et al.* [9] consider coke oven wall pressure or coke stability constraints. They propose candidate coals

TABLE 1. Choice of indices.

c	coals
b	coals coming by boat
r	coals coming by railways
k	plants
t	time periods
m	mixes
h	harbours
a	clients

to produce coke by solving first a mixed integer linear programming model (MILP) similar to the subproblems that we consider (coal quantity for one ton of mixes) and then they use a binary decision tree that ensure the production of high-quality coke. Sample results are presented on hypothetical, but realistic, data. In our model, however, the set of coals used by the company does not pose significant coke oven wall pressure or coke stability problems. So we do not take into consideration this complicating constraints.

The *main contribution of this paper* is to solve a real world problem encountered within this industry, which is a large-scale problem intractable by classical operations research methods. Defining three heuristics methods based on column generation techniques, we furnished reasonable solutions to the industrial company.

2. PROBLEM DESCRIPTION

The manager of a coke production company wants to determine the operations schedule for the next three months for the five coke production units situated in Belgium. We use t as the index for the time periods. The three time periods correspond to the months of January to March. We denote $days_t$ the number of days for month t .

These five plants, denoted by k , produce different types of coke by mixing different types of coals. The mixes are indexed by m . Different type of clients, indexed by a , are served by the company. This results in different quality constraints on the mix.

The demand is assumed to be known with accuracy for the first three months of the year. The different coals, indexed by c , are purchased from several countries and brought to Belgium by means of two transportation modes, namely boat or rail. The buying prices of the different coals and the transportation costs from producer to the plant are also given. Table 1 summarizes the indices used in the model.

2.1. PLANTS DESCRIPTION

Each production plant k is characterized by a *daily capacity of plant* – cap_k – which is the maximum amount of coal that can be handled by the plant (see

TABLE 2. Plant characteristics.

k	cap_k	$minuse_k$	$gates_k$	$mincoal_k$	$maxcoal_k$
1	2300	60%	8	5%	100%
2	2850	60%	4	15%	35%
3	1250	75%	8	10%	100%
4	3500	60%	8	10%	100%
5	3500	60%	8	10%	100%

TABLE 3. Production cost (euro per ton of coal).

$prodcost_{kt}$	$t = 1$	$t = 2$	$t = 3$
$k = 1$	11.450	11.400	11.350
$k = 2$	18.175	18.100	18.025
$k = 3$	38.150	38.000	37.850
$k = 4$	24.125	24.025	23.925
$k = 5$	11.300	11.250	11.200

Tab. 2). Since the plant can never been stopped, there is also a *minimal utilization rate for each plant*, denoted $minuse_k$ for plant k . To put the coals into the oven, there is only a limited number of entry gates, denoted $gates_k$ for plant k . This implies that the number of coal types in the mix is limited. There are also a *minimal* and a *maximal proportion* for each coal in the mix, denoted respectively $mincoal_k$ and $maxcoal_k$ for plant k . Note that these two parameters are function of the plant and are independent of the coal type: in fact, the minimal and maximal proportion depend on the charging machinery of the plants. Since the plants were constructed at different periods, the technology used is different, which explains the differences from plant to plant.

Plants also differ by their unitary production cost. We denote by $prodcost_{kt}$ the unitary production cost of plant k at time period t (see Tab. 3).

2.2. COAL CHARACTERISTICS

The firm can use 16 types of coal during this production horizon. These types differ by their characteristics such as the *ash rate*, denoted ash_c for coal c , the *sulfur rate*, denoted $sulf_c$, the *alkali rate*, denoted alk_c , the *volatile part*, denoted vol_c and the *wet part*, denoted wet_c . The wet part is the unusable part of the coal. One also distinguishes three classes among the coals: the “High Volume”, denoted HV, the “Mid Volume”, denoted MV, and the “Low Volume”, denoted LV (see Tab. 4).

TABLE 4. Coal characteristics.

c	ash_c	$sulf_c$	alk_c	vol_c	wet_c	Type
1	4.99%	.85%	.12%	17.89%	6.63%	LV
2	5.53%	.70%	.09%	17.48%	7.77%	LV
3	7.72%	.94%	.31%	22.71%	10.15%	LV
4	8.30%	.67%	.14%	21.00%	8.00%	LV
5	8.07%	.70%	.16%	23.60%	9.47%	MV
6	4.83%	.88%	.16%	30.19%	6.80%	HV
7	6.86%	1.15%	.24%	29.90%	7.54%	HV
8	6.14%	.84%	.22%	31.39%	8.66%	HV
9	6.01%	.82%	.19%	32.75%	6.29%	HV
10	6.45%	.88%	.17%	33.08%	7.20%	HV
11	7.44%	.66%	.24%	27.43%	6.45%	HV
12	7.70%	.71%	.23%	32.09%	6.81%	HV
13	7.00%	.98%	.32%	25.10%	8.00%	MV
14	7.60%	.57%	.18%	19.30%	10.14%	LV
15	5.79%	.82%	.18%	24.39%	8.99%	MV
16	5.30%	.72%	.13%	33.50%	7.90%	HV

2.3. AVAILABILITY, DELIVERY AND DEPARTURE PRICES OF THE COALS

Some of the coals are coming *by boat*. We index by b these coals (see Fig. 1). They mainly come from North America and from Australia. There are *two possible arrival harbours* that we index by h . There is an *initial coal inventory at the beginning of the year in the two harbours*, noted by $initstock_{ch}$ (see Tab. 5). We denote $expq_{ct}$ deliveries corresponding to already ordered quantities *expected for some of these coals*. The landing harbour is not yet decided.

The other coals are delivered *by rail*. We index by r these coals. They mainly come from Germany and from Eastern Europe. There is no storage capacity for these coals. Deliveries, noted $expq_{ct}$, are also expected for these coals. The final destination of these quantities is not yet decided (see Tab. 5).

The USD is the reference currency for the coals coming by boat. The euro is used for the other coals. Expected exchange rates for the next three months are given in Table 5. Prices are given at the harbour or station of departure in Table 5 where the relevant currency is also given for each coal. By multiplying the price in foreign currency by the expected exchange rate, we obtain the price of coal c at time period t in € denoted $price_{ct}$.

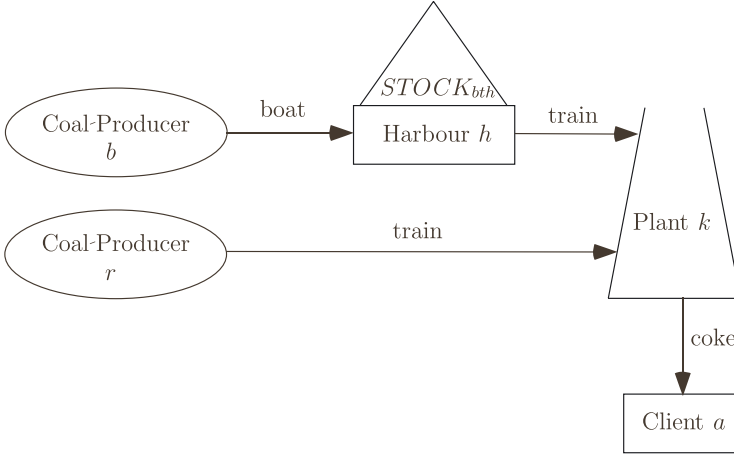


FIGURE 1. Supply chain for the input to the plants.

TABLE 5. Initial inventory, expected quantities and coal prices.

Coal (c)	<i>initstock_{ch}</i> (tons)		<i>expq_{ct}</i> (tons)			Price	
	<i>h</i> = 1	<i>h</i> = 2	<i>t</i> = 1	<i>t</i> = 2	<i>t</i> = 3		
1	<i>b</i>	558	7620	70 000	70 000	100 000	\$ 46.80
2	<i>b</i>						\$ 46.75
3	<i>t</i>			26 000	26 000	26 000	€ 55.65
4	<i>b</i>		36 655				\$ 37.75
5	<i>b</i>	42 760	22 570	40 000		60 000	\$ 45.75
6	<i>t</i>			49 600	49 600	49 600	€ 42.225
7	<i>t</i>			14 800	14 800	14 800	€ 42.225
8	<i>t</i>			10 000	10 000	10 000	€ 41.775
9	<i>t</i>			20 000	20 000	20 000	€ 43.575
10	<i>b</i>						\$ 46.65
11	<i>b</i>	5950					\$ 49.25
12	<i>b</i>						\$ 44.10
13	<i>t</i>			16 000	16 000	16 000	€ 55.125
14	<i>b</i>	43 505		40 000		60 000	\$ 40.00
15	<i>t</i>			42 400	42 400	42 400	€ 44.95
16	<i>b</i>		6450	25 000	25 000		\$ 46.80
Currency rate				<i>t</i> = 1	<i>t</i> = 2	<i>t</i> = 3	
		\$		0.975	0.925	0.875	
		€		1	1	1	

TABLE 6. Boat or rail transportation costs.

(\$/ton)	<i>boatcost_{ch}</i>								
	<i>c</i> = 1	<i>c</i> = 2	<i>c</i> = 4	<i>c</i> = 5	<i>c</i> = 10	<i>c</i> = 11	<i>c</i> = 12	<i>c</i> = 14	<i>c</i> = 16
<i>h</i> = 1	5.1	4.5	10.25	8.15	3.75	3.30	0	7.5	5.10
<i>h</i> = 2	5.1	4.5	10.25	8.15	3.75	3.30	0	7.5	5.10
(€/ton)	<i>transpcost_{hk}</i>			<i>railcost_{ck}</i> (€/ton)					
	<i>h</i> = 1	<i>h</i> = 2	<i>c</i> = 3	<i>c</i> = 6	<i>c</i> = 7	<i>c</i> = 8	<i>c</i> = 9	<i>c</i> = 13	<i>c</i> = 15
<i>k</i> = 1	4.4675	4.4675	9.155	4.67	4.67	4.67	9.155	9.155	4.67
<i>k</i> = 2	2.6375	3.25	7.39	4.2625	4.2625	4.2625	7.39	7.39	4.2625
<i>k</i> = 3	+∞	0	–	–	–	–	–	–	–
<i>k</i> = 4	3.86	3.86	8.29	6.08	6.08	6.08	8.29	8.29	6.08
<i>k</i> = 5	2.460	4.105	7.7475	2.31	2.31	2.31	7.7475	7.7475	2.31

2.4. THE TRANSPORTATION COSTS

The *transportation costs by boat to the harbour*, denoted *boatcost_{ch}*, are given in \$ per ton of coal (see Tab. 6). As can be seen, there is no difference for the transportation costs between the two arrival harbours. *Handling costs at the harbour*, denoted *dockcost_h* are the same at the two harbours due to the competition between the harbours: 3.3875 € per ton. *Transportation costs from harbours to plants*, denoted *transpcost_{hk}* are given in € per ton (see Tab. 6).

Note that there is *no transportation cost between harbour 2 and plant 3*, the plant being located in the harbour area. Consequently, plant 3 is only supplied from harbour 2. The four other plants are also supplied by rail. The *transportation costs by rail between the departure station and the plants*, denoted *railcost_{ck}*, are given in € per ton by Table 6 for all coals coming directly from the production site to the plant by rail.

2.5. CHARACTERISTICS OF THE DEMAND

The coke company must satisfy the demand of 13 clients for the next three months. We use *a* as index for the client. The *demand* of client *a* in time period *t* is denoted *dem_{at}* (see Tab. 7). In order to minimize the coke transportation cost, the plant requested by each client is given. A few of them (clients 4 and 6) can be supplied by two plants, the distances to the two plants being similar.

We explain now the *quality constraints* that the coal mix must meet. The first one concerns the *volatile rate*. The *volatile rate of the mix must be between 24% and 26%*. We denote these two parameters as *minvol* and *maxvol*. The second constraint concerns the *ash rate*. The mix has a *maximal ash rate* depending on the client. We denote *maxash_a* this quantity for client *a*. The third specification concerns the *sulfur rate*. There is a *lower and upper limit on sulfur*, denoted *minsulf_a* and *maxsulf_a*, which also depends on the client. The fourth specification concerns the *alkali rate*. There is an *upper limit on alkali*, denoted *maxalk_a*, which also depends on the client *a*. The fifth specification concerns the *Low Volume rate*. The

TABLE 7. Demand for the next three months.

dem_{at}	$t = 1$	$t = 2$	$t = 3$	Plant
Client 1	11 500	12 600	12 600	3
Client 2		4000	4000	1
Client 3	45 000	45000	45 000	1
Client 4	7000	8000	8000	2 or 3
Client 5	2000	2000	2000	3
Client 6	16 500	24 800	14 400	1 or 2
Client 7	26 700	19 400	14 800	2
Client 8		3500		2
Client 9	6000	6000	5000	1
Client 10	16 000	16 000	14 000	2
Client 11	12 000	12 000	12 000	3
Client 12	58 452	49 002	51 644	4
Client 13	68 516	68 132	71 162	5

lower and upper limits on Low Volume, denoted $minlv_a$ and $maxlv_a$, also depend on the client (see Tab. 8). For these three characteristics (ash, sulfur and alkali), there is a multiplicative coefficient from the mix to the coke: they are denoted respectively *multash*, *multsulf* and *multalk* (see Tab. 8).

For all minimal and/or maximal characteristic rates that depend on the client, we shall consider as specification for the mix of the plant *the specification of the most restrictive client* that the plant has to serve.

The sixth specification concerns the *rate in "Mid Volume"*. Here there are *lower and upper bounds on Mid Volume*, denoted $minmv$ (0.25) and $maxmv$ (0.25) respectively, which do not depend on the client. In fact, for technical reasons, there is a fixed rate of 1/4 of Mid volume coals. The seventh specification concerns the *soft rate*. There is an *upper bound on Soft coal*, denoted $maxsoft$ (0.10), which does not depend on the client. The only soft coal is coal 12. The eighth specification concerns the *Australian coal rate*. Here, for commercial reasons, there is an *upper bound*, denoted $maxaus$ (0.30), which does not depend on the client. The only Australian coal is coal 4.

Finally, note that the number of mixes that can be used for each plant per period, is limited. This maximum number of mixes, denoted $maxmix_{kt}$, is equal to 2 mixes per month for each plant. The different mixes will be indexed by m .

3. PROBLEM FORMULATION

We have already introduced the following indices (see Tab. 1): the coals are indexed by c , the plants by k , the time periods by t , the mixes by m , the harbours

TABLE 8. Specifications which depend on the client.

Client	$maxash_a$	$minsulf_a$	$maxsulf_a$	$maxalk_a$	$minlv_a$	$maxlv_a$
Client 1	10.		1.	.30	30.	100.
Client 2	10.		1.	.30	30.	100.
Client 3	10.		1.	.30	30.	100.
Client 4	10.		1.	.30	30.	100.
Client 5	10.		1.	.30	30.	100.
Client 6	9.5	.7	.9	.30	30.	100.
Client 7	9.5	.7	.9	.30	30.	100.
Client 8	10.		1.	.30	40.	50.
Client 9	10.		1.	.30	30.	100.
Client 10	10.		1.	.30	30.	100.
Client 11	10.		1.	.30	30.	100.
Client 12	9.5	.7	.9	.30	30.	100.
Client 13	9.5	.7	.9	.30	30.	100.
Multiplicative coefficient from the mix to the coke			Ash	Sulfur	Alkali	
			1.32	0.92	1.32	

by h and the clients by a . To facilitate the notation, we have also introduced the subindices b for the coals coming by boat and r for the coals coming by rail.

3.1. CHOICE OF DECISION VARIABLES

The first set of variables concerns the **production of coke in the different plants**: we denote by $COAL_{cktm}$ the quantity of coal c used in plant k at the time period t in the mix m . We introduce only for convenience the variable SUM_{ktm} , which is the total quantity of coals for plant k in the mix m at the time period t . These two quantities are given in tons. To account for the limited number of coal in the mix, we introduce the binary variables $COALPRES_{cktm}$ to indicate that coal c is present in the mix m for plant k at time period t . To account for the limited number of mixes at each time period, we introduce the binary variables $MIXPRES_{ktm}$ to indicate that mix m is used for plant k at time period t . Table 9 summarizes this choice of variables.

A second set of variables is used to **manage the deliveries and inventories of coals** (see Fig. 2).

We denote by $ORDQ_{ct}$ the new ordered quantity of coal c at time period t and assume that the orders placed in period t are also delivered in period t . Consequently, this quantity adds to the expected quantities at time period t , $expq_{ct}$ which are given (see Sect. 2). We denote by $XBOAT_{bth}$ the quantity of coal b coming by boat that is sent to harbour h at time period t and by $XRAIL_{rkt}$ the

TABLE 9. Variables for the coke production.

Variable	Definition
$COAL_{cktm}$	quantity of coal c used in the mix m at period t in mix m ,
SUM_{ktm}	total coal quantity for mix m at plan k for period t ,
$COALPRES_{cktm}$	indicator that coal c is in the mix m (binary),
$MIXPRES_{ktm}$	indicator the the mix m is used for plant k (binary).

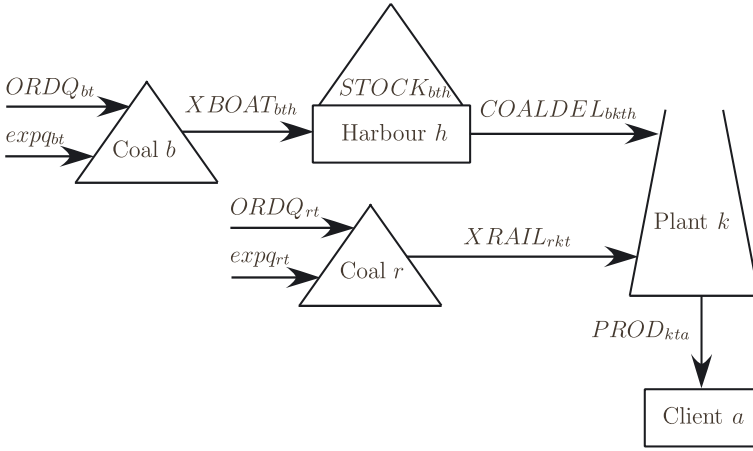


FIGURE 2. Decision variables for the coal deliveries and inventories.

quantity of coal r coming by rail that is sent to coke plant k at time period t . We let $STOCK_{bth}$ be the inventory of coal delivered by boat b at the end of time period t in harbour h and $COALDEL_{bkth}$ be the coal b deliveries from harbour h to plant k at time period t . Last we denote by $PROD_{kta}$ the tons of coke produced in plant k at time period t for client a . These variables are necessary to compute the coke transportation plan, since some clients can be supplied by more than one plant. Table 10 summarizes this choice of variables.

3.2. MATHEMATICAL EXPRESSION OF THE OBJECTIVE

The objective function is the sum of six following terms:

$$\min z = \sum_c \sum_t price_{ct}(expq_{ct} + ORDQ_{ct}) \quad (1)$$

$$+ \sum_b \sum_t \sum_h (boatcost_{bh} + dockcost_h) XBOAT_{bth} \quad (2)$$

TABLE 10. Variables for the coal deliveries and inventories.

Variable	Definition
$ORDQ_{ct}$	new ordered quantity of coal c at period t ,
$XBOAT_{bth}$	quantity of coal b sent to harbour h at the period t ,
$STOCK_{bth}$	inventory of coal b in the harbour h at the end of period t ,
$COALDEL_{bkt}$	coal deliveries from harbour h to plant k
$XRAIL_{rkt}$	quantity of coal r delivered at time t to plant k .
$PROD_{kta}$	coke produce at plant k for client a at time t .

$$+ \sum_b \sum_t \sum_h rate \cdot delprice_{bth} STOCK_{bth} \quad (3)$$

$$+ \sum_b \sum_k \sum_t \sum_h transpcost_{hk} COALDEL_{bkt} \quad (4)$$

$$+ \sum_r \sum_k \sum_t railcost_{rk} XRAIL_{rkt} \quad (5)$$

$$+ \sum_k \sum_t \sum_m prodcost_{kt} SUM_{ktm} \quad (6)$$

- the coal *purchasing cost* (1) where $price_{ct}$ is the coal price c at period t ,
- the *boat transportation cost* and the *handling cost* at the harbours (2) where $boatcost_{bh}$ is the boat transportation cost and $dockcost_h$ is the handling cost at harbour,
- the *holding inventory cost* for each quantity of coal stored at the arrival harbour (3) where $rate$ is the monthly opportunity rate (0, 5%) and $delprice_{bth}$ is the delivery price, i.e the sum of the purchasing cost, the boat transportation cost and the handling cost at the harbour:

$$delprice_{bth} = price_{bt} + boatcost_{bh} + dockcost_h$$

- the *transportation cost from harbour to plants* (4) where $transpcost_{hk}$ is the unitary transportation cost from the harbour to the plant,
- the *transportation cost by rail* (5) with $railcost_{rk}$ the unitary rail cost,
- the *production costs* (6) with $prodcost_{kt}$ the unitary production cost.

3.3. MATHEMATICAL EXPRESSION OF THE CONSTRAINTS

We have the following relations between these variables:

$$SUM_{ktm} = \sum_c COAL_{cktm} \quad (7)$$

$$COAL_{cktm} \leq COALPRES_{cktm} cap_k days_t \quad (8)$$

$$SUM_{ktm} \leq MIXPRES_{ktm} cap_k days_t \quad (9)$$

$$minuse_k cap_k days_t \leq \sum_m SUM_{ktm} \leq cap_k days_t \quad (10)$$

$$expq_{bt} + ORDQ_{bt} = \sum_h XBOAT_{bth} \quad (11)$$

$$expq_{rt} + ORDQ_{rt} = \sum_k XRAIL_{rkt} \quad (12)$$

$$\sum_h COALDEL_{bkt} = \sum_m COAL_{bktm} \quad (13)$$

$$XRAIL_{rkt} = \sum_m COAL_{rktm} \quad (14)$$

$$\sum_m MIXPRES_{ktm} \leq maxmix_{kt} \quad (15)$$

$$STOCK_{bth} - STOCK_{b,t-1,h} = XBOAT_{bth} - \sum_k COALDEL_{bkt} \quad (16)$$

$$\sum_m \sum_c (1 - wet_c) * COAL_{cktm} = \sum_a PROD_{kta} \quad (17)$$

$$\sum_{k \text{ delivering } a} PROD_{kta} \geq dem_{at} \quad (18)$$

$$\sum_c COALPRES_{cktm} \leq gates_k \quad (19)$$

$$mincoal_k COALPRES_{cktm} \leq \frac{COAL_{cktm}}{SUM_{ktm}} \quad (20)$$

$$\frac{COAL_{cktm}}{SUM_{ktm}} \leq maxcoal_k COALPRES_{cktm} \quad (21)$$

- Equation (7) computes the *sum of coals* for each mix and each time period.

- Equations (8) and (9) are the definition of the *binary variables*: namely they indicate the presence of the coal or the use of the mix. Note that cap_k is the daily coal entry capacity for plant k and $days_t$ is the number of days in time period t . This number of days is needed since capacity is a daily capacity and the variables are monthly quantities of coal entering the plant.
- Equation (10) concerns the *minimal utilization rate* and *capacities of plants*.
- Equation (11) indicates that the total amount of coal delivered by boat must be allocated between the two harbours since equation (12) indicates that the total amount of coal delivered by rail must be allocated between the plants.
- Equation (13) is the balance for each coal coming by boat at each plant since equation (14) is the balance for each coal delivered by rail.
- Equation (15) concerns the maximal number of mixes for each plant and each time period.
- Equation (16) gives the balance at the end of time period t for coal b at each harbour h since equation (17) gives the balance at the exit of the plants.
- Equation (18) is the demand satisfaction.
- Equation (19) deals with the maximal number of coal types in the mix.
- Equation (20) defines the minimal use of each coal in a mix since equation (21) concerns the maximal use of each coal in a mix.

We have two types of variables:

- the *continuous variables* which must be *nonnegative*:

$$COAL_{cktm}, SUM_{ktm}, COALDEL_{bkt}, PROD_{kta},$$

$$XBOAT_{bth}, XRAIL_{rkt}, ORDQ_{bt}, STOCK_{bth} \geq 0$$

- and the *binary variables*:

$$COALPRES_{cktm}, MIXPRES_{ktm} \in \{0, 1\}$$

Note also that the two last constraints (20) and (21) are nonlinear. All the other quality constraints on the mix are easy to write. They are listed in Appendix A.

4. SOLUTION METHOD

Except for constraints (20) and (21), all the constraints of the problem are purely linear. Note that the left hand side of constraint (20) must be multiplied by the variable COALPRES, otherwise all the coals would be present in the mix.

The solution technique followed by Sarker and Gunn [6] is to define for each ratio $COAL_{cktm}/SUM_{ktm}$ a variable, say f_{cktm} , introducing for each c, k, t and m the constraint:

$$f_{cktm}SUM_{ktm} = COAL_{cktm}.$$

They obtain a problem with a nonlinear part of bilinear type (product of variables). Heuristics for solving such problems have been proposed in the Operations Research literature. They are related to techniques for solving fractional linear programs and they resort to successive linear programming (SLP).

However, in our problem, integer variables remain, since constraints (20) and (21) are written as:

$$min_{coal_k} COALPRES_{cktm} \leq f_{cktm} \leq max_{coal_k} COALPRES_{cktm}.$$

This paper proposes three heuristic methods, based on column generation ideas. The problem formulated in Section 3 has two types of binary variables: one called $COALPRES_{cktm}$ to indicate the presence of coal c in the mix m for plant k at time period t . The other one, called $MIXPRES_{ktm}$ indicates that the mix m is used for plant k at time period t . We know that column generation procedures solve only LP problems. In our heuristic method, we keep the binary variable in the subproblem (for limiting the number of coal types in the mix) and in the main program (for limiting the number of mixes used for each plant and each time period).

The *subproblems* are defined as follows. For each plant, and for each time period, initially we consider the delivery price at the entry of the plant as the sum of the coal price and the total transportation cost to the plant. We look for a mix of **one ton** that satisfies all the quality constraints on the mix. This gives an *initial column of coal proportions* for each plant for each time period. Note that imposing the *sum of coals to be equal to one* is a simple manner to *linearize the two nonlinear constraints* (20) and (21).

In the *master problem*, we shall determine the level of utilization of these mixes in order to meet the demand of the several clients at minimal production, transportation and coal purchasing costs. By multiplying the level of use of a mix by the column of the coal proportions in the mix for a particular plant, we obtain the coal quantities needed for the plant at the time period.

To *update the delivery price at the entry* of the plant, we consider the *dual price of the balance at the entry of the plant constraints* (13) or (14) depending on the transportation mode. In fact, these dual variables give the marginal effect on the objective function of an unitary increase of the delivery of this coal to the plant.

For each subproblem, the initial coal delivery price at the entry of the plant is computed as the sum of the departure price and transportation costs. One main iteration is the solution of the 15 subproblems (5 plants \times 3 time periods) and of one main problem. At the end of the first main iteration, we update the initial delivery price at the entry of the plant by replacing the delivery price by the dual price of constraints (13) or (14).

4.1. THE SUB PROGRAMS

The *sub problems* $SP(k,t)$ where $p(c,k,t)$ is the *delivery price of the coal c* to plant k at time period t are defined as follows:

$$\begin{aligned}
 \min z &= \sum_c p(c,k,t) COAL_{ckt} \\
 \text{st } &\left\{ \begin{array}{ll}
 \sum_c COAL_{ckt} & = 1 \\
 \sum_c COALPRES_{ckt} & \leq gates_k \\
 \frac{\min_{coal_k} COALPRES_{ckt}}{COAL_{ckt}} & \leq \frac{COAL_{ckt}}{\max_{coal_k} COALPRES_{ckt}} \\
 \sum_c ash_c multash COAL_{ckt} & \leq \min_{a \text{ delivered by } k} maxash_a SUM_{kt} \\
 \sum_c sulf_c multisulf COAL_{ckt} & \geq \max_{a \text{ delivered by } k} minsulf_a SUM_{kt} \\
 \sum_c sulf_c multisulf COAL_{ckt} & \leq \min_{a \text{ delivered by } k} maxsulf_a SUM_{kt} \\
 \sum_c alk_c multalk COAL_{ckt} & \leq \min_{a \text{ delivered by } k} maxalk_a SUM_{kt} \\
 \sum_c vol_c COAL_{ckt} & \geq minvol SUM_{kt} \\
 \sum_c vol_c COAL_{ckt} & \leq maxvol SUM_{kt} \\
 \sum_{c \in LV} COAL_{ckt} & \geq \max_{a \text{ delivered by } k} minlv_a SUM_{kt} \\
 \sum_{c \in LV} COAL_{ckt} & \leq \min_{a \text{ delivered by } k} maxlv_a SUM_{kt} \\
 \sum_{c \in MV} COAL_{ckt} & \leq maxmv SUM_{kt} \\
 \sum_{c \in MV} COAL_{ckt} & \geq minmv SUM_{kt} \\
 \sum_{c \in MV} COAL_{ckt} & \leq maxsoft SUM_{kt} \\
 c \text{ is Soft} & \\
 \sum_{c \text{ is Austr.}} COAL_{ckt} & \leq maxaus SUM_{kt}
 \end{array} \right. \quad (22)
 \end{aligned}$$

Every variables must be nonnegative behalve $COALPRES$ which is binary.

The subproblems correspond to the following task: find, for each time period and each plant, the *optimal mix for one unit of coal mix*. The constraints include all the mix quality constraints, the maximal number of coals and the minimal and maximal rate for each coal in the mix. The only variables are variables $COAL_c$ and $COALPRES_c$. Recall that the objective coefficients include at the first iteration the coal purchasing cost, transportation cost, handling cost and the production cost. At the following iterations, we only consider the sum of the dual variable and the production cost.

Consequently, we need to solve as many problems as there are plants and time periods. We note $coal(c, k, t, m)$ the *proportion of coal c* used in mix m (initially, m is set to 1) at plant k at time period t .

4.2. THE MAIN PROBLEM

The *main problem* corresponds to the determination of the delivery planning of coals to the plants, the allocation of ordered quantities to the harbours, the determination of the level of use of the mixes generated by the subproblems for each plant and each time period and finally the coke transportation plan to the clients. We obtain the following *main problem*:

$$\begin{aligned}
\min z = & \sum_c \sum_t price_{ct}(expq_{ct} + ORDQ_{ct}) + \sum_b \sum_t \sum_h rate \cdot delprice_{bth} \\
& STOCK_{bth} + \sum_b \sum_t \sum_h (boatcost_{bh} + dockcost_h) XBOAT_{bth} \\
& + \sum_b \sum_k \sum_t \sum_h transpcost_{hk} COALDEL_{bkth} \\
& + \sum_r \sum_k \sum_t railcost_{rk} XRAIL_{rkt} + \sum_k \sum_t \sum_m prodcost_{kt} SUM_{ktm} \\
\text{s.t. } & \left\{ \begin{array}{l}
SUM_{ktm} \leq MIXPRES_{ktm} cap_k days_t \\
minuse_k cap_k days_t \leq \sum_m SUM_{ktm} \leq cap_k days_t \\
expq_{bt} + ORDQ_{bt} = \sum_m XBOAT_{bth} \\
expq_{rt} + ORDQ_{rt} = \sum_h XRAIL_{rkt} \\
\sum_h COALDEL_{bkth} = \sum_k coal(b, k, t, m) SUM_{ktm} \\
XRAIL_{rkt} = \sum_m coal(r, k, t, m) SUM_{ktm} \\
\sum_m MIXPRES_{ktm} \leq maxmix_{kt} \\
STOCK_{bth} - STOCK_{b,t-1,h} = XBOAT_{bth} - \sum_k COALDEL_{bkth} \\
\sum_m \sum_c (1 - wet_c) coal(b, k, t, m) SUM_{ktm} = \sum_a PROD_{kta} \\
\sum_k \sum_a PROD_{kta} \geq dem_{at}.
\end{array} \right.
\end{aligned} \tag{23}$$

Every variables must be nonnegative behalve $MIXPRES$ which is binary.

This gives a *mixed integer program* with *binary variables* $MIXPRES$ since a maximal number of different mixes for each plant at each time period must be imposed. This program can also be solved by GAMS/OSL.

At the first iteration, there are 481 rows and 556 columns. Each iteration adds 15 constraints (15 “rows”) and 30 variables (“columns”) to the main problem (see Tab. 11). The new variables correspond to the 15 new mixes binary variables

TABLE 11. Evolution of the MP objective function.

Iteration	Objective value (€)	Execution time (s)	Number of rows	Number of columns
1	91 569 578.20	3	481	556
2	79 432 349.83	4	496	586
3	75 029 945.98	11	511	616
4	71 579 163.48	48	526	646
5	69 789 373.33	49	541	676
6	68 615 626.73	522	556	706
7	68 514 039.00	52	571	736
8	68 513 503.25	409	586	766
9	68 428 567.03	119	601	796
10	68 393 561.65	4151	616	826
11	68 388 314.23	664	631	856
12	68 361 242.08	4403	646	886
13	68 358 137.05	494	661	916
14	68 349 665.80	2412	676	946
15	68 347 547.93	1865	691	976
16*	68 346 480.40	3184	706	1006

* Stopping criterion satisfied.

and the to 15 new mixes utilization levels. The new constraints correspond to the computation of the *SUM* variables corresponding to the 15 new mixes.

4.3. THE GLOBAL PROCESS

The global process includes 3 levels of inner loops in GAMS/OSL (see Brooke, Kendrick and Meeraus [1]):

- the first loop on the *main iterations*,
- the second loop on the *time periods*,
- and the third loop on the *plants*.

It can be summarized by the following procedure:

```

Initialize p(c,k,t) as the sum of the unit production cost
and the delivery price of coal;
For m = 1, 2, etc..
{
  For t = 1 to 3
  {
    For k = 1 to 5
      {Determine the optimal mix for plant k, time t

```

```

        when using  $p(c,k,t)$  as objective coefficients
        by solving the subproblem  $SP(k,t)$ ;
    Let  $coal(c,k,t,m)$  be the solution obtained;
    }
}
Determine the optimal coal purchases, mix use level and demand
satisfaction plan by solving the main problem MP
using MIP solveur GAMS/OSL;
Update  $p(c,k,t)$  to the sum of the production unit cost
and the dual variable of the balance equation for coal  $c$ 
at plant  $k$  for time period  $t$ ;
}

```

The *stopping criterion* is as follows: stop if there are no columns pricing out. Before analyzing these results, let us stress the fact that the procedure we propose is *heuristic*. In fact, with an alternative organization of the master problem and subproblem hierarchy, we can achieve a better objective function value (see Sect. 5).

Let us conclude this section with a few explanations on the “*dual prices*” used in the column generation. In fact, our master problem is a *mixed integer problem* due to the MIXPRES binary variables. It is known in mathematical programming that the dual information is only valid for a *linear program*. We explain how this dual information is generated. We have solved this MIP problem using GAMS/OSL. After the solver has found the optimal solution, GAMS fixes the levels of the integer variables and reruns the model as an LP model to obtain the *dual variables values*. As an alternative heuristic, one could also use the *dual prices of the LP relaxation of the whole MIP problem*. We refer to Section 5 for the comparison of the results of the two possible choices for generating the dual information.

5. NUMERICAL RESULTS

Table 11 presents the successive objective function values for the main program, the execution time for GAMS/OSL and the size of the main program.

Two conclusions can be taken from this table. The first one is the *important objective function decrease* (about 2 500 000 € per iteration) during the *first five iterations*. This global cost reduction continues at a lower rate during the following iterations. Secondly, the heuristic *converges in only 16 main iterations*. Recall also that, since we use a heuristic method, the solution can be suboptimal.

Now, let us say a few words about *two other heuristics* that were suggested to us to solve the problem. The first one is a variant of our column generation method that was suggested to us by Professor Yves Pochet from CORE, Université Catholique de Louvain, Belgium. Since the only difficulty in the main program is the fact that there are binary variables ($COALPRES_{cktm}$) in a product with another variable SUM_{ktm} (see Eq. 21), one can, for the main program, *fix the coals present in the mix instead of fixing their proportion in the mix*. This implies

TABLE 12. MP objective function for the second heuristic.

Iteration	Objective value (€)	Execution time (s)	Number of rows	Number of columns
1	89 378 751.35	3	1124	796
2	76 709 252.58	23	1782	1066
3	75 052 581.23	18	2440	1336
4	70 284 765.43	215	3098	1606
5	68 484 730.45	49	541	676
6	68 458 216.33	1497	4414	2146
7	68 408 431.40	543	5072	2416
8	68 374 826.85	1010	5730	2686
9	68 362 017.08	599	6388	2956
10	68 353 850.75	8165	7046	3226
11	68 348 399.20	10803	8362	3766
12	68 345 582.43	10803	646	886
13	68 342 647.50	10803	9020	4036
14*	68 341 879.48	7575	9678	4306

* Stopping criterion satisfied.

that variables sent from the lower to the upper level are not the COAL variables, but only the *COALPRES* variables. This also implies that we must impose, in the main program, *all the quality constraints on the mix* listed in appendix A.

Table 12 presents the evolution of the objective function value, the execution time for GAMS/OSL and the size of the main program for this second heuristic method.

A *slightly better solution* (68 341 879.48 € instead of 68 346 480.40 €) was obtained after 14 main iterations but *took much longer* (14 hours 28 min instead of 5 hours 6 minutes for the first heuristic). This increase in the execution time is due to the fact that the size of the main problem is increased (compare the last two columns of Tables 11 and 12). From this example, one can conclude that the second method is not competitive, as it gives a similar solution in much more time. This example also points out the fact that the proposed solution procedure is heuristic since the second solution method gives a *slightly better solution*.

Finally, we present a *third heuristic* that was suggested to us by Professor Etienne Loute from Facultés Universitaires Saint Louis, Brussels, Belgium. The only difference between this and the first method is the choice of the dual information to update the coal delivery prices for the subproblem. Instead of using the dual information given at the optimal solution of the MIP version of the master problem, the relaxed version of this master problem is solved and the dual information of this RMIP problem is then used. Table 13 presents the evolution of the objective

TABLE 13. MP objective function for the third heuristic.

Iteration	Objective value (€)	Execution time	
		RMIP (s)	MIP (s)
1	91 569 578.20	2	2
2	79 432 349.83	2	4
3	75 029 945.98	2	19
4	72 105 817.88	2	24
5	68 835 693.58	3	1071
6	68 503 629.70	2	468
7	68 433 685.88	3	1 243
8	68 400 568.13	2	75
9	68 379 477.03	2	624
10	68 358 511.35	2	270
11	68 349 689.18	2	390
12	68 349 689.18	2	385
13	68 348 287.88	2	2458
14*	68 346 804.38	2	555

* Stopping criterion satisfied.

function value, the execution time for GAMS/OSL for the RMIP and for the MIP version of the master program.

This third solution method is close to the standard column generation technique. At the end of the procedure, the integrality of variables for the main problem is restored as follows: after solving the LP-relaxation, a branch and bound procedure is performed with these columns.

A *slightly more expansive solution* (68 346 804.38 € instead of 68 346 480.40 €) was obtained after 14 main iterations but in *less time* (2 h 10 min instead of 5 h 6min with our method). All these run times were in fact acceptable to the industry since we are determining the planning for three months. The solution procedure was presented to the industrial company and helped the company in both directions. On one side, it helps the industrial company to discover the limiting constraints of its production problem. For example, the industrial company was not aware that the number of entering gates is an important constraint for the problem. In fact, if we allow as many coal type as we want in the mix, the production costs decreases. On the other side, our heuristic has helped the industrial company to decrease its production cost for the three months.

6. CONCLUSIONS

In this paper, we have formulated and solved a real world problem. It concerns the integration of the distribution and mixing of coals to satisfy the demand of

coke of a set of customers throughout a planning horizon. The modelling of the problem implies a large number of constraints with binary variables. The binary variables come from the limitation of the number of coals in the mixes at each plant at each time period and from the limitation of the number of mixes used for each plant and each time period. Some of the constraints were nonlinear. To solve this difficult problem, we have presented three heuristics based on the idea of the column generation technique. The implementation of the heuristics has helped the industrial company to reduce its production cost.

In the future, we intend to add new quality constraints on the mix which are nonlinear and that can not be linearized by the present heuristic method. It concerns upper and lower bound on the Gieseler Fluidity, which is a nonlinear function of the coal quantities in the mix (see Appendix B for a complete presentation of these new constraints).

Acknowledgements. The authors would like to thank Professor Yves Pochet of CORE for suggesting to us the other solution technique, professor Etienne Loute for valuable suggestions to improve the presentation of the paper and finally, Gráinne Corry of University of Littoral and Catherine De Wolf of MIT, for carefully reading this paper.

APPENDIX A. LINEAR QUALITY CONSTRAINTS ON THE MIX

- maximal ash rate of the mix:

$$\sum_c ash_c multash COAL_{cktm} \leq \min_{a \text{ delivered by } k} maxash_a SUM_{ktm}$$

- minimal sulfur rate of the mix:

$$\sum_c sulf_c multisulf COAL_{cktm} \geq \max_{a \text{ delivered by } k} minsulf_a SUM_{ktm}$$

- maximal sulfur rate of the mix:

$$\sum_c sulf_c multisulf COAL_{cktm} \leq \min_{a \text{ delivered by } k} maxsulf_a SUM_{ktm}$$

- maximal alkali rate of the mix:

$$\sum_c alk_c multalk COAL_{cktm} \leq \min_{a \text{ delivered by } k} maxalk_a SUM_{ktm}$$

- minimal and maximal volatiles rate of the mix:

$$minvol SUM_{ktm} \leq \sum_c vol_c COAL_{cktm} \leq maxvol SUM_{ktm}$$

- minimal Low Volume rate of the mix:

$$\sum_{c \text{ is a Low Volume}} COAL_{cktm} \geq \max_{a \text{ delivered by } k} minlv_a SUM_{ktm}$$

TABLE 14. Lower and upper bounds for $I10$ and $I20$.

Parameter	Lower limit	Upper limit
$I10$		23.
$I20$	76.5	

- maximal Low Volume rate of the mix:

$$\sum_{c \text{ is a Low Volume}} COAL_{cktm} \leq \min_a \text{ delivered by } k \text{ } maxlv_a \text{ } SUM_{ktm}$$

- minimal and maximal Mid Volume rate of the mix:

$$minmv \text{ } SUM_{ktm} \leq \sum_{c \text{ is a Mid Volume}} COAL_{cktm} \leq maxmv \text{ } SUM_{ktm}$$

- maximal rate for Soft coal in the mix:

$$\sum_{c \text{ is a Soft Coal}} COAL_{cktm} \leq maxsoft \text{ } SUM_{ktm}$$

- maximal rate for Australian coal in the mix:

$$\sum_{c \text{ is an Australian coal}} COAL_{cktm} \leq maxaus \text{ } SUM_{ktm}$$

APPENDIX B. NONLINEAR QUALITY CONSTRAINTS ON THE MIX

Two specifications are related to the coal quantities in the mix by nonlinear relations: this two indicators are named $I10$ and $I20$ and have lower and upper bounds given by Table 14.

To compute the two indicators $I20$ and $I10$, we need three variables:

- The first one is the *mix inert rate*, noted $TICM$, computed as the weighted sum of the inert rate of the coal in the mix.
- The second one is the *agglutinating power of reactive materials in the mix*, noted $RCIM$.
- The third one is the *maximal fluidity of Gieseler of the mix*, noted $LGFM$.

The indicator $I20$ is computed as the following function of $TICM$, $RCIM$ and $LGFM$:

$$I20 = -1.64 \times TICM + 209.1 \times RCIM + 58.96 \times LGFM \quad (24)$$

$$+ .0168 \times TICM^2 - 110.8 \times RCIM^2 - 13.16 \times LGFM^2 \quad (25)$$

$$- 108.66 \times RCIM \times LGFM + .818 \times LGFM \times TICM \quad (26)$$

$$+ 1.015 \times LGFM^3 + 43.34 \times (RCIM^2) \times LGFM \quad (27)$$

$$+ 6.32 \times RCIM \times LGFM^2 - .02028 \times (TICM^2) \times LGFM \quad (28)$$

$$+ .663 \times TICM \times RCIM \times LGFM - 6.1. \quad (29)$$

The indicator $I10$ is computed as the following function of $TICM$, $RCIM$ and $LGFM$:

$$I10 = 1.625 \times TICM - 200.7 \times RCIM - 61.90 \times LGFM \quad (30)$$

$$- .0159 \times TICM^2 + 105.8 \times RCIM^2 + 12.52 \times LGFM^2 \quad (31)$$

$$+ 114.83 \times RCIM \times LGFM - .654 \times LGFM \times TICM \quad (32)$$

$$- .938 \times LGFM^3 - 47.02 \times (RCIM^2) \times LGFM \quad (33)$$

$$- 6.03 \times RCIM \times LGFM^2 + .01761 \times (TICM^2) \times LGFM \quad (34)$$

$$- .734 \times TICM \times RCIM \times LGFM + 101. \quad (35)$$

REFERENCES

- [1] A. Brooke, D. Kendrick and A. Meeraus, *GAMS User's guide Release 2.25*, The Scientific Press, San Francisco (1992).
- [2] G. Desaulniers, J. Desrosiers and M.M. Solomon, *Column Generation*. Springer (2005).
- [3] D. De Wolf, *Using column generation to solve an industrial mixing problem*, CORE Discussion Paper N° 2003/42, Universit Catholique de Louvain (2003).
- [4] H. Greenberg, Analyzing the pooling problem. *ORSA J. Comput.* **7** (1995) 205–217.
- [5] L.S. Lasdon and J.C. Plummer, *SLP user guide*. XMP Corporation (1986).
- [6] R. Sarker and E. Gunn, A simple SLP algorithm for solving a class of nonlinear programs. *Eur. J. Oper. Res.* **101** (1997) 140–154.
- [7] F. Vanderbeck and L. Wolsey, An exact algorithm for IP column generation. *Oper. Res. Lett.* **19** (1996) 151–159.
- [8] F. Vanderbeck, On Dantzig-Wolfe decomposition in integer programming and ways to perform branching in a branch-and-price algorithm. *Oper. Res.* **48** (2000) 111–128.
- [9] F.J. Vasko, D.D. Newhart and A.D. Strauss, Coal blending models for optimum cokemaking and blast furnace operation. *J. Oper. Res. Soc.* **56** (2005) 235–243.
- [10] H.P. Williams, *Model building in Mathematical Programming*. John Wiley (2013).
- [11] H.P. Williams, *Model solving in Mathematical Programming*. John Wiley (1992).