

A two-stage mixed integer programming approach for optimizing the skill mix and training schedules for aircraft maintenance

De Bruecker P, Van den Bergh J, Beliën J, Demeulemeester E.





KU LEUVEN

A two-stage mixed integer programming approach for optimizing the skill mix and training schedules for aircraft maintenance

Philippe De Bruecker¹, Jorne Van den Bergh^{2,3}, Jeroen Beliën^{2,3}, Erik
Demeulemeester¹

¹ KU Leuven, Research Center for Operations Management, Leuven (Belgium),
philippe.debruecker@kuleuven.be, erik.demeulemeester@kuleuven.be

² HUBrussel, Center for Informatics, Modeling and Simulation, Brussels (Belgium),
jorne.vandenbergh@hubrussel.be, jeroen.belien@hubrussel.be

³ Affiliated researcher KU Leuven, Research Center for Operations Management,
Leuven (Belgium)

November 25, 2015

Abstract

This paper presents a two-stage mixed integer programming approach for optimizing the skill mix and training schedule at the aircraft maintenance company Sabena Technics. Of course, when all workers are trained for all skills, cheaper workforce schedules are possible. However, the training that is required to acquire all those skills can become very expensive. In the first stage of our two-staged approach, we therefore make a trade-off between the training costs and the resulting cheaper workforce schedule. As we assume that workers are unavailable to work during their training, the obtained result is only applicable in practice if the required training can be performed without endangering the current maintenance operations. In the second stage, we therefore want to find an optimal and feasible training schedule in order to obtain the desired skill mix with minimal costs. We illustrate our models with a computational experiment based on real-life data of Sabena Technics.

Keywords: Aircraft maintenance, skills, training, mixed integer programming

1 Introduction

In service industries, labor is often the largest expenditure for a company. An efficient scheduling of the workforce is therefore very important. This paper presents a two-stage mixed integer programming approach for optimizing the skill mix and training schedule at an aircraft maintenance company. In this study, we only focus on the line maintenance which takes place at the gate or parking ramp between the arrival and departure of an aircraft. Line maintenance consists of on-call assistance and routine checks of the engines, the landing gear, etc. Since different aircraft have different features and can show different problems, only adequately skilled workers should be assigned to maintain certain flights. In fact, according to aviation legislation, each type of aircraft requires its own maintenance license. Only workers that acquired the required license are allowed to maintain that aircraft. Hence, with skills we actually mean licenses in this case. As we assume that there is no hierarchical structure between the licenses, we deal with categorical skills (De Bruecker, Van den Bergh, Beliën & Demeulemeester, 2015). In accordance to the categorisation of De Bruecker et al. (2015), we also see that we consider individual skills instead of group skills as each worker that is assigned to maintain a certain aircraft must be licensed for that aircraft.

A good personnel schedule should make sure that all flights can be maintained with the available workers and their respective skills. Schedules based on the current (limited) skill pool can however be very expensive. Therefore, as a first step, we follow an integrated approach to build inexpensive personnel schedules and to design the optimal skill mix at the same time. The second step is to build the lowest cost training schedule that leads to this optimal skill mix. We illustrate our procedure using real-life data from Sabena Technics.

2 Literature review

Personnel scheduling has already been studied many times in the operations research literature. Several review articles exist that focus on general workforce planning problems (Alfares, 2004; Burke, De Causmaecker, Vanden Berghe & Van Landeghem, 2004; Ernst, Jiang, Krishnamoorthy & Sier, 2004; Kohl & Karisch, 2004; Bergh, Beliën, De Bruecker, Demeulemeester & De Boeck, 2013), but none go into detail regarding skills. To the best of our knowledge, the problem presented in this paper is unique in the literature since we integrate shift construction, skill constraints and training. Several solution approaches have already been proposed in the literature to build personnel schedules incorporating skills. Many models take into account flexible shifts or training, but none of them incorporates both.

Scheduling the workforce in shifts is usually much harder than without shifts. On top of the usual coverage constraints, shift planning adds extra constraints about the start time and duration of shifts, shift successions, team sizes, etc. Only a limited number of papers consider shift planning in combination with skills. While most papers consider shifts that are predefined (with a fixed start and end time, a fixed team size,...), only a few incorporate decisions about the composition of shifts in their model (Cai & Li, 2000; Aickelin & White, 2004; Bard & Purnomo, 2005a, 2005b; Aickelin & Dowsland, 2004; Bard & Wan, 2008; Aickelin, Burke & Jingpeng, 2009; Detienne, Péridy, Pinson & Rivreau, 2009; Avramidis, Chan, Gendreau, L'Ecuyer & Pisacane, 2010; Brunner, Bard & Kolisch, 2011; Brunner & Edenharter, 2011). To incorporate shifts in a workforce planning problem, the set covering formulation developed by Dantzig (1954) is most often used (Bard, Binici & deSilva, 2003). The number of possible shift patterns will increase dramatically in real-life problems when (lunch) breaks, flexible start times, etc. are considered (Bergh et al., 2013). Therefore, when solving optimization problems that include shifts, researchers often tend to use heuristics instead of exact solution techniques.

Incorporating skills and shifts into a personnel scheduling problem often involves assigning workers with appropriate skills to different shifts. Therefore, one must start from the current pool of workers and their current skills. Aickelin and Dowsland (2004) and Aickelin and White (2004), for example, assign nurses from the current pool of workers to shifts using a genetic algorithm. Also Cai and Li (2000) use a genetic algorithm to solve a personnel scheduling problem at a transportation company. Aickelin et al. (2009) use an evolutionary squeaky wheel optimization approach to solve the nurse scheduling problem. As in our case, Brucker, Burke, Curtois, Qu and Vanden Berghe (2010) consider a personnel scheduling problem with cyclic schedules. While we want to construct shifts and assign workers at the same time, Brucker et al. (2010) propose to use a two phased approach. They use a greedy algorithm to solve the optimization model.

Besides some popular meta-heuristics, linear programming also remains attractive to

solve complex workforce scheduling problems. Bard and Purnomo (2005a, 2005b) tackle the same kind of problem as Aickelin et al. (2009), Aickelin and Dowsland (2004) and Aickelin and White (2004) and use column generation to solve the optimization model. Brunner et al. (2011) and Brunner and Edenharter (2011) also use column generation to solve the staff scheduling problem of physicians. Bard and Wan (2008) even use pure integer linear programming to assign workers to machines and shifts.

To the best of our knowledge, studies that incorporate training and shift construction do not exist in the personnel scheduling literature. Training is a straightforward way to enhance the skills of a person in order to increase his performance (quality, speed of work and efficiency) or to allow the worker to perform a different task. The latter is referred to as cross-training. A trade-off must be made between the training cost (loss of available working time and/or monetary costs) and the possible benefits (the increased flexibility or performance). While learning happens automatically, training must be planned. Hence, decisions must be made by the workforce planning model to determine who should receive training and when this should take place.

In the work of Li and Li (2000), Huang, Lee, Song and Eck (2009), Song and Huang (2008), Marentette, Johnson and Mills (2009) and Fowler, Wirojanagud and Gel (2008), the planning decision for cross-training the workforce is incorporated in the model. The model decides on how many employees should be trained to acquire certain skills. Li and Li (2000), Song and Huang (2008) and Fowler et al. (2008) make these decisions on a weekly or monthly basis. Huang et al. (2009) and Marentette et al. (2009) only make the training decision on a yearly basis and therefore fail to provide the exact timing of the training. Tiwari, Patterson and Mabert (2009) and Zülch, Rottinger and Vollstedt (2004) neglect the training decision during the workforce planning optimization problem and analyze the bottleneck skills only afterwards. In a second phase, they provide an advice to the management about what skills should be trained. The personnel scheduling literature regarding skills apparently also lacks studies that include the precise timing of training.

To the best of our knowledge, the reduced availability or even total absence of the worker during his training period has never been incorporated in a mathematical model with skills. Training can take place on the job or externally. Training on the job is preferred in most cases because of the direct link with the company and the lower costs. However, according to Boulay and Medway (1999), training on the job is very difficult these days. For some tasks for example, it is prohibited by law to involve people without the necessary skills or qualifications. Therefore, G. Beddoe, Petrovic and Li (2009) and G. R. Beddoe and Petrovic (2006) talk about eye-training instead of on the job training. Because of high work pressure, it might also be impossible to involve people that slow down the work process.

3 Problem definition

Aviation industries are often characterized by a highly seasonal demand. Hence, the demand for aircraft maintenance also follows a seasonal pattern. Typically, aircraft maintenance companies have to build new workforce schedules twice a year, i.e., before the start of the winter as well as the summer season. A special feature of the aircraft maintenance business is the fact that every week the same set of flights must be maintained. These flight schedules are known months in advance of the start of a new season and contain information such as the STA (Scheduled Time of Arrival), the STD (Scheduled Time of Departure) and the workload. In this paper we also assume that the flight schedule specifies the required skill for each flight. An example of such a weekly flight schedule with skill requirements is shown in Table 1.

Table 1: Example of a weekly flight schedule

Flight	Company	STA	STD	Workload (man-hours)	Skill
1	AA	Monday 22:05	Tuesday 07:45	4.00	A
...
56	BA	Thursday 07:30	Thursday 11:40	6.00	A
...
100	SN33	Saturday 05:30	Saturday 10:45	4.25	B

Once the demand for maintenance is known for the next season, the construction of an optimal workforce schedule for the next season can start. Figure 1 visualizes this process. For a detailed description of the problem details and the specifics about the cycles, shifts and teams in the workforce configuration, we refer to the paper of Beliën, Demeulemeester, De Bruecker, Van den Bergh and Cardoen (2013).

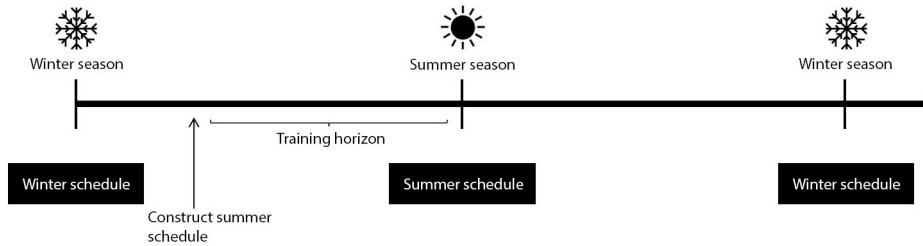
In the example shown in Figure 1, the company is currently operating based on the winter flight schedules and the winter workforce schedule. When the summer flight schedules become known, the company has to construct the summer workforce schedule. We distinguish between two important tasks in constructing an optimal summer workforce schedule.

The first task is to build an efficient schedule to make sure that all flights can be maintained with the available workers and their respective skills. We therefore have to make two major decisions. First, we have to decide on the scheduling of the shifts, and second, we have to assign the available workers to these shifts to make sure that all flights can be maintained in time with the available skills. These two decisions are integrated in the optimization model to build optimal workforce schedules.

The second task is to find the optimal skill mix. When the number of skills per worker is relatively small, too many workers would be needed to construct a feasible workforce configuration leading to very high costs. Using the current available workers and their current skills to construct the summer workforce schedule therefore only leads to a suboptimal result. Cheaper schedules can be obtained by training some of the workers. These higher skilled workers can then do some of the work for which they were not certified earlier. This can decrease the required number of staff and hence decrease costs. The workers that are not needed anymore to do the line maintenance can then be used somewhere else in the company such as in the heavy maintenance.

Because training can be very expensive, a trade-off must be made between cheaper rosters that require higher skilled workers and the training costs to obtain this higher skilled workforce. Therefore, these two tasks are incorporated in the same optimization model to build optimal workforce schedules. This model is presented in Section 4.1 and is referred to as the workforce scheduling model.

Figure 1: Overview of the seasons



While the workforce scheduling model results in a decision about the optimal skill mix and the skills that must be trained, it does not specify when this training should take place. To build the lowest cost training schedule that results in the optimal skill mix, another optimization model is constructed in Section 4.3. This model takes into account that workers are unavailable to work during their training. The goal of this model is to schedule the training during the training horizon (see Figure 1). Based on the current winter workforce schedule, this model decides who will be trained on each day in each week for each skill.

4 Methodology

4.1 Workforce scheduling model

The workforce scheduling model presented in this paper is based on the MILP formulation from Beliën et al. (2013). However, as we now incorporate skills, the model

formulation is quite different. In this section we therefore present the complete mathematical formulation of the workforce scheduling optimization model with skills. We first list the sets, along with their associated indices:

$d \in \{1, \dots, 7\}$:	days in the week
$p \in P$:	time periods in one week
$f \in F$:	flights to be serviced
$p \in P_f$:	time periods during which flight f can be serviced
$s \in S$:	skill types
$c \in C$:	cycles in the workforce configuration
$\omega \in \Omega$:	available workers
$\eta \in \{1, 2, \dots, M_c\}$:	workers in a team in cycle c , with M_c the team size in cycle c
$i \in I$:	feasible shifts
$t \in T$:	shift types $\{M, D, E, N\}$
$i \in I_t$:	feasible shifts of shift type t

The coefficients and right hand side constants are presented below:

L_{fs} :	the workload (in man-hours) of flight f of skill type s
A_{idp} :	= 1 if period p is included in shift i on day d , = 0 otherwise
B_{idp} :	the fraction of workers available to work in shift i on period p on day d
H_i :	the duration of shift i (in hours)
$\alpha_{\omega s}$:	= 1 if worker ω has skill s , = 0 otherwise
$K_s^{training} > 0$:	the training cost (for one worker) to acquire skill type s
$K_{id} > 0$:	the total cost (for one worker) of shift i on day d
W^l :	the minimum number of weeks (= teams) in a cycle
W^u :	the maximum number of weeks (= teams) in a cycle
$W^{planning}$:	the number of weeks in the planning horizon
H^l :	the minimum average number of working hours per week
H^u :	the maximum average number of working hours per week
R :	the maximum fraction of working weekends; i.e., weekends during which at least one shift is scheduled. $R = (\text{number of working weekends})/(\text{number of weeks in the cycle})$
Θ_ω^{max} :	the maximum training (in skills) worker ω can receive
M :	Big M, a large positive number
M_c :	the team size in cycle c

The decision variables are:

$a_{\omega c \eta} \geq 0 :$	> 0 if worker ω is assigned as the η -th worker of the team in cycle c , $= 0$ otherwise
$a'_{\omega c \eta s} \in \{0, 1\} :$	$= 1$ if worker ω with skill s is assigned as the η -th worker of the team in cycle c , $= 0$ otherwise
$q_{s \eta c p} \geq 0 :$	the capacity of skill type s made available by worker η in cycle c on time period p
$x_{i d c} \geq 0 :$	the number of shifts i that is scheduled during day d in cycle c
$g_{f p s} \geq 0 :$	capacity of skill type s assigned to maintain flight f during time period p
$z_{\eta c s}^{skill} \in \{0, 1\} :$	$= 1$ if worker η in cycle c requires skill s , $= 0$ otherwise
$z_{i c} \in \{0, 1\} :$	$= 1$ if shift i is used in cycle c , $= 0$ otherwise
$\tau_{\omega s} \in \{0, 1\} :$	$= 1$ if worker ω is trained to acquire skill s , $= 0$ otherwise
$n_c \in \{W^l, \dots, W^u\} :$	the number of weeks (= teams) in cycle c
$e_{i d c}^+ \in \{0, 1, \dots, W^u\} :$	the number of extra weeks needed in cycle c for day d caused by shifts of type t (with $t \in \{E, N\}$) on the preceding day
$e_{N d c}^- \in \{0, 1, \dots, W^u\} :$	the number of extra weeks needed in cycle c for day d caused by E shifts that can be compensated by an excess in N shifts on the preceding day

The workforce scheduling optimization model can be formulated as follows:

$$\text{Minimize } \sum_{c \in C} \sum_{i \in I} \sum_{d=1}^7 K_{i d} \cdot M_c \cdot x_{i d c} \cdot W^{planning} + \sum_{\omega \in \Omega} \sum_{s \in S} K_s^{training} \cdot \tau_{\omega s} \quad (1)$$

Subject to:

$$\sum_{s \in S} q_{s \eta c p} \leq \sum_{i \in I} \sum_{d=1}^7 B_{i d p} x_{i d c}, \quad \forall c \in C, \forall \eta \in \{1, 2, \dots, M_c\}, \forall p \in P \quad (2)$$

$$\sum_{f \in F} g_{f p s} \leq \sum_{c \in C} \sum_{\eta \in \{1, 2, \dots, M_c\}} q_{s \eta c p}, \quad \forall p \in P, \forall s \in S \quad (3)$$

$$\sum_{p \in P_f} g_{f p s} = L_{f s} * \frac{|P|}{24 * 7}, \quad \forall f \in F, \forall s \in S \quad (4)$$

$$q_{s \eta c p} \leq M \cdot z_{\eta c s}^{skill}, \quad \forall c \in C, \forall \eta \in \{1, 2, \dots, M_c\}, \forall s \in S, \quad \forall p \in P \quad (5)$$

$$z_{\eta c s}^{skill} \leq \sum_{\omega \in \Omega} a'_{\omega c \eta s} - (n_c - 1) + W^u \cdot (1 - z_{\eta c s}^{skill}), \quad \forall c \in C, \forall \eta \in \{1, 2, \dots, M_c\}, \forall s \in S \quad (6)$$

$$a'_{\omega c \eta s} \leq a_{\omega c \eta}, \quad \forall c \in C, \forall \eta \in \{1, 2, \dots, M_c\}, \forall s \in S, \quad (7)$$

$$\sum_{c \in C} \sum_{\eta \in \{1, 2, \dots, M_c\}} a_{\omega c \eta} \leq 1, \quad \forall \omega \in \Omega \quad (8)$$

$$\sum_{c \in C} \sum_{\eta \in \{1, 2, \dots, M_c\}} a'_{\omega c \eta s} \leq \alpha_{\omega s} + \tau_{\omega s}, \quad \forall \omega \in \Omega, \forall s \in S \quad (9)$$

$$\sum_{s \in S} \tau_{\omega s} \leq \Theta_{\omega}^{max}, \quad \forall \omega \in \Omega \quad (10)$$

$$\sum_{\omega \in \Omega} a_{\omega c \eta} = n_c, \quad \forall c \in C, \forall \eta \in \{1, 2, \dots, M_c\} \quad (11)$$

$$n_c \geq \lceil 1/R \rceil \cdot \sum_{i \in I} x_{i6c}, \quad \forall c \in C \quad (12)$$

$$n_c \geq \lceil 1/R \rceil \cdot \left(\sum_{i \in I} x_{i7c} + e_{N7c}^+ + e_{E7c}^+ \right), \quad \forall c \in C \quad (13)$$

$$\sum_{i \in I} \sum_{d=1}^7 H_i x_{idc} \geq H^l \cdot n_c, \quad \forall c \in C \quad (14)$$

$$\sum_{i \in I} \sum_{d=1}^7 H_i x_{idc} \leq H^u \cdot n_c, \quad \forall c \in C \quad (15)$$

$$x_{idc} \leq W^u \cdot z_{ic}, \quad \forall i \in I, \forall d = 1, \dots, 7, \forall c \in C \quad (16)$$

$$\sum_{i \in I_t} z_{ic} \leq 1, \quad \forall t \in T, \forall c \in C \quad (17)$$

$$\sum_{i \in I} \sum_{d=1}^7 \sum_{c \in C} A_{idp} x_{idc} \geq 1, \quad \forall p \in P \quad (18)$$

$$n_c \geq \sum_{i \in I} x_{idc} + e_{Ndc}^+ + e_{Edc}^+, \quad \forall d \in \{1, \dots, 7\}, \forall c \in C \quad (19)$$

$$e_{N(d+1)c}^+ \geq \sum_{i \in I_N} x_{idc} - \sum_{i \in I_N} x_{i(d+1)c}, \quad \forall d \in \{1, \dots, 6\}, \forall c \in C \quad (20)$$

$$e_{E(d+1)c}^+ \geq \sum_{i \in I_E} x_{idc} - \sum_{i \in I_E} x_{i(d+1)c} - e_{N(d+1)c}^-, \quad \forall d \in \{1, \dots, 6\}, \forall c \in C \quad (21)$$

$$e_{N(d+1)c}^- \leq \sum_{i \in I_N} x_{i(d+1)c} - \sum_{i \in I_N} x_{idc} + e_{N(d+1)c}^+, \quad \forall d \in \{1, \dots, 6\}, \forall c \in C \quad (22)$$

$$e_{N1c}^+ \geq \sum_{i \in I_N} x_{i7c} - \sum_{i \in I_N} x_{i1c}, \quad \forall c \in C \quad (23)$$

$$e_{E1c}^+ \geq \sum_{i \in I_E} x_{i7c} - \sum_{i \in I_E} x_{i1c} - e_{N1c}^-, \quad \forall c \in C \quad (24)$$

$$e_{N1c}^- \leq \sum_{i \in I_N} x_{i1c} - \sum_{i \in I_N} x_{i7c} + e_{N1c}^+, \quad \forall c \in C \quad (25)$$

$$x_{idc} \in \{0, 1, \dots, W^u\}, \quad \forall i \in I, \forall d \in \{1, \dots, 7\}, \forall c \in C \quad (26)$$

$$g_{fps} \geq 0, \quad \forall f \in F, \forall p \in P, \forall s \in S \quad (27)$$

$$z_{ic} \in \{0, 1\}, \quad \forall i \in I, \forall c \in C \quad (28)$$

$$n_c \in \{W^l, W^l + 1, \dots, W^u\}, \quad \forall c \in C \quad (29)$$

$$q_{s\eta cp} \geq 0, \quad \forall c \in C, \forall s \in S, \forall \eta \in \{1, 2, \dots, M_c\}, \forall p \in P \quad (30)$$

$$e_{tdc}^+ \in \{0, 1, \dots, W^u\}, \quad \forall t \in \{E, N\}, \forall d \in \{1, \dots, 7\}, \forall c \in C \quad (31)$$

$$e_{Ndc}^- \in \{0, 1, \dots, W^u\}, \quad \forall d \in \{1, \dots, 7\}, \forall c \in C \quad (32)$$

$$a_{\omega c \eta} \geq 0, \quad \forall \omega \in \Omega, \forall c \in C, \forall \eta \in \{1, 2, \dots, M_c\} \quad (33)$$

$$a'_{\omega c \eta s} \in \{0, 1\}, \quad \forall \omega \in \Omega, \forall c \in C, \forall \eta \in \{1, 2, \dots, M_c\}, \forall s \in S \quad (34)$$

$$z_{\eta cs}^{skill} \in \{0, 1\}, \quad \forall c \in C, \forall \eta \in \{1, 2, \dots, M_c\}, \forall s \in S \quad (35)$$

$$\tau_{\omega s} \geq 0, \quad \forall \omega \in \Omega, \forall s \in S \quad (36)$$

The model presented above is an extension of the MILP model presented by Beliën et al. (2013) as it incorporates skills and training. Constraints (2) to (11) can be seen as the extension, while constraints (12) to (25) can be found in the original model. The extension presented above takes into account the skills of the available workers and the skill restrictions for the maintenance jobs. The extended model allows to do two different things. First, it can be used to build personnel schedules based on the available skill pool. Second, the model allows to investigate the possible benefits of training. Therefore, when we assume a sufficient number of workers with all possible skills, the extended model will result in the same solutions as those found by the original model of Beliën et al. (2013).

Note that the extended model presented above features the same technique as the original model to avoid the non-linearity caused by the team size decision. As the team size is incorporated into the model as a parameter (M_c) instead of a decision variable, the same enumeration technique with bounding can be used to solve the model as in the paper of Beliën et al. (2013). Moreover, we also follow the same technique as Beliën et al. (2013) to reduce the computation time by replacing the binary variable x_{idwc} by the integer variable x_{idc} . Thus, instead of assigning each shift to a *particular day in a particular week*, we now limit our decision to *the number of shifts on a particular day*. Because of this simplification, we omit the index for the week (team) in our model and assume that each team in a cycle consists of the same number and type of workers. The assignment of shifts to weeks (teams) can be done afterwards. It should be clear that it is therefore not possible in the model presented above to distinguish between the workers of different teams in a cycle. With the index $\eta \in \{1, 2, \dots, M_c\}$, we can only make a distinction between the workers in the same team. All other teams in the cycle are duplicates of this team. Hence, the variable $q_{s\eta cp}$ must be seen as the sum of the capacity over all teams in cycle c of skill type s made available by worker η on time

period p with $\eta \in \{1, 2, \dots, M_c\}$. With this in mind, we can now explain the model.

The objective function (1) minimizes the total labor costs over the entire planning horizon plus the total training costs. As we assume a planning horizon of one season with 24 weeks, the weekly labor costs are multiplied by 24 ($W^{planning} = 24$). A trade-off is made between cheaper rosters that require higher skilled workers and the training costs to obtain this higher skilled workforce.

Constraint (2) distributes the capacity made available by the shifts over the different workers and their skills. Hence, it determines the capacity of skill type s made available by worker η in cycle c on time period p . Summing over s makes sure that the capacity over all skills is not larger than what is available. Constraint (3) ensures that the assigned capacity to maintain the flights is less than the available capacity. Constraint (4) then checks that there is sufficient capacity of the required skill assigned to a flight in order to be maintained in time. In Constraint (4), $\frac{|P|}{24*7}$ is a conversion factor such that the units of the RHS match the units of the LHS.

While constraints (2) to (4) determine the shifts, the required capacity and the skills, constraints (5) to (11) make sure that this can be achieved with the current available workforce and skills. According to constraint (5), $z_{\eta cs}^{skill}$ will flip to 1 when worker η in cycle c requires skill s . Constraints (6) and (11) make sure that the assigned workers to cycle c have all the skills that are required according to constraint (5). Recall that we assumed that all teams in a cycle consist of the same workers and skills. Therefore, when skill s is required for worker η in cycle c , there should be exactly n_c workers with skill s assigned to cycle c as the η -th worker of a team. Take for example the situation where skill A is required for worker η in cycle c . Since the LHS of constraint (6) is 1 for skill A, the RHS should also be at least 1. Also assume that there are 6 teams in cycle c . This means that we also need 6 workers in cycle c to work as the η -th worker of a team in order to satisfy constraints (6) and (11).

Constraints (7) and (8) make sure that a worker (ω) can only be assigned once as a team member. Therefore, we introduced the variable $a_{\omega c \eta}$. This variable is introduced to check if the worker is assigned as the η -th worker of the team in cycle c . While it is defined as a continuous variable (for efficiency reasons), it can only be zero or one because of constraints (8) and (11).

To know whether or not we need to train a certain worker, constraint (9) is used to determine the value of $\tau_{\omega s}$. Constraint (10) then limits the amount of training a worker can receive. Recall that workers that are not needed this season in the line maintenance are assigned to the heavy maintenance division of the company. This makes that many of the heavy maintenance workers are actually also qualified for the line maintenance. Therefore, when constructing the line maintenance workforce for the next season, the company cannot only choose from the workers that are currently active in the line

maintenance, but it can also consider to use some qualified workers that are currently active in the heavy maintenance division. However, as there is less idle capacity in the heavy maintenance division as compared to the line maintenance, we assume that only the workers that are currently working in the line maintenance can receive training. Therefore, Θ_{ω}^{max} will be 0 for all workers currently not working in the line maintenance.

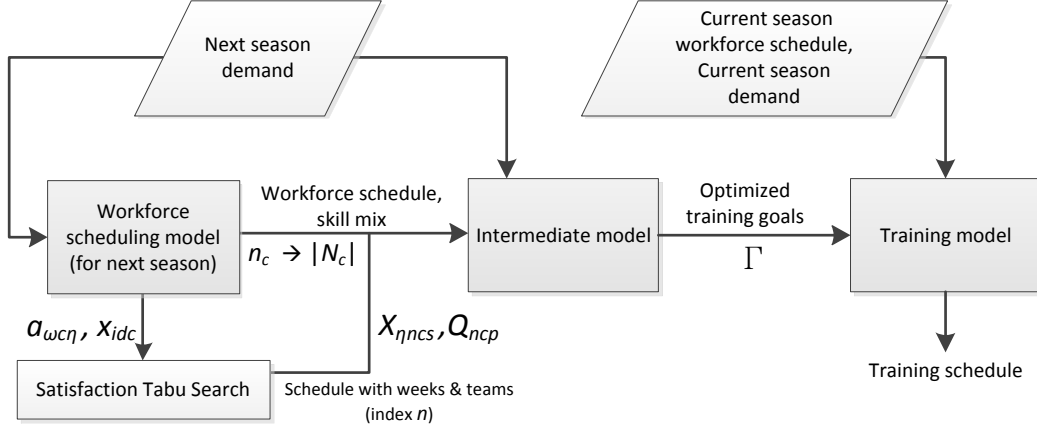
For a thorough discussion of constraints (12) to (25) we refer to the paper of Beliën et al. (2013), as they concern the same set of constraints as in the original model.

The model presented above first results in the assignment of the available workers to cycles and arbitrary teams. In a second step, we should decide on an optimal training scheme for each worker. Since we assumed that all teams are equal in one cycle, the skill requirements (and training requirements) for all teams are also equal. When skill A is required in one team, it is required in each teams of that cycle. This assumption can result in a suboptimal situation. Take the following example where there are two time periods during one week and four flights that require maintenance. Flight 1, 2 and 3 arrive and leave during the first time period. Flight 4 arrives and leaves during the second time period. Flight 1 requires one period of work of skill A, flight 2 requires one period of work of skill B, flight 3 requires two periods of work of skill C and flight 4 requires one period of work of skill C. Assume that the duration of a shift must be one time period and that we can have no more than two teams. Assuming that all skills have the same cost (1 Euro for one period), an optimal schedule is to schedule two shifts during the first time period and one shift during the second. This optimal solution has two unequal teams (i.e., with a different skill mix) of two workers and a total labor cost of 6 Euro. This solution requires that we assign one worker with skill A and one worker with skill C to the first team and one worker with skill B and one worker with skill C to the second team. There is no cheaper option that involves equal teams (i.e., with the same skill mix). Even if we only schedule one team with one worker with skill A, one worker with skill B and two workers with skill C during the two time periods, the total costs are 8 Euro.

From the previous example, we conclude that the assumption of equal teams (i.e., with the same skill mix) can lead to suboptimal solutions. Therefore, we need to refine the solution found by the workforce scheduling model before constructing the training schedule. This is the task of the intermediate model in Section 4.2. This is shown in Figure 2, which presents an overview of the three models that constitute the solution approach. Recall that the assumption of equal teams is the result of omitting the assignment of shifts to weeks because of efficiency issues (see the research of Beliën et al. (2013)). So the first step after solving the workforce scheduling model is to reconstruct a solution with weeks and, hence, with teams. Based on the value of $a_{\omega c \eta}$, each worker ω is assigned to the η -th worker of an arbitrary team n . This way, the resulting value of the decision variable $a_{\omega c \eta}$ from the workforce scheduling model is transformed into input parameter $X_{\eta n c s}$ from the intermediate model based on the value of $\alpha_{\omega s}$. $X_{\eta n c s}$

has value 1 if worker η in team n in cycle c has skill s , and has value 0 otherwise.

Figure 2: Overview of the models that constitute the solution approach



4.2 Optimize training requirements: Intermediate optimization model

As the skill requirements are not necessarily equal for all teams in an optimal solution, we use a second, intermediate mixed integer programming model to reduce the training requirements resulting from the workforce scheduling model. The most important difference with the workforce scheduling model is that the intermediate model does feature the week (team) decision which is absent in the workforce scheduling model. The use of a week (team) index in this model does not cause the same efficiency problems as in the workforce scheduling model because the most difficult decision, the shift scheduling decision, is already made. Hence, we will see that the decision variable x_{idc} from the workforce scheduling model is absent in the intermediate model. In Section 4.1 we already explained how the decision variable $a_{\omega c \eta}$ from the previous model is transformed into input parameter $X_{\eta n c s}$ for this model with the help of the satisfaction Tabu Search. Moreover, based on the output of the satisfaction Tabu Search, the decision variable x_{idc} from the previous model is transformed into input parameter $Q_{n c p}$ for this model. This can also be observed in Figure 2. Furthermore, the decision variable n_c from the previous model is now used as the input parameter $|N_c|$, which equals $|W_c|$. Of course, the team size M_c is also fixed at this time because it was optimized during the enumeration technique for solving the workforce scheduling model.

As the shift decision is already made during the workforce scheduling model, whereby the decision variable x_{idc} is transformed into input parameter $Q_{n c p}$ for this model, we know how much capacity there is available in each week, in each cycle, on each period. In the intermediate model presented below, we neglect the skill and training decisions that have been made during the workforce scheduling model as we will now reinvestigate

the skill and training requirements. The idea is that we can now make better decisions about the skill and training requirements because we can now make decisions for each team separately. As we can now make decisions for each team separately, this eliminates the need for equal skill requirements for each team in the cycle.

In the remainder of this section, we define the intermediate model. We first list the sets, along with their associated indices:

$d \in \{1, \dots, 7\}$:	days in the week
$p \in P$:	time periods in one week
$f, f' \in F$:	flights to be serviced
$f' \in \delta(f)$:	set of all flights that overlap at least during one period with flight f , with $f \in \delta(f)$
$p \in P_f$:	time periods during which flight f can be serviced
$s \in S$:	skill types
$c \in C$:	cycles in the workforce configuration
$n \in N_c$:	teams in cycle c , with $ N_c = W_c =$ the number of teams (=weeks) in cycle c
$w \in W_c$:	weeks in cycle c , with $ W_c = N_c =$ the number of weeks (=teams) in cycle c
$\eta \in \{1, 2, \dots, M_c\}$:	workers in a team in cycle c , with M_c the team size in cycle c

The coefficients and right hand side constants are presented below:

$X_{\eta ncs} \in \{0, 1\}$:	= 1 if worker η in team n in cycle c has skill s , = 0 otherwise
$W^{lem} \geq 0$:	least common multiplier of $ W_c $ for all $c \in C$
L_{fs} :	workload (in time periods p) of flight f of skill type s
$Q_{ncp} \geq 0$:	capacity made available by team n in cycle c on time period p
$K_s^{training} > 0$:	training cost for skill type s per worker
$\Theta_{\eta nc}^{max}$:	maximum training (in skills) that worker η in team n in cycle c can receive
M_c :	the team size in cycle c

The decision variables are:

- $q_{s\eta ncpw} \geq 0$: capacity of skill type s made available by worker η in team n in cycle c on time period p in week w
- $\tau_{\eta ncs} \in \{0, 1\}$: = 1 if worker η in team n in cycle c must be trained for skill s , = 0 otherwise
- $g_{fpsw} \geq 0$: capacity of skill type s assigned to maintain flight f during time period p in week w
- $m_{fs} \geq 0$: minimum capacity (over all weeks $w = 0, \dots, W^{lcm} - 1$) of skill type s that is available between the STA and STD of flight f

The intermediate optimization model can be formulated as follows:

$$\text{Minimize } \sum_{c \in C} \sum_{s \in S} \sum_{n \in N_c} \sum_{\eta \in \{1, 2, \dots, M_c\}} \tau_{\eta ncs} \cdot K_s^{training} \quad (37)$$

Subject to:

$$q_{s\eta ncpw} \leq \frac{X_{\eta ncs} + \tau_{\eta ncs}}{M_c} \cdot Q_{((n+w) \bmod (|W_c|))cp}, \quad \forall c \in C, \forall s \in S, \\ \forall p \in P, \forall w \in W_c, \\ \forall n \in N_c, \forall \eta \in \{1, 2, \dots, M_c\} \quad (38)$$

$$\sum_{s \in S} q_{s\eta ncpw} \leq \frac{Q_{((n+w) \bmod (|W_c|))cp}}{M_c}, \quad \forall c \in C, \forall p \in P, \forall w \in W_c, \\ \forall n \in N_c, \forall \eta \in \{1, 2, \dots, M_c\} \quad (39)$$

$$\sum_{s \in S} \tau_{\eta ncs} \leq \Theta_{\eta nc}^{max}, \quad \forall c \in C, \forall n \in N_c, \\ \forall \eta \in \{1, 2, \dots, M_c\} \quad (40)$$

$$\tau_{\eta ncs} + X_{\eta ncs} \leq 1, \quad \forall c \in C, \forall n \in N_c, \\ \forall \eta \in \{1, 2, \dots, M_c\}, \forall s \in S \quad (41)$$

$$\sum_{f \in F} g_{fpsw} \leq \sum_{c \in C} \sum_{n \in N_c} \sum_{\eta \in \{1, \dots, M_c\}} q_{s\eta ncpw}, \quad \forall p \in P, \forall s \in S, w = 0 \quad (42)$$

$$\sum_{p \in P_f} g_{fpsw} = L_{fs} * \frac{|P|}{24 * 7}, \quad \forall f \in F, \forall s \in S, w = 0 \quad (43)$$

$$\sum_{f' \in \delta(f)} \sum_{p \in [P_f \cap P_{f'}]} g_{f'psw} \leq m_{fs}, \quad \forall f \in F, \forall s \in S, w = 0 \quad (44)$$

$$m_{fs} \leq \sum_{p \in P_f} \sum_{c \in C} \sum_{\eta \in \{1, \dots, M_c\}} \sum_{n \in N_c} q_{s\eta ncp((w) \bmod (|W_c|))}, \quad \forall f \in F, \forall s \in S, \\ \forall w \in \{0, 1, \dots, W^{lcm} - 1\} \quad (45)$$

$$q_{s\eta ncpw} \geq 0, \quad \forall c \in C, \forall s \in S, \forall p \in P, \forall w \in W_c, \forall n \in N_c, \forall \eta \in \{1, 2, \dots, M_c\} \quad (46)$$

$$\tau_{\eta ncs} \in \{0, 1\}, \quad \forall c \in C, \forall s \in S, \forall n \in N_c, \forall \eta \in \{1, 2, \dots, M_c\} \quad (47)$$

$$g_{fpsw} \geq 0, \quad \forall f \in F, \forall p \in P, \forall s \in S, w = 0 \quad (48)$$

$$m_{fs} \geq 0, \quad \forall f \in F, \forall s \in S \quad (49)$$

Since the workforce scheduling model assumed that all teams are equal in one cycle, the training requirements are also equal for all teams in the cycle. It is however not always necessary to train all teams for the same skill. The intermediate model therefore looks to each worker in each team individually and only assigns the necessary training to a worker. With the objective function (37), the model tries to find the minimal training costs in order to execute the rosters found by the workforce scheduling model.

Constraint (38) distributes the available capacity on each time period in each week over the different workers and their skills. Hence, it determines the capacity of skill type s made available by worker η in team n in cycle c on time period p in week w . The capacity made available by team n in cycle c on time period p is therefore multiplied by the sum of the parameter $X_{\eta ncs}$ and the decision variable $\tau_{\eta ncs}$. When worker η in team n in cycle c has skill s , or when this worker is trained for skill s , this sum equals 1. Because the capacity is also distributed over the different skills, constraint (39) makes sure that the capacity over all skills is not larger than what is available. Recall that Q_{ncp} is a combination of the results of the workforce scheduling model and the reconstructed weekly schedule.

Constraint (40) then makes sure that a worker cannot be trained for more than $\Theta_{\eta nc}^{max}$ skills and constraint (41) ensures that a worker can only be trained for skill s when he lacks this skill.

Constraints (42) and (43) are similar to constraints (3) and (4) and make sure that all flights can be maintained in time with the available skills and workers. Since all teams in a cycle can be different in this model, the capacity that is available of a certain skill type on a certain time period can be different for each week. Therefore, the coverage constraints (42) and (43) should be satisfied for all weeks $w \in \{0, 1, \dots, W^{lcm} - 1\}$. Since every cycle can have a different number of teams (weeks), the total available capacity (over all cycles) follows a cyclic pattern that is repeated after W^{lcm} weeks. To avoid the large number of constraints and variables to check the coverage for all weeks $w \in \{0, 1, \dots, W^{lcm} - 1\}$, constraints (44) and (45) are used. Constraint (44) makes sure that the assigned capacity to maintain flights is not larger than the minimum capacity over all weeks that is available to maintain each flight. Constraint (45) calculates this minimum.

Suppose that a flight arrives at the beginning of time period 1 and leaves at the end of

time period 3. Also assume that in the first week, one team with one worker with skill A is scheduled to work a shift during the first period and one team with one worker with skill B is scheduled during the next two time periods. Next week, the first team will work the shift of the second team and the second team will work the shift of the first team. During the first week, the available capacity to maintain the flight is only one time period of skill type A and two time periods of skill type B. During the second week, there is only one time period of skill type B and two time periods of skill type A. This means that the minimum available capacity over all weeks of each skill type is only one time period. Hence, in a workforce schedule that is feasible for each week, the assigned work of each skill type should be no more than one time period. Note that constraint (44) checks the coverage constraint for all f time intervals, while constraint (42) checks the coverage for all p time periods. Using time intervals instead of time periods is necessary because we want to compare the assigned work g_{fpsw} during a flight interval with the minimum capacity m_{fs} that is available in this flight interval.

The model presented above results in a certain number of training goals equal to the number of workers that are required to maintain all flights in time for the next season. A training goal is defined as a collection of skills. Assume we would need two workers with four skills and that both workers already have skills A, B and C. Suppose that, according to the model, worker 1 should acquire skill D and worker 2 should acquire skill E. This results in two training goals where the first goal consists of skills A, B, C, and D, and the second goal of skills A, B, C, and E.

The next and final model uses these training goals as its input and determines the training schedule in order to obtain these goals. We refer to this model as the training model.

4.3 Finding the best training schedule: The training model

While the first two models (the workforce scheduling model and the intermediate model) are concerned with constructing and optimizing the workforce schedule of the **next** season, the training model is only concerned with constructing a feasible training schedule during the **current** season (see Figure 2). Hence, this model assumes a fixed workforce schedule which defines the capacity of each skill, for each worker, in each team, in each cycle, on each time period. This information is contained in the input parameters Q_{ncp} and X_{mcs} . Of course, the number of weeks and teams in each cycle ($|W_c| = |N_c|$) and the team size in each cycle (M_c) is also known in advance.

Furthermore, the training model assumes a set of training goals (Γ). The number of training goals contained in the set Γ equals the number of workers that are required to maintain all flights in time during the next season. This set is composed based on the result of the intermediate model (see Section 4.2). The goal of the training model

is to find a training schedule that satisfies three different criteria. First, the training schedule should not interfere with the current operations. Hence, all flights should still be maintained in time. Second, all training goals in set Γ should be satisfied at the end of the current season. Third, the right worker should be trained for the right skills such that the training costs are minimized.

In the remainder of this section, we define the training model. We first list the sets, along with their associated indices:

$d \in \{1, \dots, 7\}$:	days in the week
$d \in D^{training}$:	training days
$p \in P$:	time periods in one week
$f, f' \in F$:	flights to be serviced
$f' \in \delta(f)$:	set of all flights that overlap at least during one period with flight f , with $f \in \delta(f)$
$p \in P_f$:	time periods during which flight f can be serviced
$s \in S$:	skill types
$c \in C$:	cycles in the workforce configuration
$n \in N_c$:	teams in cycle c , with $ N_c = W_c $ = the number of teams (=weeks) in cycle c
$w \in W_c$:	weeks in cycle c , with $ W_c = N_c $ = the number of weeks (=teams) in cycle c
$w \in W^{training}$:	weeks in the training horizon
$\eta \in \{1, 2, \dots, M_c\}$:	workers in a team in cycle c , with M_c the team size in cycle c
$\gamma \in \Gamma$:	training goals

The coefficients and right hand side constants are presented below:

L_{fs} :	workload (in time periods p) of flight f of skill type s
$\alpha_{dp} \in \{0, 1\}$:	= 1 if training is possible on period p on day d , = 0 otherwise
$\lambda_{\gamma\eta nc} \in \{0, 1\}$:	= 1 if worker η in team n in cycle c already has all the skills needed for goal γ , = 0 otherwise
$Q_{ncp} \geq 0$:	the capacity made available by team n in cycle c on time period p
$K_{\gamma\eta nc}^{training} > 0$:	training cost to train worker η in team n in cycle c for goal γ
$X_{\eta ncs} \in \{0, 1\}$:	= 1 if worker η in team n in cycle c has skill s , = 0 otherwise
$D_{\gamma\eta nc} \geq 0$:	duration (in days) to train worker η in team n in cycle c to reach goal γ
$\Delta_\gamma \geq 0$:	number of workers that should reach goal γ
M_c :	the team size in cycle c

The decision variables are:

- $\tau_{\gamma\eta ncdw} \in \{0, 1\}$: = 1 if worker η in team n in cycle c is trained for goal γ on day d in week w , = 0 otherwise
- $q_{s\eta ncpw} \geq 0$: capacity of skill type s made available by worker η in team n in cycle c on time period p in week w
- $g_{fpsw} \geq 0$: capacity of skill type s assigned to maintain flight f during time period p in week w
- $m_{fs} \geq 0$: minimum capacity (over all weeks $w \in W^{training}$ in the training horizon) of skill type s that is available between the STA and STD of flight f
- $z_{\gamma\eta nc}^{choice} \in \{0, 1\}$: = 1 if goal γ is chosen for worker η in team n in cycle c , = 0 otherwise

The training model can be formulated as follows:

$$\text{Minimize } \sum_{c \in C} \sum_{\gamma \in \Gamma} \sum_{n \in N_c} \sum_{\eta \in \{1, 2, \dots, M_c\}} z_{\gamma\eta nc}^{choice} \cdot K_{\gamma\eta nc}^{training} \quad (50)$$

Subject to:

$$q_{s\eta ncpw} \leq \frac{X_{\eta ncs}}{M_c} \cdot Q_{((n+w) \bmod (|W_c|))cp} \cdot \left(1 - \sum_{d \in D^{training}} \sum_{\gamma \in \Gamma} [\tau_{\gamma\eta ncdw} \cdot \alpha_{dp}] \right), \quad \forall c \in C, \forall s \in S, \forall p \in P,$$

$$\forall w \in W^{training}, \forall n \in N_c, \forall \eta \in \{1, 2, \dots, M_c\} \quad (51)$$

$$\sum_{s \in S} q_{s\eta ncpw} \leq \frac{Q_{((n+w) \bmod (|W_c|))cp}}{M_c}, \quad \forall c \in C, \forall p \in P, \forall n \in N_c,$$

$$\forall w \in W^{training}, \forall \eta \in \{1, 2, \dots, M_c\} \quad (52)$$

$$\sum_{f \in F} g_{fpsw} \leq \sum_{c \in C} \sum_{n \in N_c} \sum_{\eta \in \{1, 2, \dots, M_c\}} q_{s\eta ncpw}, \quad \forall p \in P, \forall s \in S, w = 0 \quad (53)$$

$$\sum_{p \in P_f} g_{fpsw} = L_{fs} * \frac{|P|}{24 * 7}, \quad \forall f \in F, \forall s \in S, w = 0 \quad (54)$$

$$\sum_{f' \in \delta(f)} \sum_{p \in [P_f \cap P_{f'}]} g_{f'psw} \leq m_{fs}, \quad \forall f \in F, \forall s \in S, w = 0 \quad (55)$$

$$m_{fs} \leq \sum_{p \in P_f} \sum_{c \in C} \sum_{\eta \in \{1, 2, \dots, M_c\}} \sum_{n \in N_c} q_{s\eta ncpw}, \quad \forall f \in F, \forall s \in S, \forall w \in W^{training} \quad (56)$$

$$\begin{aligned}
\sum_{d \in D^{training}} \sum_{w \in W^{training}} \tau_{\gamma\eta n c d w} &= D_{\gamma\eta n c} \cdot z_{\gamma\eta n c}^{choice} \cdot (1 - \lambda_{\gamma\eta n c}), \quad \forall c \in C, \\
&\forall \gamma \in \Gamma, \\
&\forall \eta \in \{1, 2, \dots, M_c\}, \\
&\forall n \in N_c \tag{57}
\end{aligned}$$

$$\sum_{\gamma \in \Gamma} z_{\gamma\eta n c}^{choice} \leq 1, \quad \forall c \in C, \forall \eta \in \{1, 2, \dots, M_c\}, \forall n \in N_c \tag{58}$$

$$\sum_{c \in C} \sum_{n \in N_c} \sum_{\eta \in \{1, 2, \dots, M_c\}} z_{\gamma\eta n c}^{choice} = \Delta_{\gamma}, \quad \forall \gamma \in \Gamma \tag{59}$$

$$\tau_{\gamma\eta n c d w} \in \{0, 1\}, \quad \forall c \in C, \forall \gamma \in \Gamma, \forall \eta \in \{1, 2, \dots, M_c\}, \forall n \in N_c, \forall d \in \{1, \dots, 7\}, \\
\forall w \in W^{training} \tag{60}$$

$$q_{s\eta n c p} \geq 0, \quad \forall c \in C, \forall s \in S, \forall \eta \in \{1, 2, \dots, M_c\}, \forall n \in N_c, \forall p \in P, \\
\forall w \in W^{training} \tag{61}$$

$$g_{f p s w} \geq 0, \quad \forall f \in F, \forall p \in P, \forall s \in S, w = 0 \tag{62}$$

$$m_{f s} \geq 0, \quad \forall f \in F, \forall s \in S \tag{63}$$

$$z_{\gamma\eta n c}^{choice} \in \{0, 1\}, \quad \forall c \in C, \forall \gamma \in \Gamma, \forall \eta \in \{1, 2, \dots, M_c\}, \forall n \in N_c \tag{64}$$

The objective function (50) tries to decrease the total training costs. While the training goals are predefined by the workforce scheduling model and the intermediate model, the decision to train who, when and how is still to be made. The goal is to finish the required training during the weeks in the training horizon defined by the set $W^{training}$ with minimal costs. Recall that the training horizon is the time between the construction of the new schedule and the start of the next season (see Figure 1).

In this model, constraints (51) to (56) replace constraints (38) to (39) and constraints (42) to (45) in the intermediate model. They make sure that flights can be maintained in time with the available skills and workers. Constraint (51) is however very different from constraint (38) in the intermediate model. The capacity that is available of skill type s on time period p of team n in cycle c is multiplied by a factor that is zero or one. This factor is 0 when training takes place during that time period and is 1 otherwise. This way, the capacity decreases because a certain worker is trained at that moment. To model training on the job, the parameter α_{dp} can be fixed to a value between 0 and 1. This way, the model can account for a certain decrease in capacity due to the training on the job. Because we want to decide on the training for each week in the training horizon, we have to build constraints for all weeks $w \in W^{training}$ instead of just the weeks $w \in W_c$. This is also true for constraint (52) and (56). Constraints (53) to (55) are exactly the same as constraints (42) to (44).

Constraints (57) to (59) are specific constraints to solve the training problem. The

decision variable $z_{\gamma\eta nc}^{choice}$ can be found in all three training constraints. $z_{\gamma\eta nc}^{choice}$ is 1 when we decide that worker η in team n in cycle c will be used for goal γ and is 0 otherwise. When $z_{\gamma\eta nc}^{choice}$ is 1, it means that this worker already has all the skills for goal γ or that he will be trained to reach that goal. When a worker does not have all the skills for goal γ , and when $z_{\gamma\eta nc}^{choice}$ is 1 for this worker, he must be fully trained. This is ensured by constraint (57). When a worker should be trained for too many skills in order to reach a certain goal, $z_{\gamma\eta nc}^{choice}$ is fixed to 0. This way we can limit the number of skills a worker can be trained for.

Constraint (58) makes sure that one can only choose one goal for each worker. When the left hand side of constraint (58) is 0, the worker is not needed anymore. Of course, this worker is not just fired, but he is relocated to another division of the company (usually to the heavy maintenance division). Constraint (59) makes sure makes that the number of workers that are chosen for goal γ is equal to the required number of workers that should reach this goal.

The model presented above returns a feasible training schedule to reach all goals with minimal training costs. Recall that these costs are less than or equal to the training costs resulting from the workforce scheduling model because of the intermediate model. According to the current formulation of the training model, the model can freely decide when to schedule the training as long as there is enough capacity left to maintain all flights in time. It has however not a single incentive to finish the training as soon as possible. When the total training duration is part of the decision problem, constraint (65) can be used to define the earliest finishing time of all training. The decision variable $\phi \geq 0$ is thus defined as the earliest finishing time (in weeks) of all training required to reach all goals. The finishing time could, for instance, be minimized in the objective function.

$$\tau_{\gamma\eta mcdw} \cdot (w + 1) \leq \phi, \quad \forall c \in C, \forall \gamma \in \Gamma, \forall \eta \in \{1, 2, \dots, M_c\}, \forall n \in N_c, \quad (65)$$

$$\forall d \in \{1, \dots, 7\}, \forall w \in W^{training}$$

5 Results and discussion

5.1 Test instances

The performance of the three proposed models is tested on 20 test instances. These instances are created randomly based on real-life data dimensions from Sabena Techniques. Table 2 shows the results of the workforce scheduling model without training, the workforce scheduling model with training, the intermediate model and the training model. We solved the models with IBM optimization software CPLEX 12.4. All models were given 1500 seconds of computation time on an AMD FX 8120 eight-core processor

with a clock speed of 3.10GHz and 8GB RAM.

The first column of Table 2 shows the name of the test set. We refer to Beliën et al. (2013) for a detailed description of the test instances. The next three columns show the results of the scheduling model without training. In this model, Θ_{ω}^{max} is set to 0 and the model only tries to cover the demand with the available workers and their current skills. The weekly costs are the weekly incurred labor costs to operate the workforce schedule. Next to the costs, Table 2 also lists the number of workers that are required in the schedule as well as the optimality gap. This gap is the relative difference between the obtained solution (after 1500 seconds) and the lower bound obtained by the optimization software.

The next four columns show the results of the scheduling model with training. We assumed that each worker can be trained for at most two skills. Hence, Θ_{ω}^{max} is set to 2. Again, Table 2 shows the weekly costs incurred by the company, but this time, these costs are the sum of the weekly labor costs and the total training costs divided by 24. Recall that we assumed a planning horizon of one season with 24 weeks. The total training costs are shown in the next column, followed by the required number of workers and the optimality gap.

Table 2 also shows the savings resulting from the training. We make a difference between the pure monetary savings and the workforce reduction. The monetary savings are the differences between the weekly costs of the scheduling model without training and the weekly costs of the scheduling model with training, multiplied by 24. The reduction of the workforce can be expressed as the difference between the required number of workers in both models.

The last three columns of Table 2 present the results of the intermediate model and the training model. While the workforce scheduling models did not return an optimal solution after 1500 seconds of computation time, the intermediate model and training model did solve the problem to optimality in the allowed computation time. Therefore, the CPU times are shown instead of the optimality gaps for these two models.

The average values of each column are shown at the bottom of Table 2.

Table 2: Results of the three different skills and training models

Test Set	Scheduling model without training			Scheduling model with training				Savings		Intermediate model		Training model
	Weekly costs	Workers	Gap (%)	Weekly costs	Training costs	Workers	Gap (%)	Money	Workers	Training costs	Time (s)	Time (s)
1.1.1.1	17450	24	23.71	15154	5400	20	11.56	55105	4	4900	1.28	4.37
1.1.1.2	20869	30	15.55	19517	8000	28	9.57	32440	2	6400	3.99	5.14
1.1.1.3	17160	24	21.84	16584	3000	22	19.22	13819	2	3000	3.62	2.96
1.1.1.4	17706	26	21.49	17135	4700	24	13.15	13706	2	3900	2.54	3.48
1.1.1.5	20804	30	21.41	17569	6000	24	6.92	77656	6	6000	1.07	2.14
1.1.2.1	15651	24	27.12	13661	2000	20	16.50	47766	4	2000	1.09	3.44
1.1.2.2	19547	30	45.24	14546	10000	20	26.62	120016	10	9500	2.37	4.07
1.1.2.3	20592	30	10.74	20296	3000	30	9.44	7110	0	3000	2.03	3.50
1.1.2.4	18393	26	19.76	16483	4000	24	10.71	45844	2	4000	2.89	3.26
1.1.2.5	18267	28	22.73	16432	4000	24	14.08	44038	4	4000	0.87	3.92
1.2.1.1	13707	20	11.58	13600	0	20	10.97	2556	0	0	0.00	0.00
1.2.1.2	18759	28	21.20	16424	3000	24	9.95	56032	4	3000	0.73	2.40
1.2.1.3	11645	16	10.98	11296	1000	16	8.42	8377	0	500	0.61	2.79
1.2.1.4	13107	18	22.09	12063	3700	16	15.31	25069	2	3700	0.58	3.15
1.2.1.5	16274	22	20.91	14861	7600	20	13.39	33904	2	6800	0.69	2.36
1.2.2.1	10854	16	17.63	10837	0	16	17.50	408	0	0	0.00	0.00
1.2.2.2	11387	16	6.21	11386	0	16	6.22	26	0	0	0.71	0.00
1.2.2.3	13043	20	3.20	13004	0	20	2.88	945	0	0	0.00	0.00
1.2.2.4	10966	16	16.84	10929	0	16	16.56	888	0	0	0.00	0.00
1.2.2.5	10797	16	11.90	10797	0	16	11.87	0	0	0	1.03	0.00
Average	15849	23	18.61	14629	3270	21	12.54	29285	2	3035	1.31	2.35

5.2 The workforce scheduling model: Discussion

5.2.1 Assumptions

To test the performance of the workforce scheduling model, we designed a pool of 30 workers with three arbitrarily chosen skills. We assume that there are five different skill types and that each aircraft requires exactly one skill type. The skill required to maintain each aircraft is arbitrarily chosen such that 2% of all flights require skill A, 2% require skill B, 4% require skill C, 10% require skill D and 82% require skill E. We further assume that a week is divided in 672 periods of 15 minutes and that we only consider two cycles with at most eight teams in each cycle. The minimum team size is fixed to two workers.

As we fix Θ_{ω}^{max} to 2 in the workforce scheduling model with training, the number of skills a worker can be trained for is limited to 2. We also assume that there are two training days in each week. On these days, the training takes place between 10AM and 2PM. In order to acquire skill A, five training days are required, skills B, C and D take eight days of training and skill E takes ten days. The costs to acquire each skill type are 2000, 800, 1200, 500 and 800 Euro respectively. All assumptions are based on real-life observations.

5.2.2 Making the optimal trade-off

The workforce scheduling model is solved with CPLEX 12.4 using the enumeration approach outlined in Beliën et al. (2013) to incorporate the decision about the optimal team size. During the enumeration MILP approach, each MILP was provided with a computation time limit of 60 seconds, which is extended by another 60 seconds every time the node is revisited. To study the impact of training on the costs of the workforce schedule for the next season, we solve the model with and without training. The results are shown in Table 2.

Recall that the workforce scheduling model with training makes a trade-off between cheaper rosters that require higher skilled workers and the training costs to obtain this higher skilled workforce. This way, an optimal amount of training will be determined in order to minimize the total costs for the next season. Workers will only be trained as long as the training costs can be compensated by a decrease in the labor costs. Therefore, it is possible that the workforce scheduling model will decide that no training should take place because the potential training costs cannot be compensated by a decrease in labor costs. This is for example the case for test set 1.2.1.1 and sets 1.2.2.1 to 1.2.2.5 in Table 2. For these six test sets, no savings can be obtained by training some of the workers. Therefore, the weekly costs and the required number of workers are equal in the

model without training and the model with training. However, some small differences can be observed in Table 2 because of the slightly lower optimality gaps in the model with training.

5.2.3 The benefits of training

While training is not always beneficial, it results in substantial savings for most instances in Table 2. As shown in columns 9 and 10 in Table 2, these savings are twofold. First, the company can save money with cheaper personnel schedules. Second, some of the workers can become redundant and can be used somewhere else in the company such as in the heavy maintenance department. On average, 29285 Euro can be saved over the whole season with training. Moreover, applying the workforce scheduling model with training always results in lower (or equal) costs compared to the model without training.

While there is a strong linear relation between the monetary savings and the reduction of workers ($R^2 = 0.95$), training can also decrease costs without a reduction of workers. This is true for test sets 1.1_2.3 and 1.2_1.3 in Table 2. In these two cases, the higher skilled workforce makes it possible to reorganize the current workforce schedule and decrease the labor costs. Furthermore, a higher skilled workforce cannot only be beneficial for the next season, but it will also increase the possibility of building better and cheaper workforce schedules in the future.

5.2.4 Performance of the workforce scheduling model

While the workforce scheduling model is the most important model to minimize the costs, it is also the most time consuming of all three models. Therefore, Table 2 shows the optimality gaps instead of the required computation time. An interesting observation is made when the gaps of the scheduling model without training and the scheduling model with training are compared.

According to Table 2, the gaps of the first model are much larger than those of the second model (18.61% versus 12.54%). Furthermore, the gaps of the second model are much more similar to those of the respective MILP model presented by Beliën et al. (2013) (12.54% versus 11.01%). This means that finding a good solution without training appears to be much more difficult than finding a good solution with training. However, this statement is not true when only little training (or even no training at all) is required to find a good solution. For these cases, the gaps (and solution values) of the first model are much closer to the gaps (and solution values) of the second model. There appears to be a relatively strong linear relation between the gaps of the model without training (column 4) and the monetary savings (column 9) resulting from training ($R^2 = 0.70$). This relation is absent for the model with training ($R^2 = 0.16$). Hence, this

relation suggests that the large gaps belonging to the model without training must be seen as an indication of the possible savings that can be made when the workforce would be trained instead of an indication of poor solutions.

5.3 The intermediate model: Results and performance

Since the workforce scheduling model assumes that all teams are equal in one cycle, the training requirements are equal for all teams. It is however not always necessary to train all teams for the same skill. Therefore, the goal of the intermediate model is to find the minimal training requirements in order to perform all maintenance in time. The results are shown in columns 11 and 12 of Table 2.

Because the intermediate model minimizes the training requirements, the training costs are also minimized. The comparison of columns 6 and 11 in Table 2 reveals that the intermediate model always results in less or equal training costs. On average, the difference between the training costs found by the workforce scheduling model and those found by the intermediate model is less than 6%. However, the maximum reduction of training costs is 50%. Hence, the application of the intermediate model can be beneficial in some cases. Furthermore, solving the model always returns an optimal solution in only 1.31 seconds (on average). This makes the intermediate model an easy and possibly very beneficial model to use.

As the intermediate model results in the optimal skill requirements, it can be used (with or without the workforce scheduling model) to assess the impact of an extra maintenance contract. Tests showed that the addition of just one new flight with a new skill type already requires a lot of new training. Even if the shift schedules can remain unchanged, many workers should be trained to acquire this new skill type because of the cyclic work pattern. In a cyclic workforce schedule it is very likely that all teams in a cycle will be assigned to the same flights at least once. Therefore, when a new flight with a new skill is introduced, many workers should acquire this new skill. Hence, one must be very cautious to accept a new maintenance contract that involves training the workforce for a new skill.

5.4 Performance of the training model

The training model is very important to decide when and who will be trained in order to reach all training goals without disrupting the maintenance work. Table 2 shows the required computation time to solve the training model for each of the test instances to optimality. The objective value (the training costs) are not shown in the table because they are equal to the training costs found by the intermediate model.

Recall from Section 4.3 that as the length of the training horizon ($W^{training}$) increases, the number of decision variables in the training model also increases. Therefore, the length of the training horizon has an impact on the performance of the training model. We used a training horizon of eight weeks to solve the test sets in Table 2. While the average computation time of the training model is larger than for the intermediate model (2.35s versus 1.31s), the time required to solve the problem to optimality is still relatively small.

6 Conclusion and future research

In this paper, we present a two-stage mixed integer programming approach for optimizing the skill mix and training schedule at an aircraft maintenance company. In the first stage, a trade-off is made between cheaper rosters that require higher skilled workers and the training costs to obtain this higher skilled workforce. Finding this optimal trade-off is the goal of the workforce scheduling model.

In a second stage, the training model is used to find an optimal training schedule. This model determines the exact timing of the training for each worker and takes into account that a worker is unavailable to work during his training periods. The goal is to finish the required training during the weeks in the training horizon with minimal costs.

We successfully applied our approach to 20 different test sets based on real-life data from Sabena Technics. Experiments first demonstrate that our models succeed in finding good solutions in reasonable computation times. Second, we illustrate the benefits of training by comparing a scenario without training to a scenario with training. The results show how our approach can make a good trade-off between cheaper rosters that require higher skilled workers and the training costs to obtain this higher skilled workforce.

For future research regarding the proposed skills and training model, one can focus on the constraint that defines the earliest finishing time of all training required to reach all training goals. When the required training cannot be finished in the predefined training horizon, training costs will increase or the training model will not find a feasible result at all. A special approach can be proposed that confines the training possibilities in order to obtain a feasible training schedule. Furthermore, it can be interesting to investigate the implications of a continuous training program. Instead of a short training program dedicated to the next season, one can choose to train the workforce on a continuous basis with a long term vision.

References

- Aickelin, U., Burke, E. K. & Jingpeng, L. (2009, april). An evolutionary squeaky wheel optimization approach to personnel scheduling. *Evolutionary Computation, IEEE Transactions on*, 13(2), 433 - 443.
- Aickelin, U. & Dowsland, K. A. (2004). An indirect genetic algorithm for a nurse scheduling problem. *Computers & Operations Research*, 31, 761-778.
- Aickelin, U. & White, P. (2004). Building better nurse scheduling algorithms. *Annals of Operations Research*, 128(1-4), 159-177.
- Alfares, H. K. (2004). Survey, Categorization, and Comparison of Recent Tour Scheduling Literature. *Annals of Operations Research*, 127, 145–175.
- Avramidis, A. N., Chan, W., Gendreau, M., L'Ecuyer, P. & Pisacane, O. (2010). Optimizing daily agent scheduling in a multiskill call center. *European Journal of Operational Research*, 200(3), 822 - 832. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0377221709000332>
- Bard, J. F., Binici, C. & deSilva, A. H. (2003). Staff scheduling at the United States postal service. *Computers & Operations Research*, 30(5), 745 - 771. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0305054802000485>
- Bard, J. F. & Purnomo, H. W. (2005a). A column generation-based approach to solve the preference scheduling problem for nurses with downgrading. *Socio-Economic Planning Sciences*, 39(3), 193 - 213. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0038012104000126>
- Bard, J. F. & Purnomo, H. W. (2005b). Preference scheduling for nurses using column generation. *European Journal of Operational Research*, 164(2), 510 - 534. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0377221704000657>
- Bard, J. F. & Wan, L. (2008, January). Workforce design with movement restrictions between workstation groups. *Manufacturing & Service Operations Management*, 10(1), 24–42. Retrieved from <http://dx.doi.org/10.1287/msom.1060.0148>
- Beddoe, G., Petrovic, S. & Li, J. (2009, April). A hybrid metaheuristic case-based reasoning system for nurse rostering. *Journal of Scheduling*, 12(2), 99–119. Retrieved from <http://dx.doi.org/10.1007/s10951-008-0082-8>
- Beddoe, G. R. & Petrovic, S. (2006). Selecting and weighting features using a genetic algorithm in a case-based reasoning approach to personnel rostering. *European Journal of Operational Research*, 175(2), 649 - 671. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0377221705004674>

- Beliën, J., Demeulemeester, E., De Bruecker, P., Van den Bergh, J. & Cardoen, B. (2013). Integrated staffing and scheduling for an aircraft line maintenance problem. *Computers & Operations Research*, 40(4), 1023–1033.
- Bergh, J. Van den, Beliën, J., De Bruecker, P., Demeulemeester, E. & De Boeck, L. (2013). Personnel scheduling: A literature review. *European Journal of Operational Research*, 226(3), 367 - 385. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0377221712008776>
- Boulay, C. du & Medway, C. (1999). The clinical skills resource: A review of current practice. *Medical Education*, 33, 185–191.
- Brucker, P., Burke, E. K., Curtois, T., Qu, R. & Vanden Berghe, G. (2010, August). A shift sequence based approach for nurse scheduling and a new benchmark dataset. *Journal of Heuristics*, 16(4), 559–573. Retrieved from <http://dx.doi.org/10.1007/s10732-008-9099-6>
- Brunner, J. O., Bard, J. F. & Kolisch, R. (2011). Midterm scheduling of physicians with flexible shifts using branch and price. *IIE Transactions*, 43(2), 84–109.
- Brunner, J. O. & Edenharter, G. (2011, June). Long term staff scheduling of physicians with different experience levels in hospitals using column generation. *Health Care Management Science*, 14(2), 189-202. Retrieved from <http://ideas.repec.org/a/kap/hcarem/v14y2011i2p189-202.html>
- Burke, E. K., De Causmaecker, P., Vanden Berghe, G. & Van Landeghem, H. (2004). The state of the art of nurse rostering. *Journal of Scheduling*, 7(6), 441–499.
- Cai, X. & Li, K. N. (2000). A genetic algorithm for scheduling staff of mixed skills under multi-criteria. *European Journal of Operational Research*, 125(2), 359 - 369. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0377221799003914>
- Dantzig, G. B. (1954). Letter to the Editor - A Comment on Edie's Traffic Delays at Toll Booths. *Operations Research*, 2(3), 339-341. Retrieved from <http://or.journal.informs.org/content/2/3/339.abstract>
- De Bruecker, P., Van den Bergh, J., Beliën, J. & Demeulemeester, E. (2015). Workforce planning incorporating skills: State of the art. *European Journal of Operational Research*, 243(1), 1 - 16. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0377221714008601>
- Detienne, B., Péridy, L., Pinson, E. & Rivreau, D. (2009). Cut generation for an employee timetabling problem. *European Journal of Operational Research*, 197(3), 1178 - 1184. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0377221708003093>
- Ernst, A. T., Jiang, H., Krishnamoorthy, M. & Sier, D. (2004). Staff scheduling and

- rostering: A review of applications, methods and models. *European Journal of Operational Research*, 153, 3–27.
- Fowler, J. W., Wirojanagud, P. & Gel, E. S. (2008). Heuristics for workforce planning with worker differences. *European Journal of Operational Research*, 190(3), 724 - 740. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0377221707006017>
- Huang, H.-C., Lee, L.-H., Song, H. & Eck, B. T. (2009). SimMan - A simulation model for workforce capacity planning. *Computers & Operations Research*, 36(8), 2490 - 2497. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0305054808001974>
- Kohl, N. & Karisch, S. E. (2004). Airline crew rostering: Problem types, modeling, and optimization. *Annals of Operations Research*, 127(1-4), 223-257.
- Li, N. & Li, L. X. (2000). Modeling staffing flexibility: A case of China. *European Journal of Operational Research*, 124(2), 255 - 266. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0377221799003793>
- Marentette, K. A., Johnson, A. W. & Mills, L. (2009). A measure of cross-training benefit versus job skill specialization. *Computers & Industrial Engineering*, 57(3), 937 - 940. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0360835209000977>
- Song, H. & Huang, H.-C. (2008). A successive convex approximation method for multi-stage workforce capacity planning problem with turnover. *European Journal of Operational Research*, 188(1), 29 - 48. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0377221707004122>
- Tiwari, V., Patterson, J. H. & Mabert, V. A. (2009). Scheduling projects with heterogeneous resources to meet time and quality objectives. *European Journal of Operational Research*, 193(3), 780 - 790. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0377221707010831>
- Zülch, G., Rottinger, S. & Vollstedt, T. (2004). A simulation approach for planning and re-assigning of personnel in manufacturing. *International Journal of Production Economics*, 90(2), 265 - 277. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0925527303003414>

FACULTY OF ECONOMICS AND BUSINESS
Naamsestraat 69 bus 3500
3000 LEUVEN, BELGIË
tel. + 32 16 32 66 12
fax + 32 16 32 67 91
info@econ.kuleuven.be
www.econ.kuleuven.be

