European Journal of Operational Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/eor

Discrete Optimization

Multiskilled workforce staffing and scheduling: A logic-based Benders' decomposition approach

Araz Nasirian [a,*], Lele Zhang [b,c], Alysson M. Costa [b,c], Babak Abbasi [a]

[a] Department of Information Systems and Business Analytics, RMIT University, Melbourne, VIC 3000, Australia
[b] School of Mathematics and Statistics, The University of Melbourne, Parkville, VIC 3010, Australia
[c] ARC Industrial Transformation Training Centre in Optimisation Technologies, Integrated Methodologies, and Applications, OPTIMA, Australia

ARTICLE INFO

Keywords:
Staffing and scheduling
Multiskilled workforce
Logic-based Benders' decomposition
Two-stage stochastic programming

ABSTRACT

We study the staffing and scheduling problem of a multiskilled workforce with uncertain demand. We formulate the problem as a two-stage stochastic integer program. The first stage considers strategic decisions, including recruiting permanent staff from an available pool and training them with additional skills, and the second stage focuses on operational decisions, involving the allocation of the multiskilled workforce and the hiring of temporary staff to accommodate uncertain demand. To effectively solve problems of practical sizes, we develop a novel solution algorithm based on the logic-based Benders' decomposition (LBBD) approach, incorporating a customized analytical cut. We validate our approach through a case study using the data from a prefabrication company, demonstrating the significant cost savings achieved through workforce multiskilling. Our experimental results show that the proposed method is substantially more efficient than the latest Gurobi solver, up to 133 times faster and on average 29 times faster than directly solving the monolithic deterministic equivalent problem (MDEP).

1. Introduction

In many industries, workforce expenses constitute a significant portion of operational costs. Consequently, effective decision-making regarding the staffing and scheduling of a workforce has been repeatedly emphasized in the Operations Research literature (Kim & Mehrotra, 2015). Researchers have developed various methods to deal with the counterproductive effects of demand uncertainty on workforce costs. These methods include recruiting casual staff (Bard & Purnomo, 2005) and paying existing staff overtime (Campbell, 2017). A more effective approach to addressing demand uncertainty, especially in manufacturing and semi-manufacturing environments like prefabricated construction is multiskilling workforce (Daniels et al., 2004; Nasirian et al., 2022; Wang & Tang, 2020).

In this paper, we define a skill as a worker's ability to perform a specific task; for example, workers skilled in tiling can install floors in bathrooms and kitchens. This study considers that workers can master multiple skills and acquire extra skills through training after being recruited. In this research, *workforce staffing* involves both recruitment and training, determining how many workers with specific skill sets should be recruited and which workers should receive training in additional skills. *Workforce scheduling*, on the other hand, accounts for the allocation of the workforce to tasks based on their skills.

According to Kim and Mehrotra (2015), integrating workforce staffing and scheduling is beneficial because the ultimate objective of workforce staffing is to schedule the workforce to meet the task/skill demand. Hence, we develop a two-stage stochastic mixed integer program to integrate staffing and scheduling of a multiskilled workforce, aiming to minimize overall workforce costs. The first stage of our model addresses the staffing strategy and develops an initial solution based on imperfect information about demand. The second stage incorporates the stochastic nature of future demand, adjusting scheduling decisions to allocate the multiskilled workforce and recruit casual workers as demand is revealed. The two-stage model is logical for integrating staffing and scheduling decisions, given their different decision-making time horizons.

1.1. Contributions

One novelty of this work is the examination of staffing and scheduling a multiskilled workforce in a semi-manufacturing environment, specifically prefabricated construction. Our literature review reveals that most studies on multiskilled workforce staffing and scheduling are conducted in service contexts such as hospitals (Campbell, 2017),

* Corresponding author.
E-mail address: araz.nasirian@rmit.edu.au (A. Nasirian).

https://doi.org/10.1016/j.ejor.2024.11.033
Received 28 February 2023; Accepted 19 November 2024
Available online 29 November 2024
0377-2217/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

banks (Fragniere et al., 2010), and call centers (Wallace & Whitt, 2005). Reviews by De Bruecker et al. (2015) and Qin et al. (2015) highlight that relatively few investigations have addressed this problem in manufacturing contexts. Significant differences between the service and the manufacturing industries impact the technical aspects of mathematical model development for each sector.

The duration of providing a service often is notably shorter than the time required to manufacture a product. For example, car manufacturing may span several days, weeks or even months, whereas a route hospital checkup typically takes 15 to 30 min, and major surgical procedures extend over several hours. Nevertheless, exceptions exist, such as cancer patients who may require weeks of hospital treatment. Considering waiting times, service-related waits are generally shorter than those for products. For instance, waiting times at a bank can range from a few minutes to a few hours, while elective surgery patients may wait months or even over a year for their procedures.

Additionally, developing a multiskilled workforce in manufacturing requires more monetary and time investment compared to some service sectors, like call centers. Factors such as long planning horizons, skill affinity which encourages workforce multiskilling, and industry-specific regulations make the prefabrication sector deserving of context-specific investigations (Nasirian et al., 2019).

Another novelty of our research lies in the development of a structure-specific algorithm to solve the proposed two-stage stochastic program for the multiskilled workforce staffing and scheduling problem. A significant amount of research in this area has been conducted in a deterministic setting (De Bruecker et al., 2015). De Bruecker et al. (2015) attributed the reason for this matter to the challenge associated with solving the stochastic version of the problem.

The (two-stage) stochastic staffing and scheduling of a multiskilled workforce is a complex problem for several reasons. First, the scheduling problem is large-scale, as task combination and allocation generate numerous integer and/or binary variables. Second, converting the recourse function of a stochastic problem to its deterministic equivalent generates a large number of scenarios and consequently variables. These complexities have led earlier studies on stochastic workforce scheduling to use heuristic methods; for example, see Easton and Mansour (1999), Easton and Rossin (1996), and Bard et al. (2007). Third, a multiskilled workforce increases the problem scale by introducing more workforce alternatives. Daniels et al. (2004) showed that the deterministic multiskilled workforce scheduling problem in a flow-shop production environment is intractable for problems exceeding five tasks and three skills or four tasks and four skills. Fourth, our problem has a mixed binary integer recourse, making conventional solution methods like classic Benders' decomposition or L-shaped algorithms inapplicable. To deal with the challenge, we develop a logic-based Benders' decomposition (LBBD) approach customized to the structure of the problem. While general methods exist to modify the classic Benders' decomposition algorithm for problems with integer recourse function (Sen & Sherali, 2006), structure-specific algorithms such as the LBBD approach exhibit superior performance compared to general Benders' decomposition algorithms (Elçi & Hooker, 2022).

1.2. Organization

The rest of this paper is organized as follows. In Section 2, we present a literature review on workforce staffing and scheduling, techniques for enhancing the convergence of the Benders' decomposition algorithm, and the application of LBBD to scheduling problems. Section 3 provides the problem statement and the mathematical model for the integrated multiskilled staffing and scheduling problem. In Section 4, we investigate the properties and structure of the proposed model and elaborate on the development of the LBBD approach tailored to the specific structure of the problem. Section 5 demonstrates the cost savings achieved through workforce multiskilling and compares the performance of our model with the solution of the monolithic deterministic equivalent problem (MDEP) by Gurobi. Finally, Section 6 summarizes the paper and presents a brief rationale for future research directions.

2. Literature review

This section is divided into four parts. First, we explore the literature on skill-based workforce staffing and scheduling, focusing on their model assumptions regarding skill class and source. This is important for ensuring that the mathematical model accurately reflects real-life problem features. Second, we examine studies on skill-based workforce staffing and scheduling that incorporate uncertainty and their solution methods. Third, we provide a brief introduction to the Benders' decomposition technique and investigate strategies to enhance its convergence. Fourth, due to the limited research applying LBBD to skill-based staffing and scheduling problems, we review a broader range of scheduling problems that utilize LBBD.

2.1. Skill features

2.1.1. Skill class

De Bruecker et al. (2015) identified two skill classes, *hierarchical class* and *categorical class*, in the workforce planning and scheduling literature. The categorical skill class assumes that all workers with a given skill possess it at the same level. In other words, a worker's ability to perform a task can be represented by a binary value: 0 indicates they do not have the skill and cannot conduct the task, while 1 means they possess the skill and can conduct the task. If two workers have the same skill, their performance is indistinguishable when conducting tasks requiring that skill.

The hierarchical skill class, on the other hand, recognize differences in skill levels among workers. Then a worker's capability can be represented by a real number, typically between 0 and 1. While two workers with the same skill can both perform the associated task, the one with a lower skill level will not perform as well as the one with a higher skill level, in terms of work quality and/or speed. See the second column of Table 1.

2.1.2. Skill source

In the literature on multiskilled workforce staffing and scheduling problems, there are four primary sources to satisfy demand: recruiting permanent employees, hiring casual employees, training permanent employees, and allocating overtime to permanent employees. The third column of Table 1 categorizes the reviewed literature based on the skill sources.

Permanent workforce: This refers to a group of employees recruited by the employer on a continuous basis. Hiring and firing permanent employees is a common strategic solution to adapt the workforce level in response to changes in demand. However, this practice is often seen as unfavorable due to its high costs and adverse effects on workforce morale and motivation (De Bruecker et al., 2015). In certain industries or countries, such as the prefabrication sector in Australia, union and government regulations impose serious barriers to firing permanent workforce (Nasirian et al., 2019). According to Morris-Reade (2022), the average cost of hiring new workers in Australia is AUD$23,860 per person in 2022, and the recruitment process takes an average of 40 days.

Casual workforce: This refers to employees with no firm commitment to ongoing work from the employer. Past research found that employing casual staff increases flexibility in workforce planning. For example, Corominas et al. (2008) demonstrated increased workforce planning flexibility by employing casual staff to handle demand uncertainty associated with seasonality.

There are varying perspectives on the assumptions regarding hiring casual workers, including cost, skill levels, efficiency, and learning rates. Some researchers, such as Fragniere et al. (2010), assume that casual workers have lower skill levels than the permanent workforce and, therefore, should receive lower salaries. Another perspective, as seen by Carlos et al. (2011), is that casual staff lack efficiency compared to permanent employees and have a lower learning speed. Conversely, Heimerl and Kolisch (2010) suggest that the efficiency of casual

Table 1
Summary of workforce planning literature pertaining to skill class, skill source, context, uncertainty and methods.

Paper	Skill class	Skill source	Context	Uncertainty consideration	Solution method
Li and Li (2000)	Hl	P, T	Hospital	–	Goal programming
Stratman et al. (2004)	Hl	C	Manufacturing	Workforce cost	Simulation
Zülch et al. (2004)	Cl	C	Manufacturing	Workforce supply	Heuristic
Lee (2004)	Hl, Cl	T	General service	Customer arrival	Queuing
Wallace and Whitt (2005)	Cl	T	Call center	Demand	Simulation, Heuristic
Thompson and Goodale (2006)	Hl	P, T	Call center	Demand	Simulation, Heuristic
Corominas et al. (2008)	Hl	P	Manufacturing	–	Integer LP
Ertogral and Bamuqabel (2008)	Cl	T	Call center	Demand	Queuing, Integer LP
Song and Huang (2008)	Cl	C, T	General service	Workforce churn	Heuristic
Huang et al. (2009)	Cl	P, T	General service	Demand	Non-LP
Avramidis et al. (2010)	Cl	T	Call center	Demand	Heuristic, Simulation
Fragniere et al. (2010)	Hl	P, T	General	Demand	Mixed integer LP
Harper et al. (2010)	Hl	P, C	Hospital	Absenteeism	Simulation–optimization
Helber and Henken (2010)	Cl	T, C	Call center	Customer arrival	Simulation–optimization
Heimerl and Kolisch (2010)	Hl, Cl	P, T	Maintenance	–	Mixed integer LP
Carlos et al. (2011)	Hl	P, C	Maintenance	–	Heuristic
Sadjadi et al. (2011)	Cl	P	Different cases	Demand	Genetic Algorithm
Campbell (2011)	Hl	P, T	Hospital	Demand	Network programming, Heuristic
Campbell (2012)	Cl	P, C, O	Hospital	Demand	Network programming, Heuristic
Campbell (2017)	Cl	P, O	Hospital	Demand	Heuristic
Wang and Tang (2020)	Cl	P, T	Manufacturing	Demand	Heuristic
Naderi et al. (2023)	Hl	P, O	Home healthcare	Service and travel times	LBBD, cut propagation, SP-timed LBBD

Hl: Hierarchical; Cl: Categorical; C: casual; T: training; P: permanent; O: overtime; LP: linear programming.

workers can vary, being either more or less efficient than permanent employees.

Permanent workforce training: Training can be conducted through educational courses or on-the-job training. On-the-job training is generally less expensive, while educational courses offer higher credibility (De Bruecker et al., 2015). In some cases, regulations require educational certification, prohibiting workers from performing specific tasks without the necessary credentials (Nasirian et al., 2022). We consider such regulatory requirements in this research, as Australian laws mandate that only individuals with official certificates can work in electrical and plumbing trades.

Allocating overtime: Employers typically pay overtime at a higher rate. Overtime is a common method for addressing demand fluctuations in various sectors, including nurse scheduling. Research in this area considers three types of overtime: scheduled overtime, unscheduled overtime, and on-call overtime. *Scheduled overtime* is planned and communicated to employees in advance. This is the most desirable type of overtime and is commonly considered in research studies (Campbell, 2017). *Unscheduled overtime* is not planned in advance and is typically required to meet unexpected increases in demand. It is generally undesirable. For instance, in the healthcare sector, it can lead to errors such as dispensing incorrect medications or using inappropriate needle sticks (Campbell, 2017). *On-call overtime* is scheduled in advance, but the employee is only required to work if called in. On-call employees typically receive a standby fee for their availability and are paid at overtime rates for any hours worked (Campbell, 2012).

Considering our case study of a prefabrication factory producing bathroom and kitchen pods for construction projects, incorporating overtime for scheduling is not a favorable option due to various obstacles including (i) community disturbance, (ii) worker safety, and (iii) legal barriers. Prefabrication factories are often located near residential areas to minimize transportation expenses (Mostafavi & Cha, 2023). *Community disturbance* refers to the noise and traffic generated by industrial operations, which can lead to complaints from nearby residents (Mostafavi & Cha, 2023). In Australia, regulatory standards limit industrial operations to 7 AM to 8 PM on weekdays and 9 AM to 8 PM on weekends, ensuring minimal disruption to surrounding communities (EPAV, 2021). *Worker safety* can be compromised by overtime fatigue, which increases the risk of injuries, especially when operating heavy machinery or handling hazardous materials. Research in the United States has shown a link between working overtime and an increased risk of occupational injuries (Dembe et al., 2005). *Legal barriers* are associated with regulations on overtime. The Australian Fair Work, which oversees overtime-related matters, allows employees to decline overtime assignments for various reasons, including family responsibilities (Fair-Work-Ombudsman, 2022).

2.2. Stochastic skill-based workforce planning

Since real life is always associated with some degree of uncertainty, stochastic models can better represent these situations. Despite this, a significant amount of research on multiskilled workforce staffing and scheduling assumes deterministic settings (Wang & Tang, 2020). De Bruecker et al. (2015) reviewed the deterministic models, and therefore, here we focus solely on stochastic staffing and scheduling models.

The fourth, fifth, and sixth columns in Table 1 classify the skill-based staffing and scheduling papers based on their context, uncertainty considerations, and solution methods, respectively. Most studies such as Avramidis et al. (2010), associate stochastic programming with demand uncertainty. Some articles like Stratman et al. (2004), consider uncertainty in workforce salary and training expenses. Song and Huang (2008) link uncertainty with workforce churn, while another research stream, such as Zülch et al. (2004), considers uncertainty in the supply of qualified workforce. We observe that only a few studies in the manufacturing context consider uncertainty. Most investigations addressing uncertainty focus on service contexts, such as hospitals (Campbell, 2017) and call centers (Wallace & Whitt, 2005). Many existing studies that account for uncertainty either lack a solution method or rely on heuristic approaches.

2.3. Convergence of Benders' decomposition

Benders' decomposition tackles complex optimization problems by decomposing them into a master problem (MP) and a set of subproblems (SPs), facilitating efficient optimization for large-scale systems (Benders, 1962). Since the MP includes only a subset of the variables and constraints, it loses information regarding the components assigned to the SPs. Benders' decomposition is an iterative algorithm. Initially, the MP is solved, and the solution, including constant values for the MP variables, is communicated to the SPs. The SP solution is utilized to generate cuts, which are incorporated into the MP,

Table 2
Taxonomy for improving convergence of the Benders' decomposition algorithm.

Taxonomy	Technique
Decomposition strategy	Including complicating variable(s) in MP (Elçi & Hooker, 2022).
Solution procedure	Removing and selecting cuts (Minoux, 1986), ϵ-optimality (Geoffrion & Graves, 1974), heuristic (Poojari & Beasley, 2009), column generation (Muter et al., 2018), constraint programming, (Jain & Grossmann, 2001), branch and check (Beck, 2010; Thorsteinsson, 2001), approximation (Zakeri et al., 2000), column generation (Mercier & Soumis, 2007), parallelism (Wolf & Koberstein, 2013)
Solution generation	Valid inequalities (Elçi & Hooker, 2022), multi-cut approach (Sen et al., 2015), heuristic (Santoso et al., 2005).

thereby reintroducing the information lost during the decomposition. However, this iterative process can be computationally expensive, justifying the need for techniques to enhance the convergence of Benders' decomposition (Rahmaniani et al., 2017).

The literature survey on Benders' decomposition by Rahmaniani et al. (2017) categorizes efforts to enhance the algorithm's convergence into four taxonomies: cut generation, decomposition strategy, solution procedure, and solution generation. *Cut generation* deals with techniques for developing optimality and feasibility cuts. We will skip the discussion of cut generation techniques here, as they are the focus of Section 2.4. Table 2 summarizes the other three taxonomies, their associated techniques, and research papers that have explored these techniques.

Decomposition strategy deals with how to decompose the problem into an MP and SPs (Elçi & Hooker, 2022). *Solution procedure* identifies the algorithms used for solving the MP and SPs, applicable at both levels. Techniques at the MP level include:

- *Removing and selecting cuts:* Strategies for removing in-active cuts or only inserting active cuts. Many Benders' cuts do not contribute to the algorithm's convergence and instead make the MP huge, increasing solution time. Over 90% of the total solution time in Benders' decomposition is used for solving the MP (Rahmaniani et al., 2017). In any optimal solution to the MP, the number of active constraints is less than the number of variables (Minoux, 1986).
- *ϵ-optimality:* This technique is based on the idea that achieving optimality is not necessary for generating valid cuts. It involves solving the MP with a tolerance gap, ϵ, initially and gradually reducing ϵ to ensure global convergence (Geoffrion & Graves, 1974).
- *Heuristics:* Used for solving the MP to accelerate the improvement of the bounds (Poojari & Beasley, 2009).
- *Column generation:* Used for solving the MP if it has a specific structure (Muter et al., 2018).
- *Constraint programming:* Usually used to solve the SPs (Jain & Grossmann, 2001).
- *Branch and check:* This technique avoids solving the MP in several iterations considering the substantial time required. It generates a search tree for the MP and solves the SPs for all (Thorsteinsson, 2001) or specific (Beck, 2010) MP incumbents.

At the SP level, solution procedure techniques include:

- *Approximation:* Generates valid cuts using a sub-optimal solution for the dual of SPs (Zakeri et al., 2000).
- *Column generation:* Used for linear large-scale SPs with specific structure (Mercier & Soumis, 2007).
- *Parallelism:* Effective when there are a significant number of SPs (Wolf & Koberstein, 2013).

Solution generation deals with the techniques that facilitate obtaining the value of complicating variables. Techniques include:

- *Valid inequalities:* Adding these to the MP can decrease the solution time (Elçi & Hooker, 2022).
- *Multi-cut approach:* Inserts a cut for each sub-problem per iteration (Sen et al., 2015).
- *Heuristics:* Used as a warm-start strategy to generate an initial set of tight cuts to strengthen the relaxed MP (Santoso et al., 2005).

2.4. Application of LBBD to scheduling problems

The classic Benders' decomposition technique (Benders, 1962) and the L-shaped algorithm (Van Slyke & Wets, 1969) are two well-known methods for solving linear two-stage stochastic programs. According to Rahmaniani et al. (2017), when Benders' decomposition is applied to solve a stochastic problem, it is referred to as the L-shaped method. A limitation of the classic Benders' decomposition is that the SPs need to be a linear program or a convex nonlinear program (Geoffrion, 1972) because the Benders' cuts are derived from dual variables of the SPs. Therefore, if the SPs encompass integer variables, the conventional Benders' decomposition approach cannot be used.

Logic-based Benders' decomposition (LBBD) is an efficient method for solving integer or mixed integer stochastic problems. LBBD was introduced by Dawande and Hooker (2000) and further developed by Hooker and Ottosson (2003). A guide to applying LBBD to planning and scheduling problems is presented in Hooker (2007). In LBBD, the cuts are not generic; instead, they are tailored for each problem's requirements. Researchers can extract the Benders' cuts from the problem's structure. Following Elçi and Hooker (2022), we refer to such cuts as *analytical cuts*.

Table 3 presents an overview of the applications of LBBD and Benders' decomposition convergence enhancement strategies in planning and scheduling problems. It is evident that applying LBBD often involves one or more convergence enhancement strategies. In some cases, the algorithm's performance is improved by integrating LBBD with these strategies, as demonstrated by Li et al. (2023). In other cases, researchers have evaluated the algorithm's performance by combining LBBD with various convergence enhancement strategies (Naderi & Roshanaei, 2020) or comparing LBBD's performance against alternative decomposition techniques, like classic Benders' cuts (Elçi & Hooker, 2022).

Using constraint programming to solve SPs is a common method to deal with their complexity (Gedik et al., 2016; Naderi et al., 2022, 2021). Another technique to manage the SP complexity is limiting the SP solution time. Barzanji et al. (2020) limit each SP solution time to 5 s or 5% of the gap, whichever comes sooner. The branch and check method is frequently used alongside LBBD to deal with the MP complexity (Elçi & Hooker, 2022; Emde et al., 2020; Li et al., 2023). Additionally, using valid inequalities in conjunction with LBBD is a common technique to manage the MP complexity (Naderi & Roshanaei, 2020; Sun et al., 2019).

As discussed in Section 2.2, most workforce planning and scheduling studies either neglect uncertainty or consider it but rely on heuristic approaches. Reviewing Table 3, it becomes evident that LBBD has not yet been applied to stochastic multiskilled workforce staffing and scheduling problems, or their more general form, flexible resource planning and scheduling. Applying LBBD to stochastic planning and scheduling problems represents a novel research area. Elçi and Hooker (2022) claimed to have developed the first LBBD for a two-stage stochastic planning and scheduling problem. Therefore, developing an LBBD algorithm to address a two-stage stochastic multiskilled workforce staffing and scheduling problem, with the potential for expansion to flexible resource planning and scheduling, constitutes a significant contribution of this research.

Table 3
Applications of LBBD and Benders' decomposition convergence enhancement techniques to planning and scheduling problems.

Paper	Context	LBBD	B&C	CP	NG	H-SP	VI	WS	Cut P	Temp cut	SP R	MP R
Li et al. (2023)	Order scheduling	✓	✓									
Elçi and Hooker (2022)	Production scheduling	✓	✓		✓		✓				✓	
Naderi et al. (2022)	Process scheduling	✓		✓								
Naderi et al. (2021)	Hospital	✓		✓								
Emde et al. (2020)	Machine scheduling	✓	✓			✓	✓					
Barzanji et al. (2020)	Process scheduling	✓				✓						
Naderi and Roshanaei (2020)	Order scheduling	✓	✓					✓	✓	✓	✓	
Roshanaei et al. (2020a)	Hospital	✓	✓	✓					✓			
Heching et al. (2019)	Healthcare delivery	✓	✓	✓								
Sun et al. (2019)	Crane scheduling	✓										
Roshanaei et al. (2017b)	Hospital	✓							✓		✓	✓
Roshanaei et al. (2017a)	Hospital	✓										
Gedik et al. (2016)	Machine scheduling	✓		✓								
Verstichel et al. (2015)	Locker scheduling	✓									✓	

B&C: branch and check; CP: constraint programming; NG: no-good; H-SP: heuristic for SP; VI: valid inequalities; WS: warm start; Cut P: cut propagation; Temp cut: temporary cuts; SP R: SP relaxation; MP R: MP relaxation; SP: sub-problem; MP: master problem.

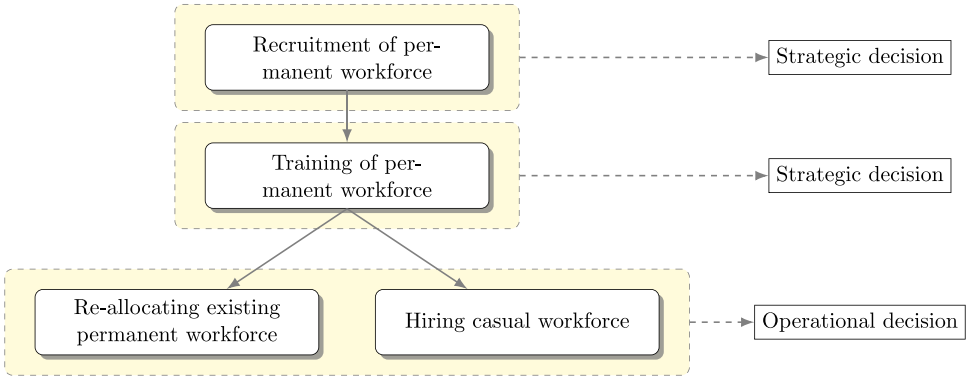


Fig. 1. Schema of the workforce planning.

3. Two-stage stochastic model

3.1. Problem description

This paper studies a skill-based workforce planning problem, which considers workforce recruitment, training, and allocation. Such workforce planning problems are relevant to industrial companies, government organizations, educational institutions, and similar entities. Henceforth, we refer to such an entity as the company. The workforce planning problem consists of two hierarchical levels of decision-making to meet the demand over a planning horizon. The main decisions of the problem, illustrated in Fig. 1, are as follows.

- *Recruitment:* At the higher level, the company decides who to recruit from a pool of potential *permanent employees*, denoted by $\mathcal{H}$. In this study, we assume that each individual (i.e., permanent candidate) $h \in \mathcal{H}$ is associated with exactly one initial (*primary*) skill j in the skill set $\mathcal{J}$, indicated by a binary parameter $S_{hj} = 1$. Thus, we have $\sum_{j\in\mathcal{J}} S_{hj} = 1, \forall h \in \mathcal{H}$.
- *Training:* After recruiting permanent workers, the company decides which staff members to train for additional (*secondary*) skill(s). The training decisions determine which skills the recruited employees will be trained to acquire.
- *Allocation:* Once workers are recruited and trained, the company allocates them to tasks to fulfill the demand. If the demand cannot be satisfied with the existing permanent workforce, the company can recruit *casual workers* to address the shortfall.

3.2. Mathematical model

In this section, we first explain the main components of our model, followed by the presentation of our mathematical model. Table 4 summarizes the key notation and symbols used in the mathematical model, with the detailed definitions provided throughout Section 3.2.

3.2.1. Demand scenarios

This study considers demand in terms of skills rather than jobs, tasks, or products. We partition the planning horizon into periods, denoted by $t \in \mathcal{T} = \{1, 2, \ldots, T\}$. We consider a scenario-based stochastic model, where each scenario ω in a finite set of scenarios Ω comprises a vector of demands $\{D_{jt\omega}\}_{j\in\mathcal{J},\, t\in\mathcal{T}}$. Here, $D_{jt\omega} \geq 0$ represents the demand of skill $j \in \mathcal{J}$ during period $t \in \mathcal{T}$ under scenario $\omega \in \Omega$. The probability of scenario $\omega \in \Omega$ occurring is P_ω, that is, $P: P_\omega \to (0, 1]$ and $\sum_{\omega\in\Omega} P_\omega = 1$.

3.2.2. Workforce recruitment and training

To track the skills of permanent workforce after recruitment and training, we introduce a binary variable, χ_{hj}, which takes the value 1 if worker h possesses skill j (either as their primary skill or a secondary skill acquired through training), and 0 otherwise. This variable, along with the parameter S_{hj} can represent recruitment and training decisions. Since each permanent worker has a unique primary skill, a worker is considered recruited if $\sum_{j\in\mathcal{J}} S_{hj}\chi_{hj} = 1$, and not recruited if $\sum_{j\in\mathcal{J}} S_{hj}\chi_{hj} = 0$. Workers cannot be trained unless they are recruited. If a recruited worker has $S_{hj} = \chi_{hj} = 1$, then j is their primary skill; if $S_{hj} = 0$, however, $\chi_{hj} = 1$, then j is a secondary skill acquired through training.

When the permanent workforce is insufficient to meet the demand, casual staff can be recruited. They are untrainable and each possesses a unique skill. So, they can be represented by their skills. The integer decision variable $\gamma_{jt\omega} \geq 0$ denotes the number of casual workers with skill j employed in period $t \in \mathcal{T}$ under scenario $\omega \in \Omega$.

Table 4
Important notation and symbols.

Symbol	Usage
Sets	
$\mathcal{H}$	Set of permanent staff candidates, $H = \lvert\mathcal{H}\rvert$
$\mathcal{J}$	Set of skills, $J = \lvert\mathcal{J}\rvert$
$\mathcal{T}$	Set of time periods, $\mathcal{T} = \{1, 2, \dots, T\}$
Ω	Set of scenarios indexed by ω
Parameters	
S_{hj}	$= 1$ if worker $h \in \mathcal{H}$ has a primary skill $j \in \mathcal{J}$; $= 0$ otherwise
Z^p_{ht}	Salary of a permanent worker (with p in the superscript) $h \in \mathcal{H}$ in period $t \in \mathcal{T}$
Z^c_{jt}	Salary of a casual worker (with c in the superscript) with skill $j \in \mathcal{J}$ in period $t \in \mathcal{T}$
Z^m_{hj}	Cost of training (with m in the superscript) worker $h \in \mathcal{H}$ to acquire skill $j \in \mathcal{J}$
$D_{jt\omega}$	Demand of skill $j \in \mathcal{J}$ in period $t \in \mathcal{T}$ under scenario $\omega \in \Omega$
P_ω	Probability of scenario $\omega \in \Omega$ to occur, $P_\omega \in (0, 1]$
Decision variables	
χ_{hj}	$= 1$ if worker $h \in \mathcal{H}$ masters skill $j \in \mathcal{J}$; $= 0$ otherwise
$\alpha_{hjt\omega}$	$= 1$ if worker $h \in \mathcal{H}$ has been allocated to conduct a task that requires skill j in period $t \in \mathcal{T}$ under scenario $\omega \in \Omega$; $= 0$ otherwise
$\gamma_{jt\omega}$	Number of casual workers hired for a task that requires skill $j \in \mathcal{J}$ in period $t \in \mathcal{T}$ under scenario $\omega \in \Omega$, $\gamma_{jt\omega} \in \mathbb{N}_0$

3.2.3. Workforce allocation

Regarding the workforce allocation, we define the binary variable $\alpha_{hjt\omega}$ to indicate whether permanent worker $h \in \mathcal{H}$ is assigned to conduct a task that requires skill $j \in \mathcal{J}$ in period $t \in \mathcal{T}$ under scenario $\omega \in \Omega$. Of course, if worker h does not possess skill j, then they cannot be assigned to perform the task, hence $\alpha_{hjt\omega} = 0$ if $\chi_{hj} = 0$. As casual staff have unique skills, they are allocated to tasks corresponding to their specific skills. The capacity allocated to perform tasks requiring skill j in period t under scenario ω is therefore $\sum_{h\in\mathcal{H}} \alpha_{hjt\omega} + \gamma_{jt\omega}$, which must satisfy the demand $D_{jt\omega}$.

3.2.4. Costs

The objective is to minimize the total cost of recruitment and training, which consists of three components:

- *Recruitment of permanent workforce:* The cost (salary) of employing a permanent worker $h \in \mathcal{H}$ during period t is Z^p_{ht}. The employment starts at the beginning of the planning horizon. Therefore, the total cost for hiring the permanent workforce is $\sum_{t\in\mathcal{T}} \sum_{h\in\mathcal{H}} Z^p_{ht} \sum_{j\in\mathcal{J}} S_{hj}\chi_{hj}$.
- *Recruitment of casual workforce:* Let Z^c_{jt} denote the cost (salary) for hiring a casual worker with skill $j \in \mathcal{J}$ during period t. The cost for the casual workforce in scenario $\omega \in \Omega$ is then $\sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} Z^c_{jt}\gamma_{jt\omega}$. The salary of a casual worker is generally higher than that of a permanent worker. In this study, following Campbell (2017), we assume an unlimited pool of casual candidates.
- *Training of permanent workforce:* The one-time cost for training worker $h \in \mathcal{H}$ to acquire a new skill $j' \in \mathcal{J}$ is $Z^m_{hj'}$. If $j \in \mathcal{J}$ is the primary skill of worker h (i.e., $S_{hj} = 1$), we set $Z^m_{hj} = 0$. The training cost for a secondary skill j' similar to the primary one j could be set lower. This could be realized by properly setting $Z^m_{hj'}$. The total cost of training is $\sum_{h\in\mathcal{H}} \sum_{j'\in\mathcal{J}} z^m_{hj'}\chi_{hj'}$,

3.2.5. Stochastic formulation

We formulate a two-stage stochastic program to address the demand uncertainty and capture the recourse decisions.

Recall that a standard two-stage stochastic program can be written as:

$$\min_{\boldsymbol{x}} \; c^T\boldsymbol{x} + \mathbb{E}[Q(\boldsymbol{x}, \xi)] \tag{1a}$$

$$\text{s.t. } A\boldsymbol{x} \le b \text{ with} \tag{1b}$$

$$Q(\boldsymbol{x}, \xi) = \min_{\boldsymbol{y}(\xi)} \{q\boldsymbol{y}(\xi) | W\boldsymbol{y}(\xi) + V\boldsymbol{x} \le h\} \tag{1c}$$

where $\boldsymbol{x}$ is the vector of the first-stage variables, $\boldsymbol{y}(\xi)$ is the vector of the second-stage variables, and ξ is a random vector, which affects the second-stage parameters, q, W, V and h. Furthermore, c, q, A, W and V are parameter matrices, and b and h are parameter vectors. In the general form (1), we did not include domain constraints $\boldsymbol{x} \ge 0$ or $\boldsymbol{y}(\xi) \ge 0$ as they can be enforced by the inequality constraints (1b) and (1c).

In this study, the strategic decisions for the recruitment and training of the permanent workforce, $\chi = \{\chi_{hj}, h \in \mathcal{H}, j \in \mathcal{J}\}$, are the first-stage (S1) variables. The operational decisions for the staff allocation, $\boldsymbol{\alpha}_\omega = \{\alpha_{hjt\omega}, h \in \mathcal{H}, j \in \mathcal{J}, t \in \mathcal{T}\}$, and the recruitment of the casual workforce, $\boldsymbol{\gamma}_\omega = \{\gamma_{jt\omega}, j \in \mathcal{J}, t \in \mathcal{T}\}$, are the second-stage (S2) recourse variables under scenario $\omega \in \Omega$. The two-stage stochastic programming model is formulated as follows.

$$\text{S1: } \min_{\chi} \sum_{h\in\mathcal{H}} \Big(\sum_{t\in\mathcal{T}} Z^p_{ht} \sum_{j\in\mathcal{J}} S_{hj}\chi_{hj} + \sum_{j\in\mathcal{J}} Z^m_{hj}\chi_{hj} \Big) + \sum_{\omega\in\Omega} P_\omega Q(\chi, \omega) \tag{2}$$

$$\text{s.t. } J \sum_{j\in\mathcal{J}} S_{hj}\chi_{hj} \ge \sum_{j\in\mathcal{J}} \chi_{hj} \qquad \forall h \in \mathcal{H} \tag{3}$$

$$\chi_{hj} \in \{0, 1\} \qquad \forall h \in \mathcal{H}, j \in \mathcal{J} \tag{4}$$

$$\text{S2: } Q(\chi, \omega) = \min_{\boldsymbol{\alpha}_\omega, \boldsymbol{\gamma}_\omega} \sum_{j\in\mathcal{J}} \sum_{t\in\mathcal{T}} Z^c_{jt}\gamma_{jt\omega} \tag{5}$$

$$\text{s.t. } \alpha_{hjt\omega} \le \chi_{hj} \qquad \forall h \in \mathcal{H}, j \in \mathcal{J}, t \in \mathcal{T}, \omega \in \Omega \tag{6}$$

$$\sum_{j\in\mathcal{J}} \alpha_{hjt\omega} \le 1 \qquad \forall h \in \mathcal{H}, t \in \mathcal{T}, \omega \in \Omega \tag{7}$$

$$\sum_{h\in\mathcal{H}} \alpha_{hjt\omega} + \gamma_{jt\omega} \ge D_{jt\omega} \qquad \forall j \in \mathcal{J}, t \in \mathcal{T}, \omega \in \Omega \tag{8}$$

$$\alpha_{hjt\omega} \in \{0, 1\} \qquad \forall h \in \mathcal{H}, j \in \mathcal{J}, t \in \mathcal{T}, \omega \in \Omega \tag{9}$$

$$\gamma_{jt\omega} \in \mathbb{N}_0 \qquad \forall j \in \mathcal{J}, t \in \mathcal{T}, \omega \in \Omega \tag{10}$$

The objective function of the first stage is given in (2), where the last component is the expected value of the second-stage cost, defined in (5). Constraint (3) guarantees that worker $h \in \mathcal{H}$ must be employed (i.e., $\sum_{j\in\mathcal{J}} S_{hj}\chi_{hj} = 1$), if any of their skills is enabled (i.e., $\sum_{j\in\mathcal{J}} \chi_{hj} > 0$). Constraint (6) ensures that the allocation of worker h to perform tasks associated with skill j is allowed only if worker h masters this skill. Constraint (7) guarantees that a worker can be allocated to at most one skill in one period. Constraint (8) ensures that the demand is fulfilled by the permanent and/or casual workforce. Let $\mathbb{N}_0$ denote the set of non-negative integers, and Constraints (4), (9), and (10) define the ranges for the decision variables.

4. Solution approach

When solving a stochastic program, the key to developing an efficient algorithm lies in leveraging the specific structure of the program. This section presents the solution approach designed based on the structure of the model developed in the previous section.

4.1. Logic-based Benders' decomposition

We apply LBBD to solve our two-stage stochastic program and develop efficient analytical cuts tailored to the proposed program. We decouple the program into a master problem, which deals with the first-stage decisions, χ, of the recruitment and training of the permanent workforce and a set of sub-problems $\{\text{SP}_{t\omega}, t \in \mathcal{T}, \omega \in \Omega\}$, which determine the second-stage variables, $(\boldsymbol{\alpha}_{t\omega}, \boldsymbol{\gamma}_{t\omega})$, of hiring casual staff and assigning the workers in each period $t \in \mathcal{T}$ under each scenario $\omega \in \Omega$. The LBBD procedure is summarized in Fig. 2, followed by a detailed description.

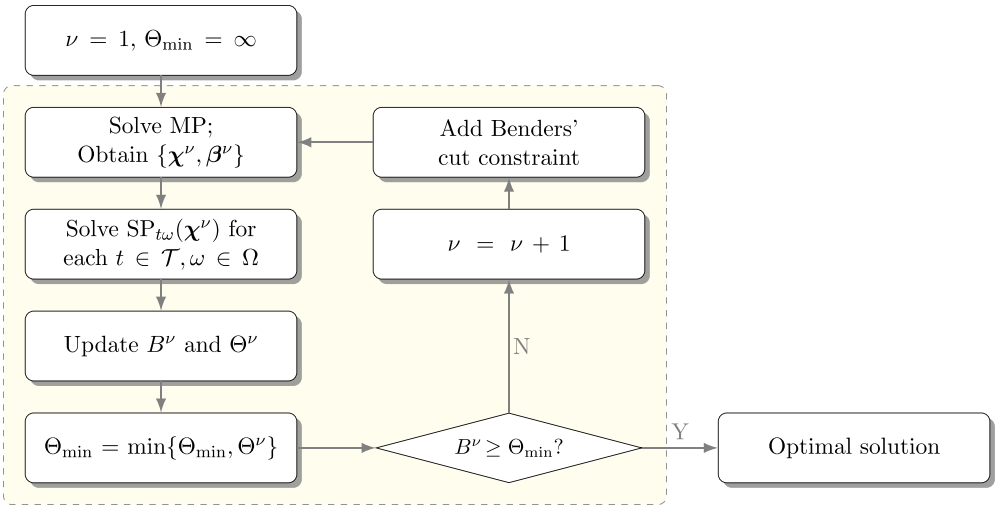


Fig. 2. Framework of the proposed LBBD algorithm.

Step (1) Set iterator $\nu = 1$, and $\Theta_{\min} = \infty$.

Step (2) Solve the following MP.

$$\text{MP: } \min_{\chi,\beta} \sum_{h\in\mathcal{H}}\Big(\sum_{t\in\mathcal{T}} Z^p_{ht}\sum_{j\in\mathcal{J}} S_{hj}\chi_{hj} + \sum_{j\in\mathcal{J}} Z^m_{hjt}\chi_{hj}\Big) + \sum_{\omega\in\Omega}\sum_{t\in\mathcal{T}} P_\omega\beta_{t\omega} \tag{11a}$$

$$\text{s.t. } \beta_{t\omega} \geq \theta^r_{t\omega}\Big(1 - \sum_{(h,j)\in\hat{\chi}^r}(1-\chi_{hj})\Big) - \zeta^r_{t\omega} \quad \forall t\in\mathcal{T}, \omega\in\Omega, r = 1,\dots,\nu-1 \tag{11b}$$

$$\beta_{t\omega} \geq 0 \qquad \forall t\in\mathcal{T}, \omega\in\Omega \tag{11c}$$

(3) and (4) (11d)

New symbols in the MP are defined as follows.

Definition 1. The variable $\theta^r_{t\omega}$, whose formula is defined in Step (3), denotes the recruitment cost of casual staff for the $SP_{t\omega}$ in iteration r, and the variable $\beta_{t\omega}$ denotes the lower bound of its cost.

Definition 2. Given the solution χ^r of the MP in iteration r, we define a set of *worker-skill pairs* (h,j): $\hat{\chi}^r = \{(h,j)|\chi^r_{hj} = 1, \forall h\in\mathcal{H}, j\in\mathcal{J}\}$.

We remark that there is a one-to-one correspondence between χ^r and $\hat{\chi}^r$, that is, one can construct χ^r ($\hat{\chi}^r$) based on $\hat{\chi}^r$ (χ^r).

Definition 3. The variable $\zeta^r_{t\omega}$ is the maximum saving achieved by any new solution χ compared to the previous one χ^r from iteration r. Its formula for calculating $\zeta^r_{t\omega}$ is given in the next section.

The objective function (11a) minimizes the cost of recruiting and training permanent staff and the lower bound on the cost of hiring casual workers. Constraint (11c) defines the domain of β. Constraint (11b) is the analytical cut, whose details are discussed in Lemma 1 and Theorem 1. Note that when $\nu = 1$, the cut is absent.

Once the MP is solved, let C^ν represent the corresponding cost for recruiting and training the permanent workforce, defined as $C^\nu = \sum_{h\in\mathcal{H}}\Big(\sum_{t\in\mathcal{T}} Z^p_{ht}\sum_{j\in\mathcal{J}} S_{hj}\chi^\nu_{hj} + \sum_{j\in\mathcal{J}} Z^m_{hj}\chi^\nu_{hj}\Big)$. If $\nu > 1$, then set the lower bound of the objective function value as $B^\nu = C^\nu + \sum_{\omega\in\Omega}\sum_{t\in\mathcal{T}} P_\omega\beta^\nu_{t\omega}$; set $B^\nu = C^\nu$ otherwise.

Step (3) For each t and ω, solve the $SP_{t\omega}$ while fixing χ at χ^ν.

$$SP_{t\omega}\text{: } \theta^\nu_{t\omega} = \min_{\alpha_{hjt\omega},\gamma_{jt\omega}} \sum_{j\in\mathcal{J}} Z^c_{jt}\gamma_{jt\omega} \tag{12a}$$

$$\text{s.t. } \alpha_{hjt\omega} \leq \chi^\nu_{hj} \qquad \forall h\in\mathcal{H}, j\in\mathcal{J} \tag{12b}$$

and (7)–(10) for the given t and ω (12c)

The recruitment cost for the casual workforce is captured by (12a). The inequality constraint (12b) links the permanent workforce's skill sets with the staff allocation.

Let $\Theta^\nu = C^\nu + \sum_{\omega\in\Omega}\sum_{t\in\mathcal{T}} P_\omega\theta^\nu_{t\omega}$ and $\Theta_{\min}$ be the lowest Θ^ν observed so far. This upper bound is compared with the lower bound B^ν. The iteration process terminates when $B^\nu \geq \Theta_{\min}$. Otherwise, a new analytical cut (11b) is generated and added to the MP, ν is incremented by 1, and the process goes to Step (2).

4.2. Analytical cut

Lemma 1. *Given a worker-skill pair set $\hat{\chi}^r$ derived from the MP solution χ^r in iteration r, we can define a new set $\hat{\chi}^+$ as $\hat{\chi}^+ = \hat{\chi}^r \cup \{(\bar{h},\bar{j})\}$, where the worker-skill pair $(\bar{h},\bar{j}) \notin \hat{\chi}^r$. The corresponding worker-skill matrix is denoted as χ^+.*

If one fixes χ at χ^+, then the optimal cost of the sub-problem $SP_{t\omega}$ satisfies

$$\theta^+_{t\omega} \geq \theta^r_{t\omega} - \max_{j\in\Gamma^r_{t\omega}} Z^c_{jt},$$

where $\Gamma^r_{t\omega} = \{j|\gamma^r_{jt\omega} > 0, \forall j\in\mathcal{J}\}$ is the set of skills for which at least one casual worker was hired in the sub-problem $SP_{t\omega}$ in iteration r.

Proof. The proof is straightforward. With this additional worker-skill pair $(\bar{h},\bar{j})$, the sub-problem $SP_{t\omega}$ can hire one fewer casual worker at most. So, the maximum cost saving in the casual recruitment would be $\max_{j\in\Gamma^r_{t\omega}} Z^c_{jt}$.

The above lemma assumes that there is at least one worker-skill pair not present in $\hat{\chi}^r$. This is because if $\nexists(h,j) \notin \hat{\chi}^r$, which means that χ^r is an all-ones matrix, then no better solution would be found in the next iteration and hence the iterative procedure in Section 4.1 should terminate.

Table 5 gives an example where this maximum saving is obtained. In this example, there are four permanent workers, #1, #2, #3 and #4, three skills, *A*, *B*, and *C*, and the corresponding casual salaries are Z^c_{jt} = 80k, 75k and 72k respectively. The demands $D_{jt\omega}$ for the three skills in period t under scenario ω are 2, 1 and 1. The MP and SP solutions in iteration r are given in the left table. Here we choose to add $(\bar{h},\bar{j}) = (2, B)$ and build $\hat{\chi}^+$ and χ^+ (see the right table). From $\gamma^r_{jt\omega}$, we have $\Gamma^r_{t\omega} = \{B, C\}$ and $\bar{j} = B = \arg\max_{j\in\Gamma^r_{t\omega}} Z^c_{jt}$. Allocating employee $\bar{h} = 2$ to skill $\bar{j} = B$ can save the cost of hiring a casual worker by 75k.

Theorem 1 explains how to derive an analytical cut based on Lemma 1. The following definitions will be used.

Table 5
Example 1: the MP and $SP_{t\omega}$ solutions χ^r and $\gamma^r_{t\omega}$ in iteration r (left), the worker-skill sets $\hat{\chi}^r$ and $\hat{\chi}^+$ and the additional worker-skill pair $(\bar{h},\bar{j})=(2,B)$ (middle), and the new solutions χ^+ and $\gamma^+_{t\omega}$ after adding $(\bar{h},\bar{j})$ (right).

χ^r_{hj}	A	B	C
1	1	1	0
2	0	0	0
3	1	0	0
4	0	0	0
$\gamma^r_{jt\omega}$	0	1	1

$\Rightarrow$ $\hat{\chi}^r=\{(1,A),(1,B),(3,A)\}$, $(\bar{h},\bar{j})=(2,B)$, $\Downarrow$, $\hat{\chi}^+=\{(1,A),(1,B),(3,A),(2,B)\}$ $\Rightarrow$

χ^+_{hj}	A	B	C
1	1	1	0
2	0	1	0
3	1	0	0
4	0	0	0
$\gamma^+_{jt\omega}$	0	0	1

Definition 4. Based on the solution $\gamma^r_{t\omega}$ to the sub-problem $SP_{t\omega}$ in iteration r, we build an ordered sequence of casual workers' salary costs $\mathcal{Z}^r_{t\omega}=\{\underbrace{Z^c_{jt},\dots,Z^c_{jt}}_{\gamma^r_{jt\omega}},\forall j\in\mathcal{J}\}$. $\mathcal{Z}^r_{t\omega,i}$ refers to the ith element of sequence $\mathcal{Z}^r_{t\omega}$, and $\mathcal{Z}^r_{t\omega,i}\geq Z^r_{t\omega,i+1}$ for $i=1,2,\dots,|\mathcal{Z}^r_{t\omega}|-1$.

Definition 5. Comparing a solution χ to the solution χ^r from iteration r, we define an upper bound of the number of additional permanent employees as $g^r=\sum_{h\in\mathcal{H}}\mathbf{1}(\sum_{j\in\mathcal{J}}(\chi_{hj}-\chi^r_{hj})>0)$, where $\mathbf{1}(\cdot)$ is an indicator function.

Definition 6. The maximum saving for the sub-problem $SP_{t\omega}$ in interaction r is $\zeta^r_{t\omega}=\sum_{i=1}^{g^r}\mathcal{Z}^r_{t\omega,i}$ if $g^r\leq|\mathcal{Z}^r_{t\omega}|$ and otherwise, $\zeta^r_{t\omega}=\sum_{i=1}^{|\mathcal{Z}^r_{t\omega}|}\mathcal{Z}^r_{t\omega,i}$.

Theorem 1. *A valid Benders' cut for period t, scenario ω, and iteration r is*

$$\beta_{t\omega}\geq\theta^r_{t\omega}\Big(1-\sum_{(h,j)\in\hat{\chi}^r}(1-\chi_{hj})\Big)-\zeta^r_{t\omega}. \tag{13}$$

Proof. We first notice that the cut becomes invalid if there exists at least one worker-skill pair $(\bar{h},\bar{j})$ such that $\chi^r_{\bar{h}\bar{j}}=1$ and $\chi_{\bar{h}\bar{j}}=0$. In that case, $\sum_{(h,j)\in\hat{\chi}^r}(1-\chi_{hj})\geq 1$, and then the right-hand side of the cut is non-positive. Otherwise, we can naively accumulate the savings for each additional worker-skill pair following Lemma 1 and obtain

$$\beta_{t\omega}\geq\theta^r_{t\omega}\Big(1-\sum_{(h,j)\in\hat{\chi}^r}(1-\chi_{hj})\Big)-\sum_{(h,j)\notin\hat{\chi}^r}\chi_{hj}\max_{j'\in\Gamma^r_{t\omega}}Z^c_{j't}. \tag{14}$$

However, the inequality (14) is not tight as the second term of the right-hand side overestimates the potential saving. Consider Example 2, which has the same permanent worker set, skill set and casual salaries as Example 1 in Table 5. The demands $D_{jt\omega}$ for the three skills during period t under scenario ω are 1, 2 and 1. The solutions of the MP and $SP_{t\omega}$ in iteration r to Example 2 as well as the new solutions are depicted in Table 6. Since $\Gamma^r_{t\omega}=\{B,C\}$ and here are four more '1' in χ compared to χ^r, the total saving by (14) is $4\times\max_{j'\in\Gamma^r_{t\omega}}Z^c_{j't}=4\times 75\text{k}$, which is far more than the actual saving. This overestimation occurs because, even if a permanent employee can acquire multiple skills, they can be allocated to at most one skill at a time. To address this, we defined g^r in Definition 5, which gives a better upper bound for the number of additional permanent employees that can be allocated to reduce the casual recruitment cost. In Example 2, by Definitions 4 and 5, $\mathcal{Z}^r_{t\omega}=\{75\text{k},75\text{k},72\text{k}\}$ and $g^r=3$. The g^r additional permanent workers can be allocated to replace the g^rth most expensive casual workers to achieve the maximum saving. In Example 2, this is realized by allocating employees #2 and #3 to skill B and employee #4 to skill C, resulting in the saving of $\zeta^r_{t\omega}=\sum_{i=1}^{3}\mathcal{Z}^r_{t\omega,i}=222\text{k}$. □

The linearization of the cut (13) is discussed in Appendix A.

4.3. Managing large master problems

Among the various approaches suggested to improve the efficiency of Benders' decomposition in Section 2.3, the method proposed by Minoux (1986) has had the most significant impact on accelerating our algorithm. Minoux (1986) argues in any optimal solution to the MP, the number of active constraints never exceeds the number of variables. To manage large-sized instances and prevent the generation of an excessive number of inactive constraints, the proposed LBBD method is implemented in two steps:

Step (i) Run the MP and the SPs alternately and iteratively for a limited number of iterations. Stop the algorithm before the number of constraints in the MP exceeds the number of variables.

Step (ii) Run the original problem, which merges the MP and SPs, inheriting all Benders' cuts generated in Step (i).

5. Results and analysis

In this section, we conduct extensive numerical experiments. Our case study focuses on a modular prefabrication factory located in Melbourne, Australia, specializing in the production of bathroom and kitchen pods. The insights gained from this study are applicable not only to similar manufacturing environments but also to other sectors that can benefit from flexible resource management, such as facilities utilizing multi-task machines.

5.1. Case study

At the time of data collection, the factory employed twenty-four single-skilled permanent workers. According to feedback from the site manager and the physical space and operational constraints, the factory could accommodate up to one extra employee per skill. That is, a maximum of 12 single-skilled permanent employees could be further recruited. There were no budget constraints limiting the recruitment of permanent staff. Table 7 lists the skills needed in the case study. The data collection process can be found in Appendix B.

5.2. Scenario generation

Instances: We study a family of 40 instances, each denoted by the format $X.Y.Z$, where X, Y and Z represent the stage of application, flexibility, and planning horizon respectively.

Stage of application: Our study examines two scenarios to assess the operational impact of our integrated staffing and scheduling model:

- *Earlier stage (E):* Represents a startup scenario where no permanent workforce exists initially. This allows for maximum flexibility in tailoring workforce recruitment and training based on demand forecasts.
- *Later stage (L):* Reflects an established company scenario where some permanent workers are already in place. Here, any new recruitment and training must consider the existing workforce structure.

Flexibility: Workforce flexibility refers to the production flexibility obtained through multiskilling. We consider two alternatives: $F2$ and $F3$, which indicate an employee can be trained to master at most two and three skills, respectively.

Planning horizon: The planning horizon represents the timeframe during which the benefits of workforce investments, including costs associated with recruiting permanent and casual employees and training

Table 6
Example 2: The MP and $SP_{t\omega}$ solutions χ^r and $\gamma^r_{t\omega}$ in iteration r (left), the worker-skill set $\hat{\chi}^r$ (middle), and the new solutions χ and $\gamma_{t\omega}$ (right).

χ^r_{hj}	A	B	C
1	1	1	0
2	0	0	0
3	0	0	0
4	0	0	0
$\gamma^r_{jt\omega}$	0	2	1

$\Rightarrow$ $\hat{\chi}^r = \{(1, A), (1, B)\}$

χ_{hj}	A	B	C
1	1	1	0
2	0	1	0
3	1	1	0
4	0	0	1
$\gamma_{jt\omega}$	0	0	0

Table 7
List of skills in the case study factory.

No.	1	2	3	4	5	6
Skill	Frame working	Caulkering	Mechanical controlling	Tiling	Plumbing	Plastering
No.	7	8	9	10	11	12
Skill	Carpentry	Electric work	Water-proofing	Glazing	Joining	Painting

Table 8
Effects of planning horizon and flexibility in different scenarios on workforce costs. (T: planning horizon; No. P: the number of permanent workers hired; No. S: the number of all skills acquired; No. M: the number of secondary skills acquired; Cost T: the total workforce cost, Cost P: permanent workforce cost; Cost M: multiskilling cost; Cost C: average casual workforce cost; *: annual cost. All costs are in million AUD).

	$E.F2$								$L.F2$							
T	No. P	No. S	No. M	Cost T*	Cost P*	Cost M	Cost M*	Cost C*	No. P	No. S	No. M	Cost T*	Cost P*	Cost M	Cost M*	Cost C*
1	32	32	0	2.13	1.76	0	0	0.37	33	33	0	2.13	1.81	0	0	0.32
2	31	36	5	2.18	1.73	0.096	0.048	0.4	31	35	4	2.18	1.74	0.084	0.042	0.4
3	28	39	11	2.21	1.58	0.24	0.08	0.55	29	35	6	2.21	1.64	0.144	0.048	0.52
4	27	40	13	2.24	1.55	0.288	0.072	0.61	29	36	7	2.26	1.7	0.168	0.042	0.51
5	27	40	13	2.28	1.59	0.288	0.0576	0.63	28	35	7	2.3	1.68	0.168	0.0336	0.59
6	27	40	13	2.32	1.64	0.288	0.048	0.64	29	37	8	2.36	1.79	0.192	0.032	0.53
7	27	40	13	2.38	1.68	0.288	0.0411	0.66	28	36	8	2.41	1.77	0.192	0.027	0.61
8	27	40	13	2.43	1.72	0.288	0.036	0.68	28	36	8	2.47	1.82	0.192	0.024	0.63
9	27	40	13	2.49	1.77	0.288	0.032	0.69	28	36	8	2.53	1.87	0.192	0.021	0.65
10	27	40	13	2.55	1.81	0.288	0.0288	0.71	28	36	8	2.6	1.92	0.192	0.0192	0.66
	$E.F3$								$L.F3$							
T	No. P	No. S	No. M	Cost T*	Cost P*	Cost M	Cost M*	Cost C*	No. P	No. S	No. M	Cost T*	Cost P*	Cost M	Cost M*	Cost C*
1	31	31	0	2.13	1.7	0	0	0.44	27	27	0	2.13	1.47	0	0	0.66
2	31	36	5	2.18	1.73	0.096	0.048	0.4	29	33	4	2.18	1.62	0.084	0.042	0.52
3	27	38	11	2.21	1.51	0.24	0.08	0.61	28	34	6	2.21	1.6	0.144	0.048	0.57
4	27	40	13	2.24	1.55	0.288	0.072	0.61	28	35	7	2.26	1.64	0.168	0.042	0.58
5	27	40	13	2.28	1.59	0.288	0.0576	0.63	29	36	7	2.3	1.74	0.168	0.0336	0.53
6	27	40	13	2.32	1.64	0.288	0.048	0.64	28	36	8	2.36	1.73	0.192	0.032	0.6
7	27	40	13	2.38	1.68	0.288	0.0411	0.66	28	36	8	2.41	1.77	0.192	0.027	0.62
8	27	40	13	2.43	1.72	0.288	0.036	0.68	28	36	8	2.47	1.82	0.192	0.024	0.63
9	27	40	13	2.49	1.77	0.288	0.032	0.69	28	36	8	2.53	1.86	0.192	0.021	0.65
10	27	40	13	2.55	1.81	0.288	0.0288	0.71	28	36	8	2.6	1.91	0.192	0.0192	0.67

the workforce, are expected to outweigh initial expenses. Based on the insights from our interviews with the site management, 10 years is the maximum length for a reasonable planning horizon. So, the planning horizon is set to be $T \leq 10$.

5.3. Computational environment

We coded the proposed LBBD algorithm using Julia programming language Version 1.6.5 with the JuMP package. The source code is available on the Github page. Julia is explicitly designed for numerical analysis and computational sciences. The algorithm was executed using Gurobi Version 10.0.1, with an optimality gap set to 10^{-4}. The computations were performed on a 5-core Intel machine with 16 GB of RAM.

5.4. Effects of planning horizon and flexibility on workforce planning

In this section, we extract some insights into the operational benefits of multiskilling in terms of workforce cost savings within a stochastic setting. Table 8 reports the results obtained from our integrated staffing and scheduling model across various instances. The table caption provides a detailed explanation of each column's meaning.

In all instances, as the planning horizon increases, the model tends to employ fewer permanent workers and invest more in multiskilling the workforce. Take the instance $E.F2$ as an example; when $T = 1$, the model recruits 32 single-skilled workers without any training. Increasing the planning horizon to $T = 2$, the model recruits 31 permanent workers and trains them in 5 extra skills. Adding one more year to the planning horizon, $T = 3$, the number of permanent staff drops to 28, while the number of secondary skills rises to 11.

Again for the instance $E.F2$, as T increases from 4 to 10, there is no change in the number of permanent and casual workers or training in secondary skills. This suggests that for this instance a planning horizon of four years is enough to reap the maximum benefit of multiskilling the workforce. Therefore, a reasonably long planning period is more practical when considering multiskilling. The planning horizon $T = 4$ leads to an annual saving of $83,500 compared with $T = 1$, in terms of the permanent and casual workforce salary.

Comparing the total workforce costs between the two flexibility alternatives $F2$ and $F3$, we see that total costs remain the same across different planning horizons within the same stage of application.

Table 9
Comparing the performance of our LBBD algorithm with Gurobi for solving MDEP with various H, J, T and $|\Omega|$. No. Con: the total number of constraints; No. Var: the total number of variables; For the LBBD, ν, No. Con* and No. Var*: the numbers of iterations, constraints and variables, respectively, by the end of Step (i) in List Step (i). The total solution time by the LBBD is composed of the solution times for the MP, the SPs, and the merged original problem. All times are in second.

Parameters				Gurobi			LBBD with cut (13)									
H	J	T	$\|\Omega\|$	No. Con	No. Var	Time	ν	No. Var*	No. Con*	No. Var	No. Con	Time MP	Time SP	Time Org	Time	Time Ratio
5	5	10	20	7055	6030	4	1	1030	55	6240	7871	0.00	0.38	0.40	0.78	5.12
5	5	10	40	14 055	12 030	14	1	2030	55	12 440	15 671	0.02	0.85	0.63	1.51	9.27
5	5	10	80	28 055	24 030	29	1	4030	55	24 840	31 271	0.02	1.40	1.82	3.24	8.95
5	5	10	160	56 055	48 030	125	1	8030	55	49 640	62 471	0.01	2.99	3.12	6.11	20.45
5	5	10	200	70 055	60 030	131	1	10 030	55	62 040	78 071	0.01	3.46	4.70	8.17	16.03
5	5	10	240	84 055	72 030	140	1	12 030	55	74 440	93 671	0.00	4.16	5.43	9.60	14.58
5	5	10	280	98 055	84 030	179	1	14 030	55	86 840	109 271	0.01	7.33	8.73	16.07	11.18
5	5	10	320	112 055	96 030	247	1	16 030	55	99 240	124 871	0.02	9.67	10.51	20.21	12.22
5	5	20	20	14 055	12 030	14	1	2030	55	12 440	15 671	0.02	4.36	2.21	6.59	2.12
5	5	20	40	28 055	24 030	49	1	4030	55	24 840	31 271	0.02	5.24	3.03	8.29	5.91
5	5	20	80	56 055	48 030	122	1	8030	55	49 640	62 471	0.014	9.42	5.90	15.34	7.95
5	5	20	160	112 055	96 030	826	1	16 030	55	99 240	124 871	0.036	17.85	13.36	31.25	26.4
5	5	20	200	140 055	120 030	1485	1	20 030	55	124 040	156 071	0.02	34.29	28.00	62.31	23.83
5	5	20	240	168 055	144 030	2408	1	24 030	55	148 840	187 271	0.015	28.81	18.53	47.36	50.84
5	5	20	280	196 055	168 030	2074	1	28 030	55	173 640	218 471	0.02	42.66	39.12	81.79	25.35
5	5	20	320	224 055	192 030	2015	2	38 440	25 671	204 850	275 287	0.13	35.64	32.85	68.63	31.68
6	6	10	20	9678	8442	13	1	1242	78	8654	10 497	0.016	1.355	0.72	2.09	6.22
6	6	10	40	19 278	16 842	38	1	2442	78	17 254	20 897	0.003	3.22	5.06	8.28	4.58
6	6	10	80	38 478	33 642	125	1	4842	78	34 454	41 697	0.015	8.05	10.34	18.41	6.78
6	6	10	160	76 878	67 242	587	1	9642	78	68 854	83 297	0.02	16.82	21.89	38.73	15.15
6	6	10	200	96 078	84 042	1112	1	12 042	78	86 054	104 097	0.01	19.41	39.10	58.52	19.00
6	6	10	240	115 278	100 842	1766	1	14 442	78	103 254	124 897	0.06	23.97	30.78	54.81	32.22
6	6	10	280	134 478	117 642	2514	1	16 842	78	120 454	145 697	0.02	27.83	49.29	77.15	32.58
6	6	10	320	153 678	134 442	3002	1	19 242	78	137 654	166 497	0.02	29.06	59.62	88.69	33.84
6	6	20	20	19 278	16 842	34	1	2442	78	17 254	20 897	0.01	2.40	2.87	5.29	6.42
6	6	20	40	38 478	33 642	109	1	4842	78	34 454	41 697	0.01	7.22	7.91	15.15	7.19
6	6	20	80	76 878	67 242	593	1	9642	78	68 854	83 297	0.01	16.48	23.24	39.73	14.92
6	6	20	160	153 678	134 442	4003	1	19 242	78	137 654	166 497	0.05	32.61	66.39	99.05	40.41
6	6	20	200	192 078	168 042	9346	1	24 042	78	172 054	208 097	0.05	35.86	62.62	98.54	94.84
6	6	20	240	230 478	201 642	13 969	1	28 842	78	206 454	249 697	0.03	47.35	109.28	156.66	89.17
6	6	20	280	268 878	235 242	18 429	1	33 642	78	240 854	291 297	0.027	55.79	156.22	212.04	86.91
6	6	20	320	307 278	268 842	26 221	1	38 442	78	275 254	332 897	0.065	52.99	142.63	195.69	133.99
7	7	10	20	12 705	11 256	25	1	1456	105	11 470	13 527	0.03	0.52	1.42	1.97	26.93
7	7	10	40	25 305	22 456	291	1	2856	105	22 870	26 927	0.01	1.20	3.85	5.05	57.62
7	7	10	80	50 505	44 856	458	1	5656	105	45 670	53 727	0.01	1.73	21.75	23.48	19.51
7	7	10	160	100 905	89 656	1778	1	11 256	105	91 270	107 327	0.01	3.28	78.68	81.97	21.69
7	7	10	200	126 105	112 056	4310	1	14 056	105	114 070	134 127	0.03	5.31	102.45	107.79	39.98
7	7	10	240	151 305	134 456	6876	1	16 856	105	136 870	160 927	0.00	6.65	160.84	167.49	41.05
7	7	10	280	176 505	156 856	10 086	2	22 470	11 327	162 484	198 949	0.16	15.51	164.57	180.24	55.95
7	7	10	320	201 705	179 256	13 760	2	25 670	12 927	185 684	227 349	0.153	34.105	257.049	291.30	47.28
7	7	20	20	25 305	22 456	101	1	2856	105	22 870	26 927	0.05	0.95	2.75	3.75	21.9
7	7	20	40	50 505	44 856	421	1	5656	105	45 670	53 727	0.00	4.32	23.09	27.41	15.93
7	7	20	80	100 905	89 656	1942	1	11 256	105	91 270	107 327	0.01	5.01	47.02	52.04	37.31
7	7	20	160	201 705	179 256	12 629	2	25 670	12 927	185 684	227 349	0.47	42.37	251.17	294.00	42.95
7	7	20	200	252 105	224 056	24 118	3	36 084	32 149	236 098	300 171	0.47	62.79	812.98	876.24	27.53
7	7	20	240	302 505	268 856	48 747	3	43 284	38 549	283 298	360 171	0.65	9.32	919.14	929.12	52.46
7	7	20	280	352 905	313 656	84 771	3	50 484	44 949	330 498	420 171	0.64	6.81	975.99	983.45	86.19
7	7	20	320	403 305	358 456	98 905	2	51 270	25 727	371 284	454 549	0.18	2.91	1572.79	1575.88	62.71

Considering $E.F2$ and $E.F3$ with $T = 3$, despite the total cost being identical, there are differences in the casual and permanent workforce costs. This is because the model has two optimal solutions. Further investigation of the model shows that no workers possess more than two skills in any instance.

Comparing the stages of application E and L, it becomes evident that implementing the integrated staffing and scheduling model before recruiting any permanent workforce offers significant advantages. First, for short planning horizons, like $T = 1$ and $T = 2$, there is no difference in the total workforce costs between E and L stages. However, as the planning horizon increases, the E instance consistently results in lower annual workforce costs. When $T = 10$, the total annual workforce cost in the E instances (i.e., $E.F2$, $E.F3$) is \$2.55 millions, which is \$50,000 lower than that of the L instance (i.e., $L.F2$, $L.F3$).

Second, adopting the integrated staffing and scheduling model before initiating any workforce recruitment significantly shortens the planning horizon required to achieve maximum cost savings. In instances $E.F2$ and $E.F3$, for $T \geq 4$, there is no alteration between the number of permanent workers hired or the number of skills they are trained in. This implies that in instances $E.F2$ and $E.F3$, a span of four years is sufficient to attain optimal staffing configuration and realize maximum workforce cost savings. In contrast, instances $L.F2$ and $L.F3$ require seven and six years, respectively, to reach optimal staffing configuration and realize the ultimate workforce cost savings.

The superiority of the earlier stage is not surprising. By applying the model in the earlier stage, the model has the full discretion to recruit the most suitable workforce and train them in appropriate skills. For example, in our case study, the model avoids recruiting two costly workers, carpenter and joiner. Instead, it trains three and five other workers to possess carpentry and joining as their secondary skills, respectively. Conversely, in the later stage, since there are already some workers with potentially undesirable skills, the model's flexibility in

Table 10
Comparing the performance of the LBBD algorithm for different numbers of iterations in Step (i) in List Step (i). ν, No. Con* and No. Var*: the numbers of iterations, constraints and variables, respectively, by the end of Step (i). The total solution time by the LBBD is composed of the solution times for the MP, the SPs, and the merged original problem. All times are in second.

		$T = 10$						$T = 20$					
$\|\Omega\|$	ν	No. Var*	No. Con*	Time MP	Time SP	Time Org	Time	No. Var*	No. Con*	Time MP	Time SP	Time Org	Time
20	1	1456	105	0.03	0.52	1.42	1.97	2856	105	0.05	0.95	2.75	3.75
	2	1670	927	0.184	0.943	0.963	2.09	3270	1727	0.03	1.39	3.19	4.61
	3	1884	1749	0.062	1.69	1.282	3.034	3684	3349	0.28	8.35	3.17	11.79
	4	2098	2571	0.078	1.836	1.304	3.218	4098	4971	0.31	10.49	4.50	15.30
40	1	2856	105	0.01	1.20	3.85	5.05	5656	105	0.00	4.32	23.09	27.41
	2	3270	1727	0.14	2.18	3.17	5.49	6470	3327	0.04	3.61	25.51	29.16
	3	3684	3349	0.11	2.96	3.53	6.60	7284	6549	0.25	14.34	25.67	40.26
	4	4098	4971	0.23	3.88	3.12	7.23	8098	9771	0.48	19.31	31.45	51.24
80	1	5656	105	0.01	1.73	21.75	23.48	11 256	105	0.01	5.01	47.02	52.04
	2	6470	3327	0.09	7.17	17.65	24.9	2870	6527	0.09	11.70	81.56	93.34
	3	7284	6549	0.21	14.07	26.04	40.32	14 484	12 949	0.37	18.46	45.58	64.42
	4	8098	9771	0.64	18.47	30.55	49.67	16 098	19 371	0.85	21.31	60.80	82.96
160	1	11 256	105	0.01	3.28	78.68	81.97	22 456	105	0.02	21.27	534.71	556.00
	2	12 870	6527	0.12	20.80	81.81	102.74	25 670	12 927	0.47	42.37	251.17	294.00
	3	14 484	12 949	0.44	26.23	110.59	137.26	28 884	25 749	0.53	41.89	265.52	307.95
	4	16 098	19 371	0.76	35.82	132.49	169.06	32 098	38 571	2.46	85.64	375.30	463.40
200	1	14 056	105	0.03	5.31	102.45	107.79	28 056	105	0.013	9.971	1517.9	1527.89
	2	16 070	8127	0.09	14.31	114.86	129.25	32 070	16 127	0.25	27.28	1207.67	1235.21
	3	18 084	16 149	0.31	24.47	136.96	161.75	36 084	32 149	0.47	62.79	812.98	876.24
	4	20 098	24 171	0.98	44.68	148.01	193.67	40 098	48 171	1.58	138.47	1101.82	1241.87
240	1	16 856	105	0.01	6.65	160.84	167.49	33 656	105	0.04	19.69	1902.52	1922.251
	2	19 270	9727	0.14	20.35	182.06	202.55	38 470	19 327	0.15	9.52	2135.68	2145.36
	3	21 684	19 349	0.34	38.95	132.24	171.53	43 284	38 549	0.65	9.32	919.14	929.12
	4	24 098	28 971	1.18	61.20	284.27	346.65	48 098	57 771	2.05	20.76	2263.9	2286.79
280	1	19 656	105	0.03	12.39	431.85	444.27	39 256	105	0.01	1.309	1758.81	1760.13
	2	22 470	11 327	0.16	15.51	164.57	180.24	44 870	22 527	0.16	3.52	1004.37	1008.06
	3	25 284	22 549	0.46	55.96	133.11	189.53	50 484	44 949	0.64	6.81	975.99	983.45
	4	28 098	33 771	1.45	82.67	916.89	1001.01	56 098	67 371	1.88	11.36	2254.09	2267.34
320	1	22 456	105	0.10	25.22	584.87	610.20	44 856	105	0.05	40.32	1955.86	1996.23
	2	25 670	12 927	0.153	34.105	257.04	291.30	51 270	25 727	0.18	2.91	1572.79	1575.88
	3	28 884	25 749	1.24	108.94	273.31	383.49	57 684	51 349	0.72	4.38	2125.14	2130.25
	4	32 098	38 571	1.99	80.23	341.21	423.44	64 098	76 971	2.28	12.52	2870.7	2885.52

training and recruitment becomes limited. More detailed results are provided in Appendix D.

5.5. Computational performance

Table 9 compares the performance of our LBBD algorithm with Gurobi in solving the MDEP. We vary the parameters including the number of permanent workers (H), the number of skills (J), the planning horizon (T), and the number of scenarios ($|\Omega|$) drawn from the sample space. The demand follows a uniform distribution with lower and upper bounds of 1 and 3, respectively. Detailed explanations of the columns can be found in the table caption. The last column 'Time Ratio' in Table 9 is the ratio of the solution time taken by Gurobi compared to our LBBD algorithm.

In all instances, the proposed LBBD algorithm consistently outperforms Gurobi for solving the MDEP. On average, our LBBD algorithm demonstrates a 29-fold speed improvement compared to Gurobi. At its peak, our algorithm achieves a 133-fold improvement over Gurobi. It is worth noting that the superiority of our algorithm increases significantly with larger values of T and $|\Omega|$.

In certain instances, the performance of Gurobi in solving the MDEP exhibits unexpected behavior. Specifically, as we increase the size of the problem, expanding the number of scenarios while keeping the parameters H, J, and T constant, we observe a decrease in the solution time. For instance, when $H = J = 5$, $T = 20$, and $|\Omega| = 240$, the solution time is 2408 s. Surprisingly, as we increase $|\Omega|$ to 280 and then to 320, the solution time unexpectedly reduces to 2074 and 2015 s, respectively. Similar unexpected behavior is observed in our LBBD algorithm. To illustrate, with $H = J = 6$, $T = 20$, and $|\Omega| = 280$, the solution time is 212.04 s. Increasing $|\Omega|$ to 320 should logically increase the solution time, but instead, it decreases to 195.69 s.

Table 10 compares the performance of the proposed LBBD algorithm when Step (i) in List Step (i) is terminated after ν iterations. We choose the instances with $J = H = 7$, $T \in \{10, 20\}$, and $|\Omega| \in \{20, 40, 80, 160, 200, 240, 280, 320\}$ as representatives. In all cases, the shortest solution time is attained when Step (i) is terminated before the number of constraints in the MP exceeds the number of variables. This finding aligns with the results of Minoux (1986) and Rahmaniani et al. (2017), who observed that when the number of constraints in the MP exceeds the number of variables, the inactive constraints further generated can enlarge the MP and slow down the solution process. As the problem size grows with larger values of both T and $|\Omega|$, more iterations are required to achieve the fastest solution time. For example, for $T = 10$ and scenario sizes $|\Omega| \in \{20, 40, 80, 160, 200, 240\}$, the fastest solution time is achieved with just one iteration. For thesis instances, employing more than one iteration actually slows down the algorithm. Nevertheless, with $|\Omega| \in \{280, 320\}$, two iterations are necessary to generate sufficient cuts for the fastest algorithm implementation.

5.6. Quality of approximation

In scenario-based stochastic programming, uncertainties are represented by a set of scenarios, each defining a specific realization of random variables along with the associated probability of occurrence. The quality of approximation can be evaluated by running the scenario-based stochastic program multiple times and examining the standard deviation of the objective function relative to the square root of sample size or employing the coefficient of variation for the objective function (Kim et al., 2022).

Table 11
Quality of approximation in scenario-based stochastic programming. $\hat{\mu}$ and $\hat{\sigma}$: the average and standard deviation of the objective function, respectively; $\hat{\sigma}/\sqrt{n}$: standard deviation of the objective function relative to the square root of sample size; $\widehat{CV}$: coefficient of variation.

	$H = J = 4$				$H = J = 5$				$H = J = 6$			
$\lvert\Omega\rvert$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\sigma}/\sqrt{n}$	$\widehat{CV}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\sigma}/\sqrt{n}$	$\widehat{CV}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\sigma}/\sqrt{n}$	$\widehat{CV}$
100	468	3.401	0.278	0.007	562	3.620	0.295	0.006	675	4.10	0.335	0.006
160	469	2.84	0.231	0.006	562	3.213	0.262	0.006	675	3.101	0.253	0.005
200	468	2.72	0.220	0.006	562	2.554	0.208	0.005	675	2.812	0.229	0.004
320	469	2.060	0.168	0.004	561	2.182	0.178	0.004	675	2.121	0.173	0.003
500	469	1.443	0.117	0.003	561	1.810	0.147	0.003	675	1.69	0.138	0.003
1000	469	1.161	0.091	0.003	561	1.180	0.096	0.002	675	1.24	0.101	0.002
2500	469	0.821	0.067	0.002	561	0.683	0.055	0.001	674	1.01	0.082	0.001
5000	469	0.566	0.046	0.001	561	0.559	0.046	0.001	674	0.57	0.046	0.001

Table 11 presents the quality of approximation for three different combinations of the numbers of workers and skills: $H = J = 4$, $H = J = 5$, and $H = J = 6$. The planning horizon is fixed at $T = 10$. The number of scenarios, $|\Omega|$, varies across the set $\{100, 160, 200, 320, 500, 1000, 2500, 5000\}$, while a constant sample size of $n = 150$ is used. Details of the table columns are explained in the caption. Across all combinations of (H, J) pairs, the accuracy of approximation consistently improves as the number of scenarios increases. Evaluating the metric $\hat{\sigma}/\sqrt{n}$, achieving a 10% level of approximation quality necessitates employing $|\Omega| = 1000$ scenarios. To achieve even higher accuracy, such as 5%, it is necessary to increase the number of scenarios to $|\Omega| = 5000$. In contrast, considering the $\widehat{CV}$ metric, a satisfactory accuracy (e.g., below 0.005) is achieved with a much smaller number of scenarios (i.e., $|\Omega| = 320$).

6. Conclusion

We developed a two-stage stochastic problem for integrated staffing and scheduling of a multiskilled workforce. The "here-and-now" decision involves recruiting permanent single-skilled workers and training them in secondary skills to become multiskilled. The "wait-and-see" decision focuses on hiring casual workers and allocating the multiskilled workforce to appropriate tasks to meet the demand.

Our integrated staffing and scheduling model employs a two-stage stochastic program with binary variables in the first stage and a combination of binary and integer variables in the second stage. Since the second stage is a mixed integer problem, we adopted the Logic-Based Benders Decomposition (LBBD) approach and developed a customized analytical cut. Our LBBD algorithm can be more than 133 times faster than the latest version of Gurobi for solving the MDEP.

We examined different planning horizons to understand the dynamics of recruiting permanent single-skilled workers, training them and hiring casual employees. As the planning horizon increases, the model favors fewer single-skilled permanent workers and more investment in training. Conversely, a short planning horizon leads to recruiting more single-skilled permanent employees and a significant amount of casual workforce costs. Our analysis suggests that recruiting single-skilled workers for every task is unnecessary. A cost-effective alternative is to avoid recruiting workers for expensive skills and instead multiskill the existing workforce to cover the demand. In our case study, the model avoids recruiting workers whose primary skills are carpentry and joining.

Advancements in robust optimization, notably in the realms of soft-constrained and almost-robust methodologies, present compelling opportunities for further research into the effective integration of uncertainty into optimization problems (Baron et al., 2019). Soft-constrained robust optimization addresses the inherent conservatism of traditional robust optimization by permitting, with associated penalties, violations of a subset of uncertain constraints. This adaptability facilitates the development of more flexible solutions in scenarios where rigid adherence to all constraints may be neither feasible nor desirable. Almost robust optimization, synthesizes concepts from stochastic and chance-constrained programming and permits a predetermined level of constraint violation at no additional cost, thereby striking a balance between the robustness and optimality of the solution.

Future research could also investigate the solution approach enhancement by adopting a multi-level LBBD approach, as suggested in the scheduling literature (Roshanaei et al., 2020b). This could potentially improve the algorithm's performance and scalability for larger and more complex problem instances (Park & Bayraksan, 2023).

CRediT authorship contribution statement

Araz Nasirian: Writing – review & editing, Writing – original draft, Visualization, Software, Project administration, Methodology, Data curation, Conceptualization. **Lele Zhang:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Alysson M. Costa:** Writing – review & editing, Writing – original draft, Methodology, Funding acquisition. **Babak Abbasi:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT and Gemini in order to improve readability and language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Acknowledgments

This research was partially funded by the Australian Government through the Australian Research Council Industrial Transformation Training Centre in Optimisation Technologies, Integrated Methodologies, and Applications (OPTIMA), Project ID IC200100009.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ejor.2024.11.033.

References

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.

Bard, J. F., Morton, D. P., & Wang, Y. M. (2007). Workforce planning at USPS mail processing and distribution centers using stochastic optimization. *Annals of Operations Research*, *155*(1), 51–78.

Bard, J. F., & Purnomo, H. W. (2005). Preference scheduling for nurses using column generation. *European Journal of Operational Research*, *164*(2), 510–534.

Baron, O., Berman, O., Fazel-Zarandi, M. M., & Roshanaei, V. (2019). Almost robust discrete optimization. *European Journal of Operational Research*, *276*(2), 451–465.

Barzanji, R., Naderi, B., & Begen, M. A. (2020). Decomposition algorithms for the integrated process planning and scheduling problem. *Omega*, *93*, Article 102025.

Beck, J. C. (2010). Checking-up on branch-and-check. In *Principles and practice of constraint programming–CP 2010: 16th international conference, CP 2010* (pp. 84–98). Springer.

Benders, J. (1962). Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, *4*(1), 238–252.

Campbell, G. M. (2011). A two-stage stochastic program for scheduling and allocating cross-trained workers. *Journal of the Operational Research Society*, *62*(6), 1038–1047.

Campbell, G. M. (2012). On-call overtime for service workforce scheduling when demand is uncertain. *Decision Sciences*, *43*(5), 817–850.

Campbell, G. M. (2017). Time-staged overtime staffing for services with updated forecasts and availabilities. *Decision Sciences*, *48*(4), 691–722.

Carlos, S., Sánchez, A., & Martorell, S. (2011). Model to study the effect of workforce on a safety equipment and its optimization. *Mathematical and Computer Modelling*, *54*(7–8), 1808–1812.

Corominas, A., Pastor, R., & Plans, J. (2008). Balancing assembly line with skilled and unskilled workers. *Omega*, *36*(6), 1126–1132.

Daniels, R. L., Mazzola, J. B., & Shi, D. (2004). Flow shop scheduling with partial resource flexibility. *Management Science*, *50*(5), 658–669.

Dawande, M., & Hooker, J. N. (2000). Inference-based sensitivity analysis for mixed integer/linear programming. *Operations Research*, *48*(4), 623–634.

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.

Dembe, A. E., Erickson, J. B., Delbos, R. G., & Banks, S. M. (2005). The impact of overtime and long work hours on occupational injuries and illnesses: new evidence from the United States. *Occupational and Environmental Medicine*, *62*(9), 588–597.

Easton, F. F., & Mansour, N. (1999). A distributed genetic algorithm for deterministic and stochastic labor scheduling problems. *European Journal of Operational Research*, *118*(3), 505–523.

Easton, F. F., & Rossin, D. F. (1996). A stochastic goal program for employee scheduling. *Decision Sciences*, *27*(3), 541–568.

Elçi, Ö., & Hooker, J. (2022). Stochastic planning and scheduling with logic-based Benders decomposition. *INFORMS Journal on Computing*, *34*(5), 2383–2865.

Emde, S., Polten, L., & Gendreau, M. (2020). Logic-based benders decomposition for scheduling a batching machine. *Computers & Operations Research*, *113*, Article 104777.

EPAV (2021). About construction noise. URL https://www.epa.vic.gov.au/for-community/environmental-information/noise/construction-noise/about-construction-noise#:~:text=The%20Regulations%20prohibit%20noise%20from,am%20and%20after%208%20pm [Accessed on November 17 2023].

Ertogral, K., & Bamuqabel, B. (2008). Developing staff schedules for a bilingual telecommunication call center with flexible workers. *Computers & Industrial Engineering*, *54*(1), 118–127.

Fair-Work-Ombudsman (2022). When overtime applies. https://www.fairwork.gov.au/employment-conditions/hours-of-work-breaks-and-rosters/hours-of-work/when-overtime-applies [Accessed on 15 October 2023].

Fragniere, E., Gondzio, J., & Yang, X. (2010). Operations risk management by optimally planning the qualified workforce capacity. *European Journal of Operational Research*, *202*(2), 518–527.

Gedik, R., Rainwater, C., Nachtmann, H., & Pohl, E. A. (2016). Analysis of a parallel machine scheduling problem with sequence dependent setup times and job availability intervals. *European Journal of Operational Research*, *251*(2), 640–650.

Geoffrion, A. M. (1972). Generalized benders decomposition. *Journal of Optimization Theory and Applications*, *10*, 237–260.

Geoffrion, A. M., & Graves, G. W. (1974). Multicommodity distribution system design by benders decomposition. *Management Science*, *20*(5), 822–844.

Harper, P. R., Powell, N., & Williams, J. E. (2010). Modelling the size and skill-mix of hospital nursing teams. *Journal of the Operational Research Society*, *61*(5), 768–779.

Heching, A., Hooker, J. N., & Kimura, R. (2019). A logic-based benders approach to home healthcare delivery. *Transportation Science*, *53*(2), 510–522.

Heimerl, C., & Kolisch, R. (2010). Scheduling and staffing multiple projects with a multi-skilled workforce. *OR Spectrum*, *32*(2), 343–368.

Helber, S., & Henken, K. (2010). Profit-oriented shift scheduling of inbound contact centers with skills-based routing, impatient customers, and retrials. *OR Spectrum*, *32*(1), 109–134.

Hooker, J. N. (2007). Planning and scheduling by logic-based Benders decomposition. *Operations Research*, *55*(3), 588–602.

Hooker, J. N., & Ottosson, G. (2003). Logic-based Benders decomposition. *Mathematical Programming*, *96*(1), 33–60.

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.

Jain, V., & Grossmann, I. E. (2001). Algorithms for hybrid MILP/CP models for a class of optimization problems. *INFORMS Journal on computing*, *13*(4), 258–276.

Kim, K., & Mehrotra, S. (2015). A two-stage stochastic integer programming approach to integrated staffing and scheduling with application to nurse management. *Operations Research*, *63*(6), 1431–1451.

Kim, S., Rasouli, S., Timmermans, H. J., & Yang, D. (2022). A scenario-based stochastic programming approach for the public charging station location problem. *Transportmetrica B: Transport Dynamics*, *10*(1), 340–367.

Lee, T. (2004). The effect of workers with different capabilities on customer delay. *Computers & Operations Research*, *31*(3), 359–381.

Li, Y., Cote, J.-F., Coelho, L. C., Zhang, C., & Zhang, S. (2023). Order assignment and scheduling under processing and distribution time uncertainty. *European Journal of Operational Research*, *305*(1), 148–163.

Li, N., & Li, L. X. (2000). Modeling staffing flexibility: A case of China. *European Journal of Operational Research*, *124*(2), 255–266.

Mercier, A., & Soumis, F. (2007). An integrated aircraft routing, crew scheduling and flight retiming model. *Computers & Operations Research*, *34*(8), 2251–2265.

Minoux, M. (1986). Mathematical programming: theory and algorithms. John Wiley & Sons.

Morris-Reade, R. (2022). The cost of hiring new workers doubles to more than $23,000. URL https://itbrief.com.au/story/the-cost-of-hiring-new-workers-doubles-to-more-than-23-000 [Accessed on 17 November 2023].

Mostafavi, A., & Cha, Y.-J. (2023). Deep learning-based active noise control on construction sites. *Automation in Construction*, *151*, Article 104885.

Muter, I., Birbil, c. İ., & Bülbül, K. (2018). Benders decomposition and column-and-row generation for solving large-scale linear programs with column-dependent-rows. *European Journal of Operational Research*, *264*(1), 29–45.

Naderi, B., Begen, M. A., & Zaric, G. S. (2022). Type-2 integrated process-planning and scheduling problem: Reformulation and solution algorithms. *Computers & Operations Research*, *142*, Article 105728.

Naderi, B., Begen, M. A., Zaric, G. S., & Roshanaei, V. (2023). A novel and efficient exact technique for integrated staffing, assignment, routing, and scheduling of home care services under uncertainty. *Omega*, *116*, Article 102805.

Naderi, B., & Roshanaei, V. (2020). Branch-relax-and-check: A tractable decomposition method for order acceptance and identical parallel machine scheduling. *European Journal of Operational Research*, *286*(3), 811–827.

Naderi, B., Roshanaei, V., Begen, M. A., Aleman, D. M., & Urbach, D. R. (2021). Increased surgical capacity without additional resources: Generalized operating room planning and scheduling. *Production and Operations Management*, *30*(8), 2608–2635.

Nasirian, A., Abbasi, B., Cheng, T., & Arashpour, M. (2022). Multiskilled workforce planning: A case from the construction industry. *Journal of Construction Engineering and Management*, *148*(5), Article 04022021.

Nasirian, A., Arashpour, M., & Abbasi, B. (2019). Critical literature review of labor multiskilling in construction. *Journal of Construction Engineering and Management*, *145*(1), Article 04018113.

Park, J., & Bayraksan, G. (2023). A multistage distributionally robust optimization approach to water allocation under climate uncertainty. *European Journal of Operational Research*, *306*(2), 849–871.

Poojari, C. A., & Beasley, J. E. (2009). Improving benders decomposition using a genetic algorithm. *European Journal of Operational Research*, *199*(1), 89–97.

Qin, R., Nembhard, D. A., & Barnes II, W. L. (2015). Workforce flexibility in operations management. *Surveys in Operations Research and Management Science*, *20*(1), 19–33.

Rahmaniani, R., Crainic, T. G., Gendreau, M., & Rei, W. (2017). The Benders decomposition algorithm: A literature review. *European Journal of Operational Research*, *259*(3), 801–817.

Roshanaei, V., Booth, K. E., Aleman, D. M., Urbach, D. R., & Beck, J. C. (2020). Branch-and-check methods for multi-level operating room planning and scheduling. *International Journal of Production Economics*, *220*, Article 107433.

Roshanaei, V., Luong, C., Aleman, D. M., & Urbach, D. R. (2017a). Collaborative operating room planning and scheduling. *INFORMS Journal on Computing*, *29*(3), 558–580.

Roshanaei, V., Luong, C., Aleman, D. M., & Urbach, D. (2017b). Propagating logic-based benders' decomposition approaches for distributed operating room scheduling. *European Journal of Operational Research*, *257*(2), 439–455.

Roshanaei, V., Luong, C., Aleman, D. M., & Urbach, D. R. (2020). Reformulation, linearization, and decomposition techniques for balanced distributed operating room scheduling. *Omega*, *93*, Article 102043.

Sadjadi, S., Soltani, R., Izadkhah, M., Saberian, F., & Darayi, M. (2011). A new nonlinear stochastic staff scheduling model. *Scientia Iranica*, *18*(3), 699–710.

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

Sen, G., Krishnamoorthy, M., Rangaraj, N., & Narayanan, V. (2015). Exact approaches for static data segment allocation problem in an information network. *Computers & Operations Research*, *62*, 282–295.

Sen, S., & Sherali, H. D. (2006). Decomposition with branch-and-cut approaches for two-stage stochastic mixed-integer programming. *Mathematical Programming*, *106*(2), 203–223.

Song, H., & Huang, H.-C. (2008). A successive convex approximation method for multistage workforce capacity planning problem with turnover. *European Journal of Operational Research*, *188*(1), 29–48.

Stratman, J. K., Roth, A. V., & Gilland, W. G. (2004). The deployment of temporary production workers in assembly operations: a case study of the hidden costs of learning and forgetting. *Journal of Operations Management*, *21*(6), 689–707.

Sun, D., Tang, L., & Baldacci, R. (2019). A benders decomposition-based framework for solving quay crane scheduling problems. *European Journal of Operational Research*, *273*(2), 504–515.

Thompson, G. M., & Goodale, J. C. (2006). Variable employee productivity in workforce scheduling. *European Journal of Operational Research*, *170*(2), 376–390.

Thorsteinsson, E. S. (2001). Branch-and-check: A hybrid framework integrating mixed integer programming and constraint logic programming. In *Principles and practice of constraint programming—CP 2001: 7th international conference, CP 2001 paphos* (pp. 16–30). Springer.

Van Slyke, R. M., & Wets, R. (1969). L-shaped linear programs with applications to optimal control and stochastic programming. *SIAM Journal on Applied Mathematics*, *17*(4), 638–663.

Verstichel, J., Kinable, J., De Causmaecker, P., & Berghe, G. V. (2015). A combinatorial Benders' decomposition for the lock scheduling problem. *Computers & Operations Research*, *54*, 117–128.

Wallace, R. B., & Whitt, W. (2005). A staffing algorithm for call centers with skill-based routing. *Manufacturing & Service Operations Management*, *7*(4), 276–294.

Wang, Y., & Tang, J. (2020). Optimized skill configuration for the seru production system under an uncertain demand. *Annals of Operations Research*, 1–21.

Wolf, C., & Koberstein, A. (2013). Dynamic sequencing and cut consolidation for the parallel hybrid-cut nested L-shaped method. *European Journal of Operational Research*, *230*(1), 143–156.

Zakeri, G., Philpott, A. B., & Ryan, D. M. (2000). Inexact cuts in Benders decomposition. *SIAM Journal on Optimization*, *10*(3), 643–657.

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.