University of Southampton

Faculty of Social, Human and Mathematical Sciences
School of Mathematics

A Mixed-Integer Model for Skill Development in a Multi-Skilled Workforce

by

Alice Lily Robins

Thesis for the degree of Doctor of Philosophy

February 2019

THIS PAGE IS LEFT INTENTIONALLY BLANK.

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SOCIAL, HUMAN AND MATHEMATICAL SCIENCES

School of Mathematics

Doctor of Philosophy

A MIXED-INTEGER MODEL FOR SKILL DEVELOPMENT IN A MULTI-SKILLED WORKFORCE

by Alice Lily Robins

Workforce planning is of strategic importance to most organisations. It is in particular challenging when tasks may require employees with specialised skills and various training activities may be proposed. In such environments, the value of dynamic skill development within a workforce cannot be ignored. We are in particular interested in increasing our understanding of how multiple skills and their development should be distributed among a pool of technicians, and how this may depend on the operational environment.

This research develops a novel approach based on mixed integer linear programming for determining an optimal strategic plan of skill development of a multi-skilled workforce when there are multiple training regimes that may be selected and various constraints on the operations of the training. For instance, additional constraints are required for skill development given uncertainty in the future demand for each type of task.

In this thesis we focus on the development of the model and its application in organisations through the development of a decision support tool. This tool provides training recommendation for employees under different training options. Each model is analysed to determine the impact of the different policies on the resultant skill gap and the run time of the model.

It is determined that the solution of these models cannot be found where the solver has difficulties proving optimality. Thus, a heuristic approach is recommended to approximate the solution. Both the exact and heuristic method are applied to a case study at Boeing for a complex maintenance line operated by a multi-skilled workforce. The problem calls for determining effective and efficient strategies for training and operational allocation of technicians.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

AUTHOR'S DECLARATION

I, **Alice Lily Robins** declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

A Mixed-Integer Model for Skill Development in a Multi-Skilled Workforce

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. None of this work has been published before submission.

Signed:

Date:

The findings presented within the case study can in no way be associated with past, current or future performance of any Boeing contract.

ACKNOWLEDGEMENTS

"It's still magic even if you know how it's done." - Terry Pratchett

This project would not have been possible without the support and guidance from many individuals. I would like to thank everyone who has helped me throughout the process and I am extremely privileged to have received this support throughout the length of the project.

I would like to thank the ESPRC and Boeing for supporting this PhD research. In particular, thank you to Boeing for giving me the opportunity to do an applicable case study and providing all the support and data I required.

I am grateful to my supervisors Professor Joerg Fliege and Dr Patrick Beullens at Southampton University, for their support and assistance throughout this process. Without their guidance and insight, this research would not have been possible. It has been a pleasure working with both on this work.

I would like to thank the Operations Analysts at Boeing for treating me as one of the team and providing advice and assistance throughout the process. I would especially like to thank John Winskill for his confidence in my ability to complete this project. Moreover, I am grateful to him and Steve Butler for their guidance throughout the years and patience with my continuous IT problems.

I am extremely grateful to my loving parents, brothers and family for believing in me and supporting me through this process but also through everything that I do. I owe my sanity to my partner Danny Wells and cannot thank him enough for his endless and unconditional support and understanding throughout the years.

Lastly, I would like to give a special thank you to my Maths teacher Chris Marchant, for introducing me to the joys of Maths and teaching me that I am capable of anything.

1. INTRODUCTION

In many industries it is essential to train employees to a sufficient level of competence in order for them to safely and efficiently carry out certain tasks. However, with a large workforce and more complex training regimes now considered, an insight into the most effective training structure would be considered beneficial.

This thesis is a collaboration between the University of Southampton and Boeing. A Training Development Model is proposed to determine the optimal allocation of employees to training to advise a workforce strategy for a selection of training policies. A generic tool is built such that a client can select the relevant training policies for their company and run the associated model.

1.1 Problem Description

Employee skill sets are influenced by the tasks they are allocated and the training they receive. Consider a problem where each member of the workforce possesses authorisations (auths), or skills, such that the corresponding auth must be obtained in order to complete a task. Thus, each task is associated with an auth.

In order to obtain an authorisation, training is required and the employee's competency must be examined. The model is devised to be strategic such that the workforce will be suitable for a long time horizon, for any task, given the company's training policies.

The motivation for this study is a case study proposed by Boeing. The company are interested in the workforce requirements for their aircraft maintenance lines. Due to the stochastic nature of events in defence, the repairs required for aircraft are not known far in advance. When choosing a suitable workforce it is important to ensure they have enough skilled employees to complete any tasks they may receive.

In order to address this case study, a generic model is proposed to solve similar workforce problems which may be faced in industry. A set of models is devised to consider the various manpower issues, including the stochastic constraints viewed in the Boeing case study.

The first model produced aims to optimally allocate the current workforce to the given tasks in order to minimise the skill gap. This is defined as the training needs analysis (TNA). If a task is not completed, the lack of skills or lack of employees can be identified. However, the needs of the workforce must also be considered in terms of their availability and skill sets. These add constraints to the allocation model. An employee is allowed to have multiple skills to define which tasks they can complete. Thus, the model is developed for a multi-skilled workforce.

Different methods of developing the skills of the workforce can be used to satisfy the perceived skill gap. These options are defined as Training Execution Methods (TEMs). The TNA and TEMs combined form the Training Model. In services such as manufacturing, maintenance and healthcare it is essential that employees have shown competency in a task before they are able to use the skill. Different policies are considered for developing the skill and measuring the competency of the employee. These policies determine the constraints in the Training Development Model. Authorisations are gained if the experience of the employee in the task exceeds a threshold value.

It is assumed that each auth will expire after a fixed amount of time, at which point an employee will need to show they have gained enough experience to regain the auth. Thus, a dynamic model is proposed. However, if not enough experience has been gained the auth will not be rewarded. A natural training method is repeating tasks. The process of repetition is used to learn a skill in most environments. This is considered as the core training method of the model and will be used with any other training policies proposed.

Repeating tasks increases an employee's experience through learning. It is assumed that the number of times a task is completed is correlated to the experience gained. If this relationship was perfectly correlated then completing the task once would define one time units worth of experience. The tasks must then be repeated until the threshold value is gained. Assigning employees to a task increases their experience whilst reducing the skill gap.

The introduction of training introduces a new objective to the Training Model to ensure the workload is equally distributed. Thus it is formulated as a multi-objective model where two objectives are considered; the skill gap and allocation fairness.

Two additional TEMs are introduced to increase experience. Here, TEM Zero, repeating tasks, can be combined with on-the-job training, training courses or both. The available training is dependent on the individual company policies.

On-the-job training (OJT) allows employees to achieve hands-on experience of the task without holding the correct auth. In manufacturing lines where the tasks are very similar this is an effective and cheap method of training.

Training courses can include a range of methods such as e-learning or residential courses. The experience gained for each employee will be dependent on the type of course. Each type of course will be associated with a fixed course duration, cost, capacity and start time, if applicable. As with on-the-job training, an employee does not need to hold the correct auth to attend the course. Training courses can be found in a variety of industries where the training may be carried out by the company or external for anyone interested in the subject, such as computer programming courses.

These three methods can be used in a combination to reflect the companies policies.

In addition to skill development through training, other criteria is considered that may affect where an employee is allocated or how the skills are developed. There are multiple constraints that can be applied to the decision support tool through model extensions. Six different extensions are presented in this thesis. A company may select a combination of the extensions to represent their requirements.

Extension One proposes an additional form of training known as compulsory training which must be satisfied in order for the authorisation to be gained. Compulsory training must be completed for the auth to be awarded, however it does not increase the experience of the employee. Thus, both the experience requirement and the compulsory requirement must be satisfied. Compulsory training could be executed as an interview to show that the skill has been learnt or by completing a specific course.

It has thus far been assumed that an employee must satisfy the same experience and training requirements independent of the previous skills of the employee. In Extension Two, the competency threshold of an employee learning the skill for the first time differs to reauthorising the skill once it has expired. It is now assumed that if an auth has been awarded, less training is required to obtain the auth in future. This would produce different requirements for compulsory training and threshold values in the previous extension.

The authorisations may also be dependent on other auths being gained. Extension Three ensures that preliminary authorisations have been completed in order to gain other auths. The dependency may be represented as a directed acyclic graph modelling precedence constraints.

Skill specialities are proposed for Extension Four such that each task may only be completed by employees of a specified level or type. Skill specialities are introduced to each task and employee. It can be used to allow tasks to have different levels or types of complexity. For instance, it can be used to define specialities of Doctors or differentiate between supervisors and technicians.

As previously stated all models consider experience through repeating tasks. It is assumed that this experience increases linearly such that the number of times the task is completed is linear to the gain in experience. In the next extension, the rate of learning is considered as a non-linear component of skill development. The learning curve is examined within our application and it's implications on the training model and results.

Finally, an extension is created to model stochastic tasks. In this way it is not known which tasks will need completing at each moment in time. The probability that each task will need completing will be dependent on the current state of the world. In healthcare, the season may affect the services that need providing to patients. For instance, in Winter a higher probability of patients requiring injections for the flu would be expected.

The stochastic nature of the tasks increase the complexity of the model. The current state of the world can create large differences in demand. The model minimises the number of tasks that cannot be completed due to lack of skills given the probabilities of the tasks occurring under different scenarios.

These training execution methods and extensions may be considered in combination to form the decision support tool. Hence, a client may choose one of 256 different combinations to accurately model their company's workforce policies and thus accurately reflect their training requirements.

Three training methods, one of which must be included in the model, and six extensions are proposed to formulate the workforce decision tool. The resultant Training Model will output the recommended training for each employee and details of their skill development. This thesis aims to look at each of the methods and determine how they can be formulated such that the policies can be combined to form one workforce model.

Each extension will present different modelling challenges. In particular Extension Five, which includes non-linear learning, and Extension Six, with stochastic tasks. A mixed-integer linear programming approach is proposed, thus a novel method is required to relate non-linear experience to the binary definition of authorisations.

The resultant training recommendations will give new insight into workforce composition. For instance, is a mixed or specialist workforce recommended? What is the impact on the total cost? Which training policy is the most effective? How does the policy determine the task allocations? This would be of commercial interest to companies who would like to adjust their policies. By considering the possible combinations of the constraints, the efficiency of the methods may be discussed.

This research aims to answer how linear programming can be used to model the training recommendations for a workforce under various training and operational constraints.

The workforce tool is applied to the case study proposed by Boeing using their training and workforce requirements. Boeing are contracted to supply the maintenance, repairs and upgrades of specified fleets using a pool of trained engineers. To effectively provide this support, Boeing must ensure their workforce are capable of satisfying the requirements of the task under their operational constraints. These constraints will be related to the TEMs and extensions proposed in this thesis. The models will consider the balance of skilled personnel required to satisfy the maintenance tasks required for the aircraft. The case study demonstrates the applicability and usability of the model within industry.

1.2 Research Questions

The aim of this thesis is to produce a generic decision support tool to compute the most effective and efficient workforce solution by recognising the skill gaps and recommending training solutions to minimise these gaps. The skills of the employees are considered as a binary value but are awarded in relation to a threshold value of experience. Hence, a novel method to model different training methods to develop skills is proposed.

Hence, this research aims to answer the following question;

RQ 1. *How can skill development of employees be optimised by translating training into a quantitative value of skill?*

The research will consider the impact of the selected policies on the allocation of employees to tasks and how, in turn, these allocations affect the employee's competency in a skill. Different policies may constrain the authorisation model and thus affect the relationship between skill and allocation. This is discussed through an analysis of each policy.

The thesis will focus on running simulations of random data for each of the policies in order to investigate different training options and combinations. Through this, an insight is provided into workforce allocation decisions and how the choice of policies impact these decisions. The model recommends the most efficient and effective means of training given the applied constraints. An analysis can determine how the choice in constraints impacts the training cost, task allocations and choice of workforce.

The resultant model is applicable to any industry where an employee can only complete a task if they have been authorised to do so. The skill development is formulated in a workforce strategy for the industry and different training policies can be included for analysis. The model allows a client to select the relevant policies to their company. A tool is created to transform the client's workforce input into training recommendations.

1.3 Organisation of Thesis

The purpose of this thesis is to explore novel approaches to modelling skill requirements of a multi-skilled workforce. The thesis is organised as follows.

First, a literature review is required to discuss ideas currently in the literature, explore opportunities and to understand each work in the context of the problem being studied here. As such, a detailed analysis will be given of papers concerning workforce allocations, training requirements and stochastic methods. Any relevant papers to the defence and aerospace industry are highlighted for relevance to the case study. A review of the literature is used to propose the method for formulating this problem. Chapter Two also contains an explanation as to how this research contributes to the current literature.

An understanding of the methods used needs to be gained. In Chapter Three, each component of the problem will be discussed in turn and possible solution methods are proposed and compared.

Given the chosen method to solve the problem, the models are proposed in Chapter Four. A detailed description of each model is given such that there are nine models in total that can be combined to form the skill development model.

In Chapter Five, data is applied to the models to analyse the behaviour of the models under different conditions. An exploration of the run time is composed, with an analysis of the cause of long run times. A sensitivity analysis is performed on relevant coefficients to determine their impact on the solution.

After exploring the models using random data, Chapter Six applies data from a Boeing case study on multi-skilled workers on a maintenance line. The solutions to these models are described in detail and their implications in terms of training needs within the company. In order for this model to be used within a case study, a decision support tool is created. The userguide for this tool is presented in this chapter.

Given the run times of the model, a heuristic approach is recommended to reduce run times. The motivation for this heuristic is discussed and the choice of method. Details of the heuristic are given and results are presented in Chapter Seven. The heuristic solutions are compared to the model solutions in order to compare run times and solution accuracy.

Finally, conclusions are drawn from this thesis in Chapter Eight. These include a summary of the models and their analysis. The answers to the research questions are provided and scope for future research is indicated.

2. LITERATURE REVIEW

To understand the contribution of the research, the current literature must first be discussed. The focus of this review is workforce strategy models, in particular looking at allocation and training. Models of skill development are reviewed and compared from various papers with a look at how this research contributes to the field. To further understand the case study for this project, military applications for the methods discussed are considered and analysed where appropriate.

In order to develop a thorough understanding of workforce management, skills are defined in detail and the methods of addressing them are compared. Applications of workforce allocation models are discussed with the associated methods. Training is examined as a core concept of skill development and again, some methods and applications of implementing training techniques are investigated.

2.1 Workforce Management

This thesis focuses on the provision of workforce resources and the relevant training. As shown in Datta, Srivastava & Roy (2013), workforce management can have an impact on the availability of the products. A lack of efficient workforce resources causes work to be redone and requires additional time to complete the job. If the job is delayed, the availability of the whole project is affected.

Two elements of the problem are considered; training need and training execution. This study analyses how employees are allocated to jobs and how the allocations affect the skills of the employees. In order to understand skills mathematically; skill classes need to be defined, the attributes of skills should be considered and modelling techniques using employee skills as the resource should be reviewed.

2.1.1 Skill Classes

A *skill* is an attribute of an employee that allows them to perform certain tasks or jobs. De Bruecker et al. (2015) defines two classes of skill; hierarchical and categorical. In the hierarchical skill class, a person is defined on a scale of how skilled they are. Thus, some employees may be more skilled than others or have a higher efficiency. Jobs require a certain level of skill to be completed, however any person with a skill level higher than the required skill level can perform the task. This is called *substitution*. The scale of skill can either be continuous or discrete.

The categorical class defines different skills but on the same level, thus the skill defines whether you can or cannot do a job but your skill level cannot be better or worse than another employee's skill level. An employee can have more than one skill to allow him to do more than one type of job, this person is referred to as being *cross-trained* or *multi-skilled*.

Some papers will consider both skill classes in one problem allowing for multi-dimensionality (Bhatnagar, Saddikutti & Rajgopalan (2007), Firat & Hurkens (2012), Heimerl & Kolisch (2010)). In this study a multi-skilled workforce is used where employees may have multiple authorisations, such that they are categorical, but are only defined by whether they can complete the task or not through the use of a skill-matrix. In order to consider skills of different levels and complexities, Extension Four defines skill specialities to address this multi-dimensionality.

Through reviewing recent literature on workforce allocation models, De Bruecker et al. (2015) defines the six most used skill determinants; age/seniority, experience, technical knowledge, qualification, job grade (nurse grades for instance) and other. Some of these are defined as hierarchical such as job grade, while others are categorical such as qualification. On the other hand experience could be considered as either.

The model should allow flexibility such that the skill types can differ for each run of the model. In regards to the skill determinants suggested in De Bruecker et al. (2015), "experience" will be used to define the technical knowledge and the competency of the employee. Skill specialities will be considered as a model extension as a method for substitution.

The proposed training model will allow cross training as described previously. The advantages and disadvantages of including this property in this model will be summarised in a later section, in addition to their implications when analysing the results of the model.

In categorical skill classes, there are advantages and disadvantages of cross training. Li & Li (2000) describes a decrease in performance when an employee is performing tasks that differ from their core role. Marentette, Johnson & Mills (2009) gives a thorough analysis of cross training, including the high costs associated with training an employee to perform multiple tasks. The cost of hiring a small number of employees with a large number of skills each should be compared with a large number of employees each with a limited skill range. Li & Li (2000) in particular considers the trade off between flexibility, efficiency and cost in their model. The decision will have a large impact on the number of available employees; Nebhard & Norman (2002) recommends a mix of flexible and specialised workers in order to minimise the disadvantages.

The problem specification for this study allows a multi-skilled workforce due to the large range in job types and demands. Thus, the negative implications are discounted for now but will be considered later when analysing the solutions.

With both substitution and cross training, there is a risk of increased turnover. The number of employees voluntarily leaving the company may increase if they disagree with the practice or are dissatisfied with their work. Thompson & Goodale (2006) discusses the problem of an unproductive workforce. They consider productivity as non-linear and model it using mathematical programming.

In order to increase flexibility in the workforce, *temporal workers* can be hired. De Bruecker et al. (2015) analyses the literature concerning temporal workers. It was found that temporal workers are considered to be less skilled and thus cost less. However, due to lack of understanding of the company, it was felt that they had a slower learning rate. It was found that most models had some constraint on the ratio of temporal to permanent employees, in this way certain tasks had to be kept internal. Overall, it was agreed that temporal workers allowed for increased flexibility, particularly with regard to high seasonal demand.

Temporal workers are not considered in this study specifically. However, by defining a temporal worker as a 'speciality', the model can be adapted to allow temporal workers to be available to complete some specified tasks.

2.2 Workforce Allocation

The first aspect of this study is the Training Needs Analysis. The workforce allocation problem defines how many employees should be hired, where they should work and which time slots they should be assigned to (De Bruecker et al., 2015). It requires an optimisation of staffing allocation over time.

The workforce allocation problem can be more complex than the resource allocation problem as each person will usually have varying skill sets and different requirements are considered. Indeed, some employees may be on part time contracts, have holiday leave or be off sick. Thompson & Pullman (2007) notes that only 8% of the 84 papers they reviewed considered the relief breaks required by workers. These conditions give constraints to the hours they can be allocated making the problem heterogeneous. Due to the strategic nature of this study, random daily relief breaks will not be included. However, our proposed model will allow for annual and sick leave as these have a larger impact on a long term strategic model.

The addition of skills to the problem also increases the complexity of the model and thus the computation time as shown by Ertogral & Bamuqabel (2008). In this paper, the model considers allocating a number of employees to types of job rather than individual workers. Here, adding skills increases the number of decision variables and constraints in the problem but is required to model the skill development.

Finally, the concept of hiring and firing employees may be adopted in order to fulfill the workforce needs. New employees can be hired in order to be allocated to jobs requiring a large workforce then fired once the demand is satisfied. This can be seen in applications with high seasonal demand. However, there is a large cost associated with regularly hiring employees, due to training, hiring and firing costs. In addition employees may leave the company for personal reasons or due to changes within the company.

Florez, Castro-Lacouture & Medaglia (2012) incorporates the hiring and firing of employees in their allocation model but minimising the fluctuations in workforce size, Fowler, Wirojanagud & Gel (2008) similarly includes the cost of hiring, firing and training in their objective function and Gomar, Haas & Morton (2002) includes penalty functions on the number of people hired, fired or swapping to a new skill. In the case of the latter, using a new skill increases their level in that skill as typical in training.

2.2.1 Workforce Allocation Methods

De Bruecker et al. (2015) defines two ways of solving a workforce scheduling problem with skills. First, the skills can be added to the constraints. Alternatively, the skills can be added to the objective function. Most commonly, skills with restrictions are modelled as constraints whereas skill based performance measures are modelled in the objective function.

In order to find a solution to the problem, many techniques can be used including mathematical programming, heuristics, discrete event simulation and queueing. Of the mathematical programming methods used, mixed-integer seems to be the most popular. For heuristics, most papers create their own methods based on the problem structure.

Four methods are suggested in Wang (2005) for solving workforce planning problems. Markov chains are used when workers can be separated into different classes based on attributes such as skill and experience. These classes are mutually exclusive such that a worker cannot belong to both classes. Thus it is unsuitable for multi-skilled workers as used in this study.

Simulation has been used to model the flow of people through various procedures. Discrete event simulation (DES) in particular is used to model a system as a discrete sequence of events in time that mark a change in state. Simulation allows modelers to incorporate all aspects of the maintenance problem by building to specific applications, thus it can deal with complexity. One advantage of simulation, according to Iwata & Mavris (2013), is the ability to implement changes at little cost in order to analyse the consequences. However, Wang (2005) notes that simulations are not optimisation tools and thus are usually used for what if scenarios rather than planning tools. They are also used to model a specific system, hence to implement new constraints or policies, new models will need creating for every option. In this thesis, this suggests 256 different simulation models will be required to model the eight training policies.

Similarly to simulation, system dynamics (SD) models the flow of items through a network. One of the most important features of system dynamics are feedback loops which allow the author to model cause and effect. The benefits and disadvantages for using SD to analyse future states after a sudden change in an organisation are explored in Winch (1999). System dynamics allows the analysis of dynamic flow networks. In military lines it is used for; forecasting responses to changes in training and recruiting, policy analysis and modelling the stages of training.

However, as explained in Garza et al. (2014), though SD can be used to model the various flows of employees through an organisation, it cannot track the individual characteristics of the employees such as the skills and skill levels. In order to do this it would need to be combined with a second method such as DES.

The final method considered by Wang (2005) is mathematical optimisation. This seems to be one of the most popular choices in the literature. Mathematical programming allows multiple constraints and multiple objectives. However, when the problem is stochastic and involves dynamic behaviour, it becomes a lot more complicated and may require a large number of variables.

2.2.2 Applications in Workforce Allocation

Though most papers on workforce allocation are applied to healthcare and manufacturing; there are a few papers applied to maintenance scheduling and a few applications in military or aircraft. Each of these areas will be considered in turn.

Carlos, Sánchez & Martorell (2011) models the maintenance planning of nuclear power plant safety equipment. In this model, both internal and external workers are considered with internal workers having different efficiency rates. The number of workers allocated to the task reduces the time the task takes to complete. Here, the decision is the number of internal and external workers to use and the maintenance interval. The problem is solved using Particle Swarm Optimisation.

Internal and external workers are also incorporated into the model by Heimerl & Kolisch (2010). Though they consider a larger range of skills for each person, again efficiency is used to define the level of the skill. The skill sets differ for internal and external employees. The cost is optimised by calculating the amount of time to be used by internal employees to complete a project using a certain skill. A mixed-binary linear program is used to solve the optimisation problem. Though external employees are not considered in our model, the cost and, in Carlos, Sánchez & Martorell (2011), the penalty associated with not assigning enough resource are noted.

Similarly, Firat & Hurkens (2012) also optimises the number of available technicians in a maintenance model. However, the technicians have multi-level skills and must remain in their groups for the duration of the day. A task can only be performed if the group assigned to it has a combined skill level above a given threshold. Tasks can be outsourced for a given cost and unavailability periods are considered for the technicians. Mixed integer linear programming (MILP) is used to solve this problem and optimise the technicians assigned to tasks and the order the tasks should be completed. In the proposed case study employees may work together within a bay, however there is no constraint on their working bay for the rest of the day. Similarly to Firat & Hurkens (2012), it is also specified that a task can only be performed if certain requirements are met. However, skill levels are used to satisfy task requirements whereas this study will consider number of assigned technicians.

To specify the skill level, Firat & Hurkens (2012) uses a binary matrix to assign to each technician. Each element in the row represents a skill and each element in the column represents the skill level. The element is equal to 1 if the skill level for that employee is less than or equal to the corresponding level in the matrix cell. In Carlos, Sánchez & Martorell (2011) the only skill considered is efficiency and this only varies between the two groups whereas in Firat & Hurkens (2012), skills vary between each worker. A similar method of defining skills will be proposed in this study where the skill matrix contains binary elements defining whether each employee has the skill or not.

For military applications two key methods are considered; simulation and mathematical programming. The difficulty in modelling military applications lies in the stochastic nature of events and thus models should take this into account.

Datta, Srivastava & Roy (2013) uses DES to determine a near optimal resource allocation of labour in the end to end repair line in a performance based environment. DES allows the whole line to be modelled and insights to be gained through experimental runs, given the stochastic elements in military repairs. The maintenance is performed over a number of lines, such that there is one for each job type (wings, engine, body). Within this, each line requires mechanical and electrical technicians. The manpower resources are selected to achieve a desired availability. The availability is calculated by counting the number of repaired aircraft at the end of each time unit, here a pulse. However, this model does not account for dynamic cross training where the employees skill ranks and trades vary with time.

The detail that can be incorporated in a DES is shown in Iwata & Mavris (2013). Here, the author models the operations and support of military aerospace vehicles. The simulation incorporates the workforce and the resources available in a multi-echelon system.

Mathematical programming also has applications in military and aerospace; Yan, Yang & Chen (2004) looks at the maintenance of commercial airliners. The problem is formulated as a MILP. The assignment of work teams to different shifts is considered, where the flights expected to arrive in each shift are known. The maintenance team are multi-skilled so can each handle a different subset of the airliners. The time to perform a maintenance check is fixed and the number of workers for each plane is fixed. Thus this is a deterministic problem analysed using the author's own heuristic.

Stolley (1969) also uses linear programming as a method to assign personnel to jobs in the German Armed Forces based on the qualifications of the individual people. The results from a series of tests are analysed and employees are thus matched to tasks dependent on their score from each test. Linear programming allows for specific constraints associated with the problem specification making it a useful tool for our model.

The next military paper is also applied to the Air Force and scheduling maintenance personnel. However, Safaei, Banjevic & Jardine (2011) only considers specialised workers, thus cross training is not considered. The paper aims to maximise the availability of the aircraft. Pre and post flight checks are required by the technicians; if a minor fault is found then it is repaired on the flight line whilst major faults require the aircraft to be sent to a job shop. The resultant mathematical program is solved using the branch and bound method.

Though Safaei, Banjevic & Jardine (2011) incorporates the stochastic nature of the repairs through probability distributions, there is no way to change the schedule once the required repairs have been realised. Additionally, the paper does not consider preventative maintenance or upgrades taking place alongside the anticipated repairs. Duffuaa & Al-Sultan (1999) uses stochastic programming with recourse to create an initial model of preventative maintenance and upgrades with emergency jobs considered, then adjusts the model once the emergency jobs have been realised. Stochastic recourse will be discussed in more detail in the next chapter.

In order to use this technique, a list of anticipated jobs is required with probabilities that they will occur at a certain time. The skill sets required for the anticipated jobs are also needed, making this problem less realistic in terms of available data. However, due to the recourse this paper is more realistic than the model presented in Safaei, Banjevic & Jardine (2011), as it is likely that workforce would be reserved for anticipated jobs. Any unused workforce is used elsewhere within the company for alternative tasks; in reality this may damage morale and efficiency of the employee.

Rather than assigning people to specific jobs, one could assign people to a resource such as a room in a hospital or maintenance bay. An example of this is seen in Datta, Srivastava & Roy (2013), where each event in the simulation is a different maintenance bay. Each bay will have different skill requirements at different times and may require more than one skill at the same time. Geerts & Vliegen (2012) considers a similar problem for hospital examination rooms in their Master's Thesis. The problem uses a MILP to satisfy the demands of the rooms using a mixed skilled workforce.

The rooms may require more than one skill and more than one person during each shift. Here, it is assumed that the demand for each room is deterministic and job lengths are less than a whole shift. Any additional free time during each shift is used for alternative tasks. In the model by Datta, Srivastava & Roy (2013), the job lengths are stochastic and determined using distributions. However, the assignment of workers to the tasks in each pulse is not done optimally, whereas assignments to tasks during shifts in Geerts & Vliegen (2012) is done optimally as the model can be solved using the Simplex method.

Bellenguez & Neron (2005) use MILP to allocate a multi-skilled, hierarchical, workforce to multiple projects. The scheduling of the jobs is taken into account however the details of the activities are not stochastic. Kuo, Leung & Yano (2014) also avoids the stochastic nature of the problem but includes availability of their multi-skilled, hierarchical employees. Chen, Liu & Wang (2014) use multi-skilled employees with project scheduling, in this case the projects are associated with a probability of occurrence thus are stochastic. The model does not include hierarchical skills or the availability of the employees.

Scheduling appears in a variety of other multi-skilled personnel papers with many applications. Tiwari, Patterson & Mabert (2009) applies MILP to the telecommunications industry, Valls, Pérez & Quintanilla (2009) considers multi-skilled, hierarchical employees and their availability in generic Service centres and Heimerl & Kolisch (2010) allocate a multi-skilled, hierarchical, workforce to multiple projects in the IT industry.

Stochastic demand, multi-skilled and hierarchical skills are considered in Gutjahr & Froeschl (2013) paper on project portfolio selection. The paper features many similarities with this study. Here a set of projects is chosen from a larger set to be completed within the time limit. Similarly to Duffuaa & Al-Sultan (1999) a recourse model is used to model stochastic demand.

Allowing scheduling of tasks ensures realistic results for short time horizons. In this study, a strategic model with a large time horizon is used. Thus, it may not be required to schedule tasks before allocating the workforce. The schedule will define the experience of the employees and, in turn, affect the future schedule, meaning allocations could not be made at the same time and a linear programming model would not be suitable. A dynamic allocation would be better but less strategic.

The training needs analysis produced in this thesis will not vary from those seen in the literature. Common components of workforce planning seen in the literature include multi-skilled workers, cross training, hiring and firing, stochastic demand, scheduling of tasks and temporal workers. Though, examples of multi-skilled workers, cross training, stochastic demand and availability constraints have been given, the skill development in this thesis through training should be considered and will define the contribution of our work.

Overall, a broad range of techniques have been applied to workforce allocation problems. There are many aspects that can be modelled that affect the technique used. Models may consider cross training, relief breaks and stochastic jobs. As defence is heavily affected by the stochastic element of emergency jobs, methods associated with stochasticity have been described in more detail. The next section will consider the second part of the model; training of the employees.

After reviewing the methods used in these workforce allocation models, a mixed integer linear programming approach has been determined as the choice method to use in this model due to it's ability to model all eight policies within one framework, rather than simulation which would require a new model for each constraint set.

2.3 Training

In addition to determining the number of workers, when they should work and where, the training of the workers should be incorporated in the model. Training could refer to cross training to learn new skills, initial training or training to improve their knowledge in a certain skill. Cross training is explained in more detail in Section 2.2.2. An example of the importance of training can be found in Cohen, Dearnaley & Hansel (1956) in regards to bus drivers. The cost of the training should be considered as well as scheduling when the training should take place. As mentioned above, cross training is considered throughout the literature with Li & Li (2000) being the only paper to include the disadvantages as well as the advantages in a planning model.

Another component to consider is training on-the-job against external training. Though training on-the-job is considered to be the cheaper option with better links to the job, du Boulay & Medway (1999) discusses why this is becoming more difficult to perform. In some jobs it is against the law to perform the job without required training. Additionally, introducing a less skilled person to the workforce could increase the time it takes to complete the job.

De Bruecker et al. (2015) points out that the literature lacks some important properties of training. These include the consequences of training such as increased flexibility and the resultant absent time while an employee is being trained. The paper also recommends more attention is paid to the negative consequences of cross training.

Though training can be used to increase an employee's knowledge, learning can also be considered. This occurs naturally through performing tasks multiple times. It can also be forgotten through lack of use. Most papers consider learning as a contribution to worker's skill level and few papers consider deterioration due to forgetting (Attia, Duquenne & Le-Lann (2014), Heimerl & Kolisch (2010), Wu, Huang & Lee (2011)).

2.3.1 Skill Competencies

In services such as maintenance, software engineering and healthcare it is essential that a person has showed competency in a skill before they are able to perform a task. One may look at the criteria that can be used to measure an employee's competency in a skill and the methods to gain them. Though many papers on training can be seen in the literature, few papers address policies to measure an employee's competency. It is generally assumed that one session of training is required for a skill to be authorised.

The literature was searched for references to certifications, competencies and skill authorisations. Though many papers on workforce problems consider using certifications or authorisations to assign workforce allocations, none consider how these workforce allocations affect the development of the certifications.

The search was limited to papers in the field of mathematics, decision science and computer science. Though other papers may be mentioned where applicable.

Many papers on certifications or authorisations look at frameworks for determining competencies and their related training and assessment criteria. This training is considered qualitatively, rather than looking at how experience is gained quantitatively. The need for these certifications has been analysed.

Ascertaining the base competencies has been completed in many fields. In psychology, Fouad et al. (2009) discusses research on competencies required to train psychologists. The essential competencies are determined qualitatively and training requirements are discussed. It is determined that competency training should be individual as each person may require more time on one type of training to reach the required level of competency. Base competencies are also identified for software engineers in Thurner, Axel & Andreas (2014) and automotive engineering in Park & Cha (2013).

An analysis of assessment tools are seen in many healthcare papers. Rekman et al. (2016), Jepsen et al. (2016) and Shin et al. (2016) all develop tools to test the competencies of staff during training. The analysis can be used to determine whether current training is sufficient to give employees authorisations. However, the competencies are only developed through training rather than developing on-the-job.

Other fields also consider new training programme initiatives that use certifications and assess its acceptance and impact through the results of surveys.

Marozzi & Bolzan (2016) produced a statistical study of training requirements in relation to a skill, such as seen in this research specification. A list of skills and associated training were considered. Surveys were carried out to rate the importance of the skills and training but do not discuss how training may be done or how to confirm the training is sufficient to obtain the skill.

Mashaud et al. (2010) proposes a study to determine the required pass rate for surgical residents undergoing proficiency-based training and retention tests. Similar to this thesis, authorisations must be reassessed after a certain amount of time. This paper reviews a training regime and assessed the skill retention at the end of the training period. Where this paper analyses measured output from a system, thus case specific, the training and experience will be modelled mathematically for this thesis.

Though measuring competency is not shown in many papers, training can be used as a method of gaining competency for a specific skill and thus recorded as a measure.

2.3.2 Training Modelling Methods

Similarly to workforce resource allocation, a wide variety of methods can be used to model the training of employees in an organisation. One of the most used methods, seen throughout the literature, is simulation. Simulation and SD are used to model the flow of people through various training procedures and are also useful for calculating the cost of training. However, as stated previously, it would be difficult to track the individual skills of the employees through the training process and make changes to the system.

Balinsky & Reisman (1972), Caputo (1969) and Mooz (1969) describe the training process as a network. They base their models on the flow of employees through the network and use optimisation techniques such as dynamic programming to determine the number of employees to enter their network to achieve the required workforce. As with SD, the individual skills cannot be tracked, however unlike SD and simulation, the problem is optimised rather than being solved through scenario creation. Additionally, the stochastic nature of the problem is captured in the network due to the uncertain transition rate between nodes. A Markovian process is usually used to model the transitions under the assumption that the current state is independent of previous states.

Many other papers use some form of linear programming to model the training methods. These have been used to determine how many new people to hire and train, the courses that individual people should attend, or the number of people to send on a specified training course. However, with very large numbers of employees, the number of decision variables required to assign specific people to specific training is very high. It is also likely that training courses will be scheduled in advance, rather than people being trained individually as required. Through different policies, a mixture of training types may be proposed, some with fixed times and others taking place individually on site.

A novel technique by Darmon (2004) uses inventory theory to determine the hiring, or reorder, point for hiring a new person in order to replace one that has left. Though in this case the leaving times (voluntary or involuntary) are deterministic, they explain that the model can be adapted to the stochastic case in order to minimise the number of times the sales team is understaffed or overstaffed.

In terms of scheduling the training events, heuristics can also be used. This can be seen in Nooriafshar (1995) and Rezaei et al. (2012).

Though simulation is a popular choice, it is not appropriate for this study. In order to obtain one tool easily adaptable to any company's training policies, a linear programming model is required such that policies can be presented as additional constraints and decision variables.

2.3.3 Applications in Training

Some of the applications of training models shall be reviewed in more detail.

Balinsky & Reisman (1972) models the flow of students from education to the workplace. The corresponding model optimises the number of people to put into education in order to obtain the required workforce for each output level. The model can include drop-outs, repeating courses and people re-entering the system later in the planning horizon.

Linear programming has also been used to model training in the construction industry. Srour, Haas & Morton (2006) looks at the problem of hiring and training a multi-skilled workforce. The workforce must be trained before the start of a certain project where the skilled demand is known. However, this model does not consider scheduling of training, assuming the training is done on-the-job or out of hours. Hence, contrary to this research, training courses do not affect the daily tasks of the employee.

Juang, Lin & Kao (2007) considers the needs of individual employees and determines the most useful courses to assign to the person given their available time for training. The aim is to fulfil the training demand through the course attributes. Similarly to Srour, Haas & Morton (2006) the scheduling of the training is not considered, however in this case the training is known to take place externally at a certain time with a certain capacity.

The stochastic nature of training can be regarded in LP models. Kim et al. (2012) considers future, unknown, jobs in the IT industry and forecasts the demand. From this, decisions are made regarding the number of employees to hire and train at the beginning of each year at a cost. If the number of employees is less than the realised demand, external workers can be hired at an additional cost. Here, the optimal mix of strategic and just in time workforce planning is analysed. In this problem, all workers have the same skill sets and can be assigned to any job. However, note that training takes place before the beginning of the time horizon, thus not affecting wokforce allocations.

Darmon (2004) also creates a stochastic model for determining when to hire new employees. As discussed above, this paper applies inventory control to a sales team in order to determine the hiring point in order to replace employees who will leave. The company does not want districts being unmanned but also wants to avoid districts with additional sales people. Hence, a balance is required.

Scheduling of the training sessions is also important in healthcare applications. Nooriafshar (1995) applies heuristics to determine when employees should be trained on-the-job. Here, the trainees count towards the required employees working that day. However, more trainees than necessary should not be supplied as there may not be enough supervision. Rezaei et al. (2012) also applies heuristic methods to schedule training. Here, the paper minimises the gap between when a person is needed for a task and when they are available and adjusts their training accordingly. In addition, training programs are not run for individuals but for groups of employees. Though in these papers the training affects the allocations in terms of time of scheduling, the training does not affect their allocations within the time horizon.

A few papers consider applications to military training. Marentette, Johnson & Mills (2009) creates a cost/benefit ratio for comparing cross training to specialised, though a mixed workforce is not considered. A case study is performed on the US Air Force to determine the benefit of cross training Air Force officers. The method identified potential cross training opportunities.

Thomas et al. (1997) uses an SD model applied to military training to show how adjustments in policy impact the number of available personnel. Wang (2007) also applies SD to military to evaluate a four rank workforce training management system. Indeed, Hafeez & Abdelmeguid (2003) shows how different aspects of workforce modeling can be combined through SD. This paper analyses the relationship between training, skills, recruitment and knowledge.

Groover (1969) applies simulation to analyse the procurement, training, retraining, promotion and retirement of personnel within eight military personnel systems in the US. Through simulation, the characteristics of the individual people, or entities, can be recorded. The requirements are compared with the inventory at the end of the run.

Mathematical programming approaches can also be used for military training, as shown in Caputo (1969), to model the training of pilots in the US Department of Defense (DoD). Here they determine the maximum number of pilots needed at one time, assuming pilots would need to be "rotated" during war scenarios to reduce combat exposure time. Rotated pilots are required to train new recruits and a supplement of pilots is required to be carried during peacetime in preparation for wartime. As with Balinsky & Reisman (1972) and Mooz (1969), the model is represented as a network with equations modelling the flow. Again, the individual attributes of the pilots are not shown. The model is used to calculate the wartime requirement for pilots and thus the number of pilots to train given the supplement of pilots.

Clough, Dudding & Price (1969) and Mooz (1969) use mathematical programming to model pilot training as well. Here, transition proportions between different states are used; training states, transport and different roles. Again as this is a Markovian process, the attributes of each pilot are not measured. This is similar to the method seen in Balinsky & Reisman (1972).

An additional quality analysed in Morgan (1969) is the renewal of contracts after a set number of years. This paper applies mathematical programming to determine the recruiting target for the year and the quota control on the number of people who can renew their contracts.

In this thesis the training process is designed as part of the workforce allocation model. A few examples of this are seen in the literature. An, Jeng et al. (2007) uses SD to model the training of personnel combined with project management such that personnel are also assigned to specific projects as they become available. However, employees have specialised skills in this model rather than multiple skills.

Liu, Yang, Li et al. (2013) applies a three stage heuristic to model the training and assignment of workers to *serus*. This is a work-cell based manufacturing technique where workers are assigned to workstations where they perform multiple tasks. Therefore the efficiency of each employee does not affect the speed of the whole production line. In this setup there are no specific projects that are allocated to the workforce but rather workers are assigned to stations where they have an unlimited stream of identical tasks. Thus, all training is performed before the start of the production process.

Finally, Dietz (1991) also considers the training and allocation of workforce to tasks. Here, personnel are assigned to various maintenance tasks on military aircraft, where the aircraft are considered identical. However, personnel are considered to be single skilled, thus the problem can be solved using queueing theory and training is only considered in terms of costs and number of people to train.

Though many papers have considered training requirements for employees as a model, there are few papers that consider the individual training regimes and their impact on the task allocation of the employees. Task allocation and training allocation will be incorporated into this study and the relationship between them will be reviewed. This relationship is dependent in both ways such that allocations affect training and training affect allocations in a dynamic manner. This could not be found in the literature. Unlike some papers in this review this study will also consider different types of training, some with fixed times rather than scheduled, others are performed on site at any time.

2.4 Learning and Forgetting

In addition to training techniques, learning may also be experienced through repeating tasks which increases an employee's experience when skills are used regularly, and forgetting, which decreases an employees experience when they are not used.

An important concept investigated in the literature is the *Learning Curve* (LC). This was first introduced in economics by Wright (1936), who produced the first mathematical model of the LC and modelled the affect of learning on the production costs in the airline industry. For this study, the rate of learning is modelled through task repetition; the more time a task is repeated, the more skilled an employee is.

Anzanello & Fogliatto (2011) state that a LC is a mathematical description of a workers' performance in repetitive tasks. In this thesis, tasks may not be repeated as frequently as on the manufacturing lines discussed in these papers. However, they are still repeated over the time horizon. Therefore, though a level of learning is expected, it should produce different results than the learning curves in the literature where the tasks are identical. An overview of learning curves will still be included in this research in order to understand industry's view on learning and how experience can be quantitatively be included in modelling. Both the Learning Curve produced by Wright (1936) and the learning curve used in this paper will use the xth completion of a task as the dependent variable. The output, however, will differ between the two models.

In general, to measure the workers' performance, the LC will include; time to produce a unit, number of units produced per time interval, cost to produce a unit and percent of non-conforming units. The parameters are estimated using non-linear regression. The LC can then be used to estimate the time to task completion, the product's life cycle, the effect of interruptions on the production rate and to assess production rate in response to policy changes. For this study, the learning rate will be used to define an employee's competency in a task and in turn assign them to tasks.

A few of the models of learning and forgetting are introduced for comparison. Initially it was determined that the time to perform a task declined at a decreasing rate as experience with the task increased. This can be viewed as a power-function where the x axis defines experience in terms of the number of times a task is completed and the y axis defines time to complete task. By reviewing experimental data and finding the best fit, Wright (1936) created the theory of the LC. The learning curve as proposed by Wright (1936) can be written as follows:

$$T_x = T_1 x^{-b} \tag{2.1}$$

where T_x is the time to produce the xth discrete unit and b is the learning curve exponent. This *learning rate* is defined as the rate to double the production output. Though Wright (1936) took this learning rate as a fixed value of 80% for all industries, it has been further researched to determine different values. As reviewed by Jaber (2006), for industrial engineering problems the value seems appropriate, however other applications have produced different rates. Many early papers determined different aspects that influenced the learning rate such as industry type (Conway & Schultz, 1959) and machine or manual labour (Hirshmann, 1964). This rate would be expected to differ for the proposed case study as each task is considered unique.

The formula for deriving the learning curve has also been discussed with many variations proposed. Table 2.1 summarises the development of the earlier models discussed in the literature.

For all models discussed in Table 2.1, the processing time for tasks is taken as variable. However, this thesis proposes experience quantitatively as the dependent variable as opposed to the processing time.

Author	Model	Details
Wright (1936)	$T_x = T_1 x^{-b}$	Defines the relationship between experience and time to produce the xth output.
Carlson (1973)	$T_x = T_1(x+B)^{-b}$	Assumes that the equivalent of B units have already been experienced.
DeJong (1957)	$T_x = T_1(M + (1-M)x^{-b})$	Where M is the factor of incompressibility, taking into account whether jobs are manual or automated.
Carlson (1973)	$T_x = T_1(M + (1-M)(x+B)^{-b})$	Combination of the previous two methods.
Glover (1966)	$\sum y_i + C = a(\sum x_i)^m$	Where x, y can represent time or quantity, a the time of the first cycle, and m is the index of the curve equal to $1-b$ for each individual i.
Thomopoulos & Lehman (1969)	$T_x = T_1 x^{-b} Q$	Model for mixed-model assembly lines, where Q is the added increase in assembly time required to process the units over an equivalent single model line.
Levy (1965)	$R(x) = P(1-e^{-\lambda x}) + (x^{-b}/T_1)e^{-\lambda x}$	$R(x)$ is the rate of production, P the maximum rate of output and j the cumulative output. The expression is multiplied by a damping factor $e^{-\lambda x}$.

Tab. 2.1: Table of Models for the Learning Curve

From the initial model by Wright (1936), there have been many developments to the LC. In addition to skill learning, one may investigate skill lose through deterioration or forgetting.

Internal and external factors can affect an employee's knowledge deterioration. Internally, an employee can forget skills due to lack of use whereas externally, advances in technology could results in taught methods becoming obsolete. Though there has been a lot of interest in the learning aspect of on-the-job training, few papers couple this with the consequences of knowledge deterioration.

Anzanello & Fogliatto (2011) defines four main learning curves; log-linear, exponential, hyperbolic and multivariate models. The log-linear model is used to define the model proposed by Wright (1936). It is the most used model for repetitive operations with manual-based operations. The flexibility of this model has allowed modifications to calculate different production factors. It can also be combined with other models, such as the Product Life Cycle Model, to build integrated production planning models.

Anzanello & Fogliatto (2011) analyses some limitations of the log-linear model proposed by Wright (1936). It was noted that with many repetitions of tasks, the production time reduces to zero, thus Hurley (1996) and Eden, Williams & Ackermann (1998) discuss the inclusion of a constant in the model. Additionally, Globerson, Levin & Shtub (1989) notes that the model does not take into account workers' prior experience. Finally, the model does not consider multiple tasks types, thus Yelle (1976) proposes a summation of LCs with different learning parameters. However, it is argued that this model leads to imprecise production rate estimates.

The log-linear approach can be integrated with scheduling techniques to determine the impact of learning on the position of jobs in a schedule as introduced in Biskup (1999). This approach has been expanded to include job dependent LCs, processing times dependent on the LC (Okolowski & Gawiejnowicz, 2010) and other scheduling scenarios.

Exponential models may also be proposed. These require a more complete set of parameters and thus give more precise estimates of production rates (Nembhard & Uzumeri, 2000). Many variations of this model are proposed including the 3-parameter, 2-parameter and Constant Time models. Many papers analyse the ability of these models to accurately fit the data set given different problem characteristics. For details of exponential models the reader is referred to Anzanello & Fogliatto (2011).

The idea of "learning from experience" is mentioned in Mincer (1962), it discusses an investment in the employees, where an individual starts at a low pay bound and gains from the experience learnt on-the-job through task repetition. However, he also notes the experience gained will be subject to the characteristics of the employee such as motivation and ability.

Learning is used to show an increase of an employee's skills. This study will combine learning with the allocation of employees to tasks. According to Pinzone et al. (2016), little research has been done in this area even with regard to learning in linear models. A few prominent papers are addressed here.

Nembhard (2001) is one of the earliest papers to consider allocations influenced by learning. In this paper, the individual learning rates of employees is considered such that employees who learn quicker are assigned to tasks with shorter production runs.

Sayin & Karabati (2007) also notes the small amount of research in this field and commits a model of workforce allocation and skill development. The model is defined as a two stage MIP where the first stage assigns employees to departments and the second stage maximises their skill development. The stages are repeated for the length of the time horizon where the skill level of the employee changes due to their assignments at each step, however the model is described only for each time unit.

For the proposed model, the incremental skill level is defined as a non-integer variable which increases if an employee is assigned to a task. However, from this the true skill level is determined as an integer value. The increase in skill is only caused by repeating tasks and only one skill is considered. The model presented in this thesis should consider alternative means of gaining skills given the multi-skilled nature of the employees.

The model given by Sayin & Karabati (2007) incorporates forgetting by noting when an employee last completed a task and allowing skills to develop at a decreased level dependent on this result. Note that this model is non-linear and considers only deterministic tasks.

The paper by Azizi & Liang (2013) combines the job allocation, training and learning components of the model to assign a multi-skilled workforce to manufacturing cells. Each employee performs one task in each time unit, the allocation changes such that the skill level for that task increases with use but decreases when they are not using it.

Skill levels are constrained by an upper and lower bound and each employee must perform all skills during the planning horizon. The objective is to minimise cost of training, flexibility and productivity lost due to lack of skill. In this model it is assumed tasks constantly require fulfilling such that the details of the tasks are not required but the models ensures the required skill levels of the employees is retained.

The learning curve is used to define the skill level of the employee. Through repeating tasks the skill level increases and if the skill is not used for an amount of time the skill level decreases. Tasks require a certain skill level to be completed.

This model shows many similarities to the proposed model in this thesis. Experience is gained through repeating tasks, however the employee does not need to be authorised in order to work on a task. Once the correct level is achieved through training, the employee can complete the task in that time unit, rather than waiting for the skill to be reviewed. Training is required only if the employee's skill level is not high enough for the allocated task. There is also a constraint on how much training an employee can do. The training model in this thesis considers other methods of obtaining skills and modelling the training logistics.

Rudek (2014) considers the learning curve and its affect on the processing time of tasks and thus the optimal allocation of the workforce. Here one employee is considered at a time with the fixed jobs they can complete. In this way the computational complexity of the problem is evaluated. It was concluded that makespan minimisation problems including learning are NP-hard.

Finally, Pinzone et al. (2016) produces a model to assign workers to tasks based on their knowledge, preferences and task demand. As with Azizi & Liang (2013), training is provided in the form of a skill gap and is given as a fixed cost and fixed skill increase for all tasks.

However, in contrast to Azizi & Liang (2013), the skill qualification increases linearly with each assignment and thus does not use the learning curve. The characteristics of the tasks are provided to define the employee's preferences for certain aspects of the task. It is solved using Quantum Theory Principles.

In the study, the requirement for different types of training is highlighted as well as future research on skill types and uncertainty. The author advises future research to focus on real life situations to analyse the short and long term implications.

The complexity of the tasks has been shown in Nembhard (2000) to have a large impact on learning and forgetting. The paper concludes that there are significant differences between experienced and inexperienced workers under different methods, machines and material. Thus, one would expect the learning curve produced from the case study to differ from the examples shown in the literature. This will be due to the large differences between tasks and the frequency in which they are performed. Experience gained will also be modelled, rather than processing time, by number of times the skill has been performed. Hence a new approach to the learning curve is proposed.

Overall, it can be seen that few papers in the literature combine the assignment of individual workers to their learning behaviour. Nembhard & Norman (2007) explains the difficulty in collecting detailed performance data for individuals. Instead most papers consider assignment with no learning or forgetting.

2.5 Research Contribution

This project was designed to contribute to current research relating to workforce allocations and training. The literature related to the subject has been examined and potential gaps were discovered. In this section, the gaps are readdressed using the results of this thesis.

In the previous sections, the literature for a model of training need and training execution has been examined. However, it is determined that though the two parts have been formed individually, there are not many papers that consider the relationship between the two. Training in the literature was mainly completed as one-off events, before the jobs are allocated. It does not dynamically influence the allocation of employee to tasks.

Thus, this thesis aims to propose a model to analyse where the gaps in the skills sets are, and concentrate on the optimal training strategy to gain skill competency and procure authorisations, thus combining training need with training execution. Training will be considered taking place during the planning horizon, thus affecting time that can be spent on tasks. The training and task allocations will be assigned simultaneously such that the training allocations determine the tasks able to be completed and the assignment to tasks determine the experience of the employee.

There are papers with dynamic training courses similar to the given problem specification, shown in Nooriafshar (1995) and Rezaei et al. (2012), who use heuristic approaches. However, they do not differentiate between each employee so do not consider their individual training needs. Papers that dynamically include training also tend to group employees together using discrete event simulation. Thus, individual employee characteristics are lost or cross training is not possible.

Here, the research will consider the individual training requirements of a multi-skilled workforce through the use of linear programming. Each employees skills are considered as decision variables and are tracked with time, making them dynamic.

In industry, it is common for learning to be assessed such that the competence of an employee indicates whether they can or cannot complete a task in future, such that skills are binary. This combination of non-linear learning and binary skills is not addressed in the literature. It was found that learning in the literature does not affect an employees authorisation to complete tasks but allows them to be more competent in their job and thus may complete the task in a shorter time. In addition, learning is not available through training events and the skill is considered as a nonlinear variable in the form of employee experience.

The learning curve is used as a training tool in Azizi & Liang (2013). Though gaining experience through repeating tasks is shown in their model, it does not relate the experience gained to the skill-matrix. The experience is considered as a non-linear value.

In this thesis, a novel skill structure is proposed within the workforce. Skills define the authorisation to complete a task and as such are binary values. This means employees must be signed off for a large variety of tasks that can vary from opening a hanger door to performing intricate repairs on engines. This dependency is not explored in any literature reviewed. The skill-matrix is applied as a means of having employees trained to complete specific tasks.

A model is proposed to satisfy this gap in the literature. The model presented in Azizi & Liang (2013) will be combined with the skill-matrix such that the learning determines which skills in the skill-matrix are awarded. Training execution methods shall be used to define how experience is gained and other constraints that may affect allocations to training or tasks. The learning curve is applied as a method of computing the skill of an employee from the number of times they complete a task. Non-linear experience is gained through training and if the experience exceeds a threshold amount, the employee is deemed competent and is awarded a binary skill. The model recognises these binary skill gaps and recommends training to satisfy the gaps using a non-linear measurement of experience.

As discussed in De Bruecker et al. (2015), the literature lacks some important properties of training. These include the consequences of training such as increased flexibility and the resultant absent time while an employee is being trained. The paper also recommends more attention is paid to the negative consequences of cross training.

A decision concerning the training and workforce allocation for each employee will be produced, subject to availability and stochastic constraints. This introduces new complexity and thus realism to the modelling of optimal workforce allocation models.

By applying linear programming, many variations of the model may be applied and include different training methods and constraints that have not been addressed in models seen previously. Different training methods and constraints are proposed to satisfy the different training options available in any workplace. Multiple training tools are allowed such as e-learning and external training courses and the different methods are compared. Other aspects such as authorisation dependency and recurrent skills are also investigated as new elements to include in training models that have not currently been investigated.

A sensitivity analysis is performed on the training execution methods. It shows the impact of the different constraints on the solution of the problem. For instance, the skill gap is shown to decrease as new training methods were introduced. Conversely, adding constraints on skill speciality of auth dependency increased the skill gap. The model in this thesis allows an analysis of the relationship between these factors that affect an employee's assignment to training.

For example, the case study shows how cross training is beneficial to companies where the types of tasks outnumbers the number of employees. The optimal solution included cross training of employees. The fairness is reduced through this method as tasks could be evenly distributed amongst a multi-skilled workforce. The individual properties of the employees indicate the impact of cross training on the employees.

The cost and impact of the different training constraints on the skill gap can be analysed using the model in this thesis. In addition, other aspects of the workforce strategy can be included. Skill speciality and authorisation criteria are not analysed in the literature with respect to training.

Pinzone et al. (2016) proposed that future research in this area should differentiate between hierarchical and categorical skills. Extension Four may be used to include hierarchical skills into the model where specialities define the level of the skill. Categorical skills are included in the multi-skilled nature of the problem.

The paper also indicated a gap between simulation and optimisation when applied to skill development. It should also consider uncertainty, training and the effect of these choices on the performance of the model. Simulation is used in the stochastic element of Extension Six to approximate the solution to the chance constraints. Monte Carlo simulation is used to produce random sets of data to input in the model to determine the behaviour of different training options.

The model will also be applied under stochastic tasks such that the probability of the tasks occurring affects the allocations and the resultant training of the employees. Different training options are included through the training methods and form the key element of our study. The results of this study analyse the effect of the training, uncertainty and skills on the performance of the models and the optimal workforce structure. The run times of the model are produced given the different model inputs and the skill gap is calculated for every model variation.

Finally, the implications of the modelling approach should be addressed through a real world case study. The model is applied to a case study at Boeing, a tool is produced to run the model in a real world application and the results of the case study are analysed to determine their implications to the company.

2.6 Summary

Given the topic of this thesis, a literature review of multi-skilled workforce problems has been discussed. The key papers are mentioned below with their relevance to the model in this thesis. Though these papers may not all be influential in their line of research, they show an aspect that is not seen in the general literature and are particularly related to the model built in this thesis. They also demonstrate key aspects in the literature and define components that need to be considered.

Cross-training is seen as an essential component of the models that will be studied. With a diverse set of tasks such that there are more tasks than employees and uncertainty in the tasks that may occur, a multi-skilled workforce will be required to allow flexibility in the allocations.

To determine which skills require development, employees must be allocated to a set of tasks. Thus, determining the skill gap subject to constraints on availability of employees and other requirements. As seen in the literature, there are many techniques to model resource requirements in an allocation model. These techniques include mathematical programming, heuristics, discrete event simulation and queueing.

After considering examples in the literature, it was determined that heuristics are used in less generalised cases; usually used in a specific problem instance. DES allows the model to be built in fine detail but does not offer an optimal solution (Wang, 2005). System dynamics, similarly to discrete event simulation, also allows for a detailed model but it cannot be used to track the individual characteristics of the employees. Neither of these models can be easily adapted when consistently changing constraints. Finally, mathematical programming has been used for many workforce allocation problems; most often as a combinatorial problem. However, the problem sizes may become quite large due to the number of variables.

Overall, it was decided that mathematical programming would be the method used for this problem formulation. It allows for a generalised model with the easy addition of new constraints subject to the requirements of the company.

Three papers are highlighted that consider workforce allocation problem instances similar to this study; Duffuaa & Al-Sultan (1999), Safaei, Banjevic & Jardine (2011) and Geerts & Vliegen (2012).

However, as well as determining the allocation of the workforce, skill development can be executed through training methods. Here, repetition of tasks, training on-the-job and training courses will be considered. Similarly to the allocation problem discussed, the methods used most commonly in the literature are DES, SD, graph theory and mathematical programming. For similar reasons mathematical programming is chosen to model the training allocation. It is noted that most training models use SD and do not follow the individual characteristics of the employees. Most papers also model training as a process that does not affect the allocation such that the training takes place out of working hours. This is proposed as unrealistic and thus the thesis will allow a model to show training and allocations simultaneously.

The learning curve proposed by Wright (1936) was introduced in this chapter as a link between training and skill allocation. Many papers have explored the use in industry and expanded on the technique. However, the problem instance here differs from those seen in the literature. Here, skills are considered as binary, thus the learning curve cannot be directly applied to predict the skill level. A linear problem would be beneficial due to the changing constraints whereas instances of the learning curve tend to be non-linear. Finally, the tasks in this problem are not repetitive as seen in production lines, thus it is expected that the learning curve will have different characteristics.

Few papers have been shown to combine both the learning curve with the idea of task allocation; Nembhard (2001), Sayin & Karabati (2007), (Azizi & Liang, 2013) and Pinzone et al. (2016). The research performed by Azizi & Liang (2013) has the most in common with the proposed problem formulation. However, the employees do not need to be authorised in order to work on a task and each task must be completed once by each employee within the model. Skills are also considered as non-binary.

This research aims to bridge the gap between binary skill development and task allocation using the skill-matrix. Thus, the learning curve may be associated to a binary variable indicating an employee's ability to complete a task. Hence, this thesis will introduce a dependency between allocations and learning. The proposed model will incorporate learning to increase the experience of the employees and some criteria to determine the skill-matrix from the experience.

The work by Azizi & Liang (2013) is expanded to show how training is completed and other constraints on the skill development are considered. The proposed model will also be applied with stochastic tasks such that the probability associated with certain tasks affects the ability to develop the skills.

3. REVIEW OF MILP METHODS

In the previous section, literature for many workforce allocation and training models has been discussed. From discussions of methods used in various papers, it was decided that the method used to create the Training Development Model would be mixed integer linear programming.

In this way, many formulations may be combined to form one problem. By reformulating the more difficult components of the model, say stochastic and non-linear, into a mixed integer formulation, one technique may be used to solve any variation of the problem specified components. This allows the model to be more generic and versatile than other methods explored previously. MILP allows constraints and variables to be easily incorporated into the model without reformulating the entire model for each combination of scenarios.

In this section a study of the different methods for solving mixed integer linear programmes is conducted. In addition to the basic approach to MILP, other modelling techniques are discussed such as stochastic programming and multi objective programming, and their solution methods explored.

3.1 Model Structure

This thesis aims to create a model of the training and allocation of workers to a set of tasks. There are two model components; the workforce allocation model (TNA) and the training execution model. Combined these form the Training Development Model and can include a selection of extensions. The completed Training Model will be used to determine the optimal training assignment.

The individual characteristics of the people in the system will be required to analyse their skill sets. Thus, discrete event simulation or system dynamics will not be used to model the flow of people. System dynamics does not allow the analysis of individual characteristics and DES can be challenging to model the training of a multi-skilled workforce when the skill sets are dynamic. In addition, neither method can be easily adapted dependent on the combination of constraints. Hence, mathematical programming will be used to optimise the training and task allocations of each employee. Due to the combinatorial nature of this problem in terms of allocation, mixed-integer techniques are proposed.

A MILP is presented as

$$\min_{X} \quad F(X) \tag{3.1}$$
$$\text{s.t.} \quad f(X) \leq 0 \tag{3.2}$$
$$X \in \mathbb{R}^n. \tag{3.3}$$

Here, X is set of decision variables of dimension size n, $F()$ is the objective function and $f()$ are the constraint functions.

3.2 Mixed Integer Linear Programmes

Integer linear programming (ILP) is used over a wide area of Operational Research. Firat & Hurkens (2012), Yan, Yang & Chen (2004) and Safaei, Banjevic & Jardine (2011) represent a small number of papers that use ILP to solve allocation models. Other applications include the knapsack, travelling salesman and facility location problem. Due to the large number of applications, many solution approaches have been discussed in the literature. However, most combinatorial optimisation problems are considered as NP-hard.

In this section two basic methods are introduced; branch & bound and cutting planes. These combine to form the basis of the two methods; branch & cut and branch & price. As branch & cut is the default solving method in IBM Cplex Optimisation Studio for solving MILP, it will be described in more detail here.

3.2.1 Branch & Bound

Branch & bound was first proposed by Land & Doig (1960) and is defined as an enumeration technique. The method essentially divides the feasible region into more manageable subdivisions and then, if required, creates further subdivisions. A mixed-integer linear problem is being considered, thus is of the form:

$$\text{minimise} \quad \boldsymbol{c}^{\boldsymbol{T}}\boldsymbol{x} + \boldsymbol{d}^{\boldsymbol{T}}\boldsymbol{y} \tag{3.4}$$
$$\text{s.t.} \quad \boldsymbol{A}\boldsymbol{x} + \boldsymbol{E}\boldsymbol{y} \leq b \tag{3.5}$$
$$\boldsymbol{x} \in \mathbb{Z}^m, \boldsymbol{y} \in \mathbb{R}^n \tag{3.6}$$

The first stage of branch & bound is to use linear relaxation. In this way, the integer variables, $\boldsymbol{x}$, are replaced with continuous variables; $\boldsymbol{z} \in \mathbb{R}^m$.

The linear relaxation can be solved using normal linear programming (LP) techniques such as the Simplex method. The solution to the linear relaxation forms a bound on the problem. In this thesis, a minimisation problem is solved so the solution of the LP forms a lower bound to our ILP.

If the solution to the LP is integer, the solution can be accepted as feasible and optimal. However, if a non integer solution is produced the solution area is divided into two regions; $z_i \leq \lfloor z_i^* \rfloor$ and $z_i \geq \lceil z_i^* \rceil$ where z_i^* is the optimal solution, for some decision variable z_i, of the linear relaxation. Thus obtaining two sub-problems:

$$
\begin{array}{rcrc}
\min & \boldsymbol{c^T z} + \boldsymbol{d^T y} & \min & \boldsymbol{c^T z} + \boldsymbol{d^T y} \\
\text{s.t.} & \boldsymbol{Az} + \boldsymbol{Ey} \leq b & \text{s.t.} & \boldsymbol{Az} + \boldsymbol{Ey} \leq b \\
& z_i \leq \lfloor z_i^* \rfloor & & z_i \geq \lceil z_i^* \rceil \\
& \boldsymbol{z} \in \mathbb{R}^m, \boldsymbol{y} \in \mathbb{R}^n & & \boldsymbol{z} \in \mathbb{R}^m, \boldsymbol{y} \in \mathbb{R}^n
\end{array}
$$

Each of these sub-problems are solved in turn. If the solution is integer and feasible, it is a feasible solution. If it is infeasible, the branch no longer requires searching and is cut. However, if it is feasible but non-integer, the solution is compared with the best feasible solution found so far. If the new solution is better than the previous solution the branching process is repeated to obtain two more sub-problems. The search continues until all branches have been explored or a specified criteria has been met.

In the case where branching can be done on more than one variable, a suitable heuristic may be applied to choose the variable to branch on.

3.2.2 Cutting Planes

Cutting planes are used to remove the LP relaxation optimal solution but keep the integer solutions intact. It is an inequality constraint which is added to the ILP to reduce the feasible region. It is defined as follows:

Definition 1. *A* cutting plane *is a linear inequality that is:*

- *Satisfied by all feasible solutions of the ILP.*
- *Violated by a feasible solution to the LP relaxation.*

In order to solve an ILP with cutting planes, some important definitions from Polyhedra theory and Linear Algebra (Nemhauser & Wolsey, 1999) are required.

Definition 2. *'Given a set $S \subseteq \mathbb{R}^n$, a point $x \in \mathbb{R}^n$ is a* convex combination *of points of S if there exists a finite set of points $\{x^i\}_{i=1}^t$ in S and a $\lambda \in \mathbb{R}_+^t$ with $\sum_{i=1}^t \lambda_i = 1$ and $x = \sum_{i=1}^t \lambda_i x^i$. The* convex hull*, Conv(S), of S is the set of all points that are convex combinations of points in S.'*

Definition 3. *A* Polyhedron *$P \subseteq \mathbb{R}^n$ is the set of points that satisfy a finite number of linear inequalities such that $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ where (A, b) is an $m \times (n + 1)$ matrix. A polyhedron is a convex set.*

Nemhauser & Wolsey (1999) shows that if $P = \{x \in \mathbb{R}_+^N : Ax \leq b\} \neq \emptyset$ and $S = P \cap \mathbb{Z}^n$, where (A, b) is an integer $m \times (n+1)$ matrix then conv(S) is a rational polyhedron, where conv() is the convex hull.

Moreover, it is shown that an integer program

$$\max\{c^T x : x \in P \cap \mathbb{Z}^n\} \tag{3.7}$$

can be solved by solving the linear programme

$$\max\{c^T x : x \in Conv(P \cap \mathbb{Z}^n\}. \tag{3.8}$$

The polyhedra can be defined by using *facets*. To define a facet, the following definition is first required.

Definition 4. *'A face of a polyhedron P is $F = \{x \in P : Ax = b\}$ where $Ax \leq b$ is some valid inequality of P.'*

The dimensions of the faces define different planes. Faces of dimension 0 are called the *extreme points*, faces of dimension 1 are called the *edges*, faces of dimension $dim(P) - 1$ are called *facets*. As the size of n increases, the number of facets grows at a greater rate to exponential.

A cutting plane is a linear inequality that is valid for P_I but not for P. The strongest cutting plane will be one defined by the facets of P_I. Thus if the inequalities for the facets can be found, these inequalities could be useful as cutting planes. For some problems, such as the knapsack and travelling salesman problem, the polyhedra are known.

The general cutting plane algorithm follows the general procedure:

1. Solve the relaxed LP to get x^*, the optimal solution

2. If x^* is integral then stop, else find a valid inequality to exclude x^*
3. Go to step 1

There are many methods to determine the inequalities to be used for cutting planes, such as Chvatal-Gomory rounding method.

3.2.3 Branch & Cut

The standard approach to branch & cut (minimisation) problems is given in the following algorithm (Crowder, Johnson & Padberg, 1983):

1. Let $x^* = \infty_+$ be the optimal solution to the IP and cx^* be the objective value.
2. Relax the IP to an LP and solve. If the LP has a feasible, integer solution $\hat{x}$, let $x^* = \hat{x}$ be the integer optimal solution and stop.
3. Branch on the solution and select one of the problem instances generated.
4. Search for strong, facet-defining inequalities that are violated by the LP solution $\hat{x}$ and add the inequality as a cutting plane to the selected problem instance.
5. Solve the relaxed problem instance to obtain $\hat{x}$. If the solution is infeasible or $c\hat{x} < cx^*$ then fathom this branch, select an alternative branch and go to step 4. Otherwise go to step 2.

If the cuts are performed after the initial branch the process is referred to as cut & branch. If the cuts are performed at every node of the branch & bound tree the process is branch & cut. There are enhancements that can be used on the branch & cut method to increase efficiency, such as reduced-cost fixing.

In addition to the methods discussed here, one could also consider Bender's Decomposition, or heuristics as solution techniques.

3.3 Multi-objective Optimisation

For some optimisation problems, it may be beneficial to create the model with more than one conflicting objective. For example, a vehicle routing problem may consider the minimisation of cost, in terms of resources, and time, in terms of delays. If more vehicles are used this would increase the cost but decrease the delay time. Thus, multiple objectives result in multiple solutions.

A *Pareto solutions* is defined as a non-dominated solution such that any improvement to one objective would negatively impact another. The set of Pareto solutions are defined as the *trade off surface* or the *Pareto front* (Collette & Siarry, 2003).

Once a set of solutions is obtained, the Pareto front can be plotted to allow decision makers to visually analyse the trade-off. A k-dimensional plot is required where k is the number of objectives in the model. In the case of 2 or 3 objectives, solutions are plotted on the graph for various trade offs between the objectives. A *Pareto-dominate* solution is defined as follows (Ngatchou, Zarei & El-Sharkawi, 2005):

Definition 5. *"A decision vector* $\vec{u} = [u_1, u_2, ..., u_p]^T$ *is said to* Pareto-dominate *the decision vector* $\vec{v} = [v_1, v_2, ..., v_p]^T$*, in a minimisation context, if and only if:*

$$\forall i \in \{K\}, f_i(\vec{u}) \leq f_i(\vec{v}), \tag{3.9}$$

$$\text{and } \exists j \in \{K\} : f_j(\vec{u}) < f_j\vec{v}. \tag{3.10}$$

where K is the set of objectives $f_i()$. Hence, a solution, $\vec{u}$, exists in the Pareto front if the objective value is better than the objective for another solution, $\vec{v}$, in all objectives and that there exists at least one objective function for which this is strictly true.

Figure 3.1 gives an illustrative example of a Pareto front. Any solution on the line is a solution within the Pareto front, any other solutions are dominated solutions. A decision maker may use this plot to determine the best trade off for them.

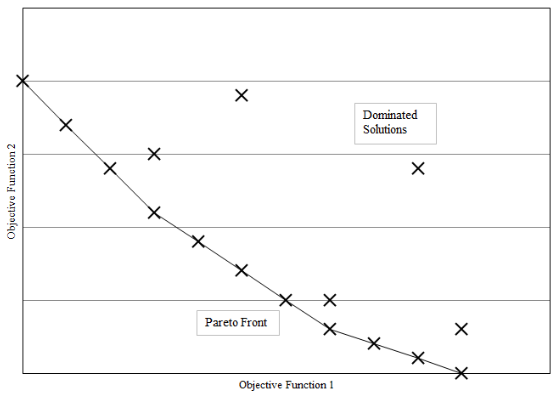

Fig. 3.1: Example Pareto front, $k = 2$

In order to create a multi-objective model, there are three families of methods that can be used; *a priori*, *progressive* and *posteriori*. Each of which define where in the method the trade off between objectives is considered.

A priori methods consider the trade off between objectives before the optimisation model is executed. Thus, one search is required to obtain a solution. Though this is the most time efficient method, it may be difficult for the decision maker to know what trade off is required prior to the optimisation. Else, they may change their decision once the optimisation has been complete, thus the optimisation will need to be rerun.

Progressive methods question the decision maker throughout the optimisation such that the search can be directed towards their required trade off. Depending on the length of the run, this method may be challenging. For longer runs of a few hours it may be difficult to continually update the model with decisions.

Finally, posteriori methods obtain a set of Pareto solutions. These solutions are then presented to the decision maker who will select the most suitable solution. Alternatively, a user may apply an utility function. Here, the utility function is applied to each solution and the solution to maximise the function is selected. Posteriori methods are considered as the most time consuming due to the number of solutions required where one search is required for each solution. For this method, the trade off surface may not be necessary where a selection of well spaced solutions may suffice.

A simple yet effective method of solving multi-objective models is through scalarising. This involves writing a multi-objective model as a single objective model such that the solutions for the single model solve the multi-objective model. As parameters must be selected before the optimisation has been executed, these are examples of a priori methods. However, the parameters may be changed and various runs produced to obtain a set of Pareto solutions.

Two scalarising methods are defined here; *linear scalarisation* and *ϵ-constraint method*.

Linear scalarisation applies a weighted value to each objective function and sets the new objective to be the weighted sum of the objective functions. Hence, the problem becomes a single objective model. The objective is expressed as;

$$\min_{x} \quad \sum_{i \in K} w_i f_i(\boldsymbol{x}) \tag{3.11}$$

where f_i defines the ith objective function, $w_i > 0$ defines the weight and $\boldsymbol{x}$ defines the decision variables. The coefficients w_i here must be set before the optimisation is executed and determine which solution is obtained from the search.

Epsilon constrained methods are used to convert the objectives into constraints such that one objective is left. Parameters ϵ_i are used as upper bounds to each constraint as follows

$$\min_{x} \quad f_j(\boldsymbol{x}) \tag{3.12}$$

$$\text{s.t.} \quad f_i(x) \leq \epsilon_i \qquad \forall i \in \{K\}, \neq j \tag{3.13}$$

where f_i defines the ith objective function, $\epsilon_i > 0$ defines the upper bounds and $\boldsymbol{x}$ defines the decision variables. Similarly to the linear scalarization, ϵ_i must be set before the optimisation is executed and the value determines the solution obtained from the search.

The values of the weights and epsilon in these two models may be determined before the execution of the model. However, the decision maker has the option to change the values to analyse the impact on the solution.

3.3.1 Goal Programming

The concept of goal programming was first introduced by Charnes & Cooper (1961). Goal programming is a multi-objective method applied when a problem requires multiple contradicting goals to be assessed. In this method one solution is derived such that the established goals are met or the smallest possible penalty is incurred. For each objective in the model, a numeric goal is formulated by introducing slack or surplus variables to represent a deviation from the goal. The solution will then minimise the weighted sum of deviations from these goals as in the linear scalarization approach defined above.

Three types of goals may be formulated:

- A one-sided goal defining the lower bound of the goal
- A one-sided goal defining the upper bound of the goal
- A two sided goal defining the specific range of the goal

Goal programming may also be defined as nonpreemptive or preemptive depending on the priority of the goals. If the goals are of the same importance, nonpreemptive techniques may be applied, if levels of importance of each goal can be defined, preemptive methods can be used.

Consider the problem containing a set of goals $f_i(\boldsymbol{x}) \leq b_i, \forall i$ where $\boldsymbol{x}$ is the vector of decision variables and b_i represents the goal value. Each goal i is associated with a penalty value p_i.

Let decision variables $y_i \in \mathbb{Q}$ represent the deviation from the goal (Ignizio, 1983).

$$\min \quad \sum_i p_i y_i \tag{3.14}$$

$$\text{s.t.} \quad f_i(\boldsymbol{x}) - b_i = y_i \qquad \forall i. \tag{3.15}$$

However, y_i must be positive in the objective function. Thus, it is specified that $y_i = y_i^+ - y_i^-, y_i^+ \geq 0, y_i^- \geq 0$. In the objective, the value of y_i is replaced with y_i^+ if the goal i is a one sided goal defining the upper limit, and y_i with y_i^- if the goal i is a one sided goal defining the lower limit. In the case of an equality goal, j, the objective will contain $p_j^+ y_j^+ + p_j^- y_j^-$, where p^+, p^- represent the penalty for exceeding the goal value and underachieving respectively.

In the constraints, the variable y_i is replaced with $y_i = y_i^+ - y_i^-$ to standardise the model.

For the preemptive case, goals are considered as being on priority levels such that higher priority goals must be satisfied before lower priority goals. Two basic methods are introduced here; sequential and streamlined.

The sequential method solves a sequence of linear programming models where the number solved is equal to the number of priority levels. In the first stage, only high priority goals are included in the model. If this produces a unique solution the method terminates with no consideration of other goals.

However, if there are multiple solutions that minimise this objective, consider the second stage of the model containing the second priority goals in the objective with the additional constraint that the first stage objective value must be equal to the objective given in the first stage of this sequence. This process is repeated for all priority levels.

The streamlined model solves one linear programme such that first priority goals are multiplied by a very large value, $M >> 0$, in the objective (Sarker & Newton, 2008).

Though more sophisticated approaches may be taken to solve a multi-objective model, the flexibility and simplicity of these methods makes them sufficient for the requirements of this thesis. For more information on other methods a reader may refer to Collette & Siarry (2003).

3.4 Stochastic Linear Programmes

Mathematical Programming has been used in Operational Research since 1939, since then the concept has adapted to include variations such as mixed integer, stochastic and bilevel programming. This section reviews applications of stochastic programming. The focus will be on papers of workforce allocation and training using stochastic programming. Important concepts of the techniques will also be defined.

A scenario can be defined as a set of conditions under which a model can be run such that different output is produced. Thus, a scenario can be defined by the different mission types. This could include different terrain types, weather conditions, risk of damage or combat type. Therefore, it is possible to differentiate between humanitarian missions, operations to warm climates and ocean operations for instance.

Consider the case when the scenario is unknown, thus probability of tasks requiring work is unknown. Bean (2011) considers three scenarios in their dissertation on inventory planning in the military. For each scenario the demand is either randomly distributed or given by a fuzzy number.

Alternatively, Markov processes can be used to model each scenario. There are given probabilities that scenario i will be followed by scenario j for all i, j. In the thesis by Ng (2003), the author aims to optimise a military supply chain over multiple periods when demand is unknown and non-stationary. The demand in each period is defined as Markovian. Here, a scenario tree is used to generate each of the scenarios. Song & Zipkin (1993) considers the case where demand varies due to the "state of the world". This is also modelled as Markovian and influences the demand. When the world is in state, or scenario, i then the demand is given by Poisson distribution λ_i. This definition of scenario will be applied in our modelling.

In order to model the unknown scenario in a MILP, consider the concept of stochastic programming. Using a set number of scenarios one can use a stochastic program to calculate the optimal workforce allocation. The simplest method for solving stochastic programming problems is to use an *Expected Value Model*. For any stochastic variables in the model, the stochastic variable is replaced with a deterministic variable equal to the expected value of the stochastic one. Though this works in many applications, the problem specified in this thesis may have a high variance. Though on average one would satisfy the demand for repairs, the cost for a delay due to lack of workforce is higher than the cost of the workforce.

Hence, more workforce would be provided than necessary.

Alternatively, consider *Robust Optimisation*. Initially established by Soyster (1973), robust optimisation allows a suboptimal solution to ensure the solution remains near optimal given uncertainty in the data. Thus, the solution can still be used if there are changes made to the data. Bertsimas & Sim (2004) argues that robust optimisation may be too conservative and suggests alternative methods that provide a trade off between robustness and the price of this robustness. Though using robust optimisation will allow enough workforce to be supplied in the *worst-case scenario*, this scenario may be rare. In which case the model would provide much more workforce than necessary for a large proportion of the time.

Robust optimisation is considered as a "hard" constraint, such that the probability of any constraint being violated is very small. However, the constraint may instead be considered as "soft", such that the decision maker may allow a non trivial chance of the constraint being violated in order to reduce costs.

If a decision maker had the opportunity to make an initial decision but adjust choices once the uncertainty has been realised, a *stochastic program with recourse* can be used, as given in Birge & Louveaux (1997). Recourse models allow the user to make initial decisions when the outcome of some event is unknown, then make second stage decisions once the outcome is known. The second stage decisions, or recourse, are used to determine a feasible solution to the problem.

$$\min \quad c^T x + E_\xi[Q(x,\xi)] \tag{3.16}$$

$$\text{s.t.} \quad Ax = b, \tag{3.17}$$

$$x \geq 0, \tag{3.18}$$

where

$$Q(x,\xi) = \min(q^T y | Wy = h - Tx, y \geq 0). \tag{3.19}$$

The stochastic program consists of two sets of decision variables. The variables x are known as the *first stage decisions* and these are taken without full information of some random events ξ, formed by the components of q^T, h^T and T. Once these decisions are made, information on some random vector ξ is received and *second stage* or *recourse* actions y take place to correct any bad effects experienced in the first stage. In the above formulation, $Q(x,\xi)$ is the objective function and A, b and c are matrices of constant numbers; the coefficients. Finally, E_ξ is the expectation with respect to ξ and W is fixed.

This form is defined as the *implicit representation* of the stochastic program. The model could also be defined by explicitly describing the second stage decision variables for all scenarios. This is called the *extensive form* of the stochastic program. A basic example of the stochastic recourse model is the famous *news vendor problem* proposed by Edgeworth (1888).

Alternatively, to consider the stochastic nature in the constraints, one could employ *probabilistic constraints* (or *chance constraints*). Here, the probability of a constraint being satisfied must be more than $\alpha\%$ (Charnes & Cooper, 1959). This concept has been applied to many stochastic programming problems. Tarim & Kingsman (2004) uses dynamic programming with probabilistic constraints to determine an optimal order quantity and replenishment period, when demand is stochastic. The constraint in this case is to obtain a level of serviceability, thus the probability that at the end of a period the net inventory will not be negative must be greater than some fixed value. A heuristic method is applied to solve it.

A chance constraint is written in the form

$$\mathbb{P}(Tx \leq D) \geq 1 - \alpha \tag{3.20}$$

where α is the confidence interval and $Tx \leq D$ is the constraint on decision variable x. Consider the cases where T or D are the stochastic variables.

There are two types of probabilistic constraints as defined in Vajda (1972); Individual and joint. For individual chance constraints, each stochastic constraint m is transformed into a chance constraint individually. Thus takes the form:

$$\mathbb{P}\left(\sum_i a_{ij}x_{ij} \geq b_j\right) \geq 1 - \alpha, \quad j = 1, ..., m. \tag{3.21}$$

For joint chance constraints, the individual chance constraints are replaced by one single joint chance constraint of the form;

$$\mathbb{P}\left(\sum_i a_{ij}x_{ij} \geq b_j, j = 1, ..., m\right) \geq 1 - \alpha \tag{3.22}$$

using the same notation as before with j equal to the number of scenarios and m equal to the number of constraints.

In order to solve the stochastic chance constrained model, it can be reformulated as a deterministic model to be solved using the methods described in Section 3.2. Due to the complexity of a probabilistic constraint, Pagnoncelli, Ahmed & Shapiro

(2009) states that little progress has been made until recently in finding a solution method. Two approaches tend to be taken; discretise the probability distribution or employ convex approximations of chance constraints.

Here, the method of *sample average approximation* (SAA) investigated in Luedtke & Ahmed (2008) and Atlason, Epelman & Henderson (2008) is discussed. The approach is to replace the distribution in the chance constraints to an empirical distribution corresponding to a random sample.

Consider chance constraints of the form

$$\min\{f(x) : x \in X, \mathbb{P}[G(x,\xi) \leq 0] \geq 1-\alpha\} \tag{3.23}$$

as specified in Luedtke & Ahmed (2008). Here $x \in X$ are the decision variables, f represents the objective function, ξ is a random vector, $G : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^m$ is a constraint mapping and α is a user specified risk parameter.

This can be restated as

$$\begin{aligned} z^*_\alpha &= \min\{f(x) : x \in X_\alpha\} \\ X_\alpha &= \{x \in X : \mathbb{P}[G(x,\xi) \leq 0] \geq 1-\alpha\} \end{aligned} \tag{3.24}$$

where it is assume z^*_α exists and is finite. The measurability of any event S taken under probability is also assumed, such as the event $\{G(x,\alpha) \leq 0\}$ for each $x \in X$.

Due to the difficulty involved in calculating $\mathbb{P}\{G(x,\xi) \leq 0\}$ for $x \in X$, it can be approximated to G_α by solving a sample approximation problem. Let $\xi_1, \xi_2, ..., \xi_N$ be an independent Monte Carlo sample of the random vector ξ. Then for fixed $\alpha \in [0,1)$ define the sample approximation problem of (3.24) as

$$\begin{aligned} \tilde{z}^N_\alpha &= \min\{f(x) : x \in X^N_\alpha\} \\ X^N_\alpha &= \{x \in X : \frac{1}{N}\sum_{i=1}^{N} \mathbb{I}(G(x,\xi_i) \leq 0) \geq 1-\alpha\} \end{aligned} \tag{3.25}$$

where

$$\mathbb{I}(t) := \begin{cases} 1 & \text{if } t \text{ is true} \\ 0 & \text{otherwise.} \end{cases} \tag{3.26}$$

It is assumed by Luedtke & Ahmed (2008) that equation (3.25) has an optimal solution except in two cases; if $X_\alpha^N = \emptyset$ then $\tilde{z}_\alpha^N = +\infty$, whereas if P_α^N is unbounded then $\tilde{z}_\alpha^N = -\infty$. This assumption holds if X is compact, $f(x)$ is continuous and $G(x, \xi)$ is continuous in x for each ξ.

The indicator function, $\mathbb{I}(t)$, is used to count the number of scenarios ξ_i where the constraint $G(x, \xi_i) \leq 0$ is not violated. This ensures the proportion of violated constraints does not exceed α.

Luedtke & Ahmed (2008) formulates (3.25) as an MILP

$$
\begin{aligned}
\min \quad & f(x) && && (3.27)\\
\text{s.t.} \quad & G(x, \xi_i) \leq M_i \delta_i && i = 1, ..., N && (3.28)\\
& \sum_{i=1}^{N} \delta_i \leq \alpha N && && (3.29)\\
& \delta_i \in \{0, 1\}, && i = 1, ..., N && (3.30)\\
& x \in X && && (3.31)
\end{aligned}
$$

where δ_j is a binary variable and M is a large positive number. Here, $z_j = 0$ when the constraint holds. Equation (3.29) is a cardinality constraint requiring the proportion of times the constraint is enforced over all scenarios.

3.5 Summary

In this chapter, the methods for solving mixed integer linear programmes have been discussed. In particular, methods relevant to the model components and extensions possible in this model. Using suitable techniques discussed in this chapter, complex features of the problem may be reformulated to make them easier to solve. In addition, it will be possible to solve them all using the same MILP solver.

For this training model, the default solution method used in IBM Cplex will be applied to the model.

4. MODEL DEVELOPMENT

In this chapter, models are proposed for the Training Needs and Training Execution Models which combined are used as a Training Development Model. As discussed in the literature, a multi-objective MILP is appropriate to model the problem with stochastic constraints on the relevant policies. Mathematical programming allows for a highly adaptable, modular model which can be customised dependent on the choice of decision variables and constraints. The method to solve the problem remain similar, independent of the chosen model methods and extensions.

In the first section the Training Needs Analysis (TNA) is formulated such that no skill development will take place. The TNA considers all possible tasks that may need completing. Tasks are assigned to employees, considering the availability of the employee and their skill sets, which are fixed. The number of unfulfilled tasks is minimised in order to analyse the skill gap in the current workforce.

Next, the TNA is expanded to include training and skill development. Various training policies are proposed and model extensions to further constrain the problem. Three Training Execution Methods (TEMs) are introduced which can affect an employee's ability to procure an auth. In addition, six model extensions are considered as additional modelling constraints. Each extension will be discussed in turn and note how they can be combined to define the workforce training policies of the company using the tool.

From these nine different model components, one training model can be created. The separate models introduce the small number of decision variables and constraints required to expand the model to meet the users needs. This makes the overall model highly adaptable and customisable using very simple methods. The extensions are easy to manage by removing or adding their associated decision variables and constraints where necessary.

The size of the training model is defined by the number of employees, e, the number of tasks, k, and the number of time units t. The notation $[e, k, t]$ will be used to define the size of each model in this thesis. Other indices applied in extensions will be defined separately.

4.1 *Training Needs Analysis*

The Training Needs Analysis incorporates the allocation of employees, $e \in E$, to tasks, $k \in K$. It is used to introduce some basic concepts that will be developed in further sections. It is assumed here that employee's authorisations are fixed for the duration of the time horizon, T.

The model requires a list of tasks K that must be completed in each time unit. The number of times each task will need completing in the time unit, t, are defined as the requirements for the task. Here it is assumed one employee is required to complete each task, thus the requirements define the number of times each employee completes the task in each time unit. Employees can be assigned to more than one task in the time unit if they have the available time.

In order to satisfy the requirements, the employee must have the associated authorisation, or skill, to complete the task. Authorisation is defined in this thesis as follows:

Definition 6. *An* authorisation *(auth) is a binary variable, $s_{ek} \in \{0, 1\}$, required by an employee e for the completion of task k. If the value of the auth is equal to one the employee is deemed competent to complete the associated task.*

Each task has an associated authorisation. The terms skill and authorisation can be used interchangeably.

The decision variable to be analysed in the TNA is the allocations of employees to tasks within each time unit. These potential allocations, $w_{ekt} \in \mathbb{Z}^+$, are subject to constraints. An employee is only able to complete a task if they have the availability and the auths. If no one is able to complete the task it is defined as 'unfulfilled'. The objective is to minimise the number of tasks as such. These unfulfilled tasks represent the skill gap.

Availability, $a_{et} \in \mathbb{Q}^+$, of each employee is modelled as a continuous variable greater than or equal to zero defining the number of hours the employee is able to work during the time unit. The processing time for each task is defined by $\eta_k \in \mathbb{Q}^+$.

In order to create this model the following assumptions are required:

- **Processing times of tasks are fixed**. If processing times vary, assume they are the average from past data.
- Each employee's **authorisations are fixed**.
- **The order of the tasks does not need to be considered** as each time

unit is large enough for tasks to be ordered efficiently without affecting the model.

- **Unexpected absences are not modelled.** To account for this, the value of the availability is set to include the average number of sick hours taken by employees in a month.
- **The size of the workforce does not change**. Thus hiring new employees or losing employees is not considered as an option.

The TNA is defined by the following mixed integer linear programme

$$\min \quad \sum_{k\in K}\sum_{t\in T}(r_k - \sum_{e\in E} w_{ekt}) \tag{4.1}$$

$$\text{s.t.} \quad w_{ekt} - r_k s_{ek} \le 0 \qquad \forall e \in E, k \in K, t \in T \tag{4.2}$$

$$\sum_{k\in K} \eta_k w_{ekt} \le a_{et} \qquad \forall e \in E, t \in T \tag{4.3}$$

$$\sum_{e\in E} w_{ekt} \le r_k \qquad \forall k \in K, t \in T \tag{4.4}$$

The objective (4.1) in the TNA model minimises the resultant slack resources caused by the under supply of workforce where $r_k \in \mathbb{Z}^+$ defines the number of times task k needs completing in each time unit t and w_{ekt} is the decision variable specifying the number of times employee $e \in E$ completes $k \in K$ in time unit $t \in T$. By determining the optimal allocation w_{ekt}, the skill gap can be minimised.

The problem is constrained such that employees can only complete a task if they have the auth, (4.2), where the auth s_{ek} is a binary variable equal to one if the employee can complete the task k.

The employee can only complete the task if they have the availability, equation (4.3). It is required that the time to process each task, given by the data $\eta_k \in \mathbb{Q}^+$, multiplied by the number of tasks they are assigned to must be less than their total availability, given by the data $a_{et} \in \mathbb{Q}^+$.

Finally, equation (4.4) ensures you cannot assign more employees than required for the task. If this constraint is relaxed, a negative skill gap would be obtained and it would not be possible to determine if tasks are still not being complete due to lack of resources.

Thus, the TNA is used to allocate a multi-skilled workforce to a set of all possible tasks given constraints on their auths and availability. This model provides the skill gap by minimising the number of tasks that do not have enough workforce to be completed.

4.2 Training Execution Methods

In the previous model, the auths were taken as fixed values given by the user. Now consider the case where skills can change through time, dynamically. It is assumed that authorisations must expire and thus be reviewed before the employee can use that skill again. The concepts of review time and review criteria are introduced in this section. This criteria must be met in order to procure the auth.

For all methods of skill development the review time must be fixed for all auths and all employees. Thus if an employee gains a skill auth at time t then it must be reviewed at time $t+\tau$ where τ is the time between reviews. To qualify for an auth, an employee must show they have gained sufficient experience since the last review to exceed the stated experience threshold for that task. For instance, if an auth is reviewed yearly, it may be required that at least six months worth of experience must be shown to qualify for the auth. It is concluded that the employee has the competency to complete that task and is thus appointed the corresponding auth.

For notation, a review number is the index defined as $z \in \overline{Z}$. The number of reviews within the time horizon is equal to $\hat{z} = \lfloor \frac{\hat{t}}{\tau} \rfloor$ where $\hat{x} = \max_x(X)$ indicates the largest value in the set for any index $x \in X$. This notation will remain consistent throughout this thesis. The sets Z_{ekz} define the review period such that they contain all times t for employee e in task k where the review, z, takes place at the beginning of the review period. For instance, if review $z = 1$ took place at $t = 6$, the first review period $Z_{ek1} = \{1, 2, 3, 4, 5\}$. The time of the first review may be any time between 1 and $\tau - 1$. For $1 < z < \hat{z}$, where $\hat{z}$ is the last review, the size of the corresponding review period is equal to τ.

During each review period the skill will remain fixed. The skill at review z is calculated from allocations made and training undertaken during the previous review period, $z - 1$, and defines the skill for the review period Z_{ekz}.

Before the first review, employees have a fixed set of auths. This is the auths held by the employee before the start of the modelling horizon and are given in the data. For new employees the value of all initial auths could be set to 0, if the candidate information is unknown. Until the first review, for $z = 1$, the skill of the employee is equal to the initial auth value for each task. Hence the value of s_{ek1} is set as an equality constraint. It is assumed that each employee has zero experience at the beginning of the time horizon, even if their initial skills is not zero. Thus, if the review period began before the beginning of the time horizon, the employee's experience gained through this time does not count towards this time. Though the model could simply be adjusted to include an initial experience, it would be difficult for the employer to determine the value of this variable unless the number of allocations since the last review are stored.

4.2.1 Description

To prove an employee is competent enough to obtain an authorisation, they must gain a certain amount of experience through training. If enough experience has been gained before the skill is reviewed, then a new auth can be allocated to that employee at the end of that review period. Hence, the employee will then have the skill to complete that authorised task for the duration of the next review period.

To obtain experience, three different methods are proposed in this thesis. The first method is considered a core method of training and is always available and thus required in any training model, this is defined as the Training Execution Method (TEM) Zero. The other two proposed TEMs are optional training techniques that can be included in the model. Each TEM defines a different method for an employee to gain experience. The combination of the TEMs with the extensions which will be discussed later is known as the Training Development Model. It should be noted that training and TEMs are not always the same thing. For example compulsory training does not increase the experience. Experience is defined as the number of times the task has been completed by an employee. As such, an employee will need to complete a task g_k times in order to show they are competent. This could also be stated in terms of time units, such that one time units worth of experience is equal to the number of times a task needs completing in a time unit.

TEM Zero describes training through repeating the same tasks throughout the time horizon. The number of times an employee is allocated to a task defines the number of time units worth of experience received by the employee. This method has no associated cost. However, as employees can only be assigned to tasks if they are authorised to do so, employees cannot gain experience with this method unless their initial skill is equal to 1.

Training Execution Method One is *on-the-job training* (OJT). OJT is widely used in companies as it is considered cheaper than other means of training. It is considered efficient as employees can train alongside their colleagues and learn how to do the specific tasks required for the job rather than generic skills. However, it can slow production and may not be effective if bad habits are passed on. Some industries that use OJT include engineering, telecommunications and manual work such as carpentry and plumbing.

An employee completing OJT can only do so if an employee with the auth is allocated to the task. The employee who requires OJT is assigned to assist with a task that is being processed in the time unit. This employee must be trained by one employee assigned to the task in a one-to-one relationship. As OJT is primarily performed by those who do not have the skill, on-the-job training does not contribute to the workforce allocated to that task and thus cannot contribute to the requirements for that task.

There is a cost associated with OJT, this can be considered as the additional pay given to the employee providing the training or the cost of a slower production rate. However unlike repeating tasks, an employee can complete OJT whether or not that employee has the associated auth.

The final TEM is *training courses*. It defines any training that does not affect the allocation to tasks or the production line. It can be hosted by an external company or by the user using the tool as long as the training does not affect the tasks being completed.

The experience gained from training courses is dependent on the type of training. Consider different types of training course, with durations, cost, maximum capacities and optional start times. A new index $n \in N$ defines each training course type. Thus, for each auth, courses like e-learning, residential courses, conferences, and reading manuals can be compared.

Each type of training course creates a gain in experience for a subset of auths. Each course can increase the experience in more than one auth. This experience is measured in terms of time units of experience. For instance, a days workshop could be equivalent to a months worth of repeating tasks. The experience gain for a course is given as an integer value equal to the number of times the task would of needed repeating in order to gain the same amount of knowledge as the course provides. The experience gained for each auth is specified in the data and can be set to zero if the course does not increase the experience for certain auths.

Training courses are used in industries where the tasks are too complex to be learnt on-the-job. This training can be specific to the task or used to teach generic skills. It is used in maintenance, manufacturing and healthcare, for instance. It is also useful when a large number of employees need to complete the same training. In this case, on-the-job training would be less useful as each employee completing OJT decreases production efficiency. Though this training is not required in order to complete a task it may be useful to increase an employee's experience on tasks they are not regularly completing. Training courses can be completed whether the employee has the skill or not.

TEMs One and Two can be chosen as necessary, subject to the needs of the company, in conjunction with TEM Zero for repeating tasks. Thus, a training policy can use repeated tasks and OJT to increase the experience and not have an option of training courses, for instance. The variables and constraints are selected as required and can easily be put into the model.

As stated, the experience obtained through repeating tasks, TEM Zero, forms the basis of all models. In this way, it is the only set of constraints that appear in all models under the assumption that at all times employees gain experience through repeatedly performing a task.

A tool is proposed for optimising the workforce strategy through training by adjusting the TNA model proposed. The purpose of the model is to output the skill gap calculated by the input requirements, $r_k \in \mathbb{Z}^+$, minus the decisions of allocations, w_{ekt}. The cost of training and the perceived fairness are also output. The fairness defines the differentiation of the number of tasks allocated by each employee. A skill profile for each employee is calculated and presented.

To integrate each of these TEMs, assumptions must be made. For on-the-job training, it is assumed:

1. **On-the-job training does not affect the tasks ability to be completed.** Thus it does not affect the processing time.
2. **One person is required to give one other person training.** The model does not need to specify which employee is training which other employee.

For training courses, it is assumed:

1. **Travel time is not considered**. Any time spent travelling should be considered in the length of the course.
2. **Course trainers are not considered**. As these consider various training courses, trainers are either not required, do not affect the work for the maintenance line or are provided by external sources. Any costs associated with course trainers is considered in the cost for the course.
3. **Courses cannot be cancelled or missed** if an employee has been assigned to it.

4.2.2 Objective

For the problem specified in the thesis, the trade off is considered between the skill gap and the level of unfairness in allocations. It is assumed that the decision maker will want to minimise the skill gap as a priority as this will have a greater impact on the costs of the system and it's effectiveness. The unfairness element may affect the level of happiness of the employees but, operationally, it is important to complete as many tasks as possible. Thus an a-priori method may be selected.

To model this, weighted values are used such that the trade off between skill gap and fairness can be considered and the value can be adjusted according to the wishes of the company. As with the TNA, the skill gap should be minimised where the skill gap refers to the total required workforce minus the total assigned workforce.

However, now that skills are being assigned to employees, a soft constraint should be considered for the fairness of the allocations. Here, fairness is dependent on the number of tasks assigned to the employees. If all tasks were assigned to one employee, they would be overworked due to the number of tasks they would need to complete in the available time. On the other hand, an employee with no tasks assigned to them may be considered as unnecessary as their time is not being sufficiently used to complete tasks. This constraint represents a goal programming constraint where the goal is to create a fair workforce solution by minimising the difference between number of tasks assigned to each member of the workforce.

Due to this fairness function, the model will prioritise assigning employees to tasks as a way of creating an even distribution of work. By minimising the skill gap over all time, the model ensures allocations are the priority of the model. The coefficient of the skill gap is set to prioritise the skill gap using linear scalarisation methods mentioned in Section 3.3.

As also discussed in Section 3.3, goal programming can be solved by introducing a decision variable to measure the deviation from the goal. This value is then minimised in the objective. The method can be applied to the fairness constraints. The decision variables defining fairness are calculated in the following constraints:

$$\frac{1}{\hat{e}-1}\sum_{e\in E:e\neq f}\sum_{k\in K}\sum_{t\in T} w_{ekt} - \sum_{k\in K}\sum_{t\in T} w_{fkt} \leq \psi_f^+ \qquad \forall f \in E \quad (4.5)$$

$$\sum_{k\in K}\sum_{t\in T} w_{fkt} - \frac{1}{\hat{e}-1}\sum_{e\in E:e\neq f}\sum_{k\in K}\sum_{t\in T} w_{ekt} \leq \psi_f^- \qquad \forall f \in E. \quad (4.6)$$

Here, the decision variable ψ_e^+ defines the positive deviation from the average allocations of employee e. Conversely, the decision variable ψ_e^- defines the negative deviation from the average allocations of employee e. By using two constraints and decision variables, the absolute value of the difference between the allocations of employee e and the average allocations of all other employees is considered in the objective, as explained in Section 3.3. The value $\hat{e}$ defines the number of employees.

The objective is given as follows, where two objectives, skill gap and fairness, are written as one using a weighted coefficient on the skill gap:

$$\min \quad \sum_{e\in E}(\psi_e^+ + \psi_e^-) + A\sum_{k\in K}\sum_{t\in T}(r_k - \sum_{e\in E} w_{ekt}) \tag{4.7}$$

where the first term indicates the fairness in terms of difference in number of allocations. The second term defines the weighted skill gap.

As with the TNA, the skill gap is calculated as the total number of times the task is required to be completed, $r_k \in \mathbb{Z}^+$, minus the number of times the employees are assigned to the tasks at each unit of time, w_{ekt}. This is summed for all tasks and all time units, such that task repetition should be the preferred method of training as it reduces the skill gap. The value $A \in \mathbb{Q}^+$ is the coefficient in the objective function used to weight skill gap against the fairness value.

To select a value of A to ensure the skill gap is always prioritised, it should be assumed that the maximum value of fairness is equal to A multiplied by a skill gap of 1. The most "unfair" allocation of skills is one employee performing all tasks.

Thus;

$$\begin{aligned} A &= \sum_{e\in E}(\psi_e^+ + \psi_e^-) \\ &= (\hat{e}-1)\frac{1}{\hat{e}-1}\sum_{k\in K}\hat{t}r_k + \sum_{k\in K}\hat{t}r_k \\ &= 2\hat{t}\sum_{k\in K} r_k. \end{aligned}$$

This ensures the skill gap is first order priority and fairness is only minimised subject to the skill gap being the minimum possible value. If weighting is specified such that fairness is first order priority then no allocations will be made such that the value of fairness is given as zero and the skill gap will be calculated as the maximum value; a trivial solution.

4.2.3 Constraints

New decision variables and data are required to define the dynamic skills. The initial skill of the employee is initialised with the data value $\Theta_{ek} = \{0, 1\}$ which equal 1 if employee e has auth k at time 1 and 0 otherwise. The skill is now defined as a decision variable rather than an input. The skill $s_{ekz} = \{0, 1\}$ is equal to 1 if employee e has auth k in review period z, or 0 otherwise. As before, the terms skill and authorisation are used interchangeably and each task k is associated with an auth.

Constraints on task allocation are similar to the constraints from the TNA, however the skill is now considered as a dynamic decision variable. As before, $a_{et} \in \mathbb{Q}^+$ is the available hours of employee e at time t, given annual leave or sick leave. For TEM Zero, the constraints are

$$w_{ekt} - r_k s_{ekz} \leq 0 \qquad \forall e \in E, k \in K, z \in \overline{Z}, t \in Z_{ekz} \tag{4.8}$$

$$\sum_{k \in K} \eta_k w_{ekt} \leq a_{et} \qquad \forall e \in E, t \in T \tag{4.9}$$

$$\sum_{e \in E} w_{ekt} \leq r_k \qquad \forall k \in K, t \in T. \tag{4.10}$$

As in the TNA model, equations (4.8), (4.9) and (4.10) define the constraints on allocations, availability and requirements respectively. Equation (4.8) introduces the set Z_{ekz} which contains all times t within the review period z for employee e in task k.

The auths are awarded by calculating the experience gained through the TEMs. This experience is then compared to the experience required for the auth to be awarded, given as $g_k \in \mathbb{Q}^+$. The experience required is defined as the number of times the task should be completed, or equivalent, in order to gain the auth. For repeating tasks, the experience is increased by one for every time a task is complete. If the experience exceeds the threshold value then the auth is rewarded.

$$s_{ek1} = \Theta_{ek} \qquad \forall e \in E, k \in K \tag{4.11}$$

$$g_k - M_k(1 - s_{ekz}) \leq \sum_{t \in Z_{ek(z-1)}} w_{ekt} \qquad \forall e \in E, k \in K, z \in \overline{Z} \tag{4.12}$$

where $M_k \gg 0$ is selected dependent on the value of k.

Given that TEM Zero is used exclusively, equation (4.11) is used to calculate the skill in the first review period and (4.12) calculates the skill in subsequent periods.

The training tool incorporates the two optional TEMs with TEM Zero. The decision variables and constraints can be removed as required with the model still functioning. Thus, a combination of TEMs can be used. The changes required to the above constraints for TEM Zero are presented. Other constraints remain in the model and are not adapted.

To incorporate the optional TEMs, two decision variables are introduced. For on-the-job training, $p_{ekt} \in \mathbb{Z}^+$ is the number of times employee e completes OJT in task k at time t. For additional training, $v_{ent} \in \{0, 1\}$ is 1 if employee e attends course n at time t. These decisions have an associated cost specified; $c_1 \in \mathbb{Q}^+$ is the cost of OJT and $c_{2n} \in \mathbb{Q}^+$ is the cost of sending an employee on course n.

First, the training provided in each optional TEM affects the available time of the employee, constraint (4.9). The availability constraint is updated to include the time spent on OJT, which is the same as the processing time for the task, $\eta_k \in \mathbb{Q}^+$, and the time spend on training courses, $\rho_n \in \mathbb{Q}^+$. This constraint can be adjusted to remove OJT or training courses from the model.

$$\sum_{k \in K} \eta_k (w_{ekt} + p_{ekt}) + \sum_{n \in N} \rho_n v_{ent} \leq a_{et} \qquad \forall e \in E, t \in T \quad (4.13)$$

The experience gained from this training is used to determine the skills in the next time period. In addition to the experience from repeating tasks indicated in equation (4.12), there is now experience gained from OJT and training courses. On-the-job training is equivalent to an experience of one every time a task is complete. For training courses this value may be more difficult to describe. The experience must be specified in equivalent terms. For instance, completing a training course may be equivalent to repeating a task x times. The sum of these values defines the amount of experience gained by the employee and must exceed the threshold value g_k in order for the auth to be awarded.

$$g_k - M_k(1 - s_{ekz}) \leq \sum_{t \in Z_{ek(z-1)}} (w_{ekt} + p_{ekt} + \sum_{n \in N} \chi_{nk} v_{ent}) \qquad \forall e \in E, k \in K, z \in \overline{Z} \quad (4.14)$$

where $\chi_{nk} \in \mathbb{Z}^+$ is the number of units of experience towards task k from course n. As stated previously, the experience gained from OJT is equal to the number of times the task is completed. The constraint on the initial skill is not affected.

In addition to the changes to the TEM Zero constraints, new constraints must also be defined. In order to model the training costs, a constraint on the available funding for training is proposed to ensure that the amount spent on training does not exceed the budget. This is a hard constraint such that the budget cannot be exceeded.

The available budget for training purposes is given by the input $b \in \mathbb{Q}^+$.

$$\sum_{e \in E} \sum_{t \in T} \left(\sum_{k \in K} c_1 p_{ekt} + \sum_{n \in N} c_{2n} v_{ent} \right) \leq b \tag{4.15}$$

The cost of OJT is given by $c_1 \in \mathbb{Q}^+$ and the cost of training course n is given by $c_{2n} \in \mathbb{Q}^+$. Thus, the left hand side defines the total amount spent on training all employees over the specified time horizon. By removing either term in the left hand side of the equation, the training options can be adjusted. There is no cost associated with completing tasks, hence it does not appear in the constraint.

Though the training selected in the optimised solution may have a cost less than the specified budget, it may not be the cheapest method to gain the optimised skill gap; there are more possible solutions. Reducing the budget may improve this cost, while the skill gap remains the same, but may affect the run time of the model. This will be explored in a later chapter.

In addition to the training costs constraints, there are constraints for each TEM required to ensure the training is carried out correctly.

$$p_{ekt} - \sum_{f \in E : e \neq f} w_{fkt} \leq 0 \qquad \forall e \in E, k \in K, t \in T \tag{4.16}$$

$$\sum_{e \in E} v_{ent} \leq u_{nt} \qquad \forall n \in N, t \in T \tag{4.17}$$

Equation (4.16) states that an employee cannot complete on-the-job training unless there is a skilled employee assigned to the task at the same time. In this way, the employee is available to assist with the training. In addition, the number of employees assigned to the task must be greater than the number of employees assigned to on-the-job training. Thus, one person cannot provide on-the-job training to multiple employees; the relationship is one-to-one. The number of the employees must be correct, however it does not matter who provides training to which employee as all employees are assumed identical.

The constraint for training courses is stated in equation (4.17). Here, the number of employees assigned to a training course at one time cannot exceed the capacity for the course, $u_{nt} \in \mathbb{Z}^+$. The capacity is set to zero if the course is not available at time t.

4.2.4 Discussion

Alternative objective functions were proposed for this model. As discussed in Section 3.3, epsilon constrained methods can be used to solve a multi objective programme by moving components of the objective into the constraints by setting a value ϵ as a bound.

Three terms could be considered in the objective function; skill gap, fairness and training cost. For each of these components, it is reasonable to consider the company aiming to minimise the value in the solution. In the model proposed above, we chose to minimise the skill gap and the fairness gap subject to training costs being less than a fixed value.

Consider each of the terms as constraints in turn. First, a constraint on the skill gap such that the number of unfulfilled tasks cannot exceed a fixed amount, ϵ. Let us consider a time unit t, where no reviews have taken place before t. In this way the skills of the employees are given by the initial skill Θ_{ek} specified before the start of the time horizon. The allocations available from these initial authorisations may result in a skill gap greater than the fixed amount with no training possible to rectify this as we are in the first review period. Thus an infeasible solution is created from the initial data, if only one value of epsilon is specified.

By defining a different epsilon for the initial review period, infeasible solutions can be avoided. However, the values of these epsilons will still require calculation. For the first review period, epsilon must be greater than the minimum number of tasks that can be completed given the initial skills. This can be calculated from the TNA, but requires an additional step to the modelling process.

The value of epsilon after the initial period is also difficult to define, as the company will need to determine the number of tasks they are allowed to miss, which in most cases would be equal to zero. In order to minimise training costs, this would require completing the minimum amount of training in order to achieve the skill gap goal. Skill gaps would be more suited as a soft constraint thus appearing in the objective.

Consider the fairness value as a constraint such that the value cannot exceed the fixed value ϵ. Again, putting a quantity to this value is more challenging as no policy will exist to specify how many allocations would be considered unfair. As a hard constraint, an allocation of no employees could be considered to procure a fairness deviation of zero; a trivial result each time. Instead, we would wish to consider the fairness as a soft constraint as it is preferential but not essential to most companies.

Finally, training costs are used as a constraint as given above. Training costs are easier to quantify and budgets are usually provided by companies. As these budgets are usually hard budgets, the training costs are more realistic as hard constraints. Unlike the skill gap, training costs as a constraint does not cause infeasible solutions. A training budget of zero may be used and still obtain a solution. In this trivial case the solution will be no allocations to training.

Hence, it was determined that a weighted multi-objective model with two objectives would give the most efficient and effective training solution in a reasonable run time. The training costs are considered as a hard constraint in the model, fairness and skill gap are considered in the objective.

4.3 Extension Policies

In addition to training, other constraints may affect how training is conducted and how auths are awarded. In this section, six extensions are considered in turn and additional constraints proposed to include in the model. A subset of these extensions may be applied to the model by choosing the constraints and decision variables as applicable.

The extensions proposed are as follows:

1. Auths may only be awarded if compulsory training is completed.
2. The requirements of an auth differ for an initial or recurrent skill.
3. Auths may depend on the completion of other auths.
4. Tasks and employees are defined by levels or speciality.
5. Experience required is a non-linear function of the total time spent on tasks.
6. Tasks have a probability of occurring in each time unit.

The decision variables and constraints are examined in turn in reference to the training model given previously. Extensions One to Five do not affect the objective function of the TEMs but may affect constraints and will provide additional constraints to the formulation. However, Extension Six will require the model to be reformulated.

In this section, previous constraints mentioned will include TEM 1 and TEM 2. In this way, the impact of the extensions on all training methods may be discussed. However, the extensions will still function is either of these TEMs are removed from the formulation.

4.3.1 Extension One - Compulsory Training

The first extension proposed is compulsory training. Though this is another training method, it does not affect the experience of the employee as seen in the TEMs. Thus, it is considered as an additional modelling constraint. Compulsory training is defined as a specific training programme that must be completed in order to obtain the related auth. This requirement may be in conjunction with the experience requirement for the auth.

Compulsory training must be completed within the review period. Each compulsory course is associated with a specific task and has a length and cost. The compulsory training can be a simple interview to prove an employee can complete the task, an e-learning course or a course held by an external company. The training specified for the task must be completed to gain the auth.

In this extension, it is assumed that compulsory training is the same for employees independent of whether or not they currently hold the auth. Thus, in order to renew or gain the auth, the same training must be completed.

Compulsory training is used in industries with strict regulations. In healthcare for instance, nurses will need to attend specific courses before being allowed to perform certain procedures. Other companies may also state certain health and safety training must take place before they can perform a task. Compulsory training may also be used to define an event that must be completed, such as a test, in order to show their competence before gaining the auth.

However, it is assumed that an employee cannot fail the compulsory course. For compulsory courses that contain a test it is assumed that either the course is enough to show competence or the experience obtained through other training methods is sufficient to show competence in this test. Hence, if an employee has gained enough experience to satisfy the experience threshold g_k, they have gained enough experience to pass the test. The value of g_k should be set accordingly.

The following assumptions are required to model the compulsory training:

1. **Travel time is not considered**. As with training courses, it is assumed that the time to travel is absorbed by the length of the course.
2. **Course trainers are not considered**. It is assumed that trainers are provided independent of the model.
3. **Courses take place at any time**. There is no constraint on when the courses may happen or how many times they may be completed.
4. **There is no capacity for these courses.** Any number of employees can

perform compulsory training at the same time.

5. **Compulsory training is always passed.** If the training consists of an interview or tests, it is assumed that the employee must have enough experience to achieve the auth and thus pass the compulsory training.

To incorporate the extension into the Training Development Model, a new binary decision variable is defined; $q_{ekt} \in \{0,1\}$ is equal to 1 if employee e is assigned to the compulsory training associated with task k at time t. The availability constraint, equation (4.9), is reformulated to include the time to complete compulsory training. This time, $\mu_k \in \mathbb{Q}^+$, is specified by the user.

$$\sum_{k\in K} \eta_k \left(w_{ekt} + p_{ekt}\right) + \sum_{n\in N} \rho_n v_{ent} + \sum_{k\in K} \mu_k q_{ekt} \leq a_{et} \qquad \forall e \in E, t \in T \quad (4.18)$$

The training costs are updated to include the cost of compulsory training, $c_{3k} \in \mathbb{Q}^+$.

$$\sum_{e\in E}\sum_{t\in T}\left(\sum_{k\in K} c_1 p_{ekt} + \sum_{n\in N} c_{2n} v_{ent} + \sum_{k\in K} c_{3k} q_{ekt}\right) \leq b \qquad (4.19)$$

One additional constraint is added to the model.

$$\pi_k s_{ekz} - \sum_{t\in Z_{ek(z-1)}} q_{ekt} \leq 0 \qquad \forall e \in E, k \in K, z \in \overline{Z} \quad (4.20)$$

Equation (4.20) ensures the compulsory training criteria is met, in the time between reviews, before an auth can be obtained. The time between reviews is given as $\tau \in \mathbb{Q}^+$, specified by the data. Thus $\pi_k \in \{0,1\}$ is equal to 1 if compulsory training is required to complete task k as given in the data. This constraint is only required for values of k such that $\pi_k = 1$, otherwise the constraint is unnecessary.

4.3.2 Extension Two - Recurrent Auths

In the previously explored TEMs and Extension One it is assumed that the criteria for achieving an auth is the same for all employees independent of whether they have had the auth before or not. The second extension researched uses two criteria for each auth. The first criteria will need to be satisfied if the employee did not have the skill at the last review, defined as an *initial* criteria. The second criteria will need satisfying if they are reviewing a skill the employee currently has, defined as a *recurrent* criteria. In this way less training can be specified if the employee already has the skill.

Though formulated in the learning curve, no paper could be found that considered the need for recurrent training when skills are binary. For learning curves the need for training decreases as more learning takes place. However, in many cases the need to complete training is dependent on an auth being reviewed or earned for the first time, hence the motivation for a mixed integer approach.

A skill is defined as recurrent, $\sigma = 2$, if the employee had it at the last review period, $s_{ek(z-1)} = 1$ and initial, $\sigma = 1$, otherwise. This is a reasonable assumption for large review times as it ensures they have received an appropriate level of training given that they have not held the skill for a large amount of time. Thus, the recurrence of the skill is memoryless and does not consider the auth ownership previous to the last review period.

To apply this extension it is assumed that skills are either recurrent or initial. No other type of skill that may affect the training requirements is defined.

The extension requires a reformulation of the current TEMs and the previous extension, Section 4.3.1, if used. First, the experience threshold is amended. The given experience threshold g_k used previously is redefined as $g_{k\sigma} \in \mathbb{Q}^+$ where $\sigma = 1$ is the initial criteria and $\sigma = 2$ is the recurrent criteria. Applying this change to equation (4.12) produces two equations:

$$\begin{aligned} g_{k1} - \boldsymbol{M}_k(1 - s_{ekz}) \leq &\sum_{t \in Z_{ek(z-1)}} (w_{ekt} + p_{ekt} \\ &+ \sum_{n \in N} \chi_{nk} v_{ent}) + M_k s_{ek(z-1)} \qquad \forall e \in E, k \in K, z \in \overline{Z} \quad (4.21) \\ g_{k2} - \boldsymbol{M}_k(1 - s_{ekz}) \leq &\sum_{t \in Z_{ek(z-1)}} (w_{ekt} + p_{ekt} \\ &+ \sum_{n \in N} \chi_{nk} v_{ent}) + M_k(1 - s_{ek(z-1)}) \qquad \forall e \in E, k \in K, z \in \overline{Z} \quad (4.22) \end{aligned}$$

where only one constraint may be binding at once due to the value of $M_k = \max_\sigma(g_{1\sigma})$. For equation (4.21), if $s_{ek(z-1)} = 1$, then the constraint becomes a non binding constraint. Similarly for equation (4.22) if $s_{ek(z-1)} = 0$. As seen in the TEMs, the achievement of the auth is dependent on the experience gained through the various TEMs exceeding the required threshold value g_{k1} or g_{k2}.

In this way, though employees may complete the same training options, they may obtain auths differently if one had previously held the auth.

The additional criteria may also affect the previous extension for compulsory training if Extension One is being used, Section 4.3.1. The compulsory training may be different for initial and recurrent auths. To reflect this, redefine the compulsory training decision variable q_{ekt} as $q_{ekt\sigma} = \{0,1\}$ and the compulsory criteria π_k as $\pi_{k\sigma} = \{0,1\}$. Equation (4.20) is transformed into two equations;

$$\pi_{k1}s_{ekz} - \sum_{t \in Z_{ek(z-1)}} q_{ekt} - s_{ek(z-1)} \leq 0 \qquad \forall e \in E, k \in K, z \in \overline{Z} \quad (4.23)$$

$$\pi_{k2}s_{ekz} - \sum_{t \in Z_{ek(z-1)}} q_{ekt} - (1 - s_{ek(z-1)}) \leq 0 \qquad \forall e \in E, k \in K, z \in \overline{Z} \quad (4.24)$$

Again, only one of these constraints can be binding at once due to the binary value of the skill. These constraints are only necessary for k such that $\pi_k = 1$.

Finally, the constraint on availability, equation (4.9) must be adjusted to consider the time to complete compulsory training given the different training times for initial or recurrent auths, $\mu_{k\sigma} \in \mathbb{Q}^+$.

$$\sum_{k \in K} \eta_k (w_{ekt} + p_{ekt}) + \sum_{n \in N} \rho_n v_{ent} + \sum_{k \in K} \sum_{\sigma=1}^{2} \mu_{k\sigma} q_{ekt\sigma} \leq a_{et} \qquad \forall e \in E, t \in T \quad (4.25)$$

The training costs are updated to include the cost of compulsory training, $c_{3k\sigma} \in \mathbb{Q}^+$ given the difference in value for initial or recurrent auths.

$$\sum_{e \in E} \sum_{t \in T} \left(\sum_{k \in K} c_1 p_{ekt} + \sum_{n \in N} c_{2n} v_{ent} + \sum_{k \in K} \sum_{\sigma=1}^{2} c_{3k\sigma} q_{ekt\sigma} \right) \leq b \qquad (4.26)$$

4.3.3 Extension Three - Skill Dependency

The next extension considered is auth dependency. This means that some auths can only be gained if the employee already has obtained other specified auths. This is useful if tasks are similar and thus training for one of those auths will allow them to gain the other auths. Alternatively, tasks may be complex and require the experience of performing other tasks. This could be seen in engineering where you would be unable to perform tasks on upgraded technology until you are able to perform tasks on the previous version. Training may still need completing for the dependent tasks but it is also compulsory that the training for the initial task should be completed.

To implement this extension, the problem may be defined as a directed acyclic graph modelling precedence constraints. It is assumed that the dependencies cannot cause the corresponding network to become cyclic. Furthermore, at least one task must have no dependencies, such that a starting node exists.

A new data value $d_{ki} \in \{0, 1\}$ is proposed, where $i \in K$. The data is equal to 1 if auth k is dependent on the employee also having auth i and 0 otherwise. The following constraint is included in the Training Development Model.

$$s_{ekz} - s_{ei\tilde{z}} \leq 1 - d_{ki} \qquad \forall e \in E, k \in K, z \in \overline{Z}, i \in K, \tilde{z} \in Z_{ekz} \cap Z_{ei\tilde{z}} \tag{4.27}$$

where the independent auth must be held in the intersection of the two skills review periods.

4.3.4 Extension Four - Skill Speciality

Previously, it was assumed that all employees are the same, such that they only differ in which tasks they can do. However, for some companies employees may differentiate by levels or by type.

Many companies use skill levels, such as supervisors and technicians, to differentiate between employees. The level of the employee may be determined by their training, education or time within the company. It allows us to also differentiate between tasks such that a task may be more complex than others. Skill levels are an example of a hierarchical skill as defined in Section 2.2.2. In this way, substitution may also be implemented by stating that a task that can be completed by an employee of a given level may also be completed by employees of higher levels.

By differentiating by type, a more complex workforce and task base can be modelled. For instance, tasks may be type specific such that workers who have the auth and are of the same type are required for specified tasks. Such as surgical tasks only being assigned to employees who have surgical skills. However other tasks may not require a specialist worker and thus can be completed by anyone who has the auth. These are defined as categorical skill groups where the auths themselves are categorical. The skills are now grouped into categories of specialities where tasks may belong to more than one of these groups if applicable.

In this extension skill speciality of the employees is fixed. It is assumed that the decisions required to promote an employee are too intricate for the training model and will not be affected by the allocations to training or tasks. In this way the decision is made at a managerial level rather than through modelling.

It is also assumed that each employee can only belong to one skill speciality. If the company uses both skill levels and types, these may be defined as different specialities. For instance, surgical intern, surgical consultant, diagnostics intern and diagnostics consultant would represent four specialities.

Though temporal workers have not been modelled specifically, they can be defined as a 'speciality'. The model will then allow temporal workers to be available to complete some specified tasks.

To implement the extension new data values are required. The value $\lambda_{el} \in \{0, 1\}$ equals 1 if employee e is of speciality $l \in L$ where L represents the set of different specialities. The value $\phi_{lk} \in \{0, 1\}$ equals 1 if task k requires completion by an employee of speciality l.

It is assumed that $\sum_{l \in L} \lambda_{el} = 1$ for all employees, however $\sum_{l \in L} \phi_{lk} \geq 1$ for all tasks. In this way an employee can only belong to one speciality however tasks may be completed by more than one type of person. This allows us to implement substitution as discussed in Section 2.2.2. It may be specified that a task can be completed by any employee that is a technician or higher for instance. In addition, a task that must be completed by two specialities may be defined twice with each requirement separately specified.

The following constraint is added to the Training Development Model:

$$w_{ekt} - r_k \sum_{l \in L} \lambda_{el} \phi_{lk} \leq 0 \qquad \forall e \in E, k \in K, t \in T. \quad (4.28)$$

Thus, constraining the tasks to only be completed by employees of a specified level or speciality.

4.3.5 Extension Five - Non-Linear Learning

Learning has been considered in training models since Wright (1936). It plays an important role in developing the skills of employees and defining the dynamic nature of an employee's capabilities. In previous sections, it is assumed that experience gained is independent of the number of times the skill is used.

In the literature, Section 2.4, the learning curve was discussed in detail, such that the time taken to complete a task decreases the more times the task is completed. However, the rate in which the time taken decreases is not linear. The model shows that the decrease is larger when the task is first completed. The decrease is smaller the more times the task is completed. This theory is shown in Figure 4.1.

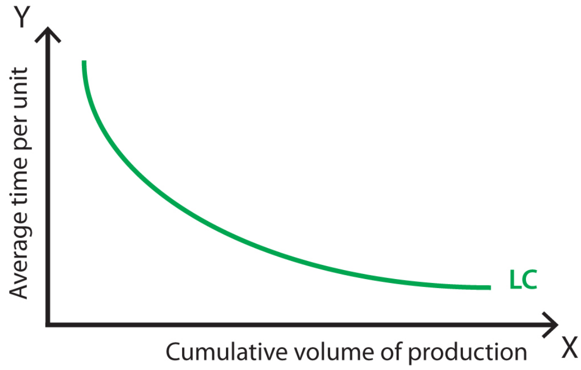

Fig. 4.1: Wright's Learning Curve (Policonomics, 2012)

Where the learning curve in the literature considers time taken to complete a task as the dependent variable, here the required experience to gain an auth is the dependent variable. In both models, the number of times a task is completed is the independent variable and the shape of the curve is assumed to remain similar due to the nature of learning.

To apply this theory to the TEMs, alterations must be proposed. Currently, the experience increases linearly each time the skill is used. However, experience can be considered as changing similarly to the time taken to complete a task. The more times a task is completed, the more competent an employee is thus less experience is required to gain the auth. Similar to Extension Two, the previous experience of the employee affects the auth requirements, however in this extension the magnitude of the effect is dependent on the quantity of experience and not on the binary value of skill in the previous time unit.

To implement this idea, experience required for an auth to be allocated, g_k, is redefined. Currently, the value here only changes depending on the task. By introducing $x \in X$ as the number of times the auth is used, a piecewise experience curve is created, $g_{kx} \in \mathbb{Q}^+$, representing a non linear function. A new decision variable $y_{ekzx} \in \{0, 1\}$ is equal to 1 if employee e has used the skill, k, x times up to the beginning of the review period z. In this way $\sum_{x \in X} g_{kx} y_{ekzx}$ gives the required experience given the previous experience of the employee.

For each company the shape of this learning curve will vary. As an example, consider the learning curve in Figure 4.2 defined as $g_{kx} = -1(0.9^{(z-x)}) + c$, where x is the number of times the skill is used. The parameters z and c are calculated by the required shape of the curve. It is assumed that when $x = 0$, y is equal to the maximum experience required. When $y = 0$, it is assumed that x is equal to the maximum number of times the task requires completing in order for an employee to be considered competent, thus requiring 0 experience to renew a skill. From these two assumptions the values of z and c may be calculated. In the following graph, $z = -17.66$ and $c = 8.47$ is required to obtain the shape. The graph is the same for all k.

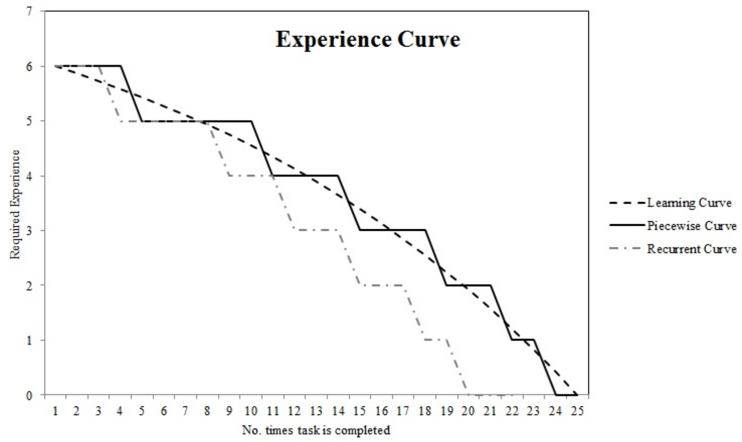

Fig. 4.2: Learning curve $g_{kx} = -1(0.9^{(-17.66-x)}) + 8.47$ using random data

To use non-linear learning in the model, it is assumed:

- All employees follow the same learning pattern
- No learning takes place before the time horizon
- The learning curve is fixed before the start of the time horizon
- Learning only affects the competence of the employee in terms of skill required
- Experience is not affected by the time between using the skill
- The position on the learning curve is dependent on the allocations over the time horizon. That is, allocations during the first review period impact the required experience in the final review period.

The skill constraint, (4.12), is written as:

$$\sum_{x \in X} g_{kx} y_{ekzx} - M_k(1 - s_{ekz}) \leq \sum_{t \in Z_{ek(z-1)}} (w_{ekt} + p_{ekt} + \sum_{n \in N} \chi_{nk} v_{ent}) \qquad \forall e \in E, k \in K, z \in \overline{Z} \tag{4.29}$$

where $M_k = \max_\sigma(g_{1\sigma})$ as defined in Section 4.3.2. Here, the experience is calculated using all TEMs. Thus, if the right hand side is greater than the value of $\sum_{x \in X} g_{kx} y_{ekzx}$ then $s_{ekz} = 1$ and the threshold experience required has been exceeded, else $s_{ekz} = 0$.

Two constraints are required to model the calculation of the decision variable y_{ekzx}, these are written as:

$$\sum_{x \in X} x y_{ekzx} = \sum_{t=1}^{\max(Z_{ek(z-1)})} w_{ekt} \qquad \forall e \in E, k \in K, z \in \overline{Z} \quad (4.30)$$

$$\sum_{x \in X} y_{ekzx} = 1 \qquad \forall e \in E, k \in K, z \in \overline{Z} \quad (4.31)$$

This ensures that the experience required is calculated from the number of times the skill has been used from the beginning of the time horizon to the time of review z.

If used in conjunction with Extension Two, recurrent skills, a second learning curve will need to be considered. Here, the curve for initial skills will differ from the curve for recurrent skills similarly to the reasoning used in Section 4.3.2; if the skill is being renewed, less experience will be required.

Various learning curves were considered for this thesis. In all cases, experience is used as the variable in the curve.

The learning curve can represent the penalty on experience such that the more times a skill is used, the less experience is gained. In this scenario, the curve will need to increase exponentially with the maximum penalty being 100%. Assume employee one has completed the task many times, this indicates a large penalty on the experience gained. Employee two is completing the task for the first time thus has no penalty on the experience gained. According to this curve, employee one will need to complete the tasks multiple times in order to gain the same experience as employee two.

Though this is logical in terms of learning, it is not a realistic when considering experience as a measure of a person's competency in gaining a skill. The model may be more applicable to experience in relation to skill levels; the ability to gain low levels is much easier than gaining higher levels.

This penalty may be considered in terms of forgetting. The penalty on experience is dependent on the number of days the skill is not used. Again, this is a logical representation of the learning process as skills are lost through lack of use. However, this is much more difficult to quantify. Does forgetting begin from the moment the skill is gained? Would the length of the gap affect the amount forgotten? Does the penalty reset every time the skill is used?

Thus, it was concluded that the learning curve should represent the experience required to gain a skill. Here, the more times a skill is used over the time horizon, less work is required to prove competency in an employee. Ideally, this should be combined with a forgetting penalty such that for long time horizons an employee cannot take advantage of experience gained from tasks many months before.

4.3.6 Extension Six - Stochastic Tasks

For industries such as healthcare and defence, the nature of the tasks are stochastic. It is not known exactly which tasks will need completing at every time unit. However, if the probability of the tasks needing completing can be estimated, a workforce can be recommended to complete a percentage of the tasks within a given confidence interval.

According to Murray (1987), defence can be difficult to model due to the complexity caused by stochastic events. The paper explains that solutions differ based on the current scenario or state of the world, for instance, war or humanitarian missions. Indeed, the scenarios are much more detailed, they are dependent on where the combat is taking place, what forces are present and what the tactical plan is. The choice of scenario affects all operations from workforce requirements to logistics. Each scenario produces different levels of uncertainty within the data.

In order to integrate the stochastic nature of the world, an optimal solution to satisfy uncertainty in the data is proposed. As a strategic model is considered, it must include all possible scenarios that could occur over the course of the model. The scenario determines the probability of each task needing completing and the requirements. The workforce allocation model will be required to reserve workforce for the tasks given these probabilities. Distributions will need to be constructed from past data to calculate the probabilities in each scenario.

The stochastic environment also allows us to consider an overestimate for the number of employees hired. It should be ensured that there are enough employees available with the correct skills at any moment over the time horizon to fulfil the unexpected and planned tasks.

Previously, the skill gap has been considered as a soft constraint. However, to minimise the risk of under supplying workforce given stochastic tasks, the skill gap is now redefined as a hard stochastic constraint.

The probabilities may not be fixed for all time but may be influenced by the state of the world. Let ξ be a random variable defining the state of the world. This has a direct impact on the probability of a task needing completing, $\mathbb{P}(k, \xi)$.

To account for the stochastic nature of our training problem, a chance (probability) constraint is included. Through allowing multiple states of the world to be considered as scenarios, an optimal solution may be produced that is acceptable in a required percentage of scenarios.

The aim is to find a workforce solution that considers the stochastic tasks for each scenario. The different techniques for solving stochastic programming problems are given in Section 3.4 and it was determined that chance constraints are the best method for formulating the model for this case.

Using the Sample Approximation Approach (SAA) as discussed in 3.4, an approximation of $\mathbb{P}(k, \xi)$ can be given using a sample of the random vector ξ. However, rather than using a random sample, a sample is selected such that each scenario $\xi_i, i = 1, .., m$ represents a specific event or scenario. For example, in the defence case study, each state of the world may be described by weather condition and mission type. Thus $\mathbb{P}(k, \xi)$ can be estimated using the probability that task k is required in scenario $\xi_i, \gamma_{k\xi_i}$, as sampled from the data. In this way the probability of each scenario occurring in the time horizon may be utilised. Though Luedtke & Ahmed (2008) uses a random sample for the scenarios, this method is more appropriate to the problem at hand and the indicator function will be weighted accordingly.

The data input is given as follows:

- $\alpha \in \mathbb{Q}^+$: Confidence parameter.
- $\beta \in \mathbb{Q}^+$: Maximum proportion of tasks that may not be complete due to lack of skill.
- $\zeta_{\xi_i} \in \mathbb{Q}, 0 \leq \zeta_{\xi_i} \leq 1$: Probability of scenario ξ_i.
- $\gamma_{k\xi_i} \in \mathbb{Q}, 0 \leq \gamma_{k\xi_i} \leq 1$: Probability of task k requiring completing in scenario ξ_i.

A solution will be required that ensures that no more than β% of the tasks are incomplete in α% of the scenarios at the last time unit $\hat{t}$, where α and β are given. This allows training to take place without affecting the optimal solution which is calculated at the end of the horizon. Thus the time horizon defines the training horizon. The number of tasks satisfied must consider the varying probabilities of tasks occurring given the state of the world and the probabilities of each scenario occurring. Hence allowing a solution that is optimal in most scenarios.

This can be written as:

$$\mathbb{P}(G(w, \xi) \leq \beta) \leq 1 - \alpha \tag{4.32}$$

where

$$G(w) = \gamma_{k\xi_i} \frac{\sum_{k \in K}(r_k - \sum_{e \in E} w_{ek\hat{t}})}{\sum_{k \in K} r_k} \tag{4.33}$$

is the proportion of incomplete tasks . The numerator defines the skill gap resulting from unfulfilled tasks and the denominator defines the total required skill for all tasks.

However, using the techniques discussed in Section 3.4, a new counting decision variable is proposed, $\delta_{\xi_i} = \{0, 1\}$, equal to 1 if the probabilistic constraint is violated and the last time unit $\hat{t}$, to count the number of times the constraint is violated. The constraints are written as:

$$\sum_{k \in K} \gamma_{k\xi_i}(r_k - \sum_{e \in E} \sum_{k \in K} w_{ek\hat{t}}) \leq \beta \sum_{k \in K} \gamma_{k\xi_i} r_k + M\delta_{\xi_i} \qquad \forall i \in 1, .., m \tag{4.34}$$

$$\sum_{i=1}^{m} \delta_{\xi_i} \zeta_{\xi_i} \leq 1 - \alpha \tag{4.35}$$

where α is the given allowed confidence interval. Thus, if $\alpha = 0.95$ the constraint may be violated in no more than 5% of the scenarios, given the weighted probability of a scenario occurring. The constant $\hat{t}$ indicates the last time unit such that the constraint only holds in the last instance of time. This is a joint probability constraint with stochastic left hand side.

The new variable δ_{ξ_i} should be added to the objective function such that the number of violations are minimised. Though the problem can still be solved, this addition reduces the run time of the model. Consequently, the multi-objective model now contains three objectives. Thus, this value will require a weighting coefficient. The value is selected to always prioritise the value of δ over the other objectives.

Due to the addition of these constraints and the loss of a soft constraint for skill gap, infeasible solutions may now be obtained.

In addition to the chance constraint, certain constraints should be adapted to allow for the probability of a task occurring. For instance, as tasks do not always appear, the availability of the employee is dependent on this probability. The new availability constraint is given as follows.

$$\sum_{k \in K} \gamma_{k\xi_i} \eta_k (w_{ekt} + p_{ekt}) + \sum_{n \in N} \rho_n v_{ent} \leq a_{et} \qquad \forall e \in E, t \in T, i \in 1, .., m \tag{4.36}$$

If an employee is assigned to a task, the processing times of the tasks that the employee can complete is examined and weighted by the probability of the tasks occurring in each scenario. For instance if all tasks had 100% chance of appearing, the employee cannot perform more tasks than can be processed in their available time. However, if all tasks had a 50% chance of appearing, they can allocate twice the number of tasks to the employee as 50% of those tasks will not be expected to occur. This constraint must be satisfied for all scenarios thus the availability cannot be violated in any circumstance.

Similarly, the experience gained from performing a task is no longer equal to one as the task may not need completing. Thus, the experience is increased by the probability of the task over all scenarios. The experience itself will also require adjustments. If a task is not expected to need completing as regularly as the tasks considered in previous extensions, then less experience should be required to prove it needs completing. The amount of experience expected should be less when the task is stochastic. Though this does not change the usage of the data g_k, the values may want to be amended to account for this difference in the nature of the tasks.

$$g_k - M_k(1 - s_{ekz}) \leq \sum_{t \in Z_{ekz}} (\gamma_{k\xi_i}(w_{ekt} + p_{ekt}) + \sum_{n \in N} \chi_{nk} v_{ent}) \qquad \forall e \in E, k \in K, z \in \overline{Z} \tag{4.37}$$

4.4 Summary

In this section three training execution methods and six model extensions have been introduced that can be incorporated in the TNA to create a decision support tool for workforce strategy. These training policies can be selected individually or as a combination to model the authorisation criteria for each task.

The three TEMs were used to describe how experience is gained through various training. The methods considered were repeating tasks, on-the-job training and training courses.

The first method uses repeating tasks to gain enough experience to succeed the threshold amount required to gain a skill. The skill gap at each unit is affected by these allocations.

The second TEM allows experience to be gained through OJT. Here, an employee must be available to supervise OJT.

The final method allows external training courses to increase the experience of the employees. Different training types can be considered, each with a different cost, duration and experience gain for each auth. New decision variables and constraints need to be added to the initial model.

In addition to the TEMs, six extensions were created to include additional constraints to the basic model.

Compulsory training can be enforced for certain auths. Here, the auth can only be gained if the training has been complete. The training has an associated cost and duration.

The criteria required to complete an auth may not be the same for someone who currently uses the skill and someone who does not. Initial and recurrent skills are defined, with different experience and compulsory training requirements.

Further constraints on the auth can be implemented in the form of dependencies. This means some auths can only be given if other auths have already been obtained.

A rate of learning is introduced to the model. Here, the increase in skill is not considered linearly, rather the rate of learning is formulated as a function of the number of times the employee has completed the task.

Finally stochastic tasks are defined such that the probability that a task will need completing in each moment of time is represented by a random variable. Chance constraints are used to model this variation.

This model allows clients to choose the methods and extensions relevant to their company in any combination. Thus, there are a possible 256 combinations that can be modelled. However, though many combinations exist, it is relatively simple to create a model that combines the necessary TEMs and extensions by combining the constraints and decision variables proposed in this section. Due to the modular nature of the mathematical programme, constraints can be simply added and adjusted as required without making changes to the solution method.

Through this, an analysis can be done to explore other training policies not currently in place without the need to change the solution method. The same software will still be used, and the tool will be adaptable to allow a user to implement the changes.

In this section it has been shown how each of these policies can be modelled into our decision support tool by changing the constraints currently in the model and adding new ones.

Next, data will be applied to each of these policies in turn and the implications will be analysed.

5. NUMERICAL EXPERIMENTATIONS AND RESULTS

In this chapter, the models are investigated further through the use of simulation. Due to the mixed-integer nature of the models, standard shadow pricing and other investigative techniques are not suitable to analyse the models. Thus, Monte Carlo simulation is used to produce samples of data to input into the models. This results section is essential in the verification and validation of the model.

The Training Development Model will be analysed, considering each training method and extension in turn as well as all models being used in combination, thus analysing 10 model combinations.

The analysis has been performed on the IRIDIS High Performance Computing Facility, with associated support services at the University of Southampton. The optimisation is executed using IBM ILOG CPLEX version 12.7.1. To obtain the random data required to perform the Monte Carlo simulation, a C++ executable has been created in Microsoft Visual Studio.

No manual alterations have been made to the Cplex parameters with the exception of the wall time for each optimisation. However, the MIP emphasis will be discussed in Section 5.1.

Through experimentation, it can be shown that in most cases, the solution determined within the first 600 seconds of the model does not differ to the solution when the model is run for a longer time period. Figure 5.1 shows the optimality gap of 10 runs over two hours. It contains five instances where the model found an optimal solution and five instances where the model could not be solved in two hours. It can be seen that in the instances where the model did not solve, the optimality gap did not change after the first few seconds, whereas the solvable problems were solved in a few seconds. Thus, it appeared sufficient to set the wall time in this section to 600 seconds, given the large difference in solution times. It is assumed that any runs longer than this are having problems proving optimality. The experimentation here was completed on problems of various sizes and not exceeding the largest problem size explored in this chapter.

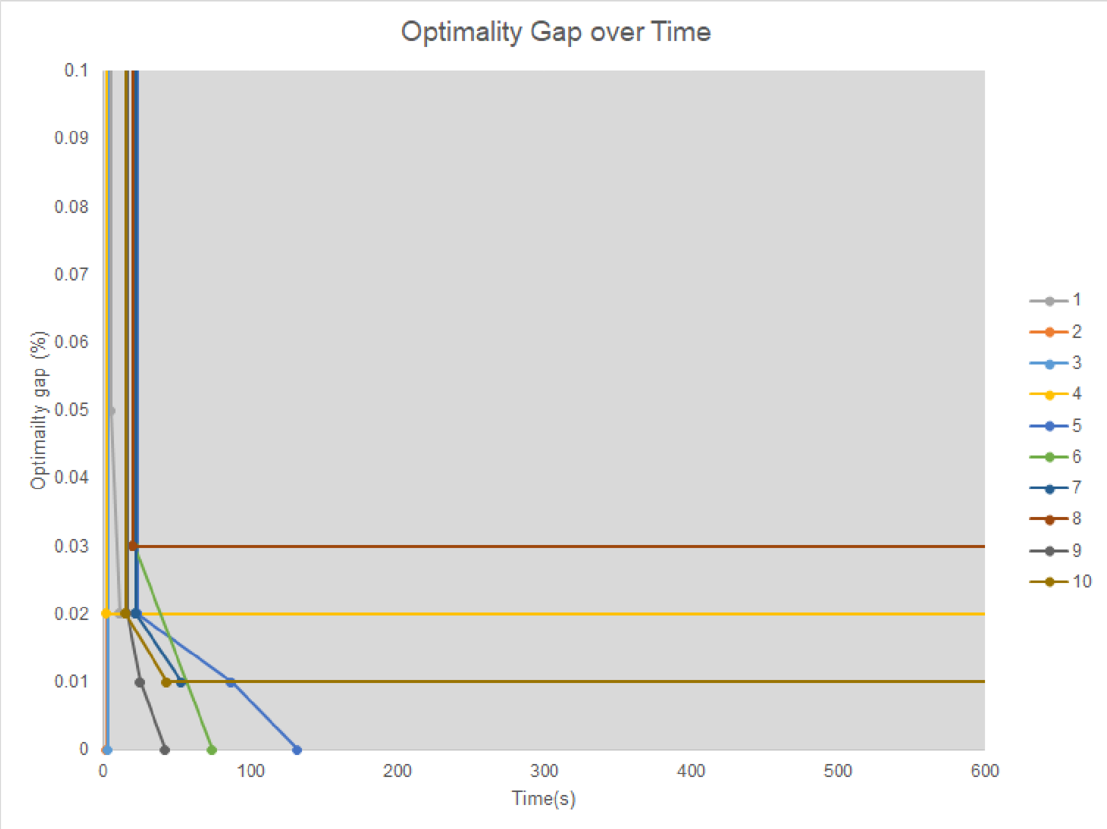

Fig. 5.1: Optimality Gap over Time

The results indicate that a solution is optimal and can be found very quickly, but the solver cannot prove it. This will be explored later in this chapter.

This wall time will be consistent in all models unless stated otherwise. A run time analysis will be presented under each model, however a comparison of model run times will be made at the end of the chapter.

Suggestions to improve this run time are provided in the next section. These adaptations will be applied to all models. The training budget will also remain constant for all models unless otherwise stated. An analysis of the training budget is completed to determine the appropriate budget for the study given the data samples used.

The first set of results are obtained from a coefficient sensitivity analysis. Here, a specific set of coefficients will be selected for each model. The values of these coefficients will be varied whilst the remaining coefficients will remain constant. In this way, the behaviour of the model under the changes to the coefficients may be examined. The study will investigate which coefficients and thus which aspects of the data have a greater affect on the solution calculated.

Shadow price analysis and reduced cost analysis are not available with mixed integer models. Though alternative methods exists, the number of model combinations available here mean these methods are difficult to implement. However, correlation and regression can be used to analyse relationships between the dependent decision variables, independent coefficients and right hand side values.

For each model, details of the data used will be given and the dimensions of the model will be stated as defined in Section 4. Due to the modelling constraints used, Extension Six is the only problem where infeasible solutions may occur for certain changes to the coefficients. Here, there may be no solution to satisfy the chance constraints. For all other models a trivial solution of no tasks and no training is possible.

The model presented in this thesis is a multi-objective model. As such, a Pareto curve can be produced for every run of the model dependent on the choice of coefficients in the objective function. An analysis of this coefficient will be completed, thus determining the relationship between the two objectives; minimising the skill gap and minimising the fairness deviation for the employees.

Finally, a run time analysis will be performed on all models using randomly selected data for all coefficients. The distributions of the run times will be compared between each model and discussed. Changes to the problem size will be made and the impact on the number of decision variables, constraints and run time will be analysed.

This analysis section is not a thorough analysis of the models. Each analysis looks at only a subset of all the problems using specific problem sizes and specific data. Thus, they do not represent the true nature of the model but act as an insight into the validity of the results, through an analysis of the output, and highlight interesting behaviours within the model.

5.1 Model Adaptations

Given the complexity and size of this model and its extensions, adaptations may be required to reduce the run times of the model. Here, changes to the model variables and run settings are discussed. These are applied before the analysis is carried out and could reduce the run time of the model.

First, note that the model is mixed-integer where some of the decision variables are binary and some are integer. The model may be adapted to remove any integer decision variables and replace them with binary decision variables. This is done through an additional index m such that

$$\begin{aligned}\sum_{m\in M} m\overline{x}_m &= x \\ \sum_{m\in M} \overline{x}_m &= 1\end{aligned} \tag{5.1}$$

where $x \in \mathbb{Z}^+$ is a decision variable that can take any positive integer value and $\overline{x}_m \in \{0,1\}$ is a binary decision variable. This technique can be used for any positive integer decision variable under the condition that the equation (5.1) is added to the model as a constraint. Though this increases the number of decision variables by a multiple of m, it has been found to decrease the run times for models. The number of indices, $\hat{m}$, should be selected such that $\hat{m} = \max x$. In this way, the minimum number of decision variables is added to the model.

Other methods exist to transform variables to binary variables. For instance, transforming $x \in \{0,n\}$ into $\log n$ binary variables by substituting x for the expression $\sum_{j=1}^{\log n} 2^j z_j$ such that $\sum_{j=1}^{\log n} z_j = 1$. However, this will require slight adjustments dependent on the value of n. In the selected method for this thesis, as above, one solution may easily be applied to all extensions and models.

Klotz & Newman (2013) explains that binary programmes benefit from some techniques in the literature that are not available to integer programmes. In the case of branch and bound, branching on a binary variable fixes the variable on the branch to be 1 but also fixes the associated variables to 0 due to the constraint that they must sum to 1. Note that other methods exist to change integer to binary, however this method could be quickly implemented into the model. This method does increase the number of decision variables in the model more than other methods.

Experimentation was completed on 200 runs of data with a problem size [5,10,6]. The results, shown in Figure 5.2 show an improvement in the average and maximum run times of the experimental data. With data shown to have difficulties proving optimality, the method was shown to make no difference. As this method has a large implication on the number of variables in the model, it is suggested that this transformation be revisited, as advances in linear programme solvers and updates to the model may affect the results.

For this chapter, any integer decision variables will be adapted in the model using the above method and an appropriate value for $\hat{m}$ will be selected. However, for understanding and neatness, the model formulations only contain the original mixed-integer model. When discussing number of decision variables, the number will include the additional variables created in this way. Run times will be affected and this technique will remain consistent through all analysis.

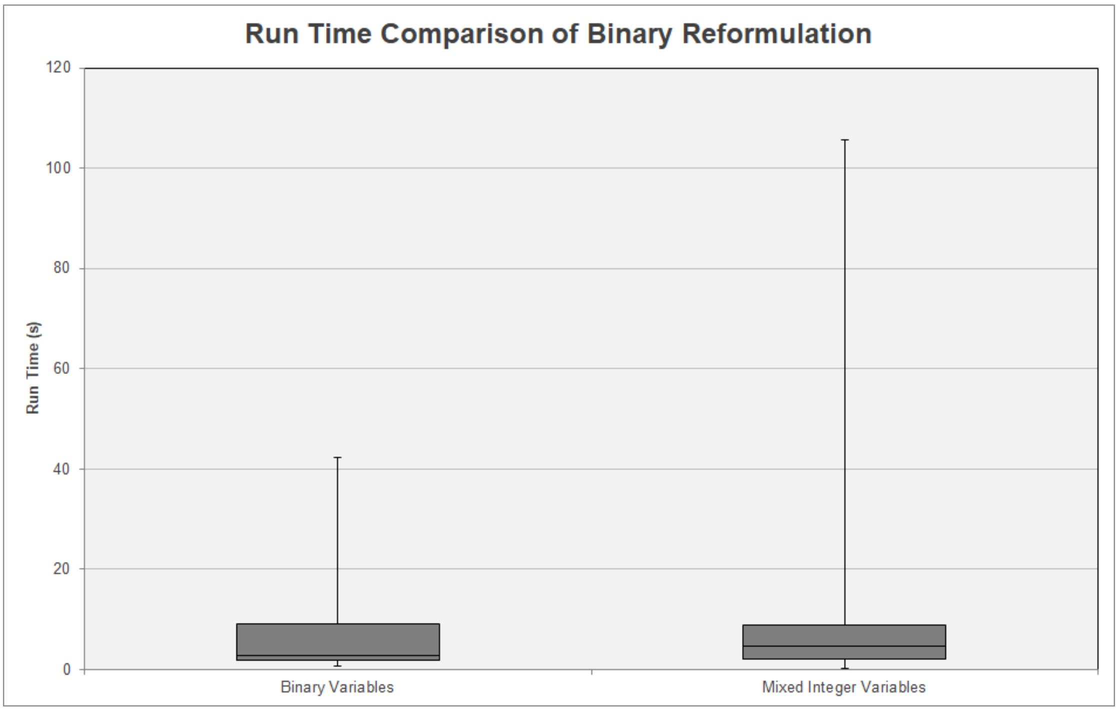

Fig. 5.2: Run time comparison of binary reformulation

In each of the models "big M" is used as a constant to indicate where a large number is required in order to ensure the constraint is satisfied. As with the value of $\hat{m}$ stated above, the value must be selected using appropriate reasoning in order to minimise the run times of the model. In this way, the value of big M will differ dependent on the constraint. Geyer, Hiller & Meinert (2010) recommends setting the value as small as possible for the constraint, such that the required result remains the same.

To solve this model, IBM Cplex is used as the solver. It is recommended that other solvers should be considered to determine if this is the best solver for the problem. Though the default settings in Cplex should provide the best method to solve this problem in a sufficient time, changing these parameters could influence the run time of the model.

First, the MIP emphasis was adjusted and the results analysed. During experimentation on the model, it was noted that some models would run indefinitely. In these instances, the output from Cplex indicated that the solver had difficulties in proving optimality rather than obtaining a solution. Indeed, the solution found after a few seconds did not differ through the remaining run time. To experiment with the MIP emphasis, a set of data is chosen in which the run time exceeds a wall time of 1200 seconds. For each optional value of the MIP emphasis, the model is run for 1200 seconds and a record is made of the objective function and the relative gap, calculated as $(bestinteger - bestnode) * objsen/(abs(bestinteger) + 1e - 10)$ (IBM, 2017). The results can be seen in Table 5.1. For more information concerning the MIP Emphasis in Cplex, the reader is referred to IBM (2017b).

Value	Description	Objective	Gap
0	default balanced	1,282	4.02%
1	feasibility	1,282	6.87%
2	optimality	1,288	2.74%
3	best bound	1,282	4.34%
4	hidden feasibility	1,282	4.09%

Tab. 5.1: Results from MIP emphasis analysis

First, consider the objective value; all but one option gives the same output. It is assumed that $1,282$ is the optimal solution. For MIP Emphasis value 2, emphasising optimality over feasibility, the objective value is higher than the other results after 20 minutes. However, the relative gap is much smaller. This indicates that this option would reduce the time needed to prove optimality but does not give the optimal solution as quickly as the other options. For the minimum relative gap after 20 minutes, the default value gives the best result. The default value finds the best balance between feasibility and optimality. Hence, the default setting will be adopted in this problem. This setting will be implemented in all tests through the remainder of this study.

In the next section, the cause of the long run times and issues proving optimality are reviewed and analysed.

5.2 Random Data for Analysis

To perform an analysis on these models, data is required for Monte Carlo simulations. This data is randomly selected from distributions based on realistic scenarios of the problem considering the data obtained from Boeing and feasible data quantities. For instance, the available hours in a day does not exceed 24. The distributions are selected to avoid infeasible or trivial solutions. The data is also selected to ensure a range of possible input is explored to test the model as completely as possible with the available techniques.

Variable	Notation	Distribution
Initial skill	Θ_{ek}	Bernoulli(0.5)
Requirements	r_k	Uniform(1,3)
Processing time	η_k	Uniform(1,20)
Exp threshold	g_{k1}	6
Review time	$max(Z_{ek1})$	Uniform$(2, \hat{t})$
Training budget	λ	1000
OJT cost	$c^{'}$	1
Training course cost	$c^{''}$	Uniform(1,100)
Training course exp	χ_{nk}	Uniform(1,3)
Training course length	ρ_n	Uniform(1,8)
Training course capacity	u_n	Uniform$(1, \hat{e})$
Comp. training cost	$c^{'''}$	Uniform(1,100)
Comp. training required	π_k	Bernoulli(0.5)
Auth dependency	d_{kk}	Bernoulli(0.5)
Employee skill level	b_{el}	$P(b_{e1} = 1) = 0.3$, $P(b_{e2} = 1) = 0$
		$P(b_{e1} = 1) = 0$, $P(b_{e2} = 1) = 0.3$
		$P(b_{el} = 1) = 0.3\ \forall l$
Task skill level	ϕ_{k1}	Bernoulli(0.5)
	ϕ_{k2}	$1 - \phi_{k1}$
Prob. of scenario	ζ_{ξ_i}	0.01 Uniform(1,20), $i \neq \hat{i}$
	$\zeta_{\xi_{\hat{i}}}$	$1 - \sum_{i=1}^{\hat{i}-1} \zeta_{\xi_i}$
Prob. of task	$\gamma_{k\xi_i}$	0.1 Uniform(1,10)

Tab. 5.2: Summary of random variables used in analysis

The distributions will be used to generate all random data in the model unless stated otherwise in the section, this allows continuity between the analyses. Table 5.2 contains details of the distributions. All possible inputs for all models are defined in this table. In most cases a Uniform distribution is used for the data to ensure the parameters are selected equally.

It is assumed that only one review period exists within the time horizon and no more than two skill specialities will be used in Extension Four. Unless stated otherwise, the number of training courses is also equal to two and five scenarios are sampled in the stochastic constraints. In addition, it is noted that the probabilities of the scenarios are set such that they sum to one in Extension Six.

For Extension Five, non linear experience is modeled with a function of a specified shape. The parameters of this function are selected under the following conditions. It is proposed that an employee with no prior experience must obtain the six time units of experience stated above. It is assumed that an employee who has completed the task 24 times is considered competent in the skill and not need any experience in the next review to obtain an auth. If the time unit was set to months, this would be approximate to two years of repeating the tasks in every time unit. It is assumed that no forgetting takes place and once an employee is competent, they cannot lose the knowledge.

It is also assumed that the curve fits the shape:

$$f(x) = -1(0.95^{(-a-x)} + b) \tag{5.2}$$

where x is the number of times the task has been completed, $f(x)$ is the experience required to gain the auth given x and a, b are parameters to be determined.

Given the assumptions above, set $f(0) = 6$ and $f(24) = 0$. Through substituting and solving these simultaneous equations it is determined that $a = 17.66$ and $b = 8.47$ to two decimal places.

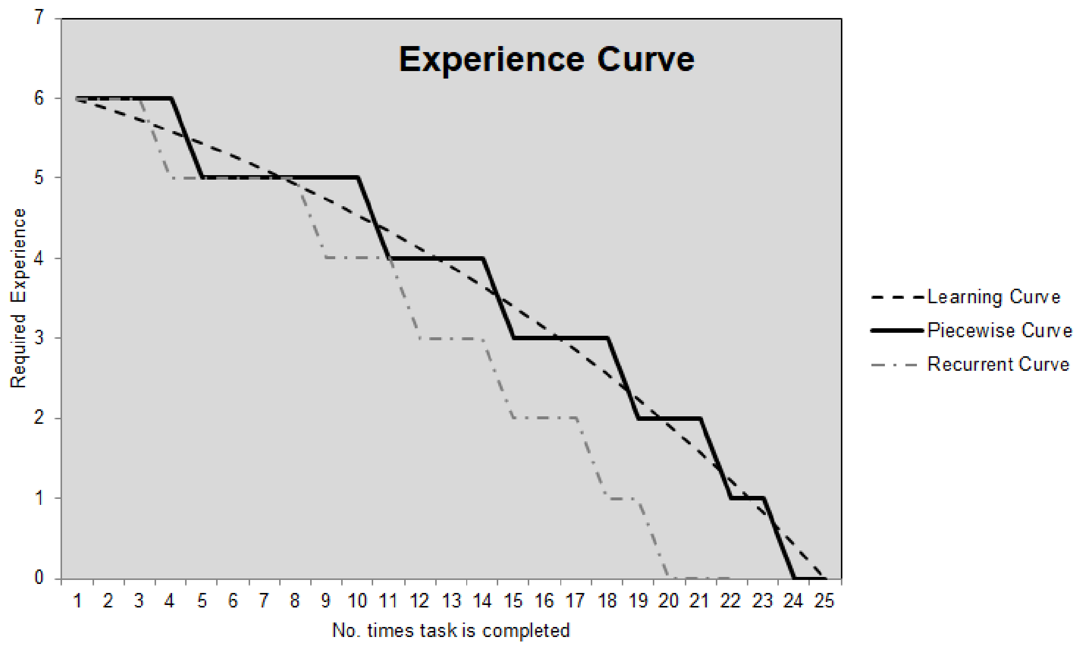

Fig. 5.3: Formulated learning curve

The learning curve will need adjusting if Extension Four is added to the model for recurrent skills. The recurrent curve is formulated from the same method, however, here it is assumed four less uses of the skill will produce a competent employee. Hence in this case the formula for the recurrent curve is given as $f(x) = -1(0.95^{(-a-x)} + b)$ where $a = 23.58$ and $b = 9.35$ to two decimal places. Using the data defined in the table, the probability of a task requiring the experience described by the curve is 90%, otherwise no experience is required independent of the number of times a task is completed by the employee.

Thus the learning curves shown in Figure 5.3 are defined for this study, where a task has a required amount of experience to show competence.

5.3 Training Budget Analysis

The training budget is a key constraint in this project. It allows a user to specify a budget to be spent on total training for all employees in the model. As discussed in Section 4.2, the training cost is not included in the objective. Thus, in order to reduce the amount of money spent on training, the user would need to decrease the available budget.

In this section, the training budget is varied to analyse the impact on the solution and on the run time of the model. The problem size is selected as [10,20,7] as defined in Section 4, such that there are 10 employees, 20 tasks and 7 time units, and all model components are included except for Extension Six; stochastic constraints. In this way, there will be no infeasible solutions to distort the output of the analysis. For each value of the training budget, 100 runs of random data are performed, following the distributions specified Section 5.2.

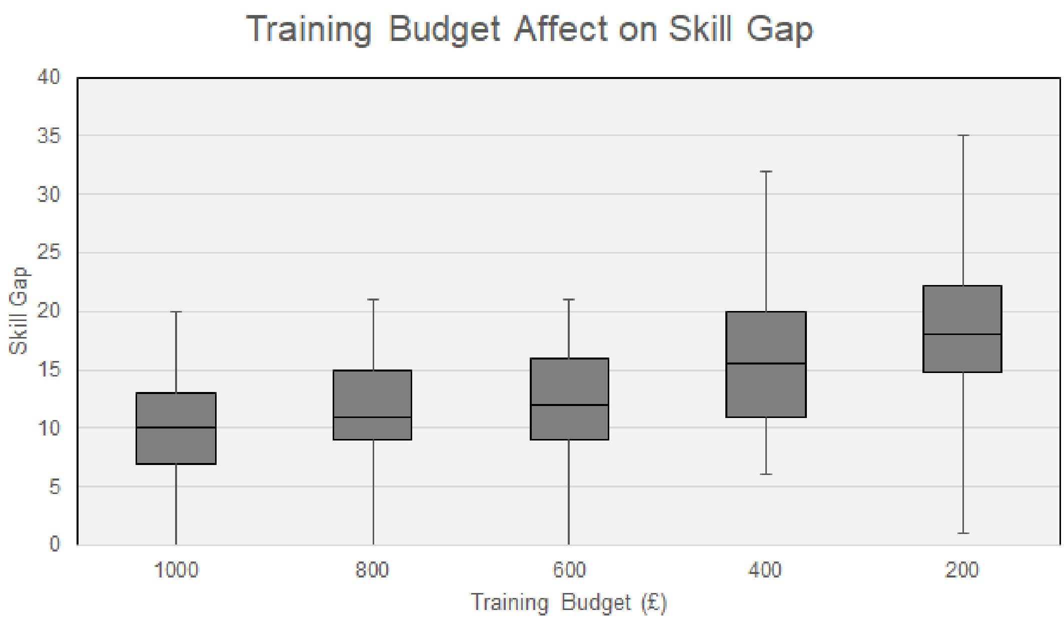

Fig. 5.4: Impact of change in the training budget on the skill gap

First, the skill gap is assessed. The results can be seen in Figure 5.4.

The average value of the skill gap increases at a larger rate as the training budget decreases. This is a reasonable result; the more training employees are able to complete, the more skills they have and hence the smaller the skill gap.

Due to the training cost appearing as a constraint as opposed to an objective function, it is likely that the training cost produced will be higher than the training cost achievable. As suggested previously, the training budget may be reduced until a suitable answer is obtained. The amount spent on training in every run of the model is always near to the value of the training budget. Thus, the fairness value, which is minimised, can be set accordingly. The various solutions available to obtain a minimimum fairness value indicates a larger solution space when the training budget is larger. In addition, as it does not appear in the objective, unnecessary training can be performed with no negative implications on the objective value. Hence, the high training costs.

For these particular distributions of data, a training budget of £600 or greater results in a similar skill gap. However, for a budget less than this, the skill gap does increase. It can be concluded that a training budget of less than £600 is more likely to be a binding constraint. Indeed, for a budget of less than £600, it is very unlikely that a skill gap of zero can be achieved. Hence, the budget is not large enough to train staff such that all tasks are complete, given the specified data ranges. For different values of data, this bound will differ.

The large interquartile range and values of the maximum skill gap indicate that the selection of data has a greater influence on the resultant skill gap than the value of the training budget. However, as the budget decreases, the interquartile range is reduced, suggesting a more constrained problem space.

Next, the run times are considered. For the previous 100 runs, the times to complete the optimisation are also produced and presented in Figure 5.5.

For a problem specification of [10,20,7], the run time for the model may exceed the wall time of 600 seconds for any training budget greater than or equal to £600. If the budget is less than £600, a smaller run time can be achieved. A small budget binds the problem such that there are less active constraints and hence less problems proving optimality. Thus, a smaller solution space results in smaller run times.

The interquartile range of run times is at a minimum when the training budget is minimised. For a training budget larger than £600, there is no correlation between run time and budget, suggesting the training budget constraint is not binding at these values.

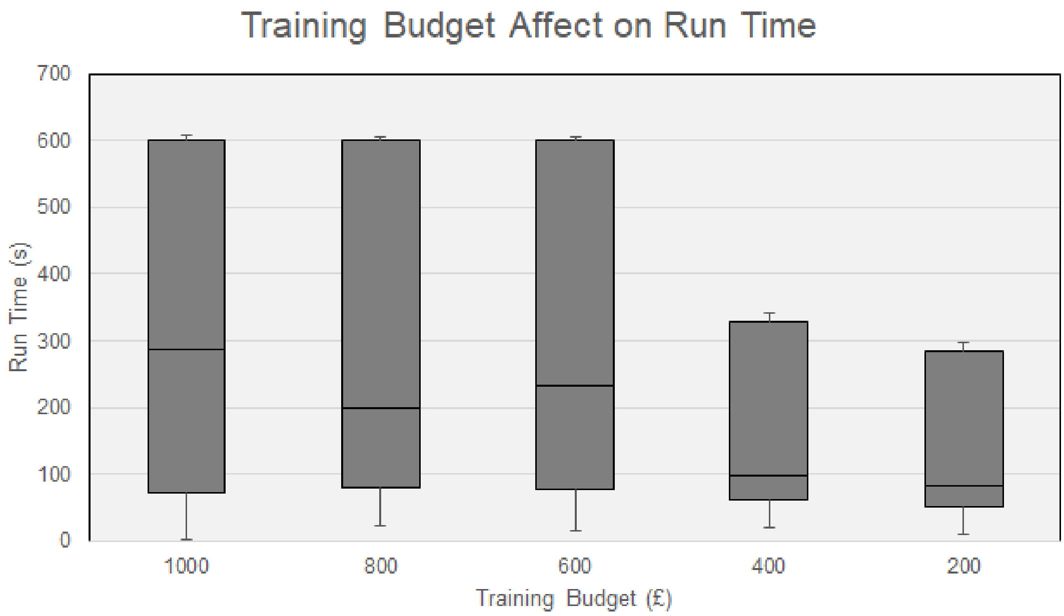

Fig. 5.5: Impact of change in the training budget on the run time

For the remainder of this study, unless otherwise stated, the training cost will be fixed to £600 if the data specified in Section 5.2 is used. As such, the budget will at times influence the solution to the model, depending on the data. However, it will not regularly reduce the solution space and distort the purpose of the analysis. When examining the behaviour of a coefficient within the model, this behaviour should not be consistently impacted by the value of the training budget.

5.4 *Coefficient Analysis*

An important step in the verification and validation of the model is a sensitivity analysis on the parameters. This analysis adjusts the values of key coefficients in each model and observes the change in the model behaviour. The data selected for the analysis is chosen such that this behaviour can be observed.

The skill gap and run time will be measured such that a correlation between the coefficient and these two outputs can be calculated. For each model, more than one coefficient is adjusted within each study, where applicable, to determine the relationship between the two key coefficients used in that model. The coefficients to be varied are as follows:

- TEM 0 - Requirements and employee initial skills
- TEM 1 - OJT cost
- TEM 2 - Training course cost and experience gained from training course

- Ext 1 - Compulsory course cost and number of tasks requiring compulsory training
- Ext 2 - Number of tasks with a reduced experience requirement and employee initial skills
- Ext 3 - Number of dependent auths and employee initial skills
- Ext 4 - Speciality of task and speciality of employee
- Ext 5 - The two parameters defining the shape of the learning curve
- Ext 6 - The values of α and β

The remaining variables in the model are fixed such that their values should not impact the results of the sensitivity analysis.

It should be noted that in each model the time to complete any component of the training is not considered as a variable input. Changing the time to complete the task is assumed to have a trivial impact on the solution of the model. For instance, increasing the time to complete a training course will decrease the amount of training an employee is able to do and thus increase the skill gap. Similarly, decreasing the availability of the employee also increases the skill gap. The relationship between the time to complete tasks, training and the availability of the employees is considered linear and trivial in terms of analysis.

However, the cost of training and the constraints on an employees ability to complete that training will have a more interesting relationship and thus should be examined.

The size of the model will remain consistent in each of the sensitivity analysis, in terms of number of employees, tasks and time units. This is set to be [5,12,7] where there are 5 employees, 12 tasks and 7 time units. In this way, the run time should remain reasonably small and a small change in the coefficients should make a significant difference to the output of the model.

5.4.1 Training Execution Method Zero

The first model to be explored is Training Execution Model Zero. Repeating tasks is the only method of gaining experience and no additional constraints are used in this model. TEM Zero acts as a base for all following investigations.

In order to determine the result of the data on the output of the model and the run time, a coefficient analysis is required. In this way it can be determined how sensitive the decision variables are to changes in the coefficients. A Monte Carlo simulation is performed for 100 runs using a base set of data but with random alterations to the coefficients of interest. The changes in these coefficients are plotted with the output decision variables to determine how the coefficients affect the decisions made in the model.

As specified in the introduction to this chapter, Section 5.4, the size of the model to be used is [5,12,7], such that there are 5 employees, 12 tasks and 7 time units. A model size of [5,12,7] with only TEM Zero active contains 4,330 decision variables and 635 constraints.

For TEM Zero, the coefficients of interest are the requirements for the tasks and the initial skills of the employees. The rest of the data is fixed as follows:

- Processing time $\eta_k = \lceil \frac{\hat{k}-k+1}{2} \rceil + 2 \quad \forall k \in K$
- Availability $a_{et} = 19$ if $e = 2$, else $a_{et} = 38 \quad \forall e \in E, t \in T$
- Renewal time $Z_{ek1} = [1, ..., \hat{t}] \quad \forall e \in E, k \in K$
- Experience threshold $g_k = 6 \quad \forall k \in K$

In this way, there is one part time employee available to provide variation in staffing types. Processing times vary per task such that a range of short and long tasks are considered. All tasks are reviewed at the end of the time horizon, this results in one review period. Reviewing tasks in this way allows more insight to be gained from the coefficients of interest. There is no training budget specified for this model as repeating tasks does not have an associated cost and no other form of training is available.

The requirements of the task, r_k, define the number of times the task needs completing within one time unit. The random values investigated here are produced by sampling random numbers from the Uniform(1,10) distribution. Thus, the minimum number of tasks that can be required in each time unit is 12, such that each task is required once. The maximum value is 120 such that each task is required 10 times in each time unit.

The initial skills of the employees, Θ_{ek}, define which tasks can be completed by each employee before the beginning of the time horizon. The random values are selected by sampling from the Bernoulli(0.5) such that the probability of each employee having each skill is equal to 0.5. Hence, it is possible to obtain a sample such that no employees have skills or all employees have all skills. For the following graphs consider the average number of tasks that can be completed by all employees before the start of the time horizon. The average is calculated as $\boldsymbol{\Theta}_{ek} = \frac{1}{\hat{e}} \sum_{e \in E} \sum_{k \in K} \Theta_{ek}$.

The initial skills of the employee form the right hand side of an equality constraint whereas the requirements are considered in two sets of constraints and the objective function. Thus, it is expected that the requirements will have a higher correlation with the skill gap than the initial skills.

First the skill gap is produced from the 100 random data samples. The skill gap output here is the gap calculated at the end of the time horizon at $\hat{t}$. It is calculated as $\sum_{k \in K}(r_k - \sum_{e \in E} w_{ek\hat{t}})$. In this way assignments made before all training has been completed are not considered. Note that this skill gap differs from the skill gap equation in the objective function such that here, skill gap is only calculated in the final time unit whereas in the objective it is summed over all time units. This definition of the skill gap will remain consistent through all of the coefficient analysis section.

Though the objective function is also composed of the fairness calculation, it is not the priority of this model. As such, the coefficient in the model A is set as a sufficiently large value.

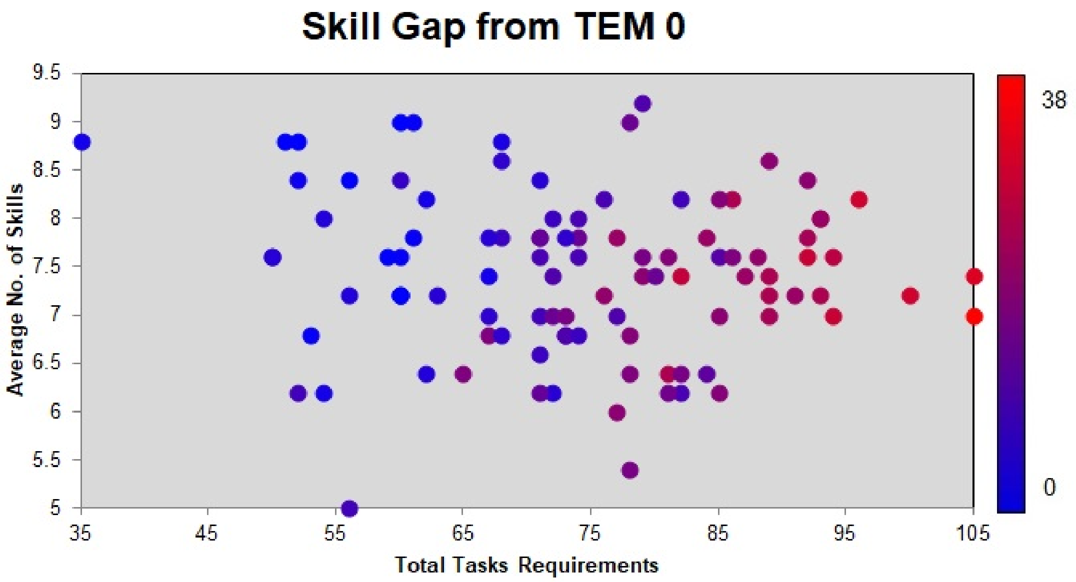

Fig. 5.6: Skill gap in TEM Zero given changes to coefficients

From Figure 5.6, a relationship between total requirements for tasks and the skill gap can be seen. This relationship seems independent of the initial skills of the employees. The more times the tasks need completing, the higher the skill gap when the number of employees and the processing times remain constant. It is expected that the more tasks an employee is able to complete, due to their initial skills, the lower the skill gap. However, the graph shows that this relationship has a smaller influence on the skill gap than the requirements value. Though an employee may be able to complete more tasks, they are still constrained by their available time. The total requirements is also a coefficient in the objective function, thus directly affecting the value of the skill gap.

These relationships can be checked by considering the coefficients separately.

Figure 5.7 shows that requirements and skill gap have a linear relationship. However, the skill input and skill gap show no correlation. It should be noted that these two plots are not independent. A change in one variable will affect the skill gap in both plots, thus noise is expected.

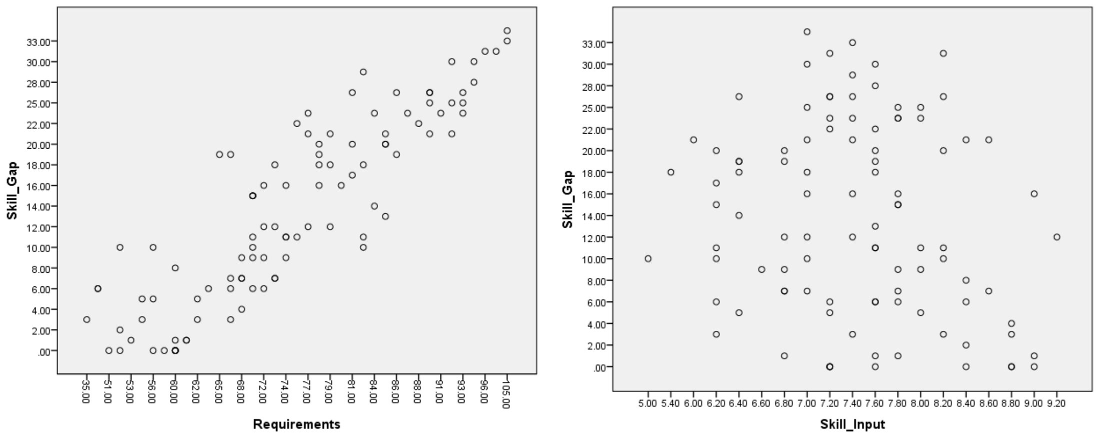

Fig. 5.7: Relationship between independent and dependent variables for TEM Zero

The skill gap produced from the model is strongly related to the requirements of the task. As the requirements are directly used in the calculation of the skill gap, the significance of the value is higher than that of the initial skills of the employee. The initial skills of the employee have little impact on the skill gap in comparison. An employee is unlikely to use their full time and skill set on the tasks that need completing.

This suggests the model is very sensitive to changes in the requirements and less sensitive to the initial skills of the employees.

If the requirements for the task are sufficiently small, the associated constraints become binding. However, for large values of requirements, the availability and skill constraints become binding as they affect the allocations and thus the size of the skill gap increases linearly with the requirements.

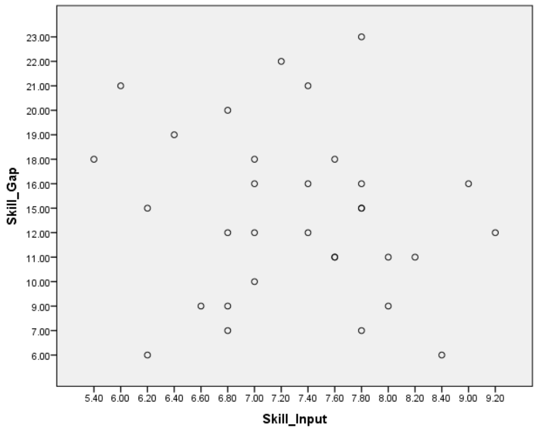

Fig. 5.8: Graph of skill input by skill gap when requirements vary between 70 and 80

Let us consider only the results given a requirement value between 70 and 80. Again, Figure 5.8 shows no significant correlation. Thus suppose that the average number of skills held by employees at the beginning of the time horizon has less significance on the resulting skill gap. It is proposed that the specific skills held by the employees and the relation with the tasks with the highest requirements may have greater influence on the value of the skill gap.

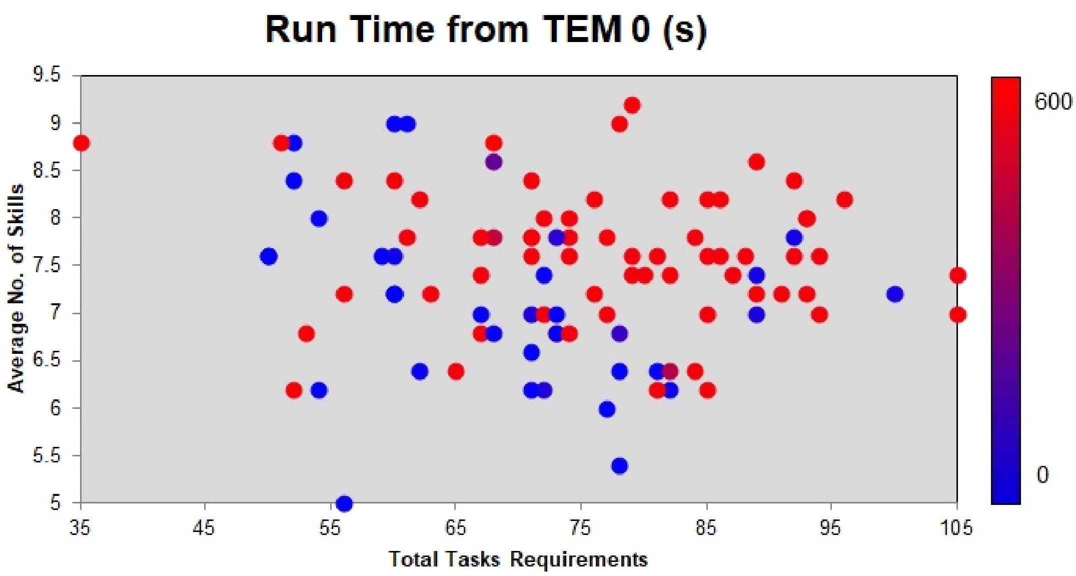

Fig. 5.9: Run time in TEM 0 given changes to coefficients

Next consider the run times for each data sample given the changes to the coefficients.

There are a large number of data points in Figure 5.9 with a run time exceeding 600 seconds. Indeed, 62% of the data points exceed the wall time of the run.

Though no clear correlation between these run times and the coefficients of interest can be seen, there is a difference in solvability dependent on the selection of data.

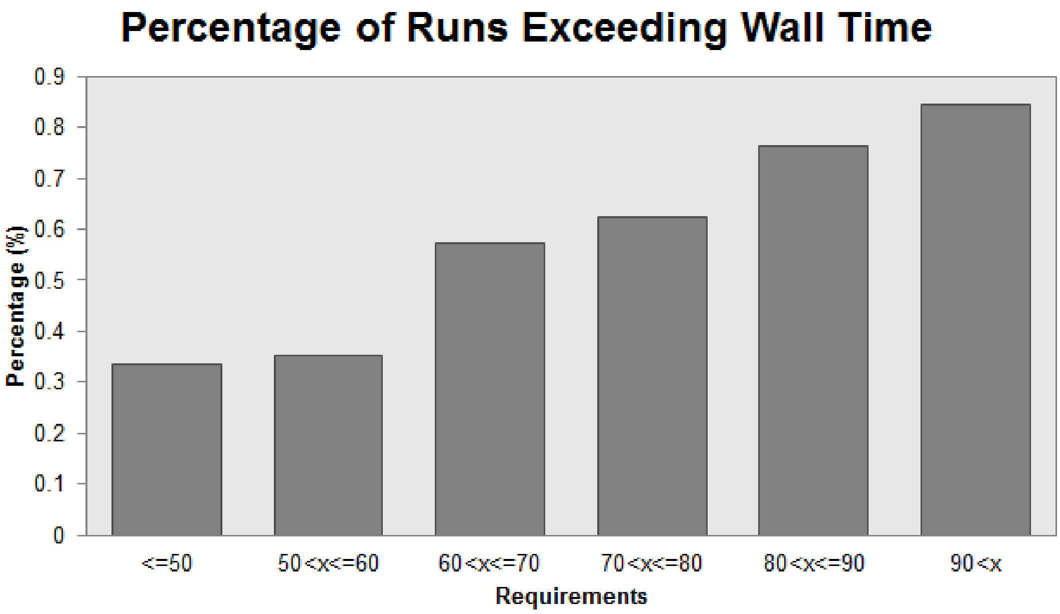

Fig. 5.10: Percentage of runs exceeding wall time

Figure 5.10 contains a histogram showing the percentage of runs that exceed the wall time set. A correlation exists between the requirements of the task and the number of tasks exceeding the wall time. Though it suggests that lower values of requirements may produce models that are solvable, correlation is not proof of causation.

5.4.2 Training Execution Model One

Training Execution Method One defines experience that can be gained through on-the-job training. As with TEM Zero, coefficients of interest and relevance to the selected model are chosen and adjusted to analyse the results of the coefficient on the decision variables. For this TEM, on-the-job training cost, c_1, is selected as the only relevant coefficient. The value is selected randomly from the Uniform(0,500) distribution. The other coefficients required for this model are fixed as follows:

- Processing time $\eta_k = \lceil \frac{\hat{k}-k+1}{2} \rceil + 2 \quad \forall k \in K$
- Availability $a_{et} = 19$ if $e = 2$, else $a_{et} = 38 \quad \forall t \in T$
- Renewal time $Z_{ek1} = [1, ..., \hat{t}] \quad \forall e \in E, k \in K, t \in T$
- Experience threshold $g_k = 6$

These values are similar to the values given in the basic TEM, with the processing time adjusted to correspond to the new value for requirements. However, the coefficients analysed in the previous section are now fixed as follows:

- Requirements $r_k = \lceil \frac{k}{3} \rceil, \forall k \in K$
- Initial skills $\Theta_{ek} = 1 \quad$ for $e < \frac{\hat{e}}{2}, k \in K$, 0 otherwise
- Training budget $= \pounds 1,000$

To apply TEM One, a training budget is required due to the cost of on-the-job training.

The base data is chosen such that half of the employees have all skills. Given that five employees are used in the model, this means two employees have all the skills. In addition, one of these employees is a part time worker, as with the previous model. All other employees have zero skills at the beginning of the time horizon. Thus, these employees will require training in order to complete any tasks.

The requirements are chosen such that a skill gap of zero cannot be obtained within the constraints. In this way, slack constraints do not intercept the analysis of the coefficients due to trivial solutions. The maximum value for any requirement is 4 for a model with 12 tasks. Using conclusions drawn from the analysis of TEM Zero, this should reduce the run time for the models such that any long run times have a lower probability of being correlated to the value of require.

With these two workers and their current skill sets only, it is possible to obtain a skill gap of 17 tasks. Thus only two employees are required to complete all but 17 of the tasks. On-the-job training can reduce this skill gap by training the remaining three employees who have no skills at the start of the time horizon. From the model description, the only method these additional employees have of gaining experience is through on-the-job training or repeating tasks. As they can only repeat tasks if they have the skill this implies that they can only gain the skill through on-the-job training.

The size of the model to be used is [5,12,7]. Here, this creates a model containing an upper bound of 4,750 decision variables and 1,055 constraints.

As before, 100 runs of randomly selected data from the above conditions are performed. The results are found in Figure 5.11.

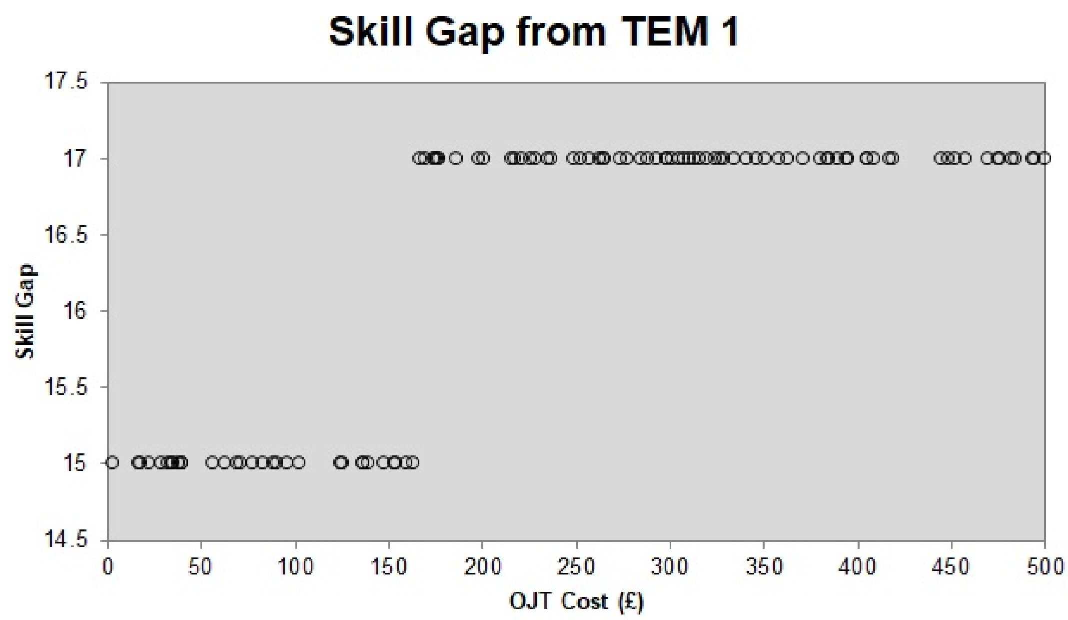

Fig. 5.11: Skill gap in TEM One given changes to coefficients

Only two possible values of the skill gap can be obtained. The skill gap obtained when no training takes place is 17 whilst the minimum skill gap available through this training is 15. The lower skill gap is only achievable for a training cost less than approximately £180.

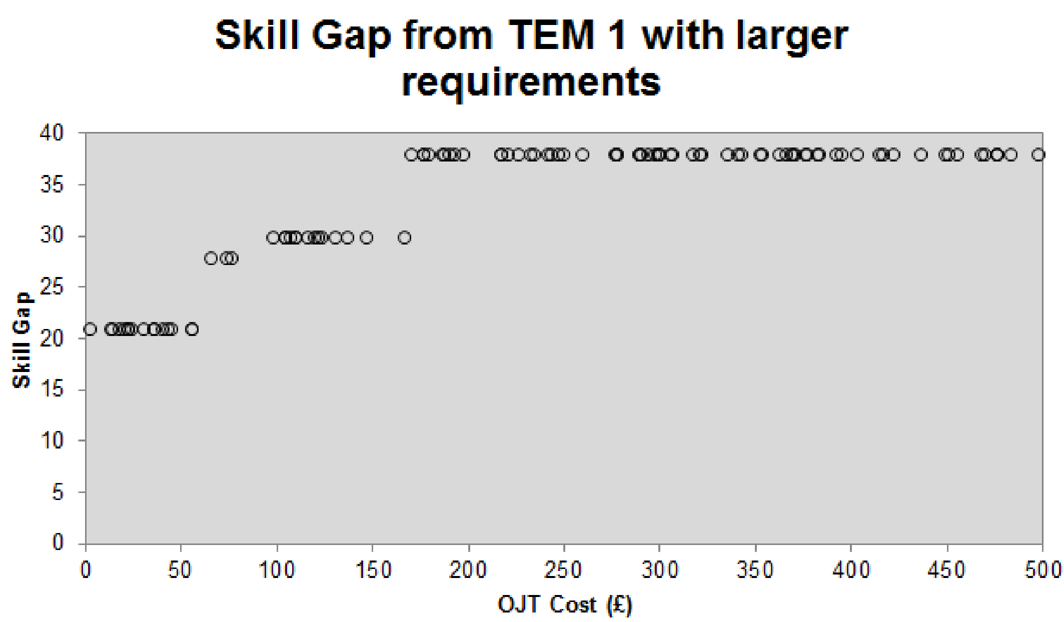

Fig. 5.12: Skill gap in TEM One given changes to larger coefficients

To obtain further insight into these results, another run of 100 random data samples is performed where the requirements are increased. In this way, processing time is set as $\eta_k = \lceil \frac{\hat{k}-k+1}{2} \rceil$ and requirements are set to $r_k = k, \forall k \in K$. Other coefficients remain the same as the previous runs. Figure 5.12 shows the behaviour of the training cost and OJT constraints with changes to the OJT cost.

From the results it can be seen that a minimum skill gap of 21 is achievable for on-the-job training costs ranging from £0 to £500. For any OJT cost greater than a certain value, the minimum skill gap obtainable is 38 which is equivalent to no on-the-job training being given. Consider one employee with no initial skill. In order to exceed the given threshold, the employee is required to complete training in a task six times. The cost of this is six times the OJT cost. To satisfy this training cost without exceeding the training budget, the total cost required to train an employee is calculated as $c_1 \times 6 < 1000$. Thus the maximum value of c_1 is equal to 166.66. If the cost exceeds this value then any training is considered too expensive and thus no employee can learn any skills. The change at this value can be seen in the graph. If the training cost is above this value, the constraint on training costs is binding such that decision variables assigning employees to OJT are forced to be zero.

The stepwise appearance also signifies the other binding moments within the model. The change in skill gap indicates a binding training budget constraint. When the skill gap does not change for a range of values of OJT cost, the training budget constraint has slack and some other constraint is binding.

The minimum skill gap achievable when training has no cost is given as 21. If no other constraint existed, all employees could be trained in all skills for no cost and the skill gap would be reduced to zero. However, as this is not seen in the results this implies that other constraints restrict the employees assignment to training. This suggests that the constraint on available time or available trainers would be more prominent than the constraint on cost.

The impact of the run time on the choice of coefficient has also been considered in Figure 5.13, using the data used to obtain Figure 5.11. Where the training costs are cheaper, the number of possible solutions increases thus the longer run times. With more money, there are fewer training combinations available for the price. Where OJT assignment is trivially 0 due to the training cost constraint, the run time is much lower and much more consistent.

For the 100 runs, there were no instances where the run time exceeds the wall time limit. Thus, there is no issues with proving optimality.

The run times appear to decrease exponentially in regards to the on-the-job training cost. This is tested using goodness of fit in IBM SPSS Statistics 24.0.

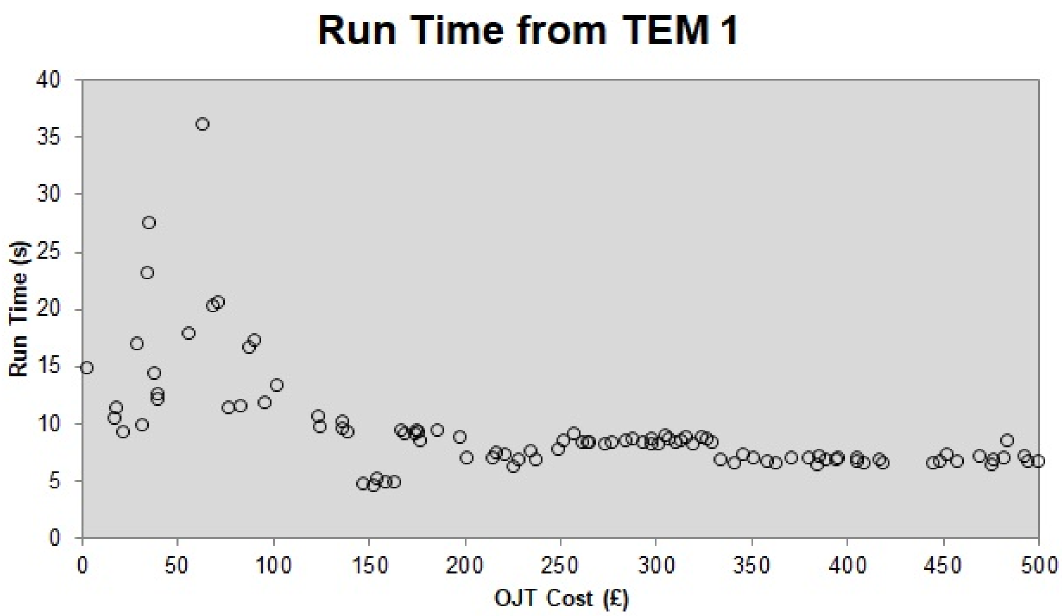

Fig. 5.13: Run time in TEM One given changes to coefficients

To satisfy the assumptions for a goodness of fit test, a correlation must exist. Using a Pearson correlation test, a value of -0.586 is given which is significant at the 0.01 level using a two-tail test. Hence, a goodness of fit test can be performed.

The independent variable is the OJT cost whilst the dependent variable is the run time. To compare possible fits, the logarithmic, quadratic, power and exponential curves are tested on the data. In terms of a control; a linear model is also applied. The R values for each test are compared. First, the linear model produces a R squared value of 0.337. For a logarithmic model, the R squared value is calculated as 0.365, which indicates a weak fit for the log model. For the quadratic model this value is given as 0.43, a slightly better result. The power curve produces a R squared value of 0.44 and finally the exponential model gives a value of 0.405.

These results indicate that the power curve is the best fit to the model. Performing an ANOVA test indicates that the model is significant in predicting the run time from the OJT cost. Thus it can be concluded that the OJT cost can be used to predict the run time of the optimisation model for TEM Zero and it can be described using a power curve.

Figure 5.14 shows the output from the Goodness of Fit test for the referred to models.

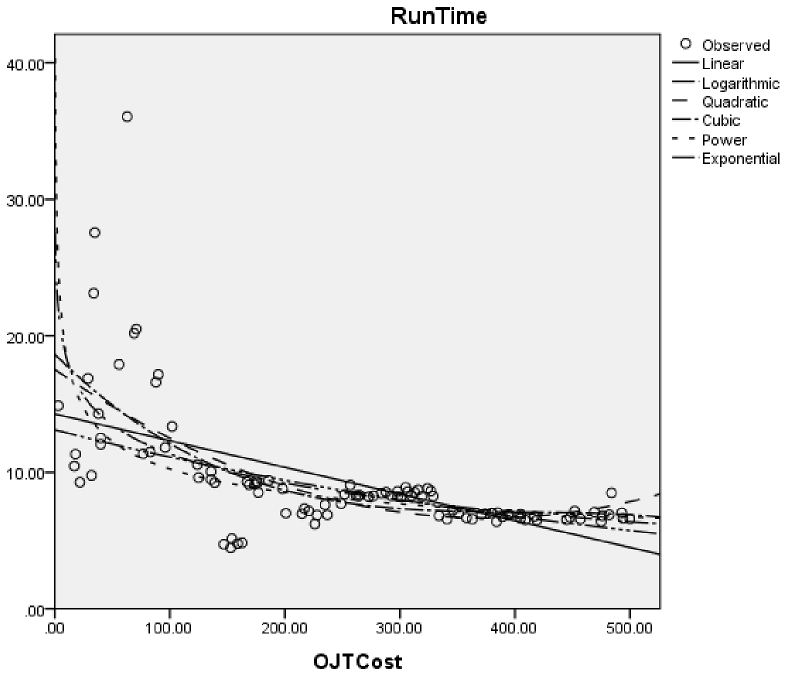

Fig. 5.14: Regression models for TEM One

5.4.3 Training Execution Model Two

The final Training Execution Method is training courses. The coefficients of interest for this model are training course costs and experience gained for each authorisation. The other variables concerning the training courses are fixed. For this analysis the number of training courses is fixed to two possible options, such that $\hat{n} = 2$.

The base data is fixed as follows:

- Processing time $\eta_k = \lceil \frac{\hat{k}-k+1}{2} \rceil + 2 \quad \forall k \in K$
- Availability $a_{et} = 19$ if $e = 2$, else $a_{et} = 38 \quad \forall t$
- Renewal time $Z_{ek1} = [1, ..., \hat{t}] \quad \forall e \in E, k \in K, t \in T$
- Experience threshold $g_k = 6 \quad \forall k \in K$
- Requirements $r_k = \lceil \frac{k}{3} \rceil, \quad \forall k \in K$
- Initial skills $\Theta_{ek} = 1 \quad$ for $e < \frac{\hat{e}}{2}, k \in K$
- Training budget = £1,000

The coefficients specified here are assigned the same values as TEM One. Two employees are given all skills, of which one employee is a part time worker, whilst all other employees have no skills. The coefficients that will be varied in this model are randomly sampled from the following distributions.

- Training course cost c_{2n} is sampled from the Uniform(0,500) distribution $\forall n \in N$
- Training course experience χ_{nk} is sampled from Uniform($1, \frac{i}{10}$) $\forall n \in N, k \in K$, where i is the run number

By producing experience dependent on the run number, this ensures a reasonable spread of results rather than the average result from the Uniform distribution. This is particularly important when varying two large sets of coefficients.

The size of the model, [5,12,7], with two training courses produces 4,400 decision variables and 649 constraints. Compared to TEM One, the size of this model is smaller and closer to the size of TEM Zero. However, it must be noted that there are a small number of training courses available.

Figure 5.15 shows the skill gap produced given different values of the coefficients. The x axis expresses the sum of the training course costs over all n, the y axis is the total experience that can be gained over all courses n and all tasks k.

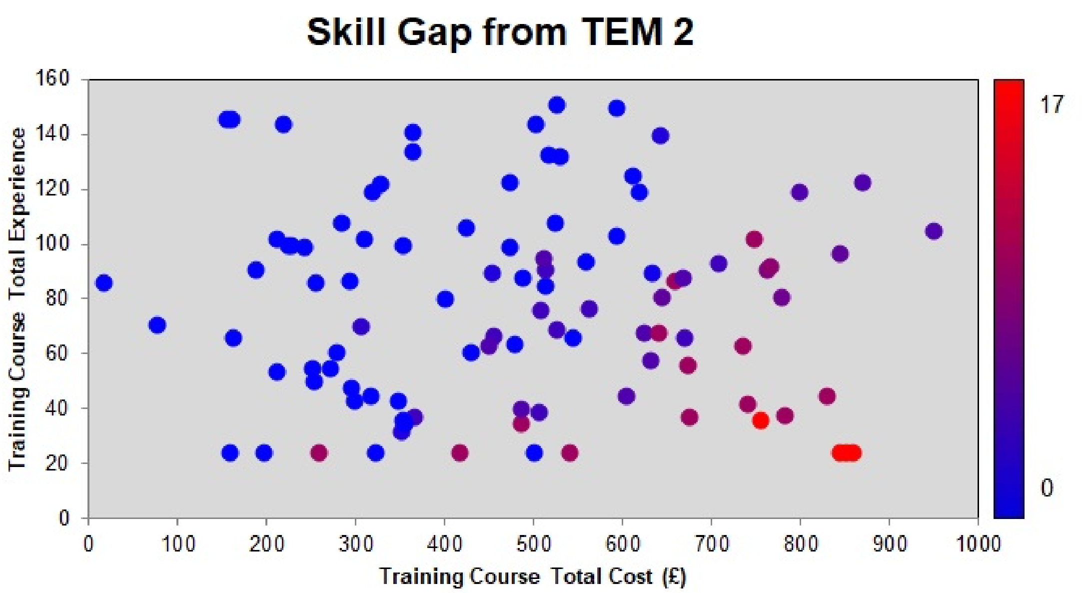

Fig. 5.15: Skill gap in TEM Two given changes to coefficients

As with TEM One the maximum skill gap possible is 17, this is equivalent to no training course being used. Thus, any improvement on this value is caused by the additional constraints provided in TEM Two. Here, it can be seen that a skill gap of 1 can be obtained. This is better than the minimum skill gap obtainable by OJT which was equal to 15. As the costs vary within the same distribution, it may be concluded that the improvement to the skill gap is caused by the potential gain in experience given by the training courses. Indeed, from the graph, a correlation can be seen between the total skill gained from the course and the outputted skill gap.

However, there is also a correlation between the cost of the course and the skill gap. Thus, the training courses that increase the experience of the employees by a small amount at a high cost produce a higher skill gap.

These correlations can be tested as before. First, the variable pairs are plotted seperately in order to justify the assumption of linearity, Figure 5.16. Note that though the pairs are plotted separately, the data is not independent. Variation in the skill gap may be caused by either of the two variables.

The linear relationship between training course cost and skill gap is clearly shown when the skill gap is not at it's minimum or maximum value. Here, noise is expected as the value of training course experience varies. The larger the total cost of the training course, the larger skill gap obtained.

However, the relationship between total experience gained through a training course and skill gap is less defined. Though it can be seen that a low experience gain results in a larger skill gap, there is a large amount of noise dependent on the cost of the training.

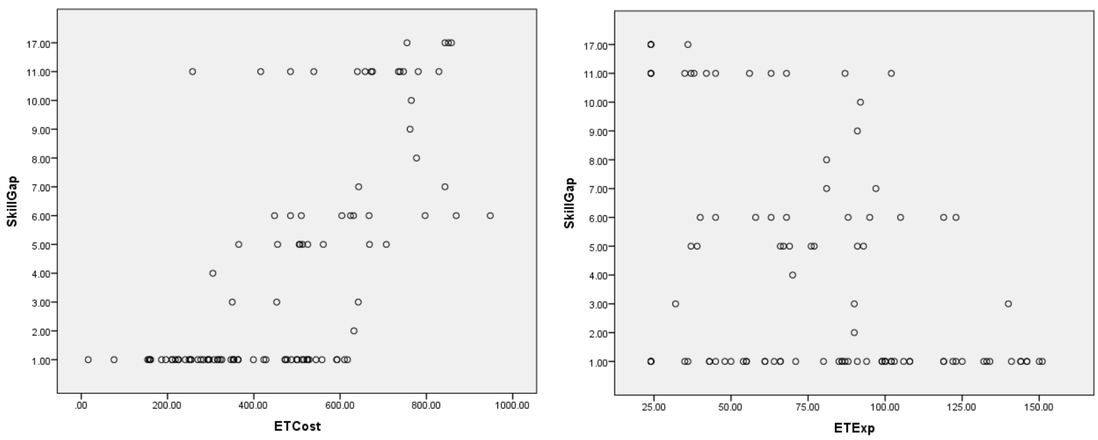

Fig. 5.16: Relationship between independent and dependent variables for TEM Two

The relationship between the two independent variables can be investigated using linear regression. First, it must be shown that individually the variables are correlated. Note that the outliers will have an impact on the results of the correlation and regression.

Performing a Pearson correlation test on training costs and skill gap creates a

Pearson correlation value of 0.678 signifying that the two variables are positively correlated. For experience gained and skill gap this value is given as -0.434, showing a negative correlation at the 0.01 significance level. However, this shows only $0.43^2 = 18\%$ of the variation in skill gap can be explained by the experience gain of the training course.

To consider which of these two variables has a more significant impact on the value of the skill gap, linear regression is used to plot a line of best fit. If these plots are both reasonable predictions of the trends, the gradient can determine which variable provides the greatest change in skill gap.

Performing the regression with both independent variables being used to predict the dependent skill gap, a best fit line of $y = 0.014x_1 - 0.51x_2 + 1.494$ is produced where y is the skill gap, x_1 is the cost and x_2 is the experience. This line has an R^2 value of 0.632 indicating that 63% of the variation in skill gap can be explained by the independent variables. From performing an ANOVA test it is determined that the regression model is statistically significant in predicting the skill gap. The values of the coefficients indicate that a small change in the experience gained from a course results in a larger change in the skill gap than for the training course cost. Contrary to the result seen in Figure 5.16.

There did not appear to be any correlation between run time and selection of variables. Most of the problem instances produced a run time that exceeded the wall time. This will be considered in a later section.

5.4.4 Extension One

As with the Training Execution Methods, the coefficients of the model extensions may also be adjusted as a method to analyse model behaviour. The first extension is compulsory training. Here, the variables of interest are the cost of the compulsory training and the number of tasks that require compulsory training.

First, the fixed data is selected as follows:

- Processing time $\eta_k = \lceil \frac{\hat{k}-k+1}{2} \rceil + 2 \quad \forall k \in K$
- Availability $a_{et} = 19$ if $e = 2$, else $a_{et} = 38 \quad \forall t$
- Renewal time $Z_{ek1} = [1, ..., \hat{t}] \quad \forall e \in E, k \in K, t \in T$
- Experience threshold $g_k = 6 \quad \forall k \in K$
- Requirements $r_k = \lceil \frac{k}{3} \rceil, \quad \forall k \in K$
- Initial skills $\Theta_{ek} = 1 \quad \forall e \in E, k \in K$

- Training budget = £1,000

Unlike the Training Execution Methods, this model requires employees to have all skills. In this extension employees cannot gain skills, other than through repeating tasks, and the skills of the employees are independent, so do not affect each others ability to gain skills. Hence, if an employee began the time horizon with no skills, there would be no method for them to gain any skills. In this way, all employees are set to have all skills initially.

All other coefficients remain the same as those given in TEM Zero. It should be noted that this could result in a skill gap lower than the lower bound on the skill gap achieved in the TEMs.

As stated above, the coefficients of interest are the compulsory training costs and requirements. These are randomly generated from the following distributions.

- Compulsory training cost c_{3k} is sampled from the Uniform(0,500) distribution $\forall k \in K$

- Compulsory requirements π_k is sampled from the Bernoulli($0.01i$) $\forall k \in K$, where i is the run number

As before, in TEM Two, using the run iteration allows more variation in the data used. The cost uses the same distribution as the cost of training in TEM One and TEM Two.

Using the problem size $[5, 12, 7]$, the model required 4,750 decision variables and 755 constraints. Though the number of constraints does not differ too much from the basic TEM Zero model, the number of decision variables has increased significantly as seen with TEM One. The number of decision variables is the same for these two models due to the index k being used in both sets of decision variables.

Figure 5.17 shows the change in the skill gap given changes to the two coefficients of interest for 100 runs of data. The total compulsory requirements and total cost over all tasks are recorded.

There appears to be a positive correlation between each variable and the skill gap. Linearity is first checked before a correlation test can be completed. Note that the two graphs are not independent, changes to one of the variables will affect the skill gap in both plots.

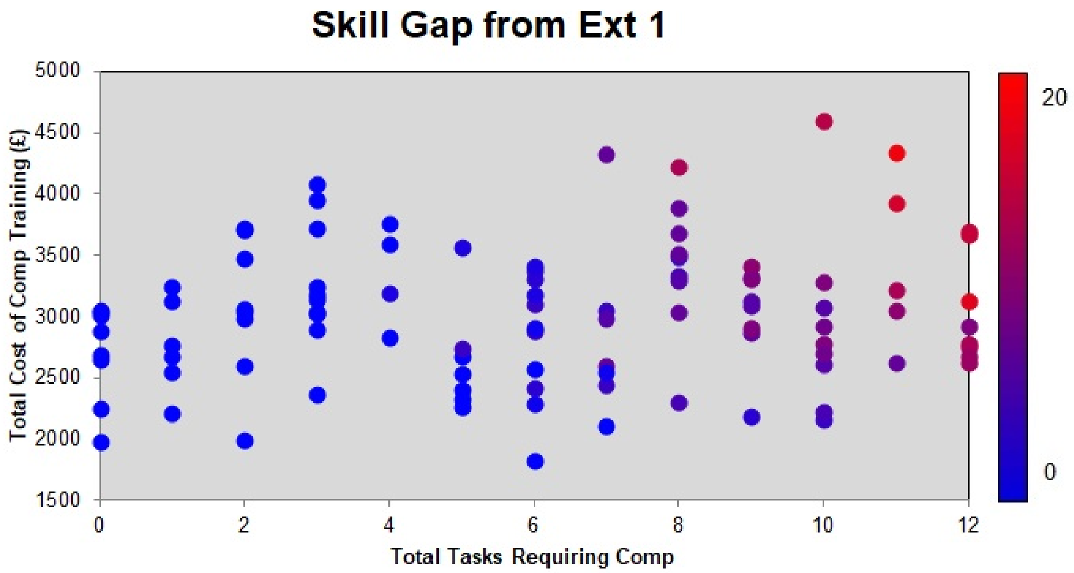

Fig. 5.17: Skill gap in Ext 1 given changes to coefficients

As seen in Figure 5.18, there does appear to be a relationship between the variables. However, the compulsory training requirements seem to have a much more defined relationship. Indeed, the relationship appears almost exponential. The correlations for the two pairs are calculated.

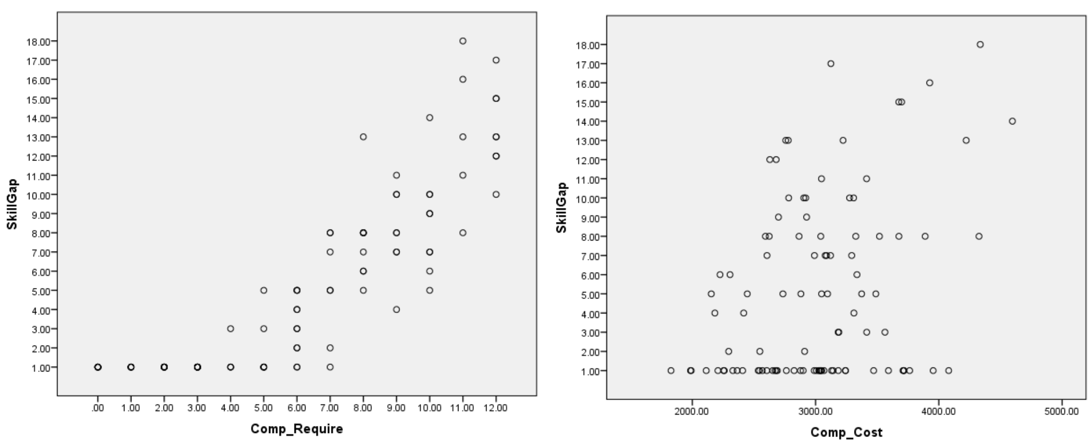

Fig. 5.18: Relationship between independent and dependent variables for Ext One

For compulsory requirements, the correlation coefficient is given as 0.864, which is a significant result. For the compulsory costs, this value is calculated as 0.354 which is much less significant. Thus, linear regression can be used to analyse the relationship between all three variables. It can be shown that 80% of the variation in the skill gap can be explained by the costs and requirements, where the coefficient of the requirements is 1.05, significantly higher than the coefficient of the costs which is 0.002. The requirements have a much larger impact on the solution in this extension. Hence, a small change in the number of tasks that require compulsory training will have a larger affect on the skill gap than changing the cost of training.

To further analyse the relationship between the skill gap and the requirements, a goodness of fit analysis is performed. This will determine the shape of the curve. Similar to the analysis performed for TEM One, the linear, quadratic, cubic and exponential curves are all investigated. The R squared values are given as:

Equation	R Squared
Linear	0.746
Quadratic	0.826
Cubic	0.829
Exponential	0.817

Tab. 5.3: Ext One goodness of fit tests for comp requirements and skill gap

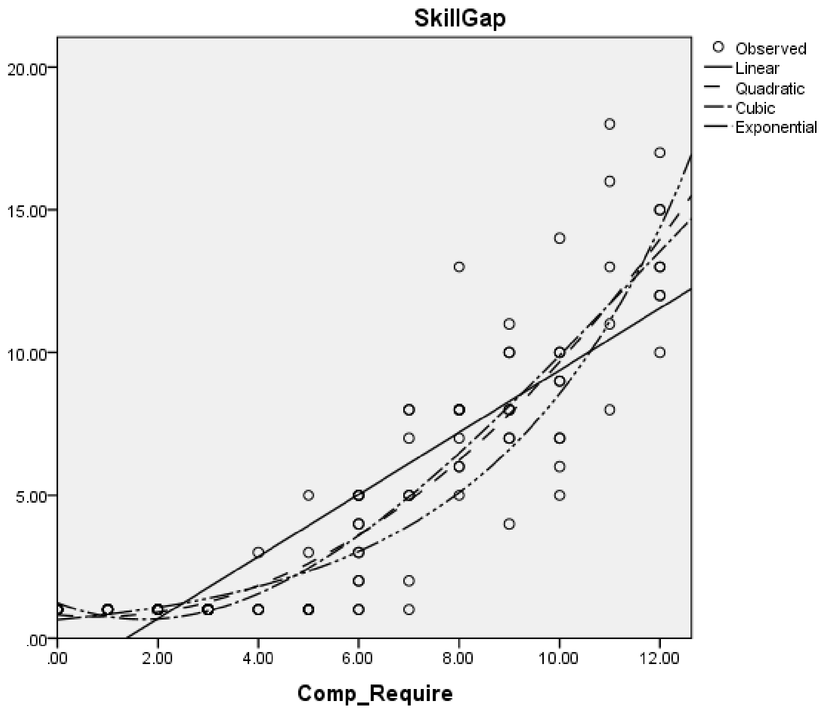

Fig. 5.19: Regression models for Ext One

From this table, the model that best fits the data is the cubic model. Here, 83% of the variation in the skill gap can be explained by the compulsory requirements. Hence, there is significant evidence to suggest the relationship is not linear. The different curves can be seen in Figure 5.19.

5.4.5 *Extension Two*

Tasks may have different requirements depending on whether they are initial or recurrent. In this case, Extension Two may be applied. Here, the initial skills of the employee will be used to determine whether a task is recurrent for that employee and required experience will differ for the two skill types. Thus, these become the variables of interest.

As with previous models, the fixed data is given as:

- Processing time $\eta_k = \lceil \frac{\hat{k}-k+1}{2} \rceil + 2 \quad \forall k \in K$
- Availability $a_{et} = 19$ if $e = 2$, else $a_{et} = 38 \quad \forall t$
- Renewal time $Z_{ek1} = [1, ..., \hat{t}] \quad \forall e \in E, k \in K, t \in T$
- Experience threshold $g_{k1} = 6 \quad \forall k \in K$
- Requirements $r_k = \lceil \frac{k}{3} \rceil + 1, \quad \forall k \in K$
- Training budget $= £1,000$

Here, the initial experience threshold for initial skills remains at 6 for all tasks. However, for recurrent skills g_{k2}, sample such that the probability that the experience required for a task is equal to 6 is $0.01i$, where i is the run number, otherwise no experience is required. The binary data for initial skills, Θ_{ek}, of the employees are sampled from Bernoulli(0.5) such that each employee has a 50% chance of having a skill at the start of the time horizon.

The requirements for the task are increased from the previous model by one to enable a larger difference in solutions. As this model allows a lower threshold to gain a skill, the resultant skill gap is lower. Hence by increasing the potential skill gap, one is able to evaluate the cause of the skill gap more efficiently.

With a problem size of [5,12,7], there are 4,330 decision variables and 755 constraints. Thus, no decision variables are required to be added to the original model, however there are approximately 100 additional constraints. There are double the number of constraints concerning skills determined from the employee experience. This accounts from the addition of recurrent skill types.

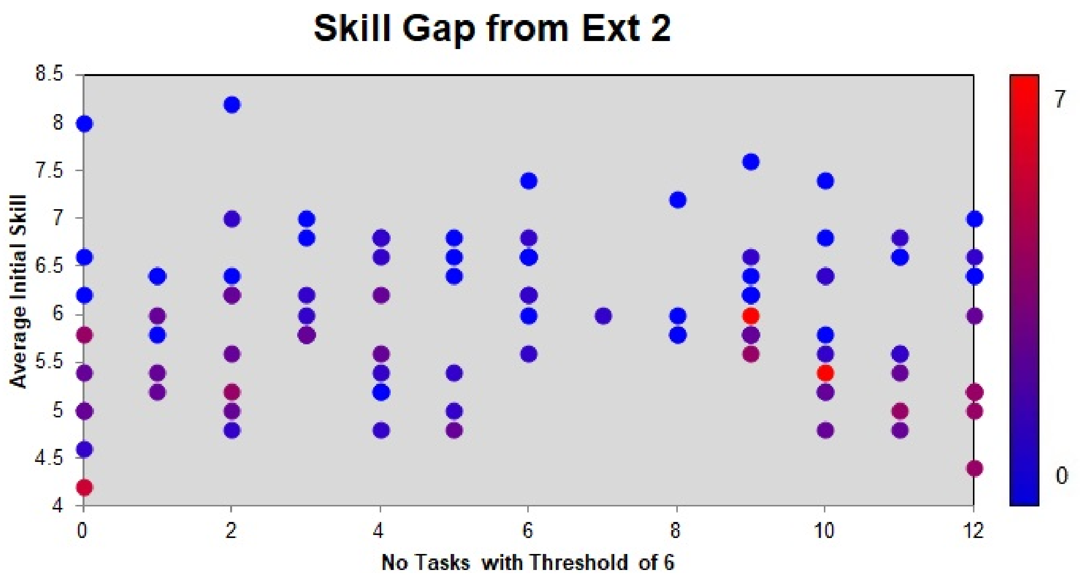

Fig. 5.20: Skill gap in Ext Two given changes to coefficients

Consider the change in skill gap subject to the change in these two variables. Figure 5.20 presents this data.

There appears to be a slight correlation between initial skills of the employee and the skill gap. The number of tasks that require a threshold of six experience does not seem to be significant in determining the solution. This could be due to availability of employees; if the employees have enough time to complete the tasks within each time unit then it does not matter if they complete it in six time units or in less time units. Both solutions would still be feasible. Repeating the analysis on this extension with longer processing times or more tasks would demonstrate this reasoning. However, if the availability of the employees was much larger, an employee may be able to complete different sets of tasks within each time unit, hence gaining experience in multiple skills before they are reviewed.

For an average initial skill of zero, no allocations can be completed. As no other training methods are offered, this means a trivial result is achieved of the maximum skill gap for the data.

The run times for the model can also be plotted as seen in Figure 5.21. Of the 100 runs, 30 instances produced a run time exceeding the limit of 600 seconds. Though no correlation can be proven, the average initial skill does seem to affect the run time. As stated before, where the average skill is set to zero, no allocations can be completed. Thus, the trivial result produces a low run time. Where constraints affect the solution, the run time increases. For an average initial skill set to the maximum number of skills; the run time always exceeds 600 seconds. Here, all skills will be recurrent and subject to the constraints on the recurrent skills.

Fig. 5.21: Run time in Ext Two given changes to coefficients

Due to the definitions, only one set of constraints, initial or recurrent, can be binding at once. In this way, the number of binding constraints is not affected by the initial skills of the employee, except when the employee has no skills. Hence the resulting run times.

5.4.6 Extension Three

The next extension to be analysed considers tasks that have dependencies. In this way, certain auths cannot be awarded unless the employee holds other specified auths. The variables of interest in this model are the average skills of the employee, as the initial auths affect the possible gained auths, and the number of auths that are dependent on other auths.

Here, two auths are suggested as non dependent auths. All other auths rely on one of these two only. By making this assumption, the number of possible connections to analyse are simplified and it ensures no circular networks are created. The number of auths that require task one is stored, where task one is a non dependent auth. Thus the number of auths that require task two is equal to the number of tasks minus the stored value minus two.

The fixed data is given as:

- Processing time $\eta_k = \lceil \frac{\hat{k}-k+1}{2} \rceil + 2 \quad \forall k \in K$
- Availability $a_{et} = 19$ if $e = 2$, else $a_{et} = 38 \quad \forall t$

- Renewal time $Z_{ek1} = [1, ..., \hat{t}] \quad \forall e \in E, k \in K, t \in T$
- Experience threshold $g_{k1} = 6 \quad \forall k \in K$
- Requirements $r_k = \lceil \frac{k}{3} \rceil, \quad \forall k \in K$
- Training budget $= \pounds 1,000$

As with Extension Two, the initial skills of the employee's, Θ_{ek}, are randomly selected from the Bernoulli(0.5) distribution. The value of the binary variable d_{ki}, equal to one if auth k requires auth $i \in \{1, 2\}$, is sampled from the Bernoulli(0.5) distribution.

Using a model size of [5,12,7] there are 4,330 decision variables and 1,643 constraints. Though no more decision variables are added to the model under this extension, the number of constraints is increased greatly. The number of constraints added to the model is equal to $\hat{k}^2\hat{t}$. As the number of tasks is the largest index in this model, and would be in most cases, this is expected to increase the number of constraints exceedingly.

Though the number of constraints is much larger than any other extension or training policy, the run time for 100 random runs of data does not exceed the wall time of 600 seconds. The maximum run time found for this problem was 60 seconds, making it an efficient model. By specifying the dependencies of each auth, the size of the solution space is greatly reduced. Indeed, the minimum skill gap is much higher than seen in previous models due to the additional constraints on the solution space.

The skill gap can be seen in Figure 5.22. The minimum possible skill gap is 16 whilst the maximum is 30.

However, due to the assumptions made, the dependency variable d_{ki} cannot equal 0 for all k, or conversely 1 for all k, in which case the extremes are not evaluated. If all tasks had zero dependencies it would produce the base model with a minimum skill gap subject to alternative constraints, thus obtain skill gaps similar to those seen in TEM Zero; much lower than the skill gaps obtained here. If all tasks were dependent on both tasks the problem would be highly constrained and the skill gap would be maximised.

This behaviour can be seen in Figure 5.22. For an average initial skill less than 6, such that an employee has half the skills, the skill gap is larger because it is unlikely that the employee will have the skill to complete task one or two and hence cannot learn any other tasks. For any number of initial skills, the large skill gaps occur when all tasks require auth one or all tasks require auth two. Here, the probability that each employee has the specified initial auth is equal to 50%. Hence, it is not likely that enough employees will have the correct auth. For an average initial skill greater than 6, therefore, the lowest skill gap occurs when there is a probability of 50% that a task will require auth one, and thus auth two. Hence, the probabilities coincide and there are more likely to be employees available to complete the associated task.

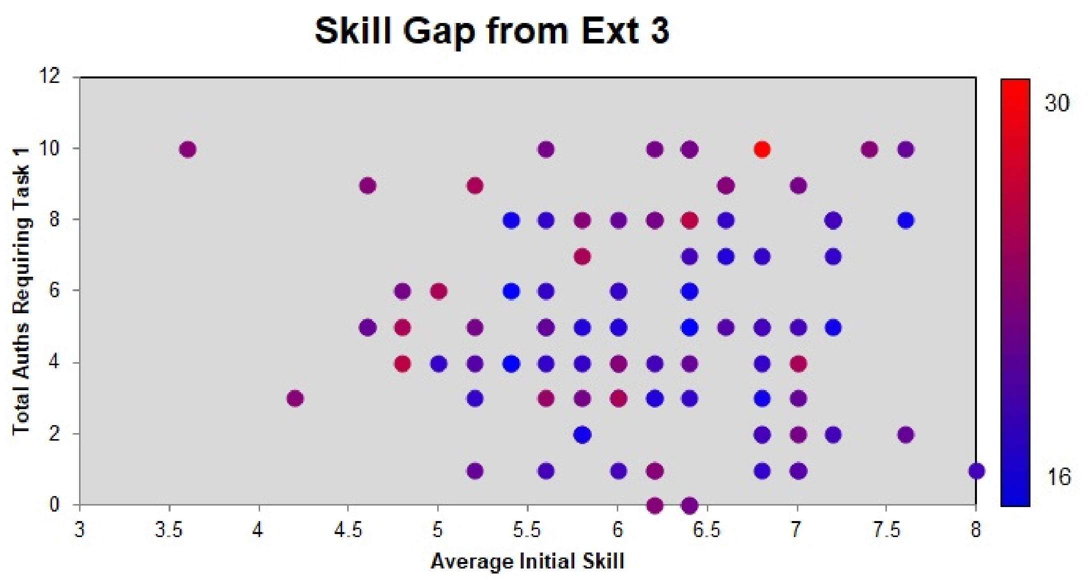

Fig. 5.22: Skill gap in Ext Three given changes to coefficients

As stated above, the average initial skill affects the skill gap such that the more skills an employee has at the beginning of the time horizon, the lower the skill gap. The dependency constraints are less likely to be binding here as an employee is more likely to have the dependent skill.

Though the relationships are suggested in the graph, there is insufficient evidence that a correlation between initial skill and skill gap or auth dependencies and skill gap exist. In particular, the relationship between auth dependencies and skill gap cannot be linear if the skill gap is maximised at the extreme values.

5.4.7 Extension Four

For Extension Four, tasks are defined with specialities or levels. As such, each task can only be completed by an employee of given levels or specialities depending on how it is defined. Thus, the two variables of interest specified in this model are the level of the task and the level of the employee.

The other data remains the same as used in previous models, with the addition of initial employee skills which are fixed such that all employees have all skills at the beginning of the time horizon. No training is available other than through repeating tasks, which is only available to those who already have the skill. As such, skills cannot be gained here. As the initial skills is of less interest in the model than the two specified coefficients, the values of the initial skills will be fixed.

Hence the following data is fixed:

- Processing time $\eta_k = \lceil \frac{\hat{k}-k+1}{2} \rceil + 2 \quad \forall k \in K$
- Availability $a_{et} = 19$ if $e = 2$, else $a_{et} = 38 \quad \forall t$
- Renewal time $Z_{ek1} = [1, ..., \hat{t}] \quad \forall e \in E, k \in K, t \in T$
- Experience threshold $g_k = 6 \quad \forall k \in K$
- Requirements $r_k = \lceil \frac{k}{3} \rceil \quad \forall k \in K$
- Initial skills $\Theta_{ek} = 1 \quad \forall e \in E, k \in K$
- Training budget $= £1,000$

The size of this model is given by [5,12,7] with $\hat{l} = 2$ where $l \in L$ is the number of levels or speciality types. Hence, a problem is obtained with 4,330 decision variables and 1,055 constraints. The number of constraints is increased by ekt due to the additional constraints on possible allocations. The number of decision variables does not increase from the base model.

Each task may have more than one specified level such that it can be completed by employees of different levels. To randomly select the level of the tasks, a general distribution is defined as follows

$$\phi_{kl} = \begin{cases} 0 \text{ for } l = 1, 1 \text{ for } l = 2 \text{ if } 0 \leq x < 0.33 \\ 1 \text{ for } l = 0, 1 \text{ for } l = 2 \text{ if } 0.33 \leq x < 0.66 \\ 1 \text{ for } l = 1, 1 \text{ for } l = 2 \text{ if } 0.66 \leq x \leq 1 \end{cases} \tag{5.3}$$

where x is randomly distributed from Uniform(0,1). As such, each task has an equal probability of requiring an employee of level one, two or either.

The level of the employee is a fixed value, b_{el}, such that one employee cannot hold multiple levels, hence first it is determined whether an employee is of level one, b_{e1}. To generate data for this sample from the Bernoulli($0.01i$) distribution where i is the iteration number. As an employee must have one skill speciality, this implies that $b_{e2} = 1 - b_{e1}$. Generating these values based on the iteration number allows a range of data values to be evaluated.

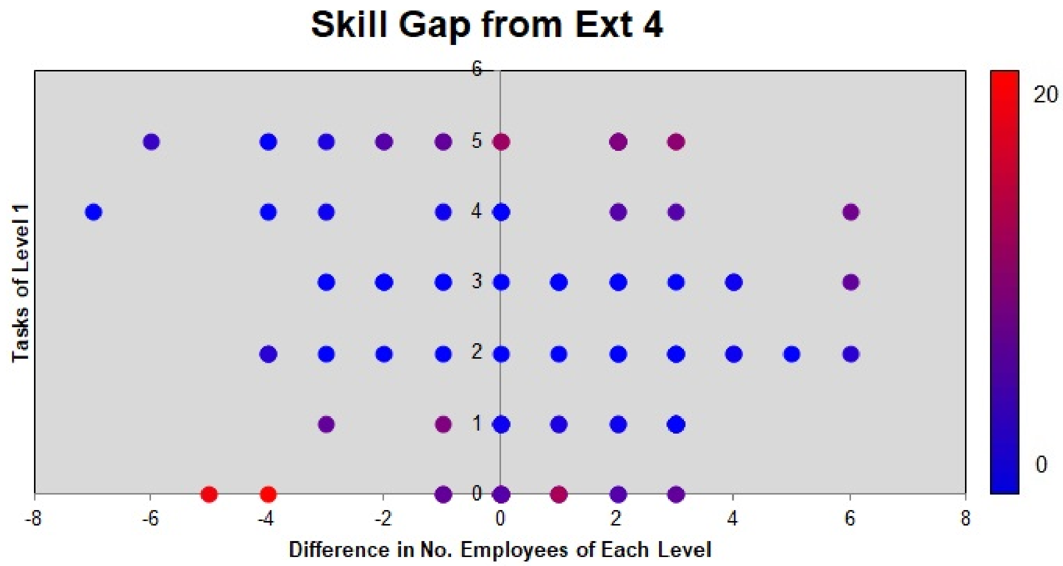

Fig. 5.23: Skill gap in Ext Four given changes to coefficients

The results can be seen in Figure 5.23.

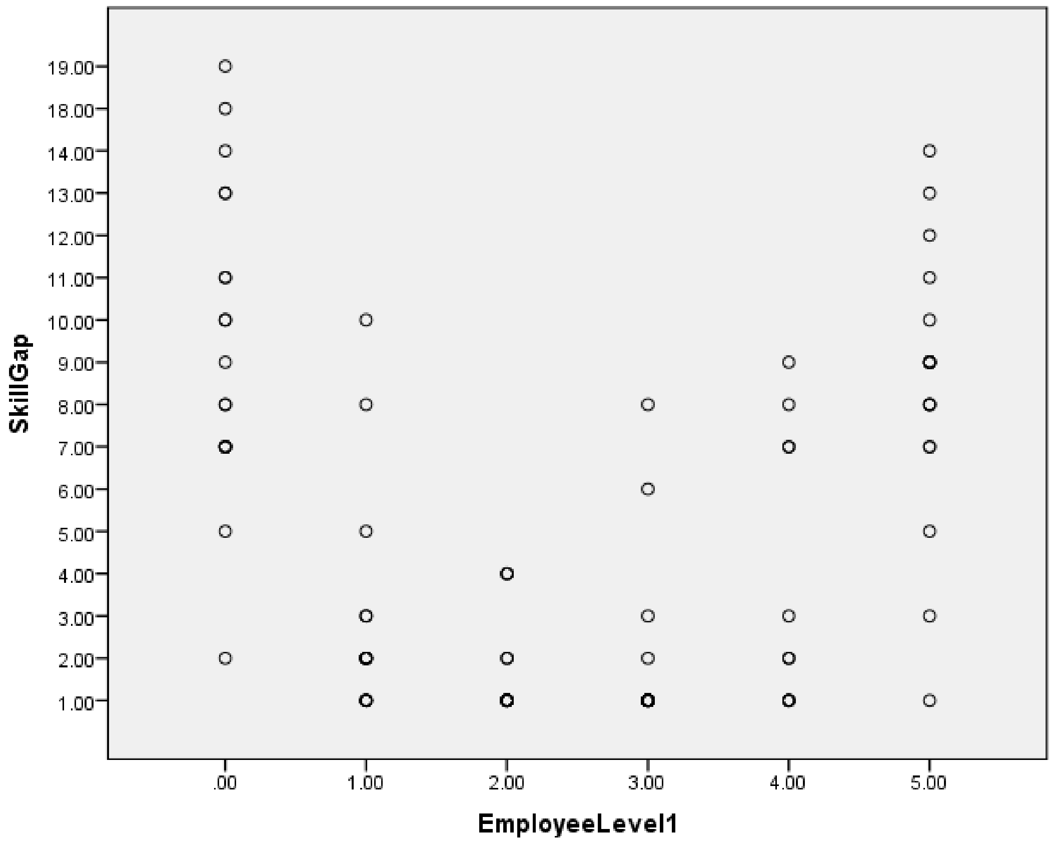

Fig. 5.24: Relationship between employee level and skill gap

As two skill levels are being evaluated, the difference is calculated as the number of tasks that require level two minus the number of tasks that require level one.

From Figure 5.23, a relationship between employees of level one, difference in task levels and skill gap can be seen. The skill gap is maximised if there are no employees of level one and more tasks require level one than level two. The skill gap is also large if all employees are level one but the number of tasks that require level one is much lower than the number of tasks that require level two. At these values, there is not enough workforce of the correct levels to satisfy the tasks that need completing, hence the larger skill gaps.

This relationship can be seen clearly in Figure 5.24. As with previous examples, the results shown are not independent of the value of the task levels. Thus, variation is expected. However, the defining bound shows the employee level creates a binding constraint on the problem. The relationship between the task level and the skill gap is less defined.

As the relationship seen in Figure 5.24 is non linear, correlation cannot be used to prove it exists. However, a line of best fit can be used where the hypothesis is that a cubic or quadratic equation will best fit this model. To analyse this, propose the linear, quadratic, cubic and exponential models and compare the R^2 values. The results can be seen in Table 5.4.

Equation	R Squared
Linear	0.003
Quadratic	0.567
Cubic	0.577
Exponential	0.001

Tab. 5.4: Ext Four goodness of fit tests for employees of level 1 and skill gap

Though the Cubic model is the best fit compared to the other lines proposed, it only explains 58% of the variation in the skill gap.

Through this analysis, it has been shown that the relationship between the employees of level one and the tasks of level one affect the skill gap obtained in the model. This has been shown in the graph. However, there is not enough evidence to suggest the skill level of the employee is sufficient in explaining the skill gap.

Finally, the run times of this model are produced. The results from the 100 runs can be seen in Figure 5.25.

No runs in this model exceeded the wall time limit of 600 seconds. Indeed, the maximum run time was 23 seconds. As with Extension Three, the constraints on the allocations has significantly restricted the solution space. Note that a problem where all employees are of level one or a model where all employees are of level two produce a much lower run time.

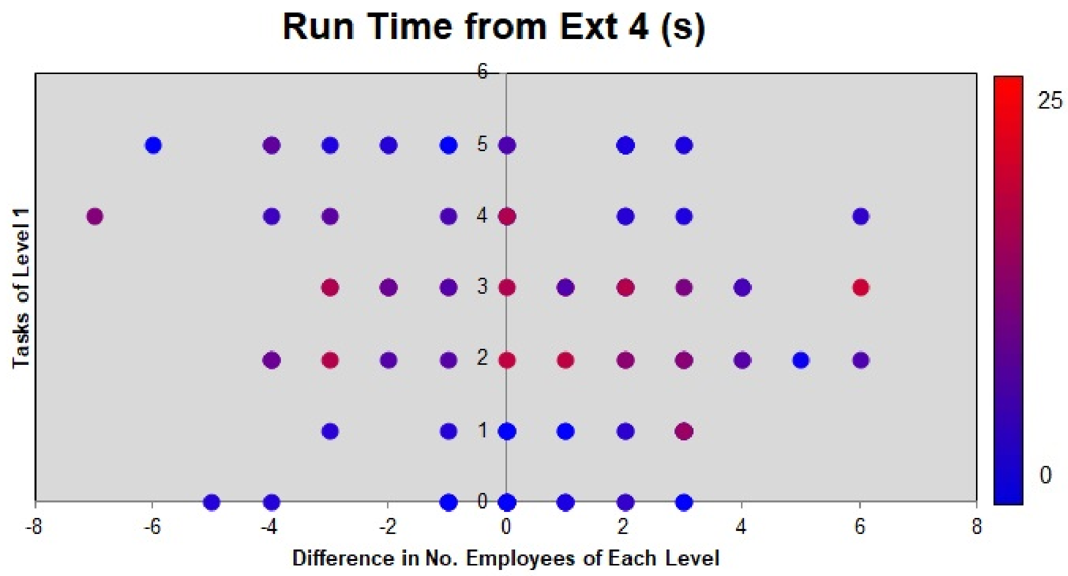

Fig. 5.25: Run times in Ext Four given changes to coefficients

If there are tasks that require level one only to be completed, however all employees in the model are of level two, this task cannot have an employee allocated to it. Thus, the problem size is reduced and the run time is reduced.

5.4.8 *Extension Five*

In Section 4.3.5,the example learning curve given in Figure 5.26 was proposed. This learning curve is created to show the expected behaviour of learning under our modelling conditions. The equation is defined as $g_{kx} = -1(0.9^{(z-x)}) + c$, where x is the number of times the skill is used.

The parameters z and c are calculated by the required shape of the curve. It is assumed that when $x = 0$, y is equal to the maximum experience required. When $y = 0$, it is assumed that x is equal to the maximum number of times the task requires completing in order for an employee to be considered competent, thus requiring 0 experience to renew a skill. From these two assumptions the values of z and c may be calculated. Hence, let $z = -17.66$ and $c = 8.47$ to obtain the shape seen in Figure 5.26.

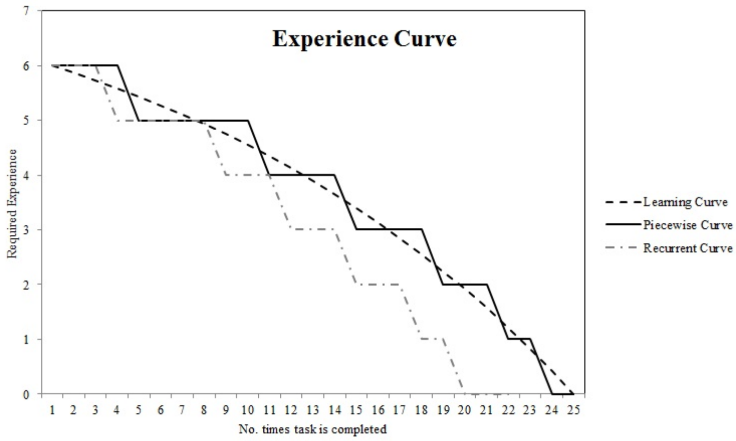

Fig. 5.26: Learning curve $g_{kx} = -1(0.9^{(-17.66-x)}) + 8.47$ using random data

Thus, two parameters are required to define the shape of the learning curve, the value of x when y is zero and conversely for y. In this analysis, the shape of the curve is adjusted using these parameters and the results analysed. It is assumed that the shape of the learning curve is fixed for all k, such that all tasks require the same level of experience.

The following data is fixed:

- Processing time $\eta_k = \lceil \frac{\hat{k}-k+1}{2} \rceil + 2 \quad \forall k \in K$
- Availability $a_{et} = 19$ if $e = 2$, else $a_{et} = 38 \quad \forall t$
- Renewal time $Z_{ek1} = [1, ..., \hat{t}] \quad \forall e \in E, k \in K, t \in T$
- Experience threshold $g_k = 6 \quad \forall k \in K$
- Requirements $r_k = \lceil \frac{k}{3} \rceil \quad \forall k \in K$
- Initial skills $\Theta_{ek} = 1 \quad \forall e \in E, k \in K$
- Training budget = £1,000

As with previous extensions, all employees begin the time horizon with all skills. As we are determining how much experience is required only, no training can take place. In addition, the allocations are not affected by previous skills of the employee.

With a problem size equal to [5,12,7], the number of decision variables is equal to 4,330 such that no variables are added to the model. The number of constraints, however, increases to 683 such that $2kz + 1$ constraints on learning are required.

The values of x, when $y = 0$, are randomly sampled from the Uniform(1,50) distribution for the 100 runs and y, when $x = 0$, are randomly distributed from the Uniform(1,10). The results are shown in Figure 5.27.

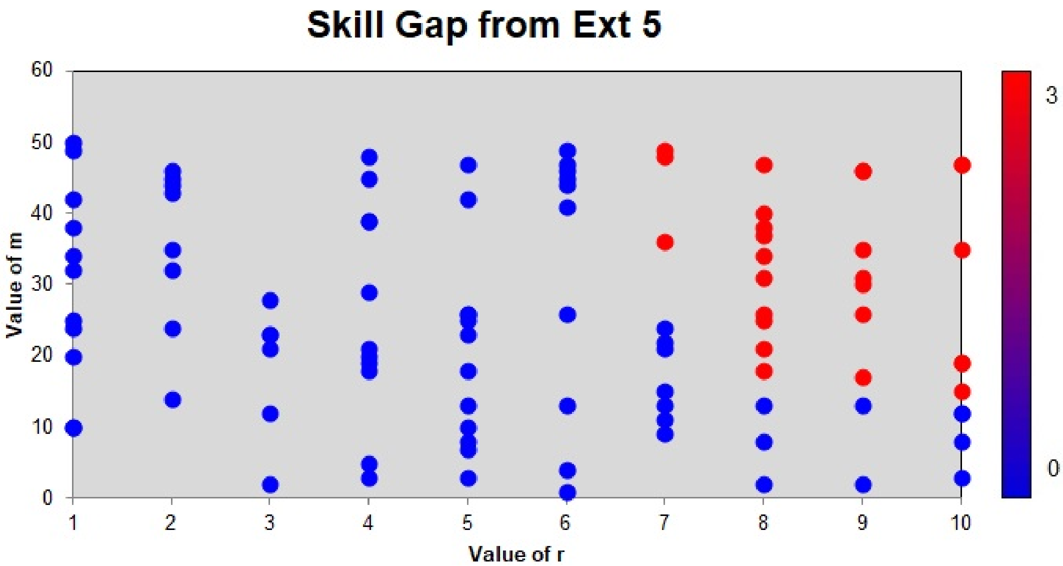

Fig. 5.27: Skill gap in Ext Five given changes to coefficients

In this analysis, only two values of the skill gap are possible solutions given the changes to the two parameters; one or three. There is a clear bound where each solution is possible. Suggesting a relationship between the values of x, y and the skill gap.

To investigate this further we analyse some of the learning curves randomly created in this experiment. The learning curves generated can be seen in Figure 5.28.

The values of x and y define where these plots cross the axis. As such, negative values of the experience may occur. Consider the solutions to these four curves. When $x = 9, y = 40$, the optimal skill gap is 3, else the skill gap is equal to 1.

Consider a task k with constraints on requirements and availability such that it may only be completed a maximum of 8 times within the time horizon by an employee. We determine the experience required according to these curves. For $x = 9, y = 40$, an employee who has completed this task 8 times, requires an experience of 9 to receive the authorisation. Hence it is not possible to obtain this auth. In comparison, the other curves require an experience of less than 8. Thus, the auth can be awarded according to the other curves.

If x and y are high, the experience required to gain a task will remain high for all values on the x-axis. The task k which can be completed the least number of times within the time horizon will determine which values of x and y produce a larger skill gap. Hence, a constraint is created such that the experience required is too large to be feasibly achieved within the time horizon.

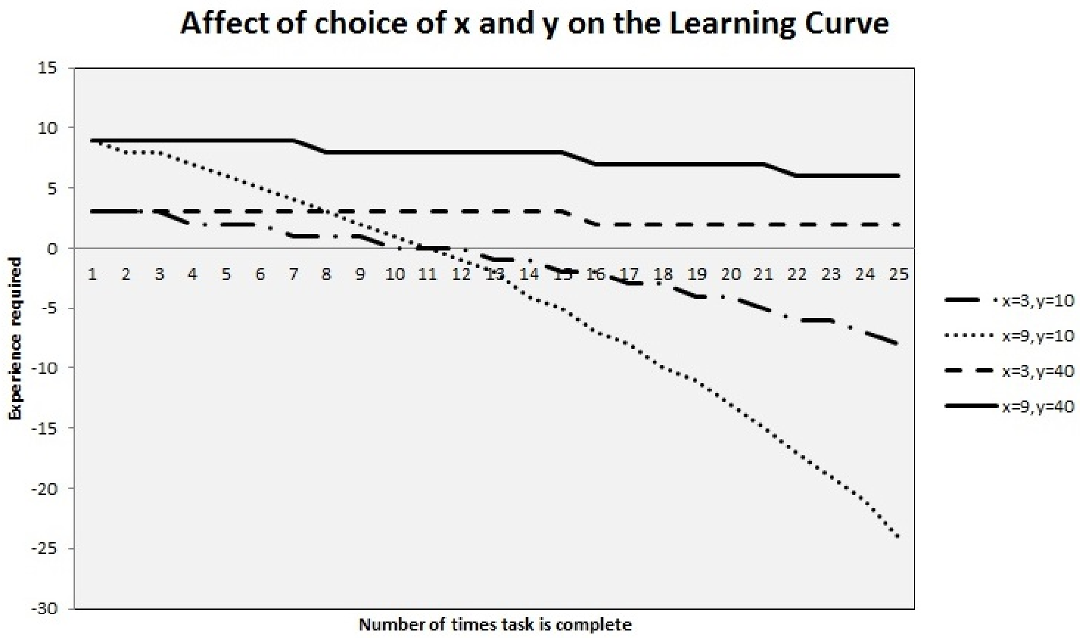

Fig. 5.28: Generated learning curves from Extension Five

However, a low value of x and a high value of y creates a low experience required for all values on the x-axis. A high value of x and a low value of y creates a large enough gradient that few tasks need completing before the experience required is at an achievable value.

Finally, consider the run times under the different coefficients. For 100 runs, there was no instance where the run time exceeded 600 seconds. However, we consider the relationship between x and the run time and y and the run time seperately. The results can be seen in Figure 5.29.

It should be noted that these results vary both x and y and calculate the run times, hence the results are not independent. However, there is a suggestion that the value of x is correlated to the run time, for small x the run times are always small. A small x indicates that for any number of task completions, the experience required is very low. Hence, the solution space is reduced and therefore the run time.

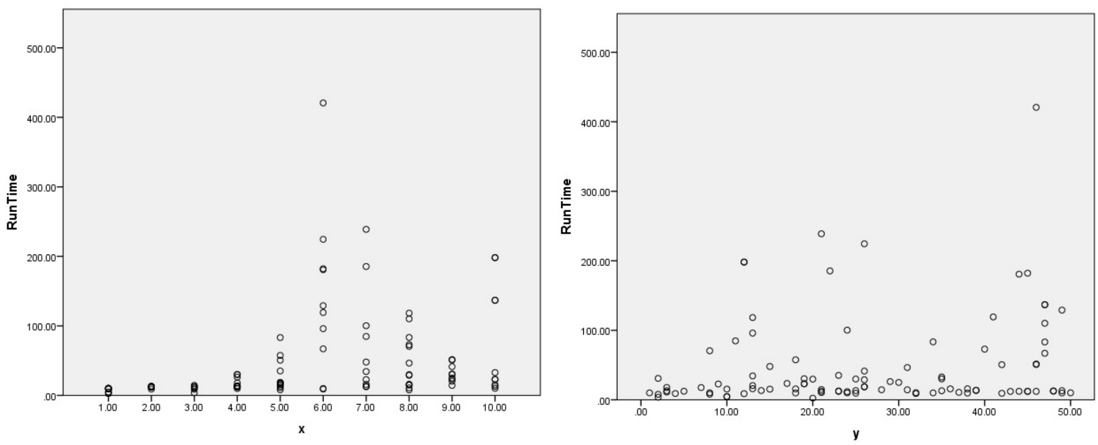

Fig. 5.29: Run time analysis for Extension Five

5.4.9 Extension Six

The final coefficient analysis looks at Extension Six and the addition of stochastic tasks. Though many variables are introduced in this extension, here we consider the limits on the chance constraints. In particular, we vary the values of α and β.

The fixed data is set as follows:

- Processing time $\eta_k = \lceil \frac{\hat{k}-k+1}{2} \rceil + 7 \quad \forall k \in K$
- Availability $a_{et} = 19$ if $e = 2$, else $a_{et} = 38 \quad \forall t$
- Renewal time $Z_{ek1} = [1, ..., \hat{t}] \quad \forall e \in E, k \in K, t \in T$
- Experience threshold $g_k = 6 \quad \forall k \in K$
- Requirements $r_k = \lceil \frac{k}{3} \rceil + 3, \quad \forall k \in K$
- Initial skills $\Theta_{ek} = 1 \quad \forall e \in E, k \in K$
- Training budget $= £1,000$

The processing time is increased by five on the previous models, as the availability constraint is now affected by probability as detailed in Section 4.3.6. The requirements are increased by three. In this way, the problem becomes less trivial. In addition, the probability of each task occurring within each time unit and the probability of the scenario, given sample average approximation is used, are chosen such that:

- $\gamma_{k\xi_i} = 1 - 0.01(\xi_i k)$; the probability of task k occurring in scenario ξ_i
- $\zeta_{\xi_i} = \frac{1}{\hat{\xi_i}}$; probability of the scenario

Thus, we must determine values for α and β. Here, α refers to the confidence interval such that the user will be confident that the optimal solution satisfies $1-\alpha\%$ of the scenarios. The percentage of tasks that must be complete in order for a scenario to be classed as satisfied is given by β. We sample $\alpha = 0.1i$ for $i =$ 1 to 10. Similarly, we sample $\beta = 0.1i$ for $i = 1$ to 10. All combinations of these values are sampled to create $10^2 = 100$ runs. There will be five scenarios sampled in this example.

The feasibility from 100 runs using this data can be found in Figure 5.30

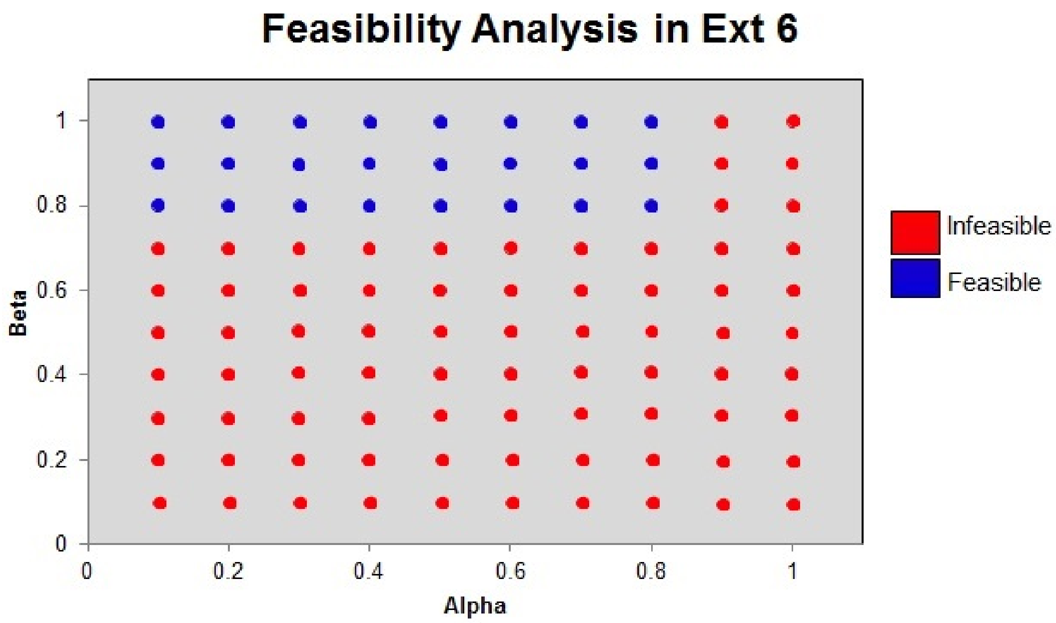

Fig. 5.30: Feasibility analysis on Ext Six given changes to coefficients

Unlike previous extensions and models, this extension can create infeasible problem sets given the data used. Thus, the results show the feasibility of the problems produced in the simulation.

The results indicate clear values of α and β after which problems sets become infeasible. In this way, for any value of α, there is an insufficient number of employees to complete more than 70% of the tasks. Similarly, no more than 80% of the scenarios can be complete with the current workforce. Given, that each scenario occurs with probability 20% under our conditions, this means that one scenario cannot be completed.

The results shown here are very dependent on the data provided. However, it indicates that there will be values of α and β in which solutions will be defined as feasible or infeasible. A much larger size problem may show more variation within the solution of the skill gap. In this instance, the skill gap remained the same for all feasible solutions.

Under this data, no run times exceeded the wall time of 600 seconds. This indicates that the additional constraints have reduced the solution space. Indeed, for infeasible solutions the run time would be much smaller. There did not appear to be any correlation between the values of α and β and the run time of the model.

5.5 Pareto Curve

Due to the multi-objective nature of the model, it contains an objective with a weighted coefficient. In previous analysis this coefficient has been set large enough to ensure the skill gap is prioritised, under the assumption that this is the most common objective of the employer. Here, the value of the coefficient is varied under random data to determine the impact the coefficient has on the decisions made. All models are combined in this analysis with the exception of Extension Six to avoid infeasible solutions.

One Pareto curve will be generated for each set of random data. The curve compares the two objectives of the model; fairness and skill gap. Here, note that the skill gap is calculated over all time as given in the objective and not as viewed in the previous results. The coefficient is varied over the range of interest. For a problem size of [5,12,7] the coefficient A, as defined in Section 4.2, is set to zero to obtain the maximum possible skill gap, the sum of the requirements. The coefficient is then iteratively changed until the solution to run i is equal to the solution for $A = M$, such that M is selected to produce the minimum possible skill gap. It is assumed that any coefficients between these values will give the same solutions due to the nature of the objective.

For this study, five sets of random data are used such that there are different maximum skill gaps in each set. The data distributions used are provided at the beginning of this chapter. However, the training budget was reduced to £500 to further constrain the problem. The resulting Pareto curves are plotted in Figure 5.31 with each colour representing a different set of data.

As expected, the maximum skill gap is obtained when the coefficient is set to zero such that the fairness objective is minimised. Conversely, the minimum skill gap is obtained when the coefficient is large in each data set.

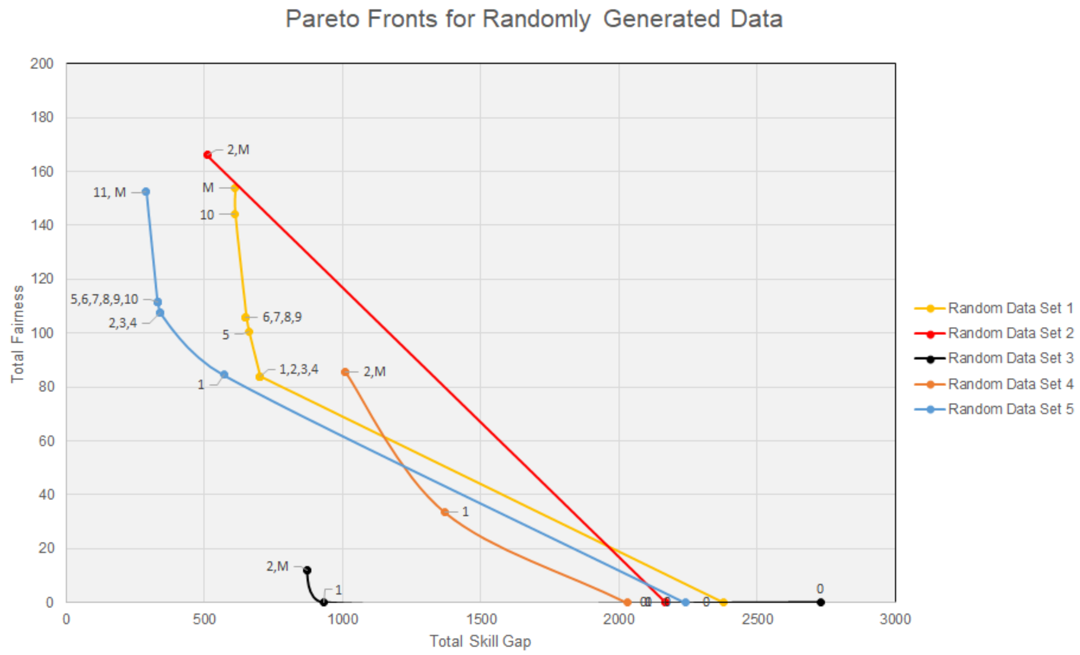

Fig. 5.31: Coefficient analysis in the objective function

For these five data sets the maximum coefficient required to ensure the skill gap has priority was 11. The average value over the runs was 5.4 and the median was 2, indicating a low coefficient is required to minimise the skill gap in most cases.

Consider the following equation for the objective function.

$$\min \quad \textstyle\sum_{e\in E}(\psi_e^+ + \psi_e^-) + A\sum_{k\in K}\sum_{t\in T}(r_k - \sum_{e\in E} w_{ekt}) \tag{5.4}$$

where the first term indicates the fairness in terms of difference in number of allocations. The second term defines the weighted skill gap.

The fairness value is calculated as the total difference between the number of tasks completed by each employee and the average number of tasks completed by all employees.

$$\frac{1}{\hat{e}}\sum_{e\in E}\sum_{k\in K}\sum_{t\in T} w_{ekt} - \sum_{k\in K}\sum_{t\in T} w_{fkt} \leq \psi_f^+ \qquad \forall f \in E \tag{5.5}$$

$$\sum_{k\in K}\sum_{t\in T} w_t - \frac{1}{\hat{e}}\sum_{e\in E}\sum_{k\in K}\sum_{t\in T} w_{ekt} \leq \psi_f^- \qquad \forall f \in E. \tag{5.6}$$

It is proposed that the fairness is maximised when one employee completes all tasks. This can be proved by calculating the fairness.

Hence, the maximum fairness is given as

$$\begin{aligned}
&\sum_{e\in E}(\psi_e^+ + \psi_e^-) \\
&= (\hat{e}-1)\frac{1}{\hat{e}-1}\sum_{k\in K}\hat{t}r_k + \sum_{k\in K}\hat{t}r_k \\
&= 2\hat{t}\sum_{k\in K} r_k
\end{aligned}$$

To ensure skill gap is given priority, a skill gap of one must exceed the maximum fairness value when multiplied by the weighted coefficient A:

$$2\hat{t}\sum_{k\in K} r_k = A \tag{5.7}$$

However, in this calculation it is assumed the only constraint on the allocation is the requirements. In which case a maximum fairness value is possible such that it is always feasible that one person can complete all tasks in each time unit. As this is not feasible given the data in this model, the value of A required in these runs is much lower than the value indicated here.

The multi-objective structure is useful in models where two or more objectives are contradicting each other. However, in this model, this is not always the case. Indeed, it is possible for the fairness value to always be minimised for any value of the coefficient. However, this is not true for the skill gap.

The run times were also recorded for the values of the coefficient. For a coefficient of zero, the run time of the model was close to 0. The solution for a zero coefficient is trivial; no employees are assigned to tasks as the skill gap does not require minimisation. Thus no training or allocations take place, thus the fairness of the employees is always zero, this can be seen in the first graph. Hence, the small run times are obtained. Otherwise, there is no correlation between the value of the coefficient and the run time of the model but instances where the model exceeded the wall time did occur.

As the problem size increases the required value of A to obtain the minimum skill gap is reduced. Examine equation (5.4); until a skill is reviewed, the skill gap will be equal to the requirements for that task minus the available workforce given that no training has been completed. In most cases, the skill will not be reviewed early in the model. If the skill is reviewed, the employee is unlikely to have shown sufficient experience to be awarded the auth. As such, the skill gap at this time will not be minimised. As the skill gap is summed over all time units, this results in a very large value of the skill gap, with the value increasing as $\hat{e}$, $\hat{k}$ and $\hat{t}$ increase. As such, for larger problem sizes, a small coefficient is significant to prioritise the skill gap. However, if the skill gap under the current workforce is equal to zero, this coefficient may not be suitable and the recommended value of A will be required.

5.6 Run Time Analysis

As shown in the coefficient analysis, for some data combinations the run time of the model can exceed the specified wall time. Through analysis of the output files, it can be seen that this is caused by an inability to prove optimality of the solution.

The run time limit is set to 600 seconds. Through further analysis of the models, it was discovered that any run that exceeds this limit does not complete after two hours even for relatively small problem sizes. However, the objective value from these runs remains the same between 600 seconds and two hours. The process takes a long time due to the difficulties in proving optimality experienced by Cplex. This could be due to a large number of nodes within the model and a large number of potential solutions to each model. For the problem sizes looked at in this analysis, if a solution can be found in less then two hours, then the solution will also be found in less than five minutes.

In this section, the run times will be closely examined to determine when this effect takes place and what size problem can be solved in a reasonable amount of time. Though the model itself is strategic such that it does not need to be run each week, for instance, it would be beneficial to have a model that runs within a day.

In this section, the run times are calculated for a variety of problem sizes. The size of the model is considered in terms of number of decision variables and constraints and how these change by adjusting the size of the number of employees, tasks and time units. Each model will also be examined independently to determine if a specific set of constraints are more likely to cause these large run times. As such, recommendations can be provided as to which extensions and models could cause increased run times.

5.6.1 Combined Problem Size Evaluation

The sizes of the model have, thus far, been calculated on an individual level. However, when the different models are combined, the model size increases at a much larger rate. Here, the dimensions are considered in turn; number of employees, number of tasks and number of time units.

All TEMs are used in this model, all extensions are also considered with the exception of Extension Six. This will not be used in this analysis due to the number of infeasible solutions created by this constraint affecting the run times. If the number of employees is too small for the number of tasks, the skill gap will be large, causing a large number of infeasible solutions. The additional dimensions introduced by the TEMs and extensions are not considered as variables in this analysis.

A thorough analysis is required into the cause of the long run times. The wall time is set to 600 seconds such that any run that exceeds this time is assumed to have problems solving optimality.

First, the runs are performed for three different problem sizes given that the only problem dimensions that will change are the number of employees and the number of tasks. Though the number of time units may also affect the model, it is believed this has a smaller impact on the run time. This assumption will be investigated later in this section. By adjusting both the number of tasks and employees simultaneously, the affect of the number of solutions on the run time is reduced. If the number of employees is much less than the number of tasks, there is a large increase in the number of solutions. The same effect is caused by a small number of employees completing a large number of tasks. The number of potential allocations is much greater. Hence, in this analysis, the problem size is kept to similar proportions and the time horizon of the model is not considered.

Other dimensions also affect the size of the model, however these are policy specific. In this analysis, the only dimensions that are investigated are the indices that are present in all models.

The data is randomly generated following the distributions specified in the introduction of this section. For 100 runs, the results can be seen in Figure 5.32. The problem sizes are [5,12,7], [10,15,7] and [15,25,7] as specified in Chapter 4.

- Problem size [5,12,7] has 5,240 decision variables and 2,785 constraints
- Problem size [10,15,7] has 13,060 decision variables and 5,825 constraints
- Problem size [15,25,7] has 32,490 decision variables and 14,820 constraints.

The other dimensions used in the extensions are given in Section 5.2. The number of decision variables and constraints is significantly larger than applying a single model of a similar size. For instance, the base model TEM Zero with a problem size of [5,12,7] contains 4,330 decision variables and 1,476 constraints.

The increase in problem size also increases drastically with a small change in any of the indices. The run times can be calculated for the three instances. The analysis is performed with 100 runs of random data and the results are plotted in Figure 5.32.

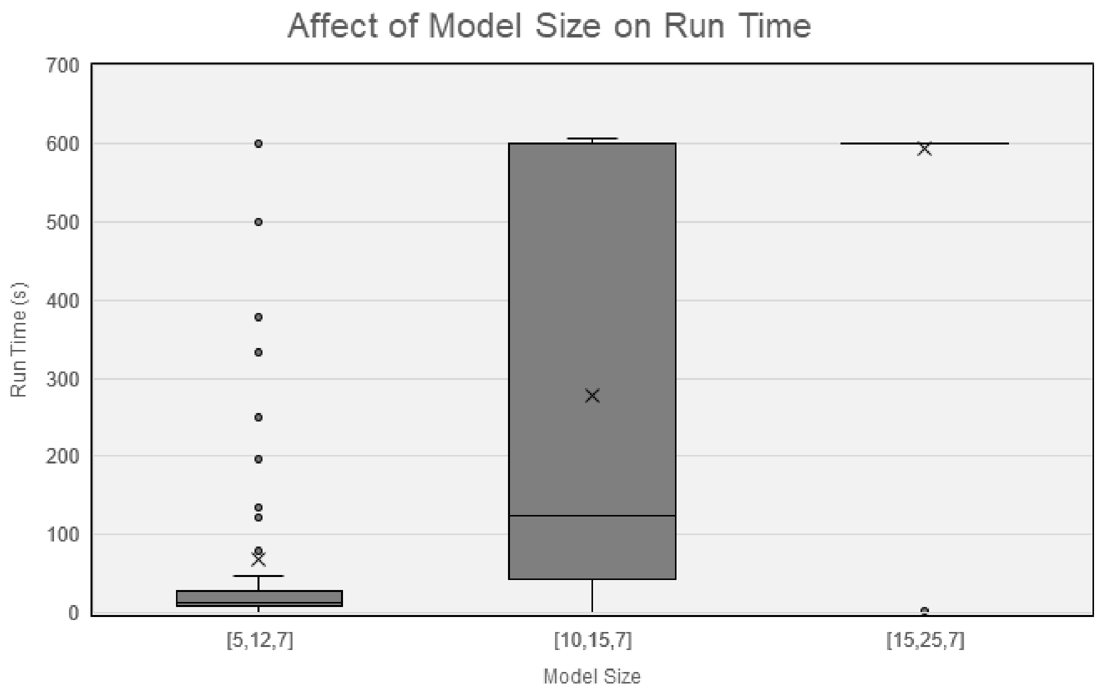

Fig. 5.32: Run times for various model sizes

The figure shows that the run time increases exponentially given the size of the problem. Given a problem size of [5,12,7], there are 5,240 decision variables and 2,785 constraints. The resulting run times under this size are mostly quick. In most cases, the model runs within a few seconds. However, under some combinations of data, the run time did exceed the wall time. This happened in 2% of the data combinations.

For a model size of [10,15,7], the number of decision variables increases to 13,060 and the number of constraints to 5,825. In this case, the number of data combinations that exceeded the wall time increased to 31%. The overall range of run times has also increased, the inter quartile range of run times is now given as 562 seconds, given the maximum run time of 600 seconds. However, the mean run time is still reasonably low at 248 seconds.

Finally, the problem size of [15,25,7] has been analysed. Here, 84% of the run times exceed the maximum time. The number of decision variables is given as 32,490 and the number of constraints is 14,820. The minimum run time is still very low, 7.3 seconds, however most combinations of data in this model have problems proving optimality.

To analyse these run times in more detail, the individual policies must be investigated to determine if one training model or extension could cause problems proving optimality. This will will be done later in this chapter.

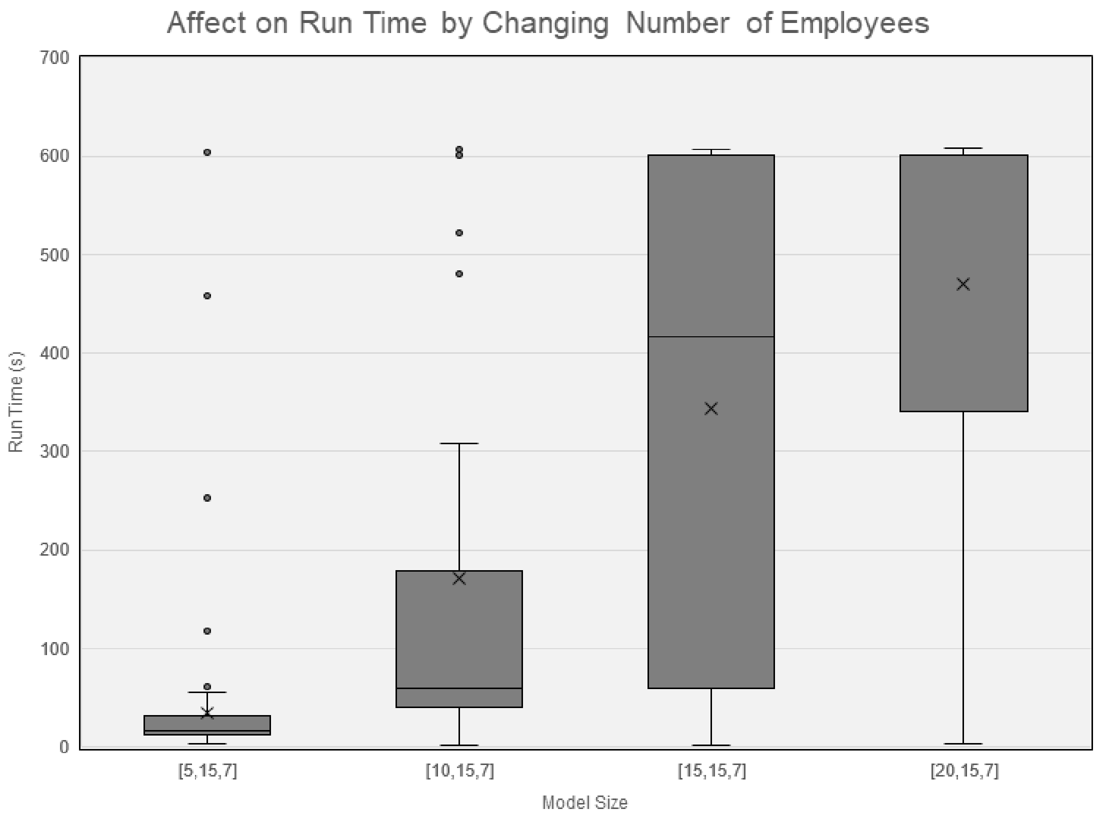

Fig. 5.33: Graph of number of employees against model run time

First, however, the problem is broken down into the three indices that remain consistent in all models; employees, tasks and time units. We have shown the problem size and run times have increased exponentially as both the number of tasks and employees has increased concurrently. Now, to further analyse the behaviour of the model, the run time is considered with each index considered independently. In this way, the behaviour of a model with a different ratios of employees to tasks and time can be analysed.

Figure 5.33 shows the run times for different problem sizes in which the number of employees is changed but the tasks and time units remain consistent.

The number of employees in the model causes significant changes to the run times. If more employees are used, there are more possible solutions as it is more likely that multiple employees will have the skill sets and availability to complete the same tasks.

When the number of tasks are greater than the number of employees, the number of possible solutions is reduced. Hence, less instances are exceeding the wall time by failing to prove optimality.

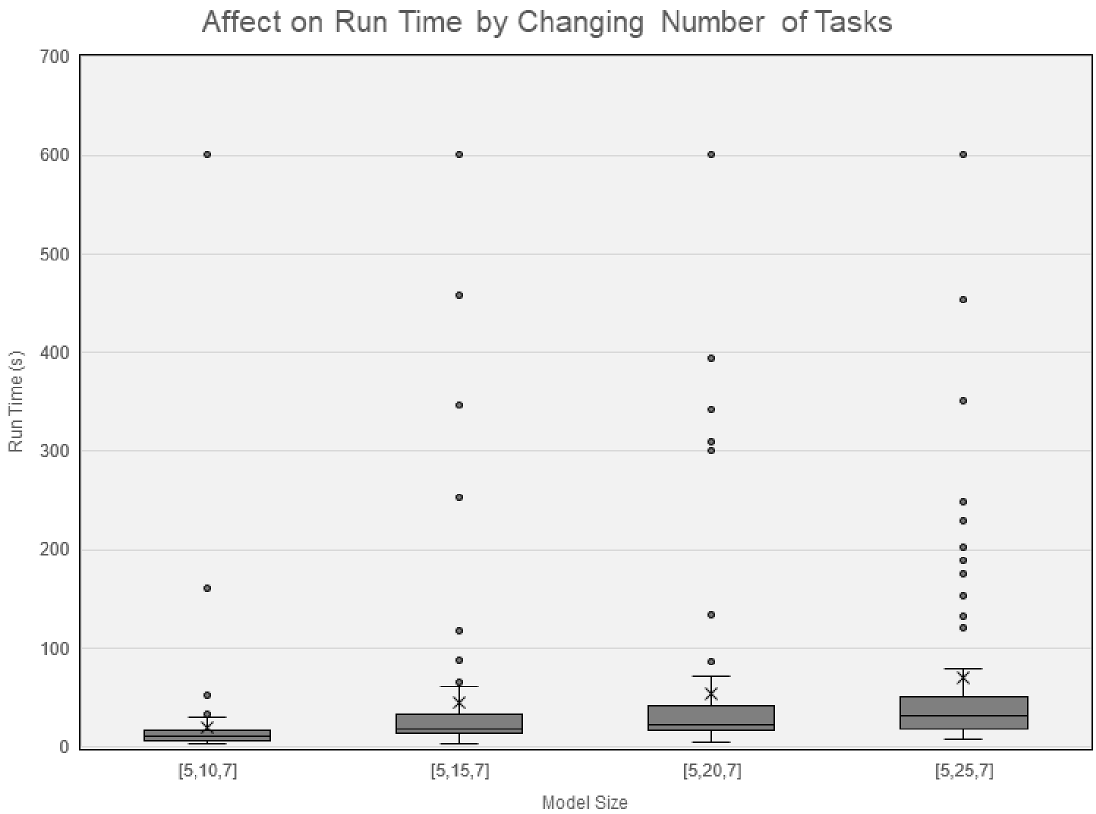

Fig. 5.34: Graph of number of tasks against model run time

Figure 5.34 similarly shows the run times for various number of tasks but consistent employees and time units.

The average run time increases with the number of tasks, as with employees, however the change is less significant. The model size also affects the skew of the data such that larger problem sizes are more likely to result in larger run times.

Though increasing the tasks creates more potential solutions to the model, it has been shown that the number of decision variables dependent on the tasks is much smaller than the number of variables dependent on employees. Hence, it is expected that the run time is less correlated to the number of tasks.

Finally, the time units are varied in Figure 5.35 with employees and tasks remaining fixed. There does not appear to be any correlation between the number of time units and the run time of the model.

The solution to a problem after one time unit is no different to the solution to the problem after x time units where x is the time before the first review. This is because, for this data set, the availability and skill sets of the employees do not change with time. Thus, a very large number of time units is required to show any significant changes to the run time. Indeed, more than one review period would be required to show a change in the run time.

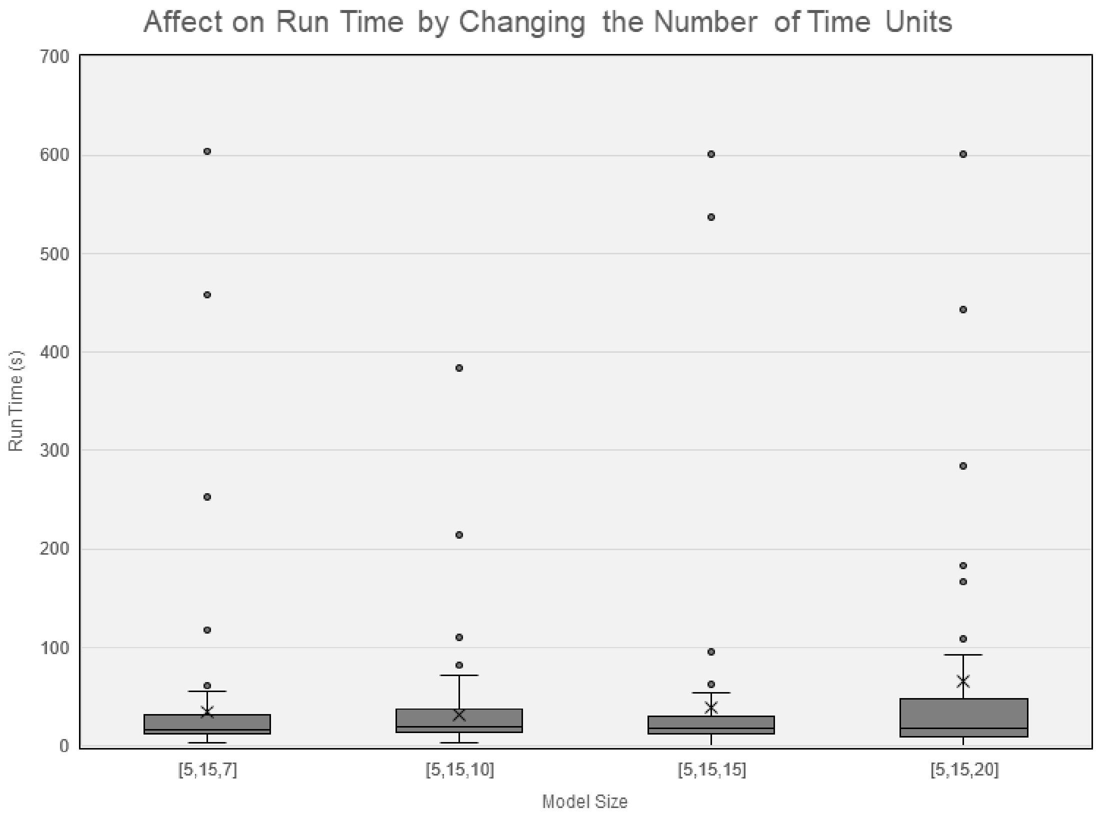

Fig. 5.35: Graph of number of time units against model run time

To further explore where the model has difficulties proving optimality, Figure 5.36 shows the percentage of runs in which the wall time is exceeded. Note, the x-axis indicates the size of the model as a rank, however the rate of increase of employees, tasks and time units are different for the relevant indices. As such, the model size is indicated as a label.

It can be seen from this figure that the number of employees has a significant impact on the number of solutions that fail to prove optimality.

By combining the output in Figures 5.36 and 5.32, it could be seen that the ratio of employees to tasks does not affect the run time or the number of solutions that fail to prove optimality. Hence, it can be concluded that the number of employees only has a significant influence on the probability of the model proving optimality for any given set of data.

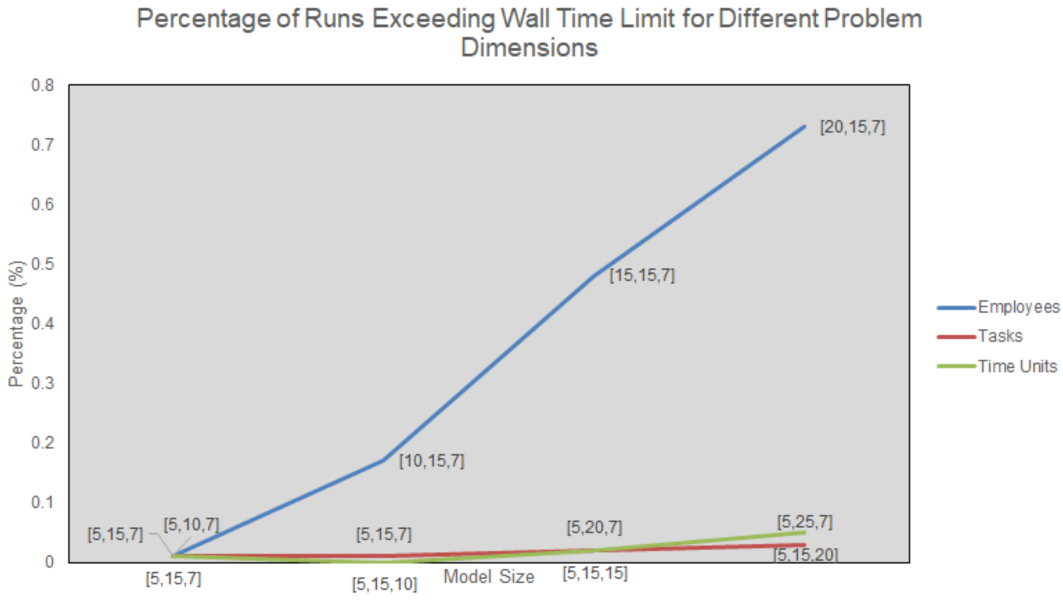

Fig. 5.36: Number of instances where the wall time was exceeded for various model sizes

5.6.2 Individual Problem Size

In the previous section, the run time and how it is affected by the size of the problem, the number of decision variables and constraints was produced. These were recorded for different numbers of employees, tasks and time units. Here, a similar analysis is conducted but for the individual models. As such, conclusions may be drawn as to the effect of the selected constraints on the solvability of the model.

Due to the model being linear, increasing any index will increase the number of decision variables and constraints at a linear rate. The rate will be recorded for each of the training execution models and extensions. As such, they can be compared and a hypothesis can be made as to which set of constraints may result in the longest run times.

The number of decision variables and constraints differ for each model used as shown in Table 5.5. The model dimensions used in this study is [10,20,7]. The number includes the additional index required to make integer decision variables into binary decision variables. Due to the data used for this model, the number of indices required to create a binary value is $m = 10$ where m is required in the binary variables w_{ektm} and p_{ektm}.

Model	Decision Variables	Constraints
TEM Zero	14,420	1,961
TEM One	15,820	3,361
TEM Two	14,560	1,975
Ext One	15,820	2,361
Ext Two	14,420	2,361
Ext Three	14,420	4,761
Ext Four	14,420	3,361
Ext Five	14,420	2,041
Ext Six	14,425	1,967

Tab. 5.5: Number of decision variables and constraints in [10,20,7] model

The rates for decision variables can be found in Figure 5.37 for any problem size. The rate of change of constraints for the same data sizes are in Figure 5.38.

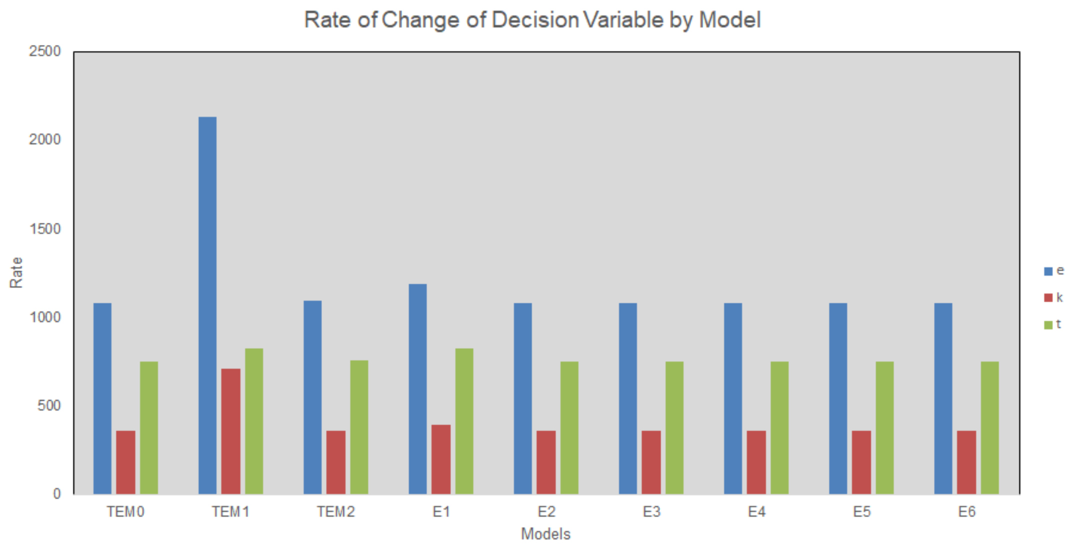

Fig. 5.37: Rate of Change of Decision Variables

First, consider the rate of increase in decision variables. The employees, e, cause the greatest increase in model size, followed by time, t. The number of tasks k has less impact on the model size in terms of number of decision variables. The number of decision variables required for the base model is calculated as $ektm+et+2e$. The change in model size described above reflects the change seen in the calculation by adding additional decision variables to this base model.

The base model, TEM Zero, contains the same number of decision variables as Extensions Two, Three, Four and Five. These models further constrain the model but do not increase the number of decisions that need making.

In contrast, the number of constraints in each model vary in a different pattern. Again, in all models except Extension Three, the number of employees has the greatest impact on the number of constraints. The number of tasks has the smallest.

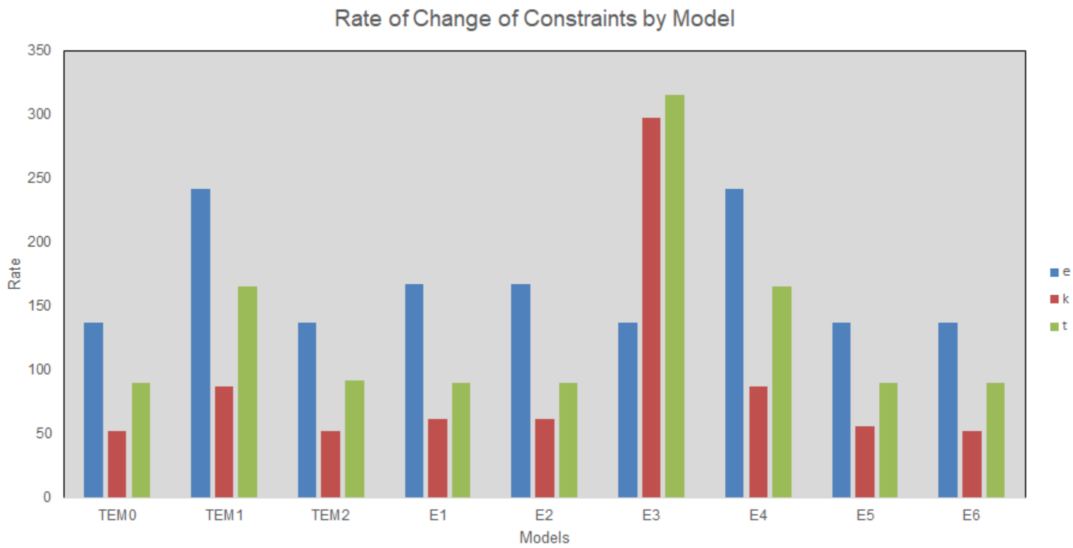

Fig. 5.38: Rate of Change of Constraints

For Extension Three, authorisations have dependencies, as such a constraint is required for each dependency combination and each time unit; k^2t.

All models have a different rate of increasing number of constraints. Extension Three has the largest constraint increase rate. TEM Two, training courses, and Extension Four, skill specialities, have the joint second largest number.

These results will be considered when looking at the run times for each model in a later section.

5.6.3 Individual Model Evaluation

The change in the number of decision variables and constraints has been calculated for each of the models and an investigation has been made into the dimensions of the model. In this section, an investigate is made into how the choice of policy affects the run time of the model. Thus, each model component will be evaluated separately. The problem size in this case will be small enough to give an idea of how the chosen model affects the results but still in a reasonable run time.

Using results from the previous section, the problem size is set as ten employees, twenty tasks over seven time units. Here, it is expected that some instances will occur where the problem size exceeds the run time limit of the model, however, the differences should mostly be caused by the selected decision variables and constraints and not by the size of the model. The number of skill specialities to be considered is two and the number of external training courses is also two. There are five scenarios required to model the stochastic tasks.

The run times for all models are produced separately and compared. The following plot shows the results.

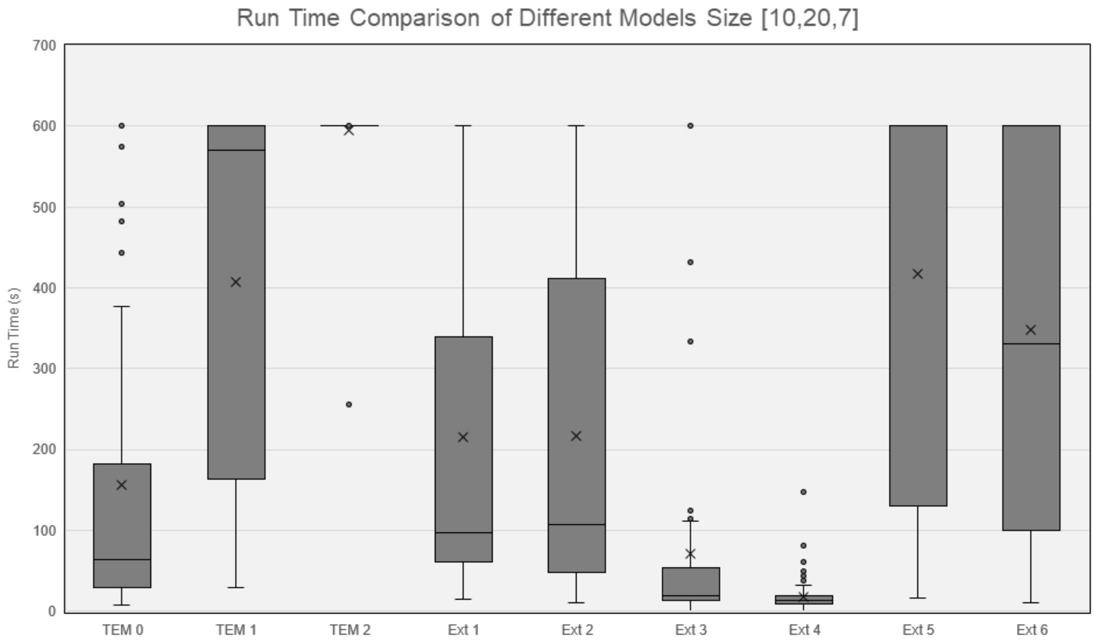

Fig. 5.39: Run times for each model at a size of [10,20,7]

It should be noted that these run times include times where the run took more than 600 seconds to complete, at this point the wall time of the model is exceeded. Table 5.6 shows the percentage of runs where the run time exceeded the wall time.

First, consider the policies with the smallest run times and the smallest percentage of runs with a run time exceeding 600 seconds. These are Extensions Three, skill specialities, and Four, auth dependencies.

Extensions Three and Four contain the same number of decision variables as the TEM Zero base model. However, they both increase the number of constraints significantly. Hence, the number of possible solutions remains the same but the problem is much more constrained, reducing the solution space. The number of possible allocations is reduced.

Model	Percentage (%)
TEM Zero	11
TEM One	49.5
TEM Two	98
Ext One	18.6
Ext Two	20.3
Ext Three	6
Ext Four	0
Ext Five	60
Ext Six	46

Tab. 5.6: Percentage of runs with a run time exceeding 600 seconds

In contrast, TEM Two increases the solution space by including an alternative allocation for employees that does not affect the requirements for the tasks without further constraining the problem. The number of decision variables has increased more than the number of constraints.

Here, 98% of the runs exceeded the wall time with issues proving optimality, due to the large solution space. There are many instances where the objective value can be similar but with very different solutions depending on which training the employee has been assigned to.

TEM One, on-the-job training, Extension Five, non linear learning, and Extension Six, stochastic tasks, can also create large run times in most instances. The constraints in these model add more solutions to the space and additional dimensions to the model.

5.7 Summary

This chapter contained an analysis of the behaviour of the model, through a completion of a coefficient analysis on each of the models and a run time analysis.

Initially, the training budgets of the model were explored. It was determined that the training budget does affect the output of the model such that the skill gap and run time depends on a specified threshold budget. This threshold indicates where the budget constraint is binding. For all values of the budget where the constraint is binding, the skill gap is increased and the run time is decreased. This threshold value can be useful when determining a starting budget for running the model. The value can then be adjusted until a solution has been presented to satisfy the customer requirements.

A coefficient analysis was performed on each model variant. First, TEM Zero was analysed by adjusting the initial skills of the employees and the requirements of the tasks. It was concluded that the requirements are much more significant in predicting the skill gap of the solution than the employee skills. They also had an impact on the run time of the model.

TEM One considered the cost of on-the-job training only. It was discovered that a value of the cost exists such that any higher cost produces the maximum skill gap such that no training is complete. The behaviour of the skill gap through changes to the cost shows the behaviour of the budget constraint. The run time is also affected by the choice in this coefficient.

For the final TEM, training courses, the experience from the training course and the cost were varied. Both variables could be seen to have a significant impact on the skill gap produced by the model. Though the results were inconclusive as to which variable was the better predictor for the skill gap. The run time did not appear to be affected by the choice in these variables.

A similar analysis was also performed on the model extensions. Extension One showed a significant change in the skill gap by adjusting either the compulsory training cost or the number of tasks requiring compulsory training. However, it was concluded that the compulsory requirements had a more significant relationship.

The variations to Extension Two demonstrated the influence of the initial skill of the employees on the skill gap when recurrent skills were used. This result differs from the result seen when only the base model, TEM Zero, is used. This indicates the relationship between recurrent skills and the initial skills of the employee.

Extension Three shows auth dependencies. By adjusting the initial skill and the number of auth dependent on task one, as opposed to task two, it can be seen that auth dependencies have an influence on the skill gap of the model when compared to other models. However, the value of the initial skill or the number of dependent auth does not have a significant relationship with the skill gap.

The skill levels demonstrate the relationship between the skill levels of the tasks when compared to the skill levels of the employees. If the number of employees who are level 1 is large but the number of tasks that require level 1 is low, then the skill gap is maximised. Similarly, for a large number of employees of level 2 where the number of tasks requiring level 2 is low.

Using the formula to define the learning curve from previous sections, the parameters of the curve can be adjusted to analyse the impact of the shape on the results. The possible solutions observed are caused by the gradient of the curve which are defined by the parameters. These values formulate the shape and determine whether a feasible solution is possible. If the shape of the curve is not formulated correctly, it is possible to get infeasible solutions or solutions in which the skill gap is always maximised.

Finally, Extension Six evaluated the feasibility of the model given different parameters on the chance constraints. Though the model used showed little difference in the skill gap, the feasibility was clearly related to the choice in parameters.

For random instances of data, the Pareto Curves were produced to indicate the behaviour of the model given changes to the weighting in the objectives. Though a calculation for the minimum value of A required to optimise the skill gap is determined, the results indicate a lower value of A often minimises the skill gap due to the relationship between the two objectives.

From the coefficient analysis, it was shown that there were instances where the solver had difficulties proving optimality. It was shown through increasing the model size that the larger the model, the more likely these instances were to occur.

Through an analysis of the employees, tasks and time units, it was concluded that the number of employees is correlated to the number of problems that are difficult to solve and the run times of the problems.

When considering this relationship for each model, the number of employees again has a significant impact on the size of the model. It was concluded that the number of employees increased the size of the problem at the highest rate in each model.

The run times for each model were also produced. In the instances where the model became more highly constrained, Extensions Three and Four, the run time of the model were greatly reduced. Where the models offered more training and allocation options, TEM Two, the run times were greatly increased and it was more likely that the optimality could not be proven. However, there was no relationship between the number of constraints and the run time of the model. Similarly with the number of decision variables.

It was concluded that problems with a larger problem space had more problems proving optimality. The instances where this occurred were discussed through the comparison of the different models. The number of employees were discovered to have the greatest impact of the size of the model and the size of the model also increased the problem space and hence the run time.

In the next chapter, the models will be applied to a case study such that the results discussed here may be applied to a real world problem.

6. APPLICATION TO CASE STUDY

6.1 Problem Development

This project is a collaboration with Boeing. A tool was requested to assist decision making for personnel on the maintenance line. The proposed Training Development Model in this thesis has included policies to address the requirements of Boeing. By selecting relevant constraints, the model can be tailored to support the training options of the maintenance workers.

The training constraints specified by Boeing in this case study correspond to some, but not all, of the policies introduced in Chapter 4. In this section, the corresponding policies are selected and combined to produce a MILP that forms the basis of the Training Development Model. An Excel based tool is designed for any application of the model such that relevant constraints may be chosen as necessary. The case study for Boeing is given as an example of the functionality of the tool and an analysis and understanding of the results is provided.

The case study and corresponding data in this section represent a subset of a problem proposed by Boeing using randomly generated data. It provides an insight into how this generic tool can be applied to a workforce setting using realistic constraints representing a real world problem. The size of the workforce has been reduced from the real world study to enable an analysis of the results to be completed and the model to run in a sufficient time.

6.1.1 Description

The purpose of this case study is to explore the skill development of a team of engineers at Boeing. The engineers provide the maintenance of a fleet of aircraft provided by Boeing for use by their customer. As such, the engineers must be trained in relevant skills in order to complete the tasks at hand. Engineering involves a set of complex tasks where it is essential that the employees are properly trained in order to reduce work place accidents. Due to the aircraft being used for military purpose, it is also required that the fleet is available as part of the contract. Any delays due to lack of manpower affect the availability of the fleet and thus should be avoided. In order to ensure the staff are correctly trained, the optimal allocation to training should be determined. As discussed in this thesis, the first step is to recognise where gaps in the workforce exist. Through the application of the TNA, any unfulfilled tasks will be ascertained.

In this section, the case study is given as a complete formulation with a description of current processes in place at Boeing. Where possible, simplifications and assumptions are stated in order to maximise the efficient running of the model. However, it must be noted that these simplifications will impact the effectiveness of the model at reproducing the real life situation.

A hanger is set up to allow aircraft to be maintained and repaired. Jobs consist of preventative maintenance tasks such as upgrades, checks or scheduled replacements. There are also emergency jobs that can appear at any time, given a probability, with unknown requirements.

In our problem specification the maintenance line for scheduled maintenance is considered only. This line consists of deterministic tasks to improve the capability of the aircraft. In this way, no stochastic tasks will be considered. This decision is primarily due to a lack of available data concerning random events. It was found that process and data changed much quicker than the length of the thesis research. Though stochastic tasks play an important role in the repairs, data has not been recorded in a suitable way to complete a full analysis.

In order to fulfill these tasks, the maintenance line is supported by a multi-skilled workforce of engineers. For each task that needs fulfilling, employees must be signed off for the associated authorisation. Considering the employees as multi-skilled allows them to have more than one auth and thus be able to complete a range of tasks. There are a large number of skill combinations for each employee, hence the individual training requirements of each employee in the hanger should be analysed using a mathematical method.

As well as performing maintenance tasks, the employees have other responsibilities such as paperwork, these are known as divergent tasks. For some employees these divergent tasks are essential and must make up a proportion of their time. However, the problem description specifies maintenance tasks only, hence paperwork does not require authorising. The divergent tasks are thus not modelled as auths but it is assumed that these tasks can be completed in any additional time throughout the day or affect the availability of the associated employee.

In addition to recognising where the skill gap lies, Boeing are interested in the training and skill development of each employee. Hence a training model is required. Constraints are specified for gaining authorisations and thus can be related to the policies discussed earlier in this thesis.

Three stages of training are required. Phase one consists of basic military training and phase two consists of learning basic maintenance skills and attending maintenance school. At this stage employees all have the same essential skills, thus the model does not consider training needs before this point. After phase two employees are sent to their respective hanger where training takes place on-the-job and through additional training courses. As such, Training Execution Methods One and Two should be included in the model as defined in Section 4.2.

Tasks are specified such that compulsory training may be required, as such Extension One will be included in the model, Section 4.3.1. In this case, the compulsory training specifies a prerequisite course that must be completed or an interview that must be attended.

Tasks have different requirements if they are being renewed, thus Extension Two should be included, Section 4.3.2. It is specified in the task description that tasks requires different amount of experience based on whether or not they are being renewed, hence they are recurrent. As discussed in this policy, the time between reviews is large enough that it may be assumed that if a person did not have the skill at the previous time step they must complete initial training even if they had the skill in the past.

Some authorisations are also dependent on others as shown in Extension Three, Section 4.3.3. This is typical in engineering where the knowledge to complete a specific engineering task may depend on current knowledge of how to complete a similar or simpler task.

To complete tasks, two types of engineer are required; avionic and mechanic. These employee types correspond to the types of task, as perceived in Extension Four, Section 4.3.4. These specialities also coincide with the training courses. Hence, for TEM Two, the number of training courses is equal to two where a training course gives you experience in all skills with the corresponding skill type. Though each course will be available to all employees, due to the skill specialities an employee will only need to complete a course if it benefits the auths which can be completed by their speciality. It is possible that some tasks can be completed by either a mechanic or an avionic specialist. However, it is not possible for an employee to have both specialities.

As stated above, the stochastic nature of tasks cannot be calculated in this case study, hence Extension Six is not used. The non-linear nature of learning is also difficult to be measured in this case study and thus it is assumed that learning is linear only and Extension Five will not be used.

Thus, the skill development analysis will be produced with TEMs Zero, One and Two describing the training methods. The Extensions; One, Two, Three and Four are also included as specified in Section 4 to describe other constraints on the training allocations.

6.1.2 Data Assumptions

Due to the complexity of this real life problem, some simplifying assumptions are required. In this section the assumptions proposed in the previous section are summarised and new assumptions are included to model the case study.

First consider the following general assumptions that affect all components of the model independent of the selected extensions:

1. **Aircraft are considered identical**. Though this seems unlikely, this is a long term strategic model and as such the details of the individual aircraft will be unknown. Indeed, due to the defence environment, the variation in possible flying hours for each aircraft will be very large and it would be challenging to forecast future damage over a long term period. Hence, it is acceptable to make this assumption.

2. **Tasks have fixed requirements and a fixed processing time**. The requirements for the tasks are given in the brief for planned tasks. The processing time are set to be the fixed amount of time allocated in the plan to complete the tasks for any aircraft. This is given as the worst case scenario for completing the task including time for potential delays. If the task is completed before the stated processing time, the additional time may be used for other non maintenance tasks, including divergent tasks. Fixing the processing time for the demand also allows the model to be more dynamic and allows a reduced problem size.

3. **Employees can be assigned to many tasks during the same time interval**, however they are constrained by their authorisations.

4. **Employees start with zero experience.** The data states that the auths are reviewed at the end of the time horizon, and the review period is the length of the time horizon. Hence, it can be surmised that the skill had been reviewed in the time unit before the model begins. Thus, the only allocations included are made during the time horizon.

5. **The maximum number of skills an employee can have is less than**

or equal to the number of tasks. If an employee had the auth for each task then they would be able to complete the total number of tasks. Though it is unlikely that they would have so many auths, the fairness constraint reflects this amount and the skill trades limit the auths each employee can gain.

6. **Experience is only gained from the last time the skill was reviewed.** In other words, the experience is a memoryless property. Thus learning a skill does not make it easier to learn the skill in future. However, recurrent skills allow the training requirements to be lowered if the employee has the skill in the current review period.

7. **Learning is linear**. It is assumed that completing a task once is equal to one time units worth of experience.

The following additional assumptions apply to the selected model extensions that are relevant to the case study:

1. **A skill is recurrent if it was used in the previous time unit only.** For instance, if a skill is lost at any time then the initial training criteria must be met for the skill to be regained. It is independent of when the skill was lost.

2. **On-the-job training does not affect the tasks ability to be completed.** Thus it does not affect the processing time. The processing time is already considered as the worst case scenario, hence includes potential delayed caused by the trainee.

3. **Course travel time is not considered**. Any time spent travelling should be included in the length of the course.

4. **Course trainers are not considered**. It is assumed trainers are provided independent of the model.

5. **Courses can take place at any time within the time horizon**. There is no constraint on when the courses may happen.

6.1.3 Model Development

The problem has been defined as a Training Development Model with TEMs Zero, One, Two and Extensions One, Two, Three and Four. From this specification, the model can be written as a MILP using the constraints and objective introduced in Chapter 4. Tables 6.1 to 6.3 defines the variables used in this model, using notation from the previous sections.

Input:	
$M >> 0$	A large number referred to as “big M”.
$A \in \mathbb{Q}$	Coefficient to define priority of objectives.
$c_1 \in \mathbb{Q}^+$	Cost of OJT training.
$c_{2n} \in \mathbb{Q}^+$	Cost of training course number n.
$c_{3k} \in \mathbb{Q}^+$	Cost of compulsory training for task k.
$b \in \mathbb{Q}^+$	Available training budget.
$r_k \in \mathbb{Z}^+$	Number of times task k will require completing in each time unit.
$a_{et} \in \mathbb{Q}^+$	Available hours of employee e at time t, given annual leave or sick leave.
$\eta_k \in \mathbb{Q}^+$	Processing time of task k.
$\Theta_{ek} \in \{0,1\}$	1 if employee e has auth k at time 0.
$\rho_n \in \mathbb{Q}^+$	Length of training course n.
$\chi_{nk} \in \mathbb{Q}^+$	Experience gained from training course n for task k.
$u_{nt} \in \mathbb{Z}^+$	Number of spaces on course n at time t
$\mu_{k\sigma} \in \mathbb{Q}^+$	Length of compulsory training k for initial, $\sigma = 1$, or recurrent, $\sigma = 1$, auths.
$g_{k\sigma} \in \mathbb{Q}^+$	Amount of experience needed for task k for initial, $\sigma = 1$, or recurrent, $\sigma = 1$, auths.
$\pi_{k\sigma} \in \{0,1\}$	Compulsory training required to complete task k for initial, $\sigma = 1$, or recurrent, $\sigma = 1$, auths.
$d_{ki} \in \{0,1\}$	1 if auth k requires the completion of auth $i \in K$.
$\lambda_{el} \in \{0,1\}$	Employee e can complete tasks of speciality l.
$\phi_{kl} \in \{0,1\}$	Task k requires speciality l.

Tab. 6.1: Data to be used in Case Study formulation

Indices:	
$e = 1, .., E$	Index of employees.
$k = 1, .., K$	Index of tasks.
$t = 1, .., T$	Index of time.
$n = 1, .., N$	Index of training courses.
$l = 1, .., L$	Index of levels.
$z \in \overline{Z}$	Index of review periods.
$t \in Z_{ekz}$	Set of times in review period z for employee e with task k.

Tab. 6.2: Indices to be used in Case Study formulation

Decision Variables:	
$w_{ekt} \in \mathbb{Z}$	Number of tasks k completed by e at t
$s_{ekz} \in \{0,1\}$	1 if employee e has auth k in review period z.
$q_{ekt\sigma} \in \{0,1\}$	1 if employee e is assigned to compulsory course k at time t for initial or recurrent auth.
$p_{ekt} \in \{0,1\}$	1 if employee e is assigned to OJT in k at time t.
$v_{ent} \in \{0,1\}$	1 if employee e completes course n at time t.
$\psi_e^+ \in \mathbb{Q}^+$	Positive deviation in fairness.
$\psi_e^- \in \mathbb{Q}^+$	Negative deviation in fairness.

Tab. 6.3: Decision Variables to be used in Case Study formulation

The first element of the problem to be defined is the objective function. As Extension Six is not used in this model, the objective for TEM Zero is used with no corrections. Hence, the objective has two components to minimise; unfulfilled tasks and fairness deviation.

$$\min \quad \sum_{e\in E}(\psi_e^+ + \psi_e^-) + A\sum_{k\in K}\sum_{t\in T}(r_k - \sum_{e\in E} w_{ekt}) \tag{6.1}$$

Due to the number of constraints, they will be introduced here in groups for input in the model. The first set of constraints model the allocations of employees to tasks and training. There are three constraints;

$$\text{s.t.} \quad w_{ekt} - r_k s_{ekz} \leq 0 \qquad \forall e \in E, k \in K, z \in \overline{Z}, t \in Z_{ekz} \tag{6.2}$$

$$\sum_{k\in K}\eta_k(w_{ekt} + p_{ekt}) + \sum_{n\in N}\rho_n v_{ent} + \\ + \sum_{k\in K}\sum_{\sigma=1}^{2}\mu_{k\sigma}q_{ekt\sigma} \leq a_{et} \qquad \forall e \in E, t \in T \tag{6.3}$$

$$\sum_{e\in E} w_{ekt} \leq r_k \qquad \forall k \in K, t \in T. \tag{6.4}$$

To model the approved allocations, equation (6.2) ensures employees can only be assigned to tasks if they have the associated skill. Equation (6.3) models the availability of each employee such that the time spent completing tasks and training does not exceed the available hours. Finally, (6.4) ensures you cannot assign more employees than required.

In accordance with the definitions of the fairness deviation. The following constraints calculate the difference between the allocations of the employees to calculate the perceived fairness of the allocations.

$$\frac{1}{\hat{e}-1}\sum_{e\in E:e\neq f}\sum_{k\in K}\sum_{t\in T} w_{ekt} - \sum_{k\in K}\sum_{t\in T} w_{fkt} \leq \psi_f^+ \qquad \forall f \in E \quad (6.5)$$

$$\sum_{k\in K}\sum_{t\in T} w_{fkt} - \frac{1}{\hat{e}-1}\sum_{e\in E:e\neq f}\sum_{k\in K}\sum_{t\in T} w_{ekt} \leq \psi_f^- \qquad \forall f \in E. \quad (6.6)$$

The initial skills of the employees are set using the equalities;

$$s_{ek1} = \Theta_{ek} \qquad \forall e \in E, k \in K \quad (6.7)$$

such that initial skill is input by the user in equation (6.7).

The skill is then calculated at the time of review. The experience gained is calculated from the TEMs; hence it is appropriate to sum the experience gained from repeating tasks, on-the-job training and training courses. In addition, the required skill is dependent on Extension Two; initial or recurrent skills.

$$g_{k1} - \boldsymbol{M}_k(1 - s_{ekz}) \leq \sum_{t\in Z_{ek(z-1)}} ((w_{ekt} \qquad \forall e \in E, k \in K,$$

$$+ p_{ekt}) + \sum_{n\in N} \chi_{nk} v_{ent}) + M_k s_{ek(z-1)} \qquad z \in \overline{Z} \quad (6.8)$$

$$g_{kx} - \boldsymbol{M}_k(1 - s_{ekz}) \leq \sum_{t\in Z_{ek(z-1)}} ((w_{ekt} \qquad \forall e \in E, k \in K,$$

$$+ p_{ekt}) + \sum_{n\in N} \chi_{nk} v_{ent}) + M_k(1 - s_{ek(z-1)}) \qquad z \in \overline{Z} \quad (6.9)$$

The compulsory training must also satisfy requirements dependent on the skill recurrency. Two sets of inequalities are defined:

$$\pi_{k1} s_{ekz} - \sum_{t\in Z_{ek(z-1)}} q_{ekt} - s_{ek(z-1)} \leq 0 \qquad \forall e \in E, k \in K, z \in \overline{Z} \quad (6.10)$$

$$\pi_{k2} s_{ekz} - \sum_{t\in Z_{ek(z-1)}} q_{ekt} - (1 - s_{ek(z-1)}) \leq 0 \qquad \forall e \in E, k \in K, z \in \overline{Z} \quad (6.11)$$

to ensure that at the time of review, the compulsory training must be complete for initial, equation (6.10), and recurrent, equation (6.11), auths.

In order for an employee to perform OJT, another employee must be completing the task in order to provide the training. The number of people who may be assigned to a training course is specified by the capacity of the course. The constraints can be written as

$$p_{ekt} - \sum_{f \in E: e \neq f} w_{fkt} \leq 0 \qquad \forall e \in E, k \in K, t \in T \quad (6.12)$$

$$\sum_{e \in E} v_{ent} \leq u_{nt} \qquad \forall n \in N, t \in T \quad (6.13)$$

Extension Three indicates that some auths can only be given if other auths are also given. This is modelled with;

$$s_{ekz} - s_{ei\tilde{z}} \leq 1 - d_{ki} \qquad \forall e \in E, k \in K, z \in \overline{Z}, \; i \in K, \tilde{z} \in Z_{ekz} \cap Z_{ei\tilde{z}} \quad (6.14)$$

Skill levels, for Extension Four, are included with the following constraints:

$$w_{ekt} - r_k \sum_{l \in L} \lambda_{el} \phi_{lk} \leq 0 \qquad \forall e \in E, k \in K, t \in T \quad (6.15)$$

Finally, the budget for training is specified to include the cost of training courses, the cost of on-the-job training and the cost of compulsory courses;

$$\sum_{e \in E} \sum_{t \in T} \left(\sum_{k \in K} c_1 p_{ekt} + \sum_{n \in N} c_{2n} v_{ent} + \sum_{k \in K} \sum_{\sigma=1}^{2} c_{3k\sigma} q_{ekt\sigma} \right) \leq b \quad (6.16)$$

Thus, the individual components of the TEMs and extensions have been combined to form a Training Development Model to determine the training allocations of engineers at a maintenance line in Boeing. This model can be used to produce workforce recommendations under the particular training constraints of the company.

6.2 Decision Support Tool

To implement all models within a company, a tool is required for inputting data, running the model and outputting the results. Though IBM CPLEX ILOG Optimisation Studio (CPLEX) is used as the linear programme solver, Microsoft Excel 2010 is used as the user interface to input data as it is more user friendly and the results can be displayed as needed. In order to complete the optimisation, the LP solver must be called from the Excel file using Visual Basic for Applications

(VBA). The process is shown in Figure 6.1.

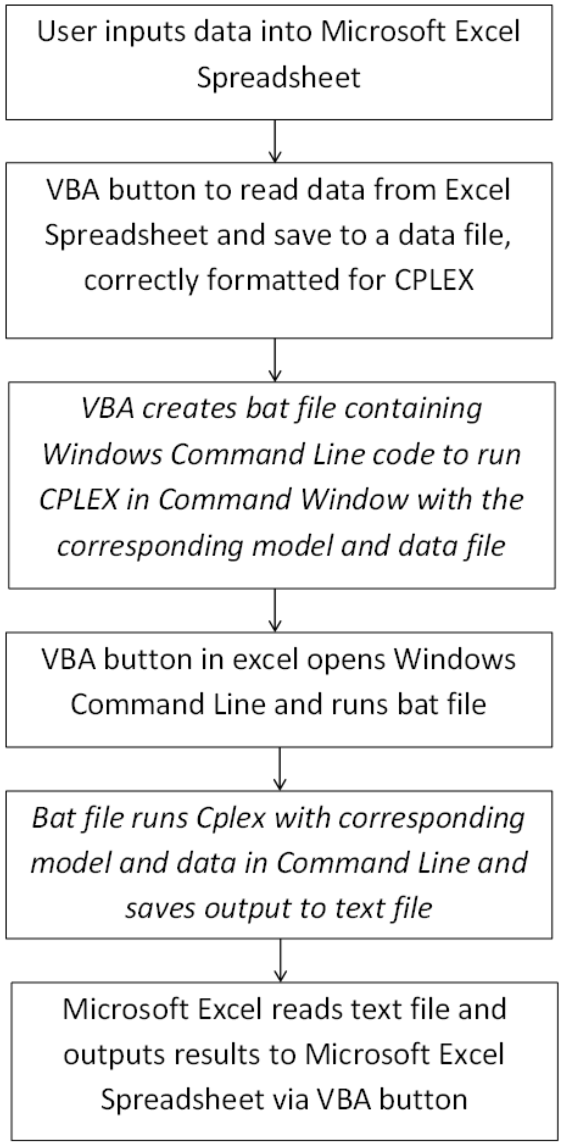

Fig. 6.1: Flowchart of process for running Cplex through Excel

The text in italics indicates tasks that Excel completes without requiring user interactions. The other boxes indicate actions that must be initiated through pressing a button or inputting data.

In the following section, a detailed user guide will be presented for using the tool.

6.2.1 User Guide

The decision support tool is a Microsoft Excel tool to recommend the allocation of employees to training programmes subject to constraints on training budget and operations. The following section is a detailed user guide to describe the operation of the tool.

The tool consists of several Excel sheets. The first sheet, “InputParam” contains the interface of the model. The model parameters are input here and the four steps to perform the optimisation can be selected. The other sheets contain the input data required for the selected policies. Figure 6.2 shows the tool upon opening. Label 1a indicates the step by step process for running the optimisation. Each step is initiated by pressing a button. Label 1b is for inputting the model parameters such as the size of the model and the file path, given for storing the model and data files. The parameters shown here are essential for the TNA and TEM Zero and hence are required for the tool to function. Similarly, Label 1d shows the available sheets upon initiating the file. These sheets must be populated with data in order to run the tool.

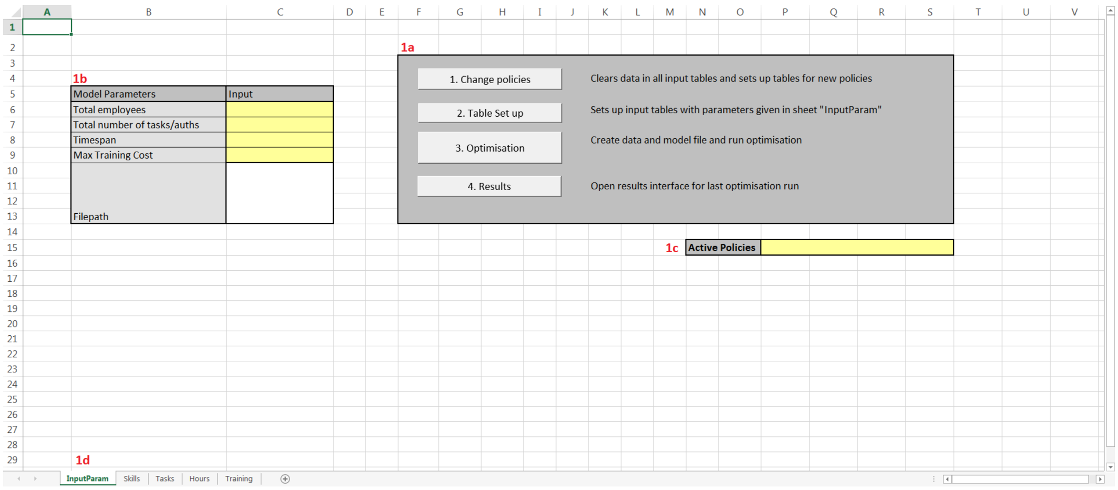

Fig. 6.2: Screenshot of Tool - Initiate

The tool is capable of running any combination of TEMs and extensions. Label 1c specifies the active policies in the model. Upon initiation, this is set to blank to indicate that no additional TEMs or extensions are active. As such, the only active constraints are those given by TEM Zero.

The model is initiated through the four steps in Label 1a. The first stage of running the model is to change the active policies. By clicking the button “1. Change Policies”, a message box appears containing information for each of the policies and the ability to select which policies are required. Figure 6.3 displays the screen upon activation of this button. The left hand side of the message box shows a list of all eight optional TEMs and extensions. By selecting each option is turn, the right hand side gives a brief description of the policy, main assumptions and the data required to implement this policy. Multiple policies may be used simultaneously depending on the operations of the company and as such there is also an option of selecting all. Once the relevant policies have been selected the user should press "Go".

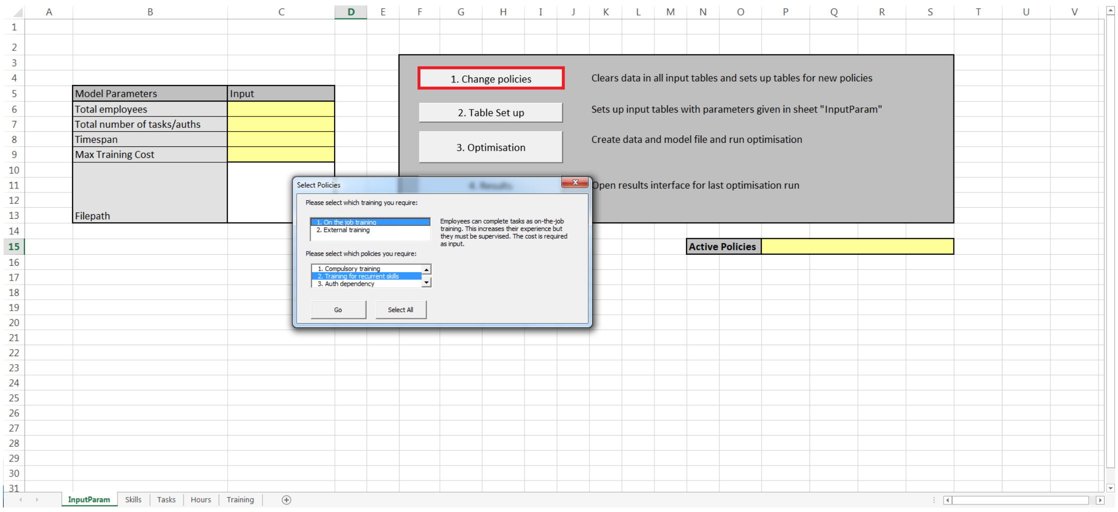

Fig. 6.3: Screenshot of Tool - Step 1

For this user guide, it is assumed that all policies have been selected. Once "Go" has ben activated, the Excel sheets are updated to include the additional information required to run the specified policies. This is seen in Figure 6.4.

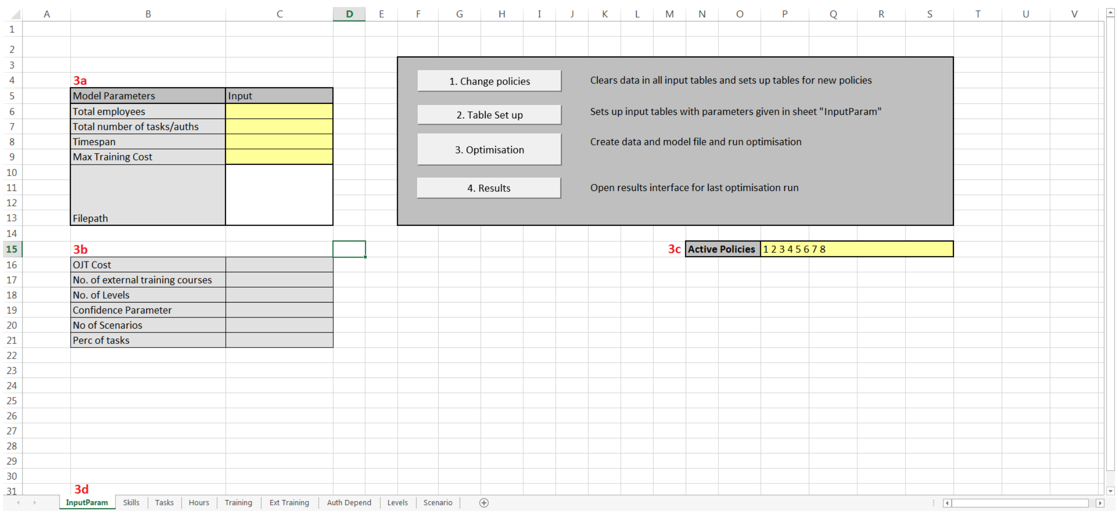

Fig. 6.4: Screenshot of Tool - Step 1 Outcome

First, label 3c shows the update to the active policies. Here, the box has been updated to include the policies selected in step 1. The numbering of the policies is as follows:

1. TEM One
2. TEM Two

3. Extension One
4. Extension Two
5. Extension Three
6. Extension Four
7. Extension Five
8. Extension Six

The TEM Zero is not specified as a number because it is an essential requirement in this model. At this step, the model parameters are now required. Hence, the user must fill in the information in Label 3a. A new box has been activated in Label 3b which contains parameters relevant to the active policies. The parameters in this list will appear dependent on the selected policies. For example, if TEM One is not selected as an active policy, the model parameter "OJT Cost" will not be present in the spreadsheet. At this stage, this information can also be populated.

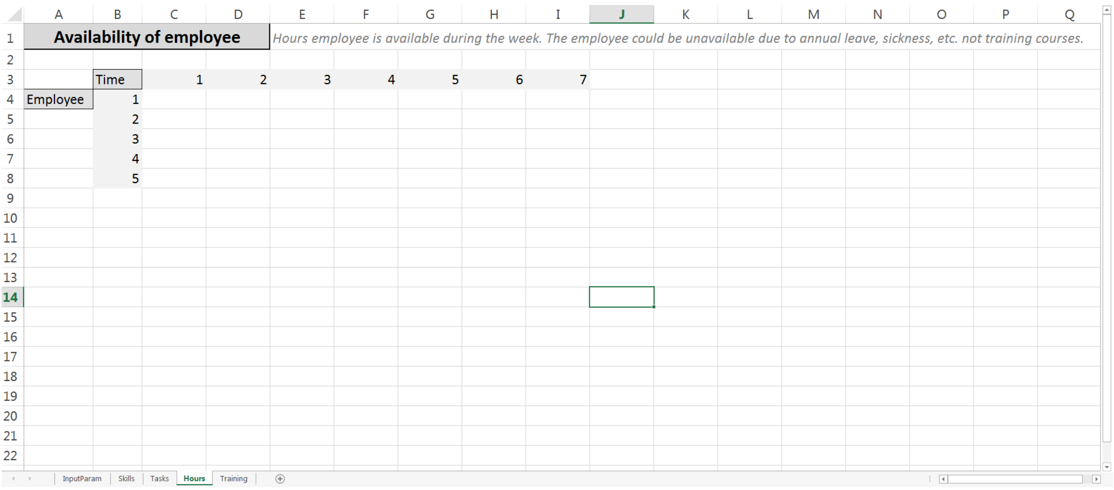

Fig. 6.5: Screenshot of Tool - Step 2

Similarly, Label 3d indicates the data sheets. These have also been updated to include data required for the active policies. The only sheets that are available are sheets that are required under the selected policies. By reselecting step 1 and changing the policies, the sheets will be reset and the model parameters will update to the new selection.

Once the "InputParam" sheet has been populated with all data, the second step can begin. By pressing the button for "2. Table Set up", the data sheets are updated given the parameters selected. For example, Figure 6.5 shows the Excel sheet called "Hours". Initially the numbers on this sheet would not be visible. Upon activating step 2, the time and employees are populated with the parameters specified in "InputParam". As such, it is now possible to fill in the available number of hours for each employee within each time unit. In this example, there are 5 employees and the duration of the model is 7 weeks.

The other sheets in this tool must also be filled in at this step. The optimisation may now be done. Selecting step "3. Optimisation" on the "InputParam" page activates the process for running the optimisation. A message box appears for "Model Size", here the size of the complete model is specified in terms of number of decision variables and constraints.

The optimisation may take some time to complete. Once the optimisation has finished, a results window will appear as shown in Figure 6.6. The results window is formed of two pages, Label 5a, "Summary" which contains an overview of the results and "Employees" which contain individual training recommendations for the employees.

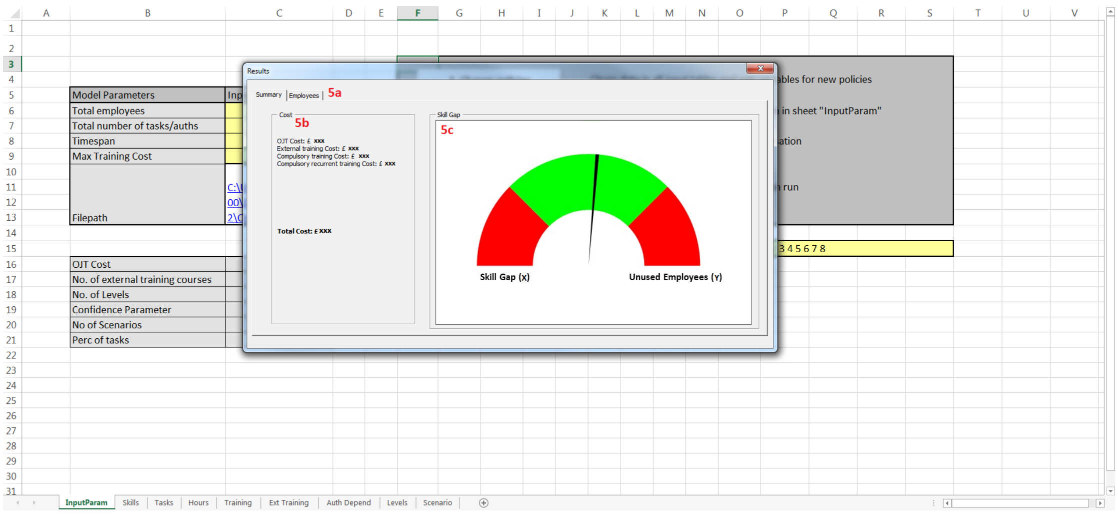

Fig. 6.6: Screenshot of Tool - Step 3/4

The summary contains a breakdown of the cost, Label 5b, for the selected policies for all employees. The total cost is also given. Label 5c contains a diagram representing the skill gap as compared the number of employees who are not required to complete a task after review. An unused employee is defined as an employee who is not assigned to any tasks over the time horizon. In this way, even if an employee is providing on-the-job training but loses their skill at review, they are still considered as a used employees.

The red areas in this gauge indicate the 25% warning interval. If the arrow is in the red area, this implies that the skill gap or number of unused employees are within 25% of the maximum possible value.

The second tab in this window is for the individual training recommendations for the employees. Figure 6.7 shows the reports for the employees. A tab appears for each employee in the model, and the skill number can be selected for further information in Label 6a. The report, Label 6b, is then generated. This report indicates if the employee will have the skill after review. If the employee has the skill then further information is provided for the training recommendations to obtain this skill.

Finally, Step 4 under "InputParam" can be used to reaccess this results page without running the model again.

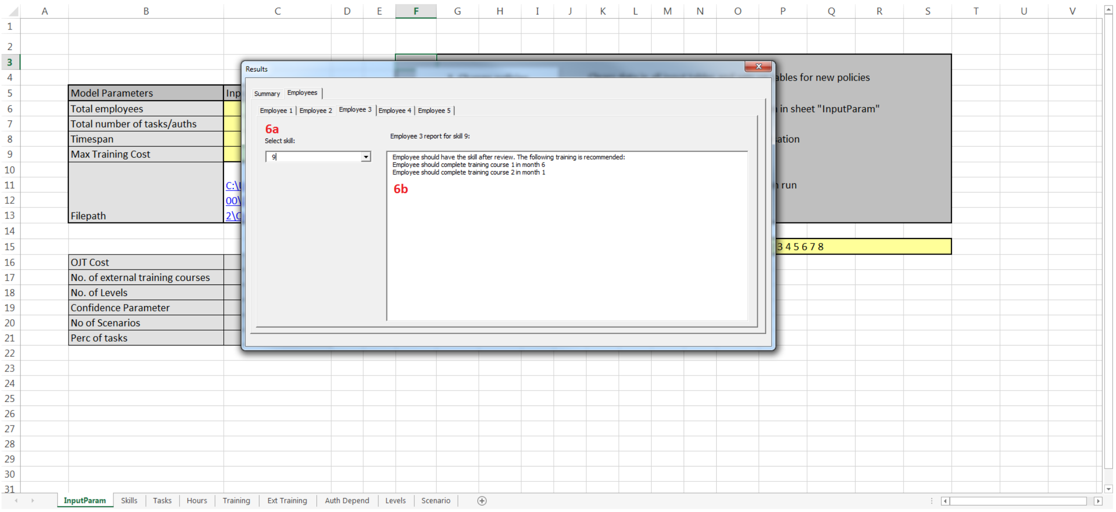

Fig. 6.7: Screenshot of Tool - Step 3/4 Part 2

6.3 *Data Analysis*

The formulation of the case study allows us to determine the relevant data to perform an analysis of the workforce requirements for Boeing. The tool has been proposed as a method of inputting this data in a real world scenario.

6.3.1 *Data Collection*

Consider the deterministic tasks at Boeing for maintaining aircraft and the employees required to complete these tasks. These tasks must be completed on each aircraft that enters the hanger for maintenance. The tasks are independent of the aircraft age and consist of checks for further repairs. For this study, only employees of Boeing are considered.

The same tasks are performed over a cycle of 20 days. The auths associated with these tasks are reviewed every six months. To ensure all skills are reviewed, consider a time period of six months such that an employee will need to be reviewed at some point within the planning horizon. In order to obtain a six month period, let each time unit equal the 20 working day cycle. Thus, six months worth of work is equal to seven time units given an average of 22 working days a month over the course of a year.

Due to the confidentiality of the data, only the fitted distributions of the data will be presented here and numbers will be rounded where necessary. Data concerning tasks was obtained as a deterministic list of all maintenance tasks requiring completion within the 20 working day cycle. Employee details were given for two sets of employees for the concerned hanger, with one year difference. Data on training was obtained from a document describing each of the auths with the associated training and the processes associated with authorisations.

First, the processing time of the tasks are required. These are modelled using an empirical distribution. Due to rounding, the processing times from the data are discrete and given in hours. The following empirical distribution is required

$$\eta_k x = \begin{cases} 1, \text{for } 0 \leq x \leq 0.36 \\ 2, \text{for } 0.36 < x \leq 0.64 \\ 3, \text{for } 0.64 < x \leq 0.75 \\ 4, \text{for } 0.75 < x \leq 0.80 \\ 5, \text{for } 0.80 < x \leq 0.81 \\ 6, \text{for } 0.81 < x \leq 0.85 \\ 8, \text{for } 0.85 < x \leq 0.88 \\ 10, \text{for } 0.88 < x \leq 0.91 \\ 16, \text{for } 0.91 < x \leq 0.93 \\ 55, \text{for } 0.93 < x \leq 0.96 \\ 60, \text{for } 0.96 < x \leq 1 \end{cases} \tag{6.17}$$

for some given task k. This can also be shown in the Figure 6.8.

As stated above, each task is required to be completed once within the 20 day period and thus once within each time unit. The availability for each employee is 7.5 hours each day, over the 20 days this equates to 150 hours. Due to no available data for working hours of employees, we must assume that all employees work all available hours and are all on full time contracts. It is also assumed that no additional tasks are required. As stated above only deterministic tasks are considered and they must be completed on every aircraft. Unscheduled repairs are not modelled.

For the training execution, data is required on the skills of the employees. The initial skill of the employees is a binary value, thus follows the Bernoulli distribution. All skills are reviewed after six months, or seven time units. As the experience is assumed to be zero at the beginning of the time horizon, and there is only one review period, it is proposed that all employees should have the opportunity to learn each skill within the time of the model. If the review time was less than seven time units then it would be impossible for an employee to gain enough experience to obtain the auth.

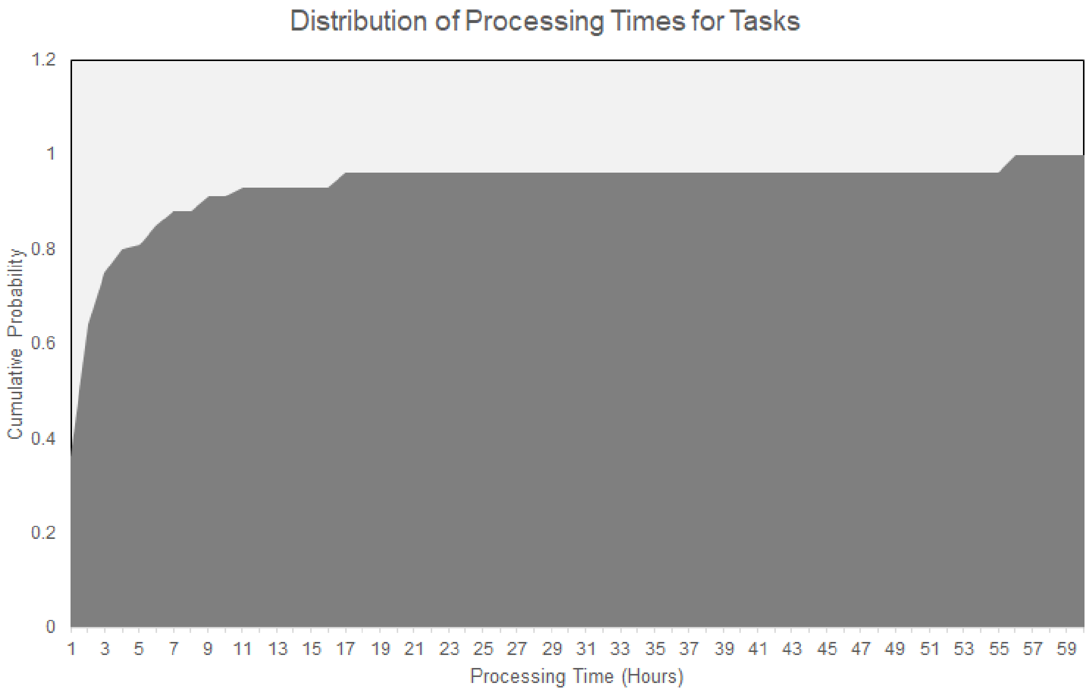

Fig. 6.8: Cumulative plot of processing time distribution

The initial skills owned by each employee is calculated such that 80% of the employees can complete task 1, a mostly essential authorisation. For other tasks, there is a 25% chance the employee will have the skill.

The experience required to gain an auth is equal to six months worth of experience where a task is completed once a month. Hence the experience required is equal to six completions of the task. If Extension Two is used, such that recurrent experience is considered then the probability that a task requires this six months of experience is equal to 90%, thus 10% of tasks do not need any experience when repeatedly obtained.

Three types of training in this model are associated with a cost; external training, on-the-job training and compulsory training. However, no data for the cost of compulsory training or on-the-job training could be obtained for this model. As a consideration to estimating OJT costs, we refer to the report made by Gay (1974) who suggests the cost of on-the-job training is dependent on the wages of the employee and the time spent training. The cost of the training is also dependent on the training completed before OJT, specifically if a training school was attended. For this case study it is assumed that all employees have attended the same training school prior to working in the hanger. No other formal training is given that may affect the OJT. The report proposes a straightforward estimate of OJT cost and applied the findings to aircraft maintenance specialists.

The true value of the external training costs cannot be enclosed in this thesis. It is assumed that on-the-job training is the cheapest training available. Hence, the on-the-job training cost is set equal to one and the other costs are scaled accordingly.

External training is split into two courses based on the different skill specialities; avionic and mechanical. Each task is defined as one or both of these values, this implies that a task can only be completed by an employee of the same skill speciality. The avionic course increases the experience for all tasks requiring an avionic technician, the course takes 40 days to complete, thus two time units of 150 hours are required to gain the required six months of experience. The mechanical course takes 30 days to complete thus, two time units of 112.5 hours are required. The course gives experience in tasks associated with the mechanical skill speciality. The adjusted cost of attending the avionic course is given as £95 per person per session and the adjusted cost of attending the mechanic course is £72, given the anonymous proposed cost of on-the-job training.

The external training can take place at any time but will need completing twice in order to gain the full experience for the course. Due to the course being divided over two time units, an employee may take half a course and complete the rest using alternative training. To force an employee to complete both halves of the training an additional constraint will be required. For use of this model it is assumed that an employee may take half the course with additional training options (OJT or repeating tasks) in order to gain the required experience. Alternatively, if it is optimal an employee may attend both halves of the training course. There is no limit on when these courses may take place, or how many employees may attend.

Compulsory training is also associated with an adjusted cost. Here, compulsory training consists of an interview for all auths to test whether sufficient training has been obtained. It is assumed that an employee sitting this interview will have obtained enough experience and pass the interview. The time to complete this interview is one hour and the adjusted cost is given as £12.

The recurrent cost is similar, however here only 15% of the tasks require a compulsory interview. The other tasks require no recurrent training. For the recurrent training, again, interviews of one hour are used. The cost of this compulsory training is slightly reduced to £10 per compulsory training session.

The total amount of money to be spent on training is £600. This value will be reconsidered within the analysis of this case study.

Extension Three uses auth dependency. In this case study, most tasks are dependent on the completion of two initial auths. In this way, 60% of the tasks require auth one to be completed first and 20% of the tasks require auth two to be gained first. Other tasks do not have any auth dependency. Due to this data item being binary, the Bernoulli distribution may be assumed. It is also assumed that a task cannot be dependent on more than one other task but can have no dependencies.

For Extension Four, skill speciality is required. As stated previously, two types of employee are available; avionic or mechanical. According to the data of the current workforce and tasks considered in this case study; the probability of a task being avionic is 20%, the probability of a task being mechanical tasks is 75% and the probability of a task being completed by either is 5%. The probability of an employee being able to complete avionic tasks is 25% and the probability of an employee being able to complete mechanical tasks is 75%. An employee can be either avionic or mechanical but cannot be both. These can be modelled from the Bernoulli distribution given the binary nature of the data, such that avionic and mechanical are considered as two binary levels of which an employee can be on that level or not.

No data could be obtained on the learning methods of the current employee set. Hence, non-linear learning in Extension Five will not be included in this case study and we may assume that all employees require the same experience to gain an auth. This assumption is reasonable due to the high turnover of employees within this field and the frequent adjustments to the tasks required with new aircraft models. In this way, it is unlikely that an employee will gain complete competence in a skill to no longer require experience to be shown.

Due to the deterministic nature of this case study, there will be no need to use Extension Six with stochastic data.

Due to the large run times associated with this model, the problem size is reduced in terms of number of employees and tasks in order to obtain solutions in a reasonable run time. An allowed 10 employees are considered for this example with 40 tasks. As stated before the number of time units required for this model is equal to 7. Due to the two course specialities; avionic and mechanical, there are two external training courses and thus two skill specialities.

6.4 Case Study Results

6.4.1 Case Study Worforce Solution

The model is run for a problem size of [10,40,7], such that there are 10 employees, 40 tasks over 7 time units, with a set of random data formulated under the conditions stated in Section 6.3.1. The solution for the generated problem can be found in Figure 6.9, as run in the decision support tool.

The skill gap is given as the calculated skill gap at the last time unit $\hat{t}$, as opposed to the skill gap definition in the objective function which is summed over all time units. Under the optimal training recommendation, four tasks could not be complete in the last time unit due to a lack of available skill in the workforce. All employees were used to generate this solution such that every employee was assigned to at least one task. As such, to achieve a skill gap of zero, the model recommends a larger workforce.

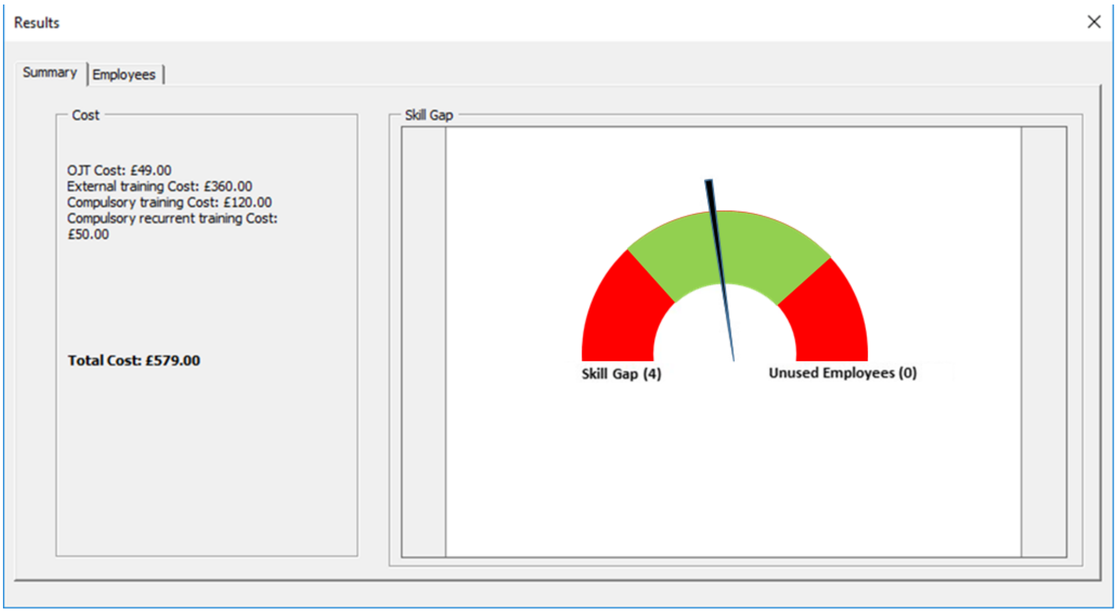

Fig. 6.9: Solution summary from case study

The low skill gap is due to the links between tasks; such as their specialities, mechanic or avionic, their dependencies and the low requirement for recurrent compulsory training. The case study suggests that some skills are essential to employees and that it is much easier, given the policies, for these employees to retain their skills.

The total amount spent on training would be £579. Of this total;

- £49 was spent on completing on-the-job training
- £360 was spent on training courses
- £120 was spent on compulsory training
- £50 was spent on recurrent compulsory training

Thus, most of the money was spent on training courses. However, the allocations to training courses are as follows:

- Employee number one completed the mechanic training course in one time unit
- Employee number two completed the mechanic training course in two time units
- Employee number eight completed the mechanic training course in two time units

From the definition of training courses in this case study, a course must be completed twice to gain enough experience to be granted the auth. Hence, the remaining experience required for employee one to gain any skill is gained through on-the-job training. As 49 instances of on-the-job training were allocated, this accounts for eight employees completing enough training to gain a skill, and employee one to complete their training for one skill.

Thus, eight employees gained skills through on-the-job training, two employees gained skills through training courses and one employee used a combination of training methods. This suggests that training courses are the most effective method of training, however due their high price and long completion times, it is not sufficient to only train employees using this method.

As every task requires compulsory training, this is expected to be high. However, only ten instances of compulsory training were completed. Given that there are a total of 40 tasks, it can be surmised that most employees completed recurrent compulsory training where they had the skill. The model prioritised assignments of tasks to employees who already had the skill as only five tasks required recurrent compulsory training. Hence the overall cost was reduced.

By multiple training options being used, it is suggested that one training type is not superior to another. In addition, by analysing the skills of the employees, it can be seen that cross training is beneficial in this scenario to reduce the skill gap. As the number of tasks exceeds the number of employees, each employee must be trained in more than one task. However, the fairness calculation is proposed to assure the workload is fairly distributed.

To further explore how much has been spent on each type of training (OJT and training courses) the budget needs to be analysed.

As the total budget available for training is £600 and the amount spent was £579, it is possible that there was not enough money available to spend on training courses. Hence, the resulting skill gap.

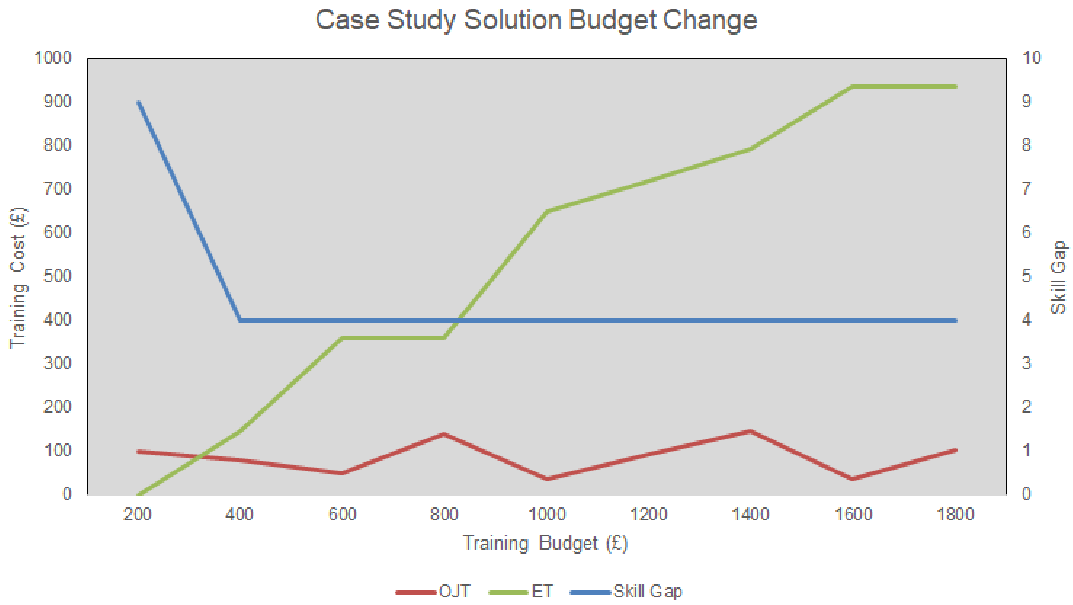

Fig. 6.10: Behaviour of solution given changes to the training budget in Boeing case study

To analyse this hypothesis, the skill gap and training costs were calculated for different training budgets. The results can be seen in Figure 6.10.

For a training budget of £400 or more, the skill gap remains at four. This indicates that this is the minimum skill gap available due to the constraints on employee availability and training operations.

The amount of money spent for on-the-job training does not vary significantly for the different available budgets whereas the training course cost increases such that the remaining budget is spent on this training. It indicates that on-the-job training is the most efficient training in this model, unlike previously hypothesised. However, the amount of OJT able to be completed is constrained. The training courses are only constrained by the cost however the skill gap does not change as this value increases.

With a training budget of £1800 the only training course used is training course two. It should be noted that course two enables an employee to learn task one and task two, both of which are essential in the auth dependency policy. Whereas training course one only enables an employee to learn task two, of these two auths. This is due to the specialities of these two tasks.

Due to the training cost not being minimised in the objective, some allocations to training are unnecessary. The amount spent on training increases but the solution does not improve. As such, the allocation to training courses in this case study indicate that employee did have available time to attend additional courses, but these were not required to improve the skill gap or fairness gap in the employee allocations. The availability constraint is currently slack for the budget specified in this experiment. Significantly increasing the budget will make the availability constraint binding.

Hence, the skill gap is constrained by OJT operations. The inefficiency in training courses could be due to the time it takes to complete a course, the levels associated with the tasks and employees or the possible allocations of the employees who can provide on-the-job training.

The fairness calculated for this case study is calculated to be 3.6 for a budget of £600. The low fairness value suggests that skills are evenly distributed between the employees. Further analysis shows that two employee only have more allocations than the other employees, though not significantly. Note that fairness is considered by the number of allocations to tasks, over the time horizon, rather than the skills owned by the employees.

The fairness may have benefited from the skill dependencies of the employees, as the distribution of skill levels was closely related to the distribution of tasks with the same level.

The run time is reasonably small due to the constraint choice; 22 seconds. Though multiple training options are available, Extensions Three and Four are beneficial to reducing the run times. This was analysed in Section 5.6.

6.4.2 Pareto Front

The solution gained for this case study assumed that Boeing prioritised the skill gap over the fairness of the employee. Here, the Pareto Front is defined for the problem such that the priority can be considered.

For the same data used in this case study, the coefficient A in the objective is set for A = 0 then increased by one until the minimum possible skill gap is achieved. The resultant skill gap and fairness is calculated for each A. The results can be seen in the Pareto Curve in Figure 6.11. The graph shows a weak Pareto efficiency at the point (280,0) as the skill gap can be decreased for the same fairness value.

As with the Pareto curves analysed in Section 5.5, a small coefficient of $A = 2$ produced the minimum values of the skill gap and fairness. However, it should be noted that the run time for these small coefficients exceeded the wall time, with the exception of $A = 0$. It is proposed that the low coefficient creates more possible solutions that minimise both the skill gap and fairness value. Hence, a larger solution space and more difficulty proving optimality.

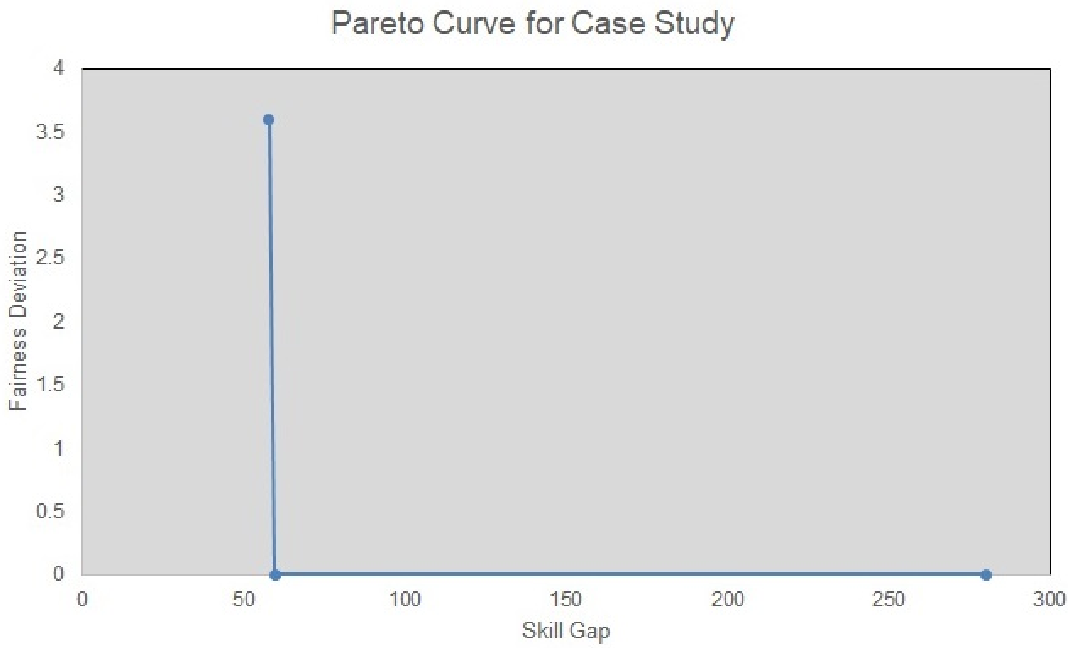

Fig. 6.11: Pareto curve for case study

Theis curve can help Boeing understand the fairness of their allocations. To be able to minimise their skill gap, the fairness value is given as 3.6, this is reasonably small. The skill gap here is 58 over the time horizon. It informs us that the total deviation from the average number of allocations is 3.6 over all employees. If Boeing would prefer to keep allocations as equal as possible, a fairness of 0 is possible with a skill gap of 60 available. Indeed, at the last time unit, the skill gap here is the same as the skill gap for a large coefficient A.

After the skill review, the fairness value is equal to zero suggesting each employee has been allocated to the same number of tasks in the last time unit. However, the allocations before the review are not equal. Allocations can only be made if the employee has the skill, before review these skills have been randomly assigned. To achieve a fairness value of zero before review, employees would leave two tasks unfulfilled. Alternatively, the fairness value before the review period may be increased to ensure the skill gap is reduced.

Due to the allocations made during the model, the fairness value after the skill review is equal to zero. This suggests that the model is fairly allocating training and tasks to ensure each employee can complete similar workloads with their new skill sets. In this way, Boeing may consider allowing for a lower fairness value and higher skill gap to ensure as many tasks are done as possible before the skill gap, with the understanding that after the skill review period all employees will have the same number of allocations.

6.5 Summary

The model built in this thesis has been applied to a case study at Boeing. The employees in question are technicians on a maintenance line where skills have to be authorised before they are applied on the aircraft. Ten employees were assigned to forty tasks in each time unit over a seven time unit horizon, or six months. The policies proposed in this thesis that are relevant to this study are:

- TEM 0 - Employees gain experience repeating tasks
- TEM 1 - Employees gain experience on-the-job
- TEM 2 - Employees gain experience through training courses
- Ext 1 - Employees need to complete compulsory training
- Ext 2 - Skill requirements are dependent on initial or recurrent training
- Ext 3 - Auths have dependencies
- Ext 4 - Employees and tasks are defined by skill specialities

To input data and run the model, a decision support tool was designed in Microsoft Excel. This allowed a familiar interface to be used. The model was then solved using IBM Cplex. A thorough userguide was provided to show how the tool was designed and can be used in a business environment.

Appropriate assumptions were made and data was randomly generated based on real world data sets. A randomly generated instance of the model was applied to the model for analysis. For the instance generated, it was possible to achieve a skill gap of four tasks at a fairness deviation of 3.6 tasks. To achieve this result, £579 was spent on training.

It could be seen that on-the-job training was the most efficient and effective training method in this model. The inefficiency in training courses could be due to the time it takes to complete the course, the levels associated with the employees and tasks or the available allocations of the employees who can provide the on-the-job training.

The impact of the training budget on the solution was discussed and a Pareto curve produced for various coefficients. It was determined that increasing the training budget from £400 did not improve the solution. Thus, this is the minimum training budget required to minimise the skill gap.

The Pareto curve indicated that fairness in the employee allocations is reasonably low. After the skill review, the fairness in allocations is equal to zero, indicating all employees are assigned to the same number of tasks. Though the fairness can be improved before the skill review, this will impact the number of tasks that are completed in this time.

For a small coefficient, the run time in this model was large. The model has problems solving optimality with this data and a small value of A. Though it was shown previously that there is no correlation between the coefficient and the run time over a random selection of models. It is still important to consider methods to avoid these instances where the run time is large.

In the next chapter, a heuristic approach is discussed to improve the run time of the model. This will be applied to random data to analyse the quality of the solution but will also be applied to this case study.

7. A HEURISTIC APPROACH

As shown in Section 5.6, the run times in this model can be very large due to problems proving optimality. As deduced from the analysis, it is not known which combination of data will cause an increase in run times before the model is solved. Although it has been shown that a wall time will allow a solution within a small percentage of the optimal solution, it is worth considering a heuristic approach to solving this model.

In this chapter, a heuristic method will be proposed to solve the model in a small time. As the model is adaptable to the customers requirements, a heuristic must be able to solve the model under any set of TEMs or extensions.

As shown in Section 5.6.2, the index that causes the greatest increase in number of decision variables and constraints is the number of employees. Similarly, in Section 5.6.1, increasing the number of employees has a larger significance on the run time than increasing the number of tasks or number of time units. This will be considered in the creation of the heuristic.

The heuristic will be used to solve each TEM and extension in turn to determine the quality of the solution. It will then be applied to different size problems to investigate the accuracy.

7.1 A Linear Programming Based Heuristic

Due to the long run times in the model for large problem instances, a heuristic approach should be considered. As discussed, the index that creates the largest increase in problem size and run time is the number of employees.

Hence, the design for this heuristic is a decomposition approach where each employee is considered as a subproblem and solved in turn. The employees can be considered in any order. Thus, each iteration is a linear programme where only the allocations to tasks and training of one employee are considered, all other allocations are considered fixed. After each iteration, the solution for the selected employee is fixed and will be used in subsequent iterations.

In order to ensure all employees are allocated to tasks the heuristic will iteratively find a solution for each employee is turn until all employees have been assigned. Once this has been completed, random shakes are introduced. An employee e is selected at random and the corresponding subproblem is solved to obtain new allocations for e, replacing the current solution.

To avoid the first employee being assigned to all tasks, a non-optimal feasible solution is accepted at each iteration by restricting the maximum number of nodes in the branch and bound trees. Thus, in the first iteration you accept the first feasible solution from the branch and bound tree. As the number of iterations increase, the maximum size of the branch and bound tree increases respectively.

Though similar to dynamic programming methods, this heuristic does not use a nesting approach to build the subproblem into the main problem. The subproblems are not solved recursively, however they explore the local neighbourhood by searching for a solution to an employee e using the solution of employee $e - 1$. Hence, it is described as a local search heuristic. Due to the issue relating to the size of the solution space, a dynamic programming approach would not be suitable here. There can be many solutions to the same problem instance.

The following pseudo code is proposed.

Start

$i := 1$; ▷ Initialisation

$\overline{e} := 1$;

while $\overline{e} \leq$ Number of employees **do** ▷ Initial solution

Solve decomposition MILP for employee $\overline{e}$, allow a maximum of i nodes on tree;

$\overline{e} := \overline{e}+1$;

$i := i+1$;

Input data of employees updated to include solutions for $\overline{e}$;

end while

while $i \leq$ Maximum number of iterations **do** ▷ Improve solution

Randomly select employee $\overline{e}$;

Set allocations of employee $\overline{e}$ to be zero such that the employee does not have any allocations;

Solve decomposition MILP for employee $\overline{e}$, allow a maximum of i nodes on tree;

$i := i+1$;

Input data of employees updated to include solutions of $\overline{e}$;

end while

if Ext Six is Selected **then** ▷ Ext Six feasibility
Check chance constraint;
if Chance constraint is violated **then**
No solution is available with the current workforce;
else
Solution is found, sum costs from each employee output to obtain total cost;
end if
else
Solution is found, sum costs from each employee output to obtain total cost;
end if
End

The decomposition MILP is designed using the original model. The variables and constraints for the base model and each of the policies are presented in turn.

7.1.1 Heuristic Model Formulation

The notation used in the heuristic model is similar to the notation used in Chapter 4. New variables are introduced to store the solutions for the employee $\overline{e}$ in the associated iteration. The fixed value of the allocations w_{ekt} is given by $\overline{w}_{ekt} \in \mathbb{Z}^+$, in a similar way the allocations to OJT, $\overline{p}_{ekt} \in \mathbb{Z}^+$ and external training, $\overline{q}_{ekt} \in \{0, 1\}$ are defined, where $e \in E$. The fixed values are initiated to be equal to zero such that no allocations or training is assigned.

The decision variables and data used in Chapter 4 are updated to remove the index e as only one employee is considered per iteration. The data used in the model must be updated to include only the relevant data for the current iteration. The fixed solutions from previous iterations are included in the model with solutions for all employees.

The formulation provided here is for all TEMs and additional constraints. As before, these components may be added or removed as required by the user allowing the heuristic to work with the flexibility of the original tool. The objective function for the employee $\overline{e}$ is stated as:

$$A \sum_{k \in K} \sum_{t \in T} (r_k - \sum_{e \in E; e \neq \overline{e}} \overline{w}_{ekt} - w_{kt}) + (\psi^+ + \psi^-) + B \sum_{i=1}^{m} \delta_{\xi_i} \tag{7.1}$$

Note that all decision variables have one less dimension for the employees. Thus the decision variable w_{ekt}, used previously to count the number of times employee e is assigned to task k at time t is substituted with the decision variable w_{kt}, equal to the number of times the selected employee for the iteration, $\overline{e}$, is assigned to task k at time t.

The calculation for the skill gap is adjusted from the original model to incorporate $\overline{w}_{ekt}$, the fixed allocations for all employee. This is stored as an input, not a decision variable. Hence, the skill gap is calculated as the required amount for the task minus the total allocations from the current employee and the allocations from employees calculated in previous iterations. The fairness calculations remain the same as the previous model, though here are used to calculate the fairness value of one employee rather than all.

The stochastic element is specified in decision variable δ_{ξ_i}. This variable counts the number of times the chance constraint is violated. By including it in the objective function here, the motivation to satisfy the chance constraint is present in the model as the chance constraints themselves will not be. The constant $B > 0$ is used to weight the stochastic element in the multi objective. To obtain a feasible solution to the problem it is recommended that $B >> A$.

As before, there are constraints on the allocations of the employees for each iteration.

$$w_{kt} - r_k s_{kz} \leq 0 \qquad \forall k \in K, z \in \overline{Z}, t \in Z_{kz} \quad (7.2)$$

$$\sum_{k \in K} \gamma_{k\xi_i} \eta_k (w_{kt} + p_{kt}) + \sum_{n \in N} \rho_n v_{nt} + $$
$$+ \sum_{k \in K} \sum_{\sigma=1}^{2} \mu_{k\sigma} q_{kt\sigma} \leq a_t \qquad \forall t \in T, i \in 1..m \quad (7.3)$$

$$\sum_{e \in \overline{E}} \overline{w}_{ekt} + w_{kt} \leq r_k \qquad \forall k \in K, z \in \overline{Z}, t \in Z_{kz} \quad (7.4)$$

The index e is removed from the data input and decision variables including the review set Z_{kz} and the decision variable for skill $s_{kt} \in \{0, 1\}$, which is equal to one if the selected employee has auth k at time t. The set $\overline{E}$ contains all employees with the exception of employee $\overline{e}$. Equation (7.4) states that the total allocation to a task, including allocations from previous iterations, must be less than the requirements for the task. Otherwise, constraints remain the same as the previous section with the adjustments for index e.

Similarly, fairness is calculated as before for each employee. The allocations include the allocations of the currently selected employee and allocations from previous iterations.

$$\sum_{k\in K}\sum_{t\in T}(\frac{1}{\hat{e}-1}\sum_{e\in\overline{E}}\overline{w}_{ekt}-w_{kt})\leq\psi^{+} \tag{7.5}$$

$$\sum_{k\in K}\sum_{t\in T}(w_{kt}-\frac{1}{\hat{e}-1}\sum_{e\in\overline{E}}\overline{w}_{ekt})\leq\psi^{-} \tag{7.6}$$

The constraint on training budget is given as follows:

$$\begin{aligned}\sum_{k\in K}\sum_{t\in T}c_1(p_{kt}+\sum_{e\in\overline{E}}\overline{p}_{ekt})+\sum_{n\in N}\sum_{t\in T}c_{2n}(v_{nt}+\sum_{e\in\overline{E}}\overline{v}_{ent})&\\+\sum_{k\in K}\sum_{t\in T}\sum_{\sigma=1}^{2}c_{3k\sigma}(qkt\sigma+\sum_{e\in\overline{E}}\overline{q}_{ekt\sigma})\leq b&\end{aligned} \tag{7.7}$$

where the cost of OJT, external training, compulsory training and recurrent compulsory training for all employees is considered.

The skill development constraints, including Extension Two, are given as follows:

$$s_{k1}=\Theta_k \qquad \forall k\in K \tag{7.8}$$

$$\begin{aligned}\sum_{x\in X}g_{kx1}y_{kzx}-M_k(1-s_{kz})\leq\sum_{t\in Z_{k(z-1)}}(\gamma_{k\xi_i}(w_{kt}+p_{kt})&\\+\sum_{n\in N}\chi_{nk}v_{nt})+M_k s_{k(z-1)}&\qquad\forall k\in K,z\in\overline{Z}\end{aligned} \tag{7.9}$$

$$\begin{aligned}\sum_{x\in X}g_{kx2}y_{k(p-1)x}-M_k(1-s_{kp})\leq\sum_{t\in Z_{k(z-1)}}(\gamma_{k\xi_i}(w_{kt}+p_{kt})&\\+\sum_{n\in N}\chi_{nk}v_{nt})+M_k(1-s_{k(z-1)})&\qquad\forall k\in K,z\in\overline{Z}.\end{aligned} \tag{7.10}$$

Here, the constraints remain the same as those previously given but with the removal of the e index. Equation (7.8) initialises the skill of the current employee and (7.9) calculates the skill from the experience.

The constraints for each TEM must also be adjusted.

As on-the-job training is dependent on the allocations of other employees, the allocations to OJT, $\overline{p}_{ekt}$, are required in the constraint. Two constraints are proposed:

$$p_{kt}-\sum_{e\in\overline{E}}\overline{w}_{ekt}\leq 0 \qquad \forall k\in K,t\in T \tag{7.11}$$

$$w_{kt} + \sum_{e \in \overline{E}} \overline{w}_{ekt} \geq \overline{p}_{ekt} \qquad \forall e \in E, k \in K, t \in T. \tag{7.12}$$

The constraints (7.11) and (7.12) ensure that employees cannot complete OJT unless another employee is also assigned to the task. Two conditions are required to satisfy this. First consider the allocations of all employees, $\overline{w}_{ekt}$, such that the currently selected employee can only be assigned to a task if another employee is available. This is implemented in equation (7.11). Second, to ensure all other OJT allocations remain feasible, the OJT allocations from all employees must be checked. If any employee has been allocated to OJT training, $\overline{q}_{ekt}$, then there must be an employee allocated to that task. This could be the selected employee w_{kt} or any other employee $\overline{w}_{ekt}$. This is implemented in equation (7.12).

External training courses require a constraint on the capacity of the course, with the index e removed. The total number of employees assigned to external training should not exceed the capacity of the course.

$$v_{nt} + \sum_{e \in \overline{E}} \overline{v}_{ent} \leq u_{nt} \qquad \forall n \in N, t \in T \tag{7.13}$$

In addition to constraints on the TEMs, constraints on each extension are also required for the heuristic.

Extension One ensures that compulsory training courses are completed for relevant skills before they can be obtained. Simultaneously, Extension Two for recurrent compulsory training is considered in the following pair of constraints.

$$\pi_{k1} s_{kz} - \sum_{t \in Z_{k(z-1)}} q_{kt1} + s_{kp} \leq 0 \qquad \forall k \in K, z \in \overline{Z} \tag{7.14}$$

$$\pi_{k2} s_{kz} - \sum_{t \in Z_{k(z-1)}} q_{kt2} + (1 - s_{k(z-1)}) \leq 0 \qquad \forall k \in K, z \in \overline{Z} \tag{7.15}$$

where $\sigma = 1$ defines an initial skill and $\sigma = 2$ defines a recurrent skill.

As allocations to compulsory training have no effect on the allocations of other employees, the only change to the constraint required in the heuristic is the removal of the e index. Similarly with the following constraints.

If the tasks form a network such that tasks depend on authorisations being achieved for other tasks then Extension Three is applied. The following constraint is used to ensure this is implemented:

$$s_{kz} - s_{i\tilde{z}} \leq 1 - d_{ki} \qquad \forall k \in K, z \in \overline{Z}, i \in K, \tilde{z} \in Z_{kz} \cap Z_{i\tilde{z}} \qquad (7.16)$$

Extension Four uses skill speciality to constrain some tasks to require a certain type of employee. The following constraint is added:

$$w_{kt} - r_k \sum_{l \in L} \lambda_l \phi_{lk} \leq 0 \qquad \forall k \in K, t \in T. \quad (7.17)$$

where λ_l defines the skill level of the employee used in the iteration.

Learning is considered to increase at a non-linear rate in extension five. The following constraints are adapted from the basic TEM.

$$\sum_{x \in X} x y_{kzx} = \sum_{t=1}^{\max(Z_{k(z-1)})} w_{kt} \qquad \forall k \in K, z \in \overline{Z} \quad (7.18)$$

$$\sum_{x \in X} y_{kzx} = 1 \qquad \forall k \in K, z \in \overline{Z} \quad (7.19)$$

The final extension allows stochastic tasks, where the probability of a task occurring in each time unit is unknown. The following constraints are adapted from the basic model.

$$\sum_{k \in K} \gamma_{k\xi_i} r_k - \sum_{k \in K} \gamma_{k\xi_i} (w_{k\hat{t}} + \sum_{e \in \overline{E}} \overline{w}_{ek\hat{t}}) \leq \beta \sum_{k \in K} \gamma_{k\xi_i} r_k + \boldsymbol{M} \delta_{\xi_i} \qquad \forall i \in 1, .., m \quad (7.20)$$

Once the heuristic has run for the specified number of iterations, the model calculates the chance constraints from Section 4.3.6.

$$\sum_{i=1}^{m} \delta_{\xi_i} \zeta_{\xi_i} \leq 1 - \alpha \qquad \forall i \in 1, .., m \quad (7.21)$$

The heuristic calculates the above equation using the solution of δ_{ξ_i} given in the last iteration. If the constraint is satisfied then the stochastic constraints are satisfied and the solution is feasible. Else there is no feasible solution to the model.

7.1.2 Discussion

Due to the iterative approach in this heuristic, the first employee assigned to the tasks will be given the largest workload in order to minimise the skill gap. Hence, the order of the employees will have an affect on the quality of the solution.

It is suggested that employees can be ordered based on the skills they initially have, the level of the employee, the available time or considering the specific tasks that the skills require. For instance, if an employee can initially complete tasks with high requirements, it may be beneficial to assign that employee first in order to minimise the skill gap. The sorting method chosen will need to fairly allocate the resources while minimising the skill gap. Hence, producing a good estimate of the exact model.

Here, the model was solved for each employee e. However, there were alternative indexes to split the model by; k and t. Though the results showed that splitting on employees was likely to be the most beneficial to decrease solution run times, there are additional reasons why e was chosen. These will be discussed here.

First splitting by tasks k. In this model a solution is found for each task in turn, assigning any applicable employee. In order to use this heuristic the allocations and availability would need to be stored. Though choosing k means a smaller model as the number of tasks is larger than the number of employees, it has been shown that the number of tasks has less influence on the run time.

Another consideration if each subproblem assigned each task k is the order of the tasks, particularly if applying Extension Three. Under the different constraints, the order of the tasks will have an impact on the results. To determine the order of the tasks, the following methods may be used:

1. If Extension Three is selected; order tasks using Kahn's Algorithm (Kahn, 1962).
2. If Extension Three is selected; order tasks by total number of dependent tasks.
3. If Extension Six is selected; order tasks by largest average probability.
4. Tasks can be ordered randomly.
5. Tasks can be ordered by requirements.
6. Tasks can be ordered by processing time.

By applying Kahn's algorithm, tasks are ordered by their dependencies. Thus if the arc $(i, j) = 1$, such that task j is dependent on task i, than task i must be ordered before task j. Where the order of two tasks does not follow this rule, they are ordered numerically.

The total number of dependent tasks can be calculated as the number of tasks that require task i. The following logic is applied, if task k requires task j and task j requires task i, then task k requires task i. Hence, the tasks are ordered in decreasing order of number of dependent tasks. It may be assumed that this order satisfies the logic used in Kahn's algorithm.

The largest average probability is used for Extension Six. In this way, tasks that have a higher probability of occurring within the time frame are given priority over other tasks. Hence, resources are reserved for these tasks as they will be allocated first.

Consider splitting by time t. This is much more difficult due to the use of the set Z_{ekz}. In addition, most constraints sum over t. The number of employees is summed over the least number of times of the three indices. If the index is used in a sum then more variables are required for the model to store the output data.

Though choosing e results in a larger model in each iteration than choosing k, less iterations are required to get a base solution and due to the allowance of number of nodes, the run time is very small.

As discussed previously, the number of employees has the greatest impact on the size of the model and the resultant run time. Thus, if the number of employees increased, the need for a heuristic approach increases. Hence, splitting on the employees seems the more reasonable choice.

In each random shake, the allocations of one employee are recalculated. It was considered reallocating the solutions to more than one employee to increase the search area. To test this theory, the heuristic was also built for two employees having their solutions reset in each random shake. The run times here were increased, as increasing the number of employees in the model increases the run time, and the solutions were not significantly improved than in the one employee model. Hence it was decided to recalculate the solutions of one employee only in the heuristic model at each shake.

7.1.3 Adaptation to Decision Support Tool

The heuristic may be added to the Excel Decision Support Tool.

Selecting step "3. Optimisation" on the "InputParam" page activates the process for running the optimisation. A message box appears for "Model Size". Here, the size of the complete model is specified in terms of number of decision variables and constraints. As such, the user is now offered a heuristic approach to reduce the run size if the model size appears too large. If "Yes" is selected, the user may specify the number of iterations to run and then the heuristic optimisation is run.

The default value for number of iterations is set to the number of employees, any number larger than this will include shakes within the heuristic, a smaller number will not allocate all employees to tasks and training. If "No" is selected, the model is run as usual. The output will remain in the same format, independent of whether the heuristic was used or not.

7.2 Heuristic Analysis

The aim of the heuristic is to produce a reasonable estimate of the solution in a shorter run time than solving the model with exact methods. In this section, the results of the heuristic are analysed.

7.2.1 Iterations

The first analysis required is the iterations. As stated previously, the minimum number of iterations required for the model is $\hat{e}$, the total number of employees. As such, each employee may receive allocations to tasks and training. Any additional iterations are used to shake the solution in order to improve the objective. As the allowed size of the branch and bound tree is increased at each iteration, it is hypothesised that the solution will improve in each iteration.

To determine how much the solution improves, the model is run for different iterations and the skill gap analysed. All TEMs and Extensions are applied to this model. A model size of [10,15,7] is analysed using the same random data as generated for Section 5.6.1 for 100 sets of data. The change in skill gap is calculated as the skill gap produced after 10 iterations, when all employees are assigned, minus the skill gap at the specified number of iterations in Figure 7.1.

The resultant plot shows a histogram with very small interquartile ranges. The crosses indicate the average and the dots are the outlying values. The number of iterations does not change the solution drastically, on average the solution chosen for ten employees remains the same for the iterations. As the heuristic will not require running for many iterations, it suggests that it will have a reasonably short run time. For each of these runs, the average run time was 2 seconds per iteration.

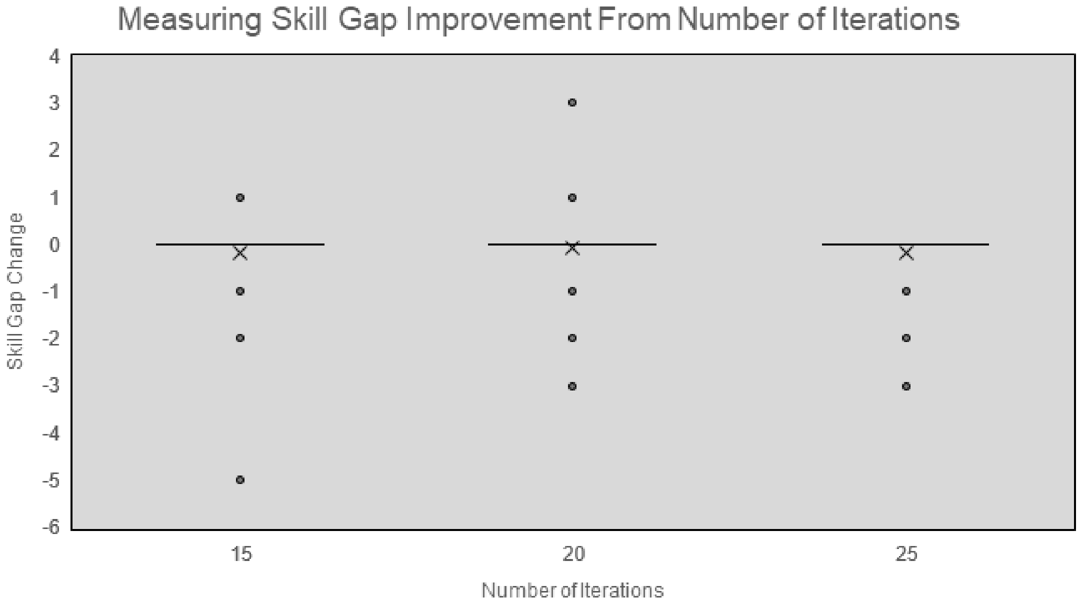

Fig. 7.1: Heuristic solution with varying number of iterations

In some instances, the skill gap is smaller than the skill gap obtained from the exact solution. In these cases the fairness calculated is larger than the exact model. This suggests the value of A in the heuristic objective should be larger.

Alternatively, by solving for each employee in turn the first employee is assigned the greatest number of tasks to minimise the objective. Hence, the last employee is assigned the least tasks, resulting in a large fairness value. Though the random shakes and limit to the tree are designed to minimise the impact of this, the results show that there are still instances where the skill gap is smaller but the fairness is larger than the exact model.

Finally, the limit on the branching tree results in solutions being accepted that may not be optimal, hence a smaller skill gap but a much larger value of fairness. The number of these occurrences decreases with the number of iterations in the model.

7.2.2 Problem Size

The run time for the model has been shown to increase as the problem size increases. Previously, this was explored for each index e, k and t. Here, the experiment is repeated but the heuristic output is shown. Due to the reasonable run time, 25 iterations were used for each run. The data used in Section 5.6.1 is used again for this analysis. As before, all TEMs and extensions are applied, with the exception of Extension Six. In this way, infeasible solutions are not possible.

Here, the run time of the heuristic is compared to the run time of the exact solution. Figure 7.2 shows the change in run time for different number of employees. The number of tasks and time units remains the same.

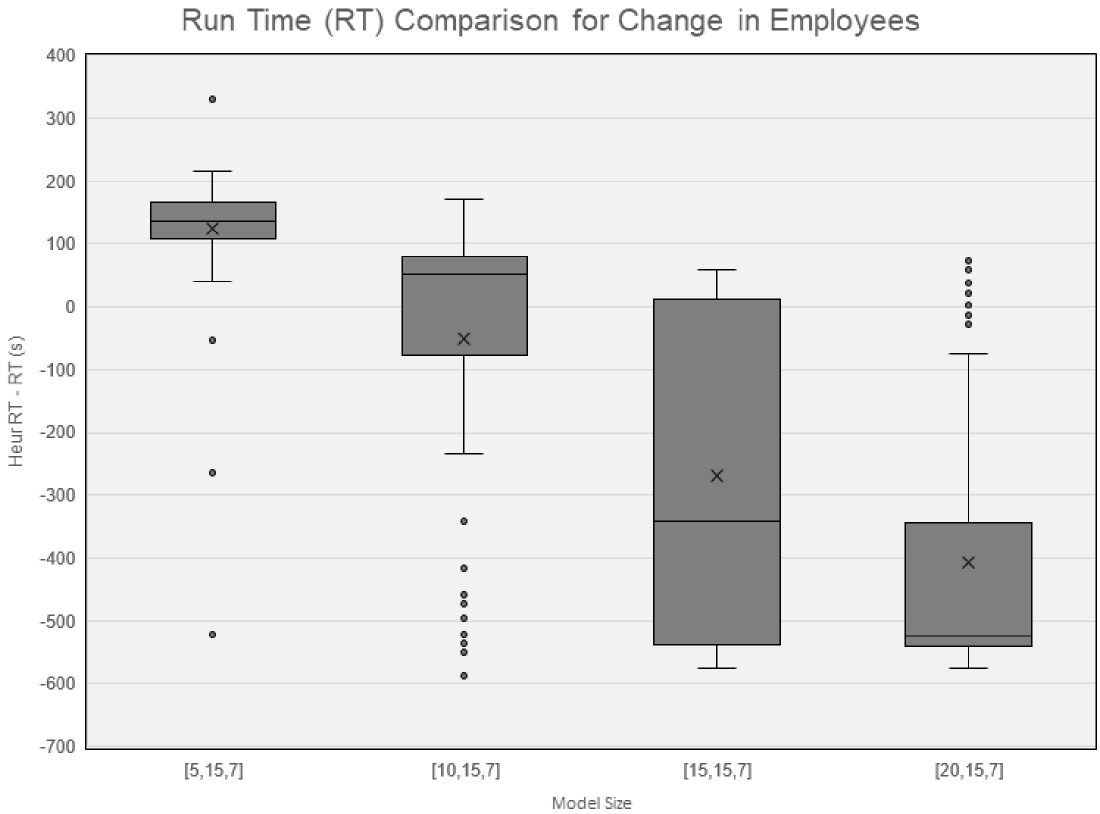

Fig. 7.2: Difference in run time by heuristic given changes to number of employees

The run times for the exact model were shown to have an increasing rate of change with the number of employees. Here, the improvement in run time is shown to decrease proportionally. The solution quality will be discussed later in this chapter.

For a small problem size of [5,15,7], the run times of the exact model are smaller than the run time of the heuristic model in most cases. In these cases, the exact model runs in a few seconds. However, in cases where the run times are very large or exceed the wall time of 600 seconds, the heuristic is shown to be beneficial.

As the problem size increases, the improvement in the run time as a result of the heuristic is much larger. For a model size of [20,15,7] the average run time of the heuristic was 67 seconds, compared with 400 seconds for the exact model.

It should be noted that the number of iterations did not change. Hence, in the smallest model size, there were 20 shakes however in the largest model size only 5 were performed.

Next the task index is explored. Figure 7.3 shows the change in run time for different number of tasks.

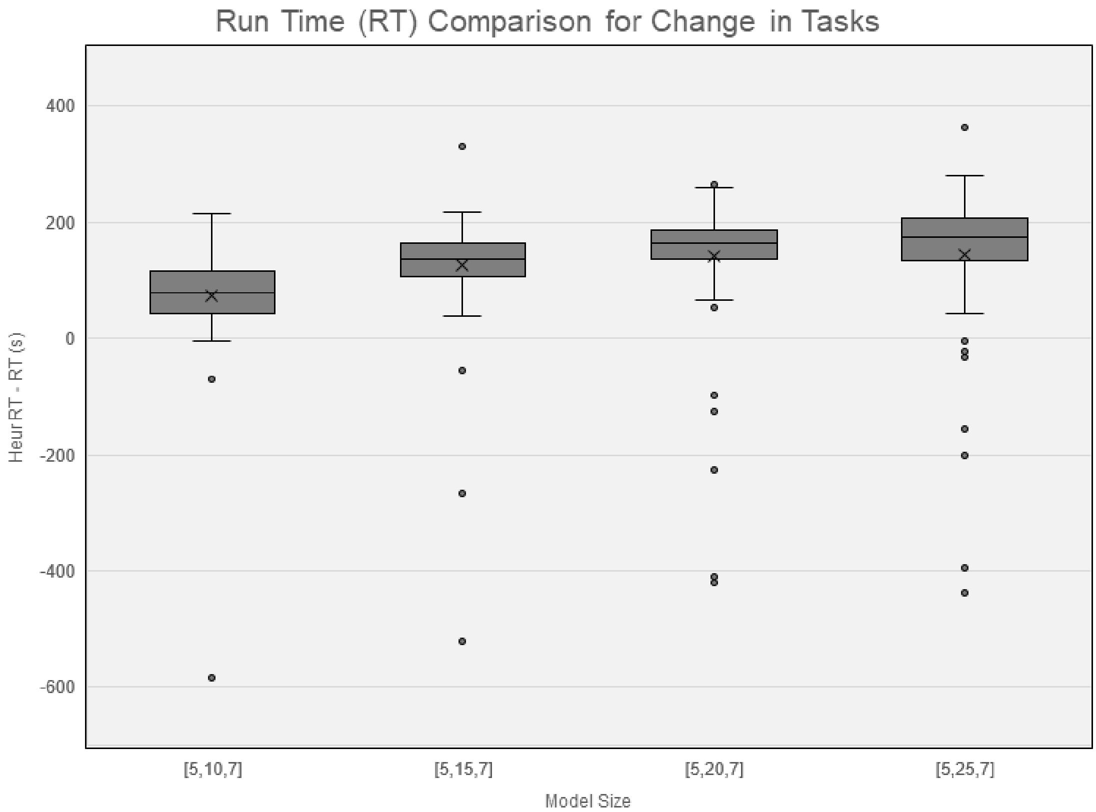

Fig. 7.3: Difference in run time by heuristic given changes to number of tasks

Previously, it was shown that the run time of the exact model did not change significantly with the number of tasks. However, the number of instances where the problem could not be solved increases as the number of tasks increases.

The heuristic shows little improvement of the run time in these cases. As the number of tasks increases, the poorer the heuristic performs in terms of run time.

However, here the number of employees is small, and it has been shown that the heuristic is less effective when there are few employees. However, the standard deviation in run times of the heuristic method is 41 for the [5,25,7] model. This is small compared to the standard deviation of 118 for the exact model.

Finally, the number of time units is increased in Figure 7.4.

Here, the heuristic can be seen as an improvement to the run time of the model as the number of time units increase. Similarly to the number of tasks, however, the heuristic does not perform as well as the exact model in most cases.

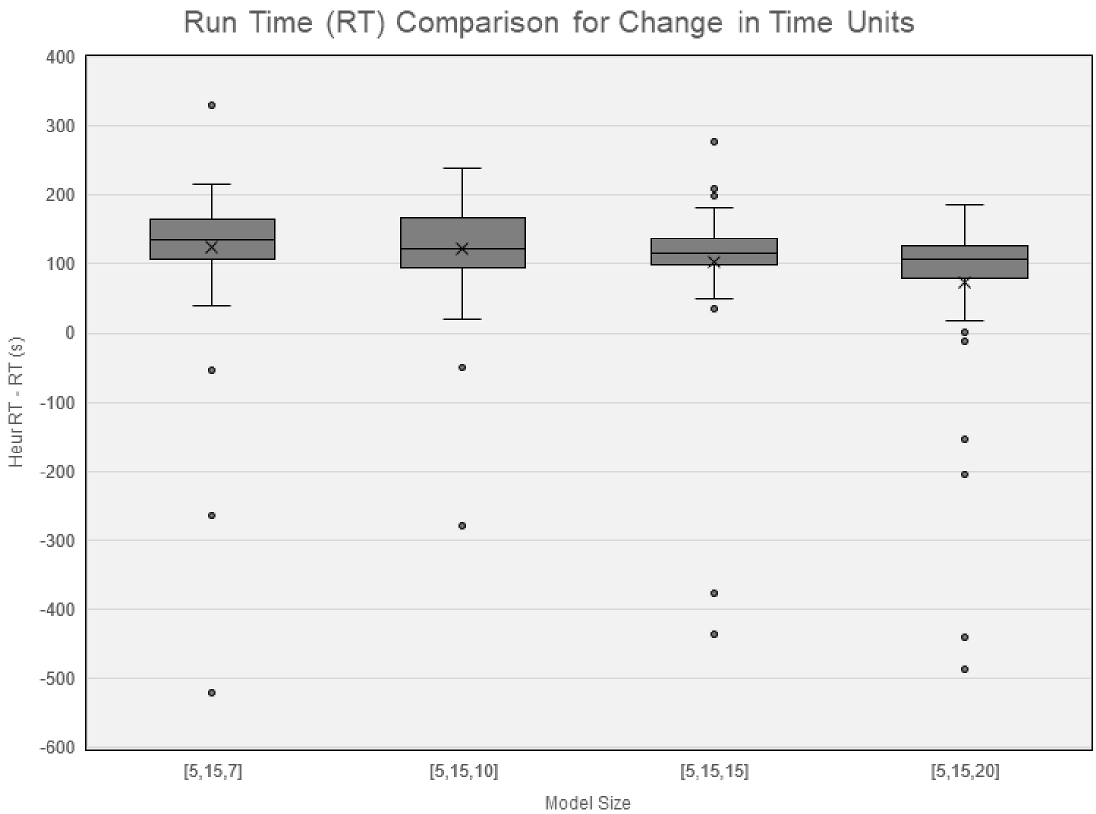

Fig. 7.4: Difference in run time by heuristic given changes to number of time units

Again, it should be noted that the number of employees has a much greater impact on the effectiveness of the heuristic.

It can be seen that the heuristic improves the run time when the number of employee and time units are large. Though the run time is less effective as the number of tasks increases, it can be seen that in all cases the standard deviation of the run times for the heuristic is much lower. These means the heuristic has a more predictable run time.

In addition, there was no instance of the heuristic run in which the run time exceeded the wall time. Thus, the heuristic approach is an effective method to remove the problems with solving optimality.

7.2.3 Heuristic Training Budget

The training budget has been shown to affect the run time and the solution to the training model. Hence, in this section the training budget analysis will be repeated with the heuristic method used.

The data used will be the same data as used in Section 5.3. As such, the model size is given as [10,20,7] and Extension Six will not be used. From the previous section, it was shown that, on average, a model with 10 employees has an improved run time when the heuristic is used. However, this was not always the case.

Here, the training budget is changed to determine the impact on the solution quality and the run time.

Figure 7.5 shows the improvement in run time by using the heuristic method.

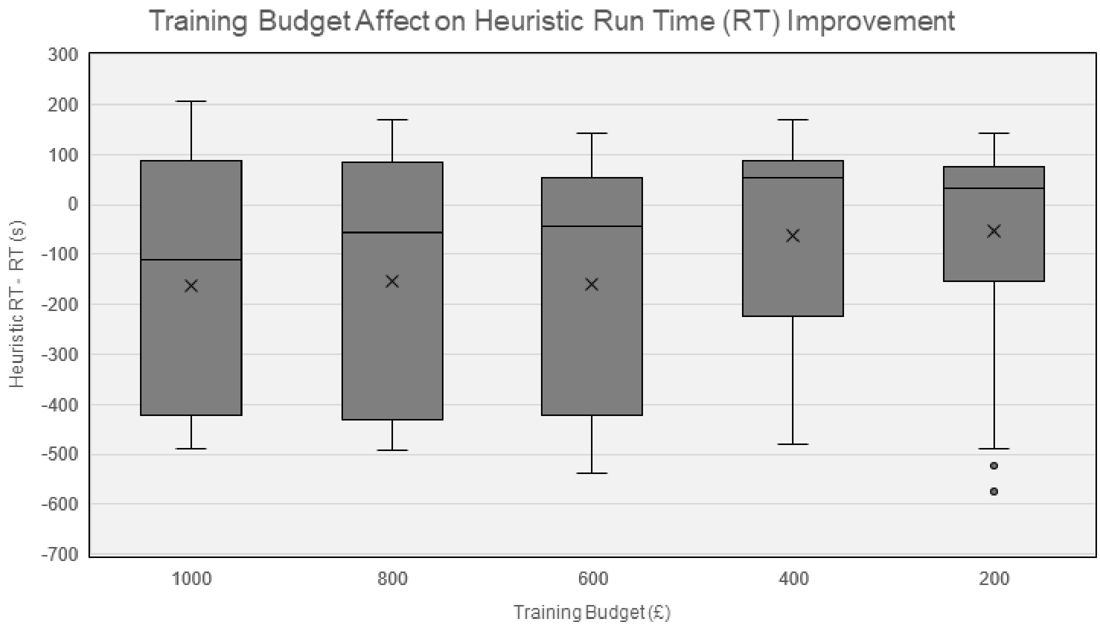

Fig. 7.5: Impact of the training budget on the heuristic run time

As specified when analysing the training budget, given the data provided, a training budget of £600 is sufficient to minimise the skill gap. At this budget, it can be seen that the heuristic is most effective.

As the training budget decreases, the problem space decreases so the possibility of getting problems that are difficult to solve also decreases. As such, the heuristic is less effective for these values. The heuristic works for all budgets to reduce the run time when the run time exceeds 600 seconds.

By analysing the standard deviation, it can be shown that the heuristic produces much more consistent run times than the exact method.

The quality of the solution may also be observed. Figure 7.6 shows the increase in skill gap through the use of the heuristic model.

There does not appear to be any significant change in the quality of the heuristic solution through changes to the training budget. The best solutions can be found when the training budget is large and thus the problem space in larger. As the run time is longer for this scenario, it has been seen that the improvement to the run time is also greater in these cases. Hence, the heuristic is effective and efficient when the training budget is large.

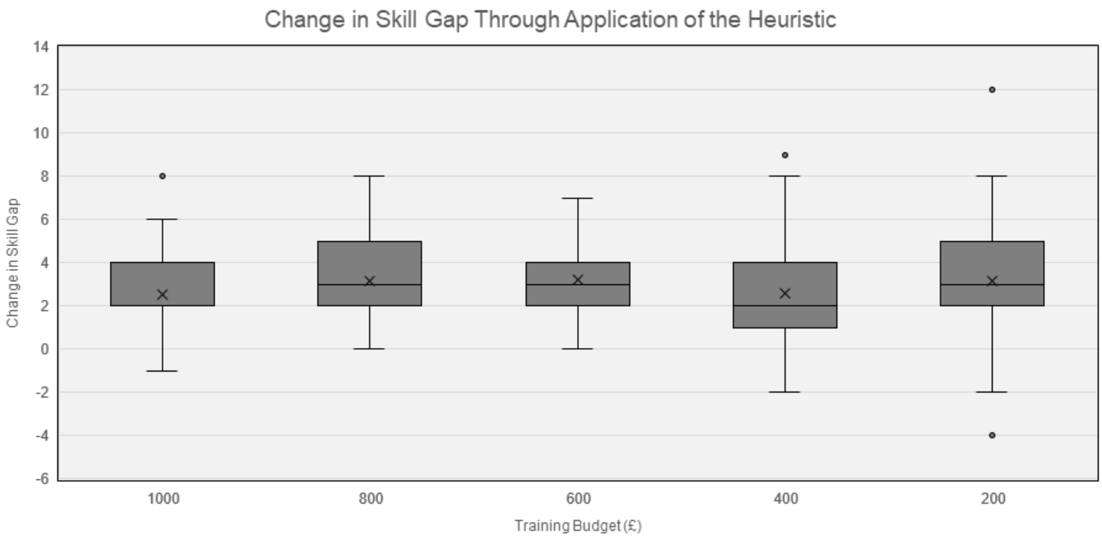

Fig. 7.6: The impact on the heuristic skill gap through changes to the training budget

The change in skill gap is reasonably small. There are 20 tasks in this model with each requiring between 1 and 3 completions in each time unit. The maximum skill gap change calculated using the heuristic is 12. This may be a significant change in the solution, however on average the skill gap change is 3 tasks which is much more reasonable as an estimate.

As seen previously, there are instances of data that produce a smaller skill gap than the exact solution. The skill gap is only a component of the objective function and by outputting this result, the quality of the solution is more difficult to compare. However, as the skill gap is determined the most useful output of the model, the aim of the heuristic is to accurately estimate this result.

7.2.4 Model Comparison

The behaviour of the heuristic has been analysed for all models combined, now the heuristic will be studied for the individual models. The size of the model is set to [10,20,7] and 25 iterations are used to ensure enough shakes were performed. Figure 7.7 calculates the change in skill gap under the different modelling constraints, calculated as heuristic solution minus exact solution.

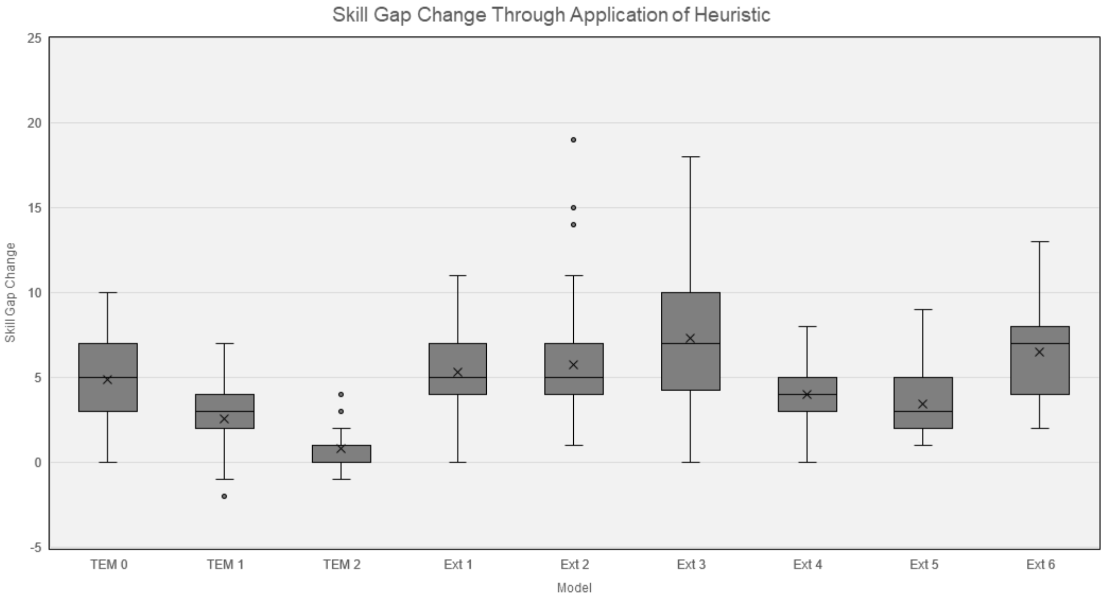

Fig. 7.7: Heuristic skill gap comparison for each model policy

Though not seen in the graph, the heuristic is least effective for estimating the solution to Extension Six. In many cases, the heuristic was poor at producing feasible solutions when the exact solution was capable. For all other model types, it is not possible to obtain infeasible solutions so this behaviour is not applicable.

The change in skill gap was largest for Extension Three, though in most of these model instances the quality of the heuristic solution is very similar. TEM Two produced the most accurate solution to the model. Overall, the largest skill gap change is given as 18 for Extension Two. This means 18 requirements for the tasks could not be completed under this solution that could have been completed given the exact solution. With only 20 tasks required approximately 1.5 times in each time unit, a skill gap of 18 is considered large. This solution would cause a significant impact if the company applied the solution.

However, on average the skill gap increase is 5 tasks. Here, the solution may be accepted by the company.

Some instances of the problem, such as in TEM Two, produce a smaller skill gap than the exact method. As discussed in previous results, these are due to the calculation of the objective differing in the two models.

The run times of the model are significantly improved by the heuristic. On average the run time is reduced in all model variations with the exception of Extension Four, skill specialities. For this model, the run time for the exact model was consistently low due to the small problem space. Similarly Extension Three shows poor run time improvement through the use of the heuristic, however the heuristic can be seen to improve the solution run time in some problem instances.

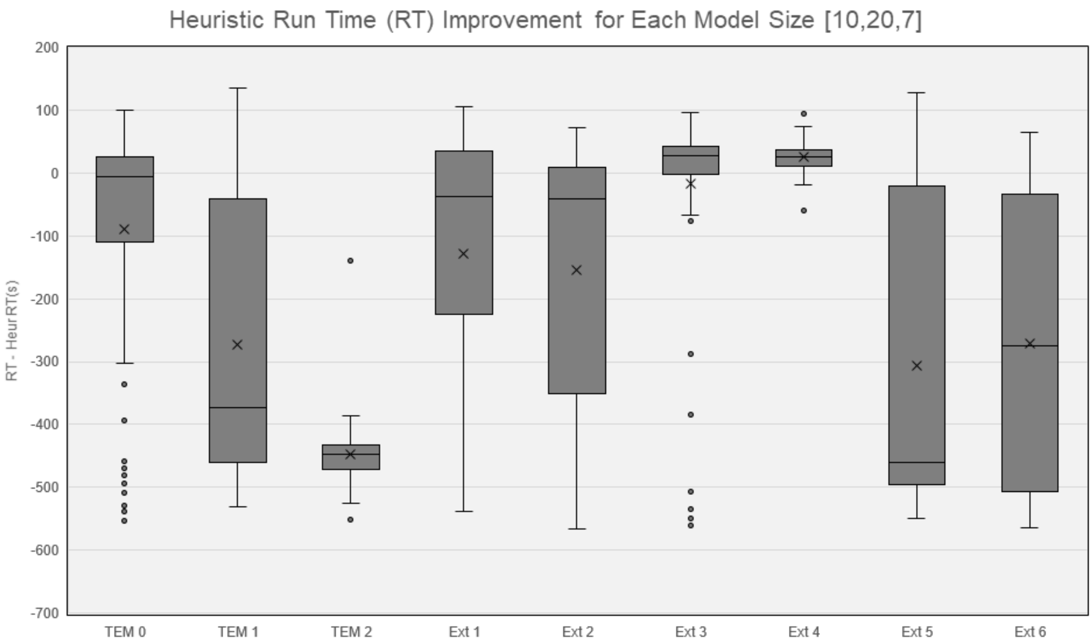

Fig. 7.8: Heuristic run time comparison for each model policy

The largest improvement can be seen in TEM Two, where all 100 runs of the model resulted in a shorter run time with the heuristic. In the exact model, the solution method was unable to prove optimality in almost all cases. Hence, the results seen here.

The heuristic model solves a smaller subproblem for each employee. The solution space is small and consistent for each iteration of the problem and thus has a short run time. Hence, the run time for the complete heuristic remains similar for each data set used. This is reflected in the results in Figure 7.8, where the run time improvement reflects the run times of the exact method.

The relationship between the quality of the solution and the resulting run time improvement can also be analysed. As mentioned, TEM Two creates the largest improvement in run time, however it also produces the most accurate heuristic solutions. Hence, if using TEM Two, the heuristic method is much more efficient and highly accurate.

Extension Three, however, produces solutions that are less accurate with little improvement to the run time. For this model, the heuristic is not effective enough at providing an accurate solution. However, the exact method here is sufficient at solving the problem in a reasonably small run time.

The other models produce varying results, though the run time is consistent in the heuristic method, the solution quality is less predictable.

7.2.5 Case Study

The decision support tool created for this model has been adapted to run the heuristic approach. Hence, the case study can be run from the same tool using the heuristic solution method. Given the results seen previously, the objective coefficient is increased to determine if the solution accuracy can be increased.

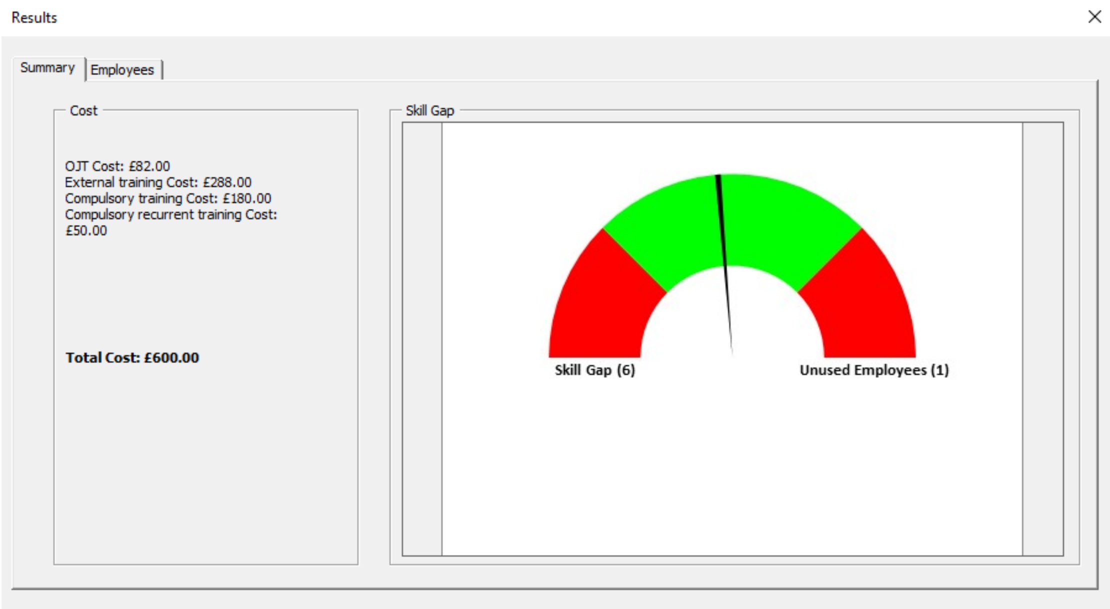

Fig. 7.9: Solution summary from case study using the heuristic approach

The data used to produce these results is the same data used in Chapter 6. With a problem size of [10,40,7], the model contains approximately 31,000 decision variables and 46,000 constraints.

The model is rerun using the heuristic approach with 25 iterations such that 15 random shakes are performed. Previously, a skill gap of four was obtained for a training cost of £579. The solution using the heuristic method can be seen in Figure 7.9.

The heuristic provided a solution in which six tasks could not be complete due to a lack of workforce. However, one employee was unused, in this case employee number 10. This is the last employee who would be allocated to tasks given the order of the heuristic iterations before the random shakes. However, there may be other reasons this employee is selected as less necessary in the solution. Though the number of skills they possesses at the beginning of the time horizon is not the lowest, they begin the model with skill 1 and not skill 2. The probability of any employee possessing skill 1 is 80% suggesting it is a common skill to hold and hence they may be less useful than other employees in the model.

The value of the skill gap is very similar to the exact model, with only two tasks difference. With 40 tasks being available in the model, the difference is not significant and this solution would be satisfactory. Though the run time of the exact method in this case study is small, the heuristic has been shown to have consistent low run times and would be beneficial if the case study had shown problems proving optimality, for instance when the coefficient in the objective was small.

As the training cost is not minimised in either model, some variation here is expected and both models are likely to produce training costs approximately equal to the training budget. Here the training cost is exactly equal to £600 with the training separated as follows:

- £82 was spent on completing on-the-job training
- £288 was spent on training courses
- £180 was spent on compulsory training
- £50 was spent on recurrent compulsory training

Comparing this to the exact solution, the results are similar. Much less money was spend on training courses and more money was spent on on-the-job training. Theoretically, this could be due to the order of the allocations and the large time requirement of these courses. It is less likely that the first employee assigned will be allocated to training courses as the objective function will be minimised if they are assigned to complete more tasks. The employee will be unable to complete tasks if they spend more time on training courses. Therefore to increase their skills, they are likely to be assigned to on-the-job training instead. The random shakes were designed to minimise the implications of allocating in order.

However, if no random shakes are used such that 10 iterations are completed, the allocations to on-the-job training and training courses remain the same. The overall skill gap is larger though, at 9 incomplete tasks.

Finally, the fairness value for each model is output. For the heuristic model, a fairness value of 8.7 is obtained. This is much higher than the fairness calculated for the exact model; 3.6. This implies that the total deviation from the average number of allocations is higher in the heuristic model. As the allocations are performed on each employee in turn, it is likely that the first employees assigned are given more tasks to complete. As one employee is unused in the model, it is possible that a small adjustment to the solution may produce a lower fairness deviation and allow the unused employee to be allocated to tasks.

7.3 Summary

It has been shown that the exact solution method has difficulty proving optimality in some data instances. The likelihood of this occurring increases with the size of the model. As such, a heuristic method is created to accurately calculate the skill gap given a decreased model run time.

Due to the run time being associated with the number of employees in the model, the heuristic approach is based on a decomposition method, where each iteration of the model solves a subproblem for each employee. As such, each employee in turn is assigned to the available tasks whilst considering the allocations of the other employees as fixed values. Once all employees have been assigned, a shake process is proposed to reassign allocations given the solution to the previous iteration. To make this shake method more effective, the solution accepted in each iteration i is limited to i branches in the branch and bound tree. As such, the solution should be improved as the number of iterations are increased and allocations should be more fairly distributed between the employees.

The constraints created for the heuristic approach were designed to make infeasible solutions not possible when optimising. In this way, the heuristic is comparable to the exact solution. As with the exact model, if no feasible training or allocations are possible, the resultant skill gap and fairness value will reflect this and affect the objective but the solution will still be valid.

In addition, it is still possible to get infeasible solutions in Extension Six. If the subproblem allows infeasible solutions, additional steps would be required in the heuristic to resolve for the employee until a feasible solution could be found.

The decision support tool created for this study was updated to include the heuristic option. The methods were then compared for difference problem sizes and each modelling component.

First, the number of iterations were analysed. It was discovered that the heuristic was capable of producing a smaller skill gap than the exact method. The likelihood of this occurring decreased with the number of iterations. Though the skill gap is not the only component of the objective function, it is the component that must be accurately calculated using the heuristic approach.

There were a few explanations provided for this behaviour. The coefficient in the objective function is different in the two models given the differences in the objectives. Hence, though the skill gap is smaller in the heuristic model, the fairness value is higher. As each employee is assigned independently, the fairness value is also affected by the allocations. The first employee assigned will be required to complete more tasks to improve the objective value. Overall, the fairness value is always larger. The limits to the branch and bound tree were designed to minimise the impact of this relationship, however it can be seen to still influence the solution quality.

Increasing the number of iterations was shown to improve the heuristic solution, however the optimal heuristic solution is achieved through a small number of iterations. This implied that the run time of the heuristic will always remain small as each iteration takes very little time to run and few iterations are required to optimise the solution.

The run time was analysed for different sizes of the problem with all model variations included. The heuristic was shown to improve the run time when the number of employees was increased. This is expected as the subproblem is defined by the number of employees, however the run time was not improved for variations to the number of tasks and time units. Here, the run times are reasonably small for the problem instances. Hence, the heuristic approach is recommended for problem sets where the number of employees may produce models that are difficult to solve.

Next, the training budget was adjusted and the heuristic approach applied. For large budgets it was shown that the run time of the model could be improved through the application of the heuristic. In addition, the accuracy of the heuristic in these problem instances was reasonably high, with an average skill gap change of three tasks. Hence, it is recommended that the heuristic is useful in problem instances where the training budget is reasonably large. For small budgets, the run time is small and hence the exact method can be used.

The models were also investigated independently to determine if a certain set of constraints affects the heuristic quality and efficiency. For the model instances, such as TEM Two, where the run times were large, the heuristic was shown to improve the run time significantly and provide accurate results. Conversely, for model instances where the run time was small for the exact method, the heuristic was less accurate.

The heuristic approach was used to produce an approximate solution to the case study proposed. The heuristic solution produced a similar recommendation to training and tasks than the exact model and thus would be sufficient to use in the case study. The method creates a decision support model, and thus it is likely that some variation in the results will occur when the solutions are implemented. Hence, the heuristic produces reasonable results.

The accuracy of the heuristic is correlated to the size of the problem space. As such, the heuristic is more accurate in models where it can significantly improve the performance of the model. It is less accurate in models where the exact method would produce a solution in minimal run times.

It could be suggested that the heuristic method is most effective when the solution space of the model is large. Here, many optimal solutions exist and thus the heuristic is more likely to estimate one of these solutions. In these instances, it has been shown that the exact method has difficulties producing an optimal solution. Hence, the heuristic method can be used to provide accurate and efficient results to the model in instances where the exact method would have problems proving optimality.

8. CONCLUSION

8.1 Model Development

The aim of this thesis is to present a novel approach to skill development in a multi-skilled workforce. The model provides support to decisions concerning training requirements and training allocations. In order to apply this model, a company must have a multi-skilled workforce where the workers are authorised to complete a set of tasks. The set of tasks must remain the same each time unit, though the probability of these tasks requiring completion each time unit may vary. Though this model has only been tested on a subset of all possible instances here, the model has been designed to bes as generic as possible. This allows it to be used by many companies and be tailored to their skill development needs.

The model is built using mixed integer linear programming with stochastic constraints where necessary. In the literature review many methods of modelling training were presented. Discrete event simulation was proposed but does not produce an optimal solution and was less useful when formulating allocation models. The specifications of the problem intuitively formulate decision variables and constraints. The flexibility of linear programming allows these additional decision variables and constraints to be added to the model depending on the requirements of the company. Whereas if DES was used, a new model would be required for each variation to the constraints. Hence, linear programming was selected.

To successfully incorporate the skill development, the model was designed to map the binary variable for authorisations with the requirement for experience to be gained in order to receive the skill. Hence, bridging the gap between the skill matrix where skills are defined as binary values and the non-linear learning curve which measures an employee's skill development. This approach has not been considered in the literature.

An employee may obtain the authorisation to complete the task if they have obtained enough experience before their auth is reviewed. This experience is provided through different training options.

Though literature exists analysing workforce allocation models, few papers consider policies to gain authorisations. A variety of policies are considered to calculate this skill and each policy is discussed to determine the effects on the development of the model, the results and the computation time.

The models can be applied in a stochastic environment, again an area not well considered in the literature. It is assumed that the probability of the tasks arriving depend on the current state of the world. To create a workforce solution to fulfil the stochastic demands it was determined that chance constraints would be most suitable.

The objective of the model is to minimise the skill gap and ensure the skills are fairly distributed amongst the workforce. To achieve this, the model must determine the optimal allocations of employees to tasks and training.

The default training is described as learning through task repetition. This was defined as TEM Zero. It is assumed that this learning is always available and must be included in the model. Each time a task is complete, one time unit of experience is gained for the associated skill. It should be noted that this training can only be completed by employees who already possess the skill. Hence, it allows an employee to retain an authorisation but not gain a new skill.

Additional training can be provided through on-the-job training (TEM One) and training courses (TEM Two). These two training methods are optional decisions and constraints that may be included in the model.

OJT allows an employee to gain experience in a skill they do not currently hold by being taught by an employee who does have the skill. The experience gained here is equivalent to completing the task without assistance (TEM Zero).

Training courses allow an employee to learn a skill they do not currently own. The definition of a training course is broad and can cover any training that does not involve allocations to a task. In other words, the training can be done remotely, away from the workplace or in a way that does not affect the current work processes. This includes online training, simulations, reading manuals and attending external training seminars. The experience gained here must be specified in equivalence to the training completed in TEM Zero.

In addition to the training that provide experience, different policies affect where an employee may be allocated to, in terms of both their skills and their training needs. Six policies were discussed and formulated as linear decision variables and constraints such that they may be encompassed in the model.

The first policy, Extension One, introduced compulsory training. This training is considered separately to the previous training options as it does not increase the experience of the employee. Moreover, this policy represents training that must be completed in order to obtain the skill. The compulsory training is not required for every task but, where specified, has an associated cost and completion time.

If the experience was not linear such that less experience is required for those employees who are reauthorising a skill, then Extension Two may be included. If this extension is used concurrently with Extension One then the compulsory training requirements for a currently authorised skill will differ from the requirement the first time the skill is learned.

With the exception of TEM Zero, the training methods described above are all associated with a cost. Hence, a budget must be specified for training needs. Though including this budget in the objective was discussed, it was decided that this would be best suited for a constraint. The remaining extensions are not associated with a cost but do affect the allocations to training courses and tasks.

Extension Three defines dependencies for the different auths. In this way, a skill may not be awarded unless the employee owns the dependent auth. This represents tasks that are complex or upgraded versions of previous tasks, or tasks that are essential to all employees.

The authorisations may only be available to employees of a certain skill level or are designated a skill speciality. Here, employees may be differentiated based on their abilities. Similarly, tasks can be grouped by their requirements. An employee can only complete a task if they possess the same level or speciality as the task. This is the definition of Extension Four.

Extension Five introduced non-linear learning to the model. This has been studied regularly through the literature, though it has not been used to determine binary skill allocations. It is assumed that an employee who has completed a task on multiple occasions requires less experience to show they are competent in a skill. Under this assumption, a graph was produced to represent the relationship and the allocations of the employee. This is used to define the experience required to gain the auth. Though the learning here is non linear, the mathematical programme is still built as a linear model.

Finally, Extension Six applies stochastic constraints to the model. Previously it was assumed that the same tasks are repeated in every time unit. Here, it is acknowledged that some industries do not perceive the same tasks but that these tasks arrive under a probability distribution. By applying chance constraints to the model, it can be specified that a proportion of tasks must be completed in a certain percentage of the possible scenarios, where the scenarios define a potential state of the world.

Each of these policies were designed to be implemented within the same model in any combination. Hence, 256 combinations of constraints are possible and an appropriate model may be selected for use by the company.

8.2 Model Analysis

Through the analysis process, these models were considered independently and concurrently to analyse the results, sensitivity and the run time. Though the analysis process is not thorough, it was designed to give a view on the model capabilities, validate the solutions by discussing the realism of the output and analyse some behaviours in the model.

The training budget was analysed first. If the training budget was small enough that the solution was affected, the run times of the model were shorter. However, by increasing the budget a threshold can be found such that the solution does not improve. Here, other constraints affect the allocations. In addition, these solutions had the longest run times. It was proposed that the solution space was largest when the training budget exceeded this threshold.

The models were analysed in turn by adjusting important coefficients in the constraints. For TEM Zero, the initial skills of the employees and the requirements of the tasks were varied. The requirements had a significant influence on the solution, causing higher skill gaps independent of the initial skills of the employees.

For TEMs One and Two the training costs were varied. As expected, a higher cost indicated a higher skill gap in the solution. For TEM One, this variation affected the run time of the model such that the run time increased with the cost. This did not happen in TEM Two. For TEM Two, the experience from the courses was also analysed and shown to impact the skill gap in the model such that the more experience gained from a course, the lower the skill gap.

The cost of compulsory training in Extension One also influenced the skill gap, however the number of tasks that required this training had a greater influence on the solution.

The initial skills of the employees were varied in both Extension Two and Three. In Extension Two a relationship was seen between these recurrent skills, the initial skills of the employee and the skill gap. Similarly, with Extension Three and the number of tasks that are dependent on other tasks. For Extension Two, this relationship was statistically significant. However, with Extension Three, though a relationship could be seen, it could not be proven as significant.

Extension Four demonstrated the relationship between the skill levels of the tasks and the skill levels of the employees. The skill gap was significantly affected by the relationship of these two values. It was proposed that the solution space was also influenced as the number of allocations would be limited by the values of these coefficients.

The parameters of the learning curve can be adjusted to analyse the solution given the shape of the curve. The parameters define the shape of the curve and determine whether a feasible solution is possible. If the shape of the curve is not formulated correctly, it is possible to get infeasible solutions or solutions in which the skill gap is always maximised. Similarly, it could be shown in Extension Six that feasibility of the model was affected by the choice in parameters.

The run time analysis indicated a problem solving the model under certain data inputs. A solution could be found in a reasonably small run time in these situations but the solver then could not prove optimality. The model could be run for 8 hours with no improvement to the solution found within the first few minutes. Though the models were evaluated individually and as a whole, there was no clear indication as to when these problem instances would be created.

However, it was noted that the number of employees affected the likelihood of the solver having issues proving optimality. In addition, if the problem space was too large, the solver would have problems. This was more likely to occur if TEMs One and Two were applied to the model.

As stated before, this analysis is not a thorough analysis of the model but provides some insight into the behaviour of the model and the validity of the results.

8.3 Case Study

For the combination of models described here to be used within a company, a tool had to be designed. This was built in Microsoft Excel with IBM Cplex used as the linear programme solver. A user guide was provided in this thesis to describe how the tool was built and how it can be used to run any combination of the TEMs and extensions.

The tool was then used for data provided by Boeing in the form of a case study. The data used in this case study was randomly produced from distributions created from real data. The results were then presented and analysed. The training budget was adjusted to evaluate the effects and a Pareto Curve was produced to show the relationship between the skill gap and the fairness value. Though this case study did not represent a specific problem instance at Boeing, it is included to demonstrate how the tool may be applied on experimental, feasible data.

The results from the case study showed how the solution from the model can help advise workforce allocation decisions. A skill gap of four tasks was achievable through the suggested allocations and a recommended training cost was provided. By reducing the cost, the solution remained the same up to a certain value. After which, the skill gap increased. This suggested the optimal training budget for that model was £400. To reduce the skill gap, a larger workforce would be required.

8.4 Heuristic

As discussed previously, the analysis within this thesis indicated a problem solving optimality within the model. This resulted in longer run times than necessary. Though it was determined that setting a wall time on the model was sufficient to ensure accurate results in a small run time, a heuristic approach was also proposed.

The heuristic was designed to solve the allocation problem for each employee in turn. It was shown that increasing the number of employees increased the run time of the model significantly. In addition, few constraints considered the relationship between employees in comparison to the relationships between tasks or time units.

Hence, for each employee, an allocation to training and tasks were made. The allocations of the other employees were fixed. The maximum number of branches available on the branch and bound tree were increased at each iteration to increase flexibility. As such, once all employees had been allocated, an employee was selected at random and their allocations were reassigned with a larger number of branches. Hence, it was proposed that each iteration would improve the solution by providing an additional opportunities for allocations in each iteration.

In practice, the heuristic was most effective at improving the run time for problems where the solution space was large. In these instances, it is believed that more optimal solutions are available, as there may be more than one, and hence it is more likely that the heuristic would estimate an optimal solution. It was determined that the heuristic is efficient and effective at solving this model when the problem could not be solved using exact methods.

8.5 Limitations and Future Research

Though it has been shown that this research satisfies the proposed research question and contributes to current literature, there are limitations to the model and future research that may be done to improve the results.

As shown in the literature, substitution and cross training impact the turnover rate in companies. As linear programming was used to formulate this model, it was not possible to include turnover without applying more stochastic techniques. If the time in which employees leave the company are known, the model would produce a solution under the knowledge that the employee will not be available in future. In reality, it is not known when the employee will leave, thus unnecessary training may be given to the employee. Future work would require turnover to be included in the model through stochastic programming methods, for instance a stochastic programme with recourse may be used. However, it should be noted that the run time of the model will be significantly increased through the addition of these constraints. Currently, the short run time of the model allows the model to calculate new recommendations when an employee is set to leave the company.

Additionally, hiring is not included in the model. However, the number of employees in the model can be set higher than the number of models currently in the workforce. They may be set with initial skills or be assigned zero skills. If one of these employees is assigned to a task this implies they have been hired into the company. This may help make hiring decisions but does not include a new employee joining the model at a random time in the run. Similarly to turnover, this has to be included with stochastic programming methods as allocations should not be affected by the random inclusion of this employee.

Though learning was incorporated as an extension to the model, little evaluation of the technique was completed. The curve provided throughout this thesis was randomly formulated given the initial and final experience of the employee. For future work, the learning curve given here should be reformulated using case study data. The true impact of training on the learning must be examined; the current model assumes that one completion of the task is equal to one unit of experience. This may not always be the relationship. In addition, a different curve could be created for each employee given that every employee is unique.

In relation to learning, the literature also introduces forgetting skills with time. A memoryless property is applied to the model such that only the previous skill changes the ability to gain a skill in Extension Two. This does replicate some properties of forgetting, however more research is required for this model. Some experimentation with forgetting was performed on the model. For instance, for every time unit the skill was not used, the experience was also decreased. However, a solution to this problem formulation could never be found due to the run times. Further research in forgetting would allow this model to analyse both skill development and skill fade.

Further analysis into the relationship between the TEMs and extensions is required. Though each model was analysed seperately, and the differences in the results discussed, there was little analysis on the results from each of the combinations of policies. As 256 combinations are possible, this analysis would take a large amount of time. However, the literature suggested that few papers analyse the influence of cross training on the employees. The model proposed allows cross training to be included with various training policies. The impact on the fairness of the allocations and the skill gap could be observed. The model could be solved with more employees than tasks and the amount of cross training recommended would suggest its effectiveness in industry. Again, each policy can be considered in this analysis.

Currently, the heuristic approach is designed such that the skill gap is being overestimated and the fairness value is usually larger. The first employee is assigned to the least training in order to reduce the skill gap in the objective. Indeed, the order of the employees has an influence on the solution of the problem. Though some ordering methods were proposed in Chapter 7, none were analysed in detail due to time constraints on the project. Future work should determine how the solution changes under these methods and if one produces a more effective heuristic.

The heuristic may also be improved by storing the best solution. It was discovered that the heuristic can produce a skill gap larger than the previous solution due to the weighting of the fairness. Thus, it is suggested that the random shakes include the option of declining a new solution if it does not produce a skill gap lower than the previous result.

However, the heuristic was not compared to any other methods in the literature. For instance, tabu search methods could be used to estimate the solution and many other techniques. The heuristic proposed was an investigation into the run time of the model and other approaches may exist that can estimate the exact method with more precision.

The model built in this thesis was designed such that additional constraints may be added to the model if needed and current constraints are easily adapted. The only constraints essential to the model are defined in TEM Zero. Hence, future work can build on the model proposed here to introduce new workforce policies as companies adapt and new training is offered.

With technological advances influencing the workplace, it will be essential to know how this influences training of employees and which opportunities should be addressed. As workplace tasks become more automated, the training to complete these tasks will change and the skill development of the employee will change. This model can be used as a template for these future analyses. Different training options can be introduced and the changes to the tasks can easily be implemented.

Due to the flexibility and generic nature of this model, it is suitable to support future workforce decisions and the analysis associated with it.

Appendices

Appendix A

LIST OF VARIABLES

Variable	Type	Definition
$e \in E$	$\mathbb{Z}^+$	Index of employees
$k \in K$	$\mathbb{Z}^+$	Index of tasks
$z \in \overline{Z}$	$\mathbb{Z}^+$	Index of review periods.
$t \in Z_{ekz}$	$\mathbb{Z}^+$	Set of times in review period z for employee e with task k.
$n \in N$	$\mathbb{Z}^+$	Index of training courses
σ	$\{1, 2\}$	Initial/recurrent
$l \in L$	$\mathbb{Z}^+$	Index of skill level
$x \in X$	$\mathbb{Z}^+$	Learning curve index
ξ_i	$1, .., m$	Random scenarios

Tab. A.1: Model Indices and sets

Variable	Type	Definition
w_{ekt}	$\mathbb{Z}$	Number of tasks k completed by e at t
s_{ekz}	$\{0, 1\}$	1 if employee e has auth k in review period z
ψ_e^+	$\mathbb{Q}^+$	Positive deviation in fairness
ψ_e^-	$\mathbb{Q}^+$	Negative deviation in fairness
p_{ekt}	$\{0, 1\}$	1 if employee e is assigned to OJT in k at time t
v_{ent}	$\{0, 1\}$	1 if employee e completes course n at time t
$q_{ekt\sigma}$	$\{0, 1\}$	1 if employee e is assigned to compulsory course k at time t for initial or recurrent auth
y_{ekzx}	$\{0, 1\}$	1 if employee e completed auth k, x times before period z
δ_{ξ_i}	$\{0, 1\}$	1 if scenario ξ_i does not satisfy the conditions of the chance constraint.

Tab. A.2: Model Decision Variables

Variable	Type	Definition
M	$>> 0$	A large number referred to as "big M"
A, B	$\mathbb{Q}$	Coefficient to define priority of objectives
c_1	$\mathbb{Q}^+$	Cost of OJT training
c_{2n}	$\mathbb{Q}^+$	Cost of training course number n
c_{3k}	$\mathbb{Q}^+$	Cost of compulsory training for task k
b	$\mathbb{Q}^+$	Available training budget
r_k	$\mathbb{Z}^+$	Number of times task k will require completing in each time unit
a_{et}	$\mathbb{Q}^+$	Available hours of employee e at time t, given annual leave or sick leave
η_k	$\mathbb{Q}^+$	Processing time of task k
Θ_{ek}	$\{0, 1\}$	1 if employee e has auth k at time 0
ρ_n	$\mathbb{Q}^+$	Length of training course n
χ_{nk}	$\mathbb{Q}^+$	Experience gained from training course n for task k
u_{nt}	$\mathbb{Z}^+$	Number of spaces on course n at time t
$\mu_{k\sigma}$	$\mathbb{Q}^+$	Length of compulsory training k for initial, $\sigma = 1$, or recurrent, $\sigma = 1$, auths
$g_{kx\sigma}$	$\mathbb{Q}^+$	Amount of experience needed for task k for initial, $\sigma = 1$, or recurrent, $\sigma = 1$, auths, having completed it x times
$\pi_{k\sigma}$	$\{0, 1\}$	Compulsory training required to complete task k for initial, $\sigma = 1$, or recurrent, $\sigma = 1$, auths
d_{ki}	$\{0, 1\}$	1 if auth k requires the completion of auth $i \in K$
λ_{el}	$\{0, 1\}$	Employee e can complete tasks of speciality l
ϕ_{kl}	$\{0, 1\}$	Task k requires speciality l
α, β	$\mathbb{Q}^+$	Probability thresholds
ζ_{ξ_i}	$\mathbb{Q}^+$	Probability of scenario ξ_i
$\gamma_{k\xi_i}$	$\mathbb{Q}^+$	Probability of task k in scenario ξ_i

Tab. A.3: Model Data

Appendix B

FULL MODEL FORMULATION

$$\min \quad \sum_{e\in E}(\psi_e^+ + \psi_e^-) + A\sum_{k\in K}\sum_{t\in T}(r_k - \sum_{e\in E} w_{ekt}) \tag{B.1}$$

$$\text{s.t.} \quad \frac{1}{\hat{e}-1}\sum_{e\in E:e\neq f}\sum_{k\in K}\sum_{t\in T} w_{ekt} - \sum_{k\in K}\sum_{t\in T} w_{fkt} \leq \psi_f^+ \qquad \forall f \in E \tag{B.2}$$

$$\sum_{k\in K}\sum_{t\in T} w_{fkt} - \frac{1}{\hat{e}-1}\sum_{e\in E:e\neq f}\sum_{k\in K}\sum_{t\in T} w_{ekt} \leq \psi_f^- \qquad \forall f \in E \tag{B.3}$$

$$w_{ekt} - r_k s_{ekz} \leq 0 \qquad \forall e \in E, k \in K, z \in \overline{Z}, t \in Z_{ekz} \tag{B.4}$$

$$\sum_{k\in K}\gamma_{k\xi_i}\eta_k(w_{ekt} + p_{ekt}) + \sum_{n\in N}\rho_n v_{ent} + {} + \sum_{k\in K}\sum_{\sigma=1}^{2}\mu_{k\sigma}q_{ekt\sigma} \leq a_{et} \qquad \forall e \in E, t \in T, i \in 1..m \tag{B.5}$$

$$\sum_{e\in E} w_{ekt} \leq r_k \qquad \forall k \in K, t \in T \tag{B.6}$$

$$s_{ek1} = \Theta_{ek} \qquad \forall e \in E, k \in K \tag{B.7}$$

$$\sum_{x\in X} g_{kx1}y_{ekzx} - \boldsymbol{M}_k(1 - s_{ekz}) \leq \sum_{t\in Z_{ek(z-1)}}(\gamma_{k\xi_i}(w_{ekt} + p_{ekt}) + \sum_{n\in N}\chi_{nk}v_{ent}) + M_k s_{ek(z-1)} \qquad \forall e \in E, k \in K, z \in \overline{Z}, i \in 1..m \tag{B.8}$$

$$\sum_{x\in X} g_{kx2}y_{ekzx} - \boldsymbol{M}_k(1 - s_{ekz}) \leq \sum_{t\in Z_{ek(z-1)}}(\gamma_{k\xi_i}(w_{ekt} + p_{ekt}) + \sum_{n\in N}\chi_{nk}v_{ent}) + M_k(1 - s_{ek(z-1)}) \qquad \forall e \in E, k \in K, z \in \overline{Z}, i \in 1..m \tag{B.9}$$

$$\sum_{e\in E}\sum_{t\in T}\left(\sum_{k\in K} c_1 p_{ekt} + \sum_{n\in N} c_{2n}v_{ent} + {} + \sum_{k\in K}\sum_{\sigma=1}^{2} c_{3k\sigma}q_{ekt\sigma}\right) \leq b \tag{B.10}$$

$$p_{ekt} - \sum_{f \in E: e \neq f} w_{fkt} \leq 0 \qquad \forall e \in E, k \in K, t \in T \tag{B.11}$$

$$\sum_{e \in E} v_{ent} \leq u_{nt} \qquad \forall n \in N, t \in T \tag{B.12}$$

$$\pi_{k1} s_{ekz} - \sum_{t \in Z_{ek(z-1)}} q_{ekt1} - s_{ek(z-1)} \leq 0 \qquad \forall e \in E, k \in K, z \in \overline{Z} \tag{B.13}$$

$$\pi_{k2} s_{ekz} - \sum_{t \in Z_{ek(z-1)}} q_{ekt2} - (1 - s_{ek(z-1)}) \leq 0 \qquad \forall e \in E, k \in K, z \in \overline{Z} \tag{B.14}$$

$$s_{ekz} - s_{ei\tilde{z}} \leq 1 - d_{ki} \qquad \forall e \in E, k \in K, z \in \overline{Z}, i \in K, \tilde{z} \in Z_{ekz} \cap Z_{ei\tilde{z}} \tag{B.15}$$

$$w_{ekt} - r_k \sum_{l \in L} \lambda_{el} \phi_{lk} \leq 0 \qquad \forall e \in E, k \in K, t \in T \tag{B.16}$$

$$\sum_{x \in X} x y_{ekzx} = \sum_{t=1}^{\max(Z_{ek(z-1)})} w_{ekt} \qquad \forall e \in E, k \in K, z \in \overline{Z} \tag{B.17}$$

$$\sum_{x \in X} y_{ekzx} = 1 \qquad \forall e \in E, k \in K, z \in \overline{Z} \tag{B.18}$$

$$\sum_{k \in K} \gamma_{k\xi_i} (r_k - \sum_{e \in E} \sum_{k \in K} w_{ek\hat{t}}) \leq \beta \sum_{k \in K} \gamma_{k\xi_i} r_k + \boldsymbol{M} \delta_{\xi_i} \qquad \forall i \in 1, .., m \tag{B.19}$$

$$\sum_{i=1}^{m} \delta_{\xi_i} \zeta_{\xi_i} \leq 1 - \alpha \tag{B.20}$$

Appendix C

CASE STUDY FORMULATION

$$\min \quad \sum_{e\in E}(\psi_e^+ + \psi_e^-) + A\sum_{k\in K}\sum_{t\in T}(r_k - \sum_{e\in E} w_{ekt}) \tag{C.1}$$

$$\text{s.t.} \quad \frac{1}{\hat{e}-1}\sum_{e\in E:e\neq f}\sum_{k\in K}\sum_{t\in T} w_{ekt} - \sum_{k\in K}\sum_{t\in T} w_{fkt} \le \psi_f^+ \qquad \forall f \in E \tag{C.2}$$

$$\sum_{k\in K}\sum_{t\in T} w_{fkt} - \frac{1}{\hat{e}-1}\sum_{e\in E:e\neq f}\sum_{k\in K}\sum_{t\in T} w_{ekt} \le \psi_f^- \qquad \forall f \in E \tag{C.3}$$

$$w_{ekt} - r_k s_{ekz} \le 0 \qquad \forall e \in E, k \in K, z \in \overline{Z}, t \in Z_{ekz} \tag{C.4}$$

$$\sum_{k\in K}\eta_k(w_{ekt} + p_{ekt}) + \sum_{n\in N}\rho_n v_{ent} + + \sum_{k\in K}\sum_{\sigma=1}^{2}\mu_{k\sigma}q_{ekt\sigma} \le a_{et} \qquad \forall e \in E, t \in T \tag{C.5}$$

$$\sum_{e\in E} w_{ekt} \le r_k \qquad \forall k \in K, t \in T \tag{C.6}$$

$$s_{ek1} = \Theta_{ek} \qquad \forall e \in E, k \in K \tag{C.7}$$

$$g_{k1} - \boldsymbol{M}_k(1 - s_{ekz}) \le \sum_{t\in Z_{ek(z-1)}}((w_{ekt} + p_{ekt}) + \sum_{n\in N}\chi_{nk}v_{ent}) + M_k s_{ek(z-1)} \qquad \forall e \in E, k \in K, z \in \overline{Z} \tag{C.8}$$

$$g_{kx} - \boldsymbol{M}_k(1 - s_{ekz}) \le \sum_{t\in Z_{ek(z-1)}}((w_{ekt} + p_{ekt}) + \sum_{n\in N}\chi_{nk}v_{ent}) + M_k(1 - s_{ek(z-1)}) \qquad \forall e \in E, k \in K, z \in \overline{Z} \tag{C.9}$$

$$\sum_{e\in E}\sum_{t\in T}\left(\sum_{k\in K} c_1 p_{ekt} + \sum_{n\in N} c_{2n}v_{ent} + + \sum_{k\in K}\sum_{\sigma=1}^{2} c_{3k\sigma}q_{ekt\sigma}\right) \le b \tag{C.10}$$

$$p_{ekt} - \sum_{f \in E: e \neq f} w_{fkt} \leq 0 \qquad \forall e \in E, k \in K, t \in T \tag{C.11}$$

$$\sum_{e \in E} v_{ent} \leq u_{nt} \qquad \forall n \in N, t \in T \tag{C.12}$$

$$\pi_{k1} s_{ekz} - \sum_{t \in Z_{ek(z-1)}} q_{ekt1} - s_{ek(z-1)} \leq 0 \qquad \forall e \in E, k \in K, z \in \overline{Z} \tag{C.13}$$

$$\pi_{k2} s_{ekz} - \sum_{t \in Z_{ek(z-1)}} q_{ekt2} - (1 - s_{ek(z-1)}) \leq 0 \qquad \forall e \in E, k \in K, z \in \overline{Z} \tag{C.14}$$

$$s_{ekz} - s_{ei\tilde{z}} \leq 1 - d_{ki} \qquad \forall e \in E, k \in K, z \in \overline{Z}, i \in K, \tilde{z} \in Z_{ekz} \cap Z_{ei\tilde{z}} \tag{C.15}$$

$$w_{ekt} - r_k \sum_{l \in L} \lambda_{el} \phi_{lk} \leq 0 \qquad \forall e \in E, k \in K, t \in T \tag{C.16}$$

Appendix D

HEURISTIC FORMULATION

$$\min \quad A\sum_{k\in K}\sum_{t\in T}(r_k - \sum_{e\in E;e\neq\bar{e}} \overline{w}_{ekt} - w_{kt}) + (\psi^+ + \psi^-) + B\sum_{i=1}^{m}\delta_{\xi_i} \tag{D.1}$$

$$\sum_{k\in K}\sum_{t\in T}(\frac{1}{\hat{e}-1}\sum_{e\in\overline{E}}\overline{w}_{ekt} - w_{kt}) \leq \psi^+ \tag{D.2}$$

$$\sum_{k\in K}\sum_{t\in T}(w_{kt} - \frac{1}{\hat{e}-1}\sum_{e\in\overline{E}}\overline{w}_{ekt}) \leq \psi^- \tag{D.3}$$

$$\text{s.t} \quad w_{kt} - r_k s_{kz} \leq 0 \qquad \forall k\in K, z\in\overline{Z}, t\in Z_{kz} \tag{D.4}$$

$$\sum_{k\in K}\gamma_{k\xi_i}\eta_k(w_{kt}+p_{kt}) + \sum_{n\in N}\rho_n v_{nt} + \\ + \sum_{k\in K}\sum_{\sigma=1}^{2}\mu_{k\sigma}q_{kt\sigma} \leq a_t \qquad \forall t\in T, i\in 1..m \tag{D.5}$$

$$\sum_{e\in\overline{E}}\overline{w}_{ekt} + w_{kt} \leq r_k \qquad \forall k\in K, z\in\overline{Z}, t\in Z_{kz} \tag{D.6}$$

$$s_{k1} = \Theta_k \qquad \forall k\in K \tag{D.7}$$

$$\sum_{x\in X} g_{kx1}y_{kzx} - M_k(1-s_{kz}) \leq \sum_{t\in Z_{k(z-1)}}(\gamma_{k\xi_i}(w_{kt}+p_{kt}) \\ + \sum_{n\in N}\chi_{nk}v_{nt}) + M_k s_{k(z-1)} \qquad \forall k\in K, z\in\overline{Z} \tag{D.8}$$

$$\sum_{x\in X} g_{kx2}y_{k(p-1)x} - M_k(1-s_{kp}) \leq \sum_{t\in Z_{k(z-1)}}(\gamma_{k\xi_i}(w_{kt}+p_{kt}) \\ + \sum_{n\in N}\chi_{nk}v_{nt}) + M_k(1-s_{k(z-1)}) \qquad \forall k\in K, z\in\overline{Z} \tag{D.9}$$

$$\sum_{k\in K}\sum_{t\in T} c_1(p_{kt} + \sum_{e\in\overline{E}} \overline{p}_{ekt}) + \sum_{n\in N}\sum_{t\in T} c_{2n}(v_{nt} + \sum_{e\in\overline{E}} \overline{v}_{ent})$$
$$+ \sum_{k\in K}\sum_{t\in T}\sum_{\sigma=1}^{2} c_{3k\sigma}(qkt\sigma + \sum_{e\in\overline{E}} \overline{q}_{ekt\sigma}) \leq b \qquad \text{(D.10)}$$

$$p_{kt} - \sum_{e\in\overline{E}} \overline{w}_{ekt} \leq 0 \qquad \forall k \in K, t \in T \qquad \text{(D.11)}$$

$$w_{kt} + \sum_{e\in\overline{E}} \overline{w}_{ekt} \geq \overline{p}_{ekt} \qquad \forall e \in E, k \in K, t \in T \qquad \text{(D.12)}$$

$$v_{nt} + \sum_{e\in\overline{E}} \overline{v}_{ent} \leq u_{nt} \qquad \forall n \in N, t \in T \qquad \text{(D.13)}$$

$$\pi_{k1}s_{kz} - \sum_{t\in Z_{k(z-1)}} q_{kt1} + s_{kp} \leq 0 \qquad \forall k \in K, z \in \overline{Z} \qquad \text{(D.14)}$$

$$\pi_{k2}s_{kz} - \sum_{t\in Z_{k(z-1)}} q_{kt2} + (1 - s_{k(z-1)}) \leq 0 \qquad \forall k \in K, z \in \overline{Z} \qquad \text{(D.15)}$$

$$s_{kz} - s_{i\tilde{z}} \leq 1 - d_{ki} \qquad \forall k \in K, z \in \overline{Z}, i \in K, \tilde{z} \in Z_{kz} \cap Z_{i\tilde{z}} \qquad \text{(D.16)}$$

$$w_{kt} - r_k \sum_{l\in L} \lambda_l \phi_{lk} \leq 0 \qquad \forall k \in K, t \in T \qquad \text{(D.17)}$$

$$\sum_{x\in X} x y_{kzx} = \sum_{t=1}^{\max(Z_{k(z-1)})} w_{kt} \qquad \forall k \in K, z \in \overline{Z} \qquad \text{(D.18)}$$

$$\sum_{x\in X} y_{kzx} = 1 \qquad \forall k \in K, z \in \overline{Z} \qquad \text{(D.19)}$$

$$\sum_{k\in K} \gamma_{k\xi_i} r_k - \sum_{k\in K} \gamma_{k\xi_i}(w_{k\hat{t}} + \sum_{e\in\overline{E}} \overline{w}_{ek\hat{t}}) \leq \beta \sum_{k\in K} \gamma_{k\xi_i} r_k + \boldsymbol{M}\delta_{\xi_i} \qquad \forall i \in 1, .., m \qquad \text{(D.20)}$$

GLOSSARY

DES	Discrete Event Simulation
DoD	Department of Defence
iid	independent identically distributed
ILP	Integer Linear Programme
LC	Learning Curve
MILP	Mixed Integer Linear Programme
OJT	On-the-job training
RAF	Royal Air Force
SAA	Sample Average Approximation
SD	System Dynamics
SMILP	Stochastic Mixed Integer Linear Programme
TEM	Training Execution Model
TNA	Training Needs Analysis
VBA	Visual Basic for Application

BIBLIOGRAPHY

L. An, J. Jeng, Y. Lee et al. (2007), "Effective workforce lifecycle management via System Dynamics modeling and simulation", *Proceedings - Winter Simulation Conference.*

M. Anzanello & F. Fogliatto, (2011), "Learning curve models and applications: Literature review and research directions", *International Journal of Industrial Ergonomics*, **41**(5), p573-583.

J. Atlason, M. Epelman, & S. Henderson, (2008), "Optimizing call center staffing using simulation and analytic center cutting plane methods", *Management Science* **54**, p295-309.

E. Attia, P. Duquenne & J-M. Le-Lann, (2014), "Considering skills evolutions in multi-skilled workforce allocation with flexible working hours", *International Journal of Production Research*, **52**(15), p4548-4573.

N. Azizi & M. Liang, (2013), "An integrated approach to worker assignment, workforce flexibility acquisition and task rotation", *Journal of the Operational Research Society*, **64**(2), p260-275.

E. Balas, (1975), "Facets of the Knapsack Polytope", *Mathematical Programming*, **8**, p146-164.

W. Balinsky & A. Reisman, (1972), "Some Manpower Planning Models Based on Levels of Educational Attainment", *Management Science*, **18**(12), p691-705.

W. Bean, (2011), "Inventory management under uncertainty : A military application", Unpublished MSc dissertation, University of Pretoria.

O. Bellenguez & E. Néron, (2005), "Lower bounds for the multi-skill project scheduling problem with hierarchical levels of skills", *Lecture Notes in Computer Science*, **3616**, p229-243.

D. Bertsimas & M. Sim, (2004), "The Price of Robustness", *Operations Research*, **52**(1), p35-53.

R. Bhatnagar, V. Saddikutti & A. Rajgopalan, "Contingent manpower planning in a high clock speed industry", *International Journal of Production Research*, **45**(9).

J. Birge & F. Louveaux, (1997), "Introduction to Stochastic Programming", US:Springer.

D. Biskup, (1999), "Single-machine scheduling with learning considerations", *European Journal of Operational Research*, **115**(1), p173-178.

J. Cao et al. (2009), "A policy-based authorization model for workflow-enabled dynamic process management", *Journal of Network and Computer Applications*, **32**, p412-422.

B. Caputo, (1969), "A Mathematical Approach to Measuring Manpower Requirements", paper presented at *Operational and Personnel Research in the Management of Manpower Systems*, Germany:NATO Science Committee.

S. Carlos, A. Sánchez, & S. Martorell, (2011), "Model to study the effect of workforce on a safety equipment and its optimization", *Mathematical and Computer Modelling*, **54**(7–8), p1808-1812.

J. Carlson, (1973), "Cubic learning curve: precision tool for labor estimating", *Manufacturing, Engineering and Management*, **71**, p22-25.

A. Charnes & W. Cooper, (1959), "Chance-Constrained Programming", *Management Science*, **6**(1), p73-79.

A. Charnes & W. Cooper, (1961), "Management Models and Industrial Applications of Linear Programming", US:John Wiley & Sons, Inc.

D. Chen, S. Liu & Y. Wang, (2014), "Modeling for project scheduling with multi-skilled workforce constraints and uncertainty", *26th Chinese Control and Decision Conference, CCDC 2014*.

V. Chvatal, (1973), "Edmonds polytopes and a hierarchy of combinatorial problems", *Discrete Mathematics*, **4**, p305-337

D. Clough, R. Dudding & W. Price, (1969), "Mathematical Programming Models of Quasi-Independent Subsystem of the Canadian Forces Manpower System", paper presented at *Operational and Personnel Research in the Management of Manpower Systems*, Germany:NATO Science Committee.

J. Cohen, E. J. Dearnaley & C. E. M. Hansel (1956), "Influence of Training on the Performance of Bus Drivers", *Journal of the Operational Research Society*, **7**(3), p67-82.

Y. Collette & P. Siarry, (2003), "Multiobjective Optimization: Principles and Case Studies", US:Springer.

R. Conway & A. Schultz, (1959), "The manufacturing progress function", *Industrial Engineering*, **10**, p120-127.

H. Crowder, E. Johnson & M. Padberg, "Solving large-scale zero-one linear programming problems", *Operations Research*, **31**, p803-834.

G. Dantzig & P. Wolfe, (1960), "Decomposition Principle for Linear Programs", *Operations Research*, **8**(1), p101-111.

R. Darmon, (2004), "Controlling sales force turnover costs through optimal recruiting and training policies", *European Journal of Operational Research*, **154**(1), p291-303.

P. Datta, A. Srivastava & R. Roy, (2013), "A simulation study on maintainer resource utilization of a fast jet aircraft maintenance line under availability contract", *Computers in Industry*, **64**(5), p543-555.

P. De Bruecker et al. (2015), "Workforce planning incorporating skills : state of the art", *European Journal of Operational Research*, **243**, p1-16.

J. DeJong, (1957), "The effect of increased skills on cycle time and its consequences for time standards", *Ergonomics*, **1**, p51-60.

D. Dietz, (1991), “Optimal Specialization and Allocation of Maintenance Manpower”, Thesis, US:The Pennsylvania State University.

C. du Boulay & C. Medway, (1999), “The clinical skills resource: A review of current practice”, *Medical Education*, **33**, p185–191.

S. Duffuaa & K. Al-Sultan, (1999), “A stochastic programming model for scheduling maintenance personnel”, *Applied Mathematical Modelling*, **23**, p385-397.

C. Eden, T. Williams & F. Ackermann, (1998), “Dismantling the learning curve: the role of disruptions on the planning of development projects”, *International Journal of Project Management*, **16**(3), p131-138.

F. Edgeworth, (1888), “The Mathematical Theory of Banking”, *Journal of the Royal Statistical Society*, **51**, p113-127.

K. Ertogral, & B. Bamuqabel, (2008), “Developing staff schedules for a bilingual telecommunication call center with flexible workers”, *Computers & Industrial Engineering*, **54**(1), p118-127.

M. Firat & C. Hurkens, (2012), “An improved mip-based approach for a multi-skill workforce scheduling problem”, *Journal of Scheduling*, **15**, p363-380.

L. Florez, D. Castro-Lacouture & A. Medaglia, (2012), “Sustainable workforce scheduling in construction program management”, *Journal of the Operational Research Society*, **64**(8), p1169-1181.

N. Fouad et al. (2009), “Competency Benchmarks: A Model for Understanding and Measuring Competence in Professional Psychology Across Training Levels”, *Training and Education in Professional Psychology*, **3**(4)

J. Fowler, P. Wirojanagud, E. Gel, (2008), “Heuristics for workforce planning with worker differences”, *European Journal of Operational Research*, **190**, p724-740.

A. Garza et al. (2014), “System Dynamics Based Manpower Modeling”, paper presented at *the 2014 Industrial and Systems Engineering Research Conference*.

R. Gay, (1974), “Estimating the Cost of On-the-Job Training in Military Occupations: A Methodology and Pilot Study”, *A Report prepared for Defense Advanced Research Projects Agency*, US:RAND

M. Geyer, B. Hiller & S. Meinert, (2010), "‘Algorithm Engineering: Bridging the Gap Between Algorithm Theory and Practice"’, Chapter 2, Germany:Springer

B. Geerts & I. Vliegen, (2012), “Model for scheduling multi-skilled personnel”, Master’s Thesis, Department of Clinical Neurophysiology, University of Twente, Enschede.

S. Globerson, N. Levin & A. Schub, (1989), “The impact of breaks on forgetting when performing a repetitive task”, *IIETransactions*, **21**(4), p376-381.

J. Glover, (1966), “Manufacturing progress functions: an alternative model and its comparison with existing functions”, *International Journal of Production Research*, **4**, p279-300

J. Gomar, C. Haas & D. Morton, (2002), "Assignment and Allocation Optimization of Partially Multiskilled Workforce", *Journal of Construction Engineering and Management*, **128**, p103-109.

R. Gomory, (1958), "Outline of an algorithm for integer solutions to linear programs", *Bulletin of the American Mathematical Society*, **64**(5), p275-278.

R. Groover, (1969), "A Generalized Entity Simulation Model of a Military Personnel System", paper presented at *Operational and Personnel Research in the Management of Manpower Systems*, Germany:NATO Science Committee.

W. Gutjahr & K. Froeschl, (2013), "Project portfolio selection under uncertainty with outsourcing opportunities", *Flexible Services and Manufacturing Journal*, **25**, p255-281.

K. Hafeez, H. Abdelmeguid, (2003), "Dynamics of human resource and knowledge management", *Journal of the Operational Research Society*, **54**(2), p153-164.

C. Heimerl, & R. Kolisch, (2010), "Scheduling and staffing multiple projects with a multi-skilled workforce", *OR Spectrum*, **32**(2), p343-368.

W. Hirshmann, (1964), "The Learning Curve", *Chemical Engineering*, **71**, p95-100

J. Hurley, (1996), "When are we going to change the learning curve lecture?", *Computers Operations Research*, **23**(5), p509-511.

IBM Knowledge Centre, (2017), "Progress reports: interpreting the node log", Retrieved 14 Nov, 2017, from `https://www.ibm.com/support/knowledgecenter/SSSA5P_12.7.0/ilog.odms.cplex.help/CPLEX/UsrMan/topics/discr_optim/mip/para/52_node_log.html`

IBM Knowledge Centre, (2017), "MIP emphasis switch", Retrieved 14 Nov, 2017,`https://www.ibm.com/support/knowledgecenter/en/SS9UKU_12.4.0/com.ibm.cplex.zos.help/Parameters/topics/MIPEmphasis.html`

J. Ignizio, (1983), "Generalized Goal Programming, An Overview", *Computers & Operations Research*, **10(4)**.

C. Iwata & D. Mavris, (2013), "Object-Oriented Discrete Event Simulation Modeling Environment for Aerospace Vehicle Maintenance and Logistics Process", *Procedia Computer Science*, **16**, p187-196.

M. Jaber, (2006), "Learning and Forgetting Models and Their Applications", *Handbook of Industrial and Systems Engineering*

R. Jepsen, (2016), "Evaluating structured assessment of anaesthesiologists ' non-technical skills", *Acta Anaesthesiologica Scandinavica*, **60**, p756-766.

Y. Juang, S. Lin & H. Kao, (2007), "An adaptive scheduling system with genetic algorithms for arranging employee training programs", *Expert Systems with Applications*, **33**(3), p642-651.

A. B. Kahn, (1962), "Topological sorting of large networks", *Communications of the ACM*, **5**(11), p558-562.

K. Kim et al. (2012), "Managing the supply and demand uncertainty in workforce recruitment: planned or just-in-time acquisition", *Journal of the Operational Research Society*, **64**(11), p1654-1663.

E. Klotz & A. Newman, (2013), "Practical Guidelines for Solving Difficult Mixed Integer Linear Programs", *Surveys in Operations Research and Management Science*, **18**(1), p18-32.

Y. Kuo, J. Leung & C. Yano, (2014), "Scheduling of multi-skilled staff across multiple locations", *Production and Operations Management*, **23**(4), p626-644.

Laerd Statistics, (2013), "Pearson's Product-Moment Correlation using SPSS Statistics", Retrieved 17th October, 2017, from `https://statistics.laerd.com/spss-tutorials/pearsons-product-moment-correlation-using-spss-statistics.php`

A. Land & A. Doig, (1960), "An Automatic Method of Solving Discrete Programming Problems", *Econometrica*, **28**(3), p497-520.

F. Levy, (1965), "Adaptation in the production process", *Management Science*, **11**, B136-B154

N. Li & L. Li, (2000), "Modeling staffing flexibility: A case of China", *European Journal of Operational Research*, **124**, p255-266.

C. Liu, N. Yang, W. Li et al. (2013), "Training and assignment of multi-skilled workers for implementing seru production systems", *Training and assignment of multi-skilled workers for implementing seru production systems*, **69**, p937-959.

J. Luedtke & S. Ahmed, (2008), "A sample approximation approach for optimization with probabilistic constraints", *SIAM Journal on Optimization*, **19**, p674-699.

K. Marentette, A. Johnson, & L. Mills, (2009), "A measure of cross-training benefit versus job skill specialization", *Computers & Industrial Engineering*, **57**(3), p937-940.

L. Mashaud et al. (2010), "Two-year skill retention and certification exam performance after fundamentals of laparoscopic skills training and proficiency maintenance", *Surgery*, **148**(2), p194-201.

M. Marozzi & M. Bolzan, (2016), "Skills and training requirements of municipal directors : a statistical assessment", *Quality & Quality*, **50**, p1093-1115.

J. Mincer, (1962), "On-the-Job Training: Costs, Returns and Some Implications", *Journal of Political Economy*, **70**(5).

W. Mooz, (1969), "The Pilot Training Study: Personnel Flow and the Pilot Model", US:Rand

R. Morgan, (1969), "Manpower Planning in the Royal Air Force; an exercise in Linear Programming", paper presented at *Operational and Personnel Research in the Management of Manpower Systems*, Germany:NATO Science Committee.

M. Murray, (1987), "Some Aspects of Operational Analysis in the Ministry of Defence", *The Journal of the Operational Research Society*, **38**(10), p875-882.

D. Nebhard & B. Norman, (2002), "Worker efficiency and responsiveness in cross- trained teams: Technical report", Madison: University of Wisconsin.

D. Nembhard & M. Uzumeri, (2000), "An Individual-Based Description of Learning within an Organisation", *IEEE Transactions on Engineering Management*, **47**(3), p370-378.

D. Nembhard, (2000), "The Effects of Task Complexity and Experience on Learning and Forgetting: A Field Study", *Human Factors: The Journal of the Human Factors and Ergonomics Society*, **42**(2), p272-286.

D. Nembhard, (2001), "Heuristic approach for assigning workers to tasks based on individual learning rates", *International Journal of Production Research*, **39**(9), p1955-1968.

D. Nembhard & B. Norman, (2007), "Cross Training in Production Systems with Human Learning and Forgetting", Chapter of *Workforce Cross Training*, US:Taylor & Francis

G. Nemhauser & L. Wolsey, (1999), "Integer and Combinatorial Optimisation", US:John Wiley & Sons.

Y. Ng, (2003), "Optimising a Military Supply Chain in the Presence of Random, Non-Stationary Demands", Master's Thesis, Naval Postgraduate School, Monterey.

P. Ngatchou, A. Zarei & A. El-Sharkawi, (2005), "Pareto Multi Objective Optimization", *Proceedings of the 13th International Conference on Intelligent Systems Application to Power Systems*, USA.

M. Nooriafshar, (1995), "A heuristic approach to improving the design of nurse training schedules", **81**(1), p50-61.

D. Oklolowski & S. Gawiejnowicz, (2010), "Exact and heuristic algorithms for parallel-machine scheduling with DeJong's learning effect", *Computers & Industrial Engineering*, **59**(2), p272-279.

B. Pagnoncelli, S. Ahmed & A. Shapiro, (2009), "Sample average approximation method for chance constrained programming: Theory and applications", *Journal of Optimization Theory and Applications*, **142**(2), p399-416

S. Park & G. Cha, (2013), "A Study on the Assessment of Key Competencies for Automotive Engineering Technology Education in Korea", *International Journal OF Engineering Education*, **29**(5), p1192-1198.

M. Pinzone et al. (2016), "A Multi-horizon, Multi-objective Training Planner: Building the Skills for Manufacturing", Chapter of: *Advances in Neural Networks: Computational Intelligence for ICT*, Switzerland:Springer International Publishing.

Policonomics, (2012), "Learning Curve", Retrieved 5 July, 2016, from `http://www.policonomics.com/learning-curve/`

J. Rekman et al. (2016), "A New Instrument for Assessing Resident Competence in Surgical Clinic : The Ottawa Clinic Assessment Tool", *Journal of Surgical Education*, **73**(4), p575-582.

M. Rezaei et al. (2012), "A heuristic method to schedule training programs for Small and Medium Enterprises", *European Journal of Operational Research*, **217**(3), p600-608.

R. Rudek, "Computational complexity of the single processor makespan minimization problem with release dates and job-dependent learning", *Journal of the Operational Research Society*, **65**, p1170-1176.

N. Safaei, D. Banjevic, & A. Jardine, (2011), "Workforce-constrained maintenance scheduling for military aircraft fleet: A case study", *Annals of Operations Research*, **186**(1), p295-316.

R. Sarker & C. Newton, (2008), "Optimization Modelling, A Practical Approach", US:Taylor & Francis Group.

S. Sayin & S. Karabati, (2007), "Assigning cross-trained workers to departments: A two-stage optimization model to maximize utility and skill improvement", *European Journal of Operational Research*, **176**, p1643-1658.

J. Shin et al. (2016), "Measuring Nontechnical Aspects of Surgical Clinician Development in an Otolaryngology Residency Training Program", *JAMA Otolaryngology-head & neck surgery*, **142**(5), p423-428.

J. Song & P. Zipkin, (1993), "Inventory Control in a Fluctuating Demand Environment", *Operations Research*, **41**(2), p351-370.

A. Soyster, (1973), "Convex Programming with Set-Inclusive Constraints and Applications to Inexact Linear Programming", *Operations Research*, **21**(5), p1154-1157.

I. Srour, C. Haas & D. Morton, (2006), "Linear Programming Approach to Optimize Strategic Investment in the Construction Workforce", *Journal of Construction Engineering and Management*, **132**(11), p1158-1166.

H. Stolley, (1969), "Models of Manpower Selection Procedure in the Armed Forces", paper presented at *Operational and Personnel Research in the Management of Manpower Systems*, Germany:NATO Science Committee.

S. Tarim & B. Kingsman, (2004), "The stochastic dynamic production/inventory lot-sizing problem with service-level constraints", *International Journal of Production Economics*, **88**, p105-119.

D. Thomas, B. Kwinn, M. McGinnis, B. Bowman,M. Entner, (1997), "The US Army enlisted personnel system: a system dynamics approach", paper presented at the *1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation.*

N. Thomopoulos,& M. Lehman, (1969), "The mixed-model learning curve", *AIIE Transactions*, **1**, p127-132

G. Thompson & J. Goodale, (2006), "Variable employee productivity in workforce scheduling", *European Journal of Operational Research*, **170**, p376-390.

G. Thompson & M. Pullman, (2007), "Scheduling workforce relief breaks in advance versus in real-time", *European Journal of Operational Research*, **181**, p139-155.

V. Thurner, B. Axel & K. Andreas,(2014), "Identifying Base Competencies as Prerequisites for Software Engineering Education", paper presented at *IEEE Global Engineering Education Conference*, p1069-1076.

V. Tiwari, J. Patterson & V. Mabert, (2009), "Scheduling projects with heterogeneous resources to meet time and quality objectives", *European Journal of Operational Research*, **193**(3), p780-790.

S. Vajda, (1972), "Probabilistic Programming", UK: Academic Press, Inc.

V. Valls, Á Péres & S. Quintanilla, (2009), "Skilled workforce scheduling in Service Centres", **193**, p791-804.

J. Wang, (2005), "A Review of Operations Research Applications in Workforce Planning and Potential Modelling of Military Training", *Technical Report*, Systems Sciences Laboratory: Australia, DSTO-TR-1688.

J. Wang, (2007), "A system dynamics simulation model for a four-rank military workforce", *Technical Report*, DSTO Defence Science and Technology Organisation: Australia, DSTO-TR-2037.

G. Winch, (1999), "Dynamic Visioning for Dynamic Environments", *The Journal of the Operational Research Society*, **50**(4), p237-243.

L. Wolsey, (1975), "Faces For a Linear Inequality in 0-1 Variables", *Mathematical Programming*, **8**, p165-178.

T. Wright, (1936), "Factors affecting the cost of airplanes", *Journal of the Aeronautical Sciences*, **3**(4), p122-128.

C. Wu, S. Huang & W. Lee, (2011), "Two-agent scheduling with learning consideration", *Computers and Industrial Engineering*, **61**(4), p1324-1335.

S. Yan, T. Yang & Y. Chen, (2004), "A model and a solution algorithm for airline line maintenance manpower supply planning with multiple aircraft type maintenance certificates", *Journal of the Chinese Institute of Engineers*, **27**(5), p719-729.

L. Yelle, (1976), "Estimating learning curves for potential products", *Industrial Marketing Management*, **5**, p147-154.