Optimising cricket team selection for the Indian Premier League

MF Alberts



orcid.org/ 0009-0000-0368-5082

Dissertation accepted in fulfilment of the requirements for the degree
Master of Engineering in Industrial Engineering
at the North-West University

Supervisor:	Prof RLJ Coetzer
Co-supervisor:	Ms C Bisset

Graduation: July 2026

"*We cannot create, we can only discover.*"
- Professor R.L.J. Coetzer,
Inaugural Lecture, 2024.

Acknowledgements

I have been privileged enough to receive tremendous support, guidance, and encouragement throughout this study. While challenging, this journey has greatly deepened my intellectual curiosity in the application of mathematical optimisation in sports. The completion of this dissertation would certainly not have been possible without the following individuals, whose unwavering belief in my work has been the sole source of motivation and inspiration. It is with sincere gratitude that I thank:

- My supervisors, Professor Roelof Coetzer and Chanel Bisset, for your encouragement, guidance, and assistance. Your continued commitment and steadfast mentorship have proven instrumental in shaping my research experience.
- My parents, Francois and Nienke Alberts, for your uncompromising love, compassion, and assistance. Your persistent support and encouragement of my passions laid the foundation for this study. It is with great honour that I dedicate this study to you.
- My brother, Gerhard Alberts, for being my constant source of laughter and comfort. Your confidence in my study has significantly contributed towards its finalisation and has instilled pride in the work presented.
- My companion, Heaven-lee Roelofse, for your continued support and encouragement. Your patience, understanding, and perseverance have played an invaluable role in the completion of this study.
- My friend, Hendrik Snyman, for your friendship, reassurance, and inspiration. Your support and advice proved invaluable during challenging stages of this study.

Above everything, I thank my Saviour, Jesus Christ, God my Father, and the Holy Spirit for the blessings bestowed upon me. It is He alone Who granted me the knowledge, talents, and discipline needed to complete this study. The completion of this study is only through His eternal love and grace.

“ [31] but those who hope in the Lord will renew their strength. They will soar on wings like eagles; they will run and not grow weary, they will walk and not be faint.”

- Isaiah 40:31, Holy Bible,
New International Version.

Publications and presentations

A significant portion of the work conducted in this study has been published and presented in pursuit of its completion. The subsequent research outputs relate to this study.

Southern African Institute for Industrial Engineering:

The paper "The application of optimisation techniques in cricket: A systematic literature review" was presented at the 34th Southern African Institute for Industrial Engineering conference in Vanderbijlpark in October 2024 and published in the proceedings. The paper investigated the considerations and shortcomings of previous research.

NWU-Purdue Joint Statistical Symposium & Workshop:

The academic talk "Utilising goal programming to optimise cricket team selection for the Indian Premier League" was delivered at the 1st NWU-Purdue Joint Statistical Symposium & Workshop in Potchefstroom in July 2025. The academic talk covered the fundamentals of this study, including the performance measures, goal programming models, and results.

Data Science, Statistics & Visualisation:

The presentation "Optimising cricket team selection for the Indian Premier League using goal programming" was delivered at the 9th Data Science, Statistics & Visualisation conference in Kruger National Park in July 2025. The presentation conveyed the fundamentals of this study, including the performance measures, goal programming models, and results.

South African Statistical Association:

The presentation "Applying goal programming to optimise cricket team selection for the Indian Premier League" was shared at the 66th South African Statistical Association conference in Vanderbijlpark in November 2025. The presentation conveyed the fundamentals of this study, including the performance measures, goal programming models, and results.

Abstract

The sport of cricket represents a complex bat-and-ball game predominantly played by the Commonwealth nations. Cricket has long since transitioned from an exclusive pastime to a multi-billion-dollar industry. The global recognition of cricket primarily stems from highly competitive Twenty20 (T20) leagues, which require rivalling franchises to execute effective team selection for success. Traditional team selection methods often involve considerable subjectivity due to several internal and external factors. Although ample research has been conducted on utilising different optimisation techniques to improve the objectivity of team selection in T20 leagues, various shortcomings exist. These include limitations in multiple objectives and constraints, squad selection, and the differentiation between auction types.

This study addresses these shortcomings by proposing a contemporary goal programming approach for optimising cricket team selection, using the 2022 Indian Premier League (IPL) mega auction as a case study. Applying a modified version of the scientific method as the research methodology, this study presents two solutions, formulated as non-preemptive and preemptive models. These solutions are developed, programmed, and solved using Elytica with the CPLEX optimisation solver. The proposed goal programming approach promotes team selection by balancing the trade-off between various conflicting and incommensurable cricketing abilities. The multifaceted nature of cricket allows players to specialise in unique cricketing abilities, such as bowling, batting, all-rounding, and wicketkeeping. This study extends existing research by presenting two new performance measures for batting and all-rounding. Utilising these statistics in conjunction with existing metrics for bowling and wicketkeeping gave performance scores for the 173 players considered for selection, which are scaled using Gaussian membership functions. The correctness of the formulation and coding for both models is verified by assessing the impact of each constraint on the results.

The importance of each cricketing ability and the corresponding number of players required for selection are computed objectively using an I-optimal crossed mixture design developed in Design-Expert. The non-preemptive goal programming model is validated by illustrating its effectiveness in selecting a team superior to the winning franchise of the 2022 IPL. The preemptive goal programming model is validated by illustrating its effectiveness in selecting a more specialised team relative to the non-preemptive model. This study contributes by presenting a unique, multipurpose approach to objective cricket team selection.

Keywords: cricket, goal programming, IPL, performance measures, team selection.

Table of contents

1 Introduction **1**

1.1 Background and rationale 1

1.2 Research problem, aim, and objectives 4

1.3 Research methodology 5

1.4 Chapter division 6

1.5 Contribution 7

1.6 Exclusions and delineations 9

2 Overview of cricket **10**

2.1 Origins of cricket 10

2.2 Gameplay and laws of cricket 11

2.3 Formats of cricket 12

2.4 Cricketing abilities and performance measures 13

2.4.1 Bowling 13

2.4.2 Batting 16

2.4.3 All-rounding 19

2.4.4 Wicketkeeping 24

2.5 Indian Premier League 25

2.6 Chapter summary 28

3 Overview of optimisation theory **29**

3.1 Introduction to optimisation theory 29
3.1.1 Background 29
3.1.2 Fundamentals of optimisation 30
3.1.3 Mathematical modelling 31
3.2 Linear programming 33
3.2.1 Origins of linear programming 33
3.2.2 Fundamentals of linear programming 34
3.2.3 Simplex method 37
3.3 Integer programming 41
3.3.1 Origins of integer programming 41
3.3.2 Fundamentals of integer programming 41
3.3.3 Branch-and-bound method 45
3.4 Goal programming 49
3.4.1 Origins of goal programming 49
3.4.2 Fundamentals of goal programming 50
3.4.3 Modified simplex method 58
3.5 Fuzzy logic 59
3.5.1 Origins of fuzzy logic 59
3.5.2 Fuzzy sets 59
3.5.3 Membership functions 62
3.6 Chapter summary 67

4 Mathematical models **68**

4.1 Model 1: Non-preemptive goal programming model 68
4.2 Model 2: Preemptive goal programming model 72
4.3 Chapter summary 73

5 Data considerations and verification **74**

5.1 Computational hardware 74

5.2 Data considerations 75

5.3 Verification 79

5.3.1 Model 1: Non-preemptive goal programming model 80

5.3.2 Model 2: Preemptive goal programming model 94

5.4 Chapter summary 99

6 Results and validation **100**

6.1 Performance measures assessment 100

6.1.1 Indexing of performance measures 100

6.1.2 Comparative analysis of combined batting measure 105

6.1.3 Comparative analysis of all-rounding measure 108

6.2 Optimisation results 111

6.2.1 Mixture design optimisation 111

6.2.2 Scenario analysis 114

6.3 Validation 119

6.3.1 Model 1: Non-preemptive goal programming model 119

6.3.2 Model 2: Preemptive goal programming model 122

6.4 Chapter summary 125

7 Conclusion **126**

7.1 Study summary 126

7.1.1 Overview 126

7.1.2 Objectives achieved 128

7.2 Limitations and recommendations 128

7.3 Practical considerations 130

A Systematic literature review **157**

B Personal and performance data **171**

C Comparison and experimental data **188**

List of figures

1.1 Modified scientific method [36] 6

2.1 Standard cricket field [7] 11

2.2 Two-dimensional framework [85] 18

2.3 Efficiency frontier [116] 21

3.1 Local and global optima [144] 31

3.2 Mathematical model classification [28] 33

3.3 Primal-dual relationship [163] 36

3.4 Simplex method [144] 38

3.5 Integer program [144] 42

3.6 Linear program [144] 42

3.7 Demonstration of P and $P^{'}$ [177] 43

3.8 Integer programming classification [169] 44

3.9 Complete search space [191] 45

3.10 Partitioned search space [191] 45

3.11 Branch-and-bound enumeration tree [157] 48

3.12 Integrality gap [197] 49

3.13 Goal programming solutions [29] 52

3.14 Infeasible linear program [28] 54

3.15 Feasible goal program [28] 54

3.16 Classical set boundary [242] 60

3.17 Fuzzy set boundary [242] 60

3.18 Mapping of fuzzy set $\underset{\sim}{A}$ [237] 61

3.19 Union of $\underset{\sim}{A}$ and $\underset{\sim}{B}$ [237] 62

3.20 Intersection of $\underset{\sim}{A}$ and $\underset{\sim}{B}$ [237] 62

3.21 Complement of $\underset{\sim}{A}$ [237] 62

3.22 Fuzzy inference system [249] 63

3.23 Membership function features [254] 63

3.24 Triangular membership function [254] 64

3.25 Trapezoidal membership function [254] 65

3.26 Gaussian membership function [272] 66

5.1 Player performance data collected 75

5.2 Membership values for bowers 77

5.3 Membership values for batsmen 77

5.4 Membership values for all-rounders 77

5.5 Membership values for wicketkeepers 77

5.6 Overview of data collection, cleaning, and transformation 78

5.7 Sub-process activities of data collection, cleaning, and transformation 79

6.1 Scatterplot of CBR and CBM 101

6.2 Scatterplot of $w_1 CBM^{'}$ and $w_2 CBR^{'}$ 103

6.3 99% confidence ellipse of all-rounders 104

6.4 Relationship between CBM ranking and BK ranking 107

6.5 Relationship between AR ranking and SB ranking 110

List of tables

1.1 Other T20 leagues [37] 8

2.1 IPL franchises [128] 25

2.2 2022 IPL group classification [128] 26

2.3 2022 IPL retention slabs [133] 27

3.1 Multi-criteria decision-making subareas [202] 50

4.1 Sets applicable to Model 1 68

4.2 Parameters applicable to Model 1 69

4.3 Constants applicable to Model 1 69

4.4 Variables applicable to Model 1 70

5.1 Player retainment characteristics. 80

5.2 Verification tests for Model 1 81

5.3 Results for Test 1 with Equation (4.1) included 81

5.4 Results for Test 1 with Equation (4.2) included 82

5.5 Results for Test 2 with Equation (4.12) included 83

5.6 Results for Test 3 with Equation (4.5) included 83

5.7 Players selected for Test 3 using Equation (4.5) 84

5.8 Results for Test 3 with Equation (5.1) included 85

5.9 Players selected for Test 3 using Equation (5.1) 85

5.10 Results for Test 4 with Equation (4.7) included 86

5.11 Players selected for Test 4 using Equation (4.7) 87

5.12 Results for Test 4 with Equations (4.8) to (4.11) included 89

5.13 Players selected for Test 4 using Equations (4.8) to (4.11) 89

5.14 Results for Test 4 with Equation (5.2) included 90

5.15 Players selected for Test 4 using Equation (5.2) 90

5.16 Results for Test 5 with Equation (4.3) included 91

5.17 Players selected for Test 5 using Equation (4.3) 92

5.18 Results for Test 5 with Equation (4.4) included 93

5.19 Players selected for Test 5 using Equation (4.4) 93

5.20 Preemptive priority factors for verification 94

5.21 Results for Scenario 1 of Model 2 95

5.22 Players selected for Scenario 1 of Model 2 95

5.23 Selection logic for Scenario 1 of Model 2 96

5.24 Results for Scenario 2 of Model 2 97

5.25 Players selected for Scenario 2 of Model 2 98

5.26 Selection logic for Scenario 2 of Model 2 98

6.1 Bowling data for identified batsmen 102

6.2 Top 20 run scorers of the 2003 ODI World Cup [289] 106

6.3 20 randomly selected all-rounders of the 2016 T20 World Cup [293] 109

6.4 Players retained by GT [302], [303] 112

6.5 Results for Run 2 of the mixture design 113

6.6 Players selected for Run 2 of the mixture design 114

6.7 Preemptive priority factors for scenario analysis 115

6.8 Results for Scenario 1 of scenario analysis 116

6.9 Players selected for Scenario 1 of scenario analysis 116

6.10 Results for Scenario 2 of scenario analysis 117

6.11 Players selected for Scenario 2 of scenario analysis 118

6.12 Players selected by GT [282] 120

6.13 Comparison of membership value totals 120

6.14 Ranking of remaining nine franchises 121

6.15 Results for validation of Model 2 123

6.16 Players selected for validation of Model 2 123

7.1 Research objectives achieved 128

A.1 Full-text summaries of quality assessment results 158

A.2 Full-text summaries of manual analysis results 167

B.1 Dataset 1 relating to bowlers [276], [282] 172

B.2 Dataset 2 relating to batsmen [276], [282] 175

B.3 Dataset 3 relating to all-rounders [276], [282] 177

(a) Bowling data for all-rounders 177

(b) Batting data for all-rounders 180

(c) Characteristic data for all-rounders 183

B.4 Dataset 4 relating to wicketkeepers [276], [282] 185

B.5 Verification subset [282] 186

C.1 Performance data for top 20 run scorers of the 2003 ODI World Cup [289] . 189

C.2 Performance data for 20 selected all-rounders of 2016 T20 World Cup [293] 190

(a) Bowling data for the 20 randomly selected all-rounders [294] 190

(b) Batting data for the 20 randomly selected all-rounders [295] 191

(c) AR data for the 20 randomly selected all-rounders 192

(d) SB data for the 20 randomly selected all-rounders 193

C.3 I-optimal crossed mixture design results 194

List of abbreviations

AHP	Analytical hierarchy process
BBC	British Broadcasting Corporation
BCCI	Board of Control for Cricket in India
BPL	Bangladesh Premier League
BPMN	Business process modelling notation
CSK	Chennai Super Kings
DC	Delhi Capitals
DEA	Data envelopment analysis
DH	Deccan Chargers
GL	Gujarat Lions
GT	Gujarat Titans
ICC	International Cricket Council
IPL	Indian Premier League
KKR	Kolkata Knight Riders
KTK	Kochi Tuskers Kerala
LPL	Lanka Premier League
LSG	Lucknow Super Giants
MI	Mumbai Indians
NP	Nondeterministic polynomial-time
ODI	One-Day International
PBKS	Punjab Kings
PWI	Pune Warriors India
RAM	Random access memory
RCB	Royal Challengers Bengaluru
RPS	Rising Pune Supergiant
RR	Rajasthan Royals
RTM	Right to Match
SLR	Systematic literature review
SRH	Sunrisers Hyderabad
T20	Twenty20

Chapter 1

Introduction

Contents

1.1 Background and rationale 1
1.2 Research problem, aim, and objectives 4
1.3 Research methodology 5
1.4 Chapter division 6
1.5 Contribution 7
1.6 Exclusions and delineations 9

1.1 Background and rationale

Cricket is a global game characterised by equality, built on the principles of sportsmanship, honour, and integrity [1], [2]. At its most basic, cricket denotes a bat-and-ball game played by various countries across distinct formats [1]. Historically, the sport enjoys an elite status due to its established history spanning multiple centuries [3]. However, cricket has long since transitioned from an exclusive pastime to a multi-billion-dollar industry [4]. Cricket assumes quasi-religious significance for select audiences [1], with approximately 2.5 billion supporters, trailing only football as the most popular sport globally [5].

Success in cricket relies on effective player selection and positioning [6]. Cricketers embody distinct strengths and weaknesses, allowing each player to fulfil a particular role within the team [7]. Cricketers perform differently under varying playing conditions, making selection even more crucial for assembling a balanced team that accommodates all facets of the game [8]. Traditional team selection methods often involve considerable subjectivity, influenced by numerous internal and external factors, including player bias, personal preference, social influence, and political pressure [6], [9], [10]. Lucrative salaries and endorsements encourage most national players to embrace cricket as their primary occupation [11]. Therefore, team selection that involves subjectivity poses a significant threat to the livelihoods of cricketers,

as contractual arrangements ultimately depend on performance [12]. Selecting cricketers based on subjectivity may lead to the inclusion of undeserving players, thereby diminishing the overall team strength [13].

The multifaceted nature of cricket enables players to specialise in unique cricketing abilities, including bowling, batting, all-rounding, and wicketkeeping [7]. The intricacy of comparing different cricketers significantly contributes to the subjectivity inherent in traditional team selection methods [14]. Lourens [15] highlighted the complexity of team selection by noting the non-comparability of the measures used to quantify player performance due to differing scales. Furthermore, the interpretation of the performance measures differs depending on the specific cricketing ability [16]. For example, lower bowling figures imply better bowling performance, while higher batting figures indicate better batting performance [16]. Thus, interdisciplinary comparisons between players with different expertise prove exceptionally challenging, especially considering the variety of cricketing abilities [17], [18].

Chand *et al.* [19] mentioned the conflicting nature of the different cricketing abilities, where increasing the selection of one ability requires a decrease in another. For example, selectors can establish a bowling-dominant team by reducing the strength of the batting department. Alternatively, selectors can opt for a batting-dominant team by weakening the strength of the bowling department. Bhattacharjee and Bhattacharjee [10] and Chakraborty *et al.* [13] highlighted the difficulty of selecting a team that suitably balances the various conflicting and incommensurable cricketing abilities. Chakraborty *et al.* [13] noted that team selection in cricket proves to be a complex decision-making problem influenced by several factors.

These factors that drive the complexity of decision-making in cricket notably contribute to the universal appeal of the sport. The global recognition of cricket predominantly stems from highly competitive Twenty20 (T20) leagues. Among these, the Indian Premier League (IPL) is recognised as the most distinguished [20]. The IPL contains ten teams, each owned by a particular franchise, competing for the annual title [21], [22]. Before the tournament begins, each franchise assembles a team by bidding on available players through an auction [22]. This auction can assume either a mini or mega status, with the latter occurring every three years [22]. Unlike mini auctions, mega auctions limit the number of players a specific franchise can retain from the previous season, forcing teams to transform and restructure [22]. Swartz [5] stated that the IPL auction has yet to be investigated in detail. Therefore, the current study focuses on team selection within T20 leagues, using the IPL auction as a case study. Specifically, the mega auction is of interest due to its fixed player retention policy. The 2022 IPL is identified as the specific tournament for team selection, being the most recent edition in which a mega auction took place before the initiation of this study.

The mega auction is a complex process with various stages, each containing uncertainties regarding player purchasing [22]. Cricketers subject to auctioning are nominated by their respective international boards and reviewed by the Board of Control for Cricket in India (BCCI) [22]. Players approved for auctioning select their base prices as the starting point for bidding [22]. Shortlisted players are then classified according to their cricketing abilities

and corresponding characteristics [22]. The purchasing process begins once the auctioneer announces the name of a player, with each bid increasing in fixed increments [22]. Subject to budgetary, auctioning, and proprietary constraints, franchises bet against each other to purchase players that would maximise their chance of winning [22]. However, most players often remain unsold during the initial stages of the auction [22]. Selected unsold players are subsequently reauctioned, allowing franchises with remaining positions and capital to strengthen their teams [22].

Team selection within the IPL auction is exceptionally subjective due to the bias inherent in purchasing prominent Indian and overseas players, who often have exceptional brand value regardless of their recent form [23]. Ahmed *et al.* [6] noted the complex task of selecting a team within a finite budget, further emphasising the challenges faced by franchises during the IPL auction. Acquiring players in such an auction environment proves challenging, as each franchise implements a different strategy [5], with constantly changing tactics due to varying circumstances. Furthermore, subtle betting contests often arise between franchise owners, introducing the possibility of overpaying or undervaluing participants.

Applying mathematical optimisation to select an impartial team proves crucial for maintaining an uncontested advantage in an ever-changing sport such as cricket. Alberts *et al.* [24] conducted a systematic literature review (SLR) investigating optimisation techniques employed in existing research for cricket team selection and batting order sequencing. The SLR [24] indicated that numerous authors have explored different optimisation techniques for optimal team selection, including integer programming [6], dynamic programming [25], genetic algorithms [26], and data envelopment analysis [27]. Appendix A lists the full-text summaries of the studies identified in the SLR by Alberts *et al.* [24].

The findings from the SLR revealed multiple gaps in research concerning the application of optimisation techniques for cricket team selection. First, the techniques applied in existing literature provide limited flexibility in the treatment of multiple objectives and constraints. Specifically, these techniques facilitate minimal consideration for the trade-off in selection between the cricketing abilities. Second, various studies have assessed team selection in the context of T20 leagues, with a particular focus on the IPL. However, most of these studies involve selecting the playing 11 as opposed to the universal squad of cricketers obtained from the auction. Third, earlier studies have failed to distinguish between mini and mega auctions, each of which stipulates unique team selection constraints. Many of these studies select teams from the baseline, neglecting important auctioning conditions, such as player retention, budgetary provisions, and nationality requirements.

Considering the limitations of previous research, this study proposes a contemporary goal programming approach to address the subjectivity associated with cricket team selection. Specifically, this study presents two solutions in the form of non-preemptive and preemptive goal programming models. Goal programming is an optimisation technique used to solve multi-objective optimisation problems by finding an acceptable trade-off between various, often conflicting goals [28]. Goal programming accurately measures the exact, under, and

overachievement of goals by applying deviational variables that capture the extent to which the target values are not attained [28]. Target values of goals define the desired performance level of the decision-maker, which may or may not be realised [28], [29]. Goal programming presents a flexible approach for solving multi-objective optimisation problems by allowing divergence from the target values, subject to achievement [28]. This flexibility enables goal programming to balance various incommensurable goals of equal or distinct priority [28].

The involvement of various internal and external factors leads to speculative team selection methods that comprise substantial subjectivity. This subjectivity often leads to an inferior selection of cricketers, negatively impacting player livelihoods and team performance. The best players must be selected to maximise the performance of the team. With considerable profits at stake and the difference between teams usually negligible, selectors cannot rely solely on personal intuition and prior experience to make informed decisions. Hence, this study proposes a unique goal programming approach that objectively selects a cricket team by balancing the trade-off between conflicting and incommensurable cricketing abilities.

1.2 Research problem, aim, and objectives

Traditional cricket team selection methods often involve considerable subjectivity, resulting in inferior player selections that negatively impact player livelihoods and team performance. Accordingly, the aim of this study is to propose a goal programming approach that enables superior and objective cricket team selection in the IPL. The objectives of the study are:

1. To provide a comprehensive overview of cricket, covering the following fundamentals:
    - Origins, gameplay, laws, and formats of cricket.
    - Cricketing abilities and performance measures.
    - IPL tournament and auction.
2. To present a descriptive overview of the optimisation theory applicable to this study, focusing on the following essential techniques:
    - Linear programming.
    - Integer programming.
    - Goal programming.
    - Fuzzy logic.
3. To propose new performance measures for batting and all-rounding, including comparisons with established metrics in the literature.
4. To propose non-preemptive and preemptive goal programming models that minimise the deviation from the target values, specifying the desired performance level of each cricketing ability.

5. To verify the correctness in formulation and coding for the non-preemptive and preemptive goal programming models.

6. To validate the effectiveness of practical cricket team selection for the non-preemptive and preemptive goal programming models.

7. To provide recommendations for future research and potential improvements relating to the work presented in this study.

1.3 Research methodology

The research methodology describes the systematic procedure used to conduct the research [30]. The current study proposes a modified version of the scientific method as the research methodology, adapted from Salminawati and Ningsih [31], Blystone and Blogett [32], and Wright [33]. The scientific method is a systematic, standardised approach that follows an iterative nature [34], allowing traversal between different stages, with numerous steps often repeated multiple times [35]. Salminawati and Ningsih [31] described the scientific method as a methodical process for gaining knowledge that involves distinct technical and working methods, procedures, and cognitive reasoning. The scientific method represents a critical, analytical, and objective approach, with findings supported by empirical evidence [31].

A core advantage of the scientific method involves its reliance on empirical experimentation and testing [32], [33]. It presents a simple way to identify errors and areas of improvement [33]. A key feature of this methodology is its ability to iteratively refine the understanding of the underlying factors that influence the problem under investigation [34]. Although the scientific method is based on empirical evidence [32], an essential drawback is its reliance on falsifiability [33]. This methodology can only be used to investigate testable phenomena, meaning that some experiences beyond objective observation cannot be tested. A further disadvantage is the ease with which the decision-maker can skew the test results to fit the perceived hypothesis [33].

The modified scientific method comprises five stages: problem identification, research, development, testing, and communication. The approach begins with problem identification, which characterises the essence of the study, underpinning its significance and contribution. The research stage entails synthesising existing literature to determine the considerations and limitations of prior work. Next, the development stage involves collecting and analysing the relevant data, followed by the development of the non-preemptive and preemptive goal programming models. Thereafter, the models are verified and validated during the testing stage, with the resultant findings reviewed to ensure the appropriateness of both solutions. The findings are presented during the communication stage, along with recommendations for future research. Figure 1.1 illustrates a process flowchart that outlines the systematic approach, specifically detailing the development and testing stages, while also highlighting the phase at which each research objective is addressed.

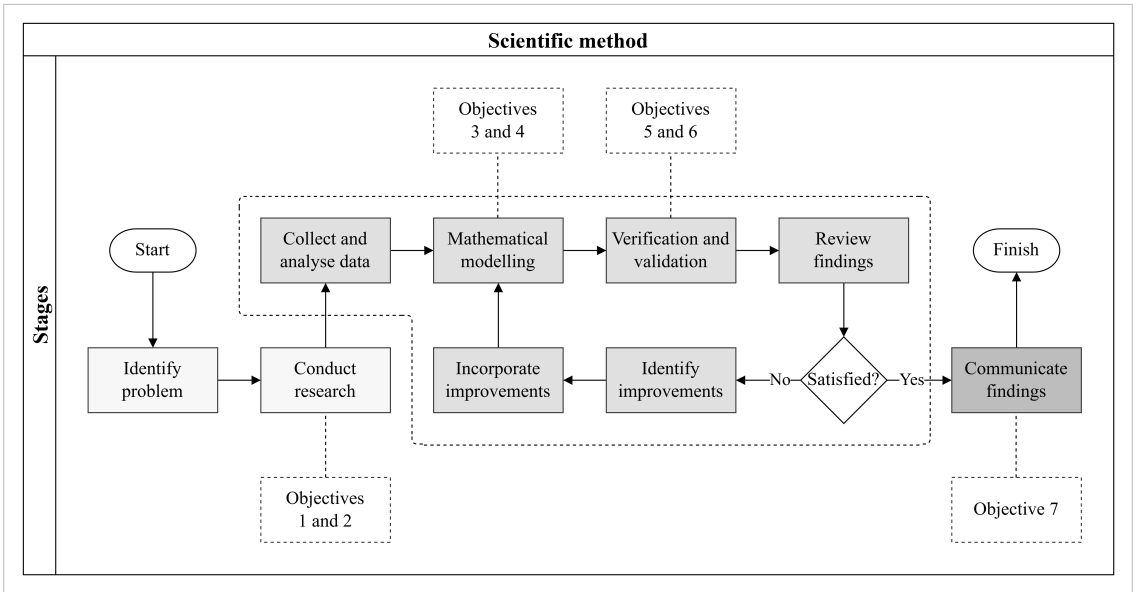


Figure 1.1: Modified scientific method [36]

1.4 Chapter division

This dissertation comprises seven chapters, a bibliography, and three appendices. Following the introduction in Chapter 1, a comprehensive overview of cricket is given in Chapter 2. The overview covers the fundamentals of cricket, including its origins, gameplay, laws, and formats. Four cricketing abilities, viz., bowling, batting, all-rounding, and wicketkeeping, are defined, accompanied by the performance measures often used in the existing literature to quantify player performance. In addition, new performance measures are presented for batting and all-rounding. Chapter 2 concludes by explaining the working principles of the IPL tournament and auction.

Chapter 3 presents a descriptive overview of the optimisation theory relevant to this study. Optimisation theory, as a facet of operations research, is explored by considering the fundamentals of optimisation and mathematical modelling. The origins, theory, and standard formulations of linear, integer, and goal programming problems are presented. In addition, the simplex, branch-and-bound, and modified simplex methods are described. Chapter 3 concludes by discussing fuzzy logic, including fuzzy set theory and membership functions.

Chapter 4 involves the development of the goal programming models. First, the formulation of the non-preemptive goal programming model is provided. Second, the formulation of the preemptive goal programming model is presented. The notation of both models is defined, explaining the purpose of each set, parameter, constant, and variable. Chapter 4 provides descriptive reasoning that articulates the functioning of the objective function, goals, and constraints associated with both models.

Chapter 5 specifies the data considerations and verification of the goal programming models. The computational hardware used in this study is defined to ensure the reproducibility

of the results. The procedure for data collection, cleaning, and transformation is discussed. Both models are verified by employing critical constraint investigations, which involve systematically evaluating the impact of each constraint on the results. Chapter 5 reveals that both models performed as expected, verifying their correctness in formulation and coding.

Chapter 6 presents the results and validation of the goal programming models. The accuracy of the new performance measures is assessed by identifying their classification capabilities and comparing their performance to existing metrics in the literature. An I-optimal crossed mixture design is constructed, along with a scenario analysis. The non-preemptive goal programming model is validated by demonstrating its effectiveness in selecting a team superior to the winning franchise of the 2022 IPL. The preemptive goal programming model is validated by demonstrating its effectiveness in selecting a more specialised team relative to the non-preemptive model. Chapter 6 reveals that both models performed accordingly, validating their effectiveness in practical cricket team selection.

Chapter 7 concludes the study by summarising the essential information from all preceding chapters. A comprehensive summary of the study is provided, highlighting its research and practical contributions. Chapter 7 describes the limitations of the study, accompanied by recommendations and potential improvements for future research.

Appendices A, B, and C provide additional information that supports the primary content of this study. Appendix A presents the full-text summaries of the SLR executed by Alberts *et al.* [24]. Appendix B details the personal and performance data of the players considered for selection. Appendix C specifies the comparison and experimental data relating to the performance measures assessment and the crossed mixture design, respectively.

1.5 Contribution

The SLR conducted by Alberts *et al.* [24] revealed that the optimisation techniques used in existing literature provide limited flexibility in team selection. The current study addresses this limitation by presenting two distinct goal programming models. The non-preemptive goal programming model balances the overall selection of cricketing abilities using weights that describe the importance of each. The preemptive goal programming model allows for the selection of specialist teams by prioritising cricketing abilities according to a predefined hierarchy. Alberts *et al.* [24] found that none of the authors employed a goal programming approach for team selection in cricket, rendering the work in this study unique. This study presents the first application of goal programming to cricket team selection.

Numerous authors have employed optimisation techniques for team selection in the IPL [24]. Alberts *et al.* [24] indicated that most of these studies focus on selecting the playing 11 rather than the universal squad of cricketers. Also, prior studies have failed to distinguish between the IPL mini and mega auctions. These findings support the claims of Swartz [5], who acknowledged that limited research has been conducted on team selection in the IPL

auction. The current study addresses these shortcomings by examining team selection in the 2022 IPL mega auction.

Despite its specific focus on the IPL, the work presented in this study applies to numerous T20 leagues. Table 1.1 lists some of the most popular T20 leagues that employ a similar auction policy for player acquisition. Both models can be expanded to the IPL mini auction with minor adjustments to the auction constraints. The application of the models is not limited to cricket and can be readily extended to other team sports. The flexibility of the proposed goal programming approach renders it ideally suited for team selection in sports that have various specialist positions.

Table 1.1: Other T20 leagues [37]

League	Country	Abbreviation	First season	Latest season
Bangladesh Premier League	Bangladesh	BPL	2012	2026
Lanka Premier League	Sri Lanka	LPL	2020	2025
SA20	South Africa	-	2023	2026

This study also contributes to the existing body of knowledge on performance measures in cricket by proposing new metrics for batting and all-rounding. Employing these metrics in conjunction with existing measures for bowling and wicketkeeping produced performance scores for all players considered. This study provides a method for capturing uncertainty in these performance scores by utilising Gaussian membership functions. Applying Gaussian membership functions enhances the precision and interpretation of performance scores by transforming them into membership values. Furthermore, this study employs an I-optimal crossed mixture design to objectively compute the importance of all cricketing abilities and the number of players required for each during selection. Using the crossed mixture design, this study proposes a distinct approach to objectifying decisions in cricket team selection that are inherently subjective.

Selecting a team that achieves an optimal trade-off between cricketing abilities proves to be a complex problem, requiring specific analytical techniques. This study demonstrates that the proposed goal programming approach bridges the gap between research and industry by overcoming the limitations of previous work, thereby promoting the objectivity of practical team selection in cricket. The contribution of this study is twofold, as it presents a unique goal programming approach applicable to research that addresses the practical problem of subjective team selection in cricket.

This study aligns closely with industrial engineering, as it applies mathematical optimisation to facilitate decision-making in cricket team selection. Industrial engineering involves structured problem-solving, efficiency, and improvement, all of which are illustrated in this study. Moreover, the combination of mathematics, systems thinking, and interdisciplinary collaboration in industrial engineering is shown by the integration of quantitative modelling and domain knowledge. This study demonstrates the significance of industrial engineering in sports analytics by improving existing processes through mathematical optimisation.

1.6 Exclusions and delineations

The following exclusions and delineations govern the extent of the research:

- This study focuses on selecting available cricketers from the 2022 IPL mega auction. Optimising the playing 11 for the selected cricketers exceeds the scope of this work.
- This study only considers the provisions of the 2022 IPL mega auction. Regulations and requirements related to other IPL auctions fall outside the scope of this work.
- This study examines four primary cricketing abilities: bowling, batting, all-rounding, and wicketkeeping. The cricketing ability of fielding is excluded due to the difficulty of quantifying its performance.
- This study does not differentiate between the various categories of all-rounders. The specialisation of all-rounders is not considered during team selection and falls beyond the scope of this work.
- This study only considers the T20 career statistics of auction participants, given its specific focus on the IPL. Data concerning One-Day Internationals (ODIs) and Test cricket are omitted.

Chapter 2

Overview of cricket

Contents

2.1 Origins of cricket 10
2.2 Gameplay and laws of cricket 11
2.3 Formats of cricket 12
2.4 Cricketing abilities and performance measures 13
2.5 Indian Premier League 25
2.6 Chapter summary 28

Chapter 2 provides a descriptive overview of cricket. It covers the fundamentals of cricket, including its origins, gameplay, laws, and formats. The chapter outlines the four cricketing abilities and the corresponding measures used to quantify player performance. Chapter 2 concludes by discussing the working principles of the Indian Premier League (IPL) and the regulations governing the 2022 mega auction.

2.1 Origins of cricket

Cricket is a global sport predominantly played by Commonwealth nations that significantly relies on physical fitness, mental strength, and strategic planning [38], [39]. The English are often credited with inventing cricket, although the exact origins of the sport are somewhat unknown [1], [40]. Some have mentioned that the sport has Celtic roots, first played by the Scottish and Irish around 500 AD [1]. Groombridge [41] and Pradhan *et al.* [42] described its potential origins in France, with Mote [43] recognising the possible ancestry of cricket in Iceland and Flanders. Regardless of the widespread speculation surrounding the origins of cricket, the earliest unambiguous reference to the sport dates back to 1597 [42].

Cricket is one of the oldest sports globally, played in over 100 countries [44], [45]. Jayawant and Kuttan [44] noted that cricket has recently experienced a significant surge in popularity.

Naha and Malcolm [46] primarily attributed this growth to the unique societal perspective presented by the cultural visibility and personal empowerment fostered by the game. Naha and Malcolm [46] further recognised that the growing interest in cricket involves a shift in the perceived character of the sport, with narratives of migration, governance, and politics emerging. The constant surge in popularity has helped cricket become a major sport with a global following [47].

2.2 Gameplay and laws of cricket

Cricket is an intricate game played between two teams, each comprising 11 members, with one substitution [48]. The game is played on a circular field containing a central 20-meter strip, known as the pitch [49]. The pitch confines a wicket at each opposing end, comprising two horizontal bails balanced on three stumps [49]. Surrounding the wicket at each end of the pitch is a crease, demarcated by white lines, which stipulates the legal area of play for both teams [50]. The field has an inner circle, which represents two semi-circles of radii 27 meters, surrounding the width of the pitch, converged by two parallel lines [7]. Figure 2.1 illustrates the configuration of a standard cricket field.

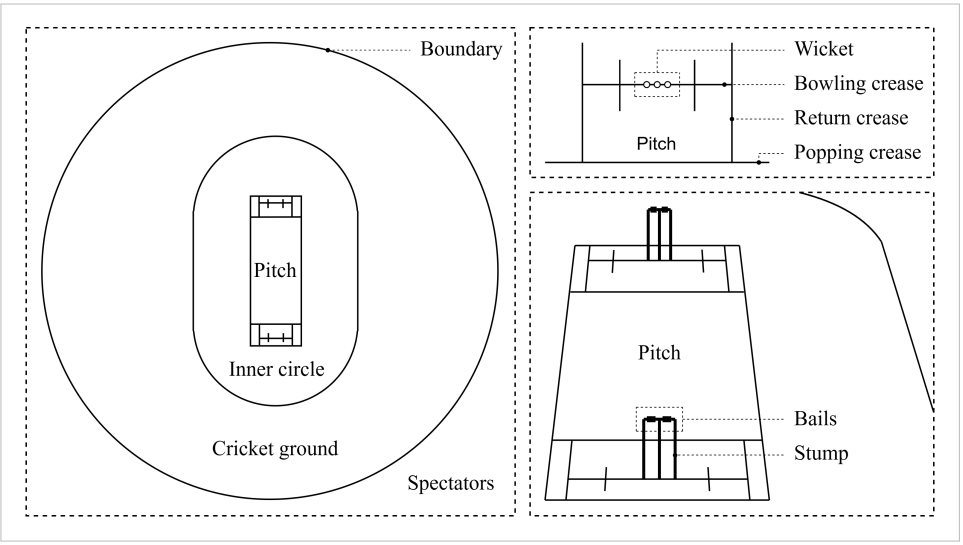


Figure 2.1: Standard cricket field [7]

International cricket games require two on-field umpires, a third umpire, a reserve umpire, and a match referee [51]. The on-field umpires impartially administer the match according to the laws of the game [48]. The third umpire represents an off-field referee who oversees the technical aspects of the game [52]. The third umpire uses technology to decide outcomes in inquiries observed by on-field umpires or players. The match referee guarantees that all participants uphold the spirit of the game by adhering to the International Cricket Council (ICC) Code of Conduct [52], [53].

The objective of cricket is to maximise the number of runs scored, such that it exceeds the total accumulated by the opposing team [49]. A coin toss, performed between the captains

of each side before the match begins, determines which team bats or bowls first [51]. The bowling team employs all eleven players to restrict the number of runs accumulated by the batting team [54]. In contrast, the batting team uses two batsmen to score runs, with the relationship between these players known as a partnership [55]. The goal of both batsmen is to score runs while defending their wickets [56]. Batsmen score runs by striking the ball and running to the crease located at the opposite end of the pitch before a player from the other team dislodges the bails of the wicket [57].

There is no limit on the number of runs a batsman can score off a specific delivery, except in the case of a boundary [58]. A boundary occurs when the ball travels beyond the boundary rope, resulting in either four or six runs [48]. Four runs are awarded to the batting team if the ball bounces at least once before crossing the boundary rope [48]. Six runs are obtained if the ball crosses the boundary rope without bouncing [48]. Teams obtain runs, known as extras, from law infringements committed by the bowling side [48]. Runs deemed as extras are only credited to the total team score and not to the runs of the on-strike batsman [59].

The period during which a particular team bats is called an innings [48]. The duration of an innings depends on the number of predefined overs stipulated by the match format [51], [56]. An over encompasses six legal deliveries [51], which involve propelling the cricket ball along the length of the pitch toward the wicket guarded by the batsman. The goal of the bowling side is to minimise the runs conceded to the batting team, with no bowler allowed to bowl two consecutive overs [49]. The bowling team can limit the number of runs scored by the batting team by either taking 10 wickets to end their innings or by completing the overs efficiently [48], [54]. The bowling team obtains wickets by dismissing the batsmen of the opposing team [48]. Once the bowling side has dismissed a cricketer, the player deemed out is replaced by the subsequent batsman in the lineup [60]. After one side has completed its innings, the other team has the opportunity to bat [54]. The team with the most runs after both sides have completed their innings is the winner [61]. A tie results if the scores of both sides are equal after the team batting second has completed its innings [62].

2.3 Formats of cricket

Malcolm and Fletcher [63] noted that perspectives on cricket as a global sport are polarised, even in nations where the game assumes quasi-religious significance. Throughout time, the public has criticised cricket for its complicated rules, slow pace, and prolonged nature [46]. Consequently, cricket has evolved into a multifaceted sport, with various forms catering to different spectators [51]. At the international level, cricket is played across three formats, distinguished by the number of overs, which are either limited or unlimited.

Unlimited overs cricket, also known as Test cricket, refers to games played without a limit on the number of overs per innings [64]. International Test cricket is played between nations that are full members of the ICC [64]. Test cricket represents the earliest form of the game, played over five days, with each team provided two separate opportunities to bat [51]. The

play for each day is typically limited to six hours, including official lunch and tea intervals of 40 and 20 minutes, respectively [64].

Limited overs cricket refers to formats that comprise a predefined number of overs, allowing play to conclude within a day [51]. Limited overs cricket relates to One-Day Internationals (ODIs) and Twenty20s (T20s) [65]. ODIs usually span a minimum of eight hours [66], with each team provided 50 overs to bat [51]. In contrast, T20s permit each team to bat for 20 overs, structured so that a game takes approximately three to four hours to complete [66]. Besides its faster completion, spectators often favour limited overs cricket over test cricket due to its superior entertainment value.

Both limited overs formats are relatively new compared to Test cricket [67]. The first ODI was played between England and Australia in 1971 as a replacement game for a cancelled Test match [68]. The first international T20 game took place in 2005, in which New Zealand suffered a 44-run loss to Australia [69]. Since its inception, interest in limited overs cricket, specifically T20 cricket, has grown significantly worldwide [70]. The T20 format has helped transform the game of cricket from a once-traditional and exclusive sport into an exciting and engaging experience for spectators [71], [72].

2.4 Cricketing abilities and performance measures

Section 2.4 describes the four cricketing abilities and the corresponding measures used to quantify player performance.

2.4.1 Bowling

Bowling is the act of delivering a cricket ball towards the batsman positioned at the opposite end of the pitch [48]. Shah [73] noted that the bowling average traditionally describes the quality of a bowler. The bowling average specifies the number of runs conceded per wicket obtained [16], with lower values indicating better bowling performance [74]. Equation (2.1) defines the bowling average AV_{BOW} [16], [75].

$$AV_{BOW} = \frac{\text{Number of runs conceded}}{\text{Number of wickets obtained}} \tag{2.1}$$

Other metrics exist for capturing bowling performance, including strike rate and economy rate [76]. The bowling strike rate represents the number of balls bowled per wicket obtained [77]. Lower strike rates indicate better bowling performance, as fewer deliveries are required to dismiss a batsman [78]. Equation (2.2) defines the bowling strike rate SR_{BOW} [16], [75].

$$SR_{BOW} = \frac{\text{Number of balls bowled}}{\text{Number of wickets obtained}} \tag{2.2}$$

The bowling economy rate is the number of runs conceded per over bowled [16]. Similar to the preceding bowling statistics, lower economy rates indicate better bowling performance [79]. Equation (2.3) defines the bowling economy rate ER_{BOW} [16], [75].

$$ER_{BOW} = \frac{\text{Number of runs conceded}}{\text{Number of overs bowled}} \tag{2.3}$$

Bhattacharjee and Saikia [76] mentioned that the traditional bowling metrics in Equations (2.1), (2.2), and (2.3) introduce different perspectives on performance. Equations (2.1) and (2.2) describe the wicket-taking capabilities of bowlers, whereas Equation (2.3) reveals the bowling expense [76]. Lemmer [80] acknowledged that although each measure has a unique significance, viewing these metrics independently can lead to a potential misperception of genuine bowling performance. For example, an economical bowler may be viewed as equal to an uneconomic bowler who has the same average [80]. Therefore, numerous authors have proposed different methods that utilise these traditional statistics to develop performance measures that capture general bowling abilities [76].

Kimber [81], van Staden [82], Barr *et al.* [83], and Saikia *et al.* [84] employed various visual methods to measure bowling performance. Kimber [81] constructed a simple yet effective contour display that simultaneously compares numerous players. Van Staden [82] extended this concept by providing hyperbolic contours that contain circles with radii corresponding to the number of wickets obtained. Barr *et al.* [83] studied the performance of bowlers in the 2007 ODI World Cup, building on the work of Barr and Kantor [85]. Saikia *et al.* [84] presented a tri-linear plot that evaluates the overall balance of bowlers using the Euclidean distance calculated from the centroid.

Beaudoin and Swartz [86] introduced a performance measure that captures the number of runs conceded relative to the total resources consumed. Gerber and Sharp [87] constructed an index to quantify performance using the bowling economy rate. Croucher [88] developed a statistic based on the principle of momentum, where the bowling average and strike rate represent mass and velocity, respectively. Cohen [89] measured bowling performance using the strike rate to determine the likelihood of dismissing the opposing team. Lemmer [80] proposed the combined bowling rate CBR, which is a performance measure that combines the bowling average, strike rate, and economy rate. Lemmer [90] revised CBR to consider the importance of each dismissal by assigning distinct weights to various wickets. Similarly, Shah [73] provided a performance measure based on the significance of the wickets obtained. Lemmer [91] presented a further metric that considers the different bowling conditions.

The CBR metric presented by Lemmer [80] is selected as the preferred measure for quantifying bowling performance due to its simplicity, consideration of career statistics, and ease of application. Saikia *et al.* [84] stated that the revised CBR provided by Lemmer [90] is more suitable for a small series of matches due to its data and computational requirements. The CBR metric utilises the harmonic mean to combine the traditional bowling statistics in Equations (2.1), (2.2), and (2.3) [80]. Lemmer [80] reported that the harmonic mean is

used to compute the average of multiple ratios, where the numerator is a constant and the denominator is variable. Hence, let r be the number of runs conceded, with w the number of wickets obtained and b the number of balls bowled [75], [84]. Equations (2.4), (2.5), and (2.6) provide the revised representations of the traditional bowling statistics [75], [84].

$$AV_{BOW} = \frac{r}{w} \tag{2.4}$$

$$SR_{BOW} = \frac{b}{w} \tag{2.5}$$

$$ER_{BOW} = \frac{r}{b/6} \tag{2.6}$$

Equations (2.4) and (2.6) show that the bowling average and economy rate share the same numerator, r, with variable denominators [80]. Lemmer [80] mentioned that the application of the harmonic mean requires converting Equation (2.5) into an equivalent expression that contains said numerator. Equation (2.7) depicts the conversion performed by Lemmer [80].

$$SR_{BOW} = \frac{b}{w} = \frac{b}{w} \times \frac{r}{r} = \frac{r}{rw/b} \tag{2.7}$$

Consequently, the traditional bowling statistics all share the same numerator, allowing the application of the harmonic mean. Equation (2.8) shows the computation of the harmonic mean H for an arbitrarily selected sample $X_1, X_2, ..., X_n$, as presented by Brettenny [75].

$$H = \frac{n}{\sum_{i=1}^{n} X_i^{-1}} \tag{2.8}$$

Applying Equation (2.8), the combined bowling rate is determined as the harmonic mean of the bowling average, strike rate, and economy rate. Equation (2.9) specifies the combined bowling rate CBR, as proposed by Lemmer [80].

$$CBR = \frac{3r}{w + b/6 + rw/b} \tag{2.9}$$

Lemmer [80] further constructed a classification framework to rank bowlers using the CBR statistic. The framework comprises ten different classes, showcasing bowlers of comparable quality [80]. Lemmer [80] determined the boundaries that distinguish the classes using the bootstrap technique to estimate the deciles of the data.

2.4.2 Batting

Batting involves using a willow bat to score runs while simultaneously defending a particular wicket [92]. Similar to bowling, Shah [73] stated that the batting average traditionally describes the quality of a batsman. Saikia *et al.* [84] acknowledged that the batting average is the most recognised performance measure in cricket. The batting average represents the number of runs scored per dismissal [93]. This metric assesses the reliability of a batsman, with higher values indicating better batting performance [94]. Equation (2.10) defines the batting average AV_{BAT} [75], [93].

$$AV_{BAT} = \frac{\text{Number of runs scored}}{\text{Number of times dismissed}} \tag{2.10}$$

Lemmer [9] mentioned that the batting average alone cannot fully capture the performance of a batsman. The rate at which a player scores runs must also be taken into consideration [9]. The batting strike rate has gained new significance due to the ever-growing interest in limited overs cricket, where success largely depends on the rate of scoring [85]. The strike rate describes the number of runs scored per 100 balls faced [93]. This metric assesses the scoring capabilities of batsmen, with higher values being preferred in limited-overs cricket [84]. Equation (2.11) defines the batting strike rate SR_{BAT} [75], [93].

$$SR_{BAT} = \frac{\text{Number of runs scored}}{\text{Number of balls faced}} \times 100 \tag{2.11}$$

The metrics provided in Equations (2.10) and (2.11) capture different batting competencies, showing distinct strengths and weaknesses [93]. Saikia *et al.* [84] noted that various cricket analysts have historically criticised these metrics for failing to accurately express the actual underlying abilities of batsmen. Therefore, several authors have employed these traditional statistics to obtain more suitable measures for quantifying batting performance [84].

Croucher [88] provided a batting index for assessing players in limited overs cricket, defined as the product of the batting average and strike rate. Gerber and Sharp [87] presented an index that relates the batting average of a single cricketer to the mean batting average of all batsmen. Kimber and Hansford [95] introduced a batting measure using a non-parametric approach, considering the total runs achieved. Barr and van den Honert [96] constructed a performance metric that combines the batting average and consistency of a player. Lemmer [97] extended this notion by providing a statistic that captures performance as the product of batting average, consistency, and strike rate. Lemmer [98] presented a metric focusing on a small series of matches, with an estimate of not-out scores. Lemmer [99] developed a method for comparing batsmen based on the conditions and importance of the runs scored.

Barr and Kantor [85] presented a performance measure for comparing batsmen, expressed as the weighted product of the batting average and strike rate. Borooah and Mangan [100] assessed batting performance using the Gini coefficient to determine playing averages, while

Damodaran [101] provided a Bayesian approach for estimating the conditional average of a batsman. Shah [73] presented a performance measure based on the batting average, taking into account the quality of the bowlers faced. Beaudoin and Swartz [86] examined batting performance using a ratio that describes the number of runs scored relative to the resources consumed. Lewis [14] introduced a similar technique for comparing batsmen based on the Duckworth-Lewis system. Saikia *et al.* [69] captured batting performance by considering the number of innings completed, average, strike rate, and percentage contribution. Van Staden [82] applied graphical displays to evaluate batsmen by introducing the survival rate as an additional statistic for performance analysis.

While various authors have used the traditional batting statistics in Equations (2.10) and (2.11) to develop performance measures, an investigation of the available literature shows that prior research has failed to recognise the applicability of the harmonic mean. Similar to bowling, the harmonic mean can be employed to quantify batting performance, as both the batting average and strike rate share the same numerator. Therefore, to the best of the author's knowledge, this study presents a new batting performance measure by extending the approach applied by Lemmer [80]. Let r denote the number of runs scored, with d the number of times dismissed and b the number of balls faced. Equations (2.12) and (2.13) provide the revised expressions of the traditional batting statistics.

$$AV_{BAT} = \frac{r}{d} \tag{2.12}$$

$$SR_{BAT} = \frac{r}{b/100} \tag{2.13}$$

Using Equation (2.8), the batting average and strike rate are combined into a single metric. Equation (2.14) specifies the combined batting measure CBM.

$$CBM = \frac{200r}{100d + b} \tag{2.14}$$

Establishing the accuracy of the measure proposed in Equation (2.14) involves comparing its results to those of an existing metric in the literature. The metric provided by Barr and Kantor [85] is selected for analysis because it employs a power coefficient α that describes the balance between batting average and strike rate. Employing this metric and adjusting α accordingly enables this study to investigate CBM against various scenarios, each with a unique importance assigned to batting average and strike rate. Barr and Kantor [85] noted that a proportional relationship exists between the batting strike rate and the probability of dismissal. Equation (2.15) specifies the probability of dismissal $P(out)$, as presented by Barr and Kantor [85]

$$P(out) = \frac{\text{Number of times dismissed}}{\text{Number of balls faced}} \tag{2.15}$$

Using the definitions for the strike rate and probability of dismissal, Barr and Kantor [85] proved that AV_{BAT} represents the quotient of SR_{BAT} and $P(out)$. Equation (2.16) shows the expression obtained by Barr and Kantor [85].

$$AV_{BAT} = \frac{SR_{BAT}}{P(out)} \tag{2.16}$$

Considering Equation (2.16), the batting characteristics of a player can be expressed in a two-dimensional space [85]. Figure 2.2 illustrates the two-dimensional framework, with the strike rate SR_{BAT} plotted on the y-axis and the probability of dismissal $P(out)$ represented on the x-axis. From Equation (2.16), rays intersecting the origin of Figure 2.2 characterise different batting averages, with players on the same line having an identical average. Figure 2.2 illustrates the two-dimensional representation of four arbitrary batsmen, denoted as A, B, C, and D.

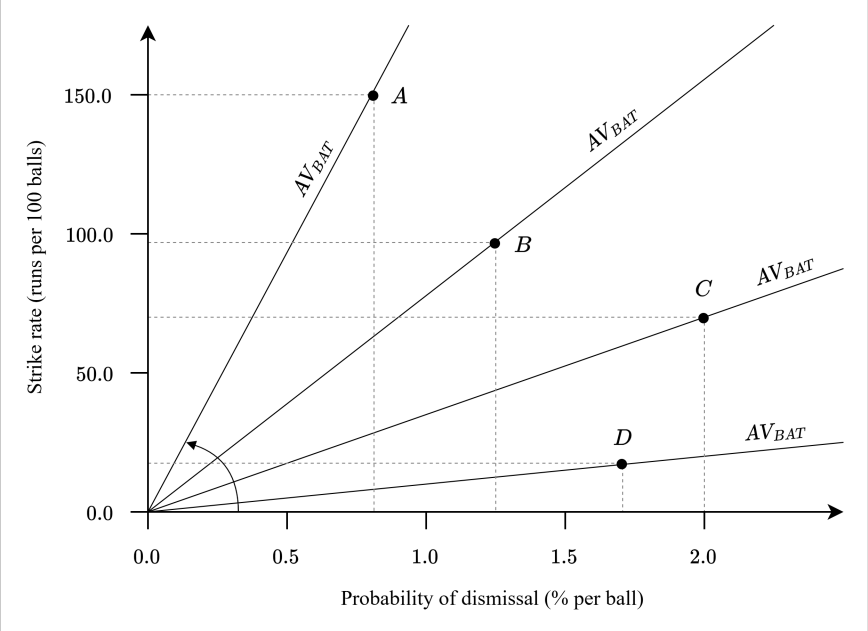


Figure 2.2: Two-dimensional framework [85]

The two-dimensional framework simultaneously captures the batting average, strike rate, and probability of dismissal [85]. From Figure 2.2, let y be the strike rate SR_{BAT} provided on the vertical axis and x the dismissal probability $P(out)$ depicted on the horizontal axis. The performance measure BK proposed by Barr and Kantor [85] represents the weighted product of the batting average and strike rate, as shown in Equation (2.17).

$$\begin{aligned} BK &= (SR_{BAT})^{\alpha}(AV_{BAT})^{1-\alpha} \\ &= y^{\alpha}(\frac{y}{x})^{1-\alpha} \\ &= \frac{y}{x^{1-\alpha}} \end{aligned} \tag{2.17}$$

From Equation (2.17), α represents the power coefficient that measures the balance between the batting average and strike rate [85]. Barr and Kantor [85] defined the feasible interval for α as $0 \leq \alpha \leq 1$. However, if $\alpha = 1$, the identity in Equation (2.17) reduces to $BK = y$, thereby eliminating the dependency on batting average as a performance statistic entirely. Therefore, the interval is revised as $0 \leq \alpha < 1$, with values for α reflecting the importance of the strike rate relative to the batting average. Lower values for α place less emphasis on the strike rate, whereas higher values place less emphasis on the batting average.

2.4.3 All-rounding

All-rounding involves regularly contributing to the team through both bowling and batting [102]. While teams expect all bowlers to bat and a select few batsmen to bowl, most players are considered specialists in only one of these abilities [102]. Wickramasinghe [102] stated that an all-rounder is a player whose bowling or batting expertise alone warrants selection. An all-rounder is sufficiently skilled to merit selection as both a bowler and a batsman, often referred to as a utility player [103]. Numerous authors have investigated the quantification of all-rounding performance using traditional bowling and batting metrics [84].

Suleman and Saeed [104] provided a performance index for capturing all-rounding abilities, described as the weighted combination of various bowling and batting factors. Gerber and Sharp [87] introduced an index based on the individual values of players, considering four categories of all-rounders. Tan and Ramachandran [105] presented a detailed mathematical expression defining all-rounding performance as a function of various bowling and batting statistics. Saikia *et al.* [54] provided a unique method for classifying all-rounders as either performers or non-performers. Van Staden [82] presented contour plots that depict bowling and batting performance, providing a holistic illustration of all-rounding capabilities.

Saikia and Bhattacharjee [103] presented a naïve Bayes classification model for examining different all-rounding classes. Wickramasinghe [102] classified all-rounders into four groups using naïve Bayes, k-nearest neighbours, and random forest. Sharp *et al.* [106] proposed an all-rounding measure representing the product function of individual bowling and batting indices. Lemmer [9] applied the metric of Sharp *et al.* [106], weighing bowling and batting equally using a power coefficient. Chakraborty and Ramakrishnan [13] used the Technique for Order of Preference by Similarity to Ideal Solution to compute the specific all-rounding scores of players.

Manage *et al.* [107] noted that all-rounders are viewed as specialist bowlers and batsmen. Hence, quantifying the performance of all-rounders requires measuring both their bowling and batting abilities [107]. The metrics in Equations (2.9) and (2.14) are collectively used to capture all-rounding performance. Quantifying the performance of all-rounders using the same metrics employed for bowling and batting ensures consistency and fairness in player comparisons. Despite the variety of measures provided in the existing literature, this study proposes a new metric that quantifies all-rounding performance using the CBR and CBM

statistics. The performance scores produced by the measures in Equations (2.9) and (2.14) are based on incompatible scales. Therefore, combining the CBR and CBM metrics into a single performance measure for all-rounders requires normalising each statistic to ensure comparability among the resulting scores.

Normalisation involves standardising features by transforming incommensurable data to a uniform scale expressed on the interval $[0, 1]$ [108], [109]. Normalisation removes the units of measurement, thereby ensuring comparability among performance scores [84]. Zirekgür and Karakaya [110] stated that min-max normalisation is widely regarded as one of the most prominent normalisation techniques. Min-max normalisation scales data by converting the minimum value to zero, the maximum value to one, and the remaining values to decimals between zero and one [110]. CBM is positively related to performance, with higher values indicating better batting. Therefore, CBM is normalised using Equation (2.18) [84].

$$CBM^{'} = \frac{CBM - CBM_{min}}{CBM_{max} - CBM_{min}} \tag{2.18}$$

From Equation (2.18), $CBM^{'}$ denotes the normalised value of CBM [111]. CBM_{min} and CBM_{max} represent the minimum and maximum values of the dataset, respectively [111].

CBR is negatively related to bowling performance, as lower values indicate better bowling. CBR is normalised using a reversed scale min-max normalisation, such that higher values signify better performance, ensuring comparability with $CBM^{'}$. CBR is normalised using Equation (2.19) [84].

$$CBR^{'} = \frac{CBR_{max} - CBR}{CBR_{max} - CBR_{min}} \tag{2.19}$$

From Equation (2.19), $CBR^{'}$ represents the scaled value of CBR [112]. CBR_{min} denotes the minimum value in the dataset, whereas CBR_{max} represents the maximum value [112].

Since normalisation produces unitless statistics [84], the standardised variables $CBM^{'}$ and $CBR^{'}$ are combined linearly through addition to obtain a descriptive measure that captures all-rounding performance. Equation (2.20) illustrates the composite measure AR, which is used to quantify the performance of all-rounders.

$$AR = w_1 CBM^{'} + w_2 CBR^{'} \tag{2.20}$$

From Equation (2.20), the weights w_1 and w_2 define the importance of batting and bowling, respectively, where $0 \leq w_1 \leq 1$ and $0 \leq w_2 \leq 1$. Selecting values for w_1 and w_2 according to personal preference proves to be subjective. This study determines the optimal combination of weights for each all-rounder using an efficient frontier, under the condition $w_1 + w_2 = 1$. Markowitz first proposed the efficient frontier theory in 1952 as a graphical representation

of financial portfolios that maximise the expected return for different risk levels [113], [114]. The utility of the efficient frontier has led to its application in various industries, including healthcare, education, and manufacturing [115].

Finding the optimal values for w_1 and w_2 involves constructing a two-dimensional plot, with $w_1CBM^{'}$ and $w_2CBR^{'}$ depicted on the x-axis and y-axis, respectively. The performance point P for each all-rounder is specified as the intercept of $w_1CBM^{'}$ and $w_2CBR^{'}$, where $w_1 = w_2 = 1$. An efficient frontier results for each all-rounder by varying the combination of w_1 and w_2, while keeping the values of $CBM^{'}$ and $CBR^{'}$ constant. The efficient frontier illustrates how AR changes for different values of w_1 and w_2. Figure 2.3 presents an efficient frontier, with each solution measured relative to the performance point P.

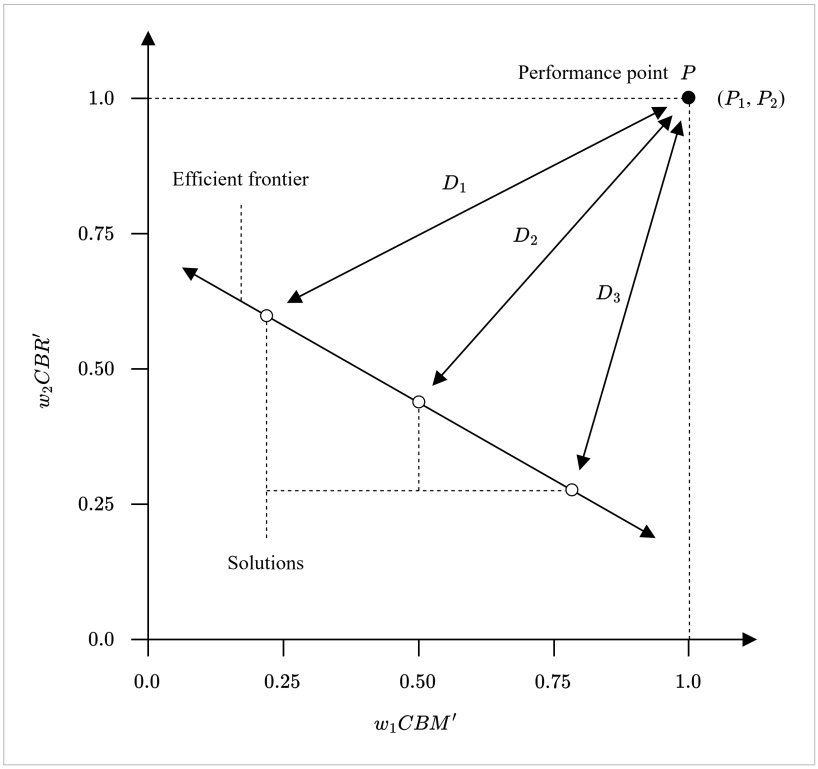


Figure 2.3: Efficiency frontier [116]

Considering Figure 2.3, the Euclidean distance D_i between the i-th solution on the efficient frontier and the performance point P is determined using Equation (2.21).

$$D_i = \sqrt{[P_1 - (w_1CBM^{'})_i]^2 + [P_2 - (w_2CBR^{'})_i]^2} \tag{2.21}$$

From Equation (2.21), $((w_1CBM^{'})_i, (w_2CBR^{'})_i)$ are the coordinates of the i-th solution, while (P_1, P_2) represent the coordinates of the performance point P. The combination of weights w_1 and w_2 that produces the solution point closest to P is considered optimal.

As before, establishing the accuracy of the measure presented in Equation (2.20) requires comparing its results to those of an existing metric in the literature. The metric proposed

by Sharp *et al.* [106] is selected for analysis because of its similarity to the statistic provided by Barr and Kantor [85] in Equation (2.17). Sharp *et al.* [106] constructed an all-rounding index using a power coefficient β to characterise the importance of the batting and bowling indices. Sharp *et al.* [106] applied the product measure of Barr *et al.* [83] as a criterion for batting. Equation (2.22) defines the product measure used to compute the batting ability of the i-th all-rounder [106].

$$U_{1i} = (AV_{BAT_i})^{\alpha}(SR_{BAT_i})^{1-\alpha} \tag{2.22}$$

Considering Equation (2.22), α denotes the power coefficient that captures the importance of the batting average and strike rate, where $0 \leq \alpha \leq 1$ [106]. Sharp *et al.* [106] employed the metric in Equation (2.22) to develop a batting index that quantifies the batting ability of the i-th all-rounder relative to all other all-rounders. Equation (2.23) defines the batting index for the i-th all-rounder [106].

$$c_{1i} = \frac{U_{1i}}{\sum_{i=1}^{m} U_{1i}} \times m \tag{2.23}$$

From Equation (2.23), m characterises the number of all-rounders [106]. Sharp *et al.* [106] adapted the product measure in Equation (2.22) as a criterion for bowling. Equation (2.24) shows the product measure used to compute the bowling ability of the i-th all-rounder [106].

$$U_{2i} = (ER_{BOW_i})^{\alpha_1}(AV_{BOW_i})^{\alpha_2}(SR_{BOW_i})^{1-\alpha_1-\alpha_2} \tag{2.24}$$

From Equation (2.24), α_1 and α_2 denote the power coefficients that specify the importance of the bowling economy rate, average, and strike rate, where $0 \leq \alpha_1 \leq 1$, $0 \leq \alpha_2 \leq 1$, and $0 \leq \alpha_1 + \alpha_2 \leq 1$ [106]. Sharp *et al.* [106] noted the dilemma of lower values for ER_{BOW_i}, AV_{BOW_i}, and SR_{BOW_i}, indicating better bowling performance compared to batting, which requires maximising the values of AV_{BAT_i} and SR_{BAT_i}. To address this, Sharp *et al.* [106] used the method of Gerber and Sharp [87] to transform the bowling performance measure. Equation (2.25) defines the bowling performance measure for the i-th all-rounder [106].

$$V_{2i} = k - \frac{U_{2i}}{\sum_{i=1}^{m} U_{2i}} \tag{2.25}$$

Considering Equation (2.25), k denotes the smallest positive integer [87] such that Equation (2.26) ensures that all bowling indices are positive [106].

$$k - \frac{U_{2i}}{\sum_{i=1}^{m} U_{2i}} > 0 \tag{2.26}$$

Subsequently, Equation (2.27) give the initial bowling index for the i-th all-rounder [106].

$$c_{2i}^{1} = \frac{V_{2i}}{\sum_{i=1}^{m} V_{2i}} \times m \tag{2.27}$$

To ensure equivalence in the spread of indices for batting and bowling, Sharp *et al.* [106] applied the scale adjustment of Lemmer [97]. The scale adjustment for the bowling index is determined by iteratively computing a new index. Equation (2.28) defines the expression of the $(p+1)$-th bowling index for the i-th all-rounder [106].

$$c_{2i}^{p+1} = c_{2i}^{p}{}^{\left(\frac{\sigma_{c_1}}{\sigma_{c_2^p}}\right)} \tag{2.28}$$

From Equation (2.28), σ_{c_1} and $\sigma_{c_2^p}$ denote the standard deviations of the batting index and the p-th iteration bowling index, respectively [106]. Equation (2.28) is employed iteratively until an acceptable level of accuracy is achieved [15]. Sharp *et al.* [106] reported that once an acceptable accuracy level is reached, the $(p+1)$-th bowling index is selected. Therefore, the expression in Equation (2.29) results [106].

$$c_{2i} = c_{2i}^{p+1} \tag{2.29}$$

Considering Equations (2.23) and (2.29), Sharp *et al.* [106] defined the all-rounding index as the weighted product of the batting and bowling indices. Equation (2.30) describes the initial all-rounding index for the i-th all-rounder [106].

$$c_{3i}^{1} = (c_{1i})^{\beta}(c_{2i})^{1-\beta} \tag{2.30}$$

From Equation (2.30), β is the power coefficient that defines the importance of the batting and bowling indices, where $0 \leq \beta \leq 1$ [106]. The same scale adjustment method used for bowling is used to spread the all-rounding index scales [106]. Again, the scale adjustment for the all-rounding index is gained by iteratively computing a new index. Equation (2.31) defines the expression of the $(p+1)$-th all-rounding index for the i-th all-rounder [106].

$$c_{3i}^{p+1} = c_{3i}^{p}{}^{\left(\frac{\sigma_{c_1}}{\sigma_{c_3^p}}\right)} \tag{2.31}$$

Considering Equation (2.31), $\sigma_{c_3^p}$ is the standard deviation of the p-th iteration all-rounding index [106]. Similar to bowling, Equation (2.31) is employed iteratively until an acceptable level of accuracy is achieved. Once an acceptable accuracy level is achieved, the $(p+1)$-th all-rounding index is selected. Adjusting the all-rounding index ensures its equivalence to

the batting and bowling indices [106]. In this study, an accuracy level of six decimal places proved sufficient for the scale adjustments involving the bowling and all-rounding indices. In the following content, the all-rounding index in Equation (2.31) is referred to as SB to ensure consistency with the other performance measures.

2.4.4 Wicketkeeping

Wicketkeeping involves catching deliveries while standing behind the wicket guarded by the batsman on strike [117]. Petersen [118] and Patel *et al.* [119] reported that wicketkeeping is arguably considered the most technical cricketing ability. Wicketkeepers play a vital role in the fielding performance of a team by taking on various responsibilities, such as limiting byes, completing dismissals, organising field placements, and capturing return throws [84]. Devising metrics capable of capturing wicketkeeping performance is challenging [84]. Saikia *et al.* [84] noted that a few authors have attempted to quantify wicketkeeping performance.

Gerber and Sharp [87] quantified the performance of wicketkeepers using an index based on the dismissal rate, defined as the average number of dismissals per game. Lemmer [120] captured wicketkeeping performance by combining the dismissal rate with a unique batting measure proposed in earlier work [97]. Lemmer [120] also demonstrated how the number of balls conceded per 1000 deliveries, identified as the bye rate, can be utilised as a third factor to capture wicketkeeping performance in Test cricket. Saikia *et al.* [84] extended the measure of Lemmer [120] to all formats, defining wicketkeeping performance as a function of batting ability, dismissal rate, and bye rate.

Saikia *et al.* [84] and Lemmer [120] cited the work of Narayanan [121] and Hemachandran [122], while the latter also recognised the research of Barry [123] and Samson [124]. The work of these authors was made available on online platforms and has since been removed. Therefore, the metrics provided by Narayanan [121], Hemachandran [122], Barry [123], and Samson [124] are summarised from the perspectives of Saikia *et al.* [84] and Lemmer [120]. Narayanan [121] constructed a points system to measure the performance of wicketkeepers, taking into account both the bye and dismissal rates [84], [120]. Hemachandran [122] ranked numerous wicketkeepers based on the product of their dismissal rate, team strike rate, and contribution to the total wickets taken [120]. Barry [123] described the bye rate as the total byes conceded per 600 deliveries [120]. Samson [124] introduced a metric that captures the performance of wicketkeepers by considering their batting average, total byes conceded per 100 balls, and the contribution of dismissals as a percentage of team wickets [120].

MacDonald *et al.* [125] mentioned that wicketkeeping denotes a specialist cricketing ability in its own right. Therefore, this study considers wicketkeeping and batting as two separate cricketing abilities, each measured independently. Numerous authors have utilised the bye rate as a statistic to measure wicketkeeping performance. Barry [123] recognised that the bye rate heavily relies on the standard of bowling and is not necessarily representative of wicketkeeping performance [120]. Lemmer [120] discussed that reliable data specifying the

number of byes conceded is not readily available. Hence, in this study, the performance of wicketkeepers is measured using a dismissal rate, defined as the number of stumpings and catches completed per innings kept. Equation (2.32) specifies the dismissal rate WK used to quantify wicketkeeping performance [87], [120].

$$WK = \frac{\text{Number of catches completed} + \text{Number of stumpings completed}}{\text{Number of innings kept}} \tag{2.32}$$

The dismissal rates proposed by Saikia *et al.* [84] and Lemmer [120] take into consideration the significance of each wicket obtained, using weights. These measures prove sufficient for capturing the performance of wicketkeepers in a small number of matches. However, since this study employs the career statistics of players, the dismissal rate in Equation (2.32) is applied to quantify wicketkeeping performance.

2.5 Indian Premier League

The T20 format presents a unique opportunity for players to earn substantial wages while gaining experience through participation in league competitions alongside other domestic and international cricketers [23]. The IPL is regarded as the most lucrative T20 league [20] and is among the most-watched sporting events globally [126]. As of 2023, The Economic Times values the IPL at approximately $10.7 billion, representing a 433% growth since its inaugural 2008 season [127].

Annually, various capitalists invest in the IPL due to its perpetual growth, global exposure, and international audience [22]. The IPL is a franchise competition, where private corporations bear ownership of teams [128]. Since its inception, 15 franchises have competed in the IPL [128], each representing an Indian city [22]. The 2022 IPL marked the 15th edition of the contest, which saw the addition of two new teams, allowing 10 franchises to compete for the championship [22]. Table 2.1 lists the 15 franchises, along with their abbreviations, years of engagement, and playing status.

Table 2.1: IPL franchises [128]

Franchise name	Abbreviation	First season	Latest season	Active
Mumbai Indians	MI	2008	2025	✓
Royal Challengers Bengaluru	RCB	2008	2025	✓
Chennai Super Kings	CSK	2008	2025	✓
Kolkata Knight Riders	KKR	2008	2025	✓
Delhi Capitals	DC	2008	2025	✓
Rajasthan Royals	RR	2008	2025	✓
Punjab Kings	PBKS	2008	2025	✓
Deccan Chargers	DH	2008	2012	✗

Kochi Tuskers Kerala	KTK	2011	2011	✗
Pune Warriors India	PWI	2011	2013	✗
Sunrisers Hyderabad	SRH	2013	2025	✓
Gujarat Lions	GL	2016	2017	✗
Rising Pune Supergiant	RPS	2016	2017	✗
Gujarat Titans	GT	2022	2025	✓
Lucknow Super Giants	LSG	2022	2025	✓

Traditionally, the IPL follows a standard round-robin approach, allowing each franchise to play against every other franchise twice, with the top four teams qualifying for the playoffs [22]. However, the 15th edition of the competition followed a pooled round-robin format, where the active teams in Table 2.1 were categorised into two groups [128]. Each franchise played twice against the teams in the same group, twice against the team in the same row of the opposing group, and once against all the remaining teams [128]. Thereafter, the top two teams in each group advanced to the playoffs for elimination [22]. Table 2.2 illustrates the group classification for the 2022 IPL.

Table 2.2: 2022 IPL group classification [128]

Group A		**Group B**	
Number	**Franchise name**	**Number**	**Franchise name**
1	Mumbai Indians	1	Chennai Super Kings
2	Kolkata Knight Riders	2	Sunrisers Hyderabad
3	Rajasthan Royals	3	Royal Challengers Bengaluru
4	Delhi Capitals	4	Punjab Kings
5	Lucknow Supergiants	5	Gujarat Titans

Franchises construct their squads by purchasing available players through an auction [22]. The auctioning procedure involves acquiring participants from a collection of overseas and Indian players [22]. Indian players are classified as either uncapped or capped [22]. Capped Indian cricketers are individuals who have previously played for the Indian national team, whereas uncapped players represent those who have yet to make their international debut [22]. Overseas players are foreign cricketers who represent rival national teams [22]. In this study, the terms team, squad, side, and franchise are used interchangeably.

IPL auctions are classified as mini or mega, depending on the specific player retention rules [22]. Mini auctions permit franchises to release or retain any number of players [22]. Mini auctions are held annually to assist franchises in addressing minor deficiencies identified in their existing squad compositions [22]. Conversely, mega auctions occur every three years, with retentions determined by the number of shortlisted participants [22]. Mega auctions facilitate team transformation and recreation by limiting the number of cricketers retained through direct retention and Right to Match (RTM) cards [22].

Direct retention allows franchises to retain a predefined number of players from the previous season without needing to repurchase them during the auction [22]. RTM cards enable sides to retain cricketers during the auction by matching the best offers proposed by competing franchises [22]. To ensure fair treatment of new teams, organisers omitted the RTM cards for the 2022 IPL mega auction [129]. The eight franchises that competed in the 2021 season were restricted to four direct retentions, including a maximum of three Indian players, two uncapped Indian players, and two overseas players [130]. Organisers permitted the two new franchises, GT and LSG, to select up to three players each before the auction from among the pool of available participants [130]. These selections included a maximum of two Indian players, one uncapped Indian player, and one overseas player [130]. The cricketers selected by GT and LSG before the auction are also referred to as retained players for consistency.

The Board of Control for Cricket in India (BCCI) enforces a fair auction by restricting the budget of each franchise, thereby eliminating the financial advantage of teams with greater resources [22]. For the 2022 IPL mega auction, each franchise had a purse of ₹90.00 crore, from which a minimum of ₹67.50 crore needed to be spent [22], [131]. In the case where a franchise assembled a team for a cost lower than the minimum spend requirement of ₹67.50 crore, the difference was paid to the BCCI as a penalty [130], [132]. A specific amount was deducted from the purse and the minimum spend requirement of the franchise, depending on the number of players retained [133]. Table 2.3 details the retention slabs enforced prior to the 2022 IPL mega auction.

Table 2.3: 2022 IPL retention slabs [133]

Retentions	Player 1	Player 2	Player 3	Player 4	Total
1	₹14.00 crore	-	-	-	₹14.00 crore
2	₹14.00 crore	₹10.00 crore	-	-	₹24.00 crore
3	₹15.00 crore	₹11.00 crore	₹7.00 crore	-	₹33.00 crore
4	₹16.00 crore	₹12.00 crore	₹8.00 crore	₹6.00 crore	₹42.00 crore

Although Table 2.3 outlines the retention slabs applicable to the 2022 IPL, franchises were allowed to pay retained players any sum mutually agreed upon [130]. The highest amount between the retention slab and the agreed payment was deducted from the purse, as well as the minimum spend requirement of the franchise [130]. For example, if a franchise retained two players following the retention slabs, ₹14.00 crore was allocated to the first and ₹10.00 crore to the second, resulting in a total deduction of ₹24.00 crore. The amount deducted to retain an uncapped player was ₹4.00 crore, regardless of the affiliated retention slab or the agreed-upon payment [130]. Retaining uncapped players allowed franchises to strategically conserve funds for the auction while also investing in emerging talent.

Utilising the available purse, franchises needed to construct a squad ranging from 18 to 25 players, including the retained cricketers [134]. Each side was allowed a maximum of eight overseas players [134] and was required to select at least 17 Indian cricketers [135]. There were no restrictions imposed on the number of specialist players a franchise could acquire, allowing teams the freedom to select any preferred combination of cricketing abilities [134].

Cricketers approved for auctioning select their base prices as the starting values for bidding [22]. Base prices incrementally range between ₹0.20 and ₹2.00 crore [22]. Shortlisted players are classified into sets based on their cricketing abilities and experience [22]. Traditionally, the auction commences with a collection of marquee players representing elite performers in T20 cricket [22]. Thereafter, sets of Indian and overseas bowlers, batsmen, all-rounders, and wicketkeepers are auctioned [22]. Repeatedly, the different franchises bet against each other to acquire players, with multiple participants often remaining unsold [22]. Selected unsold players are reauctioned through an accelerated bidding procedure, allowing franchises with remaining resources to strengthen their sides [22].

2.6 Chapter summary

Cricket is one of the oldest sports in the world [44], denoting a complex bat-and-ball game played between two teams on a circular field [48]. At the international level, teams compete in three formats: Test, ODI, and T20 [136]. The multifaceted nature of cricket enables sides to select players with diverse cricketing abilities, including bowling, batting, all-rounding, and wicketkeeping [7]. Therefore, multiple related metrics are proposed to quantify player performance, defined as CBR, CBM, AR, and WK. T20 leagues such as the IPL enable players with distinct cricketing abilities to showcase their potential [23]. The IPL is widely regarded as the most prosperous T20 league [20], with various franchises competing for the championship [22]. Annually, the IPL allows franchises to construct a squad of players by purchasing available cricketers through an auction [22]. IPL auctions are classified as mini or mega [22], with the current work focusing on team selection in the latter using the 2022 season as a case study. Subsequently, Chapter 3 presents an overview of the optimisation theory applicable to this study.

Chapter 3

Overview of optimisation theory

Contents

3.1 Introduction to optimisation theory 29
3.2 Linear programming 33
3.3 Integer programming 41
3.4 Goal programming 49
3.5 Fuzzy logic 59
3.6 Chapter summary 67

Chapter 3 introduces optimisation theory, describing the mathematical concepts necessary to understand the succeeding content. The chapter specifies the theoretical principles and standard formulations involving linear, integer, and goal programming. Finally, fuzzy logic is discussed, including fuzzy set theory and the mechanics of membership functions utilised for numerical transformation.

3.1 Introduction to optimisation theory

Section 3.1 defines the background covering operations research and mathematical optimisation, followed by the fundamentals of optimisation and mathematical modelling.

3.1.1 Background

Constantly changing social, economic, and political factors have significantly increased the complexity of modern decision-making [28]. The persistent changes brought about by the Fourth Industrial Revolution exert additional pressure on adjudicators to draw dependable conclusions under varying conditions [137]. Decision-makers cannot merely rely on personal intuition and prior experience to make informed decisions [28]. Therefore, using scientific

approaches, such as operations research, to support decision-making proves essential as it presents quantitative evidence that sustains findings [28].

Operations research originated during the Second World War as an approach for managing scarce resources [28]. During World War II, various operations research groups were formed to solve sophisticated military problems encountered by land, sea, and air expeditions [138]. The solutions produced by these groups altered strategic decision-making, eventually aiding the Allies in winning the war [138]. Following the war, researchers observed the significant wartime contributions of operations research and realised its potential to solve managerial and functional problems of non-military establishments [138]. As such, operations research has since advanced to numerous prominent fields, such as logistics, healthcare, scheduling, production planning, and transportation [139]. Operations research is regarded as a leading methodology approach for decision-making, with notable applications in various business, industrial, and managerial environments [138].

Operations research is a broad field comprising various components, including optimisation theory, of which mathematical optimisation represents a subset concerned with finding the best solution subject to several constraints [140]. Mathematical optimisation is one of the most advanced branches of operations research, especially concerning strategic and tactical incentives [141]. Sinha [142] characterised mathematical optimisation as a decision-making approach involving effectively using limited resources to accomplish some desired objective. Mathematical optimisation altered practical decision-making by promoting the efficiency of adjudicators through an improved understanding of the effects of selected outcomes [141].

3.1.2 Fundamentals of optimisation

Optimisation theory represents the science of discovering the extrema of a specific function of interest [143]. Luies [144] characterised optimisation as minimising or maximising some real-valued function by identifying the best input from a permissible set δ. Provided some function $f : \delta \rightarrow \mathbb{R}$, if the element $\boldsymbol{x}^* \, \epsilon \, \delta$ is selected such that $f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$, for all $\boldsymbol{x} \, \epsilon \, \delta$, the problem is of type minimisation [144]. Similarly, if the element $\boldsymbol{x}^* \, \epsilon \, \delta$ is selected such that $f(\boldsymbol{x}^*) \geq f(\boldsymbol{x})$, for all $\boldsymbol{x} \, \epsilon \, \delta$, the problem is of type maximisation [144]. Optimisation is the technique through which the optimal solution to a problem is obtained, given a finite set of predefined conditions [145].

Subhakala *et al.* [146] noted that optimisation can be classified as local or global, depending on the particular region where the problem conditions exist. A local optimum contains an extreme function value compared to alternative points within some remote neighbourhood [146]. The point $\boldsymbol{x}^*$ is a local minimum for some $\epsilon > 0$ such that, for all $\boldsymbol{x} \, \epsilon \, \delta$, in a distance ϵ of $\boldsymbol{x}^*$, $f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$ [144]. Conversely, the point $\boldsymbol{x}^*$ is a local maximum for some $\epsilon > 0$ such that, for all $\boldsymbol{x} \, \epsilon \, \delta$, in a distance ϵ of $\boldsymbol{x}^*$, $f(\boldsymbol{x}^*) \geq f(\boldsymbol{x})$ [144]. A global optimum adopts an extreme function value compared to all points across the entire search space [146]. The point $\boldsymbol{x}^*$ is the global minimum if $f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$, for all $\boldsymbol{x} \, \epsilon \, \delta$, and the global maximum if

$f(\boldsymbol{x}^*) \geq f(\boldsymbol{x})$, for all $\boldsymbol{x} \, \epsilon \, \delta$ [144]. Luies [144] highlighted that distinguishing between local and global optima is crucial since certain solving algorithms cannot guarantee optimality. Figure 3.1 illustrates the local and global optima for some arbitrary function $f(x)$.

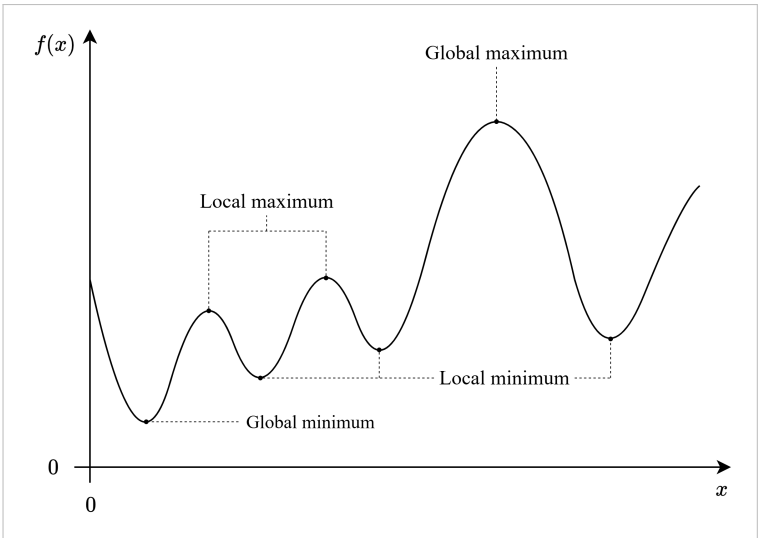


Figure 3.1: Local and global optima [144]

3.1.3 Mathematical modelling

Mathematical optimisation seeks an optimal solution by selecting the best result between multiple alternatives subject to several constraints [28]. Sharma [28] described an optimal solution as a feasible solution that results in the minimum or maximum objective function value. A feasible solution is a collection of feasible points that satisfy all model constraints [147]. Mathematical optimisation uses optimisation models, which are analytical methods, to solve intricate decision-making problems [28]. Williams [148] mentioned the importance of such models by signifying their proficiency in revealing underlying relationships between entities that would otherwise not be evident. Sharma [28] specified the following provisions for formulating a mathematical model:

- Decision variables: Denote controllable factors whose values are obtained by solving the specific problem.
- Objective function: Represents the criterion utilised to investigate alternatives based on different decision variable values to minimise or maximise performance.
- Constraints: Signify the policies and rules that restrict the use of resources, expressed as equalities or inequalities.

Mathematical models are used to analyse and understand the fundamentals of a particular system to improve its general performance [28]. Mathematical models describe functional relationships, which connect the decision variables in the objective function to those in the

model constraints [28]. A model is deemed linear if all functional relationships among the decision variables are linear [28]. Inversely, a model is considered non-linear if at least one functional relationship is non-linear [28]. Equations (3.1), (3.2), and (3.3) characterise the mathematical model of a general decision-making problem, as formulated by Sharma [28], where optimise represents either minimisation or maximisation.

$$\text{Optimise} \quad Z = f(\boldsymbol{x}) \tag{3.1}$$

$$\text{Subject to} \quad h_i(\boldsymbol{x}) \; \{\leq, =, \geq\} \; b_i, \quad \forall i = 1, 2, ..., m \tag{3.2}$$

$$\boldsymbol{x} \in \mathbb{R}_+ \tag{3.3}$$

From Equations (3.1), (3.2), and (3.3), $\boldsymbol{x}$ is the vector of decision variables $x_1, x_2, ..., x_n$, with $f(\boldsymbol{x})$ denoting the objective function for minimisation or maximisation [28]. Moreover, $h_i(\boldsymbol{x})$ represents the left-hand side of the i-th constraint and b_i the corresponding constraint requirement, where $i = 1, 2, ..., m$ [28].

Sharma [28] discussed that mathematical models are classified according to their behaviour characteristics, degrees of certainty, and solution approaches. Models are considered static or dynamic, depending on their time reference [28]. Static models represent the behaviour of a system at a particular point, neglecting changes introduced over time [28]. Conversely, dynamic models consider system state changes by treating the period as a separate variable used to determine the optimum solution [28]. Models are also categorised as deterministic or stochastic, depending on their degrees of certainty [28]. Deterministic models are those in which the parameters and functional relationships are known with certainty [28]. If the model encompasses at least one random variable, it is said to be stochastic [28]. Stochastic models better mimic practical occurrences as they consider unpredictability [28].

Sharma [28] also noted that mathematical models are normative, predictive, or descriptive, depending on their specific purpose. Normative models yield an optimal solution following a suitable course of action [149]. Predictive models estimate or forecast future events [149]. Descriptive models describe different decision processes or systems [149], typically applied to determine the severity of a problem [28]. Models are further classified according to their solution approaches: analytical and simulation [28]. Unlike simulation models, analytical models are solved using standard mathematical methods [28]. Simulation models represent systems that cannot be expressed mathematically [28].

Supporting the classification of Sharma [28], Tedeschi [150] recognised an additional categorisation that describes the time continuity of models. Tedeschi [150] acknowledged that mathematical models can be considered discrete or continuous. Discrete models necessitate variables that form part of a discrete set [151]. Alternatively, continuous models permit a continuous range of values for the related model parameters [152]. Figure 3.2 demonstrates the classification of mathematical models, adapted from Sharma [28]. Figure 3.2 includes the additional classification specified by Tedeschi [150], which is defined as the data type.

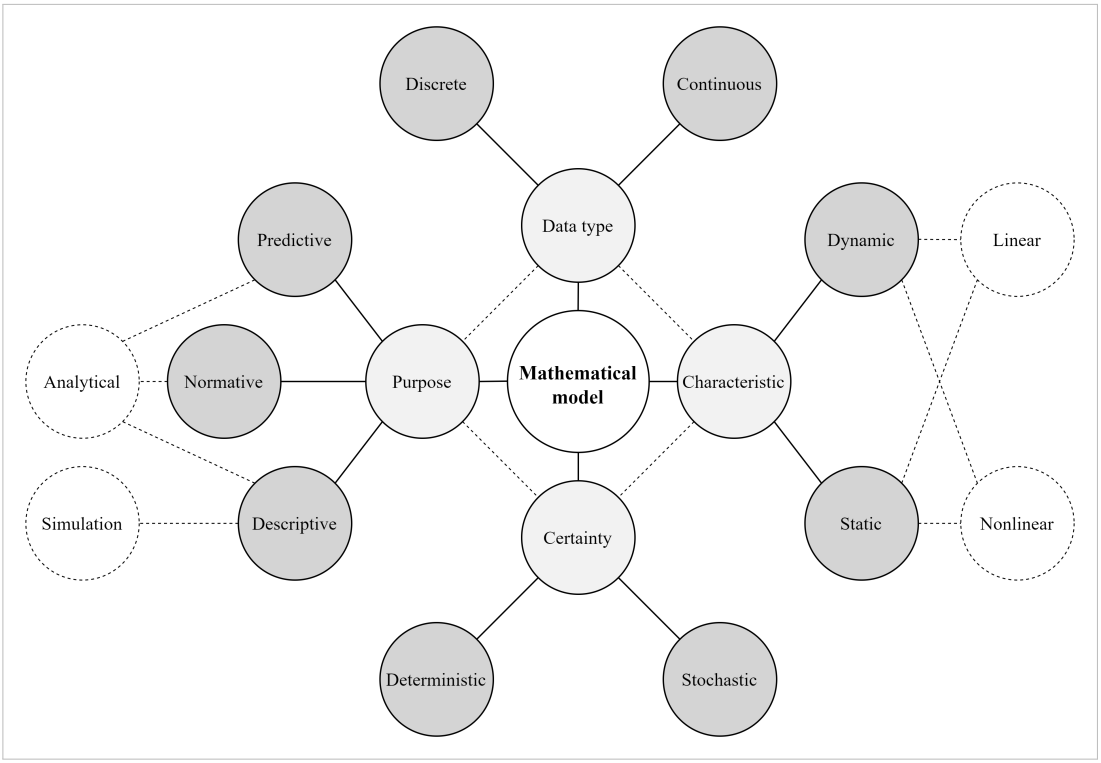


Figure 3.2: Mathematical model classification [28]

Figure 3.2 depicts various classification types involving mathematical models. This study proposes a goal programming approach for cricket team selection that balances the trade-off between several conflicting and incommensurable cricketing abilities. As such, this study utilises continuous data, providing deterministic models that follow a normative approach with an analytical solution method. Therefore, the optimisation theory applicable to this study involves a descriptive overview of essential deterministic models in linear and integer programming, followed by a detailed discussion of goal programming and fuzzy logic.

3.2 Linear programming

Section 3.2 discusses the origins of linear programming, followed by its theory and standard formulation, with a concluding overview of the simplex method.

3.2.1 Origins of linear programming

Linear programming represents the foundation of mathematical optimisation, widely viewed as one of the leading decision-making techniques for quantitative analysis of practical problems [141], [153]. Shortly after World War II, Dantzig formally defined linear programming as an analytical method to solve military logistics problems [154]. Dantzig [154] recognised that the concept of linear programming was not entirely unknown before its introduction in 1947. Prominent mathematicians, such as Fourier and de la Vallée Poussin, issued papers

covering key features of linear programming in 1823 and 1911, respectively [154]. In 1941, Hitchcock completed a paper on the famous transportation problem, which was dismissed until peer researchers rediscovered its properties in the late 1940s and early 1950s [154].

Dantzig [154] stated that the work completed on this topic before 1947 was neglected due to a global lack of interest in optimisation. However, linear programming has since become one of the leading scientific techniques for decision-making [141]. Since its inception, linear programming has been widely applied in various industries, such as agriculture, education, manufacturing, and aviation [28]. Muzhinji and Ndou [155] recognised linear programming as one of the most significant scientific developments of the 20th century, whose application altered the operation of numerous industries.

3.2.2 Fundamentals of linear programming

Linear programming is a deterministic technique that derives decision variable values that minimise or maximise a linear objective function subject to a finite set of linear constraints [156]. The feasible points to a problem are determined using a collection of linear equalities or inequalities [156]. The amount of each resource employed is proportionate to the value of the corresponding decision variable [28]. Linear programming provides a means to select the best feasible point among various alternatives [156]. The following definitions, provided by Winston and Goldberg [157], define the concept of linear programming:

Definition 3.2.1. *A function $f(x_1, x_2, ..., x_n)$ of $x_1, x_2, ..., x_n$ is deemed a linear function if, for some set of coefficients $c_1, c_2, ..., c_n$, $f(x_1, x_2, ..., x_n) = c_1x_1 + c_2x_2 + ... + c_nx_n$.*

Definition 3.2.2. *For a linear function $f(x_1, x_2, ..., x_n)$ and the number b, the inequalities $f(x_1, x_2, ..., x_n) \geq b$ and $f(x_1, x_2, ..., x_n) \leq b$ are consider linear inequalities.*

Definition 3.2.3. *A linear programming problem denotes an optimisation problem having the following characteristics:*

- *A linear objective function encompassing the defined decision variables is either minimised or maximised.*
- *The values of the decision variables satisfy a finite collection of constraints, expressed as linear equalities or inequalities.*
- *Each decision variable x_j has a corresponding sign restriction, defining the parameter as either non-negative or unrestricted in sign.*

Mathematical models cannot capture every aspect of empirical systems due to their various considerations and constantly changing characteristics [28]. Therefore, Sharma [28] noted that all mathematical models encompass certain assumptions, reducing the complexity of practical problems by providing simplified expressions that are more easily analysed. Hillier and Lieberman [158] stated that linear programming models contain several assumptions

involving the activities and data of the problem being modelled. Osagie and Icheme [159] mentioned that the assumptions of linear programming remain constant irrespective of the application domain. Linear programming problems comprise the following assumptions for modelling complex real-world systems [28], [157]:

1. Certainty: The model parameters and functional relationships are known with absolute certainty and remain constant.
2. Additivity: The objective function value and the quantity of each resource employed must equal the aggregate of the individual contribution of each decision variable.
3. Proportionality: The quantity of each resource employed and its contribution to the objective function is proportionate to the related decision variable value.
4. Divisibility: The decision variable values of the model are allowed to be continuous, with fractional solutions permitted.

The standard formulation of a linear programming problem is represented using the scalar and matrix forms [160]. All standard formulations are presented as minimisation problems. Equations (3.4), (3.5), and (3.6) specify the standard formulation of a linear programming problem expressed in scalar form [28].

$$\text{Minimise} \quad Z = \sum_{j=1}^{n} c_j x_j \tag{3.4}$$

$$\text{Subject to} \quad \sum_{j=1}^{n} a_{ij} x_j \geq b_i, \quad \forall i = 1, 2, ..., m \tag{3.5}$$

$$x_j \in \mathbb{R}_+, \quad \forall j = 1, 2, ..., n \tag{3.6}$$

From Equations (3.4), (3.5), and (3.6), c_j describes the cost coefficient associated with the non-negative decision variable x_j [28]. Likewise, a_{ij} represents the i-th constraint coefficient concerning the decision variable x_j, with b_i the corresponding constraint requirement [28]. Equations (3.7), (3.8), and (3.9) present the standard formulation of a linear programming problem expressed in matrix form [156].

$$\text{Minimise} \quad Z = \boldsymbol{c}^T \boldsymbol{x} \tag{3.7}$$

$$\text{Subject to} \quad \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b} \tag{3.8}$$

$$\boldsymbol{x} \in \mathbb{R}_+ \tag{3.9}$$

Similar to the scalar form, the matrix form encompasses a set of cost coefficients expressed by the n dimensional cost vector $\boldsymbol{c}^T = (c_1, c_2, ..., c_n)$ [156]. The cost vector $\boldsymbol{c}$ exerts a weight onto a set of decision variables, described by the non-negative n dimensional variable vector

$\boldsymbol{x} = (x_1, x_2, ..., x_n)^T$ [156]. The $m \times n$ matrix $\boldsymbol{A}$ comprises the constraint coefficients, with the m dimensional right-hand-side vector $\boldsymbol{b} = (b_1, b_2, ..., b_m)^T$ characterising the constraint requirements [156]. Equation (3.10) defines the matrix $\boldsymbol{A}$ [156].

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \tag{3.10}$$

From the standard formulations, the fundamental theorem of linear programming specifies that the optimal solution of a bounded linear function lies at a vertex of the feasible region [161]. Theorem 3.2.1 describes the fundamental theorem of linear programming, as defined by Luenberger and Ye [162].

Theorem 3.2.1. *Given a linear programming problem in standard form (3.7), (3.8), and (3.9), with* $\boldsymbol{A}$ *the* $m \times n$ *matrix of order* m*, then the following holds:*

1. *if a feasible solution exists, then a basic feasible solution exists; and*
2. *if an optimal feasible solution exists, then an optimal basic feasible solution exists.*

Each linear programming problem, described as the primal, contains a corresponding dual expression [163]. The duality theory of linear programming specifies that each problem can be evaluated from two perspectives, both leading to equivalent solutions [28]. The duality theorem investigates the relationship between the solutions of paired linear programming problems [144]. Linear programming problems constitute primal-dual pairs [163], with one program representing the primal and another the dual [144]. Sharma [28] mentioned that since the dual of the dual yields the primal, it is immaterial which of the linear programs is considered the primal. Figure 3.3 demonstrates the primal-dual relationship of a linear programming problem, with the formulation presented in Equations (3.4), (3.5), and (3.6) selected as the primal program.

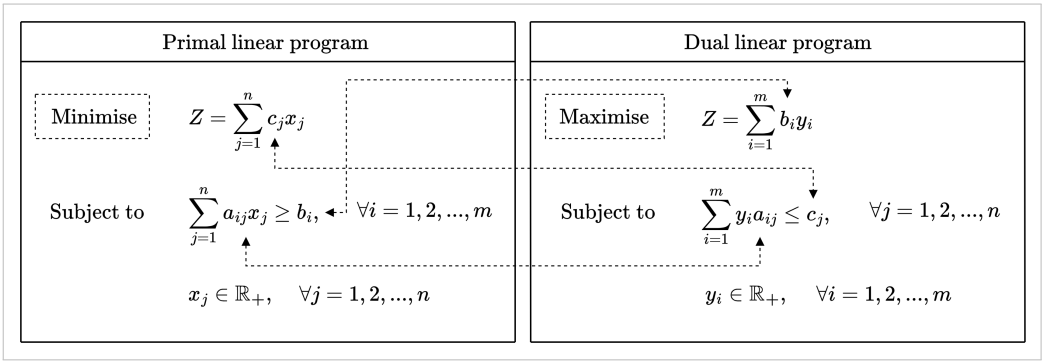


Figure 3.3: Primal-dual relationship [163]

Each feasible solution obtained for either the primal or dual program bounds the objective function value of the other [163]. Feasible solutions for the dual program denote the upper bounds of the primal program [144]. Feasible solutions for the primal program represent the lower bounds of the dual program [144]. Theorems 3.2.2 and 3.2.3 describe the duality theorem [144] and complementary slackness [163], respectively.

Theorem 3.2.2. *The duality theorem states that:*

1. *if the primal has an optimum solution, then the dual has an optimum solution;*
2. *if the primal has an unbounded solution, then the dual has an infeasible solution; and*
3. *if the primal has an infeasible solution, then the dual contains either an infeasible or unbounded solution.*

Theorem 3.2.3. *Suppose* $\boldsymbol{x} = (x_1, x_2, ..., x_n)$ *is primal feasible, whereas* $\boldsymbol{y} = (y_1, y_2, ..., y_m)$ *is dual feasible. Let* $(w_1, w_2, ..., w_m)$ *be the primal slack variables, and* $(z_1, z_2, ..., z_n)$ *be the dual slack variables. Then,* $\boldsymbol{x}$ *and* $\boldsymbol{y}$ *are deemed optimal if complementary slackness holds:*

$$x_j z_j = 0, \quad \forall j = 1, 2, ..., n \tag{3.11}$$

$$w_i y_i = 0, \quad \forall i = 1, 2, ..., m \tag{3.12}$$

Linear programming problems that comprise at most two decision variables can be solved graphically by plotting the collection of linear constraints characterising the feasible region [28]. However, most empirical problems encompass more than two decision variables [157]. Such problems can be solved using the simplex method, which is globally viewed as one of the most profound scientific developments [156].

3.2.3 Simplex method

The simplex method is the most popular technique for solving complex linear programming problems [164]. Dantzig first proposed the simplex method in 1947 as an algebraic approach that contains various geometric properties [158]. In 2000, the journal Computing in Science and Engineering included the simplex method as one of the 10 most influential algorithms of the 20th century [165]. Banciu [166] mentioned that the simplex method represents the backbone of linear programming.

Recall that the optimal solution of a linear programming problem, with a bounded objective function, lies at an extreme point of the n dimensional polytope defining the feasible region [28]. The simplex method systematically investigates these extreme points until the optimal solution is obtained [28]. Consequently, the simplex method traverses from a basic feasible solution to an adjoining vertex to better the objective function value [144]. Figure 3.4 shows the operation of the simplex method considering a linear program of type minimisation.

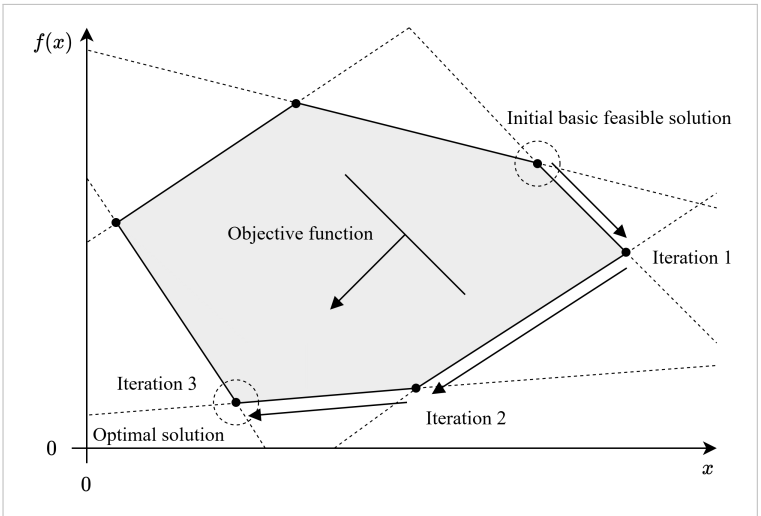


Figure 3.4: Simplex method [144]

The feasible region contains a finite number of extreme points, allowing the simplex method to improve the objective function value until the optimal solution is achieved [28]. Winston and Goldberg [157] presented the following fundamental steps that describe the operation of the simplex method:

1. Convert the linear programming problem into its standard formulation.
2. Select an initial basis for the first iteration of the algorithmic procedure.
3. Obtain a basic feasible solution from the resulting standard formulation.
4. Establish if the current solution is optimal based on the stopping criterion.
5. If the existing solution is not optimal, determine the next nonbasic variable to enter the basis and the current basic variable to leave the basis.
6. Perform elementary row operations on the resulting improved basis.
7. Determine the new basic feasible solution and continue from Step 4.

The simplex method utilises a standard linear program as input [144]. Accordingly, consider the standard formulation of a linear programming problem, as presented in Equations (3.7), (3.8), and (3.9). Hillier and Lieberman [158] reported that the algebraic approach involves solving systems of linear equations. The inequality constraints in the standard formulation are converted into comparable equality constraints by including slack variables [158]. Slack variables represent unused resources that do not influence the objective function value [28]. Incorporating slack variables yields an augmented equivalent of the standard formulation, resulting in a more convenient expression for mathematical manipulation [158]. Equation (3.13) shows the matrix $\boldsymbol{A}$, while Equation (3.14) defines the vectors $\boldsymbol{c}$, $\boldsymbol{x}$, and $\boldsymbol{b}$ associated with the augmented expression [163]. The vector of decision variables $\boldsymbol{x}$ in Equation (3.14) contains the slack variables [163].

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & 1 & & \\ a_{21} & a_{22} & \dots & a_{2n} & & 1 & \\ \vdots & \vdots & \ddots & \vdots & & & \ddots \\ a_{m1} & a_{m2} & \dots & a_{mn} & & & 1 \end{bmatrix} \tag{3.13}$$

$$\boldsymbol{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ x_{n+1} \\ \vdots \\ x_{n+m} \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \tag{3.14}$$

Each iteration of the simplex method involves establishing which m of the $n+m$ variables are considered basic [163]. Let $\mathcal{B}$ represent the set of indices relating to the basic variables, with $\mathcal{N}$ describing the set of nonbasic indices [163]. From Equation (3.8), the product $\boldsymbol{Ax}$ requires decomposition using standard matrix notation [163]. Therefore, let $\boldsymbol{B}$ denote the $m \times m$ matrix having the m basic variable columns of $\boldsymbol{A}$ [163]. Similarly, let $\boldsymbol{N}$ represent the $m \times n$ matrix comprising the n columns of $\boldsymbol{A}$ pertaining to the nonbasic variables [163]. Equation (3.15) illustrates the partitioned form of the $m \times n$ matrix $\boldsymbol{A}$ [163].

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{B} & \boldsymbol{N} \end{bmatrix} \tag{3.15}$$

Considering Equation (3.15), the matrix on the right represents a rearranged expression of $\boldsymbol{A}$ such that the columns concerning the basic variables are provided first, followed by the nonbasic variable columns [163]. Subsequently, the rows of $\boldsymbol{x}$ are rearranged analogously, as shown in Equation (3.16) [163].

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_{\mathcal{B}} \\ \boldsymbol{x}_{\mathcal{N}} \end{bmatrix} \tag{3.16}$$

From Equations (3.15) and (3.16), $\boldsymbol{Ax}$ is separated into the sum of the basic and nonbasic elements [163]. Equation (3.17) illustrates the division of $\boldsymbol{Ax}$ [163].

$$\boldsymbol{Ax} = \begin{bmatrix} \boldsymbol{B} & \boldsymbol{N} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_{\mathcal{B}} \\ \boldsymbol{x}_{\mathcal{N}} \end{bmatrix} = \boldsymbol{B}\boldsymbol{x}_{\mathcal{B}} + \boldsymbol{N}\boldsymbol{x}_{\mathcal{N}} = \boldsymbol{b} \tag{3.17}$$

Similarly, partitioning $\boldsymbol{c}$ yields the expression in Equation (3.18) [163].

$$\boldsymbol{c}^T\boldsymbol{x} = \begin{bmatrix}\boldsymbol{c}_\mathcal{B}\\ \boldsymbol{c}_\mathcal{N}\end{bmatrix}^T \begin{bmatrix}\boldsymbol{x}_\mathcal{B}\\ \boldsymbol{x}_\mathcal{N}\end{bmatrix} = \boldsymbol{c}_\mathcal{B}^T\boldsymbol{x}_\mathcal{B} + \boldsymbol{c}_\mathcal{N}^T\boldsymbol{x}_\mathcal{N} \tag{3.18}$$

Equation (3.17) is manipulated such that the basic variables $\boldsymbol{x}_\mathcal{B}$ are expressed as a function of the nonbasic variables $\boldsymbol{x}_\mathcal{N}$, considering the invertible matrix $\boldsymbol{B}^{-1}$ [163]. Equation (3.19) presents the rearranged expression [163].

$$\boldsymbol{x}_\mathcal{B} = \boldsymbol{B}^{-1}\boldsymbol{b} - \boldsymbol{B}^{-1}\boldsymbol{N}\boldsymbol{x}_\mathcal{N} \tag{3.19}$$

For each iteration of the simplex method, one nonbasic variable replaces a basic variable in the current basis to improve the objective function value [163]. Equations (3.20), (3.21), and (3.22) extend the objective function representation in Equation (3.18) using standard mathematical substitution [163].

$$\begin{aligned}\boldsymbol{c}^T\boldsymbol{x} &= \boldsymbol{c}_\mathcal{B}^T\boldsymbol{x}_\mathcal{B} + \boldsymbol{c}_\mathcal{N}^T\boldsymbol{x}_\mathcal{N} && (3.20)\\ &= \boldsymbol{c}_\mathcal{B}^T(\boldsymbol{B}^{-1}\boldsymbol{b} - \boldsymbol{B}^{-1}\boldsymbol{N}\boldsymbol{x}_\mathcal{N}) + \boldsymbol{c}_\mathcal{N}^T\boldsymbol{x}_\mathcal{N} && (3.21)\\ &= \boldsymbol{c}_\mathcal{B}^T\boldsymbol{B}^{-1}\boldsymbol{b} + (\boldsymbol{c}_\mathcal{N}^T - \boldsymbol{c}_\mathcal{B}^T\boldsymbol{B}^{-1}\boldsymbol{N})\boldsymbol{x}_\mathcal{N} && (3.22)\end{aligned}$$

From Equation (3.22), the term $(\boldsymbol{c}_\mathcal{B}^T\boldsymbol{B}^{-1}\boldsymbol{b})$ presents the objective function value of the current basis, with the reduced cost vector expressed as $(\boldsymbol{c}_\mathcal{N}^T - \boldsymbol{c}_\mathcal{B}^T\boldsymbol{B}^{-1}\boldsymbol{N})$ [167]. Subsequently, the vector $(\boldsymbol{c}_\mathcal{N}^T - \boldsymbol{c}_\mathcal{B}^T\boldsymbol{B}^{-1}\boldsymbol{N})$ characterises the nonbasic variable to enter the basis [167]. For minimisation problems, the nonbasic variable to join the basis comprises the most negative value for $(\boldsymbol{c}_\mathcal{N}^T - \boldsymbol{c}_\mathcal{B}^T\boldsymbol{B}^{-1}\boldsymbol{N})$ [167]. For maximisation problems, the nonbasic variable to join the basis contains the most positive value for $(\boldsymbol{c}_\mathcal{N}^T - \boldsymbol{c}_\mathcal{B}^T\boldsymbol{B}^{-1}\boldsymbol{N})$ [167].

In addition, the vector $(\boldsymbol{c}_\mathcal{N}^T - \boldsymbol{c}_\mathcal{B}^T\boldsymbol{B}^{-1}\boldsymbol{N})$ defines the stopping criteria for the simplex method [160]. For minimisation problems, the basis of the current iteration is deemed optimal when the vector $(\boldsymbol{c}_\mathcal{N}^T - \boldsymbol{c}_\mathcal{B}^T\boldsymbol{B}^{-1}\boldsymbol{N}) \geq 0$ [167]. For maximisation problems, the basis of the current iteration is considered optimal when the vector $(\boldsymbol{c}_\mathcal{N}^T - \boldsymbol{c}_\mathcal{B}^T\boldsymbol{B}^{-1}\boldsymbol{N}) \leq 0$ [167]. The nonbasic variable to enter the basis for each iteration assumes the value one for $\boldsymbol{x}_\mathcal{N}$ [156]. Equation (3.23) gives the expression used to select the basic variable to leave the current basis [167].

$$\boldsymbol{\delta} = \frac{\boldsymbol{B}^{-1}\boldsymbol{b}}{\boldsymbol{B}^{-1}\boldsymbol{N}\boldsymbol{x}_\mathcal{N}} \tag{3.23}$$

The basic variable comprising the smallest positive entry for $\boldsymbol{\delta}$ is selected to leave the basis [167]. After replacing the appropriate basic variable with the established nonbasic variable, the new basis is utilised as input for the succeeding iteration [167]. The process is repeated iteratively until the associative stopping criterion is met [167]. The problem is unbounded

if the test in Equation (3.23) generates all negative values for the basic variables captured by the existing basis [167].

3.3 Integer programming

Section 3.3 defines the origins of integer programming, followed by its theory and standard formulation, with a concluding discussion on the branch-and-bound method.

3.3.1 Origins of integer programming

Integer programming represents an extension of linear programming [168], closely related to combinatorial optimisation [169]. This study utilises the phrase integer programming to describe integer linear programming. In 1958, Gomory published a novel paper explaining how the simplex method can be modified accordingly to produce optimal integral solutions to linear programming problems [170]. While various authors presented integer models for different practical problems, the approach developed by Gomory represented the first finite technique to solve integer linear programs [171]. Gomory advanced prior work conducted by Dantzig, Fulkerson, and Johnson on the travelling salesman problem in 1954 by illustrating the systematic addition of cutting planes [154], [172]. Cuts are supplementary restrictions that are added to an existing collection of constraints that essentially force models to solve for integer solutions [154], [170].

Integer programming arose in response to the need observed by multiple practitioners and researchers to determine optimal solutions to empirical problems mandating integer-valued decision variables [171]. Consequently, the work of Gomory instantly received considerable attention due to its theoretical principles and perceived computational benefits, considering the technological advancements of the 1950s [170]. Pioneers such as Johnson, Balas, Land, and Doig advanced the fundamental work of Gomory [154], leading to integer programming becoming one of the most profound techniques in mathematical optimisation [173]. Eshuis and Kumar [174] regarded integer programming as a versatile technique with applications in network design, vehicle routing, facility location, and transportation.

3.3.2 Fundamentals of integer programming

Integer programming problems comprise at least one decision variable that assumes a non-fractional discrete state [171]. Integer programming concerns finding an optimum solution among all discrete points in the feasible region [171]. An integer programming problem is considered linear if relaxing the integrality restrictions on the decision variables delivers an analogous linear program [171]. Definition 3.3.1 describes a linear programming relaxation, as defined by Winston and Goldberg [157]. Similarly, Definition 3.3.2 details the integrality gap, as discussed by Adams [175].

Definition 3.3.1. *The linear program gained after omitting all the integer or binary* $(0-1)$ *constraints imposed on the decision variables is called the linear programming relaxation of the integer program considered.*

Definition 3.3.2. *The integrality gap represents the difference between the optimal solution of the integer program and its corresponding linear programming relaxation.*

Luies [144] discussed that the search space of an integer programming problem differs from that of a linear program. Figure 3.5 illustrates the search space of an integer programming problem, with the feasible points expressed as integer values [144]. Figure 3.6 demonstrates the relaxed representation of the convex polytope in Figure 3.5, with all integer restrictions removed [144]. As such, Figure 3.6 shows the feasible region for the related linear program, with the feasible points allowed to assume any real values.

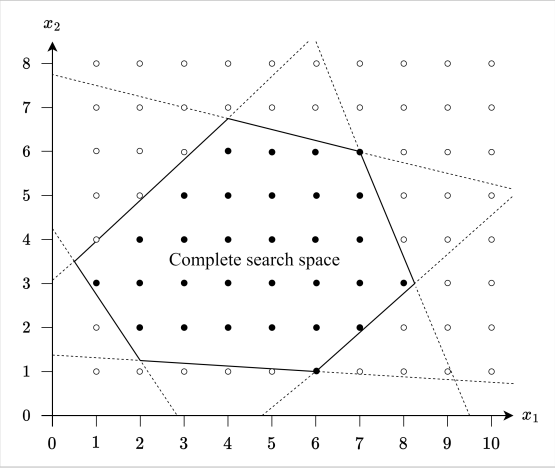


Figure 3.5: Integer program [144]

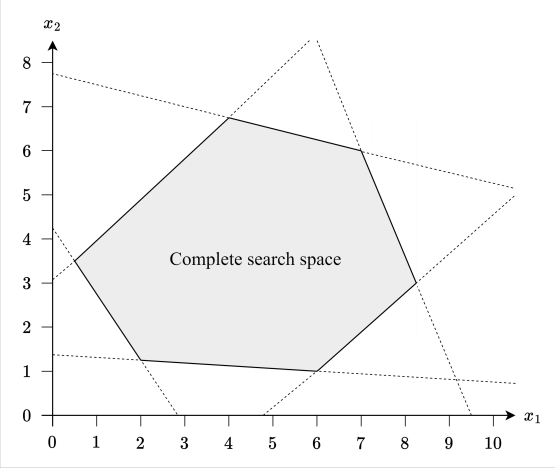


Figure 3.6: Linear program [144]

The fundamental theorem of integer programming states that for some rational polyhedron $P \subseteq \mathbb{R}^n$, the convex hull $\text{conv}(P \cap \mathbb{Z}^n)$ is also considered a rational polyhedron [176], [177]. The convex hull is the smallest convex set comprising all feasible solutions [178]. Theorem 3.3.1 introduces the fundamental theorem of integer programming, as described by Conforti *et al.* [176] and Colares [177].

Theorem 3.3.1. *Let* $P = \{(x, y) : \boldsymbol{Ax} + \boldsymbol{Gy} \geq \boldsymbol{b}\}$ *and* $S = \{(x, y) \in P : \boldsymbol{x} \in \mathbb{Z}^n\}$*, where* $\boldsymbol{A}$ *and* $\boldsymbol{G}$ *denotes rational matrices, and* $\boldsymbol{b}$ *signifies a rational vector. Then, there exist the associated rational matrices* $\boldsymbol{A}^{'}$ *and* $\boldsymbol{G}^{'}$*, with the corresponding rational vector* $\boldsymbol{b}^{'}$ *such that* $conv(S) = \{(x, y) : \boldsymbol{A}^{'}\boldsymbol{x} + \boldsymbol{G}^{'}\boldsymbol{y} \geq \boldsymbol{b}^{'}\}$.

From Theorem 3.3.1, two polyhedrons P_1 and P_2 comprise valid formulations for the same integer programming problem if $P_1 \cap \mathbb{Z}^n = P_2 \cap \mathbb{Z}^n$ [177]. Conforti *et al.* [176] mentioned that $\text{conv}(S)$ might not be a polyhedron if the matrices $\boldsymbol{A}$ and $\boldsymbol{G}$ are not rational. Figure 3.7 illustrates the concept underpinning the fundamental theorem of integer programming, where $P^{'} = \text{conv}(S) = \text{conv}(P \cap \mathbb{Z}^n)$ denotes the convex hull containing all feasible integer points located in P.

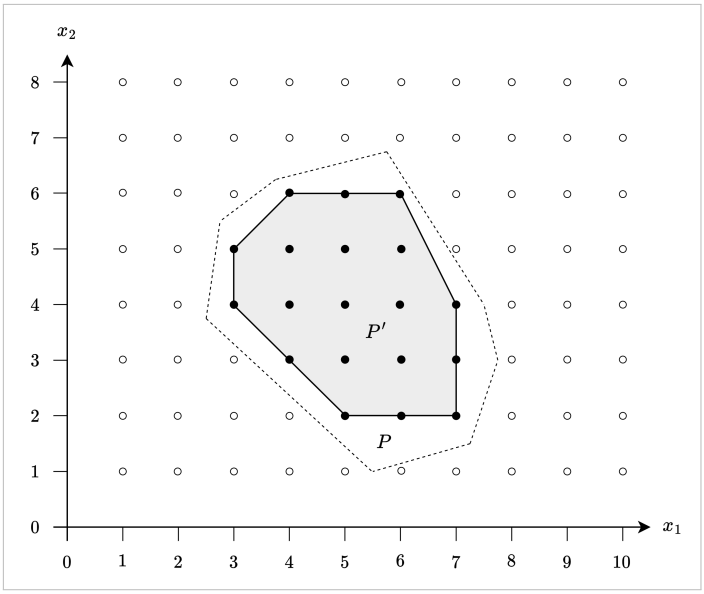


Figure 3.7: Demonstration of P and $P^{'}$ [177]

An integer programming problem is classified as pure, mixed, or binary, depending on the provisions of the system constraints [169]. An integer program is considered pure if all the decision variables are integer-valued [169]. Equations (3.24), (3.25), and (3.26) provide the standard formulation of a pure integer programming problem [179].

$$\text{Minimise} \quad Z = \sum_{j=1}^{n} c_j x_j \tag{3.24}$$

$$\text{Subject to} \quad \sum_{j=1}^{n} a_{ij} x_j \geq b_i, \quad \forall i = 1, 2, ..., m \tag{3.25}$$

$$x_j \in \mathbb{Z}_+, \quad \forall j = 1, 2, ..., n \tag{3.26}$$

A mixed integer programming problem restricts some, but not all, decision variables to be integer-valued [169]. For mixed integer programs, d_k defines the cost coefficient associated with the integer-valued decision variable y_k, where g_{ik} details the i-th constraint coefficient [169]. Equations (3.27) to (3.30) characterise the standard formulation of a mixed integer programming problem [169], [180].

$$\text{Minimise} \quad Z = \sum_{j=1}^{n} c_j x_j + \sum_{k=1}^{p} d_k y_k \tag{3.27}$$

$$\text{Subject to} \quad \sum_{j=1}^{n} a_{ij} x_j + \sum_{k=1}^{p} g_{ik} y_k \geq b_i, \quad \forall i = 1, 2, ..., m \tag{3.28}$$

$$x_j \in \mathbb{R}_+, \quad \forall j = 1, 2, ..., n \tag{3.29}$$

$$y_k \in \mathbb{Z}_+, \quad \forall k = 1, 2, ..., p \tag{3.30}$$

A binary integer programming problem bounds the decision variables to zero or one [169]. Binary integer programs involve systems that require definitive true or false outcomes [181]. Equations (3.31), (3.32), and (3.33) illustrate the standard formulation of a binary integer programming problem [28].

$$\text{Minimise} \quad Z = \sum_{j=1}^{n} c_j x_j \tag{3.31}$$

$$\text{Subject to} \quad \sum_{j=1}^{n} a_{ij} x_j \geq b_i, \quad \forall i = 1, 2, ..., m \tag{3.32}$$

$$x_j \in \{0, 1\}, \quad \forall j = 1, 2, ..., n \tag{3.33}$$

The knapsack problem is a widespread combinatorial optimisation problem [182], denoting the simplest form of integer program [183]. The knapsack problem has a single constraint with a standard formulation comparable to binary integer programming [183]. The problem mimics the practical event in which a hiker must decide the items to include in a knapsack with a predefined weight limit [169]. Each item contributes a specific value to the objective function, which involves maximising the total value among all selected items [169]. Various parallels exist between the knapsack problem and the team selection problem in the current study. Definition 3.3.3 defines combinatorial optimisation according to Chen *et al.* [184].

Definition 3.3.3. *Combinatorial optimisation concerns selecting, arranging, ordering, or grouping a collection of discrete elements from a finite set of available objects.*

Figure 3.8 illustrates a simple classification of the different integer programming problems, adapted from Chen *et al.* [169]. Specifically, Figure 3.8 depicts the underlying relationship between numerous classes of integer programming problems, focusing on pure, mixed, and binary, with linear programming relaxations and distinct knapsack problems.

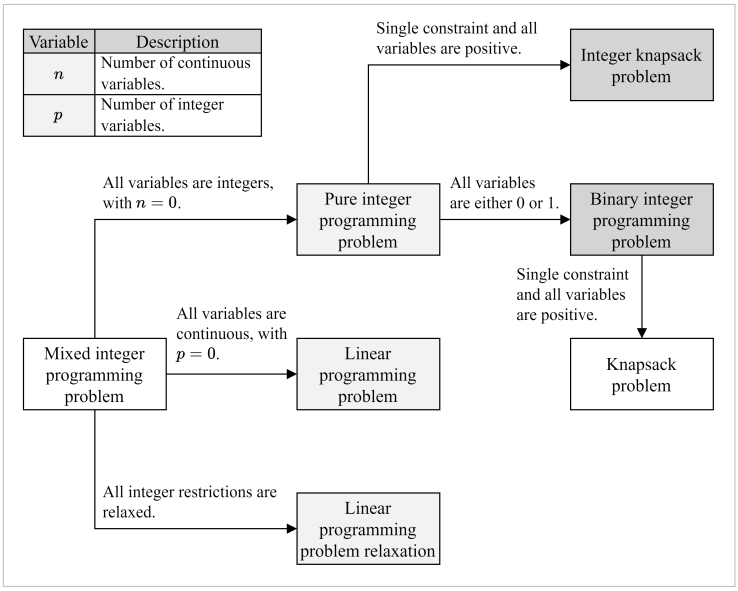


Figure 3.8: Integer programming classification [169]

Integer programs belong to a collection of nondeterministic polynomial-time (NP) complete problems that require heuristic solution procedures for moderate and large-scale systems [185]. Keskin and Gulër [185] stated that two standard exact approaches exist for solving integer programs, including the cutting plane algorithm and the branch-and-bound method [185]. While neither approach is invariably efficient computationally, Taha [186] mentioned that the branch-and-bound method is substantially more effective than the cutting plane algorithm, which gradually reduces the feasible region towards the convex hull.

3.3.3 Branch-and-bound method

Zhu and Kuno [187] and Yilmaz [188] stated that the branch-and-bound method represents the most prominent procedure for solving integer programming problems. Markowitz and Manne initially proposed the concept of the branch-and-bound method in 1957 by explicitly describing the features of the technique [189]. In 1960, Land and Doig developed the first branch-and-bound algorithm to solve pure, mixed, and binary integer programs [189]. Since its inception, the branch-and-bound method has been extensively applied to solve discrete optimisation programs, often transcending integer programming problems [189], [190].

The branch-and-bound method follows a divide-and-conquer approach [158] by partitioning complex systems into smaller, more manageable subsets [28], [167]. The branch-and-bound method divides the feasible region of the linear programming relaxation into various related subproblems [28]. The corner points of each subproblem are explored to derive an optimal solution [28]. The branch-and-bound method computes the optimum solution by implicitly enumerating the possible solutions to the linear program [157]. Figure 3.9 shows the feasible region of an arbitrary polytope [191]. Figure 3.10 depicts the operation of the branch-and-bound method by partitioning the search space in Figure 3.9.

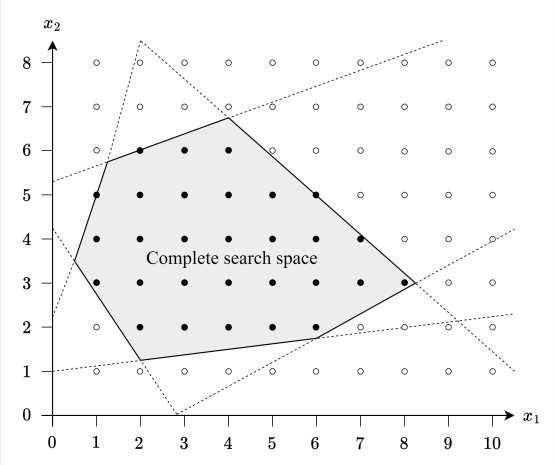


Figure 3.9: Complete search space [191]

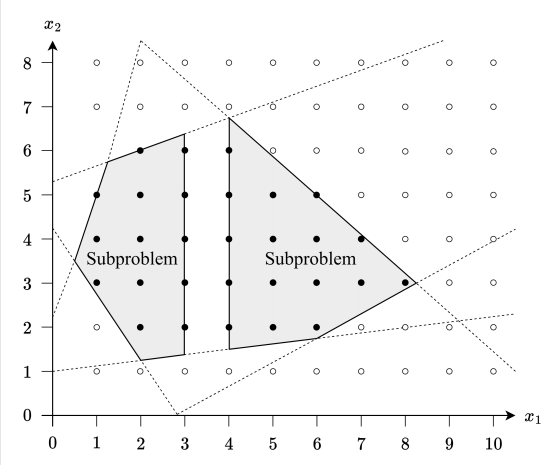


Figure 3.10: Partitioned search space [191]

The branch-and-bound method solves the linear programming relaxation of the particular integer program by iteratively branching on the decision variables having fractional values [185]. The lower and upper bounds relating to each partition are used to prune the different

tree branches, ultimately ensuring an optimum solution [185]. The following steps describe the operation of the branch-and-bound method, adapted from Sharma [28]:

1. Initialise the algorithm by solving the affiliated linear programming relaxation.
2. Generate two subproblems by branching on the variable with a fractional value.
3. Solve the subproblems using the simplex method to obtain the system bounds.
4. Review the solution of each subproblem using the fathoming conditions below:
    - Prune the tree branch if the subproblem produces an infeasible solution.
    - Return to Step 2 if the subproblem yields a related non-integer solution.
    - Continue to Step 5 if the subproblem provides a feasible integer solution.
5. Examine the objective function value of the integer solution to establish optimality.
6. Repeat the process from Step 2 until no further subproblems remain for fathoming.

Consider the standard formulation of a pure integer programming problem, as provided in Equations (3.24), (3.25), and (3.26). The branch-and-bound method starts by solving the linear programming relaxation LP_A of the integer program [28], [167], [185]. If the solution to LP_A is infeasible, the integer programming problem is infeasible [28], [167]. Likewise, if LP_A generates an unbounded solution, the integer program is considered unbounded [28], [167]. If a non-fractional solution is produced for LP_A, the optimal solution for the integer program is obtained [28]. However, the algorithm employs branching if the solution has at least one fractional decision variable [28].

Branching involves partitioning the linear programming relaxation into numerous subsets [191]. Let the objective function value Z_1 of LP_A be the initial lower bound Z_L employed for the succeeding partition [28]. Similarly, let the initial upper bound Z_U be the objective function value of the first feasible integer solution [28]. Let x_k be the non-integer decision variable with the largest fractional value for the solution of LP_A [28]. Constraining LP_A to the integer values located below and above the fractional value of x_k yields two subproblems [28]. Equation (3.34) presents the constraint applied to obtain subproblem LP_B [28].

$$x_k \leq \lfloor x_k \rfloor \tag{3.34}$$

From Equation (3.34), $\lfloor x_k \rfloor$ represents the floor value of the fractional decision variable x_k [28]. The expression in Equation (3.34) constrains x_k to the integer value directly below its fractional value [28]. Equations (3.35) to (3.38) present the resulting compact formulation for subproblem LP_B [28].

$$\text{Minimise} \quad Z = \sum_{j=1}^{n} c_j x_j \tag{3.35}$$

$$\text{Subject to} \quad \sum_{j=1}^{n} a_{ij} x_j \geq b_i, \quad \forall i = 1, 2, ..., m \tag{3.36}$$

$$x_k \leq \lfloor x_k \rfloor, \quad \forall k = 1, 2, ..., p \tag{3.37}$$

$$x_j \in \mathbb{R}_+, \quad \forall j = 1, 2, ..., n \tag{3.38}$$

Similarly, Equation (3.39) presents the constraint applied to derive subproblem LP_C [28].

$$x_k \geq \lceil x_k \rceil \tag{3.39}$$

Considering Equation (3.39), $\lceil x_k \rceil$ is the ceiling value of the fractional decision variable x_k [28]. The expression in Equation (3.39) constrains x_k to the integer value directly above its fractional value [28]. Equations (3.40) to (3.43) provide the resulting compact formulation for subproblem LP_C [28].

$$\text{Minimise} \quad Z = \sum_{j=1}^{n} c_j x_j \tag{3.40}$$

$$\text{Subject to} \quad \sum_{j=1}^{n} a_{ij} x_j \geq b_i, \quad \forall i = 1, 2, ..., m \tag{3.41}$$

$$x_k \geq \lceil x_k \rceil, \quad \forall k = 1, 2, ..., p \tag{3.42}$$

$$x_j \in \mathbb{R}_+, \quad \forall j = 1, 2, ..., n \tag{3.43}$$

Subproblems LP_B and LP_C represent the child nodes of the related parent node LP_A [167]. Let Z_2 and Z_3 be the objective function values obtained by solving subproblems LP_B and LP_C, respectively [28]. The objective function value of the best integer solution iteratively becomes the new upper bound Z_U [28]. If a subproblem yields a nonintegral solution, the process is repeated by branching on the decision variable with the largest fractional value [185]. If a specific child node delivers an infeasible solution, the branch is pruned, and the exploration of related subproblems is terminated [28], [185].

The objective function value of subproblems producing a feasible integer solution is assessed to determine its optimality [28], [185]. If the objective function value equals the lower bound Z_L, an optimum solution is obtained, with the branch-and-bound method completed [28]. If the objective function value is less than the upper bound Z_U but greater than the lower bound Z_L, the computed value becomes the new upper bound for the next partition [167]. Subproblems that produce an objective function value exceeding the upper bound Z_U are pruned and excluded from further exploration [185]. Figure 3.11 shows the working of the

branch-and-bound method for a pure integer programming problem of type minimisation [157]. The enumeration tree in Figure 3.11 employs a breadth-first search approach, with the root node describing the solution to the initial linear programming relaxation.

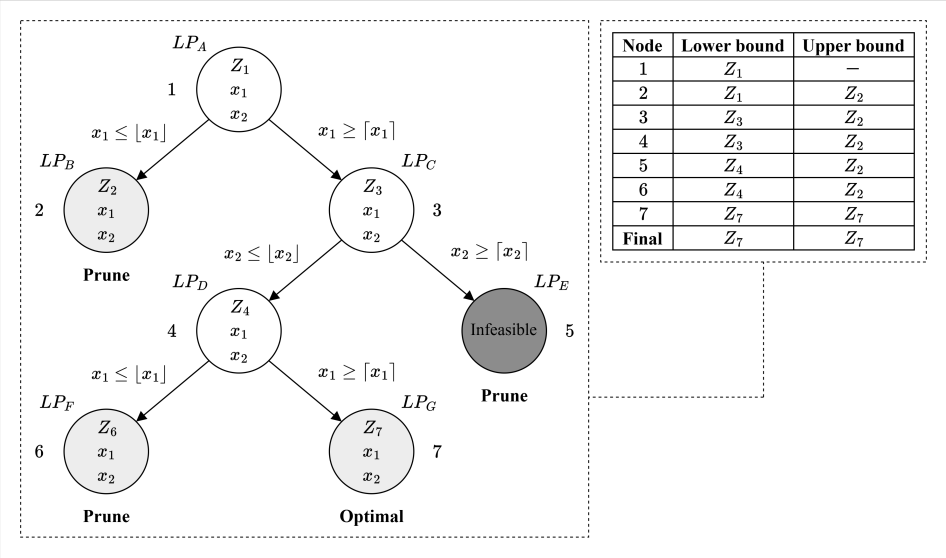


Figure 3.11: Branch-and-bound enumeration tree [157]

Considering Figure 3.11, the initial linear programming relaxation LP_A yields a non-integer solution, with the decision variable x_1 containing the largest fractional value. Therefore, Z_1 denotes the initial lower bound Z_L, with subproblems LP_B and LP_C obtained by branching on x_1. The subproblem LP_B comprises an integer solution, with Z_2 selected as the upper bound Z_U. Inversely, LP_C has a non-integer solution, with x_2 being the fractional variable. The objective function value Z_3 becomes the new lower bound Z_L, while subproblems LP_D and LP_E are obtained by branching on x_2. Again, x_1 has a fractional value for LP_D, with LP_E pruned due to infeasibility. Therefore, Z_4 replaces the existing lower bound Z_3, while branching on x_1 yields the final subproblems LP_F and LP_G that comprise integer solutions. For node 6, Z_2 remains the lower bound Z_L since LP_F delivers an objective function value exceeding that of LP_B. LP_G yields the optimal solution, with the lower and upper bounds converging to the objective function value Z_7.

The branch-and-bound method finds the optimum solution by systematically enumerating the points in the search space [144], [157]. The algorithm delivers an enumeration tree that contains sequentially ordered nodes describing related subproblems [28]. Each subproblem is explored, and the best feasible solution is recorded [144]. Branching and bounding decision variables with fractional values continue until no further nodes remain for exploration [28]. The same approach applies to integer programs of type maximisation, with the lower and upper bounds inverted [160].

From Definition 3.3.2, the integrality gap is expressed as the percentage difference between the lower and upper bounds of the problem [192], [193]. For each algorithmic iteration, the integrality gap describes the difference between the objective function values of the linear programming relaxation and the best-known integer solution [194], [195], [196]. Figure 3.12

illustrates the concept of convergence between the lower and upper bounds characterising the integrality gap.

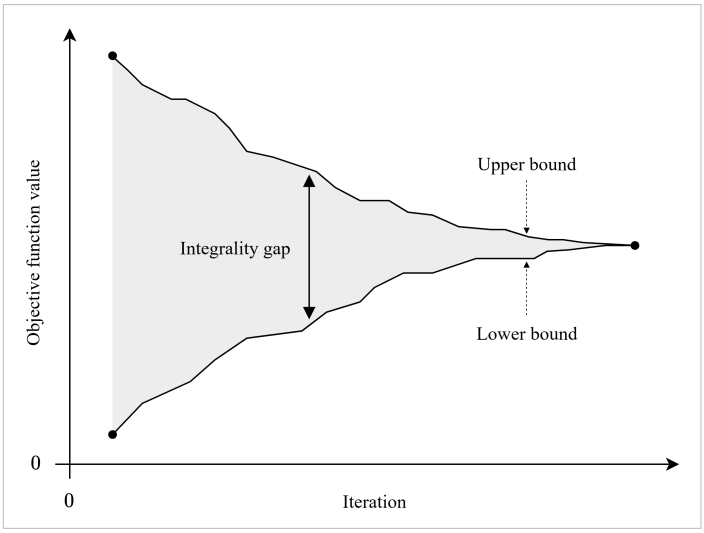


Figure 3.12: Integrality gap [197]

From Figure 3.12, an integrality gap of 0.00% implies a proven optimum solution is obtained [191]. Downs [198] mentioned that the greater the difference between the lower and upper bounds, the larger the integrality gap, causing more computational effort. Equations (3.44) and (3.45) exhibit the calculation of the integrality gap for minimisation and maximisation problems, where Z_L and Z_U represent the lower and upper bounds, respectively [199], [200].

$$\zeta_{min} = \frac{|Z_L - Z_U|}{Z_U} \times 100 \tag{3.44}$$

$$\zeta_{max} = \frac{|Z_L - Z_U|}{Z_L} \times 100 \tag{3.45}$$

3.4 Goal programming

Section 3.4 discusses the origins of goal programming, followed by its theory and standard formulation, with a concluding overview of the modified simplex method.

3.4.1 Origins of goal programming

Goal programming is a generalisation of linear programming [201], widely regarded as the oldest multi-criteria decision-making technique [29]. Schniederjans [202] noted that several researchers, such as Zoints, Steuer, Fishburn, Dyer, and Wallenius, acknowledged that the inception of goal programming led to the development of the holistic multi-criteria decision-making framework. Goal programming was first introduced in 1955 when Charnes, Cooper, and Ferguson utilised the method to determine constrained regression estimates for a linear programming problem about executive compensation [29]. Charnes and Cooper collectively

termed the technique goal programming in 1961 [203], providing an explicit representation of the approach [204]. In 1965, Ijiri advanced the rudimentary work on goal programming by including preemptive priority factors that assign unique priorities to different goals [28].

Lee and Ignizio constructed the first books exclusively covering goal programming in 1972 and 1976, respectively [204]. The initial applications of goal programming underpinned its potential to solve intricate problems encountered in various industries [203]. Yet, until the mid-1970s, research on goal programming applications was notably scarce [203]. However, the literature has since seen a considerable increase in publications involving empirical goal programming applications, primarily due to the critical work of Lee and Ignizio [203]. Goal programming has become the leading method for multi-objective optimisation [205], with applications such as resource allocation, policy analysis, and functional management [28]. Wang *et al.* [203] noted that goal programming shares various similarities with conventional compromise programming and standard reference-point methods.

3.4.2 Fundamentals of goal programming

Goal programming is globally considered the most popular multi-criteria decision-making technique [203]. Multi-criteria decision-making denotes the framework for finding the best alternatives involving problems containing numerous decision criteria [206]. Schniederjans [202] described multi-criteria decision-making as the collection of methodologies concerned with solving multiple, often conflicting, objectives. Conceptually, the relationship between multi-criteria decision-making and goal programming lies in what Zionts [207] describes as the four subareas. Table 3.1 exhibits these four subareas of multi-criteria decision-making, with their goal programming applications.

Table 3.1: Multi-criteria decision-making subareas [202]

Subarea	Description	Goal programming application
1	Multi-criteria mathematical programming.	Linear goal programming.
2	Multi-criteria discrete alternatives.	Integer goal programming.
3	Multi-attribute utility theory.	Linear goal programming. Nonlinear goal programming. Fuzzy goal programming.
4	Negotiation theory.	Interactive goal programming. Goal programming game theory.

In multi-criteria decision-making, the concept of Pareto optimality aids decision-makers in eliminating inferior solutions by only focusing on those yielding an acceptable compromise between the objectives [29]. Pareto optimality categorises the solutions to multi-objective optimisation problems as either dominated (inefficient) or non-dominated (efficient) [208]. Dominated solutions are those for which the attained level of any objective can be improved without degrading the related level of another objective [209]. Inversely, for non-dominated solutions, the achieved level of any objective cannot be improved without deteriorating the

associative level of another objective [209]. Fararah *et al.* [210] mentioned that a solution is considered Pareto optimal if no other solution can accomplish at least the same performance for all objectives while being strictly better for at least one.

The discrete set of all Pareto optimal solutions yields a Pareto efficient frontier from which the best-compromising solution can be obtained [211]. Mardle and Pascoe [211] noted that most multi-objective optimisation problems contain a Pareto efficient frontier depicting the collection of non-dominated solutions. Definition 3.4.1 defines multi-objective optimisation problems as the specific class of systems considered [212]. From Definition 3.4.1, Theorem 3.4.1 follows, which describes the concept of Pareto optimality per Szlapczynska [213].

Definition 3.4.1. *Multi-objective optimisation problems involve minimising or maximising various opposing objective functions, where improving one degenerates another.*

Theorem 3.4.1. *Without losing generality, a multi-objective optimisation problem with m decision variables and n objectives can be expressed using Equation (3.46).*

$$\boldsymbol{u} = \boldsymbol{F}(\boldsymbol{x}) = \{f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), ..., f_n(\boldsymbol{x})\} \rightarrow \text{Minimise} \tag{3.46}$$

From Equation (3.46), $\boldsymbol{x} = (x_1, x_2, ..., x_m)$ *denotes the vector of decision variables, whereas* $\boldsymbol{u} = (u_1, u_2, ..., u_n)$ *represents the related performance vector. The solution* $\boldsymbol{x}$ *containing an associative performance vector* $\boldsymbol{u}$ *is said to dominate the solution* $\boldsymbol{y}$ *comprising an affiliated performance vector* $\boldsymbol{v}$*, if both* $\boldsymbol{u}$ *and* $\boldsymbol{v}$ *meet the conditions in Equations (3.47) and (3.48).*

$$\boldsymbol{x} \prec \boldsymbol{y} \Leftrightarrow \boldsymbol{u} \prec \boldsymbol{v} \tag{3.47}$$

$$\boldsymbol{u} \prec \boldsymbol{v} \text{ iff } [\forall i \in \{1, 2, ..., n\}, u_i \leq v_i] \wedge [\exists i \in \{1, 2, ..., n\} : u_i < v_i] \tag{3.48}$$

The solution $\boldsymbol{x} \in \Omega$ *is Pareto optimal if it is non-dominated such that there exists no other solution* $\boldsymbol{x}' \in \Omega$ *for which a conjunction of* $\boldsymbol{v} \prec \boldsymbol{u}$*, Equations (3.49) and (3.50) holds.*

$$\boldsymbol{u} = \boldsymbol{F}(\boldsymbol{x}) = \{f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), ..., f_n(\boldsymbol{x})\} \tag{3.49}$$

$$\boldsymbol{v} = \boldsymbol{F}(\boldsymbol{x}') = \{f_1(\boldsymbol{x}'), f_2(\boldsymbol{x}'), ..., f_n(\boldsymbol{x}')\} \tag{3.50}$$

Pareto optimality is always considered with respect to the complete decision variable space Ω*. The collection of all non-dominated solutions represents the Pareto optimal set, which is expressed using the Pareto efficient frontier. Equations (3.51) and (3.52) define the Pareto optimal set and the Pareto efficient frontier, respectively.*

$$\boldsymbol{P}^* := \{\boldsymbol{x} \in \Omega \mid \neg \exists \boldsymbol{x}' \in \Omega : \boldsymbol{F}(\boldsymbol{x}') \prec \boldsymbol{F}(\boldsymbol{x})\} \tag{3.51}$$

$$\boldsymbol{PF}^* := \{\boldsymbol{u} = \boldsymbol{F}(\boldsymbol{x}) = \{f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), ..., f_n(\boldsymbol{x})\} \mid \boldsymbol{x} \in \boldsymbol{P}^*\} \tag{3.52}$$

Tamiz *et al.* [209] acknowledged that goal programming has received considerable criticism concerning Pareto optimality. Goal programming transforms multi-objective optimisation problems into single-objective equivalents by treating each objective as a goal comprising a target value from which deviations are minimised [214], [215]. Over-conservative selection of these target values may result in inefficient solutions [209]. Accordingly, a fundamental disadvantage of goal programming is its failure to ensure non-dominated solutions [29]. The solutions obtained using goal programming are not necessarily guaranteed Pareto optimal [29]. Despite its popularity, the prospect of inefficient solutions has led several authors to oppose the application of goal programming [209]. Figure 3.13 illustrates the possibility of goal programming delivering dominated solutions by considering two goals, g_1 and g_2.

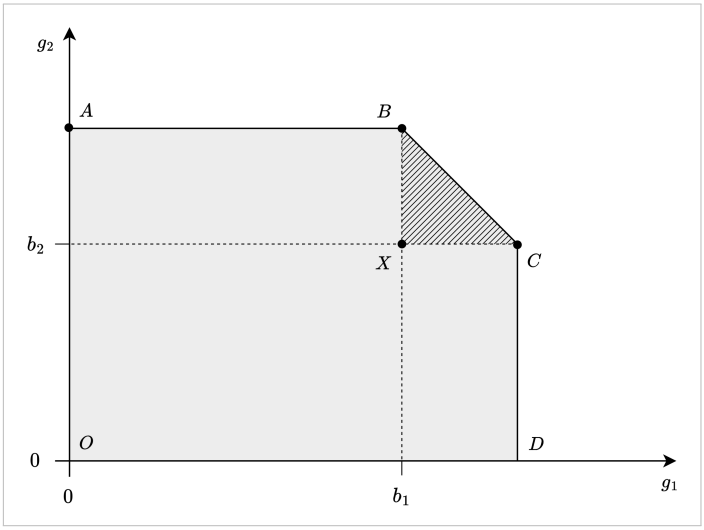


Figure 3.13: Goal programming solutions [29]

From Figure 3.13, the area $OABCD$ depicts the feasible region, with b_1 and b_2 representing the target values for g_1 and g_2, respectively [29]. Suppose the decision-maker aims to satisfy at least the target values for both goals [29]. Then, the intersection point X represents the optimum solution, as goal programming aims to minimise the deviations from the selected target values [29]. However, the solution presented by X is inefficient since it is dominated by all points relating to the triangle XBC [29]. Figure 3.13 illustrates the limitation of goal programming in producing Pareto efficient solutions [29]. However, Romero [29] mentioned that goal programming was not developed to derive non-dominated solutions. Instead, goal programming was invented to find satisfactory solutions based on the Simonian philosophy of satisficing [29]. Following this Simonian philosophy, goal programming introduces more flexibility in solving multi-objective optimisation problems.

The Simonian philosophy of satisficing enables goal programming to solve complex systems containing multiple conflicting goals in such a way that best satisfies each goal [202]. Goal programming aims to satisfy a collection of goals instead of optimising a set of objectives [216]. Therefore, goal programming seeks a satisfactory level for each goal rather than an optimal result [28]. Salas-Molina *et al.* [217] mentioned that goal programming is arguably the most popular technique that empirically applies this satisficing logic. Definition 3.4.2 defines the concept of satisficing, as described by Winter [218].

Definition 3.4.2. *Satisficing refers to the preference of decision-makers to settle for some alternative solution deemed good enough based on the available information and goals rather than striving to accomplish the optimum decision.*

Considering Definition 3.4.2, Sharma [28] formally defined goal programming as a method to solve multi-objective optimisation problems by finding a satisfactory trade-off between conflicting and incommensurable goals with different priorities [28]. Sharma [28] mentioned that a goal differs from an objective since it contains a target value subject to achievement. In contrast, an objective optimises the underlying measure of performance [28]. Hillier and Lieberman [158] claimed that there are three types of goals, defined as:

- Lower one-sided: Contains a lower limit for which underachievement is undesirable.
- Upper one-sided: Confines an upper limit for which overachievement is undesirable.
- Two-sided: Comprises a target for which under- and overachievement is undesirable.

Goal programming monitors the exact, under-, and overachievement of goals by introducing deviational variables that measure the degree to which the target values are not achieved [29]. Deviational variables are two-dimensional since they capture the negative and positive divergences from the target values, representing the equivalent of slack variables in linear programming [219]. The deviational variables that record the under- and overachievement of the target value associated with the i-th goal are expressed as [28]:

- d_i^- = Underachievement of the target value (negative deviation).
- d_i^+ = Overachievement of the target value (positive deviation).

Romero [29] stated that both deviational variables d_i^- and d_i^+ assume the value zero if the target of the i-th goal must be satisfied exactly. Equation (3.53) defines the i-th constraint, ensuring the exact target level achievement of the i-th goal [28].

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \quad \forall i = 1, 2, ..., m \tag{3.53}$$

To accommodate under- and overachievement of the target value, the expression presented in Equation (3.53) is modified to include the deviational variables d_i^- and d_i^+ [28]. Equation (3.54) provides the i-th goal constraint capturing the negative and positive divergence from the corresponding target value b_i [28].

$$\sum_{j=1}^{n} a_{ij}x_j + d_i^- - d_i^+ = b_i, \quad \forall i = 1, 2, ..., m \tag{3.54}$$

From Equation (3.54), goal programming problems encompass system and goal constraints with comparable mathematical expressions [28]. System constraints specify the restrictions imposed on the general model tangentially related to the formulated goals [28]. For system constraints, b_i denotes the i-th constraint requirement mandating absolute satisfaction [28]. Goal constraints describe the restrictions directly associated with the formulated goals [28]. For goal constraints, b_i describes the i-th target value for which under- and overachievement is captured [28]. Goal programming neglects the constant value of available resources and instead considers lower and upper target values for achievement [28]. Figures 3.14 and 3.15 visualise the system and goal constraints.

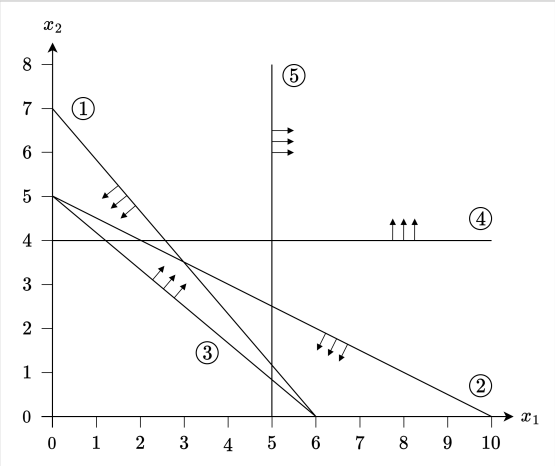


Figure 3.14: Infeasible linear program [28]

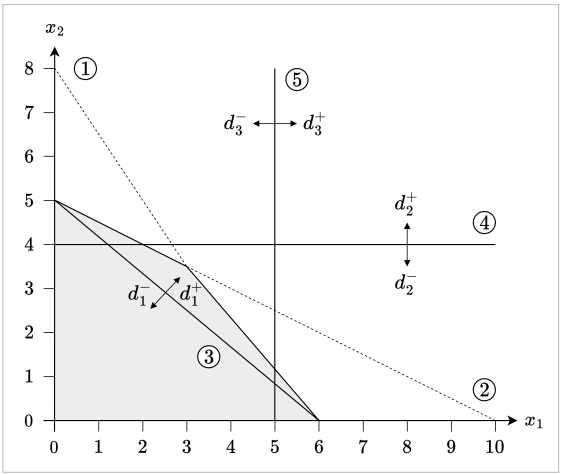


Figure 3.15: Feasible goal program [28]

Figure 3.14 depicts an infeasible linear programming problem with five system constraints. Transforming constraints (3), (4), and (5) into goals using deviational variables yields the feasible goal programming problem in Figure 3.15. Figure 3.15 shows the feasible region of the resulting goal program bounded to zero by system constraints (1) and (2). Figure 3.15 illustrates the significance of deviational variables in capturing under- and overachievement of target values. As such, Figure 3.15 demonstrates the flexibility introduced by deviational variables, allowing goal programming to balance multiple conflicting goals effectively.

Romero and Rehman [220] noted the possibility of equivalent linear and goal programming solutions. If all goals are completely achieved, such that the total deviation is zero with no trade-offs required, then the goal program is reduced to its associative linear program [29]. Therefore, the solution to a goal program in which all goals are fully satisfied is equivalent to the solution of its related linear program [29]. Romero and Rehman [220] discussed that this equivalence results from the formulation of the specific problem rather than the nature or functionality of goal programming. Schniederjans [202] mentioned that even when linear and goal programming produce comparable solutions, the latter proves superior due to its ability to replicate practical occurrences. Definition 3.4.3 specifies equivalent solutions per Lust [221], while Theorem 3.4.2 describes the equivalence theorem between linear and goal programming solutions, as derived from Romero and Rehman [220].

Definition 3.4.3. *Two solutions $\boldsymbol{x}$ and $\boldsymbol{y}$ in the feasible set X are considered equivalent if, and only if, $f(\boldsymbol{x}) = f(\boldsymbol{y})$.*

Theorem 3.4.2. *If all goals defined as g_i for $i = 1, 2, ..., m$ are precisely accomplished such that $d_i^- = d_i^+ = 0$, then the solution of the goal program is equivalent to the solution of the corresponding linear program.*

The standard formulation of a goal programming problem adopts a similar notation to that of linear programming [28]. However, in goal programming, the deviations from the target values associated with the model goals are minimised by placing the applicable deviational variables directly in the objective function [202]. Minimising the sum of deviations directly allows goal programming to sufficiently handle multiple conflicting objectives by finding a satisfactory trade-off that delivers an acceptable result [202]. Equations (3.55), (3.56), and (3.57) provide the standard formulation of a general goal programming problem [202].

$$\text{Minimise} \quad Z = \sum_{i=1}^{m} (d_i^- + d_i^+) \tag{3.55}$$

$$\text{Subject to} \quad \sum_{j=1}^{n} a_{ij} x_j + d_i^- - d_i^+ = b_i, \quad \forall i = 1, 2, ..., m \tag{3.56}$$

$$x_j, d_i^-, d_i^+ \in \mathbb{R}_+, \quad \forall i = 1, 2, ..., m, \quad \forall j = 1, 2, ..., n \tag{3.57}$$

Romero [29] stated that various methods can minimise deviations, each leading to a unique goal programming variant. Hillier and Lieberman [158] mentioned that goal programming problems are classified based on how the goals compare in importance. In non-preemptive (weighted) goal programming, the goals are equally prioritised. Conversely, in preemptive (lexicographic) goal programming, the goals are structured according to priority following a predefined hierarchical structure [158].

Preemptive goal programming treats goals based on their priority level [29]. In preemptive goal programming, a specific degree of hierarchy is considered according to the importance of each goal [28]. The priority structure guarantees that the most important goals are first achieved, often at the expense of inferior goals carrying less significance [28]. Romero [29] noted that goals are satisfied using a predefined lexicographic order, where accomplishing the goals at a higher priority level is immeasurably favoured over those at a lower priority level. Preemptive goal programming achieves each goal sequentially, following the ordinal ranking imposed on the system [28].

Sharma [28] stated that the ranking is considered absolute, meaning the preemptive priority factors contain the relationship $P_1 >> P_2 >> ... >> P_m$. Since preemptive goal programs follow the non-archimedean property [29], $P_i >> nP_{i+1}$, where $i = 1, 2, ..., m-1$, suggests that multiplication by some large number n cannot make a lower-level goal more important than a higher-level goal [28]. Hence, higher-prioritised goals will always be satisfied before lower-prioritised goals [28]. Equations (3.58), (3.59), and (3.60) demonstrate the standard formulation of a preemptive goal programming problem [202].

$$\text{Minimise} \quad Z = \sum_{i=1}^{m} P_i(d_i^- + d_i^+) \tag{3.58}$$

$$\text{Subject to} \quad \sum_{j=1}^{n} a_{ij}x_j + d_i^- - d_i^+ = b_i, \quad \forall i = 1, 2, ..., m \tag{3.59}$$

$$x_j, d_i^-, d_i^+ \in \mathbb{R}_+, \quad \forall i = 1, 2, ..., m, \quad \forall j = 1, 2, ..., n \tag{3.60}$$

From Equations (3.58), (3.59), and (3.60), P_i describes the preemptive priority factor of a non-numeric nature [28], [202]. P_i only serves as a convenient ranking symbol that indicates the importance of the i-th goal [28]. The preceding formulation only contains a preemptive ordinal ranking of the goals, without any affiliated weights. Deviational variables with the same priority level necessitate commensurability, while deviations with different priorities need not have the same unit of measure [28]. Srivastava and Gupta [222] mentioned that preemptive goal programming is used in cases where a clear priority structure exists among the various goals.

Unlike preemptive goal programming, non-preemptive goal programming considers all the goals simultaneously, expressed by a composite objective function [29]. The composite objective function minimises the weighted sum of deviations from the target values associated with the goals [29]. Non-preemptive goal programming problems necessitate no preemptive ranking [204]. Instead, all deviations are given the same priority level, with unique weights describing their importance [204]. Equations (3.61), (3.62), and (3.63) depict the standard formulation of a non-preemptive goal programming problem [202].

$$\text{Minimise} \quad Z = \sum_{i=1}^{m} (w_i^- d_i^- + w_i^+ d_i^+) \tag{3.61}$$

$$\text{Subject to} \quad \sum_{j=1}^{n} a_{ij}x_j + d_i^- - d_i^+ = b_i, \quad \forall i = 1, 2, ..., m \tag{3.62}$$

$$x_j, d_i^-, d_i^+ \in \mathbb{R}_+, \quad \forall i = 1, 2, ..., m, \quad \forall j = 1, 2, ..., n \tag{3.63}$$

From Equations (3.61), (3.62), and (3.63), w_i^- and w_i^+ represent the non-negative weights assigned to the deviational variables d_i^- and d_i^+, respectively [202]. Tamiz *et al.* [204] noted that non-preemptive goal programming is usually employed when no natural lexicographic structure exists. Non-preemptive goal programming is favoured over the preemptive variant in cases where information about the trade-off between the objectives is required [204].

Ijiri recommended combining preemptive priorities and relative weighting to overcome the limitations involving the independent use of each variant [202]. Consequently, Charnes and Cooper constructed a formulation merging traditional non-preemptive and preemptive goal programming [202]. Equations (3.64), (3.65), and (3.66) illustrate the standard formulation of a preemptive goal programming problem that contains relative weights [202].

$$\text{Minimise} \quad Z = \sum_{i=1}^{m} \sum_{k=1}^{n_i} P_i(w_{ik}^- d_i^- + w_{ik}^+ d_i^+) \tag{3.64}$$

$$\text{Subject to} \quad \sum_{j=1}^{n} a_{ij} x_j + d_i^- - d_i^+ = b_i, \quad \forall i = 1, 2, ..., m \tag{3.65}$$

$$x_j, d_i^-, d_i^+ \in \mathbb{R}_+, \quad \forall i = 1, 2, ..., m, \quad \forall j = 1, 2, ..., n \tag{3.66}$$

From Equations (3.64), (3.65), and (3.66), the weights w_{ik}^- and w_{ik}^+ represent the coefficients appended to each of the $k = 1, 2, ..., n_i$ subgoals relating to the i-th goal having the priority level P_i assigned in rank order [202]. Since goal programming minimises the total deviation, the variable d_i^- is excluded from the objective function if underachievement of the target value relating to the i-th goal is permitted [28]. Similarly, if overachievement of the target value pertaining to the i-th goal is permitted, the variable d_i^+ is omitted from the objective function [28]. To ensure precise attainment of the target value, d_i^- and d_i^+ are included in the objective function to minimise under- and overachievement concurrently [28]. Notably, weights stipulate the relative importance of goals within the same priority level but do not influence goals at different levels [28].

Tiwari and Kumar [223] and Bhargava *et al.* [224] argued that the advantage of using goal programming over alternative methods lies in its ability to mimic practical decision-making. Goal programming enables decision-makers to replicate reality by including organisational, managerial, and environmental factors using preemptive priority factors and weights [224]. Goal programming does not optimise the objectives directly and instead converts the multi-objective optimisation problem into a single-objective equivalent, which minimises the total deviation from the target values [28], [225]. Allowing deviation from the preconceived target values enables goal programming to provide a more flexible approach towards solving multi-objective optimisation problems [28]. This flexibility, in turn, allows goal programming to balance multiple conflicting goals containing equal or distinct priority levels [28].

Jiang *et al.* [225] mentioned that goal programming problems can be solved using standard linear programming methods. Similar to linear programming, goal programming problems comprising at most two decision variables can be solved graphically by plotting the feasible region using the related goal and system constraints [28]. However, Sharma [28] noted that most empirical systems necessitate balancing multiple divergent and competing objectives with more than two decision variables. Such systems, formulated as goal programs, can be solved using the simplex method [28].

Romero [29] specified that non-preemptive goal programming problems can be solved using the conventional simplex method in Section 3.2.3 since their standard formulation conforms to that of linear programs. However, Romero [29] mentioned that several algorithms exist for solving preemptive goal programs. Orumie and Ebong [226] and Lee [227] acknowledged that the modified simplex method is the most popular approach for solving preemptive goal programming problems.

3.4.3 Modified simplex method

Lee first introduced the modified simplex method in 1972 [226] to accommodate preemptive priorities and cardinal weights [228]. The modified simplex method represents the standard technique upon which most contemporary goal programming problems are based [229]. Van Horn [228] described this method as a unique multi-phase simplex algorithm applied in the presence of a preemptive set of goals. Akinnuli *et al.* [229] acknowledged that the modified simplex method solves goal programs containing preemptive priority factors using selective features of the conventional simplex method.

The underlying functionality of the modified simplex method is similar to the conventional simplex method [226]. However, the selection criteria of the modified variant establish the basic and nonbasic variables relative to the preemptive prioritisation assigned to each goal [28]. Considering Section 3.2.3, Sharma [28] and Orumie and Ebong [226] documented the following vital features distinguishing the modified simplex method from the conventional simplex method:

- Suppose there exists a goal programming problem with the standard formulation in Equations (3.64), (3.65), and (3.66). The deviational variables d_i^- and d_i^+ constitute the equivalent of slack variables in linear programming. Therefore, no slack variables are assigned to the goal constraints in the modified simplex method because they are expressed as equalities using the deviational variables. Like the conventional simplex method, slack variables are incorporated into all system constraints to transform the remaining inequalities to comparable equalities.

- In the modified simplex method, the deviational variables of the goals containing the highest priorities are chosen as the initial basic variables $\boldsymbol{x_{\mathcal{B}}}$. The remaining variables are selected as the nonbasic variables $\boldsymbol{x_{\mathcal{N}}}$. Akin to the conventional simplex method, the basic variable values are computed using the equalities by initialising the nonbasic variables to zero. The resulting basic variable values determine the values appended to the preemptive priority factors $P_1 >> P_2 >> ... >> P_m$ assigned in rank order.

- In the modified simplex method, the reduced cost vector ($\boldsymbol{c_{\mathcal{N}}^T} - \boldsymbol{c_{\mathcal{B}}^T B^{-1} N}$) is solved in terms of the different preemptive priority factors. The reduced cost vector entries for the highest priority level are first assessed, with ties among equally ranked goals arbitrarily resolved. If all entries for the highest priority level are non-negative, then the optimal solution for the particular preemptive priority is obtained, and the next lower-priority level is considered. If substituting the preemptive priority factor values leads to all entries in the reduced cost vector being non-negative, then the optimum solution to the entire problem is obtained. Akin to the conventional simplex method, the basis of the current iteration is optimal when the vector $(\boldsymbol{c_{\mathcal{N}}^T} - \boldsymbol{c_{\mathcal{B}}^T B^{-1} N}) \geq 0$.

- Like the conventional simplex method, the nonbasic variable with the most negative value for $(\boldsymbol{c_{\mathcal{N}}^T} - \boldsymbol{c_{\mathcal{B}}^T B^{-1} N})$ joins the basis if an optimal solution is not achieved. The variable to leave the basis is the one for which the ratio test $\boldsymbol{\delta}$ produces the smallest

positive value. After replacing the basic variable with the affiliated nonbasic variable, the new basis is used as input for the succeeding algorithmic iteration. The process is repeated iteratively until the discussed stopping criterion is met.

3.5 Fuzzy logic

Section 3.5 briefly defines the origins of fuzzy logic, followed by fuzzy sets and membership functions, including a concise overview of the fuzzy inference system.

3.5.1 Origins of fuzzy logic

Fuzzy logic extends conventional Boolean logic using fuzzy set theory, a generalisation of classical set theory [230], [231]. Zadeh developed fuzzy logic in 1965 [232] to mimic practical decision-making by considering imprecise and vague phenomena rather than binary true or false statements [233]. Siddiqui *et al.* [234] discussed that fuzzy logic stems from the Laws of Thought conceived by ancient philosophers such as Aristotle. Among these, the Law of the Excluded Middle specifies that each proposition is considered true or false [234]. Plato identified a third region surpassing true and false, providing the foundation for what would ultimately become fuzzy logic [234]. While modern philosophers such as Hegel, Marx, and Engels supported Plato, Lukasiewicz was the first to construct a systematic alternative to the multi-valued logic of Aristotle [234].

Fuzzy logic initially received little attention, with its practical significance only recognised nearly a decade following its inception [232]. It was not until 1974, when Mamdani utilised the technique to regulate an automated stream engine, that the value of fuzzy logic became apparent [232]. However, the literature has since seen considerable advancements in fuzzy logic [232], which is globally considered one of the most effective techniques to mimic human intelligence [235]. Ali *et al.* [232] recognised that fuzzy logic contains various applications in numerous industries, such as banking, education, healthcare, and production. Diker [236] mentioned that fuzzy logic is among the most applied artificial intelligence approaches.

3.5.2 Fuzzy sets

Fuzzy logic is a mathematical approach that deals with imprecision and uncertainty [237]. Fuzzy logic contains an inference structure, allowing the application of human intuition to artificial systems [237]. Unlike classical logic, which concerns expressions of absolute truth or falsehood, fuzzy logic involves statements containing subjective definitions [238]. Fuzzy logic replicates the decision-making abilities of humans by relying on vague and ambiguous details [238]. Sajja [239] mentioned that fuzzy logic is based on fuzzy sets, in which entities can assume partial membership. Sivanandam *et al.* [237] reported that the utility of fuzzy

sets lies in their ability to model the uncertainty associated with the imprecision stemming from a lack of information concerning a particular problem.

Fuzzy sets represent a collection of objects in which no clear boundary exists between the entities that are members of the set and those that are not [240]. Fuzzy sets are based on the concept of membership, which allows any element to be a set member to some degree [240]. Each element in a fuzzy set contains a value representing its degree of membership, expressed as a real number on the unit interval [0, 1] [240]. Values within this range denote distinct degrees of partial membership, while zero and one indicate minimum and maximum degrees of absolute membership, respectively [240]. This concept opposes classical or crisp sets, where entities are only assigned a value of zero or one [237]. Classical sets encompass elements that fulfil exact membership properties, whereas fuzzy sets contain elements that satisfy imprecise membership properties [241].

In classical sets, the transition between non-membership and membership for any element in a set is abrupt and well-defined [242]. In contrast, this transition in fuzzy sets is gradual for varying degrees of membership due to the ambiguity of the set boundaries [242]. Hence, the membership of fuzzy set elements is measured using functions that describe vagueness and uncertainty [242]. The ability of fuzzy sets to express transitions between membership and non-membership has a broad utility since it captures measurement uncertainties and vague notions conveyed in natural language [240]. Figures 3.16 and 3.17 show the difference in boundaries between classical sets and fuzzy sets, considering some universe of discourse X. Ross [242] defined the universe of discourse as the collection of all available information involving a particular problem.

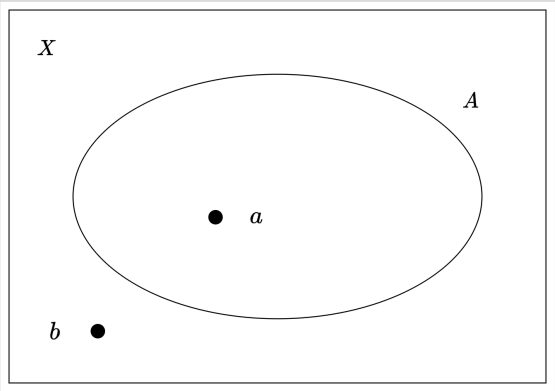


Figure 3.16: Classical set boundary [242]

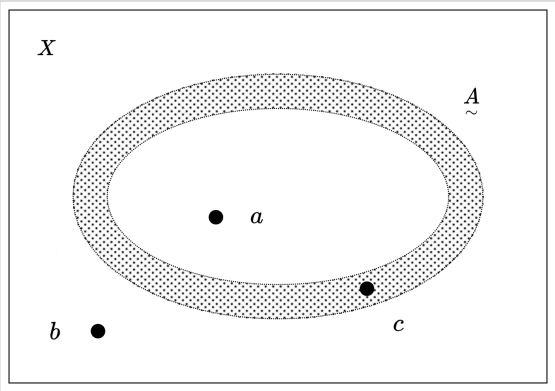


Figure 3.17: Fuzzy set boundary [242]

Figure 3.16 depicts a classical set A within the universe of discourse X [242]. From Figure 3.16, set A comprises clear boundaries, with the point a considered a member and point b regarded as a non-member [242]. Similarly, Figure 3.17 exhibits a fuzzy set $\underset{\sim}{A}$ on the same universe of discourse X, comprising vague boundaries indicated by the shaded region [242]. Considering Figure 3.17, point a denotes a full member of the set, while point b represents a non-member transcending the shaded area [242]. However, the membership of point c on the boundary region of set $\underset{\sim}{A}$ is unclear [242]. Therefore, point c contains an intermediary value on the unit interval $[0, 1]$, where one details complete membership and zero denotes

non-membership [242]. The membership of point c approaches one as it nears the centre of $\underset{\sim}{A}$ [242]. Inversely, the membership of point c approaches zero as it moves closer to the outer edge of the shaded boundary region of $\underset{\sim}{A}$ [242].

Ross [242] stated that the elements of fuzzy sets are mapped to a universe of membership values. Accordingly, if some element x in the universe is a member of the fuzzy set $\underset{\sim}{A}$, the functional mapping is expressed as $\mu_{\underset{\sim}{A}}(x) \in [0, 1]$, where $\mu_{\underset{\sim}{A}}(x)$ characterises the degree of membership for element x in the fuzzy set $\underset{\sim}{A}$ [242]. The expression $\mu_{\underset{\sim}{A}}(x)$ defines the value on the unit interval $[0, 1]$, capturing the degree to which x belongs to $\underset{\sim}{A}$ [242]. Figure 3.18 demonstrates the membership mapping for some fuzzy set $\underset{\sim}{A}$.

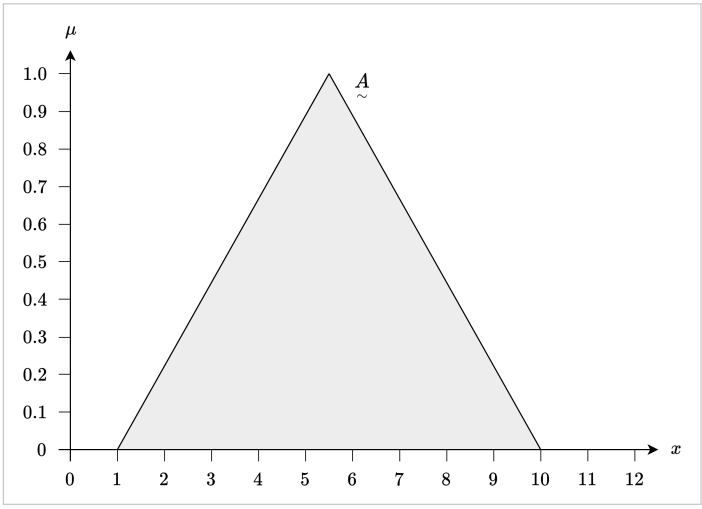


Figure 3.18: Mapping of fuzzy set $\underset{\sim}{A}$ [237]

The operations and properties of fuzzy sets are comparable to those of classical sets [243]. Consider two unique fuzzy sets $\underset{\sim}{A}$ and $\underset{\sim}{B}$ on the same universe of discourse X [242]. For a given element x, Equations (3.67), (3.68), and (3.69) illustrate the set-theoretic operations of union, intersection, and complement, respectively, considering $\underset{\sim}{A}$ and $\underset{\sim}{B}$ on X [242].

$$\mu_{\underset{\sim}{A}\cup\underset{\sim}{B}}(x) = \mu_{\underset{\sim}{A}}(x) \vee \mu_{\underset{\sim}{B}}(x) \tag{3.67}$$

$$\mu_{\underset{\sim}{A}\cap\underset{\sim}{B}}(x) = \mu_{\underset{\sim}{A}}(x) \wedge \mu_{\underset{\sim}{B}}(x) \tag{3.68}$$

$$\mu_{\overline{\underset{\sim}{A}}}(x) = 1 - \mu_{\underset{\sim}{A}}(x) \tag{3.69}$$

Bai and Wang [243] discussed that the union between fuzzy sets is computed by identifying the entity containing the highest membership value among all related elements in the sets. Similarly, the intersection between fuzzy sets is produced by selecting the object with the lowest membership value among all elements in the sets [243]. The complement of a fuzzy set describes the degree to which a specific element does not belong to the related set [244]. Figures 3.19, 3.20, and 3.21 demonstrate the union, intersection, and complement of fuzzy sets, respectively.

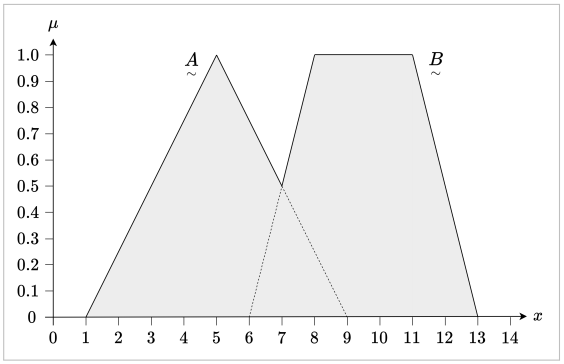


Figure 3.19: Union of $\underset{\sim}{A}$ and $\underset{\sim}{B}$ [237]

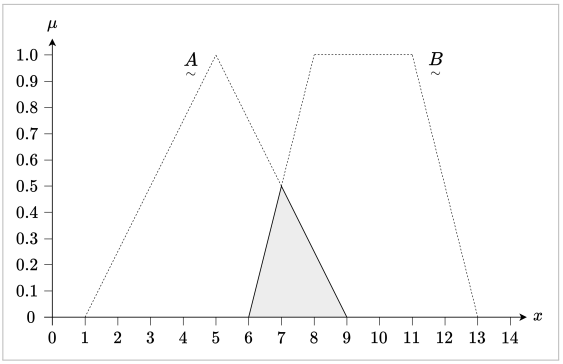


Figure 3.20: Intersection of $\underset{\sim}{A}$ and $\underset{\sim}{B}$ [237]

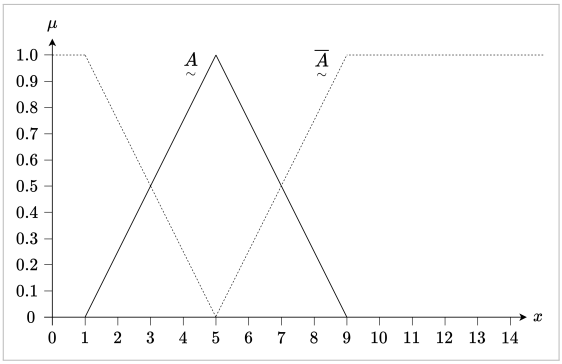


Figure 3.21: Complement of $\underset{\sim}{A}$ [237]

Ross [242] stated that the union, intersection, and complement of fuzzy sets are commonly referred to as standard fuzzy operations. The distinct properties of sets are important due to their underlying influence involving mathematical manipulation [242]. Recall that fuzzy sets embody the same properties as classical sets [242]. Therefore, classical sets are viewed as a special case of fuzzy sets, with the membership values of classical sets a subset of the unit interval [0, 1] [242].

3.5.3 Membership functions

Fuzzy sets are powerful frameworks [243] that mathematically represent the vagueness and uncertainty associated with practical problems [245]. However, Bai and Wang [243] stated that fuzzy sets alone do not present sufficient information concerning the decision-making of empirical problems. Therefore, a fuzzy inference system is needed, which utilises fuzzy sets to map multiple inputs to outputs [246], introducing a basis for general decision-making [245]. Fuzzy inference systems model the underlying relations in data based on a collection of predefined rules [247]. De Oliveira Gomez and Basilio [248] reported that standard fuzzy inference systems contain three phases: fuzzification, inference, and defuzzification.

Fuzzification concerns transforming crisp input values into fuzzy input values using membership functions representing degrees of membership in various fuzzy sets [248]. Inference involves applying fuzzy rules to the resulting fuzzy input values to generate fuzzy output values [248]. Bai and Wang [243] noted that fuzzy rules are usually expressed as a sequence

of if-then statements, where if represents the antecedent input condition and then denotes the consequent fuzzified output. Defuzzification employs unique methods to transform the fuzzy output values into crisp output values with practical applications [248]. Figure 3.22 illustrates the functionality of a standard fuzzy inference system comprising fuzzification, inference, and defuzzification.

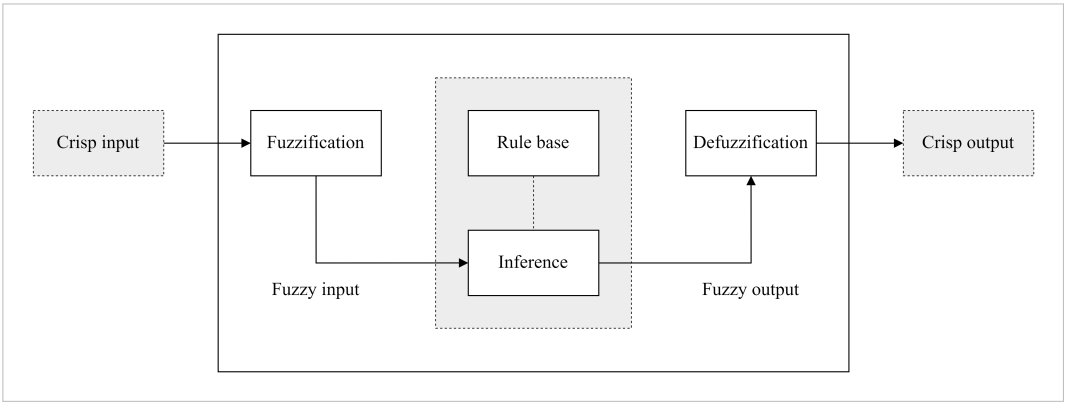


Figure 3.22: Fuzzy inference system [249]

Membership functions denote the foundation of fuzzy inference systems [250], [251]. They characterise fuzziness [252] by detailing the mathematical representation of membership in sets [242]. Specifically, membership functions are curves that map each point in the input space to a related membership value on the unit interval [0, 1] [245]. Membership functions graphically illustrate the extent of participation of each input [232], described as a method to solve problems through experience rather than intuition [253]. Figure 3.23 exhibits the features of a membership function.

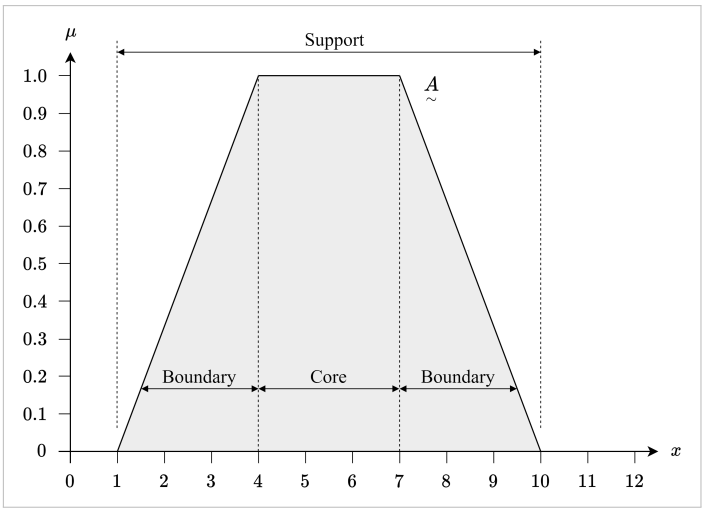


Figure 3.23: Membership function features [254]

From Figure 3.23, the core of a membership function represents the region of the universe characterised by full membership [237]. For some fuzzy set $\underset{\sim}{A}$, the core contains all elements with the expression $\mu_{\underset{\sim}{A}}(x) = 1$ [237]. The support of a membership function represents the region of the universe characterised by non-zero membership [237]. Therefore, the support contains all elements with the expression $\mu_{\underset{\sim}{A}}(x) > 0$ [237]. The boundary of a membership

function is the region of the universe characterised by non-zero but incomplete membership [237]. Hence, the boundary contains all elements with the expression $0 < \mu_{\underset{\sim}{A}}(x) < 1$ [237].

Jain and Sharma [254] mentioned that earlier estimation methods of membership functions heavily relied on statistics rather than optimisation. The shape of a membership function significantly influences system stability [255]. Dragović *et al.* [256] claimed that this shape is selected intuitively or based on the characteristics and relations of the considered data. Ali *et al.* [232] reported that the simplest membership functions are derived using straight lines. Triangular and trapezoidal membership functions have been used extensively due to their computational efficiency and basic formulas [232]. Martínez *et al.* [257] and Tripathy *et al.* [258] acknowledged that the triangular and trapezoidal variants are the most popular membership functions by virtue of their simplicity in representation.

Prajapati and Fernandez [259] stated that the triangular form is the most frequently used among all membership functions that implement straight lines. The boundary of triangular membership functions changes linearly between the lowest and highest membership grades [254]. Only a single discrete element may adopt the highest membership value in triangular membership functions [254]. A triangular membership function encompasses three separate points forming a symmetrical or asymmetrical triangle [259]. Figure 3.24 shows the shape of a triangular membership function containing the points a, b, and c.

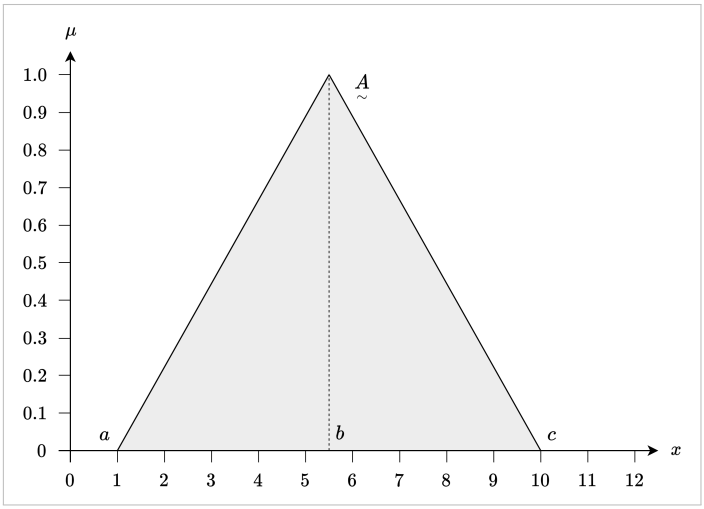


Figure 3.24: Triangular membership function [254]

From Figure 3.24, the interval between a and c represents the support of the membership function, with b the point having the highest membership value such that $\mu_{\underset{\sim}{A}}(x) = 1$ [259]. Equation (3.70) defines the expression for estimating the membership $\mu_{\underset{\sim}{A}}(x)$ of any element x in some fuzzy set $\underset{\sim}{A}$, mapped using a triangular function [260].

$$\mu_{\underset{\sim}{A}}(x) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a < x \leq b \\ \frac{c-x}{c-b}, & b < x < c \\ 0, & c \leq x \end{cases} \tag{3.70}$$

Porębski and Straszecka [261] noted that the trapezoidal membership function is considered a generalisation of the triangular form. Like triangular functions, trapezoidal membership functions have a linear boundary [254]. However, a range of elements can assume maximum membership in trapezoidal membership functions [254]. As such, trapezoidal membership functions are most suited for mapping fuzzy sets in which numerous elements may embody a membership value of one [262]. Figure 3.25 depicts the shape of a trapezoidal membership function containing the four points a, b, c, and d.

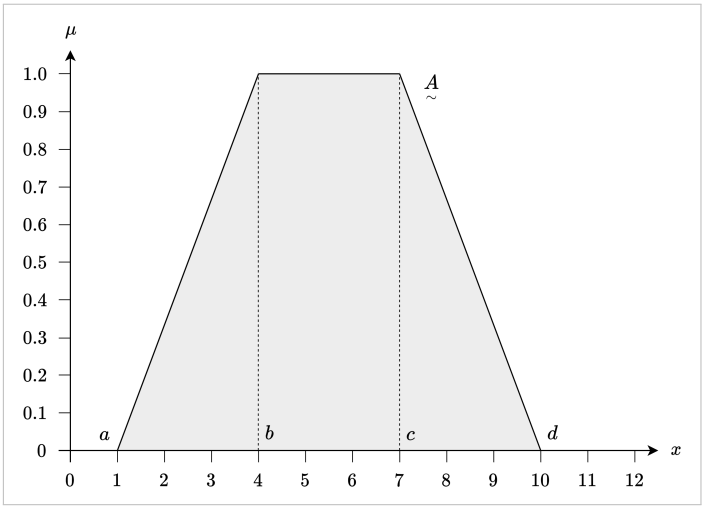


Figure 3.25: Trapezoidal membership function [254]

From Figure 3.25, trapezoidal membership functions are described by a lower constraint a, upper constraint d, lower support constraint b, and upper support constraint c, such that $a < b < c < d$ [263]. Equation (3.71) defines the expression for estimating the membership $\mu_{\underset{\sim}{A}}(x)$ of any element x in some fuzzy set $\underset{\sim}{A}$, mapped using a trapezoidal function [260].

$$\mu_{\underset{\sim}{A}}(x) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a < x < b \\ 1, & b \leq x \leq c \\ \frac{d-x}{d-c}, & c < x < d \\ 0, & d \leq x \end{cases} \tag{3.71}$$

Ali *et al.* [232] noted that since triangular and trapezoidal membership functions comprise multiple straight-line segments, they lack smoothness at the corner points specified by the system parameters. The linearity of these functions causes abrupt changes in membership values rather than gradual transitions [264]. Hameed [265] said that Gaussian membership functions address the limitations of triangular and trapezoidal functions by capturing subtle data variations through continuously differentiable curves that smoothly represent gradual shifts in membership. Due to their smoothness and concise notation, Sadollah [266] stated that Gaussian membership functions are widely used to map various fuzzy sets. Unlike the triangular and trapezoidal variants that comprise compact supports, Gaussian membership functions assign non-zero membership to all points in the universe of discourse [267].

Gaussian membership functions are non-discrete, bell-shaped curves [268], characterised by a symmetrical distribution of fuzzy set points [269]. The boundary of Gaussian membership functions is nonlinear [254], providing systems with fewer degrees of freedom and increased robustness [270]. Sonar *et al.* [271] mentioned that the continuous derivatives of Gaussian membership functions render mathematical analysis more traceable than other membership functions. Figure 3.26 illustrates the shape of a Gaussian membership function, represented by the peak c and standard deviation σ.

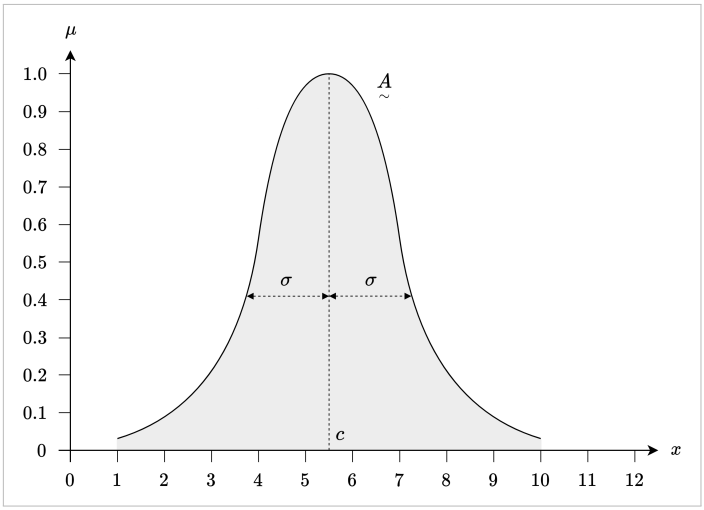


Figure 3.26: Gaussian membership function [272]

From Figure 3.26, c represents the centre of the bell-shaped curve, while σ determines its shape [267]. Xiang *et al.* [273] stated that the shape of a Gaussian membership function can approximate other membership functions, including triangular and trapezoidal. Equation (3.72) illustrates the expression for computing the membership $\mu_{\underset{\sim}{A}}(x)$ of any element x in some fuzzy set $\underset{\sim}{A}$, mapped using a Gaussian function [260], [274].

$$\mu_{\underset{\sim}{A}}(x) = e^{\frac{-(x-c)^2}{2\sigma^2}} \tag{3.72}$$

Gaussian membership functions are more computationally expensive than linear functions, which only involve single-order polynomials [274]. The gradual shifts in cricketing abilities can be captured by the smooth, bell-shaped curve of Gaussian membership functions. The continuous feature of Gaussian membership functions assists in making subtle distinctions between cricketers comprising comparable performance data, proving essential for accurate player selection. Moreover, cricket performance data tend to follow a normal distribution, where most players comprise average figures, with a select few under- and overperforming. The expected normal distribution renders the Gaussian membership function preferable to alternative functions. Hence, this study uses Gaussian membership functions to transform the cricket performance data, thereby prioritising a smooth and continuous representation of information at the expense of computational efficiency.

3.6 Chapter summary

Operations research is a broad field containing various components, including optimisation theory, of which mathematical optimisation forms a subset concerned with finding the best solution subject to several constraints [140]. Mathematical optimisation contains numerous methods, such as linear, integer, and goal programming [28]. Linear programming is widely recognised as the foundation of mathematical optimisation [141], with the simplex method deemed the most popular algorithm for solving linear programs [164]. Integer programming is an extension of linear programming [168], with the branch-and-bound method viewed as the most prominent algorithm for solving integer programs. Goal programming represents the oldest and most popular multi-criteria decision-making technique, with non-preemptive and preemptive goal programs solved using the conventional and modified simplex methods, respectively [29]. Fuzzy logic is a method that deals with ambiguity and uncertainty using fuzzy inference systems [237], which are based on membership functions [250], [251]. While numerous membership functions exist, the Gaussian membership function is selected as the preferred method to transform the cricket performance data. Next, Chapter 4 presents the non-preemptive and preemptive goal programming models.

Chapter 4

Mathematical models

Contents

4.1 Model 1: Non-preemptive goal programming model 68
4.2 Model 2: Preemptive goal programming model 72
4.3 Chapter summary 73

Chapter 4 introduces the proposed goal programming models for optimum team selection in the Indian Premier League (IPL). The chapter specifies the mathematical notation and formulation related to each model. It presents a non-preemptive solution, formulated using the budgetary, auctioning, and proprietary constraints in Section 2.5. Chapter 4 concludes by providing a corresponding preemptive solution that assigns different priority factors to the various cricketing abilities.

4.1 Model 1: Non-preemptive goal programming model

Model 1 utilises cardinal weights to produce a team that sufficiently balances the trade-off between the various conflicting cricketing abilities. During the IPL auction, each franchise must assemble a team by purchasing players from a set of available participants with unique cricketing abilities and nationalities. Therefore, let $I = \{1, 2, ..., |I|\}$ be the set of available players sold during the auction, where $i \in I$ represents each player. Let $J = \{1, 2, ..., |J|\}$ be the set of cricketing abilities, where $j \in J$ describes each ability. Let $K = \{1, 2, ..., |K|\}$ be the set of nationality classes, where $k \in K$ denotes each class. Table 4.1 defines the sets applicable to Model 1.

Table 4.1: Sets applicable to Model 1

Set	Indexing	Description
I	$i \in I$	Set representing the available players sold during the specific auction.

J	$j \in J$	Set representing the cricketing abilities applicable to the specific auction.
K	$k \in K$	Set representing the nationality classes applicable to the specific auction.

The auctioning price of the i-th cricketer is given as θ_i, with the corresponding membership value ζ_i characterising player performance. The arbitrarily large target value related to the j-th goal detailing the desired performance level of the j-th cricketing ability is described as ψ_j, while the underachieving weight w_j^- specifies the importance of the particular deviation. The number of players required and retained for the j-th cricketing ability in the squad of selected participants are defined as κ_j and τ_j, respectively. Likewise, the number of players retained for the k-th nationality class is described as μ_k. The selection condition of the k-th nationality class is given as γ_k, with δ_k the coefficient used to manipulate sign restrictions. Furthermore, σ_{ijk} represents the characteristic constant that assumes the value one if the i-th player has the j-th cricketing ability and the k-th nationality class, and zero otherwise. Table 4.2 defines the parameters applicable to Model 1.

Table 4.2: Parameters applicable to Model 1

Parameter	Indexing	Description
θ_i	$i \in I$	Parameter for the auctioning price of each player.
ζ_i	$i \in I$	Parameter for the membership value of each player.
ψ_j	$j \in J$	Parameter for the target value of each goal constraint.
w_j^-	$j \in J$	Parameter for the underachieving weight of each deviational variable.
κ_j	$j \in J$	Parameter for the number of players required for each cricketing ability.
τ_j	$j \in J$	Parameter for the number of players retained for each cricketing ability.
μ_k	$k \in K$	Parameter for the number of players retained for each nationality class.
γ_k	$k \in K$	Parameter for the selection condition of each nationality class.
δ_k	$k \in K$	Parameter for the manipulation coefficient of each nationality class.
σ_{ijk}	$i \in I$, $j \in J$, $k \in K$	Parameter for the characteristic constant of each player containing each cricketing ability and each nationality class.

The minimum and maximum number of players permitted in each team are presented as η and ι, respectively. Furthermore, β specifies the amount deducted for the number of players retained, characterised as α. The auctioning purse available to each franchise before player retention is expressed as υ, while the minimum spend requirement is described as ϵ. Table 4.3 defines the constants applicable to Model 1.

Table 4.3: Constants applicable to Model 1

Constant	Description
η	Constant for the minimum number of players permitted in each team.
ι	Constant for the maximum number of players permitted in each team.
β	Constant for the amount deducted for the number of players retained.
α	Constant for the number of players retained before the start of the auction.

υ	Constant for the auctioning purse available to each franchise before player retention.
ϵ	Constant for the minimum spend requirement of each franchise before player retention.

The binary decision variable $x_{ijk} \in \{0, 1\}$ assumes the value one if the i-th player containing the j-th cricketing ability and the k-th nationality class is chosen, and zero otherwise. The binary regulatory variable $r \in \{0, 1\}$ takes the value one if the purchasing total of the team selected satisfies the adjusted minimum spend requirement after player retention, and zero otherwise. The continuous deviational variables $d_j^-, d_j^+ \in \mathbb{R}_+$ specify the degrees of under- and overachievement of the target value ψ_j involving the j-th goal. The continuous penalty variable $p \in \mathbb{R}_+$ is the amount by which the purchasing total of the team selected lies below the adjusted minimum spend requirement following player retention. Table 4.4 defines the variables applicable to Model 1.

Table 4.4: Variables applicable to Model 1

Variable	Indexing	Domain	Description
r	-	{0,1}	Variable indicating whether the purchasing total satisfies the adjusted minimum spend requirement.
p	-	$\mathbb{R}_+$	Variable indicating the amount by which the purchasing total lies below the adjusted minimum spend requirement.
d_j^-	$j \in J$	$\mathbb{R}_+$	Variable indicating the degrees of underachievement of the target value relating to each goal.
d_j^+	$j \in J$	$\mathbb{R}_+$	Variable indicating the degrees of overachievement of the target value relating to each goal.
x_{ijk}	$i \in I$, $j \in J$, $k \in K$	{0,1}	Variable indicating whether a specific player with a particular cricketing ability and nationality class is selected.

Considering Tables 4.1 to 4.4, Equations (4.1) to (4.14) define the formulation of Model 1.

$$\text{Minimise} \quad \sum_{j \in J} w_j^- d_j^- \tag{4.1}$$

$$\text{Subject to} \quad \sum_{i \in I} \sum_{k \in K} \zeta_i x_{ijk} + d_j^- - d_j^+ = \psi_j, \quad \forall j \in J \tag{4.2}$$

$$\sum_{i \in I} \sum_{k \in K} x_{ijk} \geq \kappa_j - \tau_j, \quad \forall j \in J \tag{4.3}$$

$$\sum_{i \in I} \sum_{j \in J} \delta_k x_{ijk} \geq \delta_k \gamma_k - \delta_k \mu_k, \quad \forall k \in K \tag{4.4}$$

$$\sum_{i \in I} \sum_{j \in J} \sum_{k \in K} x_{ijk} \leq \iota - \alpha \tag{4.5}$$

$$\sum_{i \in I} \sum_{j \in J} \sum_{k \in K} x_{ijk} \geq \eta - \alpha \tag{4.6}$$

$$\sum_{i \in I} \sum_{j \in J} \sum_{k \in K} \theta_i x_{ijk} \leq \upsilon - \beta \tag{4.7}$$

$$\sum_{i \in I} \sum_{j \in J} \sum_{k \in K} \theta_i x_{ijk} + p \geq \epsilon - \beta \tag{4.8}$$

$$\sum_{i\in I}\sum_{j\in J}\sum_{k\in K}\theta_i x_{ijk} \geq r(\epsilon-\beta) \tag{4.9}$$

$$\sum_{i\in I}\sum_{j\in J}\sum_{k\in K}\theta_i x_{ijk} - \epsilon + \beta - r(\upsilon-\epsilon) \leq -p \tag{4.10}$$

$$(\upsilon-\beta)(1-r) \geq p \tag{4.11}$$

$$x_{ijk} \leq \sigma_{ijk}, \quad \forall i \in I, \quad \forall j \in J, \quad \forall k \in K \tag{4.12}$$

$$x_{ijk}, r \in \{0,1\}, \quad \forall i \in I, \quad \forall j \in J, \quad \forall k \in K \tag{4.13}$$

$$d_j^-, d_j^+, p \in \mathbb{R}_+, \quad \forall j \in J \tag{4.14}$$

Equation (4.1) presents the objective function of Model 1. The objective function minimises the sum of the weighted underachievement $w_j^- d_j^-$ across all goals. The objective function yields the combination of players for which the smallest underachieving deviation from the target value ψ_j of each goal is observed. The j-th goal corresponding to the j-th cricketing ability is outlined in Equation (4.2). Equation (4.2) specifies that for all cricketing abilities, the membership value total obtained by summing the membership value ζ_i across all players selected must equal the target value ψ_j. The deviational variables d_j^- and d_j^+ capture the under- and overachievement of the target value ψ_j for all cricketing abilities, respectively.

Equation (4.3) governs the preference of the players selected by enforcing the desired squad composition. Equation (4.3) specifies that for all cricketing abilities, the sum of all players selected is at least the difference between the number of players required κ_j and retained τ_j. Equation (4.3) ensures that the number of players purchased for each cricketing ability meets the requirements of the franchise. Equivalently, Equation (4.4) limits the number of players selected for each nationality class. Equation (4.4) indicates that for all nationality classes, the sum of the players selected must satisfy the applicable sign restriction relating to the difference between the selection condition γ_k and the number of players retained μ_k. The manipulation coefficient δ_k controls the behaviour of the inequality in Equation (4.4).

Equations (4.5) and (4.6) provide the limitations governing the number of players permitted in each squad. Equation (4.5) states that the sum of all players selected is no greater than the difference between the maximum number of players allowed ι and the number of players retained α. Equation (4.5) ensures that the number of players purchased during the auction does not surpass the maximum positions remaining after player retention. Equation (4.6) specifies that the sum of all players selected is at least the difference between the minimum number of players required η and the number of players retained α. Equation (4.6) ensures that the number of players purchased during the auction is at least the minimum positions required after player retention.

Equations (4.7) and (4.8) describe the budgetary restrictions imposed on all franchises by considering the purchasing total, characterised as the sum of the auctioning price θ_i across all players selected. Equation (4.7) stipulates that the purchasing total is less than or equal to the difference between the auctioning purse υ and the deduction amount β. Therefore, Equation (4.7) guarantees that the price of all purchased players does not exceed the purse

available after player retention. Equation (4.8) states that the sum of the purchasing total and the penalty p is at least the difference between the minimum spend requirement ϵ and the deduction amount β. Therefore, Equation (4.8) ensures that the price of the purchased players, considering the penalty, is no less than the adjusted minimum spend requirement following player retention.

Equations (4.9), (4.10), and (4.11) jointly describe the constraints governing the behaviour of the regulatory variable r and the computation of the penalty p. Equation (4.9) stipulates that the regulatory variable r assumes the value one if the purchasing total is no less than the difference between the minimum spend requirement ϵ and the deduction amount β, and zero otherwise. Equation (4.9) warrants the relevance of a penalty where a franchise fails to meet the mandatory spending condition of the specific auction. Equation (4.10) defines the penalty value for p as a function of the purchasing total, the minimum spend requirement ϵ, and the deduction amount β if the regulatory variable r is zero. Inversely, if the regulatory variable r equals one, the penalty value for p is computed by including the auctioning purse υ in Equation (4.10) while removing the minimum spend requirement ϵ entirely. Equation (4.11) applies the regulatory variable r to force the penalty p to zero, where the purchasing total is greater than or equal to the adjusted minimum spend requirement after retention. However, if the purchasing total is smaller than the adjusted minimum spend requirement after retention, Equation (4.11) subsequently limits the penalty p to the difference between the auctioning purse υ and the deduction amount β.

Equation (4.12) represents the eligibility constraint governing the selection of each player. Equation (4.12) states that for all players, cricketing abilities, and nationality classes, the value of the decision variable x_{ijk} may not exceed the corresponding characteristic constant σ_{ijk}. The role of σ_{ijk} is to associate each player with their corresponding cricketing ability and nationality class. As such, Equation (4.12) guarantees that each player can be selected at most once, and only for their specific cricketing ability and nationality class. Equations (4.13) and (4.14) specify the domains of the variables. Equation (4.13) defines the decision variable x_{ijk} and the regulatory variable r as binary, whereas Equation (4.14) describes the deviational variables d_j^- and d_j^+, and the penalty variable p as non-negative real values.

4.2 Model 2: Preemptive goal programming model

Model 2 applies preemptive priority factors to provide insight into how prioritising different cricketing abilities influences team selection. Akin to Model 1, Model 2 seeks a squad that adequately balances the trade-off between the conflicting cricketing abilities subject to the same budgetary, auctioning, and proprietary constraints in Section 2.5. However, Model 2 assigns preemptive priority factors to the different cricketing abilities based on their related importance. Thus, let $Z = \{1, 2, ...|Z|\}$ be the set of priority levels, where $z \in Z$ represents each level and $|Z| \leq |J|$. The preemptive priority factor of non-numeric nature describing the z-th priority level is defined as P_z. Therefore, the underachieving weight corresponding to the z-th priority level for the j-th cricketing ability is defined as w_{jz}^-, with the associated

deviation expressed as d_{jz}^-. All remaining sets, parameters, constants, and variables adopt the same meaning and significance as in Model 1, described in Section 4.1. Model 2 shares the same formulation as Model 1, except for the objective function. Equation (4.15) details the objective function of Model 2.

$$\text{Minimise} \quad \sum_{j \in J} P_z(w_{jz}^- d_{jz}^-) \tag{4.15}$$

The objective function minimises the sum of the weighted underachievement $w_{jz}^- d_{jz}^-$ across all goals, prioritised according to the preemptive priority factor P_z, which specifies the z-th priority level. The interpretation of the objective function in Equation (4.15) differs from that in Equation (3.64). Equation (3.64) requires each goal to have a unique priority level. Applying the logic of Schniederjans [275], Equation (4.15) permits goals to share the same priority level by distinguishing their index from that of the preemptive priority factors.

The objective function in Equation (4.15) delivers the combination of players in accordance with the preemptive priority structure for which the smallest underachieving deviation from the target value ψ_j of each goal is observed. The preemptive priority structure ensures that the most important goals are first achieved by treating each cricketing ability according to its priority level, following the preconceived lexicographic order. The constraints for Model 2 are identical to those of Model 1 in Equations (4.2) to (4.14).

4.3 Chapter summary

Models 1 and 2 enable team selection by effectively balancing the trade-off between multiple conflicting cricketing abilities. The models encompass identical formulations, with the only difference being the objective function. Model 1 exclusively employs cardinal weights, while Model 2 also incorporates hierarchical structuring through preemptive priority factors. The preemptive priority structure of Model 2 allows prioritisation of cricketing abilities, leading to more tailored selections. Although the objective functions are unique, both models are subjected to the same budgetary, auctioning, and proprietary constraints. Next, Chapter 5 discusses the data considerations and verification of Models 1 and 2.

Chapter 5

Data considerations and verification

Contents

5.1 Computational hardware 74
5.2 Data considerations 75
5.3 Verification 79
5.4 Chapter summary 99

Chapter 5 describes the data considerations and verification relating to the models proposed in Chapter 4. It discusses the computational hardware utilised to obtain the findings. The chapter articulates the data collection, cleaning, and transformation. Chapter 5 concludes by employing a critical constraint investigation to verify the correctness of the formulation and coding of both models.

5.1 Computational hardware

The models are developed, programmed, and solved using Elytica by applying the Python programming language and CPLEX optimisation solver. The necessary statistical analyses are performed using the R programming language within the RStudio integrated development environment. All data handling is conducted through Microsoft Excel. Both models are implemented using a Dell Latitude 3520 device comprising the following specifications:

- Processor: 11th Generation Intel(R) Core(TM) i7-1165G7 @ 2.80 GHz 1.69 GHz.
- Operating system: Windows 11 Pro, 64-bit system, x64-based processor.
- Installed random-access memory (RAM): 8.00 GB.

5.2 Data considerations

This study employs open-access player performance data presented by Statsguru [276], the statistical database of ESPNcricinfo [277]. The data given by ESPNcricinfo is of a practical nature and indicative of actual match results. Statsguru provides player performance data using an interactive platform that enables filtering and analysis across various cricketing formats. This platform is universally recognised for its authenticity [278], [279], [280], [281], sustaining its application within this particular study. Players requiring performance data were identified by establishing all participants sold during the 2022 Indian Premier League (IPL) mega auction using the online platform of the contest [282]. Across all ten franchises, 204 cricketers were identified as sold for a total of ₹551.70 crore. Using the IPL platform, personal data, including the purchasing franchise, cricketing ability, nationality class, and auctioned price, were collected for each player. The cricketing abilities outlined by the IPL in the 2022 mega auction were used since current player profiles may have changed due to the continuous evolution of skills. For example, a player viewed as a bowler in 2022 might have since developed sufficient batting capabilities, consequently becoming an all-rounder.

All 204 players were assigned an identification number and classified based on their related cricketing abilities to obtain four datasets corresponding to bowling, batting, all-rounding, and wicketkeeping. The performance data for all participants in each dataset were retrieved from Statsguru using appropriate search filters. The performance data was refined for each player using relevant format, time, and cricketing ability provisions. The data was filtered to include only Twenty20 (T20) statistics from domestic and international games. Assessing both domestic and international data provides a more accurate representation of the actual capabilities of players, mitigating the possibility of misinterpretations.

Since the 2022 IPL mega auction started on 12 February, the performance data was filtered until 11 February to replicate the practical availability of information. An advanced filter was applied to wicketkeepers, ensuring that the data collected only represents the games in which these players served as keepers. Figure 5.1 specifies the number of players purchased for each cricketing ability during the 2022 IPL mega auction across all 10 franchises, with the corresponding performance data collected for each participant.

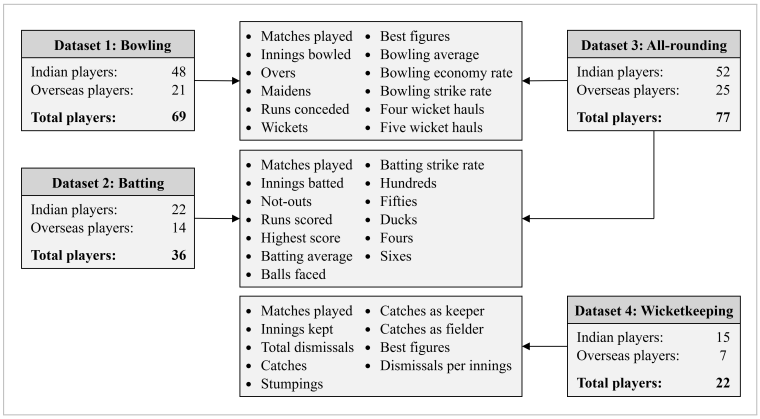


Figure 5.1: Player performance data collected

Bhattacharjee and Saikia [17] stated that to fairly gauge the genuine capabilities of players, performance should be measured across a series of games. Data concerning a small number of games may be misleading, resulting in the possibility of incorrect conclusions [17]. Saikia *et al.* [54] acknowledged that some inclusion criteria are required to identify suitable players for selection. The inclusion criteria guarantee that playing experience is considered during selection by excluding inexperienced players. Therefore, the inclusion criteria foster fairer player comparisons, which leads to a more objective selection process. After retrieving the performance data, the following inclusion criteria, obtained from Bhattacharjee and Saikia [17], were applied to ensure only appropriate players are considered:

1. Criteria employed for bowlers:
    - Players should have bowled a minimum of 5 innings.
    - Players should have delivered a minimum of 60 balls.
2. Criteria employed for batsmen:
    - Players should have batted a minimum of 5 innings.
    - Players should have faced a minimum of 60 balls.
3. Criteria employed for all-rounders:
    - Players should have conformed to the criteria for bowlers.
    - Players should have conformed to the criteria for batsmen.
4. Criteria employed for wicketkeepers:
    - Players should have kept a minimum of 5 innings.
    - Players should have dismissed a minimum of 3 batsmen.

After applying the preceding inclusion criteria, 173 players remained available for selection, of which 63, 33, 58, and 19 belong to Datasets 1, 2, 3, and 4, respectively. Datasets 1, 2, 3, and 4 consist of the cricketers available for selection for bowling, batting, all-rounding, and wicketkeeping, respectively. The performance scores were derived using the CBR, CBM, AR, and WK metrics described in Chapter 2. For each cricketing ability, the performance scores were transformed using a Gaussian membership function, where the best score in the dataset serves as the centre of the function and the sample standard deviation determines its spread. Decreasing scores of CBR indicate better performance, while increasing scores of CBM, AR, and WK specify better performance. Therefore, subtracting the minimum value for Dataset 1 and the maximum values for Datasets 2, 3, and 4 ensures that players with the best performance scores also have the highest membership values. Figures 5.2 to 5.5 demonstrate the membership values computed for the players in each dataset, visualised in ascending order.

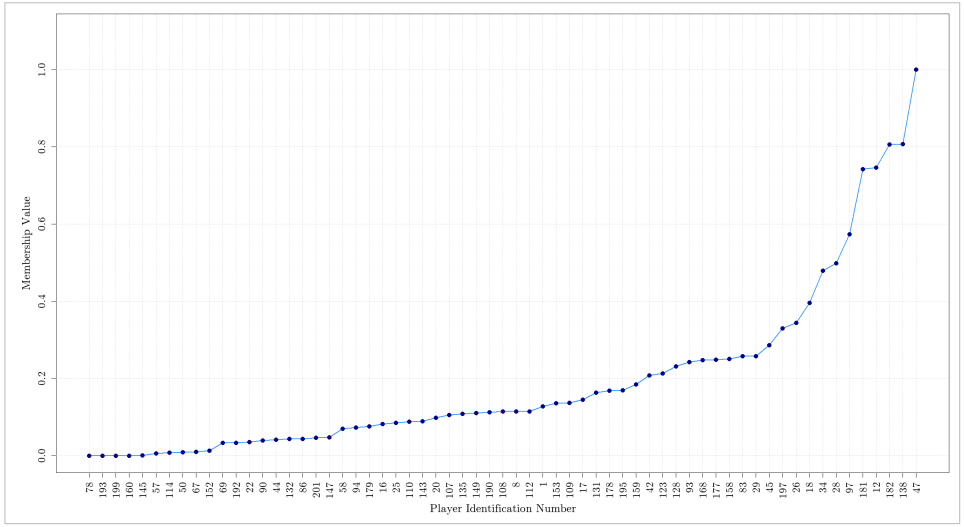


Figure 5.2: Membership values for bowers

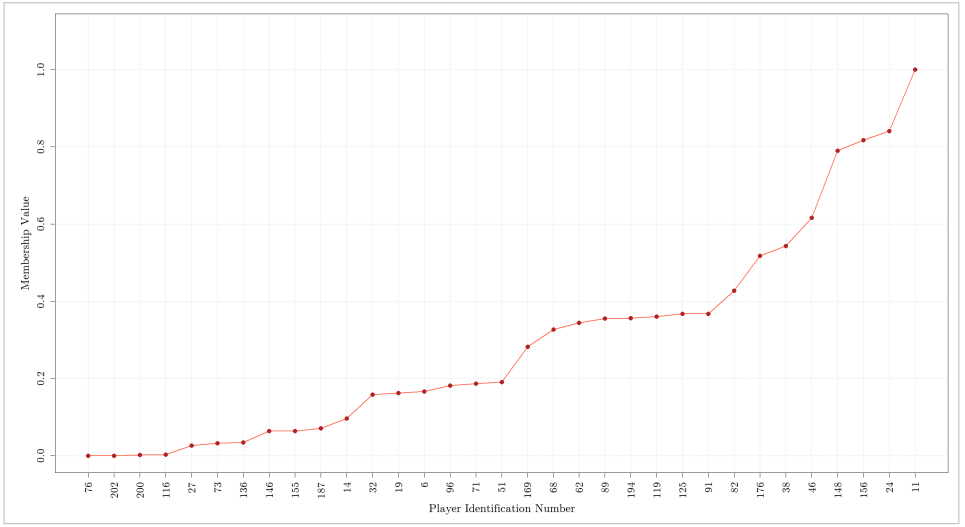


Figure 5.3: Membership values for batsmen

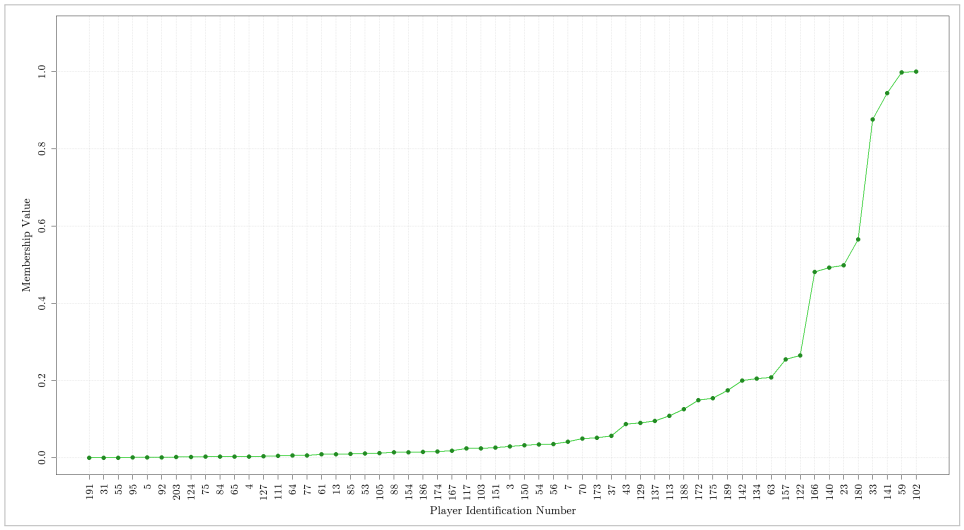


Figure 5.4: Membership values for all-rounders

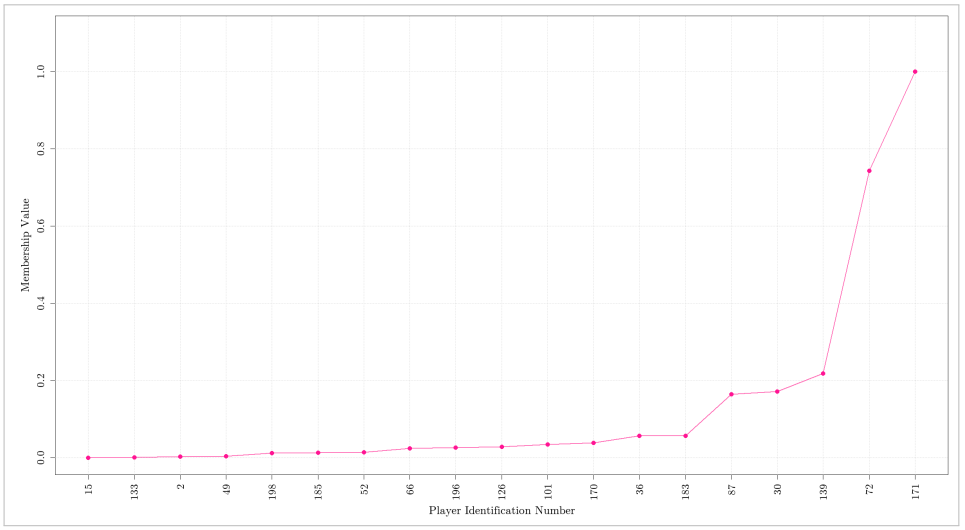


Figure 5.5: Membership values for wicketkeepers

Figures 5.2 to 5.5 show that membership values more accurately indicate superior players, highlighting relationships between cricketers that may not be evident from the performance scores alone. Applying membership functions to transform the performance scores enables player comparisons by converting the data to a uniform scale. The membership values are scaled between zero and one, which gives the advantage of players with different cricketing abilities being compared on an identical scale. Mapping the data to the unit interval [0, 1] allows Gaussian membership functions to reduce the uncertainty in the performance scores by examining the performance of each cricketer in relation to that of the best-performing player. Finally, the four datasets were combined into a single set that describes all players available for selection. The membership values were used as input to the models formulated in Chapter 4. Figure 5.6 outlines a business process modelling notation (BPMN) diagram that explains the data collection, cleaning, and transformation steps performed.

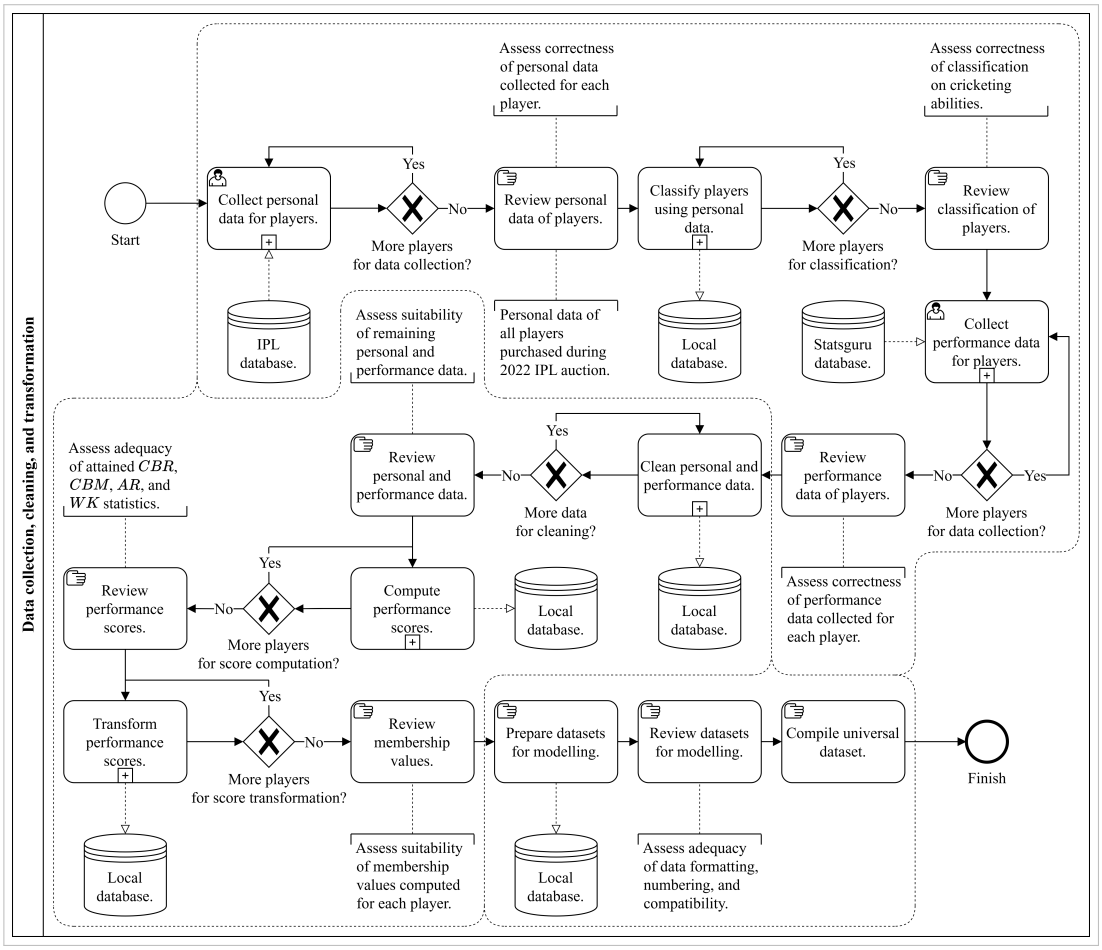


Figure 5.6: Overview of data collection, cleaning, and transformation

Figure 5.6 provides a high-level representation of the various sub-processes used to compile the finalised results. Figure 5.7 presents a related operations process chart, showing a more detailed illustration of the activities mentioned in Figure 5.6. Therefore, Figure 5.7 shows the dependencies between the different data operations, illustrating the tasks employed to ensure the suitability of the data analysis conducted.

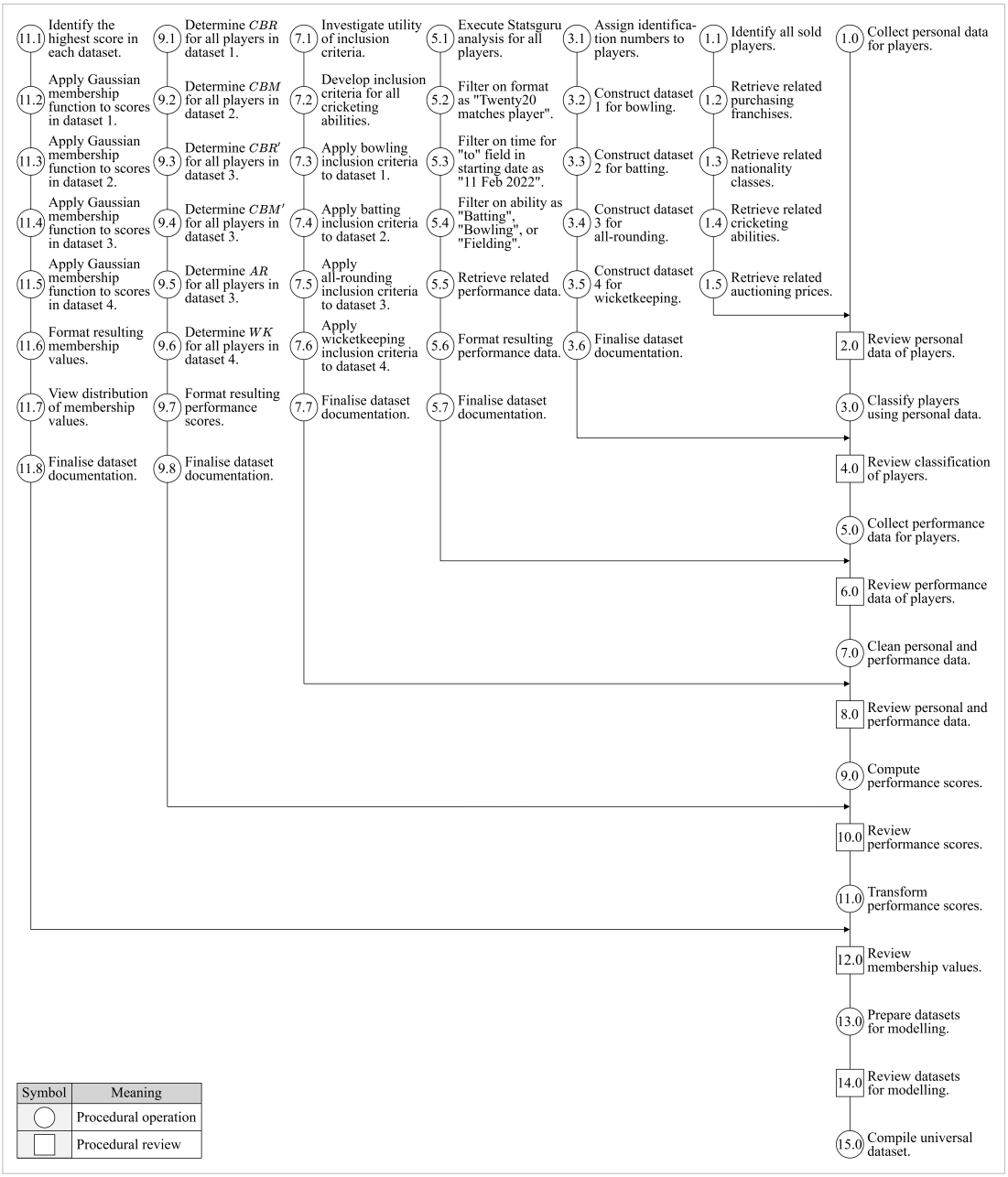


Figure 5.7: Sub-process activities of data collection, cleaning, and transformation

Figures 5.6 and 5.7 show the systematic procedure applied to obtain the results in Appendix B. Appendix B depicts the final datasets pertaining to bowling, batting, all-rounding, and wicketkeeping, for all 173 players considered for selection.

5.3 Verification

Section 5.3 verifies the correctness in formulation and coding of both models by critically investigating the influence of each constraint on the results.

5.3.1 Model 1: Non-preemptive goal programming model

Verification involves establishing the correctness of the formulation and coding relating to the models. Sample data is used to simplify the process, enabling a better understanding of the particular characteristics underpinning the models. Applying the complete dataset for verification may complicate the process unnecessarily without providing additional insights into the behaviour of the models. From the set of 173 cricketers considered for selection, 10 players were randomly chosen for each cricketing ability to construct a subset encompassing 40 players for verification. The subset used to verify the models is given in Appendix B.

The IPL allows franchises the liberty of selecting their desired squads by following the rules and regulations of the auction. Every franchise employs a unique purchasing strategy with various requirements influenced by internal and external factors. Since the provisions and expectations of each franchise fundamentally differ, several assumptions and considerations are enforced to illustrate the practical behaviour of the models. The following assumptions and considerations relate to Model 1:

1. All cricketing abilities are of equal importance, that is $w_j^- = 0.25$ for all $j \in J$.
2. All cricketing abilities have the same target value, that is $\psi_j = 100$ for all $j \in J$.
3. The minimum number of players required for each cricketing ability in the squad of 18 to 25 players is defined as follows:
   - Minimum number of bowlers: Four.
   - Minimum number of batsmen: Four.
   - Minimum number of all-rounders: Seven.
   - Minimum number of wicketkeepers: Three.
4. A maximum of four players are retained from the previous season. Table 5.1 specifies the auctioning price, cricketing ability, and nationality class for each retained player.

Table 5.1: Player retainment characteristics

Characteristic	**Player 1**	**Player 2**	**Player 3**	**Player 4**
Auctioning price	₹16.00 crore	₹12.00 crore	₹8.00 crore	₹6.00 crore
Cricketing ability	Bowling	Batting	Batting	All-rounding
Nationality class	Indian	Indian	Overseas	Overseas

Model 1 is verified by examining the impact of the constraints in Equations (4.2) to (4.12) on the results using the objective function in Equation (4.1). The constraints in Equations (4.13) and (4.14) remain fixed during verification. The verification process is not executed

in the sequential order of the formulated constraints. Instead, the constraints are included based on the nature of the results to facilitate a logical understanding of the dynamics and functionality of the model. Table 5.2 shows the constraints added for each verification test to systematically inspect their impact on the results.

Table 5.2: Verification tests for Model 1

Test	Tested formulations	Applicable equations
1	Objective function and goal constraint.	(4.1), (4.2), (4.13), and (4.14).
2	Eligibility constraint.	(4.1), (4.2), (4.12), (4.13), and (4.14).
3	Selection constraints.	(4.1), (4.2), (4.5), (4.6), (4.12), (4.13), and (4.14).
4	Budgetary and coordination constraints.	(4.1), (4.2), (4.5), (4.6), (4.7), (4.8), (4.9), (4.10), (4.11), (4.12), (4.13), and (4.14).
5	Cricketing ability and nationality class constraints.	(4.1), (4.2), (4.3), (4.4), (4.5), (4.6), (4.7), (4.8), (4.9), (4.10), (4.11), (4.12), (4.13), and (4.14).

Table 5.2 illustrates the verification process for Model 1. In each test, new constraints are introduced while maintaining the equations implemented in the previous test. For example, Test 2 implements the eligibility constraint in Equation (4.12), in addition to the equations included in Test 1, i.e., Equations (4.1), (4.2), (4.13), and (4.14). This procedure continues until Test 5, where all constraints are active and the complete model is evaluated.

Test 1: Objective function and goal constraint

The objective function in Equation (4.1) minimises the overall weighted underachievement compared to the target values for the goals. From Equation (4.1), the model is expected to select the squad of cricketers comprising the smallest total deviation to obtain the minimum objective function value. Table 5.3 illustrates the results of the model after incorporating Equation (4.1).

Table 5.3: Results for Test 1 with Equation (4.1) included

Selection results					
Objective value		0.00	Purchasing total		₹0.00 crore
Deviational variables					
Underachievement			**Overachievement**		
Goal	**Variable**	**Value**	**Goal**	**Variable**	**Value**
1	d_1^-	0.00	1	d_1^+	0.00
2	d_2^-	0.00	2	d_2^+	0.00
3	d_3^-	0.00	3	d_3^+	0.00
4	d_4^-	0.00	4	d_4^+	0.00

From Table 5.3, the model delivered an objective function value of 0.00, with no deviation observed for any of the variables. The model did not select any cricketers as demonstrated by the purchasing total of ₹0.00 crore. Since only the objective function is considered, the model excluded all players, causing the deviational variables to adopt the value 0.00. The goal constraint in Equation (4.2) is included to obtain a collection of players.

Equation (4.2) describes how the under- and overachievement are specified for the various goals. Since the target values are arbitrarily large, the model is expected to select cricketers comprising the highest membership values with the implementation of Equation (4.2). The rationale sustaining this selection logic lies in the minimisation of the underachievement of target values as a consequence of increased membership value totals. Table 5.4 illustrates the results of the model after incorporating Equation (4.2).

Table 5.4: Results for Test 1 with Equation (4.2) included

Selection results					
Objective value	74.334985		Purchasing total		₹1303.60 crore
Deviational variables					
Underachievement			**Overachievement**		
Goal	**Variable**	**Value**	**Goal**	**Variable**	**Value**
1	d_1^-	74.334985	1	d_1^+	0.00
2	d_2^-	74.334985	2	d_2^+	0.00
3	d_3^-	74.334985	3	d_3^+	0.00
4	d_4^-	74.334985	4	d_4^+	0.00

From Table 5.4, the model selected all players for each of the four cricketing abilities and the two nationality classes, which delivers an objective function value of 74.33. Each of the 40 players was selected eight times, once for each cricketing ability across both nationality classes, causing a total of 320 selections. The 320 selections yielded the purchasing total of ₹1303.60 crore. The purchasing total of ₹1303.60 crore represents the product between the cost of ₹162.95 crore for all 40 players considered and the eight times each participant was selected. In addition, the sum of the membership values relating to the 40 players equates to 12.83. Since all 40 players were selected for both nationality classes, a membership value total of 25.67 was computed for each cricketing ability, leading to an underachievement of 74.33 for each goal. The appropriate selection of target values is supported by the deviation of 0.00 computed for all variables that capture overachievement. Therefore, Tables 5.3 and 5.4 verify the correctness of Equations (4.1) and (4.2) in delivering a collection of players that minimises the total underachievement of the target values. Table 5.4 indicates that the model selected players for unrelated cricketing abilities and nationality classes. Hence, the results in Table 5.4 are not sensible as each player should be selected for their designated cricketing ability and nationality class, revealing the need for an additional constraint that governs the eligibility of the selections.

Test 2: Eligibility constraint

The eligibility constraint in Equation (4.12) guarantees that each player is only selected for the designated cricketing ability and nationality class reported. By incorporating Equation (4.12), the model is expected to select each of the 40 players only once in accordance with their corresponding cricketing ability and nationality class. Table 5.5 illustrates the results of the model after including Equation (4.12).

Table 5.5: Results for Test 2 with Equation (4.12) included

Selection results					
Objective value	96.791873		Purchasing total		₹162.95 crore
Deviational variables					
Underachievement			**Overachievement**		
Goal	**Variable**	**Value**	**Goal**	**Variable**	**Value**
1	d_1^-	96.744828	1	d_1^+	0.00
2	d_2^-	95.729428	2	d_2^+	0.00
3	d_3^-	96.415116	3	d_3^+	0.00
4	d_4^-	98.278121	4	d_4^+	0.00

From Table 5.5, the model yielded an objective function value of 96.79, with a purchasing total of ₹162.95 crore. The purchasing total of ₹162.95 crore supports the single selection of all 40 cricketers in the verification subset by precisely matching the sum of their auctioning prices. All 40 players were selected for their corresponding cricketing ability and nationality class, yielding the membership value totals of 3.26, 4.27, 3.58, and 1.72 for bowling, batting, all-rounding, and wicketkeeping, respectively. These membership value totals produced the underachieving deviations of 96.74, 95.73, 96.42, and 98.28 for the four goals, respectively. Hence, Table 5.5 verifies the correctness of Equation (4.12) in ensuring each player is only selected for their related cricketing ability and nationality class.

Test 3: Selection constraints

The upper selection constraint in Equation (4.5) ensures that the number of players selected is no more than the maximum squad positions remaining after player retention. With four players retained, the model is expected to select at most 21 cricketers to meet the maximum limit of 25 players. Table 5.6 shows the results of the model after including Equation (4.5).

Table 5.6: Results for Test 3 with Equation (4.5) included

Selection results			
Objective value	97.261220	Purchasing total	₹72.55 crore

Deviational variables					
Underachievement			Overachievement		
Goal	**Variable**	**Value**	**Goal**	**Variable**	**Value**
1	d_1^-	97.166569	1	d_1^+	0.00
2	d_2^-	95.731250	2	d_2^+	0.00
3	d_3^-	97.147060	3	d_3^+	0.00
4	d_4^-	99.000000	4	d_4^+	0.00

From Table 5.6, the model yielded an objective function value of 97.26. The model selected the 21 best players across all four cricketing abilities for a purchasing total of ₹72.55 crore, including seven bowlers, nine batsmen, four all-rounders, and a wicketkeeper. The membership value totals are 2.83, 4.27, 2.85, and 1.00, delivering the underachieving deviations of 97.17, 95.73, 97.15, and 99.00 for the four goals, respectively. Table 5.7 shows the players selected by the model.

Table 5.7: Players selected for Test 3 using Equation (4.5)

Bowling			Batting		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
18	S. Singh	Indian	24	D.A. Warner	Overseas
26	C. Sakariya	Indian	38	K.A. Hebbar	Indian
29	K. Yadav	Indian	62	S.S. Iyer	Indian
34	L.T. Ngidi	Overseas	68	A.D. Hales	Overseas
45	Y. Dayal	Indian	82	P. Singh	Indian
181	C.V. Milind	Indian	89	M.K. Pandey	Indian
197	S. Gopal	Indian	119	R.R. Buddhi	Indian
			148	D. Padikkal	Indian
			169	F. du Plessis	Overseas
All-rounding			**Wicketkeeping**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
23	M.R. Marsh	Overseas	171	A. Rawat	Indian
33	L. Yadav	Indian			
59	D.G. Nalkande	Indian			
166	P.W. Hasaranga	Overseas			

Table 5.7 indicates that among the 21 best players in the verification subset, 15 are Indian, while six are overseas. Tables 5.6 and 5.7 verify the correctness of Equation (4.5) in limiting the number of players selected to the maximum positions available after player retention.

The lower selection constraint in Equation (4.6) ensures that the number of players selected is no less than the minimum squad positions remaining after player retention. Thus, since

four players were retained, the model is expected to select at least 14 cricketers to satisfy the minimum condition of 18 players required in each squad. Applying Equation (4.6) gave the same results as in Table 5.6 and the same team as in Table 5.7. The equivalent findings resulted from the model aiming to minimise the overall underachievement by selecting the maximum number of players allowed in an attempt to maximise the membership value totals relating to the goals. To demonstrate the functionality of the constraint, the inequality in Equation (4.6) is changed to the equality in Equation (5.1).

$$\sum_{i \in I} \sum_{j \in J} \sum_{k \in K} x_{ijk} = \eta - \alpha \tag{5.1}$$

From Equation (5.1), x_{ijk} is the binary decision variable that assumes the value one if the i-th player containing the j-th cricketing ability and the k-th nationality class is selected, and zero otherwise. Moreover, η describes the minimum number of players required in the squad, while α gives the number of players retained. Equation (5.1) enabled the model to select 14 players. Table 5.8 shows the results of the model after including Equation (5.1).

Table 5.8: Results for Test 3 with Equation (5.1) included

Selection results					
Objective value		97.803795	Purchasing total		₹41.65 crore
Deviational variables					
Underachievement			**Overachievement**		
Goal	**Variable**	**Value**	**Goal**	**Variable**	**Value**
1	d_1^-	98.384153	1	d_1^+	0.00
2	d_2^-	96.683967	2	d_2^+	0.00
3	d_3^-	97.147060	3	d_3^+	0.00
4	d_4^-	99.000000	4	d_4^+	0.00

From Table 5.8, the model yielded an objective function value of 97.80. The model selected the 14 best players across all four cricketing abilities, leading to a purchasing total of ₹41.65 crore. The model selected three bowlers and six batsmen, yielding membership value totals of 1.62 and 3.32, respectively. These membership value totals produced the underachieving deviations of 98.38 and 96.68 for the two goals, respectively. The model selected the same all-rounders and wicketkeepers as in Table 5.7, with the same membership value totals and underachieving deviations as in Table 5.6. Table 5.9 lists the players selected by the model.

Table 5.9: Players selected for Test 3 using Equation (5.1)

Bowling			**Batting**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**

18	S. Singh	Indian	24	D.A. Warner	Overseas
34	L.T. Ngidi	Overseas	38	K.A. Hebbar	Indian
181	C.V. Milind	Indian	82	P. Singh	Indian
			89	M.K. Pandey	Indian
			119	R.R. Buddhi	Indian
			148	D. Padikkal	Indian
All-rounding			**Wicketkeeping**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
23	M.R. Marsh	Overseas	171	A. Rawat	Indian
33	L. Yadav	Indian			
59	D.G. Nalkande	Indian			
166	P.W. Hasaranga	Overseas			

The team in Table 5.9 differs from the one in Table 5.7. Including Equation (5.1) enabled the model to exclude the seven worst cricketers in Table 5.7, resulting in the selection of 10 Indian and four overseas players. The model removed the bowlers Sakariya, Yadav, Dayal, and Gopal, and the batsmen Iyer, Hales, and du Plessis from the squad. Thus, Tables 5.8 and 5.9 verify the correctness of Equation (4.6) in limiting the number of players selected to fill the minimum squad positions available after player retention. Equation (4.6) is again included in the model, with Equation (5.1) excluded for the rest of the verification process. Equation (5.1) only served as a convenient means to show the expected but correct results from Equation (4.6).

Test 4: Budgetary and coordination constraints

The upper budgetary constraint in Equation (4.7) ensures that the purchasing total spent to acquire the auctioning participants does not exceed the purse available after the deduction of the cost concerning the players retained. Since the total cost of all four players retained equates to ₹42.00 crore, the model is expected to select a squad of cricketers for which the purchasing total is no more than ₹48.00 crore to comply with the maximum available purse of ₹90.00 crore. Table 5.10 shows the results of the model after including Equation (4.7).

Table 5.10: Results for Test 4 with Equation (4.7) included

Selection results					
Objective value		97.376758	Purchasing total		₹45.50 crore
Deviational variables					
Underachievement			**Overachievement**		
Goal	**Variable**	**Value**	**Goal**	**Variable**	**Value**

1	d_1^-	97.166569	1	d_1^+	0.00
2	d_2^-	96.356956	2	d_2^+	0.00
3	d_3^-	97.373655	3	d_3^+	0.00
4	d_4^-	98.609850	4	d_4^+	0.00

From Table 5.10, the model yielded an objective function value of 97.38, with a purchasing total of ₹45.50 crore. The purchasing total of ₹45.50 crore registered ₹2.50 crore lower than the available purse of ₹48.00 crore after player retention. The model selected 21 cricketers, including seven bowlers, seven batsmen, four all-rounders, and three wicketkeepers. Hence, the membership value totals are 2.83, 3.64, 2.63, and 1.39, resulting in the underachieving deviations of 97.17, 96.36, 97.37, and 98.61 for the four goals, respectively. Table 5.11 lists the players selected by the model.

Table 5.11: Players selected for Test 4 using Equation (4.7)

Bowling			**Batting**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
18	S. Singh	Indian	24	D.A. Warner	Overseas
26	C. Sakariya	Indian	38	K.A. Hebbar	Indian
29	K. Yadav	Indian	68	A.D. Hales	Overseas
34	L.T. Ngidi	Overseas	82	P. Singh	Indian
45	Y. Dayal	Indian	89	M.K. Pandey	Indian
181	C.V. Milind	Indian	119	R.R. Buddhi	Indian
197	S. Gopal	Indian	148	D. Padikkal	Indian
All-rounding			**Wicketkeeping**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
23	M.R. Marsh	Overseas	30	K.S. Bharat	Indian
33	L. Yadav	Indian	139	J.M. Sharma	Indian
59	D.G. Nalkande	Indian	171	A. Rawat	Indian
157	D.J. Mitchell	Overseas			

The team in Table 5.11 differs from the one in Table 5.7. Including Equation (4.7) led to the model selecting the same bowlers while replacing the two batsmen, Iyer and du Plessis, with the two wicketkeepers, Bharat and Sharma. Also, the model swapped the all-rounder Hasaranga with a less expensive counterpart in Mitchell, yielding the selection of 16 Indian and five overseas players. As such, Tables 5.10 and 5.11 verify the correctness of Equation (4.7) in limiting the purchasing total to the purse available following player retention.

The lower budgetary constraint in Equation (4.8) ensures that the sum of the purchasing total and the relevant penalty is at least the adjusted minimum spend requirement following the price deduction for the cricketers retained. Therefore, since the cost of all four players

retained equals ₹42.00 crore, the model is expected to select a squad of cricketers for which the sum of the purchasing total and the enforced penalty is at least ₹25.50 crore to satisfy the minimum spend requirement of ₹67.50 crore. Incorporating Equation (4.8) yielded an identical team to the one in Table 5.11 and the same results as in Table 5.10. Although the model produced a penalty value of ₹0.00 crore, p remains unbounded. Therefore, even if the purchasing total satisfies the adjusted minimum spend requirement, a possibility exists for the penalty variable to assume a value greater than zero. The constraints in Equations (4.9), (4.10), and (4.11) are applied to ensure the correct calculation of the penalty value.

Equations (4.9), (4.10), and (4.11) complement one another by jointly using the regulatory variable to compute the appropriate penalty value relating to the purchasing total employed by the franchise. Since these constraints do not impact the players selected, their addition to the model yielded the same results as in Table 5.10, with an equivalent team selected as in Table 5.11. For the purpose of brevity, only the outcome of the regulatory variable and penalty value is subsequently discussed with the addition of each constraint. Following the inclusion of Equations (4.9), (4.10), and (4.11), the model is expected to allocate the value one to the regulatory variable with a penalty value of ₹0.00 crore since the purchasing total is greater than the adjusted minimum spend requirement.

Equation (4.9) establishes the suitability of a penalty by employing the regulatory variable to inspect whether the purchasing total satisfies the adjusted minimum spend requirement following the deduction of the cost of the retained players. Including Equation (4.9) yielded the value zero for the regulatory variable, while retaining the penalty value of ₹0.00 crore. Specifically, both integers zero and one prove appropriate for the regulatory variable when the purchasing total exceeds the adjusted minimum spend requirement. Consequently, the optimisation solver assigns the smallest value of zero to the regulatory variable.

Equation (4.10) determines the value of the penalty variable using its domain to force the regulatory variable to zero or one, depending on the purchasing total. Including Equation (4.10) yielded the value one for the regulatory variable with a contemporary penalty value of ₹2.50 crore. The domain of the penalty variable now prohibits the model from allocating the value zero to the regulatory variable when the purchasing total is greater than or equal to the adjusted minimum spend requirement. If the model specified the regulatory variable equal to zero, the resulting expression of $p \leq -$₹20.00 crore for Equation (4.10) would have produced an infeasible solution given the domain $p \in \mathbb{R}_+$.

Equation (4.11) employs the regulatory variable value of one to force the resulting penalty value to ₹0.00 crore. Using the value one for the regulatory variable simplifies the expression in Equation (4.11) to $p \leq$ ₹0.00 crore, constraining the penalty value to ₹0.00 crore owing to the domain $p \in \mathbb{R}_+$. Table 5.12 illustrates the results of the model after incorporating Equations (4.8) to (4.11).

Table 5.12: Results for Test 4 with Equations (4.8) to (4.11) included

Selection results					
Objective value	97.376758		Purchasing total		₹45.50 crore
Regulatory value	1		Penalty value		₹0.00 crore
Deviational variables					
Underachievement			**Overachievement**		
Goal	**Variable**	**Value**	**Goal**	**Variable**	**Value**
1	d_1^-	97.166569	1	d_1^+	0.00
2	d_2^-	96.356956	2	d_2^+	0.00
3	d_3^-	97.373655	3	d_3^+	0.00
4	d_4^-	98.609850	4	d_4^+	0.00

Similar to Table 5.10, Table 5.12 indicates that the model produced an objective function value of 97.38, with a purchasing total of ₹45.50 crore. The purchasing total of ₹45.50 crore resulted in a regulatory variable value of one, yielding a penalty value of ₹0.00 crore. Again, the model selected 21 players, namely seven bowlers, seven batsmen, four all-rounders, and three wicketkeepers. The membership value totals are 2.83, 3.64, 2.63, and 1.39 for bowling, batting, all-rounding, and wicketkeeping, respectively. These membership value totals led to the related underachieving deviations of 97.17, 96.36, 97.37, and 98.61 for the four goals, respectively. Table 5.13 shows the players selected by the model.

Table 5.13: Players selected for Test 4 using Equations (4.8) to (4.11)

Bowling			**Batting**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
18	S. Singh	Indian	24	D.A. Warner	Overseas
26	C. Sakariya	Indian	38	K.A. Hebbar	Indian
29	K. Yadav	Indian	68	A.D. Hales	Overseas
34	L.T. Ngidi	Overseas	82	P. Singh	Indian
45	Y. Dayal	Indian	89	M.K. Pandey	Indian
181	C.V. Milind	Indian	119	R.R. Buddhi	Indian
197	S. Gopal	Indian	148	D. Padikkal	Indian
All-rounding			**Wicketkeeping**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
23	M.R. Marsh	Overseas	30	K.S. Bharat	Indian
33	L. Yadav	Indian	139	J.M. Sharma	Indian
59	D.G. Nalkande	Indian	171	A. Rawat	Indian
157	D.J. Mitchell	Overseas			

From Table 5.13, the model correctly selected the same cricketers as in Table 5.11, with 16 Indian and five overseas players. Therefore, Tables 5.12 and 5.13 verify the correctness of

Equations (4.8) to (4.11) in not assigning a penalty in the case where the purchasing total satisfies the adjusted minimum spend requirement. However, to also illustrate the correct working of Equations (4.8) to (4.11) when a franchise fails to meet the adjusted minimum spend requirement, Equation (4.6) is changed to the equality in Equation (5.2).

$$\sum_{i \in I} \sum_{j \in J} \sum_{k \in K} x_{ijk} = 7 \tag{5.2}$$

Employing Equation (5.2) allows the model to select precisely seven cricketers. Table 5.14 provides the results of the model after applying Equation (5.2).

Table 5.14: Results for Test 4 with Equation (5.2) included

Selection results					
Objective value	98.552860		Purchasing total		₹18.70 crore
Regulatory value	0		Penalty value		₹6.80 crore
Deviational variables					
Underachievement			**Overachievement**		
Goal	**Variable**	**Value**	**Goal**	**Variable**	**Value**
1	d_1^-	99.258683	1	d_1^+	0.00
2	d_2^-	97.826659	2	d_2^+	0.00
3	d_3^-	98.126097	3	d_3^+	0.00
4	d_4^-	99.000000	4	d_4^+	0.00

From Table 5.14, the model yielded an objective function value of 98.55. The selected team failed to meet the adjusted minimum spend requirement with a purchasing total of ₹18.70 crore, resulting in a regulatory variable value of zero. Therefore, the penalty value equals ₹6.80 crore, which denotes the value by which the purchasing total is less than the adjusted minimum spend requirement. The model selected seven players, including a bowler, three batsmen, two all-rounders, and a wicketkeeper. The membership value totals equal 0.74, 2.17, 1.87, and 1.00 for bowling, batting, all-rounding, and wicketkeeping, respectively. These membership value totals gave the underachieving deviations of 99.26, 97.83, 98.13, and 99.00 for the four goals, respectively. Table 5.15 lists the players selected by the model.

Table 5.15: Players selected for Test 4 using Equation (5.2)

Bowling			**Batting**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
181	C.V. Milind	Indian	24	D.A. Warner	Overseas
			38	K.A. Hebbar	Indian
			148	D. Padikkal	Indian

All-rounding			Wicketkeeping		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
33	L. Yadav	Indian	171	A. Rawat	Indian
59	D.G. Nalkande	Indian			

The team in Table 5.15 differs from the one in Table 5.13. Using Equation (5.2) led to the model selecting the seven best players across all cricketing abilities, i.e., the bowler Milind, the batsmen Warner, Hebbar, and Padikkal, the all-rounders Yadav and Nalkande, and the wicketkeeper Rawat. Table 5.15 shows that among the seven best cricketers, six are Indian and one is overseas. Accordingly, Tables 5.14 and 5.15 verify the correctness of Equations (4.8) to (4.11) in enforcing an appropriate penalty in the case where the purchasing total fails to satisfy the adjusted minimum spend requirement. As previously, Equation (4.6) is included in the model, while Equation (5.2) is removed for the remainder of the verification process. Equation (5.2) merely assisted in demonstrating the expected but correct results of Equations (4.8) to (4.11).

Test 5: Cricketing ability and nationality class constraints

The cricketing ability constraint in Equation (4.3) stipulates the number of players required for each ability according to the desired squad composition. With one, two, one, and zero cricketers retained for bowling, batting, all-rounding, and wicketkeeping, respectively, the model is expected to select at least three, two, six, and three players for each corresponding cricketing ability. This selection ensures the final combination of players meets the desired team structure of four batsmen, four bowlers, seven all-rounders, and three wicketkeepers. Table 5.16 shows the results of the model after including Equation (4.3).

Table 5.16: Results for Test 5 with Equation (4.3) included

Selection results					
Objective value		97.406950	Purchasing total		₹47.90 crore
Regulatory value		1	Penalty value		₹0.00 crore
Deviational variables					
Underachievement			**Overachievement**		
Goal	**Variable**	**Value**	**Goal**	**Variable**	**Value**
1	d_1^-	97.452075	1	d_1^+	0.00
2	d_2^-	96.711953	2	d_2^+	0.00
3	d_3^-	96.738490	3	d_3^+	0.00
4	d_4^-	98.725282	4	d_4^+	0.00

From Table 5.16, the model gave an objective function value of 97.41. The model delivered a purchasing total of ₹47.90 crore and a regulatory variable value of one, leading to a penalty

value of ₹0.00 crore. Moreover, the model selected 21 players, viz., six bowlers, six batsmen, six all-rounders, and three wicketkeepers. The model appointed the minimum number of cricketers required for all-rounding and wicketkeeping while selecting three and four extra players for bowling and batting, respectively. The membership value totals are 2.55, 3.29, 3.26, and 1.27 for bowling, batting, all-rounding, and wicketkeeping, respectively. These membership value totals yielded the underachieving deviations of 97.45, 96.71, 96.74, and 98.73 for the four goals, respectively. Table 5.17 lists the players selected by the model.

Table 5.17: Players selected for Test 5 using Equation (4.3)

Bowling			**Batting**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
18	S. Singh	Indian	24	D.A. Warner	Overseas
26	C. Sakariya	Indian	38	K.A. Hebbar	Indian
29	K. Yadav	Indian	68	A.D. Hales	Overseas
34	L.T. Ngidi	Overseas	82	P. Singh	Indian
181	C.V. Milind	Indian	119	R.R. Buddhi	Indian
197	S. Gopal	Indian	148	D. Padikkal	Indian
All-rounding			**Wicketkeeping**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
23	M.R. Marsh	Overseas	36	T.L. Seifert	Overseas
33	L. Yadav	Indian	139	J.M. Sharma	Indian
59	D.G. Nalkande	Indian	171	A. Rawat	Indian
157	D.J. Mitchell	Overseas			
166	P.W. Hasaranga	Overseas			
175	M.L. Lomror	Indian			

The team in Table 5.17 differs from the one in Table 5.13. Applying Equation (4.3) caused the model to include two additional all-rounders, Hasaranga and Lomror, while excluding the bowler Dayal and batsman Pandey. Also, the model replaced the previously selected wicketkeeper Bharat with a less expensive option in Seifert, resulting in the selection of 14 Indian and seven overseas players. These selections ensured that the complete squad of 25 players met the minimum requirements of four bowlers, four batsmen, seven all-rounders, and three wicketkeepers. Therefore, Tables 5.16 and 5.17 verify the correctness of Equation (4.3) in meeting the cricketing ability requirements of the specific franchise.

The nationality class constraint in Equation (4.4) specifies the number of local and foreign players allowed in each squad. Since two Indian players were retained, the model is expected to select at least 15 local players to meet the minimum requirement of 17 Indian cricketers. Moreover, since two overseas players were retained, the model is expected to select at most six foreign players to meet the maximum limit of eight overseas cricketers. Table 5.18 gives the results of the model after applying Equation (4.4).

Table 5.18: Results for Test 5 with Equation (4.4) included

Selection results			
Objective value	97.418090	Purchasing total	₹47.90 crore
Regulatory value	1	Penalty value	₹0.00 crore
Deviational variables			
Underachievement		**Overachievement**	
Goal	**Variable**	**Value**	**Goal**	**Variable**	**Value**
1	d_1^-	97.452075	1	d_1^+	0.00
2	d_2^-	96.711953	2	d_2^+	0.00
3	d_3^-	96.738490	3	d_3^+	0.00
4	d_4^-	98.769843	4	d_4^+	0.00

From Table 5.18, the model produced an objective function value of 97.42, with the same purchasing total of ₹47.90 crore as in Table 5.16. The purchasing total of ₹47.90 crore once more produced a regulatory variable value of one, with a penalty value of ₹0.00 crore. The model selected the same bowlers, batsmen, and all-rounders as in Table 5.17, yielding the same membership value totals and underachieving deviations as in Table 5.16. The model altered the selection of the three wicketkeepers in Table 5.17, yielding a lower membership value total of 1.23 and a higher underachieving deviation of 98.77 for the goal. Table 5.19 shows the players selected by the model.

Table 5.19: Players selected for Test 5 using Equation (4.4)

Bowling			**Batting**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
18	S. Singh	Indian	24	D.A. Warner	Overseas
26	C. Sakariya	Indian	38	K.A. Hebbar	Indian
29	K. Yadav	Indian	68	A.D. Hales	Overseas
34	L.T. Ngidi	Overseas	82	P. Singh	Indian
181	C.V. Milind	Indian	119	R.R. Buddhi	Indian
197	S. Gopal	Indian	148	D. Padikkal	Indian
All-rounding			**Wicketkeeping**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
23	M.R. Marsh	Overseas	139	J.M. Sharma	Indian
33	L. Yadav	Indian	171	A. Rawat	Indian
59	D.G. Nalkande	Indian	198	V. Vinod	Indian
157	D.J. Mitchell	Overseas			
166	P.W. Hasaranga	Overseas			
175	M.L. Lomror	Indian			

The team in Table 5.19 differs from the one in Table 5.17. Applying Equation (4.4) enabled the model to replace the overseas player, Seifert, with the same-priced Indian wicketkeeper, Vinod. Selecting Vinod instead of Seifert ensured that the squad of 25 cricketers included 17 Indian and eight overseas players. Thus, Tables 5.18 and 5.19 verify the correctness of Equation (4.4) in restricting the number of cricketers selected for each nationality class. In conclusion, Tables 5.18 and 5.19 provide the final results and team selected when applying the complete model in Equations (4.1) to (4.14) to the verification subset in Appendix B.

5.3.2 Model 2: Preemptive goal programming model

Model 2 contains a comparable formulation to Model 1, with the only difference being its objective function, which has preemptive priority factors that stipulate the priority of each cricketing ability. Model 2 is verified using the following two scenarios:

- Scenario 1: Where each goal contains a unique priority level.
- Scenario 2: Where multiple goals share the same priority level, but not uniformly.

The assumptions and considerations for Model 2 are identical to those of Model 1, including the following condition:

1. Each cricketing ability contains a preemptive priority factor P_z of non-numeric nature that describes the z-th priority level. Table 5.20 details the preemptive priority factor assigned to each goal for both scenarios.

   Table 5.20: Preemptive priority factors for verification

Scenario	Bowling	Batting	All-rounding	Wicketkeeping
1	P_1	P_2	P_3	P_4
2	P_1	P_1	P_2	P_3

   From Table 5.20, each cricketing ability in Scenario 1 contains a unique priority level, with bowling assigned the highest priority (P_1), followed by batting (P_2), all-rounding (P_3), and wicketkeeping (P_4). In Scenario 2, bowling and batting share the highest priority (P_1), followed by all-rounding (P_2) and wicketkeeping (P_3).

Applying the assumptions and considerations, Model 2 is verified similarly to Model 1 by critically investigating the impact on the results caused by modifying the objective function in Equation (4.15) using the preemptive priority factors, while preserving the constraints in Equations (4.2) to (4.14). Hence, the constraints in Equations (4.2) to (4.14) remain fixed for both scenarios, with only the objective function in Equation (4.15) adjusted according to the corresponding hierarchical structure in Table 5.20.

Scenario 1: Unique priority levels

The objective function in Equation (4.15) minimises the overall weighted underachievement compared to the target values for the goals prioritised based on importance. From Equation (4.15), the model is expected to select the squad of cricketers comprising the smallest total deviation according to the hierarchical structure to obtain the minimum objective function value. Table 5.21 shows the results of the model for Scenario 1 involving Equation (4.15).

Table 5.21: Results for Scenario 1 of Model 2

Selection results					
Objective value	98.280583		Purchasing total		₹47.40 crore
Regulatory value	1		Penalty value		₹0.00 crore
Deviational variables					
Underachievement			**Overachievement**		
Goal	**Variable**	**Value**	**Goal**	**Variable**	**Value**
1	d_{11}^-	96.744828	1	d_{11}^+	0.00
2	d_{22}^-	99.029453	2	d_{22}^+	0.00
3	d_{33}^-	97.634442	3	d_{33}^+	0.00
4	d_{44}^-	99.713608	4	d_{44}^+	0.00

From Table 5.21, the model yielded an objective function value of 98.28, with a purchasing total of ₹47.40 crore. The model delivered a regulatory variable value of one and a penalty value of ₹0.00 crore. The model selected 21 players, i.e., 10 bowlers, two batsmen, six all-rounders, and three wicketkeepers. The membership value totals are 3.26, 0.97, 2.37, and 0.29 for bowling, batting, all-rounding, and wicketkeeping, respectively. These membership value totals yielded the underachieving deviations of 96.74, 99.03, 97.63, and 99.71 for the four goals, respectively. Table 5.22 lists the players selected by the model.

Table 5.22: Players selected for Scenario 1 of Model 2

Bowling			**Batting**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
1	D.L. Chahar	Indian	38	K.A. Hebbar	Indian
18	S. Singh	Indian	82	P. Singh	Indian
26	C. Sakariya	Indian			
29	K. Yadav	Indian			
34	L.T. Ngidi	Overseas			
45	Y. Dayal	Indian			
147	T.A. Boult	Overseas			
168	J.R. Hazlewood	Overseas			
181	C.V. Milind	Indian			

197	S. Gopal	Indian			
All-rounding			**Wicketkeeping**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
33	L. Yadav	Indian	36	T.L. Seifert	Overseas
54	V. Shankar	Indian	139	J.M. Sharma	Indian
59	D.G. Nalkande	Indian	198	V. Vinod	Indian
70	M. Nabi	Overseas			
157	D.J. Mitchell	Overseas			
175	M.L. Lomror	Indian			

The team in Table 5.22 differs from the one in Table 5.19 due to the preemptive priority factors of Scenario 1. Including the preemptive priority factors of Scenario 1 permitted the model to assign the highest priority to bowling, followed by batting, then all-rounding, and finally, wicketkeeping. Considering the retained cricketers, the minimum number of players required is three, two, six, and three for bowling, batting, all-rounding, and wicketkeeping, respectively. Selecting the fewest players for each cricketing ability leads to seven vacancies, resulting from the maximum number of 21 players allowed in each squad. The model filled all seven available positions with players of the highest prioritised cricketing ability, leading to the selection of all 10 bowlers in the verification subset. The model selected the minimum number of players for the remaining cricketing abilities, namely batting, all-rounding, and wicketkeeping, producing 15 Indian and six overseas players. Table 5.23 shows the selection logic followed by the model to obtain the findings in Tables 5.21 and 5.22.

Table 5.23: Selection logic for Scenario 1 of Model 2

Priority factor	**Cricketing ability**	**Total players selected**	**Purchasing total (₹ crore)**	**Remaining auctioning purse (₹ crore)**
P_1	Bowling	10	40.85	7.15
P_2	Batting	2	41.25	6.75
P_3	All-rounding	6	46.20	1.80
P_4	Wicketkeeping	3	47.40	0.60

Table 5.23 gives a breakdown of the prioritisation enforced by the model. The model selects the minimum number of players required for each lower-prioritised cricketing ability using the least expensive individuals. Therefore, the model can select the best affordable players for the current cricketing ability since the minimum number of individuals required for the remaining abilities is met. The model selected the bowlers for a purchasing total of ₹40.85 crore, leading to a remaining auctioning purse of ₹7.15 crore. Next, the model selected the best affordable batsmen, increasing the purchasing total to ₹41.25 crore while reducing the remaining auctioning purse to ₹6.75 crore. Similarly, the model selected the best affordable all-rounders, yielding a purchasing total of ₹46.20 crore with a remaining auctioning purse

of ₹1.80 crore. The remaining auctioning purse of ₹1.80 crore forced the model to pick the three cheapest wicketkeepers, leading to a final purchasing total of ₹47.40 crore.

For example, prioritising cricketing abilities for Scenario 1 resulted in the model selecting the worst bowler, Boult, with a membership value of 0.05, over the best batsman, Warner, with a membership value of 0.84. Investigating the auctioning prices of Boult and Warner revealed that the latter is also ₹1.75 crore cheaper than the former, indicating the accurate prioritisation of cricketing abilities. Moreover, the combination of wicketkeepers supports the appropriateness of the preemptive priority factors. The model ignored the membership values by selecting the three most affordable wicketkeepers since they comprise the lowest priority. Therefore, Tables 5.21, 5.22, and 5.23 verify the correctness of Equation (4.15) in applying preemptive priority factors where each goal comprises a unique priority.

Scenario 2: Shared priority levels

Modifying the objective function in Equation (4.15) to consider multiple goals at the same priority level, but non-uniformly, introduces additional insight into the appropriateness of the preemptive priority factors. Table 5.24 illustrates the results of the model for Scenario 2 involving Equation (4.15).

Table 5.24: Results for Scenario 2 of Model 2

Selection results					
Objective value	97.573566		Purchasing total		₹47.90 crore
Regulatory value	1		Penalty value		₹0.00 crore
Deviational variables					
Underachievement			**Overachievement**		
Goal	**Variable**	**Value**	**Goal**	**Variable**	**Value**
1	d_{11}^-	97.709946	1	d_{11}^+	0.00
2	d_{21}^-	96.340028	2	d_{21}^+	0.00
3	d_{32}^-	97.634442	3	d_{32}^+	0.00
4	d_{43}^-	98.609850	4	d_{43}^+	0.00

From Table 5.24, the model gave an objective function value of 97.57. The model yielded a purchasing total of ₹47.90 crore, leading to a regulatory variable value of one and a penalty value of ₹0.00 crore. The model selected 21 players, including five bowlers, seven batsmen, six all-rounders, and three wicketkeepers. The membership value totals are 2.29, 3.66, 2.37, and 1.39 for bowling, batting, all-rounding, and wicketkeeping, respectively. Consequently, these membership value totals yielded the underachieving deviations of 97.71, 96.34, 97.63, and 98.61 for the four goals, respectively. Table 5.25 lists the players selected by the model.

Table 5.25: Players selected for Scenario 2 of Model 2

Bowling			**Batting**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
18	S. Singh	Indian	24	D.A. Warner	Overseas
26	C. Sakariya	Indian	38	K.A. Hebbar	Indian
34	L.T. Ngidi	Overseas	62	S.S. Iyer	Indian
181	C.V. Milind	Indian	82	P. Singh	Indian
197	S. Gopal	Indian	89	M.K. Pandey	Indian
			119	R.R. Buddhi	Indian
			148	D. Padikkal	Indian
All-rounding			**Wicketkeeping**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
33	L. Yadav	Indian	30	K.S. Bharat	Indian
54	V. Shankar	Indian	139	J.M. Sharma	Indian
59	D.G. Nalkande	Indian	171	A. Rawat	Indian
70	M. Nabi	Overseas			
157	D.J. Mitchell	Overseas			
175	M.L. Lomror	Indian			

The team in Table 5.25 differs from the one in Table 5.22 due to the preemptive priority factors of Scenario 2. Using the preemptive priority factors of Scenario 2 enabled the model to assign the highest priority equally to bowling and batting, followed by all-rounding and then wicketkeeping. Since bowling and batting collectively contain the highest priority, the model selected the 12 best players across both cricketing abilities, leading to the occupation of the seven vacancies. The model once more selected the minimum number of players for the remaining cricketing abilities, viz., all-rounding and wicketkeeping, delivering 17 Indian and four overseas cricketers. Table 5.26 shows the selection logic followed by the model to obtain the findings in Tables 5.24 and 5.25.

Table 5.26: Selection logic for Scenario 2 of Model 2

Priority factor	**Cricketing ability**	**Total players selected**	**Purchasing total (₹ crore)**	**Remaining auctioning purse (₹ crore)**
P_1	Bowling Batting	12	37.35	10.65
P_2	All-rounding	6	42.30	5.70
P_3	Wicketkeeping	3	47.90	0.10

Table 5.26 gives an overview of the prioritisation applied by the model. The model selected the bowlers and batsmen for a purchasing total of ₹37.35 crore, with a remaining auctioning purse of ₹10.65 crore. The model identified the best affordable all-rounders, increasing the

purchasing total to ₹42.30 crore, while decreasing the remaining auctioning purse to ₹5.70 crore. The remaining auctioning purse of ₹5.70 crore provided enough monetary resources for the model to select the three best wicketkeepers, leading to a final purchasing total of ₹47.90 crore. The selection of the three best wicketkeepers stemmed from the combination of bowlers, batsmen, and all-rounders.

Selecting the 12 best players across bowling and batting delivered a remaining auctioning purse of ₹10.65 crore for the lower-prioritised cricketing abilities. The model included the six all-rounders for a purchasing total of ₹4.95 crore, yielding a remaining auctioning purse of ₹5.70 crore to select the three wicketkeepers. However, the sum of the auctioning prices for the three cheapest wicketkeepers, Sharma, Seifert, and Vinod, equals ₹1.20 crore. Thus, the portion of the remaining auctioning purse available for selecting all-rounders amounts to ₹9.45 crore, owing to their higher priority than wicketkeepers. Investigating the selected all-rounders revealed the worst-performing player as Shankar, with an associated auctioning price of ₹1.40 crore. The cheapest, better-performing alternative is the Australian, Marsh, which contains an auctioning price of ₹6.50 crore. If the model had replaced Shankar with Marsh, a purchasing total of ₹47.40 crore would have transpired. The purchasing total of ₹47.40 crore would have yielded a remaining auctioning purse of ₹0.60 crore, which proves infeasible for selecting wicketkeepers as it lies below the required minimum of ₹1.20 crore. Hence, the model selected Shankar instead of Marsh, allowing the three best wicketkeepers to be selected. Tables 5.24, 5.25, and 5.26 verify the correctness of Equation (4.15) in using preemptive priority factors where multiple goals share the same priority, but not uniformly.

5.4 Chapter summary

This study employs open-access player performance data available on Statsguru [276], the statistical database of ESPNcricinfo [277]. A total of 204 cricketers were identified as sold during the 2022 mega auction using the official platform of the IPL [282]. Next, applying a literature-based inclusion criteria delivered 173 cricketers for selection, of which 63, 33, 58, and 19 represent bowlers, batsmen, all-rounders, and wicketkeepers, respectively. From the collection of 173 cricketers, a subset of 40 players was randomly selected to verify Models 1 and 2 using a critical constraint investigation. The verification process of Model 1 involves assessing the impact of the constraints in Equations (4.2) to (4.14) on the results using the objective function in Equation (4.1). The verification process of Model 2 involves assessing the impact on the results caused by altering the objective function in Equation (4.15) using the preemptive priority factors while retaining the constraints in Equations (4.2) to (4.14). Model 2 is verified, where each goal has a different priority level and multiple goals share the same priority level, but non-uniformly. Both models performed as expected, verifying their correctness in formulation and coding. Subsequently, Chapter 6 discusses the results and validation of Models 1 and 2.

Chapter 6

Results and validation

Contents

6.1 Performance measures assessment 100
6.2 Optimisation results 111
6.3 Validation 119
6.4 Chapter summary 125

Chapter 6 presents the results and validation of the models formulated in Chapter 4. The chapter evaluates the performance measures proposed in Chapter 2. It describes the results obtained from the I-optimal crossed mixture design and the scenario analysis. The chapter concludes by using the winning team of the 2022 Indian Premier League (IPL) to validate the effectiveness of Model 1 in selecting a superior team. The resulting team serves as the baseline to validate the effectiveness of Model 2 in selecting a more specialised team.

6.1 Performance measures assessment

Section 6.1 evaluates the performance measures in Chapter 2 by examining their indexing accuracy, followed by comparative analyses of the newly proposed CBM and AR metrics.

6.1.1 Indexing of performance measures

All auctioned players were classified according to their unique cricketing abilities. Accurate classification of players is vital, as misclassification can lead to exceptional cricketers being overlooked since they need to compete against specialists as non-specialists. Evaluating the indexing accuracy of the performance measures requires examining their ability to classify

players correctly. The classification capabilities of the proposed CBM and AR metrics are examined by assessing their alignment with the IPL classification in 2022. Notably, since AR describes the weighted linear relationship between CBM' and CBR', the classification capability of CBR is also inspected.

Among the 173 players considered for selection, 154 are bowlers, batsmen, and all-rounders. In addition to the previously collected performance data, the relevant batting statistics for all bowlers and bowling statistics for all batsmen were retrieved using the same procedure as in Section 5.2. In cases where a bowler did not record any batting data, the player was excluded from consideration. Similarly, in cases where a batsman did not have any bowling data, the player was excluded from consideration. After collecting the additional data, the CBM and CBR performance scores were obtained for bowlers and batsmen, respectively. Figure 6.1 provides a scatter plot that shows the relationship between CBR and CBM for the remaining 62 bowlers, 23 batsmen, and 58 all-rounders.

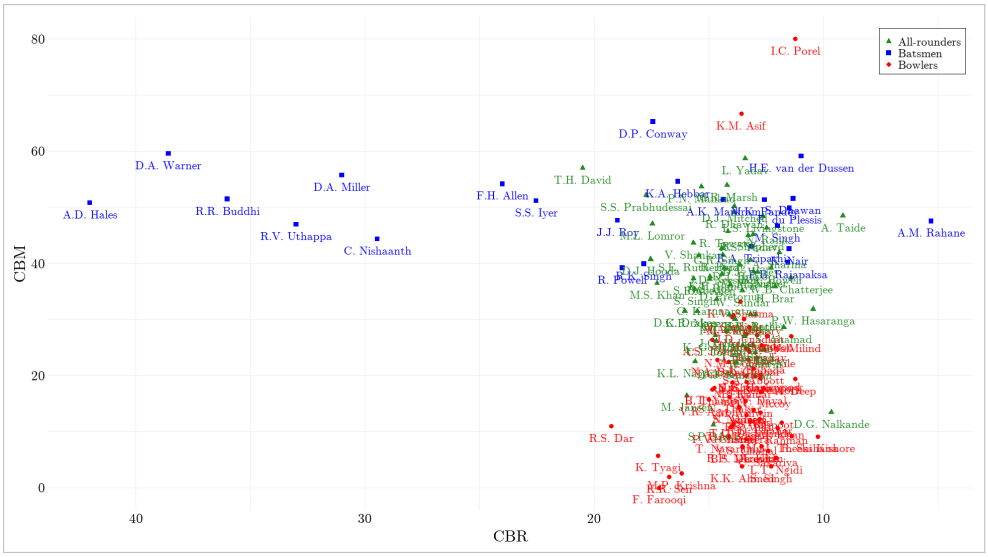


Figure 6.1: Scatterplot of CBR and CBM

Figure 6.1 illustrates the homogeneous grouping of bowlers and all-rounders, with batsmen spread across a range of CBR scores. The positioning of Hales, Warner, Buddhi, Uthappa, Miller, Nishaanth, Allen, and Iyer validates their classification as specialist batsmen, with CBR scores considerably lower than those of bowlers and all-rounders. In addition, Figure 6.1 indicates that the batsman Rahane has the lowest CBR score. Investigating the bowling data revealed that Rahane bowled only one innings, completing one over while conceding five runs and dismissing one batsman. Therefore, Rahane is an outlier due to the limited bowling data. Figure 6.1 reveals the potential for the batsmen Singh (M), Pandey, Dhawan, Rajapaksa, Nair, van der Dussen, du Plessis, Tripathi, and Markram to be categorised as all-rounders. Table 6.1 provides the bowling data related to these batsmen.

Table 6.1: Bowling data for identified batsmen

Player name	Games played	Innings bowled	Total wickets	Runs conceded	Balls bowled	CBR score
M. Singh	182	26	16	337	303	11.993540
M.K. Pandey	274	10	10	212	168	12.564440
S. Dhawan	303	6	4	66	48	11.314286
P.B. Rajapaksa	105	24	27	480	384	11.543086
K. Nair	147	10	4	81	78	11.487273
H.E. van der Dussen	140	6	3	55	54	10.959410
F. du Plessis	273	52	50	917	790	11.476624
R.A. Tripathi	104	17	12	281	222	13.133053
A.K. Markram	70	28	12	421	375	14.356841

The batsmen listed in Table 6.1 have CBR scores comparable to those of specialist bowlers. However, in cricket, there exist players referred to as part-time bowlers [283]. These players are specialist batsmen who occasionally bowl under certain circumstances [283]. Part-time bowlers are typically utilised when specialist bowlers are injured during a match or fail to dismiss the batsmen of the opposing team. To assess whether the players in Table 6.1 are all-rounders or merely batsmen who represent part-time bowlers, the percentage of games bowled, PG_{BOW}, is calculated for each cricketer using Equation (6.1).

$$PG_{BOW} = \frac{\text{Number of games bowled}}{\text{Number of games played}} \times 100 \tag{6.1}$$

Equation (6.2) provides the expression for the percentage of games batted, PG_{BAT}.

$$PG_{BAT} = \frac{\text{Number of games batted}}{\text{Number of games played}} \times 100 \tag{6.2}$$

Applying Equation (6.1) to the applicable data in Table 6.1 yielded the percentage of games bowled as 14.29%, 3.65%, 1.98%, 22.86%, 6.80%, 4.29%, 19.05%, 16.35%, and 40.00% for the players Singh, Pandey, Dhawan, Rajapaksa, Nair, van der Dussen, du Plessis, Tripathi, and Markram, respectively. Therefore, the batsmen in Table 6.1 bowled in a relatively small number of matches compared to the total number played, with Markram having the largest percentage. These percentages are noticeably lower than the average percentage of games bowled for all 58 all-rounders, which is 77.65%. Thus, comparing these percentages to the average percentage calculated for all-rounders suggests that the batsmen in Table 6.1 are part-time bowlers rather than all-rounders. These results support the classification of the players in Table 6.1 as batsmen during the 2022 IPL, while also demonstrating how CBR and CBM recognise players with expertise beyond their designated cricketing abilities.

Figure 6.1 illustrates a marginal overlap between bowlers and all-rounders, considering their CBM scores. However, the CBM scores of all bowlers lie below those of all batsmen, with the exception of Asif and Porel. Figure 6.1 shows that the bowlers Asif and Porel have the highest CBM scores. The batting data revealed that Asif batted in three innings, scoring a mere one run while facing three deliveries without being dismissed once. Similarly, Porel batted in only one innings, scoring two runs from five balls without being dismissed. Like the batsman Rahane, the bowlers Asif and Porel are outliers due to the limited batting data. These outliers reveal the need for an inclusion criterion to prevent misleading conclusions and potential misinterpretations regarding the capabilities of players. Figure 6.1 confirms the indexing accuracy of CBR and CBM by demonstrating the appropriate classification of bowlers and batsmen, respectively.

Figure 6.1 provides further insight into the classification of all-rounders. AR describes the weighted linear combination between CBR' and CBM', presented in Equations (2.19) and (2.18), respectively. Therefore, to evaluate the indexing accuracy for AR, each all-rounder is visualised by plotting their w_2CBR' against w_1CBM', where w_1 and w_2 are the optimal weights determined using the efficient frontier in Figure 2.3. Figure 6.2 presents a scatter plot of all 58 all-rounders, detailing the relationship between w_1CBM' and w_2CBR'.

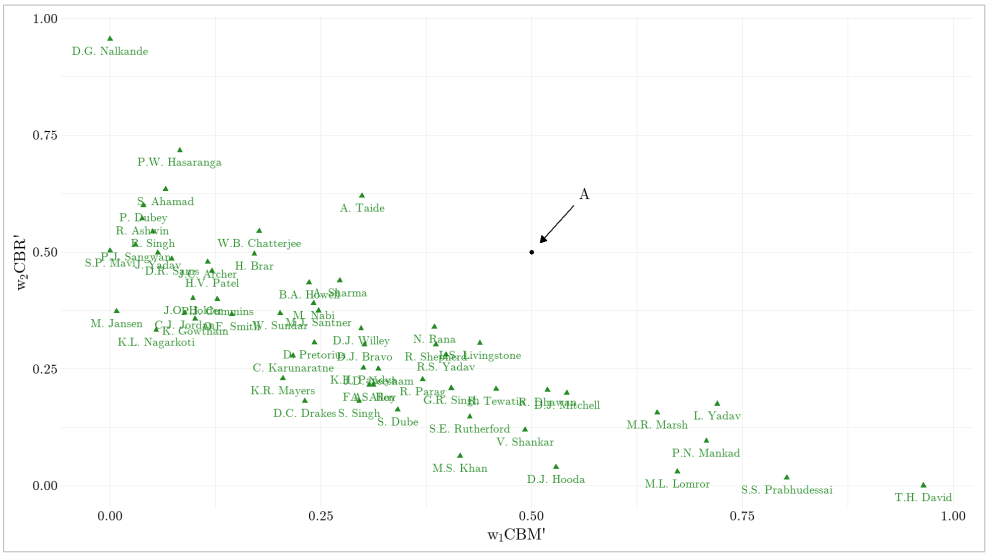


Figure 6.2: Scatterplot of w_1CBM' and w_2CBR'

From Figure 6.2, Point A denotes a reference coordinate for analysing the performance and comparison between the all-rounders. Point A defines the position of a perfectly balanced all-rounder with an identically weighted CBR' and CBM', i.e., $w_1 = w_2 = 0.50$. There is a negative relationship between CBR' and CBM'. As such, an increase in CBM' causes a decrease in CBR', and vice versa. The data in Appendix B revealed David and Nalkande as the best and second-best all-rounders, respectively. David has the second-highest CBM' value and the lowest CBR' value among all all-rounders. Nalkande has the second-highest CBR' value and the second-lowest CBM' value among all all-rounders. This further aligns

with the demonstration of CBR and CBM in Figure 6.1, considering the positioning of David and Nalkande relative to the other all-rounders.

The dominant batting ability of David and the bowling ability of Nalkande yielded absolute weights for CBM' and CBR', respectively. Specifically, the efficient frontier of David gave a weight of 0.00 for CBR' and 1.00 for CBM'. The efficient frontier relating to Nalkande produced a weight of 1.00 for CBR' and 0.00 for CBM'. Considering this absolute weight allocation, applying Equations (6.1) and (6.2) to the data outlined in Appendix B confirmed the potential misclassification of David and Nalkande as pure all-rounders. David batted in 92.94% of all 85 games played while bowling in only 32.94%. Nalkande bowled in 100.00% of all 22 games played, batting in a mere 54.55%. Considering the remaining 56 all-rounders, the average percentage of games bowled is 78.05%, while the average percentage of games batted is 73.95%. Comparing the percentages determined for David and Nalkande to the average percentages for the remaining 56 all-rounders indicates that the former is mostly a specialist batsman while the latter is mostly a specialist bowler. To further evaluate the potential misclassification of David and Nalkande, a confidence ellipse is applied.

Confidence ellipses represent a bivariate analogue of confidence intervals used for univariate data [284]. Grove and Pearson [284] discussed the usefulness of confidence ellipses in testing hypotheses involving the differences between multiple groups. Confidence ellipses facilitate an inspection of the relationships among variables [284]. The size and shape of confidence ellipses depend on the covariance matrix and the chosen significance level [285]. Confidence ellipses are based on the assumption of bivariate normality [286]. The bivariate normality of w_1CBM' and w_2CBR' is confirmed at a significance level of 0.01, with a p-value of 0.02 for the Royston multivariate normality test. After excluding the potential outliers, David and Nalkande, the test generated a p-value of 0.06, further supporting a bivariate normal distribution. Figure 6.3 illustrates a 99% confidence ellipse fitted to the 58 all-rounders.

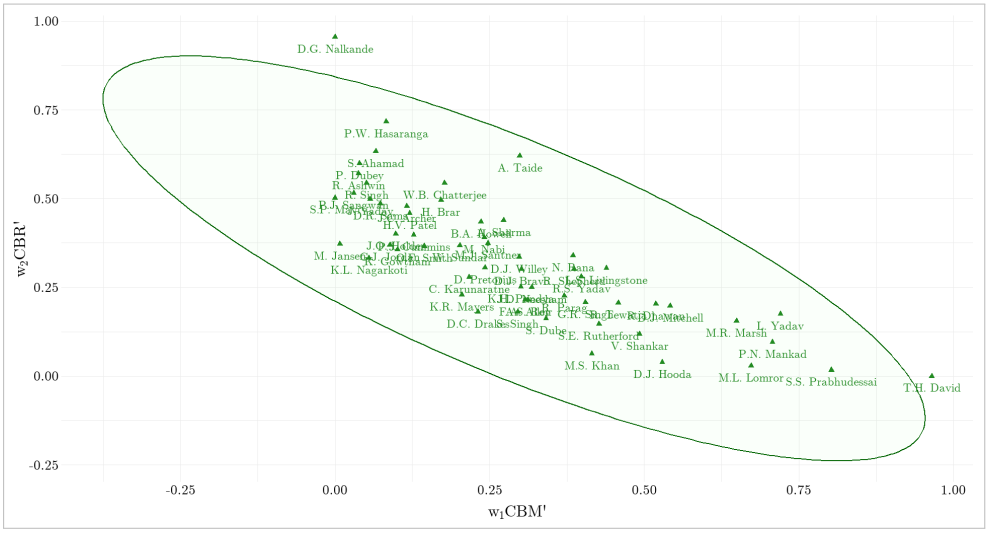


Figure 6.3: 99% confidence ellipse of all-rounders

Figure 6.3 illustrates the 99% confidence region as the green-shaded ellipse. All cricketers lie within the 99% confidence ellipse, except David and Nalkande. Accordingly, David and Nalkande are statistically different from the other players, confirming their misclassification as all-rounders. The current player profiles, available on ESPNcricinfo [277], for David and Nalkande also support the incorrect classification. ESPNcricinfo currently describes David as a specialist batsman [287] and Nalkande as a specialist bowler [288]. Hence, Figure 6.3 accurately illustrates the misclassification of these players during the 2022 IPL. Although David and Nalkande are recognised as misclassified players, they are treated as all-rounders in this study to ensure consistency with the cricketing abilities of the 2022 IPL.

Finally, Figure 6.2 indicates that Taide and Yadav stand out as significantly better players compared to the other all-rounders. The data in Appendix B also characterises Taide and Yadav as the third and fourth-best all-rounders, respectively, only trailing the misclassified players, David and Nalkande. The data in Appendix B identified Jansen and Nagarkoti as the worst and second-worst all-rounders, respectively. The most balanced all-rounders are positioned at the centre of the data cloud in Figure 6.2. These players have a strong balance between bowling and batting abilities, making them effective all-rounders. Figures 6.2 and 6.3 confirm the indexing accuracy for AR, with the suitable classification of all-rounders.

6.1.2 Comparative analysis of combined batting measure

Lemmer [120] acknowledged that numerous authors have developed performance measures that combine batting average and strike rate. The BK statistic, as presented by Barr and Kantor in Equation (2.17), serves as the baseline metric for the analysis due to its reliance on the power coefficient α. The power coefficient α details the balance between the batting average and strike rate, where decreasing values give more importance to the average, and increasing values assign more importance to the strike rate. Applying BK, and adjusting α accordingly, enables the analysis of the CBM metric in Equation (2.14) against different scenarios of varying importance imposed on the batting average and strike rate.

Barr and Kantor [85] proved the utility of BK using empirical performance data from the 2003 One Day International (ODI) World Cup. Barr and Kantor [85] applied the criterion to the top 20 run scorers of the tournament, computing multiple rankings based on distinct values for α. CBM was applied to the same performance data to determine a comparable ranking of the investigated batsmen. Table 6.2 lists the top 20 run scorers of the 2003 ODI World Cup, with the dismissal probability $P(out)$, performance scores, and rankings for each batsman. While Barr and Kantor [85] only reported the rankings where $\alpha = 0.50$ and $\alpha = 0.75$, Table 6.2 also details the ranking where $\alpha = 0.25$, providing further insight into the impact of increasing the importance of the batting average on BK. The data used to compute the results in Table 6.2 were retrieved from the original source [289] to ensure the correctness of the values. The complete dataset is provided in Appendix C.

Table 6.2: Top 20 run scorers of the 2003 ODI World Cup [289]

Player name	**Country**	**Runs scored**	***P(out)* (%)**	***BK* ($\alpha = 0.25$)**	***BK* rank ($\alpha = 0.25$)**	***BK* ($\alpha = 0.50$)**	***BK* rank ($\alpha = 0.50$)**	***BK* ($\alpha = 0.75$)**	***BK* rank ($\alpha = 0.75$)**	***CBM* score**	***CBM* rank**
S.R. Tendulkar	India	673	1.46	67.240001	4	73.898062	4	81.215399	5	72.599784	3
S.C. Ganguly	India	465	1.42	63.404977	6	69.164579	6	75.447373	9	68.131868	6
R.T. Ponting	Australia	415	1.69	59.189559	9	67.535498	8	77.058243	7	65.251572	8
A.C. Gilchrist	Australia	408	2.58	51.728802	13	65.585024	9	83.152812	4	58.832012	12
H.H. Gibbs	South Africa	384	1.05	97.175098	2	98.364580	2	99.568622	2	98.335467	2
M.S. Atapattu	Sri Lanka	382	1.55	60.877278	8	67.911782	7	75.759138	8	66.319444	7
A. Flower	Zimbabwe	332	1.53	52.706161	12	58.571011	13	65.088469	15	57.290768	13
M.L. Hayden	Australia	328	2.44	40.989996	18	51.224994	18	64.015619	18	46.524823	18
A. Symonds	Australia	326	0.56	140.724423	1	121.493027	1	104.889792	1	116.428571	1
D.R. Martyn	Australia	323	1.27	68.521324	3	72.680678	5	77.092512	6	72.178771	4
S.P. Fleming	New Zealand	321	1.87	53.636892	11	62.736489	12	73.379850	11	59.776536	11
S.T. Jayasuriya	Sri Lanka	321	1.90	47.138441	15	55.377761	15	65.057230	16	52.622951	15
R. Dravid	India	318	1.01	63.727840	5	63.855937	11	63.984291	19	63.855422	9
V. Sehwag	India	299	3.19	36.322180	19	48.536147	19	64.857275	17	41.384083	19
C.B. Wishart	Zimbabwe	293	1.75	56.160460	10	64.586974	10	74.277832	10	62.142100	10
S.B. Styris	New Zealand	268	1.90	62.938785	7	73.904676	3	86.781166	3	70.249017	5
P.A. de Silva	Sri Lanka	267	2.68	42.685068	17	54.592210	16	69.820889	12	48.589627	17
R.D. Shah	Kenya	265	2.05	35.252950	20	42.207299	20	50.533533	20	39.611360	20
B.C. Lara	West Indies	248	1.96	48.911151	14	57.878243	14	68.489311	13	54.746137	14
Y. Singh	India	240	2.49	43.073596	16	54.113928	17	67.984044	14	48.929664	16

From Table 6.2, all four criteria identified Symonds as the best batsman and Gibbs as the next best batsman. Shah was consistently selected as the worst batsman among the top 20 run scorers across all four criteria. Notwithstanding the consistency of the scenarios, the results for BK evidently illustrate the influence of varying values for α on player rankings, with significant discrepancies reported for Gilchrist, Dravid, and de Silva. Gilchrist and de Silva significantly improved in ranking, with increasing importance given to the strike rate, while Dravid considerably reduced in ranking. The variation in ranking resulting from the different values for α demonstrates the importance of balancing batting average and strike rate, providing insight into the limitations of certain players.

Since the harmonic mean computes the average of ratios [80], the ranking related to CBM is evaluated against the ranking for BK, in which equal importance is assigned to batting average and strike rate by setting $\alpha = 0.50$. The consistency between the CBM and BK rankings is investigated using the Kendall rank correlation coefficient. Abdi [290] described the Kendall rank correlation coefficient as a measure that describes the degree of similarity between different rankings gained for an equivalent collection of objects. The Kendall rank correlation coefficient takes a value ranging between [-1, 1] [291]. Values that approach one on the closed interval suggest a strong positive agreement between the two rankings, while values that approach minus one on the closed interval suggest a strong negative agreement [292]. Sazzed [292] mentioned that a Kendall rank correlation coefficient of zero implies no association and complete independence between the different rankings. The Kendall rank correlation coefficient computed for the CBM and BK rankings, equals 0.93, indicating a strong positive correlation between the rankings. The two-sided correlation test produced a p-value < 0.001, indicating that the CBM and BK rankings are statistically significant. Figure 6.4 shows a scatterplot of the CBM ranking against the BK ranking.

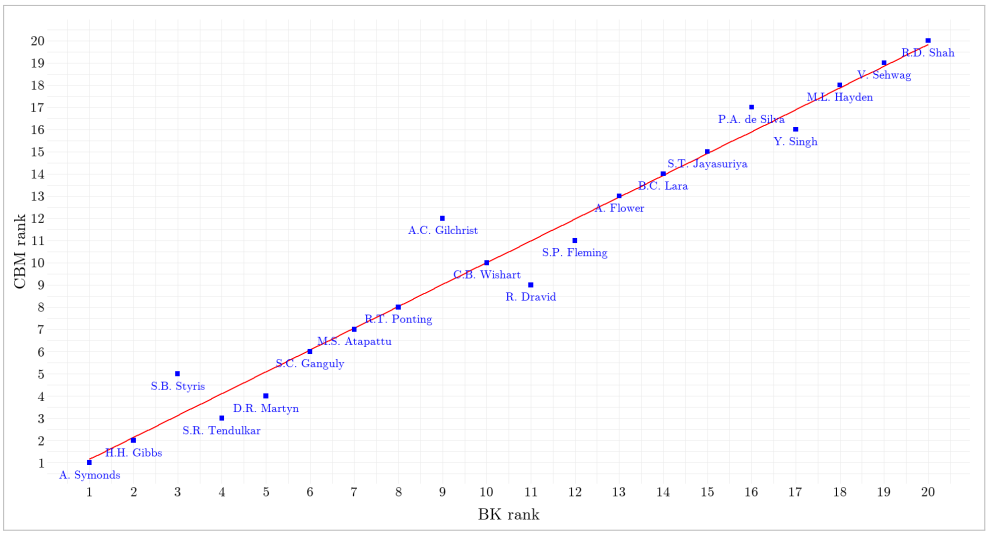


Figure 6.4: Relationship between CBM ranking and BK ranking

From Figure 6.4, the CBM ranking proves comparable to the BK ranking, with 12 players occupying the same position and five players differing by only one rank. The three batsmen,

Gilchrist, Dravid, and Styris, illustrate the greatest difference in rankings between the two approaches. The trend line indicates a strong positive correlation between the two rankings.

The results in Table 6.2 further indicate that BK is heavily dependent on the value selected for α, resulting in variations in player rankings due to the differing importance assigned to batting average and strike rate. The results provide insight into the utility of CBM and its significance in producing player rankings comparable to those of established metrics in the literature. Unlike the BK metric, CBM offers an objective approach to measuring general batting performance, without requiring the subjective selection of some power coefficient α. Consequently, CBM provides an advantage over BK by objectively capturing batting performance using the harmonic mean. The CBM metric is an accurate measure of batting performance that consolidates the batting average and strike rate, describing the ability of players to excel equally in both facets.

6.1.3 Comparative analysis of all-rounding measure

Sharp *et al.* [106] proposed an all-rounding performance measure that is similar to the BK statistic applied in Section 6.1.2. The SB metric, as characterised by Sharp *et al.* [106] in Equation (2.31), represents the baseline measure for analysis due to its comparability with BK using the power coefficient β. The power coefficient β stipulates the balance between bowling and batting. Decreasing values for β assign more importance to bowling, whereas increasing values give more importance to batting. Similar to BK, using SB and adjusting β accordingly, enables comparison of the AR measure in Equation (2.20) against different scenarios of varying importance imposed on bowling and batting.

Sharp *et al.* [106] illustrated the utility of SB using empirical performance data from the 2007 Twenty20 (T20) World Cup. Unlike the current study, Sharp *et al.* [106] did not first separate players according to their designated cricketing abilities. Specifically, the ranking obtained by Sharp *et al.* [106] for the top 10 all-rounders consists of five specialist bowlers. Although this ranking offers insight into the importance of bowlers being able to bat [106], it proves unrealistic, as these players are not considered specialist all-rounders. Failing to first separate players according to their related cricketing abilities resulted in an overlap of players across multiple rankings. For example, the individuals Afridi, Vettori, and Pathan were included in the top 10 rankings for both bowlers and all-rounders. Accordingly, this study utilises a new dataset consisting of 20 randomly selected specialist all-rounders from the 2016 T20 World Cup to evaluate the correlation between the rankings of AR and SB. Table 6.3 lists the 20 randomly selected all-rounders, along with their performance scores and rankings. While Sharp *et al.* [106] only presented the ranking where $\beta = 0.50$, Table 6.3 also describes the rankings for $\beta = 0.25$ and $\beta = 0.75$. The rankings for $\beta = 0.25$ and $\beta = 0.75$ provide further insight into the impact of increasing the importance of bowling and batting on SB, respectively.

Table 6.3: 20 randomly selected all-rounders of the 2016 T20 World Cup [293]

Player name	**Country**	**Runs scored**	**Total wickets**	***SB*** **($\beta = 0.25$)**	***SB*** **rank ($\beta = 0.25$)**	***SB*** **($\beta = 0.50$)**	***SB*** **rank ($\beta = 0.50$)**	***SB*** **($\beta = 0.75$)**	***SB*** **rank ($\beta = 0.75$)**	***AR*** **score**	***AR*** **rank**
M. Nabi	Afghanistan	105	12	1.557156	3	1.310048	7	1.039609	10	0.739830	3
S.R. Watson	Australia	96	5	1.505773	5	1.717802	2	1.767409	2	0.727067	4
S. Al Hasan	Bangladesh	129	10	1.572167	2	1.558291	3	1.414374	3	0.625966	5
D. Willey	England	41	10	1.512672	4	1.379103	5	1.176703	6	0.614587	6
C.J. Anderson	New Zealand	86	2	0.663634	15	0.666959	16	0.726342	15	0.305157	15
R. Ashwin	India	15	4	0.877261	12	0.733881	13	0.649353	16	0.367935	13
D.J. Sammy	West Indies	8	1	0.623131	18	0.373989	20	0.268705	20	0.252402	18
C.R. Brathwaite	West Indies	57	4	1.034196	8	1.445658	4	1.896902	1	0.949719	1
M. Mahmudullah	Bangladesh	86	1	0.514277	20	0.605067	18	0.789992	14	0.372264	12
T. Afzal	Hong Kong	37	2	1.216654	6	1.329645	6	1.376987	4	0.589958	7
T.A. Siriwardana	Sri Lanka	22	2	0.783363	14	0.638636	17	0.565621	18	0.272410	17
G.J. Maxwell	Australia	109	2	1.008885	10	1.081801	8	1.144376	7	0.440505	10
M.S. Rahman	Bangladesh	147	2	2.068644	1	1.787938	1	1.371691	5	0.913927	2
S. Raza	Zimbabwe	27	1	0.564522	19	0.500107	19	0.506384	19	0.241665	19
H.H. Pandya	India	16	5	0.866087	13	0.708498	15	0.616904	17	0.281581	16
S. Shinwari	Afghanistan	122	3	0.660061	16	0.765794	12	0.940660	11	0.405689	11
D.J. Bravo	West Indies	63	9	1.167180	7	0.992533	11	0.841333	12	0.488520	8
N.L. Perera	Sri Lanka	80	3	1.031214	9	1.051754	10	1.062650	9	0.334517	14
B.A. Stokes	England	56	4	0.652762	17	0.712917	14	0.834657	13	0.233111	20
S.C. Williams	Zimbabwe	78	2	1.007161	11	1.077650	9	1.138351	8	0.455241	9

Considering the dataset in Table 6.3, all players classified as all-rounders during the 2016 T20 World Cup were identified using ESPNcricinfo [293]. The bowling data related to each all-rounder in Table 6.3 were sourced from [294], with the associated batting data retrieved from [295]. The data used to obtain the results in Table 6.3 are provided in Appendix C.

From Table 6.3, when $\beta = 0.25$ and $\beta = 0.50$, SB selected Rahman as the best all-rounder. In contrast, both AR and SB, with $\beta = 0.75$, recognised Brathwaite as the best all-rounder. For $\beta = 0.25$, SB specified Mahmudullah as the worst all-rounder, while for both $\beta = 0.50$ and $\beta = 0.75$, Sammy was identified. AR recognised Stokes as the worst all-rounder, with Raza revealed as the next worst across all four criteria. Akin to Section 6.1.2, the ranking according to AR is assessed against the ranking according to SB, where equal importance is allocated to batting and bowling, with $\beta = 0.50$. Again, the similarity between the AR and SB rankings is evaluated using the Kendall rank correlation coefficient. The Kendall rank correlation coefficient for the AR and SB rankings is 0.73, indicating a strong positive correlation between the rankings. The two-sided correlation test yielded a p-value < 0.001, confirming that the AR and SB rankings are statistically significant. Figure 6.5 illustrates a scatterplot of the AR ranking against the SB ranking.



Figure 6.5: Relationship between AR ranking and SB ranking

From Figure 6.5, the AR ranking proves comparable to the SB ranking, with 14 cricketers shifting by no more than two places. The positions of Nabi, Brathwaite, Bravo, and Perera differ significantly between the two rankings. The all-rounders, Mahmudullah and Stokes, exhibit the largest variation in ranking between AR and SB, with both players varying by six positions. The trend line reveals a strong positive correlation between the two rankings.

The results in Table 6.3 reveal the dependency of SB on the value selected for β, resulting in considerable variations in player rankings due to the distinct importance of bowling and batting. The results demonstrate the utility of AR and its ability to yield player rankings

comparable to those of existing measures in the literature. AR provides several advantages over SB. First, AR does not require the subjective selection of a power coefficient β that stipulates the importance of bowling and batting. Instead, AR objectively determines the importance of bowling and batting using an efficient frontier to find the weights w_1 and w_2 for the linear combination of CBM' and CBR'. Instead of assigning a fixed importance to bowling and batting across all players, AR tailors the importance based on the CBR' and CBM' values of each cricketer. Second, AR is simpler than SB, which requires adjusting the index scales to ensure equivalence in the spread of indices across cricketing abilities.

Equations (2.23) and (2.27) enable compatibility between the batting and bowling indices with respect to their means [15]. Lourens [15] acknowledged that the comparability between these indices remains constrained by their differing standard deviations. Thus, the standard deviation of the bowling index is scaled with respect to that of the batting index. This is achieved using Equation (2.28), where the bowling index is determined as a function of the preceding index, adjusted according to the ratio of the standard deviations of the batting and bowling indices [106]. The same process is applied using Equation (2.31) to scale the standard deviation of the all-rounding index in relation to that of the batting index. The scale adjustment used by Sharp *et al.* [106] allowed the standard deviations of the bowling and all-rounding indices to scale down or up relative to that of the batting index [15].

The scale adjustment enables Sharp *et al.* [106] to compare players with varying cricketing abilities. To compare players across cricketing abilities, equivalence in the spread of indices must be maintained. Thus, adjusting the standard deviations is necessary when comparing the indices. However, the scale adjustment utilised by Sharp *et al.* [106] is not necessary for the AR statistic, as the current study only compares players who share the same cricketing ability. Comparability between the bowling and batting abilities of an all-rounder is already ensured through the normalisation of CBR and CBM. Accordingly, the AR metric is an accurate measure of all-rounding performance that combines bowling and batting abilities objectively using an efficient frontier.

6.2 Optimisation results

Section 6.2 presents the results obtained from the mixture design optimisation and scenario analysis, which illustrate the influence of prioritising different cricketing abilities.

6.2.1 Mixture design optimisation

Non-preemptive goal programming models require weights that specify the relative importance of the deviations associated with each goal. Ganesh and Reddy [296] mentioned that the selection of these weights is subjective and depends on the desires of the decision-maker.

A cricket team consists of a mixture of players with various cricketing abilities, which are selected based on the preferences of the selectors. This introduces the opportunity to select the weights and the number of players per cricketing ability objectively through a mixture design. Mixture designs denote a specialised approach in the field of experimental design. Mixture designs optimise the composition of different mixtures to obtain desired responses [297]. Responses in mixture designs denote functions of different proportions [298]. These proportions are the components of the mixture [298]. The components of mixture designs are not independent, meaning that their sum adds up to a fixed total, typically one [299].

Considering the composition of a cricket team, the components of the mixture design used in this study denote the weights of the underachieving deviations and the minimum number of players required per cricketing ability. Specifically, from Chapter 4, let $J = \{1, 2, ..., |J|\}$ represent the set of cricketing abilities, where $j \in J$ denotes each cricketing ability. Then, the weight as a proportion characterising the importance of the underachieving deviation d_j^- from the target value ψ_j related to the j-th goal is given as w_j^-, such that $w_j^- \in [0.2, 0.4]$ and $\sum_{j=1}^{4} w_j^- = 1$. The proportion of players related to the j-th cricketing ability, expressed as a ratio to the minimum number of 18 individuals required in the squad, is described as a_j, such that $a_j \in [0.2, 0.4]$ and $\sum_{j=1}^{4} a_j = 1$. Each component is constrained to the closed interval [0.2, 0.4] to ensure that no individual cricketing ability dominates the mixture. An I-optimal crossed mixture design was constructed using the Design-Expert software [300], which minimises the average prediction variance across the design space for a second-order response surface model [301]. The optimal design specified 70 experimental runs, each of which characterises a unique combination of the components.

The validation in Section 6.3.1 examines the team selected by Model 1 against the winning team of the 2022 IPL. Gujarat Titans (GT) prevailed as the winning franchise after beating Rajasthan Royals (RR) by seven wickets in the final of the tournament. GT retained three players whose auctioning prices, cricketing abilities, and nationality classes are presented in Table 6.4.

Table 6.4: Players retained by GT [302], [303]

Characteristic	S. Gill	H.H. Pandya	R. Khan
Auctioning price	₹8.00 crore	₹15.00 crore	₹15.00 crore
Cricketing ability	Batting	All-rounding	All-rounding
Nationality class	Indian	Indian	Overseas

To standardise the validation in Section 6.3.1, three players with identical characteristics to those in Table 6.4 were retained for the mixture design runs. The objective function value was computed for each run using the related proportions of weights and minimum number of players required per cricketing ability. Similar to Section 5.3, all cricketing abilities were assigned the same target value for each run, i.e., $\psi_j = 100$ for all $j \in J$. Appendix C lists the components and objective function value for each of the 70 runs conducted on the data

in Appendix B related to the 173 players considered for selection. Among the 70 runs, Run 2 yielded the best response by giving the smallest objective function value. The results are provided in Table 6.5.

Table 6.5: Results for Run 2 of the mixture design

Mixture results							
w_1	w_2	w_3	w_4	a_1	a_2	a_3	a_4
0.40	0.20	0.20	0.20	0.40	0.20	0.20	0.20
Selection results							
Objective value		95.564827		Purchasing total		₹43.00 crore	
Regulatory value		1		Penalty value		₹0.00 crore	
Deviational variables							
Underachievement			**Overachievement**				
Goal	**Variable**	**Value**	**Goal**	**Variable**	**Value**		
1	d_1^-	93.610520	1	d_1^+	0.00		
2	d_2^-	96.553529	2	d_2^+	0.00		
3	d_3^-	96.182057	3	d_3^+	0.00		
4	d_4^-	97.867506	4	d_4^+	0.00		

From Table 6.5, Run 2 generated the best proportions of 0.40, 0.20, 0.20, and 0.20 for the weights w_1, w_2, w_3, and w_4, and the minimum number of players required per cricketing ability a_1, a_2, a_3, and a_4, respectively. Thus, the proportion of players per cricketing ability translates to 7.20 bowlers, 3.60 batsmen, 3.60 all-rounders, and 3.60 wicketkeepers for the minimum number of 18 individuals required in the squad. Since the three players retained denote a batsman and two all-rounders, the model seeks a minimum of 7.20 bowlers, 2.60 batsmen, 1.60 all-rounders, and 3.60 wicketkeepers to complete the required proportion of players per cricketing ability, according to Run 2. Since the decision variable x_{ijk} is binary, the model rounds the decimal numbers to their corresponding ceiling values to satisfy the constraint in Equation (4.3). Accordingly, the model requires a minimum of eight bowlers, three batsmen, two all-rounders, and four wicketkeepers.

Table 6.5 shows that the model produced an objective function value of 95.56. The model yielded a purchasing total of ₹43.00 crore, with a regulatory variable value of one, resulting in a penalty value of ₹0.00 crore. The model selected 22 players, comprising 10 bowlers, four batsmen, four all-rounders, and four wicketkeepers. Thus, the model selected the minimum number of players for wicketkeeping, while choosing two, one, and two additional cricketers for bowling, batting, and all-rounding, respectively. Therefore, the membership value totals amount to 6.39, 3.45, 3.82, and 2.13 for bowling, batting, all-rounding, and wicketkeeping, respectively. These membership value totals gave the underachieving deviations of 93.61, 96.55, 96.18, and 97.87 for the four goals, respectively. Table 6.6 lists the players selected by the model for Run 2.

Table 6.6: Players selected for Run 2 of the mixture design

Bowling			Batting		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
12	M.M. Theekshana	Overseas	11	D.P. Conway	Overseas
18	S. Singh	Indian	24	D.A. Warner	Overseas
26	C. Sakariya	Indian	148	D. Padikkal	Indian
28	M. Rahman	Overseas	156	H.E. van der Dussen	Overseas
34	L.T. Ngidi	Overseas			
47	R. Sai Kishore	Indian			
97	M. Khan	Indian			
138	I.C. Porel	Indian			
181	C.V. Milind	Indian			
182	A. Deep	Indian			
All-rounding			**Wicketkeeping**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
33	L. Yadav	Indian	30	K.S. Bharat	Indian
59	D.G. Nalkande	Indian	72	S.P. Jackson	Indian
102	T.H. David	Overseas	139	J.M. Sharma	Indian
141	A. Taide	Indian	171	A. Rawat	Indian

Table 6.6 lists the optimal team for the 2022 IPL, selected using the optimal combination of weights and minimum number of players required per cricketing ability under retention conditions identical to those of GT. Among the 22 players, the model selected 15 Indian and seven overseas cricketers. The model selected the top players in each cricketing ability for ₹9.00 crore cheaper than the auctioning purse available following player retention, further signifying its utility in maximising team strength at minimum expense.

6.2.2 Scenario analysis

Preemptive goal programming models utilise priority levels to represent the immeasurable importance of each goal [29]. Ledwith *et al.* [304] stated that preemptive goal programming has the advantage of eliminating the need to define the degree to which one goal is favoured over another. Preemptive goal programming only requires an understanding of the specific hierarchical structure that describes the priority of each goal [304]. This advantage makes preemptive goal programming suitable for addressing practical questions faced by selectors, such as whether a bowling-dominant or a batting-dominant team is better. Ahmed *et al.* [6] noted the dilemma faced by selectors when deciding between a bowling-dominant and a batting-dominant team. Accordingly, Model 2 is used to assess whether franchises would have benefited from selecting more bowlers or batsmen during the 2022 IPL mega auction.

Similar to Section 5.3.2, the following two scenarios are used to evaluate the difference in team strength between a bowling-dominant and a batting-dominant team:

- Scenario 1: Where bowling is prioritised over all other cricketing abilities.
- Scenario 2: Where batting is prioritised over all other cricketing abilities.

The scenario analysis provides insight into the influence of prioritising bowling and batting on team selection during the 2022 IPL mega auction. To fairly gauge the favourability of bowling and batting, the same selection requirements are applied to both scenarios. The following assumptions and considerations apply to Model 2 for the scenario analysis:

1. All cricketing abilities are of equal importance, that is $w_j^- = 0.25$ for all $j \in J$.
2. All cricketing abilities have the same target value, that is $\psi_j = 100$ for all $j \in J$.
3. No players are retained for any of the cricketing abilities, that is $\tau_j = 0$ for all $j \in J$.
4. The minimum number of players required for each cricketing ability in the squad of 18 to 25 players is defined as follows:
   - Minimum number of bowlers: Five.
   - Minimum number of batsmen: Five.
   - Minimum number of all-rounders: Five.
   - Minimum number of wicketkeepers: Three.
5. Each cricketing ability contains a preemptive priority factor P_z of non-numeric nature that describes the z-th priority level. Table 6.7 details the preemptive priority factor assigned to each goal for both scenarios.

Table 6.7: Preemptive priority factors for scenario analysis

Scenario	Bowling	Batting	All-rounding	Wicketkeeping
1	P_1	P_2	P_2	P_2
2	P_2	P_1	P_2	P_2

From Table 6.7, bowling and batting contain the highest priority (P_1) in Scenarios 1 and 2, respectively. The remaining cricketing abilities for both scenarios share the same lowest priority (P_2).

Employing the assumptions and considerations, Model 2 is applied to the data in Appendix B related to the 173 players considered for selection.

Scenario 1: Bowling-dominant team

Scenario 1 prioritises bowling over batting, all-rounding, and wicketkeeping. The model is expected to fill all available positions in the team with bowlers, while picking the minimum number of batsmen, all-rounders, and wicketkeepers. Table 6.8 presents the results of the model for Scenario 1.

Table 6.8: Results for Scenario 1 of scenario analysis

Selection results					
Objective value	95.647375		Purchasing total		₹48.25 crore
Regulatory value	0		Penalty value		₹19.25 crore
Deviational variables					
Underachievement			**Overachievement**		
Goal	**Variable**	**Value**	**Goal**	**Variable**	**Value**
1	d_{11}^-	92.995034	1	d_{11}^+	0.00
2	d_{22}^-	95.938028	2	d_{22}^+	0.00
3	d_{32}^-	95.617267	3	d_{32}^+	0.00
4	d_{42}^-	98.039172	4	d_{42}^+	0.00

From Table 6.8, the model yielded an objective function value of 95.65. The selected team failed to satisfy the minimum spend requirement, with a purchasing total of ₹48.25 crore, resulting in a regulatory variable value of zero. The penalty value amounts to ₹19.25 crore, denoting the difference between the purchasing total and the minimum spend requirement. The model selected 25 cricketers, including 12 bowlers, five batsmen, five all-rounders, and three wicketkeepers. The membership value totals are 7.00, 4.06, 4.38, and 1.96 for bowling, batting, all-rounding, and wicketkeeping, respectively. Therefore, these membership value totals produced the underachieving deviations of 93.00, 95.94, 95.62, and 98.04 for the four goals, respectively. Table 6.9 lists the players selected by the model.

Table 6.9: Players selected for Scenario 1 of scenario analysis

Bowling			**Batting**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
12	M.M. Theekshana	Overseas	11	D.P. Conway	Overseas
18	S. Singh	Indian	24	D.A. Warner	Overseas
26	C. Sakariya	Indian	46	D.A. Miller	Overseas
28	M. Rahman	Overseas	148	D. Padikkal	Indian
34	L.T. Ngidi	Overseas	156	H.E. van der Dussen	Overseas
45	Y. Dayal	Indian			
47	R. Sai Kishore	Indian			

97	M. Khan	Indian			
138	I.C. Porel	Indian			
181	C.V. Milind	Indian			
182	A. Deep	Indian			
197	S. Gopal	Indian			
All-rounding			**Wicketkeeping**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
33	L. Yadav	Indian	72	S.P. Jackson	Indian
59	D.G. Nalkande	Indian	139	J.M. Sharma	Indian
102	T.H. David	Overseas	171	A. Rawat	Indian
141	A. Taide	Indian			
180	S.S. Prabhudessai	Indian			

From Table 6.9, employing the preemptive priority factors for Scenario 1 enabled the model to assign primary priority to bowling and secondary priority to batting, all-rounding, and wicketkeeping. Selecting the minimum number of players for each cricketing ability results in seven vacancies, which are limited by the maximum of 25 players permitted in the squad. The model populated all seven available positions with individuals of the highest prioritised cricketing ability, resulting in the selection of the 12 best bowlers. The model selected the minimum number of players for all remaining cricketing abilities, i.e., batting, all-rounding, and wicketkeeping. As a result of the substantial auctioning purse, ample funds remained available for the model to select the five best batsmen, the five best all-rounders, and the three best wicketkeepers. The model selected 17 Indian and eight overseas players.

Scenario 2: Batting-dominant team

Scenario 2 prioritises batting over bowling, all-rounding, and wicketkeeping. As in Scenario 1, the model is expected to fill all available places in the team with batsmen while selecting the minimum number of bowlers, all-rounders, and wicketkeepers. Table 6.10 presents the results of the model for Scenario 2.

Table 6.10: Results for Scenario 2 of scenario analysis

Selection results			
Objective value	95.682075	Purchasing total	₹50.95 crore
Regulatory value	0	Penalty value	₹16.55 crore
Deviational variables			
Underachievement		**Overachievement**	
Goal	**Variable**	**Value**	**Goal**	**Variable**	**Value**

1	d_{12}^-	96.073211	1	d_{12}^+	0.00
2	d_{21}^-	92.998648	2	d_{21}^+	0.00
3	d_{32}^-	95.617267	3	d_{32}^+	0.00
4	d_{42}^-	98.039172	4	d_{42}^+	0.00

From Table 6.10, the model yielded an objective function value of 95.68. The selected team failed to satisfy the minimum spend requirement, with a purchasing total of ₹50.95 crore, resulting in a regulatory variable value of zero. The penalty value amounts to ₹16.55 crore, representing the amount by which the purchasing total falls short of the minimum spend requirement. Again, the model selected 25 players, including five bowlers, 12 batsmen, five all-rounders, and three wicketkeepers. The membership value totals equal 3.93, 7.00, 4.38, and 1.96 for bowling, batting, all-rounding, and wicketkeeping, respectively. As such, these membership value totals yielded the underachieving deviations of 96.07, 93.00, 95.62, and 98.04 for the four goals, respectively. Table 6.11 lists the players selected by the model.

Table 6.11: Players selected for Scenario 2 of scenario analysis

Bowling			**Batting**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
47	R. Sai Kishore	Indian	11	D.P. Conway	Overseas
97	M. Khan	Indian	24	D.A. Warner	Overseas
138	I.C. Porel	Indian	38	K.A. Hebbar	Indian
181	C.V. Milind	Indian	46	D.A. Miller	Overseas
182	A. Deep	Indian	82	P. Singh	Indian
			91	E. Lewis	Overseas
			119	R.R. Buddhi	Indian
			125	S. Dhawan	Indian
			148	D. Padikkal	Indian
			156	H.E. van der Dussen	Overseas
			176	F.H. Allen	Overseas
			194	A.K. Markram	Overseas
All-rounding			**Wicketkeeping**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
33	L. Yadav	Indian	72	S.P. Jackson	Indian
59	D.G. Nalkande	Indian	139	J.M. Sharma	Indian
102	T.H. David	Overseas	171	A. Rawat	Indian
141	A. Taide	Indian			
180	S.S. Prabhudessai	Indian			

From Table 6.11, applying the preemptive priority factors for Scenario 2 enabled the model to assign primary priority to batting and secondary priority to bowling, all-rounding, and

wicketkeeping. The model once again filled all seven available positions with players of the highest prioritised cricketing ability, causing the selection of 12 batsmen. Similar to Table 6.9, the model selected the 12 best batsmen while choosing the minimum number of players for all remaining cricketing abilities, i.e., bowling, all-rounding, and wicketkeeping. Again, ample funds remained available for the model to select the best eligible players for all other cricketing abilities due to the sizable auctioning purse. Among the 12 best batsmen, seven are overseas players, leaving one position for a foreign bowler, all-rounder, or wicketkeeper. However, the overseas players Theekshana and David are among the top five bowlers and the top five all-rounders, respectively. The model selected David over Theekshana as the final overseas player due to a higher membership value. The model replaced Theekshana with the next-best Indian bowler, Khan. The model selected the five best eligible bowlers, along with the five best all-rounders and the three best wicketkeepers. The model included the same all-rounders and wicketkeepers for both scenarios, allowing for a fair comparison between the selected bowlers and batsmen.

Comparing the results in Tables 6.8 and 6.10 indicates that selecting a bowling-dominant or a batting-dominant team during the 2022 IPL would not have significantly impacted team strength. The results reveal that the bowling-dominant team delivered a slightly superior selection, resulting in a lower objective function value and purchasing total relative to the batting-dominant team. These results support the mixture design in Section 6.2.1, where Run 2 yielded the optimal proportions of weights and minimum number of players required per cricketing ability. Run 2 characterised bowling as the most important cricketing ability, requiring the highest weight and number of players.

6.3 Validation

Section 6.3 validates the effectiveness of the models by critically investigating and comparing the teams selected by each.

6.3.1 Model 1: Non-preemptive goal programming model

The optimal team selected by Model 1 in Section 6.2.1 is validated against the tournament-winning team selected by GT. The validation of Model 1 involves assessing its effectiveness in selecting a superior team compared to GT. GT fielded a squad of 23 players during the 2022 IPL, comprising three retentions and 20 auction selections. Among the 20 cricketers selected during the auction, there are seven bowlers, three batsmen, eight all-rounders, and two wicketkeepers [282]. The same individuals who failed to meet the inclusion criteria in Section 5.2 were excluded from consideration to ensure a fair comparison between the team selected by Model 1 and the one chosen by GT. GT selected two players who were initially excluded from the 173 considered for selection: the batsman Manohar and the all-rounder

Sai Sudharsan. Table 6.12 details the remaining 18 players selected by GT [282], excluding Manohar and Sai Sudharsan.

Table 6.12: Players selected by GT [282]

Bowling			**Batting**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
42	L.H. Ferguson	Overseas	46	D.A. Miller	Overseas
44	M. Shami	Indian	51	J.J. Roy	Overseas
45	Y. Dayal	Indian			
47	R. Sai Kishore	Indian			
50	A.S. Joseph	Overseas			
57	V.R. Aaron	Indian			
58	N. Ahmad	Overseas			
All-rounding			**Wicketkeeping**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
43	R. Tewatia	Indian	49	M.S. Wade	Overseas
53	J. Yadav	Indian	52	W.P. Saha	Indian
54	V. Shankar	Indian			
55	D.C. Drakes	Overseas			
56	G.R. Singh	Indian			
59	D.G. Nalkande	Indian			
61	P.J. Sangwan	Indian			

From Table 6.12, GT selected a team having 11 Indian and seven overseas players, which is dominated by bowlers and all-rounders. GT selected the best bowler, Sai Kishore, and the second-best all-rounder, Nalkande. GT included the fifth-best batsman, Miller, alongside a modest counterpart in Roy. GT also selected the moderate bowlers, Ferguson and Dayal. The remaining 12 players all have membership values smaller than 0.10, suggesting possible bias in the selections, where decisions are driven by subjective factors rather than objective performance data. Table 6.13 defines the membership value total related to each cricketing ability for the team selected by Model 1 in Table 6.6, and the team selected by GT in Table 6.12. Moreover, Table 6.13 specifies the sum of the membership value totals, which defines the collective performance of each team. Table 6.13 reveals whether Model 1 or GT yielded a superior selection for each cricketing ability and overall.

Table 6.13: Comparison of membership value totals

Number	**Cricketing ability**	**Model 1**	**GT**	**Superior**
1	Bowling	6.389480	1.618313	Model 1
2	Batting	3.446471	0.805958	Model 1

3	All-rounding	3.817943	1.172582	Model 1
4	Wicketkeeping	2.132494	0.017596	Model 1
Total		15.786388	3.614449	Model 1

From Table 6.13, Model 1 selected a notably superior team compared to GT. The sum of the membership value totals is 15.79 for the team of Model 1 and 3.61 for the team of GT. A higher membership value represents a better-performing player in a particular cricketing ability. Therefore, higher membership value totals are desirable, as they indicate that each department consists of stronger players, thereby contributing to the overall team strength. Model 1 produced a team that is approximately four times better than the one selected by GT under the same auctioning requirements and player retentions. Not only did Model 1 give a better team overall, but it also dominated the selection of all four cricketing abilities, generating the highest membership value total for each. The cricketing abilities that have the largest and smallest difference in membership value totals between the two teams are bowling and wicketkeeping, respectively. Model 1 also identified a team that is ₹6.05 crore less expensive than the one selected by GT. Table 6.13 validates the effectiveness of Model 1 in selecting a superior team compared to GT.

While the validation proves the effectiveness of Model 1 in selecting a team that dominates the winning team of the 2022 IPL, the author acknowledges its limitations. The validation of Model 1 required removing players from the team selected by GT who were not part of the 173 that met the inclusion criteria in Section 5.2. Omitting these players after selection diminished the strength of the GT team. However, it is crucial to exclude these players for a fair comparison between the sides. Even if Manohar and Sai Sudharsan were replaced by the best available batsman and all-rounder, respectively, the sum of the membership value totals for the team selected by GT remains significantly lower than that of the team selected by Model 1. Therefore, despite its limitations, the validation of Model 1 demonstrates its effectiveness in selecting a team superior to the winning franchise of the 2022 IPL.

For additional insight, the team selected by Model 1 is also compared to those selected by the remaining nine franchises that competed in the 2022 IPL. Players not part of the 173 considered for selection were once again excluded because they failed to meet the inclusion criteria in Section 5.2. Table 6.14 lists the nine franchises, along with the membership value totals that correspond to the cricketing abilities of each. Moreover, Table 6.14 provides a ranking for each franchise based on team strength, calculated by summing its membership value totals. The ranking indicates GT as the fourth-strongest team.

Table 6.14: Ranking of remaining nine franchises

Franchise	**Bowling**	**Batting**	**All-rounding**	**Wicketkeeping**	**Total**	**Rank**
RCB	2.287846	0.799522	1.433847	1.095432	5.616647	1

DC	1.699627	1.567418	1.430775	0.227901	4.925721	2
PBKS	1.652728	0.401387	2.295579	0.247466	4.597160	3
CSK	1.706532	1.425124	0.083490	0.003240	3.218385	5
RR	0.742074	1.734372	0.326411	0.000000	2.802856	6
KKR	0.043457	1.316930	0.273739	0.766690	2.400815	7
LSG	1.228565	0.903798	0.028150	0.164296	2.324809	8
MI	0.567213	0.362893	1.173106	0.034076	2.137288	9
SRH	0.691634	0.429115	0.315822	0.051329	1.487900	10

Table 6.14 indicates that the team selected by Model 1 is not only superior to the winning team of the 2022 IPL but also outperforms those selected by the remaining nine franchises. While GT won the 2022 IPL, Royal Challengers Bengaluru (RCB) had the strongest team, with the highest sum of membership value totals, which equals 5.62. However, RCB finished third in the competition. GT, as the fourth-strongest team, prevailed as the winners of the 2022 IPL, which demonstrates the uncertainties inherent in cricket, where individual player performance, match conditions, and momentum shifts often nullify statistical expectations. The case where a statistically weaker team outperforms a statistically stronger team is due to the unpredictability of cricket. Both GT and RCB participated in the 2022 IPL playoffs. The playoffs amplify this unpredictability by favouring the adaptability and high-pressure performance capabilities of teams over seasonal consistency. The ranking corresponding to the sum of the membership value totals correctly identified two of the four franchises that qualified for the 2022 IPL playoffs.

The other franchises that qualified for the 2022 IPL playoffs were RR and Lucknow Super Giants (LSG). However, Table 6.14 indicates that Delhi Capitals (DC) and Punjab Kings (PBKS) should have qualified, having the second- and third-strongest teams, respectively. Considering the strength of their teams, DC and PBKS notably underperformed, finishing fifth and sixth, respectively. In contrast, RR and LSG significantly overperformed, as they qualified for the playoffs with the sixth- and eighth-strongest sides, respectively. Table 6.14 indicates that among the 10 franchises, RCB selected a team with the strongest bowling and wicketkeeping departments. RR and PBKS identified teams with the best batting and all-rounding lineups, respectively.

6.3.2 Model 2: Preemptive goal programming model

Model 2 extends Model 1 by introducing preemptive priority factors. Therefore, the team selected by Model 2 is validated against the team selected by Model 1 in Section 6.2.1. The validation of Model 2 involves establishing its effectiveness in selecting a more specialised team compared to Model 1. Applying the preemptive priority factors of Model 2 to control the specialisation of the team provides numerous advantages over arbitrarily adjusting the

weights of Model 1. Using preemptive priority factors ensures a sensible trade-off between the different goals, rather than placing excessive emphasis on particular cricketing abilities through extreme weights. Exaggerating weights in an attempt to capture the specialisation of the team could lead to numerical instability and sensitivity in computations, resulting in unwanted selections. The hierarchy resulting from preemptive priority factors ensures that higher-priority goals are met before lower-priority goals, allowing selectors to structure the team composition around the desired specialisation.

Run 2 of the mixture design conducted identified bowling as the most important cricketing ability, followed by batting, all-rounding, and wicketkeeping, which are equally important. The validation of Model 2 replicates this hierarchy by applying preemptive priority factors, assigning primary priority to bowling and secondary priority to batting, all-rounding, and wicketkeeping. Therefore, the preemptive priority factor P_1 is assigned to bowling, with P_2 allocated to batting, all-rounding, and wicketkeeping. The target values, weights, minimum number of players required per cricketing ability, and retentions for Model 2 are identical to those of Model 1. Table 6.15 presents the results of the model.

Table 6.15: Results for validation of Model 2

Selection results					
Objective value	95.737477		Purchasing total		₹40.35 crore
Regulatory value	1		Penalty value		₹0.00 crore
Deviational variables					
Underachievement			**Overachievement**		
Goal	**Variable**	**Value**	**Goal**	**Variable**	**Value**
1	d_{11}^-	92.737163	1	d_{11}^+	0.00
2	d_{22}^-	97.343459	2	d_{22}^+	0.00
3	d_{32}^-	98.002093	3	d_{32}^+	0.00
4	d_{42}^-	97.867506	4	d_{42}^+	0.00

From Table 6.15, the model gave an objective function value of 95.74. The model yielded a purchasing total of ₹40.35 crore, resulting in a regulatory variable value of one and a penalty value of ₹0.00 crore. The model selected 22 players, viz., 13 bowlers, three batsmen, two all-rounders, and four wicketkeepers. The membership value totals are 7.26, 2.66, 2.00, and 2.13 for bowling, batting, all-rounding, and wicketkeeping, respectively. These membership value totals yielded the underachieving deviations of 92.74, 97.34, 98.00, and 97.87 for the four goals, respectively. Table 6.16 lists the players selected by the model.

Table 6.16: Players selected for validation of Model 2

Bowling			**Batting**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**

12	M.M. Theekshana	Overseas	11	D.P. Conway	Overseas
18	S. Singh	Indian	24	D.A. Warner	Overseas
26	C. Sakariya	Indian	156	H.E. van der Dussen	Overseas
28	M. Rahman	Overseas			
29	K. Yadav	Indian			
34	L.T. Ngidi	Overseas			
45	Y. Dayal	Indian			
47	R. Sai Kishore	Indian			
97	M. Khan	Indian			
138	I.C. Porel	Indian			
181	C.V. Milind	Indian			
182	A. Deep	Indian			
197	S. Gopal	Indian			
All-rounding			**Wicketkeeping**		
Number	**Player name**	**Nationality**	**Number**	**Player name**	**Nationality**
59	D.G. Nalkande	Indian	30	K.S. Bharat	Indian
102	T.H. David	Overseas	72	S.P. Jackson	Indian
			139	J.M. Sharma	Indian
			171	A. Rawat	Indian

From Table 6.16, utilising the preemptive priority factors enabled the model to assign the highest priority to bowling and the lowest priority to batting, all-rounding, and wicketkeeping. The model populated the five available positions with players of the highest prioritised cricketing ability, resulting in the selection of the 13 best bowlers. The model included the minimum number of players required for all remaining cricketing abilities, selecting the top three batsmen, the top two all-rounders, and the top four wicketkeepers. Model 2 selected the same wicketkeepers as Model 1 in Table 6.6, while replacing the batsman Padikkal and the all-rounders Yadav (L) and Taide with the three bowlers Yadav (K), Dayal, and Gopal. Among the resulting 22 cricketers, 15 are Indian, and seven are overseas players.

Comparing the results in Tables 6.5 and 6.15 indicates that Model 2 produced an inferior team, with an objective function value that is 0.17 worse than that of Model 1. However, Model 2 selected a team ₹2.65 crore more affordable than Model 1. Models 1 and 2 yielded equally strong wicketkeeping lineups, while the former accomplished stronger batting and all-rounding departments. Model 2 produced a considerably stronger bowling department, with a membership value total that is 0.87 better than that of Model 1. Although Model 1 selected a better team overall, Model 2 produced a team structured around maximising bowling performance. Tables 6.5 and 6.15 validate the effectiveness of Model 2 in selecting a more specialised team compared to Model 1.

6.4 Chapter summary

The indexing accuracy of the CBR, CBM, and AR metrics is confirmed according to the classification given by the IPL in 2022 for bowlers, batsmen, and all-rounders, respectively. The appropriateness of the proposed CBM and AR statistics is established in relation to the existing measures of BK and SB, respectively. The weights and the minimum number of players required per cricketing ability for Models 1 and 2 are computed objectively using an I-optimal crossed mixture design with 70 runs. From the output of the mixture design, Run 2 yielded the best response, i.e., the smallest objective function value. Applying Run 2, Model 1 is validated by confirming its effectiveness in selecting a superior team compared to GT. Similarly, Model 2 is validated by demonstrating its effectiveness in selecting a more specialised team compared to Model 1.

Chapter 7

Conclusion

Contents

7.1 Study summary 126
7.2 Limitations and recommendations 128
7.3 Practical considerations 130

7.1 Study summary

Section 7.1 presents a detailed overview of the study, followed by a description of how the different objectives were achieved.

7.1.1 Overview

Cricket is a global sport, classified as a complex bat-and-ball game [1], played between two opposing teams of 11 members each [48]. Cricket is among the oldest organised sports [44], with its origins commonly linked to early English development [1], [40]. The game evolved from an exclusive pastime into a multi-billion-dollar industry [4]. The global recognition of cricket is primarily due to competitive Twenty20 (T20) leagues, which require participating franchises to execute appropriate team selection for success. The nature of cricket enables players to specialise in different cricketing abilities, such as bowling, batting, all-rounding, and wicketkeeping [7]. The complexity of balancing players with distinct cricketing abilities significantly contributes to the subjectivity in traditional team selection methods.

Ample research has been conducted on utilising various optimisation techniques to enhance the objectivity of cricket team selection [24]. A systematic literature review (SLR) compiled

by Alberts *et al.* [24] identified several gaps in the existing literature. The techniques used in the existing literature introduce limited flexibility in addressing multiple objectives and constraints. In addition, numerous studies have examined team selection in the context of T20 leagues, with a focus on the Indian Premier League (IPL). Most of these studies select the playing 11 rather than the universal squad formed through the auction. Furthermore, these studies failed to distinguish between the IPL mini and mega auctions, each of which has different team selection requirements.

This study addressed these limitations by presenting a unique goal programming approach to optimise cricket team selection, using the 2022 IPL mega auction as a case study. This study followed a modified version of the scientific method, adapted from Salminawati and Ningsih [31], Blystone and Blogett [32], and Wright [33]. Using the research methodology, two solutions are devised, formulated as non-preemptive (Model 1) and preemptive (Model 2) goal programming models. These models were developed, programmed, and solved using Elytica, applying the CPLEX optimisation solver. The data supporting these models was collected from Statsguru [276], which is the statistical database of ESPNcricinfo [277]. An inclusion criteria, as proposed by Bhattacharjee and Saikia [17], was applied to ensure only suitable players were considered for selection. The performance scores of the resulting 173 players were determined using specific performance measures. This study presents two new performance measures for batting and all-rounding. These measures were used, along with existing metrics for bowling and wicketkeeping.

The performance scores were transformed using Gaussian membership functions to achieve comparability across all values. Subsequently, the correctness of the formulation and coding involving both models was verified by critically investigating the impact of each constraint on the results. For validation, an I-optimal crossed mixture design was constructed using Design-Expert [300] to determine the importance of each cricketing ability and the corresponding number of players required. The optimal design generated 70 runs, of which Run 2 gave the best response, i.e., the smallest objective function value. Model 1 was validated by illustrating its effectiveness in selecting a team superior to the winning franchise of the 2022 IPL. Not only did Model 1 generate a team that is approximately four times stronger overall, but it also dominated the selection of each cricketing ability. Model 2 was validated by illustrating its effectiveness in selecting a more specialised team compared to Model 1. Model 2 selected a significantly stronger bowling department, which is consistent with the specialisation enforced by the preemptive priority factors.

This study contributes to the existing body of knowledge involving the use of optimisation techniques in cricket team selection by addressing the limitations of prior work. This study employed a contemporary goal programming approach that has not been previously applied in the context of cricket team selection. It also introduced new performance measures for batting and all-rounding, further contributing to the literature covering cricket performance assessment. A distinctive data-scaling method using Gaussian membership functions was

applied to ensure comparability among the performance scores. Finally, this study applied a unique technique using an I-optimal crossed mixture design to further objectify decisions in cricket team selection that are inherently subjective.

7.1.2 Objectives achieved

Table 7.1 demonstrates how the different research objectives were achieved, along with the chapters in which they were addressed.

Table 7.1: Research objectives achieved

Objective	Chapter	Description	Achieved
1	2	Presented a comprehensive overview of cricket.	✓
2	3	Provided a comprehensive overview of optimisation theory.	✓
3	2	Devised performance measures for batting and all-rounding.	✓
4	4	Formulated and proposed two goal programming models.	✓
5	5	Verified the correctness of the goal programming models.	✓
6	6	Validated the effectiveness of the goal programming models.	✓
7	7	Proposed recommendations for future research.	✓

7.2 Limitations and recommendations

This study contains the following limitations and recommendations for future research:

- The dynamic structure of the IPL auction enables franchises to acquire players across multiple bidding rounds. The models developed in this study do not incorporate the uncertainty associated with the sequence in which these players are auctioned or the progressive decisions of competing franchises. This study assumes that each franchise independently selects a squad of players, regardless of the selections made by other franchises. Future research can address this shortcoming by developing a distinctive multi-stage stochastic programming model that sequentially alters the team selection as the auction rounds progress. This will ensure team selection occurs based on player availability at each stage of the auction.
- Objectively selecting the minimum number of players required per cricketing ability using an I-optimal crossed mixture design removed some of the underlying logic that governs squad composition. For example, in Table 6.6, the model included an equal number of batsmen, all-rounders, and wicketkeepers, with four cricketers selected for each. Selecting four wicketkeepers is illogical, as wicketkeeping is a highly specialised role that typically requires only one or two players. Future research can enhance the

proposed solutions by including constraints that reflect realistic selection requirement or by altering the ranges of the components in the mixture design. This will ensure a selection that not only yields an optimal team but also one that is logically sound.

- The proposed models focus on selecting a squad of players from the IPL auction, but do not address the assignment of players to specific roles within the playing 11. For example, the models select a holistic set of batsmen without distinguishing between specialised roles, such as top- and middle-order batsmen. The same limitation applies to bowlers, where no differentiation is made between spinners and seamers, as well as all-rounders, who may be bowling or batting dominant. Future research can enhance the current solutions by incorporating constraints that address the discipline-specific requirements of the franchise. This will strengthen the practicality of the models by allowing a franchise to optimise its squad based on the playing 11.

- The proposed models perform team selection solely based on historical performance data and do not consider contextual information, such as player relationships, injury history, leadership capabilities, temperament, versatility, and fitness. Future research can enhance the selection process by including contextual information to create more comprehensive player profiles. These refined profiles can help franchises complement historical performance data with qualitative considerations that usually play a crucial role in team success. This will result in performance measures that capture not only traditional performance data, but also contextual factors that impact team selection.

- This study assessed the historical performance data of cricketers in isolation, without accounting for relations among players. For example, it did not consider how player performance may be affected by team dynamics, where some players excel alongside certain team members while others do not. Future research can incorporate relational aspects such as the running assistance offered between wickets and batting or bowling partnerships. Including these factors will enable the proposed models to account for player cohesion during selection, which often plays a significant role in team success. Selecting a team based on preferred player combinations may produce a better squad than simply choosing the statistically best cricketers.

- While this study focused on the application of mathematical optimisation in cricket team selection, the SLR also revealed existing literature on machine learning. As an emerging field, machine learning provides significant potential to complement current methods by enabling models to adapt to prior auctioning decisions. Future research can consider using machine learning to predict player purchases and adjust selections accordingly. A further application worth investigating involves generating probability estimates of player values, supporting franchises in optimising team composition by evaluating projected performance across varying scenarios. A recent study performed by Adhikari *et al.* [4] demonstrates the value of using machine learning in this domain.

7.3 Practical considerations

The goal programming models proposed in this study prove relevant to the current cricket context, as they provide franchises with an objective framework for optimal team selection. However, both models are extremely sensitive to the input data, which depends on the performance measures used. Applying these performance measures requires an understanding of their operation, purpose, and limitations, which necessitates expert knowledge. Without this knowledge, the practical relevance of the models is limited. A further drawback of the current models is their discrete state. Making the models stochastic will enable franchises to alter their selections in real time based on players who have already been purchased and those who have not. Since the current models lack this feature, they can be used to select an optimal pool of players before the auction, providing franchises with an objective view of those who offer good value for money. In conclusion, while these models are valuable in the current cricket context, further refinement is needed to ensure their practical viability.

Bibliography

[1] D. Malcolm, *Globalizing cricket: Englishness, empire and identity*. Bloomsbury Publishing, 2012.

[2] S. McCalla, *Colonial education and Caribbean literature in C.L.R. James and George Lamming*. PhD thesis, St. John's University, 2020.

[3] M. Chunying and M. Ahmad, "History of Pakistani cricket team: Challenges and response, 1947-1970," *Annals of Human and Social Sciences*, vol. 4, no. 4, pp. 542–549, 2023.

[4] B. Adhikari and S. Ahamed, "Predicting BPL match winners: An empirical study using machine learning approach," in *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pp. 1–7, IEEE, 2023.

[5] T. B. Swartz, "Research directions in cricket," in *Handbook of Statistical Methods and Analyses in Sports*, pp. 461–476, CRC Press, 2017.

[6] F. Ahmed, K. Deb, and A. Jindal, "Multi-objective optimization and decision making approaches to cricket team selection," *Applied Soft Computing*, vol. 13, no. 1, pp. 402–414, 2013.

[7] S. R. Behera and V. V. Saradhi, "Learning strength and weakness rules of cricket players using association rule mining," in *International Workshop on Machine Learning and Data Mining for Sports Analytics (MLSA)*, pp. 79–92, Springer, 2021.

[8] S. Mukherjee, "Quantifying individual performance in cricket — A network analysis of batsmen and bowlers," *Physica A: Statistical Mechanics and its Applications*, vol. 393, pp. 624–637, 2014.

[9] H. H. Lemmer, "Team selection after a short cricket series," *European Journal of Sport Science*, vol. 13, no. 2, pp. 200–206, 2013.

[10] B. Bhattacharjee and D. Bhattacharjee, "Comparing performance of different team selection methods using fantasy league scoring system," *International Journal of Human Movement and Sports Sciences*, vol. 9, no. 6, pp. 1352–1361, 2021.

[11] V. Sharma, "Is cricket a gender biased game? India in relation to the world," *International Journal of Law Management & Humanities*, vol. 1, no. 5, pp. 138–149, 2018.

[12] S. Bullough, R. Millar, G. Ramchandani, and R. Coleman, "The effect of central contracts on the stability and performance of the England Test cricket team. [El efecto de contratos centrales sobre la estabilidad y el desempeño del equipo inglés de Test cricket]," *RICYDE. Revista Internacional de Ciencias del Deporte*, vol. 10, no. 35, pp. 4–15, 2014.

[13] S. Chakraborty, V. Kumar, and K. R. Ramakrishnan, "Selection of the all-time best World XI Test cricket team using the TOPSIS method," *Decision Science Letters*, vol. 8, pp. 95–108, 2019.

[14] A. J. Lewis, "Towards fairer measures of player performance in one-day cricket," *Journal of the Operational Research Society*, vol. 56, no. 7, pp. 804–815, 2005.

[15] M. Lourens, *Integer optimization for the selection of a Twenty20 cricket team.* MSc dissertation, Nelson Mandela Metropolitan University, 2009.

[16] G. Malhotra, "A comprehensive approach to predict auction prices and economic value creation of cricketers in the Indian Premier League (IPL)," *Journal of Sports Analytics*, vol. 8, no. 3, pp. 149–170, 2022.

[17] D. Bhattacharjee and H. Saikia, "On performance measurement of cricketers and selecting an optimum balanced team," *International Journal of Performance Analysis in Sport*, vol. 14, no. 1, pp. 262–275, 2014.

[18] H. Saikia, D. Bhattacharjee, and D. Mukherjee, "Decision making in cricket: The optimum team selection," in *Cricket Performance Management: Mathematical Formulation and Analytics*, pp. 201–231, Springer, 2019.

[19] S. Chand, H. K. Singh, and T. Ray, "Team selection using multi-/many-objective optimization with integer linear programming," in *2018 IEEE Congress on Evolutionary Computation (CEC)*, pp. 1–8, IEEE, 2018.

[20] J. Davis, H. Perera, and T. B. Swartz, "Player evaluation in Twenty20 cricket," *Journal of Sports Analytics*, vol. 1, no. 1, pp. 19–31, 2015.

[21] P. Ghanty, I. Kar, M. Singh, P. Chakraborty, and R. Baranwal, "Indian Premier League cricket score prediction: A machine learning technique," *PANCHAKOTesSAYS*, vol. 14, no. 2, pp. 38–45, 2023.

[22] V. Chittibabu and M. Sundararaman, "Base price determination for IPL mega auctions: A player performance-based approach," *Journal of Sports Analytics*, vol. 9, no. 1, pp. 77–97, 2023.

[23] N. Chauhan, P. Jain, P. Gupta, N. Kumar, and N. Bajaj, "A study of the application of operations research in the valuation of players in IPL," *International Journal of Advance Research, Ideas and Innovations in Technology*, vol. 4, no. 5, pp. 719–723, 2018.

[24] M. F. Alberts, C. Bisset, and R. L. J. Coetzer, "The application of optimisation techniques in cricket: A systematic literature review," in *34th Southern African Institute for Industrial Engineering (SAIIE) Conference*, pp. 1–22, SAIIE, 2024.

[25] S. R. Clarke and J. M. Norman, "Dynamic programming in cricket: Choosing a night watchman," *Journal of the Operational Research Society*, vol. 54, no. 8, pp. 838–845, 2003.

[26] S. S. Sathya and S. Jamal, "Applying genetic algorithm to select an optimal cricket team," in *Proceedings of the International Conference on Advances in Computing, Communication and Control (ICAC)*, pp. 43–47, Association for Computing Machinery, 2009.

[27] G. R. Amin and S. K. Sharma, "Cricket team selection using data envelopment analysis," *European Journal of Sport Science*, vol. 14, no. sup1, pp. S369–S376, 2014.

[28] J. K. Sharma, *Operations research: Theory and applications.* Laxmi Publications, 2017.

[29] C. Romero, *Handbook of critical issues in goal programming.* Pergamon Press, 2014.

[30] M. J. Coy, "Research methodologies: Increasing understanding of the world," *International Journal of Scientific and Research Publications*, vol. 9, no. 1, pp. 71–77, 2019.

[31] S. Salminawati and T. Ningsih, "Science structure as a scientific method concept (the teaching of elementary-age children)," *Jurnal Basicedu*, vol. 6, no. 2, pp. 2278–2285, 2022.

[32] R. V. Blystone and K. Blodgett, "WWW: The scientific method," *CBE—Life Sciences Education*, vol. 5, no. 1, pp. 7–11, 2006.

[33] G. Wright and T. Lavery, "Scientific method," *TechTarget*, Feb. 10, 2023. [Online]. Available: `https://www.techtarget.com/whatis/definition/scientific-method`. [Accessed: Mar. 17, 2024].

[34] R. P. Hooper, "Applying the scientific method to small catchment studies: A review of the Panola Mountain experience," *Hydrological Processes*, vol. 15, no. 10, pp. 2039–2050, 2001.

[35] D. G. Feitelson, "From repeatability to reproducibility and corroboration," *ACM SIGOPS Operating Systems Review*, vol. 49, no. 1, pp. 3–11, 2015.

[36] A. Cowen, "Steps of the scientific method," *ScienceBuddies*. [Online]. Available: `https://www.sciencebuddies.org/science-fair-projects/science-fair/steps-of-the-scientific-method`. [Accessed: Mar. 15, 2024].

[37] H. Saikia, D. Bhattacharjee, and D. Mukherjee, "Franchisee cricket and cricketers' auction," in *Cricket Performance Management: Mathematical Formulation and Analytics*, pp. 23–36, Springer, 2019.

[38] G. McLeod, S. O'Connor, D. Morgan, A. Kountouris, C. F. Finch, and L. V. Fortington, "Prospective reporting of injury in community-level cricket: A systematic review to identify research priorities," *Journal of Science and Medicine in Sport*, vol. 23, no. 11, pp. 1028–1043, 2020.

[39] D. N. Pardiwala, N. N. Rao, and A. V. Varshney, "Injuries in cricket," *Sports Health*, vol. 10, no. 3, pp. 217–222, 2018.

[40] D. Birley, *A social history of English cricket*. Aurum Press, 2013.

[41] N. Groombridge, *Sports criminology: A critical criminology of sport and games*. Policy Press, 2016.

[42] B. Pradhan, K. Banerjee, and S. Bhowmick, "Relation of selected anthropomeric parameters with performance of off drive in cricket," *International Journal of Yogic, Human Movement and Sports Sciences*, vol. 3, no. 2, pp. 143–147, 2018.

[43] A. Mote, *The glory days of cricket: The extraordinary story of Broadhalfpenny Down*. Robson Books, 1997.

[44] S. Jayawant and V. Kuttan, "Assessment and comparison of agility in cricketers depending on their playing position - An observational pilot study," *International Journal of Science and Healthcare Research*, vol. 7, no. 3, pp. 61–65, 2022.

[45] V. V. Sunkaranarayanan, *Towards a time-lapse prediction system for cricket matches*. MSc dissertation, The University of British Columbia, 2011.

[46] S. Naha and D. Malcolm, "Cricket at the beginning of the long twenty-first century," *Sport in Society*, vol. 24, no. 8, pp. 1267–1273, 2021.

[47] F. Mustafa, "Cricket and globalization: Global processes and the imperial game," *Journal of Global History*, vol. 8, no. 2, pp. 318–341, 2013.

[48] J. Knight and T. Dunmore, *Cricket for dummies*. John Wiley & Sons, 2023.

[49] A. T. Scanlan, D. M. Berkelmans, W. M. Vickery, and C. O. Kean, "A review of the internal and external physiological demands associated with batting in cricket," *International Journal of Sports Physiology and Performance*, vol. 11, no. 8, pp. 987–997, 2016.

[50] S. A. Massab, M. M. Uddin, M. A. Haleem, M. A. Bani, and M. Furkhan, "Pitch analysis for cricket ground," *International Journal of Innovative Research in Technology*, vol. 9, no. 12, pp. 1262–1267, 2023.

[51] M. W. Farooq, M. Farooq, and A. Rauf, "Critical review on Test match format of cricket," *Global Social Sciences Review*, vol. 6, no. 2, pp. 121–129, 2021.

[52] M. Kowsher, M. A. Alam, M. J. Uddin, F. Ahmed, M. W. Ullah, and M. R. Islam, "Detecting third umpire decisions & automated scoring system of cricket," in *2019 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2)*, pp. 1–8, IEEE, 2019.

[53] J. M. Adie, I. Renshaw, R. Polman, and D. L. Mann, "Home advantage and LBW decisions: The contribution of umpiring to the home-ground (dis)advantage in cricket," *Journal of Science and Medicine in Sport*, vol. 25, no. 6, pp. 520–523, 2022.

[54] H. Saikia, D. Bhattacharjee, and U. K. Radhakrishnan, "A new model for player selection in cricket," *International Journal of Performance Analysis in Sport*, vol. 16, no. 1, pp. 373–388, 2016.

[55] R. M. Tripathy, A. Bagchi, and M. Jain, "Complex network characteristics and team performance in the game of cricket," in *International Conference on Big Data Analytics*, pp. 133–150, Springer, 2013.

[56] T. Krishnamohan, "Computing the runs that should be scored every over when chasing a target in limited-overs cricket using the A* algorithm," *Applied Artificial Intelligence*, vol. 35, no. 15, pp. 2087–2101, 2021.

[57] T. B. Swartz, P. S. Gill, D. Beaudoin, and B. M. DeSilva, "Optimal batting orders in one-day cricket," *Computers & Operations Research*, vol. 33, no. 7, pp. 1939–1950, 2006.

[58] T. Krishnamohan, "Maximizing the runs scored by a team in cricket using genetic algorithm," *Journal of Electrical and Computer Engineering Research*, vol. 3, no. 3, pp. 7–12, 2023.

[59] M. Singh, "What are extras in cricket? Everying you need to know," *Zupee*. [Online]. Available: `https://www.zupee.com/blog/what-are-extras-in-cricket/`. [Accessed: Jul. 24, 2025].

[60] P. J. Rani, A. V. Kamath, A. Menon, P. Dhatwalia, D. Rishabh, and A. Kulkarni, "Selection of players and team for an Indian Premier League cricket match using ensembles of classifiers," in *2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)*, pp. 1–6, IEEE, 2020.

[61] S. Chowdhury, K. A. Islam, M. M. Rahman, T. S. Raisa, and N. M. Zayed, "One Day International (ODI) cricket match prediction in logistic analysis: India VS.

Pakistan," *International Journal of Human Movement and Sports Sciences*, vol. 8, no. 6, pp. 543–548, 2020.

[62] P. Allsopp, *Measuring team performance and modelling the home advantage effect in cricket*. PhD thesis, Swinburne University of Technology, 2005.

[63] D. Malcolm and T. Fletcher, "The International Cricket Council World Cup: A "second class" megamediasport event?," in *Sport, Media and Mega-events*, pp. 115–129, Routledge, 2017.

[64] S. Prakash and S. Bhalla, "History of men Test cricket: An overview," *International Journal of Physiology, Nutrition and Physical Education*, vol. 6, no. 1, pp. 174–178, 2021.

[65] S. Ray and S. Roychowdhury, "Cricket mix optimization using heuristic framework after ensuring Markovian equilibrium," *Journal of Sports Analytics*, vol. 7, no. 3, pp. 155–168, 2021.

[66] V. K. Borooah and J. Mangan, "Mistaking style for substance: Investor exuberance in the 2008 Indian Premier League auction," *Journal of Sports Economics*, vol. 13, no. 3, pp. 266–287, 2012.

[67] S. Mondal, D. Plumley, and R. Wilson, "The evolution of competitive balance in men's international cricket," *Managing Sport and Leisure*, vol. 28, no. 5, pp. 554–573, 2023.

[68] A. Kaluarachchi and A. S. Varde, "CricAI: A classification based tool to predict the outcome in ODI cricket," in *2010 Fifth International Conference on Information and Automation for Sustainability*, pp. 250–255, IEEE, 2010.

[69] H. Saikia, D. Bhattacharjee, and A. Bhattacharjee, "Is IPL responsible for cricketers' performance in Twenty20 World Cup," *International Journal of Sports Science and Engineering*, vol. 6, no. 2, pp. 96–110, 2012.

[70] K. M. Shah and A. A. Saeed, "Innovating a centuries-old sport: How emerging data analytics tools are redefining cricket," *Harvard Data Science Review*, no. 5.2, pp. 1–20, 2023.

[71] C. Rumford, *Twenty20 and the future of cricket*. Routledge, 2013.

[72] C. Watson, "Test match special, Twenty20 and the future of cricket," *Sport in Society*, vol. 14, no. 10, pp. 1383–1394, 2011.

[73] P. Shah, "New performance measure in cricket," *IOSR Journal of Sports and Physical Education*, vol. 4, no. 3, pp. 28–30, 2017.

[74] P. Premkumar, J. B. Chakrabarty, and S. Chowdhury, "Key performance indicators for factor score based ranking in One Day International cricket," *IIMB Management Review*, vol. 32, no. 1, pp. 85–95, 2020.

[75] W. Brettenny, *Integer optimisation for the selection of a fantasy league cricket team.* MSc dissertation, Nelson Mandela Metropolitan University, 2010.

[76] D. Bhattacharjee and H. Saikia, "Visualizing bowling performance in cricket using contour plot," *MOJ Sports Medicine*, vol. 6, no. 3, pp. 97–100, 2023.

[77] K. Mody, D. Malathi, and J. D. Jayaseeli, "An artificial neural network approach for classifying cricket batsman's performance by Adam Optimizer and prediction by Derived Attributes," in *2021 Smart Technologies, Communication and Robotics (STCR)*, pp. 1–7, IEEE, 2021.

[78] N. R. Das, I. Mukherjee, G. Paul, and R. Priya, "A multiple criteria decision making approach for ranking cricket captains," in *2023 IEEE Guwahati Subsection Conference (GCON)*, pp. 1–8, IEEE, 2023.

[79] J. Thorley, "Age-related changes in the performance of bowlers in Test match cricket," *International Journal of Sports Science & Coaching*, vol. 16, no. 5, pp. 1138–1151, 2021.

[80] H. H. Lemmer, "The combined bowling rate as a measure of bowling performance in cricket," *South African Journal for Research in Sport, Physical Education and Recreation*, vol. 24, no. 2, pp. 37–44, 2002.

[81] A. Kimber, "A graphical display for comparing bowlers in cricket," *Teaching Statistics*, vol. 15, no. 3, pp. 84–86, 1993.

[82] P. J. Van Staden, "Comparison of cricketers' bowling and batting performances using graphical displays," *Current Science*, vol. 96, no. 6, pp. 764–766, 2009.

[83] G. D. I. Barr, C. G. Holdsworth, and B. S. Kantor, "Evaluating performances at the 2007 Cricket World Cup," *South African Statistical Journal*, vol. 42, no. 2, pp. 125–142, 2008.

[84] H. Saikia, D. Bhattacharjee, and D. Mukherjee, "Quantifying performance of cricketers," in *Cricket Performance Management: Mathematical Formulation and Analytics*, pp. 37–94, Springer, 2019.

[85] G. D. I. Barr and B. S. Kantor, "A criterion for comparing and selecting batsmen in limited overs cricket," *Journal of the Operational Research Society*, vol. 55, no. 12, pp. 1266–1274, 2004.

[86] D. Beaudoin and T. Swartz, "The best batsmen and bowlers in one-day cricket," *South African Statistical Journal*, vol. 37, no. 2, pp. 203–222, 2003.

[87] H. Gerber and G. D. Sharp, "Selecting a limited overs cricket squad using an integer programming model," *South African Journal for Research in Sport, Physical Education and Recreation*, vol. 28, no. 2, pp. 81–90, 2006.

[88] J. S. Croucher, "Player ratings in one-day cricket," in *Proceedings of the Fifth Australian Conference on Mathematics and Computers in Sport (5M&CS)*, pp. 95–106, Bond University, 2000.

[89] G. L. Cohen, "Cricketing chances," in *Proceedings of the Sixth Australian Conference on Mathematics and Computers in Sport (6M&CS)*, pp. 1–13, Bond University, 2002.

[90] H. H. Lemmer, "A method for the comparison of the bowling performances of bowlers in a match or a series of matches," *South African Journal for Research in Sport, Physical Education and Recreation*, vol. 27, no. 1, pp. 91–103, 2005.

[91] H. H. Lemmer, "Individual match approach to bowling performance measures in cricket," *South African Journal for Research in Sport, Physical Education and Recreation*, vol. 34, no. 2, pp. 95–103, 2012.

[92] M. A. I. Hossain and A. Tareq, "Biomechanical analysis of cricket batting of BKSP cricketers," *Journal of Physical Education Research*, vol. 2, no. 2, pp. 7–13, 2015.

[93] S. Verma, V. Pandey, M. Pant, and V. Snasel, "A balanced squad for Indian Premier League using Modified NSGA-II," *IEEE Access*, vol. 10, pp. 100463–100477, 2022.

[94] O. G. Stevenson and B. J. Brewer, "Finding your feet: A Gaussian process model for estimating the abilities of batsmen in Test cricket," *Journal of the Royal Statistical Society Series C: Applied Statistics*, vol. 70, no. 2, pp. 481–506, 2021.

[95] A. C. Kimber and A. R. Hansford, "A statistical analysis of batting in cricket," *Journal of the Royal Statistical Society Series A: Statistics in Society*, vol. 156, no. 3, pp. 443–455, 1993.

[96] G. D. I. Barr and R. Van den Honert, "Evaluating batsmen's scores in Test cricket," *South African Statistical Journal*, vol. 32, no. 2, pp. 169–183, 1997.

[97] H. H. Lemmer, "A measure for the batting performance of cricket players," *South African Journal for Research in Sport, Physical Education and Recreation*, vol. 26, no. 1, pp. 55–64, 2004.

[98] H. H. Lemmer, "Measures of batting performance in a short series of cricket matches," *South African Statistical Journal*, vol. 42, no. 1, pp. 65–87, 2008.

[99] H. H. Lemmer, "The single match approach to strike rate adjustments in batting performance measures in cricket," *Journal of Sports Science & Medicine*, vol. 10, no. 4, pp. 630–634, 2011.

[100] V. K. Borooah and J. E. Mangan, "The "Bradman class": An exploration of some issues in the evaluation of batsmen for Test matches, 1877-2006," *Journal of Quantitative Analysis in Sports*, vol. 6, no. 3, pp. 1–19, 2010.

[101] U. Damodaran, "Stochastic dominance and analysis of ODI batting performance: The Indian cricket team, 1989-2005," *Journal of Sports Science & Medicine*, vol. 5, no. 4, pp. 503–508, 2006.

[102] I. Wickramasinghe, "Classification of all-rounders in the game of ODI cricket: Machine learning approach," *Athens Journal of Sports*, vol. 7, no. 1, pp. 21–34, 2020.

[103] H. Saikia and D. Bhattacharjee, "On classification of all-rounders of the Indian Premier League (IPL): A Bayesian approach," *Vikalpa*, vol. 36, no. 4, pp. 51–66, 2011.

[104] M. T. Suleman and M. A. Saeed, "Option on human performance," *SSRN*. [Online]. Available: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1474039`. [Accessed: Jul. 28, 2025].

[105] A. Tan and R. Ramachandran, "Ranking the greatest all-rounders in Test cricket," *The Journal of the Cricket Society*, vol. 22, no. 4, pp. 7–10, 2006.

[106] G. D. Sharp, W. J. Brettenny, J. W. Gonsalves, M. Lourens, and R. A. Stretch, "Integer optimisation for the selection of a Twenty20 cricket team," *Journal of the Operational Research Society*, vol. 62, no. 9, pp. 1688–1694, 2011.

[107] A. B. W. Manage, R. C. Kafle, and D. K. Wijekularathna, "Classification of all-rounders in limited over cricket - A machine learning approach," *Journal of Sports Analytics*, vol. 6, no. 4, pp. 295–306, 2020.

[108] M. Gheorghe and R. Petre, "The importance of normalization method for mining medical data," *International Journal of Computers & Technology*, vol. 14, no. 8, pp. 6014–6020, 2015.

[109] R. M. Prabhu, G. Hema, S. Chepure, and M. N. Guptha, "Logistics optimization in supply chain management using clustering algorithms," *Scalable Computing: Practice and Experience*, vol. 21, no. 1, pp. 107–114, 2020.

[110] M. Zirekgür and B. Karakaya, "FPGA implementation of deep learning model utilizing different normalization algorithms for COVID-19 diagnosis," *Niğde Ömer Halisdemir University Journal of Engineering Sciences*, vol. 13, no. 3, pp. 905–916, 2024.

[111] A. Masitha, M. K. Biddinika, and Herman, "Preparing dual data normalization for KNN classfication in prediction of heart failure," *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 4, no. 3, pp. 1227–1234, 2023.

[112] T. Guo, Y. Chen, and M. Kriegel, "Novel multi-level optimization of district heating systems: Managing spatial scale and equipment portfolio design," *Energy and Buildings*, vol. 317, pp. 1–18, 2024.

[113] D. Rachmat and A. B. Nugroho, "Portfolio determination and Markowitz efficient frontier in five Indonesian industrial sector," *Journal of Business and Management*, vol. 2, no. 1, pp. 116–131, 2013.

[114] A. Ganguly and P. Prakash, "Investment and retirement planning - A conceptual analysis," *International Journal of Professional Business Review*, vol. 8, no. 8, pp. 1–16, 2023.

[115] M. Benchener and L. Li, "Using efficiency frontier to inform operational decision-making," *SN Applied Sciences*, vol. 3, no. 5, pp. 1–10, 2021.

[116] G. Perez-Verdin, J. C. Monarrez-Gonzalez, A. Tecle, and M. Pompa-Garcia, "Evaluating the multi-functionality of forest ecosystems in northern Mexico," *Forests*, vol. 9, no. 4, pp. 1–14, 2018.

[117] R. M. V. Ishara and T. G. D. K. Sumanathilaka, "Wicket keeping technique analysis with deep neural networks and pose estimations," in *Proceedings of the International Conference on Innovations in Info-business & Technology (ICIIT)*, pp. 111–113, Informatics Institute of Technology, 2022.

[118] C. J. Petersen, *Elite cricket match and training demands and performance enhancement in hot/humid environments.* PhD thesis, The University of Western Australia, 2010.

[119] V. Patel, P. S. Vijayakumar, and A. U. Sahana, "To develop and validate an integrated yoga module for the sport of cricket," *International Journal of Physical Education, Sports and Health*, vol. 8, no. 4, pp. 294–299, 2021.

[120] H. H. Lemmer, "Performance measures for wicket keepers in cricket," *South African Journal for Research in Sport, Physical Education and Recreation*, vol. 33, no. 3, pp. 89–102, 2011.

[121] A. Narayanan, "Test wicketkeepers - An analysis," *The ESPNcricinfo Blog.* [Online]. Available: `https://blogs.cricinfo.com/itfigures/archives/wicketkeepers/`. [Accessed: Jun. 29, 2024].

[122] B. Hemachandran, "Cricket's best wicket keepers," *Cricket Discussion Forum.* [Online]. Available: `https://www.cricforum.com/viewtopic.php?t=15492`. [Accessed: Jun. 29, 2024].

[123] D. Barry, "Analysing wicketkeepers by byes conceded," *Wicketkeepers.* [Online]. Available: `https://blogs.cricinfo.com/itfigures/archives/wicketkeepers/`. [Accessed: Jun. 29, 2024].

[124] A. Samson, "The bat and ball brimborion: Is Mark Boucher the new Jack Blackham?," *A Blog of Cricket Stats and More.* [Online]. Avail-

able: https://batandballbrimbirio.blogspot.com/2007/10/comparing-wicketkeepers-of-different.html. [Accessed: Jun. 29, 2024].

[125] D. MacDonald, J. Cronin, and P. Macadam, "Key movements and skills of wicket-keepers in One Day International cricket," *International Journal of Sports Science & Coaching*, vol. 13, no. 6, pp. 1156–1162, 2018.

[126] C. Goyal, "IPL-2017 cross country cluster analysis," *International Journal of Computer Science Trends and Technology*, vol. 5, no. 4, pp. 117–122, 2017.

[127] J. Farooqui, "IPL brand value rise 28% to $10.7 billion in 2023: Brand finance," *The Economic Times*. [Online]. Available: https://economictimes.indiatimes.com/news/sports/ipl-brand-value-rises-28-to-10-7-billion-in-2023-brand-finance/articleshow/105967823.cms?from=mdr. [Accessed: Jan. 24, 2024].

[128] S. S. Birunda, P. Nagaraj, B. J. A. Jebamani, B. Revathi, and V. Muneeswaran, "A structured analysis on IPL 2022 matches by approaching various data visualization and analytics," in *2023 International Conference on Computer Communication and Informatics (ICCCI)*, pp. 1–9, IEEE, 2023.

[129] J. Mathews, "Contract of agency in IPL," *Jus Corpus Law Journal*, vol. 3, no. 3, pp. 18–24, 2023.

[130] ESPNcricinfo, "IPL 2022 auction purse swells to INR 90 crore," *ESPNcricinfo*. [Online]. Available: https://www.espncricinfo.com/story/ipl-2022-auction-purse-swells-to-inr-90-crore-1286198. [Accessed: Jul. 18, 2025].

[131] V. Dikshit, "FAQs: All you wanted to know about the IPL 2022 auction," *ESPNcricinfo*. [Online]. Available: https://www.espncricinfo.com/story/faqs-all-you-wanted-to-know-about-the-ipl-2022-auction-1300565. [Accessed: Jun. 22, 2024].

[132] A. Bansal, "5 things all franchises have to keep in mind while bidding at the IPL 2021 auction," *Sportskeeda*. [Online]. Available: https://www.sportskeeda.com/cricket/news-ipl-2021-auction-5-rules-to-follow. [Accessed: Jul. 18, 2025].

[133] N. Gollapudi, "Dhoni, Kohli, Rohit, Bumrah retained by IPL franchises for 2022," *ESPNcricinfo*. [Online]. Available: https://www.espncricinfo.com/story/ipl-2022-ms-dhoni-virat-kohli-rohit-sharma-jasprit-bumrah-retained-by-ipl-franchises-1291658. [Accessed: Jun. 21, 2024].

[134] G. Gokul and M. Sundararaman, "Determining the playing 11 based on opposition squad: An IPL illustration," *Journal of Sports Analytics*, vol. 9, no. 3, pp. 191–203, 2023.

[135] P. Chandiramani, "Transfer of players in Indian Premier League - A critique," *Legal Spectrum Journal*, vol. 1, no. 2, pp. 1–7, 2021.

[136] H. Saikia and D. Bhattacharjee, "Survival ability of Indian and overseas batsmen on the cricket pitch in Indian Premier League," *MOJ Sports Medicine*, vol. 2, no. 4, pp. 113–116, 2018.

[137] V. V. Goncharov, "The Fourth Industrial Revolution: Challenges, risks and opportunities," *Eruditio*, vol. 2, no. 6, pp. 95–106, 2020.

[138] S. I. Gass and A. A. Assad, "History of operations research," in *Transforming Research into Action*, pp. 1–14, INFORMS, 2011.

[139] M. Persson, *On the improvement of healthcare management using simulation and optimisation.* PhD thesis, Blekinge Institute of Technology, 2010.

[140] S. Russenschuck, "Mathematical optimization techniques," in *International Roxie Users Meeting and Workshop*, pp. 60–72, CERN, 1999.

[141] S. P. Bradley, A. C. Hax, and T. L. Magnanti, "Mathematical programming: An overview," in *Applied Mathematical Programming*, pp. 1–37, Addison-Wesley, 1977.

[142] S. M. Sinha, *Mathematical programming: Theory and methods.* Elsevier, 2006.

[143] A. D. Snider, *Basics of optimization theory.* Springer, 2023.

[144] R. Luies, *Towards optimal route planning for solar-powered battery electric vehicles.* PhD thesis, North-West University, 2021.

[145] L. R. Foulds, *Optimization techniques: An introduction.* Springer, 1981.

[146] S. Subhakala, S. Bhuvana, and R. Radhakrishnan, "An optimized CBIR using particle swarm optimization algorithm," *International Journal of Computer Science and Business Informatics*, vol. 11, no. 1, pp. 40–55, 2014.

[147] C. Lewis, *Linear programming: Theory and applications.* Whitman College, 2008.

[148] H. P. Williams, *Model building in mathematical programming.* John Wiley & Sons, 2013.

[149] P. S. H. Leeflang, D. R. Wittink, M. Wedel, and P. A. Naert, "Specifying models according to intended use," in *Building Models for Marketing Decisions*, pp. 123–155, Springer, 2000.

[150] L. O. Tedeschi, "The prevailing mathematical modelling classifications and paradigms to support the advancement of sustainable animal production," *Animal*, vol. 17, no. Sup5, pp. 1–13, 2023.

[151] R. G. Parker and R. L. Rardin, *Discrete optimization.* Academic Press, 1988.

[152] S. J. Wright, "Continuous optimization (nonlinear and linear programming)," in *The Princeton Companion to Applied Mathematics*, pp. 281–293, Princeton University press, 2015.

[153] H. Suprajitno, "Solving multiobjective linear programming problem using interval arithmetic," *Applied Mathematical Sciences*, vol. 6, no. 80, pp. 3959–3968, 2012.

[154] G. B. Dantzig, "Linear programming," *Operations Research*, vol. 50, no. 1, pp. 42–47, 2002.

[155] K. Muzhinji and N. Ndou, "A cash crop combination for maximum net income: A case study of a small-scale cash crop farmer in Vhembe, Limpopo province, South Africa," *African Journal of Food, Agriculture, Nutrition and Development*, vol. 22, no. 7, pp. 21056–21071, 2022.

[156] E. K. P. Chong and S. H. Żak, *An introduction to optimization.* John Wiley & Sons, 2013.

[157] W. L. Winston and J. B. Goldberg, *Operations research: Applications and algorithms.* Thomson: Brooks/Cole, 2004.

[158] F. S. Hillier and G. J. Lieberman, *Introduction to operations research.* McGraw-Hill Education, 2015.

[159] G. N. Osagie and M. O. Icheme, "Optimization of resources with the aid of linear programming model in Steve Shoe Production Centre, Enugu, Nigeria," *International Journal of Research - Granthaalayah*, vol. 6, no. 9, pp. 123–136, 2018.

[160] C. Bisset, *A stochastic programming approach for marketing campaign optimisation.* MSc dissertation, North-West University, 2022.

[161] F. Tardella, "The fundamental theorem of linear programming: Extensions and applications," *Optimization*, vol. 60, no. 1–2, pp. 283–301, 2011.

[162] D. G. Luenberger and Y. Ye, *Linear and nonlinear programming.* Springer, 2008.

[163] R. J. Vanderbei, *Linear programming: Foundations and extensions.* Springer, 2020.

[164] M. R. Hussain, M. Qayyum, and M. E. Hussain, "Effect of seven steps approach on simplex method to optimize the mathematical manipulation," *International Journal of Recent Technology and Engineering*, vol. 7, no. 5, pp. 34–43, 2019.

[165] S. A. Edalatpanah, "A paradigm shift in linear programming: An algorithm without artificial variables," *Systematic Analytics*, vol. 1, no. 1, pp. 1–10, 2023.

[166] M. Banciu, "Dual simplex," in *Wiley Encyclopedia of Operations Research and Management Science*, pp. 1–13, John Wiley & Sons, 2010.

[167] J. Van Niekerk, *Mixed integer linear programming for unit commitment and load dispatch optimisation.* MSc dissertation, North-West University, 2019.

[168] J. Pfeifle, "Positive plücker tree certificates for non-realizability," *Experimental Mathematics*, vol. 33, no. 1, pp. 69–85, 2024.

[169] D.-S. Chen, R. G. Batson, and Y. Dang, *Applied integer programming: Modeling and solution.* John Wiley & Sons, 2010.

[170] M. Jünger, T. M. Liebling, D. Naddef, G. L. Nemhauser, W. R. Pulleyblank, G. Reinelt, G. Rinaldi, and L. A. Wolsey, *50 Years of integer programming 1958-2008: From the early years to the state-of-the-art.* Springer, 2010.

[171] H. A. Taha, *Integer programming: Theory, applications, and computations.* Academic Press, 2014.

[172] M. Padberg, "Mixed-integer programming - 1968 and thereafter," *Annals of Operations Research*, vol. 149, no. 1, pp. 163–175, 2007.

[173] C. Kastning, *Integer programming and related areas: A classified bibliography.* Springer, 1976.

[174] R. Eshuis and A. Kumar, "An integer programming based approach for verification and diagnosis of workflows," *Data & Knowledge Engineering*, vol. 69, no. 8, pp. 816–835, 2010.

[175] K. P. Adams, *An approach to real-time adaptive decision making in dynamic distributed systems.* PhD thesis, Virginia Polytechnic Institute and State University, 2005.

[176] M. Conforti, G. Cornuéjols, and G. Zambelli, "Polyhedral approaches to mixed integer linear programming," in *50 Years of integer programming 1958-2008: From the early years to the state-of-the-art*, pp. 343–385, Springer, 2010.

[177] R. Colares, *Exploring combinatorial aspects of the stop number problem.* PhD thesis, Université Clermont Auvergne, 2019.

[178] J. Cho and A. Papavasiliou, "Exact mixed-integer programming approach for chance-constrained multi-area reserve sizing," *IEEE Transactions on Power Systems*, vol. 39, no. 2, pp. 3310–3323, 2024.

[179] I. Hossain and M. B. Hasan, "A decomposition technique for solving integer programming problems," *GANIT: Journal of Bangladesh Mathematical Society*, vol. 33, pp. 1–11, 2013.

[180] L. A. Wolsey, *Integer programming.* John Wiley & Sons, 2021.

[181] S. N. S. Ahamad and N. H. A. Ghani, "A binary integer programming model for personnel scheduling: A case study at fast-food restaurant in Johor," *Journal of Advanced Research in Applied Sciences and Engineering Technology*, vol. 30, no. 3, pp. 334–347, 2023.

[182] A. Giudici, P. Halffmann, S. Ruzika, and C. Thielen, "Approximation schemes for the parametric knapsack problem," *Information Processing Letters*, vol. 120, pp. 11–15, 2017.

[183] T. C. Hu, L. Landa, and M.-T. Shing, "The unbounded knapsack problem," in *Research trends in combinatorial optimization*, pp. 201–217, Springer, 2009.

[184] M. X. Cheng, Y. Li, and D.-Z. Du, *Combinatorial optimization in communication networks*. Springer, 2006.

[185] M. E. Keskin and M. G. Güler, "Influence maximization in social networks: An integer programming approach," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 26, no. 6, pp. 3383–3396, 2018.

[186] H. A. Taha, *Operations research: An introduction*. Pearson Education, 2007.

[187] Y. Zhu and T. Kuno, "A disjunctive cutting-plane-based branch-and-cut algorithm for 0-1 mixed-integer convex nonlinear programs," *Industrial & Engineering Chemistry Research*, vol. 45, no. 1, pp. 187–196, 2006.

[188] N. K. Yilmaz, *Path planning of autonomous underwater vehicles for adaptive sampling*. PhD thesis, Massachusetts Institute of Technology, 2005.

[189] W. Cook, "Markowitz and Manne + Eastman + Land and Doig = Branch and bound," in *Optimization Stories: 21st International Symposium on Mathematical Programming*, pp. 227–238, Zuse Institute Berlin, 2012.

[190] T. Van Niekerk, *Exact approaches towards solving generator maintenance scheduling problems*. MSc dissertation, North-West University, 2019.

[191] D. Chen, *A branch and cut algorithm for the halfspace depth problem*. MSc dissertation, Liaoning University, 2003.

[192] E. Lobato, P. Sánchez-Martín, M. Rajabdorri, and L. Sigrist, "Practical solutions to limit computational burden of UC in island power systems," in *2024 20th International Conference on the European Energy Market (EEM)*, pp. 1–6, IEEE, 2024.

[193] M. Kormilitsin, R. Chirkova, Y. Fathi, and M. F. Stallmann, "Plan-based view and index selection for query-performance improvement," *ResearchGate*. [Online]. Available: `https://www.researchgate.net/publication/228554427_Plan-based_view_and_index_selection_for_query-performance_improvement`. [Accessed: Sep. 19, 2024].

[194] C. D. Sánchez, *The rank pricing problem: A mixed-integer linear optimization approach*. PhD thesis, Universidad de Murcia, 2021.

[195] J. P. Brooks, "Support vector machines with the ramp loss and the hard margin loss," *Operations Research*, vol. 59, no. 2, pp. 467–479, 2011.

[196] H. Stadtler, "Linear and mixed integer programming," in *Supply Chain Management and Advanced Planning: Concepts, Models, Software, and Case Studies*, pp. 473–483, Springer, 2005.

[197] L. N. Hoang, "Optimization by integer programming," *Science4All*. [Online]. Available: `https://www.science4all.org/article/integer-programming/`. [Accessed: Sep. 7, 2024].

[198] B. T. Downs, *A robust dual-based solution algorithm for the maximal covering problem*. PhD thesis, University of Cincinnati, 1991.

[199] C. M. O. Pimentel, F. P. e. Alvelos, and J. M. Valério de Carvalho, "Comparing Dantzig-Wolfe decompositions and branch-and-price algorithms for the multi-item capacitated lot sizing problem," *Optimization Methods & Software*, vol. 25, no. 2, pp. 299–319, 2010.

[200] A. Moura, J. Oliveira, and C. Pimentel, "A mathematical model for the container stowage and ship routing problem," *Journal of Mathematical Modelling and Algorithms in Operations Research*, vol. 12, no. 3, pp. 217–231, 2013.

[201] I. M. Hezam, S. A. H. Taher, A. Foul, and A. F. Alrasheedi, "Healthcare's sustainable resource planning using neutrosophic goal programming," *Journal of Healthcare Engineering*, vol. 2022, pp. 1–23, 2022.

[202] M. J. Schniederjans, *Goal programming: Methodology and applications*. Kluwer Academic Publishers, 1995.

[203] J. Wang, D. Wang, and A. Li, "Goal programming and its variants," in *Encyclopedia of Decision Making and Decision Support Technologies*, pp. 410–417, IGI Global, 2008.

[204] M. Tamiz, D. F. Jones, and E. El-Darzi, "A review of goal programming and its applications," *Annals of Operations Research*, vol. 58, no. 1, pp. 39–53, 1995.

[205] T. J. Stewart, "Goal programming and cognitive biases in decision-making," *Journal of the Operational Research Society*, vol. 56, no. 10, pp. 1166–1175, 2005.

[206] H. Taherdoost and M. Madanchian, "Multi-criteria decision making (MCDM) methods and concepts," *Encyclopedia*, vol. 3, no. 1, pp. 77–87, 2023.

[207] S. Zionts, "Some thoughts on research in multiple criteria decision making," *Computers & Operations Research*, vol. 19, no. 7, pp. 567–570, 1992.

[208] H. Borzęcka, "Multi-criteria decision making using fuzzy preference relations," *Operations Research and Decisions*, vol. 22, no. 3, pp. 5–21, 2012.

[209] M. Tamiz, D. Jones, and C. Romero, "Goal programming for decision making: An overview of the current state-of-the-art," *European Journal of Operational Research*, vol. 111, no. 3, pp. 569–581, 1998.

[210] E. Fararah, E. A. Nasr, and H. A. Mahmoud, "Multi-objective optimization modeling of integrated supply chain for solid waste treatment," *International Journal of Industrial Engineering*, vol. 30, no. 1, pp. 147–165, 2023.

[211] S. Mardle and S. Pascoe, "A review of applications of multiple-criteria decision-making techniques to fisheries," *Marine Resource Economics*, vol. 14, no. 1, pp. 41–63, 1999.

[212] B. Xue, L. Cervante, L. Shang, and M. Zhang, "A particle swarm optimisation based multi-objective filter approach to feature selection for classification," in *PRICAI 2012: Trends in Artificial Intelligence - 12th Pacific Rim International Conference on Artificial Intelligence*, pp. 673–685, Springer, 2012.

[213] J. Szlapczynska, "Multiobjective weather routing with customised criteria and constraints," *The Journal of Navigation*, vol. 68, no. 2, pp. 338–354, 2015.

[214] S. T. Şahin and S. Sahin, "Determination of sample size selecting from strata under nonlinear cost constraint by using goal programming and Kuhn-Tucker methods," *Gazi University Journal of Science*, vol. 24, no. 2, pp. 249–262, 2011.

[215] E. Xevi and S. Khan, "Computer software for nodal network modelling, simulation and optimisation of water resources management," in *MODSIM05: International Congress on Modelling and Simulation*, pp. 1992–1997, Modelling & Simulation Society of Australia and New Zealand, 2005.

[216] H. A. Abbas Amein, *Computational intelligence techniques for decision making with applications to the dairy industry*. PhD thesis, Queensland University of Technology, 2000.

[217] F. Salas-Molina, J. A. Rodríguez-Aguilar, and F. Bistaffa, "Shared value economics: An axiomatic approach," *arXiv*. [Online]. Available: `https://arxiv.org/pdf/2006.00581`. [Accessed: Sep. 12, 2024].

[218] S. G. Winter, "Satisficing," in *The Palgrave Encyclopedia of Strategic Management*, pp. 1509–1513, Macmillan Publishers, 2018.

[219] B. Sinha and N. Sen, "Goal programming approach to tea industry of Barak Valley of Assam," *Applied Mathematical Sciences*, vol. 5, no. 29, pp. 1409–1419, 2011.

[220] C. Romero and T. Rehman, "Goal programming and multiple criteria decision-making in farm planning: An expository analysis," *Journal of Agricultural Economics*, vol. 35, no. 2, pp. 177–190, 1984.

[221] T. Lust, *New metaheuristics for solving MOCO problems: Application to the knapsack problem, travelling salesman problem and IMRT optimization.* PhD thesis, Université de Mons, 2010.

[222] P. W. Srivastava and T. Gupta, "Optimum time-censored modified ramp-stress ALT for the Burr Type XII distribution with warranty: A goal programming approach," *International Journal of Reliability, Quality and Safety Engineering*, vol. 22, no. 3, pp. 1–23, 2015.

[223] S. Tiwari and A. Kumar, "Material requirement planning (MRP) using goal programming method," *International Journal of Advanced Scientific Research and Management*, vol. 3, no. 6, pp. 225–228, 2018.

[224] A. K. Bhargava, S. R. Singh, and D. Bansal, "Production planning using fuzzy metagoal programming model," *Indian Journal of Science and Technology*, vol. 8, no. 34, pp. 1–9, 2015.

[225] Y. Jiang, J. Wang, and L. Song, "Quantitative analysis of cost savings and occupants' preferences in grid-interactive smart home operation," *Science and Technology for the Built Environment*, vol. 29, no. 9, pp. 845–857, 2023.

[226] U. C. Orumie and D. W. Ebong, "Comparisons of linear goal programming algorithms," *Mathematical Theory and Modeling*, vol. 3, no. 13, pp. 9–23, 2013.

[227] S. M. Lee, "Goal programming," in *Mathematical Programming for Operations Researchers and Computer Scientists*, pp. 101–122, CRC Press, 2020.

[228] J. C. Van Horn, *A goal programming model for resource allocation in higher education.* PhD thesis, University of Nebraska, 1978.

[229] B. O. Akinnuli, P. K. Farayibi, and O. O. Ojo, "Manufacturing procurement cost allocation as dominant factor under limited available manufacturing equipment budget," *International Journal of Mathematics in Operational Research*, vol. 17, no. 2, pp. 278–299, 2020.

[230] A. A. Kindo, G. Kaladzavi, S. Malo, G. Camara, T. M. Y. Tapsoba, and D. T. Kolyang, "Fuzzy logic approach for knowledge modeling in an ontology: A review," in *2020 IEEE 2nd International Conference on Smart Cities and Communities (SCCIC)*, pp. 1–8, IEEE, 2020.

[231] E. Herrera-Viedma, "Fuzzy sets and fuzzy logic in multi-criteria decision making. The 50th anniversary of Prof. Lotfi Zadeh's theory: Introduction," *Technological and Economic Development of Economy*, vol. 21, no. 5, pp. 677–683, 2015.

[232] O. A. M. Ali, A. Y. Ali, and B. S. Sumait, "Comparison between the effects of different types of membership functions on fuzzy logic controller performance," *International*

Journal of Emerging Engineering Research and Technology, vol. 3, no. 3, pp. 76–83, 2015.

[233] J. K. Kumari and K. Ramu, "A high step-up DC converter for renewable energy resources using fuzzy logic technique," *International Journal of Scientific Engineering and Technology Research*, vol. 3, no. 41, pp. 8311–8318, 2014.

[234] S. A. Siddiqui, A. Q. Ansari, and S. Agarwal, "A journey through fuzzy philosophy," *Pranjana: The Journal of Management Awareness*, vol. 6, no. 2, pp. 29–33, 2003.

[235] S. Singh, S. Singh, and V. K. Banga, "Design of fuzzy logic system framework using evolutionary techniques," *Soft Computing*, vol. 24, no. 6, pp. 4455–4468, 2020.

[236] F. Diker, "Evaluation of the efficiency of university classrooms with the fuzzy logic," *Facta Universitatis, Series: Architecture and Civil Engineering*, vol. 17, no. 4, pp. 341–357, 2019.

[237] S. N. Sivanandam, S. Sumathi, and S. N. Deepa, *Introduction to fuzzy logic using MATLAB*. Springer, 2007.

[238] R. K. Bisht, V. Bhatt, and B. C. Joshi, "Fuzzy logic approach for evaluation of general proficiency of students," in *AIP Conference Proceedings*, pp. 1–8, AIP Publishing, 2022.

[239] P. S. Sajja, "Computer aided development of fuzzy, neural and neuro-fuzzy systems," *International Journal of Engineering and Applied Computer Science*, vol. 2, no. 1, pp. 10–17, 2017.

[240] Z. Guozheng, "Research on supplier selection based on fuzzy sets group decision," in *2009 Second International Symposium on Computational Intelligence and Design*, pp. 529–531, IEEE, 2009.

[241] E. Petritoli and F. Leccese, "Precise Takagi-Sugeno fuzzy logic system for UAV longitudinal stability: An Industry 4.0 case study for aerospace," *ACTA IMEKO*, vol. 9, no. 4, pp. 106–113, 2020.

[242] T. J. Ross, *Fuzzy logic with engineering applications*. John Wiley & Sons, 2010.

[243] Y. Bai and D. Wang, "Fundamentals of fuzzy logic control—fuzzy sets, fuzzy rules and defuzzifications," in *Advanced Fzzy Logic Technologies in Industrial Applications*, pp. 17–36, Springer, 2006.

[244] V. Novitzká, W. Steingartner, and V. Richnavská, "Fuzzy logic for educational purposes," *IPSI Transactions on Internet Research*, vol. 18, no. 2, pp. 82–87, 2022.

[245] J. Kaur and P. Sethi, "Evaluation of fuzzy inference system in image processing," *International Journal of Computer Applications*, vol. 68, no. 22, 2013.

[246] P. Smith and S. Greenfield, "Towards refined autism screening: A fuzzy logic approach with a focus on subtle diagnostic challenges," *Mathematics*, vol. 12, no. 13, pp. 1–28, 2024.

[247] C. Fuchs, A. Wilbik, and U. Kaymak, "Towards more specific estimation of membership functions for data-driven fuzzy inference systems," in *2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pp. 1–8, IEEE, 2018.

[248] A. De Oliveira Gomes and J. C. Basilio, "A fuzzy inference model to identify the current industry maturity stage in the transformation process to Industry 4.0," *IEEE Transactions on Automation Science and Engineering*, vol. 21, no. 2, pp. 1607–1622, 2024.

[249] S. Rizvi, J. Mitchell, A. Razaque, M. R. Rizvi, and I. Williams, "A fuzzy inference system (FIS) to evaluate the security readiness of cloud service providers," *Journal of Cloud Computing: Advances, Systems, and Applications*, vol. 9, no. 42, pp. 1–17, 2020.

[250] H. Behbahani, H. Ziari, A. Amini, V. N. M. Gilani, and R. Salehfard, "Investigation of un-signalized roundabouts delay with adaptive-network-based fuzzy inference system and fuzzy logic," *Computational Research Progress in Applied Science & Engineering*, vol. 2, no. 4, pp. 140–149, 2016.

[251] A. Guidara, *Policy decision modeling with fuzzy logic: Theoretical and computational aspects*. Springer Nature, 2021.

[252] E. Bilgin Sari, "Fuzzy based failure mode and effect analysis towards to risks of autonomous maintenance activities: As a TPM implementation," *EGE Academic Review*, vol. 21, no. 1, pp. 17–27, 2021.

[253] S. Hrehova, J. Husár, and V. Hladkỳ, "Possibility of using matlab application to propose fuzzy computer model," in *IOP Conference Series: Materials Science and Engineering*, pp. 1–9, IOP Publications, 2021.

[254] A. Jain and A. Sharma, "Membership function formulation methods for fuzzy logic systems: A comprehensive review," *Journal of Critical Reviews*, vol. 7, no. 19, pp. 8717–8733, 2020.

[255] Z.-D. Xu, Y.-Q. Guo, J.-T. Zhu, and F.-H. Xu, *Intelligent vibration control in civil engineering structures*. Academic Press, 2017.

[256] I. Dragović, N. Turajlić, D. Pilčević, B. Petrović, and D. Radojević, "A Boolean consistent fuzzy inference system for diagnosing diseases and its application for determining peritonitis likelihood," *Computational and Mathematical Methods in Medicine*, vol. 2015, pp. 1–10, 2015.

[257] J. J. Martínez, J. A. Padilla-Medina, S. Cano-Andrade, A. Sancen, J. Prado, and A. I. Barranco, "Development and application of a fuzzy control system for a lead-acid battery bank connected to a DC microgrid," *International Journal of Photoenergy*, vol. 2018, no. 1, pp. 1–14, 2018.

[258] S. Tripathy, D. Sahoo, M. Satpathy, and M. Mutyam, "Fuzzy fairness controller for NVMe SSDs," in *Proceedings of the 34th ACM International Conference on Supercomputing*, pp. 1–12, Association for Computing Machinery, 2020.

[259] S. Prajapati and E. Fernandez, "Performance evaluation of membership function on fuzzy logic model for solar PV array," in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, pp. 609–613, IEEE, 2020.

[260] M. H. Azam, M. H. Hasan, S. J. A. Kadir, and S. Hassan, "Prediction of sunspots using fuzzy logic: A triangular membership function-based fuzzy c-means approach," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 2, pp. 357–362, 2021.

[261] S. Porębski and E. Straszecka, "Membership functions for fuzzy focal elements," *Archives of Control Sciences*, vol. 26, no. 3, pp. 395–427, 2016.

[262] B. Vallamsundar, *Congestion control for adaptive satellite communication systems with intelligent systems.* MSc dissertation, University of Waterloo, 2007.

[263] S. H. Khairuddin, M. H. Hasan, and M. A. Hashmani, "Integration of cluster centers and Gaussian distributions in fuzzy c-means for the construction of trapezoidal membership function," *Mathematics and Statistics*, vol. 8, no. 5, pp. 559–565, 2020.

[264] C. S. Reddy and K. V. S. N. Raju, "An improved fuzzy approach for COCOMO's effort estimation using Gaussian membership function," *Journal of Software*, vol. 4, no. 5, pp. 452–459, 2009.

[265] I. A. Hameed, "Using Gaussian membership functions for improving the reliability and robustness of students' evaluation systems," *Expert Systems with Applications*, vol. 38, no. 6, pp. 7135–7142, 2011.

[266] A. Sadollah, "Introductory chapter: Which membership function is appropriate in fuzzy system?," in *Fuzzy logic based in Optimization Methods and Control Systems and its Applications*, IntechOpen, 2018.

[267] J. Z. Sąsiadek, S. Ulrich, and A. Krzyżak, "Direct fuzzy adaptive control and nonparametric identification of robot manipulator with elastic joints," in *2017 22nd International Conference on Methods and Models in Automation and Robotics (MMAR)*, pp. 119–124, IEEE, 2017.

[268] R. Prinz, *Synthesizing the sit-to-stand movement using fuzzy logic-based control and a simple biomechanical model.* MSc dissertation, University of Victoria, 2005.

[269] Z. Zhang, J. Tang, J. Zhang, and T. Zhang, "Research on energy hierarchical management and optimal control of compound power electric vehicle," *Energies*, vol. 17, no. 6, pp. 1–22, 2024.

[270] Z. Li, J. Wang, B. Li, J. Gao, and X. Tan, "GPS/INS/Odometer integrated system using fuzzy neural network for land vehicle navigation applications," *The Journal of Navigation*, vol. 67, no. 6, pp. 967–983, 2014.

[271] A. D. Sonar, R. H. Chile, V. H. Kondekar, R. Pushpan, S. B. Thigale, V. Rane, and A. Vibhute, "Dual parameter control of drum boiler using formula-based fuzzy controller," *Journal of Electrical Systems*, vol. 20, no. 2, pp. 2547–2557, 2024.

[272] N. M. Thoppil, V. Vasu, and C. S. P. Rao, "On the criticality analysis of computer numerical control lathe subsystems for predictive maintenance," *Arabian Journal for Science and Engineering*, vol. 45, no. 7, pp. 5259–5271, 2020.

[273] S. Xiang, Y. Lv, Y. Li, and L. Qian, "Reliability analysis of failure-dependent system based on Bayesian network and fuzzy inference model," *Electronics*, vol. 12, no. 4, pp. 1–23, 2023.

[274] A. Amanatiadis, I. Andreadis, and K. Konstantinidis, "Design and implementation of a fuzzy area-based image-scaling technique," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 8, pp. 1504–1513, 2008.

[275] M. J. Schniederjans, "Goal programming," in *Encyclopedia of Operations Research and Management Science*, pp. 337–338, Springer, 2001.

[276] Statsguru, "Statistics/Statsguru," *Statsguru*. [Online]. Available: `https://stats.espncricinfo.com/ci/engine/stats/index.html`. [Accessed: Mar. 10, 2024].

[277] ESPNcricinfo, "ESPNcricinfo," *ESPNcricinfo*. [Online]. Available: `https://www.espncricinfo.com/`. [Accessed: Mar. 10, 2024].

[278] A. H. Sajib, S. H. Limon, A. I. Naser, and G. Saha, "Analysing factors impacting Bangladesh men's T20 cricket performance," *Scientific Journal of Sport and Performance*, vol. 3, no. 1, pp. 79–94, 2024.

[279] S. R. Behera and V. V. Saradhi, "Cricket player profiling: Unraveling strengths and weaknesses using text commentary data," in *2019 18th International Conference on Machine Learning and Applications (ICMLA)*, pp. 1–30, IEEE, 2023.

[280] M. K. Manju and A. O. Philip, "Novel method for ranking batsmen in Indian Premier League," *Data Science and Management*, vol. 6, no. 3, pp. 158–173, 2023.

[281] L. M. Jayalath, "Hawk Eye technology used in cricket," *South Asian Research Journal of Engineering and Technology*, vol. 3, no. 2, pp. 55–67, 2021.

[282] Indian Premier League, "TATA IPL Auction - 2022," *Indian Premier League*. [Online]. Available: `https://www.iplt20.com/auction/2022`. [Accessed: Oct. 7, 2024].

[283] V. Sankaranarayanan, *Towards a time-lapse prediction system for cricket matches.* MSc dissertation, University of British Columbia, 2014.

[284] M. Grove and J. Pearson, "Visualisation and permutation methods for archaeological data analysis," *Archaeological and Anthropological Sciences*, vol. 6, no. 4, pp. 319–328, 2014.

[285] The Comprehensive R Archive Network, "Confidence ellipse," *CRAN*. [Online]. Available: `https://cran.r-project.org/web/packages/ConfidenceEllipse/vignettes/confidence-ellipse.html`. [Accessed: Mar. 19, 2024].

[286] J. G. Owen and M. A. Chmielewski, "On canonical variates analysis and the construction of confidence ellipses in systematic studies," *Systematic Zoology*, vol. 34, no. 3, pp. 366–374, 1985.

[287] ESPNcricinfo, "Tim David," *ESPNcricinfo*. [Online]. Available: `https://www.espncricinfo.com/cricketers/tim-david-892749`. [Accessed: Mar. 13, 2025].

[288] ESPNcricinfo, "Darshan Nalkande," *ESPNcricinfo*. [Online]. Available: `https://www.espncricinfo.com/cricketers/darshan-nalkande-1111917`. [Accessed: Mar. 13, 2025].

[289] ESPNcricinfo, "Most runs in ICC World Cup, 2002/03," *ESPNcricinfo*. [Online]. Available: `https://www.espncricinfo.com/records/tournament/batting-most-runs-career/icc-world-cup-2002-03-865`. [Accessed: Nov. 25, 2024].

[290] H. Abdi, "The Kendall rank correlation coefficient," *Encyclopedia of Measurement and Statistics*, vol. 2, pp. 508–510, 2007.

[291] Y. Shuai, L. Xiaopeng, G. Ziyuan, L. Dongsheng, C. Xiaolei, and C. Wenbin, "New principle of pilot protection based on Kendall's rank correlation coefficient," in *2021 2nd International Conference on Clean Energy and Electric Power Engineering (ICCEPE)*, pp. 1–7, IOP Publishing, 2021.

[292] S. Sazzed, "An investigation of the performances of simple gene selection methodologies for cancer classification," in *2021 IEEE 21st International Conference on Bioinformatics and Bioengineering (BIBE)*, pp. 1–6, IEEE, 2021.

[293] ESPNcricinfo, "World T20 2015/16 - Squads," *ESPNcricinfo*. [Online]. Available: `https://www.espncricinfo.com/series/world-t20-2015-16-901359/squads`. [Accessed: Mar. 30, 2025].

[294] ESPNcricinfo, "Most wickets for World T20, 2015/16," *ESPNcricinfo*. [Online]. Available: `https://www.espncricinfo.com/records/tournament/bowling-most-wickets-career/world-t20-2015-16-10291`. [Accessed: Mar. 30, 2025].

[295] ESPNcricinfo, "Most runs in World T20, 2015/16," *ESPNcricinfo*. [Online]. Available: `https://www.espncricinfo.com/records/tournament/batting-most-runs-career/world-t20-2015-16-10291`. [Accessed: Mar. 30, 2025].

[296] T. Ganesh and P. R. S. Reddy, "Testing the consistency of subjective weights in goal programming - The Analytical Hierarchy Process approach," *American Journal of Applied Mathematics and Statistics*, vol. 2, no. 3, pp. 92–95, 2014.

[297] L. P. Kolluru, R. P. Gala, P. N. Shastri, and R. V. Ubale, "Design of experiments: A valuable "Quality by Design" tool in formulation development," in *Nanoparticulate Vaccine Delivery Systems*, pp. 61–91, Jenny Stanford Publishing, 2015.

[298] S. Beg, "Mixture designs and their applications in pharmaceutical product development," *Design of Experiments for Pharmaceutical Product Development*, vol. 1, pp. 87–96, 2021.

[299] G. Squeo, D. De Angelis, R. Leardi, C. Summo, and F. Caponio, "Background, applications and issues of the experimental designs for mixture in the food sector," *Foods*, vol. 10, no. 5, pp. 1128–1158, 2021.

[300] StatEase, "DesignExpert," *StatEase*. [Online]. Available: `https://www.statease.com/software/design-expert/`. [Accessed: Nov. 20, 2025].

[301] A. C. Atkinson and A. N. Donev, *Optimum experimental designs*. Oxford University Press, 1992.

[302] ESPNcricinfo, "FAQs: All you wanted to know about the IPL 2022 auction," *ESPNcricinfo*. [Online]. Available: `https://www.espncricinfo.com/story/faqs-all-you-wanted-to-know-about-the-ipl-2022-auction-1300565`. [Accessed: Apr. 12, 2025].

[303] ESPNcricinfo, "GT Indian Premier League 2022 Squad," *ESPNcricinfo*. [Online]. Available: `https://www.espncricinfo.com/series/indian-premier-league-2022-1298423/gujarat-titans-squad-1300880/series-squads`. [Accessed: May. 18, 2025].

[304] M. C. Ledwith, B. J. Hufstetler, and M. A. Gallagher, "Stochastic preemptive goal programming to balance goal achievements under uncertainty," *Journal of Multi-Criteria Decision Analysis*, vol. 28, no. 1–2, pp. 85–98, 2021.

[305] B. Kitchenham, O. P. Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering - A systematic literature review," *Information and Software Technology*, vol. 51, no. 1, pp. 7–15, 2009.

[306] B. Kitchenham, "Systematic reviews," in *10th International Symposium on Software Metrics*, p. 12, IEEE, 2004.

[307] A. Jha, A. K. Kar, and A. Gupta, "Optimization of team selection in fantasy cricket: A hybrid approach using recursive feature elimination and genetic algorithm," *Annals of Operations Research*, vol. 325, no. 1, pp. 289–317, 2023.

[308] R. Jadhav, B. Pawar, N. Bhat, S. Kawale, and A. Gawai, "Predicting optimal cricket team using data analysis," in *2021 International Conference on Emerging Smart Computing and Informatics (ESCI)*, pp. 269–274, IEEE, 2021.

[309] M. A. Rauf, H. Ahmad, C. N. Faisal, S. Ahmad, and M. A. Habib, "Extraction of strong and weak regions of cricket batsman through text-commentary analysis," in *2020 IEEE 23rd International Multitopic Conference (INMIC)*, pp. 607–612, IEEE, 2020.

[310] M. Shetty, S. Rane, C. Pandita, and S. Salvi, "Machine learning-based selection of optimal sports team based on the players performance," in *2020 5th International Conference on Communication and Electronics Systems (ICCES)*, pp. 1267–1272, IEEE, 2020.

[311] S. Vijayabalaji and P. Balaji, "Best'11 strategy in cricket using MCDM, rough matrix and assignment model," *Journal of Intelligent & Fuzzy Systems*, vol. 39, no. 5, pp. 7431–7447, 2020.

[312] A. Adhikari, A. Majumdar, G. Gupta, and A. Bisi, "An innovative super-efficiency data envelopment analysis, semi-variance, and Shannon-entropy-based methodology for player selection: Evidence from cricket," *Annals of Operations Research*, vol. 284, no. 1, pp. 1–32, 2020.

[313] S. I. Shanto and N. Awan, "A sequential principal component-based algorithm for optimal lineup and batting order selection in One Day International cricket for Bangladesh," *International Journal of Performance Analysis in Sport*, vol. 19, no. 4, pp. 567–583, 2019.

[314] H. Perera, J. Davis, and T. B. Swartz, "Optimal lineups in Twenty20 cricket," *Journal of Statistical Computation and Simulation*, vol. 86, no. 14, pp. 2888–2900, 2016.

[315] D. Bhattacharjee and H. Saikia, "An objective approach of balanced cricket team selection using binary integer programming method," *OPSEARCH*, vol. 53, no. 2, pp. 225–247, 2016.

[316] D. Bhattacharjee and H. Saikia, "Selecting the optimum cricket team after a tournament," *Asian Journal of Exercise & Sports Science*, vol. 10, no. 2, pp. 77–91, 2013.

[317] W. J. Brettenny, D. G. Friskin, J. W. Gonsalves, and G. D. Sharp, "A multi-stage integer programming approach to fantasy team selection: A Twenty20 cricket study," *South African Journal for Research in Sport, Physical Education and Recreation*, vol. 34, no. 1, pp. 13–28, 2012.

[318] F. Ahmed, A. Jindal, and K. Deb, "Cricket team selection using evolutionary multi-objective optimization," in *Swarm, Evolutionary, and Memetic Computing SEMCCO)*, pp. 71–78, Springer, 2011.

[319] J. M. Norman and S. R. Clarke, "Optimal batting orders in cricket," *Journal of the Operational Research Society*, vol. 61, no. 6, pp. 980–986, 2010.

[320] J. M. Norman and S. R. Clarke, "Dynamic programming in cricket: Optimizing batting order for a sticky wicket," *Journal of the Operational Research Society*, vol. 58, no. 12, pp. 1678–1682, 2007.

[321] B. Bukiet and M. Ovens, "A mathematical modelling approach to one-day cricket batting orders," *Journal of Sports Science & Medicine*, vol. 5, no. 4, pp. 495–502, 2006.

[322] R. Chaudhary, S. Bhardwaj, and S. Lakra, "A DEA model for selection of Indian cricket team players," in *2019 Amity International Conference on Artificial Intelligence (AICAI)*, pp. 224–227, IEEE, 2019.

[323] I. Kumarasiri and S. Perera, "Optimal One Day International cricket team selection by genetic algorithm," *International Journal of Sciences: Basic and Applied Research*, vol. 36, no. 4, pp. 213–221, 2017.

[324] D. Das, "Moneyballer: An integer optimization framework for fantasy cricket league selection and substitution," *Debarghyadas*. [Online]. Available: `https://debarghyadas.com/files/IPLpaper.pdf`. [Accessed: Mar. 20, 2024].

[325] A. G. Kamble, R. V. Rao, A. Kale, and S. P. Samant, "Selection of cricket players using analytical hierarchy process," *International Journal of Sports Science and Engineering*, vol. 5, no. 4, pp. 207–212, 2011.

[326] S. N. Omkar and R. Verma, "Cricket team selection using genetic algorithm," in *International Congress on Sports Dynamics*, pp. 1–9, Citeseer, 2003.

Appendix A

Systematic literature review

Appendix A provides the full-text summaries of the studies identified by Alberts *et al.* [24] in the systematic literature review (SLR). Alberts *et al.* [24] followed an adapted protocol from Kitchenham *et al.* [305] and Kitchenham [306]. Table A.1 lists the summaries of the 26 research results obtained from the quality assessment phase of the protocol. The authors of these studies referenced related studies, which were identified through a manual analysis. Alberts *et al.* [24] defined the manual analysis as a procedure that involves inspecting the relevance of cited work by evaluating titles and abstracts. Table A.2 details the summaries of the 10 studies resulting from the manual analysis. The numbering convention in Table A.2 resumes from Table A.1 to ensure consistency with the work of Alberts *et al.* [24].

Table A.1: Full-text summaries of quality assessment results

Number	Year	Title	Full-text summary	Reference
1	2023	Optimization of team selection in fantasy cricket: A hybrid approach using recursive feature elimination and genetic algorithm.	This study follows a unique method, considering both player performance and team selection in a fantasy environment. This study uses random forest recursive feature elimination to recognise and and order player performance indicators. Mean decrease accuracy is applied to weigh the relevant indicators, subsequently generating batting, bowling, and all-rounder indices for fitness evaluation. The virtual fantasy team selection is performed by applying a prominent genetic algorithm using NSGA-II. Finally, the algorithmic team is tested against the first-ranked team in multiple playing contests associated with the 2019 One Day International (ODI) World Cup. The team selected by the genetic algorithm repeatedly outperformed the first-ranked squad, with a selection equivalent to the tournament Dream Team. This study is significant, as it introduces an algorithm for team selection that integrates multiple factors.	Jha *et al.* [307].
2	2022	A balanced squad for Indian Premier League using modified NSGA-II.	This study treats team selection for an Indian Premier League (IPL) squad as a complex knapsack problem. This paper proposes two algorithms, BNSGA-II and INSGA-II, variants of NSGA-II, to solve the problem. This study incorporates specialist and celebrity players with exceptional brand value, distinguishing it from similar work. Performance is estimated using a batting performance factor and the combined bowling rate outlined in alternative literature. The appropriateness of the method was validated using data from the 2020 IPL season, confirming a sufficiently balanced side. In addition, BNSGA-II proved to be a superior algorithm following a comparative study involving existing team selection NSGA-II approaches and other prominent algorithms, such as RNSGA-II, NSDE, and MOPSO-CD. The contribution of this study lies in its proposal of a new approach for balanced team selection, which is solved using BNSGA-II.	Verma *et al.* [93].
3	2021	Predicting optimal cricket team using data analysis.	An automated system enables data processing and decision-making based on predefined conditions, thereby eliminating discrepancies introduced by manual errors. This study provides a system that synthesises vast amounts of data to produce a Twenty20 (T20) team that is the most suitable for a scenario. This study reveals approximately 35 distinct player performance parameters required for effective team selection, weighted based on their importance. These performance measures include playing venue, opposition, batting statistics, bowling statistics, and team arrangement. This study examines the performance and behaviour of the system using player statistics and team data from 2008 to 2018, thereby revealing the impact of various conditions on team selection. The automated system introduces desirable benefits for selectors, as the proposed performance metrics significantly reduce the difficulties associated with screening cricketers for selection.	Jadhav *et al.* [308].

4	2020	Extraction of strong and weak regions of cricket batsman through text-commentary analysis.	A text-based mining approach is provided for extracting information indicative of strong and weak regions of batsmen. Therefore, this study employs a regular expression matching procedure on T20 cricket text commentary, with each line parsed into tokens based on standard cricket terminology. Thus, deliveries are mapped to pitch lengths, allowing the model to determine the corresponding runs scored using a particular fielding region. This paper proposes a unique region-wise strike rate calculation, which has not been previously performed. The data considered in the study comprises all international T20 players between 2008 and 2019, after removing end-of-over text and discussion commentary. Accordingly, this study shows the algorithmic performance by highlighting the strong and vulnerable playing areas for the 2019 top ten international batsmen. This study presents the region-wise strike rate of each batsman, with strong and weak zones designated by the algorithm.	Rauf *et al.* [309].
5	2020	Machine learning-based selection of optimal sports team based on the players performance.	A machine learning algorithm is proposed for team selection that takes into account various factors, including match conditions, historical contests, applicable grounds, pitch scenarios, and competing teams. These factors result in a model with higher prediction accuracy, distinguishing it from prior research. This study utilises data from ODI games played by India, including comprehensive player statistics, competitors, and venues. This study examines the performance of numerous procedures, namely logistic regression, a support vector machine classifier, a decision tree, and a random forest, with the latter proving to be the most accurate. Applying the test data, the random forest method ultimately gave prediction accuracies of 76%, 67%, and 95% for batsmen, bowlers, and all-rounders, respectively. This study allocates an equal importance to all-rounders, a shortcoming identified as a limitation of the current literature.	Shetty *et al.* [310].
6	2020	Best'11 strategy in cricket using MCDM, rough matrix and assignment model.	This study presents Best'11, a unique methodology for optimising cricket teams and batting orders. The approach employs multiple-criteria decision-making, rough matrices, and integer programming through an assignment model. First, the research applies rough matrix theory, using an estimated ranking parameter, to 32 players for team selection involving T20, ODI, and test formats. Second, this paper uses an assignment model to optimise the batting order of the resulting test team. This study analyses the performance of the model by comparing the predicted team to the actual Indian 2019 ODI World Cup squad. The model considered batting and bowling indices of players, with the statistics weighted according to their importance. The model yielded a 93% match to the empirical Indian squad selected for the 2019 ODI World Cup, emphasising the significance of the approach.	Vijayabalaji and Balaji [311].

7	2020	An innovative super-efficiency data envelopment analysis, semi-variance, and Shannon-entropy-based methodology for player selection: Evidence from cricket.	This study presents a distinctive selection method that considers player efficiency, consistency, and significance in team strategy. This study employs an adjusted data envelopment analysis approach to measure the efficiency of players. Next, this study employs a semi-variance procedure to develop consistency indices, which are then clustered into an individual index using Shannon entropy. This study introduces a value index that describes player importance and overall fitness. The efficiency, consistency, and value indices are used to obtain a performance index for each player, representing their overall cricketing ability. The performance of the developed model is evaluated by deriving an all-star ODI team comprising players from 1971 to 2015. Using a holistic comparative analysis, the proposed team is superior to the all-time sides compiled by the International Cricket Council (ICC) in 2011 and the British Broadcasting Corporation (BBC) in 2015. Finally, this study demonstrates the flexibility of the model by reflecting decision-making processes involving multiple scenarios and team combinations, which are necessary to accommodate diverse pitches and conditions.	Adhikari *et al.* [312].
8	2019	Decision making in cricket: The optimum team selection.	This study introduces numerous approaches for cricket team selection, building on previous work conducted by the same authors. This study presents a preliminary approach that normalises the performance scores of the cricketing abilities. These values are subsequently multiplied by weights and totalled to determine the performance value concerning each cricketer. Thereafter, the values are converted into associated indices, used to rank players based on their specialities. Accordingly, using data from the 2012 Bangladesh Premier League (BPL), the study selects a 15-player side, of which 12 selections proved to be similar to the Bangladesh team. The second method uses a binary integer programming model to select players, considering the cricketing ability of the captain. The second method differs from similar models proposed in previous research in that it excludes fielding as a cricketing ability and treats specific performance metrics differently. This study validates the model by selecting a fantasy team for the 2018 IPL, led by the winning captain of the tournament.	Saikia *et al.* [18].
9	2019	A sequential principal component-based algorithm for optimal lineup and batting order selection in one day international cricket for Bangladesh.	This study presents a statistical approach to player ordering, introducing a sequential principal component algorithm that incorporates both batting and bowling performance. As such, this paper employs principal component analysis to identify the most suitable player for a particular position, considering only the Bangladeshi ODI contracted players from 2015 to 2017. The proposed method seeks to determine the optimal batting order by utilising consecutive principal components for each batting position. Notably, the algorithm selects the only wicketkeeper available due to contractual limitations imposed by the Bangladesh Cricket Board. A principal component analysis applied to bowling allowed for the selection of various bowlers. Again, this study utilised principal component analysis to determine the optimal batting positions for the selected cricketers. Similarly, this study utilises repetitive principal component analysis to obtain the specific batting positions of top- and middle-order batsmen based on the selected bowling combination.	Shanto and Awan [313].

10	2018	Team selection using multi-/many-objective optimization with integer linear programming.	This study presents a distinct multi-objective integer programming model for selecting a collection of optimum teams. The proposed model minimises five objectives, including the combined bowling average, batting average, cost, player star power, and fielding average. First, this study illustrates the scalability of the approach by demonstrating its behaviour in dissimilar scenarios. Two, three, and five objective formulations are used to illustrate the flexibility of the model. Second, this paper illustrates the use of the model in the context of creating a partial team with predefined members. Third, this study presents insight into player value by utilising a hypervolume measure to sequence players within a team, considering their performance in relation to their initial cost. Data from 129 players in the 2011 IPL are used to test the performance of the model, with comparisons drawn to the champion squad. The results revealed captivating details, including the contradiction between different playing objectives.	Chand *et al.* [19].
11	2016	A new model for player selection in cricket.	This paper claims that existing team selection models offer limited flexibility for subjective factors, such as playing conditions. This study provides a performance measure that considers the efficiency of players. Subsequently, a composite index is formulated, to specify the performance value of each player. The model determines the distribution of the performance values given dissimilar expertise, ultimately providing a subset of available players. This study divides the index into three intervals, labelling players as poor performers, average performers, and performers. The 2012 IPL performers, as selected by the approach, are compared to the players selected for the Indian team following the tournament. The comparison revealed that the Indian selection board included five non-performing players. Thus, the proposed approach is significant as it simplifies team selection by reducing the sampling pool to only the top performers.	Saikia *et al.* [54].
12	2016	Optimal lineups in Twenty20 cricket.	This study employs relative value statistics to optimise a Twenty20 cricket team based on batting and bowling lineups. This study characterises the objective function as the maximisation of the run differential, a number representing the difference between the runs scored and conceded. Simulated annealing is applied to optimise the lineup such that the greatest run differential results. Following numerous iterations, the algorithm ultimately identifies the optimal lineup with the most significant run differential. The algorithm is tested by comparing the proposed optimal lineup to the practical lineups used by India and South Africa during multiple Twenty20 World Cup tournaments. While the algorithm presents player selections similar to the actual squads, it specifies essential variations among the lineups that challenge traditional reasoning. This study is paramount as it introduces a quantitative basis for selecting lineups that diverge from conventional methods.	Perera *et al.* [314].

13	2016	An objective approach of balanced cricket team selection using binary integer programming method.	This study provides a binary integer programming model for team selection that treats the captain as a distinct player. This study investigates how the specialities of the captain impact the selection of the remaining team members. First, this study draws on the literature to develop a performance measurement procedure that defines the quality of players. Second, this paper presents an integer programming approach that employs the aforementioned indices for empirical team selection. The objective function of the proposed model concerns maximising team performance by applying the computed player indices. This study validates the model using data from the 2012 IPL to select the optimal team for each franchise, given a predetermined set of criteria. The validation demonstrates that the two franchises that reached the final contain the smallest difference in team strength when comparing the proposed optimal side to the actual side. Using the captain of the winning franchise, the study also computes an all-star team for the tournament, which has a total strength of 6.66.	Bhattacharjee and Saikia [315].
14	2014	On performance measurement of cricketers and selecting an optimum balanced team.	This study introduces a composite index that quantifies player performance, enabling comparisons across multiple cricketing abilities. This paper proposes a binary integer program that identifies a squad of 15 cricketers from a finite set of available players. The players available for selection are identified using a particular criterion applied to IPL tournaments. The model considers specialities, including opening batsmen, middle-order batsmen, all-rounders, seam bowlers, spin bowlers, and wicketkeepers, as the constraints governing performance maximisation. The team selected by the model is compared to the Indian squad selected for the Twenty20 World Cups following the 2009, 2010, and 2012 IPL seasons. The team selected by the model recognised the suboptimal selection of the 2009 and 2010 World Cup teams, with score differences of 3.37 and 3.87, respectively. The difference in performance score between the optimal team and the actual Indian team proved to be the smallest for the 2012 World Cup, with a value of 1.62 to optimal.	Bhattacharjee and Saikia [17].
15	2013	Team selection after a short cricket series.	This study utilises binary integer programming to select a cricket team after a short series. This paper proposes suitable performance measures for player evaluation that consider the conditions under which the cricketers performed. To maximise performance, the model requires comparable scaling. The performance measures are converted into indices, with different specialities used as constraints on the solution to maximise performance. The study demonstrates the unreliability of traditional performance measures for team selection by employing the ICC Champions Trophy data from the 2009 ODI series. The performance measures produced a significantly different player ranking compared to other statistics that excluded the performance conditions. The circumstances considered in this study concern the conditions under which a player scores runs and the pedigree of the batsman dismissed by a bowler. The contribution of this study is significant, as it provides performance measures that take into account playing conditions.	Lemmer [9].

16	2013	Multi-objective optimization and decision making approaches to cricket team selection.	This study proposes a unique approach to team selection, utilising a genetic algorithm supported by a complex decision-making approach. The study uses a genetic algorithm, specifically NSGA-II, to identify a trade-off frontier with various solutions to the team selection problem. The trade-off frontier is used to construct the decision-making criteria employed for final selection, incorporating factors such as captain success rate and wicketkeeping performance. The objective of the resultant model is to maximise batting, bowling, and fielding performance, with the latter characterised as a secondary objective that decides between equivalent solutions. NSGA-II gave 35 high-performing teams from 129 players, with each team ultimately compared to the winning side of the 2011 IPL. The comparison yielded intriguing results, showing that multiple algorithmic teams generated by NSGA-II proved superior to the winning team. Therefore, the contribution of this study is evident, as it provides selector boards with a variety of team choices that support diverse strategies.	Ahmed *et al.* [6].
17	2013	Selecting the optimum cricket team after a tournament.	A dream team refers to a fictitious team that comprises the best-performing players for a selected tournament. This study proposes a binary integer programming model that facilitates the selection of a dream team based on player performance scores. This paper uses data from the 2012 IPL to derive performance scores that consider bowling, batting, and wicketkeeping abilities. The scores are then used to maximise fantasy team performance, subject to competition requirements. The provisions required a minimum of one player per franchise, role-specific selection, and a maximum of four foreigners per side. Subject to these constraints, the optimal team produced a performance value of 16.21, which is 0.53 better than the team submitted by the best public participant. This study demonstrates the distinction between practical and fantasy selection, revealing disparities in certain factors, such as opposition strength and pitch conditions.	Bhattacharjee and Saikia [316].
18	2012	A multi-stage integer programming approach to fantasy team selection: A Twenty20 cricket study.	This study provides a multi-stage integer programming model for fantasy team selection in the 2008 IPL. This paper considers a typical fantasy tournament, which is divided into four stages, with each phase allowing three changes to the existing team. The model allows for team selection across four stages, maximising the total points scored based on player performance scores. The model identifies the optimum team for each tournament phase by using players with the highest performance scores. Except for the initial team, which uses career statistics, the model selects a side for a specific stage, considering only the data available from the prior phase. The model is evaluated by comparing the final algorithmic output to simulated squads with both the same and different initial lineups. The comparison placed the score of the optimal team within 0.5% and 0.02% of the simulated results for the scenarios. This study is significant as it demonstrates the relevance of optimisation techniques across multiple stages, using only the data available from the prior phase.	Brettenny *et al.* [317].

19	2011	Cricket team selection using evolutionary multi-objective optimization.	This study uses a genetic algorithm, specifically NSGA-II, to optimise the composition of a cricket team. This study utilises data from the 2011 IPL, with a specific focus on bowling and batting averages as measures of player performance. The approach comprises two objectives, involving the maximisation of batting and bowling performance, subject to squad and tournament constraints. Subsequently, the trade-off solution reveals a team with better bowling and batting performance compared to the winning franchise of the 2011 IPL. A sensitivity analysis conducted on the budget constraint indicates that capital limitations have a greater influence on batting-based teams than on bowling-based sides. Following a knee-point approach, this study presents an optimal team that strikes a satisfactory compromise between bowling and batting, considering fielding, wicketkeeping, and captaincy as key decision-making criteria. This study gives a systematic analysis that enables selectors to derive an optimal side using self-selected criteria based on tactical preferences.	Ahmed *et al.* [318].
20	2011	Integer optimisation for the selection of a Twenty20 cricket team.	This study proposes a binary integer programming model for team selection in T20 cricket. This paper utilises indices to obtain coefficients that define player performance across different cricketing abilities. These indices are employed to find an optimal team using the proposed model. This study outlines a fixed set of requirements to reduce the likelihood of selecting an outlier, i.e., a player who appears to excel in a given role but is actually not a specialist. The model is tested using data from the 2007 T20 World Cup. The resulting team is compared to the ODI World XI, which was gained from another source in the literature. The comparison revealed few similarities between the teams, emphasising the impact of cricketing formats on strategy and selection.	Sharp *et al.* [106].
21	2010	Optimal batting orders in cricket.	This study utilises dynamic programming to optimise batting orders for multiple cricketing formats. Two scenarios are assessed, revealing the influence of modified batting orders on both limited-overs and unlimited-overs cricket. The sample data reveals that sequencing players in a decreasing order by batting average is advantageous for a large number of available overs. In contrast, for 40 overs or less, the study suggests that teams order their batting lineups based on the run rate. This study demonstrates that altering the batting lineup, especially in limited-overs cricket, depends mainly on the type of cricketer that is dismissed, as well as the calibre of the player. This study demonstrates how modifying the batting order affects the expected runs, using examples from both limited-overs and unlimited-overs cricket. This study is significant as it provides evidence of the advantages of varying batting orders based on game circumstances rather than using predefined lineups.	Norman and Clarke [319].

22	2009	Applying genetic algorithm to select an optimal cricket team.	This study utilises a genetic algorithm to produce a cricket team using a fitness function to achieve optimality. The fitness of a team depends on various factors, including individual player talent, the diversity of batting styles, the ability to run between wickets, the variety of bowlers, partnerships, and squad composition. First, the algorithm obtains an initial population of 50 teams by repeatedly selecting 11 arbitrary players from a finite set of 50 Indian cricketers. Second, the method computes the fitness value for each entity, utilising crossover and mutation to generate a new population. The process is repeated iteratively, obtaining higher fitness values with each cycle. The model exhibits a persistent increase in the fitness function value. The teams selected by the algorithm share various similarities with the 2009 Indian ODI team, with a few alternative players chosen to enhance squad balance and performance.	Sathya and Jamal [26].
23	2007	Dynamic programming in cricket: Optimizing batting order for a sticky wicket.	This study presents a stochastic dynamic programming model to evaluate the applicability of using lower-order batsmen in challenging playing conditions. The formulated model is assessed using two test cricket scenarios, where the objective function involves maximising runs. In the first scenario, it is assumed that the difficult conditions finish before the end of play, meaning that batsmen need not play themselves in multiple times. In contrast, the second case assumes that when play resumes in normal circumstances, the batsmen need to reset and play themselves in again. These scenarios demonstrate that using lower-order batsmen is often more effective when the challenging conditions persist until the end of play. However, if the challenging conditions are temporary, it is beneficial to use a specialist batsman instead. This study reveals the value of alternating batting lineups based on playing conditions and game scenarios.	Norman and Clarke [320].
24	2006	A mathematical modelling approach to one-day cricket batting orders.	This paper offers a Markov chain approach for analysing the runs accumulated by different batting orders. This study exhibits the application of the model in predicting the anticipated performance of a specific sequence of cricketers. Both straightforward and streamlined procedures are considered for studying the distribution and expected number of runs scored given dissimilar batting orders. However, the straightforward process proves challenging due to the computational time required, making the streamlined method desirable because it exploits smaller subsets. This study employs practical data from existing Australian players to estimate the probability of experiencing a specific state, describing the different outcomes associated with a delivery. As such, the sum of the product between the likelihood of occurrence and the related runs for a particular state reveals the expected score for a proposed batting order. This study contributes by illustrating optimal player sequencing based solely on current playing conditions using Markov chains.	Ovens and Bukiet [321].

25	2006	Optimal batting orders in one-day cricket.	This study introduces optimal batting orders within the context of the ODI format using simulated annealing. The simulation estimates the mean number of runs scored by a particular lineup as the objective function. This study employs a Bayesian log-linear method to estimate the characteristics of different Indian players. The model gave estimates for the attributes by averaging the results of the Markov chain. Using the estimates, this study reviews the simulation approach by assessing 71 matches played by India between 1998 and 2003. This paper compares the actual runs accumulated across these games to the simulated runs. In addition, this study shows a bar plot for the 71 games, illustrating the agreement between both the simulated and practical dismissals. Finally, this study provides two simulated batting lineups, both comprising an estimated mean number of runs greater than the average of the order used by the Indian team for the 2003 World Cup final. This study opposes traditional wisdom by recommending unprecedented batting lineups.	Swartz *et al.* [57].
26	2003	Dynamic programming in cricket: Choosing a night watchman.	This study provides a dynamic programming approach for decision-making involving the likely use of lower-order batsmen as an alternative to specialist batsmen. This study presents two models that challenge traditional beliefs by suggesting recommendations for using cricketers as diverse resources. The first model assumes players are only dismissed at the end of an over. The second model allows for player dismissal either halfway through or at the end of an over. Using these models, this study examines the night watchman as a defensive and an attacking metric by independently considering stonewalling and slogging strategies. This study suggests that a night watchman is advantageous when limited overs remain, particularly when the team still has various specialist and non-specialist batsmen. In contrast, if few batsmen remain, this study suggests using the next recognised batsman in the lineup rather than a night watchman.	Clarke and Norman [25].

Table A.2: Full-text summaries of manual analysis results

Number	Year	Title	Full-text summary	Reference
27	2019	A DEA model for selection of Indian cricket team players.	This study utilises a unique data envelopment analysis (DEA) approach to select cricket teams and compute player efficiencies. The method involves a ratio of the sum of weighted outputs and inputs, providing a measure of the proximate efficiency of each player. Historical performances are assessed to determine a DEA score for each player. The DEA scores are used to identify an Indian test team that consists of various specialist bowlers and batsmen. The selected side is compared to the actual Indian side chosen for the 2018 series against England. The comparison revealed various similarities among the squads. This study extends the application of DEA beyond the T20 format by assessing team selection in test cricket.	Chaudhary *et al.* [322].
28	2017	Optimal one day international cricket team selection by genetic algorithm.	This study investigates player selection using a genetic algorithm, which converges iteratively to an optimal solution. This paper assesses 30 Sri Lankan players categorised as batsmen, wicketkeepers, spin bowlers, pace bowlers, and all-rounders. Career and Australian-based statistics are applied to derive a fitness function that incorporates the mean batting average, bowling average, strike rate, number of wickets, prior experience, and win-loss ratio. This paper identifies the applicable players, enabling the algorithm to produce 100 teams, of which 20 denote the initial population. The fitness values for each team are determined, with the process repeated until the margin between the best and worst cricketers is smaller than five. After 350 iterations, the algorithm proposed a Sri Lankan side for the 2015 ODI World Cup, with a combined fitness value of 133.76. This study is unique in that it identifies player performance in relation to specific countries, a key factor in team selection.	Kumarasiri and Perera [323].
29	2014	Moneyballer: An integer optimization framework for fantasy cricket league selection and substitution.	This study introduces a binary integer programming model for optimal team selection in fantasy cricket. This work presents a unique financial perspective on model development and compares its utility to the monetary Markowitz and Black-Litterman models. This study examines both static and dynamic forecasting, with practical testing conducted using the 2014 IPL, which featured 178 cricketers and 37 playing days. The model includes platform, match, and team-based constraints. The model selected a side that ranks within the top 97.23% of an actual fantasy league tournament with 431,984 participants. However, to eliminate the shortcomings of practical testing, this study considered a live backtest, in which the team placed within the top 99.53% of participants. This study claims to be the first to employ optimisation techniques in a practical fantasy tournament, automate data analysis, and propose a backtesting framework.	Das [324].

30	2014	Cricket team selection using data envelopment analysis.	This study presents a DEA model to evaluate player efficiency in various cricketing abilities. The procedure simultaneously considers numerous measures used to quantify performance. This paper utilises four and five metrics to measure bowling and batting performance, respectively. Moreover, nine performance metrics are used for all-rounders, whereas five batting-based metrics describe the wicketkeeping abilities of cricketers. This paper utilises data from the 2011 IPL to calculate DEA scores for each player across all cricketing abilities. The DEA scores are used to rank the cricketers, showing their corresponding efficiency. This study utilises the ranking to select a 15-player squad comprising five bowlers, six batsmen, two all-rounders, and two wicketkeepers.	Amin and Sharma [27].
31	2011	Selection of cricket players using analytical hierarchy process.	This study employs an analytical hierarchy process (AHP) that utilises historical performance data to rank and select players. AHP is a decision-making technique that solves complex multi-attribute problems by drawing pairwise comparisons between numerous entities. This study considers six Indian players, identifying batsmen, bowlers, all-rounders, and wicketkeepers as the four primary model attributes. Matrix theory is used to determine the relative importance of each attribute in relation to the ultimate objective of balanced team selection. A pairwise comparison is conducted among the six selected players based on the degree to which they meet each attribute. The player index for each entity is computed, describing the overall performance of each cricketer. The indices are used to rank the players, ultimately revealing the best cricketers for selection.	Kamble *et al.* [325].
32	2010	Integer optimisation for the selection of a fantasy league cricket team.	This study proposes a binary integer programming model that enables the selection of fantasy sides for a T20 tournament. This study evaluates the suitability of existing measures for capturing player performance. The analysis indicates that the most descriptive measures consider both the fantasy score and the recent form of cricketers. Considering the 2008 IPL, this study employs a prospective approach, utilising the proposed performance measures to construct a fantasy side for each stage of the tournament. The model selects an optimal squad across several tournament stages by enforcing team selection changes that maximise the total fantasy score. The model selects the initial fantasy squad using statistics from the 2008 IPL, with the team subsequently modified at each stage using tournament-specific data. This study achieved 24,621 fantasy points, indicating that the optimal team formulation consists of four bowlers, four batsmen, two all-rounders, and one wicketkeeper.	Brettenny [75].

33	2009	Integer optimization for the selection of a Twenty20 cricket team.	This study advances prior research by selecting a South African T20 squad using a binary integer programming model. Specifically, this study utilises the work conducted by Gerber and Sharp [87] and Lemmer [80], [97]. This study scales different performance measures, allowing for comparability among values using indices that represent player performance. Subsequently, this study presents a model that maximises the sum of the index coefficients for the selected players. The resulting team comprises 15 cricketers, including six bowlers, six batsmen, two all-rounders, and one wicketkeeper. This study examines the similarities and differences between the selected team and those presented in the mentioned studies. The chosen team also proves stronger than the actual 2007 South African World Cup squad.	Lourens [15].
34	2006	Selecting a limited overs cricket squad using an integer programming model.	This study proposes a binary integer programming model that determines the optimal combination of cricketers. The performance coefficients for 32 South African players are calculated using indices that represent their capabilities in bowling, batting, fielding, all-rounding, and wicketkeeping. This study uses international data for those players who have played at least 15 games for the Proteas, and domestic data for all other players. The model maximises the sum of the indices for all players considered, with constraints added based on tactical decision-making. Using the indices, the model determines an optimal ODI team comprising 15 players, with mandatory, optional, and flexibility constraints applied. The model contains generalised constraints, enabling teams to adjust player provisions based on their strategic incentives.	Gerber and Sharp [87].
35	2004	A criterion for comparing and selecting batsmen in limited overs cricket.	This study introduces a two-dimensional graphical representation for examining player performance within the risk-return framework commonly utilised in portfolio analysis. The representation places the strike rate on the vertical axis and the dismissal probability on the horizontal axis. The graphic also portrays the batting average, which is defined as the quotient between the strike rate and the dismissal probability. This study combines batting average and strike rate to construct a criterion for player ranking, predominantly considering one-day and test formats. This criterion is then used to rank the top 20 batsmen for the 2003 World Cup. The proposed representation denotes a unique approach for gaining insight into player performance.	Barr and Kantor [85].

36	2003	Cricket team selection using genetic algorithm.	This study presents a genetic algorithm that considers player fitness in cricket team selection. The algorithm generates the initial population, followed by an evaluation of the chromosome fitness, where each chromosome represents an entity within the population. Next, the algorithm uses the one-cut point method to perform crossover, replacing the least fit individuals. The mutation stage is then performed using a probability of 5.00%, followed by a reevaluation of the fitness. This study evaluates the performance of the algorithm using two scenarios related to tournament and national team selection. The tournament selection yielded 15 players from a set of 100, of which 13 players matched the actual cricketers playing for India. Similarly, for the national selection, the algorithm only selected two members that did not correspond with the 2003 Indian World Cup side.	Omkar and Verma [326].

Appendix B

Personal and performance data

Appendix B provides the personal and performance data for the 173 cricketers considered for selection. Although all performance data outlined in Figure 5.1 were collected, only the relevant statistics are displayed in Tables B.1 through B.4. Table B.1 specifies the bowling dataset, recognising Sai Kishore as the best bowler. Table B.2 details the batting dataset, where Conway is identified as the best batsman. Table B.3 defines the all-rounding dataset, with secondary Tables B.3(a), B.3(b), and B.3(c) describing the specific bowling, batting, and characteristic data of these players, respectively. Tables B.3(a) and B.3(b) reveal Taide and Yadav (L) as the best bowling and batting all-rounders, respectively. Moreover, Table B.3(c) identifies David as the best overall all-rounder. Table B.4 defines the wicketkeeping dataset, which recognises Rawat as the best wicketkeeper. As discussed in Section 5.2, the datasets in Tables B.1 to B.4 were combined into a single, universal set for team selection. Table B.5 defines the verification subset, which encompasses the 40 players used to assess the correctness of the formulation and coding of both models.

Table B.1: Dataset 1 relating to bowlers [276], [282]

Player number	Player name	Franchise	Nationality	Auctioned price (₹ crore)	Innings bowled	Total wickets	Runs conceded	Balls bowled	*CBR* score	Membership value
1	D.L. Chahar	CSK	Indian	14.00	115	131	3133	2467	13.265458	0.127294
8	A.F. Milne	CSK	Overseas	1.90	133	151	3645	2854	13.343224	0.114314
12	M.M. Theekshana	CSK	Overseas	0.70	43	46	949	948	11.385790	0.745516
16	K.M. Asif	CSK	Indian	0.20	20	21	505	373	13.575475	0.081572
17	T.U. Deshpande	CSK	Indian	0.20	31	44	955	658	13.170788	0.144565
18	S. Singh	CSK	Indian	0.20	20	24	488	393	12.271427	0.395560
20	M. Choudhary	CSK	Indian	0.20	12	16	367	259	13.453329	0.097710
22	S.N. Thakur	DC	Indian	10.75	122	145	3661	2563	14.093683	0.035183
25	K.K. Ahmed	DC	Indian	5.25	69	87	2082	1537	13.548345	0.084958
26	C. Sakariya	DC	Indian	4.20	37	51	1054	836	12.417902	0.344227
28	M. Rahman	DC	Overseas	2.00	176	238	4678	3856	12.001004	0.498417
29	K. Yadav	DC	Indian	2.00	98	122	2653	2109	12.694409	0.257871
34	L.T. Ngidi	DC	Overseas	0.50	83	117	2207	1663	12.050459	0.478970
42	L.H. Ferguson	GT	Overseas	10.00	95	112	2520	2004	12.882594	0.207702
44	M. Shami	GT	Indian	6.25	117	136	3451	2487	14.005384	0.040954
45	Y. Dayal	GT	Indian	3.20	15	15	332	276	12.600660	0.285506
47	R. Sai Kishore	GT	Indian	3.00	38	43	748	821	10.246110	1.000000
50	A.S. Joseph	GT	Overseas	2.40	36	41	1132	761	14.841262	0.008444
57	V.R. Aaron	GT	Indian	0.50	85	87	2531	1769	14.996778	0.006079
58	N. Ahmad	GT	Overseas	0.30	33	33	898	745	13.679026	0.069627
67	U.T. Yadav	KKR	Indian	2.00	149	150	4323	3134	14.750215	0.010184
69	T.G. Southee	KKR	Overseas	1.50	212	245	6395	4663	14.125644	0.033273
78	R.S. Dar	KKR	Indian	0.20	6	4	211	136	19.256188	0.000000

83	A. Khan	LSG	Indian	10.00	48	65	1385	1068	12.695044	0.257689
86	M.A. Wood	LSG	Overseas	7.50	40	46	1168	849	13.972191	0.043320
90	P.V. Chameera	LSG	Overseas	2.00	80	82	2188	1670	14.032595	0.039096
93	A.S. Rajpoot	LSG	Indian	0.50	87	105	2263	1751	12.748439	0.242737
94	S. Nadeem	LSG	Indian	0.50	138	114	3335	2937	13.650347	0.072784
97	M. Khan	LSG	Indian	0.20	26	33	638	540	11.815625	0.572939
107	M. Ashwin	MI	Indian	1.60	85	82	2111	1749	13.403969	0.104904
108	T.S. Mills	MI	Overseas	1.50	150	176	4181	3207	13.344283	0.114145
109	J.D. Unadkat	MI	Indian	1.30	157	195	4476	3379	13.210374	0.137143
110	R.P. Meredith	MI	Overseas	1.00	55	71	1659	1187	13.522037	0.088349
112	M. Markande	MI	Indian	0.65	43	43	1076	877	13.343045	0.114343
114	B. Thampi	MI	Indian	0.30	68	69	1997	1428	14.847816	0.008330
123	K. Rabada	PBKS	Overseas	9.25	133	180	3914	2977	12.863420	0.212487
128	R.D. Chahar	PBKS	Indian	5.25	70	84	1898	1549	12.792849	0.230738
131	V.G. Arora	PBKS	Indian	2.00	12	12	302	260	13.078916	0.162931
132	N. Ellis	PBKS	Overseas	0.75	44	50	1293	962	13.976511	0.043006
135	S. Sharma	PBKS	Indian	0.50	146	161	3994	3201	13.381948	0.108243
138	I.C. Porel	PBKS	Indian	0.25	22	30	537	480	11.221593	0.806419
143	B.S. Dhanda	PBKS	Indian	0.20	22	26	637	489	13.517804	0.088904
145	M.P. Krishna	RR	Indian	10.00	54	48	1672	1157	16.170268	0.000358
147	T.A. Boult	RR	Overseas	8.00	149	176	4462	3276	13.918859	0.047363
149	Y.S. Chahal	RR	Indian	6.50	219	243	5956	4720	13.371252	0.109895
152	N.A. Saini	RR	Indian	2.60	64	53	1711	1398	14.629507	0.012980
153	N.M. Coulter-Nile	RR	Overseas	2.00	142	170	3981	3087	13.215197	0.136258
158	O.C. Mccoy	RR	Overseas	0.75	34	45	936	686	12.721266	0.250275

159	K.C. Cariappa	RR	Indian	0.30	38	37	946	840	12.978517	0.184872
160	K.R. Sen	RR	Indian	0.20	17	12	495	362	16.733870	0.000074
168	J.R. Hazlewood	RCB	Overseas	7.75	65	82	1820	1473	12.732703	0.247084
177	J.P. Behrendorff	RCB	Overseas	0.75	92	106	2380	1956	12.727783	0.248454
178	S. Kaul	RCB	Indian	0.75	124	147	3389	2674	13.051801	0.168662
179	K.V. Sharma	RCB	Indian	0.50	134	121	3144	2525	13.620289	0.076215
181	C.V. Milind	RCB	Indian	0.25	52	83	1427	1123	11.396700	0.741317
182	A. Deep	RCB	Indian	0.20	21	26	482	450	11.222448	0.806115
190	B. Kumar	SRH	Indian	4.20	203	207	5338	4470	13.353947	0.112607
192	T. Natarajan	SRH	Indian	4.00	49	50	1354	1032	14.123745	0.033384
193	K. Tyagi	SRH	Indian	4.00	14	13	491	313	17.216062	0.000017
195	S.A. Abbott	SRH	Overseas	2.40	106	134	2886	2039	13.049044	0.169253
197	S. Gopal	SRH	Indian	0.75	79	91	1901	1515	12.460522	0.329981
199	F. Farooqi	SRH	Overseas	0.50	5	3	135	99	17.167630	0.000020
201	J. Suchith	SRH	Indian	0.20	66	57	1609	1320	13.931558	0.046373

Table B.2: Dataset 2 relating to batsmen [276], [282]

Player number	Player name	Franchise	Nationality	Auctioned price (₹ crore)	Innings batted	Total dismissals	Runs scored	Balls faced	*CBM* score	Membership value
6	R.V. Uthappa	CSK	Indian	2.00	271	247	7042	5293	46.957623	0.166679
11	D.P. Conway	CSK	Overseas	1.00	110	86	3765	2938	65.262611	1.000000
14	C. Nishaanth	CSK	Indian	0.20	26	22	593	475	44.336449	0.096179
19	S.P. Senapati	CSK	Indian	0.20	24	22	637	521	46.821022	0.162265
24	D.A. Warner	DC	Overseas	6.25	312	273	10308	7317	59.554554	0.840112
27	R. Powell	DC	Overseas	2.80	95	77	1771	1339	39.185751	0.026356
32	M. Singh	DC	Indian	1.10	160	126	3594	2800	46.675325	0.157651
38	K.A. Hebbar	DC	Indian	0.20	39	32	1117	893	54.580992	0.543300
46	D.A. Miller	GT	Overseas	3.00	323	221	7718	5595	55.735692	0.615501
51	J.J. Roy	GT	Overseas	2.00	262	248	7088	4949	47.652022	0.190457
62	S.S. Iyer	KKR	Indian	12.25	156	131	4180	3249	51.134626	0.343939
68	A.D. Hales	KKR	Overseas	1.50	329	302	9284	6348	50.804422	0.327010
71	A.M. Rahane	KKR	Indian	1.00	202	182	5395	4496	47.541417	0.186519
73	R.K. Singh	KKR	Indian	0.55	56	42	981	716	39.910496	0.032168
76	A. Tomar	KKR	Indian	0.40	5	5	65	89	22.071307	0.000047
82	P. Singh	KKR	Indian	0.20	27	25	829	649	52.651635	0.427247
89	M.K. Pandey	LSG	Indian	4.60	251	195	6315	5098	51.345638	0.354996
91	E. Lewis	LSG	Overseas	2.00	188	176	5518	3797	51.577324	0.367346
96	M. Vohra	LSG	Indian	0.20	105	97	2813	2170	47.396799	0.181456
116	A. Singh	MI	Indian	0.20	29	27	503	471	31.725008	0.002444
119	R.R. Buddhi	MI	Indian	0.20	12	7	222	163	51.448436	0.360449
125	S. Dhawan	PBKS	Indian	8.25	300	270	8775	7032	51.569111	0.366904
136	P.B. Rajapaksa	PBKS	Overseas	0.50	91	81	1912	1420	40.168067	0.034483

146	S.O. Hetmyer	RR	Overseas	8.50	111	92	2339	1784	42.589221	0.064000
148	D. Padikkal	RR	Indian	7.75	54	49	1829	1340	58.621795	0.789929
155	K. Nair	RR	Indian	1.40	133	117	2973	2261	42.590072	0.064013
156	H.E. van der Dussen	RR	Overseas	1.00	132	108	4129	3172	59.103922	0.816430
169	F. du Plessis	RCB	Overseas	7.00	256	226	6984	5408	49.871465	0.281768
176	F.H. Allen	RCB	Overseas	0.80	51	48	1537	875	54.167401	0.517755
187	R.A. Tripathi	SRH	Indian	8.50	100	86	2215	1700	43.009709	0.070802
194	A.K. Markram	SRH	Overseas	2.60	64	54	1732	1343	51.371793	0.356380
200	P.K. Garg	SRH	Indian	0.20	32	29	518	451	30.916144	0.001822
202	R. Samarth	SRH	Indian	0.20	21	20	275	291	24.006984	0.000112

Table B.3: Dataset 3 relating to all-rounders [276], [282]

(a) Bowling data for all-rounders [276], [282]

Player number	Player name	Franchise	Nationality	Auctioned price (₹ crore)	Innings bowled	Total wickets	Runs conceded	Balls bowled	CBR score	CBR' value
3	D.J. Bravo	CSK	Overseas	4.40	492	569	13752	10049	13.649598	0.603533
4	S. Dube	CSK	Indian	4.00	53	33	1052	723	15.661241	0.426275
5	C.J. Jordan	CSK	Overseas	3.60	254	265	7354	5166	14.676322	0.513063
7	M.J. Santner	CSK	Overseas	1.90	118	119	2916	2466	13.042791	0.657003
13	D. Pretorius	CSK	Overseas	0.50	131	108	2815	2077	14.062315	0.567166
23	M.R. Marsh	DC	Overseas	6.50	89	66	1733	1254	14.196752	0.555320
31	K.L. Nagarkoti	DC	Indian	1.10	16	12	399	289	15.599315	0.431732
33	L. Yadav	DC	Indian	0.65	41	32	870	754	13.412834	0.624396
37	P. Dubey	DC	Indian	0.50	22	20	446	375	12.588597	0.697025
43	R. Tewatia	GT	Indian	9.00	75	54	1566	1265	14.164155	0.558193
53	J. Yadav	GT	Indian	1.70	69	44	1407	1330	13.519572	0.614991
54	V. Shankar	GT	Indian	1.40	60	33	1123	821	15.671795	0.425345
55	D.C. Drakes	GT	Overseas	1.10	23	22	707	457	16.043669	0.392577
56	G.R. Singh	GT	Indian	0.50	19	8	262	222	14.437531	0.534104
59	D.G. Nalkande	GT	Indian	0.20	22	43	549	454	9.650515	0.955917
61	P.J. Sangwan	GT	Indian	0.20	96	100	2602	2032	13.774050	0.592567
63	N. Rana	KKR	Indian	8.00	49	29	651	565	12.472792	0.707229
64	P.J. Cummins	KKR	Overseas	7.25	111	124	3225	2479	13.851472	0.585745
65	S.P. Mavi	KKR	Indian	7.25	33	31	910	667	14.799915	0.502172
70	M. Nabi	KKR	Overseas	1.00	315	300	7251	6262	12.863623	0.672790
75	C. Karunaratne	KKR	Overseas	0.50	43	34	957	688	14.650929	0.515300

77	A.S. Roy	KKR	Indian	0.20	31	19	694	587	14.946513	0.489254
84	J.O. Holder	LSG	Overseas	8.75	157	153	4241	3275	14.184540	0.556396
85	K.H. Pandya	LSG	Indian	8.25	136	101	3131	2560	14.424264	0.535273
88	D.J. Hooda	LSG	Indian	5.75	54	17	913	703	17.530172	0.261593
92	K. Gowtham	LSG	Indian	0.90	67	48	1560	1266	14.710185	0.510079
95	K.R. Mayers	LSG	Overseas	0.50	34	19	651	491	15.496952	0.440752
102	T.H. David	MI	Overseas	8.25	28	7	489	324	20.498900	0.000000
103	J.C. Archer	MI	Overseas	8.00	118	153	3446	2700	12.950451	0.665139
105	D.R. Sams	MI	Overseas	2.60	69	87	1992	1408	13.436705	0.622293
111	F.A. Allen	MI	Overseas	0.75	57	41	1308	1009	14.959039	0.488151
113	R.S. Yadav	MI	Indian	0.50	19	16	446	402	13.280233	0.636080
117	R. Singh	MI	Indian	0.20	5	3	76	66	13.062500	0.655266
122	L.S. Livingstone	PBKS	Overseas	11.50	69	67	1404	1070	12.639260	0.692560
124	M.S. Khan	PBKS	Indian	9.00	6	2	103	80	17.254537	0.285880
127	O.F. Smith	PBKS	Overseas	6.00	32	36	889	596	14.108773	0.563073
129	H. Brar	PBKS	Indian	3.80	31	30	672	626	12.105359	0.739606
134	R. Dhawan	PBKS	Indian	0.55	96	87	2373	1905	13.880623	0.583176
137	B.A. Howell	PBKS	Overseas	0.40	135	161	3393	2857	12.287961	0.723515
140	P.N. Mankad	PBKS	Indian	0.20	28	20	611	426	15.315145	0.456772
141	A. Taide	PBKS	Indian	0.20	12	10	134	150	9.150228	1.000000
142	W.B. Chatterjee	PBKS	Indian	0.20	22	21	418	402	11.417040	0.800257
150	R. Ashwin	RR	Indian	5.00	262	264	6548	5682	12.964313	0.663918
151	R. Parag	RR	Indian	3.80	40	23	717	593	14.374232	0.539681
154	J.D. Neesham	RR	Overseas	1.50	161	158	4015	2670	14.329191	0.543650
157	D.J. Mitchell	RR	Overseas	0.75	63	59	1387	897	13.882510	0.583010

166	P.W. Hasaranga	RCB	Overseas	10.75	79	113	1775	1653	10.444453	0.885958
167	H.V. Patel	RCB	Indian	10.75	116	142	3234	2440	13.166415	0.646110
172	S. Ahamad	RCB	Indian	2.40	35	35	715	660	11.726651	0.772976
173	D.J. Willey	RCB	Overseas	2.00	184	203	4646	3527	13.170944	0.645710
174	S.E. Rutherford	RCB	Overseas	1.00	15	8	257	185	15.436410	0.446087
175	M.L. Lomror	RCB	Indian	0.95	35	8	489	397	17.460010	0.267775
180	S.S. Prabhudessai	RCB	Indian	0.30	9	2	127	102	17.729015	0.244071
186	W. Sundar	SRH	Indian	8.75	96	74	2129	1887	13.532061	0.613890
188	R. Shepherd	SRH	Overseas	7.75	39	53	1123	759	13.062309	0.655283
189	A. Sharma	SRH	Indian	6.50	22	16	320	276	11.917956	0.756119
191	M. Jansen	SRH	Overseas	4.20	13	9	349	270	15.952260	0.400632
203	S. Singh	SRH	Indian	0.20	16	10	332	240	15.603133	0.431396

(b) Batting data for all-rounders [276], [282]

Player number	Player name	Franchise	Nationality	Auctioned price (₹ crore)	Innings batted	Total dismissals	Runs scored	Balls faced	CBM score	CBM' value
3	D.J. Bravo	CSK	Overseas	4.40	402	285	6729	5292	39.825994	0.603190
4	S. Dube	CSK	Indian	4.00	62	47	1020	768	37.307974	0.550126
5	C.J. Jordan	CSK	Overseas	3.60	149	89	1307	1071	26.216026	0.316376
7	M.J. Santner	CSK	Overseas	1.90	93	58	1308	994	38.504563	0.575343
13	D. Pretorius	CSK	Overseas	0.50	120	93	1934	1385	36.200281	0.526783
23	M.R. Marsh	DC	Overseas	6.50	128	95	3241	2510	53.971690	0.901295
31	K.L. Nagarkoti	DC	Indian	1.10	9	5	66	87	22.487223	0.237795
33	L. Yadav	DC	Indian	0.65	37	19	698	480	58.655462	1.000000
37	P. Dubey	DC	Indian	0.50	16	9	126	123	24.633431	0.283024
43	R. Tewatia	GT	Indian	9.00	67	43	1170	822	45.685279	0.726668
53	J. Yadav	GT	Indian	1.70	38	22	313	268	25.364668	0.298434
54	V. Shankar	GT	Indian	1.40	87	65	1714	1356	43.635438	0.683470
55	D.C. Drakes	GT	Overseas	1.10	11	9	158	103	31.505484	0.427845
56	G.R. Singh	GT	Indian	0.50	100	77	1985	1600	42.688172	0.663507
59	D.G. Nalkande	GT	Indian	0.20	12	10	73	85	13.456221	0.047477
61	P.J. Sangwan	GT	Indian	0.20	43	21	260	236	22.260274	0.233013
63	N. Rana	KKR	Indian	8.00	128	114	3184	2355	46.295892	0.739536
64	P.J. Cummins	KKR	Overseas	7.25	59	34	576	431	30.070478	0.397604
65	S.P. Mavi	KKR	Indian	7.25	12	9	54	64	11.203320	0.000000
70	M. Nabi	KKR	Overseas	1.00	275	221	4933	3545	38.471437	0.574645
75	C. Karunaratne	KKR	Overseas	0.50	41	21	410	342	33.579034	0.471543
77	A.S. Roy	KKR	Indian	0.20	24	14	304	215	37.647059	0.557272
84	J.O. Holder	LSG	Overseas	8.75	121	91	1422	1111	27.852316	0.350859

85	K.H. Pandya	LSG	Indian	8.25	111	79	1753	1302	38.100413	0.566826
88	D.J. Hooda	LSG	Indian	5.75	116	91	2172	1565	40.731364	0.622270
92	K. Gowtham	LSG	Indian	0.90	53	41	610	390	27.171492	0.336511
95	K.R. Mayers	LSG	Overseas	0.50	66	52	946	815	31.454697	0.426775
102	T.H. David	MI	Overseas	8.25	79	55	1908	1197	56.980738	0.964707
103	J.C. Archer	MI	Overseas	8.00	61	32	551	373	30.842429	0.413872
105	D.R. Sams	MI	Overseas	2.60	58	46	682	451	27.004554	0.332993
111	F.A. Allen	MI	Overseas	0.75	54	35	745	502	37.231384	0.548512
113	R.S. Yadav	MI	Indian	0.50	23	18	494	398	44.949955	0.711172
117	R. Singh	MI	Indian	0.20	8	7	98	72	25.388601	0.298939
122	L.S. Livingstone	PBKS	Overseas	11.50	155	141	4095	2834	48.364238	0.783124
124	M.S. Khan	PBKS	Indian	9.00	39	26	547	401	36.454515	0.532140
127	O.F. Smith	PBKS	Overseas	6.00	16	11	192	147	30.793905	0.412849
129	H. Brar	PBKS	Indian	3.80	18	7	149	132	35.817308	0.518712
134	R. Dhawan	PBKS	Indian	0.55	82	46	1423	1200	49.068966	0.797975
137	B.A. Howell	PBKS	Overseas	0.40	126	91	2101	1622	39.190450	0.589797
140	P.N. Mankad	PBKS	Indian	0.20	30	23	763	543	53.675695	0.895057
141	A. Taide	PBKS	Indian	0.20	20	17	507	391	48.493544	0.785849
142	W.B. Chatterjee	PBKS	Indian	0.20	15	12	262	200	37.428571	0.552667
150	R. Ashwin	RR	Indian	5.00	96	59	795	696	24.105518	0.271899
151	R. Parag	RR	Indian	3.80	48	41	1019	807	41.532505	0.639153
154	J.D. Neesham	RR	Overseas	1.50	143	104	2376	1735	39.159456	0.589144
157	D.J. Mitchell	RR	Overseas	0.75	117	88	2722	2055	50.152004	0.820799
166	P.W. Hasaranga	RCB	Overseas	10.75	68	55	992	726	31.866367	0.435450
167	H.V. Patel	RCB	Indian	10.75	74	55	949	631	30.957429	0.416295

172	S. Ahamad	RCB	Indian	2.40	24	18	293	250	28.585366	0.366307
173	D.J. Willey	RCB	Overseas	2.00	155	124	2943	2086	40.632335	0.620183
174	S.E. Rutherford	RCB	Overseas	1.00	72	55	1348	1009	41.419573	0.636773
175	M.L. Lomror	RCB	Indian	0.95	58	47	1367	1108	47.073003	0.755913
180	S.S. Prabhudessai	RCB	Indian	0.30	20	14	443	299	52.148323	0.862869
186	W. Sundar	SRH	Indian	8.75	61	42	866	730	35.131846	0.504267
188	R. Shepherd	SRH	Overseas	7.75	22	12	315	195	45.161290	0.715626
189	A. Sharma	SRH	Indian	6.50	41	32	791	569	41.973998	0.648457
191	M. Jansen	SRH	Overseas	4.20	9	8	71	63	16.454229	0.110657
203	S. Singh	SRH	Indian	0.20	27	21	424	298	35.362802	0.509134

(c) Characteristic data for all-rounders

Player number	Player name	w_1 value	w_2 value	AR score	Membership value
3	D.J. Bravo	0.50	0.50	0.603362	0.029001
4	S. Dube	0.62	0.38	0.503063	0.003093
5	C.J. Jordan	0.28	0.72	0.457990	0.000947
7	M.J. Santner	0.43	0.57	0.621889	0.041309
13	D. Pretorius	0.46	0.54	0.548590	0.009140
23	M.R. Marsh	0.72	0.28	0.804422	0.498266
31	K.L. Nagarkoti	0.23	0.77	0.387127	0.000118
33	L. Yadav	0.72	0.28	0.894831	0.875996
37	P. Dubey	0.14	0.86	0.639065	0.056396
43	R. Tewatia	0.63	0.37	0.664332	0.086600
53	J. Yadav	0.19	0.81	0.554845	0.010515
54	V. Shankar	0.72	0.28	0.611195	0.033756
55	D.C. Drakes	0.54	0.46	0.411622	0.000250
56	G.R. Singh	0.61	0.39	0.613040	0.034968
59	D.G. Nalkande	0.00	1.00	0.955917	0.997907
61	P.J. Sangwan	0.13	0.87	0.545825	0.008586
63	N. Rana	0.52	0.48	0.724029	0.207906
64	P.J. Cummins	0.32	0.68	0.525540	0.005356
65	S.P. Mavi	0.00	1.00	0.502172	0.003025
70	M. Nabi	0.42	0.58	0.631569	0.049328
75	C. Karunaratne	0.46	0.54	0.495172	0.002534
77	A.S. Roy	0.56	0.44	0.527344	0.005590
84	J.O. Holder	0.28	0.72	0.498846	0.002782
85	K.H. Pandya	0.53	0.47	0.551996	0.009867
88	D.J. Hooda	0.85	0.15	0.568168	0.014071
92	K. Gowtham	0.30	0.70	0.458008	0.000948
95	K.R. Mayers	0.48	0.52	0.434043	0.000483
102	T.H. David	1.00	0.00	0.964707	1.000000
103	J.C. Archer	0.28	0.72	0.594785	0.024465
105	D.R. Sams	0.22	0.78	0.558647	0.011437
111	F.A. Allen	0.56	0.44	0.521953	0.004915
113	R.S. Yadav	0.56	0.44	0.678132	0.107869
117	R. Singh	0.17	0.83	0.594690	0.024419
122	L.S. Livingstone	0.56	0.44	0.743276	0.264610
124	M.S. Khan	0.78	0.22	0.477963	0.001622
127	O.F. Smith	0.35	0.65	0.510495	0.003720

129	H. Brar	0.33	0.67	0.666711	0.090007
134	R. Dhawan	0.65	0.35	0.722796	0.204579
137	B.A. Howell	0.40	0.60	0.670028	0.094936
140	P.N. Mankad	0.79	0.21	0.803017	0.492192
141	A. Taide	0.38	0.62	0.918623	0.944041
142	W.B. Chatterjee	0.32	0.68	0.721029	0.199873
150	R. Ashwin	0.14	0.86	0.609035	0.032383
151	R. Parag	0.58	0.42	0.597375	0.025766
154	J.D. Neesham	0.54	0.46	0.568217	0.014085
157	D.J. Mitchell	0.66	0.34	0.739951	0.254176
166	P.W. Hasaranga	0.19	0.81	0.800362	0.480771
167	H.V. Patel	0.29	0.71	0.579463	0.017877
172	S. Ahamad	0.18	0.82	0.699775	0.149095
173	D.J. Willey	0.48	0.52	0.633457	0.051035
174	S.E. Rutherford	0.67	0.33	0.573847	0.015884
175	M.L. Lomror	0.89	0.11	0.702218	0.154394
180	S.S. Prabhudessai	0.93	0.07	0.819554	0.564790
186	W. Sundar	0.40	0.60	0.570041	0.014648
188	R. Shepherd	0.54	0.46	0.687868	0.125168
189	A. Sharma	0.42	0.58	0.710901	0.174351
191	M. Jansen	0.07	0.93	0.380334	0.000095
203	S. Singh	0.58	0.42	0.476484	0.001560

Table B.4: Dataset 4 relating to wicketkeepers [276], [282]

Player number	Player name	Franchise	Nationality	Auctioned price (₹ crore)	Innings kept	Total dismissals	Catches keeping	Stumpings keeping	WK score	Membership value
2	A.T. Rayudu	CSK	Indian	6.75	37	25	22	3	0.675676	0.003165
15	N. Jagadeesan	CSK	Indian	0.20	22	11	8	3	0.500000	0.000074
30	K.S. Bharat	DC	Indian	2.00	61	58	47	11	0.950820	0.171666
36	T.L. Seifert	DC	Overseas	0.50	132	113	93	20	0.856061	0.056234
49	M.S. Wade	GT	Overseas	2.40	145	99	87	12	0.682759	0.003610
52	W.P. Saha	GT	Indian	1.90	180	137	104	33	0.761111	0.013986
66	S.W. Billings	KKR	Overseas	2.00	123	98	77	21	0.796748	0.024346
72	S.P. Jackson	KKR	Indian	0.60	33	38	32	6	1.151515	0.742344
87	Q. de Kock	LSG	Overseas	6.75	206	195	153	42	0.946602	0.164296
101	I.P. Kishan	MI	Indian	15.25	61	50	42	8	0.819672	0.034076
126	J.M. Bairstow	PBKS	Overseas	6.75	88	71	56	15	0.806818	0.028276
133	P. Singh	PBKS	Indian	0.60	30	18	15	3	0.600000	0.000706
139	J.M. Sharma	PBKS	Indian	0.20	40	39	31	8	0.975000	0.218484
170	K.D. Karthik	RCB	Indian	5.50	285	236	176	60	0.828070	0.038389
171	A. Rawat	RCB	Indian	3.40	24	31	25	6	1.291667	1.000000
183	L.S. Sisodia	RCB	Indian	0.20	7	6	2	4	0.857143	0.057044
185	N. Pooran	SRH	Overseas	10.75	144	109	90	19	0.756944	0.013075
196	G.D. Phillips	SRH	Overseas	1.50	76	61	51	10	0.802632	0.026580
198	V. Vinod	SRH	Indian	0.50	8	6	1	5	0.750000	0.011674

Table B.5: Verification subset [282]

Player number	Player name	Nationality	Auctioned price (₹ crore)	Cricketing ability	Membership value
1	D.L. Chahar	Indian	14.00	Bowling	0.127294
18	S. Singh	Indian	0.20	Bowling	0.395560
26	C. Sakariya	Indian	4.20	Bowling	0.344227
29	K. Yadav	Indian	2.00	Bowling	0.257871
34	L.T. Ngidi	Overseas	0.50	Bowling	0.478970
45	Y. Dayal	Indian	3.20	Bowling	0.285506
147	T.A. Boult	Overseas	8.00	Bowling	0.047363
168	J.R. Hazlewood	Overseas	7.75	Bowling	0.247084
181	C.V. Milind	Indian	0.25	Bowling	0.741317
197	S. Gopal	Indian	0.75	Bowling	0.329981
24	D.A. Warner	Overseas	6.25	Batting	0.840112
38	K.A. Hebbar	Indian	0.20	Batting	0.543300
62	S.S. Iyer	Indian	12.25	Batting	0.343939
68	A.D. Hales	Overseas	1.50	Batting	0.327010
82	P. Singh	Indian	0.20	Batting	0.427247
89	M.K. Pandey	Indian	4.60	Batting	0.354996
119	R.R. Buddhi	Indian	0.20	Batting	0.360449
148	D. Padikkal	Indian	7.75	Batting	0.789929
169	F. du Plessis	Overseas	7.00	Batting	0.281768
200	P.K. Garg	Indian	0.20	Batting	0.001822
23	M.R. Marsh	Overseas	6.50	All-rounding	0.498266
33	L. Yadav	Indian	0.65	All-rounding	0.875996
54	V. Shankar	Indian	1.40	All-rounding	0.033756
59	D.G. Nalkande	Indian	0.20	All-rounding	0.997907
63	N. Rana	Indian	8.00	All-rounding	0.207906
70	M. Nabi	Overseas	1.00	All-rounding	0.049328
150	R. Ashwin	Indian	5.00	All-rounding	0.032383
157	D.J. Mitchell	Overseas	0.75	All-rounding	0.254176
166	P.W. Hasaranga	Overseas	10.75	All-rounding	0.480771
175	M.L. Lomror	Indian	0.95	All-rounding	0.154394
30	K.S. Bharat	Indian	2.00	Wicketkeeping	0.171666
36	T.L. Seifert	Overseas	0.50	Wicketkeeping	0.056234
52	W.P. Saha	Indian	1.90	Wicketkeeping	0.013986
87	Q. de Kock	Overseas	6.75	Wicketkeeping	0.164296
101	I.P. Kishan	Indian	15.25	Wicketkeeping	0.034076

139	J.M. Sharma	Indian	0.20	Wicketkeeping	0.218484
170	K.D. Karthik	Indian	5.50	Wicketkeeping	0.038389
171	A. Rawat	Indian	3.40	Wicketkeeping	1.000000
185	N. Pooran	Overseas	10.75	Wicketkeeping	0.013075
198	V. Vinod	Indian	0.50	Wicketkeeping	0.011674

Appendix C

Comparison and experimental data

Appendix C presents the comparison and experimental data pertaining to the performance measures assessment and mixture design optimisation, respectively. Table C.1 specifies the performance data for the top 20 run scorers of the 2003 One Day International (ODI) World Cup. Table C.2 describes the performance data for the 20 randomly selected all-rounders of the 2016 Twenty20 (T20) World Cup. The secondary Tables C.2(a), C.2(b), C.2(c), and C.2(d) define the bowling, batting, AR, and SB data corresponding to the 20 all-rounders, respectively. From Table C.2(d), the batting criterion U_1 was obtained for each all-rounder by assigning equal importance to batting average and strike rate by setting $\alpha = 0.50$. The bowling criterion U_2 was determined analogously for each all-rounder by assigning the same emphasis to bowling economy rate, average, and strike rate, through $\alpha_1 = \alpha_2 = 0.33$. The bowling index c_2 was found for each all-rounder, where the sample standard deviation s_{c_2} converged with the sample standard deviation s_{c_1} of the batting index c_1 to an acceptable accuracy of six decimal places. Similarly, the all-rounding index SB for $\beta = 0.25$, $\beta = 0.50$, and $\beta = 0.75$ was obtained for each all-rounder, where the sample standard deviation s_{SB} converged with the sample standard deviation s_{c_1} of the batting index c_1 to an acceptable accuracy of six decimal places. Table C.3 gives the results of the I-optimal crossed mixture design, detailing the components and response associated with each of the 70 runs.

Table C.1: Performance data for top 20 run scorers of the 2003 ODI World Cup [289]

Player name	**Country**	**Runs scored**	***P(out)* (%)**	**Innings batted**	**Not outs**	**Balls faced**	**Batting average**	**Batting strike rate**	***BK* ($\alpha = 0.25$)**	***BK* ($\alpha = 0.50$)**	***BK* ($\alpha = 0.75$)**	***CBM* score**
S.R. Tendulkar	India	673	1.46	11	0	754	61.18	89.26	67.240001	73.898062	81.215399	72.599784
S.C. Ganguly	India	465	1.42	11	3	565	58.13	82.30	63.404977	69.164579	75.447373	68.131868
R.T. Ponting	Australia	415	1.69	10	2	472	51.88	87.92	59.189559	67.535498	77.058243	65.251572
A.C. Gilchrist	Australia	408	2.58	10	0	387	40.80	105.43	51.728802	65.585024	83.152812	58.832012
H.H. Gibbs	South Africa	384	1.05	6	2	381	96.00	100.79	97.175098	98.364580	99.568622	98.335467
M.S. Atapattu	Sri Lanka	382	1.55	10	3	452	54.57	84.51	60.877278	67.911782	75.759138	66.319444
A. Flower	Zimbabwe	332	1.53	7	0	459	47.43	72.33	52.706161	58.571011	65.088469	57.290768
M.L. Hayden	Australia	328	2.44	11	1	410	32.80	80.00	40.989996	51.224994	64.015619	46.524823
A. Symonds	Australia	326	0.56	5	3	360	163.00	90.56	140.724423	121.493027	104.889792	116.428571
D.R. Martyn	Australia	323	1.27	8	3	395	64.60	81.77	68.521324	72.680678	77.092512	72.178771
S.P. Fleming	New Zealand	321	1.87	8	1	374	45.86	85.83	53.636892	62.736489	73.379850	59.776536
S.T. Jayasuriya	Sri Lanka	321	1.90	10	2	420	40.13	76.43	47.138441	55.377761	65.057230	52.622951
R. Dravid	India	318	1.01	10	5	496	63.60	64.11	63.727840	63.855937	63.984291	63.855422
V. Sehwag	India	299	3.19	11	0	345	27.18	86.67	36.322180	48.536147	64.857275	41.384083
C.B. Wishart	Zimbabwe	293	1.75	7	1	343	48.83	85.42	56.160460	64.586974	74.277832	62.142100
S.B. Styris	New Zealand	268	1.90	7	2	263	53.60	101.90	62.938785	73.904676	86.781166	70.249017
P.A. de Silva	Sri Lanka	267	2.68	8	0	299	33.38	89.30	42.685068	54.592210	69.820889	48.589627
R.D. Shah	Kenya	265	2.05	9	0	438	29.44	60.50	35.252950	42.207299	50.533533	39.611360
B.C. Lara	West Indies	248	1.96	6	0	306	41.33	81.05	48.911151	57.878243	68.489311	54.746137
Y. Singh	India	240	2.49	10	3	281	34.29	85.41	43.073596	54.113928	67.984044	48.929664

Table C.2: Performance data for 20 selected all-rounders of 2016 T20 World Cup [293]

(a) Bowling data for the 20 randomly selected all-rounders [294]

Player name	Country	Innings bowled	Total wickets	Runs conceded	Balls bowled	Overs bowled	Maidens bowled	Bowling average	Bowling economy rate	Bowling strike rate
M. Nabi	Afghanistan	7	12	164	162	27.00	1	13.67	6.07	13.50
S.R. Watson	Australia	4	5	103	84	14.00	0	20.60	7.36	16.80
S. Al Hasan	Bangladesh	6	10	166	138	23.00	0	16.60	7.22	13.80
D. Willey	England	6	10	159	126	21.00	0	15.90	7.57	12.60
C.J. Anderson	New Zealand	5	2	84	72	12.00	0	42.00	7.00	36.00
R. Ashwin	India	5	4	115	90	15.00	0	28.75	7.67	22.50
D.J. Sammy	West Indies	2	1	31	18	3.00	0	31.00	10.33	18.00
C.R. Brathwaite	West Indies	6	4	145	108	18.00	0	36.25	8.06	27.00
M. Mahmudullah	Bangladesh	4	1	55	36	6.00	0	55.00	9.17	36.00
T. Afzal	Hong Kong	3	2	51	48	8.00	0	25.50	6.38	24.00
T.A. Siriwardana	Sri Lanka	4	2	63	42	7.00	0	31.50	9.00	21.00
G.J. Maxwell	Australia	4	2	61	48	8.00	0	30.50	7.63	24.00
M.S. Rahman	Bangladesh	2	2	16	18	3.00	0	8.00	5.33	9.00
S. Raza	Zimbabwe	3	1	44	36	6.00	0	44.00	7.33	36.00
H.H. Pandya	India	5	5	143	84	14.00	0	28.60	10.21	16.80
S. Shinwari	Afghanistan	7	3	138	132	22.00	0	46.00	6.27	44.00
D.J. Bravo	West Indies	6	9	190	144	24.00	0	21.11	7.92	16.00
N.L. Perera	Sri Lanka	4	3	86	50	8.20	0	28.67	10.32	16.67
B.A. Stokes	England	6	4	179	112	18.40	0	44.75	9.59	28.00
S.C. Williams	Zimbabwe	3	2	61	54	9.00	0	30.50	6.78	27.00

(b) Batting data for the 20 randomly selected all-rounders [295]

Player name	Country	Innings batted	Total dismissals	Runs scored	Balls faced	Highest score	Fifties scored	Sixes scored	Fours scored	Batting average	Batting strike rate
M. Nabi	Afghanistan	7	7	105	80	52	1	3	8	15.00	131.25
S.R. Watson	Australia	4	2	96	64	44*	0	4	9	48.00	150.00
S. Al Hasan	Bangladesh	7	4	129	102	50*	1	5	10	32.25	126.47
D. Willey	England	3	2	41	31	21	0	4	1	20.50	132.26
C.J. Anderson	New Zealand	5	5	86	88	34	0	1	8	17.20	97.73
R. Ashwin	India	2	1	15	22	10	0	0	1	15.00	68.18
D.J. Sammy	West Indies	3	3	8	13	6	0	0	0	2.67	61.54
C.R. Brathwaite	West Indies	3	1	57	28	34*	0	7	1	57.00	203.57
M. Mahmudullah	Bangladesh	5	4	86	73	49*	0	1	9	21.50	117.81
T. Afzal	Hong Kong	3	1	37	29	31*	0	1	3	37.00	127.59
T.A. Siriwardana	Sri Lanka	3	3	22	20	15	0	1	2	7.33	110.00
G.J. Maxwell	Australia	4	4	109	84	31	0	4	8	27.25	129.76
M.S. Rahman	Bangladesh	7	6	147	119	44	0	4	17	24.50	123.53
S. Raza	Zimbabwe	3	3	27	30	15	0	0	1	9.00	90.00
H.H. Pandya	India	2	2	16	14	15	0	1	2	8.00	114.29
S. Shinwari	Afghanistan	5	5	122	95	43	0	3	11	24.40	128.42
D.J. Bravo	West Indies	4	4	63	66	28	0	2	3	15.75	95.45
N.L. Perera	Sri Lanka	4	4	80	55	40	0	4	7	20.00	145.45
B.A. Stokes	England	5	4	56	33	15	0	3	5	14.00	169.70
S.C. Williams	Zimbabwe	3	3	78	58	53	1	0	9	26.00	134.48

(c) *AR* data for the 20 randomly selected all-rounders

Player name	***CBR*** **score**	***CBM*** **score**	***CBR′*** **value**	***CBM′*** **value**	w_1 **value**	w_2 **value**	***AR*** **score**
M. Nabi	9.619117	26.923077	0.793165	0.259810	0.10	0.90	0.739830
S.R. Watson	12.295595	72.727273	0.574973	0.805419	0.66	0.34	0.727067
S. Al Hasan	11.059543	51.394422	0.675738	0.551307	0.40	0.60	0.625966
D. Willey	10.935590	35.497835	0.685843	0.361951	0.22	0.78	0.614587
C.J. Anderson	15.428571	29.251701	0.319565	0.287548	0.45	0.55	0.305157
R. Ashwin	14.308756	24.590164	0.410855	0.232021	0.24	0.76	0.367935
D.J. Sammy	16.252427	5.111821	0.252402	0.000000	0.00	1.00	0.252402
C.R. Brathwaite	15.893099	89.062500	0.281695	1.000000	0.93	0.07	0.949719
M. Mahmudullah	19.348534	36.363636	0.000000	0.372264	1.00	0.00	0.372264
T. Afzal	12.618557	57.364341	0.548644	0.622419	0.56	0.44	0.589958
T.A. Siriwardana	15.750000	13.750000	0.293361	0.102896	0.11	0.89	0.272410
G.J. Maxwell	14.591362	45.041322	0.387816	0.475630	0.60	0.40	0.440505
M.S. Rahman	7.081967	40.890125	1.000000	0.426182	0.15	0.85	0.913927
S. Raza	16.054054	16.363636	0.268574	0.134029	0.20	0.80	0.241665
H.H. Pandya	15.593250	14.953271	0.306140	0.117229	0.13	0.87	0.281581
S. Shinwari	14.714055	41.008403	0.377814	0.427591	0.56	0.44	0.405689
D.J. Bravo	12.701950	27.038627	0.541846	0.261187	0.19	0.81	0.488520
N.L. Perera	15.642684	35.164835	0.302110	0.357984	0.58	0.42	0.334517
B.A. Stokes	18.479312	25.866051	0.070861	0.247219	0.92	0.08	0.233111
S.C. Williams	13.801676	43.575419	0.452193	0.458169	0.51	0.49	0.455241

(d) SB data for the 20 randomly selected all-rounders

Player name	U_1 ($\alpha = 0.50$)	c_1	U_2 ($\alpha_{1,2} = 0.33$)	k	V_2	c_2	SB ($\beta = 0.25$)	SB ($\beta = 0.50$)	SB ($\beta = 0.75$)
M. Nabi	44.370598	0.873575	10.387048	1.00	0.969793	1.697837	1.557156	1.310048	1.039609
S.R. Watson	84.852814	1.670594	13.655111	1.00	0.960290	1.318569	1.505773	1.717802	1.767409
S. Al Hasan	63.864517	1.257374	11.824673	1.00	0.965613	1.519606	1.572167	1.558291	1.414374
D. Willey	52.070052	1.025162	11.489871	1.00	0.966586	1.559433	1.512672	1.379103	1.176703
C.J. Anderson	40.998891	0.807192	21.955834	1.00	0.936150	0.685920	0.663634	0.666959	0.726342
R. Ashwin	31.980107	0.629629	17.053322	1.00	0.950407	1.011064	0.877261	0.733881	0.649353
D.J. Sammy	12.810252	0.252210	17.931841	1.00	0.947852	0.943560	0.623131	0.373989	0.268705
C.R. Brathwaite	107.719875	2.120804	19.903178	1.00	0.942120	0.807487	1.034196	1.445658	1.896902
M. Mahmudullah	50.327693	0.990859	26.280011	1.00	0.923575	0.484732	0.514277	0.605067	0.789992
T. Afzal	68.707275	1.352718	15.742627	1.00	0.954219	1.120472	1.216654	1.329645	1.376987
T.A. Siriwardana	28.401878	0.559180	18.124142	1.00	0.947293	0.929373	0.783363	0.638636	0.565621
G.J. Maxwell	59.464375	1.170743	17.738503	1.00	0.948415	0.958034	1.008885	1.081801	1.144376
M.S. Rahman	55.013367	1.083111	7.268482	1.00	0.978863	2.156096	2.068644	1.787938	1.371691
S. Raza	28.460499	0.560334	22.647427	1.00	0.934139	0.649077	0.564522	0.500107	0.506384
H.H. Pandya	30.237158	0.595313	16.993954	1.00	0.950580	1.015790	0.866087	0.708498	0.616904
S. Shinwari	55.977439	1.102092	23.328616	1.00	0.932158	0.614651	0.660061	0.765794	0.940660
D.J. Bravo	38.773820	0.763384	13.880054	1.00	0.959635	1.295704	1.167180	0.992533	0.841333
N.L. Perera	53.935989	1.061899	17.020352	1.00	0.950503	1.013686	1.031214	1.051754	1.062650
B.A. Stokes	48.741744	0.959634	22.904058	1.00	0.933393	0.635895	0.652762	0.712917	0.834657
S.C. Williams	59.131647	1.164192	17.738503	1.00	0.948415	0.958034	1.007161	1.077650	1.138351

Table C.3: I-optimal crossed mixture design results

Run	w_1	w_2	w_3	w_4	a_1	a_2	a_3	a_4	Response
1	0.366887	0.233113	0.200000	0.200000	0.299368	0.292871	0.207761	0.200000	95.661920
2	0.400000	0.200000	0.200000	0.200000	0.400000	0.200000	0.200000	0.200000	95.564827
3	0.200000	0.400000	0.200000	0.200000	0.400000	0.200000	0.200000	0.200000	95.847446
4	0.282031	0.278215	0.234245	0.205509	0.200000	0.400000	0.200000	0.200000	95.792081
5	0.232076	0.359769	0.208154	0.200000	0.299478	0.200000	0.200000	0.300522	95.993345
6	0.299883	0.200000	0.200000	0.300117	0.220000	0.200000	0.358740	0.221260	95.981293
7	0.210145	0.200000	0.295620	0.294235	0.200000	0.222611	0.200000	0.377389	96.298992
8	0.200000	0.303989	0.200000	0.296011	0.200000	0.222574	0.357658	0.219768	95.985116
9	0.298954	0.200000	0.301046	0.200000	0.350511	0.240346	0.200000	0.209144	95.720821
10	0.221660	0.200000	0.200000	0.378340	0.203893	0.300965	0.295141	0.200000	96.249450
11	0.366259	0.211467	0.200000	0.222274	0.302287	0.200000	0.200000	0.297713	95.942921
12	0.200000	0.301503	0.298497	0.200000	0.237117	0.200000	0.362883	0.200000	95.753870
13	0.400000	0.200000	0.200000	0.200000	0.200000	0.400000	0.200000	0.200000	95.735120
14	0.216064	0.366022	0.200000	0.217914	0.301778	0.298222	0.200000	0.200000	95.730672
15	0.200000	0.200000	0.200000	0.400000	0.400000	0.200000	0.200000	0.200000	96.342563
16	0.400000	0.200000	0.200000	0.200000	0.200000	0.200000	0.200000	0.400000	96.144484
17	0.365662	0.200000	0.210360	0.223978	0.200000	0.302872	0.297128	0.200000	95.710753
18	0.200000	0.400000	0.200000	0.200000	0.200000	0.200000	0.200000	0.400000	96.210851
19	0.200000	0.200000	0.200000	0.400000	0.200000	0.200000	0.200000	0.400000	96.633273
20	0.200000	0.200000	0.400000	0.200000	0.200000	0.400000	0.200000	0.200000	95.728850
21	0.215361	0.206205	0.291144	0.287290	0.400000	0.200000	0.200000	0.200000	96.040646
22	0.259156	0.255104	0.285740	0.200000	0.281093	0.318907	0.200000	0.200000	95.791084
23	0.200000	0.299758	0.200000	0.300242	0.220098	0.200000	0.220796	0.359106	96.311727

24	0.200000	0.200000	0.228220	0.371780	0.201649	0.200000	0.298214	0.300137	96.377810
25	0.289451	0.291513	0.211814	0.207223	0.400000	0.200000	0.200000	0.200000	95.795745
26	0.200000	0.200000	0.400000	0.200000	0.200000	0.200000	0.200000	0.400000	96.158954
27	0.200000	0.299038	0.300962	0.200000	0.209327	0.218723	0.200000	0.371950	96.151029
28	0.293719	0.289264	0.202818	0.214199	0.200000	0.200000	0.200000	0.400000	96.331912
29	0.200000	0.248765	0.268459	0.282776	0.314862	0.208544	0.203946	0.272648	96.093848
30	0.200000	0.200000	0.302431	0.297569	0.200000	0.229603	0.370397	0.200000	95.993559
31	0.215000	0.200000	0.371000	0.214000	0.296170	0.200000	0.298503	0.205326	95.722909
32	0.200000	0.200000	0.400000	0.200000	0.200000	0.200000	0.400000	0.200000	95.657487
33	0.200000	0.400000	0.200000	0.200000	0.200000	0.400000	0.200000	0.200000	95.623168
34	0.306255	0.200000	0.293745	0.200000	0.204000	0.344742	0.200000	0.251258	95.901265
35	0.200000	0.200000	0.200000	0.400000	0.200000	0.200000	0.400000	0.200000	96.315287
36	0.297777	0.200000	0.200000	0.302223	0.355644	0.219214	0.225141	0.200000	95.989531
37	0.250000	0.250000	0.250000	0.250000	0.250000	0.250000	0.250000	0.250000	96.010725
38	0.200000	0.303681	0.296319	0.200000	0.354286	0.232512	0.209039	0.204164	95.842136
39	0.200000	0.200000	0.400000	0.200000	0.399905	0.200000	0.200095	0.200000	95.768278
40	0.273143	0.275414	0.251444	0.200000	0.320783	0.200000	0.221572	0.257645	95.894071
41	0.200000	0.305323	0.294677	0.200000	0.207779	0.350073	0.242149	0.200000	95.752922
42	0.200000	0.228550	0.356155	0.215295	0.200000	0.200000	0.298455	0.301545	95.971978
43	0.288328	0.291581	0.207288	0.212803	0.200000	0.200000	0.400000	0.200000	95.844460
44	0.298917	0.200000	0.301083	0.200000	0.230526	0.200000	0.215023	0.354451	96.112494
45	0.200000	0.200000	0.363341	0.236659	0.306772	0.293228	0.200000	0.200000	95.796171
46	0.200000	0.210000	0.390000	0.200000	0.204376	0.293456	0.210512	0.291657	95.924227
47	0.200000	0.225492	0.200000	0.374508	0.204000	0.297126	0.200000	0.298874	96.381970
48	0.221270	0.200000	0.214774	0.363956	0.300160	0.200000	0.200000	0.299840	96.355776

49	0.200000	0.220000	0.216000	0.364000	0.300453	0.299547	0.200000	0.200000	96.215575
50	0.346959	0.253041	0.200000	0.200000	0.200000	0.305000	0.200000	0.295000	95.962041
51	0.278635	0.273045	0.200000	0.248319	0.265386	0.215432	0.200000	0.319182	96.092592
52	0.200000	0.214000	0.227000	0.359000	0.300745	0.200000	0.299255	0.200000	96.206459
53	0.355659	0.215855	0.228486	0.200000	0.298169	0.200000	0.301831	0.200000	95.678122
54	0.400000	0.200000	0.200000	0.200000	0.200000	0.200000	0.400000	0.200000	95.649107
55	0.200000	0.365910	0.214269	0.219821	0.200000	0.301203	0.200000	0.298797	95.962841
56	0.200000	0.301784	0.200000	0.298216	0.364000	0.200000	0.210000	0.226000	96.148103
57	0.200000	0.200000	0.299512	0.300488	0.200000	0.364831	0.212870	0.222299	96.108816
58	0.200000	0.400000	0.200000	0.200000	0.200000	0.200000	0.400000	0.200000	95.706422
59	0.301861	0.200000	0.200000	0.298139	0.200000	0.227048	0.222698	0.350255	96.273064
60	0.200000	0.200000	0.200000	0.400000	0.200000	0.400000	0.200000	0.200000	96.311390
61	0.200000	0.218000	0.355000	0.227000	0.300491	0.200000	0.200000	0.299509	96.044610
62	0.245694	0.354306	0.200000	0.200000	0.200000	0.304112	0.295888	0.200000	95.695460
63	0.222944	0.360340	0.200000	0.216716	0.298145	0.200000	0.301855	0.200000	95.732619
64	0.200000	0.298399	0.200000	0.301601	0.204720	0.370812	0.224138	0.200330	96.002591
65	0.301755	0.200000	0.200000	0.298245	0.226443	0.359557	0.200000	0.214000	95.999697
66	0.236528	0.200000	0.363472	0.200000	0.200000	0.302345	0.297655	0.200000	95.702971
67	0.297304	0.200000	0.300176	0.202520	0.200000	0.200000	0.372000	0.228000	95.843562
68	0.214398	0.372863	0.212739	0.200000	0.200000	0.200000	0.299326	0.300674	95.928458
69	0.364672	0.232746	0.202583	0.200000	0.200000	0.200000	0.302319	0.297681	95.918532
70	0.200000	0.322344	0.277656	0.200000	0.203817	0.260589	0.282141	0.253453	95.863882