OPTI-GOAL

USING OPTIMIZATION TO BEAT 99.995% OF FANTASY PREMIER LEAGUE MANAGERS WORLDWIDE

Managed by

Dilan SriDaran & Ethan Fahimi

8 December 2023

1 Introduction

We introduce an optimization framework to automate Fantasy Premier League management, overcoming typical human biases. Our approach ranks 463rd out of 10M+ active managers as at 8 December 2023. In this paper, we offer insights into the strategic elements of our approach, which not only provide a novel perspective on optimization in fantasy sports, but also pave the way for future developments in broader sports management.

1.1 Background

The English Premier League (EPL) stands as the top professional football league in England and is widely considered one of the most competitive and popular football leagues globally. Comprising 20 teams, each club competes in a round-robin format, playing a total of 38 matches in a season – 19 at home and 19 away.

Fantasy Premier League (FPL) is one of the world's largest fantasy football games, boasting over 10.3 million users, herein "Managers," for the 2023/24 season. Managers select "Players" each week, who score points based on their actions in real-life EPL matches. Each player is assigned a price in the game based on their perceived ability. For example, E. Haaland – the highest performer last season – is assigned a price of £14.0, whilst F. Pellistri – a rookie – is assigned a price of £4.5. Managers are allocated a fixed £100 budget to assemble a "Squad" of 15 players, comprising goalkeepers, defenders, midfielders, and forwards, across all 38 "Gameweeks." Buying all the best players is unfeasible due to budget constraints, and therefore success hinges on balancing high-performing players with cost-effective options. Managers then select a "Starting XI" from their squad to score points each week, and designate a "Captain" to earn double points. These points are based on real-life actions in EPL matches, such as goals, assists, clean sheets, and saves. The objective is to accumulate the highest point total across the season.

Managers can make player changes, "Transfers," each Gameweek to optimize their squad. This may reflect changing perceptions about players in response to recent form or injuries, or deliberate strategies to exploit favorable fixtures. For example, a forward player may have lower expected points in a Gameweek when they play against Manchester City – one of the best teams in the league – than when they play against bottom-of-the-table Burnley. Managers also have a total of five chips at their disposal over the course of a season: one "Triple Captain" (where captains earn triple points for a Gameweek), one "Bench Boost" (where all squad players, and not just the starting XI) score points for a Gameweek), one "Free Hit" (where unlimited transfers are allowed for a Gameweek, but the squad reverts to its initial composition at the end of the week), and two "Wildcards" (where unlimited transfers are allowed for a Gameweek, and the squad remains at this composition at the end of the week). The intricacies of FPL rules are discussed further in Section 3.

1.2 Scope

The objective of this project is to create an automated strategy to manage an FPL team over the 2023/24 season using optimization. The mixed-integer optimization model takes as input an externally sourced dataset of expected points by player by Gameweek for the entire 2023/24 season, set as at 10 August 2023, one day prior to the season commencing. The model then determines the optimal long-term strategy, including squad, team, transfers, and chip use by Gameweek for the entire season, reflecting the game's rules and scoring system. Whilst mathematically formulable, we made the decision not to incorporate the Free Hit chip into our framework. This is because this chip is often best saved for unforeseen circumstances that cannot be predicted or observed in the expected points input data, such as injuries and match cancellations due to severe weather events or fixture congestion.

The current season is underway, having commenced on 11 August 2023, and is due to complete in May 2024. This report evaluates the model on the completed 15 (of 38) Gameweeks as at the 8 December 2023, using the other 10,390,499 active managers as benchmarks.

2 Data

The optimization framework developed uses expected points data by player by Gameweek, which have been externally sourced from FPL Form (https://fplform.com). This dataset spans all 714 Premier League players, and was determined as at 10 August 2023, prior to the league season commencing. There is therefore no data leakage or use of prospective information to shape antecedent decisions. There is limited visibility regarding the calculation of these expected points, and we leave as a future refinement the task of creating our own data-driven framework to derive our own expected points.

Figure 1 illustrates predicted points for the top three players by total expected points over the entire season, split by position. These include goalkeepers such as A. Becker (Liverpool, Price: £6.0), defenders such as K. Trippier (Newcastle, £5.5), midfielders such as M. Salah (Liverpool, £12.5), and forwards such as E. Haaland (Manchester City, £14.0).

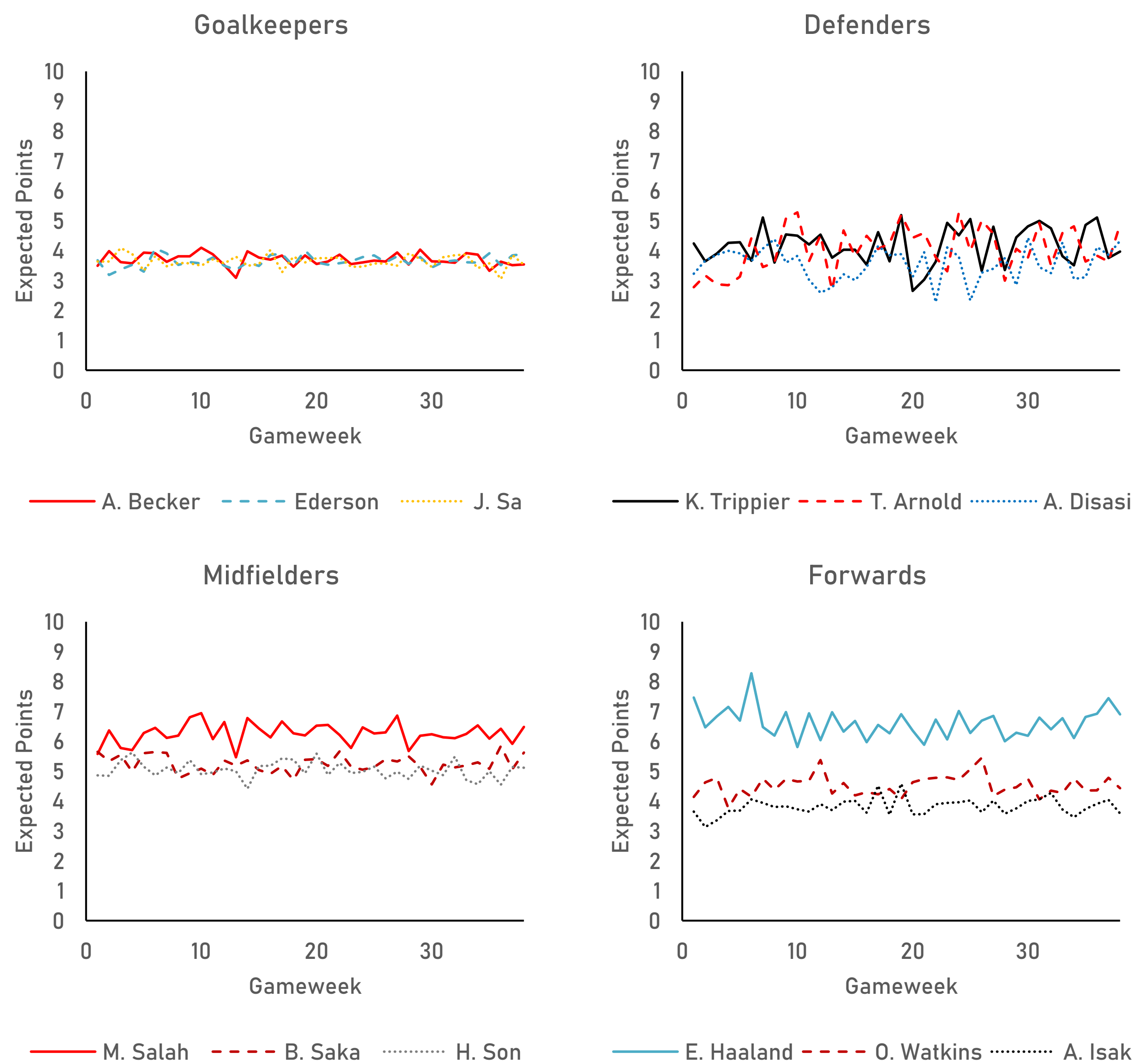


Figure 1: Expected points by high-profile players, split by goalkeepers, defenders, midfielders, and forwards.

Expected points per player vary by Gameweek, which may reflect a number of factors:

- **Opposition strength:** Players have higher expected points when playing against weaker teams. For example, a defender may be more likely to keep a clean sheet – and therefore score more points – when playing against a weaker team, such as Burnley, than a stronger attacking team, such as Manchester City.

- **Home or away:** Teams and players typically perform better when playing at home rather than away. For example, 46% of EPL matches are won by home teams, compared to just 34% for away teams, with the remaining 21% of matches drawn.[1]

- **Rescheduled fixtures:** Fixture changes are not uncommon in the EPL. For example, due to logistical issues, the match between Luton City and Burnley was postponed from Gameweek 2 to 7. These teams therefore had no scoring opportunities in Gameweek 2, but 2 matches from which to score points in Gameweek 7.

Key Finding: Expected points vary by Gameweek
Optimization has the potential to optimally determine transfer strategies and chip usage each week based on expected points variations by Gameweek

Table 1 demonstrates a key distinction between total expected points and value. The mechanism for points scoring in FPL is renowned for being biased in favor of attacking players. Therefore, total expected points are typically higher for forwards and midfielders. To compensate for this bias, goalkeepers and defenders are on average priced lower, such that "value" is more comparable between positions.

Table 1: Top players by expected points and by value.

Top 10 Players by Expected Points			
Player	Position	Price (£)	Expected Points
E. Haaland	FWD	14.0	251
M. Salah	MID	12.5	238
B. Mbeumo	MID	7.0	229
B. Saka	MID	8.5	199
M. Odegaard	MID	8.5	192
H. Son	MID	9.0	192
B. Fernandes	MID	8.5	189
J. Bowen	MID	7.0	187
J. Maddison	MID	8.0	182
O. Watkins	MID	8.0	172

Top 10 Players by Value[2]			
Player	Position	Price (£)	Value
B. Mbeumo	MID	7.0	32.7
A. Areola	GK	4.0	31.0
M. Turner	GK	4.0	29.8
B. Leno	GK	4.5	29.5
J. Pickford	GK	4.5	29.1
N. Neto	GK	4.5	28.9
M. Flekken	GK	4.5	28.8
J. Prowse	MID	6.0	28.6
R. Sanchez	GK	4.5	28.6
M. Cash	DEF	4.5	28.5

Key Finding: Expected points, price, and value vary by position
Optimization has the potential to identify the best subset of players each week to maximize expected points, subject to salary constraints, by balancing high scoring and high value players

[1] Source: SoccertStats, https://www.soccerstats.com, accessed 4 December 2023.
[2] "Value" is defined as total expected points per unit of price.

3 Optimization Framework

Let $p \in [P]$ denote the p^{th} Player, $w \in [38]$ denote the w^{th} Gameweek, and $k \in [20]$ denote the k^{th} Club.

3.1 Data

We define the following data inputs:

b_p	$\forall p \in [P]$	Price of Player p
z_{pw}	$\forall p \in [P], w \in [38]$	Expected points of Player p in Gameweek w
y_{pk}	$\forall p \in [P], k \in [20]$	1 if Player p belongs to Club k; else 0
g_p	$\forall p \in [P]$	1 if Player p is a goalkeeper (GK); else 0
d_p	$\forall p \in [P]$	1 if Player p is a defender (DEF); else 0
m_p	$\forall p \in [P]$	1 if Player p is a midfielder (MID); else 0
f_p	$\forall p \in [P]$	1 if Player p is a forward (FWD); else 0

3.2 Decision Variables

We define the following decision variables:

s_{pw}	$\forall p \in [P], w \in [38]$	1 if Player p is selected in Squad in Gameweek w; else 0
x_{pw}	$\forall p \in [P], w \in [38]$	1 if Player p is selected in Team in Gameweek w; else 0
b_w	$\forall w \in [38]$	1 if Bench Boost is played in Gameweek w; else 0
t_w	$\forall w \in [38]$	1 if Triple Captain is played in Gameweek w; else 0
c_{pw}	$\forall p \in [P], w \in [38]$	Captaincy multiplier applied to Player p in Gameweek w
n_{pw}	$\forall p \in [P], w \in [38]$	1 if Player p is transferred into Squad in Gameweek w; else 0
q_w	$\forall w \in [38]$	1 if Wildcard is played in Gameweek w; else 0
N_w	$\forall w \in [38]$	Number of points-penalizable transfers in Gameweek w

3.3 Objective Function

The objective is to maximize the expected points earned over the entire 38-week season:

$$\max \sum_{w=1}^{38} \sum_{p=1}^{P} z_{pw} \times (x_{pw} + c_{pw}) - \sum_{w=2}^{38} 4 \times N_w$$

The first component captures the expected total points earned by all players selected in the team across all weeks, including an uplift for the selected captain. The second component incorporates the 4-point penalty per "Penalizable Transfer" from Gameweek 2 to 38.

3.4 Constraints

Squad Selection

Each week, a squad of 15 players is needed, comprising exactly 2 GKs, 5 DEFs, 5 MIDs, and 3 FWDs:

$$\sum_{p=1}^{P} g_p \times s_{pw} = 2, \sum_{p=1}^{P} d_p \times s_{pw} = 5, \sum_{p=1}^{P} m_p \times s_{pw} = 5, \sum_{p=1}^{P} f_p \times s_{pw} = 3, \forall w \in [38]$$

Managers are allocated a fixed £100 budget to build their squad each week:

$$\sum_{p=1}^{P} b_p \times s_{pw} \leq 100, \forall w \in [38]$$

A maximum of three players can be chosen from each of the 20 Premier League clubs:

$$\sum_{p=1}^{P} y_{pk} \times s_{pw} \leq 3, \forall w \in [38], k \in [20]$$

Team Selection

For consistency, players selected into the team (to score points) must be members of the squad:

$$x_{pw} \leq s_{pw}, \forall w \in [38], p \in [P]$$

Once a season, a Bench Boost chip can be played:

$$\sum_{w=1}^{38} b_w = 1, \forall w \in [38]$$

In most weeks, only 11 players can be chosen in the team, however if the Bench Boost is played, all 15 players can score points for that week:

$$\sum_{p=1}^{P} x_{pw} \leq 11 + 4b_w, \forall w \in [38]$$

Regardless of the Bench Boost, the team must include at least 3 DEFs, 2 MIDs, and 1 FWD:

$$\sum_{p=1}^{P} d_p \times x_{pw} \geq 3, \sum_{p=1}^{P} m_p \times x_{pw} \geq 2, \sum_{p=1}^{P} f_p \times x_{pw} \geq 1, \forall w \in [38]$$

When the Bench Boost is not played versus played, exactly 1 GK and 2 GKs must be selected respectively:

$$\sum_{p=1}^{P} g_p \times x_{pw} = 1 + b_w, \forall w \in [38]$$

Captaincy

Each week, one player is selected as captain and earns double points. Once a season, a Triple Captain chip can be played, which results in the selected captain scoring triple points for that week only:

$$\sum_{w=1}^{38} t_w = 1, \forall w \in [38]$$

For consistency, the captain must be a member of the team, and the assigned multiplier must be capped at 1 if the Triple Captain is not played, and 2 if the Triple Captain is played:

$$c_{pw} \leq x_{pw} + t_w, \forall w \in [38], p \in [P]$$

No additional constraints are required to ensure that the whole captaincy multiplier is assigned to a single player. This will occur naturally as the solver will always be incentivized to assign the whole multiplier to the player in the team with the highest expected points for the week.

Transfer Management

Each week after the first week, managers carry forward their squad from the prior week, but are allowed to make changes to their players. Players transferred into the squad are identified using:

$$n_{pw} \geq s_{pw} - s_{p,w-1}, \forall w \in 2, \dots, 38, p \in [P]$$

$$n_{pw} \geq 0, \forall w \in 2, \dots, 38, p \in [P]$$

Each Gameweek, one transfer is allowed for "free," with subsequent changes costing 4 points per transfer. However, if a Wildcard is played, unlimited "free" transfers are permitted in that week. One Wildcard can be played within each half of the season:

$$\sum_{w=2}^{19} q_w = 1, \sum_{w=20}^{38} q_w = 1$$

The number of penalizable transfers is determined using:

$$N_w \geq \left(\sum_{p=1}^{P} n_{pw}\right) - (1 + 14q_w), \forall w \in [38]$$

$$N_w \geq 0, \forall w \in [38]$$

Chip Use

No more than one chip can be played in a single Gameweek w:

$$q_w + b_w + t_w \leq 1, \forall w \in [38]$$

3.4 Notes on Computation

Our formulation is extremely computationally intensive; given 714 players and 38 weeks, there are 100,00+ decision variables and 25,000+ potential constraints. The model defined above and applied to the data in its entirety was not tractable, and therefore, the following adjustments were made:

1. **Trimming players:** A heuristic approach was used to trim the initial list of 714 players to 223 players. The approach followed was to select the top 8 GKs, 12 DEFs, 15 MIDs, and 12 FWDs for each distinct price, based on total expected points for the whole season. This eliminated extremely poor-performing players that the optimization algorithm would have had an extremely low likelihood of selecting.

2. **Relaxing integrality constraints:** Integrality constraints were relaxed, such as for the binary decision variables x_{pw}, where it was determined that despite relaxing the constraints, the solution would be always guaranteed to remain integral.

3. **Splitting problem:** The problem was split into two subproblems and the solutions "glued" together. First, we solved the model for the first 19 weeks, and then for the second 19 weeks, with a starting squad fixed as the ending squad from the first half. The Wildcard lends itself well to such a split, as this chip can be played once in each season half. However, care had to be taken for the Bench Boost and Triple Captain. These chips can only be used once a season, i.e., if used in the first 19 weeks, the chip cannot be used in the second 19 weeks. We ran four separate combinations of the model, corresponding to the unique use of these two chips across each half, and selected the results with the highest overall expected points.

4 Results

4.1 Model Performance

Table 2 summarizes the performance of our model as at 8 December 2023 (end of Gameweek 15), relative to various benchmarks derived from 10,390,499 active players. This includes two current MBAn students relying on non-optimization heuristic approaches. Our approach performs extremely competitively, currently ranking 463rd out of the 10M+ players in the world, representing the top 0.005% of all players.

Table 2: Performance of model relative to benchmarks.

Manager	Points Earned
Top 1	1,046
Top 100	1,002
Our Model (Rank: 463)	986
Top 1K	979
Top 100K	916
Dilan SriDaran (MIT MBAn)	880
Martin Bogaert (MIT MBAn)	867
Top 1M	864
Top 2M	839
Top 4M	798

It is noted that our approach has not yet played any of its chips. By contrast, 81% of managers ranked ahead of us have played one or more chips (Appendix B), the use of which has inflated their performance relative to our model. Early chip usage allows some managers to dominate the early-to-mid season rankings, but in many instances is not a viable strategy for long-term, sustained success. There are only 86 managers globally that have not yet played a chip, yet have accumulated more points than our model.

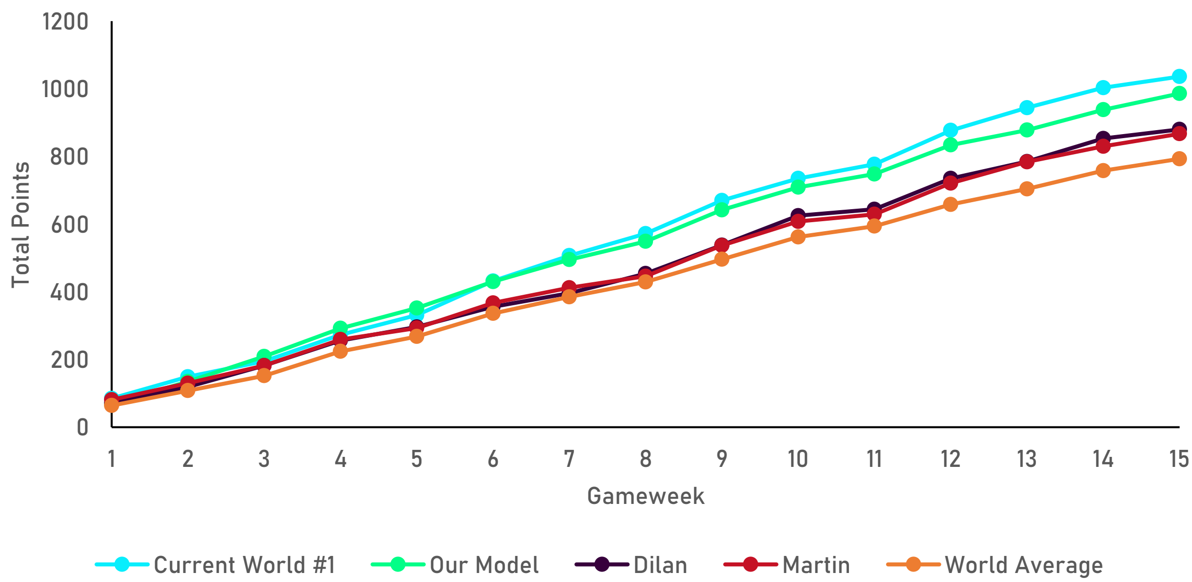


Figure 2: Total points earned by manager by Gameweek, up to and including Gameweek 15.

4.2 Strategic Insights

A comprehensive breakdown of all model decisions is provided in Appendix A. A summary of notable elements of our strategy are outlined below, which helps to understand our strong performance:

1. **Balancing core squad of high-, medium-, and low-cost players:** Our approach typically picks the same core of players each week, with a number of players selected across the whole season. These include high-cost and popular players, such as M. Salah (£12.5) and E. Haaland (£14), medium-cost players such as B. Saka (£8.5) and O. Watkins (£8), and low-cost options such as J. Ward Prowse (£6) and B. Mbeumo (£7). Whilst the high- and medium-cost players mentioned are typically highly popular options for most managers, J. Ward-Prowse and B. Mbeumo were less popular, being selected by just 10% and 1% of managers respectively at the start of the season. This demonstrates the ability of optimization to find "hidden gems," by balancing price and cost.

2. **Alternating captain strategy:** Most human managers tend to select the same captain every week, with E. Haaland – last season's top scoring player – being the most popular option. By contrast, our model typically alternates captaincy between the two best players, E. Haaland and M. Salah, based on which player has the more favorable fixture for that week.

3. **Emotionless decisions:** Most human managers tend to exhibit strong recency bias. For example, E. Nketiah – an initially relatively unpopular player – scored a huge 17 points in Gameweek 10. Over 0.5 million managers subsequently transferred him into their squads after that week, yet he scored just 4 points in total in the subsequent 3 weeks. By contrast, our approach demonstrates patience and long-term thinking, and persists with players even if they have one-off poor matches.

4. **Emphasis on goalkeeper and defender transfers:** Most human managers tend to use their transfers on headline-grabbing midfielders and forwards. By contrast, our approach tends to pick the same midfielders and forwards each week and instead emphasizes using transfers to target goalkeepers and defenders with favorable fixtures. This makes intuitive sense as goalkeepers and defenders score the bulk of their points from clean sheets (which are team-based actions), whereas midfielders and forwards score their points from goals and assists (which are individual-based actions). Therefore, good forwards will likely be able to score many points regardless of their opposition, whereas goalkeepers and defenders are only likely to score highly when they play worse opposition.

5. **Long-term transfer planning:** Most human managers tend to plan transfers only a couple of weeks ahead, and typically need to take multiple transfer penalties to shape their squad each week. By contrast, our approach uses optimization to create a globally optimal strategy that considers the totality of the season. We are therefore able to create strong squads each week, navigating the entire season without having to make a single penalizable transfer. A notable example of this is transferring in J. Trafford into our squad in Gameweek 3, despite only "needing" him to be in our team from Gameweek 5 onwards, as there were no better transfers to make in Gameweek 3 (refer to Appendix A). Another example of strategic transfers was transferring in C. Morris into the team for a single week in Gameweek 7, as his team, Luton, had 2 fixtures in that week.

5 Conclusion

5.1 Impact

We have introduced a mixed-integer optimization model to automate FPL team management. This framework overcomes typical human biases and prioritizes long-term strategic planning and efficient budget allocation. This performs exceptionally strongly, achieving a rank of 463rd out of over 10 million players as at 8 December 2023.

Our findings not only provide a novel perspective on sports analytics and optimization in fantasy sports, but also pave the way for future enhancements in broader sports management. Many aspects of FPL have parallels in real-world sports, for example FPL's budget cap translates to real-world salary caps. Another example is FPL's transfer penalty system, which could be likened to real-world player transfers. There are usually short-term teething costs when teams make multiple transfers at a point in time as players need time to build chemistry, however these could potentially build long-term benefits if performed correctly.

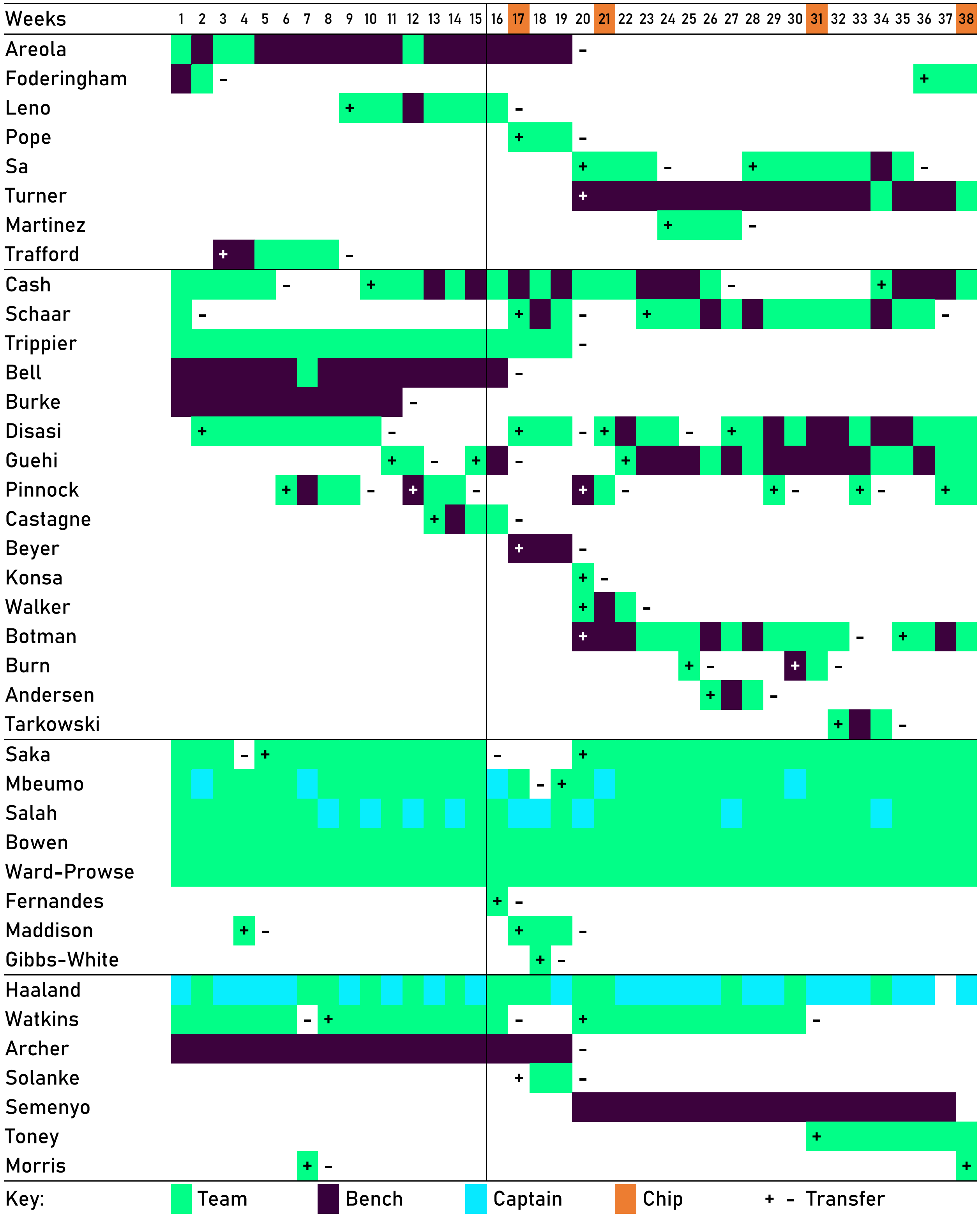
5.2 Limitations and Next Steps

The current optimization framework relies on externally sourced expected points data that were not robustly validated. We have noted some unusual results within this data and believe that our framework could be improved by constructing our own independent and transparent points-prediction engine.

Further, a key limitation of the current approach is that it relies on static predicted points set at the start of the season to determine a transfer strategy for the entirety of the season. This is clearly seen in Figure 2, where our model actually outperforms the manager currently ranked first in the world over the first five weeks of the season. However, over time as player form and injuries develop, the predicted points we use become less "relevant," and this manager, who has access to real-time information, starts to outperform our model. Our model could be enhanced by developing our own points-prediction engine that could be re-run each week based on all historic information up to that point. We could then iteratively re-solve our strategy each week, using a starting team defined by the solution from the prior week, but using updated points expectations for the remainder of the season. This approach would allow the capture of useful information, such as unforeseen player form/performance, injuries, and suspensions.

Appendix A Model Solution

Table 3: Full breakdown of model strategy by week, with players segmented by GK, DEF, MID, and FWD.



Appendix B Chip Usage

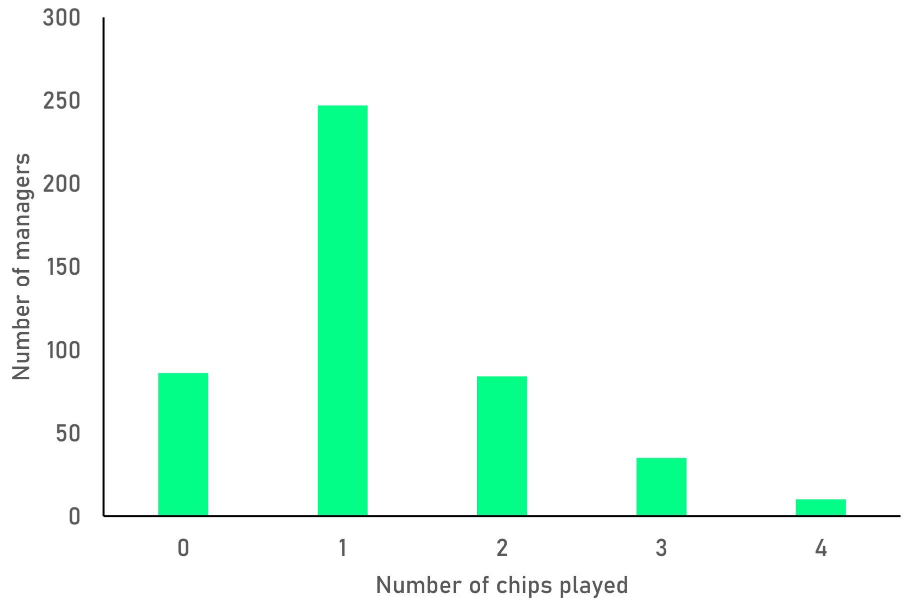


Figure 3: Distribution of chip usage for managers ranked ahead of our model as at 8 December 2023.