Integration of Forecasting, Scheduling, Machine Learning, and Efficiency Improvement Methods into the Sport Management Industry

Caleb J. McGrady

A Senior Thesis submitted in partial fulfillment
of the requirements for graduation
in the Honors Program
Liberty University
Spring 2020

Acceptance of Senior Honors Thesis

This Senior Honors Thesis is accepted in partial fulfillment of the requirements for graduation from the Honors Program of Liberty University.

____
Andy Ham, Ph.D.
Thesis Chair

____
Frank Tuzi, Ph.D.
Committee Member

____
James H. Nutter, D.A.
Honors Director

____
Date

Abstract

Sport management is a complicated and economically impactful industry and involves many crucial decisions: such as which players to retain or release, how many concession vendors to add, how many fans to expect, what teams to schedule, and many others are made each offseason and changed frequently. The task of making such decisions effectively is difficult, but the process can be made easier using methods of industrial and systems engineering (ISE). Integrating methods such as forecasting, scheduling, machine learning, and efficiency improvement from ISE can be revolutionary in helping sports organizations and franchises be consistently successful. Research shows areas including player evaluation, analytics, fan attendance, stadium design, accurate scheduling, play prediction, player development, prevention of cheating, and others can be improved when ISE methods are used to target inefficient or wasteful areas.

Keywords: sport management, ISE, forecasting, scheduling, machine learning, efficiency improvement

Integration of Forecasting, Scheduling, Machine Learning, and Efficiency Improvement Methods into the Sport management Industry

The Institute of Industrial and Systems Engineering (IISE) claims, “Industrial and systems engineers make things better in any industry” (About IISE, 2020, para. 1). One industry it can be immensely beneficial for is sport management, and although the benefits are significant, the current utilization of ISE methods is limited in many organizations. Sports is a multibillion-dollar industry, and it only makes sense to explore the benefits of using ISE methods to maximize the potential of sport management industry.

This research studies the effects of four integral ISE methods on the tasks involved in the sport management field. Methods of forecasting, scheduling, machine learning, and improving efficiency will be studied to determine what benefits may be reaped. In industries where ISE has been implemented, results include increased efficiency, better quality, and decreased waste. In the sport management where competition is fierce and resources are limited, the integration of ISE methods may be the difference between success and failure.

Integration Methods

ISE has the ability to make a direct impact on almost any industry. This is due to the diverse set of skills, tools, methods, and techniques that have been developed. While the integration into sport management may seem strange at first, there are actually many methods of integration that provide opportunities for significant change and improvement.

Forecasting

The first method, forecasting, involves the collecting of data, modeling of data, and predicting of future results. Many methods can be applied, and the most accurate method will vary based on the structure of the data including the presence of trends or cyclic nature. The best method will also depend on the whether the desired forecast is short term or long term. Forecasting is a valuable method with widespread uses for planning and decision making. While this application is mainly utilized in manufacturing and logistics organizations, if applied correctly could be greatly beneficial in sport management and sports operations as well.

Player evaluation. One way to introduce forecasting is the continual evaluation of players. This mirrors one of the central tenants of lean manufacturing, continuous improvement, which is the attempt to always be looking for new or better methods to achieve the same level of quality and performance while reducing cost and/or waste. One of biggest areas of player evaluation is player statistics. Each sport has different statistics that measure greatness, but it is true for all sports that the best players produce quality results each year. Each year massive contracts are shelled out during the offseason, while most of these contracts end up paying exorbitant amounts of money for results that are less than stellar. It would be beneficial for the upper level management personnel of teams to forecast contribution rather than making contract offers based on the most recent season's performance, the player's name, or the player's legacy. Many times, a viable option can be found to provide similar impact for a lower contract. The Oakland Athletics of Major League Baseball (MLB) are an example of an organization that has applied this method successfully in their famous implementation of Bill James' "Moneyball" approach (Wassermann, 2005). In a research study completed by Douglas Hwang (Hwang, 2012), an attempt was made to predict the next year's impact of National Basketball Association (NBA) players who were to become free agents. Regarding the potential impact of an accurate forecast of player performance, Hwang (2012) suggested:

> NBA franchises in many ways "risk" their future on certain players; this includes monetary commitments, salary cap restrictions, impact on overall franchise value, and

> team marketability. Instead of determining player value based on an individual players' past results (scoring average, eFG%, TS%, PER, Win Shares, +/-), it would be useful to determine player value by projecting future performance. This is especially important when examining long-term player contracts, and negating short-term bias in the statistically significant "contract" year performance. (p. 1)

He achieved significantly accurate results, shown in Table 1, by fitting the parameters of a Weibull Distribution to a gamma distribution (Hwang, 2012). He makes this remark regarding the output of this study in terms of importance to NBA franchises:

> Using probability models to forecast the performance of NBA players is an excellent way to evaluate the future "portfolio assets" for the team, franchise, and business. Investing in NBA players is expensive, and instead of determining value from the "gut," or from what the "market" thinks, the value should be determined by using a combination of quantitative analysis with the right qualitative analysis. (Hwang, 2012, p. 6)

Table 1: *Projections vs Actual for 2010-2011 season for seven 2010 summer free agents.*

	Actual Pts/Game 10-11 Season	**Pts/Game Model Prediction**	**Difference**
Lebron James	26.7	26.2	-0.5
Dwyane Wade	25.5	24.1	-1.4
Chris Bosh	18.7	20.2	1.5
Dirk Nowitzki	23.0	22.4	-0.6
Joe Johnson	18.2	20.1	1.9
Amare Stoudemire	25.3	26.5	1.2
Carlos Boozer	17.5	16.4	-1.1

Each year, teams go through a cycle of player evaluation which can be summarized into four main phases that use player evaluations in slightly different ways. These are contract negotiation, free agency, new player draft, and trades. The cycle generally follows this pattern although there can be overlap between the timelines of each phase.

Contract negotiation. The contract negotiation phase begins as soon as a team's season is over. In this phase, teams have the ability to interact with players currently on their roster in ways such as restructuring contracts, offering contract extensions, and offering new contracts. This is the team's chance to retain the talent they already possess but is also one of the most critical and costly periods of the offseason as teams must operate within the salary cap of their respective league and offering contracts to players who are about to become free agents can become absurdly expensive. To be most effective, teams should invest in forecasting prior to this period to determine which players should be prioritized to keep and which can be let go. Some sports, such as the MLB, also have arbitration periods to settle contract negotiation disputes. This would be a prime focus area for forecasting to provide justification for the amount of compensation that is being offered. A mix of short term and long-term forecasts are necessary, while trends are likely to also play a large role in determining future value.

Free agency. The next phase, free agency, begins at some point towards the end of the contract negotiation phase. Each sport is different, but most have a set date that teams can start contacting free agents. If teams implement the forecasting element into their roster plans, they would have the knowledge of which players they should target based on predicted value and contribution as well as the maximum compensation they should offer. Most forecasting in the free agency phase would be long term as most players who reach free agency are looking for mega contracts while trends would become a major factor.

Draft. The most critical phase, the draft, for teams to become or remain elite several years down the road. Teams have one chance each year to draft new players for their team and making the correct picks is often the difference between competing for championships or not. While players in the draft would be the most difficult to forecast due to factors like limited statistics and different competition levels in college, forecasting can still be used to determine what positions the team will need in upcoming seasons and which player should be picked. The draft phase would use a mix of short term and long-term forecasts when determining which players will excel and which positions need upgraded.

Trades. The trade phase can overlap with any of the other phases and provides teams with the ability to instantly make changes in personnel. Most trades are designed to give one team an immediate boost in performance while providing the other team with resources to improve at a later time. This phase has the potential to use both short term and long-term forecasting in the same action. The team looking for performance increase would need a short-term forecast whereas the rebuilding team would need a long-term forecast to determine what trade compensation presents enough value. This phase also provides a significant amount of intrigue as most trades that make a team better can end up crippling the team in the future when they give up future draft picks and/or young player talent. The ability to accurately forecast the performance that will be gained as well as what will be lost gives team leadership the insight to determine if the trade is valuable enough.

Analytics. In recent years, sport management personnel have begun to notice the effects analytics can have if utilized correctly. Statistics such as Wins Above Replacement (WAR) in the MLB and plus/minus in the NBA, National Hockey League (NHL), and Major League Soccer (MLS) are designed to try and show individual player contribution and effect when they are in the game in contests involving team sports. While it is debated if these statistics are all-inclusive or merely supplemental in determining which players have the greatest impact, it can be agreed that statistics such as these can provide valuable insight into the effect a player has on the game and their team. Martijn Hosten (2017) says this:

> Imagine the possible gains by making game-changing decisions based on data instead of gut feelings. And what if athletes could extend and improve their career by predicting injuries and preventing them by a small tweak in their training programs? The answer to all of this is sports analytics. (para. 2)

If forecasting was able to be used in juncture with these statistics before and during competition, many sport management decisions could be impacted and improved which might affect the outcome.

Organization decisions. For sport franchises, players are a valuable and limited resource. Often, decisions are made concerning the usage and utilization of this resource. Unforeseen circumstances such as injuries could be catastrophic if there does not exist a plan to alleviate this risk. Hosten (2017) mentions the Sports Performance Platform of Microsoft as a tool designed for this purpose and says this of its purpose:

> It’s a tool that includes machine learning and AI in its algorithms to come up with data-driven decisions for athletes and teams. In short, it provides solutions for the locker room, performance lab and the side-line. The platform contains algorithms to prevent injuries, make game-changing decisions based on each athlete’s game-day availability and alter the training scheme to keep athletes performing. For instance, the platform predicts the recovery time while taking into account factors like distance sprinted, temperature,… (para. 3)

He also presented the Danish football (soccer) club, FC Midtjylland, as an organization that has been drastically changed after adopting a predictive analytic approach where they went from being close to bankruptcy to winning their first championship title (Hosten, 2017). He has this to say about their new approach:

> Each player is now analyzed twice a month and receives an individual training plan. Analytical experts provide the coach with insights during half-time to alter the game plan based on in-game statistics. The management now also uses insights from analytical models that suggest new players. They thus completely changed their approach from using their heart to using their brain. (Hosten, 2017, para. 4)

It is situations such as this where methods stemming from ISE can have game changing impacts on an industry that is still warming up to the idea of change.

Coaching decisions. Next, examples of impact contributed to predictive analytics usage during contests are also prevalent. Coaches often have to make split-second decisions that can determine the outcome. Having predictive analytics available to help make these decisions can be the difference between winning and losing in these moments. In the MLB, it is becoming popular for managers to call upon "specialist" pitchers to come in and get one out or make position player substitutions in key situations due to predictions and statistics that say there are high chances of success. Hosten mentions other sports where predictive analytics is helping make decisions within games. He mentioned IBM's "SlamTracker" technology for tennis that can predict match winners based on tendencies in certain situations, predictive models in the NBA that can determine if a player will try to score or pass in a given situation, and AI and analytical data that is used in Formula 1 to make race-changing decisions (Hosten, 2017). As predictive analytics technologies become more advanced coaches will lean on them even more to make the most difficult decisions.

Fan attendance. Another area where forecasting can be applied to maximize profit is fan attendance. Using past attendance statistics, it is possible to forecast, within a reasonable amount of confidence, the expected attendance for future seasons. While attendance has many determining factors (e.g. team success, weather, etc.) knowing a reasonable expectation can allow for greater revenue retention. The attendance expectation will affect the organization's gameday operations. For example, an accurate attendance forecast can determine the optimal ticket prices, number of concession vendors and staff, commitment from sponsors for advertisement profit, frequency of souvenirs and promotions, and many other crucial management decisions. Another area where accurate attendance forecasts become useful is the construction of a new stadium. Old stadiums are constantly being renovated or demolished in favor of new stadiums. Knowing an accurate attendance prediction several years into the future can be massively impactful in the location, size, cost, and expected profit of a new stadium as well as the necessary characteristics that need to go into the stadium design.

In a study, this author proposed to forecast the annual attendance of a recently formed MLB team using attendance data for all MLB teams dating back to 1890. A comparative study between linear regression and dynamic regression forecasting using ARIMA was chosen. The data was formatted as a time series and used analytics to confirm an expected trend in the data, shown in Figure 1. The data followed a linear trend indicating the potential to obtain an accurate forecast. Then, the Colorado Rockies were arbitrarily chosen, and the linear regression and dynamic regression were calculated. The regressions were based on the trend component, the average MLB attendance and the attendances of the four closest teams by market size.

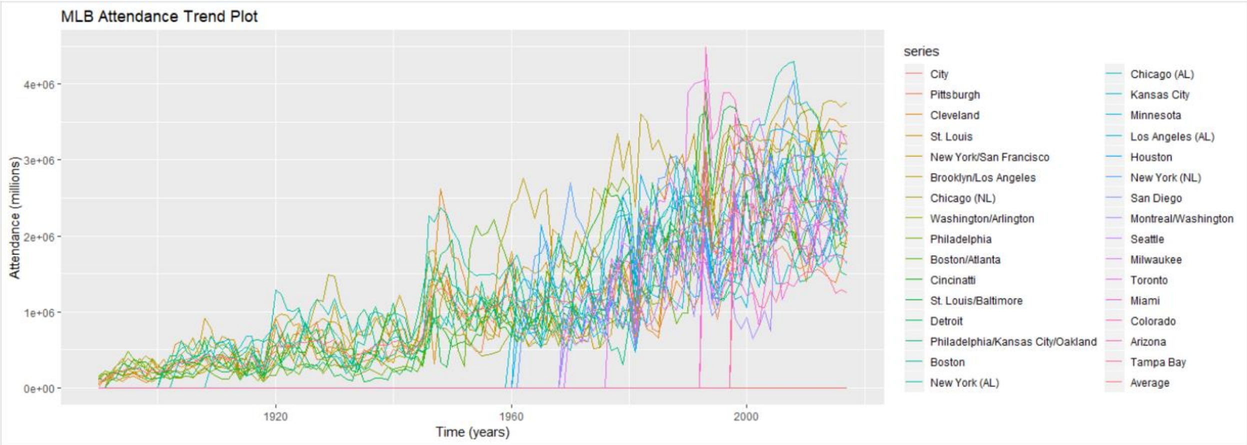

Figure 1: Trend plot of MLB team annual attendance (1890-2017).

Next, this author analyzed the residuals of the linear regression and dynamic regression models and determined the linear regression model (Figure 2) was a better forecast for this data set than the dynamic forecast. (Figure 3) It was concluded it would be possible to forecast annual attendance to an extent for future seasons, although this author predicted teams with more years

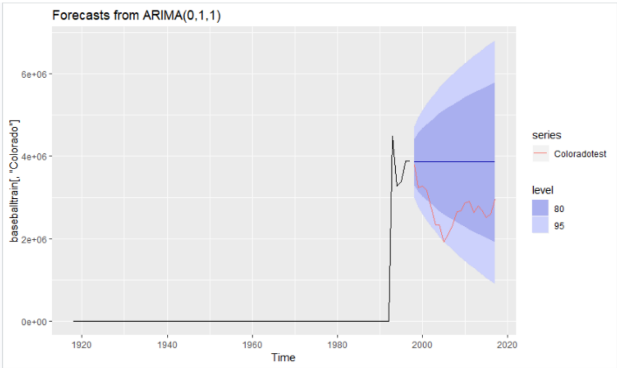

Figure 2: Linear regression forecast for Colorado Rockies.

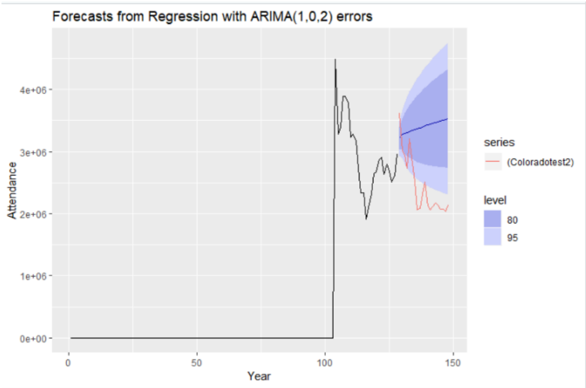

Figure 3: Dynamic regression forecast for Colorado Rockies.

of attendance data would receive better results after comparing the forecasts for Colorado and the forecasts for the average MLB attendance, using the same methods of linear regression and dynamic regression (Figure 4, 5).

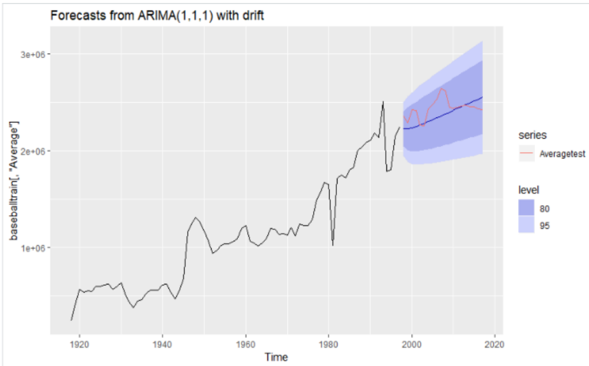

Figure 4: Linear regression forecast for average MLB attendance.

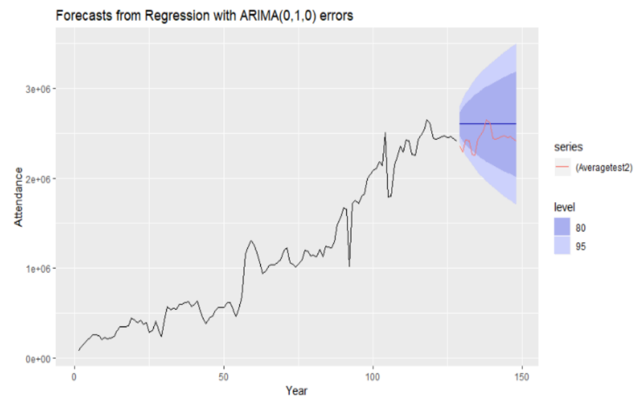

Figure 5: Dynamic regression forecast for average MLB attendance.

Ticket prices. Using an accurate forecast of fan attendance such as these methods presented has the potential to help sports organizations to capitalize on a main source of revenue, ticket sales. As anyone who has attended a sporting event can attest, the prices of tickets can vary significantly at any given time. With all of this variability, it would be impossible to estimate the annual revenue that would be generated from ticket sales. Although, an accurate model that can continually predict attendance numbers based on data gained after each game such as team record, previous attendance, fan feedback, and others is a conceivable option in the future that could provide some consistency and confidence in revenue to be generated. In some ways, this concept has already been adopted and is being used successfully through dynamically priced tickets which works similar to how industries like rental cars, hotels, and airlines set their prices. In an article published on Forbes, Patrick Rishe (2012) says, "In the ever competitive sports industry where teams constantly explore ways to maximize their revenues streams, it should not come as a surprise that we should expect dynamic pricing in sports to take off" (para.2). He also presents some takeaways he had after speaking with sources from the sports industry who are involved in dynamic pricing. He says:

- There is greater fan acceptance of real-time pricing in sports.

> - Ticket pricing technologies have advanced to the point where it has become logistically more efficient to implement dynamic pricing in sports.
> - Dynamic pricing boosts the revenue maximization goals of sports organizations.
> - Dynamic pricing incentivizes consumers to purchase season-tickets in order to secure a greater sense of price certainty in the face of real-time pricing.
> - Dynamic pricing allows consumers the flexibility to acquire significant savings on low demand games. (Rishe, 2012, para. 4)

As to the growth of dynamic ticket pricing, Patrick states the market leader, Qcue, went from one professional sports team using their services to over thirty teams just three years later, and he mentions teams using this service can see revenue increases ranging from 5-30% depending on the demand (Rishe, 2012). Rishe (2012) concludes with:

> the feedback from sports industry executives is that dynamic pricing allows them to more accurately price both high-demand and low-demand games, generating more revenue for the organization while creating savings opportunities for season-ticket holders and added purchasing flexibility for all consumers. (p. 29)

As this example shows, there can be significant revenue gains when analytics and forecasting are used to set ticket prices as well as the potential benefit of an enhanced home field advantage if more fans are attending the game in person.

Concession vendors. Another area teams may find benefit from an accurate fan attendance forecast comes in concessions. Concessions pull in a majority of gameday revenue if ticket sales are left out. Knowing how many fans are likely to attend any given game can help teams plan what types and how many concession vendors should be available. If analytics are brought into consideration, the concession vendors available can be tailored to the correct demographic of fans. An example of how analytics are being implemented is presented by Jaime Zuluaga. He writes:

> Many sports venues are now implementing data collection through concession factors like digital menu boards for dynamic pricing, heat map technology to analyze where to cut down on wait times, and mobile apps for in-seat ordering options. Data like this helps managers understand if their prices are in line with attendee demand, peak times of the game, and which menu items drive profitability. (Zuluaga, 2017, para. 4)

Current analytics usage in the concessions area is limited, but examples like this show the possibility for concession optimization in the future.

Souvenirs and promotions. A similar analytical approach to concessions may be used for souvenirs and promotions. Zuluaga suggests pulling data on the past game's apparel and souvenir revenue, the highest selling items by player or position, price ranges yielding the highest purchase volume can guide future merchandise shipments and allow for the optimal number of employees to be placed at each merchandise stand (Zuluaga, 2017). He also says data like social media mentions, how many attendees leave before the game ends, which rival teams tend to drive up ticket prices, and tracking the team's mobile app engagement during games offer a few ways to find opportunities for promotional giveaways, fan-to-play interaction, and coupon offers (Zuluaga, 2017).

Stadium design. One cost that can become a major burden for sports organizations is the construction or renovation of their stadium or facilities. If not properly prepared and planned for, there is a high opportunity cost that may result such as lost revenue from tickets and the inability to host championships or other events. An accurate forecast of revenue, fan attendance, and opportunity cost can provide the rationale for making the decision to expand or renovate. This information is critical when the team does not own their stadium and has a contract with the city to use the stadium. They must then rely on the city's willingness to use taxpayer funding for the proposed capital expenditure. One example of this is the St. Louis Rams moving to Los Angeles in 2015. Adam Crane (2016) presents the three reasons for the move being a new stadium, fan involvement, and a larger market. All three of these can be attributed to the team having a forecast for the future.

Assuming an organization has obtained the ability to begin construction or renovations on a facility, there are key areas where accurate forecasts can be applied with analytics to maximize the potential of the new facility. These areas are often overlooked by many people but could be the difference between constructing a stadium that exceeds expectations or one that lives in infamy.

Parking. A main goal of industrial engineering is optimization, and parking is an area where great enhancements can be made when constructing a new facility. Using an accurate forecast of fan attendance and an estimate of how many parking spaces are needed with the space available for parking can allow for the best parking structure layout to be achieved. The formula to calculate the size for one space is $PW = \frac{SW}{Sine\ \theta}$ where PW is parking width, SW is stall width, and θ is the angle of the parking space (James A. Tompkins, 2010). The known parking dimensions and required aisle space can then be used to determine the maximum number of spaces for a proposed layout.

Seating. A primary goal in the design of a stadium is to maximize the number of seats that will be filled. In other words, it does no good to build a stadium that holds 100,000 people if there will generally only be around 20,000 in attendance. This goal accomplishes two main desires which are maximizing revenue and creating a distinct home field advantage. On the topic of determining the number of seats for a stadium design, Jim Swords (2013) says:

> Facility owners should have a total seat count in mind well before design work begins on new construction or a renovation. This should not be an arbitrary number picked out of thin air, which actually does happen. Ideally, seating capacity numbers should be thoroughly researched and derived from a market feasibility study conducted by the university or athletic department and carefully tailored to the primary team tenant and its audience. (para. 9)

Although, the method for determining the optimal number of seats in a stadium is not perfect. Swords (2013) addressed this fact:

> No single one-size-fits-all formula or mathematical equation will allow you to easily determine the seating needs of an arena; it's something that requires extensive research and an intuitive sense of the current and projected fan base, as well as the desire to think creatively about the type of seating provided. (para. 13)

Therefore, it follows reason that using a forecasting method along with analytics of past seasons and similar teams as part of the market feasibility study has the potential to provide for the most accurate seating count that should be allotted in the design plans.

Facilities. Once the number of seats has been determined, it is important to determine the layout and number of facilities such as restrooms, dining options, exits, etc. that should be available. This can be accurately estimated using the forecast for fan attendance and guidelines for facility planning in a similar method to parking space optimization. Table 2 shows the guidelines presented in Facilities Planning, 4th Edition regarding restrooms in large facilities (James A. Tompkins, 2010). Other guidelines exist for the various facilities that may be needed, and an accurate forecast can help sport management design the optimal stadium by limiting the space needed for facilities.

Table 2: *Facilities planning guidelines for restrooms in large assembly designs.*

Assembly, Other Than Religious, and Schools					
Water Closets	Occupants	Urinal	Male Occupants	Lavatories	Occupants
1	1–100	1	1–100	1	1–100
2	101–200	2	101–200	2	101–200
3	201–400	3	201–400	3	201–400
4	401–700	4	401–700	4	401–700
5	701–1100	5	701–1100	5	701–1100
One additional water closet for each 600 occupants in excess of 1100.		One additional urinal for each 300 occupants in excess of 1100.		One additional lavatory for each 1500 occupants in excess of 1100. Such lavatories need not be supplied with hot water.	

Scheduling

The next ISE method, scheduling, is an important aspect of an organization's sport management department. It is a complicated process as the team's schedule is constructed. A significant number of constraints are used in this process to obtain quality schedules. The complexity is magnified when other events are also scheduled by the organization to increase the profit the team is able to accumulate.

Constraints and objectives. Each organization has specific constraints they must meet when constructing a schedule. Each sport is different in the number of games and ways the basic schedule is constructed. In some leagues, such as the NFL and MLB, teams are given the opponents they must play each season based on a rotating schedule sequence, while teams in other leagues, such as the National Collegiate Athletic Association (NCAA) or high school leagues, have the freedom to choose opponents in addition to the required games against conference opponents. In cases that fall under the classification of the first example, the goal of the scheduling process is to achieve a level of fairness when schedules are compared to each other. In the case of the second example, it is more open to the teams to determine how difficult their schedule may be with potential rewards from tougher schedules. Regardless of these slight differences, there are some main constraints that apply to all schedules. First, the playing surface must be available on the day of the event. Second, there are sometimes constraints given by sponsors or broadcast companies (e.g. TV, radio, etc.). There are also some main objectives of any given schedule. For example, travel distances or times will want to be minimized, lengthy trips for successive games will want to be minimized, teams having several home games in a row should be reduced or eliminated to produce an even schedule, and in the case where multiple teams share a stadium, they cannot play at home on the same day. Other preferences may also be specified for each team (Prusty, 2017).

Scheduling methods. Creating a feasible schedule for a team comes with many variables and constraints and can be accomplished in a variety of ways. One common method is to use integer programming. There are two ways to utilize integer programming: mixed integer linear programming (MILP) and binary programming (Prusty, 2017). If integer programming does not allow for the complexity required, other methods may be utilized to different levels of success. These include constraint programming and metaheuristic approaches like simulated annealing or genetic algorithms. The correct method to use may be different depending on sport, league, level of competition, or organization.

In research this author completed, a study was proposed to come up with a constraint programming algorithm that would simulate the scheduling process used for Liberty University's intramural sports leagues. The 4v4 Flag Football league for the spring season of 2019 was arbitrarily chosen with the goal of producing a feasible schedule based on the constraints specific to the intramural sports season. IBM OPL CPLEX, an optimization programming language, was used to perform the simulation. Specific constraints applied in this simulation were 12 teams playing a combined total of 84 games throughout 7 weeks, each team playing 3 teams twice, games being played on 2 nights per week with six game slots, each team playing one game per night when games are played, each team playing every other team at least once and no more than twice, and games between opponents playing twice could not be in successive days games are played. The limitations with this simulation included not assigning times to games, evening the time slots of each team's games. The resulting schedule is shown in Table 3 and serves as an example of how ISE optimization methods can be used to create viable schedules for sports operating with many complex constraints.

Table 3: *Feasible schedule for Liberty University 4v4 Flag Football intramural sports league.*

Teams (size 12)	Weeks (size 14)													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	7	12	9	6	3	2	8	11	12	8	5	3	4	10
2	5	11	6	10	7	1	9	8	5	12	3	9	8	4
3	4	9	11	8	1	12	7	6	10	6	2	1	10	5
4	3	5	10	11	9	6	12	7	8	10	11	7	1	2
5	2	4	7	12	10	11	6	9	2	7	1	8	6	3
6	11	8	2	1	12	4	5	3	9	3	7	10	5	7
7	1	10	5	9	2	8	3	4	11	5	6	4	12	6
8	10	6	12	3	11	7	1	2	4	1	9	5	2	9
9	12	3	1	7	4	10	2	5	6	11	8	2	11	8
10	8	7	4	2	5	9	11	12	3	4	12	6	3	1
11	6	2	3	4	8	5	10	1	7	9	4	12	9	12
12	9	1	8	5	6	3	4	10	1	2	10	11	7	11

Machine Learning

As technology advances, machine learning and artificial intelligence are becoming increasingly more effective and widely used. Gavin Edwards (2018) describes machine learning as "tool for turning information into knowledge" and says, "machine learning is undeniably one of the most influential and powerful technologies in today's world" (para. 1, 2). Machine learning is most effective at finding trends and relationships in industries with large and complex data. Certainly, the sports industry fits these criteria and could be significantly improved through machine learning algorithms.

Play prediction. One crucial aspect of coaching during games is calling plays. The situation and flow of the game dictate the plays called, and the uncertainty associated with the success of any given play can make it extremely difficult to make the correct decisions throughout the entire course of a game. The correct decisions can also vary from person to person based on the perception of the situation. This is a prime area to implement machine learning to help make these crucial decisions as there are tons of games and plays to look back upon and use as the dataset. Also, analytics are advanced enough to provide accurate probabilities for the expected success of any given play in a situation. One specific situation where coaches are starting to lean towards analytics for their decision making is fourth down where the team can punt, attempt a field goal, or go for it to try and keep possession. Brian Burke (2013) of the *New York Times* has developed an algorithm to determine the correct play call for any fourth down situation:

> NYT 4th Down Bot works on the premise that for most of the game, a coach should be trying to maximize the point spread between his team and his opponent – that is, scoring as many points as possible while suppressing his opponent's opportunities for points. (para. 4)

Burke presents this graphic (Figure 6) showing the disparity between actual plays called and the mathematical play call for all fourth downs since 2002 (Brian Burke, 2013). The algorithm works by calculating the expected points[1], for each situation before 10 minutes remaining in the

[1] Expected points and its application to fourth downs is not new. It was created in 1971 by former N.F.L. quarterback Virgil Carter and Robert E. Machol and has been improved and refined in various ways since, notably with the book "The Hidden Game of Football" and David H. Romer's signature 2002 paper (Brian Burke, 2013).

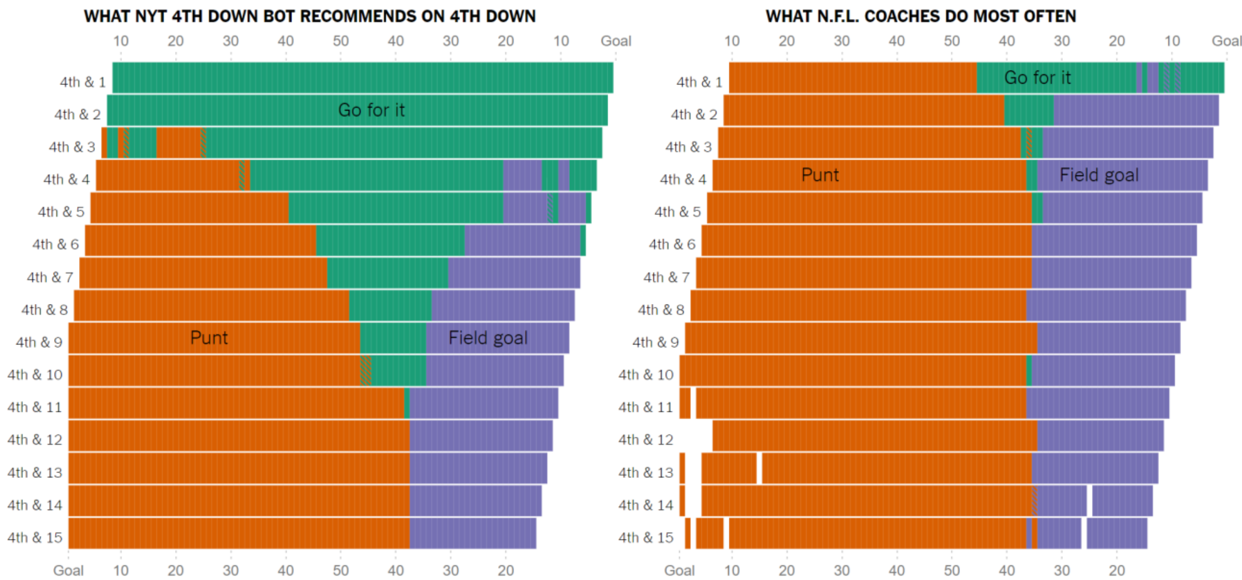

Figure 6: NYT 4th Down Bot recommendation vs. NFL coaches' tendency for fourth down play calls since 2002.

fourth quarter, and win percentage[2], for each situation after 10 minutes remaining in the fourth quarter, of each outcome and chooses the option with the highest value.[3] The algorithm shows the most disparity between analytics and NFL coaches is the decision to not go for it on fourth down, although statistics show teams are becoming more aggressive and following the analytics. For fourth-and-1 decisions, from 1994-2004 teams went for it 28% of the time, from 2005–2014, teams went for it 35% of the time, and from 2015-2016, teams went for it more than 40 percent of the time (Stuart, 2017). This trend is mirrored by a rise in overall points per season leading to the conclusion coaches should trust analytics in more situations throughout the game in an effort to be more successful.

Player development. Another possible target for machine learning usage is player development and wellness. Atif Kureishy (2019) has this to say of the potential of machine learning usage for player development and wellness:

---

[2] Win percentage measures how often teams who punted, attempted a field goal or went for a first down won the game (Brian Burke, 2013).
[3] Formula: net value = (success rate * success value) + (failure rate * failure value) (Brian Burke, 2013).

> Artificial Intelligence & Deep Learning can identify latent patterns, known as biomarkers, in EEG that change over time, and are associated with player performance and wellness. These biomarkers can be used to measure and achieve the personalized objectives of each player and team. By combining high quality data from different modalities (e.g. EEG, sleep, heart rate variability, performance statistics) and forming a player performance & wellness profile, we can identify these biomarkers and better understand causal relationships between the physical and mental aspects of sports. Pervasive Data Intelligence enables the automation of biomarker discovery through representation learning, i.e. finding higher order feature spaces that relate to specific phenomenon. By applying Pervasive Data Intelligence, a Sports Team Doctor now has a path to utilize these biomarkers per player performance & wellness profile. Artificial Intelligence & Deep Learning allow organizations to maximize player performance while minimizing player risk through better insights from 100% of performance and wellness data, at the scale that only automation can deliver. (para. 16)

The benefits could be profound, and future issues such as the chronic traumatic encephalopathy (CTE) epidemic in the retired player community of the NFL could be mitigated if signs and predictive data gave notification years in advance.

Protection against fraud and cheating. Teams are always looking for a competitive edge, but sometimes organizations cross lines by breaking established rules in pursuit of winning. Examples of cheating incidents occurring in the last few years include the Russian doping scandal spanning several years and Olympics, any of the numerous scandals involving the New England Patriots in the NFL, the recent sign stealing scandal involving the Houston Astros and Boston Red Sox, and many others. These incidents leave lasting impressions on anyone involved with the respective sport, taint the integrity and accuracy of statistics, and can lead to significant penalties and repercussions. It is possible though, that a machine learning approach could help tame the scandals by recognizing patterns associated with cheating behavior and providing the ability to stop cheating before it can cause a significant impact. Malek Murison says the World Anti-Doping Agency (WADA) plans to use artificial intelligence to detect doping incidents in sporting events such as the Olympics and other world competitions (Murison, 2018). Murison (2018) writes:

> WADA has a huge amount of data at its disposal, from biological passports to athletes' test results and activities. But modern doping methods are often so subtle that cheating is difficult to spot, even in the cold light of day. Now WADA plans to use AI's ability to spot the behavior patterns that dopers leave in their wake and flag up suspicious events – information that human observers might miss. (para. 6, 7)

According to Murison (2018), the sophisticated algorithms to be used by WADA can detect even the slightest anomalies in performance which often go unnoticed, and the knowledge provided by this artificial intelligence will allow for WADA to focus their efforts on correct individuals or teams who may be cheating although, he says a system like this should be handled with care to prevent biased targeting of individuals or teams. Applications of technology such as this will change the sports world for the better, while creating for fairer competition and greater enjoyment by fans.

Efficiency Improvement

One building block of lean manufacturing in industrial engineering is improving efficiency and decreasing waste. The same goal can and should be applied in sport management.

Improving the efficiency of how the organization is run will allow for more effort and attention to be put towards the main goals of making money and winning championships.

Data envelopment analysis. An interesting and useful method of determining efficiency is data envelopment analysis (DEA). DEA uses a linear programming approach to compare relative efficiencies of selected decision-making units (DMU) with any number of inputs and outputs by assigning appropriate weights to each input or output and developing a function to determine the efficiency. This ability to present multi-dimensional data in one efficiency metric can become greatly beneficial in selecting areas where improvement can be made to increase efficiency. In research this author conducted, a DEA approach was studied to determine the relative efficiencies of MLB teams for the 2010-2019 decade. After several iterations and attempts, five important team statistics were chosen to be used as inputs. These were market size, 100 win seasons, decade win percentage, average team wins above replacement (WAR), and decade payroll with one output comprised of a formula to calculate a score for postseason success (Table 4)[4] (Call, 2016). Due to limitations of the software being used for the DEA calculation, it was not possible to compare all 30 MLB teams at once, so only teams who had a postseason score of 15 or higher were chosen because the determination of efficiency would be predicated on success in the postseason. The results of this study (Table 5) showed Houston, Kansas City, San Francisco, and St. Louis to be the most efficient teams of the decade with Boston, Los Angeles (NL), and Chicago (NL) completing the top 7. To determine the accuracy of the results, they were compared with articles from two reputable sports journals, The Ringer and Bleacher Report. The article from Ben Lindbergh of The Ringer had Houston and San Francisco as the top two teams of the decade with Boston, Chicago (NL), Los Angeles, St. Louis,

[4] Postseason Score=Playoff Appearances+3*LCS Appearances+5*World Series Appearances+7*World Series Championships

and Washington comprising the top 7 in no particular order (Lindbergh, 2019). Zachary Rymer of Bleacher Report has the top 7 teams as San Francisco, Boston, St. Louis, Los Angeles (NL), Washington, Houston, and Chicago (NL) (Rymer, 2019). When comparing this study with these two independent articles, there is only one team they include in the top 7 that this study excluded and vice versa. Therefore, it follows that DEA could be a very beneficial tool for sport franchises to determine their efficiency when using the correct data.

Table 4: *Inputs and output data for DEA of MLB teams 2010-2019.*

Team	Market Size	100 win seasons	Decade Win Percentage	Average team WAR	Decade Payroll		Postseason Score
ARI	3,251,876	0	0.49	30.6	$886,373,096		2
ATL	4,112,198	0	0.52	32.0	$986,967,131		5
BAL	7,608,070	0	0.466	29.5	$1,051,928,720		6
BOS	5,819,100	1	0.538	43.2	$1,818,759,631		34
CHC	9,157,540	1	0.504	33.7	$1,350,717,231		25
CHW	9,157,540	0	0.459	29.3	$1,018,274,203		0
CIN	1,979,202	0	0.478	30.4	$1,005,964,957		3
CLE	2,945,831	1	0.528	39.3	$899,763,842		12
COL	2,581,506	0	0.464	29.4	$1,036,554,665		2
DET	5,456,428	0	0.483	32.9	$1,449,203,188		18
HOU	4,669,571	3	0.487	33.4	$874,385,140		30
KC	1,776,062	0	0.468	30.9	$936,762,067		25
LAA	16,373,645	0	0.507	34.9	$1,473,822,189		1
LAD	16,373,645	2	0.567	42.8	$1,822,238,150		29
MIA	3,878,380	0	0.436	22.8	$749,448,114		0
MIL	1,689,572	0	0.509	30.8	$929,598,577		9
MIN	2,968,806	1	0.471	29.5	$1,036,481,600		3
NYM	21,199,865	0	0.49	29.9	$1,161,633,572		10
NYY	21,199,865	2	0.569	44.6	$2,047,552,304		19
OAK	7,039,362	0	0.518	38.8	$725,582,299		5
PHI	6,188,463	1	0.486	28.0	$1,452,696,599		5
PIT	2,358,695	0	0.489	27.3	$739,516,946		3
SD	2,813,833	0	0.456	22.2	$751,510,767		0
SF	7,039,362	0	0.507	31.5	$1,526,419,962		49
SEA	3,554,760	0	0.468	30.0	$1,129,543,348		0
STL	2,603,607	1	0.555	39.0	$1,248,522,247		38
TB	2,395,997	0	0.531	42.4	$664,836,059		4
TEX	5,221,801	0	0.52	37.7	$1,230,561,929		21
TOR	4,682,897	0	0.49	34.3	$1,103,011,624		8
WAS	7,608,070	0	0.543	38.1	$1,297,288,059		20

Table 5*: Results from DEA of MLB teams 2010-2019.*

DMU No.	DMU Name	Input-Oriented CRS Efficiency	Sum of lambdas	RTS	Optimal Lambdas with Benchmarks			
1	BOS	0.72438	0.746	Increasing	0.512	SF	0.234	STL
2	CHC	0.56943	0.572	Increasing	0.160	HOU	0.412	SF
3	DET	0.43556	0.424	Increasing	0.115	KC	0.309	SF
4	HOU	1.00000	1.000	Constant	1.000	HOU		
5	KC	1.00000	1.000	Constant	1.000	KC		
6	LAD	0.52921	0.592	Increasing	0.592	SF		
7	NYY	0.34550	0.388	Increasing	0.388	SF		
8	SF	1.00000	1.000	Constant	1.000	SF		
9	STL	1.00000	1.000	Constant	1.000	STL		
10	TEX	0.54405	0.476	Increasing	0.097	KC	0.379	SF
11	WAS	0.48025	0.408	Increasing	0.408	SF		

Conclusion

Sport management is a daunting task with varying degrees of complexity, and because of this, the integration and utilization of ISE methodology can be greatly effective in maximizing profit and effectiveness while minimizing time invested. The usage of ISE has the potential to revolutionize how the sport management industry is operated with potentially invaluable results.

The potential impact of ISE methods is not limited to the areas of forecasting, scheduling, machine learning, and efficiency improvement included in this study. Other areas may stand to improve greatly if given the chance through integration of ISE methods.

References

About IISE. (2020). Retrieved from Institute of ISE: https://www.iise.org/details.aspx?id=282

Brian Burke, K. Q. (2013, November 28). *How coaches and the NYT 4th down bot compare*. Retrieved from New York Times 4th Down Bot: http://www.nytimes.com/newsgraphics/2013/11/28/fourth-downs/post.html

Call, B. J. (2016, July 14). *Moneyball: Which major league baseball team is the most efficient?* Retrieved from *Dallas Business Journal*: https://www.bizjournals.com/dallas/news/2016/07/14/moneyballwhich-major-league-baseball-team-is-the.html

Crane, A. (2016, February 9). *3 reasons why the Saint Louis Rams moved to Los Angeles*. Retrieved from Samford University: https://www.samford.edu/sports-analytics/fans/2016/3-reasons-why-the-saint-louis-rams-moved-to-los-angeles

Edwards, G. (2018, November 18). *Machine learning - an introduction*. Retrieved from *Towards Data Science*: https://towardsdatascience.com/machine-learning-an-introduction-23b84d51e6d0

Hosten, M. (2017, September 17). *Artificial intelligence and predictive analytics in sports: a blessing for some, a nightmare for others*. Retrieved from We Are 4C: https://weare4c.com/blog/2017-09-04-artificial-intelligence-and-predictive-analytics-in-sports-a-blessing-for-some-a-nightmare-for-others

Hwang, D. (2012). Forecasting NBA player performance using a weibull-gamma statistical timing model. *MIT Sloan Sports Analytics Conference*.

Kureishy, A. (2019, September 12). *How artificial intelligence & deep learning change the game*. Retrieved from teradata: https://www.teradata.com/Blogs/How-Artificial-Intelligence-Deep-Learning-Change-the-Game

Lindbergh, B. (2019, November 1). *Who was the MLB team of the decade?* Retrieved from *The Ringer*: https://www.theringer.com/mlb/2019/11/1/20942803/team-of-the-decade-2010s-houston-astros-boston-red-sox-chicago-cubs-san-francisco-giants

Murison, M. (2018, March 21). *Anti-doping agency to use AI to catch sports drug cheats*. Retrieved from Internet of Business: https://internetofbusiness.com/wada-ai-catch-drug-cheats/

Prusty, M. (2017, October 23). A study on the application of operation research in sports and sport management. *Medium*.

Rishe, P. (2012, January 6). *Dynamic pricing: The future of ticket pricing in sports*. Retrieved from *Forbes*: https://www.forbes.com/sites/prishe/2012/01/06/dynamic-pricing-the-future-of-ticket-pricing-in-sports/#72daf25b600f

Rymer, Z. D. (2019, December 31). *Ranking the most successful MLB franchises of the decade*. Retrieved from Bleacher Report: https://bleacherreport.com/articles/2868809-ranking-the-most-successful-mlb-franchises-of-the-decade#slide2

Stuart, C. (2017, November 10). *Smart coaches don't punt*. Retrieved from Slate: https://slate.com/culture/2017/11/nfl-coaches-are-making-better-fourth-down-decisions.html

Swords, J. (2013, July). *Seating an important component of a sports venue renovation*. Retrieved from *Athletic Business*: https://www.athleticbusiness.com/stadium-arena/seating-an-important-component-of-a-sports-venue-renovation.html

Tompkins, J. A. (2010). *Facilities Planning*, 4th Edition. Hoboken, NJ: Wiley.

Wassermann, E. C. (2005). An examination of the moneyball theory: A baseball statistical Analysis. *The Sport Journal*.

Zuluaga, J. (2017, March 29). *4 ways sport management can use data to drive sales*. Retrieved from SAP: https://blogs.sap.com/2017/03/29/4-ways-sports-management-can-use-data-to-drive-sales/