

Trabajo Fin de Grado
Grado en Ingeniería de las Tecnologías
Industriales

Análisis de datos en el deporte. Aplicación DEA a la
eficiencia en la NBA

Autora: Marta Mejías Osuna

Tutor: Gabriel Villa Caro

Dpto. Organización Industrial y Gestión de Empresas I
Escuela Técnica Superior de Ingeniería
Universidad de Sevilla

Sevilla, 2019



Trabajo Fin de Grado
Grado en Ingeniería de las Tecnologías Industriales

Análisis de datos en el deporte. Aplicación DEA a la eficiencia en la NBA

Autora:

Marta Mejías Osuna

Tutor:

Gabriel Villa Caro

Dpto. Organización Industrial y Gestión de Empresas I

Escuela Técnica Superior de Ingeniería

Universidad de Sevilla

Sevilla, 2019

Trabajo Fin de Grado: Análisis de datos en el deporte. Aplicación DEA a la eficiencia en la NBA

Autora: Marta Mejías Osuna

Tutor: Gabriel Villa Caro

El tribunal nombrado para juzgar el Proyecto arriba indicado, compuesto por los siguientes miembros:

Presidente:

Vocales:

Secretario:

Acuerdan otorgarle la calificación de:

Sevilla, 2019

El Secretario del Tribunal

A mi familia

A mis maestros

La Investigación Operativa (IO) y su relación con la toma de decisiones es un área que ha ido teniendo cada vez más importancia en el campo de la organización industrial. El hecho de contar con determinadas herramientas e información para tomar decisiones en relación a una unidad productiva concreta ha ganado interés debido a que de esta forma los beneficios obtenidos van a ser mayores que en el caso de no contar con dichas herramientas. El número de aplicaciones ha ido creciendo desde que se empezó a desarrollar esta disciplina, llegando al sector deportivo.

En dicho sector también se ha visto una evolución, puesto que cada vez son más las aplicaciones de estas técnicas en distintos aspectos de esta actividad, la cual implica un movimiento económico importante para cualquier sociedad. En uno de los deportes en los que se pueden aplicar estas técnicas es en el baloncesto, donde se va a realizar un estudio de eficiencia utilizando la metodología DEA y basándonos en los tiempos de ataque y defensa de un equipo en un partido concreto, pudiéndose implementar este estudio en la estrategia elegida a la hora de afrontar una competición.

Abstract

Operations Research (OR) and its relation with decision making has become increasingly important in industrial management. OR brings the possibility of having tools and reliable information to make better decisions related to a DMU. In case of use the advantages that OR provides you, you can get higher profits and this is what make this method increasingly interesting for most of the businesses. The number of applications has increased since the first moment the method started being developed, reaching the sports field.

Inside the sports area we can also appreciate a development as the increasing of the applications in different issues of this activity, which means a significant economic movement inside the society. Basketball is one of the sports in which we can apply these techniques. In this work we are going to make an efficiency analysis taking into account defensive and offensive times of each team in each match played. We are going to use DEA to solve the problem and the objective is to get information from the anlysis of data to improve the way they play and the strategy of each team.

| | |
|--|------------|
| Resumen | ix |
| Abstract | xi |
| Índice | xii |
| Índice de tablas y gráficas | xiv |
| Índice de figuras | xv |
| 1 Objetivos | 1 |
| 2 La Investigación Operativa y el deporte | 3 |
| 2.1 <i>Investigación Operativa en el deporte: Antecedentes y tecnología.</i> | 6 |
| 2.2 <i>Empresas dedicadas al análisis de datos deportivos</i> | 9 |
| 2.3 <i>El baloncesto y el análisis de datos</i> | 11 |
| 3 Introducción al Análisis por Envoltura de Datos (DEA) | 14 |
| 3.1 <i>Conceptos previos</i> | 14 |
| 3.2 <i>Introducción al DEA.</i> | 14 |
| 3.3 <i>Modelos más representativos</i> | 17 |
| 3.3.1 Modelos con retorno de escala constante | 17 |
| 3.3.2 Modelos con retorno de escala variable | 19 |
| 3.3.3 Modelo aditivo. | 20 |
| 3.3.4 Modelos con entradas y salidas no discrecionales. | 22 |
| 3.3.5 Entradas y salidas no deseadas. | 23 |
| 3.3.6 Modelos en red | 24 |
| 4 Medida de eficiencia de los partidos de la NBA | 26 |
| 4.1 <i>Introducción</i> | 26 |
| 4.2 <i>Modelo gráfico</i> | 26 |
| 4.3 <i>Modelo matemático</i> | 32 |
| 5 Análisis de resultados | 35 |
| 5.1 <i>Eficiencia media. Gráficas generales.</i> | 38 |
| 5.2 <i>Partidos eficientes.</i> | 39 |
| 5.3 <i>Comparativa por conferencias.</i> | 43 |
| 5.3.1 Conferencia oeste. | 43 |
| 5.3.2 Conferencia este. | 48 |
| 5.4 <i>Final de la NBA. Comparativa de los finalistas.</i> | 53 |
| 5.5 <i>Análisis individual de los tiempos de defensa y ataque.</i> | 55 |
| 6 Resumen y conclusiones | 63 |
| Bibliografía | 65 |
| Glosario | 68 |
| Anexo | 70 |

ÍNDICE DE TABLAS Y GRÁFICAS

| | |
|--|----|
| Tabla 2-1. Aplicaciones de la Investigación Operativa | 6 |
| Tabla 5-1. Divisiones y conferencias de la NBA | 35 |
| Gráfica 5-1. Eficiencia media de defensa de la NBA (2017-2018) | 38 |
| Gráfica 5-2. Eficiencia media de ataque de la NBA (2017-2018) | 39 |
| Gráfica 5-3. Número de partidos eficientes en defensa (2017-2018) | 40 |
| Gráfica 5-4. Número de partidos eficientes en ataque (2017-2018) | 41 |
| Gráfica 5-5. Número de partidos eficientes (2017-2018) | 42 |
| Gráfica 5-6. Eficiencia defensiva. HOU – UTA | 43 |
| Gráfica 5-7. Línea de tendencia de eficiencia defensiva. HOU – UTA | 44 |
| Gráfica 5-8. Eficiencia defensiva. NOP – GSW | 45 |
| Gráfica 5-9. Línea de tendencia defensiva. NOP – GSW | 45 |
| Gráfica 5-10. Eficiencia ofensiva. HOU – UTA | 46 |
| Gráfica 5-11. Línea de tendencia ofensiva. HOU – UTA | 46 |
| Gráfica 5-12. Eficiencia ofensiva. NOP – GSW | 47 |
| Gráfica 5-13. Línea de tendencia de eficiencia ofensiva. NOP – GSW | 47 |
| Gráfica 5-14. Eficiencia defensiva. TOR – CLE | 48 |
| Gráfica 5-15. Gráfica de eficiencia defensiva. TOR – CLE | 49 |
| Gráfica 5-16. Eficiencia defensiva. PHI - BOS | 49 |
| Gráfica 5-17. Línea de tendencia de eficiencia defensiva. PHI – BOS | 50 |
| Gráfica 5-18. Eficiencia ofensiva. TOR – CLE | 51 |
| Gráfica 5-19. Línea de tendencia de eficiencia ofensiva. TOR – CLE | 51 |
| Gráfica 5-20. Eficiencia ofensiva. PHI – BOS | 52 |
| Gráfica 5-21. Línea de tendencia de eficiencia ofensiva. PHI – BOS | 53 |
| Gráfica 5-22. Comparativa eficiencia defensiva GSW-CLE | 53 |
| Gráfica 5-23. Comparativa eficiencia ofensiva GSW-CLE | 54 |
| Gráfica 5-24. Comparativa de tiempo de ataque real y eficiente. GSW | 57 |
| Gráfica 5-25. Comparativa de tiempo de defensa real y eficiente. GSW | 57 |
| Gráfica 5-26. Comparativa de puntos marcados GSW | 58 |
| Gráfica 5-27. Comparativa de puntos en contra GSW | 59 |
| Gráfica 5-28. Comparativa de tiempo de ataque real y eficiente. CLE | 60 |
| Gráfica 5-29. Comparativa de tiempo de defensa real y eficiente. CLE | 60 |
| Gráfica 5-30. Comparativa de puntos en contra CLE | 61 |
| Gráfica 5-31. Comparativa puntos marcados CLE | 61 |

ÍNDICE DE FIGURAS

| | |
|--|----|
| Figura 2-1. Fases de un estudio de Investigación de Operaciones. | 4 |
| Figura 2-2. Empresas dedicadas al análisis de datos deportivos. | 10 |
| Figura 2-3. Evolución del volumen de tiros triples 2001-2017. | 12 |
| Figura 3-1. Ejemplo general de una unidad productiva o entidad cualquiera. | 15 |
| Figura 3-2. Estructura interna de una DMU de un modelo DEA en red | 25 |
| Figura 4-1. Variables del modelo de la NBA. | 27 |
| Figura 4-2. Modelo gráfico. Fase 2.1. | 30 |
| Figura 4-3. Modelo gráfico definitivo. | 31 |
| Figura 5-1. Playoffs temporada 2017-2018 | 36 |

1 OBJETIVOS

Los problemas asociados a la Investigación de Operaciones (IO) y sus distintas aplicaciones han ido ganando en interés e importancia a lo largo de la historia, tal y como veremos en los apartados posteriores.

Por una parte, el objetivo de este trabajo es mostrar la importancia de las técnicas de la Investigación de Operaciones en el sector deportivo y cómo se extiende su uso por parte de organizaciones deportivas, directores deportivos y entrenadores, entre otros. Presentaremos cómo las ligas más importantes de distintos deportes cuentan con profesionales dedicados al análisis de datos y al desarrollo de nuevas tecnologías que faciliten información de los deportistas y de los equipos y nos centraremos en el baloncesto, concretamente en la NBA. A partir de los datos de una temporada concreta aplicaremos el análisis por envoltura de datos (DEA) como técnica de esta área (IO) para estudiar la eficiencia de los distintos equipos de la NBA.

Por otra parte, como acabamos de indicar, la técnica que va a emplearse en este estudio es el análisis por envoltura de datos (DEA), lo que nos va a permitir, a partir de los conceptos básicos y los modelos originales, avanzar en el modelado DEA aplicado a partidos de baloncesto.

Se ha escogido esta temática para este Trabajo de Fin de Grado debido, por una parte, al interés académico (la relación de la IO con el deporte y las múltiples aplicaciones del DEA son aspectos que motivan un creciente interés de estudio por distintas universidades) y, por otra parte, a su creciente importancia en el mundo empresarial y de la ingeniería de la organización.

El trabajo comienza mostrando el origen de la Investigación Operativa y su evolución, además del empleo de estas técnicas en distintas áreas, siendo el deporte una de ellas. A continuación se hace una descripción a lo largo del tiempo del uso de la IO en este sector, mostrando diversas aplicaciones en el mismo, lo que permite apreciar el movimiento económico que puede llegar a tener la relación del análisis de datos con el mundo del deporte. En este sentido analizaremos distintas empresas cuyo objetivo es la recopilación de información para su posterior comercialización a equipos de las ligas más importantes.

Para terminar este apartado se introduce la relación existente entre el análisis de datos con el baloncesto, concretamente con la NBA, que es la liga sobre la que aplicaremos el modelo DEA. Se mostrarán algunos ejemplos donde se ha aplicado y se indicarán referencias de algunas personas que han tenido una influencia importante en la expansión de su uso.

En el apartado siguiente se introduce la metodología concreta que va a aplicarse para este estudio, que es DEA. En él se exponen los principales conceptos y los modelos que tienen relación con este trabajo.

En el apartado 5 se explica el problema concreto que vamos a tratar y el modelo que lo resuelve. Se expone de forma detallada el análisis de eficiencia que se quiere realizar, describiendo las variables del problema y apoyándonos en lo que llamamos el modelo gráfico, para pasar a explicar la función objetivo y las restricciones del modelo matemático utilizado.

A continuación se realiza el análisis de resultados en el apartado 6. Se empieza haciendo un análisis

más general comparando todos los equipos, para terminar con un análisis individual de los mejores equipos de la temporada estudiada. Tras el análisis de resultados, en el apartado 7 se muestra un resumen y las conclusiones del trabajo.

Por último, se finaliza con la bibliografía empleada, el glosario y el anexo donde se encuentra el código del modelo que se ha implementado en LINGO14.

2 LA INVESTIGACIÓN OPERATIVA Y EL DEPORTE

En este apartado se pretende introducir el problema que vamos a tratar posteriormente, que es el estudio de eficiencia en la NBA usando DEA, es decir, una aplicación de la Investigación Operativa para un caso concreto del sector deportivo. Antes de comenzar, se va a hacer una introducción a esta disciplina, desde un punto de vista más general y, posteriormente, nos centraremos en el caso concreto de la relación entre Investigación de Operaciones y deporte, particularizando para el baloncesto, que es el objeto del presente trabajo.

En todos los ámbitos de las actividades que realizamos se deben tomar decisiones de mayor o menor importancia y éstas se pueden tomar desde diferentes puntos de vista. Hoy en día los problemas o situaciones a los que nos enfrentamos son cada vez más complejos, cada vez tenemos que tener en cuenta más variables y cada vez la competitividad en todos los campos es mayor, y por ello las decisiones que se tomen son claves para estar en una buena posición frente a los demás. Debido a esto, se comenzaron a desarrollar técnicas y nuevas disciplinas que nos permitiesen abordarlo desde una perspectiva menos cualitativa, lo que nos iba a proporcionar mejores resultados.

La Investigación Operativa o Investigación de Operaciones (IO) es una materia basada en la aplicación de técnicas analíticas avanzadas para una mejora en la gestión de los recursos y en la toma de decisiones, consiguiendo así una mayor efectividad en las operaciones de una determinada organización. Al emplear estas técnicas se pueden hacer predicciones mucho más precisas y aportar conocimiento y datos reales a las decisiones que se tienen que tomar, las cuales se apoyan así en una base cuantitativa. [Steffens, 2004]

El momento exacto en el que empezaron a aplicarse técnicas relacionadas con esta disciplina no está claro, ya que desde mucho antes del origen oficial de la Investigación Operativa como tal, se realizaron estudios y se aplicaron métodos que podríamos considerar dentro de este campo. Sin embargo, el inicio de lo que conocemos como Investigación de Operaciones se atribuye a los servicios militares durante la segunda guerra mundial. Durante este conflicto los recursos eran escasos y surgió la necesidad de asignarlos con la mayor eficiencia posible. Fue por ello por lo que en Inglaterra se le encarga a un grupo de científicos que desarrollen políticas óptimas para el nuevo sistema de detección militar conocido como radar y así poder hacer frente a la Alemania nazi. [Taha, 2004]

Tras apreciar los excelentes resultados, Inglaterra los siguió aplicando. Asimismo, estas técnicas fueron adoptadas también por Estados Unidos al unirse a la guerra y años más tarde formó un grupo de trabajo que se dedicaba a la mejora de los procesos de planificación a gran escala, el cual estaba constituido, entre otros, por George Bernard Dantzig, el responsable del algoritmo del método Simplex.

Una vez terminada la guerra, y viendo el éxito alcanzado, el uso de la Investigación Operativa comenzó a extenderse más allá del campo militar y los métodos aplicados fueron adaptados para mejorar la productividad y la eficiencia en el sector civil. Como la evolución industrial estaba muy presente, los distintos sectores comenzaron a ser más competitivos, por lo que el hecho de contar con estas nuevas técnicas para optimizar los recursos resultó de gran importancia.

La motivación por parte de los investigadores que trabajaron durante la guerra, unida a la llegada de

las computadoras, dieron lugar a la evolución y al crecimiento de esta disciplina. Al mismo tiempo que ésta se desarrollaba, las técnicas de computación progresaron, con la creación de paquetes de software que permitieron reducir considerablemente el tiempo de resolución de los problemas, los cuales eran cada vez más complejos y por tanto los modelos más completos, por lo que los algoritmos en muchas ocasiones precisaban de un número de iteraciones cuya resolución no se podía plantear sin la ayuda de los programas informáticos.

Un estudio de Investigación de Operaciones está formado por distintas partes, las cuales podemos clasificar en 6 fases (ver Figura 3-1) y que están explicadas a continuación [Castillo, 2015].

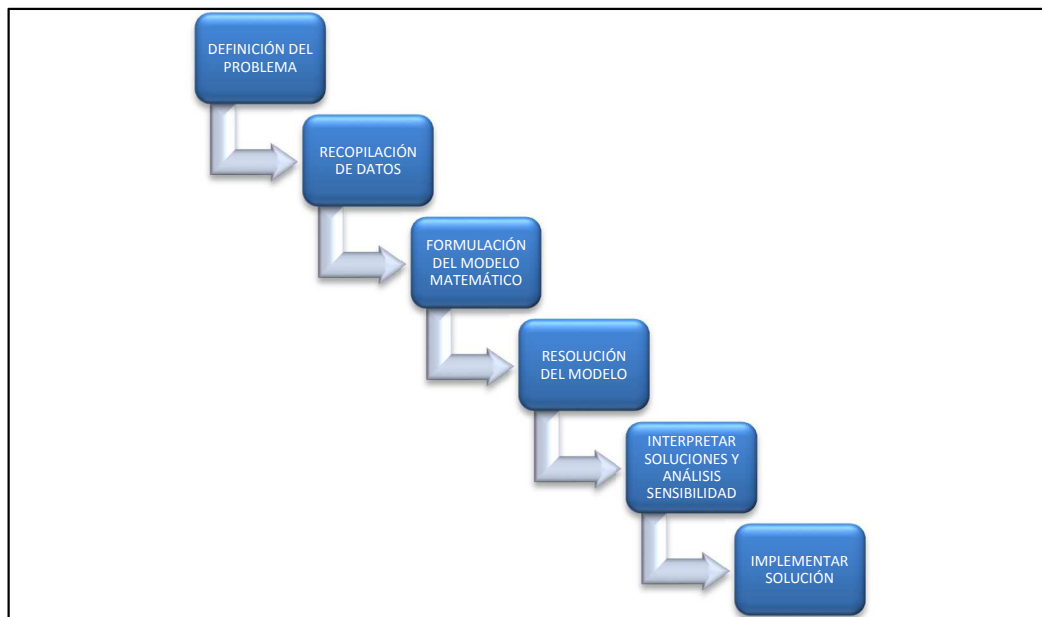


Figura 2-1. Fases de un estudio de Investigación de Operaciones.

1. **Definición del problema.** En primer lugar, la definición y descripción del problema. En esta fase se busca conocer el grado de complejidad del mismo y las variables que en él intervienen. Es importante definir claramente nuestro objetivo y el entorno en el que se desarrolla.
2. **Recopilación de datos.** A continuación es necesario la recopilación de datos relevantes para nuestro problema. Según las variables seleccionadas y la descripción anteriormente realizada, necesitaremos un mayor o menor número de datos. Es importante que éstos estén adecuadamente ordenados para facilitar las fases posteriores.
3. **Formulación del modelo matemático.** Formulación de un modelo matemático que sirva para describir el problema y se acerque lo máximo posible a la realidad del mismo. Esto implica plantear ecuaciones matemáticas que formarán la función o funciones objetivo del problema y el conjunto de restricciones correspondiente.
4. **Resolución del modelo.** Obtener soluciones a partir de dicho modelo empleando distintas técnicas y métodos matemáticos que nos permitan encontrar soluciones factibles. Para poder llegar a éstas, es importante que los datos que se hayan recogido sean apropiados para el problema en cuestión.

5. **Interpretar soluciones y análisis de sensibilidad.** Una vez hayamos obtenido alguna solución factible hay que tener en cuenta que éstas son resultados matemáticos y es necesario interpretarlos. En ocasiones es interesante realizar un análisis de sensibilidad para ver cómo se comporta el modelo modificando levemente las especificaciones. De esta forma se pueden detectar errores de modelado o en las restricciones impuestas. Es necesario ser consciente de los puntos débiles del modelo y bajo qué condiciones funciona mejor, incluso puede darse la circunstancia de que bajo algunas de ellas no sea válido.
6. **Implementar solución.** Finalmente tenemos que implementar los resultados, es decir, traducir la solución matemática final obtenida en instrucciones y pautas para aquellas personas que trabajen con el sistema, de forma que entiendan los resultados sin tener que entender cómo funciona el modelado del mismo.

Desde que comenzó su estudio, esta disciplina ha tenido un gran impacto en la mejora de la eficiencia de un gran número de organizaciones por todo el mundo, mejorando así la productividad económica de muchos países. La relevancia de esta materia puede verse, por ejemplo, en la creación de organizaciones como la IFORS (International Federation of Operational Research Societies), la cual agrupa asociaciones de Investigación Operativa de diferentes países con el objetivo de impulsar esta disciplina, patrocinando asambleas internacionales, proporcionando diferentes medios para el intercambio de información entre las diferentes sociedades, animando a los distintos países a la creación de sociedades en torno a la Investigación de Operaciones y promoviendo el crecimiento de las áreas de aplicación de la misma, haciendo mejorar las ya existentes y promoviendo nuevas. [Hillier, 2010]

Actualmente, la Investigación de Operaciones se utiliza a nivel mundial, tanto en el sector público como en el privado, además de ser un área académicamente activa. El alcance de esta disciplina puede verse en las distintas áreas en las que se utiliza y las numerosas aplicaciones dentro de cada una de ellas. En la tabla 2-1 se exponen algunas de las aplicaciones:

| Área | Aplicación | Referencia |
|--------------------|---|---------------------|
| Agricultura | Gestión de riegos | [Lyeme et al, 2012] |
| Industria | Control de inventarios a lo largo de la cadena de suministros | [Hillier, 2010] |
| Industria | Previsión de la demanda | [Lyeme et al, 2012] |
| Industria | Transporte | [Lyeme et al, 2012] |
| Industria | Mantenimiento | [Lyeme et al, 2012] |
| Industria | Programación y secuenciación | [Lyeme et al, 2012] |

| | | |
|-----------------------|--|--------------------------|
| Agricultura | Simulación para estudiar el rendimiento de los cultivos | [Hernández et. al, 2009] |
| Finanzas | Gestión de líneas de crédito y tasas de interés | [Hillier, 2010] |
| Finanzas | Desarrollar un sistema de apoyo para la toma de decisiones para asesores de inversión. | [Hillier, 2010] |
| Sector público | Gestión de rutas para la recolecta y eliminación de basura. | [Hillier, 2010] |
| Sector público | Planificar rutas para el transporte ferroviario | [Hillier, 2010] |

Tabla 2-1. Aplicaciones de la Investigación Operativa.

Además de estas aplicaciones, existen muchas otras en distintos campos. Las técnicas de la IO fueron desarrollándose, y es en 1954 cuando parece ser que aparece el primer artículo en el que se relacionaba esta disciplina con el deporte. Fue en el recién creado periódico de *Operations Research society of America* [Wright, 2009]. Mottley era un militar que encontró ciertas similitudes entre el combate en tierra y determinados deportes como el béisbol o el baloncesto. El deporte forma parte de la cultura, la sociedad y la vida de la mayoría de los países. La toma de decisiones es algo presente en la mayoría de actividades que realizamos diariamente y la práctica deportiva no es una excepción. Así, esta disciplina se usa para realizar las programaciones de las temporadas, de forma que se puedan minimizar el número de viajes y esto suponga un ahorro de recursos a la hora de tener que jugar con los distintos equipos, como puede ocurrir en deportes tales como el fútbol, béisbol o baloncesto. También es usado para la organización de los horarios, la previsión de resultados o para la toma de decisiones en el ámbito de la técnica deportiva y la estrategia a seguir. Esto último es en lo que se centra este trabajo y se irá explicando detenidamente en los siguientes apartados.

2.1 Investigación Operativa en el deporte: Antecedentes y tecnología.

Para abordar la primera parte de este apartado nos hemos basado en el estudio realizado por Mike Wright [Wright, 2009], profesor del departamento de *Management Science* de la universidad de Lancaster. Muchas de las investigaciones relativas a la Investigación Operativa en el deporte siguieron la visión de Mottley en lo que se refiere a centrarse en las cuestiones de estrategia y táctica deportiva. De esta forma, Silverman y Schwartz, en 1973 consideraron bajo qué circunstancias un equipo de fútbol americano debía perder a propósito. Por otra parte, Freeze, en 1974 hizo uso de la simulación para encontrar el mejor orden de bateo para un equipo de béisbol. Ladany, en 1975 aplicó el análisis

probabilístico para determinar la mejor altura inicial para un saltador de pértiga. Esto fue modificado en Hersh y Ladany años más tarde, en 1989, cuando las reglas cambiaron. Otros trabajos publicados en las décadas de los setenta y los ochenta sobre la táctica deportiva hablan de otras aplicaciones en distintos deportes, tales como:

- Halterofilia: decidir de la mejor manera posible qué pesos elegir (Lilien, 1970).
- Salto de longitud: a qué distancia de la línea de salto es conveniente saltar para cumplir el objetivo (Sphicas y Ladany, 1977).
- Dardos: qué números deben ser el objetivo (Köler, 1982).

Ese trabajo continuó y fue cada vez más complejo debido en parte al gran desarrollo de las técnicas de computación y a la velocidad a la que éstas crecían. Así, comenzaron a existir cada vez más aplicaciones de la programación dinámica, las cuales podían requerir un tiempo prolongado en problemas de mayor tamaño.

Muchos de estos estudios sobre programación dinámica fueron desarrollados por Steve Clarke y John Norman en cuestiones tácticas relativas al críquet (Clarke y Norman, 1999, 2003; Norman y Clarke, 2007) y en otros deportes tales como el tenis (Norman, 1985) y el fútbol australiano (Clarke y Norman, 1988). El artículo publicado por Clarke sobre los índices de puntuación óptimos en críquet, también basado en formulación de programación dinámica, fue reconocido como una de las publicaciones más importantes del JORS (Journal of the Operational Research Society) en ese año. Washburn (1991) aplicó la programación dinámica en hockey sobre hielo, a la hora de decidir el momento de sustituir un jugador de campo por un portero.

Otras técnicas aplicadas incluyen una combinación de técnicas estadísticas y de optimización. Un ejemplo de esto es el trabajo de Swartz en 2006, que usó modelos loglineales, cadenas de Markov, métodos de búsqueda metaheurística y simulación Monte Carlo para determinar el mejor orden de bateo en críquet.

Muchos de los ejemplos expuestos abarcan decisiones que tienen que tomarse en tiempo real, durante el partido. Sin embargo, hay muchas decisiones estratégicas y tácticas que se llevan a cabo a la hora de entrenar a un deportista o un equipo. De esta forma, Boon y Sieksma (2003) produjeron un sistema de ayuda a la toma de decisiones que serviría a los entrenadores y directores deportivos a la hora de saber evaluar el potencial de cada jugador y su contribución real en el equipo, de forma que sirviese a la hora de seleccionar los jugadores para un partido, al mismo tiempo que servía para buscar y seleccionar a nuevos jugadores. Cooper et al (2008) hizo uso de la envoltura de análisis de datos (DEA) con el mismo propósito, pero aplicado al baloncesto en lugar de al fútbol.

Otros estudios relacionados con este campo son los relativos al fútbol americano (Brimberg et al, 1999), baloncesto (Annis, 2006), voleibol (Lee y Chin, 2004), ciclismo (Scarf y Grehan, 2005), vela (Golding, 2002), pentatlón (Ladany, 1975) y natación (Nowak et al, 2006), entre otros muchos.

Aquí finaliza la información obtenida del estudio del profesor Mike Wright y como podemos comprobar, desde que se descubrió la posibilidad de aplicar las técnicas de la IO al deporte ha habido muchos estudios acerca de ello y cada vez se han descubierto más aplicaciones y al final podemos llegar a la conclusión de que no existe un deporte que no pueda beneficiarse de esta disciplina. Debido a esto, cada vez se ha ido obteniendo más beneficio de la Investigación de Operaciones en este sector, en gran parte por parte de los seleccionadores, entrenadores, directores deportivos y todas aquellas

personas que, a nivel profesional tienen que tomar decisiones estratégicas dentro del equipo para conseguir el mayor rendimiento posible, tanto de los componentes de un equipo de forma individual, como de todos ellos como grupo.

Cada vez fueron más los estudios acerca de las aplicaciones de la Investigación Operativa dentro del mundo del deporte y cómo la toma de decisiones es un aspecto fundamental en este sector para conseguir destacar sobre el resto. Noelia Hernández, en un artículo fechado en el 2018 en la plataforma Nobbot [Hernández-b, 2018], afirmaba la importancia de las tácticas en el deporte y describía el análisis de datos deportivos como un nuevo factor que ayuda en gran medida a conseguir los objetivos deseados, haciendo uso de tecnologías como el aprendizaje automático, la inteligencia artificial o el big data.

Una persona de referencia en este área es Michelle McKenna-Doyle, vicepresidenta ejecutiva y CIO (Chief Information Officer) de la NFL (National Football League). Es la primera persona en la historia de la NFL en alcanzar este título y es la responsable de la estrategia tecnológica [Hamburg, s.f.]. Su trabajo como CIO consiste en analizar el máximo beneficio que se puede obtener de las nuevas tecnologías. La NFL es la mayor organización deportiva americana y hace uso de una tecnología basada en una serie de sensores, colocados tanto en los propios jugadores como en el campo de juego, encargados de recopilar una gran cantidad de datos que se almacenan en la plataforma Next Gens Satats y se analizan en tiempo real. De esta forma tienen la capacidad de ofrecer una experiencia más completa a los espectadores y, por otra parte, mejora la calidad del juego haciendo uso de estos datos para analizar y mejorar el rendimiento. [Michelle McKenna-Doyle, 2017].

Otra aplicación la encontramos en el medio “Innovation Enterprise”, en un artículo donde se ocupa de la importancia de la tecnología en el ámbito deportivo y cómo ha evolucionado sobre todo en los últimos años [Oppermann, s. f.]. Éste artículo trata de cómo la tecnología ha hecho posible proporcionar información en tiempo real en muchos casos para entrenadores o científicos relacionados con el deporte. Esto ha provocado que la tecnología pase de los laboratorios al campo de juego. El ejemplo que se trata en este artículo es de un nuevo sistema denominado Adidas MiCoach que se utiliza tanto en entrenamientos como en competiciones oficiales, grabando información acerca de cada deportista, como ritmos cardíacos, velocidad, altura de saltos u otras opciones que pudiesen interesar según el deporte en cuestión. Esto es posible gracias al uso de GPS en miniatura que, junto con giroscopios y acelerómetros se implementan en un pequeño sistema que se incorpora a la ropa de los deportistas o incluso a las zapatillas. Esto permite controlar mucho mejor el entrenamiento, evitando lesiones en muchos casos y sacando su máximo rendimiento.

En relación a la tecnología en el ámbito deportivo también se puede destacar el artículo “The Evolution of Technology & Sports “, de Bellyupsports [Bellyupsports, 2018] en el que se habla del avance de las tecnologías en el mundo del deporte como un hecho que ha beneficiado tanto a entrenadores (tal y como hemos expuesto a lo largo de todo el trabajo, incluido en el análisis de nuestros propios resultados), como a espectadores, deportistas y árbitros, llevando a vivir el deporte como una experiencia más completa para todos. Entre las tecnologías y aplicaciones que se describen en dicho artículo encontramos el Photo Finish, tecnología incorporada a la ropa de los deportistas para medir diferentes aspectos que faciliten el entrenamiento, como veíamos en el ejemplo anterior, repetición instantánea, análisis de datos y estadísticas de los diferentes deportes, tecnología de radar o técnicas de imagen virtual (usada sobre todo en deportes como la natación o el tenis).

Otro artículo que podemos destacar es “La tecnología hace del deporte una disciplina científica” de el Economista [García, 2018], en el que se se trata el uso de la informática para estudiar los movimientos del cuerpo y la obtención rápida de información que permitan ayudar a la predicción y control de los resultados. En este artículo se habla de la gran dimensión social de la industria del deporte, comentada también en la descripción del problema y cómo se han desarrollado técnicas deportivas y nuevos métodos de entrenamiento basados en esta información adicional que se puede conseguir gracias a estas tecnologías, las cuales mejoran también la capacidad de gestionar la toma de decisiones por parte de entrenadores y cuerpos técnicos. Elementos como el Big Data, los wearables o la realidad virtual y aumentada, han provocado una gran transformación en la manera de ver y vivir el deporte.

En un artículo para el MIT Sloan Sports Conference sobre el análisis espacial y gráfico para la NBA se habla de la introducción de nuevas técnicas analíticas (CourtVision) con el objetivo de mejorar el baloncesto profesional [Goldsberry, 2012]. Este sistema consistía en un conjunto de técnicas analíticas diseñadas para visualizar y cuantificar aspectos espaciales de la NBA con una precisión mayor de la que se había tenido hasta entonces. Se propone una nueva forma para medir las marcas de los distintos jugadores que daban información adicional a la que se había tenido hasta el momento.

Un deporte que merece especial atención dentro de este apartado es el béisbol, ya que dentro del análisis estadístico siempre ha ocupado un lugar importante. En 1947 se contrató al primer estadístico en béisbol, y años más tarde, hacia 1980, Bill James empezó a hablar de una nueva disciplina que él llamó *sabermetrics* [Silverman, 2009], la cual se definía como la búsqueda del conocimiento objetivo del béisbol y comenzó a publicar trabajos acerca de ella. El *sabermetrics* le debe su nombre a la “Society for American Baseball Research” y consiste en el uso del análisis estadístico para determinar el rendimiento de cada jugador a partir de los datos recogidos durante un partido. Esta disciplina fue extendiéndose y en ella se basa el libro Moneyball. Este libro supuso un fenómeno importante y trata del director general del Oakland Athletics, Bill Beane, y cómo se hizo uso de este análisis estadístico para tomar una serie de decisiones que determinaron un cambio en el rumbo de este equipo dentro de la liga. Actualmente, todos los equipos hacen uso en mayor o menor medida del análisis estadístico y el *sabermetrics* ha servido de inspiración para otros deportes.

2.2 Empresas dedicadas al análisis de datos deportivos

Hoy en día el deporte en general supone un fenómeno social con una gran capacidad de convocatoria y que por tanto implica un gran movimiento económico en torno a él. Una prueba de la importancia que tiene en la sociedad es que existen estudios universitarios que abarcan distintas áreas relativas a la industria de este sector, incluyendo en algunos casos el análisis de datos y su creciente aplicación en el mismo. En universidades internacionales como las de Columbia o Georgetown y en universidades españolas como las de Valladolid, Segovia o Murcia, entre otras, se pueden encontrar cursos e incluso masters que abarcan temas de gestión deportiva y el análisis de datos y Big Data aplicado a este sector.

La industria del deporte tiene muchos sectores que pueden explotarse económicamente y aunque siempre ha existido, en los últimos años ha crecido aún más y cada vez es mayor la cantidad de recursos

económicos que moviliza y el número de personas que siguen algún evento deportivo, ya sea por televisión o en directo. Se explota principalmente en Estados Unidos, concretamente en las ligas nacionales de fútbol americano (NFL – National Football League), béisbol (MLB – Major League Baseball) y baloncesto (NBA – National Basketball Association) [Marentes s.f.].

Todo esto ha provocado que las ligas de los diferentes deportes sean cada vez más competitivas, pasando a hacer uso de las innovaciones tecnológica para ganar ventaja unos equipos frente a otros. Tal y cómo hemos visto en el apartado anterior, en muchos casos se hace uso a día de hoy del análisis de datos o de técnicas de IO para ser más eficientes y mejorar el rendimiento, y el interés en este tipo de técnicas ha hecho que se genere una industria dedicada al análisis de datos deportivos, para venderlos a los directores deportivos de los equipos y que así los entrenadores puedan hacer uso de esta información y sacar ventaja con respecto a los demás.

Actualmente existen empresas que han llevado a cabo una gran inversión en nuevas tecnologías que permiten recopilar gran volumen de datos y analizarlos, con el objetivo de comercializar con organizaciones y equipos de diferentes deportes. A continuación se muestran algunos ejemplos y a qué se dedican cada una de ellas (Ver figura 2-2).



Figura 2-2. Empresas dedicadas al análisis de datos deportivos.

- **STATS:** es una de las compañías líderes a nivel mundial en tecnología y deporte. Se fundó en 1981 y fue la primera empresa dedicada en exclusiva a unir ambas materias. Poco a poco fue creciendo hasta tener clientes por todo el mundo. Trabajan ofreciendo datos e información acerca del juego tanto a fans como a los entrenadores de diferentes equipos en cualquier deporte, con el objetivo de poder sacar ventaja respecto a los demás gracias a esta información que se les proporciona.
- **PFF (Pro Football Focus):** se centra exclusivamente en el fútbol americano. Se encargan de recoger y analizar datos para la NFL (National Football League), NCAA (National Collegiate Athletic Association) y CFL (Continental Football League), además de ofrecer este servicio a

representantes de los jugadores o incluso a empresas de comunicación.

- **GREENUPDATA:** cuenta con un gran equipo de analistas y trabaja con más de la mitad de las ligas profesionales. Se dedican al análisis de mercados deportivos, analizando a unos 1.170 equipos y a más de 23.000 jugadores por todo el mundo. Tienen como objetivo ofrecer a las distintas organizaciones la posibilidad de transformar los datos y el análisis de los mismos en oportunidades.
- **NBN23:** es una empresa española dedicada al baloncesto que hace uso de la tecnología Real Time Tracking, lo que supone una revolución a la hora de entender el baloncesto. Se consigue una gran cantidad de información a tiempo real que permite conocer datos como las aceleraciones y deceleraciones de los jugadores, además del estudio de la disposición del equipo y de cada jugador concreto, viendo su impacto tanto ofensivo como defensivo y la zona de influencia que tiene cada jugador. Toda esta tecnología consigue en definitiva vivir el deporte de una forma más completa. Esta herramienta busca el desarrollo de los jugadores de baloncesto de forma que se puedan prevenir lesiones, maximizar su rendimiento y así elevar el potencial de cada uno de ellos. Por todo esto, facilita y mejora el papel de los preparadores físicos, además de que sirve para realizar nuevas y mejores tácticas en función de los datos que se analicen.
- **SYNERGY SPORTS TECHNOLOGY:** empresa americana que a partir de una tecnología propia consigue conectar un gran conjunto de datos y fragmentos de vídeo de manera que pueda proporcionar estadísticas y análisis de los distintos equipos, jugadores y del juego que se esté desarrollando. Ofrece estas estadísticas en tiempo real, ofreciendo la oportunidad a los entrenadores de resolver cuestiones que puedan mejorar la estrategia seguida. Se encargan de seleccionar la información relevante de cada juego e intentan facilitar lo máximo posible la interpretación de los datos recogidos y analizados de cara al cliente. Equipos de la NBA, WNBA (Women's National Basketball Association), NCAA, equipos olímpicos y otras ligas internacionales hacen uso de la tecnología que les ofrece SYNERGY.

Aquí sólo se han expuesto algunos ejemplos de un mercado que está en crecimiento, puesto que los diferentes deportes cada vez son más conscientes de la necesidad de usar estas nuevas herramientas para conseguir una ventaja real sobre los demás y ser más competitivos.

2.3 El baloncesto y el análisis de datos

Tal y como hemos visto a lo largo de este apartado, las diferentes técnicas de la Investigación Operativa han hecho subir la competitividad en las ligas más importantes de muchos deportes y ha provocado que suba el nivel de las mismas. Este trabajo se centrará en el baloncesto, concretamente en la NBA.

La NBA es la liga de baloncesto más popular del mundo y también es la más importante en términos tanto económicos como deportivos. El hecho de incorporar la Investigación Operativa y el análisis de datos suspuso un cambio a la hora de jugar el baloncesto, ya que se comenzó a ver desde una nueva perspectiva.

Un ejemplo de esto es Daryl Morey, el director general del Houston Rockets, un equipo que cambió

completamente su forma de juego a la incorporación del mismo a la organización [The Economist, 2018]. Daryl Morey es un experto informático que llegó al Houston rockets en el año 2006. En él buscaban a alguien que fuese capaz de ver a los jugadores desde otra perspectiva y así, poco más de un año después consiguió ser el director general de este equipo. Ese mismo año (2006) fundó, junto a Jessica Gelman, la MIT Sloan Sports Analytic Conference, la cual es un ciclo de conferencias que tienen como objetivo hablar de la importancia del movimiento *sabermetrics* en el deporte profesional, estudiando las posibilidades que ofrece y hasta dónde puede llegar. Este evento fue ganando en importancia y popularidad dentro de este sector, a medida que fue siendo más aceptado el uso de estas técnicas de análisis de datos y de Investigación Operativa dentro del mundo deportivo.

Como podemos ver en el artículo escrito por Andrés Monje para la plataforma Kiaenzona [Monje, 2017]. La llegada de Morey supuso un cambio significativo para los Houston Rockets y el baloncesto en general. Convencido de la importancia de integrar el análisis de datos en el deporte, provocó que los Houston Rockets fuesen de los primeros equipos en instalar un sistema de localización de vídeo, con el que se conseguía extraer datos directamente del terreno de juego durante el partido para su posterior análisis. Gracias a esto, llegaron a la conclusión de la importancia de los tiros triples frente a los normales y comenzó así una nueva forma de jugar de este equipo, seguida de un aumento del número de partidos ganados y como consecuencia de su notable mejoría dentro de la clasificación general. Esto se debe a que en una liga de élite como es la NBA el hecho de contar con una ventaja así, implica una gran diferencia a la hora de tomar buenas decisiones y obtener mejores resultados. Así, poco a poco, los demás equipos adoptaron una forma de juego parecida a la de los Houston Rockets en cuanto a la importancia de centrarse en los tiros triples por encima de los normales. En la figura 2-3 se puede ver el efecto que ha tenido este cambio de estrategia en la NBA. En la gráfica de esta figura se observa cómo ha ido creciendo el volumen de tiros triples de forma considerable.

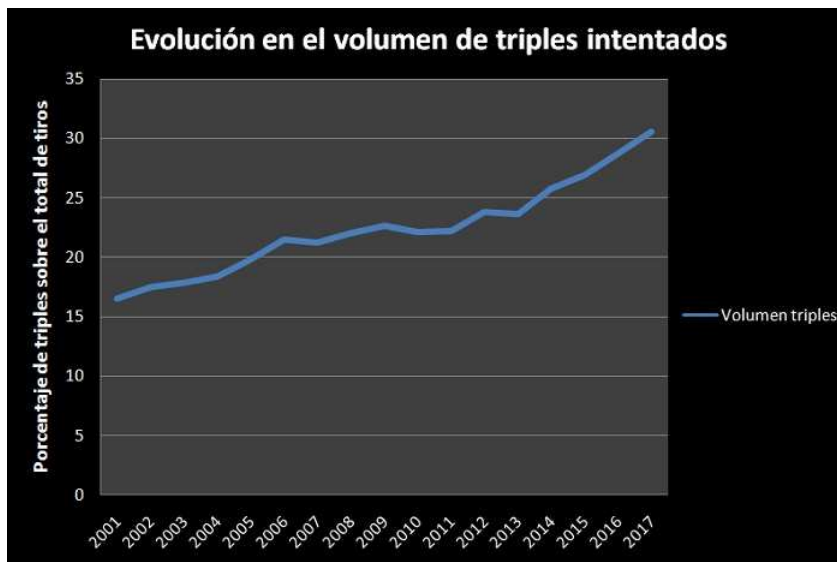


Figura 2-3. Evolución del volumen de tiros triples 2001-20017

(Ref: -NBA Stats)

Actualmente, la mayoría de equipos de la NBA cuentan con un analista de datos que trabaja con el entrenador y el director deportivo, de forma que las decisiones que se tomen con respecto a las

estrategias de juego, tanto del equipo como de cada jugador individualmente, estarán respaldadas por una serie de datos con los que anteriormente no se contaba. Así, se maximiza la productividad y se obtienen, sin ninguna duda, mejores resultados.

Además, todo este análisis de datos también permite acercar el juego a los fans y que éstos vivan el baloncesto de una forma más completa y se impliquen más, tanto con los equipos como con los jugadores, de forma individual. Así, los espectadores desde sus casas tienen acceso a todo tipo de datos y estadística casi a tiempo real y esto hace que se sientan parte del juego, teniendo la posibilidad de vivir una experiencia mucho más inmersiva, algo que antes no era posible.

El director actual de la NBA, Adam Silvers, en una entrevista ofrecida a Daniel Pink [Pink, 2017], trata la cuestión de cómo la tecnología había transformando y seguía transformando la NBA. Adam Silvers está de acuerdo en que el análisis de datos es parte fundamental de prácticamente todas las actividades que realizamos. En lo referente al baloncesto se usa tanto en partidos como en entrenamientos para mejorar el rendimiento y entre otras aplicaciones, sirve para seleccionar nuevos jugadores para un equipo. Una de las preguntas de Daniel trataba sobre las diferencias que se pueden apreciar a la hora de ver un partido debido al uso de estas nuevas tecnologías. Uno de los aspectos en los que afecta es que es posible calcular y controlar el descanso necesario para cada jugador, de forma que, por ejemplo, las grandes estrellas de un equipo son sentadas en el banquillo, siguiendo las órdenes del entrenador en ocasiones que el espectador no lo esperaba, debido a que las aplicaciones y los datos que usan como apoyo así lo indicaban.

En este apartado se ha presentado la evolución que ha tenido la Investigación Operativa dentro del mundo del deporte, concretamente, dentro del baloncesto. Además, hemos visto algunas empresas importantes dedicadas al análisis de datos deportivos y el movimiento económico que esto supone, gracias a que esta información es requerida por parte de los preparadores físicos, directores generales y entrenadores de los diferentes equipos para mejorar la eficiencia de los mismos.

Este trabajo se va a centrar en el estudio de la eficiencia de la NBA, basándonos en los resultados obtenidos en la temporada 2017-2018. El objetivo será determinar la eficiencia de los equipos en relación a la estrategia elegida en cada partido. Lo haremos teniendo en cuenta la cantidad de tiempo que el equipo ha jugado de forma defensiva o en forma de ataque. Estos tiempos se determinarán a partir de unos parámetros que se explicarán en mayor profundidad en los siguientes apartados, además de explicar más detalladamente tanto el problema de estudio como los métodos utilizados para resolverlo.

3 INTRODUCCIÓN AL ANÁLISIS POR ENVOLTURA DE DATOS (DEA)

Para resolver el problema que se ha introducido vamos a utilizar la metodología conocida como data envelopment analysis (DEA) o análisis por envoltura de datos. Se comenzará explicando la metodología de forma general, los conceptos claves y los modelos más comunes que tengan relación con el modelado posterior que se se va a hacer del problema de eficiencia de partidos de baloncesto.

3.1 Conceptos previos

Antes de entrar a describir qué es el DEA, cómo se utiliza y cuáles son los modelos más habituales, se van a definir un par de conceptos básicos:

- **Eficiencia:** la eficiencia de una organización consiste en crear un determinado producto usando la menor cantidad de recursos posibles.
- **Eficacia:** realizar una determinada tarea impuesta en el tiempo previsto y de manera correcta.

Puede darse el caso de que una entidad sea eficiente pero no sea eficaz y viceversa. La eficiencia es una definición a largo plazo, mientras que la eficacia es algo inmediato, una definición a corto plazo. Si se mantiene la eficiencia se conseguirá la eficacia y por ello, cualquier organización debería tender a ser eficiente.

3.2 Introducción al DEA.

El DEA fue introducido por primera vez por Charnes, Cooper y Rhodes [Charnes et al, 1978], basándose en el trabajo seminal de Farrel [Farrell, 1957]. Además de Farrel, otros autores establecieron los cimientos que más tarde servirían de base al desarrollo del DEA, como por ejemplo Aigner y Chu [Aigner, 1968] y Afriat [Afriat, 1972]. Posteriormente fue ampliado y mejorado por Banker, Charnes y Cooper en 1984 [Banker, 1984]. La metodología DEA está basada en la programación lineal que mide la **eficiencia relativa** de **unidades productivas** que funcionan de forma similar.

Una unidad productiva es cualquier entidad (empresa, hospital, país, línea de fabricación, organización, club deportivo, etc.) que, a partir de ciertos recursos o entradas genera determinados productos o salidas. Este concepto podemos verlo de forma visual en la figura 3-1.



Figura 3-1. Ejemplo general de una unidad productiva o entidad cualquiera

Para explicar el significado de eficiencia relativa se presenta una definición de productividad. Podemos definir la productividad de una entidad como la relación existente entre los productos generados y los recursos empleados para obtenerlos.

Esta definición es reducida por Farrel (1957) al introducir la expresión matemática:

$$Productividad = \frac{Producción\ creada}{Recurso\ consumido} = \frac{Salida}{Entrada}$$

El hecho de evaluar la productividad solo tiene sentido cuando las unidades productivas con las que estemos trabajando sean flexibles a la hora de tomar decisiones y puedan modificar la cantidad de recursos que se estén empleando o los productos que se obtienen. No es necesario que exista total libertad o flexibilidad a la hora de tomar estas decisiones con respecto a todos y cada uno de los recursos o productos, pero sí con respecto a algunos de ellos. Debido a esto, nos referimos a las unidades productivas como DMU (Decision Making Unit) y de esta forma se le añade el calificativo de decisora, imprescindible para poder evaluar su productividad.

Para evaluar la productividad de una DMU real con varias entradas y salidas hay que seguir una serie de pasos:

1. Determinar qué factores o recursos son importantes para obtener los resultados que se buscan en la DMU que se esté estudiando y decidir la trascendencia de cada uno de ellos.
2. Estudiar las causas por las que se han elegido unos recursos en lugar de otros y establecer qué criterios se van a seguir para medir cada uno de ellos. Esto es necesario porque a la hora de pasar al siguiente paso podemos encontrarnos con dificultades para valorar ciertos criterios.
3. Una vez analizados los recursos elegidos y establecidos los criterios, hay que medir cada recurso para poder determinar cómo ha influido en conseguir los resultados de la unidad productiva.
4. Una vez hecho este primer paso se procede a analizar la productividad y para ello necesitamos una expresión que agrupe entradas (recursos) y salidas (productos) de diferente naturaleza y, por tanto, con unidades de medida diferentes. Esto se soluciona con lo que se conoce como entradas y salidas virtuales, lo cual consiste en escalar estas entradas y salidas asociándoles un

peso y consiguiendo un resultado adimensional e independiente de la escala que se haya utilizado para medir cada recurso y cada producto. Esto nos lleva a la expresión:

$$Productividad = \frac{\text{Suma ponderada de salidas}}{\text{Suma ponderada de entradas}}$$

A continuación se van a definir otros conceptos importantes antes de pasar a describir los modelos más representativos. Cualquier modelo DEA se caracteriza por su tecnología (CRS/VRS) y por su orientación (entrada/salida) y esto es lo que se va a exponer a continuación:

- **Retorno de escala constante (CRS):** consiste en considerar que cualquier unidad productiva puede alcanzar la productividad de las eficientes. Al tener como unidad de referencia la de mayor productividad, la eficiencia que se calcula es la eficiencia global.

Las DMUs posibles se reducen al conjunto:

$$T_{CRS} = \{(\vec{x}, \vec{y}): \exists \vec{\lambda} \geq 0, \vec{\lambda}X \leq \vec{x}; \vec{\lambda}Y \geq \vec{y}\}$$

$\vec{\lambda}$: vector con el mismo número de componentes que las DMUs que formen parte del problema.

X, Y : matrices correspondientes a las entradas y salidas de las DMUs. Estas matrices tienen el mismo número de filas que DMUs tenga el problema. X tiene el mismo número de columnas que entradas, mientras que Y tiene el mismo número de columnas que salidas.

- **Retorno de escala variable (VRS):** consiste en considerar que algunas unidades productivas cuyo tamaño sea diferente al de las eficientes, podrían no tener la capacidad de conseguir la productividad de éstas. En este caso lo que habría que hacer es referir cada una de ellas a la productividad mayor entre las de su tamaño, pasando a realizar el estudio mediante la eficiencia técnica en lugar de la eficiencia global.

El conjunto de los puntos admisibles del problema se reduce a:

$$T_{VRS} = \{(\vec{x}, \vec{y}): \exists \vec{\lambda} \geq 0, \vec{\lambda}X \leq \vec{x}; \vec{\lambda}Y \geq \vec{y}; \vec{\lambda}\vec{e}^T = 1\}$$

En este caso la suma de las componentes del vector $\vec{\lambda}$ tiene que sumar la unidad.

- **Orientación de entrada:** consiste en que una DMU alcance la productividad de la unidad de referencia reduciendo la cantidad de recursos que emplea.
- **Orientación de salida:** se refiere al hecho de que una DMU consiga la productividad de la unidad de referencia aumentando la cantidad de producto conseguido.

3.3 Modelos más representativos

Para exponer algunos de los principales modelos DEA vamos a clasificarlos en dos grupos: de retorno de escala constante y de retorno escala variable. Además, aparte de los modelos incluidos en estas dos clasificaciones se van a presentar otros que se aplican a situaciones no consideradas en esta primera clasificación. [Cooper, 2000].

3.3.1 Modelos con retorno de escala constante

Todos ellos tienen en común que a la hora de estudiar la eficiencia relativa, las unidades toman como referencia a la DMU de mayor productividad de entre las que forman parte del estudio.

Los modelos que se presentan a continuación se obtienen a partir del modelo RATIO. Éste es no lineal, su función objetivo es un cociente y se basa en la libertad a la hora de que las unidades productivas escojan los valores de los pesos que optimicen su eficiencia.

3.3.1.1 Modelo CCR-INPUT

El modelo RATIO es un modelo no lineal de difícil resolución y de esta forma surgen nuevos modelos. El objetivo de CCR-INPUT es convertir los cocientes del modelo RATIO en expresiones lineales y para ello debemos tener en cuenta estas dos consideraciones:

- Maximizar un cociente es equivalente a maximizar el numerador, manteniendo el denominador constante.
- El resultado de un cociente será menor a la unidad siempre que el numerador sea menor que el denominador.

A partir de esto se consigue un modelo lineal con $n+1$ restricciones y $s+m$ cotas en lo que sería la forma multiplicadora del mismo. Sin embargo, es más habitual trabajar con la forma envolvente, que consiste en utilizar las variables del dual y que facilitan el análisis de los resultados obtenidos.

$$MIN: \theta_j - \varepsilon \left[\sum_{k=1}^s h_k^+ + \sum_{i=1}^m h_i^- \right]$$

s.a:

$$\sum_{j=1}^n x_{ij} \lambda_j = \theta_j x_{iJ} - h_i^- \quad i = 1, 2, \dots, m$$

$$\sum_{j=1}^n y_{kj} \lambda_j = y_{kJ} + h_k^+ \quad k = 1, 2, \dots, s$$

$$\lambda_j, h_i^-, h_k^+ \geq 0$$

θ_j libre

Las n variables λ_j corresponden a las n primeras restricciones del problema primal, θ_j corresponde a la restricción que queda y h_i^- y h_k^+ son las variables de holgura y corresponden a las cotas del problema primal ($s+m$). Para resolver este modelo hay que pasar por dos fases:

1. Minimizar θ_j con la máxima prioridad, dejando a un lado las holguras. Esta solución nos da un valor θ_j^* que se utilizará en la siguiente fase. Podríamos entender θ_j como el valor por el que hay que multiplicar los recursos para ser eficiente. Si $\theta_j^* = 0,45$ significa que hay que reducir las entradas en un 55% para alcanzar la eficiencia.
2. Maximizar las variables de holgura con la solución del modelo anterior (θ_j^*).

Si al resolver el modelo $\theta_j^* = 1$ y $(h_k^+)^*, (h_i^-)^- \neq 0$ para alguna entrada o alguna salida, esto da lugar a una proyección paralela al eje correspondiente a la variable de holgura que no es nula. Sin embargo, si $\theta_j^* = 1$ y $(h_k^+)^*, (h_i^-)^- = 0$, no se produce ninguna proyección, lo que implica que la unidad es eficiente porque se proyecta sobre sí misma.

Las proyecciones que se calculan representan la unidad en la que debería transformarse cada DMU_j para considerarse eficiente. Para llegar a esta eficiencia se deben reducir las entradas y de ahí que este problema se plantee con orientación de salida.

3.3.1.2 Modelo CCR-OUTPUT

En este caso se va a obtener un modelo con orientación de salida. Partiendo de RATIO se va a minimizar el denominador de la expresión de la función objetivo, manteniendo constante el numerador. Al igual que en el caso anterior, se trabaja con la representación dual del modelo, la cual queda de la siguiente forma:

$$\begin{aligned}
 \text{MAX: } & \gamma_j + \varepsilon \left[\sum_{k=1}^s h_k^+ + \sum_{i=1}^m h_i^- \right] \\
 \text{s.a:} & \\
 & \sum_{j=1}^n x_{ij} \lambda_j = x_{ij} - h_i^- \quad i = 1, 2, \dots, m \\
 & \sum_{j=1}^n y_{kj} \lambda_j = \gamma_j y_{kj} + h_k^+ \quad k = 1, 2, \dots, s \\
 & \lambda_j, h_i^-, h_k^+ \geq 0 \\
 & \gamma_j \text{ libre}
 \end{aligned}$$

La solución se alcanza también en dos fases:

1. Maximizar γ_j dejando a un lado las holguras. Esta solución nos dejará un valor γ_j^* que usaremos en la fase siguiente. Esta nueva variable consiste en una amplificación radial que debe tener lugar en las salidas para proyectarse sobre la frontera eficiente.

2. Maximizar las holguras. En este caso hacemos uso del valor γ_j^* hallado anteriormente.

Al igual que en el caso anterior, si $\gamma_j = 1$ y $(h_k^+)^*, (h_i^+)^- \neq 0$ para alguna entrada o salida, esto implica una proyección paralela al eje correspondiente al término que no es nulo. En caso contrario, si $\gamma_j = 1$ y $(h_k^+)^*, (h_i^+)^- = 0$, no tiene lugar ninguna proyección, lo que implica que la DMU_j se proyecta sobre ella misma y por tanto es eficiente.

3.3.2 Modelos con retorno de escala variable

Cuando el tamaño de las unidades productivas que formen parte del estudio no sea el mismo, pueden surgir problemas con los modelos que se acaban de plantear. Para trabajar con un problema que se plantee con retorno de escala variable, algo bastante común en un estudio real, se plantean otros modelos que se van a explicar a continuación.

3.3.2.1 Modelo BCC-INPUT

En este caso partimos del modelo RATIO linealizado, añadiendo algún elemento, ya sea una restricción o una variable, que lo limite a la hora de comparar cada DMU_j, comparando a ésta exclusivamente con las de su tamaño en lugar de compararla con todas las del problema. A partir de esto, la forma envolvente del modelo quedaría:

$$MIN: \quad \theta_j - \varepsilon \left[\sum_{k=1}^s h_k^+ + \sum_{i=1}^m h_i^- \right]$$

sa:

$$\sum_{j=1}^n x_{ij} \lambda_j = \theta_j x_{ij} - h_i^- \quad i = 1, 2, \dots, m$$

$$\sum_{j=1}^n y_{kj} \lambda_j = y_{kj} + h_k^+ \quad k = 1, 2, \dots, s$$

$$\sum_{j=1}^n \lambda_j = 1$$

$$\lambda_j, h_i^-, h_k^+ \geq 0$$

$$\theta_j \text{ libre}$$

La tercera restricción es la que obliga a que la proyección de la unidad productiva se efectúe sobre la frontera eficiente de las unidades de su tamaño. La frontera eficiente en este caso está formada por más unidades que en los modelos anteriores.

Este modelo, a pesar de ser con orientación de entrada, permite también que se modifiquen las salidas.

Si tenemos un caso real en el que las salidas no puedan modificarse no se puede aplicar. El hecho de que tenga orientación de entrada implica que la reducción radial solamente se permite para las entradas. Al igual que en el modelo CCR-INPUT, θ_j es la eficiencia relative de cada DMU_j.

3.3.2.2 Modelo BCC-OUTPUT

Es análogo al modelo anterior pero en este caso con orientación de salida. La formulación del modelo quedaría de la siguiente forma:

$$MAX: \quad \gamma_j + \varepsilon \left[\sum_{k=1}^s h_k^+ + \sum_{i=1}^m h_i^- \right]$$

sa:

$$\sum_{j=1}^n x_{ij} \lambda_j = x_{ij} - h_i^- \quad i = 1, 2, \dots, m$$

$$\sum_{j=1}^n y_{kj} \lambda_j = \gamma_j y_{kj} + h_k^+ \quad k = 1, 2, \dots, s$$

$$\sum_{j=1}^n \lambda_j = 1$$

$$\lambda_j, h_i^-, h_k^+ \geq 0$$

$$\gamma_j \text{ libre}$$

Al igual que ocurría con el caso anterior, pueden ser modificadas tanto las entradas como las salidas. En este caso la proyección radial solo puede considerarse con las salidas, al ser un modelo con orientación de salida. Una vez que se haya proyectado de forma radial se pasarían a las proyecciones rectangulares.

3.3.3 Modelo aditivo.

Es un modelo que sirve para problemas que no tienen una orientación concreta. Además, puede aplicarse para retornos de escala constante y variable. Otra diferencia con respecto a los casos anteriores es que no cuenta con primera fase, es decir, no existe la proyección radial (reducción radial de las entradas o aumento radial de las salidas) y sólo va a ejecutar la segunda fase descrita en los modelos CCR y BCC. El objetivo va a ser siempre maximizar las holguras, de ahí que no influya la orientación del problema. Por tanto, al realizar la representación analítica se va a prescindir de la variable implicada en la variación radial de entradas o salidas.

A partir de todo esto, para modelos CRS resultaría de la siguiente forma:

$$MAX: \quad \sum_{k=1}^s h_k^+ + \sum_{i=1}^m h_i^-$$

sa:

$$\sum_{j=1}^n x_{ij} \lambda_j = x_{iJ} - h_i^- \quad i = 1, 2, \dots, m$$

$$\sum_{j=1}^n y_{kj} \lambda_j = y_{kJ} + h_k^+ \quad k = 1, 2, \dots, s$$

$$\lambda_j, h_i^-, h_k^+ \geq 0$$

Para modelos VRS, la formulación sería:

$$MAX: \quad \sum_{k=1}^s h_k^+ + \sum_{i=1}^m h_i^-$$

sa:

$$\sum_{j=1}^n x_{ij} \lambda_j = x_{iJ} - h_i^- \quad i = 1, 2, \dots, m$$

$$\sum_{j=1}^n y_{kj} \lambda_j = y_{kJ} + h_k^+ \quad k = 1, 2, \dots, s$$

$$\sum_{j=1}^n \lambda_j = 1$$

$$\lambda_j, h_i^-, h_k^+ \geq 0$$

La eficiencia que se calcula se mide mediante las variables de holgura. Si una unidad es eficiente, se va a obtener el mismo resultado a partir del modelo BCC que del modelo aditivo (en el caso de retorno de escala variable; ocurriría exactamente lo mismo entre el modelo CCR y el modelo BCC para problemas con retorno de escala constante). En ambos casos se obtiene la misma frontera eficiente. Sin embargo, si una DMU_j es ineficiente, el nivel de eficiencia que se obtiene puede variar según qué modelo se esté aplicando.

A la hora de resolver el problema hay que tener en cuenta que cuando se realizan las proyecciones rectangulares siempre hablamos de reducción de las entradas, ya que DEA no nos permite aumentarlas debido a que implicaría un cambio de escala del problema. Por el mismo motivo, en el caso de las salidas solamente se considera el aumento de ellas.

3.3.4 Modelos con entradas y salidas no discretionales.

Para comenzar este apartado se va a presentar en primer lugar el concepto de entrada y salida no discrecional. La única forma de saber si una entrada o salida es o no discrecional es conocer la entidad a la que va dirigido el estudio.

Se dice que una entrada o salida es no discrecional cuando la unidad productiva no tiene la capacidad de modificarla en el problema que se esté estudiando. Existen problemas reales donde la cantidad de recursos que se consumen no se puede controlar, al igual que existen otros problemas donde las salidas son fijas.

En un problema dado no pueden ser todas las entradas y salidas no discretionales, ya que si no se puede modificar ninguna entrada ni salida, no se podría modificar tampoco de ninguna forma la eficiencia de la unidad productiva, por tanto, no tendría sentido compararla con otras.

Esta cuestión que se acaba de exponer hay que implantarlo en los modelos. Para llevarlo a cabo se puede dividir el conjunto de entradas y salidas en otros dos subconjuntos, en función de tener entradas/salidas discretionales o no discretionales.

$$I = I_D \cup I_{ND}$$

$$O = O_D \cup O_{ND}$$

Donde I se refiere al conjunto formado por el total de las entradas, incluyendo las discretionales (indicadas con el subíndice D) y las no discretionales (indicadas con el subíndice ND) y O se refiere al conjunto total de las salidas, incluyendo también todas ellas.

Vamos a introducir estos conjuntos en un modelo para ver cómo quedaría. En este caso elegimos el CCR-INPUT y en su forma dual se representaría así:

$$MIN: \theta_J - \varepsilon \left[\sum_{k \in O_D} h_k^+ + \sum_{i \in I_D} h_i^- \right]$$

s.a:

$$\sum_{j=1}^n x_{ij} \lambda_j = \theta_J x_{iJ} - h_i^- \quad \forall i \in I_D$$

$$\sum_{j=1}^n y_{kj} \lambda_j = y_{kJ} + h_k^+ \quad \forall k \in O_D$$

$$\sum_{j=1}^n x_{ij} \lambda_j \leq x_{iJ} \quad \forall i \in I_{ND}$$

$$\sum_{j=1}^n y_{kj} \lambda_j \geq y_{kJ} \quad \forall k \in O_{ND}$$

$$\lambda_j, h_i^-, h_k^+ \geq 0$$

θ_j libre

La variable θ_j no afecta a las entradas no discretionales. A la hora de resolverlo debemos tener en cuenta que a las entradas no discretionales no se les aplica la segunda fase del modelo, explicada anteriormente en este apartado. Esto quiere decir que las únicas holguras que se maximizan en la función objetivo son aquellas que pertenecen al conjunto de las discretionales, lo que implica que no tiene lugar la proyección radial ni rectangular de aquellos recursos o salidas que no pueden ser modificados.

3.3.5 Entradas y salidas no deseadas.

Para explicar el tratamiento en DEA a las entradas y salidas no deseadas nos vamos a basar en una aplicación del análisis de productividad no paramétrica a la industria canadiense de celulosa y papel [Hailu, 2001].

El artículo donde se expone este estudio amplía el enfoque de Chavas-Cox al análisis no paramétrico incorporando salidas no deseadas para proporcionar una representación más completa de la tecnología de producción.

En la última década se ha experimentado un aumento de interés en la aplicación de la eficiencia y medidas para cambiar la productividad teniendo en cuenta las salidas no deseadas. Las herramientas analíticas empleadas se pueden agrupar en tres categorías generales: clasificación numérica, paramétricas y no paramétricas.

Los modelos no paramétricos empleados en la literatura en la estimación del rendimiento medioambiental están basados en los modelos de retorno de escala variable del análisis por envoltura de datos (DEA) de Banker, Charnes y Cooper. Este artículo presenta representaciones no paramétricas de los límites internos y externos incorporando salidas no deseadas. Los métodos que se proponen se basan en modificaciones de los modelos no paramétricos de Varian-Banker-Maindiratta ampliado por Chavas y Cox. Se aplicará a un análisis de series temporales de datos en la industria canadiense del papel y la celulosa.

Los planteamientos no paramétricos pueden ser modificados para incorporar las salidas no deseadas en el análisis. Vamos a suponer un proceso productivo en el que se empleen N entradas o recursos para obtener M salidas deseadas y K salidas no deseadas. Tendríamos los vectores $x \in R^N$, $(p \in R^N)$, $(v \in R^M)$, $(r \in R^M)$, $(w \in R^K)$, $(q \in R^K)$, siendo estos vectores las entradas, valor de las entradas, salidas deseadas, valor de las salidas deseadas, salidas no deseadas y valor de las salidas no deseadas, respectivamente. Para introducir las salidas no deseadas se propone tratarlas como entradas, como vemos reflejado a continuación:

$$TE_s^{EYI} = \text{Min } \theta$$

Sujeto a:

$$v^s \leq \sum_{t \in T} z^t v^t$$

$$w^s \geq \sum_{t \in T} z^t w^t$$

$$\theta x^s \geq \sum_{t \in T} z^t x^t$$

$$z^t \in R_+^1, t \in T$$

3.3.6 Modelos en red

Otro modelo de interés en este trabajo son los modelos en red. Para su explicación vamos a basarnos en un estudio de Herbert F. Lewis para la universidad de Nueva York: "Network DEA: efficiency analysis of organizations with complex internal structure". [Lewis, 2004].

En los modelos vistos hasta ahora DEA consideraba cada DMU como una "caja negra" donde las entradas entraban y las salidas salían, sin tener en cuenta ningún paso intermedio, por lo que resulta difícil encontrar la causa exacta de la ineficiencia de una DMU sin saber qué está ocurriendo exactamente en el proceso a la hora de obtener unos resultados u otros. En este artículo se muestra cómo poder observar en el interior de esa "caja negra" para hallar las causas de ineficiencia de una DMU concreta.

La solución que se propone es trabajar con un modelo DEA en red el cual consiste en una DMU que, al mismo tiempo, es una red de sub-DMUs. Cada una de éstas tendrá sus entradas y salidas. Las entradas de alguna de ellas podrían ser las salidas de otra y viceversa. En este caso debemos tener en cuenta que el diseño del sistema de producción es más complejo y que cada una de las sub-DMUs tendrá una tecnología diferente.

Anteriormente se habían hecho otras consideraciones acerca de los modelos DEA en red, como por ejemplo las de Färe y Whittaker [Färe, 1995] y Färe y Grosskopf [Färe, 2000], los cuales presentaron modelos similares, que definen el conjunto de referencia para la DMU como el conjunto de combinaciones lineales convexas de las mismas. En nuestro modelo definimos el conjunto de referencia para la DMU basado en las hipotéticas Sub-DMUs que se identifican con cada Sub-DMU. En la figura 3-2 se puede observar la estructura interna de una DMU de un modelo DEA en red.

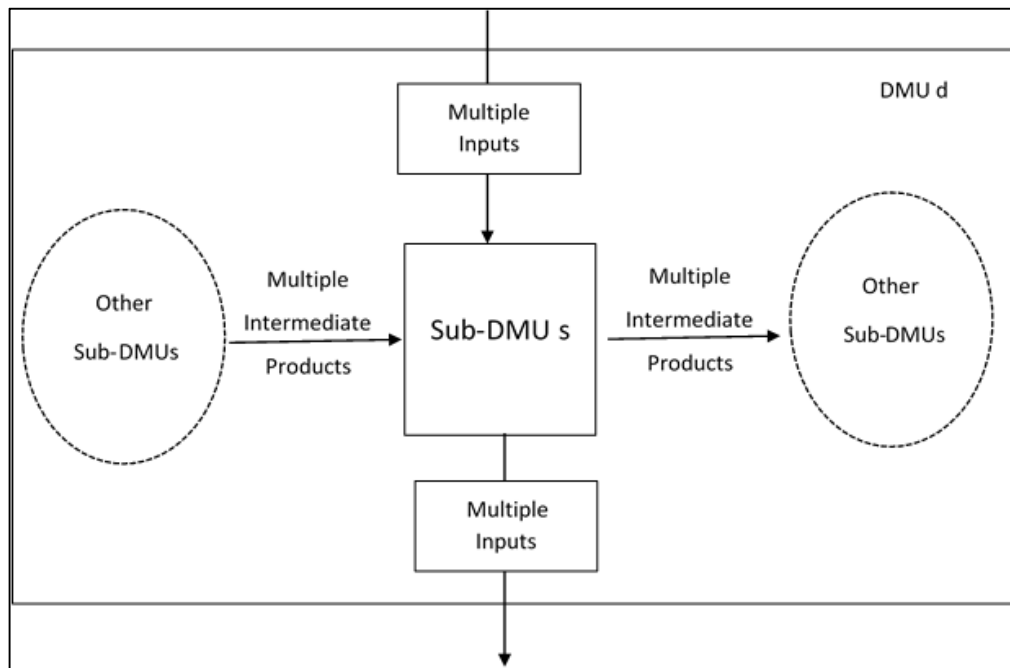


Figura 3-2. Estructura interna de una DMU de un modelo DEA en red
(Ref: Lewis, 2004)

En el estudio realizado por Lewis se aplica este modelo a la Major League Baseball (MLB), que es la liga de béisbol americana. Y nosotros en este trabajo vamos a usar también el concepto de modelos en red para estudiar la eficiencia en partidos de la NBA en defensa y en ataque, teniendo en nuestro caso dos Sub-DMUs.

4 MEDIDA DE EFICIENCIA DE LOS PARTIDOS DE LA NBA

4.1 Introducción

Tras explicar los conceptos más relevantes del DEA y sus modelos más característicos presentamos a continuación el problema concreto que se va a tratar. Tal y como se ha introducido anteriormente, el objetivo del trabajo es realizar un análisis de eficiencia en partidos de baloncesto.

En los apartados anteriores hemos visto la importancia del análisis de datos y de la investigación operativa dentro del mundo deportivo y cómo cada vez se desarrollan más tecnologías centradas en este sector y los equipos de las ligas más importantes del mundo invierten en esta información. En este caso, el estudio de eficiencia referido anteriormente se concreta en relación al tiempo de defensa y ataque de cada equipo en cada partido de baloncesto jugado, por lo que cada DMU estará representada por cada equipo de la NBA en cada partido. Con los datos recopilados de toda la temporada se obtendrán unos tiempos eficientes de defensa y de ataque para cada equipo en cada uno de los partidos jugados. La liga elegida para llevar a cabo este estudio es la NBA.

Para llevarlo a cabo nos hemos centrado en la temporada 2017/2018 y hemos recopilado los datos necesarios de todos los partidos jugados en dicha temporada.

4.2 Modelo gráfico

En primer lugar, se va a exponer el modelo de forma gráfica o abstracta y se puede dividir en dos bloques: uno de ellos es el correspondiente a la defensa de un equipo y el otro el correspondiente al ataque del mismo, ya que hemos dicho que nuestro modelo se basa en calcular la eficiencia en relación al tiempo empleado en ataque y defensa de un equipo en los distintos partidos. Estos dos bloques estarían realacionados precisamente por este tiempo de ataque/defensa.

Teniendo en cuenta esto, se va a exponer el modelo de forma conceptual en lo que podríamos considerar dos fases. En la primera fase se presentan las variables de cada uno de los bloques (defensa y ataque), de forma que podremos diferenciar aquellas que afectan a los puntos anotados por un equipo y las que influyen en los puntos anotados por el equipo contrario. En la segunda fase se van a colocar estas entradas y salidas de forma adecuada para poder aplicar DEA y teniendo en cuenta dos conceptos: la relación directa e indirecta de las entradas y salidas, y el hecho de tener entradas y salidas no deseadas. Estos conceptos se explicarán antes de mostrar cómo quedaría finalmente el modelo.

Por tanto, la figura 4-1 nos va a facilitar la visión general del modelo y las variables elegidas para llevar a cabo este estudio:

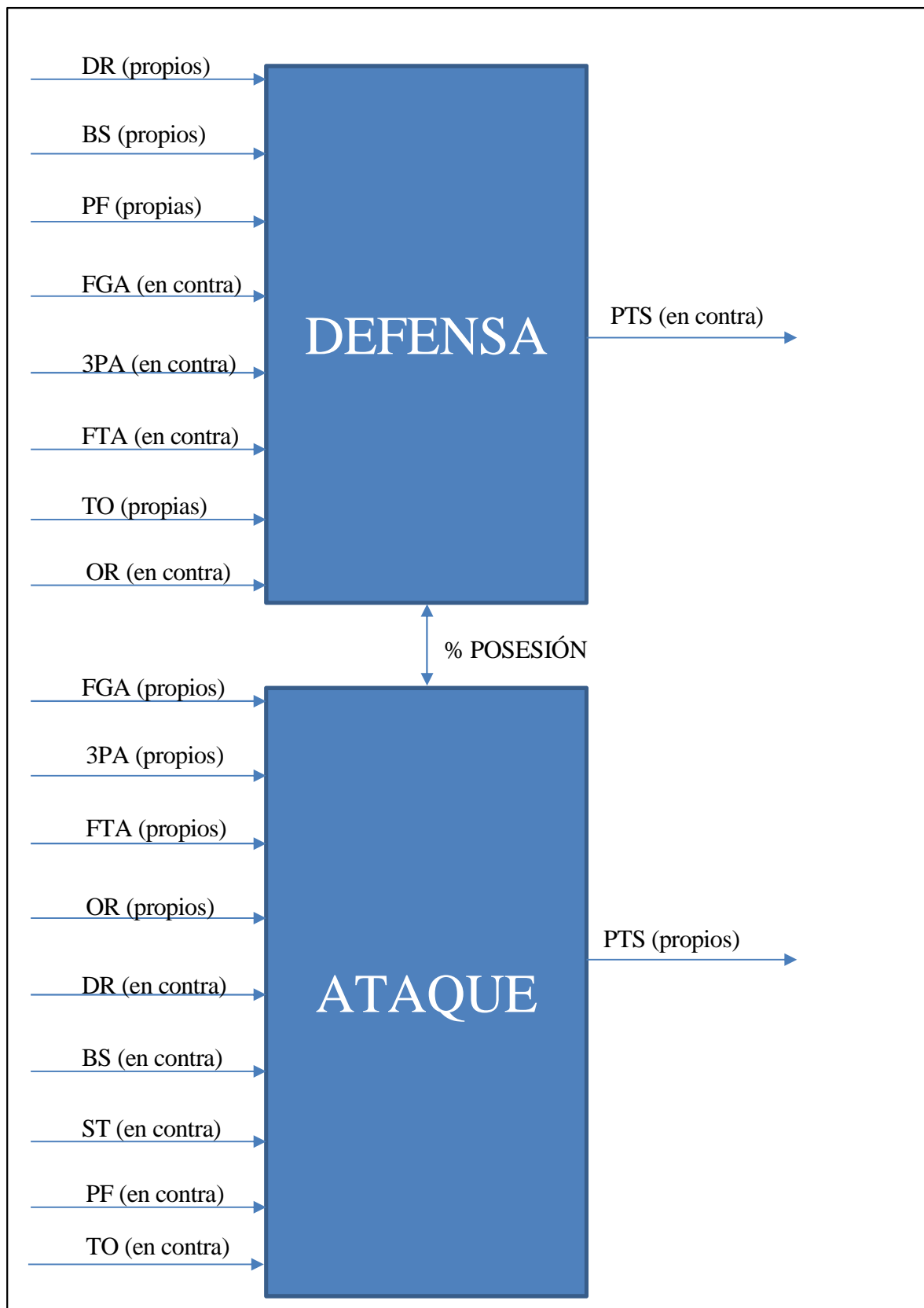


Figura 4-1. Variables del modelo

En la figura 4-1 se muestran las variables consideradas en este trabajo, las cuales vamos a describir a continuación. En la mayoría de ellas se han usado las siglas en inglés. Para realizar el estudio, tal y como se ha dicho anteriormente, se han usado los datos relativos a la temporada 2017/2018, tanto de la liga regular como de los playoffs, hasta la final de la temporada.

- **DR (Defensive Rebound):** son los rebotes defensivos. Un rebote defensivo consiste en atrapar el balón tras un lanzamiento del equipo contrario. Es un concepto clave en baloncesto y se podría decir que una buena defensa debe terminar con la captura del rebote. El número de rebotes defensivos propios se encuentra en el bloque de la defensa y afecta a los puntos en contra, de forma que cuantos más rebotes defensivos propios tengamos, menor será el número de puntos en contra. Por otra parte, el número de rebotes defensivos en contra (los del equipo contrario) se encuentra en el bloque correspondiente al ataque y la relación con la salida es inversamente proporcional, es decir, a mayor número de rebotes defensivos del equipo contrario, menor será el número de puntos propios anotados.
- **BS (Blocked Shots):** son los bloqueos o taponos. El tapón es una acción que consiste en que el defensor desvía el tiro de un atacante del equipo contrario sin tocar al jugador rival, es decir, legalmente. En este caso, el defensor no puede llegar a tocar las manos del atacante para que no se considere la acción una falta personal. En el bloque de la defensa hay que tener en cuenta que cuanto más bloqueos o taponos propios tengamos (los que un equipo le hace al otro), menor será el número de puntos en contra. Del mismo modo, en el bloque de ataque, el número de bloqueos o taponos en contra (los que un equipo “recibe”) afecta negativamente a los puntos que anota un equipo.
- **PF (Personal Foul):** falta personal. Es una infracción que consiste en el contacto entre dos jugadores de equipos contrarios. Cuando un jugador contacta ilegalmente con un adversario, causándole una desventaja u obteniendo una ventaja se considera una falta personal. Cada jugador tiene permitido un máximo de 6 faltas personales por partido. En el bloque de la defensa, podemos decir que el aumento de las faltas personales podría suponer una disminución del número de puntos en contra. Por otra parte, con respecto al ataque, las faltas personales del contrario también hacen que disminuya el número de puntos que se anoten.
- **FGA (Field Goals Attempted):** tiro de campo. Son los intentos de puntuar de un equipo, ya sea tirando desde la distancia correspondiente a los 2 puntos o a los 3 puntos. En la defensa, a mayor número de intentos de tiro del adversario, mayor será el número de puntos en contra (nuestra salida). En lo correspondiente al ataque, el número de intentos de tiro propios afectará positivamente a los puntos anotados, siendo ésta una relación directamente proporcional (a mayor número de intentos de tiro, mayor número de puntos anotados por un equipo).
- **3PA (Three Points Attempted):** intentos de tiros de tres puntos. Tal y como su nombre indica, son los intentos de puntuar, por parte de un equipo, a una distancia a partir de la cual se consideren los tres puntos. En la defensa, se entiende que a mayor número de estos intentos de tiro del adversario, mayor será el número de puntos en contra. De la misma manera, en lo que corresponde al ataque, cuanto mayor sean los intentos de tiro de tres puntos propios de un equipo, mayor será también el número de puntos anotados por el mismo.
- **FTA (Free Throws Attempted):** intentos de tiros libres. Este número abarca todas las oportunidades de un equipo de anotar un tiro libre. Este tiro es el de menos puntuación (un

punto) y, al igual que ocurría con FGA y 3PA, en la defensa a mayor número de intentos de tiros libres del contrario, mayor será el número de puntos anotados por el mismo y en el ataque, cuanto más intentos de tiros libres tenga un equipo, mayor será el total de puntos anotados al final del partido.

- **TO (Turnovers):** pérdidas de balón. En baloncesto, una pérdida de balón tiene lugar cuando un equipo pierde la posesión del mismo antes de tirar a canasta para intentar anotar algún punto para su equipo. En relación a la defensa, podemos decir que cuanto mayor es el número de pérdidas de balón de un equipo, mayor podría ser el número de puntos en contra de dicho equipo. De la misma forma, en lo relativo al ataque, las pérdidas de balón del contrario favorecen los puntos que anote nuestro equipo.
- **OR (Offensive Rebound):** son los rebotes ofensivos. El rebote ofensivo es muy importante en el ataque, ya que puede conceder una segunda oportunidad tras cometer algún fallo. En lo relativo a la defensa, un mayor número de rebotes ofensivos en contra implica un mayor número de puntos anotados por el adversario. Por otra parte, con respecto al ataque, un mayor número de rebotes ofensivos propios va a facilitar un número mayor de puntos anotados por el equipo.
- **ST (Steals):** son los robos de balón de un equipo al contrario. Provoca que el equipo contrario pierda la posesión del balón. Se acreditan al primer jugador que haya causado la pérdida de balón en una jugada concreta. En lo relativo al ataque, los robos del balón por parte de nuestro adversario, dificulta que anotemos puntos.
- **PTS (Points):** es el total de puntos anotados al final del partido. En la defensa tenemos en cuenta los puntos anotados por el equipo contrario y en el ataque los puntos que anota el propio equipo.
- **%TIEMPO:** es la variable más importante del modelo, ya que relaciona los dos bloques entre sí. Es el porcentaje de tiempo con respecto al total que un equipo está atacando y que lo vamos a calcular a partir de una fórmula relativa a la posesión de balón que veremos posteriormente.

Hasta aquí la primera fase de la que hablábamos, en la que hemos visto cuáles serían las entradas y las salidas de nuestro problema y hemos explicado todas las variables que vamos a tener en cuenta. Sin embargo, en DEA éstas no serían las entradas y salidas en la formulación, ya que tenemos que considerar dos aspectos fundamentales que van a ser explicados a continuación y que nos van a permitir reordenar las variables de la figura 4-1.

Por un lado, en la primera parte de esta segunda fase hay que analizar cuál es la relación directa de nuestro problema, es decir, cuando aplicamos DEA lo habitual es que a mayor número de recursos invertidos se incremente la cantidad de productos. Ésta sería la relación natural y, en este caso, los recursos estarían todos en el lugar correspondiente a las entradas y los productos en el lugar correspondiente a las salidas.

En nuestro caso esta relación directa vamos a establecerla tanto para el bloque correspondiente a la defensa, como para el correspondiente al ataque.

En el caso de la defensa, nuestro objetivo es que el número de puntos en contra sea lo menor posible. Esto implica que aquellas entradas que favorezcan la disminución del número de puntos en contra se

considerará que tienen una relación directa con las salidas y, por tanto, las dejaremos en las entradas. Sin embargo, aquellas entradas que provoquen un aumento del número de puntos en contra tendrán que ser consideradas como salidas por tener una relación inversa. Estas relaciones ya han sido mencionadas al describir cada una de las variables, por lo que podemos deducir que en el caso de la defensa aquellas entradas que tienen una relación inversa con la salida son FGA (en contra), 3PA (en contra), FTA (en contra), TO (propios), OR (en contra).

De la misma forma, en el bloque que corresponde al ataque, podemos decir que nuestro objetivo sería que el número de puntos anotados sea el máximo posible. Así, aquellas entradas que favorezcan el total de puntos propios quedarán colocadas donde están, mientras que aquellas que provoquen la disminución de número de puntos anotados se considerará que tienen una relación inversa a la establecida y tendremos que considerarlas como salidas. Tal y como ocurría con la defensa, estas relaciones han sido explicadas anteriormente, por lo que podemos considerar como salidas las entradas: DR (en contra), BS (en contra), ST (en contra), PF (en contra).

En ambos bloques podríamos decir que aquellas entradas que tienen una relación inversa se tratan como si fuesen entradas no deseadas y por eso pasan a ser salidas del modelo definitivo. Teniendo en cuenta todo esto, el modelo de la figura 4-1 quedaría ahora de la siguiente forma:

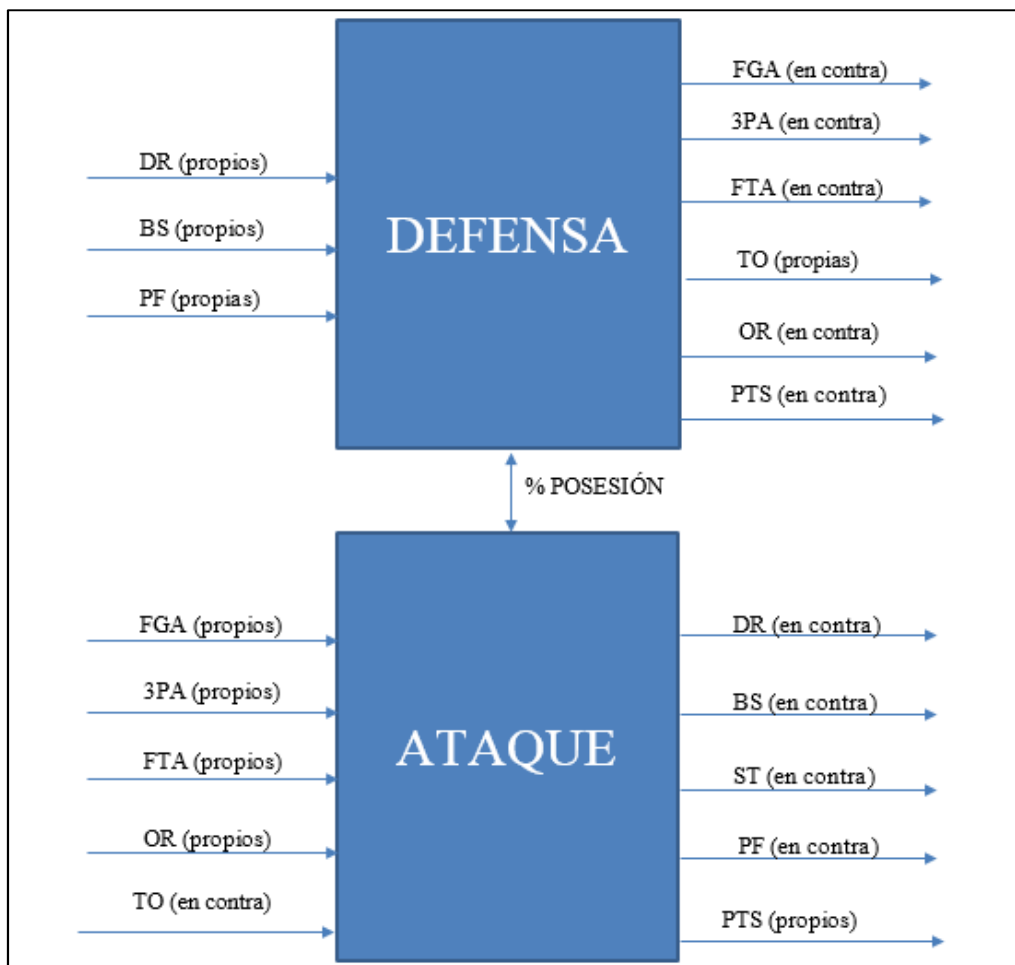


Figura 4-2. Modelo gráfico. Fase 2.1.

En la segunda parte de esta segunda fase hay que tener en cuenta las entradas y salidas no deseadas. Cuando tenemos una entrada no deseada, ésta la colocamos como una salida y, de la misma forma cuando tenemos una salida no deseada, la colocamos como entrada. En este caso donde tenemos que tener en cuenta esta consideración es en la defensa, puesto que la salida que habíamos considerado en la primera fase del modelo era el número de puntos en contra (PTS (en contra)) y ésta se considera una salida no deseada, puesto que queremos que disminuya en lugar de aumentar (que es lo que se suele buscar en una salida). Teniendo en cuenta esto, debemos hacer una transformación más, colocando el total de puntos en contra como una entrada en el bloque correspondiente a la defensa y teniendo el modelo definitivo con el que podemos trabajar según el DEA:

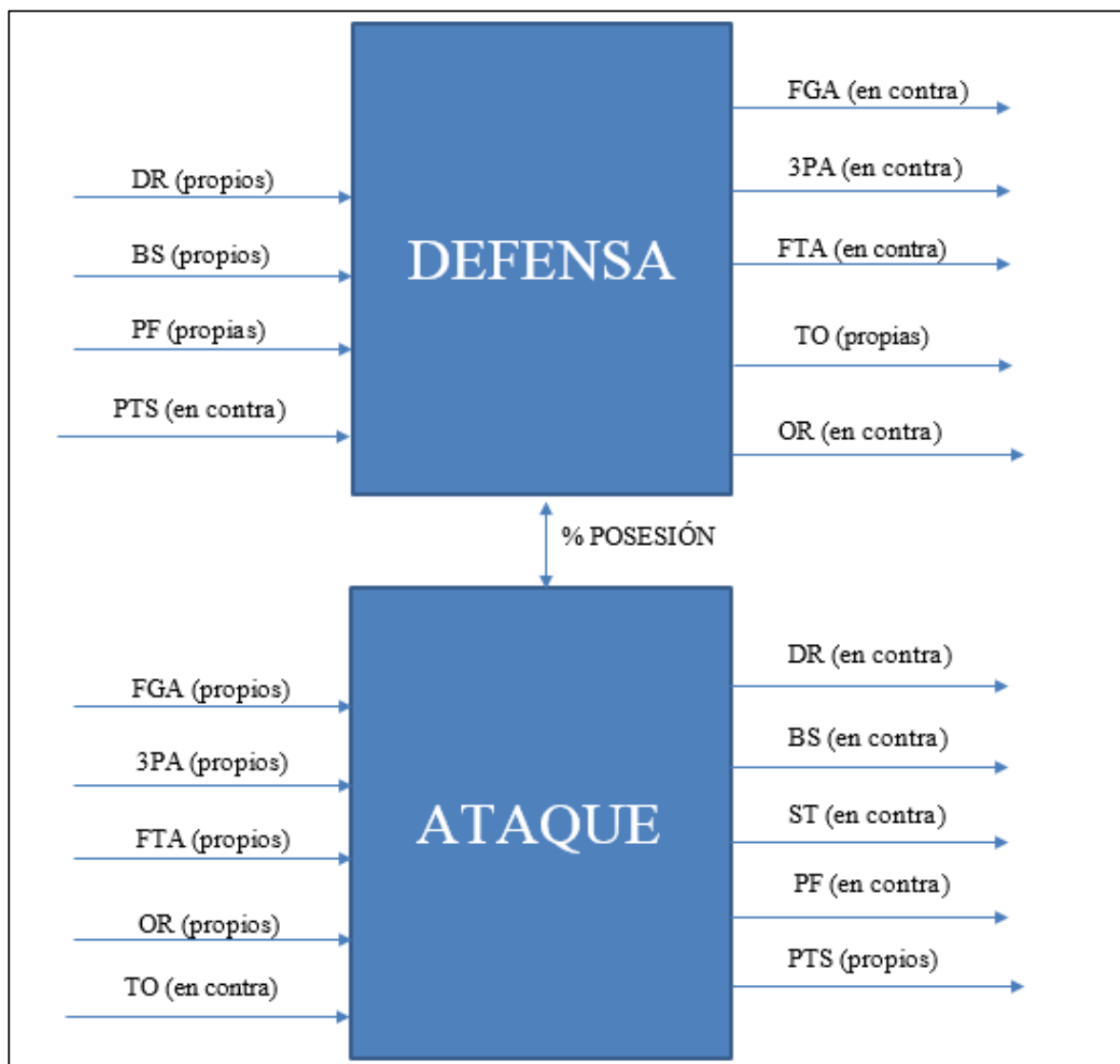


Figura 4-3. Modelo gráfico definitivo

Así, hemos llegado a este modelo final que se puede observar en la figura 4-3. Este modelo es el que vamos a aplicar, considerando como entradas DEA todas aquellas variables a la izquierda de los bloques de defensa y ataque, y como salidas todas aquellas variables a la derecha de dichos bloques. A continuación se presenta el modelo matemático del problema.

4.3 Modelo matemático

Una vez que hemos visto cómo quedaría el modelo definitivo que vamos a usar, pasamos a la traducción del mismo a un modelo matemático que es el que posteriormente se va a implantar.

$$\text{Max } s_0^{\text{off}} + s_0^{\text{def}}$$

s.t.

$$\sum_{j=1}^n \lambda_j^{\text{def}} x_{ij}^{\text{def}} \leq x_{io}^{\text{def}} \quad (1)$$

$$\sum_{j=1}^n \lambda_j^{\text{def}} y_{kj}^{\text{def}} \geq y_{ko}^{\text{def}} \quad (2)$$

$$\sum_{j=1}^n \lambda_j^{\text{at}} x_{ij}^{\text{at}} \leq x_{io}^{\text{at}} \quad (3)$$

$$\sum_{j=1}^n \lambda_j^{\text{at}} y_{kj}^{\text{at}} \geq y_{ko}^{\text{at}} \quad (4)$$

$$\sum_{j=1}^n \lambda_j^{\text{def}} \text{points}_j^{\text{def}} = \text{tpoints}_0^{\text{def}} \quad (5)$$

$$\text{tpoints}_0^{\text{def}} = \text{points}_0^{\text{def}} - s_0^{\text{def}} \quad (6)$$

$$\sum_{j=1}^n \lambda_j^{\text{at}} \text{points}_j^{\text{at}} = \text{tpoints}_0^{\text{at}} \quad (7)$$

$$\text{tpoints}_0^{\text{at}} = \text{points}_0^{\text{at}} + s_0^{\text{at}} \quad (8)$$

$$\sum_{j=1}^n \lambda_j^{\text{def}} \text{pos}_j^{\text{def}} = \text{tpos}_0^{\text{def}} \quad (9)$$

$$\sum_{j=1}^n \lambda_j^{\text{at}} \text{pos}_j^{\text{at}} = \text{tpos}_0^{\text{at}} \quad (10)$$

$$\text{tpos}_j^{\text{def}} + \text{tpos}_j^{\text{at}} = 100 \quad (11)$$

$$\sum_{j=1}^n \lambda_j^p = 1 \quad p = \text{at, def} \quad (12)$$

$$\lambda_j^p \geq 0 \quad p = \text{at, def}; \forall j \quad (13)$$

$$s_0^p \geq 0 \quad p = \text{at, def} \quad (14)$$

$$\text{tpoints}_0^p, s_0^p \text{ integer} \quad p = \text{at, def} \quad (15)$$

Aquí se presenta el modelo matemático que se va a implementar para obtener los resultados. El modelo cuenta con una función objetivo y 19 restricciones, ya que las ecuaciones (12), (13), (14) y (15) serían dobles, al considerar el caso defensivo (def) y ofensivo o de ataque (at).

Tenemos un problema de tipo MILP (Mixed Integer Linear Programming), lo que se conoce como programación lineal entera-mixta. Estos problemas se caracterizan porque al menos una de las variables tiene la restricción de ser entera, mientras que el resto no tienen por qué serlo. En este caso, las variables que tienen que ser enteras son s_0^{at} y s_0^{def} y como consecuencia $tpoints_0^{at}$ y $tpoints_0^{def}$. Esto se debe a que están relacionadas con los puntos de un equipo de baloncesto, tanto a favor como en contra, por lo que en ningún caso podríamos obtener un valor que no fuese entero en estas variables a la hora de resolver el modelo.

Los modelos que hemos descrito en apartados anteriores eran una representación de una tecnología concreta. Sin embargo, el modelo que nos ocupa tiene la peculiaridad de que trabaja con dos tecnologías. Éste es el motivo por el que, desde que empezamos a describir cómo íbamos a modelar el problema, hablábamos de dos bloques, uno defensivo y otro ofensivo. Cada uno de estos bloques representaría una tecnología diferente, las cuales se relacionan a través de la variable temporal, por lo que ésta es fundamental. El tiempo nos va a permitir calcular la posesión eficiente de cada equipo en cada partido, lo que permitiría aplicar estos resultados de forma real en un equipo concreto por parte de un director deportivo a la hora de elaborar la estrategia a seguir, cómo se describió en el apartado de modelos en red.

Si consideramos la primera representación gráfica que se hizo del modelo, en ella todas las variables empleadas en el problema se consideraban entradas, excepto los puntos en contra y los puntos anotados por un equipo. Partiendo de esta representación, podemos decir que todas las entradas van a ser no discretas, puesto que son datos que obtenemos en función de cómo hayamos jugado. Vamos a considerar discretas tanto los puntos en contra (en la tecnología referente a la defensa) como los puntos anotados por un equipo (en la tecnología correspondiente al ataque). La variable de holgura correspondiente a los puntos en contra será s_0^{def} y la variable de holgura correspondiente a los puntos anotados será s_0^{at} . Nuestro objetivo, tal y como podemos observar en la función objetivo del modelo, será maximizar la suma de ambas variables de holgura. De esta forma, lo que buscaremos será maximizar cada uno de estos valores. Así, al maximizar s_0^{at} y teniendo en cuenta la restricción (8) se pretenderá que los puntos anotados por un equipo sean los máximos posibles y, de la misma manera, al maximizar s_0^{def} y considerando la restricción (6), con esta variable de holgura se perseguirá disminuir los puntos en contra. Esta holgura viene del interés de analizar qué ha podido ocurrir en un equipo en un partido concreto cuando, habiendo empleado recursos similares al otro, se ha obtenido diferente resultado.

El modelo, al implementarlo, realizará una iteración por cada equipo en cada partido, es decir, cada partido jugado durante la temporada implica dos iteraciones del modelo, llegando hasta un total de 2.624 iteraciones. El hecho de que existan dos tecnologías hace que tengamos dos vectores lambda. Tal y como puede observarse en las restricciones del modelo, existe uno para el ataque (λ_j^{at}) y otro para la defensa (λ_j^{def}). Además, debemos tener en cuenta que el sumatorio de las componentes de ambos vectores tiene que resultar la unidad tanto para la defensa, como para el ataque (restricción 12), porque se considera retorno de escala variable en ambas sub-DMUs (VRS).

Otro aspecto a destacar del modelo es que las ecuaciones (9) y (10) determinarán el tiempo óptimo de posesión. Por otra parte, la ecuación número (11) tiene en cuenta el porcentaje de esta posesión de balón que se ha mencionado con anterioridad cuando se ha explicado el modelo. Esto significa que el tiempo de defensa y de ataque en porcentaje siempre va a sumar 100, de forma que está repartido entre las dos opciones.

Por otra parte, las ecuaciones (5) y (7) son análogas, con la única diferencia de que una está referida a la defensa y la otra al ataque.

5 ANÁLISIS DE RESULTADOS

Antes de comenzar vamos a hacer una introducción de la liga que se va a analizar y cómo se desarrolla de forma general una temporada.

La NBA es una liga de baloncesto privada de Estados Unidos. Está formada por 30 equipos que están divididos en dos conferencias, las cuales se dividen a su vez en divisiones. La temporada está dividida en dos fases.

La primera fase es la temporada regular, donde cada equipo disputa un total de 82 partidos, 41 como local y 41 como visitante. Cada equipo se enfrenta con los oponentes de su misma división en cuatro ocasiones, con los de otras divisiones de su conferencia entre tres o cuatro veces y contra los de otra conferencia dos veces al año. Una vez terminada la temporada regular, los ocho mejores de cada conferencia se clasifican para jugar los Playoffs, siendo ésta la segunda fase.

Los Playoffs consisten en tres rondas de competición entre 16 equipos, 8 por conferencia. En cada conferencia se juegan lo que podríamos llamar cuartos de final de conferencia entre los 8 mejores resultantes de la liga regular, pasando 4 equipos de cada conferencia a las semifinales de conferencia y quedando los dos mejores para la final de conferencia. El ganador de cada conferencia disputará la final de la NBA. En los Playoffs la eliminatoria es al mejor de siete partidos, siendo así en todas las fases, incluyendo la final de la NBA entre los mejores de cada conferencia.

A continuación, en la tabla 5-1 se muestra cómo está organizada la liga por divisiones y conferencias.

| CONFERENCIA DEL ESTE | | |
|----------------------------|------------------------------|-----------------------------|
| División del Atlántico | División Central | División del Sudeste |
| Boston Celtics (BOS) | Chicago Bulls (CHI) | Atlanta Hawks (ATL) |
| Brooklyn Nets (BKN) | Cleveland Cavaliers (CLE) | Charlotte Bobcats (CHA) |
| New York Knicks (NYK) | Detroit Pistons (DET) | Miami Heat (MIA) |
| Philadelphia 76ers (PHI) | Indiana Pacers (IND) | Orlando Magic (ORL) |
| Toronto Raptors (TOR) | Milwaukee Bucks (MIL) | Washington Wizards (WAS) |
| CONFERENCIA DEL OESTE | | |
| División del Suroeste | División del Noroeste | División del Pacífico |
| Dallas Mavericks (DAL) | Denver Nuggets (DEN) | Golden State Warriors (GSW) |
| Houston Rockets (HOU) | Minnesota Timberwolves (MIN) | Los Angeles Clippers (LAC) |
| Memphis Grizzlies (MEM) | Portland Trail Blazers (POR) | Los Angeles Lakers (LAL) |
| New Orleans Pelicans (NOP) | Oklahoma City Thunder (OKC) | Phoenix Suns (PHX) |
| San Antonio Spurs (SAS) | Utah Jazz (UTA) | Sacramento Kings (SAC) |

Tabla 5-1. Divisiones y conferencias de la NBA

En la temporada 2017-2018 al finalizar la liga regular los 8 mejores de cada conferencia que se clasificaron para los playoffs fueron aquellos que se pueden ver en la figura 5-1. En esta misma figura se puede observar cómo se desarrollaron los playoffs y qué equipos fueron pasando de fase hasta llegar a la final de la temporada.



Figura 5-1. Playoffs temporada 2017-2018

Tras ver cómo se organizan los equipos y la temporada de la NBA pasamos al análisis de los resultados obtenidos.

Los datos empleados son los correspondientes a la temporada jugada en 2017-2018. Éstos son las variables elegidas que se han explicado en el apartado anterior, por lo que se ha recogido el total de rebotes defensivos (propios y en contra), rebotes ofensivos (propios y en contra), faltas personales (propias y en contra), tiros de campo (propios y en contra), intentos de tiro de tres puntos (propios y en contra), intentos de tiros libres (propios y en contra), robos de balón (propios), pérdidas de balón (propias y del contrario) y los puntos marcados y los que ha anotado el equipo contrario.

Al recoger estos datos el resultado es un Excel de 2.624 filas, siendo cada una de ellas una iteración del modelo explicado y cuyo código podemos encontrar en el anexo de este documento.

Una vez implementado el modelo en LINGO14, éste tarda en resolverse 40 minutos, tiempo de ejecución que varía en función del ordenador que se utilice.

Una vez ejecutado, los resultados se van a registrar en una tabla en un archivo de Excel donde recogeremos la información obtenida tras la resolución del problema.

Los datos que se obtienen directamente de esta resolución es el porcentaje de tiempo de ataque y defensa eficiente por parte de cada equipo en cada partido, medido a partir de la posesión de balón anteriormente explicada. Además, obtendremos el valor de las holguras s_0^{def} y s_0^{at} y, por tanto, de los puntos que deberían haber anotado ($tpoints_0^{at}$) y los que deberían haber marcado el equipo contrario ($tpoints_0^{def}$) para alcanzar esa eficiencia.

La media de la eficiencia se ha calculado a partir de las holguras, de la siguiente forma:

$$\begin{aligned} \text{➤ Eficiencia en defensa} &= \frac{1}{1 + \frac{s_0^{def}}{pts_{def}}} \\ \text{➤ Eficiencia en ataque} &= \frac{1}{1 + \frac{s_0^{at}}{pts_{at}}} \end{aligned}$$

Donde s_0^{def} y s_0^{at} son las holguras del problema, pts_{def} son los puntos en contra que tiene cada equipo en cada partido y pts_{at} son los puntos que ha anotado cada equipo.

De esta forma, si el valor de la holgura en defensa o ataque es cero, el resultado de esta expresión será 1 y, por tanto, implicará que ese equipo en ese partido concreto ha sido eficiente en defensa o en ataque. Puede darse la situación de que un equipo jugando un partido determinado sea eficiente en ataque y no en defensa o viceversa. En cualquier otro caso en el que la holgura tenga un valor diferente de cero, el resultado de estas expresiones sería distinto de 1 e implicaría que el equipo no ha sido eficiente en defensa, ataque o ambos.

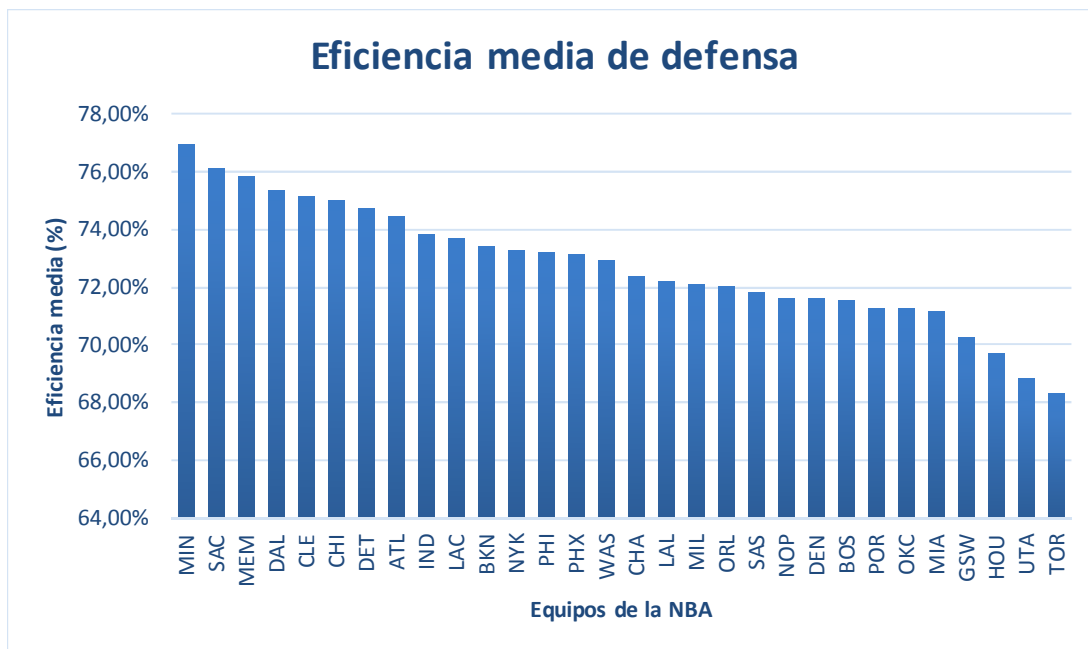
Sólo en el caso en el que un equipo sea eficiente en defensa y en ataque al mismo tiempo (resultados de ambas expresiones igual a 1) consideraremos que el equipo ha jugado de forma eficiente.

Debido a la gran cantidad de datos almacenados, en este capítulo se van a agrupar los resultados obtenidos de forma que se pueda realizar una comparativa general de todos los equipos y se analizarán individualmente los resultados de aquellos que sean más relevantes.

5.1 Eficiencia media. Gráficas generales.

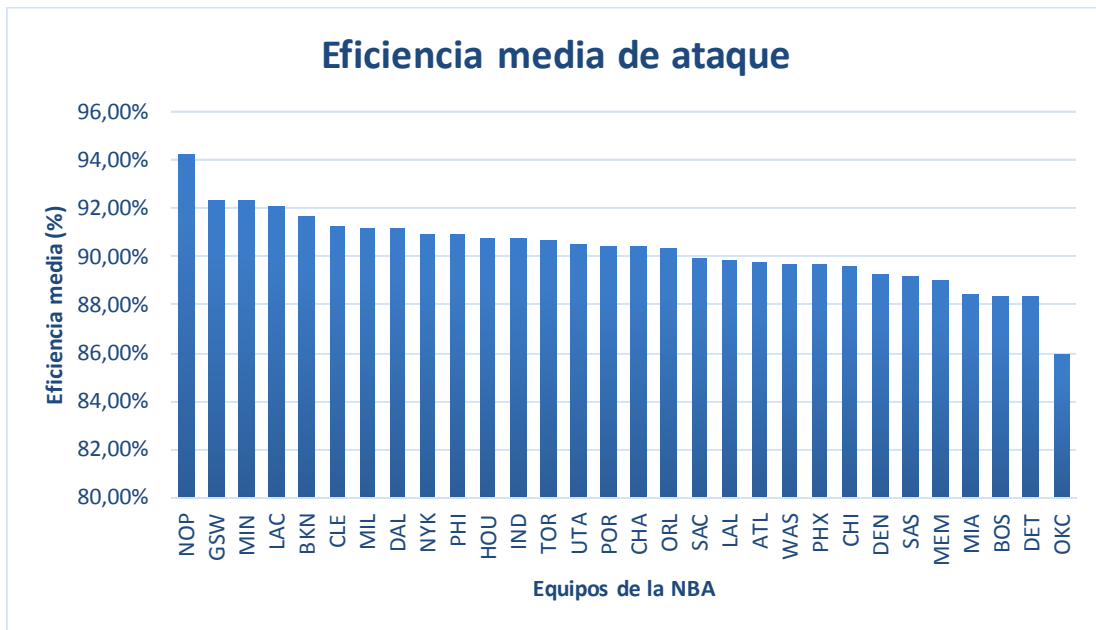
En este primer apartado se van a mostrar varias gráficas que son una comparativa de la eficiencia media de todos los equipos de la NBA, tanto en defensa como en ataque, de forma que se pueda tener una visión general de los resultados obtenidos.

Para ello, se ha calculado un promedio de la eficiencia en defensa y en ataque de todos los equipos que forman parte de la NBA, teniendo en cuenta la eficiencia calculada en cada partido para cada uno de ellos. En la gráfica 5-1 podemos ver el promedio de la eficiencia de cada uno de los equipos ordenados de mayor a menor. Vemos que los que han obtenido un promedio más elevado de eficiencia en defensa son los Minnesota Timberwolves (MIN), con un promedio del 76.97 %, mientras que los que han obtenido un promedio de eficiencia en defensa más bajo, tal y como se puede observar en dicha gráfica, son los Toronto Raptors (TOR), cuyo valor es del 68.34 %.



Gráfica 5-1. Eficiencia media de defensa de la NBA (2017-2018)

Por otra parte, la gráfica 5-2 es una gráfica similar, con la diferencia de que en este caso se muestra la eficiencia media en ataque durante la temporada de cada uno de los equipos, en lugar de la eficiencia en defensa. Estos promedios están ordenados de mayor a menor, siendo en este caso los que poseen una mayor eficiencia media en ataque los New Orleans Pelicans (NOP), con un valor del 94.24%, mientras que los que poseen una menor eficiencia media de ataque son los Oklahoma City Thunder (OKC), con un promedio de eficiencia ofensiva del 85.92%.



Gráfica 5-2. Eficiencia media de ataque de la NBA (2017-2018)

El promedio de eficiencia en ataque es superior al promedio de eficiencia en defensa. La diferencia es significativa tanto en el equipo con mayor eficiencia como en el que menos eficiencia media tiene y como consecuencia se verá reflejado en el número de partidos en los que estos equipos han sido eficientes, lo que veremos en el siguiente apartado.

5.2 Partidos eficientes.

En este caso vamos a mostrar unas gráficas que comparan de nuevo todos los equipos de la NBA. Se van a agrupar todos los partidos en los que cada equipo ha sido eficiente.

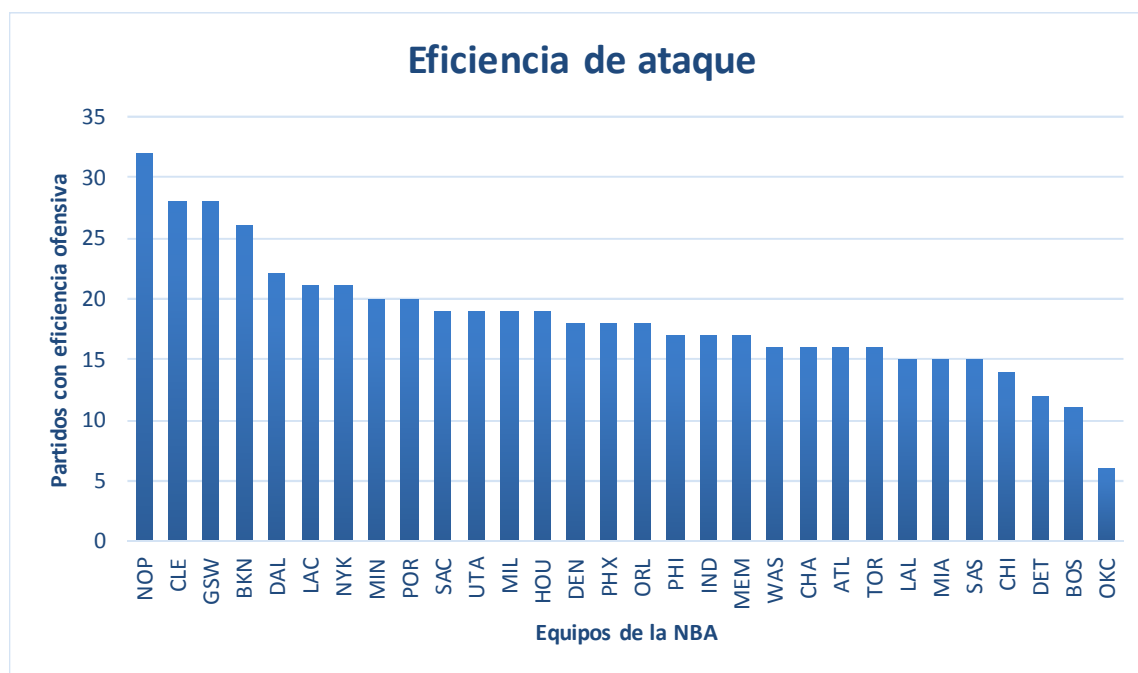
La gráfica 5-3 muestra en orden decreciente el número de partidos en los que cada uno de los equipos ha sido eficiente en defensa según el modelo planteado y en la gráfica 5-4, se muestra, en orden decreciente también, el número de partidos en los que cada equipo ha sido eficiente desde el punto de vista ofensivo.



Gráfica 5-3. Número de partidos eficientes en defensa (2017-2018)

En la gráfica 5-3 observamos que los equipos que han sido en más ocasiones eficientes en defensa son los Brooklyn Nets (BKN) y Detroit Pistons (DET), con un total de 10 partidos, lo cual tiene sentido si lo relacionamos con la gráfica 5-1, ya que ambos equipos se encuentran en la primera mitad de la gráfica y, por tanto, su promedio de eficiencia en defensa está entre los mejores. Sin embargo, a pesar de ser de los equipos que en más ocasiones han sido eficientes en defensa, no tienen el promedio de eficiencia defensiva más elevado, lo que implica que cuando estos equipos no son eficientes desde el punto de vista defensivo alcanzan valores de eficiencias menores que otros equipos de la liga. Al mismo tiempo se observa cómo los OKC no han resultado ser eficientes desde el punto de vista defensivo en ningún partido a lo largo de toda la temporada, lo que influye en que tengan uno de los promedios en eficiencia media de defensa más bajos de la NBA.

En relación a la gráfica 5-4, podemos ver cómo los que tienen una mayor frecuencia de eficiencia en ataque son los New Orleans Pelicans (NOP), lo que se corresponde con la gráfica 5-2, al ser éste el equipo con el promedio más alto de eficiencia ofensiva. En este caso NOP suma un total de 32 partidos donde el equipo ha sido eficiente en ataque, mientras que los OKC vuelven a ser el equipo con el número de partidos más bajos en los que ha jugado de forma eficiente en ataque, sumando un total de 6 partidos en los que se puede considerar que ha sido eficiente. Aquí de nuevo vemos la diferencia entre la eficiencia de ataque y de defensa dentro de los diferentes equipos.



Gráfica 5-4. Número de partidos eficientes en ataque (2017-2018)

Hasta ahora los resultados y las gráficas que se han comentado muestran por separado los resultados de la eficiencia de los distintos equipos desde el punto de vista ofensivo y defensivo. Para hablar de la eficiencia de un equipo en un partido determinado se ha establecido el siguiente criterio: aquellos equipos que en un partido determinado han sido eficientes en defensa y en ataque se les ha otorgado una valoración de 1, lo que implica un 100% de eficiencia. En caso contrario se le valora con un cero, puesto que implica que no han sido eficientes, independientemente de que lo hayan sido en ataque o defensa por separado. Solo se valorará con un 1 la eficiencia simultánea en defensa y ataque.

A continuación, en la gráfica 5-5 se va a mostrar el total de partidos en los que cada equipo que forma parte de esta competición ha sido eficiente. Los equipos que más veces han sido eficientes, tanto en defensa como en ataque, según el modelo planteado serían los Cleveland Cavaliers (CLE) y BKN, con un total de seis partidos. Teniendo en cuenta la cantidad de partidos jugados en esta competición en una temporada completa pueden sorprender las cifras de eficiencia que se han obtenido, puesto que de más de mil partidos jugados obtenemos unos resultados en los que solamente en 69 ocasiones un equipo ha sido eficiente en un partido concreto. Si volvemos a separar la información en defensa y ataque, hemos obtenido un total de 551 casos de eficiencia en ataque frente a 138 de eficiencia defensiva.



Gráfica 5-5. Número de partidos eficientes (2017-2018)

Hasta ahora hemos visto una serie de gráficas donde se han mostrado de forma general los resultados de eficiencia obtenidos del modelo empleado para este problema. A pesar de ser unos resultados agrupados en promedios y comparativas de todos los equipos podemos empezar a sacar las primeras conclusiones. En primer lugar, hay una diferencia significativa entre la eficiencia obtenida en ataque y en defensa, lo que se ha podido observar tanto en promedios de eficiencia como en cantidad de partidos eficientes desde un punto de vista u otro. Esta información podría resultar de interés para los entrenadores o directores deportivos de los distintos equipos, ya que a partir del análisis de resultados se podrían mejorar los aspectos más débiles de un equipo con el objetivo de obtener mejores resultados en un futuro.

Tanto en el apartado 5.1 como en el apartado 5.2 se ha hecho un análisis general comparando todos los equipos de la NBA. Hay un dato que nos ha parecido de especial interés, puesto que Los Golden State Warriors (GSW) son los ganadores de la temporada que hemos estudiado y, sin embargo, el número de partidos eficientes tanto en defensa como en ataque es de los más bajos, lo que podemos comprobar en la gráfica 5-5.

Una vez visto este dato hemos analizado el resto de gráficas y la posición de este equipo con respecto a los demás y podemos observar cómo los Warriors están bien posicionados en eficiencia media de ataque y, como consecuencia, en cantidad de partidos en los que han sido eficientes desde el punto de vista ofensivo, posicionándose en estas gráficas al principio de las mismas. Sin embargo, en la gráfica donde se compara la eficiencia media defensiva, los Warriors no destacan por tener un buen promedio de eficiencia en defensa en comparación con los demás, lo que implica que hay pocos partidos en los que el equipo ha jugado de manera eficiente desde el punto de vista defensivo y esto ha podido provocar que el número de partidos eficientes en ataque y defensa sea de los más bajos, a pesar de ser los ganadores de la temporada. Esto nos da información acerca de la estrategia seguida por este equipo

y su forma de juego y cómo es probable que prioricen el ataque frente a la defensa. Realmente esto ocurre en la mayoría de los equipos, como ya se ha mencionado, ya que existe una diferencia importante en los valores obtenidos de eficiencia en cada caso.

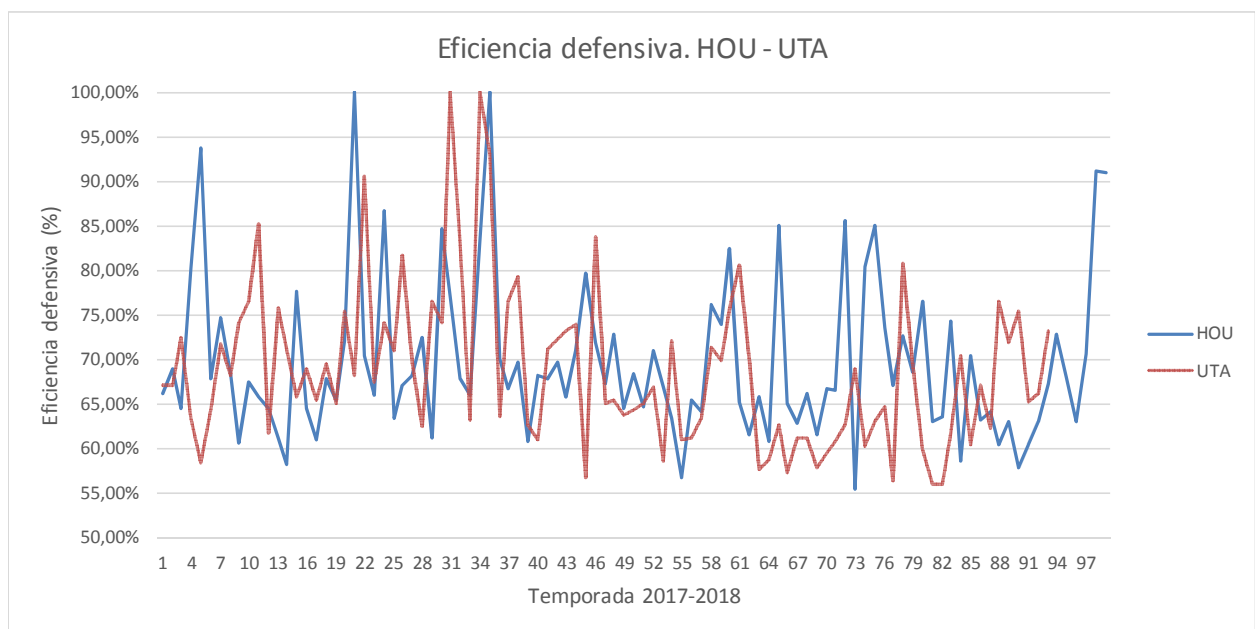
Esta información podría ser útil a la hora de establecer nuevas estrategias, de forma que el entrenador de un equipo como los Golden State Warriors podría plantearse cómo mejorar la eficiencia defensiva del equipo y obtener aún mejores resultados o haciendo uso de los recursos de los que se dispone de una forma más eficiente y que quizás provocase menos desgaste en los jugadores o poder obtener las victorias incluso con más ventaja de la que se tiene actualmente.

5.3 Comparativa por conferencias.

En este apartado se pretende mostrar una comparativa de la eficiencia a lo largo de la temporada por conferencias. Concretamente vamos a realizar esta comparativa con aquellos equipos que lograron pasar a las semifinales de conferencia, siendo éstos los cuatro mejores equipos de cada conferencia.

5.3.1 Conferencia oeste.

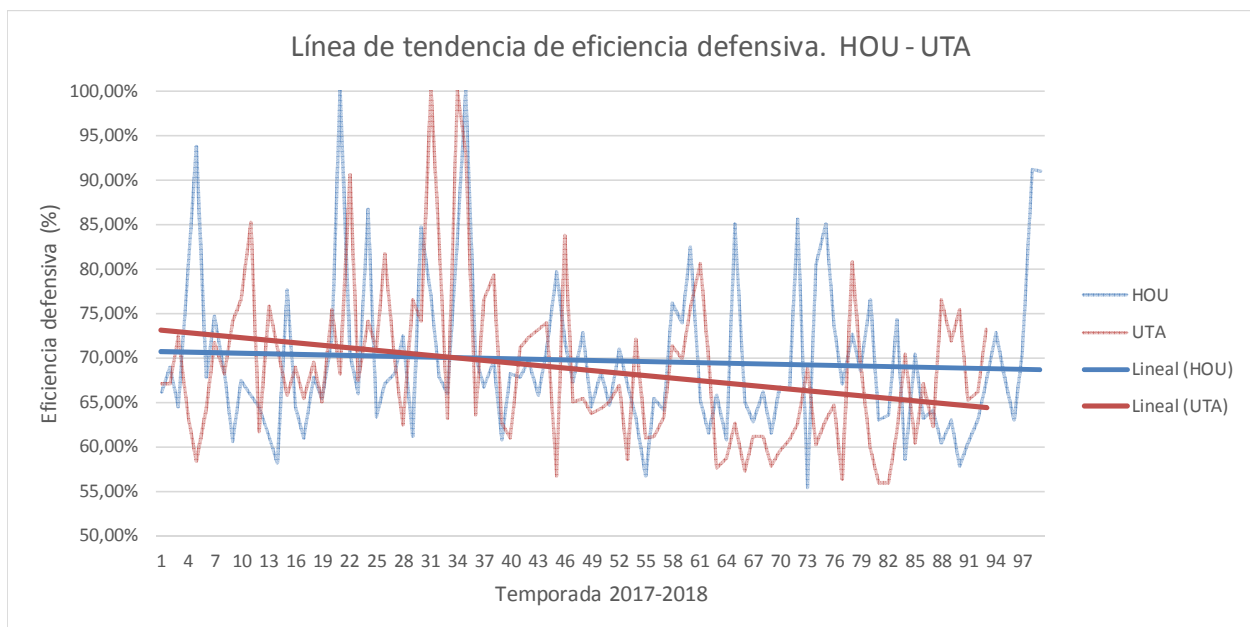
Para ver los resultados de forma más clara se van a agrupar según los equipos que se enfrentaron durante esta fase de los playoffs. Así, en la gráfica 5-6 se puede comparar la eficiencia defensiva a lo largo de la temporada de los Houston Rockets (HOU) y Utah Jazz (UTA).



Gráfica 5-6. Eficiencia defensiva. HOU – UTA

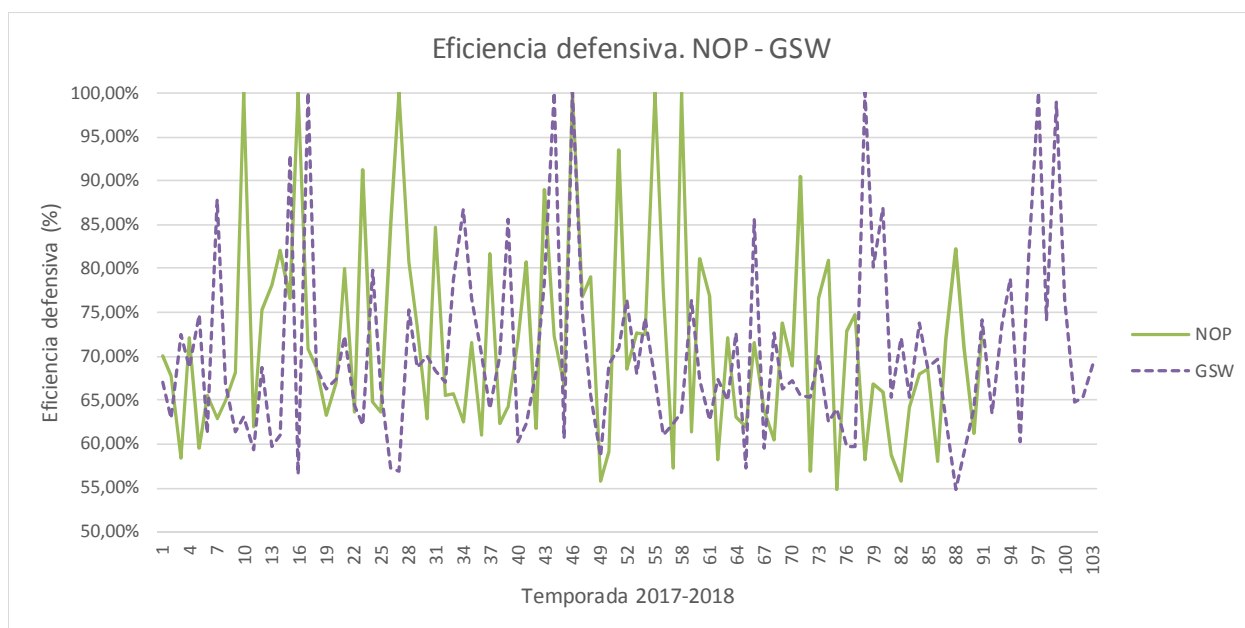
En esta gráfica podemos comprobar que la línea correspondiente a los HOU llega hasta el final, a diferencia de la de UTA y esto se debe a que los primeros son los que llegaron hasta la final de conferencia y se enfrentaron contra el otro vencedor de las semifinales de esta conferencia. Según la gráfica ambos equipos son eficientes desde el punto de vista defensivo en tan sólo dos ocasiones, por lo que no habría ninguna diferencia entre ambos. Además, podemos observar en qué momento de la temporada la eficiencia defensiva de cada uno de los equipos está por encima, y se aprecia que la de los Houston Rockets está en más ocasiones por encima que la de su contrincante, lo que se corresponde con la gráfica 5-1, donde podemos ver que la eficiencia media de HOU es superior a la de UTA.

Para complementar esta información, en la gráfica 5-7 se observa la línea de tendencia de la eficiencia defensiva a lo largo de la temporada de ambos equipos (HOU y UTA). Con estas líneas de tendencia vemos cómo la eficiencia defensiva de los Houston Rockets mantiene una línea de tendencia constante a lo largo de la temporada, mientras que la de Utah Jazz tiene una tendencia decreciente, lo que provoca que su promedio sea inferior a la de HOU.



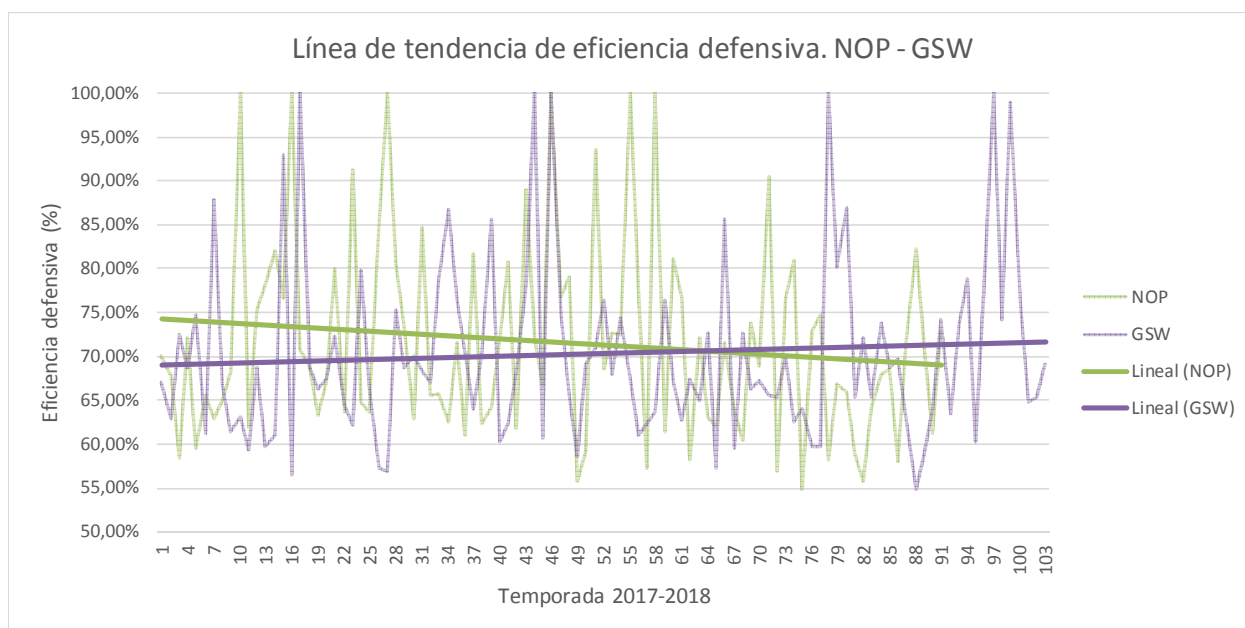
Gráfica 5-7. Línea de tendencia de eficiencia defensiva. HOU – UTA

En las gráficas siguientes se realiza un análisis similar a éste con los otros dos equipos que se enfrentaron en la semifinal de conferencia. En este caso, en la gráfica 5-8 se compara la eficiencia defensiva de los Golden State Warriors (GSW) y los New Orleans Pelicans (NOP). En ella podemos ver cómo el número de partidos en los que ambos equipos han sido eficientes en defensa es similar, siendo los NOP tan sólo eficientes desde el punto de vista defensivo en una ocasión más que los GSW. En este caso, es la línea correspondiente a los GSW la que llega al final de la gráfica, por ser los que pasaron a la final de conferencia, en la que se enfrentaron con los HOU.



Gráfica 5-8. Eficiencia defensiva. NOP – GSW

A continuación, en la gráfica 5-9 podemos observar la tendencia de la eficiencia defensiva de ambos equipos.

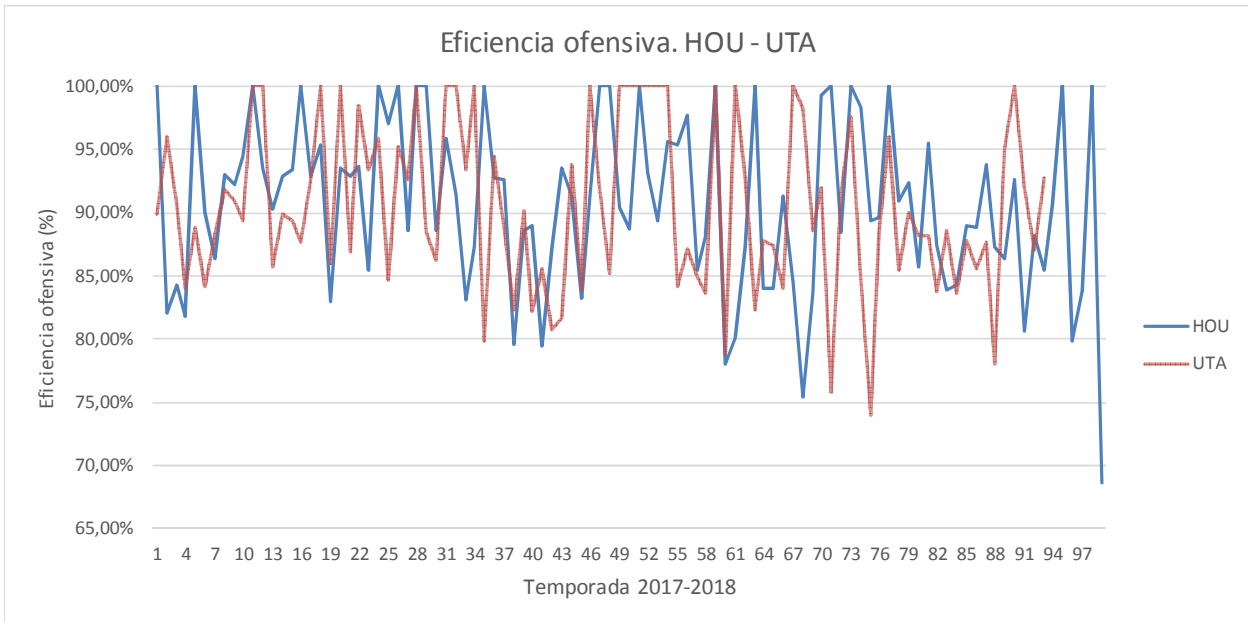


Gráfica 5-9. Línea de tendencia defensiva. NOP - GSW

En esta gráfica se observa la diferencia entre ambos equipos a lo largo de toda la temporada. A pesar de que el promedio de eficiencia defensiva es ligeramente superior en el caso de los NOP, tal y como se puede ver en la gráfica 5-1, su tendencia a lo largo de la temporada es decreciente, al contrario que en el caso de los GSW.

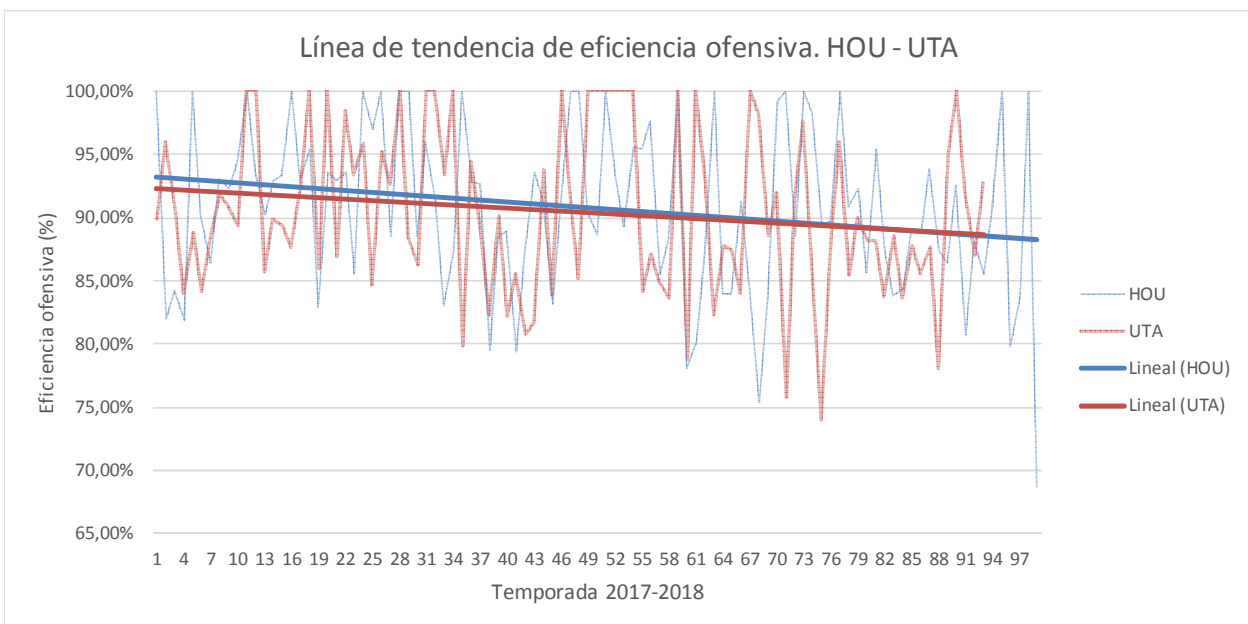
Tras estas observaciones de la eficiencia defensiva de la conferencia oeste pasamos a realizar el mismo análisis de la eficiencia ofensiva de los mismos equipos. En la gráfica 5-10 podemos ver la eficiencia ofensiva de los HOU y UTA. En ella se observa cómo la eficiencia ofensiva tiene valores superiores a

la defensiva y que se alcanza la eficiencia en muchas más ocasiones. En el caso de UTA vemos que llega a ser eficiente desde el punto de vista ofensivo de forma constante durante varios partidos hacia la mitad de la temporada. Podemos ver que durante la mayor parte de la temporada ambos equipos presentan una eficiencia superior al 80% y que los HOU alcanzan el valor de eficiencia más bajo.



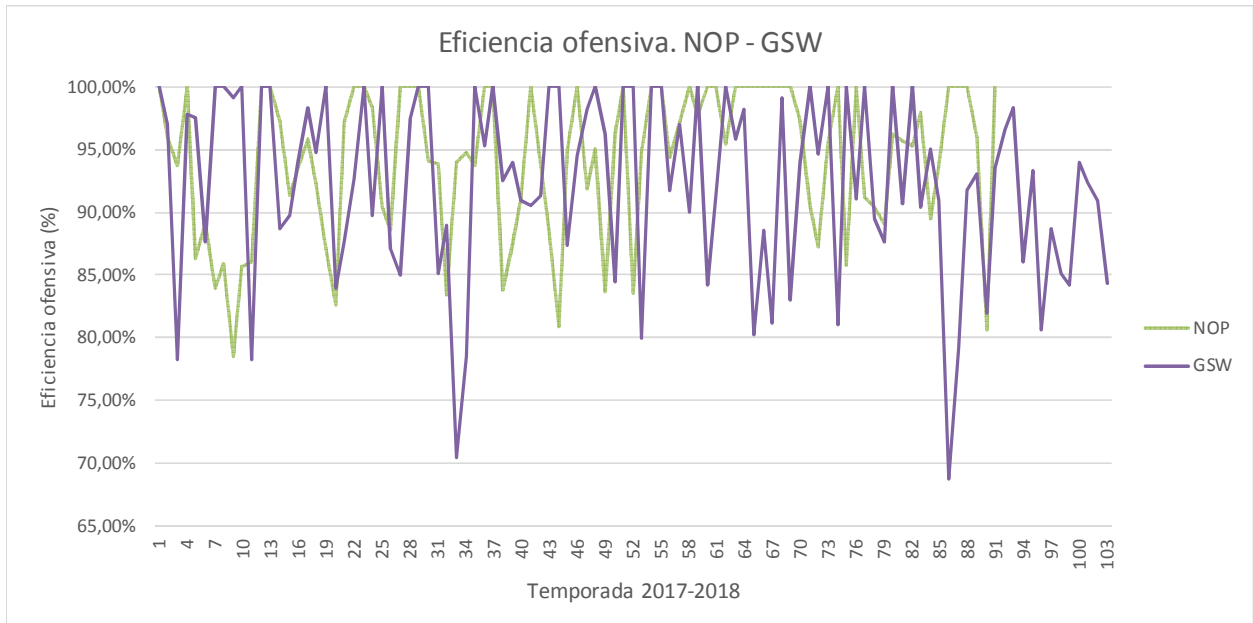
Gráfica 5-10. Eficiencia ofensiva. HOU – UTA

Tras analizar la eficiencia ofensiva vamos a ver en la gráfica 5-11 las líneas de tendencia correspondientes. En este caso, en ambos equipos la tendencia es decreciente y no hay una gran diferencia entre ambas, siendo ligeramente superior la de HOU a lo largo de la mayor parte de la temporada y llegando a unirse al final de la misma.



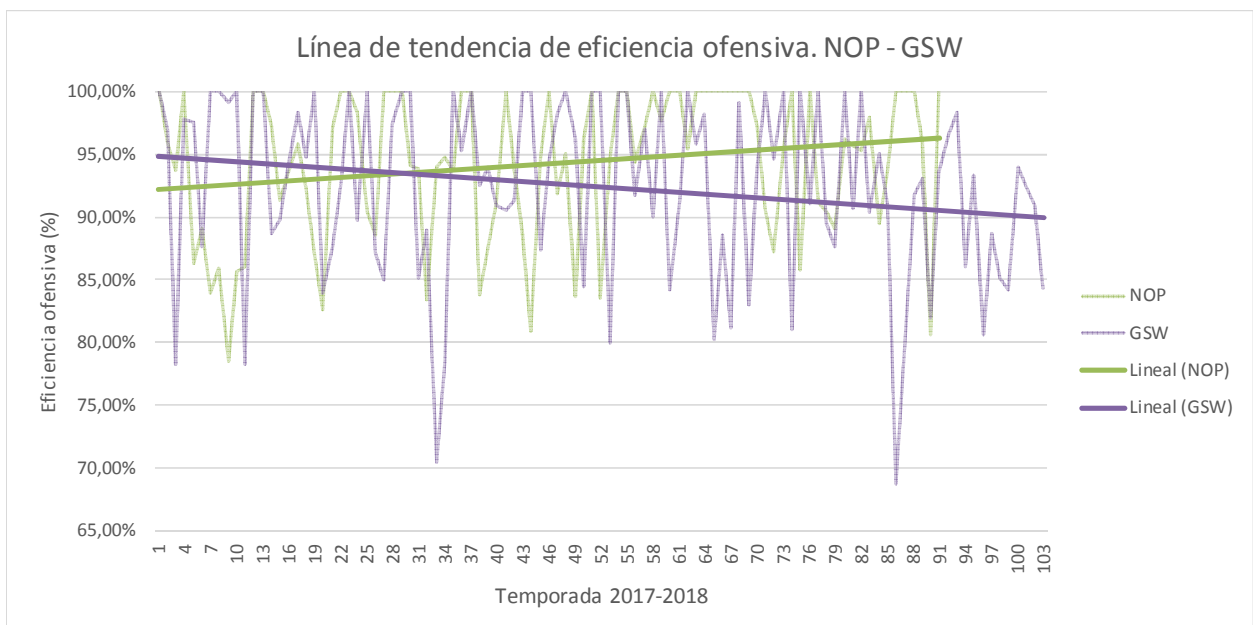
Gráfica 5-11. Línea de tendencia ofensiva. HOU - UTA

En las gráficas 5-12 y 5-13 se observa el mismo análisis de eficiencia ofensiva, pero en este caso de NOP y GSW.



Gráfica 5-12. Eficiencia ofensiva. NOP – GSW

De nuevo observamos que la mayoría de partidos se encuentran por encima del 80% de eficiencia y en muchas ocasiones se alcanza la eficiencia. No destaca ninguno de los equipos sobre el otro, ya que, según el momento de la temporada en el que estén, puede ser superior GSW o NOP.

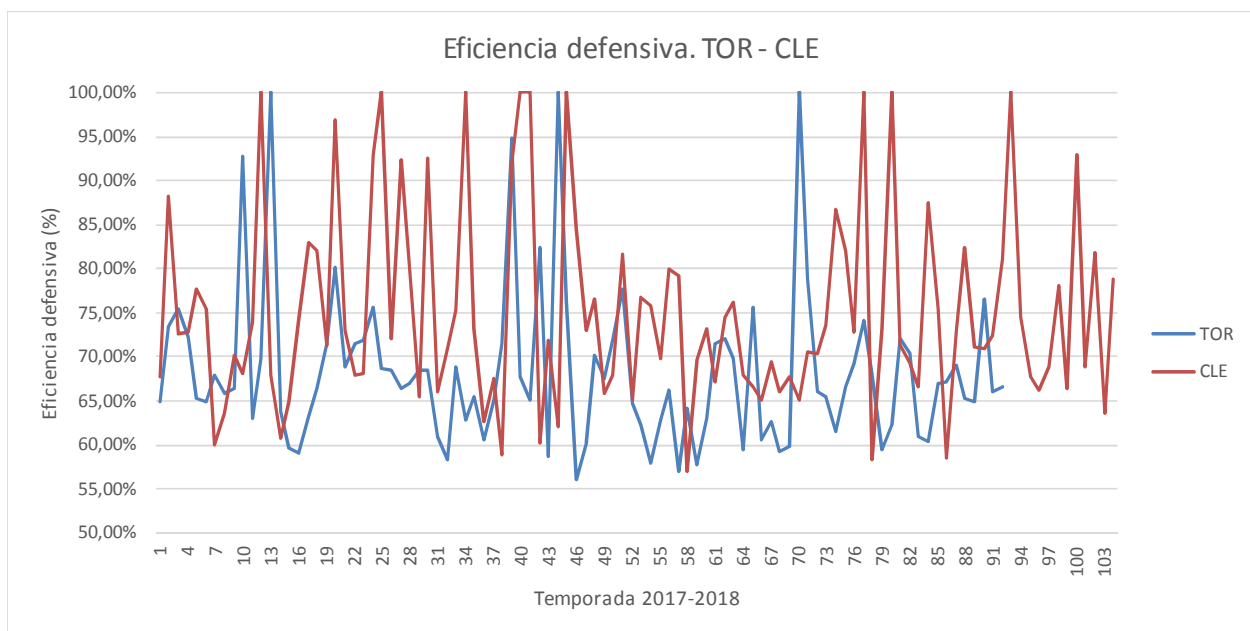


Gráfica 5-13. Línea de tendencia de eficiencia ofensiva. NOP - GSW

Para ver la evolución de la eficiencia ofensiva a largo de la temporada, en la gráfica 5-13 se puede ver las líneas de tendencia de ambos equipos. La tendencia de NOP es creciente, mientras que la de GSW es decreciente, quedando esta línea durante la mayor parte de la temporada por debajo de la de NOP. Esto es un aspecto que debemos destacar, ya que los Warriors son los ganadores de la temporada a pesar de que la tendencia de la eficiencia ofensiva a lo largo de la misma haya sido decreciente. Esta información podría resultar de interés a la hora de analizar la temporada con el objetivo de mejorar algunos aspectos a la hora de afrontar la siguiente, de forma que la eficiencia ofensiva tuviese una tendencia más favorable.

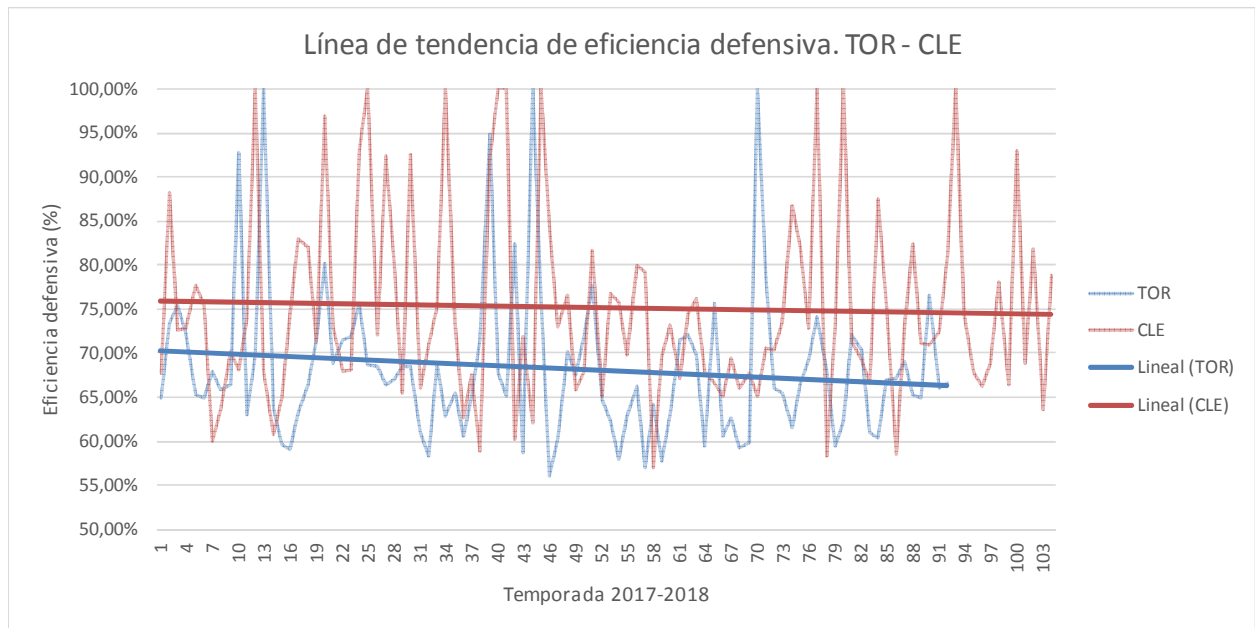
5.3.2 Conferencia este.

Al igual que ocurría con la conferencia oeste, se van a mostrar los resultados de forma que se puedan comparar la trayectoria de los equipos que se enfrentaron en las semifinales de conferencia. En primer lugar se va a analizar el aspecto defensivo de los cuatro equipos. En la gráfica 5-14 se puede comparar la eficiencia defensiva de los Toronto Raptors (TOR) y Cleveland Cavaliers (CLE). Todos los valores de eficiencia están por encima del 55% y en este caso vemos claramente que CLE supera en eficiencia defensiva a TOR durante la mayor parte de la temporada, teniendo muchos más partidos donde ha jugado de forma eficiente desde el punto de vista defensivo.



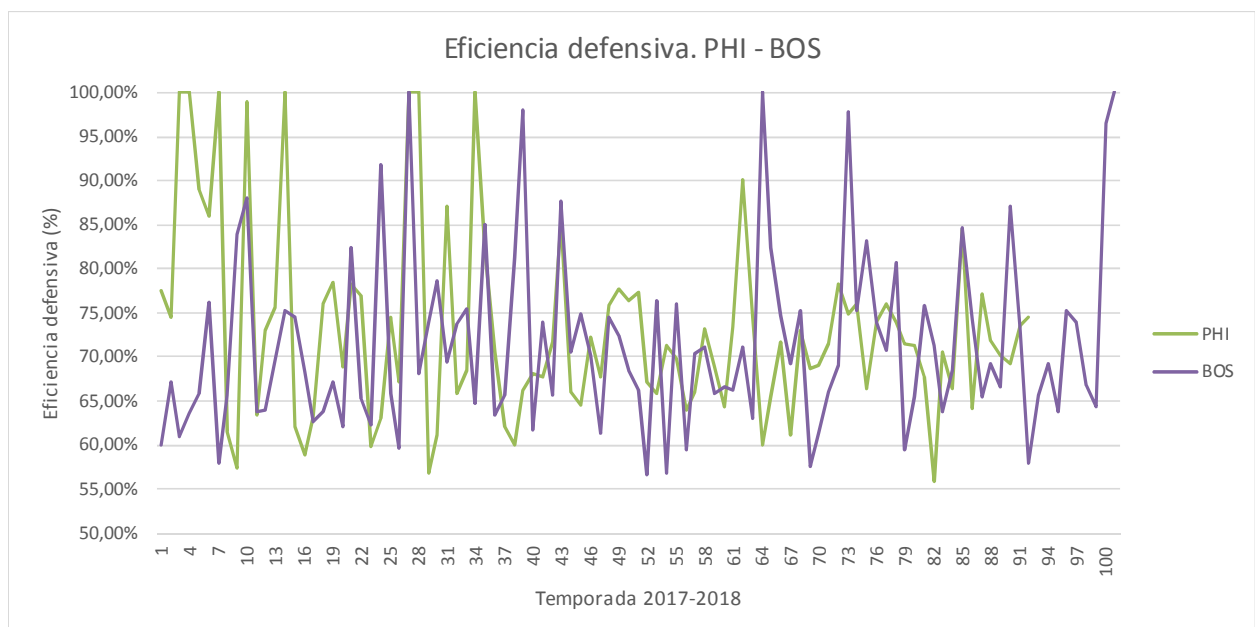
Gráfica 5-14. Eficiencia defensiva. TOR – CLE

Además, esta información se puede confirmar con la que se obtiene de la gráfica 5-15, donde se muestran las líneas de tendencia correspondientes a la eficiencia en defensa de ambos equipos durante la temporada. En esta gráfica se aprecia a simple vista una gran diferencia en las líneas de tendencia. A pesar de que la tendencia de ambas es similar y levemente decreciente, la línea correspondiente a CLE queda muy por encima de la de TOR. La estrategia a seguir en este caso puede servir de ejemplo, puesto que los Cleveland Cavaliers son los que jugaron la final de conferencia, pasando también esa fase de los playoffs y consiguiendo jugar la final de la temporada contra los GSW.



Gráfica 5-15. Gráfica de eficiencia defensiva. TOR – CLE

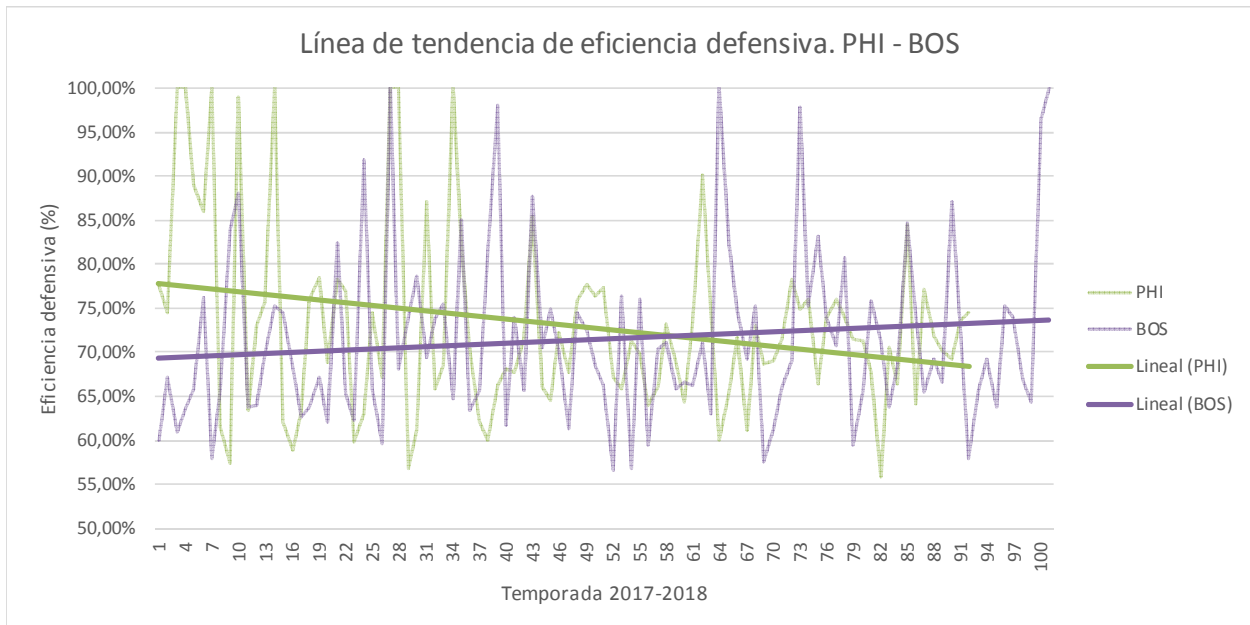
A continuación, se va a exponer el mismo análisis para los otros dos equipos de la conferencia este que jugaron la semifinal. En este caso, en la gráfica 5-16 se puede observar la trayectoria con respecto a la eficiencia defensiva de Philadelphia Sixers (PHI) y Boston Celtics (BOS) y en la gráfica 5-17 las líneas de tendencia asociadas a dichas eficiencias.



Gráfica 5-16. Eficiencia defensiva. PHI – BOS

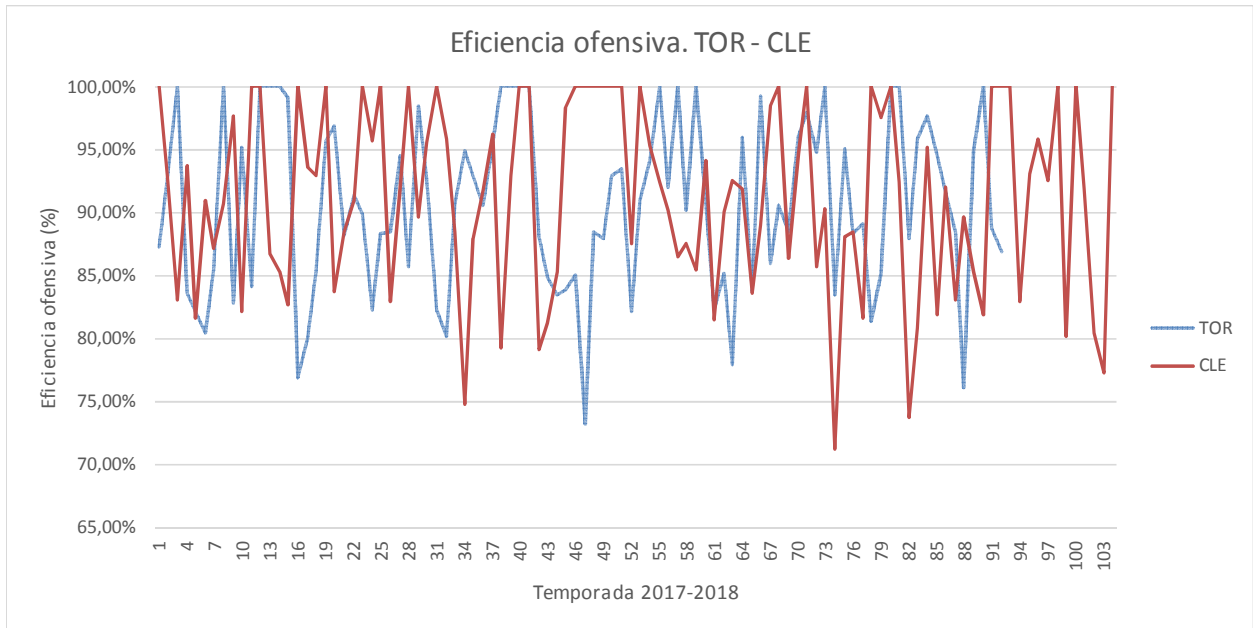
En la gráfica 5-16 podemos observar que, al igual que en los otros dos equipos, la eficiencia defensiva no baja del 55%. En este caso PHI ha sido eficiente en defensa en más ocasiones que BOS, a pesar de ser estos últimos los que pasaron a la final de conferencia. En relación a las líneas de tendencia (Gráfica

5-17), en este caso sí hay un cruce entre ambas, empezando PHI por encima de BOS pero terminando de forma contraria. Esto se corresponde con la gráfica anterior, puesto que el momento en el que los Philadelphia Sixers acumulan más partidos eficientes es precisamente al principio de la temporada, mientras que al final no hay ninguno (su línea de tendencia es decreciente). Con los Boston Celtics ocurre exactamente lo contrario, ya que su tendencia creciente implicaría una mejora de la eficiencia defensiva lo largo del transcurso de la temporada y es al final de la misma donde se observan mejores valores de eficiencia defensiva.



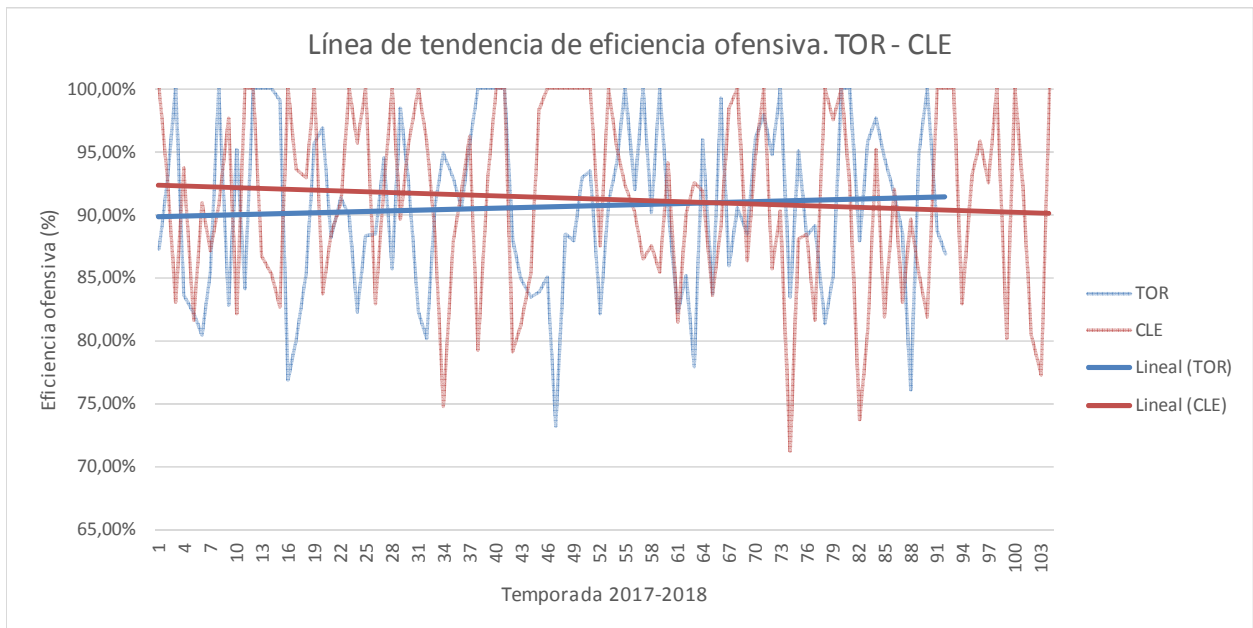
Gráfica 5-17. Línea de tendencia de eficiencia defensiva. PHI – BOS

Con respecto a la eficiencia de ataque, vamos a comenzar con la trayectoria de los Toronto Raptors y Cleveland Cavaliers. En la gráfica 5-18 se observa que, al igual que ocurría en la conferencia oeste, los valores de eficiencia ofensiva son superiores a los de eficiencia defensiva. Además, también hay una diferencia significativa en el número de partidos eficientes desde los dos puntos de vista. En este caso, ningún valor de eficiencia baja del 70% y podemos observar que CLE ha sido eficiente en ataque en más ocasiones que TOR, lo que podría haber provocado que fuesen los que pasasen a la final de conferencia.



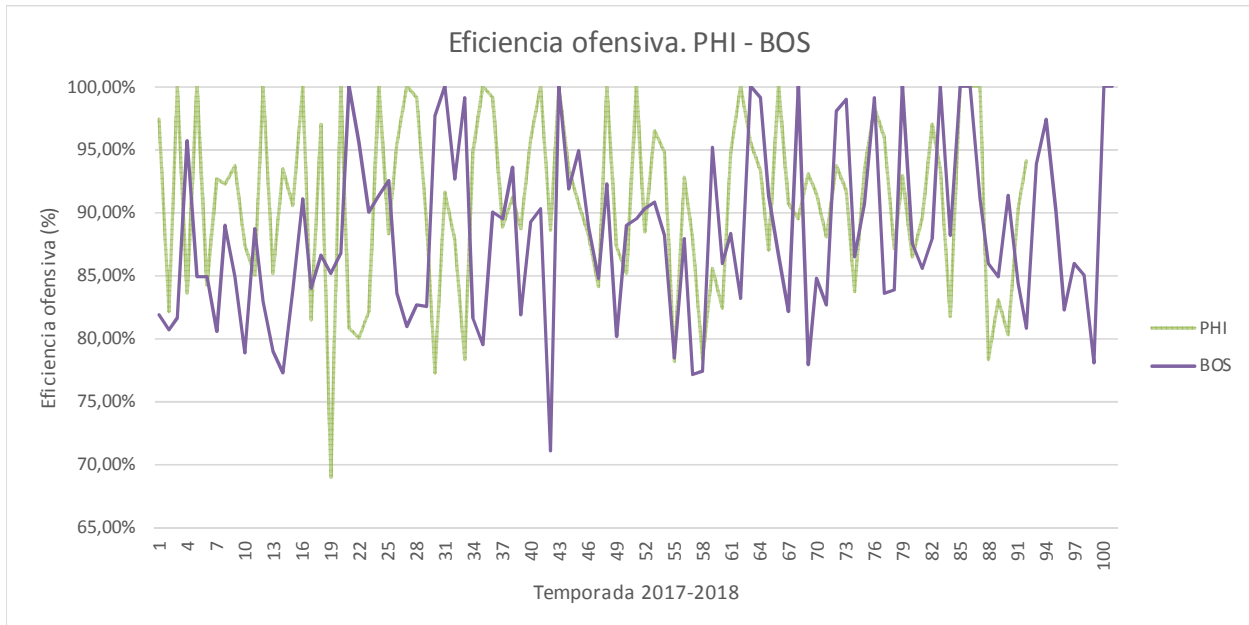
Gráfica 5-18. Eficiencia ofensiva. TOR – CLE

Sin embargo, la línea de tendencia correspondiente a CLE es ligeramente decreciente, mientras que la de TOR es creciente, llegando a estar por encima de la de CLE, como se puede observar en la gráfica 5-19. Hay que tener en cuenta que Cleveland Cavaliers registra un mayor número de eficiencias teniendo en cuenta la fase final de la temporada, en la que los Toronto Raptors quedaron fuera de la competición al ser vencidos por CLE en la semifinal de conferencias



Gráfica 5-19. Línea de tendencia de eficiencia ofensiva. TOR – CLE

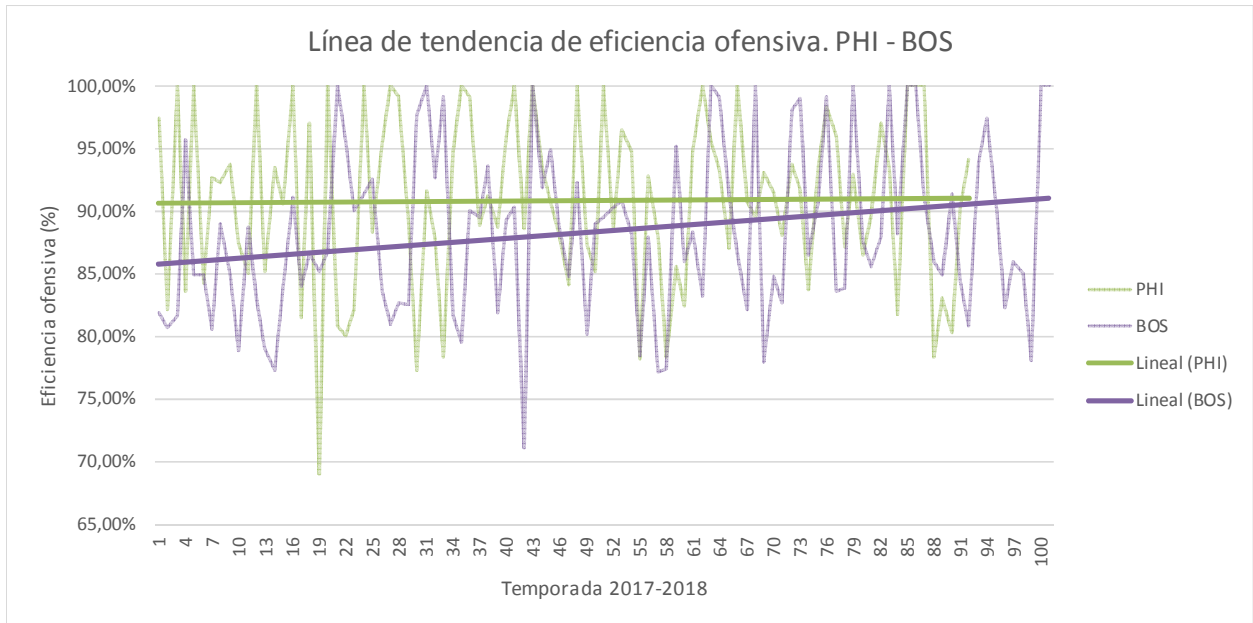
Con respecto a los otros dos equipos que jugaron la otra semifinal de conferencia, en la gráfica 5-20 se puede observar la eficiencia ofensiva de PHI y BOS. En ella vemos de nuevo que la eficiencia ofensiva no baja del 70%, dejando claro una vez más la importancia del ataque en un deporte como el baloncesto y, concretamente, en una liga como la NBA.



Gráfica 5-20. Eficiencia ofensiva. PHI - BOS

Además, se observa que el número de partidos eficientes de PHI es superior al de BOS, a pesar de ser estos últimos los que llegaron hasta la final de conferencia. En la gráfica 5-21 vemos las líneas de tendencia asociadas a estas eficiencias de ataque.

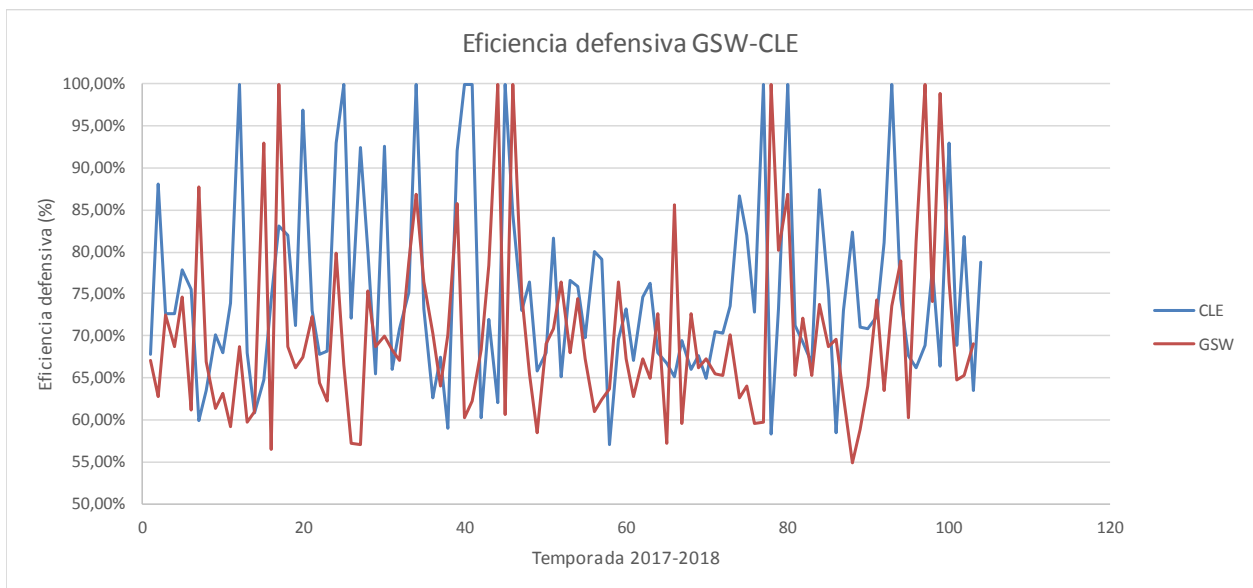
En ella vemos cómo, a pesar de que la eficiencia de ataque referente a BOS es creciente a lo largo de toda la temporada, siempre es inferior de la de PHI, que se mantiene constante y por encima de la de su contrincante, lo que implicaría mejores valores de eficiencia ofensiva y, por tanto un promedio de eficiencia de ataque más elevado que el de BOS. Si volvemos a la gráfica 5-2 podemos comprobar que la eficiencia media de ataque de PHI es superior a la de BOS. Aunque la diferencia no sea muy significativa confirmamos los resultados observados en esta gráfica.



Gráfica 5-21. Línea de tendencia de eficiencia ofensiva. PHI - BOS

5.4 Final de la NBA. Comparativa de los finalistas.

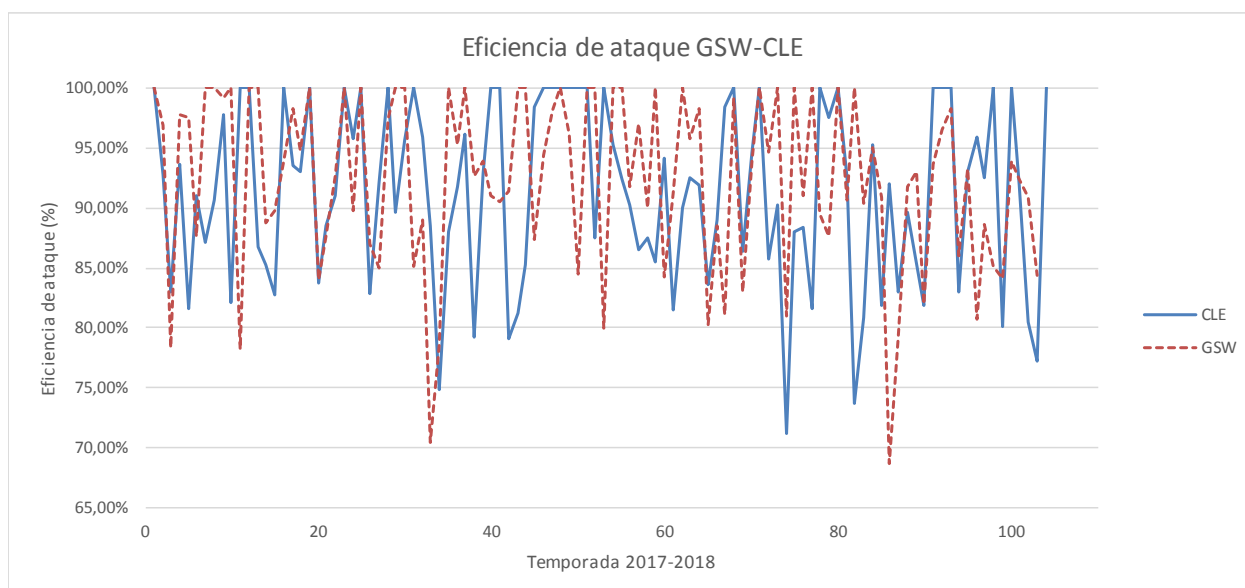
En la figura 5-1 podemos ver que la final de la NBA de la temporada 2017-2018 la disputaron Cleveland Cavaliers (equipo ganador de la conferencia este) y Golden State Warriors (equipo ganador de la conferencia oeste). En este apartado se va a realizar una comparativa similar a la del apartado anterior, pero teniendo en cuenta únicamente los equipos que jugaron la final.



Gráfica 5-22. Comparativa eficiencia defensiva GSW-CLE

En la gráfica 5-22 se muestra la evolución a lo largo de la temporada de la eficiencia en defensa obtenida para los GSW y los CLE. En ella podemos ver cómo los CLE alcanzan la eficiencia defensiva en más puntos que los GSW a lo largo de la temporada, según el modelo que hemos aplicado. Esto puede resultar interesante ya que, a pesar de que los Warriors fueron el equipo ganador de la temporada, en eficiencia defensiva están por delante los Cleveland Cavaliers.

Por otra parte, en la gráfica 5-23 se puede observar cómo se ha ido desarrollando la temporada en lo relativo a eficiencia en ataque por parte de estos dos equipos.



Gráfica 5-23. Comparativa eficiencia ofensiva GSW-CLE

En esta gráfica vemos que a lo largo de la temporada ambos equipos han sido eficientes durante gran parte de la misma desde el punto de vista ofensivo. De hecho, a simple vista nos costaría identificar si alguno ha sido más veces eficiente que el otro, ya que en muchos momentos podemos ver que ambos son igual de eficientes en el mismo momento. Si volvemos a la gráfica 5-4 comprobamos que ambos equipos cuentan con el mismo número de partidos eficientes. Se puede destacar que al final de la temporada la eficiencia en ataque de los Warriors es considerablemente menor que la de los Cleveland Cavaliers, siendo GSW los ganadores de la liga. Esta información podría ser útil a la hora de estudiar la estrategia seguida por GSW a lo largo de la temporada y poder comparar momentos donde el equipo estaba siendo eficiente con aquellos otros donde se ve una clara disminución de la eficiencia del equipo en lo relativo al ataque.

Sin embargo, debemos tener en cuenta que aunque en este gráfico se comparan exclusivamente estos dos equipos, el modelo calcula la eficiencia teniendo en cuenta la totalidad de partidos jugados, tanto en la liga regular como en los playoffs y comparándose con todos ellos, por lo que para intentar encontrar el origen de esta disminución de eficiencia ofensiva en el final de la temporada no podemos centrarnos exclusivamente en las diferencias de juego entre los Cleveland Cavalier y los Golden State Warriors.

Toda esta información que se va obteniendo puede tener gran utilidad a la hora de elaborar estrategias

para la siguiente temporada, ya que se pueden comparar los resultados de partidos en los que el equipo ha jugado de manera eficiente con otros partidos en los que el equipo no ha sido eficiente y a partir de ahí establecer nuevas formas de juego, de manera que se puedan aprovechar mejor los recursos y obtener mejores resultados, incluso habiendo ganado.

En muchas ocasiones el equipo victorioso no actúa de forma eficiente según el modelo con el que hemos trabajado. De hecho, en el caso de los Golden State Warriors sus promedios de eficiencia, tanto defensiva como general, no son de los más altos de la NBA, lo que podría llevar a hacer un análisis de cómo se está jugando y qué recursos se podrían aprovechar mejor para ser más eficientes y obtener aún mejores resultados.

Por otra parte, otro aspecto que hemos estado observando de forma recurrente en las gráficas que se han mostrado hasta ahora, es la poca importancia que parece tener la estrategia defensiva para los equipos, lo que puede llevar a la conclusión de que el baloncesto se juega de forma que se prioriza el ataque y la posesión de balón. Sin embargo, en algunas ocasiones se podría estudiar la posibilidad de conseguir mejores resultados mejorando la eficiencia defensiva en lugar de priorizar el ataque y la posesión de balón.

5.5 Análisis individual de los tiempos de defensa y ataque.

En el apartado 4 se expuso tanto el modelo que se iba a implementar como las variables necesarias para dicho modelo. Una de estas variables es la posesión de balón, considerándose ésta como tiempo de ataque del equipo. Debido a que el análisis de eficiencia que se está llevando a cabo está basado en la eficiencia en cuanto a tiempo de defensa y ataque de cada partido por parte de los dos equipos, éste era un dato fundamental que necesitábamos para llevar a cabo la aplicación de este modelo y que no encontramos recogido de forma directa en las fuentes de donde se recopilaron el resto de datos.

Existen varias fuentes en internet donde podemos encontrar una gran cantidad de datos de los dos equipos que juegan un partido determinado. Además, existen distintos tipos de datos, ya que podemos encontrarlos por equipo, por jugador y otro tipo de estadísticas ya realizadas en dichas fuentes. Sin embargo, en ninguna de ellas se lograba hallar información relativa a este porcentaje de tiempo con respecto al total del partido que se dedicaba al ataque o a la defensa.

Debido a ello, tras buscar en distintas fuentes e incluso intentar ponernos en contacto con alguna de ellas (incluyendo la página web oficial de la NBA) sin ningún resultado, nos hemos basado en un estudio publicado en la revista argentina EFDeportes y realizado en conjunto por Alberto Lorenzo Calvo y Miguel Ángel Gómez Ruano, ambos de la Universidad Politécnica de Madrid y Antonio Jaime Sampaio de la Universidade de Tras-os-Montes e Alto Douro de Portugal [Lorenzo, et. al, 2003].

Dicho estudio se realizó con el objetivo de estudiar las posesiones y su utilización por los equipos durante los partidos de baloncesto. Está basado en el análisis de 10 partidos de la Liga ACB y Euroliga, recogiendo información de qué segundo de la posesión es en el cual se finalizaba el ataque de cada equipo y si éste terminaba con un tiro de dos o tres puntos, o si terminaba con una falta recibida o la pérdida del balón. Todo esto se estudiaba de forma independiente por cuartos, por equipos y por intervalos de tiempo.

Para llevarlo a cabo, el estudio hace una revisión al mismo tiempo de distintos estudios estadísticos en los que se plantean fórmulas y definiciones que pueden ser de interés. Entre estos estudios estadísticos

resulta de especial interés para este trabajo el de Sampaio y Janeira (2001), los cuales analizan qué datos llevan a un equipo a ser más eficaz que otros y determinan, entre otras, una fórmula que permite calcular la posesión de balón de un equipo en función de otros datos recogidos al finalizar un partido de baloncesto. Esta fórmula es la siguiente:

$$PB = LCT + \frac{LTL}{2} + PDB$$

Donde:

- PB: posesión de balón.
- LCT: lanzamientos intentados.
- LTL: lanzamientos tiros libres.
- PDB: pérdidas de balón. En este caso tenemos en cuenta las pérdidas de balón del equipo contrario, ya que son las que van a provocar que la posesión de balón de un equipo aumente y, por tanto, el porcentaje de tiempo que se considera que dicho equipo está atacando.

Esta fórmula, reescrita con las siglas de las variables de nuestro modelo quedaría de la siguiente manera:

$$\%TIEMPO = FGA + \frac{FTA}{2} + TO \text{ (en contra)}$$

En esta fórmula nos hemos basado para conseguir el porcentaje de tiempo de ataque que necesitamos para cada equipo en cada partido, considerando el tiempo de posesión de balón como tiempo en el que un equipo está atacando. Se ha calculado la posesión de balón de cada equipo en cada partido y, posteriormente, el porcentaje al que corresponde del total. Este resultado es el que vamos a llamar tiempo real a la hora de desarrollar este apartado.

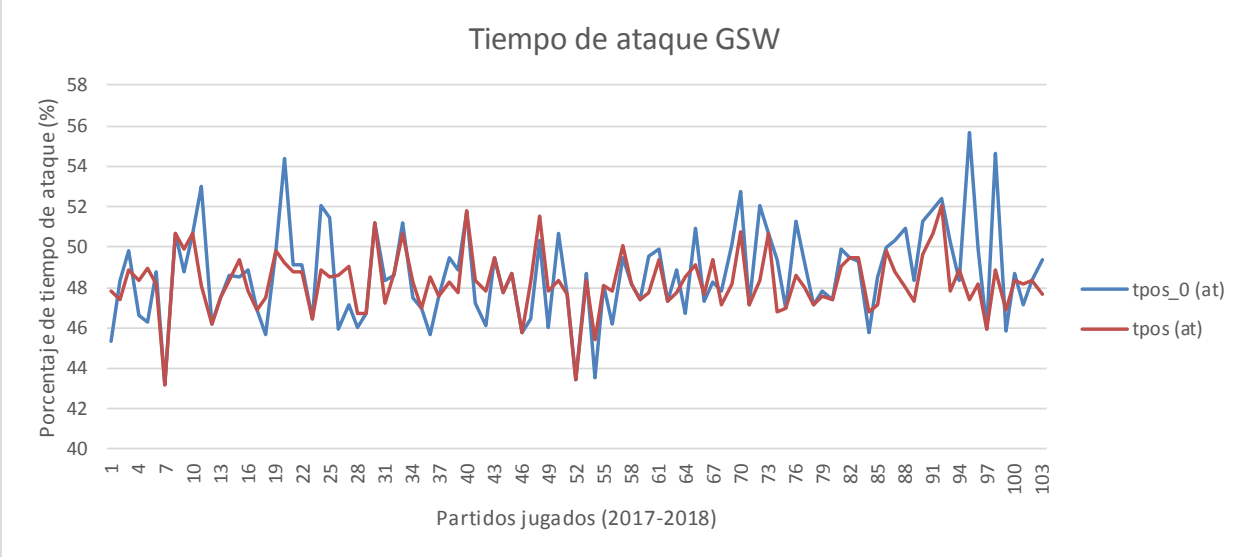
Hasta ahora, en el análisis de los resultados del problema se han mostrado promedios de eficiencia de ataque, de defensa y general, así como la frecuencia de dichas eficiencias por parte de cada equipo que forma parte de la NBA. Sin embargo, en ningún momento se ha hablado acerca de estos tiempos óptimos de defensa y ataque que también calcula el modelo.

Tal y como se explicó en el apartado 6, este modelo trabaja en dos bloques aparentemente independientes que se relacionan gracias a una variable de tiempo de defensa o ataque en cada partido y en este apartado se va a hacer un análisis de este tiempo óptimo obtenido a partir del modelo y vamos a compararlo con los datos de tiempo de ataque y defensa reales que resultaron en cada caso.

Vamos a hacer esta comparación con los mejores en cada conferencia, es decir, aquellos que jugaron la final de la NBA.

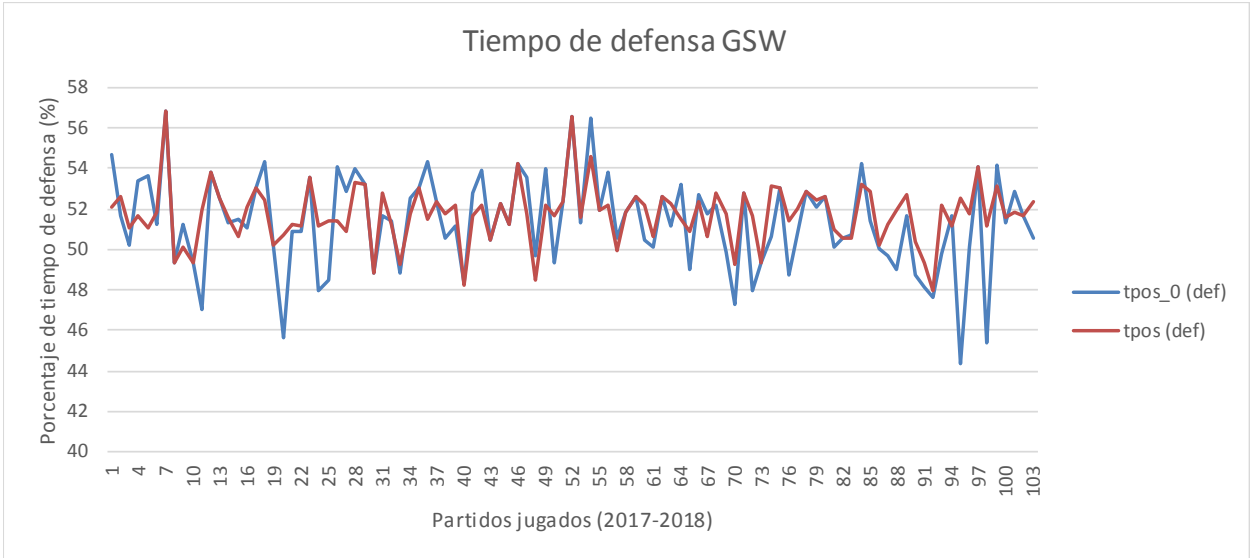
En primer lugar vamos a analizar a los Golden State Warriors. En la gráfica 5-24 se puede observar una gráfica que abarca todos los partidos jugados por los Warriors. En esta gráfica “tpos_0 (at)” nos indica los porcentajes de tiempo de ataque de los GSW a lo largo de la temporada y calculados a partir de la posesión de balón. Por otra parte, “tpos (at)” indica el porcentaje de tiempo óptimo en cada momento de la temporada calculado a partir del modelo que se ha implementado. En muchos casos

vemos que pueden coincidir, mientras que en otros el tiempo dedicado a ataque o la posesión de balón debería haber sido menor que el que se ha empleado realmente para lograr la eficiencia. Destaca sobre todo el final de la temporada, donde el tiempo de ataque eficiente es considerablemente menor al que fue realmente, lo que podría servir para analizar esos partidos y qué diferencias ha habido con otros en los que sí se alcanzaba la eficiencia ofensiva.



Gráfica 5-24. Comparativa de tiempo de ataque real y eficiente. GSW

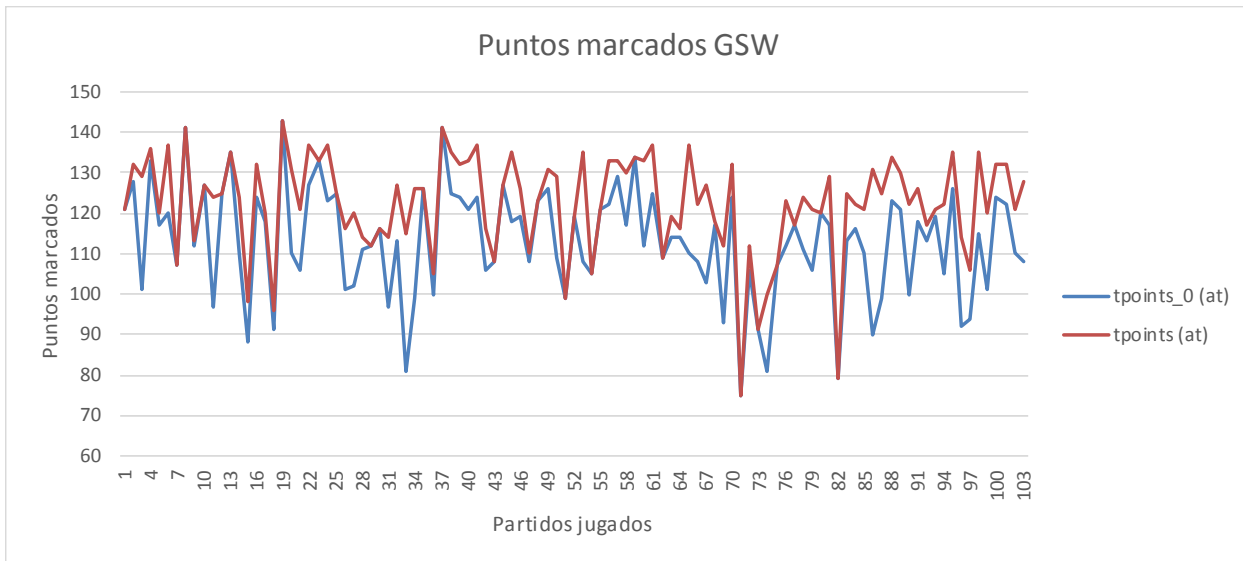
Por otra parte, la gráfica 5-25 es análoga a la anterior, pero considerando en este caso el porcentaje de tiempo total empleado en defensa. De nuevo, la etiqueta “tpos_0 (def)” está referida al porcentaje de tiempo real empleado en defender, mientras que “tpos (def)” es el tiempo calculado a partir de nuestro modelo que resultaría eficiente en cada partido. Realmente ambos porcentajes de tiempo (el real y el eficiente) no difieren en gran medida en muchos momentos de la temporada, a pesar de que la eficiencia defensiva que hemos calculado y analizado en otras gráficas durante este análisis de resultados no era muy elevada.



Gráfica 5-25. Comparativa de tiempo de defensa real y eficiente. GSW

Para complementar esta información hemos hecho uso de las holguras, las cuales son s_0^{at} y s_0^{def} , de forma que hemos calculado los puntos que deberíamos haber marcado y los puntos que deberíamos haber tenido en contra en cada partido. A partir de este resultado hemos realizado dos gráficas. En primer lugar, vamos a exponer la gráfica 5-26, donde se muestra una comparativa de los puntos reales obtenidos en cada partido, indicados con la etiqueta “tpoints_0 (at)” y los puntos que, según los resultados de eficiencia obtenidos tras implementar el modelo, deberíamos haber obtenido.

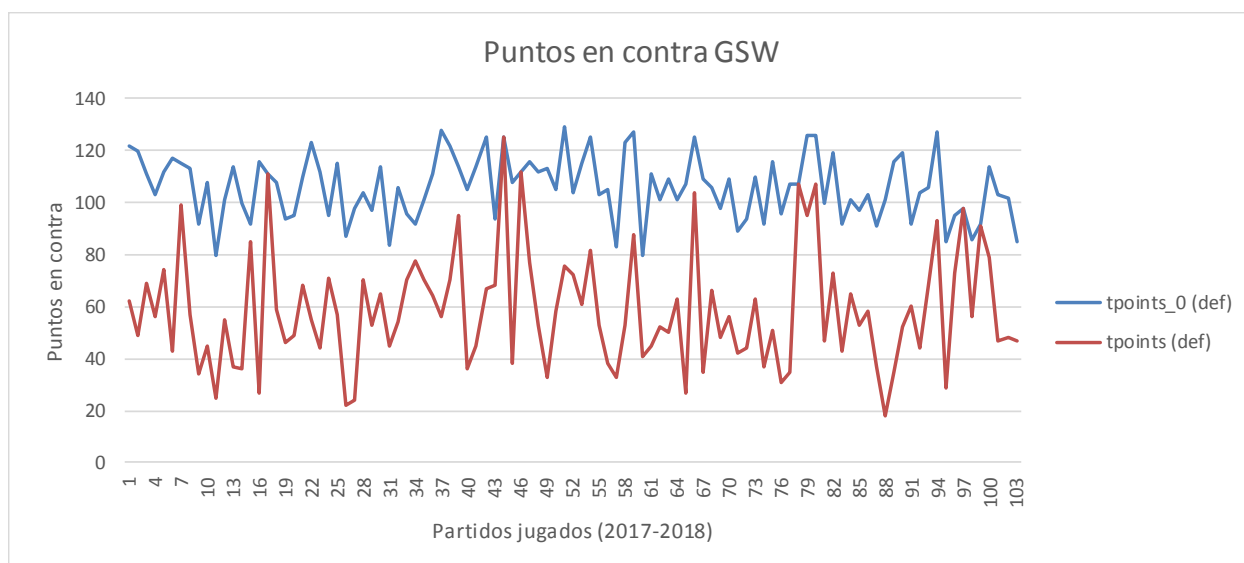
Estos puntos que deberíamos haber obtenido están indicados con la etiqueta “tpoints (at)” y los hemos calculado de la siguiente forma: $tpoints (at) = tpoints_0 (at) + s_0^{at}$.



Gráfica 5-26. Comparativa de puntos marcados GSW

En esta gráfica ambas líneas coincidirán siempre que el valor de la holgura s_0^{at} sea cero y, por tanto, que el equipo haya sido eficiente en ataque en ese partido. En caso contrario, el valor de la holgura siempre será distinto de cero y “tpoints (at)” quedará por encima de “tpoints_0 (at)”.

Esto mismo lo hemos analizado para los puntos en contra en cada partido jugado por los Warriors, tal y como se observa en la gráfica 5-27. En este caso, la etiqueta “tpoints_0 (def)” hace referencia a los puntos en contra de los Warriors en cada partido, mientras que “tpoints (def)” son los puntos que debería haber tenido en contra a lo largo de la temporada y calculados de nuevo a partir de las holguras de la siguiente forma: $tpoints (def) = tpoints_0 (def) - s_0^{def}$.

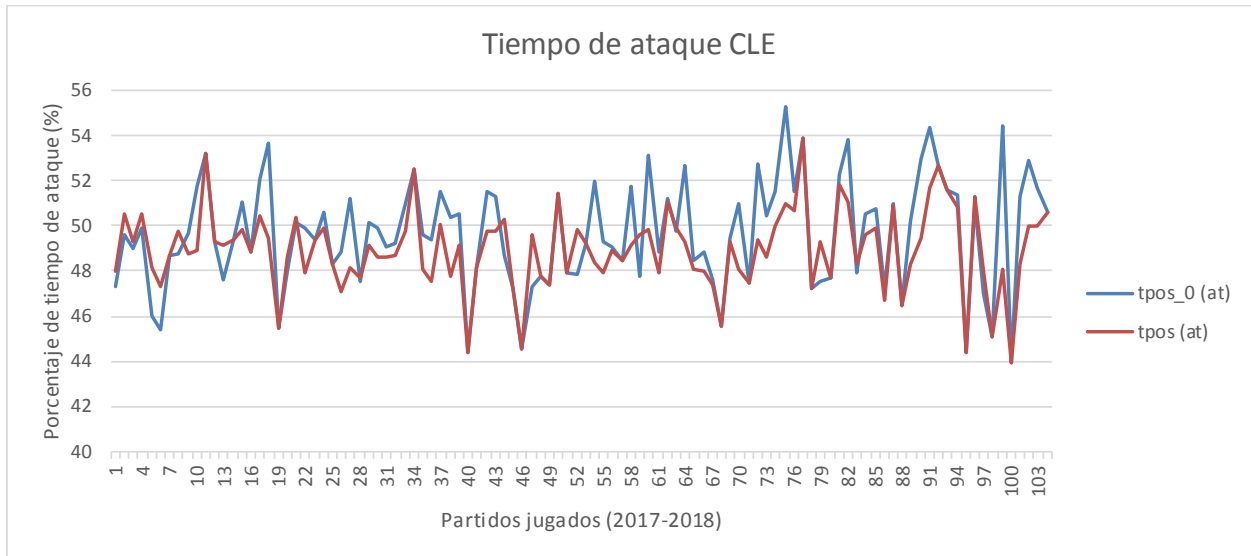


Gráfica 5-27. Comparativa de puntos en contra GSW

En este caso, sí vemos una diferencia notable entre los puntos que el equipo contrario ha marcado y los que debería haber marcado en cada partido para alcanzar la eficiencia. Debido a esta diferencia entre ambos valores en la gran mayoría de partidos jugados, la eficiencia defensiva calculada y mostrada en gráficos anteriores es tan baja, a pesar de haber visto en la gráfica 5-25 que el tiempo eficiente de defensa no se alejaba mucho del realmente empleado en ello. Esta información puede ser muy útil por parte de los entrenadores para trabajar en aprovechar mejor y de forma más eficiente los momentos en los que la posesión del balón la tiene el equipo contrario.

Una vez hecho este análisis para los Golden State Warriors, hemos realizado las mismas gráficas para los Cleveland Cavaliers. Las etiquetas usadas en las gráficas han seguido el mismo criterio que en las de los Warriors.

Así, en la gráfica 5-28 se puede ver de nuevo una comparación entre el tiempo real en el que el equipo estaba atacando (tpos (at)) y el tiempo que debería haberse destinado al ataque, según las variables consideradas y el modelo implementado. Podemos ver que ambas líneas no difieren en gran medida. Incluso al final de la temporada, vemos que a pesar de que el tiempo de ataque eficiente debería haber sido menor al que fue realmente, la diferencia es menor que en el caso de los Warriors.



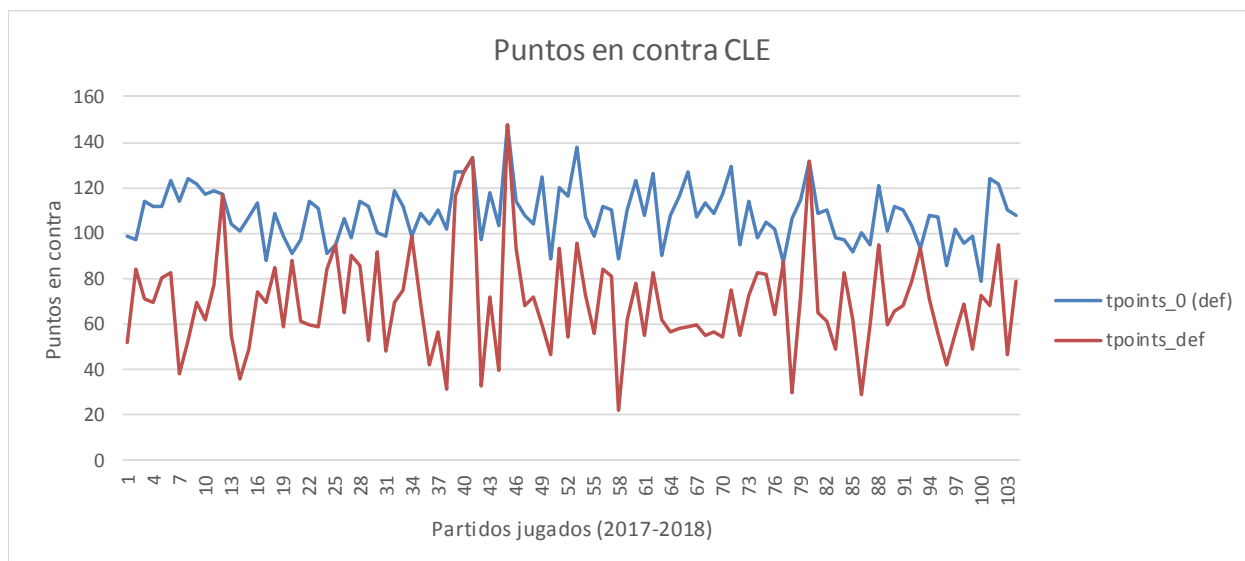
Gráfica 5-28. Comparativa de tiempo de ataque real y eficiente. CLE

A continuación, se muestra una gráfica similar a la anterior, pero en este caso se compara el tiempo de defensa real con el que se considera que es eficiente según los resultados obtenidos.



Gráfica 5-29. Comparativa de tiempo de defensa real y eficiente. CLE

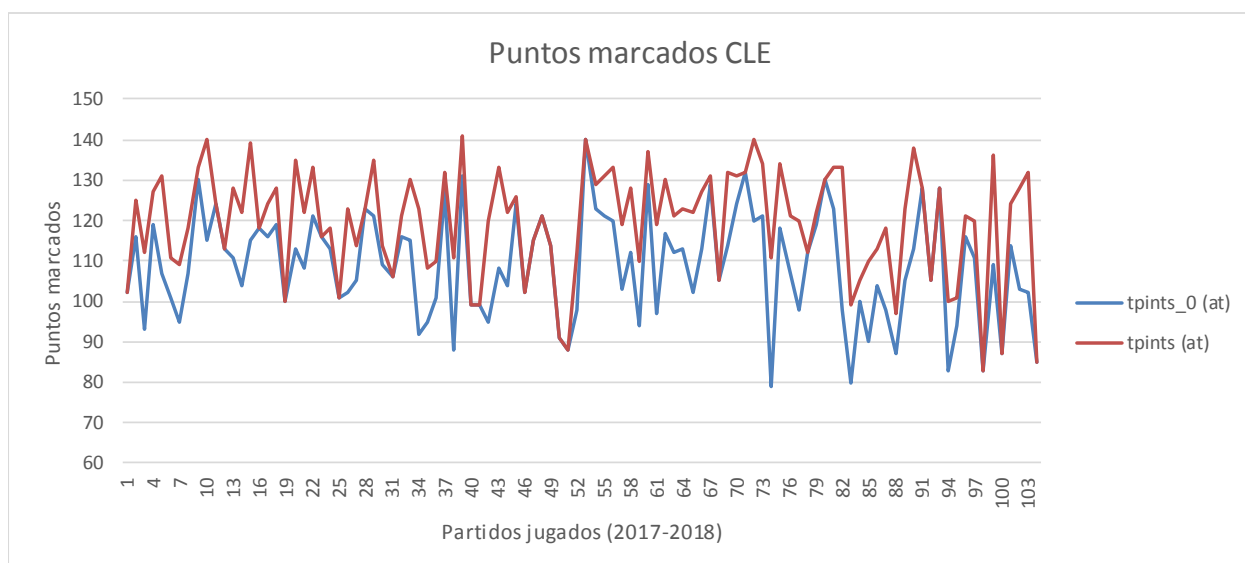
Como podemos observar en la gráfica 5-29 en muchos momentos de la temporada el porcentaje de tiempo en el que el equipo está defendiendo coincide con el eficiente. En este aspecto se asemeja a lo que ocurría con la gráfica de los Warriors. Esta información puede resultar útil a la hora de perfeccionar la estrategia a seguir en determinados momentos claves o determinados partidos de la temporada donde el resultado no haya sido el esperado. Sin embargo, en el caso de los Cleveland Cavaliers el promedio de eficiencia defensiva es notablemente superior que el de los Golden State Warriors, por lo que debe notarse esta discrepancia a la hora de analizar la diferencia entre los puntos en contra recogidos como datos durante toda la temporada y los que han sido calculados por el modelo. Esta información se observa en la gráfica 5-30, que se muestra a continuación.



Gráfica 5-30. Comparativa de puntos en contra CLE

En este caso, tal y como se observa en esta gráfica, la diferencia entre los puntos en contra reales y los que se deberían haber obtenido es importante, pero en comparación con los Warriors y con la gráfica 5-27 esta diferencia se ha reducido notablemente, lo que provoca que los Cleveland Cavaliers se muestren superiores a los Golden State Warriors en cuanto a eficiencia defensiva. Como consecuencia de esto, debido a que la eficiencia en ataque es similar en ambos equipos, esta diferencia en la eficiencia defensiva provoca que en el gráfico de eficiencia media los Cleveland Cavaliers presenten un mejor comportamiento. Debemos comentar que siempre que las dos líneas coincidan en algún punto implica el valor nulo de la holgura s_0^{def} y como consecuencia la eficiencia en defensa del equipo en el partido correspondiente.

Por último, vamos a ver la diferencia de puntos que han marcado los Cleveland Cavaliers a lo largo de la temporada 2017-2018 y los que deberían haber marcado si hubiesen sido eficientes. Tal y como ocurre con la gráfica anterior, si en la gráfica 5-31 las dos líneas coinciden en algún punto significa que en ese partido el equipo ha sido eficiente y por tanto el valor de la holgura s_0^{at} es nulo.



Gráfica 5-31. Comparativa puntos marcados CLE

La diferencia de puntos marcados con los que deberían haber obtenido los Cleveland Cavaliers es bastante inferior a lo que ocurría con los puntos en contra. Esto es algo que se repite a lo largo de todo este apartado, puesto que esto sólo demuestra una vez más que la eficiencia en ataque va a ser superior a lo largo de la temporada que la eficiencia en defensa. De hecho, en este caso, en muchos puntos de la gráfica ambas líneas coinciden, lo que significa que el equipo es eficiente en ataque según el modelo usado en gran parte de los partidos jugados.

Para terminar, podría resultar extraño que teniendo ambos equipos un promedio en eficiencia en ataque similar, el que tiene un promedio de eficiencia defensiva superior no sea el ganador de la temporada, ya que esto es lo que los diferencia según los resultados obtenidos.

Esto puede deberse por una parte a que influyan además de las variables consideradas, otros factores externos a este modelo, y por otra parte a que los valores relacionados con la eficiencia defensiva realmente son considerablemente más bajos que los resultados obtenidos en eficiencia en ataque, por lo que el hecho de que un equipo tenga un valor mucho más alto que otro en eficiencia defensiva realmente no suponga gran diferencia a la hora de los resultados reales. Sin embargo, esto es algo que cada equipo se podría plantear, tal y como se ha ido indicando, ya que potenciar la eficiencia defensiva y mejorar este aspecto puede ayudar a conseguir los resultados esperados o conseguir los mismos resultados de manera más eficiente, lo que podría implicar regular la actividad de determinados jugadores, o potenciar la actividad en el equipo de otros que a priori no se considerasen importantes.

6 RESUMEN Y CONCLUSIONES

Los objetivos de este trabajo eran, por una parte, mostrar la creciente importancia del análisis de datos en el mundo deportivo y, por otra, resolver un problema de eficiencia en la NBA usando DEA como herramienta de resolución y, de esta forma, mostrar un ejemplo de cómo la tecnología y el deporte están relacionados y qué información puede llegar a obtenerse con un análisis de datos adecuado.

En primer lugar, durante el desarrollo de este documento y antes de plantear nuestro problema concreto de la NBA se han mostrado numerosos casos en los que la tecnología puede servir de apoyo al sector deportivo. En la introducción al problema hemos visto una presentación de la investigación operativa y de muchas de sus aplicaciones, de las que destacamos la deportiva. Se han expuesto numerosos casos de aplicaciones concretas en distintos deportes, además de ver su crecimiento y el alcance que tiene actualmente, presentando diferentes artículos y revistas de interés donde se pueden encontrar estudios sobre la tecnología y el deporte desde distintos puntos de vista. Se ha mostrado esta integración, además de ser conscientes de que existe un amplio mercado alrededor del análisis de datos aplicado a este sector. Así, existen empresas que se dedican exclusivamente a este campo y equipos y organizaciones de los diferentes deportes integran cada vez a más ingenieros, matemáticos y estadísticos en su personal que elaboren tareas que complementen a los entrenadores y a la gestión de los equipos de las ligas más importantes.

Teniendo en cuenta esto, hemos realizado un análisis de eficiencia de la NBA de la temporada 2017-2018. En él se ha aplicado un modelo con las variables más apropiadas y haciendo uso de la metodología DEA se han obtenido unos resultados que hemos analizado posteriormente. Como hemos podido comprobar, a partir de estos resultados se pueden obtener múltiples y diversos análisis. Además de realizar algunas comparaciones generales de todos los equipos, también se puede hacer un análisis más detallado de cada uno de ellos en relación a los tiempos óptimos de ataque y defensa y a los puntos que se deberían haber obtenido siendo eficientes. En este caso, hemos hecho este análisis individual con los dos equipos que jugaron la final de la temporada (GSW y CLE). A partir de esta información, un entrenador podría comparar distintos partidos en los que la eficiencia ha sido diferente y analizar la estrategia seguida con el objetivo de mejorar resultados y obtener conclusiones en cuanto a la forma de juego conveniente en cada caso y con cada contrincante. Como hemos dicho, este análisis igual que se ha realizado con estos dos equipos se podría realizar con cualquier otro de la NBA. Esto es un ejemplo más de cómo la tecnología y una herramienta de análisis de eficiencia puede ayudar a mejorar los resultados deportivos de un equipo concreto. Al igual que se ha realizado este análisis de eficiencia existen numerosas herramientas y cada vez tecnología más avanzada que hacen del deporte un sector cada vez más competitivo, ya que si no se hace uso de estos avances es difícil ser competitiva en las distintas ligas profesionales.

En este caso concreto, tras realizar el análisis de resultados, una de las principales conclusiones sería la importancia que se le da al ataque por encima de la defensa. En los gráficos generales vemos cómo el promedio de eficiencia ofensiva tiene una diferencia de un 20% tanto en el valor más alto como en el más bajo con respecto al promedio de eficiencia defensiva. Además, esto puede verse reflejado a lo

largo de todo el análisis de resultados. Una vez finalizado éste hemos contrastado esta información y se ha observado que en los últimos años se habla en la NBA de la explosión ofensiva y cómo es cada vez más difícil elaborar una buena estrategia defensiva que haga frente a las estrategias de ataque llevadas a cabo, ya que es lo que están priorizando los equipos de esta liga de baloncesto. Esta información podemos encontrarla en revistas como ESPNDEPORTES, MARCA o incluso en artículos publicados en la página oficial de la NBA.

El modelo aplicado podría servir de apoyo para mejorar estas estrategias defensivas estudiando partidos en los que se considera que un equipo ha sido eficiente desde este punto de vista y comparándolo con otros en los que no lo haya sido. Además, es un modelo que podría usarse recopilando los datos en tiempo real gracias a las nuevas tecnologías. De esta manera, se puede aplicar para modificar de manera inmediata la estrategia a seguir durante la temporada.

Se han mostrado las posibilidades de la tecnología dentro del mundo deportivo y cómo es una industria que está desarrollándose actualmente, además de las posibilidades que tiene concretamente el estudio de eficiencia que se ha realizado en este trabajo, del que se pueden obtener numerosos resultados que pueden ayudar a la hora de decidir estrategias y entrenar a un equipo de la NBA.

Por último, este modelo podría aplicarse a otros deportes, adaptándolo en cada caso según conveniese, con sus variables específicas. En fútbol americano podría resultar de gran interés debido a la forma de juego, ya que es un deporte en el que la decisión de elegir atacar o defender en cada momento es fundamental y cuentan con un equipo diferente para cada caso, de forma que se podría analizar con mayor detalle los factores que influyen en la eficiencia defensiva u ofensiva.

Además, se podrían estudiar otros análisis de eficiencia usando nuevos modelos que incluyesen otras variables que en este caso no se han contemplado y comparar los resultados obtenidos en los diferentes casos. Por último, también se pueden emplear otras herramientas de Investigación Operativa, como el análisis multivariante y aplicar los resultados obtenidos.

BIBLIOGRAFÍA

Referencias en artículos

- **[Afriat, 1972]** Afriat, S. Efficiency Estimation of Production Functions. *International Economic Review*, 13, 568-598. Wiley, 1972.
- **[Aigner, 1968]** Aigner, D.J. and Chu, S.F. On Estimating the Industry Production Function. *American Economic Review*, 58, 826-839. American Economic Association, 1968.
- **[Banker, 1984]** RD Banker, A Charnes and WW Cooper. "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis." *Management Science* 30(9): 1078-1092. 1984.
- **[Charnes et al, 1978]** A. Charnes, WWCooper, EL Rhodes. Measuring the Efficiency of Decision Making Units. *EJOR* 2: 429-444. Elsevier, 1978.
- **[Cooper, 2000]** W. W. Cooper, L. M. Seiford, K. Tone. *Data Envelopment Analysis*. Kluwer Academic Publishers, 2000.
- **[Färe, 1995]** R.Färe, G. Whittaker. An intermediate input model of dairy production using complex survey data. *Journal of Agricultural Economics* 1995; 46(2):201–23.
- **[Färe, 2000]** R. Färe, S. Grosskopf. Network DEA. *Socio-Economic Planning Sciences* 2000; 34: 35 – 49.
- **[Farrell, 1957]** M. J. Farrell. The measurement of productive efficiency. *Journal of the Royal Statistical Society*, Vol. 120, pp.253-290, 1957.
- **[Hailu, 2001]** A. Hailu, T. S. Veeman. Non-parametric Productivity Analysis with Undesirable Outputs: An Application to the Canadian Pulp and Paper Industry. *American Journal of Agricultural Economics*, Volume 83, Issue 3, pp 605-616, 2001.
- **[Hernández et. al, 2009]** N.Hernández, C.F. Soto, C.A. Caballero. Modelos de simulación de cultivos. Características y usos. *Cultivos Tropicales*. cultrop v.30 n.1 Ediciones INCA, 2009.
- **[Hillier, 2010]** F.S. Hillier, G. J. Lieberman. *Introduction to operations research*. Ninth Edition. Mc Graw-Hill, 2010.
- **[Lewis, 2004]** H. F. Lewis, T. R. Sexton. Network DEA: Efficiency analysis of organizations with complex internal structure. *Computers & Operations Research*. 31. 1365-1410, 2004
- **[Lyeme et al, 2012]** H. Lyeme, M. Seleman. *Introduction to Operations Research. Theory and Applications*. LAP Lambert Academic Publishing, 2012.
- **[Steffens, 2004]** L. Steffens, *Executive Guide to Operations Research*. The institute for Operations Research and the Management Sciences (INFORMS), 2004.
- **[Taha, 2004]** H. A. Taha. *Investigación de Operaciones*. 7º edición. PEARSON EDUCACIÓN, 2004
- **[Wright, 2009]** M. Wright. 50 years of OR in sport. *Journal of the Operational Research Society*, Volume 60, Supplement 1, pp S161–S168. Springer, 2009.

- **[Goldberry, 2012]** K.Goldsberry. CourtVision: New Visual and Spatial Analytics for the NBA. Harvard University, 2012

Referencias en páginas web

- **[Bellyupsports, 2018]** Bellyupsports. The Evolution of Technology & Sports, 2018.
URL: <https://bellyupsports.com/2018/11/the-evolution-of-technology-sports/>
Último acceso: 17 de Abril de 2019.
- **[Castillo, 2015]** Y. A. Castillo. Orígenes de la Investigación de Operaciones, 2015.
URL: <https://www.monografias.com/trabajos104/origenes-investigacion-operaciones/origenes-investigacion-operaciones.shtml>
Último acceso: 15 de Febrero de 2019.
- **[García, 2018]** N. García. La tecnología hace del deporte una disciplina científica. El Economista, 2018.
URL: <https://www.economista.es/salud-innovacion/noticias/9097351/04/18/No-hay-deporte-que-no-haya-sufrido-una-revolucion-tecnologica.html>
Último acceso: 19 de Abril de 2019.
- **[Hamburg, s.f.]** J. Hamburg. How a Woman Revolutionizes the NFL, s.f
URL: <https://www.rollins.edu/magazine/spring-2014/michelle-mckenna-doyle.html>
Último acceso: 7 de Marzo de 2019.
- **[Hernández-b, 2018]** N. Hernández. El análisis de datos ya decide el resultado de competiciones deportivas, 20018
URL: <https://www.nobbot.com/futuro/analisis-datos-resultado-competiciones-deportivas/>
Último acceso: 6 de Marzo de 2019.
- **[Lorenzo, et. al, 2003]** A. Lorenzo Calvo, M.A. Gómez Ruano, A. Jaime Sampaio. Análisis descriptivo de las posesiones de 24 segundos en baloncesto. Revista digital. Año 9 - N° 67, 2003.
URL: <https://www.efdeportes.com/efd67/balonc.htm>
Último acceso: 27 de Mayo de 2019.
- **[Marentes s.f.]** L. Marentes. La industria deportiva. Periodismo y marketing deportivo, s.f.
URL: <https://luismarentes.wordpress.com/la-industria-deportiva/>
Último acceso: 5 de Marzo de 2019.
- **[Michelle McKenna-Doyle, 2017]** Michelle McKenna-Doyle. How the NFL Can Take ‘Next Gen Stats’ to the Next Level, 2017
URL: <https://www.youtube.com/watch?v=gjDLN3qJudA>
Último acceso: 18 de Febrero de 2019.

- **[Monje, 2017]** A. Monje. Cómo el modelo de los Rockets está transformando la NBA moderna, 2017
URL: <https://kiaenzona.com/nba/james-harden-imperio-la-analitica-andres-monje-66714/>
Último acceso: 18 de febrero de 2019
- **[Oppermann, s. f.]** P.Oppermann. How Technology is Changing Contact Sports, s. f.
URL: <https://channels.theinnovationenterprise.com/articles/how-technology-is-changing-contact-sports>
Último acceso: 26 de Mayo de 2019.
- **[Pink, 2017]** D. Pink, A. Silver. How Data Science is Transforming the Game of Basketball, 2017.
URL: <https://heleo.com/conversation-how-data-science-is-transforming-the-game-of-basketball/14179/>
Último acceso: 25 de Febrero de 2019.
- **[Silverman, 2009]** J. Silverman. How Sabermetrics Works, 2009.
URL: <https://entertainment.howstuffworks.com/sabermetrics.htm>
Último acceso: 2 de Marzo de 2019.
- **[The Economist, 2018]** The Economist. How data transformed the NBA , 2018
URL: <https://www.youtube.com/watch?v=oUvvhkXyOA>
Último acceso: 19 de Febrero de 2019.

GLOSARIO

ATL: Atlanta Hawks
BKN: Brooklyn Nets
BOS: Boston Celtics
BS: Blocked Shots
CFL: Continental Football League
CHA: Charlotte Hornets
CHI: Chicago Bulls
CIO: Chief Information Officer
CLE: Cleveland Cavaliers
CRS: Constant Return to Scale
DAL: Dallas Mavericks
DEA: Data Envelopment Analysis
DEN: Denver Nuggets
DET: Detroit Pistons
DMU: Decision Making Unit
DR: Defensive Rebound
FGA: Free Goals Attempted
FTA: Free Throws Attempted
GSW: Golden State Warriors
HOU: Houston Rockets
IFORS: International Federation of Operational Research Societies
IND: Indiana Pacers
IO: Investigación Operativa
JORS: Journal of the Operational Research Society
LAC: LA Clippers
LAL: Los Angeles Lakers
LCT: Lanzamientos intentados
LTL: Lanzamientos tiros libres
MEM: Memphis Grizzlies
MIA: Miami Heat
MIL: Milwaukee Bucks
MIN: Minnesota Timberwolves
MIT: Massachusetts Institute of Technology
MLB: Major League Baseball

NCAA: National Collegiate Athletic Association
NBA: National Basketball Association
NFL: National Football League
NOP: New Orleans Pelicans
NYK: New York Knicks
OKC: Oklahoma City Thunder
OR: Operation Research
OR: Offensive Rebounds
ORL: Orlando Magic
PB: Posesión de balón
PDB: Pérdidas de balón
PF: Personal Foul
PHI: Philadelphia Sixers
PHX: Phoenix Suns
POR: Portland Trail Blazers
PTS: Points
SAC: Sacramento Kings
SAS: San Antonio Spurs
ST: Steals
UTA: Utah Jazz
TO: Turnovers
TOR: Toronto Raptors
VRS: Variable Return to Scale
WAS: Washington Wizards
WNBA: Women's National Basketball Association
3PA: Three Points Attempted

ANEXO

➤ Código del modelo que se ha implementado en LINGO:

SETS:

!conjuntos primitivos;

```
ENTRADADEF/1,2,3/:XRDEFSOL;
SALIDADEF/1,2,3,4,5/:YRDEFSOL;
ENTRADAATAQ/1,2,3,4,5/:XRATAQSOL;
SALIDAATAQ/1,2,3,4/:YRATAQSOL;
```

```
DMU/1..2624/:POINTSDEF,POINTSATAQ,POSDEFSOL,POSATAQSOL,POSDEF,POSATAQ;
```

!conjuntos derivados;

```
ITERACION(DMU):LANDADEF,LANDAATAQ,SDEF,SATAQ;
DMU_ENTRADADEF(DMU,ENTRADADEF):XDEF;
DMU_ENTRADAATAQ(DMU,ENTRADAATAQ):XATAQ;
DMU_SALIDADEF(DMU,SALIDADEF):YDEF;
DMU_SALIDAATAQ(DMU,SALIDAATAQ):YATAQ;
```

ENDSETS

DATA:

```
XDEF,YDEF,XATAQ,YATAQ,POINTSDEF,POINTSATAQ,POSDEF,POSATAQ = @OLE
('C:\PROYECTO\DATOS_NBA.xlsx','XDEF','YDEF','XATAQ','YATAQ','POINTSDEF','POIN
TSATAQ','POSDEF','POSATAQ');
```

ENDDATA

SUBMODEL NBA:

!Funcion objetivo;

```
MAX= SRDEF+SRATAQ;
```

!Restricciones;

```
@FOR (ENTRADADEF(I):
    @SUM(DMU(J):
        LANDADEF(J)*XDEF(J,I)) = XRDEFSOL(I));

@FOR (ENTRADAATAQ(I):
    @SUM(DMU(J):
        LANDAATAQ(J)*XATAQ(J,I)) = XRATAQSOL(I));

@FOR (ENTRADADEF(I):
    XRDEFSOL(I) <= XDEF(R,I));
@FOR (ENTRADAATAQ(I):
    XRATAQSOL(I) <= XATAQ(R,I));

@FOR (SALIDADEF(K):
    @SUM(DMU(J):
        LANDADEF(J)*YDEF(J,K)) = YRDEFSOL(K));
```

```

@FOR (SALIDAATAQ(K):
    @SUM(DMU(J):
        LANDAATAQ(J)*YATAQ(J,K)) = YRATAQSOL(K));

@FOR (SALIDADEF(K): YRDEFSOL(K) >= YDEF(R,K));
@FOR (SALIDAATAQ(K): YRATAQSOL(K) >= YATAQ(R,K));

@SUM(DMU(J): LANDADEF(J)*POSDEF(J)) = POSRDEFSOL;

@SUM(DMU(J): LANDAATAQ(J)*POSATAQ(J)) = POSRATAQSOL;

POSRDEFSOL+POSRATAQSOL= 100;

@SUM (DMU(J): LANDADEF(J)*POINTSDEF(J))= POINTSDEF(R)-SRDEF;

@SUM (DMU(J): LANDAATAQ(J)*POINTSATAQ(J))= POINTSATAQ(R)+SRATAQ;

@SUM(DMU(J):LANDADEF(J))=1;

@SUM(DMU(J):LANDAATAQ(J))=1;

@GIN (SRDEF);
@GIN (SRATAQ);

ENDSUBMODEL

CALC:
@SET( 'TERSEO', 2);
!@SET( 'STAWIN', 0);

@FOR (ITERACION(ITER):

R=ITER;

@SOLVE(NBA);

SDEF(ITER)=SRDEF;
SATAQ(ITER)=SRATAQ;
POSDEFSOL(ITER)=POSRDEFSOL;
POSATAQSOL(ITER)=POSRATAQSOL;

!@IFC(R#EQ#31:

@BREAK;

!);

);
ENDCALC

```