



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

**MÉTODOS CUANTITATIVOS PARA LA GENERACIÓN DE RANKINGS DE
SELECCIONES DE FÚTBOL Y SU APLICACIÓN A LA CONFECCIÓN DE
GRUPOS BALANCEADOS EN LA COPA DEL MUNDO**

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN GESTIÓN DE OPERACIONES
MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

SEBASTIÁN ALEJANDRO CEA BONTÁ

PROFESOR GUÍA:

DENIS SAURÉ VALENZUELA

MIEMBROS DE LA COMISIÓN:

GUILLERMO DURÁN

MARCELO OLIVARES ACUÑA

MARIO GUAJARDO ANDRADES

SANTIAGO DE CHILE

2016

"Métodos cuantitativos para la generación de rankings de selecciones de fútbol y su aplicación a la confección de grupos balanceados en la copa del mundo"

El presente trabajo se centra en la confección de distintas metodologías de creación de ratings, mediante técnicas estadísticas para el ordenamiento de las selecciones de fútbol internacionales con el fin de mejorar el ya existente y conocido mundialmente Ranking FIFA.

Los métodos utilizados en este trabajo integran los datos de partidos de fútbol entre selecciones adultas desde el año 2005 al 2013 y que permitieron la asignación de los cabezas de serie para el Mundial de fútbol acontecido en Brasil durante los meses de Junio y Julio del 2014. Para el estudio de las variables más importantes que determinan el resultado de un encuentro se utilizaron 4 modelos predictivos. El principal enfoque utilizado fue el uso de modelos Logit Multinomiales.

El modelo que permita explicar de mejor manera la data utilizando el mínimo número de variables será la primera metodología que se utilizará como propuesta al ranking FIFA actual. El modelo de predicción que mejor se ajustó a la data fue el basado en 2 atributos: Diferencia de ranking entre los equipos, y el factor localía. De esta manera variables como pertenecer a cierta confederación o la localía continental no fueron incluidas en las propuestas finales. Además de lo recién señalado se realizan modificaciones estructurales al modelo FIFA existente, que permiten corregir ciertos defectos.

Luego de proponer metodologías diferentes a la actual, se evalúan los resultados de cada una de ellas bajo métricas de comparación. Primero que todo se debe notar que una metodología será evaluada de acuerdo al ranking que de ella se desprenda. Estos rankings serán comparados mediante la métrica "mean squared error" (MSE) a partir de Rankings de Referencias.

En la segunda parte de esta tesis se propone un modelo lineal que sea capaz de generar propuestas para mundiales con fase de grupos mucho más equilibradas que los que se han visto en las últimas ediciones mundialistas, en las cuales la disparidad en el nivel de los equipos intergrupos ha sido tema recurrente.

AGRADECIMIENTOS

Comienzo agradeciendo de manera general a todas las personas que han sido importantes en mi vida.

Agradezco a compañeros y amigos de básica, a profesores que con tanto cariño comenzaron con mi etapa de formación. Agradezco a los amigos que dejé de niño, y a los que crecieron conmigo, en especial a estos últimos porque ya son más de 20 años de amistad. Gracias, en particular a mi primo y mejor amigo de niño, Jorge. Gracias por las conversas, por las peleas, por las risas, por los partidos y por esa gran amistad.

Gracias a los amigos de universidad, a los que comenzaron conmigo en primer año y a los que conocí en el proceso. Gracias por todas las juntas, las palabras de apoyo y por aliviar los días más difíciles. Gracias a los amigos de la Feria, proyecto único, que me hizo conocer a gente que sé me acompañarán por mucho tiempo de mi vida. En particular gracias al "Patio" por llenar este último año de muy buenos momentos.

Gracias a los amigos de los distintos equipos de fútbol en los que he participado, gracias por compartir conmigo la misma pasión por la redondita.

A los que me acompañan desde el colegio, y que siguieron también este camino "Industrial". Lucho y Pablo, gracias por los buenos momentos, por las entregas a última hora, por los envíos a las 23:59. Gracias porque muchas veces aunque dejáramos las cosas para el final, sabíamos que se podía. Gracias por la amistad incondicional.

Gracias a mis profesores de la comisión. Gracias por guiarme en este último proceso universitario, por sus recomendaciones y por aportar en gran medida a este trabajo.

Gracias a mis familiares, tíos, primos, especialmente a los que se dieron el tiempo de escucharme y aconsejarme cuando lo necesité.

Agradezco a mis hermanos por cuidarme desde nuestro primer colegio, por comprarme colación en los recreos, por enseñarme cosas que los viejos no tenían tiempo de enseñar. Gracias por no defenderme cuando no me lo merecía, ya que de esos momentos también aprendí. Les agradezco por hacerme el camino más sencillo.

Agradezco a mis padres, porque este éxito es de ellos y por ellos. Gracias por entender los malos humores en ciertos periodos, los días sin llegar a casa en época de exámenes y por estar siempre preocupados por mí. Gracias por el sacrificio hecho con mis hermanos y conmigo y por enseñarme el gran valor de la familia.

Finalmente agradezco a las personas que por diferentes circunstancias ya no están, a las personas que pasaron dejando aprendizajes y a los que ya partieron. Un beso grande al cielo.

Tabla de Contenido

1.	INTRODUCCIÓN.....	1
2.	SITUACIÓN ACTUAL.....	5
2.1.	LA FIFA.....	5
2.2.	RANKING FIFA.....	6
2.2.1.	PARTICULARIDADES.....	8
3.	DEFINICIÓN DE LOS PROBLEMAS A RESOLVER.....	10
4.	OBJETIVOS.....	12
4.1.	Objetivo General.....	12
4.2.	Objetivos Específicos.....	12
5.	METODOLOGÍA UTILIZADA.....	13
6.	TRABAJOS Y LITERATURA ASOCIADA.....	15
6.1.	MÍNIMOS CUADRADOS.....	15
6.2.	MÁXIMA VEROSIMILITUD.....	16
6.3.	SISTEMA ELO.....	17
7.	DATOS.....	20
8.	MODELOS PROPUESTOS.....	22
8.1.	MODELO N°1: MODELO BASADO EN RANKING.....	23
8.2.	MODELO N°2: MODELO BASADO EN RANKING Y LOCALÍA.....	25
8.3.	MODELO N°3 MODELO BASADO EN RANKING, LOCALÍA NACIONAL Y CONTINENTAL.....	27

8.4.	MODELO N°4: MODELO BASADO EN RANKING, LOCALÍA Y FACTOR CONFEDERACIÓN	29
9.	BOOKMAKERS Y EL FACTOR LOCALÍA.	31
10.	SIMULACIÓN DE METODOLOGÍAS Y PROPUESTA SELECCIONADA	39
10.1.	METODOLOGÍA 1: CON LOCALÍA INCLUIDA Y VARIABLE CONFEDERACIÓN ELIMINADA.....	39
10.2.	METODOLOGÍA 2: CON PARTIDOS AMISTOSOS OMITIDOS.....	40
10.3.	METODOLOGÍA 3: REESTRUCTURACIÓN DE PUNTOS POR VICTORIA	40
10.4.	METODOLOGÍA 4: SIN DEPRECIACIÓN DE PUNTAJE	41
10.5.	MÉTODOS DE COMPARACIÓN DE RANKINGS Y RESULTADOS.....	42
11.	ASIGNACIÓN JUSTA Y MUNDIAL MÁS COMPETITIVO	54
11.1.	MODELO DE BALANCE DE FUERZA ENTRE GRUPOS.....	58
11.1.1.	Conjuntos.....	58
11.1.2.	Datos	58
11.1.3.	Variables de Decisión	59
11.1.4.	Restricciones	59
11.1.5.	Función Objetivo	60
11.2.	RESULTADOS DE OPTIMIZACIÓN LINEAL MIXTA	60
12.	CONCLUSIONES Y RECOMENDACIONES	64
13.	Bibliografía.....	69
14.	ANEXOS	71
14.1.	ANEXO 1: Países por Confederación	71

14.2. ANEXO 2: Resultados de Ranking según Metodología	73
14.3. ANEXO 3: Resultados Parámetros estimados.....	81
14.4. ANEXO 4: Resultados Modelo de optimización con F.O 1.....	85
14.5. ANEXO 5: Resultados Modelo de optimización con F.O 2.....	87

Índice de tablas

Tabla 1: Ejemplo puntajes FIFA.....	7
Tabla 2: Puntaje Chile Agosto 2014.....	7
Tabla 3: Ranking FIFA Octubre 2013.....	9
Tabla 4: Partidos de Fútbol de Selecciones nacionales adultas 2005- Oct 2013.....	20
Tabla 5: Distribución de resultados disputados con Localía	21
Tabla 6: Resultados, Modelo N°1	24
Tabla 7: Resultados, Modelo N°2	25
Tabla 8: Copas mundiales disputadas en Sudamérica y Europa	27
Tabla 9: Resultados, Modelo N°3	28
Tabla 10: Resultados, Modelo N°4.....	29
Tabla 11: Resumen de resultados.....	30
Tabla 12: Ejemplo de probabilidades implícitas a partir de cuotas	32
Tabla 13: Frecuencia de Resultados para partidos entre 2009-2010, jugados con localía.....	36
Tabla 14: Posición Final Copa del Mundo Brasil 2014.....	43
Tabla 15. Resultados Metodologías propuestas, Ranking de Referencia basado en resultados Mundial 2014.	43
Tabla 16. Parámetros en base a partidos 2005-2013 para los países participantes del Mundial Brasil 2014.....	45
Tabla 17. Resultados simulaciones de torneos con sistema Round Robin, en base a datos históricos	47
Tabla 18. Resultados Metodologías propuestas, Ranking de Referencia basado en simulaciones en base a data histórica.	48
Tabla 19. Ejemplos partidos Mundial 2014, Cuotas-Probabilidades y Lambdas	48

Tabla 20. Parámetros en base a cuotas de casas de apuestas partidos Copa Mundial Brasil 2014.....	49
Tabla 21. Resultados simulaciones de torneos con sistema Round Robin, en base a casas de apuestas.....	51
Tabla 22. Resultados Metodologías propuestas, Ranking de Referencia basado en casas de apuestas	52
Tabla 23. Resumen de resultados. Comparación de Metodologías.....	52
Tabla 24: Pots oficiales mundial Brasil 2014.....	54
Tabla 25: Draw final, Mundial Brasil 2014.....	57
Tabla 26: Suma de Ranking relativo y Ranking FIFA por grupo.....	57
Tabla 27: Resultado Óptimo modelo de optimización con F.O. 1, Opción 1 ...	61
Tabla 28: Resultado modelo de optimización con F.O. 2, Opción 1	62
Tabla 29: Países por Confederación.....	71
Tabla 30: Resultados Metodología de Rating 1.....	73
Tabla 31: Resultados Metodología de Rating 2.....	75
Tabla 32: Resultados Metodología de Rating 3.....	77
Tabla 33: Resultados Metodología de Rating 4.....	79
Tabla 34. Parámetros estimados, datos años 2009-2013.	81
Tabla 35: Resultado modelo de optimización con F.O. 1, Opción 2	85
Tabla 36: Resultado modelo de optimización con F.O. 1, Opción 3	86
Tabla 37: Resultado modelo de optimización con F.O. 1, Opción 4.....	86
Tabla 38: Resultado modelo de optimización con F.O. 2, Opción 2	87
Tabla 39: Resultado modelo de optimización con F.O. 2, Opción 3.....	87
Tabla 40: Resultado modelo de optimización con F.O. 2, Opción 4	87
Tabla 41: Resultado modelo de optimización con F.O. 2, Opción 5	88
Tabla 42: Resultado modelo de optimización con F.O. 2, Opción 6	88

Tabla 43: Resultado modelo de optimización con F.O. 2, Opción 7	88
Tabla 44: Resultado modelo de optimización con F.O. 2, Opción 8	89
Tabla 45: Resultado modelo de optimización con F.O. 2, Opción 9	89
Tabla 46: Resultado modelo de optimización con F.O. 2, Opción 10	89

Índice de ilustraciones

Ilustración 1: Distribución de diferencias en probabilidad de victoria entre equipo 1 y 2 según su diferencia en ranking, partidos jugados en territorio neutral, 2009-2010.....	33
Ilustración 2: Distribución de diferencias en probabilidad de victoria entre equipo 1 y 2 según su diferencia en ranking, partidos jugados con el equipo 1 actuando de local 2009-2010.....	33
Ilustración 3: Distribución de diferencias en probabilidad de victoria entre equipo 1 y 2 según su diferencia en ranking entre -100 y 100. Partidos jugados con el equipo 1 actuando de local 2009-2010.....	34
Ilustración 4: Distribución de empate según diferencia en probabilidad de victoria entre equipo 1 y 2, partidos jugados en territorio neutral en los años 2009-2010.....	37
Ilustración 5: Distribución de empate según diferencia en probabilidad de victoria entre equipo 1 y 2, partidos jugados con equipo 1 de local en los años 2009-2010.....	38
Ilustración 6: : Comparación draw oficial vs draw optimizado (1).....	61
Ilustración 7: Comparación draw final vs draw optimizado (2)	63

1. INTRODUCCIÓN

Los resultados de los eventos deportivos son una materia largamente estudiada desde hace décadas. Una característica principal de estos eventos es la incertidumbre intrínseca de sus resultados, lo cual lo hace una materia muy interesante de estudiar. La gran dificultad para predecir los resultados deportivos no impide que existan modelos que permitan obtener una distribución de probabilidad sobre ellos.

Es por esto que expertos y /o apostadores en temas deportivos han intentado acercarse lo mejor posible a estas distribuciones de probabilidad deportivas mediante el uso y análisis de datos históricos. Si bien la gran parte de los modelos predictivos se basan en el análisis de registros previos, la idea principal es poder comparar a dos equipos y/o jugadores que se enfrentan mediante una medida de superioridad o supremacía de uno sobre otro. El resultado de este enfoque es llamado sistema de rating.

Los ratings de equipos, países y/o jugadores es un tema cada vez más relevante en el área deportiva. A medida que han pasado los años los ratings han permitido determinar quienes participan en torneos de mayor envergadura o el cuadro final de una competición.

Ratings deportivos tienen la función de representar de manera numérica el poder o fuerza de un equipo en determinado deporte, en cierto periodo de tiempo. Esto permite comparar de manera automática a dos equipos enfrentándose en un partido y poder establecer predicciones sobre el resultado del mismo. Dada esta función y responsabilidad que recae sobre los distintos ratings se ha generado un debate sobre si estos, en base a la metodología que lo estructuran, son capaces de cumplirla.

Entre los principales ratings deportivos se encuentran el rating ATP (tenis), NCAA College Basketball Ratings y Rating FIFA (Selecciones de fútbol). Este último será el principal objeto de estudio dada la importancia que ha tomado en los últimos años al ser el responsable de la designación de los mejores sembrados en las copas mundiales. Al igual como el ranking de un sitio web en Google puede determinar el fracaso o éxito de cualquier negocio, el éxito de un país participante en la mayor cita planetaria del balompié está fuertemente relacionado con su ranking FIFA.

La Copa Mundial de la FIFA, también conocida como Copa Mundial de Fútbol es el principal torneo internacional de fútbol masculino a nivel de selecciones nacionales en el mundo. Este torneo se disputa desde el año 1930 teniendo lugar en Uruguay.

La realización de una Copa del Mundo envuelve mucho más que sólo los partidos a disputarse, ya que conlleva un gran movimiento económico tanto para las selecciones que la disputarán, como también, en mayor medida, para el anfitrión. En la última cita planetaria Brasil gastó una cifra récord en remodelación de estadios, infraestructura de aeropuertos, autopistas y hoteles invirtiendo un total aproximado de US\$11.000 millones. Pero la organización de un mundial no trae consigo sólo gastos ya que de acuerdo a cifras oficiales el turismo generó un crecimiento de más del 80% en el país y el desempleo disminuyó considerablemente gracias a la creación de 3,5 millones de trabajos temporales. Por su parte los ingresos de la FIFA también fueron los más altos históricamente alcanzando los US\$4.000 millones por concepto de transmisión de partidos y anuncios publicitarios.

En cuanto a los montos recibidos por los equipos, por solo asistir al mundial recibieron US\$ 8 millones; por pasar a octavos de final, US\$ 9 millones y los cuartos finalistas, US\$14 millones. De igual forma, las selecciones semifinalistas se embolsaron US\$ 20 millones y el tercer puesto US\$ 22 millones. El equipo finalista US\$ 25 millones y el campeón, Alemania, se llevó el premio mayor que ascendió a US\$ 35 millones [1].

En lo que respecta al sistema de torneo utilizado por la FIFA para los torneos mundiales, éste ha cambiado a través de los años; si bien estructuralmente todas las ediciones han constado de una fase clasificatoria y una ronda final, la cantidad de selecciones que llegan a esta última y su modalidad ha tenido ciertas variantes. Cabe notar que toda selección participante debe ser miembro tanto de FIFA como de algunas de las 6 confederaciones continentales existentes.

- Fase Clasificatoria: Cada una de las 6 confederaciones organizan un sistema de elección de sus representantes a través de encuentros deportivos. El número de representantes de cada confederación es definido previamente por la FIFA a través de la entrega de cupos, algunos de los cuales son completos, equivalentes a un equipo en la fase final y otros son compartidos, en los que un equipo debe definir su clasificación a la ronda final ante un representante de otra confederación en un proceso denominado generalmente repechaje.
- Fase Final: En sus inicios la fase final consistía en la eliminación directa de todos los participantes que habían salido victoriosos en la fase clasificatoria, hasta determinar al campeón. Hasta el mundial de 1994 disputado en Estados Unidos 24 equipos llegaban a la fase final y eran distribuidos en 6 grupos de 4 equipos cada uno, de los cuales

los 2 mejores de cada grupo avanzaban a la siguiente etapa, acompañados de los 4 mejores terceros. Ya en esta etapa hubo campeonatos que utilizaban el sistema de eliminación directa hasta dirimir al campeón, y otros que utilizaban una nueva fase de grupos para determinar a los finalistas. Para aumentar la probabilidad de ver partidos más atractivos pasada la etapa de grupos se determinaban 5 cabezas de serie.

La Copa Mundial de 1998 disputada en Francia fue la primera a la que asistieron 32 selecciones. La primera fase constó de ocho grupos con cuatro integrantes cada uno, liderados por un cabeza de serie. Las mejores dos escuadras avanzaban a octavos de final y al igual que en los mundiales anteriores, a partir de esta etapa los partidos se disputaban en modalidad de eliminación directa. Las principales razones por las cuales FIFA ha aumentado el cupo de participantes ha sido la masificación de este mundial hacia más países y la mayor integración para los países africanos y asiáticos. Es de esperar que para las próximas versiones los cupos para la Copa del Mundo se sigan ampliando y se pueda llegar a jugar con 40 o 48 naciones.

Otro hecho relevante en cada mundial es la designación de los cabezas de serie respectivos, tema que ha sido muy discutido en todas las ediciones. Los mejores sembrados hasta el mundial de Sudáfrica 2010 eran siempre el equipo anfitrión y los demás eran escogidos tanto por el Ranking FIFA como por los resultados de los países en ediciones anteriores, mientras que para el mundial de Brasil 2014 se determinó que los 7 cabezas de serie fueran designados exclusivamente por el Ranking FIFA. Cabe notar que el hecho de ser cabeza de serie tiene una directa implicancia en las oportunidades de éxito en el torneo, ya que es muy probable que éstos puedan avanzar por lo menos la fase de grupos, ya que se evitan tener un partido con los rivales más fuertes a inicios de campeonato.

Como fue señalado anteriormente, dado que el éxito de un país en el mundial está fuertemente influenciado por ser o no cabeza de serie, y esta decisión proviene directamente del Ranking FIFA, es que los ojos del mundo han estado puestos en cómo se realiza este último. Este tema es el eje principal a tratar en las próximas secciones.

Abordar el tema de generación de ratings y en particular analizar la confección del Ranking FIFA actual puede constituir un gran aporte en el área deportiva. Ser capaz de detectar errores y/o entregar opciones de mejoras puede beneficiar en primera medida la competitividad de los torneos a nivel

mundial y asimismo entregar reconocimiento justo a los países mejores sembrados.

Dentro de este contexto los temas principales de esta tesis son por un lado generar metodologías estadísticas para la creación de un Ranking robusto y que sean implementables por la FIFA, y por otra parte, generar un modelo que permita presenciar Copas Mundiales de fútbol mucho más equitativas, entendiendo este calificativo como un mayor balance en las fuerzas de los equipos que conforman los grupos.

2. SITUACIÓN ACTUAL

2.1. LA FIFA

La Federación Internacional de Fútbol Asociación, mundialmente conocida como FIFA, es el ente controlador del fútbol a nivel mundial. Su fundación data de Mayo de 1904 en la ciudad de Zurich, Suiza. FIFA es la encargada de la realización de la Copa Mundial de Fútbol, los otros campeonatos del mundo en sus diferentes categorías y los torneos olímpicos entre otros. Este organismo está compuesto por 209 asociaciones nacionales (países) y está organizada de la siguiente manera:

- el Congreso (órgano legislativo),
- el Comité Ejecutivo (órgano ejecutivo),
- la Secretaría General (órgano administrativo), y
- las Comisiones (que asisten al Comité Ejecutivo).

Las actividades de la FIFA persiguen el objetivo final de mejorar la experiencia de todos aquellos que deseen jugar al fútbol. Por ende, para fomentar el desarrollo de sus actividades, debe formar y capacitar a las personas responsables de proporcionar dicha experiencia. Sumado a lo anterior, es posible destacar que la misión de este ente se define como *"desarrollar el fútbol en todas partes y para todos, emocionar al mundo a través de sus inspiradoras competiciones y edificar un futuro mejor mediante el poder del fútbol"* [2].

Como se mencionó anteriormente, la FIFA es la encargada de la realización de las Copas Mundiales de Fútbol, torneo internacional de fútbol masculino a nivel de selecciones más importante del mundo. Este torneo se desarrolla cada 4 años en una sede escogida por el Comité ejecutivo de la FIFA. Hasta la fecha se cuenta con 20 campeonatos del mundo, y tan sólo 8 países han logrado ser los vencedores; Brasil en 5 ocasiones, Alemania e Italia con 4 triunfos, Argentina y Uruguay cada uno con 2 trofeos, mientras que Inglaterra, Francia y España completan el selecto grupo con tan sólo un título.

2.2. RANKING FIFA

La clasificación Mundial de la FIFA, también conocida como ranking FIFA es un sistema de clasificación de todas las selecciones asociadas a la FIFA, y ha sido utilizada desde el año 1993.

Desde sus comienzos esta clasificación se ha basado en los resultados de los partidos oficiales disputados por las 209 selecciones. En el año 1999 se introducen cambios ya que la FIFA detectó factores que inciden en los resultados de un partido. De esta manera se creó una metodología que permitía incluir estos factores al momento de clasificar a los equipos. Los principales cambios fueron; diferenciación de acuerdo a condición de local o visita, cantidad de goles a favor y en contra, importancia del partido y fuerza de la confederación.

El año 2006 se introdujeron nuevos cambios que simplificaban la metodología de cálculo pero que eliminaban variables fundamentales en el desarrollo de un partido. Entre los cambios realizados se encuentran; eliminación del factor localía, cambio en la cantidad de años que se toman en consideración y la cantidad de goles realizados en un partido. De esta manera el método actual es el siguiente.

Los puntos de un partido se calculan mediante la fórmula [3]:

$$Puntos = P * I * F * C \quad \text{donde}$$

P= Puntos por Victoria: 3 si es un triunfo, 1 en caso de empate, 0 en caso de derrota

I= Importancia del partido: 4 si es Copa del Mundo, 3 si es fase final de torneo de Confederación (Copa América, Copa de oro, Eurocopa, etc.) o Copa Confederaciones, 2.5 si es eliminatoria mundialista o en el ámbito de confederación, y 1 si es un partido amistoso o torneo menor de confederación

F= Fuerza del equipo contrincante: 200-ranking del oponente. En caso que el oponente tenga un ranking inferior al 150°, la fuerza del contrincante se fija en un mínimo de 50. Al líder de la clasificación se le asigna un valor máximo de 200

C= Fuerza de la Confederación: La fuerza de una confederación se calcula de acuerdo con el número de victorias que ha obtenido en las últimas

tres ediciones de la Copa Mundial de la FIFA. Los valores antes del mundial Brasil 2014 eran los siguientes:

$$CONMEBOL/UEFA = 1 \quad CONCACAF = 0,88 \quad AFC/CAF = 0,86 \quad OFC = 0,85$$

El valor que tomara el parámetro C será la media entre los valores de los equipos participantes en un encuentro. Para que quede aún más claro se muestra el siguiente ejemplo de los puntos obtenidos por dos equipos en un partido del mundial Sudáfrica 2010.

Tabla 1: Ejemplo puntajes FIFA

	HONDURAS	VS	CHILE
Fecha: 16-06-2010			
Tipo de partido Mundial			
Posición en ranking	38		18
P: Resultado del partido	0		3
I: Importancia del partido	4		4
F: Fuerza del competidor	182		162
C: Fuerza de la confederación	0,94	(0,88+1)/2	0,94
Puntos=P I F C	0		1227

Ya explicado cómo se calculan los puntos obtenidos en un partido, el rating de un equipo en particular se calcula a partir de todos los puntos sumados en los últimos 4 años en base a sus promedios por partido depreciados por año como se muestra a continuación.

$$PUNTOS_{TOTAL} = \overline{PUNTOS}_t + 0,5 * \overline{PUNTOS}_{t-1} + 0,3 * \overline{PUNTOS}_{t-2} + 0,2 * \overline{PUNTOS}_{t-3}$$

A continuación se ejemplifica el caso de Chile para el mes de Agosto de 2014.

Tabla 2: Puntaje Chile Agosto 2014

Año	Promedio	Peso	Resultante
2014	673,61	100%	673,61
2013	391,47	50%	195,73
2012	469,7	30%	140,91
2011	447,07	20%	89,41
		Total	1099,66

A partir de este rating, se obtiene el ordenamiento del mayor a menor y mes a mes de los 209 países asociados a la FIFA entregando esto como resultado el tan conocido ranking FIFA. Este ranking es utilizado principalmente para obtener los mejores sembrados en el campeonato de fútbol de selecciones más importante del mundo, la Copa Mundial.

2.2.1. PARTICULARIDADES

Como se mostró anteriormente la metodología utilizada para la obtención del puntaje FIFA o rating FIFA es bastante sencilla. Al igual como muchos otros rating (ATP, tenis. FAPL, Premier League entre otros) este rating permite sembrar a los mejores en un determinado torneo, en este caso la Copa Mundial. Este rating permite generar un ordenamiento de los países otorgando el #1 al equipo que posea el mayor rating. Los 7 mejores sembrados más el país anfitrión serán los que encabezarán cada uno de los grupos en la competencia más importante a nivel de selecciones. Este hecho les otorga de paso una gran ventaja, dado que las chances de pasar la primera fase o fase de grupos se ven incrementadas dado que evitan partidos difíciles en la ronda inicial. Esta gran importancia que ha ido adquiriendo ha generado que los ojos de todo el mundo estén muy pendientes de su metodología y de los problemas que puede presentar. A continuación se presentan algunas particularidades que generan suspicacias en cuanto a la validez de la metodología utilizada.

- **Partidos Amistosos**

Dada la metodología actual de promediar todos los partidos sumado a la escasa cantidad de puntos que entrega un partido amistoso, para los países con mejor rating no es conveniente disputar este tipo de partidos. Se mostrará esto con un ejemplo. En Agosto del 2014 Chile tenía un promedio en el último año de 673,61 puntos (Ver Tabla 2). Jugar un partido amistoso contra el mejor equipo del mundo y vencerlo otorga el máximo puntaje posible para estos encuentros, equivalente a 600 puntos ($3 * 1 * 200 * 1 = 600$). Por lo tanto derrotar al mejor equipo en ese momento disminuye su rating y a su vez probablemente lo hace retroceder en el escalafón mundial. Un hecho muy importante en el pasado mundial fue que 3 de los 7 países (Colombia, Bélgica y Suiza) que fueron cabezas de serie lo lograron en mayor medida por jugar menos partidos amistosos que el resto de los equipos que estaban disputando esas plazas.

Tabla 3: Ranking FIFA Octubre 2013

Posición	País	Puntos
1	España	1513
2	Alemania	1311
3	Argentina	1266
4	Colombia	1178
5	Bélgica	1175
6	Uruguay	1164
7	Suiza	1138
8	Países Bajos	1136
9	Italia	1136
10	Inglaterra	1080

En Octubre del año 2013, momento en el cual se decidieron los cabezas de serie del mundial del año entrante Colombia y Suiza habían jugado sólo 3 partidos amistosos en el último año y Bélgica lo había hecho en 4 ocasiones, mientras que Holanda e Inglaterra habían disputado 5 partidos de esta categoría e Italia lo había hecho en 6 oportunidades. Este pequeño detalle fue clave en el ordenamiento de los países en el mes de octubre de 2013, y por ende, en la designación final de los cabezas de serie del mundial de Brasil 2014.

- **Calendarización de Confederaciones**

No todas las confederaciones juegan los campeonatos de su confederación (Copa América, Copa de Oro, Eurocopa, Copa de Asia, Copa Africana, Copa de las Naciones de la OFC) en las mismas fechas. Dado que este tipo de torneos entrega mayor puntaje y dada la depreciación del mismo por año, es probable que los países pertenecientes a confederaciones que juegan su torneo más cerca del próximo mundial se vean beneficiados. Es decir a igualdad de puntajes en los demás torneos, es probable que el campeón de la Copa América jugada en el año 2011, se encuentre en desventaja frente al campeón de la Eurocopa 2012.

- **Equipos Fuertes versus equipos débiles**

Con el sistema actual, dado un tipo de partido cualquiera empatar con los mejores equipos entrega menos puntos que ganarle a equipos de la medianía de Ranking. Por ejemplo empatar con el n° 1 del ranking en un

partido amistoso entrega como máximo 200 puntos, mientras que ganarle a un equipo ubicado en la posición 100 reporta como mínimo 255 puntos (estos puntajes dependen de las confederaciones que disputen el partido)

- **Variables importantes omitidas**

El puntaje de un partido toma en consideración la dificultad del mismo, ya que a medida que se enfrenta con un rival más difícil las probabilidades de vencerlo son menores, por lo tanto, derrotarlo es más meritorio y conlleva una mayor puntuación. Sin embargo existen otras variables igual de relevantes que son omitidas por la fórmula de la puntuación. La más llamativa que no es considerada es la relacionada con la localía. Este factor ha sido objeto de estudio en los más variados deportes y ha sido ampliamente abordado en los últimos 30 años. Particularmente en el fútbol se han encontrado evidencias de que este factor es clave en los resultados del equipo local, ya que estadísticamente los resultados en todas las ligas del mundo son mejores para los dueños de casa. Dentro de los principales hallazgos encontrados se encuentran ciertos factores que inciden en estos resultados; Downward y Jones [4] relacionan esta ventaja de localía a la presión ejercida sobre los árbitros y el aliento de los aficionados, mientras R. Pollard [5] atribuye esta ventaja al mayor conocimiento del campo de juego por parte del equipo local.

Un segundo factor no tomado en cuenta y que será motivo de análisis es, para los partidos que no son disputados por un equipo jugando de local, identificar cómo influye para un equipo jugar en su mismo continente. Esta variable está relacionada principalmente a temas de aclimatación y tiempos de viaje y será analizada en esta investigación debido principalmente a los resultados obtenidos en los mundiales, en los cuales en casi la totalidad de ellos ha resultado campeón un equipo perteneciente al continente organizador (excepto cuando han sido en territorio asiático o africano).

3. DEFINICIÓN DE LOS PROBLEMAS A RESOLVER

Actualmente el Ranking confeccionado por la FIFA ha generado más de algún debate. En realidad lo que mucha gente refuta es la dificultad para calcular el puntaje real de cada país, esto debido principalmente a que el Ranking no refleja de manera correcta el nivel actual de los países. Sin embargo esta dificultad para hacer coincidir un alto rendimiento de un equipo con un respectivo alto ranking en el escalafón mundial se debe a los errores

en la metodología existente, algunos ya mencionados en el apartado anterior.

El foco de esta tesis es el estudio del sistema actual utilizado por FIFA para el Ranking de fútbol de los 209 países asociados a alguna confederación y el desarrollo de modelos que permitan agregar y/o eliminar variables significativas que influyen en un determinado resultado. A su vez y dado que este Ranking es utilizado como input para la confección del draw principal de la Copa Mundial, esta tesis intentará mostrar que a partir de programación lineal se pueden obtener mundiales mucho más equitativos, entendido esto como mayor paridad de fuerzas de equipos inter-grupos.

La idea principal es generar propuestas para generar ratings y a partir de ello nuevos rankings, considerando la inclusión de variables que puedan resultar significativas o eliminación de alguna ya existente. Algunos cambios estructurales también serán testeados con la finalidad de lograr el ranking más robusto posible.

Dado que actualmente no se tiene una metodología formal para medir u ordenar rankings en fútbol, lo primero será generar alternativas de Rankings Referenciales, los cuales serán generados a partir de:

- Resultados del Mundial
- Data histórica de los equipos
- Datos de casas de apuestas

Estos rankings serán utilizados para ordenar los resultados obtenidos mediante las distintas metodologías propuestas.

El segundo problema a abordar será el desequilibrio existente en la designación de grupos en las últimas ediciones de Copas Mundiales. Esta falencia se modelará como un Problema de Programación lineal para entregar resultados que cumplan con las restricciones ya establecidas por FIFA, pero con grupos balanceados de acuerdo a su ranking.

4. OBJETIVOS

4.1. Objetivo General

- ✓ El objetivo general de esta tesis es contribuir a la generación de rankings consistentes de acuerdo a los resultados obtenidos, principalmente en lo que respecta al sistema de clasificación de las 209 selecciones de fútbol masculino pertenecientes a la FIFA a partir de datos desde el año 2005 al 2014. También encontrar un método que permita establecer prioridad dentro de los diferentes rankings generados y observar como este input puede ser utilizado para tener eventos deportivos mucho más atractivos para los espectadores.

4.2. Objetivos Específicos

- ✓ Desarrollar un marco conceptual que permita comprender el Ranking FIFA y las variables que actualmente influyen en su creación.
- ✓ Analizar estadísticamente las variables utilizadas actualmente y estudiar la inclusión de nuevas variables relevantes.
- ✓ Complementar análisis con datos de casas de apuestas, ya que las probabilidades implícitas de éstas son mundialmente aceptadas como confiables.
- ✓ Elaborar nuevas propuestas para ordenar a las selecciones adultas de fútbol y comparar de maneras distintas estos nuevos rankings.
- ✓ Obtener conclusiones que permitan generar conocimiento y recomendaciones para futuros ciclos competitivos

5. METODOLOGÍA UTILIZADA

El paso inicial para comenzar esta tesis es definir los problemas que se enfrentan y que se intentan resolver. Estos problemas son principalmente 2, en primer lugar corregir ciertos problemas que tiene actualmente el ranking FIFA con la finalidad de generar una metodología robusta para el ranking de selecciones de fútbol adultas y el segundo, definido por el desequilibrio existente entre los grupos percibidos en la mayoría de los mundiales.

Ya definidos los problemas a resolver, se recopilaron los datos necesarios que permitieron identificar lo existente actualmente, las variables que están siendo utilizadas y las que están siendo omitidas.

En tercer lugar se procede al estudio de los trabajos y literatura existentes, relacionados con el tema, con la finalidad de revisar que es lo que se ha hecho con respecto a metodologías de ranking deportivo y analizar si es posible utilizar los desarrollos realizados previamente para entregar mayor solidez a esta tesis.

Posteriormente se generan 4 modelos Logit Multinomiales. Cada uno de ellos con variables explicativas diferentes tratando de captar la mejor manera de modelar las probabilidades del resultado de un partido. La mejor opción entre ellas, es decir, el modelo con las variables que mejor ajuste a la base de datos será la primera opción a considerar para la creación de una metodología de rating robusta para la FIFA. Para corroborar el modelo Logit Multinomial utilizado y las relaciones entre sus variables se procede a analizar el modelo estadístico subyacente de las distintas casas de apuestas.

Con los datos ya obtenidos se realiza una simulación del sistema utilizado actualmente por FIFA para el cómputo de puntos tanto para cada partido, como para la confección mensual de su ranking. Una vez realizada esta simulación, se proponen diferentes metodologías para determinar los puntos obtenidos por un equipo en un partido determinado. Estas metodologías son evaluadas por el ranking que generan y son comparadas entre sí por la métrica error cuadrático medio (MSE). Esta métrica indica qué tan alejado se encuentra el ordenamiento de los equipos de un Ranking Referencial. En esta tesis se proponen 3 distintos rankings referenciales, y se entregan los resultados de la evaluación de las metodologías.

Establecida ya la mejor metodología para la confección de ranking se aborda posteriormente el problema de desequilibrio entre los grupos en las Copas Mundiales. Para analizar el segundo tema se modela el problema de desequilibrio existente entre las fuerzas inter-grupos en la fase inicial de la Copa del Mundo. Este problema será modelado como un problema de programación lineal mixta.

Finalmente se presentan los resultados más relevantes con sus principales análisis y conclusiones y se realizarán recomendaciones que de estos se desprenden.

6. TRABAJOS Y LITERATURA ASOCIADA

Sistemas de rankings deportivos (sport ratings system) han sido objeto de estudio en la literatura estadística desde la década de los 70. Estos sistemas analizan los resultados de competiciones en deporte para proveer puntuaciones para cada equipo o jugador de forma objetiva. Por lo tanto, los rankings son una consecuencia de la ordenación de estas puntuaciones, donde al mayor puntaje se le otorga un ranking igual a 1 [6].

Sistemas de rankings deportivos son modelos matemáticos que proveen ratings a equipos, países y/o jugadores y que permiten medir la fuerza o nivel de uno en comparación a otro.

Primero que todo se debe hacer una separación entre rating y ranking. Como describe Langville y Meyer [7] "Un rating de ítems asigna un puntaje numérico a cada uno de ellos. Cuando una lista de rating se ordena crea una lista de ranking".

La literatura revisada estudia principalmente tres metodologías distintas de sistemas de rankings deportivos; mínimos cuadrados, máxima verosimilitud y ELO. El primer paso para saber si alguno de estas metodologías tiene características que puedan aportar al desarrollo de esta tesis es entender cada una de ellas.

6.1. MÍNIMOS CUADRADOS

En los estudios de Kenneth Massey [8] se plantea un modelo de mínimos cuadrados para estimar ratings que permitan predecir resultados de partidos. Además de poseer un fuerte fundamento matemático esta metodología tiene la ventaja de ser muy fácil de implementar e interpretar. Además tiene la gran ventaja de poder ser modificada agregando nuevos parámetros de una manera sencilla sin tener que cambiar su estructura fundamental.

Bajo esta metodología se asume que la diferencia en marcador entre un equipo ganador y perdedor se deben principalmente a su diferencia en rating. Más aún la metodología de mínimos cuadrados intenta que la diferencia en puntos o goles entre dos equipos o jugadores provenga de una función lineal de las fuerzas de ambos competidores. De esta manera si Y es

una variable aleatoria que representa el margen de victoria para un partido específico entre dos equipos o jugadores A y B [9].

$$E[Y] = \text{ranking}_a - \text{ranking}_b = r_a - r_b$$

Si bien muchos factores pueden influir en el margen de victoria observado, el valor esperado $E[Y]$ entrega un resultado promedio si el juego se jugara muchas veces.

Una de las grandes ventajas de este modelo es que se pueden incluir factores como por ejemplo la localía de una manera muy sencilla, por otro lado el modelo otorga demasiada importancia a la diferencia de score en un partido siendo que un encuentro se pueden presentar demasiados factores que posibiliten una mayor o menor diferencia, por lo cual siempre el que tenga mayor diferencia en goles o puntuación encabezará el ranking bajo este método, lo cual en la realidad no siempre es así.

6.2. MÁXIMA VEROSIMILITUD

Usando el método de máxima verosimilitud, David Harville [10] intenta encontrar los ratings de cada equipo que permitan maximizar la probabilidad de que ocurran los resultados que se dieron en una determinada competición.

Bajo este enfoque La probabilidad de ver que en un campeonato ocurran ciertos eventos es la probabilidad de que ocurran cada uno de esos eventos, de manera independiente. En este modelo a cada equipo i se le asigna un rating r_i que será utilizado para predecir el resultado esperado entre él y su oponente j con la probabilidad de que i derrote a j dada por:

$$P_{ij} = \frac{r_i}{r_i + r_j} \text{ de igual manera } P_{ji} = \frac{r_j}{r_i + r_j} = 1 - P_{ij}$$

Dada la no negatividad de la probabilidad se asume de inmediato que $r_i \geq 0$ para todo equipo i .

Si se analizan torneos donde los equipos se enfrentan en más de una vez, por ejemplo n_{ij} veces, la probabilidad de que i derrote a j en x_{ij} oportunidades estará dada por:

$$F_{ij}(x_{ij}, n_{ij}, P_{ij}) = \binom{n}{x} P_{ij}^x (1 - P_{ij})^{n-x}$$

De la misma forma la verosimilitud de un campeonato o una serie de partidos estará dada por:

$$F = \prod_{i \neq j} F_{ij}(x_{ij}, n_{ij}, P_{ij})$$

Cuando se intenta maximizar esta función una manera de hacerlo es realizando derivadas parciales sobre cada uno de los ratings de cada equipo e igualándolas a cero y resolver tantas ecuaciones como equipos estén participando, sin embargo se hace muy difícil de resolver dada la forma de la función. Para solucionar este problema se puede derivar con respecto al logaritmo natural de F y dado que ambas derivadas tienen el mismo signo ambas se maximizan en el mismo punto.

Un problema en este tipo de metodología es que si un equipo no pierde ningún partido se puede dar el caso de que su rating sea infinito, si bien esto es poco probable, existen casos excepcionales donde esto ocurre.

6.3. SISTEMA ELO

El sistema Elo fue inventado por Arpad Elo y en sus comienzos su función principal fue mejorar el sistema de clasificación de ajedrez imperante hasta esos momentos. Desde ahí en más, dada su alta sofisticación fue implementado en muchos otros deportes.

Una característica singular de este tipo de rating es su constante actualización a medida que los jugadores o equipos van disputando encuentros. Esta característica permite hacer una distinción entre derrotar a equipos débiles o fuertes (rating alto o bajo respectivamente), entregando una mayor puntuación a equipos y/o jugadores que vencen a oponentes más difíciles que a los que, por el contrario, vencen a equipos más débiles.

Otra característica importante de este sistema es que asume que la diferencia en scoring en un partido disputado se puede representar de buena manera como una regresión logística entre los rating de los equipos que disputan el encuentro.

El sistema Elo como se utiliza actualmente, requiere que todos los equipos participantes comiencen con un rating inicial. Mientras los equipos

van disputando partidos su rating previo es utilizado para determinar el nuevo rating. Matemáticamente el rating del equipo i luego de disputa un encuentro contra el equipo j será:

$$Rating_{2i} = Rating_{1i} + K (g_{ij} - \mu_{ij}) \quad (1)$$

Donde $Rating_{1i}$ representa el rating del equipo i antes de disputar el partido contra el equipo j , el rating posterior será $Rating_{2i}$.

Otra variable de la ecuación (1) es g_{ij} la cual en la mayoría de los usos que tiene esta metodología pueden ser asignados a tres posibles valores, uno, un medio o cero dependiendo si el equipo ha ganado empatado o perdido contra el equipo j respectivamente.

Por su parte μ_{ij} tiene la función de entregar mayor o menor peso al mérito de obtener un resultado. A medida que su valor es menor significará que es más meritorio obtener un empate o victoria por lo tanto su rating aumentará en mayor medida que cuando μ_{ij} sea mayor. La gran característica de esta variable es que se asume es una función logarítmica de su diferencia en rating. De esta manera μ_{ij} se calcula mediante la siguiente función:

$$\mu_{ij} = \frac{1}{1 + 10^{-d_{ij}/400}}$$

En esta ecuación $d_{ij} = Rating_{1i} - Rating_{2i}$ refleja la diferencia en rating entre los 2 equipos antes de disputar el partido. Una gran ventaja de esta variable μ_{ij} es que al cumplir la funcionalidad de corregir por mérito se le puede agregar de manera sencilla variables importantes tales como el efecto localía aumentando el valor d_{ij} al equipo que actúa de local.

La constante K funciona principalmente para ajustar los cambios relativos entre antiguos y nuevos ratings. A mayores valores para la variable K un resultado favorable en un encuentro se reflejará de inmediato en un aumento considerable en su rating, por el contrario para valores de K pequeños las variaciones en rating serán menores. Estos valores de K pueden variar por campeonato o por fases dentro de un mismo campeonato. Por ejemplo en el ranking ATP (tenis) los puntos por ganar un encuentro en las fases finales son cada vez mayores en comparación con las iniciales, y por otro lado, los torneos grandes como los Grand Slam entregan

mayor puntaje que cualquier otro torneo en igualdad de fase. Lo mismo ocurre en fútbol donde los partidos amistosos entregan mucho menor puntaje al Ranking FIFA que los partidos enmarcados en competencias internacionales.

En resumen, de forma primera se debe destacar que los sistemas de máxima verosimilitud y mínimos cuadrados no hacen diferencia en cuanto a importancia de partidos sino que simplemente se ajustan a los resultados obtenidos independiente se esté jugando un partido amistoso o un torneo de mayor envergadura. Por su parte el sistema Elo es mucho más parecido al que implementa actualmente la FIFA ya que por una parte genera diferencias por importancia de torneos y también tiene la característica de ajustar los puntajes de un partido por dificultad del mismo (ya sea por jugar contra equipos más difíciles o si se actúa de local). Otra importante diferencia es que los sistemas de máxima similitud y de mínimos cuadrados no actualizan rating partido a partido sino que muestran un rating general final dada una serie de partidos, además requieren de una relación entre todos los equipos existentes ya sea por haber jugado un partido entre ellos o por jugar contra algún competidor en común.

En esta tesis se utilizarán componentes tanto del método de máxima verosimilitud como del sistema Elo pero desde una mirada diferente. Por un lado, a partir de la data histórica de partidos entre selecciones, se estimarán parámetros de ataque y defensa para cada país utilizando el criterio de máxima verosimilitud. Con estos parámetros se simularán campeonatos que permitan generar un ranking referencial, que servirá como base para la comparación de metodologías propuestas.

Por otro lado, dado que el sistema Elo es el más completo de los sistemas analizados ya que es capaz de integrar tanto diferencia de ranking como factor localía en su modelo, y cualquier variable que se considere relevante, se intentará adaptar el Ranking FIFA acercándose a la idea del Ranking Elo. Se analizarán los datos disponibles con tal de entregar una respuesta numérica, con base estadística, a la importancia del factor localía en los resultados de un partido, con el fin de agregar esta variable como propuesta de reestructuración al actual Ranking FIFA. Al mismo tiempo esta tesis intentará encontrar una relación de equivalencia entre esta variable mencionada y la diferencia de ranking de los equipos al momento de disputar un encuentro, con la finalidad de integrarlas de manera conjunta al sistema de puntajes utilizado en la actualidad.

7. DATOS

La base de datos está compuesta por 8049 partidos, los cuales constituyen la totalidad de partidos jugados entre los años 2005 y Octubre del 2013, entre selecciones adultas de fútbol asociadas a alguna de las 6 confederaciones continentales que son miembros de FIFA.

Además de los equipos que disputan el partido, la base de datos cuenta con información para determinar si el partido se jugó en territorio neutral o no, el resultado final del partido, el tipo de competencia que se disputaba y la posición en el ranking de los equipos al momento de disputar el encuentro. Una breve descripción sobre la data se muestra en la Tabla 4

Tabla 4: Partidos de Fútbol de Selecciones nacionales adultas 2005- Oct 2013

	Amistosos	Clasificatorias Mundial	Fase Final Torneo de Confederación	Copa del Mundo	Total
Con Localía	2716	3079	189	10	5994
Neutros	1170	282	485	118	2055
Total	3886	3361	674	128	8049

Como se puede apreciar el mayor porcentaje de partidos de la data se encuentra distribuido entre partidos amistosos y clasificatorios al mundial alcanzando alrededor de un 90%. El restante 8,5% consisten en partidos de la fase final de los torneos continentales y el 1,5% a los 2 mundiales que están incluidos en la base de datos (Alemania 2006 y Sudáfrica 2010).

En cuanto a la relación de partidos jugados en territorio neutral o con alguno de los dos equipos siendo local, se puede ver fácilmente que de la totalidad de partidos cerca del 75% corresponden a estos últimos. Dado que este tipo de partidos representan un alto porcentaje de la base de datos y además es el factor principal a analizar por su importancia en el resultado de un partido, se realizó un análisis de éstos.

Revisando la Tabla 5 se puede apreciar de manera inmediata que el porcentaje de partidos ganados por el equipo dueño de casa es significativamente mayor que el del equipo visitante, lo cual entrega un

indicio que la localía es un factor clave en el resultado de un partido. Este tema se abordará en los modelos de la siguiente sección.

Tabla 5: Distribución de resultados disputados con Localía

	Local	Empate	Visita	Total
N° de Partidos	3043	1398	1553	5994
Porcentaje	50,8%	23,3%	25,9%	100%

Dado que actualmente, el rating FIFA en su metodología toma en consideración la posición en el ranking de los equipos que disputan un partido pero no así el efecto de jugar de local, una de los principales resultados que se pretende obtener es además de responder si estadísticamente este último factor es relevante en el resultado de un partido, de ser efectivo, poder compararlo cuantitativamente con la diferencia de ranking de los equipos.

8. MODELOS PROPUESTOS

En la literatura existen dos métodos principales para predecir el resultado de un evento deportivo, por un lado están los modelos estadísticos y por otro la evaluación o juicio de expertos. En cuanto a los primeros, numerosos modelos han sido utilizados, cada uno dependiendo de lo que se quiera modelar.

Si la idea principal es determinar las probabilidades de triunfo en un partido en el cual la existencia de un empate no existe, como en el tenis, los enfoques principalmente utilizados han sido los modelos Logit o Probit. En el caso de posibilidad de la existencia de empate como en el básquetbol o el fútbol los principales modelos utilizados han sido Probit ordenados y Logit Multinomial. Por su parte si las posibilidades de triunfo están basadas en diferencia en score o goles convertidos los modelos principalmente utilizados son regresiones de Poisson o modelos Binomial Negativa, los cuales son capaces de trabajar de muy buena manera con datos de naturaleza discreta [10].

En esta tesis el principal enfoque a utilizar será el modelo Logit Multinomial. Este modelo es utilizado para predecir las probabilidades de diferentes resultados posibles de una variable dependiente distribuida categóricamente, dado un conjunto de variables independientes. La gran ventaja de este modelo es que es utilizado principalmente para problemas multiclasas, es decir, para elecciones con más de 2 resultados posibles.

En este modelo se asume que la probabilidad de que la variable dependiente Y_i sea observada en una categoría s de un total de S posibles está dada por [11]:

$$P(Y_i = s) = \frac{e^{\beta'_s x_i}}{\sum_{k=1}^S e^{\beta'_k x_i}}$$

En deportes donde existe la posibilidad de empate, este tipo de modelo es ampliamente utilizado. Específicamente en el fútbol con el resultado de un partido tomando los posibles valores 1, 0 ó -1 en caso de triunfo, empate o derrota respectivamente, covariables denotadas por X y tomando el empate

como una categoría de referencia el modelo asume que Y tiene una distribución Multinomial que se expresa como sigue:

$$p_1 = P(Y = 1 | X_1, X_{-1}) = \frac{\exp(\beta^T X_1)}{1 + \exp(\alpha + \beta^T X_1) + \exp(\alpha + \beta^T X_{-1})}$$

$$p_0 = P(Y = 0 | X_1, X_{-1}) = \frac{1}{1 + \exp(\beta^T X_1) + \exp(\beta^T X_{-1})}$$

$$p_{-1} = P(Y = -1 | X_1, X_{-1}) = \frac{\exp(\beta^T X_{-1})}{1 + \exp(\beta^T X_1) + \exp(\beta^T X_{-1})}$$

Cabe mencionar que esta ecuación se construye a partir de la componente determinística de la función de utilidad de cada opción:

$$V_i = \beta^T X_i + \varepsilon_i$$

Donde ε_i distribuye Gumbel o valor extremo tipo I y la matriz X contiene todas las covariables que intentan explicar el resultado de un partido y que son características de cada equipo en un encuentro determinado.

8.1. MODELO N°1: MODELO BASADO EN RANKING.

El modelo más sencillo para representar las probabilidades de un partido está dado por considerar que éste se puede explicar solamente por la diferencia en Ranking de los dos equipos que disputan un encuentro. El modelo es el siguiente:

$$V_i = \beta_0 + \beta_{difstrong} * difstrong_i$$

Cabe mencionar que dado que se tienen los datos del ranking FIFA de los países al momento del partido se hizo una transformación de esta variable tal como lo hace el actual modelo FIFA. La fuerza o strong del equipo será:

$$Strong_i = 200 - ranking_i$$

De esta manera un número de Ranking menor equivale directamente a un equipo con mayor strong.

En este modelo $difstrong_i$ es una variable entera que representa la diferencia existente entre el ranking del equipo i y su oponente, es decir si el equipo i con un strong de 60 disputa un encuentro contra un equipo de strong 100 esta variable tomará el valor de -40. Por otra parte si se mira desde el punto de vista del oponente esta variable tomará el valor de 40.

La estimación de parámetros para el modelo de regresión logística se realiza con el método de máxima verosimilitud, utilizando el software R Studio, ya que la estimación por mínimos cuadrados no es capaz de producir estimadores insesgados de mínima varianza para los parámetros. Este método, como lo dice su nombre, pretende maximizar la función de verosimilitud, es decir, ser capaces de encontrar los $k+1$ parámetros para los cuales la probabilidad de observar los datos sea la más alta posible, donde k son las cantidad de características de un equipo al momento del match.

Los resultados de este modelo se presentan a continuación:

Tabla 6: Resultados, Modelo N°1

Variable Independiente	Valor Coeficiente	Error
β_0	0,3071	0,0274
<i>Difstrong</i>	0,0131	0,0003
Log Verosimilitud: -7.704		
AIC: 15.412		BIC: 15.426

Como era de esperar el signo del parámetro $difstrong$ es positivo, por lo que un equipo con un strong mayor tendrá mayores probabilidades de ganar que el equipo adversario. Además ésta es una variable significativa en el modelo ya que en el intervalo $[\beta_{difstrong} - 1,96 * error, \beta_{difstrong} + 1,96 * error]$ no se encuentra el valor 0.

El signo y valor del intercepto indica que ante partidos con equipos de similar ranking, siempre será más probable que el encuentro termine con alguno de los 2 ganando a que el resultado final sea un empate. Adicionalmente al comparar ambos estimadores se puede apreciar que las probabilidades de empatar o ganar de un equipo se igualan cuando el equipo presenta $\frac{\beta_0}{Difstrong} \approx 23$ unidades menos en la variable strong en comparación al adversario.

Si bien estos resultados demuestran por un lado que el ranking de los equipos no está del todo errado, ya que funciona como un buen predictor de lo que efectivamente ocurre, no dice nada sobre la evolución del ranking de un equipo en el tiempo, es decir, no se sabe si los cambios en el ranking de un determinado país se explican de buena forma por los resultados obtenidos por él.

8.2. MODELO N°2: MODELO BASADO EN RANKING Y LOCALÍA

El segundo modelo a considerar incluye la variable localía. Como se explicó anteriormente, se justifica el factor localía tanto por una mejora del desempeño del equipo local, así como también una disminución en el nivel de juego del equipo visitante, esto ya sea por el conocimiento de la superficie de juego, apoyo de la hinchada y/o presión hacia los árbitros, entre otros.

El modelo a considerar será el siguiente:

$$V_i = \beta_0 + \beta_{difstrong} * difstrong_i + \beta_{localia} * localia_i + \beta_{visita} * visita_i$$

En este modelo al igual que en el anterior $difstrong_i$ es una variable entera que representa la diferencia existente entre el ranking del equipo i y su oponente.

Por su parte las variables $localia_i$ y $visita_i$ son binarias y representan si el equipo que está disputando el encuentro lo disputa en su recinto o no. Si un partido lo está disputando el equipo i de local contra el equipo j , para ese partido las variables $localia_i$ y $visita_j$ tendrán un valor de 1, mientras que $localia_j$ y $visita_i$ estarán asignados con un valor 0. Es importante notar que en la base de datos existen partidos que se disputan en territorio neutral, por lo tanto ambas variables toman el valor 0 para los dos equipos en disputa.

Los resultados de este modelo se presentan a continuación:

Tabla 7: Resultados, Modelo N°2

Variable Independiente	Valor Coeficiente	Error
β_0	0,3019	0,0519
<i>Difstrong</i>	0,014	0,0003
<i>Localia</i>	0,3464	0,0614
<i>Visita</i>	-0,5058	0,0645
Log Verosimilitud: -7.409		
AIC: 14.828 BIC: 14.856		

Dado los resultados del modelo, es posible inferir que las 3 variables presentes son estadísticamente significativas, debido a que el intervalo de confianza de todos los parámetros a un 95% de confianza no incluyen el valor 0.

En cuanto a la variable Localía, se puede ver que su signo es positivo, hecho que era de esperarse ya que los datos mostrados en secciones anteriores mostraban una tendencia clara a mejores resultados para los equipos que disputaban el partido en su recinto. Por otro lado la variable Visita toma un valor negativo, e incluso en módulo mayor que la variable Localía, por lo cual se puede inferir que al jugar uno de los equipos de local son más preponderante los factores que impiden un buen funcionamiento del equipo visitante, sean estos cansancio por el vuelo, mal arbitraje, o escaso conocimiento del campo de juego, que los que juegan a favor del elenco dueño de casa.

Otro de los resultados importantes que se pretendían obtener, corresponde a la comparación de cuanta diferencia de ranking equivale el factor localía, lo que es posible obtener despejando la siguiente ecuación.

$$\beta_{Localia} - \beta_{Visita} = 2 * x * \beta_{Difstrong}$$

Reemplazando con el resultado de cada parámetro se tiene:

$$0,3464 + 0,5058 = 2 * 0,014 * x$$

$$x = \frac{0,8522}{0,028} = 30,4$$

Lo anterior indica que los partidos más parejos dentro del ámbito de selecciones adultas se producen cuando el equipo local tiene un strong30 unidades más bajo en comparación al equipo visitante.

Al ser este factor Localía una variable tan determinante en el resultado de un partido, es de esperar que los modelos o metodologías de construcción de ratings robustos sí lo tengan considerado como un factor relevante.

Finalmente, el mejor criterio para decidir si un modelo ajusta mejor que otro es el criterio de información Bayesiana (BIC) ya que castiga el número de parámetros adicionales de un modelo sobre otro de manera más fuerte que el criterio de información de Akaike (AIC). La fórmula del BIC es la siguiente:

$$BIC = -2Ln\hat{L} + kLn(n) \text{ en este caso } k=4 \text{ y } n=8049$$

El valor del BIC aumenta ya sea por la cantidad de variables o por la varianza no explicada de la variable dependiente, por lo tanto un menor BIC muestra un modelo que ajusta de mejor manera a los datos, que es lo que ocurre con este modelo en comparación al anterior.

8.3. MODELO N°3 MODELO BASADO EN RANKING, LOCALÍA NACIONAL Y CONTINENTAL

La variable localía continental fue tomada en consideración debido a los resultados obtenidos en los torneos más importantes a nivel de selecciones adultas, de los cuales 13 de 14 fueron ganados por equipos del mismo continente del país organizador, principalmente cuando se disputaban en territorio sudamericano o europeo.

Tabla 8: Copas mundiales disputadas en Sudamérica y Europa

Año	País Organizador	Campeón	Subcampeón
1930	Uruguay	Uruguay	Argentina
1934	Italia	Italia	Checoslovaquia
1938	Francia	Italia	Hungría
1950	Brasil	Uruguay	Brasil
1954	Suiza	Alemania Federal	Hungría
1958	Suecia	Brasil	Suecia
1962	Chile	Brasil	Checoslovaquia
1966	Inglaterra	Inglaterra	Alemania Federal
1974	Alemania federal	Alemania federal	Países Bajos
1978	Argentina	Argentina	Países Bajos
1982	España	Italia	Alemania Federal
1990	Italia	Alemania Federal	Argentina
1998	Francia	Francia	Brasil
2006	Alemania	Italia	Francia

Fuente: Elaboración propia, en base a información obtenida de la página:
www.sobrefutbol.com/torneos/mundial_detalle.htm

El modelo a considerar será el siguiente:

$$V_i = \beta_0 + \beta_{difstrong} * difstrong_i + \beta_{localia} * localia_i + \beta_{visita} * visita_i + \beta_{continente} * Continente_i$$

Como se puede apreciar a este modelo se le agrega la variable binaria $Continente_i$, que tomará el valor 1 si el partido se disputa en el continente del equipo i y 0 en caso contrario. Es importante señalar que el factor de variable continental se considerará solamente para los partidos en que no exista localía nacional. Es decir en el caso que la variable $localia_i$ tome el valor 1 la variable $Continente_i$ tomará un valor de 0.

Los resultados de este modelo se presentan a continuación:

Tabla 9: Resultados, Modelo N°3

Variable Independiente	Valor Coeficiente	Error
β_0	0,1478	0,0895
<i>Difstrong</i>	0,0140	0,0003
<i>Localía</i>	0,5005	0,0954
<i>Visita</i>	-0,3517	0,0974
<i>Continente</i>	0,2049	0,0975
Log Verosimilitud: -7.407		
AIC: 14.825 BIC: 14.860		

La Tabla 9 muestra que la variable Continente sí es una variable significativa y posee signo positivo, por lo cual un equipo que actúa en su continente tendrá mayores probabilidades de ganar que su equipo contrincante, el cual juega fuera de su continente. Además como era previsible, cuantitativamente el efecto de localía continental es mucho menor que el efecto de localía nacional.

Si bien la variable Continente es significativa, se puede ver que el BIC de este modelo es mayor que el BIC del modelo N°2. Es decir, el aumento de log verosimilitud de este modelo no compensa la agregación de un parámetro extra. Por lo tanto el modelo N°2 continúa siendo el que mejor ajusta a la data.

8.4. MODELO N°4: MODELO BASADO EN RANKING, LOCALÍA Y FACTOR CONFEDERACIÓN

Este modelo pretende identificar, si al categorizar a los países por la confederación de la cual provienen se obtiene un mejor ajuste. La variable Confederación se incluye, dado que en la actual metodología para la construcción del Ranking FIFA es utilizada.

El modelo a considerar será el siguiente:

$$V_i = \beta_0 + \beta_{difstrong} * difstrong_i + \beta_{localia} * localia_i + \beta_{visita} * visita_i + \beta_{AFC} * AFC_i + \beta_{OFC} * OFC_i + \beta_{CAF} * CAF_i + \beta_{CONMEBOL} * CONMEBOL_i + \beta_{UEFA} * UEFA_i$$

Donde AFC_i , OFC_i , CAF_i , $CONMEBOL_i$ y $UEFA_i$ representan variables binarias que indican si el equipo pertenece o no a esa confederación. En este modelo se omitió la confederación CONCACAF, por lo tanto será tomada como la confederación de referencia.

Los resultados de este modelo se presentan a continuación:

Tabla 10: Resultados, Modelo N°4

Variable Independiente	Valor Coeficiente	Error
β_0	0,1478	0,0895
<i>Difstrong</i>	0,0140	0,0003
<i>Localía</i>	0,3601	0,0954
<i>Visita</i>	-0,4670	0,0974
<i>AFC</i>	0,0611	0,060
<i>OFC</i>	0,4770	0,2012
<i>CONMEBOL</i>	0,3136	0,1021
<i>UEFA</i>	-0,1454	0,0631
<i>CAF</i>	0,1705	0,0633
Log Verosimilitud: -7.391		
AIC: 14.800 BIC: 14.862		

Al analizar estos resultados se puede observar que la variable AFC no es significativa en el modelo, es decir, estadísticamente no es distinta de 0 por lo tanto es igual a la variable omitida, que en este caso es CONCACAF, por lo cual países de ambas confederaciones se comportan de manera

similar. Por otro lado países de la CONMEBOL, OFC y CAF presentan una mejor performance que CONCACAF y países de la UEFA tienen peores resultados.

Lo importante a dilucidar es si la inclusión de las variables que representan a las confederaciones ayuda a mejorar el ajuste del modelo. Nuevamente para responder esto se debe observar el criterio de información bayesiana. Dado que en este modelo se agregan más variables se debe revisar si el aumento en la log verosimilitud es lo suficiente para determinar que el modelo mejora.

Como se comentó anteriormente basta con comparar el BIC de este modelo y el del modelo N°2. Los resultados del último modelo muestran que el ajuste no mejora, ya que su BIC es mayor. Dado esto el Modelo N°2 continúa siendo aquel que mejor explica la data con el mínimo número de parámetros.

En la Tabla 11 se muestra un resumen de las principales características y resultados de los modelos.

Tabla 11: Resumen de resultados

	Log Verosimilitud	N° Parámetros estimados	AIC	BIC
Modelo N°1	-7704	2	15412	15426
Modelo N°2	-7409	4	14828	14856
Modelo N°3	-7407	5	14825	14860
Modelo N°4	-7391	9	14800	14862

9. BOOKMAKERS Y EL FACTOR LOCALÍA.

En este apartado se intentará cuantificar cómo las distintas casas de apuestas integran el factor localía en sus cuotas de cada partido para posteriormente compararla con el resultado encontrado en la sección anterior. De la totalidad de datos incluidos en la base, se tomó una muestra de 1653 partidos, correspondientes a los encuentros disputados entre los años 2009-2010. De estos partidos se obtuvieron los promedios de las cuotas asignadas por las casas de apuestas más populares a nivel mundial para 1306 partidos [12]. Los partidos a los cuales no se encontraron los datos de apuestas son principalmente encuentros amistosos de países africanos o asiáticos de segundo orden en lo que respecta a fútbol.

Una característica muy importante de las cuotas que fijan las casas de apuestas es que están relacionadas directamente a las probabilidades implícitas de observar un resultado, ya sea victoria del local, empate o victoria visitante. Esta relación se especifica a continuación.

Denotemos por $O_{ik}^L, O_{ik}^E, O_{ik}^V$ las cuotas que ofrece la casa de apuesta k para la victoria del equipo local, empate, y victoria visita en el partido i respectivamente. La probabilidad implícita $P_{ik}^L, P_{ik}^E, P_{ik}^V$ estará dada por [13]:

$$P_{ik}^L = \frac{1}{O_{ik}^L} * \frac{1}{Margen_{ik} + 1}$$

$$P_{ik}^E = \frac{1}{O_{ik}^E} * \frac{1}{Margen_{ik} + 1}$$

$$P_{ik}^V = \frac{1}{O_{ik}^V} * \frac{1}{Margen_{ik} + 1}$$

Donde $Margen_{ik} = \frac{1}{O_{ik}^L} + \frac{1}{O_{ik}^E} + \frac{1}{O_{ik}^V} - 1$ representa el margen de ganancia de la casa de apuesta k en el partido i.

Para dejar más claros estos conceptos se presentan ejemplos de partidos de la base de datos.

Tabla 12: Ejemplo de probabilidades implícitas a partir de cuotas

Partido	Cuotas			Margen	Probabilidades Implícitas		
	Local	Empate	Visita		Local	Empate	Visita
Iraq-Syria	1,9	3,28	3,81	9,4%	48,1%	27,9%	24,0%
Panamá-Perú	2,6	3,24	2,54	8,7%	35,4%	28,4%	36,2%
Honduras-Chile	2,91	3,18	2,27	9,9%	31,3%	28,6%	40,1%
Italia-Serbia	1,7	3,46	5,27	6,7%	55,1%	27,1%	17,8%
Ghana-Sudán	1,15	6,08	13	11,1%	78,3%	14,8%	6,9%

De aquí se puede observar que las distintas casas de apuestas operan con un margen de ganancia que oscila entre un 6% y un 12% y que intrínsecamente las cuotas de apuestas reflejan la probabilidad de que un resultado específico suceda.

Lo importante en este caso será poder demostrar si las casas de apuestas utilizan el factor localía como un dato importante para asignar las cuotas a cada partido, y si de ser así, ser capaces de comparar esta magnitud con la diferencia existente entre los rankings de los países que disputan un encuentro. Para esto, primero se analizan los 1306 partidos con apuestas encontradas y se realiza un análisis de regresión por separado. Por una parte, a los encuentros disputados con uno de los dos equipos siendo local (987 partidos), y los restantes, que fueron jugados en campo neutral (319 partidos).

La variable dependiente en el análisis de regresión será la diferencia en probabilidad de que el ganador sea el equipo 1 sobre que el ganador del encuentro sea el equipo 2. La variable dependiente será exclusivamente la diferencia en ranking entre el equipo 1 y el equipo 2 al momento de disputarse el encuentro.

El valor de R^2 mostrado en cada una de las ilustraciones es simplemente una medida estadística para medir qué tan cercana es la línea de tendencia a los datos encontrados en la realidad. Cabe mencionar que una perfecta relación entre datos y línea de tendencia se obtiene con un $R^2 = 1$. Como se puede apreciar en todas las ilustraciones existe una buena relación entre la diferencia en probabilidad de ganar entre el equipo 1 y 2 y su diferencia en ranking, en la cual casi un 65% de la varianza de la data puede ser explicada por la ecuación de mejor ajuste para el primero, y un 80% para el segundo.

Ilustración 1: Distribución de diferencias en probabilidad de victoria entre equipo 1 y 2 según su diferencia en ranking, partidos jugados en territorio neutral, 2009-2010

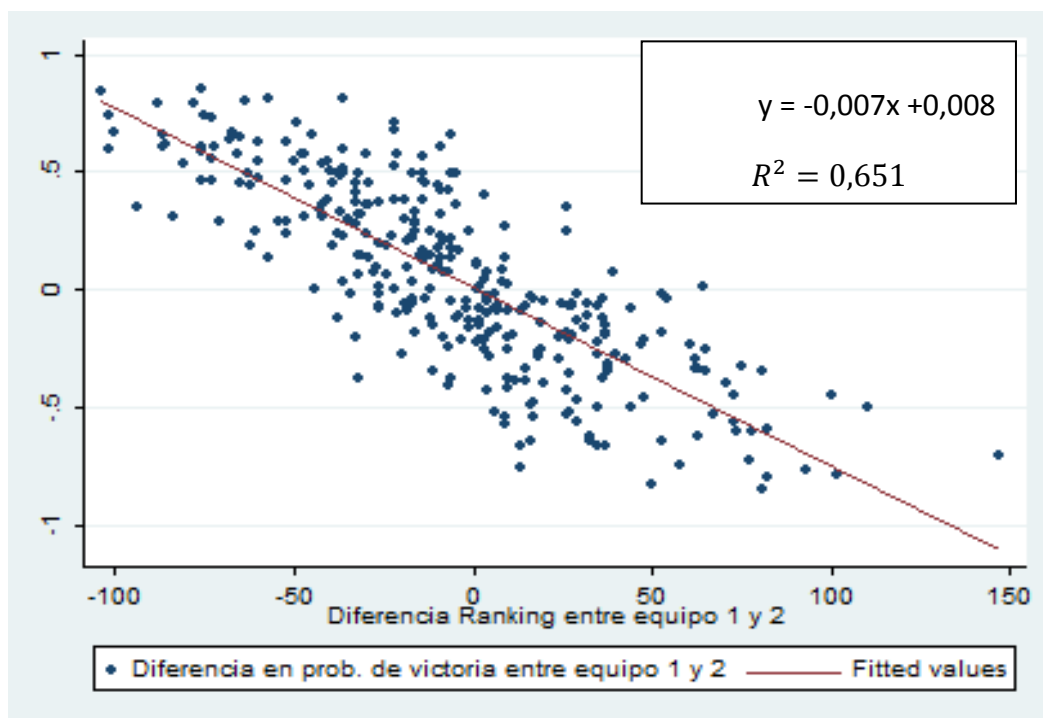
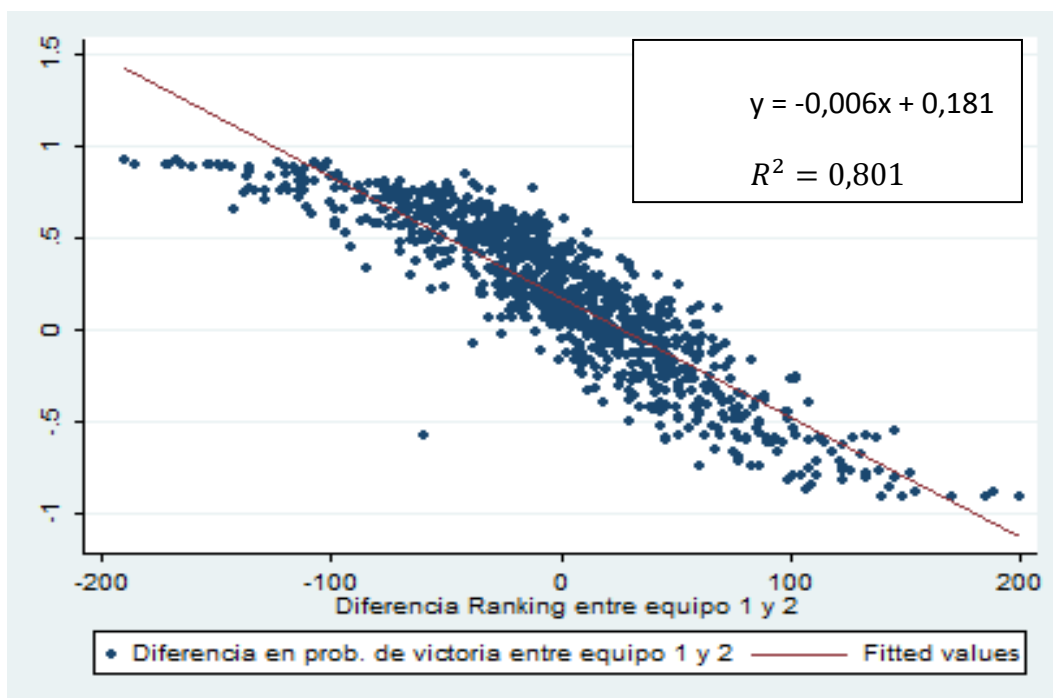


Ilustración 2: Distribución de diferencias en probabilidad de victoria entre equipo 1 y 2 según su diferencia en ranking, partidos jugados con el equipo 1 actuando de local 2009-2010.



Lo importante acá es identificar el resultado de las regresiones lineales que separan los partidos disputados con localía y en territorio neutral.

La línea de tendencia para los partidos jugados en territorio neutral es la siguiente:

$$y = -0,007x + 0,008 \quad (1)$$

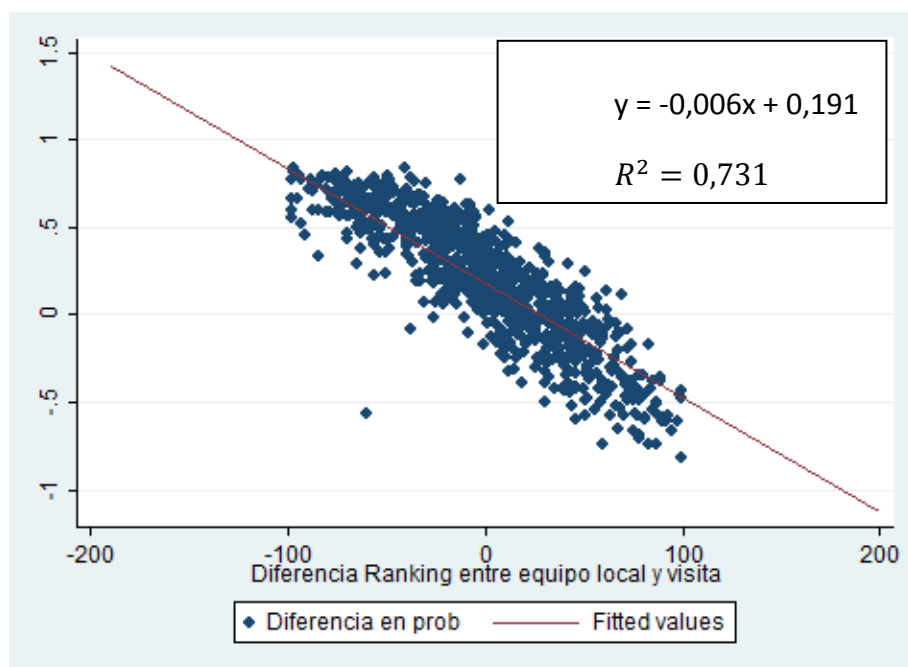
Mientras que para los partidos donde el equipo 1 actuaba de local la regresión es la siguiente.

$$y = -0,006x + 0,181 \quad (2)$$

Donde y es la diferencia en probabilidad de que el equipo 1 gane el partido por sobre que el equipo 2 alcance la victoria expresado como número decimal, y x es la diferencia en Ranking FIFA entre el equipo 1 y 2 al momento de efectuarse el partido.

En la ilustración 2 se puede observar claramente que cuando la diferencia de Ranking se encuentra entre -100 y 100 la relación lineal se hace más evidente. Por otro lado para valores extremos de esta variable se observa que no existen cambios significativos en la probabilidad de éxito del equipo local sobre el equipo visitante.

Ilustración 3: Distribución de diferencias en probabilidad de victoria entre equipo 1 y 2 según su diferencia en ranking entre -100 y 100. Partidos jugados con el equipo 1 actuando de local 2009-2010.



De estas 2 ecuaciones se desprende de manera sencilla cualquier diferencia de probabilidad de victoria entre 2 países ya sea se esté jugando en territorio neutral o con uno de ellos de local.

Nuevamente se analizarán algunos de los ejemplos mostrados anteriormente. En el partido de Iraq-Syria jugado en territorio iraquí los equipos estaban en la posición 100° y 107° respectivamente, luego dada la ecuación 2 el local tendrá un 22,3% más de probabilidad de lograr la victoria por sobre la visita, lo cual está muy cerca del 24,1% que entregan las apuestas. Por otro lado en el partido de Honduras-Chile jugado en Estados Unidos, territorio neutral, los equipos estaban en las posiciones 40° y 30° respectivamente, por lo cual dada la ecuación 1 en este caso el equipo 2 tendrá un 6,2% más de lograr la victoria en comparación al equipo 1.

Como se puede apreciar sólo conociendo los rankings de los equipos y si el partido se disputa en territorio neutral o no se puede obtener una muy buena estimación sobre la ventaja de un equipo sobre otro en las apuestas. Pero hay otros detalles muy interesantes y que se desprenden al comparar ambas ecuaciones. Por una parte se puede apreciar que las casas de apuestas parecieran sí utilizar el dato de la posición en el Ranking FIFA para estimar las cuotas de cada equipo al momento de un partido, por lo cual si bien este Ranking tiene fallas en la omisión de variables importantes y problemas estructurales, de igual manera es un buen predictor para obtener las probabilidades de triunfo de un equipo u otro.

Otro aspecto relevante es que era de esperar que la constante en la ecuación 1 fuera muy cercana a 0, ya que en un partido que se juega en territorio neutral es lógico pensar que ninguno de los 2 equipos debiera tener mayor probabilidad de ganar que el contrincante si es que poseen un ranking similar. El valor obtenido, cercano a un 1%, se atribuye directamente a la limitada cantidad de data para los encuentros jugados en territorio neutral, muy baja (319) en comparación a cuando se disputa un partido con uno de los equipos actuando de local (987). Por lo cual es de esperar que a medida que se puedan agregar más datos esta constante se acerque cada vez más a 0.

No obstante los detalles más importantes, que representan la razón por la cual se realizó este análisis, corresponden a revisar como las casas de apuesta incluyen el factor localía al momento de asignar las cuotas para cada uno de los resultados posibles, y a su vez, la asignación de probabilidades para éstos. Adicionalmente se pretende comparar este factor con su equivalente en diferencia de Ranking.

Comparando las constantes de las ecuaciones 1 y 2 se puede ver que el hecho de jugar de local otorga un $18,1\% - 0,8\% = 17,3\%$ de ventaja en probabilidad para los equipos que juegan de local, y dado que se espera que con mayor cantidad de datos la constante de la ecuación 1 se acerque más a 0, se puede decir que las casas de apuestas sí integran este dato como una variable importante y que su magnitud podría estar más cercana a un 18%. Estos resultados se acercan mucho a las estadísticas de estos compromisos mostrados en la Tabla 13, en la cual la diferencia porcentual entre victorias del local y de la visita es cercana a un 21%.

Tabla 13: Frecuencia de Resultados para partidos entre 2009-2010, jugados con localía.

Resultados	Frecuencia	Porcentaje (%)
Empate	218	22,09
Local	492	49,85
Visita	277	28,06
Total	987	100

Para resolver la equivalencia de este factor localía con una determinada diferencia de Ranking FIFA se puede obtener este resultado simplemente revisando el punto donde la recta de tendencia intersecta el eje x en la ecuación 2, ya que de esta manera se obtiene cual es la diferencia de ranking que equilibra un partido que se está jugando con el equipo 1 actuando de local. El resultado encontrado es $30,16 \approx 30$ similar a los 30,4 encontrados con el modelo Logit Multinomial obtenido en los resultados del capítulo 8.2 de esta tesis.

Otro aspecto no menos importante es que la probabilidad de empate que asignan las casas de apuestas se puede inferir de manera directa de la diferencia en probabilidad de victoria del equipo 1 por sobre el equipo 2, ya sea en partidos en territorio neutral o con localía. Es decir que a partir de las regresiones 1 y 2 se podrá obtener la diferencia en probabilidad de victoria

entre los equipos que disputan un partido, y con ello se puede obtener de manera inmediata la estimación de la probabilidad de empate. Algo importante que se puede ver a partir de la regresión cuadrática mostrada en las ilustraciones 4 y 5 es que la probabilidad de empate toma su valor máximo cuando los equipos están completamente equiparados, es decir, cuando la probabilidad de triunfo es similar para ambos.

Por otra parte, no existe mayor cambio en la probabilidad de empate, se esté desarrollando el partido o no en territorio neutral, ya que en ambos casos el valor máximo es cercano a un 29%.

Finalmente, dado que la probabilidad de empate depende exclusivamente de la diferencia de probabilidad de triunfo entre los 2 equipos, y que ésta a su vez depende principalmente del factor localía y de la diferencia de ranking entre ellos, se puede afirmar que las cuotas asignadas al resultado empate por las distintas casas de apuestas proviene de elegir al empate como un resultado de referencia al igual que en el modelo Logit realizado.

Ilustración 4: Distribución de empate según diferencia en probabilidad de victoria entre equipo 1 y 2, partidos jugados en territorio neutral en los años 2009-2010.

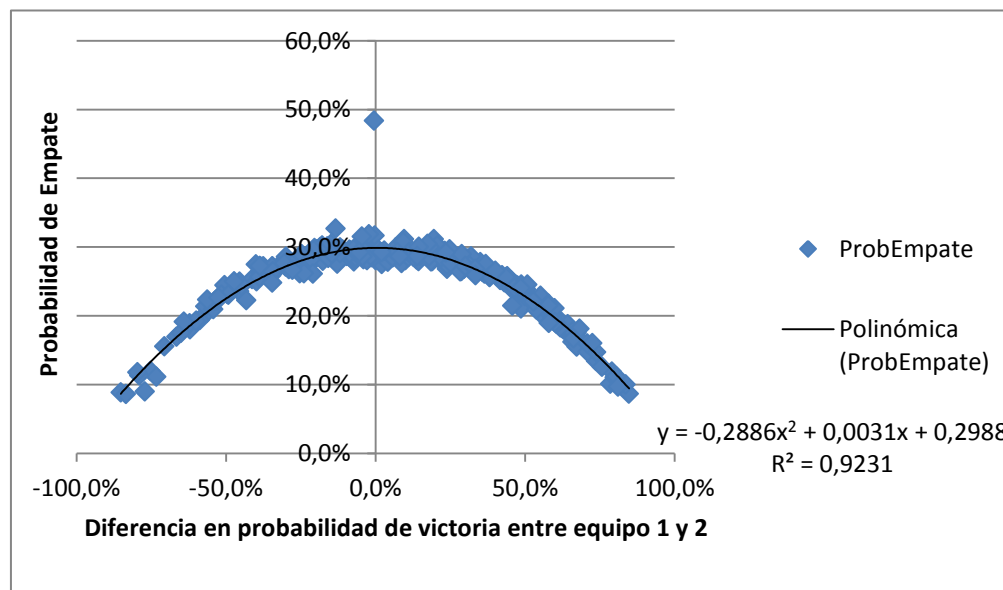
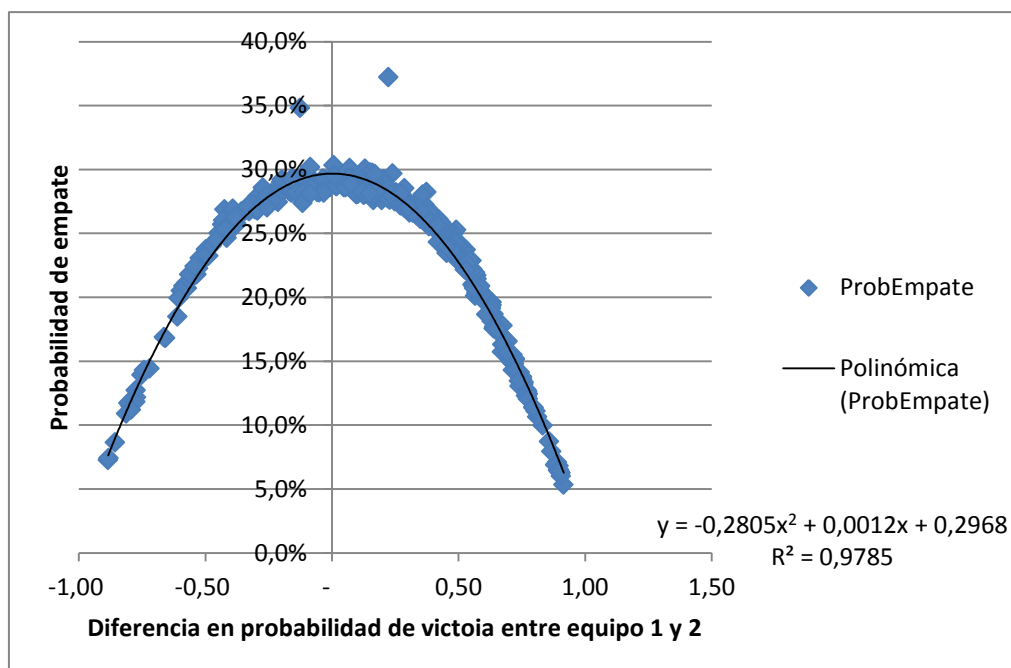


Ilustración 5: Distribución de empate según diferencia en probabilidad de victoria entre equipo 1 y 2, partidos jugados con equipo 1 de local en los años 2009-2010.



10. SIMULACIÓN DE METODOLOGÍAS Y PROPUESTA SELECCIONADA

A partir de los resultados encontrados en los capítulos anteriores, se procedió a generar distintas metodologías de construcción de ratings para las 209 selecciones adultas de fútbol. Todas las propuestas formuladas a continuación fueron implementadas utilizando el software Java.

10.1. METODOLOGÍA 1: CON LOCALÍA INCLUIDA Y VARIABLE CONFEDERACIÓN ELIMINADA.

Dado los resultados obtenidos en los modelos de la sección anterior, un modelo más robusto que el que tiene actualmente la FIFA se logra agregando la variable localía y eliminando la variable confederación. Por lo cual la primera propuesta de metodología de rating será la siguiente.

El puntaje de un equipo que disputa un partido fuera de su país será

$$Puntos = P * I * F$$

Mientras que el puntaje para un equipo jugando de local será:

$$Puntos = P * I * (F - 30)$$

P= 3 si es un triunfo, 1 en caso de empate, 0 en caso de derrota.

I= 4 si es Copa del Mundo, 3 si es fase final de torneo de Confederación (Copa América, Copa de oro, Eurocopa, etc.) o Copa Confederaciones, 2.5 si es eliminatoria mundialista o en el ámbito de confederación, y 1 si es un partido amistoso o torneo menor de confederación.

F=200-ranking del oponente. La resta de 30 será para equilibrar el hecho de que esté disputando su partido como local. Se debe dejar en claro que el mínimo valor posible para esta variable se mantiene en 50.

El ranking generado por esta metodología para Octubre del año 2013 se muestra en la Tabla 30 Anexos sección 2. (Se muestran los 200 mejores equipos). Esta fecha como ya se ha explicado es relevante dado que este ranking fue el utilizado para establecer los cabezas de serie del mundial del año siguiente.

10.2. METODOLOGÍA 2: CON PARTIDOS AMISTOSOS OMITIDOS.

Como se comentó anteriormente el actual sistema de generación de rating FIFA tiene serios errores estructurales. Dentro de los más graves se encuentra el hecho de que jugar amistosos puede ser perjudicial para el puntaje acumulado de un equipo dado que entrega muy pocos puntos incluso si terminan en triunfo, dado que se promedian con los puntos obtenidos hasta ese momento.

Por lo tanto una alternativa de mitigar este error estructural es que ningún partido amistoso sume puntos. Esta alternativa no se aleja de la realidad ya que en muchos de estos partidos los equipos deciden probar formaciones alternativas o no presentarse con sus mejores jugadores, lo cual desvirtúa el resultado del partido. Además en muchos otros deportes de alta competencia estos tipos de partidos no suman ni restan al score oficial. En esta metodología la omisión de los partidos amistosos se suma a la eliminación del factor Confederación y a la agregación del factor localía.

El ranking generado para Octubre del 2013 se muestra en la Tabla 31, Anexos Sección 2 (Se muestran los 200 mejores equipos).

10.3. METODOLOGÍA 3: REESTRUCTURACIÓN DE PUNTOS POR VICTORIA

Otro importante error estructural que se presenta bajo la metodología actual, es que existe demasiada diferencia en puntaje obtenido en un partido si se empata con los mejores del mundo en comparación a si se obtiene un triunfo ante equipos de medianía de tabla, en igualdad de formato de partido (ambos amistosos, o eliminatoria o mundial).

Mientras empatar con el mejor equipo entrega 200 puntos, derrotar al 100° del mundo entrega más de 250 puntos, lo cual no se condice con el mérito de obtener ambos resultados.

Es por esto que bajo esta metodología se pretende reasignar el puntaje por victoria, empate o derrota con tal de disminuir el error recién explicado. Para esto los puntajes asignados fueron 0, 1 y 2 para la derrota, empate y victoria respectivamente.

De esta manera alcanzar un empate con equipos del top ten del ranking será equiparable o muchas veces superior a derrotar a equipos débiles, que están de media tabla hacia abajo.

En este modelo, la reestructuración de puntos por victoria va acompañada de la omisión de amistosos y de factor confederación, y de la agregación del factor localía.

El ranking generado para octubre del año 2013 se muestra en la Tabla 32, Anexos Sección 2 (Se muestran los 200 mejores equipos).

10.4. METODOLOGÍA 4: SIN DEPRECIACIÓN DE PUNTAJE

El último error percibido en la estructura actual del rating FIFA es el hecho de que exista depreciación de puntaje siendo que no todas las confederaciones juegan sus partidos o torneos importantes en las mismas fechas. De esta manera, por ejemplo, equipos europeos sacan ventaja de países sudamericanos dado que su torneo continental se juega sólo 2 años antes del próximo mundial, mientras que la Copa América se realiza 3 años antes. Es así como es más probable que ante igualdad de resultados, un equipo europeo tenga más puntaje y sea cabeza de serie en el mundial.

Una manera simple de corregir este error es no depreciar el puntaje por año. Esta propuesta se basa en que la metodología de construcción de Rating FIFA es de las pocas que presenta en su estructura esta característica.

Si se compara con el tenis, ambos deportes acumulan puntaje durante un periodo de tiempo, un año en caso del tenis y cuatro en caso del fútbol. Estos periodos se deben principalmente a que los ciclos de realización de torneos son anuales para el deporte blanco, mientras que para el fútbol el ciclo de torneos se completa cada 4 años.

Al no depreciar los puntajes por año, el score que presenta un país en cualquier momento reflejará de mejor manera su performance de los últimos años, abarcando de esta forma todos los torneos posibles en que una país puede participar.

Los resultados para octubre del año 2013 se muestran en la Tabla 33, Anexos Sección 2 (Se muestran los 200 mejores equipos).

10.5. MÉTODOS DE COMPARACIÓN DE RANKINGS Y RESULTADOS

Habiendo generado ya 4 propuestas de metodologías de creación de rating para selecciones de fútbol adultas, se necesita tener un criterio que permita discernir cuál de ellas es la mejor.

Un primer acercamiento para responder esto es que las metodologías debieran ser evaluadas por el ranking que generan. Una de las formas más comunes en la literatura para evaluar rankings ha sido mediante la minimización del error cuadrático medio (MSE).

El problema mayor acá es cómo definir la minimización del MSE para este tipo de resultados. Una buena aproximación es encontrar un Ranking Referencial que permita obtener numéricamente el MSE. Teniendo este ranking referencial la métrica de decisión queda como sigue:

$$MSE_j = \sum_{i \in N} \left(\frac{(Ranking_{ij} - RankingReferencial_i)^2}{N} \right)$$

Donde MSE_j es el error cuadrático medio de la metodología j , $Ranking_{ij}$ es el ranking del equipo i bajo la metodología j y $RankingReferencial_i$ es la posición del equipo i bajo un Ranking de Referencia escogido. N es la cantidad de países incluidos en el modelo.

El mejor resultado será entonces la metodología que reporte un menor MSE, es decir, la metodología que construya rankings que más se asemejen al Ranking de Referencia.

Si la idea principal de un ranking es mostrar qué equipo es más probable que resulte triunfador en un campeonato, el resultado del mundial puede ser una buena propuesta de Ranking Referencial. El ordenamiento final de los países en el mundial Brasil 2014 se presenta en la Tabla 14.

Tabla 14: Posición Final Copa del Mundo Brasil 2014

Pos.	País	Pos.	País	Pos.	País	Pos.	País
1	Alemania	9	Chile	17	Ecuador	25	Ghana
2	Argentina	10	México	18	Portugal	26	Inglaterra
3	Holanda	11	Suiza	19	Croacia	27	Corea del Sur
4	Brasil	12	Uruguay	20	Bosnia y Herzegovina	28	Irán
5	Colombia	13	Grecia	21	Costa de Marfil	29	Japón
6	Bélgica	14	Argelia	22	Italia	30	Australia
7	Francia	15	Estados Unidos	23	España	31	Honduras
8	Costa Rica	16	Nigeria	24	Rusia	32	Camerún

Fuente: Elaboración propia basada en información obtenida de la página: Obtenido de http://www.elgraficochile.cl/las-posiciones-finales-del-mundial-de-brasil-2014/prontus_elgrafico/2014-07-13/225031.html

Para poder obtener el MSE asociado a cada metodología se extrajo el ranking generado por cada una en Octubre del año 2013. Posteriormente se filtró y generó un ordenamiento relativo entre los 32 países que participaron en el mundial. La Tabla 15 muestra el MSE de cada una de las metodologías propuestas y el Ranking FIFA existente a Octubre del año 2013.

Tabla 15. Resultados Metodologías propuestas, Ranking de Referencia basado en resultados Mundial 2014.

Modelo	MSE
Ranking FIFA, Octubre 2013	1983
Metodología 1: CON LOCALÍA INCLUIDA Y VARIABLE CONFEDERACIÓN ELIMINADA	2337
Metodología 2: CON PARTIDOS AMISTOSOS OMITIDOS.	2638
Metodología 3: REESTRUCTURACIÓN DE PUNTOS POR VICTORIA	2828
Metodología 4: SIN DEPRECIACIÓN DE PUNTAJE	3436

Estos resultados muestran que bajo el ranking de referencia ordenado de acuerdo al resultado del Mundial, el Ranking FIFA supera a todas las metodologías propuestas. Sin embargo, se debe tener en consideración que el resultado del Mundial se encuentra fuertemente influenciado por los cabezas de serie designados, los cuales provienen directamente del Ranking

FIFA. Es debido a esto que es necesario encontrar sistemas de referencias que eliminen este sesgo, para obtener resultados más robustos.

Se deja como propuesto pero no desarrollado en esta tesis la evaluación de las distintas metodologías en los resultados obtenidos en la Eurocopa 2012, en la cual los cabezas de serie no fueron designados por el Ranking FIFA sino que por el Coeficiente UEFA¹. Se espera que los resultados evaluando este torneo muestren una mejoría de las metodologías propuestas por sobre el Ranking FIFA.

Por otra parte, si la idea principal de un ranking es reflejar qué equipo llega de mejor manera a una competencia internacional, sin esperar los resultados obtenidos en él, resulta lógico pensar en utilizar la data histórica de los últimos años para generar un Ranking de Referencia. Para la obtención de este ranking se implementó un sistema de simulación de torneos, que se detalla a continuación.

En cada simulación de torneo, la predicción de cada partido está basada en un análisis de regresión. Para esto se tiene que, dado un partido entre los equipos A y B, la cantidad de goles de cada uno es modelada como una variable aleatoria con distribución Poisson. Luego, la tasa con la cual el equipo A hace goles depende de 3 variables:

- La capacidad ofensiva del equipo A
- La capacidad defensiva del equipo B
- Si el equipo actúa de local, neutral o de visita.

Formalmente el modelo queda de la siguiente manera:

$$X_{AB} \sim \text{Poisson}(\lambda_{AB})$$

$$X_{BA} \sim \text{Poisson}(\lambda_{BA})$$

$$\text{Ln}(\lambda_{AB}) = \beta + a_A + d_B + \text{localia} + \text{visita}$$

$$\text{Ln}(\lambda_{BA}) = \beta + a_B + d_A + \text{localia} + \text{visita}$$

¹ <http://es.uefa.com/memberassociations/uefarankings/country/>.

Donde a_A es el parámetro ofensivo del equipo A, d_B es el parámetro defensivo del equipo B, y viceversa. Por su parte, tanto el parámetro localía como visita están relacionados al territorio donde se disputa el partido. Estas 2 últimas variables son modeladas de modo tal que no cambian según los equipos que disputan el partido. Además se debe dejar en claro que en un partido jugado en territorio neutral ninguno de los parámetros asociados a localía o visita influyen en el resultado.

Dado que el modelo se sustenta en que X_{AB} y X_{BA} son variables independientes, la probabilidad de los resultados de un partido será la multiplicación de la probabilidad de goles del equipo A por la probabilidad de goles del equipo B.

$$P(X_{AB} = m, X_{BA} = n) = \frac{e^{-\lambda_{AB}} \lambda_{AB}^m}{m!} * \frac{e^{-\lambda_{BA}} \lambda_{BA}^n}{n!}$$

Para estimar los parámetros $a_A, d_B, localia$ y $visita$ se utiliza el método de máxima verosimilitud. Para esto se utilizan los datos de todos los partidos desde el año 2009 hasta el año 2013, entre los 100 equipos que históricamente han ocupado las mejores posiciones en el Ranking FIFA, dentro de los cuales se encuentran los 32 participantes de la Copa Mundial. Dado que son 100 equipos, serán $2 * 100 + 2 = 202$ los parámetros a estimar.

A continuación se muestra un extracto de los resultados obtenidos, los cuales consisten en los parámetros de los 32 países integrantes del Mundial. La totalidad de los resultados de este modelo se pueden ver en la sección anexos sección 14.3.

Tabla 16. Parámetros en base a partidos 2005-2013 para los países participantes del Mundial Brasil 2014

País	Ataque	Defensa
Algeria	0,3147	0
Argentina	0,7608	-0,3229
Australia	0,5053	0
Belgium	0,5075	0
Bosnia-Herzegovina	0,6125	0
Brazil	0,8399	-0,2978
Cameroon	0,3681	0

Chile	0,6265	0
Colombia	0,54	-0,3192
Costa Rica	0,3757	0
Croatia	0,5325	-0,2653
Ecuador	0,6017	-0,2875
England	0,6239	-0,3391
France	0,5256	-0,4031
Germany	0,9013	0
Ghana	0,4109	-0,2303
Greece	0,2329	0
Honduras	0,4263	0
Iran	0	0
Italy	0,5395	-0,2463
IvoryCoast	0,6097	0
Japan	0,5635	0
KoreaRepublic	0,5569	0
Mexico	0,5351	-0,2441
Netherlands	0,7376	-0,3213
Nigeria	0,5111	0
Portugal	0,5169	0
Russia	0,4378	-0,3762
Spain	0,8069	-0,3109
Switzerland	0,5627	0
UnitedStates	0,4999	0
Uruguay	0,6344	-0,2552

Una vez obtenidos estos parámetros es necesario determinar qué formato de torneo se utilizará en las simulaciones para obtener un ordenamiento adecuado de los equipos.

Existen estudios que han analizado la influencia del diseño y estructura de un torneo en los resultados obtenidos por los distintos equipos. En particular el estudio de Scarfy Yusof [14], plantean que la probabilidad que el mejor equipo (el que presenta los mejores parámetros) salga campeón, varía considerablemente dependiendo del formato de torneo, siendo un torneo con formato Round Robin (todos contra todos) el sistema que maximiza la correlación entre el mejor equipo y resultado final. Tomando esto en consideración, se realizaron 100.000 simulaciones de Monte Carlo de torneos en territorio neutral con formato Round Robin entre los 32 equipos de la Tabla 16.

En la Tabla 17 se muestran los resultados de las simulaciones, ordenados de mayor a menor de acuerdo a la probabilidad de salir campeón.

Tabla 17. Resultados simulaciones de torneos con sistema Round Robin, en base a datos históricos

Orden	País	Prob. Campeón
1	Brazil	0,2445
2	Spain	0,20491
3	Argentina	0,15333
4	Netherlands	0,12498
5	Germany	0,05303
6	England	0,05265
7	France	0,0327
8	Uruguay	0,03048
9	Ecuador	0,02946
10	Colombia	0,02009
11	Croatia	0,01184
12	Russia	0,01076
13	Italy	0,01047
14	Mexico	0,00958
15	Chile	0,00193
16	Bosnia- Herzegovina	0,00189
17	Ghana	0,00189
18	IvoryCoast	0,00156
19	KoreaRepublic	0,00076
20	Switzerland	0,00072
21	Japan	0,00068
22	Belgium	0,00042
23	Australia	0,00032
24	Nigeria	0,00031
25	UnitedStates	0,00027
26	Portugal	0,00026
27	Honduras	0,00012
28	Cameroon	3,00E-05
29	Costa Rica	3,00E-05
30	Algeria	2,00E-05
31	Greece	1,00E-05
32	Iran	0

Tomando el ordenamiento de la Tabla 17 como un segundo Ranking de Referencia, los resultados de las metodologías propuestas de acuerdo a su MSE se muestran a continuación.

Tabla 18. Resultados Metodologías propuestas, Ranking de Referencia basado en simulaciones en base a data histórica.

Modelo	MSE
Ranking FIFA, Octubre 2013	2120
Metodología 1: CON LOCALÍA INCLUIDA Y VARIABLE CONFEDERACIÓN ELIMINADA	2602
Metodología 2: CON PARTIDOS AMISTOSOS OMITIDOS.	2332
Metodología 3: REESTRUCTURACIÓN DE PUNTOS POR VICTORIA	2236
Metodología 4: SIN DEPRECIACIÓN DE PUNTAJE	1834

De acuerdo a los resultados obtenidos, se tiene que utilizando el Ranking de Referencia basado en la data de los últimos 4 años, la metodología N°4 ajusta de mejor manera que el Ranking FIFA. Es importante notar que 3 de las 4 metodologías propuestas mejoraron significativamente su ajuste, en comparación al MSE obtenido usando los resultados del Mundial.

Un último Ranking de Referencia a generar será el proveniente de los datos de las cuotas asignadas por las casas de apuestas a cada partido disputado en el Mundial Brasil 2014. Como se explicó en el capítulo 9, las cuotas entregadas por las casas de apuestas explican de manera intrínseca las probabilidades asignadas al triunfo del equipo 1, triunfo del equipo 2 y empate. Aún más, en sus estudios Nils Rudi y Shanker [15] demuestran que para cualquier distribución de resultados (p_{-1}, p_0, p_1) existe un único par $(\lambda_{-1}, \lambda_1)$ capaz de replicarla. Para ejemplificar se muestra en la Tabla 19, 3 encuentros disputados en el mundial, con sus respectivas cuotas, y tasas que las generan.

Tabla 19. Ejemplos partidos Mundial 2014, Cuotas-Probabilidades y Lambdas

Partido		Cuotas			Probabilidades			Lambdas	
Equipo 1	Equipo 2	1	E	2	1	E	2	lambda1	lambda2
Brazil	Croatia	1,31	5,1	11,68	0,730	0,188	0,082	1,9949	0,4984

Spain	Chile	1,57	4,41	5,51	0,609	0,217	0,174	1,9212	0,9149
England	Italy	2,5	3	3,21	0,382	0,319	0,298	1,0181	0,8603

De la misma forma que antes, en cada simulación de torneo, la predicción de cada partido está basada en un análisis de regresión. La variable dependiente será la tasa obtenida para cada partido, mientras que las variables independientes serán:

- La capacidad ofensiva del equipo A
- La capacidad defensiva del equipo B
- Si el equipo actúa de local, neutral o de visita.

Formalmente el modelo queda de la siguiente manera:

$$X_{AB} \sim \text{Poisson}(\lambda_{AB})$$

$$X_{BA} \sim \text{Poisson}(\lambda_{BA})$$

$$\ln(\lambda_{AB}) = \beta + a_A + d_B + \text{localia} + \text{visita}$$

$$\ln(\lambda_{BA}) = \beta + a_B + d_A + \text{localia} + \text{visita}$$

Como se mencionó anteriormente, para estimar los parámetros a_A , d_B , localia y visita se utiliza el método de máxima verosimilitud. Para esto se utilizan los datos de las casas de apuestas de todos los partidos preparatorios entre equipos participantes del Mundial desde noviembre del 2013 hasta el inicio del Mundial, más todos los partidos disputados en la copa del mundo, contabilizando 100 partidos en total.

Dado que son 32 equipos, serán $2 * 32 + 2 = 66$ los parámetros a estimar. A continuación se muestran los parámetros obtenidos.

Tabla 20. Parámetros en base a cuotas de casas de apuestas partidos Copa Mundial Brasil 2014

Localía = 0,212	Visita = -0,155	$\beta = 0,279$
País	Ataque	Defensa
Algeria	0	0
Argentina	0,514	-0,73
Australia	0	0
Belgium	0,324	-0,528

Bosnia-Herzegovina	0	0
Brazil	0,598	-0,676
Cameroon	0	0
Chile	0,404	-0,466
Colombia	0,348	-0,606
Costa Rica	0	0
Croatia	0	-0,453
Ecuador	0	-0,331
England	0,388	-0,548
France	0,509	-0,684
Germany	0,53	-0,822
Ghana	0,279	0
Greece	0	-0,2155
Honduras	0	0
Iran	-0,219	0
Italy	0,334	-0,549
IvoryCoast	0,199	-0,327
Japan	0,189	-0,297
KoreaRepublic	0	-0,258
Mexico	0	-0,403
Netherlands	0,422	-0,626
Nigeria	0	-0,195
Portugal	0,451	-0,368
Russia	0,215	-0,322
Spain	0,648	-0,782
Switzerland	0	-0,3892
UnitedStates	0	-0,346
Uruguay	0,3894	-0,496

Tal como se hizo anteriormente, se utilizaron estos parámetros para realizar 100.000 simulaciones de Monte Carlo de torneos con formato Round Robin entre los equipos de la Tabla 20.

En la Tabla 21 se muestran los resultados de las simulaciones, ordenados de mayor a menor de acuerdo a la probabilidad de salir campeón. Cabe destacar que bajo los parámetros de la Tabla 20, existen países con probabilidad nula de salir campeones, como el caso de Nigeria, Honduras y Costa Rica, entre otros. Para ordenar a estos equipos se utilizó el criterio de ordenarlos desde el último lugar por la probabilidad de terminar últimos en el campeonato.

Tabla 21. Resultados simulaciones de torneos con sistema Round Robin, en base a casas de apuestas

Orden	País	Prob. Campeón	Prob. Último
1	Spain	0,48147	0
2	Germany	0,15296	0
3	Argentina	0,11611	0
4	France	0,06809	0
5	Netherlands	0,04037	0
6	England	0,03463	0
7	Brazil	0,02897	0
8	Portugal	0,01717	0
9	Chile	0,01707	0
10	Colombia	0,01431	0
11	Italy	0,01241	0,00001
12	Uruguay	0,01078	0
13	Belgium	0,00544	0,00001
14	UnitedStates	5,00E-05	0,00185
15	Switzerland	4,00E-05	0,00242
16	Ghana	3,00E-05	0,00512
17	IvoryCoast	3,00E-05	0,00354
18	Mexico	3,00E-05	0,00345
19	Russia	2,00E-05	0,00393
20	Croatia	1,00E-05	0,00357
21	Japan	1,00E-05	0,00688
22	Ecuador	0	0,0075
23	Greece	0	0,00804
24	KoreaRepublic	0	0,0155
25	Algeria	0	0,08633
26	Australia	0	0,08673
27	Nigeria	0	0,08702
28	Honduras	0	0,08749
29	Bosnia- Herzegovina	0	0,08768
30	Cameroon	0	0,08805
31	Costa Rica	0	0,08899
32	Iran	0	0,32589

La Tabla 22 muestra el MSE de cada uno de las metodologías propuestas, usando este ranking como referencia.

Tabla 22. Resultados Metodologías propuestas, Ranking de Referencia basado en casas de apuestas

Modelo	MSE
Ranking FIFA, Octubre 2013	1101
Metodología 1: CON LOCALÍA INCLUIDA Y VARIABLE CONFEDERACIÓN ELIMINADA	1439
Metodología 2: CON PARTIDOS AMISTOSOS OMITIDOS.	1204
Metodología 3: REESTRUCTURACIÓN DE PUNTOS POR VICTORIA	1026
Metodología 4: SIN DEPRECIACIÓN DE PUNTAJE	1252

Como se puede observar en la Tabla 22, cuando se utilizan las cuotas de casas de apuestas de los partidos previos y durante el Mundial como un ranking de referencia, el menor MSE y por lo tanto la mejor metodología es la n°3.

Como se pudo apreciar, para 3 Rankings de Referencias generados de diferentes maneras, no existe una única metodología transversal mejor que las demás. Mientras el Ranking FIFA se ajustó de mejor manera a los resultados del Mundial, la eliminación de variables, la omisión de partidos amistosos, la reestructuración de puntos y la omisión de depreciación anual generan Rankings que ajustan mejor ya sea a la data histórica de partidos, como a los datos reportados por las casas de apuestas.

Tabla 23. Resumen de resultados. Comparación de Metodologías

Metodología/Ranking Referencial	Resultados Mundial 2014	Data Histórica	Casas de Apuestas
Ranking FIFA, Octubre 2013	1983	2120	1101
Metodología 1: CON LOCALÍA INCLUIDA Y VARIABLE CONFEDERACIÓN ELIMINADA	2337	2602	1439
Metodología 2: CON PARTIDOS AMISTOSOS OMITIDOS.	2638	2332	1204
Metodología 3: REESTRUCTURACIÓN DE PUNTOS POR VICTORIA	2828	2236	1026
Metodología 4: SIN DEPRECIACIÓN DE PUNTAJE	3436	1834	1252

En la Tabla 23 se presenta un resumen de los resultados de las 4 metodologías cuando se comparan bajo diferentes criterios. Si bien el Ranking FIFA no es la metodología ganadora absoluta, pareciera que en

comparación a las metodologías propuestas ajusta bastante bien independiente la Referencia que se escoja.

Un hecho a destacar es la mejora considerable en el ajuste de todas las metodologías, incluyendo la del Ranking FIFA cuando se utiliza como referencia un Ranking en base a casas de apuestas. Esto se puede observar comparando los números de la última columna de la Tabla 23, con los generados por los otros 2 Rankings Referenciales. Estos resultados demuestran que el uso de las casas de apuestas, en particular de las probabilidades que entregan para el resultado de cada partido, es una información valiosa que hasta el momento no está siendo utilizada en ninguna metodología de creación de ranking a nivel deportivo. La idea principal que aparece luego de analizar estos resultados es que se debe intentar identificar de la mejor manera posible los modelos que ocupan las casas de apuestas y las variables que se utilizan para la designación de sus cuotas. Obtener data que permita estudiar en profundidad estos modelos sería de gran utilidad para la mejora de los modelos propuestos en esta tesis.

Por el momento, una solución a considerar es generar una metodología de construcción de rankings en base a las casas de apuestas. Es decir, que los puntajes asignados a cada partido sea una función proveniente de las cuotas asignadas por éstas, ya que son las que entregan la mejor aproximación al mérito (baja probabilidad, mayor mérito, mayor puntaje), de obtener un resultado específico en un partido.

11. ASIGNACIÓN JUSTA Y MUNDIAL MÁS COMPETITIVO

Teniendo ya propuestas concretas que mejoran el sistema de generación de rating FIFA, el segundo problema en discusión es generar mundiales que sean justos y equitativos. En el último mundial, las fuerzas de cada grupo fueron muy dispares. Mientras hubo un grupo con 3 campeones mundiales (Inglaterra, Italia y Uruguay), existen grupos muy débiles tales como los encabezados por Bélgica y Suiza. Esto se debe principalmente al formato que se utiliza para decidir el cuadro principal del mundial, que se realiza principalmente con el objetivo de respetar las restricciones geográficas dentro de cada grupo.

Para el sorteo del Mundial Brasil 2014, los países se dispusieron en 4 pots distintos, el primero de ellos constituido por los 8 cabeza de serie, el segundo por los equipos sudamericanos no cabezas de serie y los países africanos, el tercero constituido por los países norteamericanos y asiáticos y el último por todos los países europeos que no fueron primer sembrados, tal como muestra la Tabla 24.

Tabla 24: Pots oficiales mundial Brasil 2014

Pot 1	Pot 2	Pot 3	Pot 4
11 Brasil (1)	12 Chile (12)	13 USA (13)	8 Holanda (9)
1 España (2)	17 Costa de Marfil (17)	24 México (23)	9 Italia (10)
2 Alemania (3)	22 Ecuador (21)	31 Costa Rica (24)	10 Inglaterra (11)
3 Argentina (4)	23 Ghana (22)	34 Honduras (27)	14 Portugal (14)
4 Colombia (5)	32 Argelia (25)	44 Japón (28)	15 Grecia (15)
5 Bélgica (6)	33 Nigeria (26)	49 Irán (29)	16 Bosnia (16)
6 Uruguay (7)	59 Camerún (32)	56 Corea (30)	18 Croacia (18)
7 Suiza (8)	Equipo de pot 4 (Italia)	57 Australia (31)	19 Rusia (19)
			21 Francia (20)

Fuente: Elaboración propia basada en información obtenida de la página: Obtenido de <http://ultra.zone/2014-FIFA-World-Cup-Group-Stage-Draws>

Dado que países dentro de un mismo pot no pueden pertenecer a un mismo grupo, se aprecia de inmediato que este sistema es injusto para algunos países. Por ejemplo Estados Unidos se ve altamente afectado con este sistema ya que comparte en el pot 3 con equipos mucho más débiles y

con los cuales estará impedido de enfrentarse en la fase de grupos. Además estará obligado a enfrentarse a un cabeza de serie y a otro equipo europeo (con un ranking entre 8 y 21) por lo cual es muy probable que el sorteo lo deje en un grupo muy difícil.

Otro país perjudicado fue Italia, el cual fue sorteado del pot 4 hacia el pot 2 quedando con países más débiles como lo son Argelia, Nigeria y Camerún. De esta manera ambos países mencionados quedaron en los grupos más fuertes, y mientras Italia finalmente quedó eliminado en primera ronda, Estados Unidos clasificó segundo en su grupo, pero quedando eliminado en Octavos de Final.

Si bien, como se mencionó anteriormente, el objetivo principal de este formato de sorteo mediante 4 pots es asegurar las restricciones geográficas dentro de cada grupo, se mostró que éste perjudica a algunos países antes que se realice el sorteo dado que equipos con mejor ranking tienen mayor probabilidad de situarse en grupos más difíciles.

Se mostrará que existen estudios actualmente que proponen mejoras al sistema actual y posteriormente se entregará una propuesta diferente en esta tesis que corrija el desbalance existente, manteniendo las restricciones geográficas.

En su análisis, Guyon [17] entrega 3 alternativas para generar de distinta manera el sorteo para el mundial de fútbol. La primera es distribuir los países en tómbolas o pots de la misma manera como lo hace la UEFA para la Champions League, en la cual se generan 4 pots donde en el primero se agrupan los 8 mejores, en el segundo los segundos 8 mejores y así sucesivamente hasta llegar al cuarto pot donde se agrupan los países enumerados del 25 al 32. Cada equipo de la primera tómbola es enviado a un grupo distinto desde el A hasta el H, quedando de esta manera como cabezas de serie. Se prosigue asignando a los países de la tómbola 2, donde un país es asignado a un grupo si cumple con las restricciones geográficas y además posibilita la formación de grupos factibles teniendo en consideración los países de los pots 3 y 4, es decir, que no genere infactibilidad geográfica dado los equipos que aún no son sorteados. Esta factibilidad la entrega en pocos segundos un algoritmo computacional. Se prosigue de la misma manera con las últimas dos tómbolas.

La segunda alternativa propuesta consiste en nuevamente generar 4 pots listados de modo tal que los primeros 8 países (los más fuertes) se encuentren en el primer pot y los últimos 8 en el cuarto pot. Lo nuevo de esta propuesta es que, teniendo asignados cada equipo del primer pot a un

grupo distinto, se asignan continentes a esos grupos provenientes de los demás pots con tal de asegurar factibilidad geográfica, y posterior a esto, se asignan países de los respectivos continentes. Dado que existen muchas maneras de asignar los continentes a los grupos se debe randomizar sobre estas posibilidades para obtener un cuadro final.

La última alternativa es añadir una restricción en donde se ordenen los equipos de acuerdo a su ranking relativo (RR) entre los 32 participantes. Nuevamente se ordenan en 4 pots de 8 equipos de acuerdo a su ranking relativo, y se determina 8 grupos totalmente balanceados de la siguiente manera:

- El equipo listado como n° 1 y el equipo listado n° 16 serán asignados a un mismo grupo.
- El equipo listado como n° 2 y el equipo listado n° 15 serán asignados a un mismo grupo.
- El equipo listado como n° 3 y el equipo listado n° 14 serán asignados a un mismo grupo, y así sucesivamente

Lo mismo ocurre para el equipo n°17 y n°32, con el n°18 y n°31 y con el resto de los países. Posteriormente estos ocho grupos que contienen a los últimos 16 equipos son agrupados con los 8 grupos de los equipos más fuertes.

Sin embargo los grupos inducidos por este sistema es probable que no cumplan con las restricciones geográficas, por lo que se debe generar una regla de distorsión mínima al balance de los grupos con tal de respetar las restricciones geográficas.

Habiendo explicado estas 3 alternativas, se procede a generar una propuesta nueva que permita arreglar las falencias de desbalance actuales respetando las restricciones del problema, sin la necesidad del uso de pots. El desarrollo de la propuesta se presenta a continuación.

Para estimar la dificultad de un grupo se define primeramente la fuerza de cada equipo como la posición que ocupan ya sea en el Ranking FIFA o en el Ranking Relativo del Mundial. Los 8 primeros clasificados serán cabezas de serie. Los demás equipos se integrarán a cada uno de los 8 grupos respetando las restricciones geográficas que impone el sistema actual, en las cuales no pueden haber más de 2 países europeos en un mismo grupo y que no haya más de uno de los demás continentes.

La función objetivo a minimizar será la diferencia de fuerza entre el grupo más débil y el grupo más fuerte, fuerza definida por la suma individual de las fuerzas de cada país integrante del grupo. Con esto se logran los grupos más balanceados posibles, evitando así, los problemas que se han generado en todos los últimos mundiales, donde la existencia de grupos de "la muerte" (fuerza baja, alta dificultad) viene acompañado de grupos muy débiles.

Tabla 25: Draw final, Mundial Brasil 2014

Grupo A	Grupo B	Grupo C	Grupo D	Grupo E	Grupo F	Grupo G	Grupo H
11 Brasil (1)	1 España (2)	4 Colombia (5)	6 Uruguay (7)	7 Suiza (8)	3 Argentina (4)	2 Alemania (3)	5 Bélgica (6)
18 Croacia (18)	8 Holanda (9)	15 Grecia (15)	31 Costa Rica (24)	22 Ecuador (21)	16 Bosnia (16)	14 Portugal (14)	32 Argelia (25)
24 México (23)	12 Chile (12)	17 Costa de Marfil (17)	10 Inglaterra (11)	21 Francia (20)	49 Irán (29)	23 Ghana (22)	19 Rusia (19)
59 Camerún (32)	57 Australia (31)	44 Japón (28)	9 Italia (10)	34 Honduras (27)	33 Nigeria (26)	13 USA (13)	56 Corea (30)

Fuente: Elaboración propia basada en la información obtenida de "Rethinking the FIFA World Cup final draw" (Guyoni, 2014)

El número que antecede al país es su Ranking FIFA respectivo en el mes de Octubre del 2013, mes en el cual se determinaron los cabezas de serie. Entre paréntesis aparece el ranking relativo entre las 32 selecciones mundialistas. Brasil por ser local obtiene el primer lugar.

Tabla 26: Suma de Ranking relativo y Ranking FIFA por grupo

	Grupo A	Grupo B	Grupo C	Grupo D	Grupo E	Grupo F	Grupo G	Grupo H
Suma de Ranking relativo ()	74	54	65	52	76	75	52	80
Suma de Ranking FIFA	112	78	80	56	84	101	52	112

Fuente: Elaboración propia basada en la información obtenida de "Rethinking the FIFA World Cup final draw" (Guyoni, 2014)

Como se aprecia en las Tabla 26 el formato de sorteo posibilitó que existiera un gran desbalance entre los diferentes grupos, obteniéndose algunos con mayor dificultad (grupo B, D y G), y otros grupos más débiles (grupo H, A, E y F). Este desbalance se intentará corregir mediante la modelación de un problema de programación lineal.

11.1. MODELO DE BALANCE DE FUERZA ENTRE GRUPOS

Para resolver el problema se plantea un modelo de programación entera mixta, con variables binarias y enteras.

A continuación se presenta el modelo detallando tanto sus variables como sus restricciones.

11.1.1. Conjuntos

Dado que la principal restricción que se impone en el sorteo es cumplir con la restricción geográfica se definen los siguientes grupos

CONMEBOL := {BRASIL, ARGENTINA, COLOMBIA, CHILE, ECUADOR, URUGUAY}

AFRICA := {CAMERÚN, NIGERIA, COSTA DE MARFIL, ARGELIA, GHANA}

CONCACAF := {USA, HONDURAS, MÉXICO, COSTA RICA}

ASIA := {JAPÓN, IRÁN, COREA, AUSTRALIA}

UEFA:= {ESPAÑA, ALEMANIA, SUIZA, BÉLGICA, ITALIA, FRANCIA, HOLANDA, INGLATERRA, BOSNIA, CROACIA, GRECIA, PORTUGAL, RUSIA}

I:= {ESPAÑA, ALEMANIA, SUIZA, BÉLGICA, ITALIA, FRANCIA, HOLANDA, INGLATERRA, BOSNIA, CROACIA, GRECIA, PORTUGAL, RUSIA, JAPÓN, IRÁN, COREA, AUSTRALIA, USA, HONDURAS, MÉXICO, COSTA RICA, CAMERÚN, NIGERIA, COSTA DE MARFIL, ARGELIA, GHANA, BRASIL, ARGENTINA, COLOMBIA, CHILE, ECUADOR, URUGUAY} {Todos}

CS.= {BRASIL, ESPAÑA, ALEMANIA, ARGENTINA, SUIZA, BÉLGICA, URUGUAY, COLOMBIA} {Cabezas de serie}

G := {GRUPO 1, GRUPO 2, GRUPO 3, GRUPO 4, GRUPO 5, GRUPO 6, GRUPO 7, GRUPO 8}

11.1.2. Datos

P_i: Ranking FIFA del equipo i al momento de elegir los cabezas de serie

R_i : Ranking Relativo del equipo i al momento de elegir los cabezas de serie

11.1.3. Variables de Decisión

$$X_{ig} \begin{cases} 1 & \text{si equipo } i \text{ es asignado a grupo } g \\ 0 & \text{si no} \end{cases}$$

Y_g : Fuerza de Ranking FIFA del grupo g

Z_g : Fuerza de Ranking Relativo del grupo g

11.1.4. Restricciones

- Todos los grupos deben estar conformados por 4 equipos

$$\sum_{i \in I} X_{ig} = 4 \quad \forall g \in G$$

- Sólo un cabeza de serie por grupo

$$\sum_{i \in CS} X_{ig} = 1 \quad \forall g \in G$$

- Todos los países deben ser asignados a un grupo

$$\sum_{g \in G} X_{ig} = 1 \quad \forall i \in I$$

- Definición de Fuerza de Ranking FIFA de cada grupo

$$\sum_{i \in I} P_i X_{ig} = Y_g \quad \forall g \in G$$

- Definición de Fuerza Relativa de cada grupo

$$\sum_{i \in I} R_i X_{ig} = Z_g \quad \forall g \in G$$

- Definición de Fuerza de Ranking FIFA de grupo mínima y máxima

$$w_{min} \leq Y_g$$

$$w_{max} \geq Y_g$$

- Definición de Fuerza de Ranking Relativo de grupo mínima y máxima

$$v_{min} \leq Z_g$$

$$v_{max} \geq Z_g$$

- Respetar condiciones geográficas.

$$\sum_{i \in UEFA} X_{ig} \leq 2 \quad \forall g \in G$$

$$\sum_{i \in CONMEBOL} X_{ig} \leq 1 \quad \forall g \in G$$

$$\sum_{i \in ASIA} X_{ig} \leq 1 \quad \forall g \in G$$

$$\sum_{i \in AFRICA} X_{ig} \leq 1 \quad \forall g \in G$$

$$\sum_{i \in CONCACAF} X_{ig} \leq 1 \quad \forall g \in G$$

- Naturaleza de las variables

$$X_{ig} \in \{0,1\} Y_g, Z_g \geq 0 \quad w_{min}, w_{max}, v_{min}, v_{max} \geq 0$$

11.1.5. Función Objetivo

2 opciones:

Minimizar Diferencia de Fuerza de ranking FIFA entre los grupos.

$$\min w_{max} - w_{min} \quad (1)$$

o Minimizar Diferencia de Fuerza Relativa entre los grupos.

$$\min v_{max} - v_{min} \quad (2)$$

11.2. RESULTADOS DE OPTIMIZACIÓN LINEAL MIXTA

A continuación se presentan los grupos que cumplen todas las restricciones del modelo anteriormente descrito y que generan balance, y por lo tanto, grupos mucho más equilibrados de acuerdo al ranking obtenido. Los

resultados se obtuvieron utilizando el solver CPLEX del lenguaje AMPL, en menos de 5 segundos.

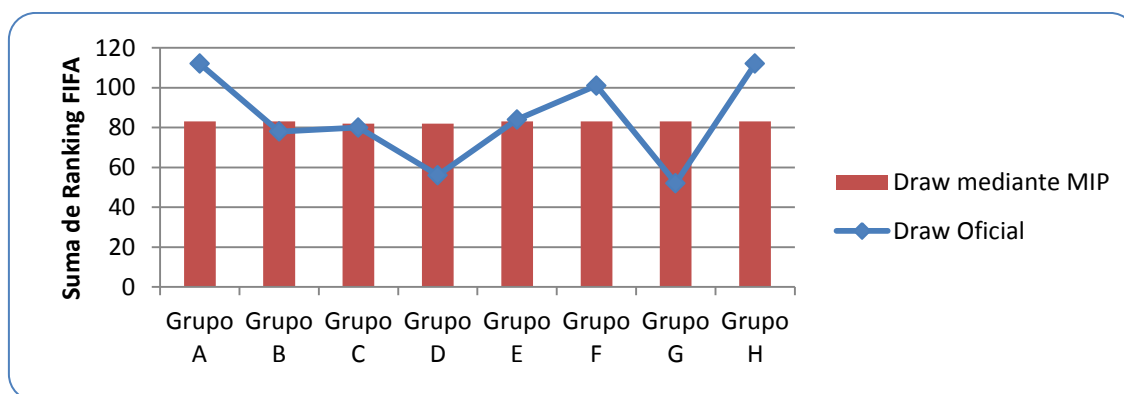
El primer resultado se obtuvo optimizando 1, es decir, balanceando los grupos de acuerdo al Ranking FIFA con el cual se presentaron al mundial. Los resultados se muestran en la Tabla 27.

Tabla 27: Resultado Óptimo modelo de optimización con F.O. 1, Opción 1

Grupo1	Grupo2	Grupo3	Grupo4	Grupo5	Grupo6	Grupo7	Grupo8
Alemania	España	Suiza	Uruguay	Bélgica	Brasil	Colombia	Argentina
Chile	Irán	Ecuador	Honduras	Australia	México	Corea	Japón
Camerún	Costa Marfil	Argelia	Ghana	Usa	Nigeria	Italia	Costa Rica
Inglaterra	Bosnia	Francia	Rusia	Holanda	Grecia	Portugal	Croacia
Fuerza	83	83	82	82	83	83	83

Se aprecia que para la creación de grupos más equitativos no es necesario el uso de pots, sino que basta con un modelo de programación lineal mixta que cumpla con las restricciones geográficas impuestas por la FIFA. Como se puede ver en la Tabla 27, el mundial más equitativo en términos de Ranking FIFA agrupa a equipos como Uruguay y España con países mucho más accesibles en primera ronda. La diferencia entre el draw oficial y la solución se puede ver más clara en la ilustración N°6.

Ilustración 6: : Comparación draw oficial vs draw optimizado (1)



Es posible ver una homogeneidad entre las fuerzas de cada grupo, ya que la función objetivo (diferencia entre el grupo más fuerte y más débil) tomó el valor de 1.

El hecho que el sorteo del draw oficial sea tan emocionante a nivel mundial se debe principalmente a que contiene una componente importante de aleatoriedad, lo cual podría ser un obstáculo para implementar este sistema. Sin embargo, con el fin de mantener un sorteo aleatorio, se pueden obtener las 5 ó 10 mejores soluciones y finalmente randomizar sobre ellas. De igual manera, dado que la composición de los grupos es la solución importante, el orden inter e intra grupo también puede ser obtenido de manera aleatoria, es decir, el grupo de Argentina no necesariamente debe ser el grupo 8, y dentro del mismo Japón no necesariamente debe ser el segundo equipo. Las 3 mejores soluciones restantes se encuentran en sección anexos.

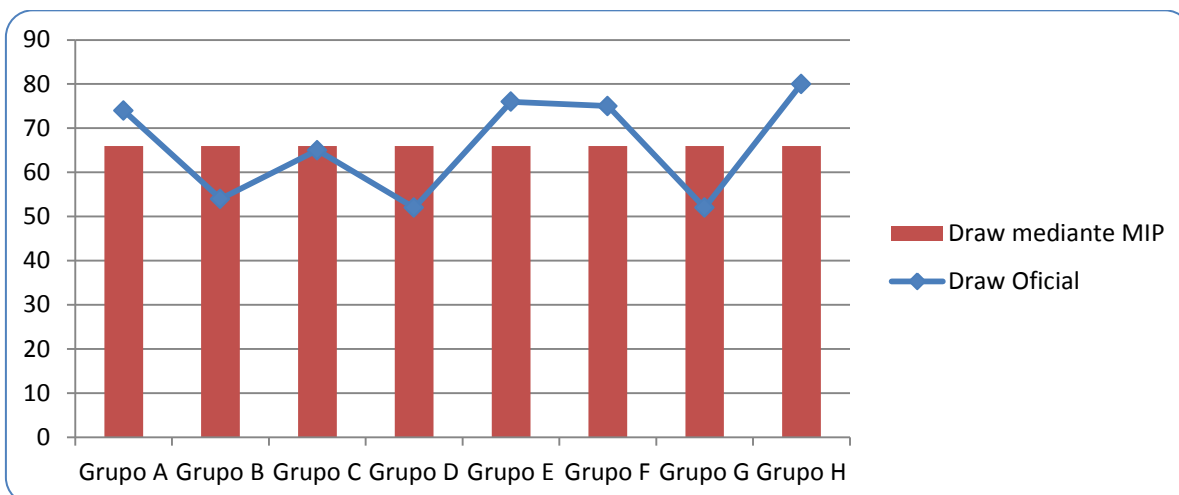
El segundo resultado se obtuvo optimizando la función objetivo (2), es decir, balanceando los grupos de acuerdo al Ranking Relativo con el cual se presentaron al mundial. Los resultados se muestran en la Tabla 28.

Tabla 28: Resultado modelo de optimización con F.O. 2, Opción 1

Grupo1	Grupo2	Grupo3	Grupo 4	Grupo5	Grupo6	Grupo7	Grupo8
Brasil	Uruguay	Argentina	Suiza	España	Colombia	Bélgica	Alemania
Camerún	Corea	Australia	Chile	Irán	Costa Rica	Ecuador	Rusia
Italia	Francia	Bosnia	Japón	USA	Nigeria	Argelia	Honduras
México	Holanda	Grecia	Croacia	Ghana	Inglaterra	Portugal	Costa Marfil
Fuerza	66	66	66	66	66	66	66

Al igual que para la solución anterior, no es necesario el uso de pots para obtener un draw mucho más equitativo. De hecho esta solución entrega grupos totalmente balanceados, ya que la función objetivo toma un valor de 0. La diferencia entre el draw oficial y esta solución se puede ver más clara en la ilustración N°7.

Ilustración 7: Comparación draw final vs draw optimizado (2)



Una conclusión importante que se puede despejar de la ilustración N°6 es cuales equipos fueron perjudicados por el actual sistema y cuales fueron beneficiados. Todos los grupos en los cuales los puntos se encuentran por debajo del gráfico de barra fueron altamente perjudicados, esto sucedió con los grupos B, D y G. Por otro lado los grupos A, E, F y H fueron claramente beneficiados.

12. CONCLUSIONES Y RECOMENDACIONES

Para comenzar con el análisis de distintas metodologías de creación de ratings se debía entender cómo operaba actualmente la FIFA a través de su Ranking, y conocer todo el proceso que involucra este sistema. A partir de los datos y del conocimiento adquirido se pudieron vislumbrar ciertas falencias que poseía su metodología, las cuales radicaban principalmente en la omisión de variables claves en el desarrollo de un partido, así como también falencias estructurales.

La primera de las falencias reside en que el puntaje entregado en un partido no refleja de buena manera cuán meritorio fue para cada equipo obtenerlo. Si bien la metodología de FIFA entrega mayor puntaje por jugar contra equipos con mejor ranking, no hace referencia alguna a otras variables tan relevantes, como lo es por ejemplo el factor localía. En este mismo ámbito FIFA incluye en su metodología una componente continental para otorgar mayor o menor puntaje a un encuentro. Esta variable también fue motivo de estudio.

Por otro lado, la construcción del Rating FIFA tiene graves falencias estructurales. La principal de ellas es que permite a países sacar ventaja de una buena posición evitando jugar partidos amistosos en el año previo a la designación de los cabezas de series de la Copa del Mundo. Esto debido principalmente a la escasa entrega de puntos de este tipo de partidos en comparación con cualquier otro tipo de formato, combinado con la fórmula de promediar los puntajes de los juegos disputados.

En esta tesis se muestra con datos que esta fue la principal razón por la cual países como Suiza y Colombia fueron parte de los 8 países sembrados del Mundial disputado en Brasil el año 2014. Otro problema estructural se da por la depreciación del promedio de puntos de los últimos 4 años. Este sistema genera disparidad entre países de distintos continentes debido a que no todos juegan su torneo continental en el mismo año, por ende, ante igualdad de resultados en los demás torneos, es probable que los equipos de confederaciones que jueguen sus torneos más cerca del próximo mundial se vean beneficiados.

Todos los problemas o falencias explicados anteriormente dejan entrever que el sistema actual de ordenamiento de las selecciones adultas de

fútbol tiene mucho potencial de mejora. Esto, acompañado de la importancia que tiene pertenecer al selecto grupo de los 8 primeros antes de los mundiales hace imperiosa la generación de propuestas que permitan mitigar los problemas que se han presentado en la actualidad.

Como primer paso para la indagación de variables relevantes se realizó una caracterización general de los partidos contenidos en la base de datos. Los resultados mostraron una tendencia clara a mejores performances de los equipos cuando juegan en su recinto, ya que de la totalidad de encuentros más del 50% terminaron con triunfo del equipo local. Esto fue tomado como una primera señal para indagar sobre la relevancia de esta variable con modelos estadísticos.

Avanzando hacia la generación de propuestas concretas se generaron 4 modelos estadísticos con el fin de explicar de la mejor manera posible los resultados de los partidos de la data. Estos modelos tienen el enfoque Logit Multinomial, el cual posibilita, mediante el uso de variables explicativas, la generación de distribuciones de probabilidades para la ocurrencia de un evento multiclases, como lo es el resultado de un juego de fútbol. Cada uno de los modelos fue realizado con distintas variables explicativas, comenzando con la más sencilla que explica un partido solamente por la diferencia de rankings entre los equipos contendores y agregando variables como localía Nacional, localía Continental, y el efecto de pertenecer a una Confederación.

Los resultados muestran lo fundamental que es la variable localía, ya que entrega una gran ventaja al equipo que posea dicho atributo. Este factor acompañado con la diferencia de Ranking entre equipos fueron los únicos 2 que aportan a ajustar de mejor manera la data. Al mismo tiempo se desechó la idea de que la pertenencia a una determinada confederación aporte a determinar el resultado de un partido. Esta misma conclusión se tuvo para la variable localía continental.

Otro resultado importante obtenido a partir del modelo basado en Diferencia de Ranking y Localía es que se pudo determinar una relación de equivalencia entre ambas variables. Mediante el modelo de predicción N°2 se obtuvo que el efecto de jugar de local es comparable con tener una diferencia de ranking de 30 puestos a favor. Esta ventaja no menor que genera la localía es un indicio claro que cualquier metodología de creación de rating que se considere robusta debe incluirla.

Otra manera de cuantificar la variable localía se obtuvo analizándolas cuotas de las casas de apuestas para un subconjunto de los partidos de la base de datos. A partir de una transformación de las cuotas de las casas de apuestas, se pudo obtener la probabilidad implícita de cada resultado posible de todos los partidos seleccionados. Se consideraron los partidos internacionales disputados en los años 2009 y 2010 haciendo una separación entre los disputados en territorio neutral y los que no. Esta separación permitió obtener información valiosa que se comenta a continuación.

El buen ajuste de las regresiones lineales hace pensar que el ranking de los equipos sí es un buen indicador para asignar probabilidades de éxito en un partido. Además para los equipos actuando de local las casas de apuesta conceden cerca de un 18% más de probabilidad de ganar para éstos en comparación con el equipo visitante.

Otro resultado importante fue corroborar que la probabilidad de empate queda muy bien determinada a partir de la diferencia en probabilidad de triunfo entre los 2 equipos, lo cual muestra que las casas de apuestas también utilizan este resultado como una variable de referencia.

El último resultado, y quizás el más relevante fue la equivalencia entre los factores localía y diferencia en ranking observado para las casas de apuestas. Al igual que con el Logit Multinomial, la regresión lineal muestra que la localía equivale a aproximadamente una diferencia de 30 puestos en el ranking.

Por otra parte, como resultado de las metodologías propuestas se concluye lo siguiente:

Ningún sistema de rating puede captar de manera perfecta qué equipo es relativamente mejor que cualquier otro, especialmente cuando se refiere a fútbol, en el cual el resultado de un partido puede estar influenciado por variables difíciles de controlar. Sin embargo, dado que el ordenamiento de los equipos es clave en los resultados de la Copa Mundial, se hace necesario estudiar el sistema actual y generar propuestas que permitan mejorar.

Se generaron 4 propuestas de creación de rating, incluyendo la variable localía, eliminando la variable confederación, omitiendo partidos amistosos, reestructurando los puntos por partido y cambiando el sistema de depreciación por año. Un punto importante a abordar fue la comparación de las distintas metodologías propuestas. Para discriminar cual era mejor que otro se utilizó la métrica "mean squared error" (MSE) a partir de Rankings de Referencias. Estos Rankings de Referencia se crearon de 3 maneras diferentes:

- A partir de los resultados observados en el Mundial Brasil 2014
- A partir de simulaciones de torneos, con parámetros obtenidos de la data de los últimos 4 años antes previos al Mundial.
- A partir de simulaciones de torneos, con parámetros obtenidos de las cuotas asignadas por casas de apuestas para partidos preparatorios y del Mundial.

Los resultados demuestran que no existe una única metodología transversal que supere a las demás, dado el ranking que generan. Por un lado el ranking FIFA actual si bien tiene claras falencias, que fueran evidenciadas en esta tesis, no está del todo alejado de los Rankings Referenciales obtenidos, y funciona como un buen input en la predicción de los resultados.

Lo que queda muy claro a partir de los resultados obtenidos, y que va en línea con mejorar el sistema actual, es la necesidad de acercarse a los modelos utilizados por las casas de apuestas para la generación de sus pronósticos. Se evidenció que todos los modelos propuestos, incluso el ranking actual, ajustan de mejor manera, y que esta mejora es significativa, cuando se utiliza este tipo de información para generar un Ranking Referencial.

En cuanto al tema de mundiales con asignación a grupos equilibrados, la problemática surgió de revisar los sorteos de las últimas ediciones mundialistas, en los cuales la disparidad de fuerzas entre grupos era siempre un tema de discusión.

Modelando los datos como un problema de programación lineal mixta se obtienen de manera sencilla grupos mucho más equilibrados en cuanto a

fuerza, ya sea esta fuerza proveniente del ranking FIFA o del ranking relativo de los equipos en el mundial. El modelamiento como programación lineal mixta entregó resultados que cumplieron con todas las restricciones impuestas por la FIFA, sin la necesidad de uso de pots y con un tiempo de proceso muy bajo.

Ya que la incertidumbre del sorteo mundialista genera una componente alta de atractivo, se propone generar las 5 ó 10 mejores soluciones posibles y generar una metodología para randomizar el draw final entre ellas.

Si bien los resultados del problema abordado se obtuvieron utilizando el Ranking FIFA actual, es de esperar que aplicando mejoras a éste, algunas planteadas en esta tesis, se obtenga la mejor solución posible, de manera tal, que en las Copas del Mundo venideras se observen campeonatos muchos más justos, con grupos cada vez más equilibrados.

Por último, se declaran cumplidos los objetivos de esta tesis, ya que se plantearon metodologías que superan, bajo ciertos criterios, los resultados obtenidos hasta hoy. Por otra parte se encontraron maneras de comparar los resultados obtenidos. Por su parte el problema de Mundiales más justos fue abordado de manera exitosa, se lograron obtener resultados que demuestran que existen soluciones factibles que mejoran el sistema actual, cumpliendo con todas las restricciones necesarias y en muy poco tiempo.

13. Bibliografía

- D. Didonato, «ESPN Deportes,» 5 Diciembre 2013. [En línea]. Available:
1] http://espndeportes.espn.go.com/blogs/index?nombre=o_blog_2014&entryID=1973717.
[Último acceso: 18 Agosto 2015].
- Fédération internationale de Football association, «FIFA,» 1994. [En línea].
2] Available: <http://es.fifa.com/about-fifa/>. [Último acceso: 2 Septiembre 2015].
- F. I. d. F. Association, «FIFA.com,» 2014. [En línea]. Available: <http://es.fifa.com/fifa-world-ranking/procedure/men.html>. [Último acceso: 4 Septiembre 2015].
- Downward y Jones, «effects of crowd size on referee decisions: analysis of then FA
4] Cup,» 2007.
- R. Pollard, «Home advantage in soccer: A retrospective analysis,» *Journal of Sports
5] Sciences*, 1986.
- J. Hoppenbrouwer y M. Winkels, «Sports Ratings,» Universidad de Amsterdam,
6] Amsterdam, 2013.
- A. Langville y C. Meyer, *Who's #1? The Science of Rating and Ranking*, 2013.
7]
- K. Massey, «Statistical Models Applied to the Ratings of Sport Teams,» 1997.
8]
- R. T. Stefani, «Football and Basketball Predictions Using Least Squares,» 1997.
9]
- J. Del Corrala y J. Prieto Rodríguez, «Are differences in ranks good predictors for
10] Grand Slam tennis matches?,» *International Journal of Forecasting*, 26 Julio 2010.
- S. Akhtar, P. Scarf y Z. Rasool, «Rating players in test match cricket,» University of

11] Malakand, Pakistan; and 2University of Salford, UK, Malakand;Salford.

Betexplorer.com, «BetExplorer.com,» [En línea]. Available:
12] <http://www.betexplorer.com/soccer/>. [Último acceso: 19 Noviembre 2015].

H. Nyberg, «A Multinomial Logit-based Statistical Test of Association Football
13] Betting Market Efficiency,» Junio, Helsinki, 2014.

P. Scarf y M. Yusof, «A numerical study of tournament structure and seeding policy
14] for the soccer World Cup Finals,» *Statistica Neerlandica*, pp. 43 - 57, 2011.

N. Rudi y B. S. Uppari, «In-play Football (Soccer) Prediction,» 2015.
15]

14. ANEXOS

14.1. ANEXO 1: Países por Confederación

Tabla 29: Países por Confederación

N°	AFC	CAF	CONCACAF	CONMEBOL	OFC	UEFA
1	Afghanistan	Algeria	Anguilla	Argentina	American Samoa	Albania
2	Australia	Angola	Antigua and Barbuda	Bolivia	Cook Islands	Andorra
3	Bahrain	Benin	Aruba	Brazil	Fiji	Armenia
4	Bangladesh	Botswana	Bahamas	Chile	New Caledonia	Austria
5	Bhutan	Burkina Faso	Barbados	Colombia	New Zealand	Azerbaijan
6	Brunei	Burundi	Belize	Ecuador	Papua New Guinea	Belarus
7	Burma	Cameroon	Bermuda	Paraguay	Samoa	Belgium
8	Cambodia	Cape Verde	British Virgin Islands	Peru	Solomon Islands	Bosnia-Herzegovina
9	China PR	Central African Republic	Canada	Uruguay	Tahiti	Bulgaria
10	China Taiwan	Chad	Cayman Islands	Venezuela	Tonga	Croatia
11	East Timor	Comoros	Costa Rica		Vanuatu	Cyprus
12	Guam	Congo	Cuba			Czech Republic
13	Hong Kong	Congo DR	Curacao			Denmark
14	India	Djibouti	Dominica			England
15	Indonesia	Egypt	Dominican Republic			Estonia
16	Iran	Equatorial Guinea	El Salvador			Faro Islands
17	Iraq	Eritrea	Grenada			Finland
18	Japan	Ethiopia	Guatemala			France
19	Jordan	Gabon	Guyana			Georgia
20	Korea DPR	Gambia	Haiti			Germany
21	Korea Republic	Ghana	Honduras			Greece
22	Kuwait	Guinea	Jamaica			Hungary
23	Kyrgyzstan	Guinea-Bissau	Mexico			Iceland
24	Laos	Ivory Coast	Montserrat			Ireland Republic
25	Lebanon	Kenya	Nicaragua			Israel
26	Macau	Lesotho	Panama			Italy
27	Malaysia	Liberia	Puerto Rico			Kazakhstan

28	Maldives	Libya	St Kitts and Nevis	Latvia
29	Mongolia	Madagascar	St Lucia	Liechtenstein
30	Nepal	Malawi	St Vincent and the Grenadines	Lithuania
31	Oman	Mali	Suriname	Luxembourg
32	Pakistan	Mauritania	Trinidad and Tobago	Macedonia
33	Palestine	Mauritius	Turks and Caicos Islands	Malta
34	Philippines	Morocco	United States	Moldova
35	Qatar	Mozambique	Virgin Islands	Montenegro
36	Saudi Arabia	Namibia		Netherlands
37	Singapore	Niger		Northern Ireland
38	Sri Lanka	Nigeria		Norway
39	Syria	Rwanda		Poland
40	Tajikistan	Sao Tomé and Príncipe		Portugal
41	Thailand	Senegal		Romania
42	Turkmenistan	Seychelles		Russia
43	United Arab Emirates	Sierra Leone		San Marino
44	Uzbekistan	Somalia		Scotland
45	Vietnam	South Africa		Serbia
46	Yemen	South Sudan		Slovakia
47		Sudan		Slovenia
48		Swaziland		Spain
49		Tanzania		Sweden
50		Togo		Switzerland
51		Tunisia		Turkey
52		Uganda		Ukraine
53		Zambia		Wales
54		Zimbabwe		

14.2. ANEXO 2: Resultados de Ranking según Metodología

Tabla 30: Resultados Metodología de Rating 1

1	Spain	51	Cameroon	101	Azerbaijan	151	Nicaragua
2	Argentina	52	Senegal	102	Central African Republic	152	Hong Kong
3	Germany	53	Turkey	103	Zimbabwe	153	St Lucia
4	Uruguay	54	Paraguay	104	Tajikistan	154	Thailand
5	Colombia	55	Zambia	105	Niger	155	Singapore
6	Ivory Coast	56	South Africa	106	Guatemala	156	Sri Lanka
7	Italy	57	Cuba	107	Liberia	157	Bangladesh
8	United States	58	United Arab Emirates	108	Guyana	158	Pakistan
9	Brazil	59	Guinea	109	Canada	159	Kyrgyzstan
10	Belgium	60	Togo	110	Turkmenistan	160	Bermuda
11	Netherlands	61	New Zealand	111	Estonia	161	Sao Tomé and Príncipe
12	Ghana	62	Morocco	112	Malawi	162	Puerto Rico
13	Chile	63	Jordan	113	Moldova	163	Liechtenstein
14	Mexico	64	Hungary	114	New Caledonia	164	Indonesia
15	Switzerland	65	Wales	115	Georgia	165	Dominica
16	England	66	Norway	116	Bahrain	166	China Taiwan
17	Portugal	67	Iceland	117	St Vincent and the Grenadines	167	Burma
18	Nigeria	68	Jamaica	118	Qatar	168	Guam
19	Greece	69	Montenegro	119	Kenya	169	Malaysia
20	Algeria	70	Albania	120	Burundi	170	Nepal
21	Mali	71	Slovakia	121	Rwanda	171	Solomon Islands
22	Honduras	72	El Salvador	122	Mozambique	172	Yemen
23	Costa Rica	73	Austria	123	Antigua and Barbuda	173	Laos
24	Ecuador	74	Gabon	124	Lebanon	174	Curaçao
25	Bosnia-Herzegovina	75	Finland	125	Afghanistan	175	Barbados
26	Croatia	76	Trinidad and Tobago	126	Tanzania	176	Mauritius
27	Panama	77	Congo DR	127	Namibia	177	Vanuatu
28	France	78	Bolivia	128	Equatorial Guinea	178	Swaziland
29	Denmark	79	Dominican Republic	129	Sudan	179	Tonga
30	Ukraine	80	Sierra Leone	130	Syria	180	American Samoa
31	Russia	81	Ireland Republic	131	Suriname	181	Aruba
32	Japan	82	Congo	132	Grenada	182	Faro Islands

33	Sweden	83	Uganda	133	Lithuania	183	Guinea-Bissau
34	Czech Republic	84	Haiti	134	Tahiti	184	Fiji
35	Tunisia	85	Bulgaria	135	Latvia	185	Madagascar
36	Burkina Faso	86	Israel	136	Palestine	186	Mongolia
37	Romania	87	Poland	137	Lesotho	187	Montserrat
38	Peru	88	Benin	138	Belize	188	Bahamas
39	Serbia	89	Iraq	139	Philippines	189	Brunei
40	Venezuela	90	Oman	140	Gambia	190	East Timor
41	Egypt	91	Angola	141	Chad	191	Samoa
42	Iran	92	Korea DPR	142	Mauritania	192	Comoros
43	Australia	93	Belarus	143	Luxembourg	193	Virgin Islands
44	Libya	94	Ethiopia	144	Malta	194	Papua New Guinea
45	Scotland	95	China PR	145	Cyprus	195	Cayman Islands
46	Uzbekistan	96	Saudi Arabia	146	India	196	Eritrea
47	Korea Republic	97	Botswana	147	St Kitts and Nevis	197	Somalia
48	Slovenia	98	Macedonia	148	Kazakhstan	198	Cambodia
49	Armenia	99	Northern Ireland	149	Maldives	199	Andorra
50	Cape Verde	100	Kuwait	150	Vietnam	200	Djibouti

Tabla 31: Resultados Metodología de Rating 2

1	Brazil	51	Iraq	101	Guatemala	151	Mauritania
2	Spain	52	Algeria	102	Sudan	152	Kazakhstan
3	Argentina	53	Turkey	103	Kenya	153	Kyrgyzstan
4	Germany	54	Serbia	104	Zimbabwe	154	Liechtenstein
5	UnitedStates	55	South Africa	105	Mozambique	155	Cyprus
6	Chile	56	China PR	106	Palestine	156	Nepal
7	Netherlands	57	Slovenia	107	Tanzania	157	StKitts and Nevis
8	Uruguay	58	Hungary	108	Philippines	158	St Lucia
9	Italy	59	Armenia	109	Qatar	159	Nicaragua
10	IvoryCoast	60	Morocco	110	Liberia	160	Puerto Rico
11	Colombia	61	Scotland	111	Lesotho	161	Indonesia
12	Belgium	62	Saudi Arabia	112	Sierra Leone	162	Bangladesh
13	Ghana	63	Norway	113	Tajikistan	163	Yemen
14	Japan	64	Bahrain	114	Canada	164	Burma
15	Mexico	65	Libya	115	Moldova	165	Guam
16	France	66	UnitedArabEmirate s	116	NorthernIreland	166	China Taiwan
17	Switzerland	67	Slovakia	117	Azerbaijan	167	Dominica
18	Portugal	68	Jamaica	118	Benin	168	Sao Tom? and Prncipe
19	England	69	Togo	119	DominicanRepublic	169	Bermuda
20	Greece	70	Guinea	120	Macedonia	170	Maldives
21	Denmark	71	Finland	121	Namibia	171	Laos
22	Bosnia- Herzegovina	72	Bulgaria	122	Georgia	172	Sri Lanka
23	Iran	73	Angola	123	Guyana	173	Cura?ao
24	KoreaRepublic	74	Iceland	124	Niger	174	Barbados
25	Nigeria	75	Congo DR	125	Rwanda	175	Vanuatu
26	Zambia	76	Cape Verde	126	St Vincent and the Grenadines	176	Faroelands
27	Cameroon	77	Israel	127	Latvia	177	Solomon Islands
28	Ecuador	78	Kuwait	128	Turkmenistan	178	Swaziland
29	Costa Rica	79	Trinidad and Tobago	129	Burundi	179	Tonga
30	Egypt	80	New Zealand	130	Afghanistan	180	American Samoa
31	Tunisia	81	Haiti	131	Central AfricanRepublic	181	Fiji
32	Burkina Faso	82	Belarus	132	Singapore	182	Mauritius
33	Ukraine	83	Montenegro	133	Lithuania	183	Guinea-Bissau
34	Uzbekistan	84	El Salvador	134	Antigua and Barbuda	184	Montserrat
35	Croatia	85	Botswana	135	Suriname	185	Aruba

36	Venezuela	86	Ethiopia	136	Pakistan	186	Mongolia
37	Honduras	87	Syria	137	New Caledonia	187	Bahamas
38	Australia	88	IrelandRepublic	138	India	188	Madagascar
39	Peru	89	Lebanon	139	Belize	189	Samoa
40	Sweden	90	Cuba	140	Thailand	190	VirginIslands
41	Russia	91	Bolivia	141	Gambia	191	Eritrea
42	Romania	92	Malawi	142	Vietnam	192	Comoros
43	Mali	93	Congo	143	Chad	193	CaymanIslands
44	Uganda	94	Poland	144	Grenada	194	Andorra
45	Jordan	95	Korea DPR	145	Tahiti	195	Somalia
46	CzechRepublic	96	Estonia	146	Equatorial Guinea	196	Papua New Guinea
47	Oman	97	Austria	147	Hong Kong	197	Cambodia
48	Panama	98	Wales	148	Luxembourg	198	Brunei
49	Paraguay	99	Albania	149	Malta	199	East Timor
50	Senegal	100	Gabon	150	Malaysia	200	Cook Islands

Tabla 32: Resultados Metodología de Rating 3

1	Brazil	51	Algeria	101	Wales	151	Mauritania
2	Spain	52	Iraq	102	Mozambique	152	Tahiti
3	Argentina	53	Morocco	103	Austria	153	Kazakhstan
4	Germany	54	South Africa	104	Gabon	154	Nepal
5	Chile	55	Turkey	105	Tanzania	155	Nicaragua
6	Italy	56	Serbia	106	Sierra Leone	156	Indonesia
7	Uruguay	57	China PR	107	Lesotho	157	Malta
8	UnitedStates	58	Armenia	108	Zimbabwe	158	Kyrgyzstan
9	Netherlands	59	Bahrain	109	Palestine	159	StKitts and Nevis
10	Colombia	60	Slovenia	110	Qatar	160	St Lucia
11	IvoryCoast	61	Hungary	111	Canada	161	Bangladesh
12	Belgium	62	Norway	112	Azerbaijan	162	Yemen
13	England	63	Scotland	113	Liberia	163	Puerto Rico
14	Ghana	64	Libya	114	Philippines	164	Burma
15	France	65	Guinea	115	Benin	165	China Taiwan
16	Japan	66	Slovakia	116	Namibia	166	Guam
17	Switzerland	67	Bulgaria	117	NorthernIreland	167	Sao Tom? and Prncipe
18	Portugal	68	Jamaica	118	Moldova	168	Laos
19	Mexico	69	Congo DR	119	Georgia	169	Maldives
20	Denmark	70	Togo	120	Tajikistan	170	Dominica
21	Zambia	71	Finland	121	Macedonia	171	Cura?ao
22	Greece	72	Angola	122	Rwanda	172	Bermuda
23	Cameroon	73	Saudi Arabia	123	Latvia	173	Barbados
24	Nigeria	74	Bolivia	124	DominicanRepublic	174	Sri Lanka
25	Iran	75	Israel	125	Niger	175	FaroelIslands
26	KoreaRepubli c	76	Kuwait	126	St Vincent and the Grenadines	176	Solomon Islands
27	Bosnia- Herzegovina	77	UnitedArabEmirate s	127	Lithuania	177	Vanuatu
28	Ecuador	78	Iceland	128	Burundi	178	Swaziland
29	Tunisia	79	Montenegro	129	Guyana	179	Tonga
30	Venezuela	80	Cape Verde	130	Singapore	180	American Samoa
31	Costa Rica	81	Malawi	131	Turkmenistan	181	Fiji
32	Ukraine	82	Lebanon	132	Afghanistan	182	Mauritius
33	Australia	83	IrelandRepublic	133	Central AfricanRepublic	183	Guinea-Bissau
34	Peru	84	Poland	134	Antigua and Barbuda	184	Mongolia
35	Uzbekistan	85	El Salvador	135	Suriname	185	Aruba
36	Burkina Faso	86	Belarus	136	Luxembourg	186	Madagascar
37	Croatia	87	New Zealand	137	Pakistan	187	Montserrat
38	Honduras	88	Ethiopia	138	Belize	188	Samoa

39	Egypt	89	Syria	139	India	189	Bahamas
40	Paraguay	90	Botswana	140	Hong Kong	190	Eritrea
41	Oman	91	Trinidad and Tobago	141	Equatorial Guinea	191	VirginIslands
42	Russia	92	Kenya	142	Gambia	192	Andorra
43	Sweden	93	Haiti	143	New Caledonia	193	Comoros
44	Jordan	94	Congo	144	Chad	194	Somalia
45	Uganda	95	Estonia	145	Cyprus	195	Papua New Guinea
46	CzechRepubli c	96	Guatemala	146	Thailand	196	CaymanIslands
47	Senegal	97	Korea DPR	147	Malaysia	197	Cook Islands
48	Panama	98	Albania	148	Vietnam	198	Cambodia
49	Romania	99	Sudan	149	Grenada	199	Brunei
50	Mali	100	Cuba	150	Liechtenstein	200	East Timor

Tabla 33: Resultados Metodología de Rating 4

1	Brazil	51	Algeria	101	Trinidad and Tobago	151	Nepal
2	Spain	52	Slovakia	102	Wales	152	India
3	Argentina	53	Hungary	103	Qatar	153	Puerto Rico
4	Germany	54	Ukraine	104	Macedonia	154	Hong Kong
5	Uruguay	55	China PR	105	Benin	155	Tahiti
6	Netherlands	56	Turkey	106	Canada	156	Indonesia
7	Chile	57	Senegal	107	Congo	157	Equatorial Guinea
8	IvoryCoast	58	Panama	108	Austria	158	Pakistan
9	Italy	59	Burkina Faso	109	Azerbaijan	159	Bermuda
10	Japan	60	Montenegro	110	Zimbabwe	160	Barbados
11	UnitedStates	61	Kuwait	111	Tanzania	161	Cura?ao
12	England	62	Slovenia	112	Central AfricanRepublic	162	Malta
13	Colombia	63	Jamaica	113	Tajikistan	163	Bangladesh
14	KoreaRepubli c	64	Belarus	114	NorthernIreland	164	Nicaragua
15	Greece	65	Guinea	115	Haiti	165	Faroelands
16	Mexico	66	IrelandRepublic	116	Rwanda	166	Burma
17	Paraguay	67	Libya	117	Moldova	167	Sao Tom? and Prncipe
18	Switzerland	68	Botswana	118	Palestine	168	St Lucia
19	Ghana	69	Angola	119	Niger	169	Vanuatu
20	France	70	Morocco	120	Cuba	170	China Taiwan
21	Portugal	71	Malawi	121	Lithuania	171	Laos
22	Zambia	72	New Zealand	122	Latvia	172	Kyrgyzstan
23	Denmark	73	Syria	123	Namibia	173	Mauritania
24	Iran	74	Armenia	124	Philippines	174	Guinea-Bissau
25	Venezuela	75	Bulgaria	125	Guyana	175	Maldives
26	Russia	76	Estonia	126	Lesotho	176	Guam
27	Croatia	77	Scotland	127	Liberia	177	Solomon Islands
28	Sweden	78	Poland	128	Burundi	178	Fiji
29	Belgium	79	Bahrain	129	DominicanRepubli c	179	Swaziland
30	Bosnia- Herzegovina	80	Korea DPR	130	Thailand	180	Madagascar
31	Ecuador	81	Gabon	131	Chad	181	Tonga
32	Honduras	82	Finland	132	St Vincent and the Grenadines	182	American Samoa
33	Australia	83	Sudan	133	Turkmenistan	183	Sri Lanka
34	Peru	84	Israel	134	Antigua and Barbuda	184	Bahamas
35	Nigeria	85	Togo	135	Gambia	185	VirginIslands

36	Uzbekistan	86	Bolivia	136	Singapore	186	Comoros
37	Jordan	87	Iceland	137	Vietnam	187	Mauritius
38	Egypt	88	Saudi Arabia	138	Suriname	188	Aruba
39	Serbia	89	Albania	139	StKitts and Nevis	189	CaymanIslands
40	CzechRepublic	90	Cape Verde	140	Luxembourg	190	Eritrea
41	Tunisia	91	Lebanon	141	Liechtenstein	191	Montserrat
42	Costa Rica	92	Georgia	142	Yemen	192	Mongolia
43	Norway	93	Ethiopia	143	Afghanistan	193	Samoa
44	Cameroon	94	Congo DR	144	Cyprus	194	Cambodia
45	Iraq	95	Sierra Leone	145	Belize	195	Papua New Guinea
46	Mali	96	Guatemala	146	New Caledonia	196	Somalia
47	Romania	97	UnitedArabEmirates	147	Grenada	197	Andorra
48	South Africa	98	El Salvador	148	Dominica	198	Cook Islands
49	Oman	99	Kenya	149	Kazakhstan	199	Brunei
50	Uganda	100	Mozambique	150	Malaysia	200	East Timor

14.3. ANEXO 3: Resultados Parámetros estimados

Tabla 34. Parámetros estimados, datos años 2009-2013.

Ataque	Coefficiente	Error Estándar	t	P> t	Intervalo de Confianza	
algeria	.3147315	.1488112	2.11	0.035	.0229357	.6065274
angola	.200089	.1542308	1.30	0.195	-.1023337	.5025118
argentina	.7608125	.1211855	6.28	0.000	.5231865	.9984384
armenia	.5459245	.1462798	3.73	0.000	.2590924	.8327565
australia	.5053499	.1359409	3.72	0.000	.2387908	.771909
austria	.468473	.1414916	3.31	0.001	.1910297	.7459162
azerbaijan	.0537619	.1492872	0.36	0.719	-.2389672	.346491
belarus	.3880652	.140992	2.75	0.006	.1116016	.6645288
belgium	.5075456	.1281021	3.96	0.000	.2563572	.7587339
benin	.138582	.165985	0.83	0.404	-.186889	.4640529
bolivia	.3858257	.1432747	2.69	0.007	.1048861	.6667653
bosnia-herzegovina	.6125127	.1345551	4.55	0.000	.3486709	.8763546
botswana	.042184	.1590284	0.27	0.791	-.269646	.3540141
brazil	.8399448	.1188401	7.07	0.000	.6069177	1.072.972
bulgaria	.3183624	.1479388	2.15	0.031	.0282771	.6084476
burkina faso	.2467892	.1486867	1.66	0.097	-.0447624	.5383407
cameroon	.3681584	.1385697	2.66	0.008	.0964446	.6398722
cape verde	.0370426	.1775995	0.21	0.835	-.3112026	.3852878
Chile	.6265665	.1212069	5.17	0.000	.3888984	.8642345
china pr	.379861	.1415639	2.68	0.007	.1022761	.6574458
colombia	.5400733	.1310794	4.12	0.000	.2830469	.7970997
Congo	.1484875	.1867644	0.80	0.427	-.2177286	.5147037
congodr	.3607768	.1619894	2.23	0.026	.0431405	.678413
costa rica	.3757103	.1264276	2.97	0.003	.1278052	.6236153
croatia	.5325755	.1319095	4.04	0.000	.2739213	.7912297
Cuba	.095858	.1857577	0.52	0.606	-.2683841	.4601002
czechrepublic	.4652057	.1376502	3.38	0.001	.1952949	.7351165
denmark	.3572253	.1247033	2.86	0.004	.1127014	.6017491
dominicanrepublic	.1551563	.2412549	0.64	0.520	-.3179073	.6282198
ecuador	.6017255	.1307462	4.60	0.000	.3453523	.8580987
Egypt	.5254646	.1338021	3.93	0.000	.2630995	.7878298
el salvador	.3598724	.1360917	2.64	0.008	.0930177	.6267271
england	.6239589	.1228727	5.08	0.000	.3830245	.8648933
estonia	.4048321	.1369845	2.96	0.003	.1362267	.6734375
ethiopia	.2143205	.1756789	1.22	0.223	-.1301587	.5587996
finland	.3231816	.1367547	2.36	0.018	.0550266	.5913365
france	.5256966	.127104	4.14	0.000	.2764653	.7749279
gabon	.1778789	.1462249	1.22	0.224	-.1088455	.4646032
georgia	.210813	.146221	1.44	0.149	-.0759039	.4975299

germany	.9013044	.1211242	7.44	0.000	.6637985	113.881
ghana	.4109157	.1293453	3.18	0.002	.1572896	.6645419
greece	.2329047	.1339366	1.74	0.082	-.0297243	.4955337
guinea	.3127618	.1458364	2.14	0.032	.0267991	.5987244
haiti	.3033838	.1880963	1.61	0.107	-.0654439	.6722115
honduras	.4263472	.1284563	3.32	0.001	.1744641	.6782302
hungary	.3892589	.1376812	2.83	0.005	.1192873	.6592304
iceland	.4535714	.1435428	3.16	0.002	.1721061	.7350366
iran	.2346646	.1424334	1.65	0.100	-.0446254	.5139545
irelandrepublic	.4417554	.1291094	3.42	0.001	.1885919	.6949189
israel	.2531802	.1388181	1.82	0.068	-.0190207	.5253811
italy	.5395582	.1219436	4.42	0.000	.3004457	.7786707
ivorycoast	.6097321	.1306613	4.67	0.000	.3535254	.8659388
jamaica	.3869332	.1487803	2.60	0.009	.0951981	.6786684
japan	.5635156	.1291085	4.36	0.000	.3103538	.8166774
jordan	.2762383	.1446933	1.91	0.056	-.0074828	.5599595
korearepublic	.5569544	.1292964	4.31	0.000	.3034241	.8104846
libya	.3511415	.1698444	2.07	0.039	.0181029	.6841801
macedonia	.2863743	.1403572	2.04	0.041	.0111555	.5615932
mali	.5684686	.1431919	3.97	0.000	.2876914	.8492457
mexico	.535112	.1190641	4.49	0.000	.3016456	.7685784
moldova	.33125	.1555099	2.13	0.033	.026319	.636181
montenegro	.2583249	.1372788	1.88	0.060	-.0108576	.5275074
morocco	.1850501	.1512498	1.22	0.221	-.1115273	.4816276
netherlands	.7376483	.1218899	6.05	0.000	.498641	.9766555
new zealand	.2937957	.1674132	1.75	0.079	-.0344757	.6220672
nigeria	.5111449	.1351841	3.78	0.000	.2460698	.77622
northernireland	.1371402	.1592985	0.86	0.389	-.1752195	.4494999
norway	.3771822	.1322517	2.85	0.004	.117857	.6365074
oman	.2545249	.1526258	1.67	0.096	-.0447507	.5538006
panama	.4254432	.1300938	3.27	0.001	.1703494	.6805371
paraguay	.4729164	.127535	3.71	0.000	.22284	.7229928
peru	.3564645	.1324211	2.69	0.007	.0968072	.6161218
poland	.334909	.129068	2.59	0.010	.0818266	.5879914
portugal	.5169273	.1273937	4.06	0.000	.2671279	.7667267
romania	.4362433	.132245	3.30	0.001	.1769313	.6955553
russia	.4378846	.1330295	3.29	0.001	.1770343	.6987349
scotland	.4085868	.14722	2.78	0.006	.1199112	.6972624
senegal	.2943548	.1510527	1.95	0.051	-.0018363	.5905458
serbia	.4918427	.1308487	3.76	0.000	.2352686	.7484167
sierra leone	.2040258	.1939752	1.05	0.293	-.1763296	.5843812
slovakia	.2844383	.1348486	2.11	0.035	.020021	.5488556
slovenia	.3881799	.1368661	2.84	0.005	.1198065	.6565533
southafrica	.2876672	.1382266	2.08	0.038	.0166262	.5587081

spain	.8069658	.1184293	6.81	0.000	.5747443	1.039.187
sweden	.6570332	.1267454	5.18	0.000	.4085051	.9055612
switzerland	.5627907	.1354101	4.16	0.000	.2972724	.828309
togo	.020519	.154223	0.13	0.894	-.2818885	.3229265
trinidad and tobago	.3869084	.1597989	2.42	0.016	.0735673	.7002494
tunisia	.4005963	.1406755	2.85	0.004	.1247534	.6764393
turkey	.4739327	.131885	3.59	0.000	.2153266	.7325388
uganda	.2511598	.1586926	1.58	0.114	-.0600119	.5623315
ukraine	.5948383	.130762	4.55	0.000	.3384341	.8512424
unitedarabemirates	.4612047	.1686706	2.73	0.006	.1304677	.7919416
unitedstates	.4999294	.1188664	4.21	0.000	.2668506	.7330081
uruguay	.6344092	.122399	5.18	0.000	.3944036	.8744148
uzbekistan	.2703775	.153818	1.76	0.079	-.0312359	.5719909
venezuela	.4369909	.1284482	3.40	0.001	.1851238	.688858
wales	.2635136	.1432621	1.84	0.066	-.0174013	.5444285
zambia	.1925751	.1487196	1.29	0.195	-.0990411	.4841914

defensa						
algeria	.0510443	.1470864	0.35	0.729	-.2373694	.339458
angola	-.0337032	.1440593	-0.23	0.815	-.3161811	.2487748
argentina	-.3229733	.1234476	-2.62	0.009	-.565035	-.0809115
armenia	.0505302	.1337074	0.38	0.706	-.2116494	.3127098
australia	-.0431714	.1350649	-0.32	0.749	-.3080127	.22167
austria	.0141886	.1311107	0.11	0.914	-.2428993	.2712764
azerbaijan	.0408672	.1335919	0.31	0.760	-.2210859	.3028204
belarus	-.171217	.1340455	-1.28	0.202	-.4340595	.0916255
belgium	-.1220198	.1303487	-0.94	0.349	-.3776136	.1335739
benin	.124303	.1605292	0.77	0.439	-.1904699	.4390759
bolivia	-.0815693	.1282666	-0.64	0.525	-.3330803	.1699417
bosnia-herzegovina	-.0257336	.1345698	-0.19	0.848	-.2896043	.238137
botswana	.1822591	.1542669	1.18	0.238	-.1202345	.4847527
brazil	-.2978304	.1303153	-2.29	0.022	-.5533586	-.0423021
bulgaria	-.2033453	.1299947	-1.56	0.118	-.4582448	.0515542
burkina faso	.0320148	.1390261	0.23	0.818	-.2405939	.3046235
cameroon	-.0450147	.1398651	-0.32	0.748	-.3192686	.2292391
cape verde	-.0218593	.1751263	-0.12	0.901	-.365255	.3215364
chile	-.1452792	.1247705	-1.16	0.244	-.3899349	.0993765
china pr	-.006441	.1406372	-0.05	0.963	-.2822088	.2693268
colombia	-.3192416	.1398609	-2.28	0.023	-.5934873	-.044996
congo	.1606908	.1709651	0.94	0.347	-.1745454	.495927
congodr	.1296646	.1522234	0.85	0.394	-.1688221	.4281512
costa rica	-.1826012	.123184	-1.48	0.138	-.4241461	.0589437
croatia	-.2653776	.1329149	-2.00	0.046	-.5260033	-.004752
cuba	.2038036	.1568257	1.30	0.194	-.1037075	.5113146

czechrepublic	.015473	.140025	0.11	0.912	-.2590945	.2900405
denmark	-.0949744	.1267247	-0.75	0.454	-.3434621	.1535132
dominicanrepublic	.2161709	.2458568	0.88	0.379	-.2659164	.6982581
ecuador	-.2875479	.1296501	-2.22	0.027	-.5417717	-.033324
egypt	.0879581	.1382415	0.64	0.525	-.1831122	.3590284
el salvador	-.0328569	.1267283	-0.26	0.795	-.2813516	.2156377
england	-.3391673	.1290031	-2.63	0.009	-.5921225	-.0862121
estonia	.1115291	.1237381	0.90	0.367	-.1311022	.3541604
ethiopia	.1300526	.1553174	0.84	0.402	-.1745008	.4346061
finland	-.0361884	.1283984	-0.28	0.778	-.2879578	.2155809
france	-.4031619	.1268467	-3.18	0.001	-.6518887	-.1544351
gabon	-.0856558	.1427948	-0.60	0.549	-.3656544	.1943428
georgia	-.001528	.1368517	-0.01	0.991	-.2698731	.2668171
germany	-.1623218	.1241473	-1.31	0.191	-.4057556	.0811119
ghana	-.2303614	.1287548	-1.79	0.074	-.4828297	.0221068
greece	-.1692179	.1364895	-1.24	0.215	-.4368527	.0984169
guinea	.1707702	.1462868	1.17	0.243	-.1160756	.4576161
haiti	.15534	.1475801	1.05	0.293	-.1340418	.4447219
honduras	-.1468626	.1224263	-1.20	0.230	-.3869217	.0931966
hungary	-.0369787	.1316572	-0.28	0.779	-.2951382	.2211808
iceland	.1261096	.1334535	0.94	0.345	-.1355721	.3877913
iran	-.1894381	.1498165	-1.26	0.206	-.4832051	.104329
irelandrepublic	-.1656298	.1296598	-1.28	0.202	-.4198727	.088613
israel	.0869763	.1351571	0.64	0.520	-.1780458	.3519984
italy	-.2463513	.123689	-1.99	0.047	-.4888864	-.0038162
ivorycoast	.0704795	.1405462	0.50	0.616	-.2051099	.3460689
jamaica	-.0523791	.1371992	-0.38	0.703	-.3214054	.2166473
japan	-.0438407	.130307	-0.34	0.737	-.2993526	.2116712
jordan	.1005197	.146958	0.68	0.494	-.1876423	.3886816
korearepublic	-.0812484	.1302357	-0.62	0.533	-.3366206	.1741238
libya	-.1765135	.1622634	-1.09	0.277	-.4946869	.14166
macedonia	-.0793335	.1311178	-0.61	0.545	-.3364352	.1777681
mali	-.1426851	.1357889	-1.05	0.293	-.4089461	.1235759
mexico	-.2441737	.1216253	-2.01	0.045	-.4826622	-.0056852
moldova	.0716475	.1328352	0.54	0.590	-.1888218	.3321168
montenegro	.0928251	.1401769	0.66	0.508	-.1820402	.3676903
morocco	.0120044	.1507323	0.08	0.937	-.2835584	.3075671
netherlands	-.3213897	.1294013	-2.48	0.013	-.5751257	-.0676537
new zealand	.0840296	.1508827	0.56	0.578	-.211828	.3798872
nigeria	-.1762895	.1347138	-1.31	0.191	-.4404425	.0878634
northernireland	.1717222	.1329593	1.29	0.197	-.0889905	.4324348
norway	-.1535358	.1313304	-1.17	0.242	-.4110545	.103983
oman	-.0292945	.1421285	-0.21	0.837	-.3079866	.2493977
panama	-.143369	.1284744	-1.12	0.265	-.3952875	.1085496

paraguay	-.1157848	.1287897	-0.90	0.369	-.3683215	.136752
peru	-.2006549	.1315586	-1.53	0.127	-.458621	.0573112
poland	-.1073383	.1248185	-0.86	0.390	-.3520881	.1374114
portugal	-.1667979	.1368218	-1.22	0.223	-.4350843	.1014885
romania	-.0149886	.1323172	-0.11	0.910	-.2744423	.2444651
russia	-.3762129	.1381475	-2.72	0.007	-.6470989	-.105327
scotland	-.0240593	.1363099	-0.18	0.860	-.2913421	.2432234
senegal	-.055209	.1502229	-0.37	0.713	-.3497729	.2393549
serbia	-.1990181	.1295761	-1.54	0.125	-.4530968	.0550606
sierra leone	.0472897	.1809306	0.26	0.794	-.3074874	.4020667
slovakia	-.1243999	.1292012	-0.96	0.336	-.3777435	.1289438
slovenia	-.0487403	.1339228	-0.36	0.716	-.3113423	.2138617
southafrica	-.2732043	.1314436	-2.08	0.038	-.5309449	-.0154636
spain	-.3109396	.1307754	-2.38	0.017	-.56737	-.0545092
sweden	-.1160111	.1294393	-0.90	0.370	-.3698215	.1377993
switzerland	-.1179182	.1431561	-0.82	0.410	-.3986252	.1627889
togo	.1686162	.144344	1.17	0.243	-.1144202	.4516525
trinidad and tobago	.232559	.1418863	1.64	0.101	-.0456581	.5107761
tunisia	-.0336477	.137233	-0.25	0.806	-.3027404	.2354451
turkey	-.1495318	.1306872	-1.14	0.253	-.4057893	.1067256
uganda	-.0109149	.1604545	-0.07	0.946	-.3255415	.3037116
ukraine	-.0994888	.1311383	-0.76	0.448	-.3566308	.1576532
unitedarabemirates	.0634587	.1526854	0.42	0.678	-.2359338	.3628511
unitedstates	-.0031666	.1196397	-0.03	0.979	-.2377616	.2314284
uruguay	-.2552713	.1223777	-2.09	0.037	-.4952351	-.0153075
uzbekistan	-.0744681	.139402	-0.53	0.593	-.347814	.1988777
venezuela	-.0250169	.1245437	-0.20	0.841	-.269228	.2191941
wales	-.0100615	.1309126	-0.08	0.939	-.2667609	.246638
zambia	-.1272542	.1400546	-0.91	0.364	-.4018796	.1473712
local	.081537	.0246066	3.31	0.001	.0332872	.1297868
visita	-.0731372	.0256712	-2.85	0.004	-.1234745	-.0228
_cons	.0897033	.1423192	0.63	0.529	-.1893626	.3687692

14.4. ANEXO 4: Resultados Modelo de optimización con F.O 1

Tabla 35: Resultado modelo de optimización con F.O. 1, Opción 2

Grupo1	Grupo2	Grupo3	Grupo4	Grupo5	Grupo6	Grupo7	Grupo8
Alemania	España	Suiza	Uruguay	Bélgica	Brasil	Colombia	Argentina

	Inglaterra	Portugal	Rusia	Camerún	Argelia	Costa Rica	Japón	Irán
	Usa	Corea	Ghana	Holanda	México	Nigeria	Costa Marfil	Bosnia
	Australia	Chile	Honduras	Italia	Ecuador	Francia	Croacia	Grecia
Fza.	82	83	83	82	83	83	83	83

Tabla 36: Resultado modelo de optimización con F.O. 1, Opción 3

Grupo1	Grupo2	Grupo3	Grupo4	Grupo5	Grupo6	Grupo7	Grupo8
Alemania	España	Suiza	Uruguay	Bélgica	Brasil	Colombia	Argentina
Holanda	Grecia	Francia	Honduras	Italia	México	Japón	Corea
Usa	Costa Rica	Nigeria	Ghana	Australia	Argelia	Costa Marfil	Inglaterra
Camerún	Irán	Ecuador	Rusia	Chile	Bosnia	Croacia	Portugal
Fza.	82	83	83	82	83	83	83

Tabla 37: Resultado modelo de optimización con F.O. 1, Opción 4

Grupo1	Grupo2	Grupo3	Grupo4	Grupo5	Grupo6	Grupo7	Grupo8
Alemania	España	Suiza	Uruguay	Bélgica	Brasil	Colombia	Argentina
Italia	Holanda	Ecuador	México	Inglaterra	Honduras	Irán	Japón
Camerún	Costa Marfil	Argelia	Nigeria	Corea	Ghana	Bosnia	Costa Rica
Usa	Australia	Francia	Rusia	Chile	Grecia	Portugal	Croacia
Fza.	83	83	82	82	83	83	83

14.5. ANEXO 5: Resultados Modelo de optimización con F.O 2

Tabla 38: Resultado modelo de optimización con F.O. 2, Opción 2

Grupo1	Grupo2	Grupo3	Grupo4	Grupo5	Grupo6	Grupo7	Grupo8
Brasil	Uruguay	Argentina	Suiza	España	Colombia	Bélgica	Alemania
Nigeria	Honduras	Japón	Holanda	Ghana	Irán	Bosnia	Italia
Francia	Croacia	México	Costa Rica	Corea	Costa Marfil	Australia	Camerún
Rusia	Portugal	Inglaterra	Argelia	Chile	Grecia	USA	Ecuador
Fza.	66	66	66	66	66	66	66

Tabla 39: Resultado modelo de optimización con F.O. 2, Opción 3

Grupo1	Grupo2	Grupo3	Grupo4	Grupo5	Grupo6	Grupo7	Grupo8
Brasil	Uruguay	Argentina	Suiza	España	Colombia	Bélgica	Alemania
Corea	Costa Rica	Irán	Ecuador	Inglaterra	Camerún	Costa Marfil	Japón
Italia	Bosnia	Croacia	México	Honduras	Francia	Australia	Usa
Argelia	Rusia	Grecia	Portugal	Nigeria	Holanda	Chile	Ghana
Fza	66	66	66	66	66	66	66

Tabla 40: Resultado modelo de optimización con F.O. 2, Opción 4

Grupo1	Grupo2	Grupo3	Grupo4	Grupo5	Grupo6	Grupo7	Grupo8
Brasil	Uruguay	Argentina	Suiza	España	Colombia	Bélgica	Alemania
Japón	Costa Rica	Camerún	Portugal	Holanda	México	Croacia	Grecia
Costa Marfil	Argelia	Inglaterra	Usa	Nigeria	Ghana	Corea	Honduras
Francia	Italia	Rusia	Australia	Irán	Bosnia	Chile	Ecuador
Fza	66	66	66	66	66	66	66

Tabla 41: Resultado modelo de optimización con F.O. 2, Opción 5

Grupo1	Grupo2	Grupo3	Grupo4	Grupo5	Grupo6	Grupo7	Grupo8
Brasil	Uruguay	Argentina	Suiza	España	Colombia	Bélgica	Alemania
Japón	Australia	Irán	Grecia	Holanda	Nigeria	Croacia	Portugal
Honduras	Costa Marfil	Usa	Ghana	Camerún	Rusia	Corea	Argelia
Italia	Inglaterra	Francia	Ecuador	México	Bosnia	Chile	Costa Rica
Fza	66	66	66	66	66	66	66

Tabla 42: Resultado modelo de optimización con F.O. 2, Opción 6

Grupo1	Grupo2	Grupo3	Grupo4	Grupo5	Grupo6	Grupo7	Grupo8
Brasil	Uruguay	Argentina	Suiza	España	Colombia	Bélgica	Alemania
Irán	México	Australia	Rusia	Italia	Japón	Bosnia	Francia
Honduras	Argelia	Costa Marfil	Nigeria	Corea	Grecia	Camerún	Ghana
Holanda	Inglaterra	Australia	Usa	Costa Rica	Croacia	Chile	Ecuador
Fza	66	66	66	66	66	66	66

Tabla 43: Resultado modelo de optimización con F.O. 2, Opción 7

Grupo1	Grupo2	Grupo3	Grupo4	Grupo5	Grupo6	Grupo7	Grupo8
Brasil	Uruguay	Argentina	Suiza	España	Colombia	Bélgica	Alemania
Corea	Ghana	Honduras	Costa Marfil	Inglaterra	Japón	Bosnia	Portugal
Grecia	Croacia	Argelia	Irán	Camerún	Costa Rica	Usa	Nigeria
Francia	Rusia	Italia	Chile	Ecuador	Holanda	Australia	México
Fza	66	66	66	66	66	66	66

Tabla 44: Resultado modelo de optimización con F.O. 2, Opción 8

Grupo1	Grupo2	Grupo3	Grupo4	Grupo5	Grupo6	Grupo7	Grupo8
Brasil	Uruguay	Argentina	Suiza	España	Colombia	Bélgica	Alemania
Usa	Corea	Honduras	Ghana	Italia	Australia	Bosnia	Holanda
Camerún	Grecia	Costa Marfil	Costa Rica	Nigeria	Inglaterra	México	Argelia
Francia	Portugal	Croacia	Chile	Japón	Rusia	Ecuador	Irán
Fza	66	66	66	66	66	66	66

Tabla 45: Resultado modelo de optimización con F.O. 2, Opción 9

Grupo1	Grupo2	Grupo3	Grupo4	Grupo5	Grupo6	Grupo7	Grupo8
Brasil	Uruguay	Argentina	Suiza	España	Colombia	Bélgica	Alemania
Australia	Usa	Corea	Argelia	Portugal	Camerún	Inglaterra	México
Bosnia	Nigeria	Costa Marfil	Costa Rica	Ecuador	Italia	Ghana	Japón
Croacia	Francia	Grecia	Holanda	Irán	Rusia	Honduras	Chile
Fza	66	66	66	66	66	66	66

Tabla 46: Resultado modelo de optimización con F.O. 2, Opción 10

Grupo1	Grupo2	Grupo3	Grupo4	Grupo5	Grupo6	Grupo7	Grupo8
Brasil	Uruguay	Argentina	Suiza	España	Colombia	Bélgica	Alemania
Australia	México	Japón	Costa Rica	Inglaterra	Honduras	Costa Marfil	Irán
Grecia	Nigeria	Bosnia	Argelia	Ecuador	Francia	Corea	Ghana
Rusia	Italia	Croacia	Holanda	Camerún	Portugal	Usa	Chile
Fza	66	66	66	66	66	66	66