Applications of Operations Research and Statistics to Sports Analytics

Mario Guajardo

mario.guajardo@nhh.no

Deparment of Business and Management Science,
NHH Norwegian School of Economics, Bergen, Norway

Denis Sauré

denis.saure@dii.uchile.cl

Department of Industrial Engineering
University of Chile

ELAVIO 2017, Argentina

The speaker (10 years younger)

Rosiane

Gustavo & Yennifer

MARRIED

Luciana

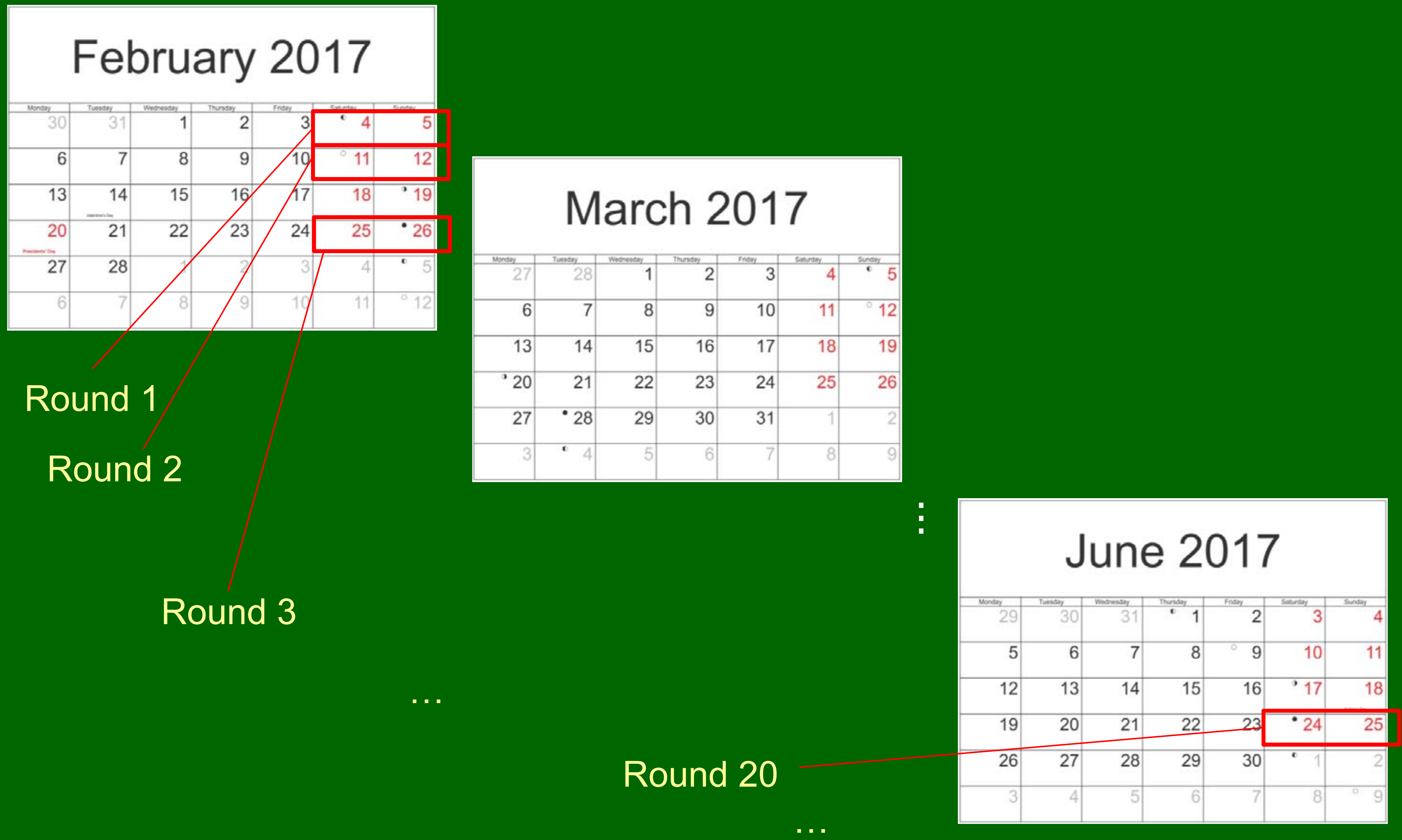
XII ELAVIO, Itaipava, Brazil, 2007

$$\sum_{j} x_{ijk} = 1$$

$$\sum_j$$ 

?

SCHEDULE OF A SPORT LEAGUE



“ SCHEDULE ” (or “FIXTURE”)

1ª jornada	2ª jornada	3ª jornada
9 octubre 2015	13 octubre 2015	13 noviembre 2015
Colombia - Perú	Paraguay - Argentina	Chile - Colombia
Chile - Brasil	Brasil - Venezuela	Argentina - Brasil
Argentina - Ecuador	Ecuador - Bolivia	Ecuador - Uruguay
Venezuela - Paraguay	Perú - Chile	Bolivia - Venezuela
Bolivia - Uruguay	Uruguay - Colombia	Perú - Paraguay

4ª jornada	5ª jornada	6ª jornada	7ª jornada	8ª jornada
17 noviembre 2015	junio 2016	junio 2016	septiembre 2016	septiembre 2016
Colombia - Argentina	Chile - Argentina	Colombia - Ecuador	Colombia - Venezuela	Chile - Bolivia
Paraguay - Bolivia	Brasil - Uruguay	Paraguay - Brasil	Paraguay - Chile	Brasil - Colombia
Brasil - Perú	Ecuador - Paraguay	Argentina - Bolivia	Argentina - Uruguay	Venezuela - Argentina
Venezuela - Ecuador	Bolivia - Colombia	Venezuela - Chile	Ecuador - Brasil	Perú - Ecuador
Uruguay - Chile	Perú - Venezuela	Uruguay - Perú	Bolivia - Perú	Uruguay - Paraguay

9ª jornada	10ª jornada	11ª jornada	12ª jornada	13ª jornada
octubre 2016	octubre 2016	marzo 2017	marzo 2017	junio 2017
Paraguay - Colombia	Colombia - Uruguay	Colombia - Chile	Chile - Uruguay	Colombia - Bolivia
Brasil - Bolivia	Chile - Perú	Paraguay - Perú	Argentina - Colombia	Paraguay - Ecuador
Ecuador - Chile	Argentina - Paraguay	Brasil - Argentina	Ecuador - Venezuela	Argentina - Chile
Perú - Argentina	Venezuela - Brasil	Venezuela - Bolivia	Bolivia - Paraguay	Venezuela - Perú
Uruguay - Venezuela	Bolivia - Ecuador	Uruguay - Ecuador	Perú - Brasil	Uruguay - Brasil

14ª jornada	15ª jornada	16ª jornada	17ª jornada	18ª jornada
junio 2017	septiembre 2017	septiembre 2017	octubre 2017	octubre 2017
Chile - Venezuela	Chile - Paraguay	Colombia - Brasil	Colombia - Paraguay	Paraguay - Venezuela
Brasil - Paraguay	Brasil - Ecuador	Paraguay - Uruguay	Chile - Ecuador	Brasil - Chile
Ecuador - Colombia	Venezuela - Colombia	Argentina - Venezuela	Argentina - Perú	Ecuador - Argentina
Bolivia - Argentina	Perú - Bolivia	Ecuador - Perú	Venezuela - Uruguay	Perú - Colombia
Perú - Uruguay	Uruguay - Argentina	Bolivia - Chile	Bolivia - Brasil	Uruguay - Bolivia

15th round Argentinean Football First Division 2016/17

March 3rd @ San Lorenzo’s stadium

San Lorenzo	vs	Belgrano

This match is played at this venue and on this date as result of a

DECISION

Nº Teams	Nº Schedules
2	1
4	6
6	720
8	31,449,600

...

MANY TEAMS: 16 TO 20

More than
- 200 games
- 30 rounds

HOW TO SELECT ONLY 1 SCHEDULE FROM MORE THAN MILLIONS OF ALTERNATIVES

OUTLINE

1) Background
2) Template schedules
3) League schedules
4) Implementation/Solution
5) Referee Assignment

Today

Friday

OUTLINE

1) Background

2) Template schedules

3) League schedules

4) Implementation/Solution

5) Referee Assignment

SPORTS ANALYTICS

Descriptive analytics

what's going on?

Predictive analytics

what will happen?

***Prescriptive* analytics**

what should be done?

BACKGROUND

SPORTS SCHEDULING

- How to schedule a tournament?
- Real problems are hard: many criteria, combinatorial structure, large dimension.
- Practical and theoretical development for about 40 years (Campbell and Chen 1976, Wright 2009, Kendal et al. 2009, Trick and Rasmussen 2008).
- TTP: Traveling Tournament Problem (Easton et al. 2001).

BACKGROUND

OR IN SCHEDULING SPORTS LEAGUES

- Basketball: USA (Nemhauser & Trick 1997), New Zealand (Wright 2004), Argentina (Durán et al. 2016).

- Cricket: Australia (Willis & Terrill 1994), England (Wright 1994), New Zealand (Wright 2005), World Cup (Armstrong & Willis 1993).

- Ice Hockey: USA (Ferland & Fleurent 1991), Finland (Kyngäs & Nurmi 2009).

- Table Tennis: Germany (Knust 2009).

- Volleyball: Holland (van Weert & Schreuder 1992), Argentina (Bonomo et al. 2012).

OR IN SCHEDULING FOOTBALL LEAGUES

- Holland (Schreuder 1992)
- Germany and Austria (Bartsch et al. 2006)
- Chile (Durán et al. 2007)
- Denmark (Rasmussen 2008)
- Belgium (Goossens and Spieksma 2009)
- Norway (Flatberg et al. 2009)
- Honduras (Fiallos et al. 2010)
- Brazil (Ribeiro and Urrutia 2011)
- Ecuador (Recalde et al. 2013)

WHY IS IT IMPORTANT TO IMPROVE THE SCHEDULES?

- **ECONOMIC REASONS**
  - Rising attendance to stadiums.
  - Reducing costs (travelling, hotels, TV).
- **SPORT REASONS**
  - Sport fairness.
  - Home/Away balance.
- **CONTRIBUTION TO SPECTACLE**
  - Important games on appropriate dates.
  - More attractive tournaments to the fans and media.

OUTLINE

1) Background

2) Template schedules

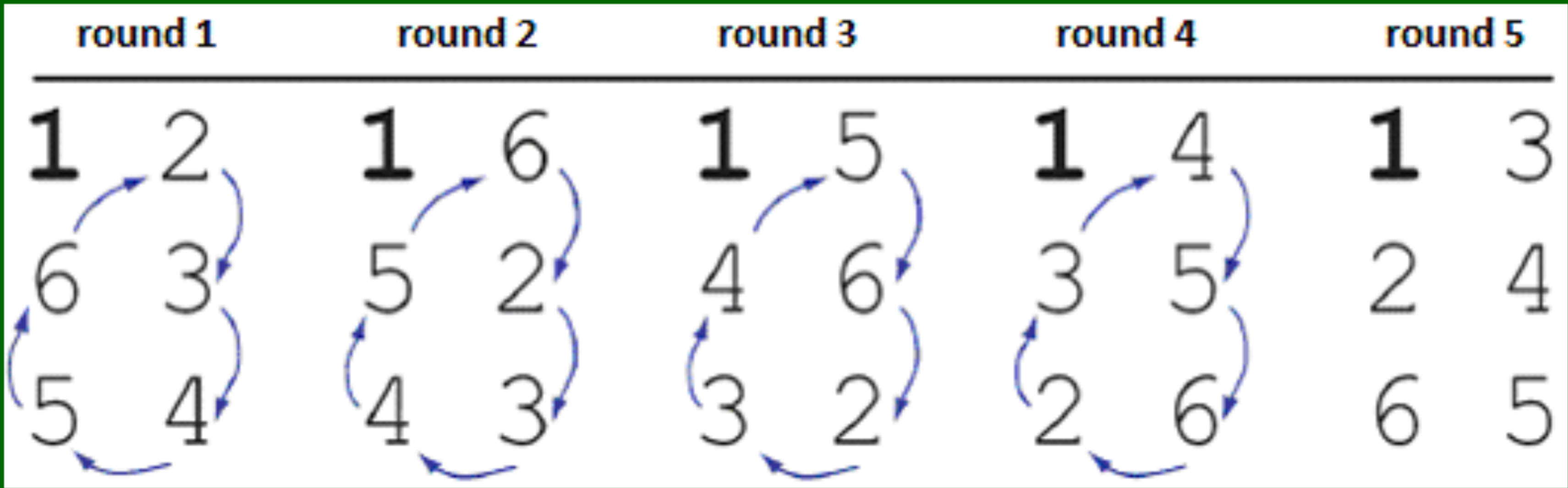
3) League schedules

4) Implementation/Solution

5) Referee Assignment

TRADITIONAL METHOD

Assigns teams randomly to a draw on a pre-established schedule template, based on a “circular” procedure.



CANONICAL SCHEDULE

EXAMPLE

18 teams, each team must play every other team over the 17-week season

Week 1	
1	2
3	18
4	17
5	16
6	15
7	14
8	13
9	12
10	11

CANONICAL SCHEDULE

EXAMPLE

18 teams, each team must play every other team over the 17-week season

Week 1	
1	2
3	18
4	17
5	16
6	15
7	14
8	13
9	12
10	11

Week 2	
1	3
4	2
5	18
6	17
7	16
8	15
9	14
10	13
11	12

CANONICAL SCHEDULE

EXAMPLE

18 teams, each team must play every other team over the 17-week season

Week 1		Week 2		Week 3	
1	2	1	3	1	4
3	18	4	2	5	3
4	17	5	18	6	2
5	16	6	17	7	18
6	15	7	16	8	17
7	14	8	15	9	16
8	13	9	14	10	15
9	12	10	13	11	14
10	11	11	12	12	13

CANONICAL SCHEDULE

EXAMPLE

18 teams, each team must play every other team over the 17-week season

Week 1	
1	2
3	18
4	17
5	16
6	15
7	14
8	13
9	12
10	11

Week 2	
1	3
4	2
5	18
6	17
7	16
8	15
9	14
10	13
11	12

Week 3	
1	4
5	3
6	2
7	18
8	17
9	16
10	15
11	14
12	13

…

Week 17	
1	18
2	17
3	16
4	15
5	14
6	13
7	12
8	11
9	10

CANONICAL SCHEDULE

Competition	Canonical 2008/09	Canonical 1994/95
Austria	yes	(yes)
Belgium	no	(yes)
Cyprus	yes	(yes)
Czech Rep.	no	(yes)
England	no	(no)
France	no	(no)
Germany	no	(yes)
Hungary	yes	(yes)
Ireland	no	(yes)
Italy	no	(no)
Luxembourg	yes	(yes)
Malta	yes	(yes)
Netherlands	no	(no)
N. Ireland	yes	(no)
Norway	no	(yes)
Poland	no	(yes)
Portugal	yes	(yes)
Romania	yes	(yes)
Russia	yes	(no)
Scotland	no	(no)
Slovakia	yes	(yes)
Spain	yes	(yes)
Switzerland	yes	(no)
Turkey	yes	(yes)
Wales	no	(no)

- A survey on 25 European football leagues concludes that the **popularity** of the canonical schedule still holds (Goossens & Spieksma 2012)
- 16 out of 25 leagues used it in the season 1994/95
- 13 out of 25 leagues used it in the season 2008/09
- Norwegian Tippeligaen used the canonical schedule until 2007

SCHEDULE OF THE TIPPELIGAEN 2007

Teams randomly assigned to a pre-established **canonical schedule** template.

Round	1	2	3	4	5	6	7	8	9	10	11	12	13
Team1	2	3	4	5	6	7	8	9	10	11	12	13	14
Team 2	1	14	3	4	5	6	7	8	9	10	11	12	13
Team 3	13	1	2	14	4	5	6	7	8	9	10	11	12
Team 4	12	13	1	2	3	14	5	6	7	8	9	10	11
Team 5	11	12	13	1	2	3	4	14	6	7	8	9	10
Team 6	10	11	12	13	1	2	3	4	5	14	7	8	9
Team 7	9	10	11	12	13	1	2	3	4	5	6	14	8
Team 8	14	9	10	11	12	13	1	2	3	4	5	6	7
Team 9	7	8	14	10	11	12	13	1	2	3	4	5	6
Team 10	6	7	8	9	14	11	12	13	1	2	3	4	5
Team 11	5	6	7	8	9	10	14	12	13	1	2	3	4
Team 12	4	5	6	7	8	9	10	11	12	13	1	2	3
Team 13	3	4	5	6	7	8	9	10	11	12	14	1	2
Team 14	8	2	9	3	10	4	11	5	12	6	13	7	1

Table taken from T. Flatberg, E.J. Nilssen, M. Stølevik. *Scheduling the topmost football leagues of Norway*. Euro 2009, Bonn.

THE NORWEGIAN CASE

League schedule –Tippeligaen 2007

- **1. round:**
  Stabæk - Brann
  Lillestrøm - Fredrikstad
  Lyn - Sandefjord
  Aalesund - Start
  Viking - Rosenborg
  Tromsø - Vålerenga
  Strømsgodset - Odd Grenland
- **2. round:**
  Vålerenga- Stabæk
  Rosenborg - Tromsø
  Start - Viking
  Sandefjord - Aalesund
  Fredrikstad - Lyn
  Odd Grenland - Lillestrøm
  Brann - Strømsgodset
- **3. round:**
  Stabæk - Rosenborg
  Brann - Vålerenga
  Lyn - Odd Grenland
  Aalesund - Fredrikstad
  Viking - Sandefjord
  Tromsø - Start
  Strømsgodset - Lillestrøm
- **4. round**
  Rosenborg - Brann
  Start - Stabæk
  Sandefjord - Tromsø
  Fredrikstad - Viking
  Odd Grenland - Aalesund
  Lillestrøm -Lyn
  Vålerenga - Strømsgodset
- **5. round:**
  Stabæk - Sandefjord
  Brann - Start
  Vålerenga - Rosenborg
  Aalesund - Lillestrøm
  Viking - Odd Grenland
  Tromsø - Fredrikstad
  Strømsgodset – Lyn
- **6. round:**
  Start - Vålerenga
  Sandefjord - Brann
  Fredrikstad - Stabæk
  Odd Grenland - Tromsø
  Lillestrøm - Viking
  Lyn - Aalesund
  Rosenborg - Strømsgodset

T. Flatberg, E.J. Nilssen, M. Stølevik (2009). *Scheduling the topmost football leagues of Norway*. http://folk.uio.no/trulsf/pub/euro2009.pdf

THE NORWEGIAN CASE

League schedule –Tippeligaen 2007

- **1. round:**
  1 - 2
  9 - 7
  10 - 6
  11 - 5
  12 - 4
  13 - 3
  14 - 8
- **2. round:**
  Vålerenga- Stabæk
  Rosenborg - Tromsø
  Start - Viking
  Sandefjord - Aalesund
  Fredrikstad - Lyn
  Odd Grenland - Lillestrøm
  Brann - Strømsgodset
- **3. round:**
  Stabæk - Rosenborg
  Brann - Vålerenga
  Lyn - Odd Grenland
  Aalesund - Fredrikstad
  Viking - Sandefjord
  Tromsø - Start
  Strømsgodset - Lillestrøm
- **4. round:**
  Rosenborg - Brann
  Start - Stabæk
  Sandefjord - Tromsø
  Fredrikstad - Viking
  Odd Grenland - Aalesund
  Lillestrøm -Lyn
  Vålerenga - Strømsgodset
- **5. round:**
  Stabæk - Sandefjord
  Brann - Start
  Vålerenga - Rosenborg
  Aalesund - Lillestrøm
  Viking - Odd Grenland
  Tromsø - Fredrikstad
  Strømsgodset – Lyn
- **6. round:**
  Start - Vålerenga
  Sandefjord - Brann
  Fredrikstad - Stabæk
  Odd Grenland - Tromsø
  Lillestrøm - Viking
  Lyn - Aalesund
  Rosenborg - Strømsgodset

T. Flatberg, E.J. Nilssen, M. Stølevik (2009). *Scheduling the topmost football leagues of Norway*. http://folk.uio.no/trulsf/pub/euro2009.pdf

THE NORWEGIAN CASE

Alanzinho

Player of the year, Tippeligaen 2007 & 2008

Result – Tippeligaen 2007

		S	V	U	T	+	–	D	P
1	Brann	26	17	3	6	59	39	20	**54**
2	Stabæk	26	14	6	6	53	35	18	**48**
3	Viking	26	14	5	7	50	40	10	**47**
4	Lillestrøm	26	12	8	6	47	28	19	**44**
5	Rosenborg	26	12	5	9	53	39	14	**41**
6	Tromsø	26	12	4	10	45	44	1	**40**
7	Vålerenga	26	10	6	10	34	34	0	**36**
8	Fredrikstad	26	9	9	8	37	40	-3	**36**

T. Flatberg, E.J. Nilssen, M. Stølevik. *Scheduling the topmost football leagues of Norway*. Euro 2009, Bonn.

THE NORWEGIAN CASE

October 28th - 2007

INTEGER PROGRAMMING MODEL

$$\min \sum_{i,j \in T} c_{i,j}$$

Minimizes carry-over effect

s.t.

$$
\begin{array}{ll}
\sum_{r \in R} x_{i,j,r} = 1, & i,j \in T, i \neq j \\
\sum_{j \in \Pi_{i,r}} x_{i,j,r} = 1, & i \in T, r \in R \\
x_{i,j,r} = x_{j,i,r}, & i,j \in T, r \in R \\
c_{ij} \geq \sum_{r \in R} y_{i,j,r} - 1, & i,j \in T \\
c_{ij} \geq 1 - \sum_{r \in R} y_{i,j,r}, & i,j \in T \\
y_{i,j,r} \geq x_{i,k,r} + x_{j,k,r+1} - 1, & i,j \in T, r \in R \\
\sum_{j \in T} y_{i,j,r} = 1, & i \in T, r \in R \\
y_{i,j,r} \in \{0,1\}, & i,j \in T, r \in R \\
x_{i,j,r} \in \{0,1\}, & i,j \in T, r \in R
\end{array}
$$

League constraints and logical constraints

T. Flatberg, E.J. Nilssen, M. Stølevik. *Scheduling the topmost football leagues of Norway*. Euro 2009, Bonn.

THE NORWEGIAN CASE

Tippeligaen - 2008

Team\round	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
1	3	9	4	14	5	2	8	10	6	11	7	12	13	6	10	5	9	7	11	8	2	3	12	13	14	4
2	8	6	10	12	11	1	7	9	13	4	3	5	14	8	7	9	11	12	3	13	1	14	4	10	5	6
3	1	7	9	4	6	5	14	8	11	12	2	13	10	7	8	11	12	13	2	4	14	1	10	9	6	5
4	13	11	1	3	9	6	5	14	8	2	10	7	12	5	9	10	6	8	7	3	13	12	2	14	11	1
5	7	13	12	9	1	3	4	6	14	10	8	2	11	4	14	1	10	9	6	11	12	13	7	8	2	3
6	11	2	7	13	3	4	9	5	1	14	12	10	8	1	13	14	4	10	5	9	7	8	11	12	3	2
7	5	3	6	8	14	10	2	11	12	13	1	4	9	3	2	13	14	1	4	10	6	9	5	11	8	12
8	2	10	11	7	12	13	1	3	4	9	5	14	6	2	3	12	13	4	14	1	10	6	9	5	7	11
9	12	1	3	5	4	14	6	2	10	8	13	11	7	14	4	2	1	5	10	6	11	7	8	3	12	13
10	14	8	2	11	13	7	12	1	9	5	4	6	3	12	1	4	5	6	9	7	8	11	3	2	13	14
11	6	4	8	10	2	12	13	7	3	1	14	9	5	13	12	3	2	14	1	5	9	10	6	7	4	8
12	9	14	5	2	8	11	10	13	7	3	6	1	4	10	11	8	3	2	13	14	5	4	1	6	9	7
13	4	5	14	6	10	8	11	12	2	7	9	3	1	11	6	7	8	3	12	2	4	5	14	1	10	9
14	10	12	13	1	7	9	3	4	5	6	11	8	2	9	5	6	7	11	8	12	3	2	13	4	1	10

T. Flatberg, E.J. Nilssen, M. Stølevik. *Scheduling the topmost football leagues of Norway*. Euro 2009, Bonn.

Result – Tippeligaen 2008

		S	V	U	T	+	–	D	P
1	Stabæk	26	16	6	4	58	24	34	**54**
2	Fredrikstad	26	14	6	6	38	28	10	**48**
3	Tromsø	26	12	8	6	36	23	13	**44**
4	Bodø/Glimt	26	12	6	8	37	38	-1	**42**
5	Rosenborg	26	11	6	9	40	34	6	**39**
6	Viking	26	11	6	9	38	32	6	**39**
7	Lyn	26	11	5	10	38	34	4	**38**
8	Brann	26	8	8	9	36	36	0	**33**

T. Flatberg, E.J. Nilssen, M. Stølevik. *Scheduling the topmost football leagues of Norway*. Euro 2009, Bonn.

DOES THE CARRY-OVER EFFECT EXIST?

D. Goossens, F. Spieksma (*Journal of Sports Economics*, 2011) used data from 30 seasons (~ 10,000 matches) to measure whether carryover effects have an influence on the outcome of football matches. The title of their paper is:

The carryover Effect Does Not Influence Football Results.

RONALDO

Golden Ball Best Player Award World Cup 1998

FORLÁN

Golden Ball Best Player Award World Cup 2010

https://www.youtube.com/watch?v=OAui7fpUkwc

Timeline

2005/1	2005/2	2007	2013	2015	2015-2017
1st Division	**U-19 Division**	**2nd Division**	**3rd, U-15, U-16 and U-17 Divisions**	**Chile Cup**	**South American World Cup Qualifiers**

Chile did not qualify to 2006 World Cup

Schedule proposal for 2010 World Cup rejected by CONMEBOL

(Made in ELAVIO 2007)

Chile qualified to 2010 World Cup

Chile qualified to 2014 World Cup

CONMEBOL agreed new schedule to be decided by a random draw…

FIFA Soccer World Cup

6 continental confederations

Qualification phase: ~200 teams

Final phase: 32 teams

CONMEBOL's South American World Cup Qualifiers

- 10 teams compete for 4.5 slots in the finals
- Double round robin: every team plays twice against every other team, once at home and once away
- 2 years
- 18 rounds grouped into 9 double rounds

Round 1 Round 2

1st double round

8 – 13 Oct

Round 3 Round 4

2nd double round

12 – 17 Nov

...

Breaks within double rounds

	Round 1	Round 2	Round 3	Round 4	
ARGENTINA	Home	Away	Home	Away	...
BOLIVIA	Away	Home	Away	Away	

break

1st double round

2nd double round

Inconvenient for the teams/players

Breaks within double rounds

	Round 4	Round 5	Round 6	Round 7
CHILE	Home	Away	Away	Home …

break

3rd double round

Nov 2011 — June 2012 — Sept 2012

Almost a year without a home game

Inconvenient for the fans/local sponsors

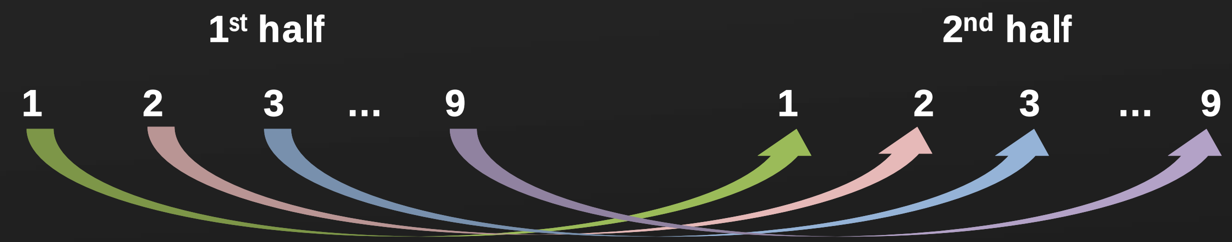
Home vs Away start of a double round

	Round 1	Round 2	Round 3	Round 4	
ARGENTINA	Home	Away	Home	Away	...
BOLIVIA	Away	Home	Away	Away	

1st double round (Rounds 1–2) — 2nd double round (Rounds 3–4)

For logistical reasons, starting a double round with a home game is usually preferred

Traditional mirrored **schedule** (5 World Cups 1998-2014)



Round 1	Round 10
Argentina – Chile	Chile – Argentina
Uruguay – Bolivia	Bolivia – Uruguay
Ecuador – Venezuela	Venezuela – Ecuador
Peru – Paraguay	Paraguay – Peru
Colombia – Brazil	Brazil – Colombia

1) Could we find a mirrored schedule without breaks within double rounds?

- A mirrored schedule with no breaks in double rounds for CONMEBOL’s qualifiers is infeasible
- Minimum: 16
- Old schedule 2002-2014: 18
- Tradition: mirrored hard to change (same schedule used for 4 last World Cups)

2) Could we find a schedule *as mirrored as possible* without breaks within double rounds?

Symmetric schedules

- Symmetric double round-robin tournament schedule: matches ordered with respect to some "structure", usually the second half with regard to the first one.
- 20 out of 25 European football leagues follow a symmetric scheme (Goossens & Spieksma 2012).
- 15 of them use a mirrored format.

Competition	Symmetry
Austria	English + English
Belgium	Mirror
Cyprus	Mirror
Czech Rep.	French
England	None
France	French
Germany	Mirror
Hungary	Mirror
Ireland	Mirror
Italy	Mirror
Luxembourg	French
Malta	Mirror
Netherlands	None
N. Ireland	Mirror
Norway	None
Poland	Mirror
Portugal	Mirror
Romania	Mirror
Russia	French
Scotland	None
Slovakia	Mirror
Spain	Mirror
Switzerland	Mirror + Inverted
Turkey	Mirror
Wales	None

Integer programming model

Decision variables

$$x_{i,j,k} = \begin{cases} 1 & \text{team } i \text{ plays at home against team } j \text{ in round } k, \\ 0 & \sim . \end{cases}$$

$i \in I$ (set of teams), $j \in I$ ($i \neq j$), $k \in K$ (set of rounds)

$$w_{i,k} = \begin{cases} 1 & \text{team } i \text{ has an away double round break in round } k, \\ 0 & \sim . \end{cases}$$

$$y_{i,k} = \begin{cases} 1 & \text{team } i \text{ has a H-A sequence in round } k, \\ 0 & \sim . \end{cases}$$

$i \in I$, $k \in K_{Odd} = \{1,3,5,\ldots\}$

$$x_{ijk} = \begin{cases} 1 & \text{if team } i \text{ plays at home against team } j \text{ in round } k \ (i \neq j) \\ 0 & \text{otherwise} \end{cases}$$

$i \in I$ (set of teams), $j \in I$ $(i \neq j)$, $k \in K$ (set of rounds)

Basic constraints

- Double round-robin split in two single-round robin halves: team i and team j play against each other once in each half of the tournament

$$\sum_{k \in K : k \leq n-1} (x_{i,j,k} + x_{j,i,k}) = 1 \quad \forall i \in I, j \in I : i \neq j$$

$$\sum_{k \in K : k > n-1} (x_{i,j,k} + x_{j,i,k}) = 1 \quad \forall i \in I, j \in I : i \neq j$$

Basic constraints

$$x_{ijk} = \begin{cases} 1 & \text{if team } i \text{ plays at home against team } j \text{ in round } k\ (i \neq j) \\ 0 & \text{otherwise} \end{cases}$$

$i \in I$ (set of teams), $j \in I$ $(i \neq j)$, $k \in K$ (set of rounds)

- Home-away balance with opponent: team i plays at home against team j exactly once

$$\sum_{k \in K} x_{i,j,k} = 1 \quad \forall i \in I, j \in I : i \neq j$$

- Compactness: every team plays exactly one game per round

$$\sum_{i \in I : i \neq j} (x_{i,j,k} + x_{j,i,k}) = 1 \quad \forall j \in I, k \in K$$

Double rounds constraints

$$y_{ik} = \begin{cases} 1 & \text{if team } i \text{ plays at home in round } k \text{ and away in round } k+1 \\ & i \in I, k \in K_{odd} = \{1,3,5,\ldots\} \\ 0 & \text{otherwise} \end{cases}$$

H-A ... H-A ...

- Balance of home-away sequences: every team starts a double round with a home game at least 4 and at most 5 times

$$\sum_{k \in K_{odd}} y_{i,k} \geq \frac{n}{2} - 1 \quad \forall i \in I$$

$$\sum_{k \in K_{odd}} y_{i,k} \leq \frac{n}{2} \quad \forall i \in I$$

- Logical relationships between variables x and y

$$\sum_{j \in I: i \neq j} (x_{i,j,k} + x_{j,i,k+1}) \leq 1 + y_{i,k} \quad \forall i \in I, k \in K_{odd}$$

$$y_{i,k} \leq \sum_{j \in I: i \neq j} x_{i,j,k} \quad \forall i \in I, k \in K_{odd}$$

$$y_{i,k} \leq \sum_{j \in I: i \neq j} x_{j,i,k+1} \quad \forall i \in I, k \in K_{odd}$$

$$w_{i,k} = \begin{cases} 1 & \text{team } i \text{ has an away double round break in round } k, \\ 0 & \sim . \end{cases}$$

Objective function

- Minimize breaks within double rounds

$$\min z = \sum_{i \in I} \sum_{k \in K_{odd}} w_{i,k}$$

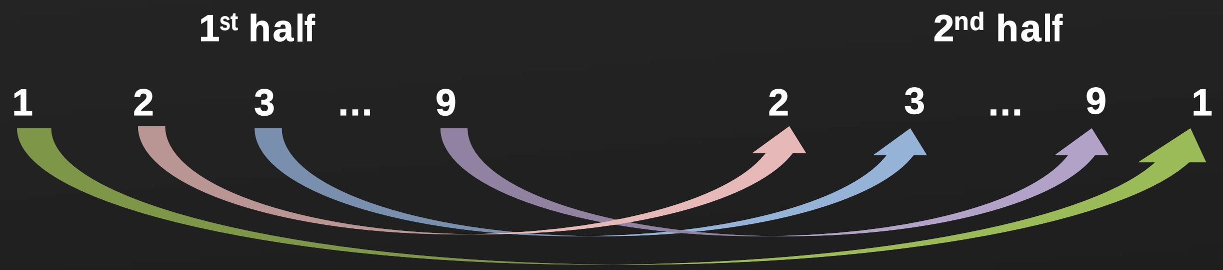
- Logical relationships

$$\sum_{j \in I : i \neq j} (x_{j,i,k} + x_{j,i,k+1}) \leq 1 + w_{i,k}, \quad \forall i \in I, k \in K_{odd},$$

$$w_{i,k} \leq \sum_{j \in I : i \neq j} x_{j,i,k}, \quad \forall i \in I, k \in K_{odd},$$

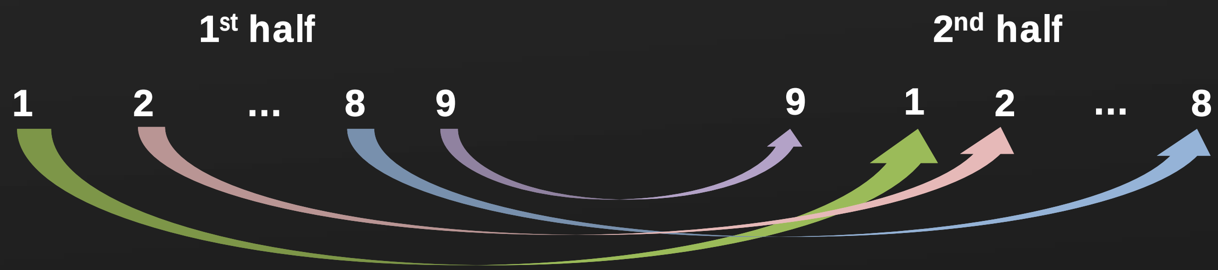
$$w_{i,k} \leq \sum_{j \in I : i \neq j} x_{j,i,k+1}, \quad \forall i \in I, k \in K_{odd}.$$

Symmetry: French scheme



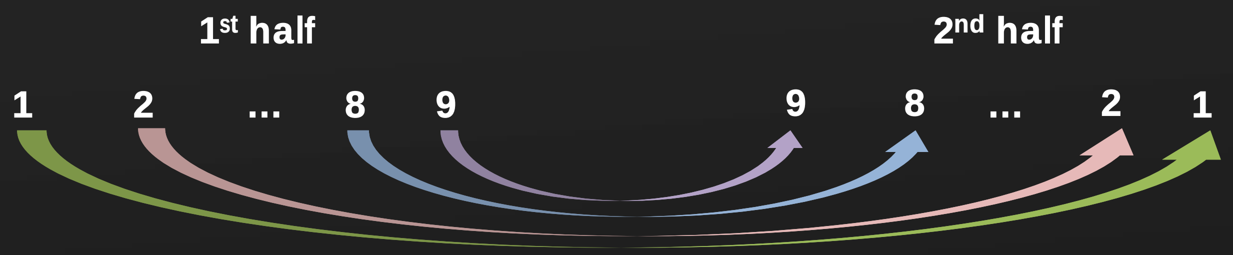
$$
\begin{aligned}
x_{i,j,1} &= x_{i,j,|K|} \quad \forall i \in I, j \in I : i \neq j \\
x_{i,j,k} &= x_{i,j,k+n-2} \quad \forall i \in I, j \in I, k \in K : i \neq j, 2 \leq k \leq n-1
\end{aligned}
$$

Symmetry: English scheme



$$
\begin{aligned}
x_{i,j,n-1} &= x_{i,j,n} \quad \forall i \in I, j \in I : i \neq j \\
x_{i,j,k} &= x_{i,j,k+n} \quad \forall i \in I, j \in I, k \in K : i \neq j, 2 \leq k \leq n-2
\end{aligned}
$$

Symmetry: Inverted scheme



$$x_{i,j,k} = x_{i,j,2n-1-k} \quad \forall i \in I, j \in I, k \in K : i \neq j, 1 \leq k \leq n-1$$

Symmetry: back-to-back scheme

$$x_{i,j,k} = x_{j,i,k+1} \quad \forall i \in I, j \in I, k \in K_{odd} : i \neq j$$

Symmetry: Min-Max separation scheme

- Every team plays against every other team at most once in c consecutive rounds

$$\sum_{\bar{k}\in K: k\leq \bar{k}\leq k+c-1} (x_{i,j,\bar{k}} + x_{j,i,\bar{k}}) \leq 1 \quad \forall i \in I, j \in I, k \in K : i \neq j, k \leq |K| - c + 1$$

- Every team plays against every other team at least once in d consecutive rounds

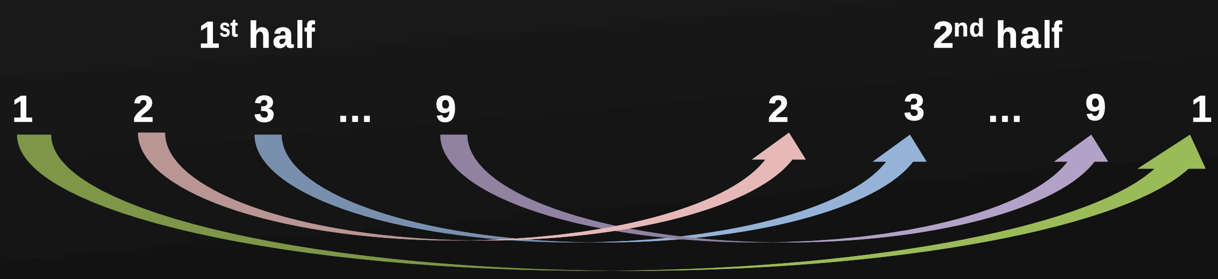
$$\sum_{\bar{k}\in K: k\leq \bar{k}\leq k+d-1} (x_{i,j,\bar{k}} + x_{j,i,\bar{k}}) \geq 1 \quad \forall i \in I, j \in I, k \in K : i \neq j, k \leq |K| - d + 1$$

Implementation

- AMPL/CPLEX 12.5, Intel Core 2 Duo 2.26GHz

  ~1700 binary variables

  ~1700 constraints
- Solutions found quickly
- Several proposals discussed with the Chilean ANFP officials
- They chose one of the template schedules that we generated according to the ***French scheme***



CONMEBOL meeting

http://conmebol.com/

Mirrored

Same schedule

Decision: new schedule by a random draw

Our schedule proposal
unanimously
agreed by the 10 CONMEBOL's countries

Qualifiers for the 2018 World Cup Russia

1998 **2002** **2014** **Jan-2015** **May-2015** **Oct-2015/Oct-2017**

TEAM NAME

TEAM NUMBER

2018 FIFA World Cup Russia™
Preliminary Competition Format & Draw Procedures
SOUTH AMERICAN ZONE

FIFA WORLD CUP RUSSIA 2018

Shaun Botterill/Getty Images Europe (zimbio.com/)

Ronaldo
Best World Cup
Player Award 1998

SCHEDULE

Round 1		Round 2	
Team 1	Team 9	Team 3	Team 4
Team 2	Team 5	Team 5	Team 7
Team 4	Team 6	Team 6	Team 8
Team 7	Team 3	Team 9	Team 2
Team 8	Team 10	Team 10	Team 1

...

Round 17		Round 18	
Team 1	Team 3	Team 9	Team 1
Team 2	Team 6	Team 5	Team 2
Team 4	Team 9	Team 6	Team 4
Team 7	Team 10	Team 3	Team 7
Team 8	Team 5	Team 10	Team 8

1st double-round

9th double-round

OPERATIONS RESEARCH

Shaun Botterill/Getty Images Europe (zimbio.com/)

Forlán
Best World Cup
Player Award 2010

Template schedule 2018 World Cup

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	9	@ 10	@ 2	4	@ 8	6	7	@ 5	@ 3	10	2	@ 4	8	@ 6	@ 7	5	3	@ 9
2	5	@ 9	1	@ 10	4	@ 7	@ 3	8	@ 6	9	@ 1	10	@ 4	7	3	@ 8	6	@ 5
3	@ 7	4	@ 9	8	@ 6	5	2	@ 10	1	@ 4	9	@ 8	6	@ 5	@ 2	10	@ 1	7
4	6	@ 3	5	@ 1	@ 2	8	10	@ 7	@ 9	3	@ 5	1	2	@ 8	@ 10	7	9	@ 6
5	@ 2	7	@ 4	9	10	@ 3	@ 6	1	8	@ 7	4	@ 9	@ 10	3	6	@ 1	@ 8	2
6	@ 4	8	10	@ 7	3	@ 1	5	@ 9	2	@ 8	@ 10	7	@ 3	1	@ 5	9	@ 2	4
7	3	@ 5	@ 8	6	@ 9	2	@ 1	4	@ 10	5	8	@ 6	9	@ 2	1	@ 4	10	@ 3
8	10	@ 6	7	@ 3	1	@ 4	9	@ 2	@ 5	6	@ 7	3	@ 1	4	@ 9	2	5	@ 10
9	@ 1	2	3	@ 5	7	@ 10	@ 8	6	4	@ 2	@ 3	5	@ 7	10	8	@ 6	@ 4	1
10	@ 8	1	@ 6	2	@ 5	9	@ 4	3	7	@ 1	6	@ 2	5	@ 9	4	@ 3	@ 7	8

- Top team constraints: no team plays consecutive matches against Argentina and Brazil.

$$\sum_{j \in I_S} (x_{i,j,k} + x_{j,i,k} + x_{i,j,k+1} + x_{j,i,k+1}) \leq 1 \quad \forall\, i \in I \setminus I_S, k \in K : k < |K|$$

New schedule 2018 World Cup

Number	Team	1	2	3	4	5	6	7	8	9
4	ARG	ECU	@ PAR	BRA	@ COL	@ CHI	BOL	URU	@ VEN	@ PER
8	BOL	URU	@ ECU	VEN	@ PAR	COL	@ ARG	PER	@ CHI	@ BRA
5	BRA	@ CHI	VEN	@ ARG	PER	URU	@ PAR	@ ECU	COL	BOL
2	CHI	BRA	@ PER	COL	@ URU	ARG	@ VEN	@ PAR	BOL	@ ECU
1	COL	PER	@ URU	@ CHI	ARG	@ BOL	ECU	VEN	@ BRA	@ PAR
6	ECU	@ ARG	BOL	URU	@ VEN	PAR	@ COL	BRA	@ PER	CHI
3	PAR	@ VEN	ARG	@ PER	BOL	@ ECU	BRA	CHI	@ URU	COL
9	PER	@ COL	CHI	PAR	@ BRA	VEN	@ URU	@ BOL	ECU	ARG
10	URU	@ BOL	COL	@ ECU	CHI	@ BRA	PER	@ ARG	PAR	VEN
7	VEN	PAR	@ BRA	@ BOL	ECU	@ PER	CHI	@ COL	ARG	@ URU

…

Old schedule 2002-2014 World Cups

Team	1	2	3	4	5	6	7	8	9
ARG	CHI	@VEN	BOL	@COL	ECU	@BRA	PAR	@PER	URU
BOL	@URU	COL	@ARG	@VEN	CHI	PAR	@ECU	@BRA	PER
BRA	@COL	ECU	@PER	URU	@PAR	ARG	@CHI	BOL	@VEN
CHI	@ARG	PER	@URU	PAR	@BOL	@VEN	BRA	COL	@ECU
COL	BRA	@BOL	VEN	ARG	@PER	@ECU	URU	@CHI	PAR
ECU	VEN	@BRA	@PAR	PER	@ARG	COL	BOL	@URU	CHI
PAR	@PER	URU	ECU	@CHI	BRA	@BOL	@ARG	VEN	@COL
PER	PAR	@CHI	BRA	@ECU	COL	@URU	VEN	ARG	@BOL
URU	BOL	@PAR	CHI	@BRA	VEN	PER	@COL	ECU	@ARG
VEN	@ECU	ARG	@COL	BOL	@URU	CHI	@PER	@PAR	BRA

...

Comparison

	Schedule 2002-2014					Schedule 2018				
Team	B_h	B_a	B	H-A	A-H	B_h	B_a	B	H-A	A-H
ARG	0	0	0	9	0	0	0	0	5	4
BOL	2	2	4	2	3	0	0	0	5	4
BRA	0	0	0	0	9	0	0	0	4	5
CHI	1	1	2	1	6	0	0	0	5	4
COL	1	1	2	6	1	0	0	0	5	4
ECU	1	1	2	4	3	0	0	0	4	5
PAR	1	1	2	3	4	0	0	0	4	5
PER	1	1	2	6	1	0	0	0	4	5
URU	1	1	2	4	3	0	0	0	4	5
VEN	1	1	2	1	6	0	0	0	5	4
Total	9	9	18	36	36	0	0	0	45	45

B_h: number of home breaks within double rounds
B_a: number of away breaks within double rounds
B: total number of breaks within double rounds
H-A: number of double rounds started with a home game
A-H: number of double rounds started with an away game

OUTLINE

1) Background

2) Template schedules

3) League schedules

4) Implementation/Solution

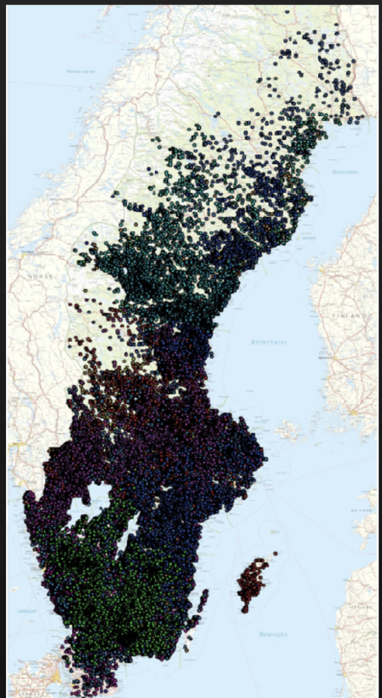
5) Referee Assignment

Cooperative game theory & logistics

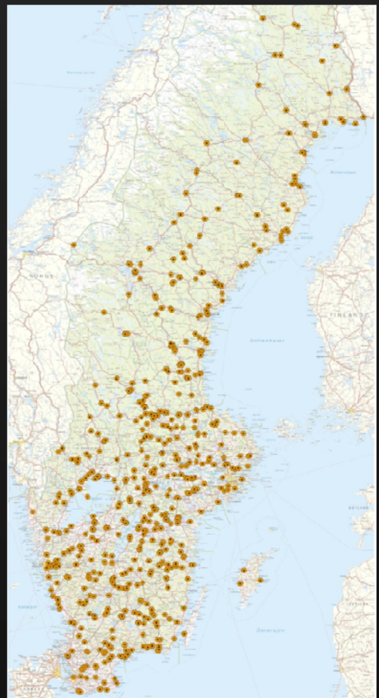
In Sweden, several companies supply forest fuels from harvest areas to heating plants where they are used as bioenergy source.

Which coalitions *should* form?

How to allocate costs if they cooperate?



Supply



Demand

Large-scale problems in forest transportation

UNSOLVED PROBLEM

- 18 teams
- 17 rounds
- Compact single round robin tournament
- 9 venues
- Schedule such that:

  Rounds 1 to 9 : All teams play in all venues

  Rounds 9 to 17 : All teams play in all venues

OUTLINE

1) Background

2) Template schedules

3) League schedules

4) Implementation/Solution

5) Referee Assignment

SCHEDULING THE FIRST DIVISION

- 16 - 20 teams, 2 tournaments per year. In each tournament, every team plays once against every other team.

- **MAIN GOAL:** *ATTRACTIVE* SCHEDULE OF GAMES.

- **OTHER GOALS:** Sport fairness, contributing to the financial benefit of teams and TV.

- **LIMITATIONS:** Fulfilling several conditions required by the ANFP, teams and TV, which are related to different issues and some of them imply a trade-off.

Marcelo Salas

INTEGER LINEAR PROGRAMMIG MODEL

Decides which games will be played in every round and which teams will play at home, fullfiling ANFP's requirements.

1st Round Jan-28		
O'HIGGINS	-	COLO COLO
EVERTON	-	U. CATOLICA
HUACHIPATO	-	ANTOFAGASTA
COBRESAL	-	COBRELOA
D. CONCEPCION	-	D. LA SERENA
PALESTINO	-	S. WANDERERS
D. PUERTO MONTT	-	A. ITALIANO
U. DE CHILE	-	U. ESPAÑOLA
S. MORNING	-	RANGERS
COQUIMBO	-	U. DE CONCEPCION

2nd Round Feb-4		
RANGERS	-	U. DE CHILE
D. LA SERENA	-	COBRESAL
COBRELOA	-	COQUIMBO
A. ITALIANO	-	EVERTON
ANTOFAGASTA	-	PALESTINO
U. DE CONCEPCION	-	S. MORNING
S. WANDERERS	-	O'HIGGINS
U. ESPAÑOLA	-	D. CONCEPCION
COLO COLO	-	HUACHIPATO
U. CATOLICA	-	D. PUERTO MONTT

⋮

19th Round Jun-03		
U. DE CONCEPCION	-	COLO COLO
HUACHIPATO	-	U. CATOLICA
EVERTON	-	COBRELOA
PALESTINO	-	D. LA SERENA
D. PUERTO MONTT	-	S. WANDERERS
COBRESAL	-	A. ITALIANO
ANTOFAGASTA	-	U. ESPAÑOLA
S. MORNING	-	O'HIGGINS
COQUIMBO	-	RANGERS
U. DE CHILE	-	D. CONCEPCION

DECISION VARIABLES

$$x_{ijk} = \begin{cases} 1 & \text{if team } i \text{ plays at home against team } j \text{ in round } k \\ 0 & \text{otherwise} \end{cases}$$

$i, j \in I$: Set of teams $\{\text{Team 1, Team 2, Team 3, ...}\}$

$k \in K$: Set of rounds $\{\text{Round 1, Round 2, Round 3, ...}\}$

CONSTRAINTS

Basic schedule constraints

1. Each team plays each of the others once over the course of the the tournament.

$$\sum_{k \in K} x_{ijk} + x_{jik} = 1 \qquad \forall i, j \in I,\ i \neq j$$

2. Each team plays one game each round either at home or away.

$$\sum_{\substack{j \in I \\ (j \neq i)}} x_{ijk} + x_{jik} = 1 \qquad \forall i \in I, k \in K$$

3. Lower and upper bounds to balance the number of home games.

$$\frac{n}{2} - 1 \leq \sum_{\substack{j \in I \\ (j \neq i)}} \sum_{k \in K} x_{ijk} \leq \frac{n}{2} \qquad \forall i \in I$$

IP MODEL

AUXILIARY DECISION VARIABLES

$$y_{ik} = \begin{cases} 1 & \text{if team } i \text{ plays at home in rounds } k \text{ and } k+1 \\ 0 & \text{otherwise} \end{cases}$$

$$w_{ik} = \begin{cases} 1 & \text{if team } i \text{ plays away in rounds } k \text{ and } k+1 \\ 0 & \text{otherwise} \end{cases}$$

$i \in I$: Set of teams
$k \in K$: Set of rounds

CONSTRAINTS

Home and away game constraints

4. Each team plays at most one away break.

$$\sum_{k<|K|} y_{ik} \leq 1 \qquad \forall i \in I$$

Relationship between variables x and y.

$$\sum_{j \in I} x_{ijk} + x_{ij(k+1)} \leq 1 + y_{ik} \qquad \forall i \in I, k < |K|$$

$$y_{ik} \leq \sum_{j \in I} x_{ijk} \qquad \forall i \in I, k < |K|$$

$$y_{ik} \leq \sum_{j \in I} x_{ij(k+1)} \qquad \forall i \in I, k < |K|$$

CONSTRAINTS

Home and away game constraints

5. Each team plays at most one home break.

$$\sum_{k<|K|} w_{ik} \leq 1 \qquad \forall i \in I$$

$$\sum_{j \in I} x_{jik} + x_{ji(k+1)} \leq 1 + w_{ik} \qquad \forall i \in I, k < |K|$$

$$w_{ik} \leq \sum_{j \in I} x_{jik} \qquad \forall i \in I, k < |K|$$

$$w_{ik} \leq \sum_{j \in I} x_{ji(k+1)} \qquad \forall i \in I, k < |K|$$

6. If a team plays at home (away) in an "adjustment round" (rounds 1, 16, 18), it must play away (home) in the following round. This is to provide a balanced start and end of the tournament for all teams.

$$\sum_{j \in I} x_{ijk} + x_{ij(k+1)} = 1 \qquad \forall i \in I, k \in A$$

Note till here conditions are as in template schedule (non-dependent on the identity of the teams)

CONSTRAINTS

Geographic conditions: Weekend - Weekend

7. When a North (South) team plays 2 consecutive away games, neither of them will be in the South (North).

$$\sum_{i \in South} x_{ijk} + x_{ij(k+1)} \le 1 - w_{jk} \quad \forall j \in North, k \notin WeekDay$$

$$\sum_{i \in North} x_{ijk} + x_{ij(k+1)} \le 1 - w_{jk} \quad \forall j \in South, k \notin WeekDay$$

8. When a North (South) team plays 2 consecutive away games, at least one of the games will be in the North (South).

$$w_{ik} \le \sum_{j \in North} x_{jik} + x_{ji(k+1)} \quad \forall i \in North, k \notin WeekDay$$

$$w_{ik} \le \sum_{j \in South} x_{jik} + x_{ji(k+1)} \quad \forall i \in South, k \notin WeekDay$$

4,200 km

CONSTRAINTS

Geographic conditions: Weekend - Weekend

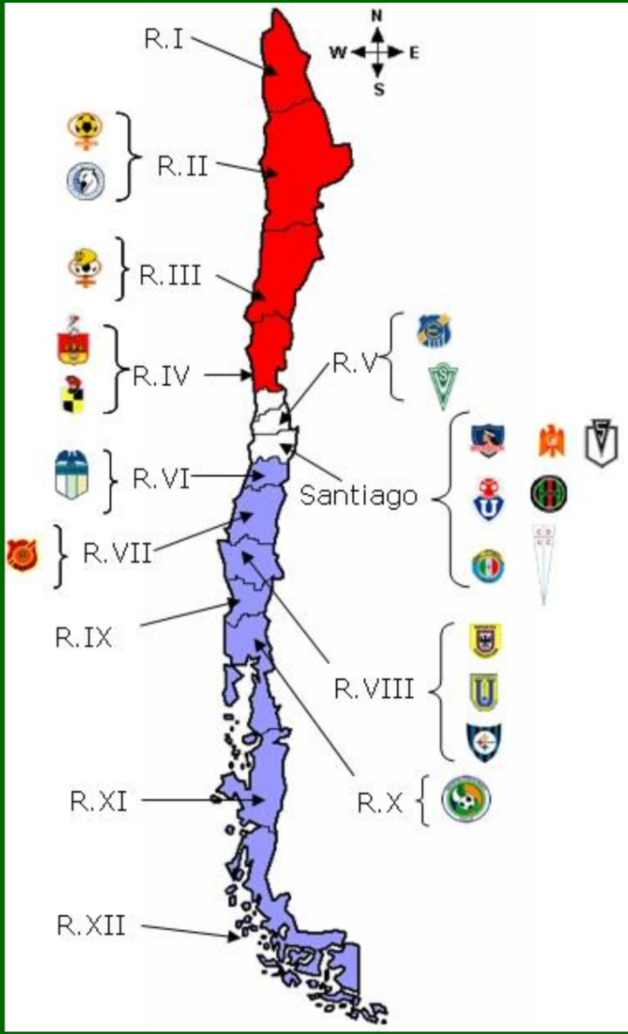
9. When a Center team plays 2 consecutive away games, at least one of the games will be in the Center.

$$w_{ik} \leq \sum_{j \in Center} x_{jik} + x_{ji(k+1)} \qquad \forall i \in Center, k \notin WeekDay$$

10. Each North (South) team shall play at least once at home against a North (South) team.

$$\sum_{k \in K} \sum_{j \in North} x_{ijk} \geq 1 \qquad \forall i \in North$$

$$\sum_{k \in K} \sum_{j \in South} x_{ijk} \geq 1 \qquad \forall i \in South$$



4,200 km

CONSTRAINTS

Geographic conditions: Wednesday - Weekend

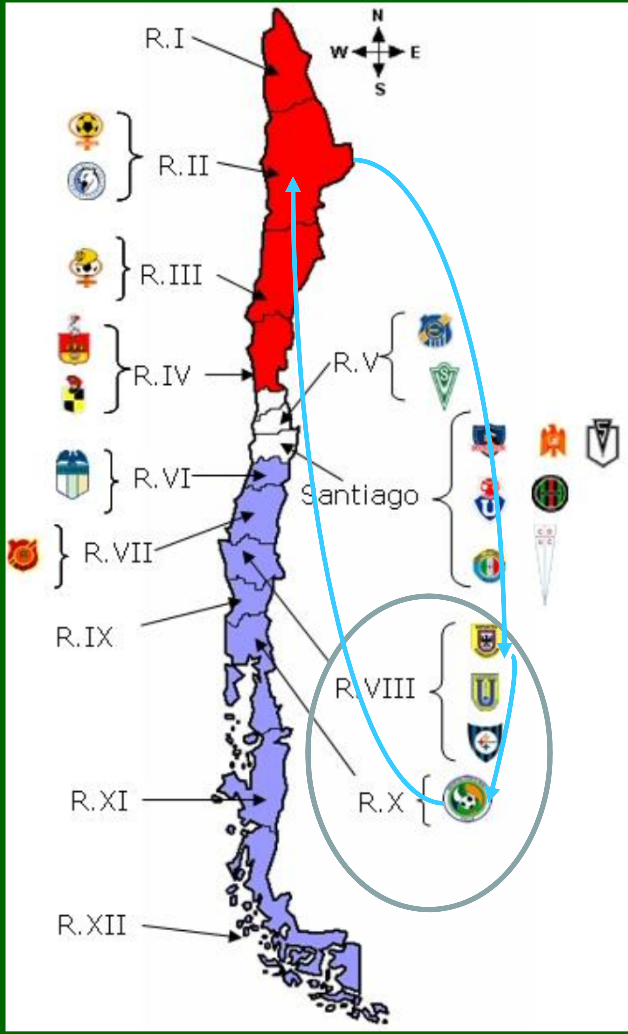
$$v_{ik} = \begin{cases} 1 & \text{if team } i \text{ has a good trip involving midweek round } k \\ 0 & \text{otherwise} \end{cases}$$

11. Logical conditions.

$$\sum_{j \in Good(i)} \left(x_{ji(k+1)} + 2x_{jik} + x_{ji(k-1)} \right) \geq 3v_{ik} \qquad \forall k \in WeekDay,\ i \in I$$

12. Minimum number of good trips.

$$\sum_{i \in I} \sum_{k \in K} v_{ik} \geq m$$



4,200 km

CONSTRAINTS

Constraints on highly popular teams

13. If a team plays at home (away) against *Colo Colo*, it plays away (at home) against *Universidad de Chile* (fairness and balance of revenue between the Opening and Closing tournaments).

$$\sum_{k \in K} x_{hik} + x_{hjk} = 1 \qquad \forall h, i = COLO, j = UCH$$

14. Each of the 3 popular teams plays exactly one classic game at home.

$$\sum_{k \in K} x_{hik} + x_{jik} = \sum_{k \in K} x_{hjk} + x_{ijk} \qquad \forall h = CATO, i = COLO, j = UCH$$

15. The 3 classic games between these popular teams must be played in one of the possible rounds for classic games contained in set $Kcla$.

$$\sum_{(i,j) \in PopT} \sum_{k \in Kcla} x_{ijk} = 3$$

CONSTRAINTS

Regional classic games

16. Regional classic games are held in some round contained in KRg.

$$\sum_{k \in KRg} x_{ijk} + x_{jik} = 0 \qquad \forall (i, j) \in RgC$$

Constraints on strong teams

17. No team may play 2 consecutive games against a strong team.

$$\sum_{j \in StrT} x_{ijk} + x_{jik} + x_{ij(k+1)} + x_{ji(k+1)} \leq 1$$

$$\forall i \in I, \; k < |K|$$

Complementary constraints

H - A - H - A . . .
A - H - A - H . . .

H - A - A - H . . .
A - H - H - A . . .

18. When one team of a complementary pair is playing at home, the other team of the pair plays away.

$$\sum_{h \in I} x_{ihk} + x_{jhk} = 1 \qquad \forall (i, j) \in CpT, k \in K$$

CONSTRAINTS

Stadium availability

19. Teams which not have their stadium available in some rounds should be scheduled for away games in the corresponding dates.

$$\sum_{i \in I} x_{ijk} = 1 \qquad \forall (j,k) \in SA$$

Games in Santiago

20. The number of games held in Santiago in each round cannot be less than $Smin$ or more than $Smax$.

$$Smin \leq \sum_{i \in Stgo} \sum_{j} x_{ijk} \leq Smax \qquad \forall k \in K$$

Tourism-related constraints

21. Each team located in a tourist area plays at home against at least one of the popular teams in one of the summer rounds.

$$\sum_{j \in Pop} \sum_{k \in KSu} x_{ijk} \geq 1 \qquad \forall i \in TouristT$$

CONSTRAINTS

TV Conditions

22. When a popular team plays in the North (South), neither of the 2 other popular teams can play in the South (North).

$$\sum_{i \in North} x_{ijk} \leq 1 - \sum_{i \in South} x_{ihk}$$

$$\forall j, h \in PopT, j \neq h, k \in K$$

International Competitions

23. In rounds consecutive to an away game in international competitions, it is usually better for teams to play at home or close.

$$\sum_{j \in I} x_{ijk} + \sum_{j \in Ng(i)} x_{jik} = 1 \qquad \forall (i,k) \in IC$$

Closing: Home-Away condition

24. If team *i* played at home against *j* in the Opening, team *i* must play away against *j* in the Closing tournament.

$$\sum_{k \in K} x_{jik} = 1 \qquad \forall (i,j) \in Op(i,j)$$

OPENING TOURNAMENT 2005

- **20 teams**, **19 rounds**, 4 groups of 5 teams each one.
- Regular season: All teams play against each other one game (single round robin).
- Playoffs: 2 best teams of each group advance to the the playoffs where a champion is decided.

GROUP A	GROUP B	GROUP C	GROUP D
1A) Colo Colo	**1B) Cobreloa**	**1C) Unión Española**	**1D) Univ. Católica**
2A) Huachipato	**2B) Coquimbo**	**2C) Dep. Concepción**	**2D) Univ. de Chile**
3A) San Felipe	3B) La Serena	3C) Univ. Concepción	3D) Cobresal
4A) Melipilla	4B) Wanderers	4C) Palestino	4D) Everton
5A) Audax Italiano	5B) Puerto Montt	5C) Temuco	5D) Rangers

OBJECTIVE FUNCTION

Maximizing the concentration of games between teams in the same group toward the final rounds of the tournament.

$$Max \quad \sum_{k\in K} \sum_{e\in E} \sum_{i\in t(e)} \sum_{j\in t(e)} k \cdot x_{ijk}$$

$t(e)$ denotes the set of teams in group e, $e \in \{1,2,3,4\}$

For example: If team 1 and team 2 belong to the same group, it is better if they play each other in round 17 or 18 (when the match will probably be decisive for the qualification to the playoff) than in round 1 or 2.

ADAPTING THE MODEL

- Conditions change from year to year: teams playing international tournaments, Wednesday rounds, new teams promoted to the 1st Division and other teams relegated to the 2nd Division…
- 2007-2015: the Closing had to be the mirror of the Opening tournament.
- 2008-2013: there were no groups and the 8 teams which scored more points during the regular season qualified to the playoffs.
- Since 2013, there are no playoffs.
- 2015: the objective function minimized the number of winter month games in the colder southern regions to reduce cancellations due to bad weather.
- 2017: A proposal to have one longer tournament per year instead of two shorter tournaments is approved for 2018.

OTHER OBJECTIVE FUNCTIONS

Maximizing the number of good trips

$$Max \sum_{k \in K} \sum_{i \in I} v_{ik}$$

- Decision variables

$$v_{ik} = \begin{cases} 1 & \text{if team } i \text{ has a good trip involving midweek round } k \\ 0 & \text{otherwise} \end{cases}$$

OTHER OBJECTIVE FUNCTIONS

Minimizing the distances travelled by the teams

$$\sum_{i \in I} \sum_{j \in I} \sum_{h \in I} \sum_{k \in K} d_{ij} z_{hijk}$$

- Decision variables

$$z_{hijk} = \begin{cases} 1 & \text{if from round } k \text{ to } k+1 \text{ team } h \text{ travels from the city of team } i \text{ to the city of team } j \\ 0 & \text{otherwise} \end{cases}$$

- Parameters

d_{ij} : distance between the city of team i and the city of team j

"TTP": Traveling tournament problem (Easton et al. 2001, Bonomo et al. 2008)

IP MODEL

TTP: Traveling tournament problem (Easton et al. 2001)

Input: n, the number of teams;
D an n by n integer distance matrix;
L, U integer parameters.

Output: A double round robin tournament on the n teams such that

– The length of every home stand and road trip is between L and U inclusive, and

– The total distance traveled by the teams is minimized.

Challenge Traveling Tournament Instances http://mat.tepper.cmu.edu/TOURN/

- NL10. 10 teams Data set
  Feasible Solution: 68691 (Rottembourg and Laburthe June 2001), 66369 (Dorrepaal, June 21 2002), 66037 (Cardemil, July 2 2002), 64597 (Zhang August 6, 2002), 61608 (Zhang August 19, 2002), 59583 (Van Hentenryck January 14, 2003), 59436 (Langford, June 13, 2005 solution)
  Lower Bound: 56506 (Waalewijn July 2001), 57500 (Easton June 2002), 57817 (Irnich and Schrempp, June 24 2008), 58831 (Uthus, Riddle, and Guesgen, Feb 11 2009), 59436 (Uthus, Riddle, and Guesgen December 11, 2009)

- NL12. 12 teams Data set
  Feasible Solution: 143655 (Rottembourg and Laburthe May 2001), 125803 (Cardemil, July 2 2002), 119990 (Dorrepaal July 16, 2002), 119012 (Zhang, August 19 2002), 118955 (Cardemil, November 1 2002), 114153 (Anagnostopoulos, Michel, Van Hentenryck and Vergados January 14, 2003), 113090 (Anagnostopoulos, Michel, Van Hentenryck and Vergados February 26, 2003), 112800 (Anagnostopoulos, Michel, Van Hentenryck and Vergados June 26, 2003), 112684 (Langford February 16, 2004), 112549 (Langford February 27, 2004), 112298 (Langford March 12, 2004), 111248 (Anagnostopoulos, Michel, Van Hentenryck and Vergados May 13, 2004), 110729 (Van Hentenryck and Vergados, May 30 2007).
  Lower Bound: 107483 (Waalewign August 2001), 107494 (Melo, Ribeiro, and Urrutia July 15 2006), 107548 (Mitchell, Trick and Waterer July 31 2008), 108244 (Uthus, Riddle, and Guesgen, Feb 11 2009), 108629 (Uthus, Riddle, and Guesgen January 6, 2010)

GAP

- Application to schedule the Argentinean Volleyball League (Bonomo et al. 2012)

OUTLINE

1) Background

2) Template schedules

3) League schedules

4) Implementation/Solution

5) Referee Assignment

SOLVING THE MODEL

- Exact methods to solve this formulation are not easy to compute given its complex structure and size:

  8,000 binary variables

  3,000 constraints

- Implementation: GAMS or AMPL (modelling software), and CPLEX (solver).

- Computer running more than 1 day and not even a feasible solution was found.

- Over the years, we have used a variety of strategies: constraint programming, LP relaxation based approaches, local search procedures, cuts.

Pattern-based decomposition approach

- Pattern:

Round 1	Round 2	Round 3	...			
H	A	H	A	A	H	A

H: Home
A: Away

Pattern-based decomposition approach

- Phase I :

Round 1	Round 2	Round 3				
H	A	H	A	A	H	A
A	H	A	H	H	A	H
H	A	A	H	A	H	A
A	H	A	H	A	A	H
			⋮			
A	H	A	H	A	H	A
			⋮			
H	A	H	A	H	A	A

Pattern-based decomposition approach

- Phase II :

Round 1	Round 2	Round 3					
H	A	H	A	A	H	A	Team 1
A	H	A	H	H	A	H	Team 2
H	A	A	H	A	H	A	Team 3
A	H	A	H	A	A	H	Team 4
⋮							
A	H	A	H	A	H	A	Team 8

Pattern-based decomposition approach

- Phase III :

Round 1	Round 2	Round 3				
H	A	H	A	A	H	A
H	A	H	A	A	H	A
H	A	A	H	A	H	A
A	H	A	H	A	A	H
			⋮			
A	H	A	H	A	H	A

Pattern-based decomposition approach

- Phase III :

Round 1	Round 2	Round 3				
H	A	H	A	A	H	A
H	A	H	A	A	H	A
⋮						
A	H	A	H	A	H	A

SOLVING METHODOLOGY

The pattern-based approach leads to decrease time to find feasible and close-to-optimum solutions dramatically.

IMPROVING SOLUTION

Local Search Procedures

- Swapping patterns
- Using more patterns than the number of teams
- Un-fixing patterns for some teams

CODE EXAMPLE (in AMPL)

```
#---- MODEL SOUTH AMERICAN QUALIFIERS - MAY 2013 - MMOTO ------#
set Teams;
set Rounds;
set FechasImpares within Rounds;

var x {Teams, Teams, Rounds} binary;    #1 if team i plays at home against j in round k, zero otherwise
var v {Teams, FechasImpares} binary;    #1 if team i plays L-V in round k and k+1, k impar
```

⋮

```
subject to

LocaloVisita{j in Teams, k in Rounds}:  #all team plays one match each round
        sum{i in Teams: i<>j} (x[i,j,k] + x[j,i,k]) = 1;

TodosContraTodos1{i in Teams, j in Teams: i<>j}:   #all teams play against each other once en 1ra rueda
        sum{k in Rounds:k<=9} (x[i,j,k] + x[j,i,k]) = 1;

TodosContraTodos2{i in Teams, j in Teams: i<>j}:   #all teams play against each other once en 2da rueda
        sum{k in Rounds:k<=18 and k>=10} (x[i,j,k] + x[j,i,k]) = 1;

TodosContraTodos1_Local{i in Teams, j in Teams: i<>j}:   #all teams play against each other once at home
        sum{k in Rounds} x[i,j,k] = 1;
```

CLUSTERS, SERVERS

- Submit jobs to a cluster of computers via SSH

- Online server, use **for free!**

  Pro versions of many languages & solvers

  http://www.neos-server.org/neos/solvers/index.html

FINAL SCHEDULE

Fulfills all conditions and schedules the most attractive games in appropriate rounds.

	Round 17	Round 18	Round 19
UCH	@UDC	RNGS	@CBSAL
COLO	HCH	@SFLP	MLPLL
CBLOA	LSRN	@WDRS	PMNTT
UDC	UCH	@CONCE	PLTN
CATO	@CBSAL	EVRT	@RNGS
AUDAX	WDRS	@HCH	SFLP
WDRS	@AUDAX	CBLOA	@CQMB
HCH	@COLO	AUDAX	@UE
UE	CONCE	@PLTN	HCH
CQMB	PMNTT	@LSRN	WDRS
TMC	PLTN	@PMNTT	CONCE
EVRT	RNGS	@CATO	LSRN
PMNTT	@CQMB	TMC	@CBLOA
SFLP	@MLPLL	COLO	@AUDAX
LSRN	@CBLOA	CQMB	@EVRT
RNGS	@EVRT	@UCH	CATO
PLTN	@TMC	UE	@UDC
CBSAL	CATO	@MLPLL	UCH
CONCE	@UE	UDC	@TMC
MLPLL	SFLP	CBSAL	@COLO

Example: During the last 3 rounds of the 2005 Opening Tournament, 24 games between teams of the same group were played (this is the maximum possible number for 3 rounds because at most 8 games between teams of the same group can be played in a given round).

OR IN PRACTICE

- **25 CHILEAN 1st DIVISION TOURNAMENTS:** All Opening and Closing tournaments between 2005-2017.

- **17 CHILEAN 2nd DIVISION TOURNAMENTS:** All tournaments between 2007-2017.
- **OTHERS:** Youth divisions, 3^{rd} Division, Chile Cup, South American Qualifiers to the 2018 FIFA World Cup Russia.

OUTLINE

1) Background

2) Template schedules

3) League schedules

4) Implementation/Solution

5) Referee Assignment

REFEREE ASSIGNMENT

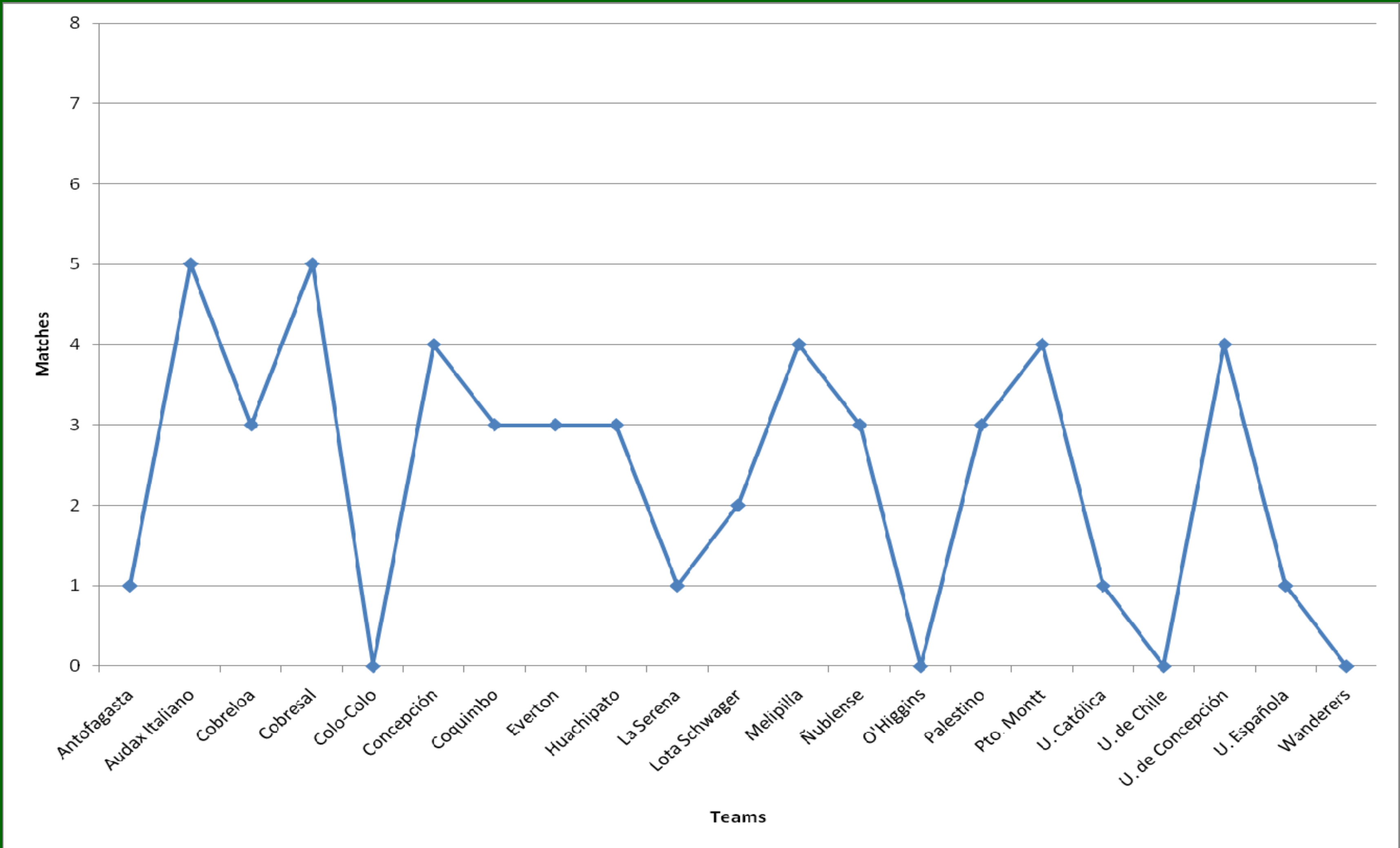
- **Theoretical work:**
  Room squares, Dinitz and Stinson (2005); Traveling Umpire Problem (TUP), Trick and Yildiz (2006, 2011, 2012); Referee Assignment Problem (RAP), Duarte et al. (2005); Spanish report, Gil and Rojas (2007); Turkish League, Yavuz et al. (2008); Turkish League, Atan and Hüseyinoglu, (2015).

- **Applications:** USA Baseball, Evans (1988); English cricket, Wright (1991); USA Tennis, Farmer et al. (2007); USA Baseball, Trick et al. (2012**); Argentinean Basketball**, Durán et al. (2017).

- **The literature on applications of sports scheduling techniques to real-world referee assignment problems is scarce relative to the applications to match scheduling.**

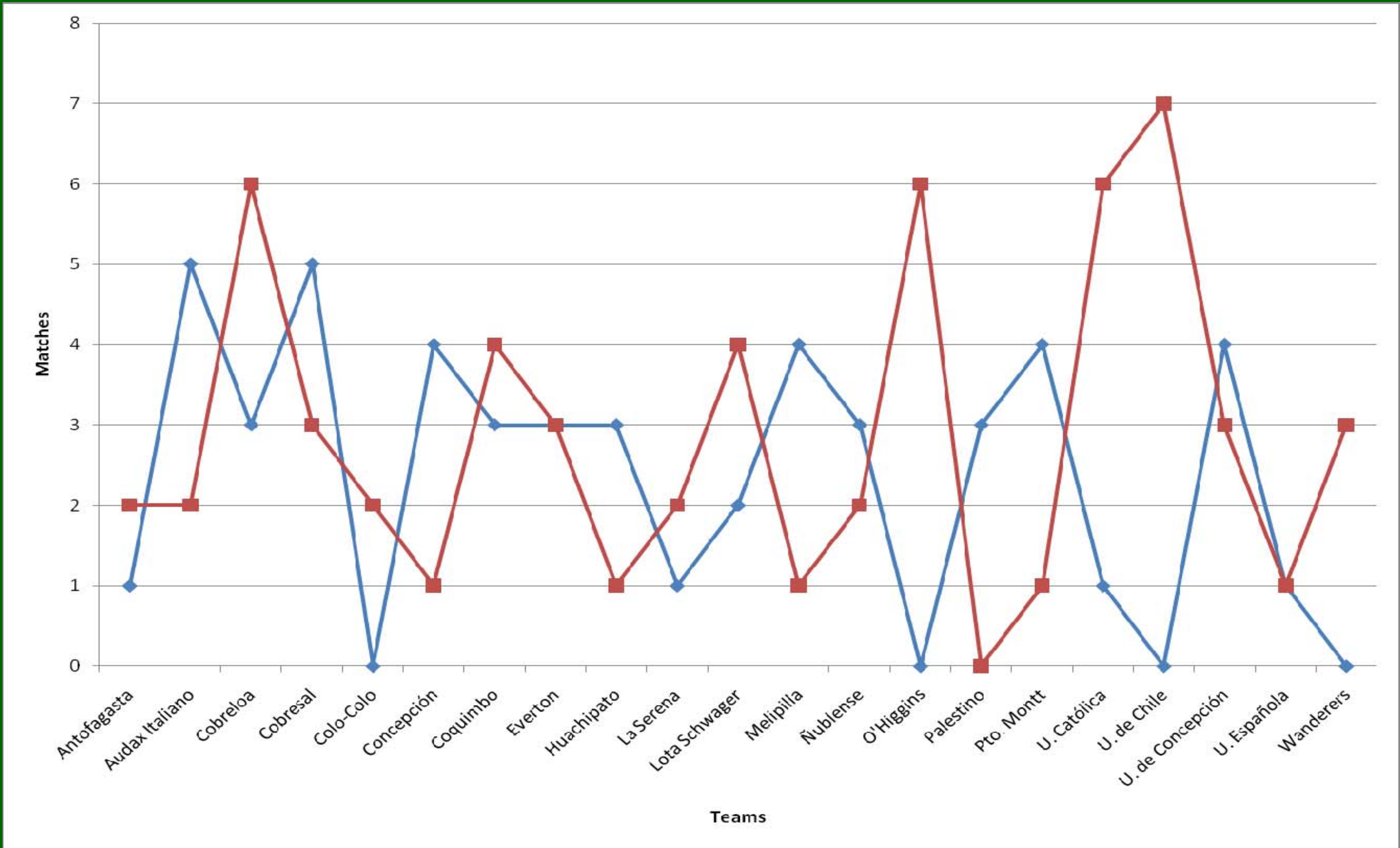
DISADVANTAGES OF THE TRADITIONAL ASSIGNMENT

FREQUENCY OF REFEREES vs TEAMS



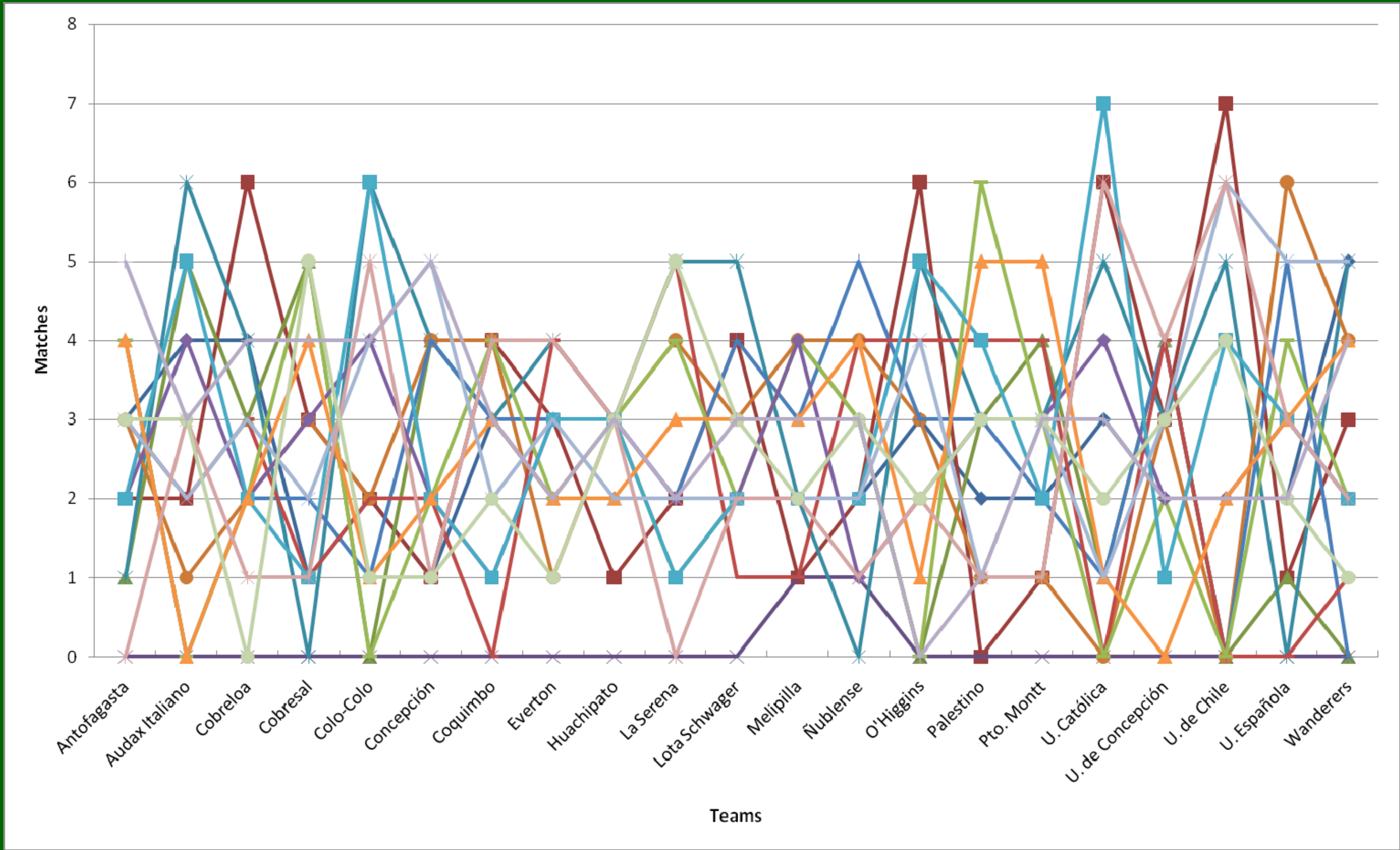
DISADVANTAGES OF THE TRADITIONAL ASSIGNMENT

FREQUENCY OF REFEREES vs TEAMS



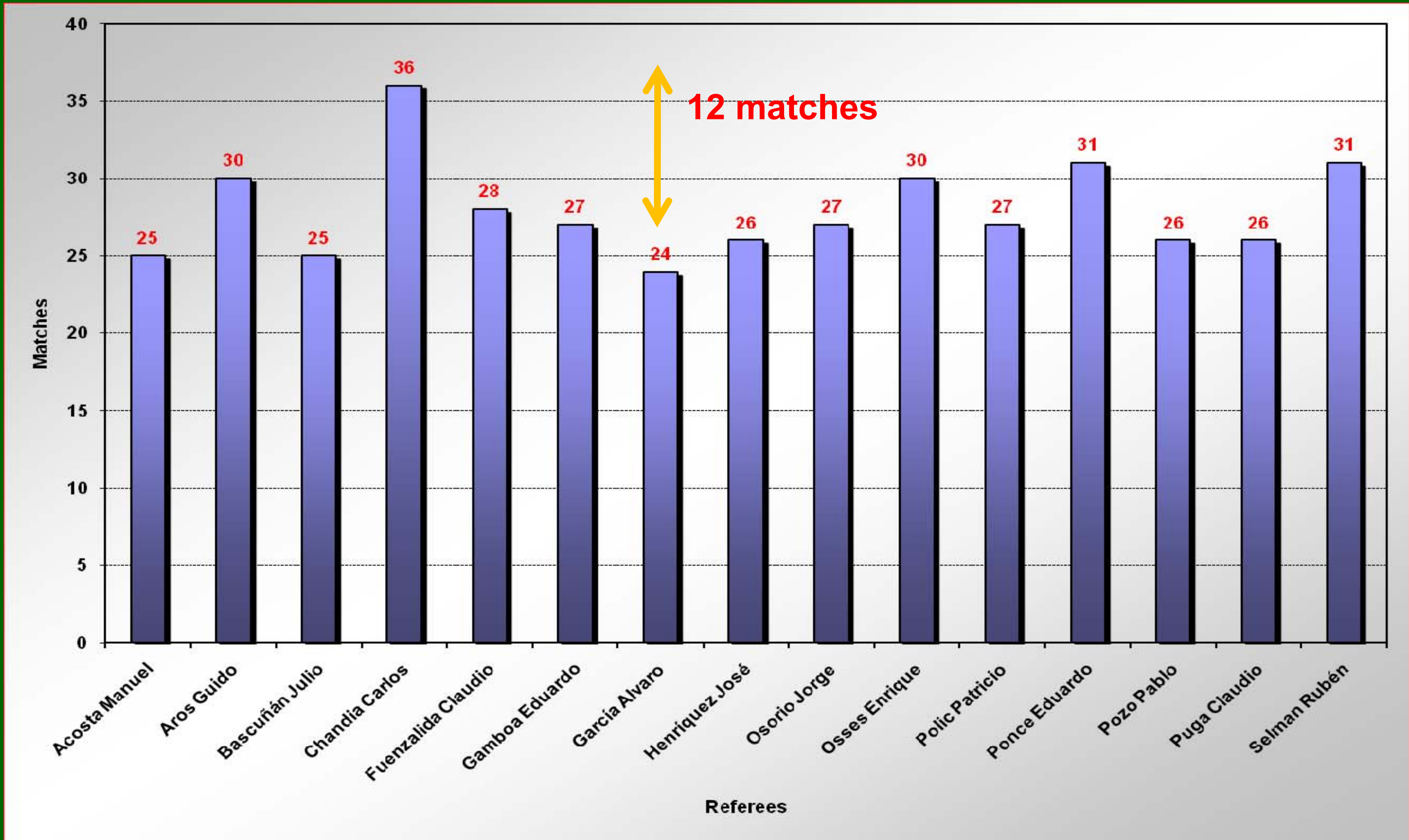
DISADVANTAGES OF THE TRADITIONAL ASSIGNMENT

FREQUENCY OF REFEREES vs TEAMS



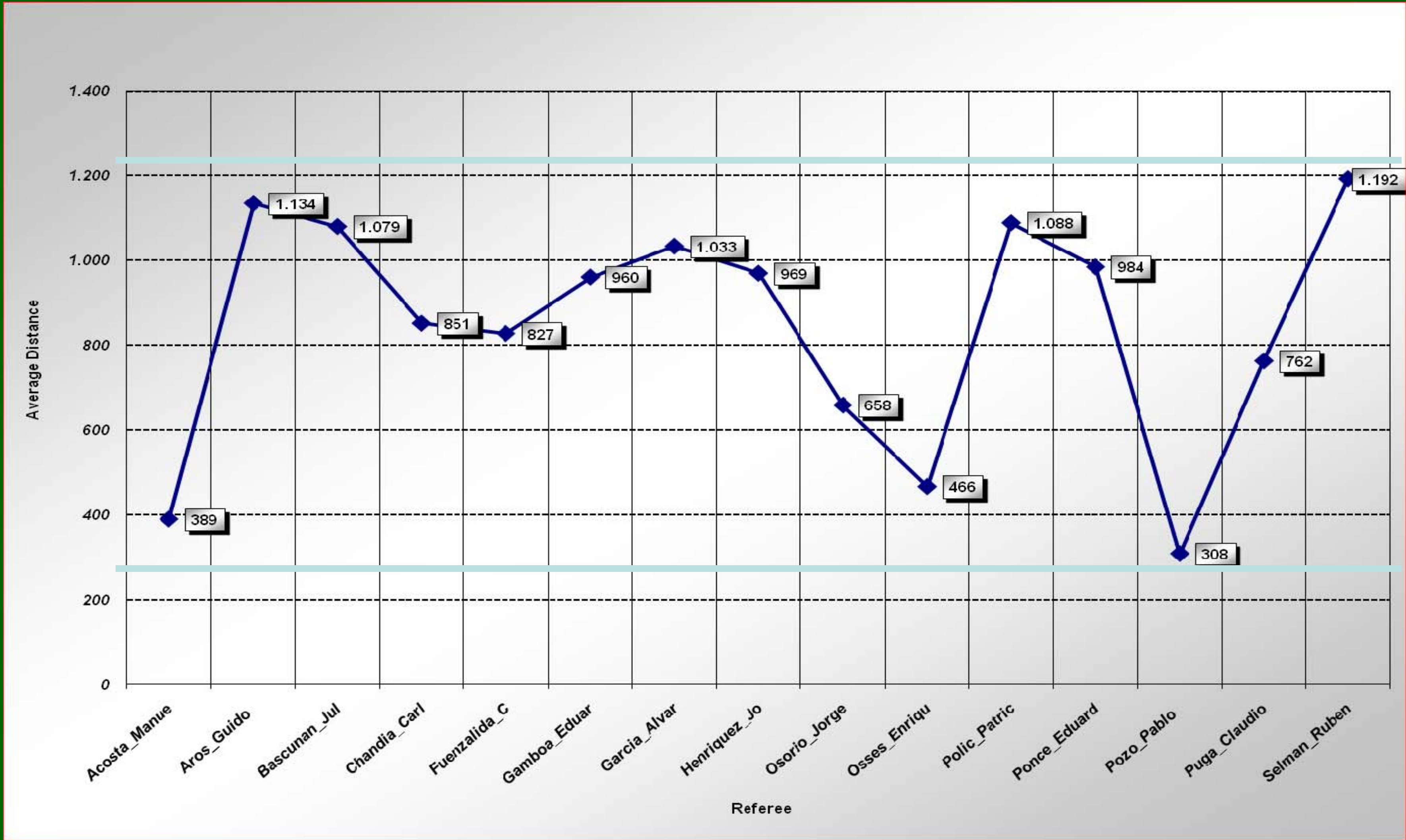
DISADVANTAGES OF THE TRADITIONAL ASSIGNMENT

TOTAL NUMBER OF GAMES BY EACH REFEREE



DISADVANTAGES OF THE TRADITIONAL ASSIGNMENT

AVERAGE TRAVEL DISTANCES



PROBLEM DEFINITION

Assigning referees to a schedule of matches satisfying a number of conditions

- Balancing the frequency of referees vs teams.
- Balancing the total number of games assigned to each referee.
- Balancing the average travel distances of the referees.
- Taking into account the experience of the referees to officiate some particular games.
- Considering some pre-defined assignments and unavailable referees per round.
- Desirably meeting a target number of games to be officiated for each referee.

The number of alternative assignments… hard to tackle the problem manually

N° of Teams	N° of Referees	Type of Tournament	N° of Posible Solutions (without any constraint)
2	1	SRR	1
2	2	SRR	2
2	2	DRR	4
4	2	SRR	8
4	2	DRR	64
4	3	DRR	46,656
6	4	DRR	6.34×10^{13}
6	5	DRR	6.04×10^{17}
10	6	DRR	2.70×10^{51}
10	8	DRR	7.81×10^{68}
21	16	DRR	5.36×10^{438}

SRR: Single round robin; DRR: Double round robin

IP MODEL

INTEGER LINEAR PROGRAMMIG MODEL

Given the whole schedule of games of the tournament, the model decides which referee will be assigned to each game, satisfying all conditions.

1st Round Jan-28		
O'HIGGINS	-	COLO COLO
EVERTON	-	U. CATOLICA
HUACHIPATO	-	ANTOFAGASTA
COBRESAL	-	COBRELOA
D. CONCEPCION	-	D. LA SERENA
PALESTINO	-	S. WANDERERS
D. PUERTO MONTT	-	A. ITALIANO
U. DE CHILE	-	U. ESPAÑOLA
S. MORNING	-	RANGERS
COQUIMBO	-	U. DE CONCEPCION

Carlos Chandía

.
.
.

Guido Aros

DECISION VARIABLES

$$x_{a,p} = \begin{cases} 1 & \text{if the referee } a \text{ is assigned to the game } p \\ 0 & \sim \end{cases}$$

Δ_a = Absolut value of the diference between the target and the actual number of games assigned to the referee a.

CONSTRAINTS

- Basic constraints

1) One referee is assigned to each game.

$$\sum_{a=1}^{|A|} x_{a,p} = 1 \qquad \forall\, p \in P.$$

2) Each referee can be assigned to at most one game per round.

$$\sum_{p=1}^{|P|} rounds^{p,f} \cdot x_{a,p} \leq 1 \qquad \forall\, a \in A,\ f \in F.$$

IP MODEL

- Referee-team balance constraints

3) Minimum number of times that the referee *a* must be assigned to matches where the team *e* plays.

$$\sum_{p=1}^{|P|} plays^{p,e} \cdot x_{a,p} \geq MINP^{a,e} \qquad \forall\, a \in A,\ e \in E.$$

4) Maximum number of times that the referee *a* must be assigned to matches where the team *e* plays.

$$\sum_{p=1}^{|P|} plays^{p,e} \cdot x_{a,p} \leq MAXP^{a,e} \qquad \forall\, a \in A,\ e \in E.$$

5) In D consecutive rounds, the same referee can be assigned to the same team at most one time.

$$\sum_{d=0}^{D-1} \sum_{p=1}^{P} plays^{p,e} \cdot rounds^{p,f+d} \cdot x_{a,p} \leq 1 \qquad \forall\, a \in A,\ \forall f \leq |F|-D+1,\ \forall e \in E.$$

IP MODEL

- Season match assignment balance constraints

6) Minimum number of total season match assignments for each referee.

$$\sum_{p=1}^{|P|} x_{a,p} \geq MINT^a \qquad \forall\, a \in A.$$

7) Maximum number of total season match assignments for each referee.

$$\sum_{p=1}^{|P|} x_{a,p} \leq MAXT^a \qquad \forall\, a \in A.$$

IP MODEL

- Average travel distance balance constraints

8) A bound on the difference between the *average* distances travelled by the referee *a* and by the referee r.

$$\frac{1}{T^a}\sum_{p=1}^{|P|} DIST^{a,p}\cdot x_{a,p} - \frac{1}{T^r}\sum_{p=1}^{|P|} DIST^{r,p}\cdot x_{r,p} \leq DISTMAX \quad \forall\, a,r \in A.$$

- Bounding the inactivity of the referees

9) Maximum number of consecutive rounds for which a referee may have no assignment.

$$\sum_{s=0}^{S^a}\sum_{p=1}^{|P|} rounds^{p,f+s}\cdot x_{a,p} \geq 1 \qquad \forall\, a \in A,\ f \leq |F| - S^a.$$

IP MODEL

- Referee category and match importance level

10) The category of a referee assigned to a game must fulfill the level required for the game.

$$\sum_{a \in A_{High}} x_{a,p} = 1 \qquad \forall\, p \in P_{High}.$$

$$\sum_{a \in A_{Med} \cup A_{High}} x_{a,p} = 1 \qquad \forall\, p \in P_{Med}.$$

11) A same referee can not be assigned to two consecutive games of the highest level.

$$x_{a,p} + x_{a,\hat{p}} \leq 1 \qquad \forall\, a \in A,\ p, \hat{p} \in PN(p).$$

IP MODEL

- Fixed assignments

12) Referee *a* can not be assigned to game *p*.

$$x_{a,p} = 0 \qquad \forall\,(a,p) \in NOFIX.$$

13) Referee *a* must be assigned to game *p*.

$$x_{a,p} = 1 \qquad \forall\,(a,p) \in FIX.$$

IP MODEL

OBJECTIVE FUNCTION

Minimize the sum over all referees of the absolute value of the difference between the target and the actual number of games assigned to each referee (RAP, Duarte et al. 2007)

$$\min g = \sum_{a \in A} \Delta_a$$

- Constraints to compute the value of the variables Δ_a

$$\sum_{p \in P} x_{a,p} + \Delta_a \geq T^a \qquad \forall\, a \in A.$$

$$\sum_{p \in P} x_{a,p} - \Delta_a \leq T^a \qquad \forall\, a \in A.$$

SOLVING THE IP MODEL

- The IP model contains about 6,700 variables and 23,000 constraints.
- Implementation in AMPL/CPLEX, optimal solution in 14 to 72 minutes.
- Performance in solution time can be improved by a decomposition approach that solves two problems sequentially: 1) pattern generation; 2) pattern-based assignment (Alarcón et al. 2014).

  Optimal solution found in 2 seconds to 5 minutes.

PATTERNS FOR ASSIGNING REFEREES

REFEREE PATTERNS

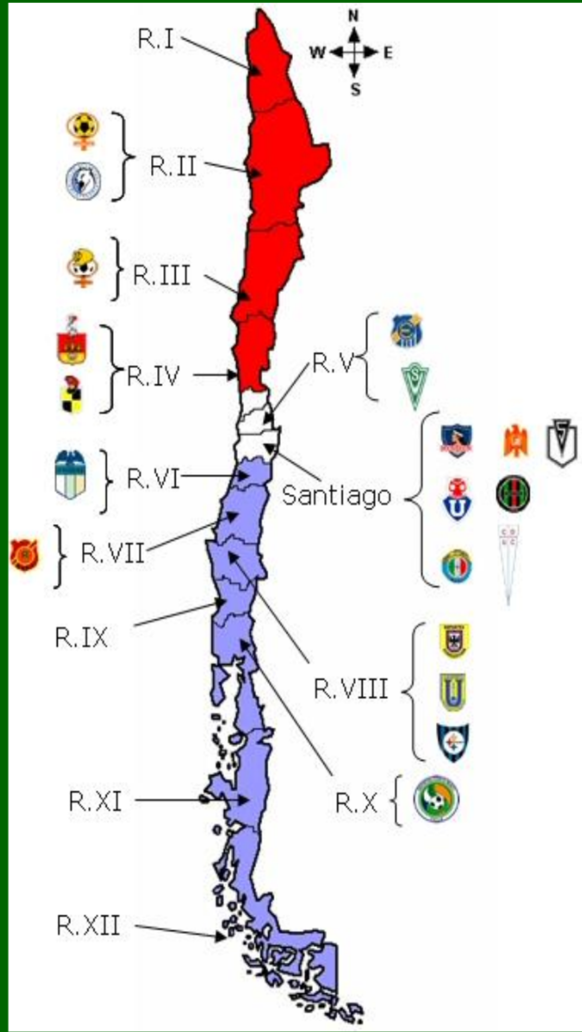
- The patterns we implement for the referee assignment indicate the set of games to which a referee can be assigned in each round.

- Given the particular geography of Chile (length: 4,200 kms), we define these sets of games based on the location of the venues where they are going to be played. Any other arbitrary criteria to define the sets may also be suitable.

$$P= \overbrace{(N, S, C, F, N\ldots)}^{\text{Nº of rounds}}$$

4 *clusters*

N: North C: Center S: South F: Free



SOLVING METHODOLOGY

1) PATTERN-GENERATION MODEL (GP Model)

Generates the patterns for each referee by solving an IP model that considers a modified version of some constraints of the original problem.

$$P = (\overbrace{N, S, C, F, N\ldots}^{\text{Nº of rounds}})$$

$$\overbrace{P_1 \quad P_2 \quad P_3 \quad P_4 \ldots}^{\text{Nº of referees}}$$

$$\downarrow \quad \downarrow \quad \downarrow \quad \downarrow$$

2) PATTERN-BASED ASSIGNMENT MODEL (AP Model)

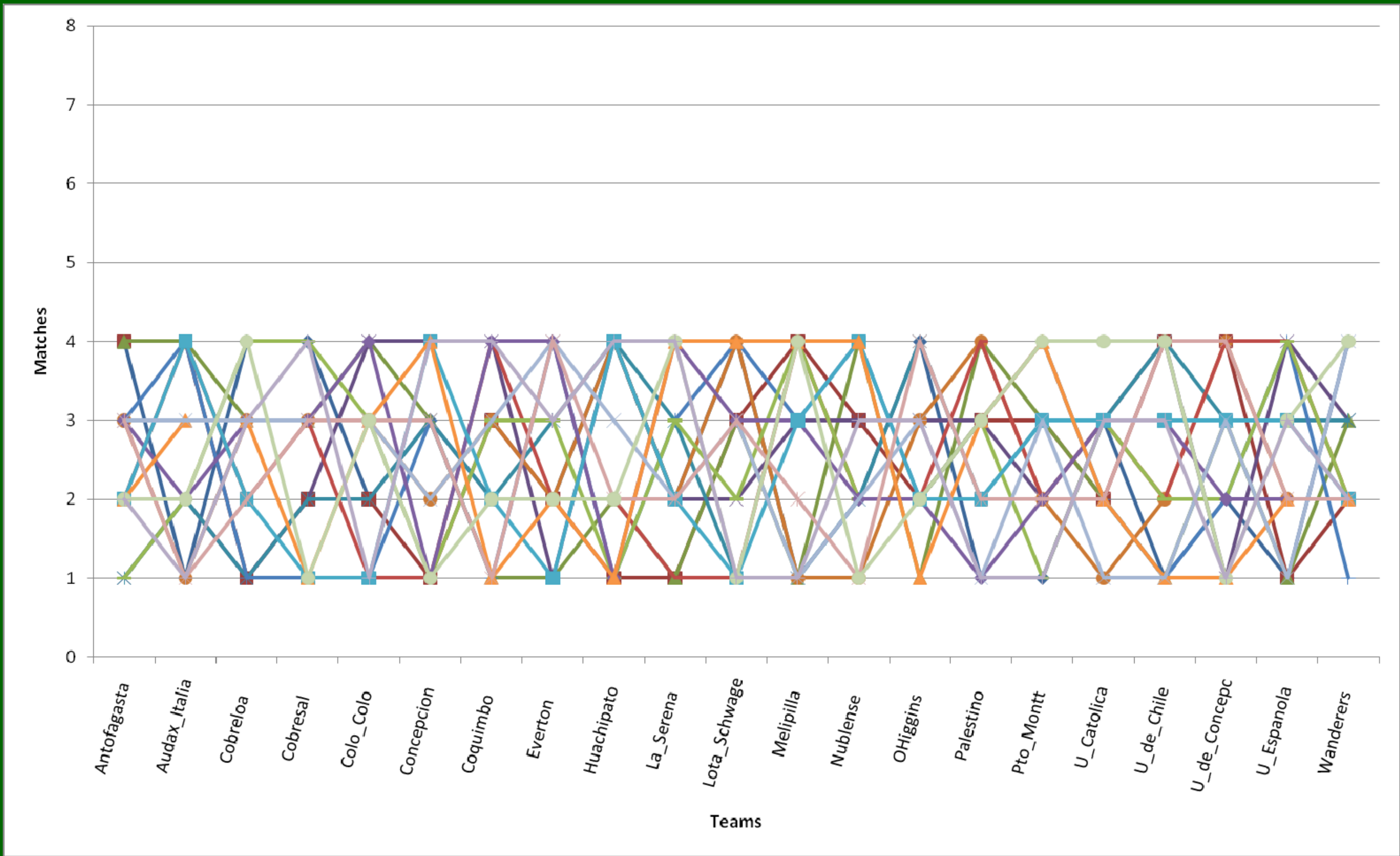
Another IP model that incorporates the rest of the conditions and assigns the referees to the games of the tournament, considering the patterns generated in the previous stage.

RESULTS

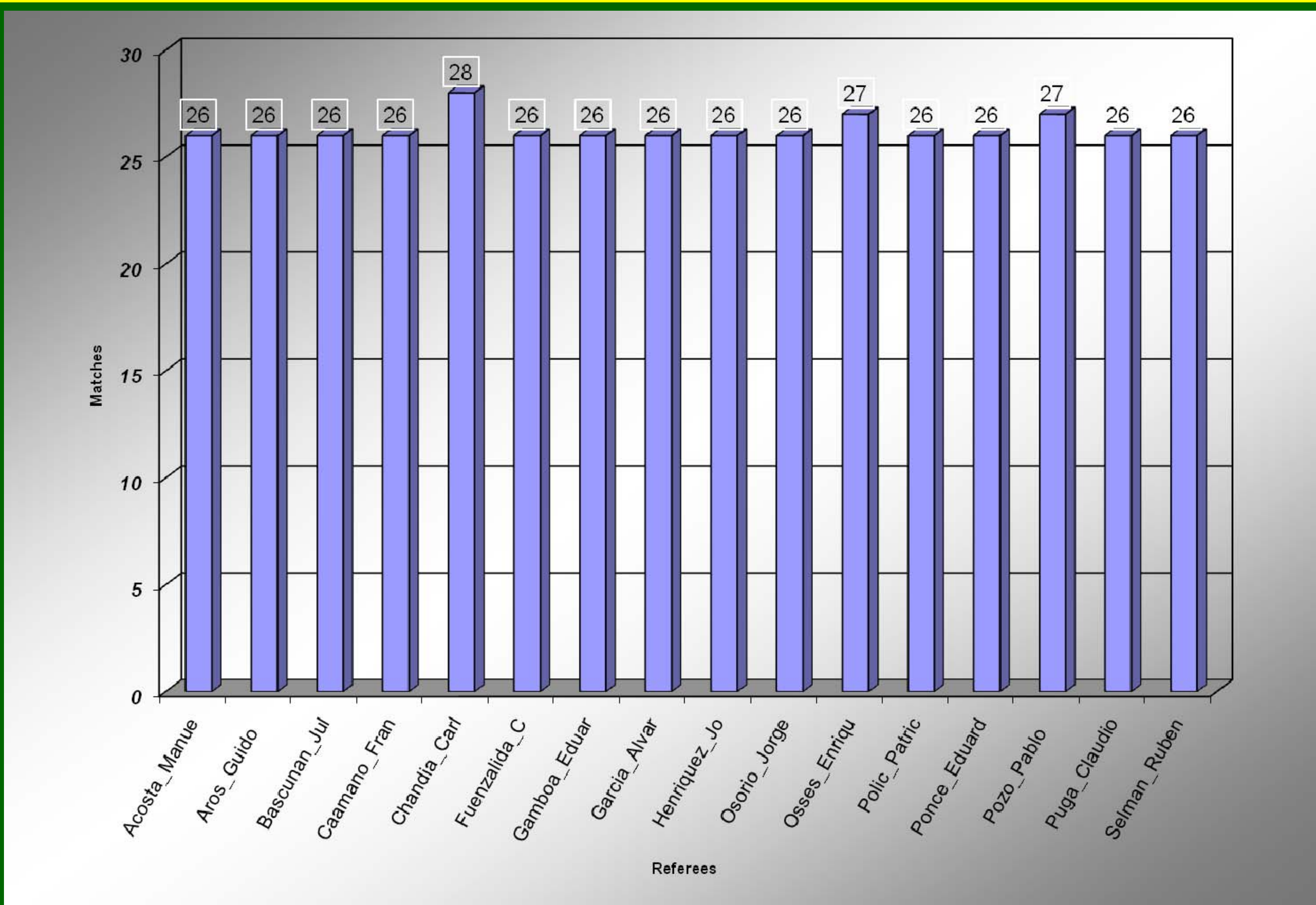
OR ASSIGNMENT vs TRADITIONAL

- Balance improvement
  - Frequency of referees vs teams
  - Total number of games assigned to each referee
  - Average travel distances of the referees

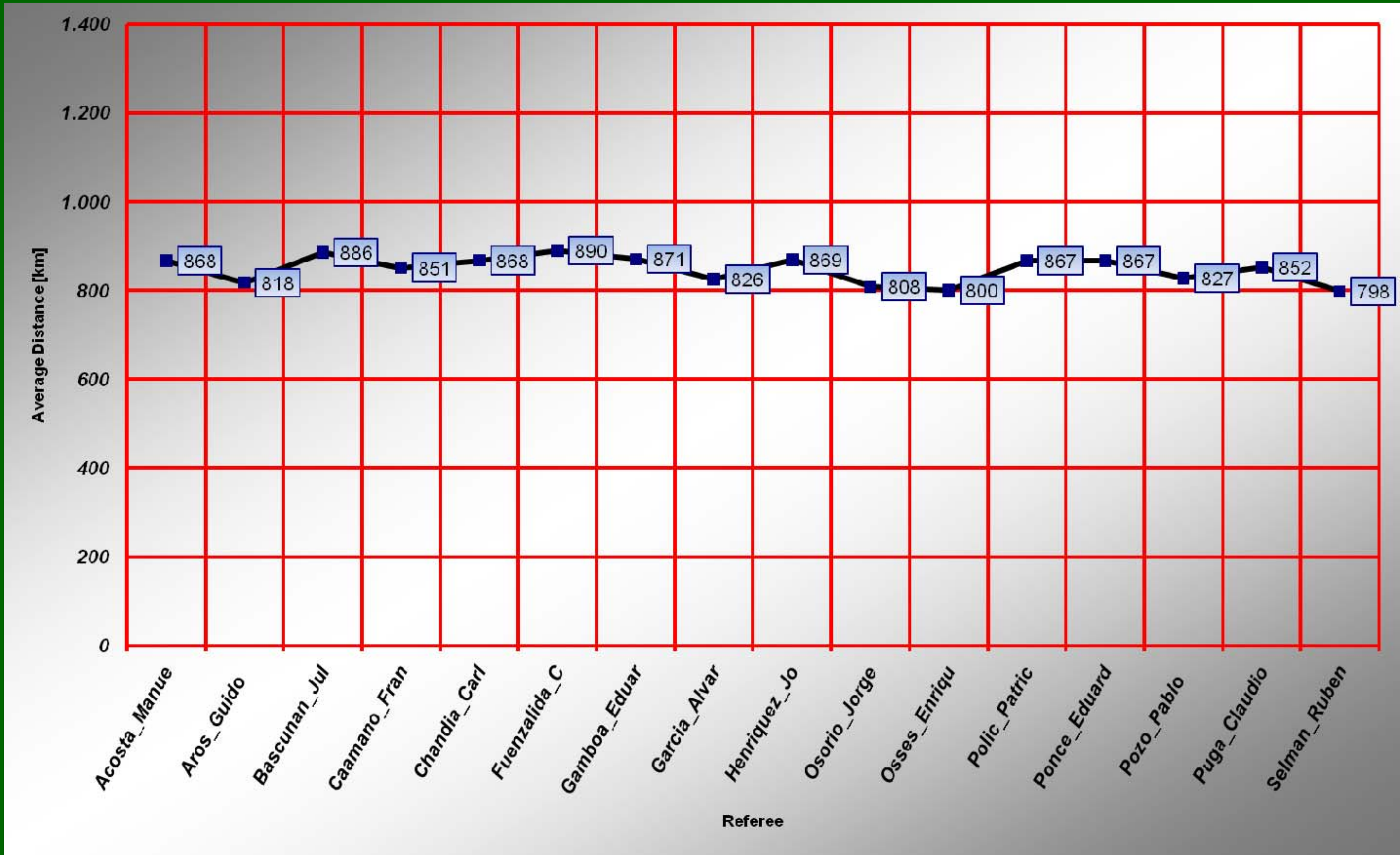
RESULTS: FREQUENCY OF REFEREES vs TEAMS



RESULTS: TOTAL NUMBER OF GAMES BY EACH REFEREE



RESULTS: AVERAGE TRAVEL DISTANCES



ARGENTINEAN BASKETBALL LEAGUE

An IP approach is currently being used by the league to assign referees in the 2016–17 season (Facu et al., work-in-progress)

$$\min: \overbrace{\sum_{i\in\mathbb{A}} \sum_{\substack{(s,t,k,m,n)\in\mathbb{P}\\ k>0\\ m>0}} v_{istkm}\cdot d_{km}}^{\text{viaje entre equipos}} + \overbrace{\sum_{i\in\mathbb{A}} \sum_{\substack{(s,t,k,m,n)\in\mathbb{P}\\ k>0\\ m=0}} v_{istkm}\cdot r_{ik}}^{\text{viajes que llegan a un domicilio}} +$$

$$+\underbrace{\sum_{i\in\mathbb{A}} \sum_{\substack{(s,t,k,m,n)\in\mathbb{P}\\ k=0\\ m>0}} v_{istkm}\cdot r_{im}}_{\text{viajes que salen de un domicilio}} + \underbrace{\sum_{i\in\mathbb{A}} \sum_{\substack{(s,t,k,m,n)\in\mathbb{P}\\ k>0\\ m>0\\ zA_i\neq zE_k}} s\cdot v_{istkm}\cdot c}_{\text{costo de hotel}}$$

Reduction in travel distances and costs ~ 25 to 30%.

QUALITATIVE BENEFITS

- Conditions made known to all concerned
  - ✓ Transparency
- All referees and teams are taken into account
  - ✓ Fairness
- Clearly defined criteria
  - ✓ Objectivity

IN AN *IDEAL* SETTING: REFEREE ASSIGMENT WOULD NOT BE AN ISSUE;

REALITY **IS DIFFERENT… AND TEAMS COMPLAIN…**

ALI BENNACEUR

“Y… no es tan clara” (*Willy* Durán)