
IMMC 2025
BATORY HS

OBJECTIVES

Generalization

OBJECTIVES

Sport

Teams

Scheduling

Generalization

Sport

Key aspects:

- Global fanbase
- Playing frequency
- Competition level across continents
- No dominant nations

Fan Volume Index (FVI)

Match Fairness Index (MFI)

Fan Volume Index (FVI)

$$\boldsymbol{Fan\ Volume\ Index\ (FVI)} = n \times f$$

Where:

- n – number of fans of the sport worldwide
- f – number of times per week the sport can be played

Match Fairness Index (MFI)

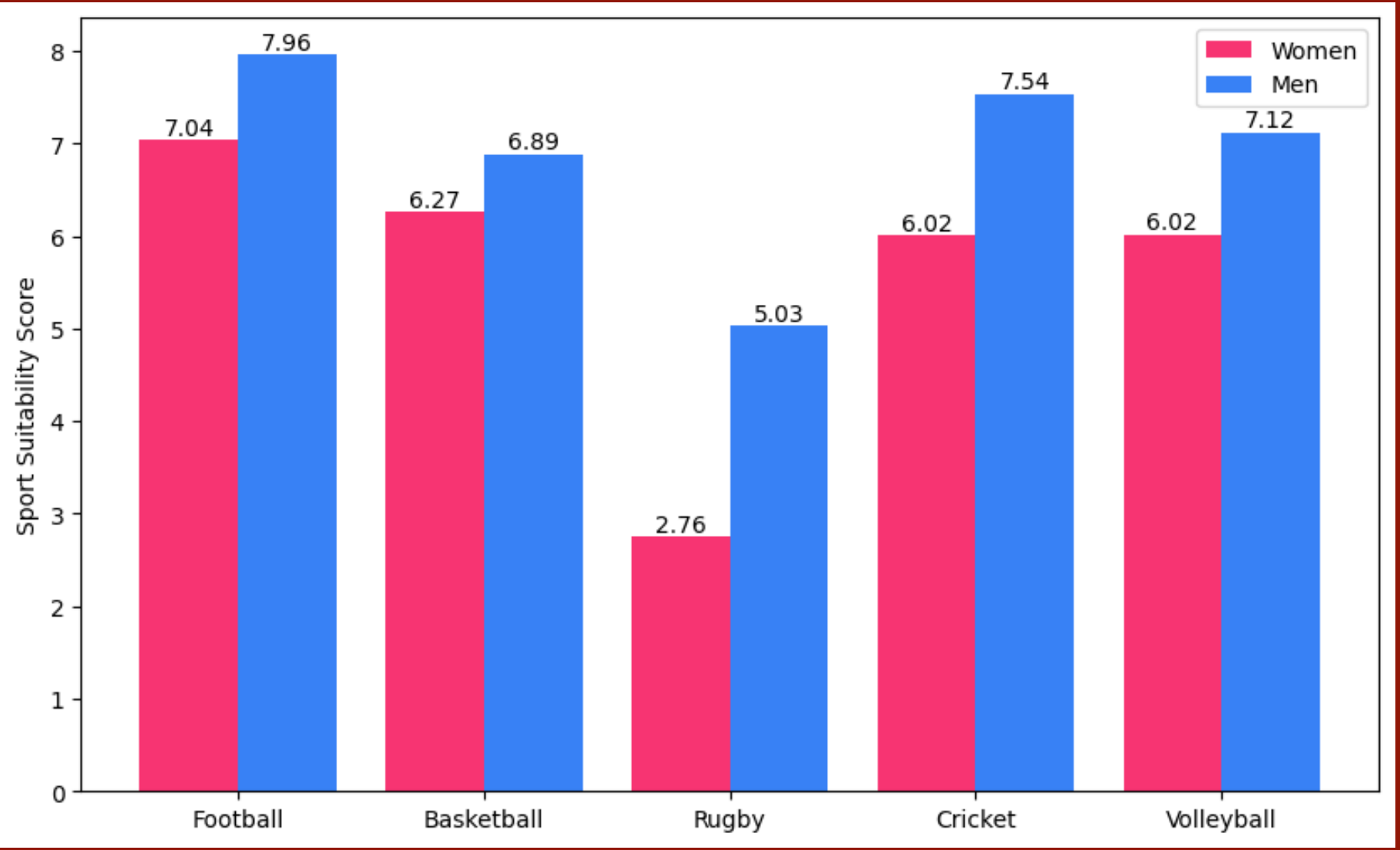
$$\boldsymbol{Match\ Fairness\ Index\ (MFI)} = \frac{v}{\sigma_{continents}}$$

Where:

- v– number of different winners in quadrennial world championships
- $\sigma_{continents}$ – standard deviation of the number of teams in the top 50 global ranking per continents

$$\boldsymbol{Sport\ Suitability\ Score\ (SSS)} = \ln(FVI \times MFI)$$

SSS Results


Women
Men
Sport Suitability Score
0
1
2
3
4
5
6
7
8
7.04
7.96
6.27
6.89
2.76
5.03
6.02
7.54
6.02
7.12
Football
Basketball
Rugby
Cricket
Volleyball

Sport

- Global fanbase
- Playing frequency
- Competition level across continents
- No dominant nations

Fan Volume Index (FVI)

Match Fairness Index (MFI)

OBJECTIVES

Sport

Teams

Scheduling

Generalization

Teams

ELO rating system

National league matches

International competition matches

Team win probability

$$E_{T1} = \frac{1}{1 + 10^{\frac{R_{T1} - R_{T2}}{400}}}$$

Where:

- E_{T1} - probability of Team 1 winning
- R_{T1} - ELO ranking of Team 1
- R_{T2} - ELO ranking of Team 2

Taking into account different league strengths

$$\boldsymbol{league\ factor\ (lf)} = \frac{s_{league}}{s_{benchmark}}$$

Where:

- s_{league} - league relative strength score
- $s_{benchmark}$ - lowest strenght score in the ranking

International competitions bonus

$$\boldsymbol{international\ factor\ (if)} = \ln\left(\frac{p_{competition}}{1000}\right)$$

Where:

- $p_{competition}$ - prize value of the competition

Chosen teams

- Real Madrid
- Bayern Munchen
- Manchester City
- Dinamo Zagreb
- Paris Saint-Germain
- Benfica
- SE Palmeiras
- RB Salzburg
- Juventus
- Al Ahly
- PSV Eindhoven
- Qarabak FK
- TP Mazembe
- LDU Quito
- Johor FC
- Ulsan Hyundai FC
- Auckland City FC
- UANL Tigres
- CD Olimpia
- Melbourne Heart FC

Teams

ELO rating system

National league matches

International competition matches

OBJECTIVES

Sport

Teams

Scheduling

Generalization

Scheduling

Schedule's goal

- Minimize Travel distance
- Minimize Travel distance inequalities

$$Minimize\ Cost = W_{Total} \cdot TotalTravel + W_{StdDev} \cdot StdDev(Travel)$$

- Constraints:
  - Play Everyone Twice (Double Round Robin)
  - One Game Per Week
  - Teams can't stay home or away forever

Financial analysis

Costs considered:

- **Hotel stays**
- **Emissions**
- **Flights**
- **Match organization**

$$Team + Environment\ (total) = 2{,}785.3364 \times d + 1{,}473{,}070 \times n + 1{,}180 \times w\ [USD]$$

Obtaining a DRR schedule

Step 1:

CP-SAT solver - a constraint programming tool (CP) that uses different satisfiability (SAT) techniques to obtain fitting results:

- **The two teams teams play in either of their respective cities**
- **Every team hosts a match with all the teams once a season**
- **Enforce the max consecutive home / away games – set to 3**

Step 2:

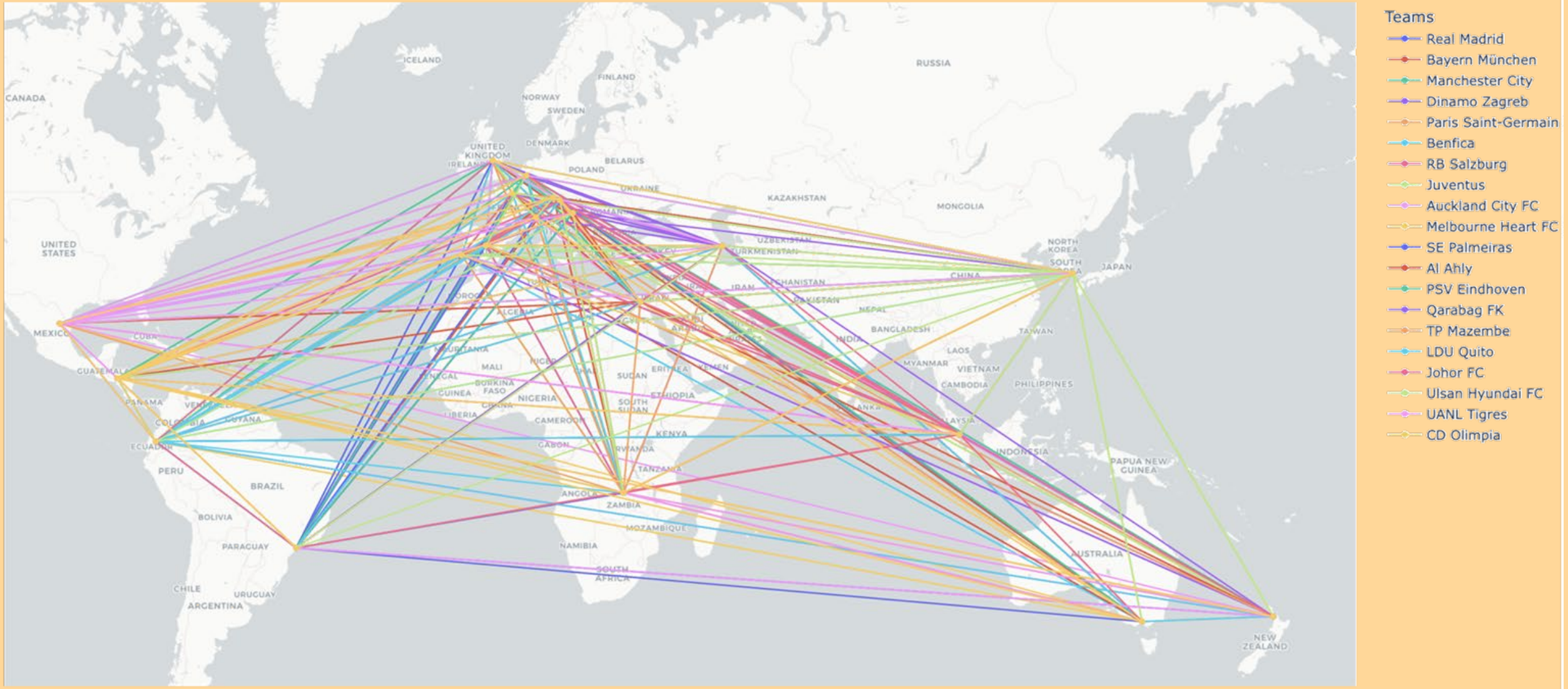
Simulated Annealing (SA) – a probabilistic technique used for finding the global optimum of a function

Optimization results

Metric	Initial Feasible Schedule	Optimized Schedule
Total Travel Distance (km)	4,935,726	4,501,332
Std. Dev. of Team Travel (km)	61,532	13,455
Spread of Team Travel (km)	244,367	42,645
Min Team Travel (km)	165,372	209,740
Max Team Travel (km)	409,739	252,385
Combined Cost (Eq. 10)	70,613,073	17,803,059

Table 1: Comparison of Initial (CP-SAT) and Optimized (SA) Schedule Travel Metrics

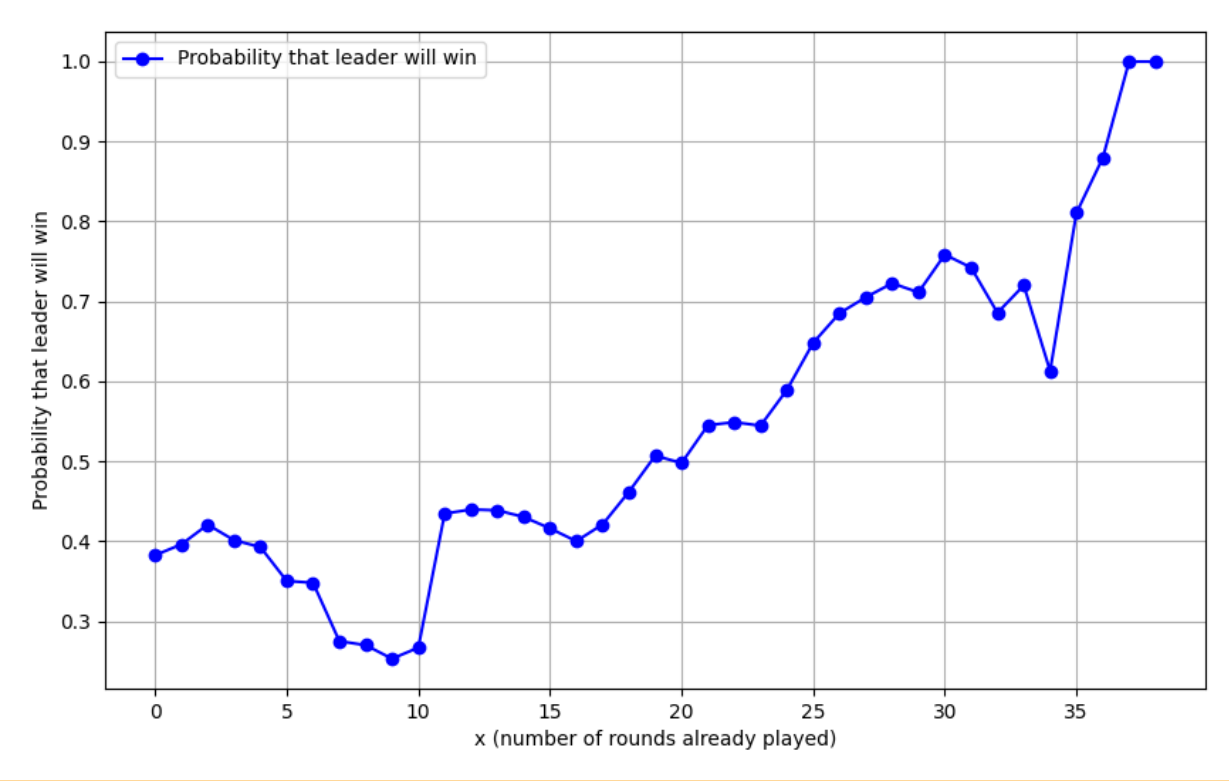
Travel path visualization



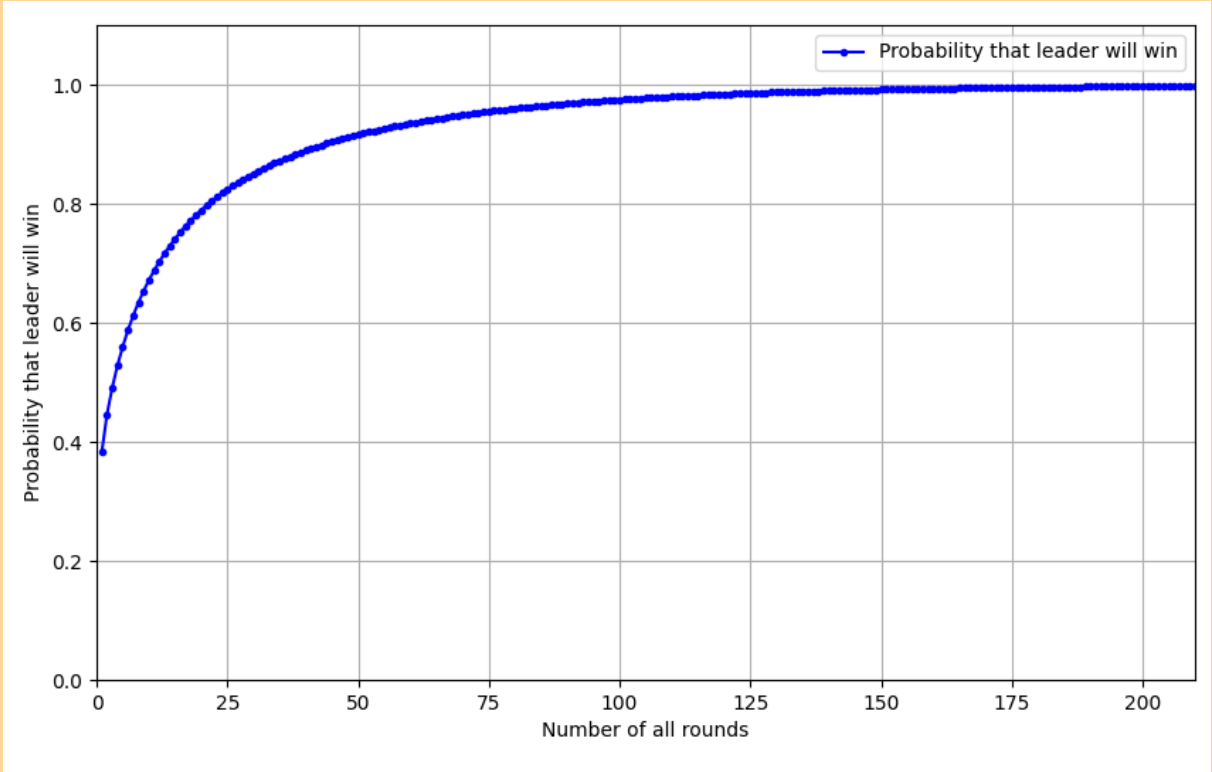
Is the win certain?

- **Assuming that the team's final score undergoes normal distribution we can calculate expected value of it's final score**
- **Using a cumulative distribution function we can predict robability that the team with the highest elo rating will win**

Probability that L will win after each round (10 000 simulations)



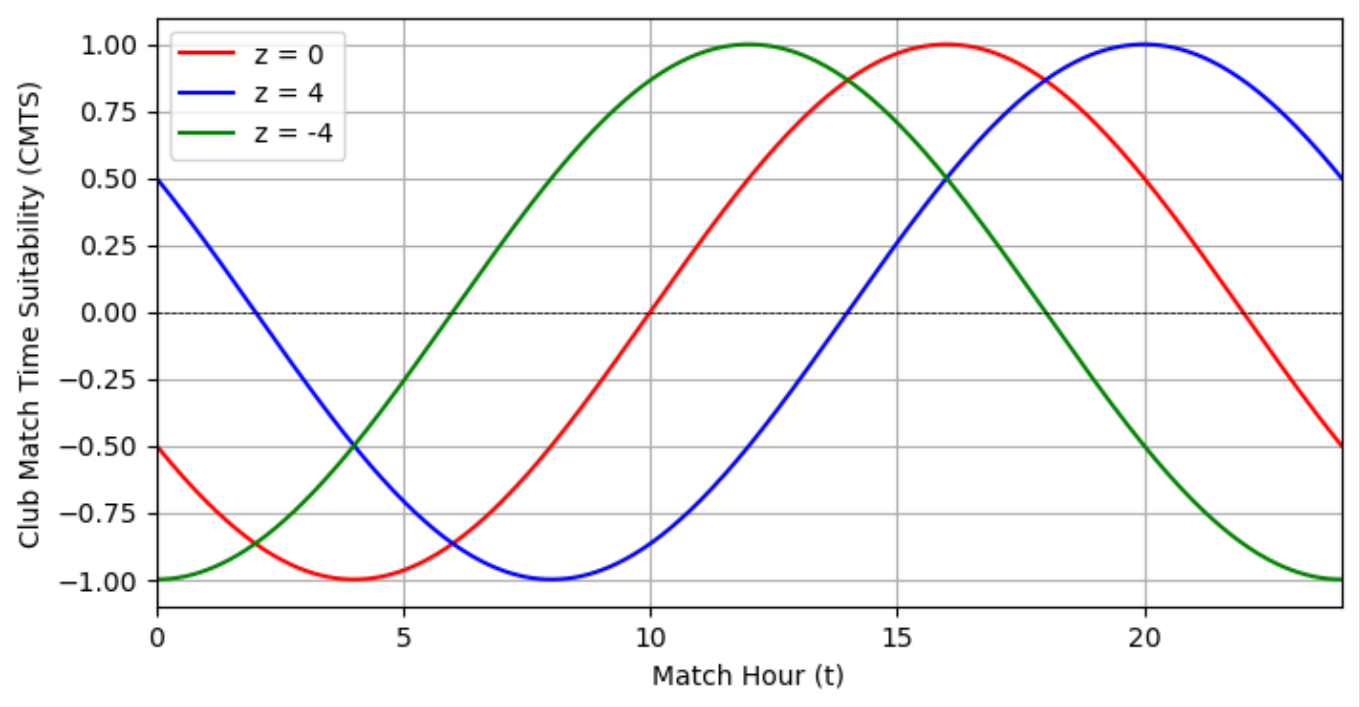
Probability that L will win in case of x numer of rounds



Generating match hours

- Human body can shift its biological clock by 1 hour a day
- Team gets to the city they play in 4 days before their match
- 2 teams playing experience the same level of discomfort

$$Club\ Match\ Time\ Suitability\ (CMTS) = -\cos\left(\frac{2\pi}{24} \times (t - 4 - z)\right), 0 \leq t < 24$$



OBJECTIVES

Sport

Teams

Scheduling

Generalization

Generalization

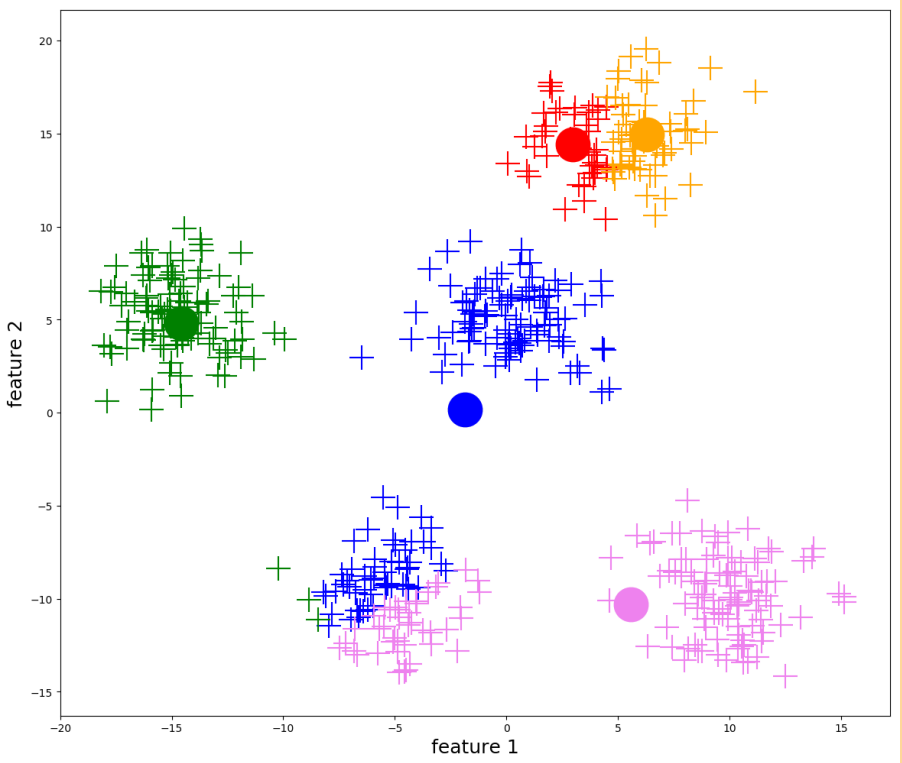
Extending the GSL

Added teams:

- **Shaktar Donetsk**
- **Mamelodi Sundowns**
- **Celtic FC**
- **Esperance Tunis**

Double Round Robin takes too much time

K-means algorithm



1. Random centroids

2. Assign points closest to centroids

3. Arithmetic mean of centroids create new centroid

4. Iterate steps 2 and 3

Our clustering algorithm

Key assumption: each cluster has 4 teams

$$\overrightarrow{v_i} = (x_i, y_i, z_i, w \times s_i)$$

Vector | Geography | Weight × score

Distance between cities i and j in 4D

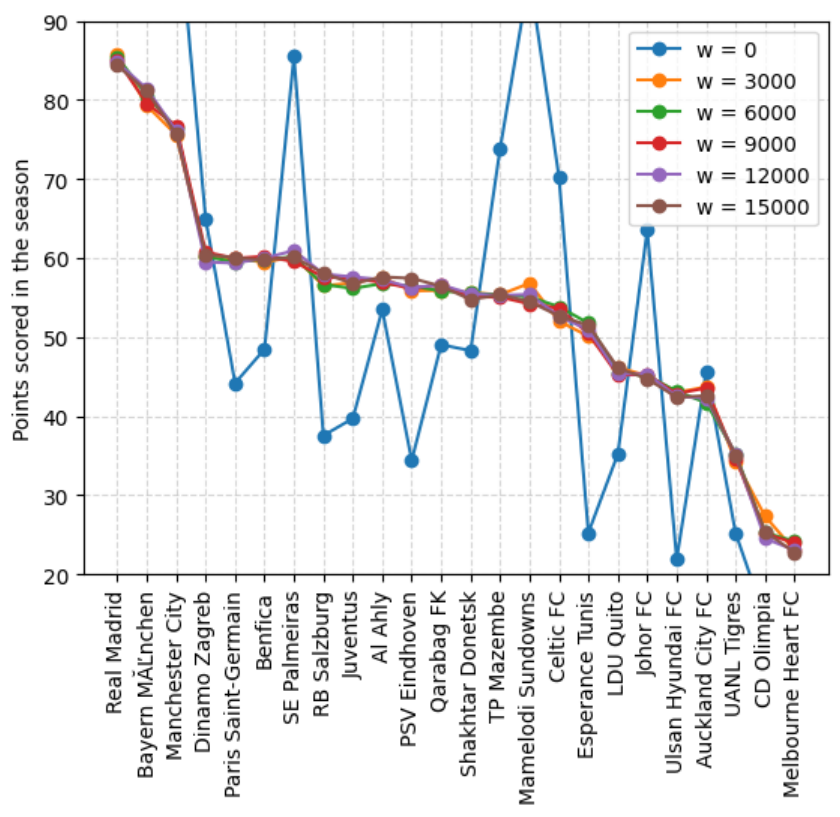
$$D(v_i, v_j) = \sqrt{\left(d\_geo(x_i, y_i, z_i, x_j, y_j, z_j)\right)^2 + w^2(s_i - s_j)^2}$$

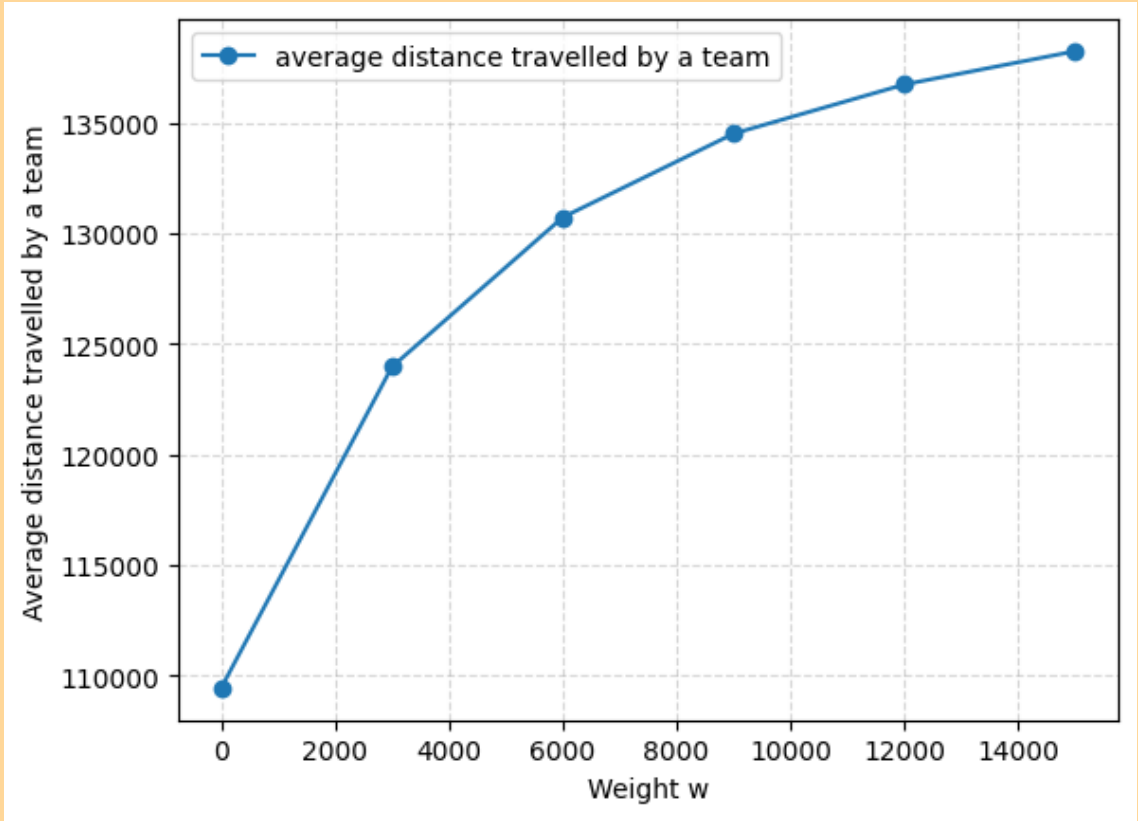
Where:

$$d\_geo(x_i, y_i, z_i, x_j, y_j, z_j) = 6371 \arccos\left(\frac{x_i x_j + y_i y_j + z_i z_j}{6371}\right)$$

Choosing the weight w

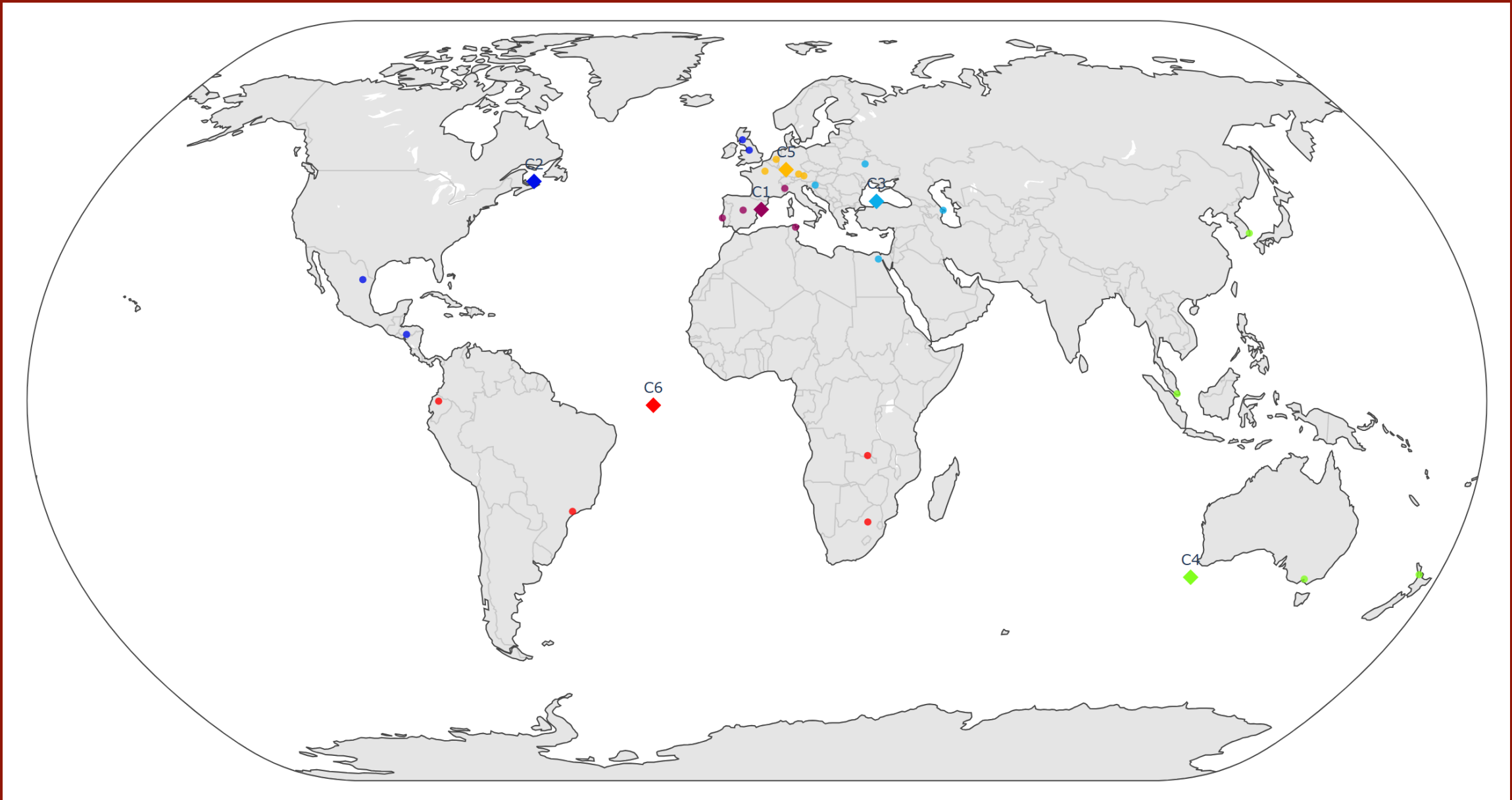
Trade-off between goals





w = 9000

Sample clusters



K – means results

Metric	Optimized Double Robin Schedule	Balanced K-means schedule
Total Travel Distance (km)	4,501,332	3,376,193
Distance per team (km)	225,067	140,674
Mean ELO difference	380	192
Number of match days	38	36
Number of teams (km)	20	24
Estimated total cost ($)	65 mln	70 mln.

Increase in cost is caused by an increase in numer of teams

OBJECTIVES

Sport

Teams

Scheduling

Generalization

Extending the model to other sports

- Our model can be extended to other zero-sum games – one's person loss is equivalent to another person's gain.
- Key requirements;
  - Sufficient datasets (historical games, league rakings, etc.),
  - Adjusting the ELO system according to the rules of the sport,
  - Ensuring that the sport fits all the key assumptions of the model


THANK YOU
FOR YOUR
ATTENTION