



Streaky Good Models

Incorporating Short-Run Trends into NBA
Predictions

Harry Jain
Advisors: James Glenn
& Kevin O'Neill

Win Predictions in the NBA

Elo Ratings

- Provides a mathematical metric for calculating win probability that is updated from each game result

$$E_A = \frac{1}{10^{\frac{-(R_A - R_B + Adj)}{400}} + 1} \quad R'_A = R_A + K(S_A - E_A)$$

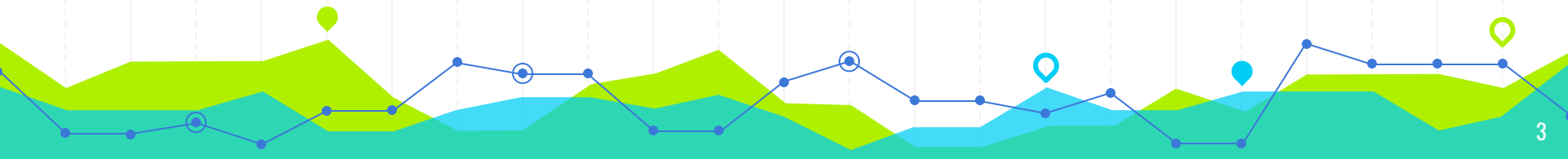
- Used by FiveThirtyEight

Adjusted Point Differentials

- Predicts average margin of victory
- Bayesian framework based on priors like offensive and defensive efficiency, schedule strength, preseason expectations, etc.
- Used in ESPN's BPI and Basketball Reference's SRS

Motivation

- While these aforementioned frameworks do a generally great job of predicting games, they are slow to react to team trends
 - E.g. injuries, changes in team chemistry, etc.
 - Based on the assumption that the expected performances of a team changes *slowly* over time
- Yet streaks are clearly very predictive of single-game results

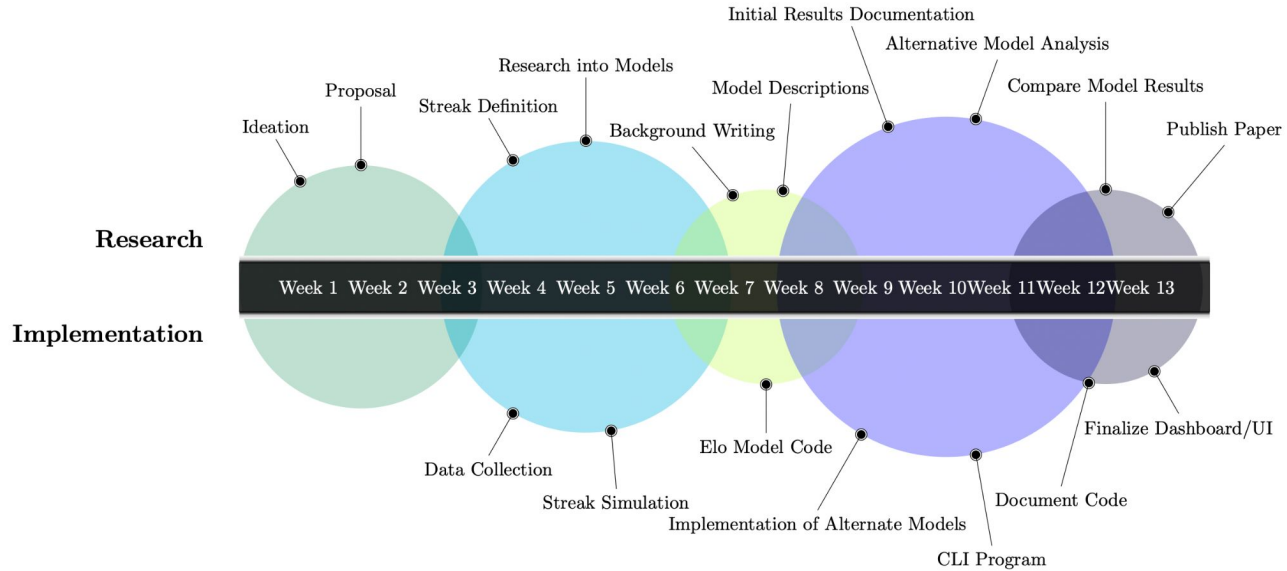


Outline of Project

1. **Collection of data** ✓
 - a. **Game and player data** ✓
2. **Analysis of streaks** ✓
 - a. **Compare distribution of streaks vs. “control ratings”** ✓
 - i. **Final Elo score from FiveThirtyEight** ✓
 - ii. Elo score from the first 20 games of the season
3. **Incorporating streaks into win predictions**
 - a. Update k factor of Elo score dynamically
 - i. Bayesian priors
 - ii. Logistic regressions
 - iii. Neural networks
 - b. Consider alternate models trained to be streak-sensitive

$$R'_A = R_A + K(S_A - E_A)$$

Timeline



Technology Stack and Deliverables



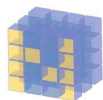
PostgreSQL

Database

Predictive Models



TensorFlow



NumPy

matplotlib



Pandas



Seaborn

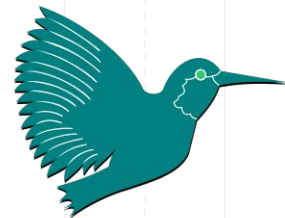


web development,
one drop at a time

Interactive
Dashboard



Written
Report



The L^AT_EX Project

Case Study: 2021-22 Phoenix Suns

- Question: Are streaks significant?
- Method: Simulate win/loss streaks under the assumption of constant Elo ratings
 - Monte Carlo simulation
- Metrics:
 - Mean length of the streaks
 - Distribution of streak length

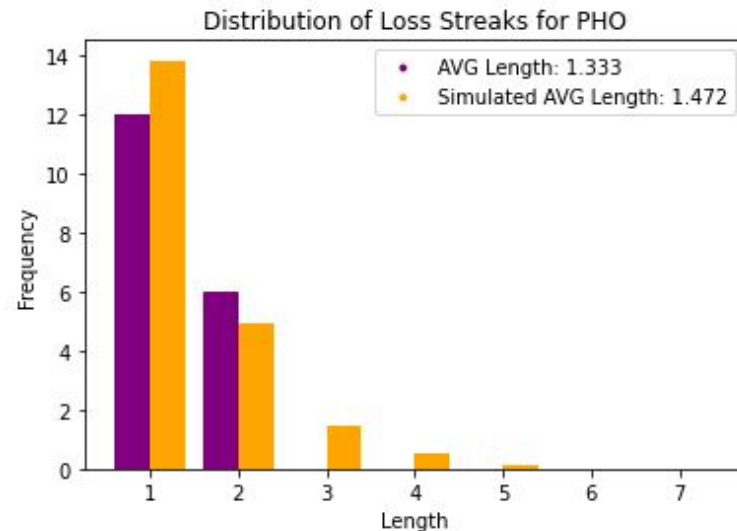
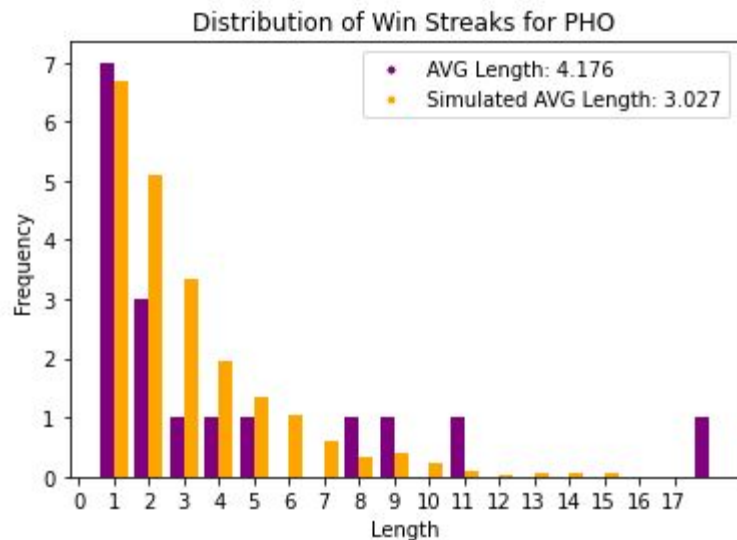


FiveThirtyEight's Elo Ratings

CURRENT RATING	FULL-STRENGTH RATING	TEAM	CONFERENCE	PLAYOFFS	CONF. SEMIS	CONF. FINALS	FINALS	CHANCE OF WINNING FINALS
1753	1773	Celtics 2-0	East	✓	94%	73%	55%	40%
1688	1707	Bucks 1-1	East	✓	87%	24%	14%	8%
1652	1703	Suns 1-1	West	✓	72%	48%	32%	15%
1607	1687	Nets 0-2	East	✓	6%	3%	1%	0.6%
1662	1686	Clippers	West	—	—	—	—	—
1669	1684	76ers 3-0	East	✓	99%	50%	14%	7%
1568	1683	Nuggets 0-3	West	✓	2%	0.5%	0.2%	<0.1%
1677	1677	Heat 2-0	East	✓	88%	46%	15%	7%
1681	1673	Warriors 3-0	West	✓	98%	58%	29%	11%
1545	1663	Mavericks 2-1	West	✓	60%	27%	13%	4%
1649	1649	Jazz 1-2	West	✓	40%	17%	7%	2%
1655	1648	Grizzlies 2-1	West	✓	81%	37%	15%	5%
1599	1622	Hawks 0-2	East	✓	12%	3%	0.6%	0.2%
1546	1604	Pelicans 1-1	West	✓	28%	8%	2%	0.4%
1586	1586	Timberwolves 1-2	West	✓	19%	4%	1%	0.1%
1576	1576	Raptors 0-3	East	✓	1%	0.4%	<0.1%	<0.1%
1406	1575	Lakers	West	—	—	—	—	—
1191	1557	Trail Blazers	West	—	—	—	—	—
1494	1550	Bulls 1-1	East	✓	13%	0.3%	<0.1%	<0.1%

CURRENT RATING	FULL-STRENGTH RATING	TEAM	CONFERENCE	PLAYOFFS	CONF. SEMIS	CONF. FINALS	FINALS	CHANCE OF WINNING FINALS
1771	1771	Celtics 4-0, 4-3, 4-3, 2-4	East	✓	✓	✓	✓	—
1711	1711	Bucks 4-1, 3-4	East	✓	✓	—	—	—
1693	1693	Heat 4-1, 4-2, 3-4	East	✓	✓	✓	—	—
1693	1688	Nets 0-4	East	✓	—	—	—	—
1691	1685	Suns 4-2, 3-4	West	✓	✓	—	—	—
1682	1682	Clippers	West	—	—	—	—	—
1664	1677	76ers 4-2, 2-4	East	✓	✓	—	—	—
1604	1676	Nuggets 1-4	West	✓	—	—	—	—
1704	1673	Warriors 4-1, 4-2, 4-1, 4-2	West	✓	✓	✓	✓	✓
1651	1651	Mavericks 4-2, 4-3, 1-4	West	✓	✓	✓	—	—
1637	1637	Grizzlies 4-2, 2-4	West	✓	✓	—	—	—
1636	1636	Jazz 2-4	West	✓	—	—	—	—
1624	1624	Hawks 1-4	East	✓	—	—	—	—
1621	1612	Pelicans 2-4	West	✓	—	—	—	—
1599	1599	Raptors 2-4	East	✓	—	—	—	—
1582	1582	Timberwolves 2-4	West	✓	—	—	—	—
1408	1576	Lakers	West	—	—	—	—	—
1192	1551	Trail Blazers	West	—	—	—	—	—
1474	1535	Bulls 1-4	East	✓	—	—	—	—

Case Study: 2021-22 Phoenix Suns



Case Study: 2021-22 Phoenix Suns

- Difficulties
 - Lots of moving parts \Rightarrow no clear “baseline” for comparison
 - Elo scores, team trends, opponent trends
 - Different significance for different teams
 - Some teams are more “streaky”
 - Lower performing teams have larger losing streaks
- Conclusions
 - Ratings have to be dynamic
 - Longer streaks are more likely to continue than their statistic probability



Bibliography

- [1] NBA Logo. url: <https://www.nba.com>.
Data Collection Elo Model Code Streak Simulation
Implementation of Alternate Models
- [2] Cal Ripken Quote. AZ Quotes. url: <https://www.azquotes.com/quote/712065>.
- [3] Bethlehem Shoals. The 25 Most Unbreakable Streaks in Sports. GQ. url: <https://www.gq.com/gallery/greatest-sports-streaks>.
- [4] Raghav Mittal. What is an ELO Rating? Medium. url: <https://medium.com/purple-theory/what-is-elo-rating-c4eb7a9061e0>.
- [5] Nate Silver and Reuben Fischer-Baum. How We Calculate NBA Elo Ratings. FiveThirtyEight. url: <https://fivethirtyeight.com/features/how-we-calculate-nba-elo-ratings/>.
- [6] How Our NFL Predictions Work. FiveThirtyEight. url: <https://fivethirtyeight.com/methodology/how-our-nfl-predictions-work/>.
- [7] ESPN Analytics. ESPN's Basketball Power Index. ESPN. url: https://www.espn.com/nba/story/_/page/Basketball-Power-Index/espn-nba-basketball-power-index.
- [8] Mike Lynch. SRS Calculation Details. Sports Reference. url: <https://www.sports-reference.com/blog/2015/03/srs-calculation-details/>.
- [9] Sharon Katz. How ESPN's NFL Football Power Index was developed. ESPN. url: https://www.espn.com/nfl/story/_/id/13539941/how-espn-nfl-football-power-index-was-developed-implemented.
- [10] Sascha Wilkens. "Sports prediction and betting models in the Machine Learning Age: The case of tennis". In: Journal of Sports Analytics 7.2 (2021), pp. 99–117. doi: 10.3233/jsa-200463.

THANKS!

Any questions?

You can also email:
harry.jain@yale.edu



bit.ly/3Fd3wIU