

2024 ATP Season Analysis & Simulation Report

1. Introduction

This report summarizes the analytical findings extracted from the cleaned **ATP 2024 match dataset** and the outcomes generated using the custom **Elo-based match simulator** developed in this project.

Combined, these methods provide a comprehensive view of player performance, surface specialization, ranking stability, and simulated season outcomes.

The analysis integrates:

- Actual 2024 match results
- Elo progression throughout the season
- Surface-specific win indicators
- Model-based season simulations
- Probabilistic tournament outcome distributions

2. Data Overview

The dataset ([atp_matches_2024.csv](#)) contains all ATP-level matches played in 2024, cleaned and standardized through preprocessing pipeline.

The data includes:

- Player names and match identifiers
- Match winner/loser fields
- Serve percentage metrics
- Break-point statistics

- Surface type
- Match-level Elo for both players
- Tournament category and round indicators

After cleaning, the dataset supports:

- Elo calculations
- Surface-specific performance comparisons
- Outcome simulations using season engine

3. Player Performance Metrics

3.1 Overall Performance Leaders

Across the full 2024 dataset, three players consistently dominated statistical and simulation models:

1. Novak Djokovic

- Highest computed Elo rating
- Best win/loss efficiency
- Top performer in serve reliability and break conversion
- Lowest error rate among top-ranked players
- Most stable rating trajectory over the season

2. Carlos Alcaraz

- Slightly fewer wins than Djokovic but with higher peak performance moments
- Strongest clay-court performance index
- Highest volatility range (largest positive and negative Elo deltas)

- High upside potential in the simulation engine

3. Jannik Sinner

- One of the strongest hard-court performers
- Most consistent rating curve
- Top “overperformance vs higher-ranked opponent” indicator

These three players formed a statistically identifiable top tier.

4. Surface-Level Insights

4.1 Hard Courts

- Highest average Elo gains across players
- Most consistent match outcome patterns
- Strongest correlation between pre-match Elo and actual outcome
- Djokovic and Sinner dominated this surface

4.2 Clay Courts

- Highest volatility and largest spread in performance
- Strong specialization effect: certain players performed disproportionately better
- Alcaraz displayed dominant clay patterns
- Several mid-ranked players were consistently “clay overperformers”

4.3 Grass Courts

- Smallest number of data points → highest simulation uncertainty

- Elo changes were more abrupt due to fewer matches
 - Upset probability remained the highest of all surfaces
-

5. Elo Dynamics

The Elo progression throughout the season revealed:

- Djokovic maintained the most stable rating trajectory
- Alcaraz showed the steepest up/down cycles
- Sinner exhibited minimal variance, suggesting high reliability
- Surfaces influenced Elo movement far more dramatically in clay and grass seasons

The Elo system also aligned strongly with simulation results, confirming model reliability.

6. Season Simulation Engine

The season simulation engine executed **100 independent season runs**, each progressing tournament by tournament using probabilistic Elo-based outcome predictions.

6.1 Average Finish Positions

Across all 100 simulations, the average expected finishing order was:

1. **Novak Djokovic**
2. **Carlos Alcaraz**
3. **Jannik Sinner**

Djokovic consistently generated the highest expected title count and the highest simulated end-of-year Elo.

6.2 Title Distributions

- Djokovic: most simulated tournament wins
- Alcaraz: strong title count, especially on clay
- Sinner: frequent deep runs and stable top-3 finishes
- Other players occasionally broke into the top 5 but rarely sustained it across simulations

6.3 Rating Volatility Across Simulations

- Djokovic: minimal rating drop between simulations
 - Alcaraz: widest range of possible outcomes due to style volatility
 - Sinner: narrow distribution, consistently strong
-

7. Comparative Interpretation

7.1 Dominance Indicators

Djokovic's combination of consistency, serve quality, and low unforced-error rates created the strongest statistical profile in the dataset.

Even in simulations, where randomness exists, he maintained top placement.

7.2 Momentum and Breakout Signals

Alcaraz's seasonal peaks suggest that he remains highly likely to challenge for the top spot, particularly in clay-heavy stretches.

His model behavior indicates the highest upside but also higher risk.

7.3 Reliability & Stability

Sinner's season displayed exceptional reliability across surfaces.

The simulator repeatedly placed him in late tournament rounds due to his low Elo variance.

7.4 Surface Specialists

The clay season revealed the largest group of surface specialists, while hard courts favored well-rounded players with strong serve metrics.

8. Key Quantitative Findings

- Djokovic holds the **highest average Elo** and the **lowest simulation variance**, confirming his statistical dominance.
- Alcaraz shows the **largest Elo swings** and the **highest clay-court win probability**.
- Sinner is the **most consistent performer**, rarely dropping out of the simulated top three.
- Grass remains the most unpredictable surface in both real and simulated data.
- Hard courts drive the largest share of Elo progression due to season structure.

9. Conclusion

The combined 2024 ATP dataset and Elo-based simulation framework confirm a clear performance hierarchy led by Djokovic, followed by Alcaraz and Sinner.

Real-world data and modeled outcomes align tightly, reinforcing that the top tier of players is well-separated from the rest of the field.