Modeling Rotational Efficiency in Elite Men's Volleyball Using a Probabilistic Rally Simulation Framework

Olivia Czech Marzouk

All data, CSVs, plots, simulation scripts, and other materials used—or ultimately unused—in this project are available in the accompanying GitHub repository: https://github.com/oliviacm08/volleyball-rotation-analysis

Introduction

Rotational efficiency, which refers to the effectiveness of a team in scoring or securing side-outs from each of its six rotations, is one of the most influential yet understudied aspects of men's volleyball. By analyzing rotations, coaches can identify favorable matchups and formations to optimize scoring opportunities. Traditional evaluations rely on match film, manual coding, or set-based summaries, but these observations are often shaped by opponent-specific factors, short-term streaks, and limited sample sizes. Even at the highest levels, a rotation may appear strong or weak due to temporary mismatches rather than consistent performance trends. Simulations offer an alternative approach: by modeling rallies probabilistically, rotation-level tendencies can be explored across thousands of hypothetical scenarios, independent of opponent-specific or situational noise. This project constructs a mathematically grounded rally simulator, validates it against real match data, and uses it to assess rotation efficiency, providing insights that complement conventional performance analyses. Notably, standard rotation efficiency often emphasizes attack point performance, which could make a rotation with a strong scorer appear highly effective even if its serving performance is suboptimal. In contrast, this project defines rotation efficiency more broadly, focusing on all first-touch and baseline actions to evaluate the overall structural effectiveness of each rotation.

Data Collection and Simulation Rationale

To identify rotation-level patterns and assess efficiency across elite men's volleyball, it is first necessary to establish a comprehensive match database. The primary dataset for this project consists of rally-by-rally recordings from top international teams. Specifically, seven matches were analyzed from varying international events and competitions recognized by the International Volleyball Federation (FIVB), totaling over 1,300 rallies. Furthermore, to guarantee match-level and skill consistency, every team appearing in the sampled matches has been listed, in accordance with the official FIVB men's volleyball world rankings, in the top 10 over the course of the last five years. For each rally, detailed information that was considered possibly relevant for the assessment was collected, including the serving team, receiving team, rotation, serve type, serve result, serve zone, pass quality, and the outcome of the rally. A mismatch check, comparing serving sequences and rally outcomes, was also performed to corroborate the validity of the dataset.

While real match data provides an empirical foundation for analyzing rotation-level performance, it is inherently constrained by sample size, opponent-specific dynamics, and short-term variability. Directly calculating probabilities from these matches could offer some insight, but such estimates are limited to the specific conditions under which the observed rallies occurred. A simulation framework overcomes these limitations by generating a much larger number of hypothetical rallies and sets, allowing rotation-level tendencies to emerge independently of temporary streaks or matchup-specific effects. By iteratively modeling rallies based on conditional probabilities derived from the real data, simulations enable robust and generalizable conclusions about rotation performance, offering a controlled environment in which the structural effectiveness of each rotation can be examined systematically.

First Attempted Simulation (Full-State Model)

To implement this simulation, the project initially adopted a probabilistic framework inspired by Markov chains, in which the state of a rally—initially defined by the serving team, current rotation, and all recorded metrics—determined the likelihood of the next outcome. In this fully specified "full-state" model, every variable recorded in the raw dataset (serve type, serve zone, pass quality, serve result, etc.) could theoretically serve as part of the state, allowing the simulator to generate rallies by repeatedly sampling outcomes based on observed conditional probabilities.

Mathematically, the full-state approach can be expressed as a discrete-time Markov chain with a state space

$$\mathcal{X} = \{\text{serving team}\} \times \{\text{rotation}\} \times \{\text{serve type}\} \times \{\text{serve zone}\} \times \{\text{pass quality}\} \times \{\text{serve result}\} \times \ldots$$

and transition matrix

$$\mathbf{P}_{ij} = \mathbb{P}(X_{t+1} = x_j \mid X_t = x_i), \quad x_i, x_j \in \mathcal{X}.$$

Sampling from $\boldsymbol{P}$ generates rallies consistent with the empirical conditional probabilities. However, because many metric combinations are rare or unobserved, $\boldsymbol{P}$ is extremely sparse and high-dimensional, with most entries either zero or based on very few observations. This sparsity produces two major issues: unrealistic transitions—where rare or zero-probability states are either artificially inflated during sampling or skipped entirely, leading to rally sequences that do not reflect real volleyball dynamics—and statistical instability—where high-dimensional sparsity amplifies sampling noise, causing simulated distributions of pass quality, serve zones, and serve results to diverge sharply from real match data. For instance, a team serving in rotation 4—the rotation where the opposite hitter typically serves—was sometimes assigned a FLOAT serve, which is nearly absent in the raw data.

Thus, comparing the simulated rallies with real match data confirmed that, while computationally correct, the full-state approach did not reliably reproduce realistic patterns, emphasizing the necessity of selecting and adapting metrics carefully within the simulation framework.

Metrics

To ensure the simulation accurately reflected rotation-level performance, two primary sets of metrics were incorporated: rotation win probabilities and rotation-specific serve type probabilities.

For a given team T in rotation R, the rotation win probability was defined as:

$$P_{\text{win}}(T, R) = \frac{\text{Points won by team T while serving in rotation R}}{\text{Total rallies served by team T in rotation R}}$$

Similarly, the probability of a particular serve type S (FLOAT, JUMP, or HYBRID) for team T in rotation R was computed as:

$$P_{\text{serve}}(T, R, S) = \frac{\text{Number of serves of type } S \text{ by team } T \text{ in rotation } R}{\text{Total serves by team } T \text{ in rotation } R}.$$

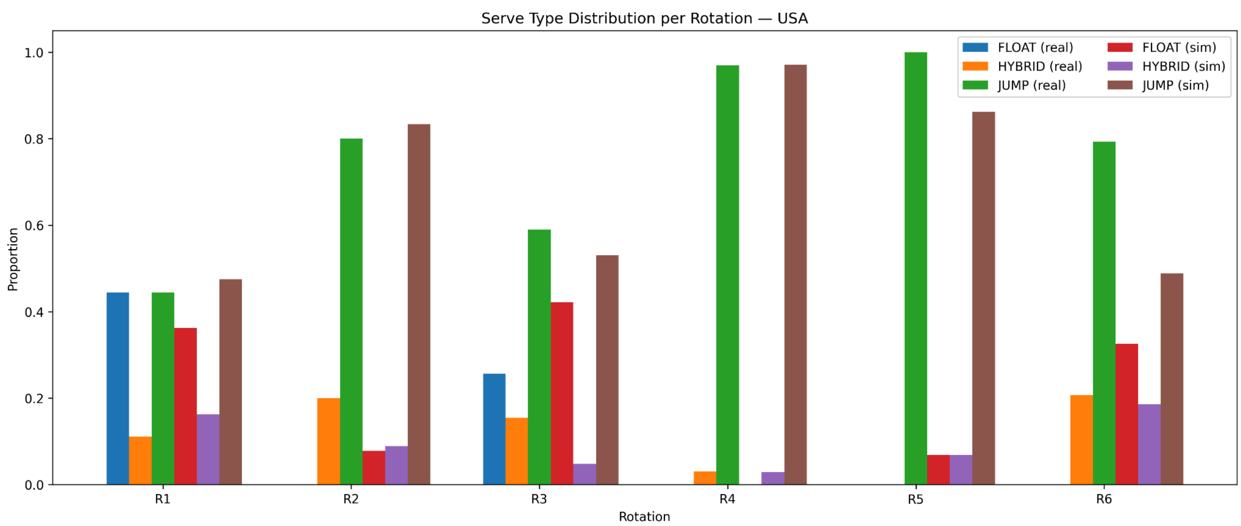
These metrics were used to generate pre-sampled sequences of rally states. During the simulation, the next rally state was drawn according to the current rotation and the corresponding pre-generated serve type sequence, creating a semi-Markov framework in which transitions were conditioned on both rotation and observed serving behavior.

To assess the fidelity of the simulation, the rotation efficiency comparison CSV was examined.

rotation	Efficiency_sim	Efficiency_real
1	0.3375	0.354978354978355
2	0.3266666666666666	0.30493273542600896
3	0.3175182481751825	0.3305084745762712
4	0.3095238095238095	0.3125
5	0.2664835164835165	0.2850678733031674
6	0.335149863760218	0.29357798165137616

Simulated efficiencies closely matched those observed in real matches for the set as a whole. For example, in rotation 1, the simulated efficiency was 0.338 versus 0.355 in real data, while in rotation 3, the simulated value was 0.317 versus 0.331 in observed rallies. Maximum absolute differences across all rotations were modest, confirming that rotation-level performance patterns were preserved.

Serve type distributions were similarly compared using the serve type distribution per team per rotation CSV. For instance, Brazil's R2 JUMP serve probability shifted slightly from 0.857 (real) to 0.894 (simulated), while Italy's R4 demonstrated perfect alignment with all simulated serves matching the observed type. The largest observed discrepancy was 0.621 for FRA R5 (see CSV), reflecting a rare deviation likely due to limited sample size, but general patterns remain well-reproduced. A representative serve type distribution plot for the United States is provided, illustrating the close alignment between simulated and observed proportions across rotations.



These comparisons indicate that the pre-generated serve sequences effectively captured rotation-specific serving behavior and preserved realistic variability across rotations.

* Complete rotation-specific serve probabilities for all teams and rotations are provided in the supplementary material or available at the project repository, allowing full transparency without overloading the main text.

In addition to rotation win probabilities and rotation-specific serve types, several other variables were initially considered for inclusion in the simulation, namely pass quality, serve zone, and serve result. These metrics were recorded in the raw data due to their potential influence on rally outcomes and were mathematically formalized as possible states within the simulation framework. However, even when incorporated into the semi-Markov model, they produced distributions that diverged substantially from observed match patterns. Despite technically correct implementation, the simulated rallies frequently generated unrealistic sequences—such as improbable combinations of serve zones and pass qualities—highlighting the limitations of these metrics for probabilistic modeling. Further attempts to apply these metrics successfully were also explored outside the revised model, ultimately leading to the conclusion that these discrepancies are partly due to their inherent subjectivity, high player-specific variability, and sensitivity to small sample sizes, confirming that including them would compromise the accuracy of the simulation as a whole, regardless of implementation approach.

Simulation Methodology and Execution

With these two carefully selected metrics, the simulation is now positioned to generate realistic rally sequences, providing a solid foundation for the subsequent modeling and analysis stages of the project. Building on this foundation, the next step was to formalize the rally-generation process itself. The simulator was designed to operate at the rally level, beginning each sequence with a designated serving team and rotation. For every rally, the model first referenced the pre-sampled serve type sequence associated with that team–rotation pair, ensuring that serving behavior remained consistent with observed tendencies. Once the serve type was assigned, the outcome of the rally was determined by sampling from the corresponding rotation win probability distribution, which governed whether the serving or receiving team secured the point. After each rally, rotations were updated according to standard volleyball rotation rules, and the process repeated until a full set was generated.

To be precise, the rally generator operates as a discrete-time stochastic process, approximating a semi-Markov model in which the state at rally t is defined as:

$$X_t = (\text{serving team}, \text{rotation}).$$

Unlike a full-state Markov chain—which would incorporate pass quality, serve zone, and other micro-actions—the state space is deliberately restricted to rotation-level descriptors to ensure statistical stability and avoid sparsity-related artifacts. Other recorded metrics, while not part of the state, inform the model probabilistically through the conditional distributions defined by the two selected metrics.

At each rally, the outcome is determined by sampling a Bernoulli variable:

$$Y_t \sim \text{Bernoulli}(p_{T,R}),$$

where $p_{T,R}$ is the empirically observed rotation win probability for the serving team T in rotation R. If $Y_t = 1$, the serving team wins the rally and retains serve; if $Y_t = 0$, the receiving team wins, becomes the new serving team, and its rotation index is incremented modulo 6:

$$R' = (R + 1) mod\ 6,$$

with the modulo mapped to the standard rotation labels $\{R1, ...,R6\}$ The resulting transition kernel is:

$$X_{t+1} = \begin{cases} (T, R), & Y_t = 1, \\ (T, R), & Y_t = 0, \end{cases}$$

where T denotes the opposing team.

To maintain realistic serving behavior, serve types are pre-generated for each team–rotation pair. For a given team T and rotation R, a sequence

$$S_{T,R} = (s_1, s_2, ..., s_n), \quad n \geq 100$$

is constructed by drawing independently from the empirical serve type distribution $\mathbb{P}(s \mid T, R)$, with $s \in \{JUMP,\ FLOAT,\ HYBRID\}$. Rally t then uses the t-th element of this sequence modulo n:

$$s_t = S_{T,R}[\,t \bmod n\,].$$

This approach produces a deterministic-but-randomized "serve tape" for each rotation, preventing unrealistic fluctuations while preserving probabilistic consistency.

Simulated sets are generated iteratively, rally by rally, until standard volleyball conditions are met: the first team to reach 25 points with at least a two-point margin wins the set. For every rally, the simulator records the tuple:

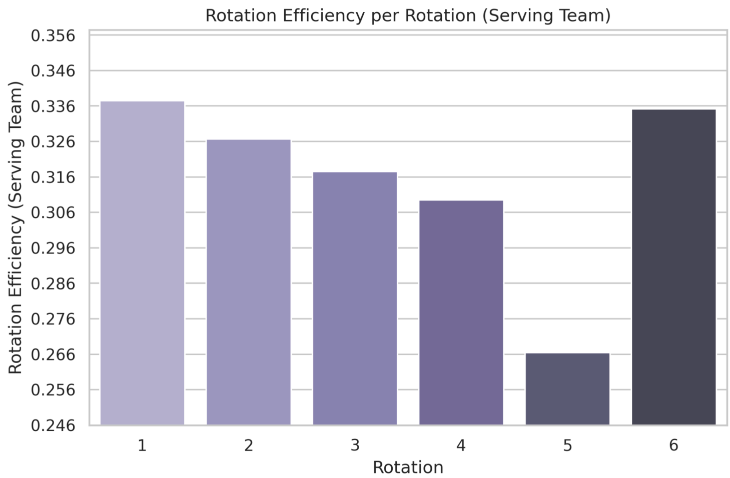
$$(\text{serving team}_t, \text{rotation}_t, \text{serve type}_t, \text{result}_t, \text{score}_t),$$

yielding a full rally-by-rally dataset that preserves sequential structure and rotation-specific behavior. For each pair of teams, multiple independent sets (e.g., five) are simulated, generating both rally-level logs and set summaries. Aggregated across all matches, these outputs provide a comprehensive dataset for rotation-level and set-level statistical analysis.

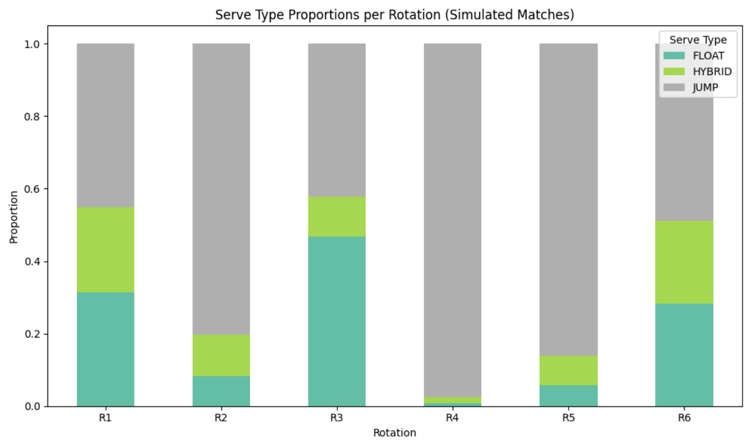
By drawing all stochastic elements from empirical rotation-level distributions, the semi-Markov simulator produces realistic, large-scale datasets that reflect expected performance characteristics of each rotation while remaining independent of opponent-specific dynamics. This framework enables systematic exploration of rotation efficiency, supporting downstream analyses, visualizations, and validation against observed match data.

Results

Simulation of elite men's volleyball matches revealed clear differences in rotation efficiency. Rotations 1 and 6 were the most effective, achieving efficiencies around 0.33–0.34, followed closely by Rotation 2. Contrawise, Rotations 4 and 5 were the least effective, with Rotation 5 falling to 0.26. These results, derived directly from the semi-Markov simulation, provide a robust assessment of rotation-level performance, independent of opponent-specific dynamics or short-term streaks.



Serve-type distributions provide important context for these differences. Rotations 1 and 6 exhibited a balanced mix of float and hybrid serves, Rotation 2 displayed a moderate mixture, and Rotations 4 and 5 were dominated by jump serves. Hybrid and float serves, disproportionately frequent in the high-efficiency rotations, appear to offer an optimal balance of technical difficulty for the receiver and stability for the server. By contrast, rotations dominated by jump serves—particularly R4—showed higher variability: although jump serves can produce aces at higher rates, they also introduced more errors, lowering net rotation efficiency.



A rotation-specific analysis highlights the practical implications of these patterns. Rotation 1's strong efficiency is notable given that setters are often perceived as weaker servers; in the simulation, this rotation benefits from hybrid serves and is frequently supplemented in practice by serving specialists providing high-velocity, targeted jump serves. Rotation 2's solid efficiency aligns with real-match findings that outside hitters commit fewer service errors than opposites, a pattern supported by raw error distributions in the dataset. Rotation 6 exhibits a hybrid profile, blending setter-type technical serves with the physical power characteristic of middle blockers, producing both unpredictability and force that contribute to its effectiveness. In contrast, Rotations 4 and 5, dominated by jump serves, displayed lower and more volatile efficiency. While the simulation captures these rotation-level tendencies, raw data on serve errors, aces, and in-play performance provide additional independent support, reinforcing the observed efficiency rankings and the role of serve-type distributions.

Discussions and final Evaluation

While the current simulation successfully captures rotation-level tendencies using rotation win probabilities and serve-type distributions, several additional metrics collected in the raw dataset were not incorporated due to the aforementioned limitations. Metrics such as pass quality, side-out efficiency, serve zone, and individual serve outcomes could provide further insight into the micro-dynamics that influence rotation effectiveness. For example, understanding how pass quality interacts with specific serve types could reveal subtle advantages in rotations that the current model cannot capture. Similarly, including detailed side-out efficiency metrics would allow evaluation of rotations under high-pressure situations, complementing the rotation-level results derived from serve initiation alone.

Future work should aim to integrate these excluded metrics into a more comprehensive modelling framework. Developing a probabilistic model that captures serve-to-receive transitions, or implementing a Markov process that accounts for rally progression, could make it possible to evaluate rotations not only by serve initiation but by their full impact on play continuity. Such approaches would allow the assessment of how serve type influences reception quality, side-out probability, and rally outcomes, directly addressing the current study's limitations.

Mathematical and computational methods capable of handling complex, high-variance metrics should be explored. Probabilistic hierarchical models, for example, could account for player-specific tendencies while pooling information across rotations and matches, mitigating the effects of sparse observations. Markov-type frameworks that incorporate rally progression could simulate sequences of actions beyond the serve, linking rotation efficiency to full-rally outcomes rather than point initiation alone.

By expanding the model in these ways, future studies could provide a more comprehensive understanding of rotational advantage in high-level volleyball. This would enhance the practical utility of the findings for coaching and tactical planning, allowing teams to identify rotation-specific strengths and weaknesses more precisely and to optimize rotation strategies in a data-driven manner. That said, the current model's outputs and framework are sufficient to support decisions based on these findings, offering meaningful contributions to the strategization and optimization of men's elite-level volleyball.