Modeling the limits of human performance: simulating the decathlon

Paul-Hieu V. Nguyen* James Smoliga † Benton Lindaman ‡ Sameer K. Deeshpande §

Abstract

The decathlon is an interesting competition to study from a statistical perspective for several reasons, such as its multi-event nature and unique scoring system, which are distinct from other athletic contests. Most previous analyses on the decathlon are descriptive; they focus on grouping decathlon events with principal component analysis or studying the trade-offs between athlete specialization. In this paper, we introduce a Bayesian compositional model that incorporates an athlete's age as well as their preceding-event performances from the same decathlon to incorporate the decathlon's unique dependence structure. We simulate decathlon performances to compare predictive performance across multiple models. We find that the compositional model performs similarly in predictive performance compared to our baseline model, while providing further insight to the decathlon's individual events. Additionally, we use the compositional model to generate age curves for individual athletes, evaluate realistic maximal scores, and study the relationship between various decathlon events.

1 Introduction

1.1 Motivation: Age curves for decathletes

The decathlon is a ten event track-and-field competition that takes place over two days. The order of events is consistent for each decathlon. On the first day, athletes compete in the 100m, long jump (LJ), shot put (SP), high jump (HJ), and 400m. On day two, they perform the 110m hurdles, discus throw (DT), pole vault (PV), javelin throw (JT), and 1500m. An athlete earns points for each event performance based on a scoring table developed by the World Athletics, and the overall score for the decathlon is the sum of each event's points.

Understanding how athletes improve or decline over time is crucial to training and performance evaluation; aging curves, which depict the relationship between athletic performance and age, offer a valuable tool for this purpose. They provide insights into an athlete's career trajectory and expected peak performance, which can guide competition goals and training program design. Figure 1 displays decathlon performances for two athletes, Ashton Eaton (green) and Harrison Williams (purple) up to age 25. These performances raise natural questions: did Eaton surpass expectations following his strong early career trajectory? How might Harrison perform as he continues to age? What are reasonable forecasts for each athlete's future decathon performances? What is a realistic maximum score for each athlete?

Table 1 depicts the world record achieved by Ashton Eaton at 27 years old in 2015 (now broken by Kevin Mayer). We also display Eaton's "virtual best", a method of calculating the theoretical maximum score an athlete could achieve by combining their personal bests in each of the decathlon events across multiple competitions. However, an athlete's virtual best is an imperfect estimate for their maximal score, as it assumes an athlete could achieve all personal bests in a single competition.

Another way to address these questions is to directly model total decathlon points as a function of age. However, because the decathlon is constructed from ten distinct events, aggregating scores may obscure important variation. Performance in individual events may peak at different ages, motivating a more granular approach: modeling each event separately and computing total points from the individual predictions. An alternative approach takes into account the structure of the decathlon. The order of events in the decathlon is fixed, and performance in earlier events could be predictive for later ones, through event similarity, fatigue, or momentum. To capture this structure, we introduce a compositional Bayesian model that explicitly conditions each event's performance on both age and performances in preceding events.

*Department of Statistics, University of Wisconsin–Madison. pvnguyen5@wisc.edu
†Department of Rehabilitation Science, Tufts University. james.smoliga@tufts.edu
‡Department of Rehabilitation Science, Tufts University. benton.lindaman@tufts.edu
§Department of Statistics, University of Wisconsin–Madison. sameer.deshpande@wisc.edu

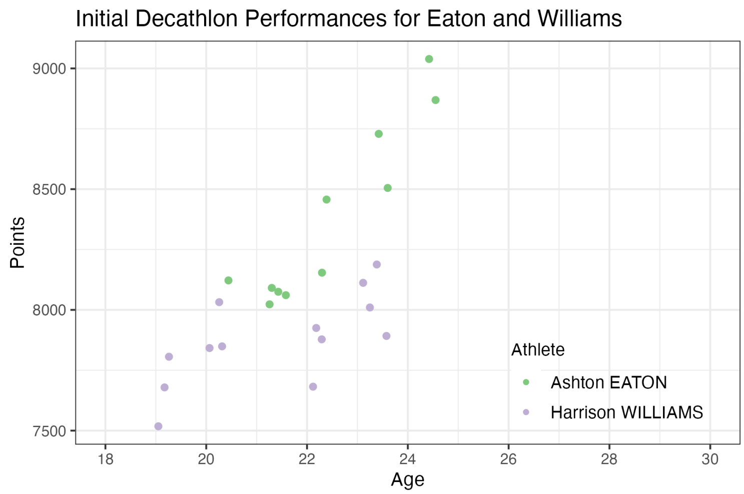

Figure 1: Selected decathlon performances from Ashton Eaton and Harrison Williams before age 25.

Table 1: Breakdown of maximal decathlon performances. Real best denotes Ashton Eaton's decathlon world record (2015). Virtual best combines Ashton Eaton's event-specific personal bests from each of his decathlons.

Event	Real Best	Virtual Best	Event	Real Best	Virtual Best
100m	10.23	10.21	Hurdles	13.69	13.35
Long jump	7.88	8.23	Discus Throw	43.34	47.36
Shot put	14.52	15.40	Pole vault	5.20	5.40
High jump	2.01	2.11	Javelin throw	63.63	66.64
400m	45.00	45.00	1500m	257.50	254.00
Total	9045	9543			

1.2 Previous Research

The multi-event nature of the decathlon presents a rich dataset with opportunities to analyze how age impacts each individual event and how performances in different events relate to each other. The majority of prior research has been descriptive in nature, for example grouping decathlon events (Cox and Dunn, 2002; Woolf et al., 2007; Walker and Caddigan, 2015) and athlete archetypes (Dziadek et al., 2022; Kenny et al., 2005; Van Damme et al., 2002). Several studies have used clustering or dimension reduction to characterize the structure of decathlon events. Cox and Dunn (2002) applied hierarchical clustering to group decathlon events, broadly seperating the track disciplines and field disciplines, with javelin as an outlier. Woolf et al. (2007) used cluster analysis based on personal best performances of elite decathletes and suggest track athletes may have a scoring advantage. Another popular method to analyze the decathlon is principle component analysis (PCA). Park and Zatsiorsky (2011); Dziadek et al. (2022) both used PCA to identify latent structures across events. Dziadek et al. (2022) specifically tracked how the structure of the decathlon shifts over an athlete's career and found that athletes broadly shift from generalists to specializing in particular events. Modeling performance trade-offs and specialization is a recurring topic of interest within the decathlon literature. Van Damme et al. (2002) found evidence of antagonistic traits, as well as tradeoffs between specialist and generalist phenotypes. Aoki et al. (2015) found physiological differences between athletes specializing in sprinting and jumping, though their analysis is not directly focused on the decathlon. Walker and Caddigan (2015); Kenny et al. (2005) argue that there is no evidence of event trade-offs when

Table 2: Descriptive statistics for decathlon events, age, and overall points.

Variable	Mean	Median	SD	Variable	Mean	Median	SD
100m	11.21	11.21	0.30	110mH	15.11	15.06	0.62
LJ	6.99	6.99	0.34	DT	39.60	39.72	4.79
SP	13.18	13.22	1.33	PV	4.47	4.50	0.37
HJ	1.93	1.93	0.09	JV	53.64	53.58	6.91
400m	50.52	50.42	1.53	1500m	284.96	284.00	14.98
Age	24.13	23.47	3.31	Points	7379.37	7362.00	511.39

looking at the subpopulation of elite decathlon performers: top athletes perform uniformly well across all events, and they found positive correlation between all decathlon events.

Fewer studies have explored the decathlon from a predictive standpoint. Battles et al. (2025) use gamma regression to model an athlete's career best using early-career decathlon performances. Battles et al. (2025) identified pole vault, javelin, long jump, and shot put as especially informative in predicting future potential. Wimmer et al. (2011) is the only application of Bayesian modeling we observed in the decathlon literature. Wimmer et al. (2011) use semi-parametric latent variable models to estimate event groupings and model the effects of age, season, and year on decathlon results. For a review of Bayesian methods within the larger sports analytics literature, we refer the reader to Santos-Fernandez et al. (2019). Age-performance relationships have also been explored outside of the decathlon. Researchers have sought to identify ages for peak performance in hockey (Schuckers et al., 2023; Brander et al., 2014), baseball (Fair, 2008), golf (Baker and McHale, 2023), and the triathlon (Villaroel et al., 2011). Griffin et al. (2022) use a Bayesian analysis to model sprinting and weightlifting events.

1.3 Our contributions

We introduce a compositional Bayesian model to account for the multi-event, dependent nature of the decathlon. In this method, we model an athlete's future scores in a decathlon as a function of the athlete's age and their preceding-event scores within the same decathlon. We compare multiple models, with varying levels of flexibility and granularity in an extensive set of experiments. We show that compositional models accurately predict actual decathlon scores. The compositional model also allows for greater interpretability, enabling researchers to study relationships between events. Using our probablistic models, for individual athletes, we can obtain personalized decathlon and event-specific age curves to develop training programs, set goals for competititions, and model potential for future success.

2 Modeling choices

2.1 Data

For the decathlon data, we used the raw data from the following address: `https://github.com/Battles186/DecathlonCareerBest.git`, which was constructed by Battles et al. (2025). The dataset includes complete men's decathlon performances from 2001-2022 obtained from World Athletics, a publicly available database. World Athletics contains decathlon scores of greater than 7000 for 2001-2008, greater than 6600 for 2009, and above 6400 for the years 2010-2022. We also filtered to keep only athletes with at least four performances with greater than 6400 points, so that our sample included only athletes with a sustained high level of performance. This is slightly less strict than Battles et al. (2025), who only keep athletes competing on record for at least four seasons. Increasing the sample size may better reflect a larger range of development trajectories. Our final dataset includes 8668 decathlon performances and 1007 unique athletes. Table 2 contains descriptive statistics for age, overall points, and each decathlon event.

2.2 Model

2.2.1 Baseline model

Now, we describe the set of models we use to predict decathlon performance. Let i denote the athlete, and j the decathlon observation. For each decathlon j, we observe the age of the athlete, $X_{age_{i,j}}$, vector of events, $\boldsymbol{X}_{i,j}$, and total number of points, $P_{i,j}$. We standardize each variable to have zero mean and standard deviation of one. We can now directly model total decathlon points with our baseline model, using only age as our sole covariate:

$$P_{i,j} = \alpha_i + \sum_d \beta_d \cdot \phi_d(X_{age_{i,j}}) + \epsilon_{i,j}, \quad \epsilon_{i,j} \sim \mathcal{N}\left(0, \sigma^2\right) \tag{1}$$

with α_i a random intercept for each athlete. We let ϕ be a fixed basis. In our experiments, we test both a cubic spline basis with interior knots placed at age deciles and a simple cubic polynomial basis. We assume the following default priors:

$$\begin{aligned}
&\alpha_i \sim \mathcal{N}\left(\mu_\alpha, \sigma_\alpha^2\right) \\
&\mu_\alpha \sim \mathcal{N}\left(0, 1\right) \\
&\sigma_\alpha^2, \sigma^2 \sim \text{Inv. Gamma}\left(2, 1\right) \\
&\boldsymbol{\beta} \sim \mathcal{N}\left(\mathbf{0}, I\right)
\end{aligned} \tag{2}$$

2.2.2 Simple model: changing response

Because the decathlon consists of ten distinct events, aggregating scores may obscure important variation. For example, performance in events likely peak at different ages, motivating a more granular approach. In our simple models, we predict each event separately and compute total points from the individual predictions. Again, let i denote the athlete, j the decathlon observation, and e the individual event. For each of the ten decathlon events, we model individual event scores using X_{age} as a covariate, independent of the other decathlon events. We then calculate the number of points earned for a given event score with $G_e(\cdot)$, an event specific formula further described in the Github repository. We sum the points earned for each event to determine the overall decathlon score.

$$\begin{gathered}
Y_{i,j,e} = \alpha_{i,e} + \sum_d \beta_{d,e} \cdot \phi_d(X_{age_{i,j}}) + \epsilon_{i,j,e}, \quad \epsilon_{i,j,e} \sim \mathcal{N}\left(0, \sigma_e^2\right) \\
P_{i,j} = \sum_{e=1}^{10} G_e(Y_{i,j,e})
\end{gathered} \tag{3}$$

Now, $\alpha_{i,e}$ represents a random intercept for each athlete specific to event e. Again, we use the cubic spline and cubic polynomial basis expansion for ϕ, as in Section 2.2.1, and we use the same priors as Equation (2). Using the simple rather than the baseline model, we model the decathlon with increased granularity. Furthermore, we will be able to track how age affects each event seperately. While an athlete's ability in certain types of events may degrade as they grow older, performance in other events may improve with age.

2.2.3 Compositional model: adding preceding events

In the compositional models, we take advantage of the fact that the order of events in the decathlon is fixed and predetermined. Unlike the simple models, where we model individual events independently, in our compositional model, we model individual events using age *and all previous events*. Similarly to the simple models, we compute and sum the predicted event scores to calculate the overall decathlon points.

$$
\begin{gathered}
Y_{i,j,e} = \alpha_{i,e} + \sum_{d} \beta_{d,e} \cdot \phi_d(X_{age_{i,j}}) + \sum_{m=1}^{e-1} \gamma_{m,e} \cdot Y_{i,j,m} + \epsilon_{i,j,e}, \quad \epsilon_{i,j,e} \sim \mathcal{N}\left(0, \sigma_e^2\right) \\
P_{i,j} = \sum_{e=1}^{10} G_e(Y_{i,j,e})
\end{gathered} \tag{4}
$$

The priors on α and β remain as in Equation (2). We place $\mathcal{N}(0,1)$ priors on each γ_m as well.

For each model previously described, we investigate gains from incorporating open-event data to inform our models, specifically the relationship between age and performance. For each of the 10 events in the decathlon, we identified the top 26 (and in some cases top 50) competitors in the corresponding open events. Then, for each athlete, we recorded their best performance each year they competed in. To predict decathlon event e, we first model the corresponding open event with a cubic regression or spline. We then used the estimated age coefficients as the means for the prior distribution and compared open-event based priors with the default priors. In our experiments, we found that utilizing additional open-event data did not greatly influence the predictive results; in some cases, the spline models performed worse. Thus, we continue using the default priors in Equation (2).

3 Experimental results

We first evaluated the reliability of the compositional model using simulated data with known underlying structure. In these experiments, we focus on the capacity of the compositional model for parameter recovery. In Section 3.2, we fit the models, including priors based on open-event data, onto the real decathlon data to evaluate each model's predictive proficiency.

3.1 Parameter recovery with known $\boldsymbol{\beta}$

First, we demonstrate that the compositional model recovers known inter-event dependence when it exists and quantifies uncertainty appropriately. We specify a known structure of linear dependence between events; for example, long jump (LJ) depends on 100m, shotput (SP) depends on LJ and 100m, and so on. For each event, we estimate coefficients for the age polynomials and preceding events using the original decathlon data. We then simulate 200 datasets with 8668 observations each, treating the estimated coefficients as the true parameters.

We then fit the cubic, compositional model to the 200 simulated datasets and extracted posterior distributions over the β coefficients corresponding to age and preceding events. For each coefficient, we computed the proportion of 95% posterior intervals containing the true value used in simulation. We direct the reader to the Github repository for tabulations. The model generally recovers the true parameters well, with most coverage proportions near or above the 95% level. The combination with the lowest coverage is javelin and pole vault; even so, the coverage is quite high at 90.5%. Overall, the high coverage across virtually all predictors suggests that the model is well-calibrated for recovering inter-event relationships.

3.2 Model Comparisons with Real Data

We compared the out of sample accuracy on real decathlon data for each model under two frameworks. The first framework is a 'general' case, where we created ten 90%-10% training/testing splits, with observations placed in the test split at random. In 'tail' framework, we create ten training/testing splits by selecting 10% of the decathletes and using their last decathlon observation as the test set, and all other observations in the training dataset.

Notably, most models have very similar predictive performances for both prediction of events and overall points. We direct the reader to the Github repository for tabulations for each individual event and model. For points, the cubic and spline baseline models have a mean SMSE of 0.234. Each other model using the default priors had a mean SMSE of 0.235. Generally, the models with open-event matched priors performed similarly, though some spline combinations performed slightly worse, with mean SMSE's of 0.236 and 0.238. The results for the 'tail' observations follow this general trend. Overall, we see that modeling the age-event relationship with a cubic regression, rather than a spline, tends to perform better. The simple and compositional models have similar predictive results, compared to the baseline models. While predictive performance is similar across models, the more complex formulations offer added interpretability and flexibility, further discussed in Section 3.3. Despite its greater structural complexity, the compositional model performs on par with the simpler alternatives, even without any explicit tuning to optimize predictive accuracy. The compositional model does not sacrifice generalization – SMSE values are comparable to both the simple and baseline models. This is a key strength: more complexity did not result in worse performance. Additionally, the compositional model more accurately captures the uncertainty of the overall decathlon compared to the simple model, as seen in Figure 4, which displayes the posterior predictive intervals for decathlon performances by Eaton and Williams. Thus, the added structure of the compositional model maintains predictive accuracy while bringing interpretability benefits through capturing realistic inter-event relationships.

3.3 Studying inter-event relationships with the compositional model

Importantly, the compositional model reproduces the inter-event correlations observed in the real data within its posterior predictive distribution, whereas the other models tend to understate (simple) or ignore these dependencies entirely (baseline). To demonstrate, we generate 2,000 simulated decathlon datasets from the posterior predictive distribution of the simple and compositional model, using the same athletes and ages as in the observed data. Note that the baseline model does not produce event-specific estimates. For each posterior predictive draw, we compute the correlation between every pair of events and compare these posterior predictive correlations with the empirical correlation. We plot the posterior predictive correlations for the simple and compositional model in Figure 2. The results highlight how the compositional model accurately captures the event correlation, whereas the simple model routinely underestimates the relationship strength. In the Github repository, we display tables with the emperical correlations, and posterior predictive intervals for the simple and compositonal models. These realistic correlations arise naturally from the model's sequential structure, which allows performance in earlier events to influence later events in ways that match the patterns seen in actual decathlon results.

To better understand the relationship between events, we can examine the posterior distributions of selected coefficients. Figure 3 displays the distributions of two example relationships: the effect of the 100m sprint on long jump performance (left), and the effect of the long jump on javelin throw performance (right). The posterior mean of each distribution is depicted with a solid red line, and the 2.5 and 97.5 percentiles are marked with the dotted line. For instance, an improvement of 1 second in the 100m, *holding age constant*, is associated with an expected increase of .54 meters in the long jump. Similarly, an improvement of 1 meter in the long jump, *holding age and all other previous events constant,* is associated with an improvement of approximately 1.02 meters in javelin.

These associations should not be interpreted as being strictly causal. For instance, we would not expect an athlete who only trains for the long jump to directly improve their javelin performance. Rather, these coefficients can be viewed as reflecting broader trends in performance among elite decathletes. Training speed, for example, could enhance multiple events, such as the 100m, 110m hurdles, and 400m. A positive coefficient between the 100m and long jump may arise from shared underlying attributes rather than a direct transfer of skill from one event to another. This is particularly relevant given that our dataset consists only of top decathlon performances, where athletes must be competitive across all events. Specialists who excel in only one or two events but underperform in the remaining events are unlikely to appear in this data. Consequently, these coefficients may be better interpreted as indicators of general performance patterns at the elite level rather than isolated event-on-event effects.

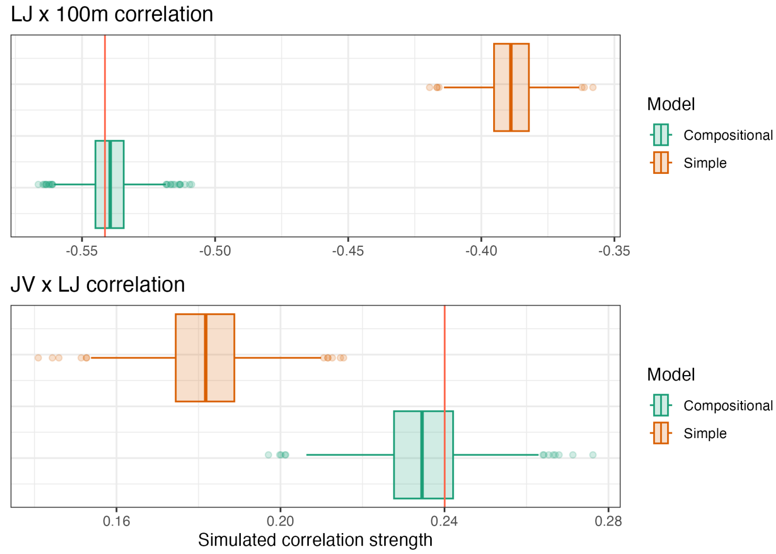

Figure 2: Boxplots depicting the posterior predictive correlations for the simple and compositional models over 2000 simulated datasets between the 100m and long jump (top) and javelin and long jump (bottom). For each model, the 25th, 50th, and 75th percentile is marked in the boxplot. The empirical correlation from the observed data is plotted in red.

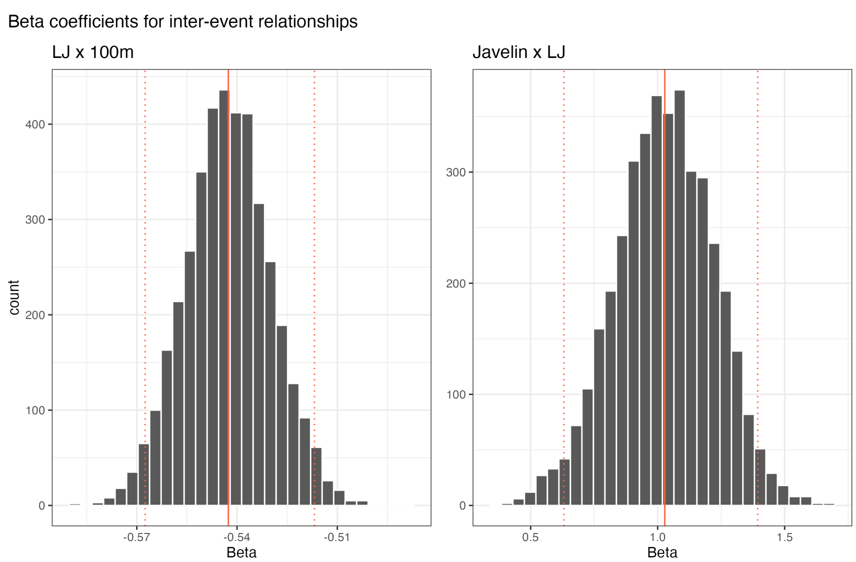

Figure 3: Histogram of posterior draws of β for prediction of long jump (left) and javelin (right). Specifically, the coefficients associated with the 100m when predicting LJ and LJ when predicing javelin, respectively, are plotted. The posterior mean is marked with a solid red line, and the 0.025 and 0.975 quantiles are denoted with the dotted red lines.

4 Forward simulation

We use forward simulation to generate realistic, full decathlon performances for individual athletes by drawing from the posterior distribution of our Bayesian models. Using the simulations, we explore decathlon outcomes at different ages to identify performance peaks of individual events and to study inter-event dynamics.

4.1 Simulation procedure

To evaluate realistic decathlon trajectories, we simulate full decathlon performances for individual athletes across a range of ages. For each athlete i, we draw simulated event results using posterior samples from the fitted model. At each simulation step, we use posterior draws of athlete-specific intercepts, $\alpha_{i,e}$, age effects, β_e, residual standard deviations, σ_e, and, in the compositional model, regression coefficients, γ_e, that capture dependencies between events. For a given age, x_{age}, ranging from 19-30, and posterior sample, (s), we simulate performances sequentially over events $e \in \{1, \ldots, 10\}$ as follows. For each event, e, first sample a posterior draw of all parameters, $\alpha_{i,e}^{(s)}, \beta_e^{(s)}, \gamma_e^{(s)}, \sigma_e^{(s)}$. Then simulate the event performance for athlete i from a Gaussian distribution. For the simple model:

$$Y_{i,e}^{(s)} \sim \mathcal{N}\left(\alpha_{i,e}^{(s)} + \sum_d \beta_{d,e}^{(s)} \cdot \phi_d(x_{age_{i,j}}), \sigma_e^{(s)}\right) \tag{5}$$

In the compositional model, the simulation of the 100m, $(e = 1)$, remains the same. For all other events, we use the simulated performances in previous events from draw s, denoted, $Y_{i,m}^{(s)}$ for $m \in \{1, \ldots, e-1\}$. This allows us to capture and propagate inter-event dependencies throughout the decathlon.

$$Y_{i,e}^{(s)} \sim \mathcal{N}\left(\alpha_{i,e}^{(s)} + \sum_d \beta_{d,e}^{(s)} \cdot \phi_d(x_{age_{i,j}}) + \sum_{m=1}^{e-1} \gamma_{m,e}^{(s)} \cdot Y_{i,m}^{(}s), \sigma_e^{(s)}\right) \tag{6}$$

After simulating all ten events, we compute a total decathlon score using the official scoring tables. Repeating this process across posterior draws produces a distribution over decathlon trajectories for a specific athlete at each age. This simulation framework provides athlete-specific age curves for both the overall decathlon and each individual event using a coherent and interpretable generative model. The compositional model is particularly valuable for generating simulations that reflect realistic event interplay.

4.2 Simulation example: Eaton and Williams

As an example, we simulate 4000 decathlon performances using the compositional model for Ashton Eaton and Harrison Williams for each $x_{age} \in \{19, 19.1, \ldots, 29.9, 30\}$. In total, we simulate 444,000 decathlon performances for each athlete. Consider Figure 4, which displays 95% posterior predictive intervals for Ashton Eaton's and Harrison Williams' points for the baseline, simple, and compositional models across the range of ages. The posterior mean is depicted with a solid line. While the overall shapes of the age curves are consistent, the simple model displays noticably tighter intervals. Each model predicts a peak at roughly the age 26. Interestingly, the simple model's intervals occasionally miss some of Eaton's later performances, possibly reflecting the model's lack of modeling for inter-event dependencies. In contrast, the compositional model produces wider intervals that better capture performance variability.

4.3 Maximal event performance

While the baseline model performed well in terms of SMSE, increasing model complexity offers important interpretability benefits without sacrificing predictive accuracy. For example, we can break age curves down and identify peak ages for each individual event using the more complex models. One may initially expect

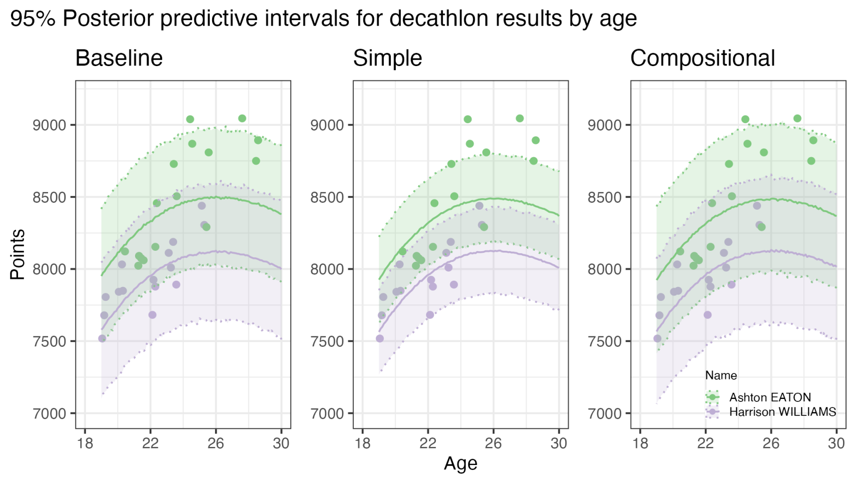

Figure 4: Age curve for decathlon points for baseline (left) and simple (middle) and compositional (right) models, specifically for Ashton Eaton, green, and Harrison Williams, purple. The shaded areas represent the 95% posterior predictive interval for each respective model, and the solid line depicts the posterior mean

that identifying the peak age for athletic performance simply requires minimizing or maximizing the age effect function, $\sum_d \beta_{d,e} \cdot \phi_d(x_{age})$, which captures the marginal contribution of age on performance. Since this function is modeled using a cubic polynomial (or spline), one could compute the derivative analytically and solve for the age at which performance is optimized. This approach aligns with the traditional understanding of age-performance relationships and is straightforward when modeling each event independently.

However, within our compositional framework, the posterior predictive distribution does not only reflect the isolated age terms, but also how age indirectly influences event performance through preceding events. Specifically, for later events in the decathlon, age affects performance through both the age coefficients and indirectly through preceding event performances, which themselves are functions of age. These earlier events then enter as covariates when modeling downstream events, introducing additional dependencies that are not captured by studying $\sum_d \beta_{d,e} \cdot \phi_d(x_{age})$ alone. As a result, the performance for a given event can change meaningfully from what would be inferred by considering only the β coefficients directly related to age. By leveraging the full compositional posterior distribution, we incorporate both direct and indirect age effects and obtain accurate representations of when peak decathlon performance is likely to occur.

To estimate the maximal performance age for each event, we simulate event trajectories using posterior draws from the compositional model. For each posterior sample, we evaluate the predicted performance, without added observational noise, across a grid of ages from 19-30. We then propagate these predictions through the compositional model, using earlier predicted performances without noise, as inputs to subsequent events. For each posterior draw, we identify the age that yields the best expected performance in a given event. Repeating this across posterior draws produces a posterior distribution over the ages of maximal performance in each decathlon event. Table 3 indicates summary statistics for the optimal ages for each decathlon event. In Figure 5 we present the age curves for the 400m and javelin for Eaton and Williams from ages 19-30. We shade the 95% credible intervals for simulated scores and plot observed scores for each athlete. We also shade in red the 25% to 75% quartiles for optimal age in each event from Table 3. We note that throwing events generally tend reach their optimal age (27-29) substantially later than sprinting events (24-26). We also find that certain events exhibit much more concentrated peak-age estimates, indicating that their optimal ages are highly consistent across posterior draws. For example, the 25-75% quartile range for optimal age for the 400m is half of a year, whereas the optimal age range for the shot put is almost one full year.

Table 3: The mean, 25%, and 75% quantiles for age resulting in optimal performance in each decathlon event across 4000 decathlon age curves. Optimal performance is defined as the shortest time for sprints, furthest distance for throws and jumps, and most points for the overall decathlon

Event	Mean	Q25	Q75	Event	Mean	Q25	Q75
100m	24.32	24.10	24.55	110mH	26.08	25.80	26.40
Long Jump	25.24	24.90	25.60	Discus Throw	29.65	29.05	30.00
Shot Put	28.90	28.45	29.40	Pole Vault	27.47	27.10	27.85
High Jump	24.16	23.80	24.55	Javelin Throw	27.16	26.60	27.75
400m	23.58	23.35	23.85	1500m	22.72	21.85	23.35
Overall	26.00	24.6	27.5				

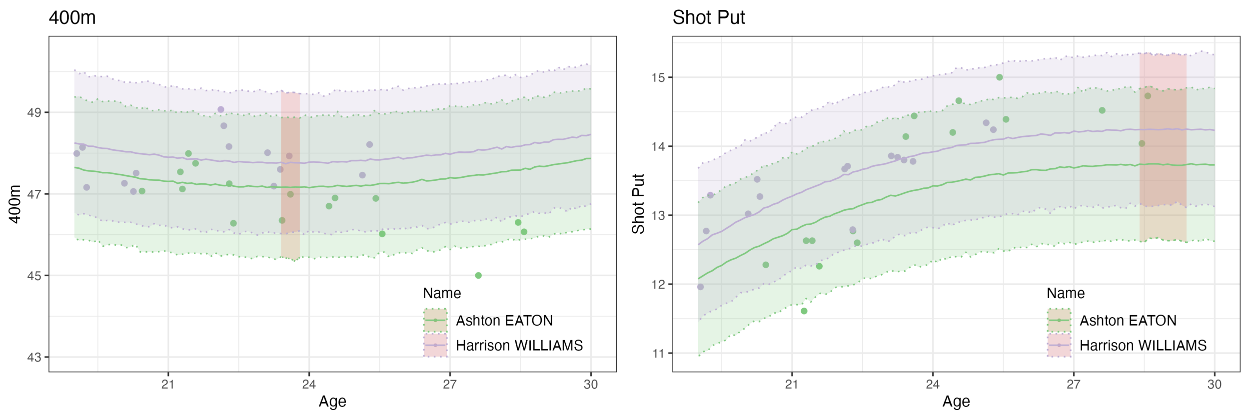

Figure 5: Age curves for shot put and 400m for Eaton and Williams. The shaded areas represents the 95% credible interval for event performance, and actual observations are plotted. The solid line depicts the posterior mean. The red area shades the predicted maximal ages for event performance, respectively.

5 Discussion

In this paper, we have devceloped a new, compositional Bayesian model that provides granular and interpretable results while matching the predictive performance of our baseline model. We evaluated the compositional model using both synthetic and real-world data experiments, and we produced both overall decathlon and event-specific age performance curves. The model quantified uncertainty effectively and provided insight for inter-event relationships. With this model, we can identify peak performance ages for each event and assess how performance in one discipline relates to another in a clear and coherent manner. Ultimately, our model provides targeted insight into how earlier events help explain variation in later ones while preserving predictive accuracy.

There are several extentions to our proposed methodology. Future work could incorporate imputation methods from Schuckers et al. (2023) to address potential selection bias. One concern is that only the most successful athletes are observed in the data at the youngest and oldest ages. While we believe this bias is less severe in track and field than in professional team sports, we believe this is still worth investigating. Another natural extension is to apply the compositonal model to other competitions with multiple events, such as the triathlon or heptathlon. Further work can incorporate additional data and covariates, such as season and environmental factors, which can impact performance. Lastly, we consider alternative modifications to the model structure, such as tree-based methods like Bayesian additive regression trees (BART) to better capture non-linearities and interactions.

References

Aoki, K., Kohmura, Y., Sakuma, K., Koshikawa, K., and Naito, H. (2015). Relationships between field tests of power and athletic performance in track and field athletes specializing in power events. *International Journal of Sports Science & Coaching*, 10(1):133–144.

Baker, R. D. and McHale, I. G. (2023). A flexible mixed model for age-dependent performance: application to golf. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 72(5):1260–1275.

Battles, P., Noble, T. J., and Chapman, R. F. (2025). Predicting high-performance decathlon career best. *Experimental Physiology*.

Brander, J. A., Egan, E. J., and Yeung, L. (2014). Estimating the effects of age on nhl player performance. *Journal of Quantitative Analysis in Sports*, 10(2):241–259.

Cox, T. F. and Dunn, R. T. (2002). An analysis of decathlon data. *Journal of the Royal Statistical Society Series D: The Statistician*, 51(2):179–187.

Dziadek, B., Iskra, J., Mendyka, W., and Przednowek, K. (2022). Principal component analysis in the study of the structure of decathlon at different stages of sports career. *Polish Journal of Sport and Tourism*, 29(4):21–28.

Fair, R. C. (2008). Estimated age effects in baseball. *Journal of Quantitative Analysis in Sports*, 4(1).

Griffin, J. E., Hinoveanu, L. C., and Hopker, J. G. (2022). Bayesian modelling of elite sporting performance with large databases. *Journal of Quantitative Analysis in Sports*, 18(4):253–268.

Kenny, I. C., Sprevak, D., Sharp, C., and Boreham, C. (2005). Determinants of success in the olympic decathlon: Some statistical evidence. *Journal of Quantitative Analysis in Sports*, 1(1).

Park, J. and Zatsiorsky, V. M. (2011). Multivariate statistical analysis of decathlon performance results in olympic athletes (1988-2008). *World Academy of Science, Engineering and Technology*, 5(5):985–988.

Santos-Fernandez, E., Wu, P., and Mengersen, K. L. (2019). Bayesian statistics meets sports: a comprehensive review. *Journal of Quantitative Analysis in Sports*, 15(4):289–312.

Schuckers, M., Lopez, M., and Macdonald, B. (2023). Estimation of player aging curves using regression and imputation. *Annals of Operations Research*, 325(1):681–699.

Van Damme, R., Wilson, R. S., Vanhooydonck, B., and Aerts, P. (2002). Performance constraints in decathletes. *Nature*, 415(6873):755–756.

Villaroel, C., Mora, R., and Gonzalez-Parra, G. C. (2011). Elite triathlete performance related to age. *Journal of Human Sport and Exercise*, 6(2):363–373.

Walker, J. A. and Caddigan, S. P. (2015). Performance trade-offs and individual quality in decathletes. *Journal of Experimental Biology*, 218(22):3647–3657.

Wimmer, V., Fenske, N., Pyrka, P., and Fahrmeir, L. (2011). Exploring competition performance in decathlon using semi-parametric latent factor models. *Journal of Quantitative Analysis in Sports*, (4).

Woolf, A., Ansley, L., and Bidgood, P. (2007). Grouping of decathlon disciplines. *Journal of Quantitative Analysis in sports*, 3(4).