Bayesian inference of the impulse-response model of athlete training and performance

Kangyi Peng[1,2], Ryan T. Brodie[2,3,4], Tim B. Swartz[1,2], David C. Clarke[2,3,4,*]

[1]Department of Statistics and Actuarial Science, [2]SFU Sports Analytics Group, [3]Department of Biomedical Physiology and Kinesiology, Simon Fraser University, Burnaby BC, V5A 1S6, Canada; [4]Canadian Sport Institute Pacific, Pacific Institute for Sport Excellence, 4371 Interurban Road, Victoria BC, V9E 2C5 Canada.

*Corresponding author:

Email: dcclarke@sfu.ca

Tel: 778-782-9777

Fax: 778-782-3040

Running title: Bayesian inference of IR model

Keywords: athletic performance, mathematical modeling, Bayes Theorem, Banister impulse-response model, computer simulation.

Word count:

Abstract

The Banister impulse-response (IR) model is the most studied mathematical model relating athletic performance to training. The model features important limitations that impair its practical usefulness, in particular difficulties in obtaining precise parameter estimates and future performance predictions. Here we develop a Bayesian implementation of the IR model, which formalizes the combined use of prior knowledge and fitting to data, and which helps overcome its foremost limitations. We report three methodological contributions: 1) we reformulated the model to facilitate the specification of informative priors, 2) we compiled published estimates of the IR model parameter values and used them to specify the prior distributions, and 3) we developed a method that enabled the JAGS software to be used while enforcing parameter constraints. We then applied the model to the training and performance data of a national-class middle-distance runner, demonstrating proof-of-principle the superiority of the Bayesian approach to nonlinear least squares in terms of the precision and plausibility of the parameter estimates. The Bayesian approach enabled several insights into the athlete's performance, and the method can be iteratively applied in perpetuity as new training and performance data are collected. We conclude that the Bayesian implementation of the IR model overcomes the foremost challenges to its usefulness for athlete monitoring.

1. Introduction

Maximizing athletic performance depends primarily on athletes undertaking appropriate training loads at appropriate times. Understanding the quantitative relationship between training and performance is thus of interest to athletes and their advisors. Mathematical models that predict performance from training have been proposed, with the most studied being the Banister impulse-response (IR) model (Clarke & Skiba, 2013). The IR model expresses performance at time *t* as the sum of the initial or baseline performance capacity, P_0, the positive training effects, and the negative training effects (Equation 1).

$$P_t = P_0 + K_1 \sum_{s=0}^{t-1} e^{-\frac{t-s}{\tau_1}} * W_s - K_2 \sum_{s=0}^{t-1} e^{-\frac{t-s}{\tau_2}} * W_s \quad (1)$$

K_1 and K_2 are terms that express the change in performance per unit training accomplished, τ_1 and τ_2 are constants that describe the decay rates of the positive and negative training effects over time, and W_s is the training accomplished at time = s. The model presents an intuitive framework for understanding the dynamic response to training (Clarke & Skiba, 2013). This form of the model features five adjustable parameters ($P_0, K_1, K_2, \tau_1, \tau_2$) that are typically estimated by fitting the model to data from maximal-effort performances using maximum-likelihood approaches. The model has been used to analyze and predict performance and optimize training in various sports such as cycling, running, swimming, weightlifting, and track and field events.

Despite its promise, the model features several noteworthy limitations. First, its use can be burdensome in terms of time and effort. Training load data W_s must be rigorously collected, which is facilitated by available wearable and portable technologies such as bicycle-mounted power meters and GPS wristwatches. High-quality performance P_t data must likewise be regularly collected, and this requirement particularly challenges the model's widespread use. For example, athletes may compete too infrequently to accumulate sufficient data from competitions, or they may be reluctant to devote training time to performance tests. Even if sufficient performance data are accumulated, the signal-to-noise ratio in these data is typically low for experienced athletes because their performance levels tend to be relatively stable. Accordingly, the parameters of the model are often poorly estimated (Busso & Thomas, 2006; Hellard et al., 2006). The aforementioned data challenge is difficult to overcome: IR model estimation methods employed to date are entirely data driven, with no formal way to incorporate other knowledge into the framework. Approaches to overcome these challenges are therefore sought.

An analogous challenge has been successfully addressed by anti-doping organizations in implementing the Athlete Biological Passport (ABP). The ABP is a framework developed to monitor suspicious changes in biomarkers of doping over time. The effectiveness of the ABP is challenged by the relatively infrequency of athlete testing and the measured variables being influenced by both biological and nuisance technical factors. Stewards of the ABP resolved this challenge by employing Bayesian methods, in which prior probability distributions ("priors") based on population averages define the normal ranges for the measured variables for a given athlete, and these ranges are updated using data collected from the athlete. With every test, the ranges become increasingly athlete specific. The ABP has been successful in reducing doping prevalence. More recently, a Bayesian framework was proposed to monitor suspicious changes in performance, as part of an emerging "performance passport" approach to anti-doping (Hopker et al., 2020). By formalizing the judicious use of prior information, Bayesian approaches are useful when data are sparse, and athletes, coaches, and sport scientists can contribute their knowledge to the specification of the priors. Despite the promise of Bayesian approaches, the IR model has yet to be specified in a Bayesian framework and applied in practice.

The purpose of this study is to cast the IR model in a Bayesian framework and to apply it to data from an elite middle-distance runner. We report three main contributions: first, we reformulated the model to enhance our ability to specify informative prior distributions for the model parameters. Second, we developed a generalizable procedure for imposing parameter constraints that enabled the computations to be conducted using JAGS software. Third, we demonstrate the superiority of Bayesian inference compared to a commonly used nonlinear regression procedure. We conclude that the Bayesian inference approach provides a theoretically and empirically superior approach for applying the IR model to the longitudinal monitoring and prediction of athletic training and performance.

Methods

2.1 Study design & participant

The study design was observational; we used previously collected training-load and performance data to fit the models. Ethical approval was obtained from the Simon Fraser University Office of Research Ethics. A Canadian national-level middle-distance runner volunteered to participate in the study and provided informed consent. The athlete provided a season's worth of training and performance data, spanning September 1, 2017 to July 28, 2018 (301 days), during which time 259 workouts were documented.

2.2 Training and performance data

Commented [DC1]: How many days from initial workout to last training session or performance prior to the break?

Commented [DC2R1]: I feel that we should not include the final performance data point, there was no training data collected on the preceding 9 days, which is unrealistic.

The daily training loads W_s were recorded as the individualized training impulse (TRIMPi; Manzi, Iellamo, Impellizzeri, D'Ottavio, & Castagna, 2009). An athlete-specific multiplying factor was used to represent the nonlinear effect of intensity on training load. The value was generated from the relationship between blood lactate levels and the fraction of heart-rate reserve measured during an incremental treadmill exercise test.

Commented [DC3]: Report this function (Ryan)

The athlete trained on 259 days during the season but TRIMPi were measured only for 173 of those days, likely because the athlete did not wear the heart-rate chest strap for all workouts. We therefore imputed the TRIMPi values in the following manner. First, we assumed that the TRIMPi were missing at random, and we observed that they were linearly associated with the distances run (km) during the workouts recorded by the GPS wristwatch. We used linear regression to quantify the relationship between TRIMPi and distance run, with distance run specified as the explanatory variable and log(TRIMPi) as the response variable. The TRIMPi were log transformed to ensure the validity of the normality assumption of the linear regression. Second, we used single imputation (Zhang, 2016), in which random errors are added to the predicted values from the regression model, to ensure that the imputed values had similar variation as the observed data.

Performance P_t was expressed as IAAF points achieved in sanctioned races. This approach was used because the athletes raced over different distances (e.g., 800 m, 1,500 m, and one mile), whose times and velocities are not straightforwardly comparable. Referring to equation (1), the data are denoted $P = (P_1,...,P_N)$ where N measurements were recorded.

Model estimation: nonlinear least squares

The parameters of the IR model (Equation 1) were estimated using nonlinear least squares. This procedure finds the combination of parameter values that minimize the sum-of-squares of the residual values corresponding to the modeled and measured performances (Johnson & Frasier, 1985). The method was implemented in R using the "nlminb" function. Confidence intervals for the parameter values were computed using a bootstrap method (Efron & Tibshirani, 1986).

2.3 Model formulation

We cast the IR model in a stochastic framework by reformulating the original version of the model (Equation 1) as follows:

$$P_t = P_0 + K_1 \sum_{s=0}^{t-1} e^{-\frac{t-s}{\tau_1}} * W_s - \theta * K_1 \sum_{s=0}^{t-1} e^{-\frac{t-s}{\tau_2}} * W_s + \varepsilon_t = \mu_t + \varepsilon_t \quad (2)$$

where μ_t is the expected value of performance, ε_t are the unobserved errors, and θ is an unknown constant greater than one that relates K_1 to K_2. The μ_t and ε_t terms allow the model to be probabilistically assessed. In section 2.4, we discuss distributional assumptions concerning ε_t. We rewrote K_2 as $\theta * K_1$ because θ is a parameter for which we have greater prior knowledge, and it is less dispersed than K_2. The corresponding physiology imposes the restriction $\theta > 1$.

To enhance the interpretability of the IR model, two derived parameters are commonly calculated, t_n and t_g. t_n is the day after which training has a net negative influence on performance at time t, and t_g is the day where training has the highest positive influence on performance at time t. t_n and t_g are computed from the following formulae:

$$t_n = \frac{\tau_1 \tau_2}{\tau_1 - \tau_2} ln(\frac{K_2}{K_1}) \quad (3)$$

$$t_g = \frac{\tau_1 \tau_2}{\tau_1 - \tau_2} ln(\frac{K_2}{K_1}\frac{\tau_1}{\tau_2}) \quad (4)$$

Using these equations, θ can be rewritten entirely in terms of τ_1, τ_2, t_n, and t_g, which are parameters for which we have the best prior knowledge.

$$\theta = K_2/K_1 = (\tau_1/\tau_2)^{(\frac{1}{t_g/t_n - 1})} \quad (5)$$

Overall, the IR model parameters $(P_0, K_1, K_2, \tau_1, \tau_2)$ are reformulated as $(P_0, K_1, \theta, \tau_1, \tau_2)$.

2.4 Bayesian implementation of the IR model

Bayesian approaches are being increasingly used in sports science and are particularly useful for applications involving elite athletes (Hecksteden et al., 2022; Santos-Fernandez, Wu, & Mengersen, 2019). Primers on the use of Bayesian approaches are available elsewhere (van de Schoot et al., 2021; Van de Schoot et al., 2014). In the Bayesian approach, a *posterior probability distribution* is obtained from the *prior distribution* and the *likelihood function*. To cast the IR model in a Bayesian framework, we first assume that the model parameters are random variables that conform to particular probability distributions. The prior density $\pi(\Theta)$ encodes background knowledge regarding the model parameters. The likelihood function $f(P|\Theta)$ specifies the information from the data. The posterior density describes the updated

Commented [DC4]: Deleted text: In the IR model (1), parameters K1 and K2 are linear coefficients of positive training effects and negative training effects, respectively.

Unlike the other three parameters P_0, τ_1 and τ_2, K1 and K2 do not have formal physiological interpretations. Introducing informative prior knowledge on K1 and K2 is challenging since different training loads and performance measurements can affect their scale.

Commented [DC5]: Deleted text

The Bayesian statistics was first described by T. Bayes in 1774 (Bayes & Price, 1763; Stigler, 1986).

probability associated with the model parameters (given the data) and is proportional to the product of the prior distribution and likelihood function, as follows:

$$\pi(\Theta|P) \propto f(P|\Theta)\pi(\Theta) \tag{6}$$

where Θ refers to the parameters in the IR model, including the variance parameters associated with the random error term ε_t. The vector P = (P_1, ..., P_N) is the performance data.

Next, we assumed that the observed performances P_1,..., P_N are recorded daily, although this assumption is not necessary in practice. We then assumed that the performances are correlated in time; specifically, the performance on day t is related to the performance on day t-1, t-2, and so on with decreasing correlation. To encode this assumption, we assumed that the error terms $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N$ conformed to the multivariate normal distribution $MVN_N(0, \Sigma)$, and we modeled the athlete's performances [P_1, P_2 ,..., P_N| $P_0, K_1, \theta, \tau_1, \tau_2, \Sigma$] as $MVN_N(\mu, \Sigma)$, where μ = (μ_1, μ_2 ,..., μ_N). Note that μ_t is the expected performance on day t (Equation 2). The parameter Σ is the variance-covariance matrix of the multivariate normal distribution, where the i,jth term of Σ is equal to $\sigma^2 * \rho^{|i-j|}$, $0<\rho<1$. The parameter σ^2 is the variance term, and $\rho^{|i-j|}$ is the correlation between performances on day i and day j. When i = j, $\rho^{|i-j|}$ is maximized and is equal to 1; for i ≠ j, $\rho^{|i-j|}$ decreases as |i-j| increases. This stipulation reflects the notion that performances closer in time to one another are expected to be more similar. This idea has been used in the analysis of substitution times in soccer (Silva & Swartz, 2016). This parametrization is appealing due to its simplicity because the N(N+1)/2 parameters in Σ are reduced to two parameters (ρ, σ). Using the density function of the multivariate normal distribution, the likelihood function of the data is therefore expressed as follows:

Commented [DC6]: Ken, is this reference correct?

$$f(P|\Theta) = f(P|P_0, K_1, \theta, \tau_1, \tau_2, \Sigma(\sigma,\rho)) \propto det(\Sigma(\sigma,\rho))^{-\frac{1}{2}} e^{-\frac{1}{2}(P-\mu)^T \Sigma(\sigma,\rho)^{-1}(P-\mu)} \tag{7}$$

2.5 Prior elicitation

The prior density $\pi(\Theta)$ expresses our prior beliefs regarding the model parameters (Van de Schoot et al., 2014); it does not depend on the data. Prior elicitation involves specifying the probability distribution to which the parameter is expected to conform. The certainties of the priors are encoded in the widths of the distributions: for parameters whose values are well established, relatively strong priors are assigned, whereas the priors for parameters whose values are less certain, more diffuse priors are assigned.

We elicited the prior density of Θ = ($P_0, K_1, \theta, \tau_1, \tau_2$, σ, ρ), which includes the five model parameters (Equation 2) and the two parameters related to the error distribution. We made the standard assumption that the priors are statistically independent. This assumption enabled us to simplify the prior density $[P_0, K_1, \theta, \tau_1, \tau_2, \sigma, \rho]$ as the product $[P_0][K_1][\theta][\tau_1][\tau_2][\sigma][\rho]$. We then assigned priors to $[P_0]$, $[K_1]$, $[\theta]$, $[\tau_1]$, $[\tau_2]$, $[\sigma]$, and $[\rho]$. P_0 is the initial performance of the athlete and we let $[P_0]$ ~ Normal (p_0, σ_{p0}) with hyper-parameters p_0 and σ_{p0} which are later specified. The parameters K_1 and θ have continuous values and express the average change in IAAF scores per unit positive training effect (K_1) and per unit negative training effect ($\theta * K_1$). The specification of prior information about K_1 is challenging due to the inconsistent measurements in performance and training load in different sports. Different performance and training load measurements lead to different K_1 scales. We therefore assigned a flat prior with a large range to K_1 as $[K_1]$ ~ Uniform(0, 10). The interpretation of θ is well understood and expressed via Equation 5. Therefore, we assigned a strong prior to θ as $[\theta]$ ~ Normal(4.137, 6), truncated(1,∞). The parameters τ_1 and τ_2 are time constants that respectively describe the temporal decays of the positive and negative training effects. Values reported in the literature spanned 4 to 169 and 1 to 69 for τ_1 and τ_2, respectively. Based on our experience with the model, we found these ranges to be excessively wide, such that we set the following constraints: $5 < \tau_1 < 60$ and $3 < \tau_2 < 60$. Therefore, we assigned normal priors $[\tau_1]$ ~ Normal $(50, 38)$, truncated (5,60), and $[\tau_2]$ ~ Normal $(13, 12)$, truncated (3,60). The parameter ρ is a correlation coefficient and $0<\rho<1$. The variability of ρ was assigned as $[\rho]$ ~ Beta(10, 1), where $E(\rho) = 0.91$. This reflects the assumption that the performance on day i and day i - 1 are positively correlated. For σ, we assigned the standard Jeffreys reference prior $[\sigma] \propto 1/\sigma$. And for P_0, we assigned $[P_0]$ ~ Normal (1000, 20).

Using the likelihood function (Equation 7), the posterior density was expressed as the following product:

$$[P_0, K_1, \theta, \tau_1, \tau_2, \sigma, \rho | P] \propto f(P|\Theta)[P_0][K_1][\theta][\tau_1][\tau_2][\sigma][\rho] \quad (8)$$

The specified values for the parameters and hyper-parameters are based on information from published studies featuring the IR model. Specifically, we curated studies from our personal libraries and by identifying papers that cited the original Banister et al. (1975) study and Clarke and Skiba (2013). Altogether, we compiled 40 studies, from which we extracted approximately 100 sets of estimated parameters. Of these, 57 parameter sets adhered to the assumptions of the model, which were used to inform the priors (Supplementary information https://github.com/kenp666/IR-model).

2.6 Computation

Commented [DC7]: Ken:
- from Ryan's table, how were model parameter values from grouped data treated compared to values from individual fits?
- Was a weighting scheme used to favour certain parameter sets over others?

Commented [8R7]: Dave, each study was considered as an object and the average/variance of each parameter was calculated. So no weighting scheme was used.
Also, the sample size was missing for some of the studies. We can consider weighting scheme on sample size, the prior mean will change

The posterior density (Equation 8) is complex and intractable, such that it is challenging to gain insights into the model parameters directly. We therefore used Markov Chain Monte Carlo (MCMC) simulation to generate samples of model parameters from the posterior distribution. We implemented MCMC using the JAGS package in R, the code for which is provided as Supplementary Information at https://github.com/kenp666/IR-model. A challenge with implementing MCMC is that the IR model parameters have the following constraints: $0 < K_1 < K_2$, $\theta > 1$, $5 < \tau_1 < 60$, and $3 < \tau_2 < 60$, which cannot be straightforwardly enforced within the JAGS software. Accordingly, an extra step in the sampling procedure was introduced to implement the constraints. For each iteration in the MCMC simulation, we first checked whether the constraints were satisfied. If they were satisfied, then the simulation results were retained; if not, then the generated variates were discarded and the sampling was repeated. This procedure slowed the computation time, such that 1,000 iterations for adaptation and 5,000 iterations were run for each model. The posterior means served as point estimates of the parameters and the lower (2.5%) and upper (97.5%) quantiles of the posterior distributions served as the lower and upper bounds of the 95% credible interval estimates.

Commented [DC9]: Ken & Tim: I reworded this passage to be less vague. Please confirm that the new wording is correct.

Commented [DC10]: This passage is still unclear to me. Is this the "burn in"?

Commented [DC11]: Ken/Tim: please verify wording

Predicted IAAF points, P, can also be generated from the MCMC simulations. The procedure to simulate the IAAF points involved three steps (A Gelman et al., 2013). First, (P_0, K_1, θ, τ_1, τ_2, σ, ρ) are sampled from the posterior distribution. Second, P is sampled from the multivariate normal distribution [$P|P_0$, K_1, θ, τ_1, τ_2, σ, ρ], as expressed in Equation 6, using the sampled values of (P_0, K_1, θ, τ_1, τ_2, σ, ρ) and the training loads as the inputs to the IR model. This process provided a single variate P from the predictive distribution. Third, steps 1 and 2 were repeated to approximate the predictive distribution of P. The prediction of P was iterated 5,000 times, resulting in a distribution of P(t) trajectories. The 2.5% and 97.5% quantiles were computed for P(t) to estimate the 95% prediction interval of IAAF points. Standard diagnostic checks were performed to assess convergence (Andrew Gelman & Rubin, 1992).

Results

Training and performance data

The runner completed 259 workouts, the distances for which were 18.7 ± 12.6 km. The athlete's distances were linearly associated with the TRIMPi values (r = 0.65), and the missing TRIMPi were imputed as described in the Methods (Figure 1).

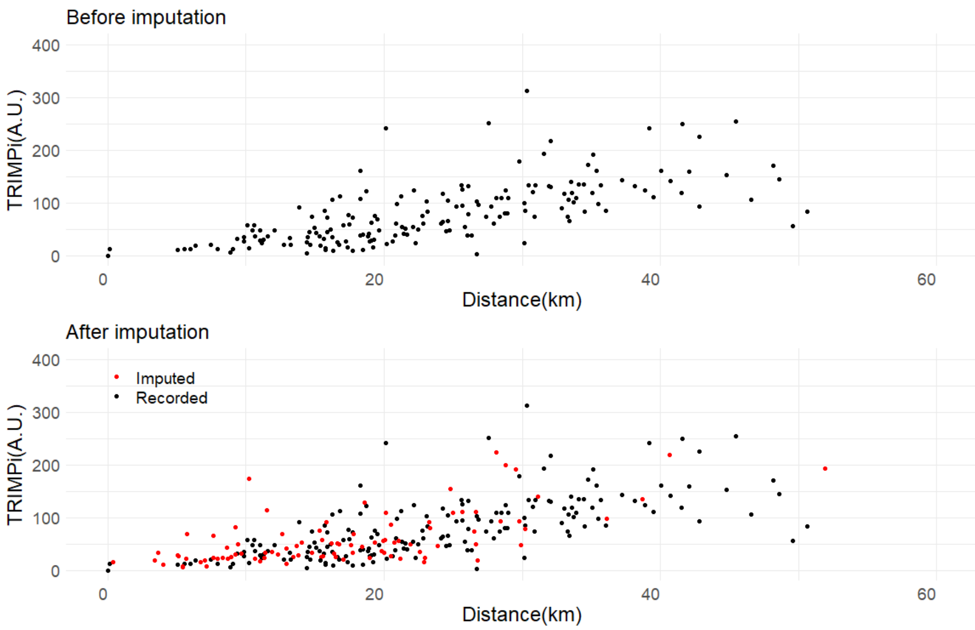

Figure 1. Scatterplots of TRIMPi versus distance run (km). A. Scatterplot of the measured values of TRIMPi and workout distance (km). B. Measured values (black points) overlaid with the imputed values of TRIMPi (red points).

The TRIMPi values (observed and imputed) are plotted by day in Figure 2. The athlete competed in 13 races, 10 of which were 800 m (outdoor), two were 1,500 m (outdoor), and one was 1 mile (indoor). The athlete's IAAF scores ranged from 1,018 to 1,109 points.

Commented [12]: Dave, maybe we want to keep the imputation part simple and don't show this plot? I can color code the Figure 1 to show which days were imputed

Commented [DC13R12]: I don't understand how such scattered red dots arose? If the TRIMPi values were imputed from the linear regression, then they should be located along the regression line, no? Is noise introduced somehow?

Commented [DC14R12]: Change "workout distance" to "distance" and the axis scale to km, not metres. Also, the plot axis titles need units, i.e., TRIMPi (A.U.) and Distance (km).

Commented [DC15R12]: Finally, I noticed you left in some outliers, e.g., the 70 km workout and the 400 TRIMPi workout. Shouldn't these be eliminated prior to fitting a linear regression to be used for imputation?

Commented [DC16R12]: On the top plot, I think we should show the regression line, report the equation, and the R^2 value.

Commented [DC17]: Report race times?

Commented [18R17]: Reported in new graph

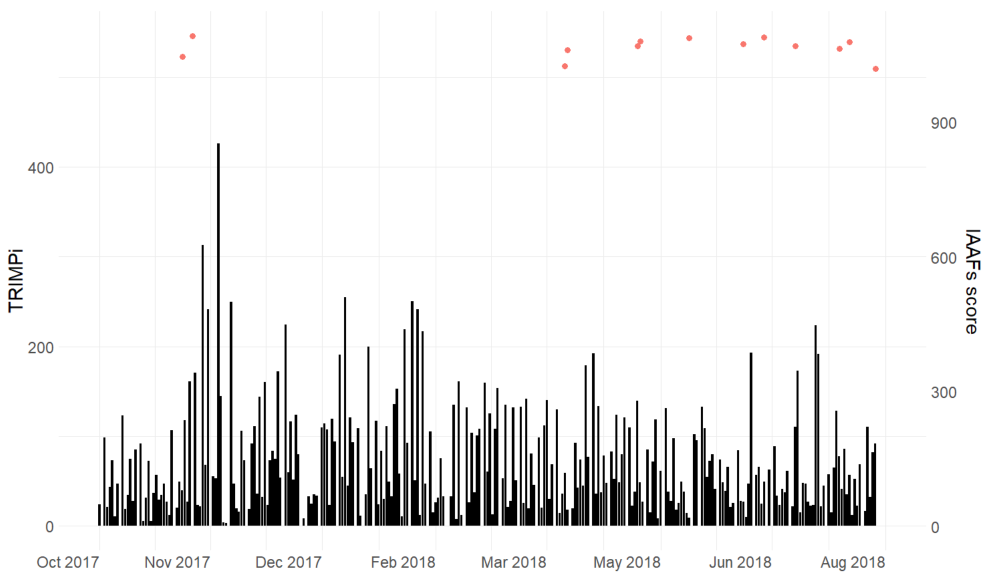

Figure 2. Training load and performances in season 2018. The black bars are the daily training loads (TRIMPi) and the red points are athlete's performances (IAAF score) in 14 races.

Model fitting using non-linear least squares

We first fitted the IR model (Equation 1) using non-linear least squares. We observed that the model-predicted IAAF scores followed the trend of the true IAAF scores (Figure 3), and the method provided plausible estimates for P_0, K_1 and K_2. However, the estimated parameter values featured wide confidence intervals (Table 1). The estimated values of the well-understood parameters t_n and t_g were 84 and 155, respectively. The plausibility of these values of is questionable.

Commented [DC19]: Does this figure include the imputed TRIMPi values?

Commented [DC20]: Increase the font sizes of all text elements in the plot, the text has to be legible for if the width of the plot is reduced by half.

Also, move the IAAF score label to the right so it does not overlap with the numbers. The title can be deleted because it is communicated in the figure legend.

Truncate the plot at the end of the Jul 2018 data. The athlete seems to have taken time off after that point and there are no remaining performances.

Commented [DC21]: Deleted text: Therefore, the parameter estimates based on only 14 IAAF observations are not able to provide practical suggestions for the athlete.

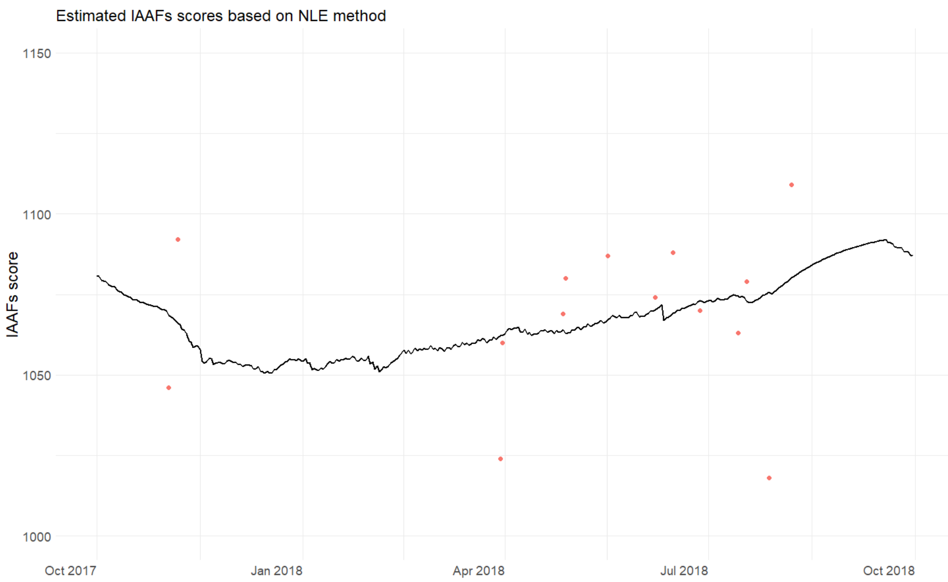

Figure 3. Modeled performance (black line) compared to performance data (IAAF scores, red points). The IR model was fitted using the non-linear least squares method.

Table 1. Estimated parameters and 95% confidence intervals from the nonlinear least-squares procedure.

Parameter	Estimate	95% Confidence interval
P_0	1078	1,022, 1,799
K_1	0.056	-2.82, 16.15
K_2	0.068	-2.80, 16.17
τ_1	77	-731,491, 154
τ_2	65	-5,803, 129
t_n	84	19, 617
t_g	155	-167,864, 334

Model fitting using the Bayesian approach

The Bayesian approach led to predicted IAAF scores and corresponding 95% predictive intervals that captured most of observed IAAF scores. In addition, the fitted curve was smoother and less scattered than the one from the non-linear least square estimates (Figure 4). Model diagnostics are reported in the Appendix.

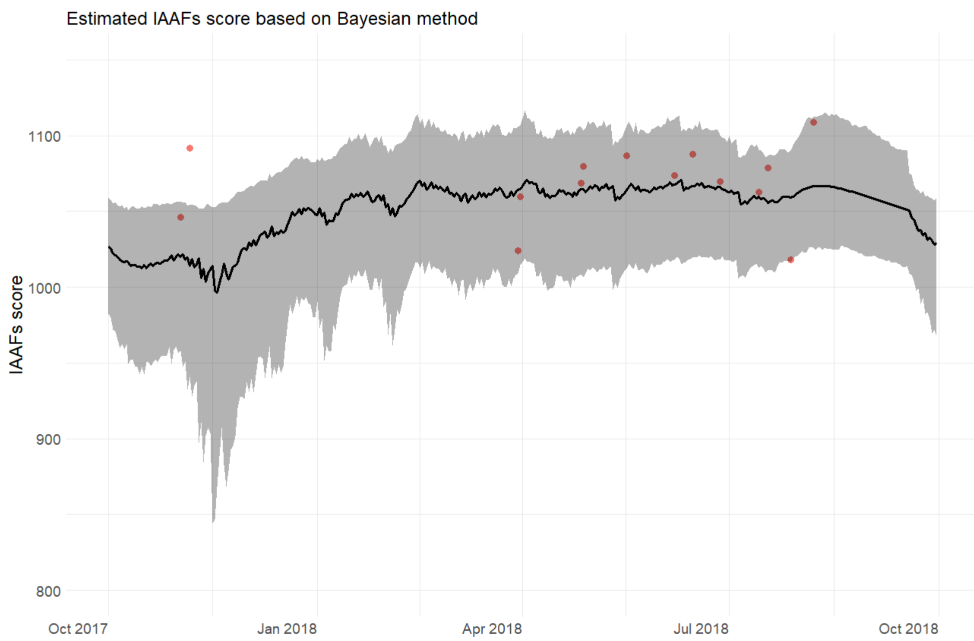

Figure 4. Modeled performance (black line is the point estimates, grey shadow is the posterior predictive interval) compared to performance data (IAAF scores, red points). The IR model was fitted using Bayesian method.

Commented [DC22]: There are still 14 points in this graph, there should only be 13.

The Bayesian approach led to parameter estimates that adhered to the IR model constraints and posterior intervals of reasonable widths (Table 2). The width of the posterior intervals of parameters t_n and t_g were still wide, but their estimates were more believable, and the fitted model can be used to suggest the taper strategy for this athlete.

Table 2. Estimated parameters and 95% credible intervals from the Bayesian version of the IR model.

Parameter	**Estimate (posterior mean)**	**Credible interval (2.5%, 97.5%)**
P_0	1028	986, 1,063

k_1	0.028	0.0024, 0.68
θ	2.75	1.16, 5.30
τ_1	50	45, 55
τ_2	13	9, 16
t_n	16	2, 29
t_g	39	25, 55
ρ	0.83	0.61, 0.97
σ	43	25, 86

Comparing the prior and posterior distributions for parameters P_0, K_1, θ, τ_1, τ_2, σ revealed that the data had a strong influence on the posterior distributions (Figure 5). In particular, the posterior standard deviations of the parameters τ_1, τ_2, and θ were markedly reduced, while the posterior mean of P_0 was shifted to the right. The posterior density of K_1 was narrow despite its less-informative prior, which was flat and had a wide range.

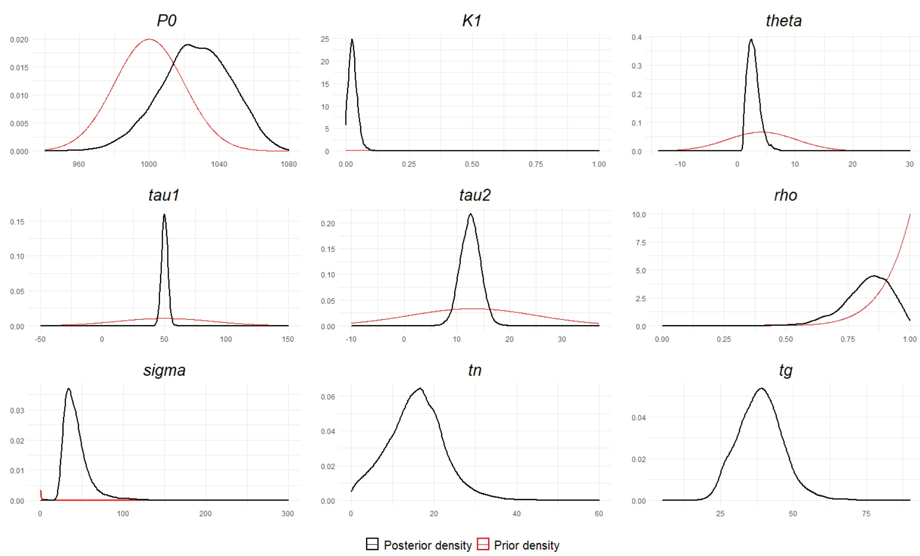

Figure 5. Comparison of the prior (red) and posterior (black) densities for the IR model parameters.

Discussion

In this study, we applied Bayesian methods to the IR model involving an elite middle-distance runner. We made several methodological advances, including a reformulated the model to facilitate the specification

Commented [DC23]: The font size of the plot titles is excessive, while the font size for the legend at the bottom of the figure is way too small.

Commented [DC24R23]: Also, remove the word "Parameter" from the plot titles.

Commented [DC25]: Placeholder text: For example, coaches should maximize the training load of this athlete at about five weeks before the competition; reduce the training load at about two weeks before the competition. Another important piece of information is the estimated value of correlation parameter ρ = 0.84 (0.61, 0.98), this value still suggests that athlete's performances are correlated in time although the posterior mean is lower than the prior mean(0.91).

of informative priors, we compiled published data to specify the priors, and we developed a method that enabled the JAGS software to be used while enforcing logical parameter constraints. We applied the Bayesian approach to the data from a national-class middle-distance runner, and compared the fits to those obtained using nonlinear least squares. Our proof-of-principle results demonstrate that the Bayesian approach method is superior to nonlinear least squares because the estimated parameter values from the former were more precise, well behaved, and believable. Bayesian inference led to actionable insights whereas the nonlinear least squares approach did not.

Bayesian methods offer theoretical and practical advantages compared to frequentist methods such as nonlinear least squares, and are finding increasing use in the sport science literature (Hecksteden et al., 2022; Hopker et al., 2020). First, the nonlinear least squares approach relies solely on the data for fitting models, such that the procedure will work poorly when the data are sparse, the results are highly sensitive to noise. The present data set featured relatively few performances, because middle-distance runners tend to compete sporadically, and there was considerable variability in the data that could have unduly influenced the model fit. Bayesian methods, by contrast, can still be used when data are sparse and can be more robust to noise depending on the strengths of the priors. A second advantage of Bayesian methods is the ability to iterate the procedure as new data become available, which is particularly useful for longitudinal athlete monitoring. In this case, the posterior distribution from the 2018 season could be specified as the prior distribution for the following season. Specifically, the results from Table 2 would be used to specify the priors as follows: $[K_1]$ ~ Normal(0.038, 0.02); [θ] ~ Normal(μ_θ = 2.78, σ_θ = 1.1), truncated(1,∞); $[\tau_1]$ ~ Normal($\mu_{\tau1}$ = 49, $\sigma_{\tau1}$ = 2.6), truncated (5,60); $[\tau_2]$ ~ Normal($\mu_{\tau2}$ = 13, $\sigma_{\tau2}$ = 1.9), truncated (3,60); $[P_0]$~Normal (p_0 = 1025, σ_{p0} = 20); $[\sigma^2] \propto$ inverse gamma(0.001, 0.001); and [ρ]~Beta (10, 1). Note that the P_0 of the following season is the predicted performance from the end of the 2018 season, and the prior for K_1 is specified as a normal distribution. In addition, the widths of the new priors are less than those of the 2018 season and are therefore more informative.

The results of the analysis can be interpreted to provide practical interpretations and advice for the athlete, from both retrospective and predictive standpoints. From a retrospective standpoint, the predicted performance from the Bayesian model demonstrated that the athlete improved in the early part of the season, from December to April, and then maintained their performance level during the competition period (April to July). Such a pattern might be expected if the athlete is competing frequently and has less opportunity for high volumes of training. For this athlete, the training loads during the competition phase may have been insufficient to support an increase in performance in the latter part of the season. From a predictive standpoint, the estimated t_g of 39 days suggests that training quantity and quality should be maximized approximately 40 days before the main competition, which could inform the scheduling of future training camps. The estimated t_n of 16 days suggests that the athlete would most benefit from tapers lasting approximately 2.5 to 3 weeks. The Bayesian model could be used to predict the effects of different training programs simulated as TRIMPi profiles over time (Clarke & Skiba, 2013).

While the Bayesian approach addresses some of the foremost challenges limiting the usefulness of the IR model, particularly those relevant to parameter estimation, it does not overcome all of them. For example, owing to diffuse priors and sparse few data, the prediction intervals may be overly wide for some applications. While the Bayesian approach may be less sensitive to noise in the data, the quality of training and performance data still matters. Here our data set featured missing training data, which we addressed using imputation, but it would have been preferable if we had more complete training data. We also quantified performance using IAAF points, because this metric enables races of different distances to be used as performance data. However, IAAF points for middle-distance races are a function of race time and placing, and such races are not always run as well-paced maximal efforts. Race tactics, such as a conservatively paced first half, can confound the performance data. Further research is needed to propose and evaluate improved performance metrics. Finally, the Bayesian approach does not overcome the theoretical shortcomings of the IR model, such as the assumption that performance is solely a function of training load. That said, the Bayesian approach is general and can be applied to future improved versions of the IR model.

In summary, we have developed here a Bayesian implementation of the Banister IR model. We made several methodological contributions and showed proof-of-principle by applying the approach to analyze the training and performance data from a national-class middle-distance runner. The Bayesian approach outperformed nonlinear least squares and provided actionable insights for the athlete. The Bayesian implementation of the IR model helps to overcome several of the IR model's foremost challenges that have heretofore impaired its practical usefulness.

Acknowledgements

This study was funded by an Own The Podium-Mitacs Accelerate Fellowship to RTB and DCC. The authors have no conflicts of interest to disclose.

References

Banister, E. W., Calvert, T. W., Savage, M. V., & Bach, T. (1975). A Systems Model of training for athletic performance. *Australian Journal of Sports Medicine*, *7*, 57–61.

Busso, T., & Thomas, L. (2006). Using mathematical modeling in training planning. *International Journal of Sports Physiology and Performance*, *1*(4), 400–405.

Clarke, D. C., & Skiba, P. F. (2013). Rationale and resources for teaching the mathematical modeling of athletic training and performance. *Advances in Physiology Education*, *37*(2), 134–152. https://doi.org/10.1152/advan.00078.2011

Efron, B., & Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, 54–75.

Gelman, A, Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis* (3rd ed.). Boca Raton, FL: CRC Press.

Gelman, Andrew, & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 457–472.

Hecksteden, A., Forster, S., Egger, F., Buder, F., Kellner, R., & Meyer, T. (2022). Dwarfs on the Shoulders of Giants: Bayesian Analysis With Informative Priors in Elite Sports Research and Decision Making. *Frontiers in Sports and Active Living*, *4*.

Hellard, P., Avalos, M., Lacoste, L., Barale, F., Chatard, J. C., & Millet, G. P. (2006). Assessing the limitations of the Banister model in monitoring training. *J Sports Sci*, *24*(5), 509–520.

Hopker, J., Griffin, J., Brookhouse, J., Peters, J., Schumacher, Y. O., & Iljukov, S. (2020). Performance profiling as an intelligence-led approach to antidoping in sports. *Drug Testing and Analysis*, *12*(3), 402–409.

Johnson, M. L., & Frasier, S. G. (1985). [16] Nonlinear least-squares analysis. In *Methods in enzymology* (Vol. 117, pp. 301–342). Elsevier.

Manzi, V., Iellamo, F., Impellizzeri, F., D'Ottavio, S., & Castagna, C. (2009). Relation between Individualized Training Impulses and Performance in Distance Runners. *MEDICINE AND SCIENCE IN SPORTS AND EXERCISE*, *41*(11), 2090–2096. https://doi.org/10.1249/MSS.0b013e3181a6a959

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

Silva, R. M., & Swartz, T. B. (2016). Analysis of substitution times in soccer. *Journal of Quantitative Analysis in Sports*, *12*(3), 113–122.

van de Schoot, R., Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M. G., … others. (2021). Bayesian statistics and modelling. *Nature Reviews Methods Primers*, *1*(1), 1–26.

Van de Schoot, R., Kaplan, D., Denissen, J., Asendorpf, J. B., Neyer, F. J., Van Aken, M. A. G., … Van Aken, M. A. G. (2014). A gentle introduction to Bayesian analysis: Applications to developmental research. *Child Development*, *85*(3), 842–860.

Zhang, Z. (2016). Missing data imputation: focusing on single imputation. *Annals of Translational Medicine*, *4*(1).

Appendix

Bayesian model diagnostics

We checked the convergence of MCMC simulation with trace plots (Figure A1). The plot shows that the two chains (black and red) rapidly converged in the first few iterations, and both chains converged to similar estimates. The Gelman Rubin Diagnostic ($\hat{R}$, "shrink factor") value were low for each parameter, and they approached values equal to 1 within approximately 2,000 iterations, which further indicated that the MCMC converged (Figure A2). Therefore, we used the first 2,000 iterations as the "burn in" and retained the last 3,000 iterations as our sample of the posterior.

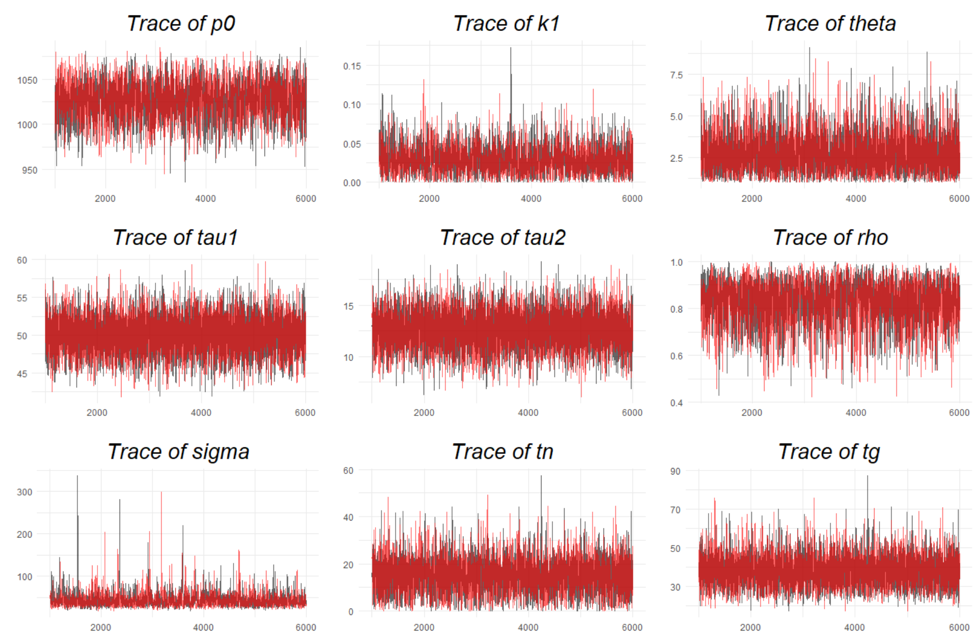

Figure A1. IR model parameter diagnostics. Trace plots for two chains (red and black lines) for the indicated IR model parameters.

Commented [DC26]: I assume that R_hat ought to be expressed this way?

Commented [DC27]: Ken: What is the "bandwidth"?

Commented [DC28R27]: Standardize the order of presentation when reporting parameters: Po, K1, theta, tau1, tau2, rho, sigma, tn, tg. It helps the reader to know what to expect.

Commented [29R27]: Dave, the density curve is smoothed from the histogram. And the bandwidth is related to plotting(how smooth is it).
I suggests not show density curve here. Since they were already shown in result section

Commented [DC30]: What is the y-axis, the parameter values?

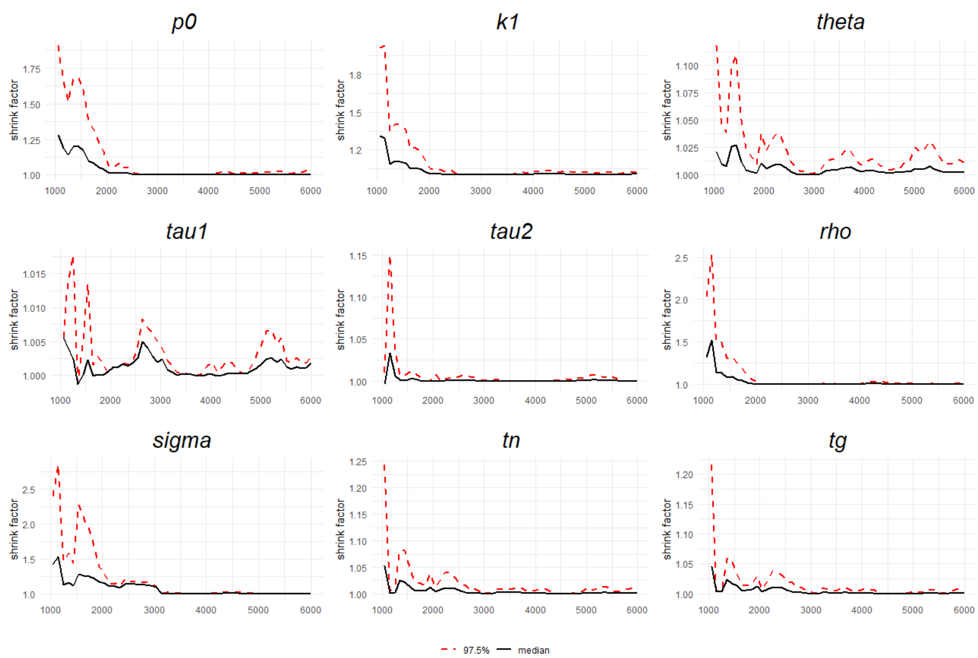

Figure A2. IR model diagnostics. Gelman Rubin Diagnostic ($\hat{R}$, y-axis) are plotted as a function of the iteration number (x-axis). The red dashed line represents the upper bound of $\hat{R}$ (97.5% quantile) and the black solid line represents the point estimates for $\hat{R}$. $\hat{R}$ values equal to ~1 imply that the MCMC algorithm converged.