

Bayesian Approaches for Critical Velocity Modelling of Data from Intermittent Efforts

Ken Peng, Dave C. Clarke and Tim B. Swartz *

Abstract

In sports science, critical power and related critical velocity models have been widely investigated, and are being increasingly applied to field-based team sports. A challenge associated with these models is that laboratory experiments which yield accurate measurements of maximal sustainable velocity are expensive. Alternatively, inexpensive field data (from training and matches) are being used to fit such models. However, the intermittent nature of efforts in field-based sports implies that the dependent variable concerning maximum sustainable velocity is reliably calibrated only for short time durations. This paper develops methods where field data based on short time durations is combined with prior knowledge to fit the three-parameter critical velocity model. This is accomplished in a Bayesian framework for which Markov chain methods are required for model fitting and inference.

Keywords : correlated observations, informative priors, Markov chain Monte Carlo, **three-parameter critical velocity model**, wearable devices.

*K. Peng is an MSc candidate (Department of Statistics and Actuarial Science), T. Swartz is Professor (Department of Statistics and Actuarial Science), and D. Clarke is Associate Professor, (Department of Biomedical Physiology and Kinesiology), Simon Fraser University, 8888 University Drive, Burnaby BC, Canada V5A1S6. Clarke and Swartz have been partially supported by the Natural Sciences and Engineering Research Council of Canada. The work is also partially funded through a Canadian Statistical Sciences Institute Collaborative Research Team (CANSSI CRT). The authors thank Daniel Stenz who provided tracking data used in the construction of Table 1. **We also thank four anonymous reviewers whose comments helped improve the paper.** Approval for the use of these data for research was obtained from the Simon Fraser University Research Ethics Board.

1 INTRODUCTION

Critical power and critical velocity models are widely researched models in sports science. Critical power models investigate the relationship between the maximum intensity (power) that an athlete can sustain in a given time duration versus the time duration. Critical velocity models investigate the relationship between the maximum speed that an athlete can sustain in a given time duration versus the time duration. Knowledge of critical power and critical velocity relationships are useful for optimizing athletic training programs and pacing strategies. Critical velocity and critical power models have been studied across various sports including cycling, running, rowing and swimming (Poole et al. 2016). Critical power models have also been identified as potential tools in the detection of doping in sport (Puchowicz et al. 2018) and have been extended to whole-body exercise (Moritani et al. 1981).

The two-parameter critical power model was first introduced by Monod and Scherrer (1965) who referred to it as the CP model. The CP model has since been utilized in the context of critical velocity where it is known as the CV-2 model. The original CP model has been generalized in various directions to improve model fit. For example, Jones and Vanhatalo (2017) consider activities where exercise intensity varies. In these activities, CP models have been introduced that allow for the reconstitution of power during lower intensity periods. A limitation of the CV-2 model (or any CV model with an asymptote) is that it yields an infinite maximal sustainable velocity as the time duration $t \rightarrow 0$. Another limitation of the CV-2 model is that it implies that athletes can maintain a maximal sustainable velocity for infinite periods of time. Our focus will mainly concern the CV-3 model proposed by Morton (1996) which provides a more realistic relationship between maximum sustainable speed and time duration.

In this work, we are explicit about the statistical assumptions that arise in the context of critical velocity models. In particular, we explore the types of data that arise and how this impacts the reliability of calibration. Statistical independence and correlation are investigated, as well as the distributional assumptions associated with measurement error. All of these topics impact the estimation of parameters in critical velocity models. Incorrect model

assumptions can lead to biased parameter estimates. Statistical insight has proven valuable to problems arising in sports science and sports analytics (Albert et al. 2017).

Careful consideration of the features and limitations of critical velocity data leads to modelling opportunities. A primary contribution of this paper is the development of the CV-3 model in a Bayesian setting. In this case, the Bayesian framework is applied by synthesizing prior knowledge from laboratory settings with “inexpensive” data obtained from training sessions. By inexpensive, we mean that the data are conveniently obtained from athletes during training sessions through the inconspicuous use of wearable devices. Because field-based sports tend to feature intermittent efforts (e.g. repeated sprints interspersed with walking and jogging), it is recognized that maximal speeds when extracted from training data are only reliable for short time durations. Consequently, in our proposed method, only short time duration data are utilized, and the information shortfall is supplemented by prior information gathered from gold-standard laboratory-based testing. Ideally, such data would be obtained from the same group of athletes; however, in the absence of such data, data collected from an independent sample of athletes can be used instead.

Therefore, a main focus of the paper concerns situations where no laboratory measurements are available and critical velocity parameters for an athlete cannot be estimated. Yet, by introducing unintrusive training data, reliable critical velocity parameter estimates may be obtained. Our careful consideration of model assumptions (including the correlations between observations) is taken into account. Another contribution of the paper is the development of a framework to study the adequacy of the “maximal moving average” formula (Lord et al. 2020) as an estimator of maximal sustainable velocity. In Appendix A, we describe how this problem can be addressed via simulation studies where the design of the training session impacts the adequacy of the estimation procedure. We confirm the intuition that the maximal moving average formula does not work well for prolonged durations.

In a sports science application, Peng et al. (2021) have proposed Bayesian methods and elicited priors in the context of the impulse-response model. Clarke and Skiba (2013) emphasize the pedagogical importance of performance modelling in exercise physiology; discussion is provided in the context of critical power models and the impulse-response model.

In Section 2, we discuss how data are typically collected in the context of fitting critical velocity models. We specifically highlight the pitfalls of measuring maximal sustainable velocities for prolonged durations using wearable devices. In Section 3, the CV-2 and CV-3 models are presented from a statistical perspective. The assumptions related to the models are clearly specified, and the impact and relevance of these assumptions are discussed. In Section 4, Bayesian formulations of the CV-3 model are presented. The motivation is that prior knowledge based on laboratory testing can be combined with field data to provide adequate estimates of the CV-3 parameters. Computational and inferential aspects are discussed. The methods are illustrated via an example in Section 5. Section 6 provides a short discussion and outlines future research opportunities.

2 DATA IN CRITICAL VELOCITY MODELS

Critical velocity data are collected in various formats. Gold standard data which are considered the most reliable measurements are obtained in laboratory settings. Such data are naturally time consuming and expensive. **And of course, the reliability of such data is dependent on athletes providing full effort.**

Laboratory data are collected using various schemes including treadmill tests and shuttle run tests. For example, in a shuttle run test, two physical endpoints are constructed m metres apart, and an athlete is requested to run continuously and alternatively, from one endpoint to the other within s seconds. The athlete is asked to continue the exercise as long as possible, and upon exhaustion, completes x traversals (**i.e. from one end to the other**). The response variable $y_t = m/s$ is the maximum speed (measured in metres per second) that the athlete sustains during the time duration $t = xs$ seconds. **For example, suppose that the endpoints are spaced $m = 35$ metres apart. Suppose further that the athlete runs between endpoints in $s = 5.0$ seconds, and accomplishes this $x = 10$ times prior to collapsing. Then the athlete will have run for a duration of $t = xs = 10(5.0) = 50$ seconds, maintaining a maximal sustainable velocity of $y_t = y_{50} = m/s = 35/5.0 = 7.0$ metres/second.**

Due to fatigue, it is intuitive that y_t ought to be modelled as a decreasing function of

t . It is also clear that an athlete cannot be asked to repeat the exercise for different values of t without recovery periods. Therefore, the collection of y_t for different values of t may take place over multiple sessions. This contributes to the expense of gold standard data collection. Laboratory testing also distracts from training, and may not be welcomed by athletes and coaches.

A less expensive data collection method involves the use of wearable sensing systems during training and matches. Prominent sensing systems are based on GPS (global positioning system) technology for which there are various commercial providers. For a given athlete, a “wearable” sensing device provides positioning data from which velocities are obtained. Typically, GPS sensors capture data at a sample rate of 10-20 Hz. Let

$$v_1, v_2, \dots, v_N \tag{1}$$

denote the **instantaneous** velocity data during a training session where the observations are dependent. For illustration, and without loss of generality, v_i in (1) refers to the instantaneous speed of the athlete in metres/second at second i of training, and the training session consists of N seconds. For example, in a 1.5-hour training session, $N = 5400$. Therefore, sensing data may be considered “big data”, and are inexpensive in the sense that athletes simply carry out normal activities. Wearable devices can provide much more information in addition to instantaneous speed measurements. For example, location data, acceleration data and physiological measurements may also be recorded through wearables. A review of various wearable devices including an assessment of measurement accuracy is given by Lutz et al. (2020). Since measurement errors may be present in the recording of instantaneous speeds, pre-processing is recommended in which a maximum velocity threshold is determined for the raw data in (1). The threshold may be set by looking at the data, identifying outliers and assessing what may be reasonable upper bound for a given athlete. Pre-processing techniques for GPS data are discussed in detail by Abbruzzo, Ferrante and de Cantis (2019).

In the context of critical velocity models, an immediate question concerns the relationship between the field measurements v_i in (1) to maximal sustainable velocities. A common

approach defines a response variable y_t referred to as the maximal moving-average (Lord et al. 2020) as

$$y_t = \max \left\{ \sum_{i=1}^t v_i/t, \sum_{i=2}^{t+1} v_i/t, \dots, \sum_{i=N-t+1}^N v_i/t \right\} \quad (2)$$

where again for illustration, $t = 1, \dots, N$ is the time duration measured in seconds. Each component in (2) is an average speed taken over a period of t seconds. What is not entirely clear is why y_t in (2) is thought to be a good approximation of maximal sustainable speed for time duration t . For large values of t , one could realistically imagine an athlete slowing down and taking a rest for portions of every component interval; in this case y_t in (2) underestimates maximal sustainable velocity. On the other hand, for small values of t , one might expect that there is some period of time duration t where an athlete is running at a nearly constant speed, nearly as fast as possible, and the average over the period reasonably captures maximal sustainable velocity. **For example, coaches could design sessions with specific drills that are expected to generate maximal sustainable velocities.** This provides the motivation for the proposed methods that follow. Specifically, we will assume that training (field) data y_t in (2) collected through wearable devices provides reliable information on maximal sustainable speeds for short time durations t . Again, we emphasize that unlike laboratory data, field data are easy to collect.

In Appendix A, we use point processes and simulation studies to explore the adequacy of the maximal moving average formula (2) as an approximation to maximal sustainable velocity. This appears to be the first investigation of the adequacy of this widely used approximation. As anticipated, the formula is not a good approximation for prolonged durations. **An additional message from Appendix A is that the adequacy of the maximal moving average formula deteriorates with activities (matches or training) having frequent rest intervals.**

3 MODEL DEVELOPMENT

Since the CV-2 model (Monod and Scherrer 1965) provides the foundation for all critical power and critical velocity models, we begin by examining modelling issues with respect to this simple model. We then transition to the CV-3 model (Morton 1996) which provides a more realistic description of the relationship between maximum sustainable velocity and time duration. The CV-3 model is **examined** in Section 4, where short-duration field data are supplemented with prior knowledge under the Bayesian framework.

3.1 The CV-2 Model

The simplest of the critical velocity models is a decreasing rational function referred to as the CV-2 model. Using the notation presented in Section 2, we write the CV-2 model for a particular athlete as

$$y_t = \theta_1 + \frac{\theta_2}{t} + \epsilon_t . \quad (3)$$

Using conventional statistical notation, the two unknown parameters in (3) are the quantities of interest and are expressed using the Greek symbols $\theta_1 > 0$ and $\theta_2 > 0$, respectively. The variable y_t is **the observed response**, the time duration t is the observed covariate, and the unobserved ϵ_t term denotes random error. For probabilistic assessment of the model, distributional assumptions concerning ϵ_t need to be introduced. In the above formulation, t is fixed (i.e. not random) and y_t inherits randomness from the error term. Burnley and Jones (2018) provide a detailed discussion of the physiological aspects of model (3). Using laboratory data, Patoz et al. (2021) consider inferential aspects associated with the CV-2 model. In particular, they investigate the bias associated with various weighted least squares fitting procedures.

In the sports science literature, θ_1 is typically referred to as *critical velocity* and is written as *CV*. In the critical power setting, θ_1 is referred to as *critical power* and is written as *CP*. Mathematically, θ_1 is the value of the limiting asymptote corresponding to large time

durations t . Immediately, we see a theoretical failing in the CV-2 model (3), since it is physiologically impossible for athletes to sustain a threshold of performance for unbounded periods of time. For moderate time durations, the critical velocity model may be both practical and adequate.

In the sports science literature, the parameter θ_2 is typically written as D' in the critical velocity setting and W' in the critical power setting. In the critical velocity setting, the expected value of ϵ_t is zero, i.e. $E(\epsilon_t) = 0$, and we can therefore express the parameter θ_2 as a product of time and speed via $\theta_2 = t(E(y_t) - \theta_1)$. It follows that θ_2 has an interpretation as the expected distance that can be travelled in excess of the distance travelled at critical velocity for any time duration t . We also note that in the sports science literature, the recognition and the assumptions concerning ϵ_t are generally not made explicit. Typically, the ϵ_t are assumed to be independent, normally distributed with mean zero and a common variance. However, as previously discussed, with data arising from training sessions, there is dependence in the response variable (2). Such assumptions are important as they impact statistical inference involving the parameters θ_1 and θ_2 .

An immediate limitation of the fitted CV-2 model $\hat{y}_t = \hat{\theta}_1 + \hat{\theta}_2/t$ is that the estimated instantaneous maximal speed of the athlete $\lim_{t \rightarrow 0} \hat{y}_t$ is infinite. Importantly, this limitation is not just a mathematical curiosity corresponding to $t = 0$. If we take a typical estimate $(\hat{\theta}_1, \hat{\theta}_2) = (4m/sec, 200m)$ obtained from Dicks et al. (2018), we observe that even for time durations as large as $t = 20$ seconds, the estimated sustainable velocity for the athlete would be 14.0 m/sec corresponding to an impossibly fast 50.4 km/hour.

3.2 The CV-3 Model

Again, we prefer to use conventional notation for statistical models where the response variable is on the left hand side of the equation and model parameters are denoted using Greek symbols. Doing so and re-parametrizing, Morton's three-parameter CV-3 model (Morton 1996) can be expressed as

$$y_t = \theta_1 + \theta_2/(t + \theta_3) + \epsilon_t . \tag{4}$$

The parametrization (4) provides a simplification of what one typically sees in the sports science literature and aids in the interpretation of parameters. For example, the θ_3 parameter provides a leftward shift of the CV-2 model (3) by θ_3 units in the variable t . Moreover, as $\theta_3 \rightarrow 0$, the CV-3 model reduces to the CV-2 model. In this parametrization, $\theta_1 > 0$ maintains the same interpretation as in the CV-2 model and θ_2 determines the curvature. Although the CV-3 model (4) is not linear in the parameters θ_1 , θ_2 and θ_3 and consequently does not benefit from linear model theory, the parametrization has various appealing properties. First, as logic involving maximal sustainable speeds would suggest, it can be shown that $\hat{y}_t = \hat{\theta}_1 + \hat{\theta}_2/(t + \hat{\theta}_3)$ is a decreasing function of t . Second, the instantaneous maximal speed $\hat{y}_0 = \hat{\theta}_1 + \hat{\theta}_2/\hat{\theta}_3$ is necessarily finite, and unlike the CV-2 model, the CV-3 model provides sensible maximal sustainable speeds for short time durations t . In the parametrization specified in equation (4), we impose the restrictions $\theta_1 > 0$, $\theta_2 > 0$ and $\theta_3 > 0$.

In Figure 1, we demonstrate the role of the parameters θ_1 , θ_2 and θ_3 in (4) by plotting various CV-3 models. We see that θ_1 alters the vertical displacement, whereas θ_2 and θ_3 impact the curvature.

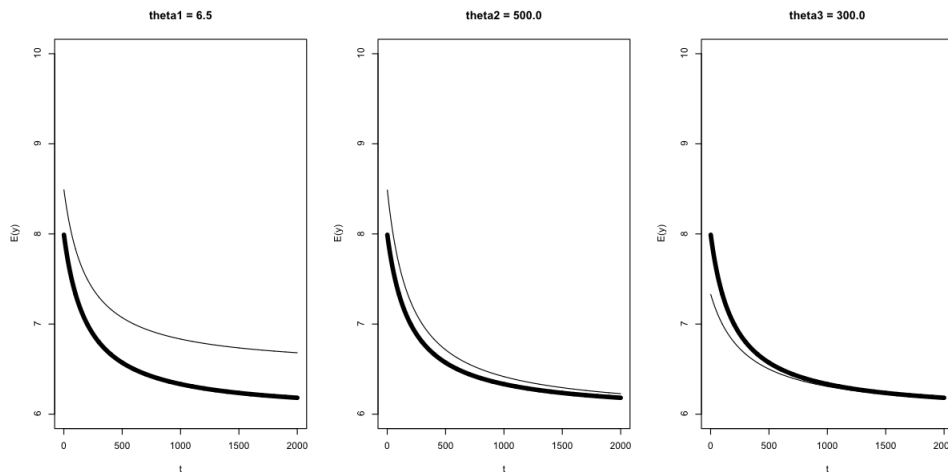


Figure 1: Plots of the expected CV-3 model given by $E(y_t) = \theta_1 + \theta_2/(t + \theta_3)$. The bold line corresponds to realistic settings of the parameters, $\theta_1 = 6.0$, $\theta_2 = 400.0$ and $\theta_3 = 200.0$. The thin lines correspond to a change in a single parameter as indicated by the main captions.

4 A BAYESIAN VERSION OF THE CV-3 MODEL

In Bayesian frameworks, classical statistical models are augmented with prior knowledge of parameters through the use of prior distributions. The synthesis of the two components results in a posterior distribution which gives the complete probabilistic description of the parameters.

4.1 Initial Model

Let $[A | B]$ generically denote the density function corresponding to A given B . Then based on (4), and the introduction of the parameter σ which characterizes variability in the data, the Bayesian paradigm expresses the posterior density of $(\theta_1, \theta_2, \theta_3, \sigma)$ as

$$[\theta_1, \theta_2, \theta_3, \sigma | y_1, \dots, y_n] \propto [y_1, \dots, y_n | \theta_1, \theta_2, \theta_3, \sigma] [\theta_1, \theta_2, \theta_3, \sigma] \quad (5)$$

where y_1, \dots, y_n are field data as specified in (2). **Note that n is the number of maximal velocity observations that have been included in the model.** The variable n is chosen sufficiently small such that y_t accurately measures maximal sustainable velocity for time duration t , $t = 1, \dots, n$. The term $[y_1, \dots, y_n | \theta_1, \theta_2, \theta_3, \sigma]$ is referred to as the likelihood, where the precise form of the likelihood is determined by the distributional assumptions assigned to the ϵ_t terms in (4). The term $[\theta_1, \theta_2, \theta_3, \sigma]$ specifies the prior distribution concerning the parameters θ_1 , θ_2 , θ_3 and σ .

We begin with some strong but standard assumptions involving the initial model. First, we assume that the error terms $\epsilon_1, \dots, \epsilon_n$ are independent with $\epsilon_t \sim \text{Normal}[0, \sigma^2]$. The consequence of the distributional assumption is that the field data $\underline{y} = (y_1, \dots, y_n)^T$ are conditionally independent and that $y_t \sim \text{Normal}[\theta_1 + \theta_2/(t + \theta_3), \sigma^2]$. We further assume that $\underline{\theta} = (\theta_1, \theta_2, \theta_3)^T$ is statistically independent of σ , where we assign Jeffreys reference prior $[\sigma] \propto 1/\sigma$. We assign the multivariate normal distribution $\underline{\theta} \sim \text{Normal}_3[\underline{\mu}, \Sigma]$ where $\underline{\mu} = (\mu_1, \mu_2, \mu_3)^T$ and Σ are specified hyperparameters. **Instead, it would also be possible to consider a truncated multivariate normal distribution where the constraints $\theta_1 > 0$, $\theta_2 > 0$**

and $\theta_3 > 0$ are imposed. Under these assumptions, the posterior density (5) reduces to

$$[\theta_1, \theta_2, \theta_3, \sigma \mid \underline{y}] \propto \exp \left\{ \sum_{t=1}^n -\frac{1}{2\sigma^2} (y_t - \theta_1 - \theta_2/(t + \theta_3))^2 \right\} \exp \left\{ -\frac{1}{2} (\underline{\theta} - \underline{\mu})^T \Sigma^{-1} (\underline{\theta} - \underline{\mu}) \right\} \sigma^{-(n+1)}. \quad (6)$$

There is a remaining detail that needs to be overcome in the specification of the posterior density (6). That is, we require the specification of the mean vector $\underline{\mu}$ and the variance-covariance matrix Σ , which convey prior opinion concerning the parameter vector θ . Imagine that we have estimates $(\hat{\theta}_{11}, \hat{\theta}_{12}, \hat{\theta}_{13}), (\hat{\theta}_{21}, \hat{\theta}_{22}, \hat{\theta}_{23}), \dots, (\hat{\theta}_{m1}, \hat{\theta}_{m2}, \hat{\theta}_{m3})$ for m athletes whose estimates were obtained from gold standard laboratory testing. Recall again that such estimates are expensive but are thought to be reliable. It would be sensible if the athlete in question resembled this population of athletes. For example, if the athlete was a high-level rower, you would prefer laboratory data involving rowers of the same sex, comparable competitive levels and similar ages. In this case, all that one needs to do is average the estimates to obtain $\underline{\mu}$, i.e. $\mu_j = (1/m) \sum_{i=1}^m \hat{\theta}_{ij}$, for $j = 1, 2, 3$. Similarly, $\Sigma = (\sigma_{ij})$ is specified by calculating sample variances and sample covariances; $\sigma_{ij} = (1/m) \sum_{k=1}^m (\hat{\theta}_{ki} - \mu_i)(\hat{\theta}_{kj} - \mu_j)$. **Of course, other approaches for the specification of the hyperparameters $\underline{\mu}$ and Σ are possible.**

4.2 A More Complex Model

We now weaken some of the assumptions introduced in the initial model.

We first recognize that the observations y_1, \dots, y_n obtained through training sessions are not likely to be conditionally independent. Referring to (2), it is **likely** that the interval $v_k, v_{k+1}, \dots, v_{k+t-1}$ which yields y_t overlaps with the interval which yields y_{t-1} . **Consequently**, the field data y_t and $y_{t+\Delta}$ are more likely to be similar when the time difference Δ is small.

This insight may be modelled by introducing greater structure on the likelihood term in (5). Specifically, we model $[y_1, \dots, y_n \mid \theta_1, \theta_2, \theta_3]$ as a multivariate $\text{Normal}_n[\underline{\lambda}, W]$ density where $\lambda_t = \theta_1 + \theta_2/(t + \theta_3)$ and W is defined as a first order autoregressive covariance matrix

whose (i, j) th element is $\sigma^2 \rho^{|i-j|}$ and the correlation between the i th term y_i and the j th term y_j is $\rho^{|i-j|}$. We refer to $\rho \in (0, 1)$ as the correlation parameter. The consequence is that field data y_t and $y_{t+\Delta}$ are more positively correlated when the time difference Δ is small. The proposed approach has the benefit of “smoothing” the data y_1, \dots, y_n . This idea has been used in the analysis of substitution times in soccer (Silva and Swartz 2016). The additional parameter ρ is assumed independent of the remaining parameters and we assign $\rho \sim \text{Uniform}(0, 1)$.

The net effect of the additional modelling is that the posterior now takes the form

$$\begin{aligned}
[\theta_1, \theta_2, \theta_3, \sigma, \rho \mid \underline{y}] &\propto [\underline{y} \mid \theta_1, \theta_2, \theta_3, \sigma, \rho] [\theta_1, \theta_2, \theta_3, \sigma, \rho] \\
&\propto [\underline{y} \mid \theta_1, \theta_2, \theta_3, \sigma, \rho] [\theta_1, \theta_2, \theta_3] [\sigma] [\rho] \\
&\propto \exp \left\{ -\frac{1}{2}(\underline{y} - \underline{\lambda})^T W^{-1}(\underline{y} - \underline{\lambda}) - \frac{1}{2}(\underline{\theta} - \underline{\mu})^T \Sigma^{-1}(\underline{\theta} - \underline{\mu}) \right\} \sigma^{-(n+1)} \quad (7)
\end{aligned}$$

where we note that $\underline{\lambda}$ is a function of $(\theta_1, \theta_2, \theta_3)$ and the matrix W is a function of (σ, ρ) .

4.3 Computation and Inference

Based on the synthesis of data and prior knowledge, the posterior distributions (6) and (7) fully describe the uncertainty in the parameters. However, as written, these densities are complex, are not recognizable as familiar parametric distributions and are not readily interpretable. Therefore, it would be instructive to obtain posterior summaries of the parameters of interest θ_1 , θ_2 and θ_3 . Examples of posterior summaries are posterior means and posterior standard deviations which themselves are not analytically tractable quantities. The remedy is to generate the parameters from the posterior distributions using Markov chain Monte Carlo (MCMC) methods.

Here, MCMC is carried out using JAGS software (Plummer 2003). A feature of JAGS is that the experimenter does not have to derive the full conditional distributions required to implement Gibbs sampling (a basic version of MCMC). Rather, the experimenter needs only to provide the data, and specify both the sampling and prior distributions. With this

information, MCMC variate generation is carried out by JAGS in the background. Upon practical convergence of the MCMC algorithm, the experimenter is provided with posterior samples from the posterior distribution. The samples are then used for inferential purposes. For example, a user may average the samples of a parameter to approximate its posterior mean. It is also possible to create density plots based on the samples to approximate posterior density functions.

5 EXAMPLE

It is good to provide an overview of our problem. Imagine an athlete for whom we want estimates of critical velocity parameters, but there is no laboratory data. Perhaps the best that can be done is to simply reason that **their** parameters are similar to the average values of comparable athletes. Our belief is that we can do better than that. We can take data from unintrusive training sessions and use the reliable short duration measurements to modify the average values. The method for doing this is not arbitrary - it is the well-established Bayesian approach.

The test case that we consider is based on field data arising as in (1) and converted to y_t measurements as specified in (2). In this case, the field data corresponds to a single 90 minute soccer match by an anonymous forward in the Chinese Super League (CSL) during the 2019 season. The data were acquired via cameras and optical recognition software by the provider TRACAB.

The measurements y_t are recorded in Table 1 and correspond to time durations $t = 1, 2, \dots, 15$ seconds. In soccer matches, players do not typically run full out for 15 consecutive seconds. The sport involves frequent starts, changes of direction, and periods of rest. Therefore, we expect the measurements in Table 1 to be less reliable for larger values of t . Note that $y_1 = 9.41$ m/sec nearly represents maximal speed and corresponds to 33.9 km/hr. For reference, the maximum speed attained by Usain Bolt over a 20 metre section at the **IAAF** World Challenge in Zagreb 2011 was 43.7 km/hr (Coh et al. 2018). We utilize short time durations y_1, \dots, y_5 as our measurements in fitting CV-3 models. Note that the y_t are

decreasing, although according to the construction in (2), this does not mathematically need to be the case. **Therefore, with limited data (i.e. $n = 5$ data points), we are attempting to fit useful 3-parameter models. What makes this possible is the Bayesian construction where we supplement field data with informative prior knowledge.** And, we repeat, a main feature of the approach is that the field data are obtained inexpensively.

t	y_t	t	y_t	t	y_t
1	9.41	6	7.17	11	5.62
2	8.92	7	6.69	12	5.42
3	8.50	8	6.37	13	5.38
4	8.08	9	6.12	14	5.34
5	7.67	10	5.88	15	5.28

Table 1: Values y_t from (2) measured in m/sec for a forward during a CSL match.

The next step involved in the implementation of the Bayesian analyses requires the specification of the hyperparameters $\underline{\mu}$ and Σ which are common to both models (6) and (7). Referring to Morton (1996), we can reparametrize the fitted parameters corresponding to the six long distance runners in Table 8 (middle three columns). Then following the discussion at the end Section 4.1, we obtain hyperparameters

$$(\mu_1, \mu_2, \mu_3) = (5.82, 396.00, 198.21)^T \quad (8)$$

and

$$\Sigma = \begin{pmatrix} 0.11 & -22.86 & -8.85 \\ -22.86 & 18891.60 & 13208.13 \\ -8.85 & 13208.13 & 12833.37 \end{pmatrix}. \quad (9)$$

From (8), the prior estimate of critical velocity $\mu_1 = 5.82$ m/sec corresponds to 21.0 km/hr. This value likely overestimates the player's critical velocity and highlights the importance of

having good prior information. Using (8), we also obtain the prior estimate $\mu_1 + \mu_2/\mu_3 = 7.81$ m/sec of instantaneous velocity which corresponds to 28.1 km/hr. Using the relationship between covariances and correlation, we calculate from (9) that there are meaningful prior correlations between the parameters. Specifically, we obtain $\text{Corr}(\theta_1, \theta_2) = -0.50$, $\text{Corr}(\theta_1, \theta_3) = -0.24$ and $\text{Corr}(\theta_2, \theta_3) = 0.85$. Recognizing that reliable prior information is important and that the long distance runners studied by Morton (1996) may differ from soccer players, we introduce a conservative procedure and inflate Σ in (9) using $k\Sigma$ where $k > 0$ is a tuning parameter. An inflated value of Σ introduces greater variability in the prior distribution and consequently puts more inferential emphasis on the data. For illustration, we set $k = 1.5$.

With observations y_1, \dots, y_5 from Table 1 and prior knowledge as expressed in (8) and (9), we fit **three models; the classical model $y_t \sim \text{Normal}[\theta_1 + \theta_2/(t + \theta_3), \sigma^2]$ using maximum likelihood estimation, the initial Bayesian model (6) and the enhanced Bayesian model (7).** For the latter two models, computation was carried out in JAGS based on 20,000 iterations where practical convergence of the chain was quickly achieved. From the MCMC output, 5,000 iterations were discarded as burn-in, and 15,000 iterations were used for inference. The MCMC iterations required less than 1.0 seconds of computation on a laptop computer. Note that a feature of simulation output is that it facilitates estimation of functionals. For example, the instantaneous velocity $\text{MaxV} = \theta_1 + \theta_2/\theta_3$ is a functional of interest as it is a function of the parameters θ_1, θ_2 and θ_3 . The fact that MaxV is relatively complex and does not have an apparent closed-form estimator is unimportant. In an MCMC setting, we can readily estimate the instantaneous velocity MaxV via the generated output of the primary parameters θ_1, θ_2 and θ_3 .

The estimated parameters are provided in Table 2. From Table 2, we immediately observe that maximum likelihood estimation (MLE) based on only 5 observations provides unrealistic CV-3 estimates. For example, the estimate of θ_1 is negative. Therefore, it is apparent that field data on its own with only five data points cannot be used to estimate CV-3 parameters. **It seems that auxiliary information is necessary in this application.** Accordingly, we supplement the field data with prior knowledge in a Bayesian framework leading to the

estimates corresponding to Model (6) and Model (7). An important observation is that the field data impact estimation. Although the parameter estimates in Model (6) and Model (7) are in keeping with prior knowledge, there are differences between the posterior estimates and the prior means given in (8). We also observe that Model (6) and Model (7) provide comparable and believable values. For example, instantaneous maximal velocities of 8.39 m/sec and 8.27 m/sec correspond to 30.2 km/hr and 29.8 km/hr, respectively. Relative to the magnitude of the data in Table (1), the estimates of σ in Model (6) and Model (7) appear large. This may be explained by the discrepancy between prior knowledge and the data, a difference that is reduced with more reliable prior information as specified in (8) and (9). When comparing the estimates involving Model (6) and Model (7), it is difficult to ascertain the truth. However, the correlation $\rho = 0.66$ in Model (7) provides strong evidence that the observations y_1, \dots, y_5 are dependent. This suggests that the enhanced Model (7) is the preferred model. **The JAGS code corresponding to Model (7) in Table 2 is provided in Appendix B.**

Parameter	MLE	Model (6)	Model (7)
θ_1	-3.26	5.92 (000.41)	5.90 (000.38)
θ_2	320.38	365.28 (126.82)	376.79 (137.45)
θ_3	24.29	154.84 (068.14)	175.96 (089.32)
MaxV	9.93	8.39 (000.41)	8.27 (0.68)
σ	0.01	0.84 (000.42)	0.99 (0.64)
ρ			0.66 (0.25)

Table 2: Estimates of CV-3 model parameters (and standard errors) based on maximum likelihood estimation, the initial Bayesian model (6) and the enhanced Bayesian model (7) when $n = 5$ observations y_t are utilized. The parameter $\text{MaxV} = \theta_1 + \theta_2 / \theta_3$ denotes maximum instantaneous velocity.

An important talking point throughout this investigation is that field data do not provide reliable measurements y_t for larger time durations t . To check the sensitivity of the estimates in Table 2, we repeat the analyses based on a larger sample size of $n = 15$ observations y_1, \dots, y_{15} . Table 3 provides the estimation results with the larger sample size. We

now observe that CV-3 estimates are unreasonable in each of MLE, Model (6) and Model (7) contexts. For example, the critical velocity θ_1 is too low under MLE, θ_2 and θ_3 are inconsistently large in Model (6), and MaxV is impossibly large for Model (7).

Parameter	MLE	Model (6)	Model (7)
θ_1	1.04	5.55 (000.36)	5.98 (00.36)
θ_2	106.32	500.70 (169.01)	158.25 (62.53)
θ_3	11.49	316.57 (112.38)	15.94 (04.20)
MaxV	10.30	7.14 (000.30)	15.61 (01.42)
σ	0.19	1.48 (000.30)	4.39 (03.43)
ρ			1.00 (00.00)

Table 3: Estimates of CV-3 model parameters (and standard errors) based on maximum likelihood estimation, the initial Bayesian model (6) and the enhanced Bayesian model (7) when $n = 15$ observations y_t are utilized. The parameter $\text{MaxV} = \theta_1 + \theta_2/\theta_3$ denotes maximum instantaneous velocity.

A drawback of the illustration in this section is that long distance runners (upon which the prior is based) most likely have different critical velocity characteristics than the soccer player of interest. In practice, it is important that the prior describes a realistic probabilistic description of the athlete of interest. We have tried to mitigate the shortcomings of our dataset by using a more diffuse prior (i.e. using $k\Sigma$ instead of Σ with $k = 1.5$). In practice, how might a user select k ? As a tuning parameter, one might consider various choices of $k > 0$ and compare various posterior estimates as reported in Table 2. If the parameter estimates are non-sensical and too precise (i.e. small standard errors), this suggests that the prior is dominating the data and that k needs to be larger. However, if k becomes too large, then the inferences are essentially based only on the data y_1, \dots, y_n , and with small n , it is difficult to reliably fit a three-parameter model. If one was really confident in the prior, then choosing $k < 1$ may be appropriate; this would have the effect of moving posterior estimates closer to prior estimates.

In our Bayesian formulation, the specification of reliable prior information provides a key role, and we have outlined an approach using parameter estimates from other athletes. Of

course, prior information may arise from other sources including expert knowledge, and this knowledge may be incorporated into the prior distributions.

6 DISCUSSION

The main contribution of this paper is the estimation of CV-3 parameters with inexpensive field data. The challenge is that field data are unreliable for larger time durations t . Therefore, only minimal field data are used (based on smaller time durations) where the data are supplemented with prior information.

Moving forward, how might one improve estimation? An obvious strategy involves the improvement of prior knowledge, and this may be accomplished by setting hyperparameters based on athletes who are more similar to the athlete in question. Another avenue is through improved models. We have provided careful investigation of model assumptions including the conditional independence of the data y_t . One may consider alternative models that have been proposed in the critical power literature. Another possibility involves the design of training sessions where athletes are required to move at maximal speeds for longer time durations. This would permit the inclusion of data y_t for larger values of t .

In the context of field data, we have also investigated the use of the ubiquitous maximal moving average formula (2) as an approximation to maximal sustainable velocity (see Appendix A). We have observed that the formula may not be ideal for larger time durations. There is also the suggestion that modified training schedules may help improve the approximation.

7 REFERENCES

- Abbruzzo, A., Ferrante, M. and De Cantis, S. (2019). A pre-processing and network analysis of GPS tracking data. *Spatial Economic Analysis*, 16(2), 217-240.
- Albert, J.A., Glickman, M.E., Swartz, T.B. and Koning, R.H., Editors (2017). *Handbook of Statistical Methods and Analyses in Sports*, Chapman & Hall/CRC Handbooks of Modern

Statistical Methods, Boca Raton.

- Burnley, M. and Jones, A.M. (2018). Power-duration relationship: Physiology, fatigue and the limits of human performance. *European Journal of Sports Science*, 18(1), 1-12.
- Clarke, D.C. and 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.
- Coh, M., Hebert-Losier, K., Stuhec, S. and Babic V. (2018). Kinematics of Usain Bolt's maximal sprint velocity. *Kinesiology*, 50, 100-101.
- Dicks, N.D., Joe, T.V., Hackney, K.J. and Pettitt, R.W. (2018). Validity of critical velocity concept for weighted sprinting performance. *International Journal of Exercise Science*, 11, 900-909.
- Jones, A.M. and Vanhatalo, A. (2017). The 'critical power' concept: Applications to sports performance with a focus on intermittent high-intensity exercise. *Sports Medicine*, 47(Supplement 1), 65-78.
- Lord, C., Blazevich, A.J., Abbiss, C.R., Drinkwater, E.J. and Ma'ayah, F. (2020). Comparing maximal mean and critical speed and metabolic powers in elite and sub-elite soccer. *International Journal of Sports Medicine*, 41(4), 219-226.
- Lutz, J., Memmert, D., Raabe, D., Dornberger, R. and Donath, L. (2020). Wearables for integrative performance and tactic analyses: Opportunities, challenges, and future directions. *International Journal of Environmental Research and Public Health*, 17(1):59.
- Monod, H. and Scherrer, J. (1965). The work capacity of a synergic muscular group. *Ergonomics*, 8, 329-338.
- Moritani, T., Nagata, A., Devries, H.A. and Muro, M. (1981). Critical power as a measure of physical work capacity and anaerobic threshold. *Ergonomics*, 24, 339-350.
- Morton, H. (1996). A 3-parameter critical power model. *Ergonomics*, 39, 611-619.

- Patoz, A., Spicher, R., Pedrani, N., Malatesta, D. and Borrani, F. (2021). Critical speed estimated by statistically appropriate fitting procedures. *European Journal of Applied Physiology*, To appear.
- Peng, K., Brodie, R.T., Swartz, T.B. and Clarke, D.C. (2021). Bayesian inference for the impulse-response model of athletic training and performance.
- Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*, Vienna, Austria.
- Poole, D.C., Burnley, M., Vanhatalo, A., Rossiter, H.B. and Jones, A.M. (2016). Critical power: An important fatigue threshold in exercise physiology. *Medicine & Science in Sports & Exercise*, 48(11), 2320-2334.
- Puchowicz, M.J., Mizelman, E., Yogev, A., Koehle, M.S., Townsend, N.E. and Clarke, D.C. (2018). The critical power model as a potential tool for anti-doping. *Frontiers in Physiology*, 9, Article 643.
- Silva, R. and Swartz, T.B. (2016). Analysis of substitution times in soccer. *Journal of Quantitative Analysis in Sports*, 12(3), 113-122.
- Vandewalle, H. (2018). Modelling of running performances: Comparisons of power-law, hyperbolic, logarithmic, and exponential models in elite endurance runners. *Biomed Research International*, Article ID 8203062, 23 pages.

8 APPENDIX A

In the context of field data, we investigate the adequacy of the widely used maximal moving average formula (2) given by

$$y_t = \max \left\{ \sum_{i=1}^t v_i/t, \sum_{i=2}^{t+1} v_i/t, \dots, \sum_{i=N-t+1}^N v_i/t \right\}$$

as an approximation to maximal sustainable velocity for time duration t . In (2), v_i refers to the instantaneous speed of the athlete measured in metres/second at second i of training. In a training session of duration N seconds, we therefore have measurements v_1, v_2, \dots, v_N .

We begin the investigation under the assumption that the CV-3 model (4) is an accurate description of maximal sustainable velocity. Therefore, we define the expected maximal sustainable velocity for time duration t as

$$E(y_t) = \theta_1 + \theta_2/(t + \theta_3) . \quad (10)$$

Our approach will compare realizations of formula (2) via simulation experiments with the benchmark expected maximal sustainable velocities (10). If there is agreement, it is suggestive that the maximal moving average formula (2) is a good proxy for maximal sustainable velocity.

Our realizations from formula (2) arise from a behavioural model based on a point process intended to resemble a training session. In this behavioural model, we assume that athletes exert maximal effort (i.e. velocity) for tasks - training sessions may be designed accordingly. Further, we assume that athletes approach each task so that nearly constant effort is applied throughout the task. In this framework, we have m start-stop intervals for the tasks where m is odd, $s_0 = 1$, $s_{m+1} = N$ and the times $s_1 < s_2 < \dots < s_m$ are generated according to a point process. This leads to active task periods

$$(s_0, s_1), (s_2, s_3), (s_4, s_5), \dots, (s_{m-1}, s_m)$$

and rest periods

$$(s_1, s_2), (s_3, s_4), (s_5, s_6), \dots, (s_m, s_{m+1}) .$$

Suppose now that time $i \in (s_j, s_{j+1})$. **According to the CV-3 sampling model (4) with independence of y_1, \dots, y_n ,** and following the behavioural assumptions, we randomly generate

the instantaneous speed v_i according to

$$v_i \sim \begin{cases} 0 & \text{j odd (rest period)} \\ \text{Normal}(\theta_1 + \theta_2/(t_i + \theta_3), \sigma^2) & \text{j even} \end{cases} \quad (11)$$

where $t_i = s_{j+1} - s_j$.

The generated v_i 's in (11) then permit the calculation of the maximal moving average y_t given by (2). However, we are interested in the general properties of the maximal moving average formula. Therefore, we repeat the simulation procedure M times. Specifically,

(1) generate $s_1, s_2, \dots, s_m \rightarrow$ generate $v_1, \dots, v_N \rightarrow$ calculate $y_t^{(1)}$

(2) generate $s_1, s_2, \dots, s_m \rightarrow$ generate $v_1, \dots, v_N \rightarrow$ calculate $y_t^{(2)}$

.

.

.

(M) generate $s_1, s_2, \dots, s_m \rightarrow$ generate $v_1, \dots, v_N \rightarrow$ calculate $y_t^{(M)}$

The iterative procedure allows us to calculate the average

$$\bar{y}_t = \frac{1}{M} \sum_{k=1}^M y_t^{(k)} \quad (12)$$

for $t = 1, 2, \dots, N$.

The simulation procedure attempts to set realistic values. For the CV-3 model, we set $\theta_1 = 6.0$, $\theta_2 = 400.0$ and $\theta_3 = 200.0$ (see equation (8)). We also consider two settings of the CV-3 parameter $\sigma = 0.1$ and $\sigma = 0.5$. Note that σ takes into account variability in measurements due to the athlete and also variability in the measurement device. We let $N = 5400$ which corresponds to a 1.5 hour training session and we let $m = 181$ and $m = 21$ which leads to training sessions with 90 and 10 tasks, respectively.

The point process is the component of the simulation for which different settings may be chosen. We illustrate with a very simple point process. In particular, we generate m independent points according to the discrete uniform distribution on $(1, N)$. We then reorder the points such that $s_1 < s_2 < \dots < s_m$. Although one particular set of points (s_1, s_2, \dots, s_m) may lead to unrealistic start-stop times, the average pattern over $M = 100$ iterations is intended to be reasonable. Of course, one may consider generating alternative point processes according to particular training sessions. For example, one may introduce constraints in the point-process such that there is a minimum rest period between tasks. Part of our intention in this exploratory analysis is the development of a framework for which the maximal moving average formula may be investigated.

Again, we are interested in the comparison of the simulated \bar{y}_t values in (12) according to the behavioural model and the expected sustainable velocity values $E(y_t)$ given by (10). Figure 2 provides plots of $E(y_t) - \bar{y}_t$ versus t under the four settings $(\text{tasks}, \sigma) = (90, 0.5)$, $(90, 0.1)$, $(10, 0.5)$ and $(10, 0.1)$. The simulations required approximately 13 hours of computation on a laptop computer. When the differences are tightly scattered about zero, this is evidence that the maximal moving average formula (2) provides accurate estimation of maximal sustainable velocity. When σ is large (i.e. $\sigma = 0.5$), we observe that \bar{y}_t tends to overestimate the intended estimand $E(y_t)$ by roughly 1.0 m/sec when t is small. This may occur when we have a measuring device that is inaccurate. The best approximations occur when there are few tasks (i.e. 10) and σ is small (i.e. $\sigma = 0.1$). In this case (the bottom right hand plot), the maximal moving average seems to be good for reasonably wide intervals of t , say $1 < t < 500$. As anticipated, it is clear in all four cases that the approximation is inadequate for large values of t and can be biased by roughly 2.0 m/sec. The explanation is that athletes tend to have periods of rest that reduce the \bar{y}_t values. Consequently, increased numbers of periods of rest tend to negatively affect the approximation. A message here is that training sessions ought to be designed to elicit good approximations.

Finally, the reader may notice that whereas the simulation study with $(\text{task}, \sigma) = (10, 0.1)$ provides reasonable approximations for the interval $t \in (1, 500)$ seconds, our example in Table 1 used duration intervals $t = 1, \dots, 5$ seconds. The reason for the large discrepancy is that

the proposed behavioural model in Appendix A was based on a training session with the prescribed point process. In the example, data was collected from an actual match where players rarely sprint continuously for extended periods.

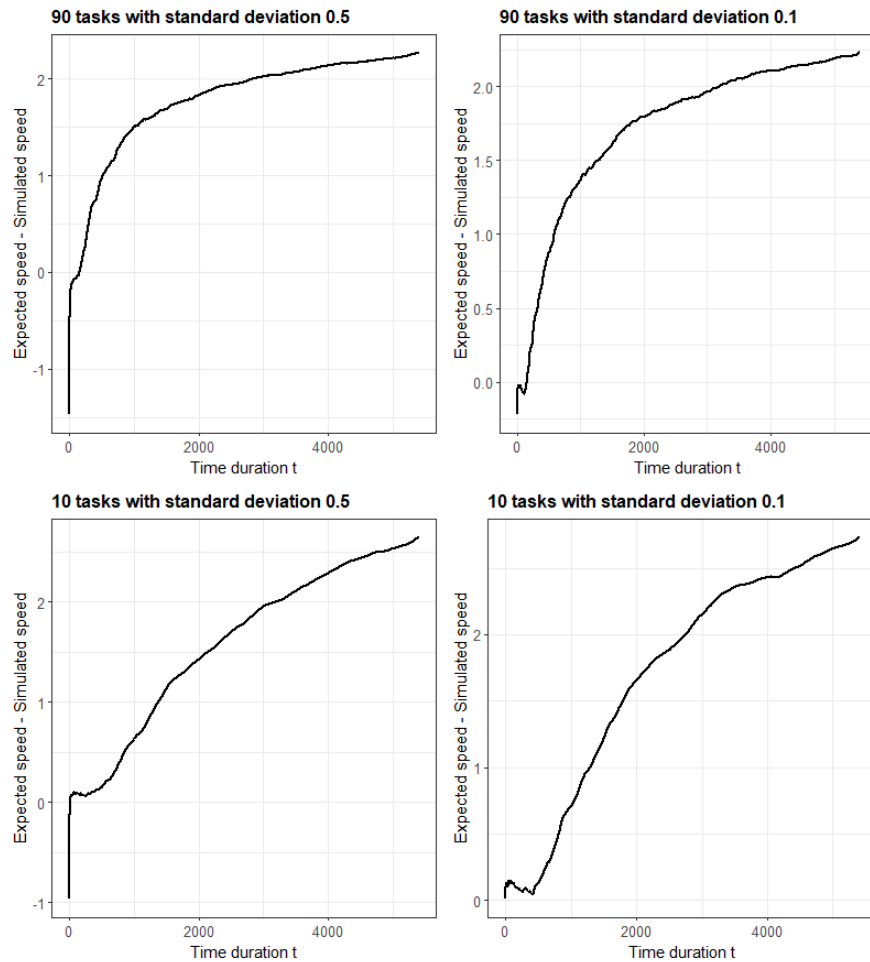


Figure 2: Plots of $E(y_t) - \bar{y}_t$ versus time duration t .

9 APPENDIX B

Below we provide the JAGS code used in fitting the preferred model (7) associated with the data and output described in Table 2. The simpler models described in this paper are easily deduced from the code corresponding to model (7).

```
library(R2jags) # download JAGS at https://sourceforge.net/projects/mcmc-jags/files/
library(MASS)

##### enter the data in Table 1 #####
time <- 1:5
yt <- c(9.41, 8.92,8.50, 8.08,7.67)

##### produce model (7) estimates given in Table 2 #####

# define the parameters and the data:
n = length(yt)
data = list("yt" =yt,"t"=time, "n" = n) # identify data

# model definition:
cat(paste("model{

yt[1:n] ~ dnorm(mu[1:n],s[,])
  for (i in 1:n) {
    for (j in 1:n){
      s1[i,j] <- (1/tau)*(pow(pp,abs(i-j)))
    }
    mu[i] <- theta[1]*t[i]/(t[i]+theta[3])
              + theta[2]/(t[i]+theta[3]) + theta[1]*theta[3]/(t[i]+theta[3])
  }
  s[1:n,1:n] <- inverse(s1[,])

theta[1:3] ~ dnorm(thetamu[1:3],thetasigma1[,])
k = 1.5 # tuning paramter in Section 5
```

```

# prior means and covariances from equations (8), (9) in Section 5
thetamu[1] <- 5.815
thetamu[2] <- 396
thetamu[3] <- 198.212
thetasigma[1,1] <- 0.11163
thetasigma[1,2] <- -22.862
thetasigma[1,3] <- -8.846309
thetasigma[2,1] <- -22.862
thetasigma[2,2] <- 18891.6
thetasigma[2,3] <- 13208.13
thetasigma[3,1] <- -8.846309
thetasigma[3,2] <- 13208.13
thetasigma[3,3] <- 12833.37
thetasigma1[1:3,1:3] <- inverse(k*thetasigma[,]) # JAGS uses the inverse of the variance
pp ~ dunif(0,1)
tau ~ dgamma(1.0E-3, 1.0E-3) # JAGS doesn't allow improper priors
sigma <- 1/sqrt(tau)
Smax = theta[1]+theta[2]/theta[3]
      }",sep = ""), file="CV3Model2.txt")

# random initial values:
inits <- function () {list(theta = mvrnorm(1,c(5.815,396,198.212),
      matrix(c(0.11163,-22.862,-8.846309,-22.862,18891.6,13208.13,-8.846309,13208.13,12833.37),
      nrow = 3)), pp = runif(1, min = 0, max = 1),tu = 1)}

# generate sample from posterior:
jags.m <- jags.model(file = "CV3Model2.txt",data = data, inits=inits, n.chains=3,n.adapt = 1000)
post_sample = coda.samples(jags.m,
      c("theta","sigma","Smax","pp"), # which parameters to be monitored
      n.iter = 20000)

summary(window(post_sample, start=16001,end = 20000))

```