Estimation of self-excitation branching processes models in assessing subsequent injuries.

John Worrall[1,2]*, Paul Pao-Yen Wu[1,2†], Liam Toohey[3†] and Kerrie Mengersen[1,2†]

*Corresponding author(s). E-mail(s): john.a.worrall@hdr.qut.edu.au;

Contributing authors: p.wu@qut.edu.au; Liam.Toohey@ausport.gov.au; k.mengersen@qut.edu.au;
†These authors contributed equally to this work.

Abstract

Maximum likelihood estimation of self-exciting processes is often unstable and biased given limited observations. We explore alternative methods by viewing the estimation as a branching process and propose a computationally efficient latent variable formulation in the context of a novel application to modelling subsequent sporting injuries. By using the challenging problem of determining which injuries are connected, either directly or indirectly, we demonstrate that this causal structure can be found probabilistically parametrically. We employ methods to examine the evidence for the existence of subsequent injury clustering within professional Australian Football League(AFL) players, and highlight the broader benefits of using such methods for self-exciting point processes in a small sample context.

Keywords: Hawkes process, Point processes, reinjury, injury forecasting, stochastic clustering, Expectation-Maximisation, Markov chain Monte Carlo.

1 Introduction

A mathematical model introduced by Alan G. Hawkes (1971) describes the sequential arrival of events as a non-Markovian process with a self-exciting nature, where events occur with frequencies fluctuating over time increases the likelihood of further events. The Hawkes process (HP) has wide application, such as seismology (Ogata, 1981; Rasmussen, 2013); crime analysis (Yang et al., 2018; Zhuang & Mateu, 2019); traffic incidents (Kalair et al., 2020; Li et al., 2018); terrorism (Porter & White, 2010; White et al., 2012); finance (Bacry et al., 2015); infectious diseases (Browning et al., 2021; Kelly et al., 2019); and social media trends (Hall & Willett, 2016; Zhang et al., 2020).

In a HP, the self-exciting nature of the data is modelled through the conditional intensity function which governs the expected arrival rate of events. An important characteristic of

*[1]School of Mathematical Sciences, Queensland University of Technology, Brisbane, Australia, [2]QUT Centre for Data Science, [3] Australian Institute of Sport, Bruce, Australia

this intensity function is the triggering kernel, where there has been much recent research in representing these triggering kernels, including investigation of the underlying assumptions defined through parametric functions (Chen et al., 2021; Chiang et al., 2021). This has motivated new and improved estimation methods including expectation maximisation (EM), stochastic declustering/reconstructing algorithms (Marsan & Lengliné, 2008; G. O. Mohler et al., 2011; Zhou et al., 2020; Zhuang & Mateu, 2019) and new Bayesian perspectives lending themselves readily to handling parameter uncertainty and support for smaller sample contexts (Donnet et al., 2020; Markwick, 2020; Ross, 2021). In this paper, we further improve estimation algorithms by combining the branching structure of the HP to both efficiently estimate parameters and provide a casual structure for reinjury in sport.

Injuries, irrespective of their severity, can trigger other injuries. The initial injury to a predisposed athlete may be succeeded by several subsequent injuries, which in turn may activate their own local injury rate, resulting in a type of self-excitation process. Determining the triggering of these cascading injuries and their connectivity is a complex and challenging problem, with evolutionary characteristics, clustering effects, and interactions between individual players collectively having a great effect on injury rate. In team sports, injuries have a significant influence on seasonal performance (Drew et al., 2017) and incur a substantial financial cost. The causes of sport injuries are multifactorial with characteristics such as injury history, age, current fitness, and level of performance as factors (Bittencourt et al., 2016; Meeuwisse et al., 2007). In addition, characteristics of individuals differ between players, with internal and external factors known to affect the risk of injury to an athlete. Where some players may therefore have lower than average injury rates, while others have higher than average ones. Controlling for these known risk factors is vital in injury prevention and of particular interest in sport science, where existing frameworks such as training load and performance time are informed through the use of statistical models(Bahr & Holme, 2003; Hägglund et al., 2005; J. J. Stares et al., 2019).

Previous published analyses of the frequency of injury incidents have employed linear regression models(J. Stares et al., 2018; Zou, 2004) and generalised mixed models Casals et al. (2015) with the latter author accommodating for overdispersion with random effects to model injuries per season of wrestler with the number of matches per season as offset and both weights and winning as the explanatory variables. These authors compare the performance of parameter estimation via likelihood, extended likelihood, and Bayesian analysis to highlight estimation methods in the epidemiology of sport injuries. However, none of these papers consider the causal structure or the long term injury dependency incurred through the recurrence of injury, furthermore within sports epidemiology, the studies on reinjury has not addressed as a Poisson process. For more accurate modelling, we consider the Hawkes process for each individual's injuries and its long-term effects for subsequent injury.

In this study we focus on the use of efficient declustering estimation methods to first solve the causal structure in self-exciting point processes and secondly model parameter uncertainty within small sample problems through Bayesian approaches. With these methods, we provide a novel application to injury modelling that provide suitable analysis and prediction for future individual players' seasonal injuries thereby presenting the self-exciting point process as a good model for the reproducing injury rates. The paper is organised as follows. Section 2 introduces the Hawkes process and its branching interpretation, along an application to an injury dataset from a professional Australian Football League (AFL) club. This is followed by Section 3 defining the estimation methods. Firstly a stochastic declustering approach is described, by constructing an unobservable branching variable with the direct use of background and triggering probabilities. Secondly, an efficient Gibbs scheme is defined which exploits the conditional independence via introduced latent variable and results in weakly dependent uncertainty parameters estimates for rare events. Finally, an adaptive

parallel Griddy Gibbs method is outlined, that improves on the previous scheme with less computational requirement and automatic weighting for improved approximation. Results are presented in Section 4 and is further discussed in Section 5.

2 Self-exciting process

We can consider injuries as a point process with no overlapping events, characterised by a conditional intensity function, $\lambda^*(t)$. The conditional probability density $f(t|\mathcal{H}_t)$ and the corresponding cumulative distribution function $F(t|\mathcal{H}_t)$ of next arrival defines $\lambda^*(t)$, with $*$ borrowed from Daley & Vere-Jones (2003) used to represent conditioning on the history up to time t. An equivalent and more intuitive definition is the expected number of events per unit time, modelled as

$$\lambda^*(t) = \frac{f^*(t)}{1 - F^*(t)} = \lim_{h \to 0} \frac{\mathbb{E}[N(t+h) - N(t)|\mathcal{H}_t]}{h}. \tag{1}$$

In quantifying the degree of self-excitation, the following Hawkes process expression is designed to capture the influences of all previous events in a conditional intensity value. The univariate conditional intensity function is thus expressed,

$$\lambda^*(t) = \mu + \int_0^t \phi\,(t-s)\,dN(s) \tag{2}$$

where $\mu > 0$ is the non-negative background intensity and $\phi(\cdot) : (0, \infty) \to (0, \infty]$ is the excitation function. Although the structure of the conditional intensity function is quite flexible, a simpler parametric exponential kernel, $\phi(\tau) = \alpha\beta e^{-\beta(\tau)}$ where $\alpha, \beta, \tau > 0$ is assumed to support the low number of events and model complexity. This is the case for the following sport injury study further described below. The homogeneous Poisson component is then represented by constant μ, the excitation term α and the decay component β . Letting the subscript p indicate player

$$\lambda_p^*(t) = \mu_p + \sum_{t > t_i} \alpha\beta e^{-\beta(t_i - t_j)} \quad . \tag{3}$$

With this parameterisation the HP can be viewed as a Poisson branching process (Alan G Hawkes & Oakes, 1974), where each event is a parent and gives birth to an offspring according to a non-stationary process with intensity $\phi(\tau)$. The critical parameter (branching ratio) is then given as the average number of direct offspring per ancestor $\rho = \int_0^t \phi(s)ds$, with a the stationary Hawkes process as $\rho < 1$. The model require the following assumptions,

Assumption 1 (Stationarity)**.** The process has asymptotically stationary increments and a *kernel* that satisfies the assumption condition of stability.

Assumption 2 (IID)**.** The process has event occurrences that follow a independent and identically distributed sequence of Poisson random variables.

2.1 AFL injury data

Australian football is a contact sport played over four 20 minute quarters between two teams consisting of 18 players (4 interchange players). The AFL preseason training phase is generally between late November and February with the competition phase lasting from March to August and an additional four weeks of finals in September. As a collision sport, the league and clubs monitor injury incidences closely with particular focus on prevention.

Injury data were collected from a single AFL club over four seasons(2016-2020), with all consenting participants anonymised before analysis. Ethical approval was obtained from the Human Research Ethics Committee (RA/4/1/6362). Figure(1) shows the frequency of injury occurrences per day during the study period with all 63 participants (left) and the frequency of associated injuries (right). It is notable that 84% of the players have recorded six or more injuries. The distribution of players inter arrival injuries times is shown in figure(2).

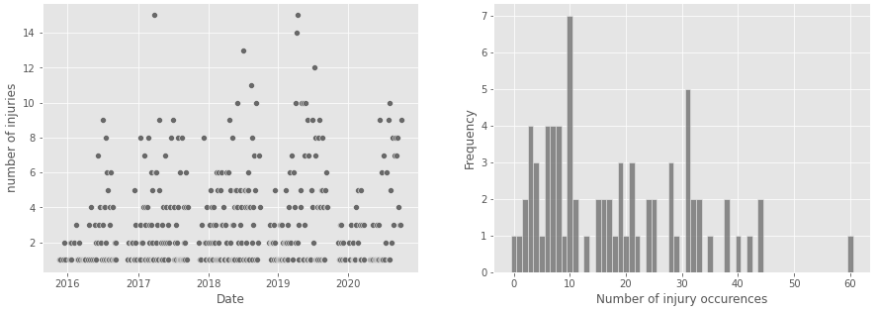

Figure 1 Daily injury events (left) and number of injury occurrences(right) for AFL players.

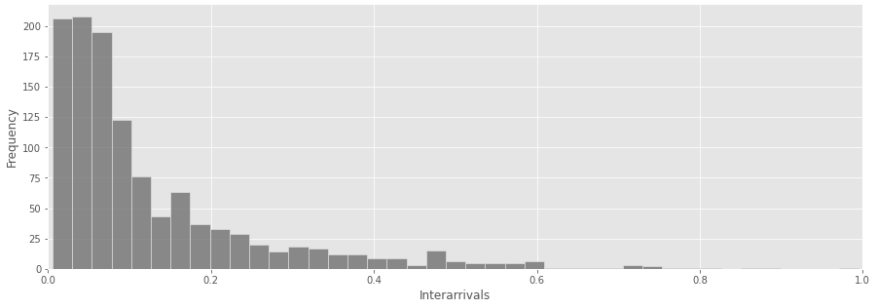

Figure 2 Distribution of interarrival injury times for all players.

3 Estimation methods

The HP model parameters can be estimated through the EM algorithm and MCMC schemes where the likelihood function is derived.

Likelihood function

The self-exciting process was implemented, as outlined above, to model kernel and background intensity with likelihood for event given as,

$$p(t_1 \ldots, t_n|\ \mu, \phi) = \prod_{i=1}^{N} \left(\mu + \sum_{t_j < t_i} \phi(t_i - t_j) \right) . \exp \left(- \int_0^T \mu + \sum_{t_i < t} \phi(t - t_i) dt \right) \tag{4}$$

in which is a vector of parameters $\Theta = \{\mu, \alpha, \beta\}$ to be estimated on $[0, T]$, with observations as realisation of $t_1, ..., t_n$.

3.1 Stochastic Declustering

In an attempt to differentiate between 'true' background events and triggered events the Model Independent Stochastic Declustering (MISD) method derives the causual structure via latent branching variable. Early work by Marsan & Lengliné (2008) defined a homogeneous background constant μ and later extended for the more general case of varying $\mu(t)$ (Lewis & G. Mohler, 2011).The representation of the HP as a branching process reduces both baseline and triggering kernel into a density estimation problem. The augmented probability of observations $(t_1, ..., t_n)$ and branching structure B with two independent components is given by

$$p(t_1, ..., t_n, B|\mu, \phi(\tau)) = \underbrace{u^{b_{ii}} \exp(-uT)}_{u} \cdot \underbrace{\prod_{i=2}^{N} \prod_{j=1}^{i-1} \phi(t_i - t_j)^{b_{ij}} \prod_{i=1}^{N} \exp \left(- \int_0^T \phi(\tau) d\tau \right)}_{\phi(t)} . \tag{5}$$

The recovered parameters are then updated via stochastic declusting (expectation) step, where p_{ij} is replaced by the expectation $\mathbb{E}[p_{ij}] = b_{ij}$, representing the probability that event i is caused by event j

$$p_{ij}^k = \frac{\alpha^k \beta^k e^{-\beta^k (t_i - t_j)}}{u^k + \sum_{j=1}^{i-1} \alpha^k \beta^k e^{-\beta^k (t_i - t_j)}} , \qquad p_{ii}^k = \frac{u^k}{u^k + \sum_{j=1}^{i-1} \alpha^k \beta^k e^{-\beta^k (t_i - t_j)}} \tag{6}$$

This allows for the construction of a P matrix(7), with events caused by the background rate (diagonal elements) or another event (non-diagonal elements).

$$P = \begin{bmatrix} p_{11} & 0 & \ldots & \ldots & 0 \\ p_{21} & p_{22} & 0 & \ldots & 0 \\ p_{31} & p_{32} & p_{33} & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ p_{d1} & p_{d2} & p_{d3} & \ldots & p_{dn} \end{bmatrix} \tag{7}$$

where $\sum_{j=1}^{i} p_{ij} = 1$ as event i must be caused by previous events or itself.

The maximisation step then updates parameters with current matrix of probabilities such that,

$$u^{k+1} = \frac{\sum_{i=1}^{n} p_{ii}^k}{T}, \quad \alpha^{k+1} = \frac{\sum_{i>j} p_{ij}^k}{n}, \quad \beta^{k+1} = \frac{\sum_{i>j} p_{ij}^k}{\sum_{i>j} (t_i - t_j) p_{ij}^k} \tag{8}$$

Iteration is continued until convergence of parameter estimates with causal structure represented via a lower triangular matrix. Finally, branching variable are determine given the expected probabilities, where

$$b_{ii} = \begin{cases} 1 & \text{if event i is a background event} \\ 0 & \text{if event i is a offspring event} \end{cases}, \quad b_{ij} = \begin{cases} 1 & \text{if event i is caused by event j} \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

3.2 Gibbs Scheme

In the previous described method, applying superposition principle of Poisson processes where Hawkes process is viewed as many different non-homogeneous Poisson processes layered up on each other. For each event t_i a new Poisson process with intensity is spawned with intensity $\phi(t-t_i)$ between $[t_i, T]$. Under this same principle we generalise event classification into being spawned from the background rate μ or from another event during the excitation. Branching latent variables in equation(9) allows the likelihood to be split and providing an efficient MCMC sampler. Event times being partitioned into the following appropriate sets (Ross, 2021) $S_0, ...S_n$ where S_0 contains all events from the background and S_j contains offspring from event t_i.

$$S_j = \{t_i, b_{ij} = j\}, \quad 0 \leq j < n,$$

From equation(5) the introduce sets from branching structure can be rewritten (conditional on knowing events from background and excitation) as

$$p(t_i, .., t_n|B, \theta) = e^{-\mu T}\mu^{|S_0|}\prod_{j=1}^{n}\left(e^{-\alpha\Phi(T-t_j)}\alpha^{|S_j|}\prod_{t_i\in S_j}\phi(t_i - t_j)\right). \tag{10}$$

where $\Phi(t) = \int \phi(t)dt$. Therefore given the latent variable B the MCMC scheme sequentially samples three parameters (μ, α, β) that are weakly dependent.

Background

For all events in S_0, sample a new value μ from first term of equation(10),

$$p(t_i, .., t_n|B, \theta) = e^{-\mu T}\mu^{|S_0|}.$$

The conjugate Gamma(α_μ, β_μ) prior allows for direct samples from the posterior conditional distribution as,

$$p(\mu|t_i, .., t_n, B, \theta) = Gamma(\alpha_\mu + |S_0|, \beta_\mu + T) \tag{11}$$

Excitation

For all events in S_j, sample a new value α from middle term of equation(10) ,

$$\begin{aligned} p(t_i, .., t_n|B, \theta) &= \prod_{j=1}^{n} e^{-\alpha\Phi(T-t_j)}\alpha^{|S_j|} \\ &= \alpha^{\sum_{j=1}^{n}|S_j|}e^{-\alpha\sum_{j=1}^{n}\Phi(T-t_j)} \\ p(t_1, .., t_n|B, \theta) &= \alpha^{\sum_{j=1}^{n}|S_j|}e^{-\alpha\tilde{\Phi}}, \end{aligned}$$

where Φ is the cumulative distribution of the normalised probability density $\phi(t)$. Defining our prior the α as a gamma distribution and exploiting the conjugate gives the posterior conditional distribution of α as

$$p(\alpha|t_1,..,t_n,B,\theta) = Gamma\left(\sum_{j=1}^{n}|S_j|\alpha_\alpha, \tilde{\Phi}+\beta_\alpha\right) \tag{12}$$

where α_α, β_α are the hyperparameters in the prior for α. In practice $\tilde{\Phi} \approx n$ as $\Phi(T-t_j) = 0$ for the events that are far away from the boundary.

Decay

For all events in S_j, sample a new value β from the last term of equation(10),

$$\begin{aligned} p(t_1,..,t_n|\theta) &= \prod_{j=1}^{n}\prod_{t_i\in S_j}\phi(t_i-t_j) \\ &= \beta^{N_{S_j}} exp^{-\beta\sum_{i=1}^{N_{S_j}}\tau_i}, \end{aligned}$$

where $\tau_i = t_i - t_j$ is the time of a offspring events relative to its parent for all events with a parent that isn't the background. The conjugate is a gamma distribution and therefore

$$p(\beta|\tau_1,..,\tau_n,B) = Gamma\left(N_{S_j}+\alpha_\beta, \sum_{i}^{n}\tau+\beta_\beta\right) \tag{13}$$

where $\alpha_\beta, \beta_\beta$ are the parameters of the prior.

Algorithm

Using these posterior distribution, the parameters of the Hawkes process can be updated from its full conditional distribution. Gibbs sampler is given below.

Algorithm 1: Sampling a Hawkes process using the latent variables

Data: $t_i, ..., t_n$
Result: N samples of μ, α, β, B
initialization;
while $i \leq N$ **do**
  Sample new latent variables B (Eq.9);
  Sample new μ (Eq.11);
  Sample new α (Eq.12);
  Sample new β (Eq.13);
  i++ ;

3.3 Adaptive Griddy Gibbs

Our motivation for an adaptive parallel griddy-Gibbs (APGG) sampler is to improve the previous Gibbs scheme with respect to efficiency and mixing properties, by extending the original griddy Gibbs (Ritter & Tanner, 1992) to incorporate parallel computation and latent variables in the approximation.

The original griddy-Gibbs scheme approximates a cumulative distribution function either as a piecewise linear or piecewise constant function of the full condition distribution based on a grid of points. The inverse transform method is used to generate random variables with approximately the right distribution. The scheme is summarised as follows,

1. Define the range of θ_j be dividing into G grid points $\theta_j^1, \ldots, \theta_j^G$.
2. Evaluate the full conditionals density $p(\theta_j^g|.)$ at each grid point to obtain weights, $w_j^1, \ldots, w_j^G$ where $w_j^g = p(\theta_g|.)$ for $g = 1, \ldots, G$
3. Normalise w_j^g to obtain $\tilde{w_j^g} = \frac{w_j^g}{\sum_{g=1}^G w_j^g}$ and cumulative sum $\tilde{W_j^g} = \sum_{h=1}^g w_j^h$ defined as the empirical CDF $F(\theta_j^g)$.
4. Sample $x \sim U\,(0,1)$ and use the approximate inverse CDF method to obtain parameter via $x \leq F(\theta_j^g) \implies F^{-1}(x) \leq \theta_j^g$.

Our hybrid implementation samples the background intensity parameter μ via APGG while the remaining parameters (α, β) obtained via the Gibbs sampler. We improve on both step 1 and 2, firstly dividing weighted grid points non-uniformly via adaptive quadrature and setting further limits on grid points through the latent variable. Secondly, evaluations at $p(\mu|.)$ are done independently and simultaneously via parallel computation.

4 Results

This section presents results for the subsequent injury case study described in section 2.1 assumed self-exciting parameters, θ=(μ, α, β and B). Firstly, the estimated parameters of MISD are discussed in the context of better understanding the effects of direct triggering injuries. This is followed by an efficient MCMC approach to estimating rare events of injuries at an individual player level and supporting uncertainty in parameter estimates. Next, an adaptive griddy Gibbs method improves on previous scheme in terms of convergence through adaptive grid and in parallel execution. Finally, the fitted model is used to forecast future player injuries within the next season.

4.1 MISD

Each player is fitted to the Hawkes process using the MISD algorithm described in section 3.1 to estimate each conditional intensity function. Summary results are provided in table(1) including the diagonal mass of the probability matrix P, estimated background rate, average number of offspring events, and average number of offspring events occurring in the first two months for players exhibiting self-excitation.

The following 14 players have the leading estimated subsequent injuries, although the majority of events are probabilistically treated as background events, with background rate

making up roughly 52% to 98% of observed injuries. The estimated background rate for players 12, 13, 16, 2, 32, 34, 40, 41, 47, 52 and 6 are estimated to be between 0.20 to 0.64 injuries events per month while the background rate for players 54, 42, 25 are substantially larger with an estimated monthly rate of injuries of 0.97, 0.94 and 0.81 respectively.

Players	Diagonal Mass	Background Rate	No. of offspring	No. of offspring in 2 months	Per. of offspring in 2 months
Player 12	92.42%	0.6422	0.087	0.0435	50%
Player 13	81.76%	0.4517	0.2258	0.129	57.14%
Player 16	51.67%	0.2023	0.6667	0.6667	100%
Player 2	87.68%	0.4886	0.125	0.0938	75%
Player 25	87.43%	0.8116	0.1163	0.1163	100%
Player 32	84.77%	0.2826	0.2	0.2	100%
Player 34	83.5%	0.535	0.1081	0.1081	100%
Player 40	62.71%	0.3062	0.3333	0.2857	85.71%
Player 41	85.04%	0.4432	0.1333	0.1	75%
Player 42	80.23%	0.9352	0.1282	0.0769	60%
Player 47	78.46%	0.3404	0.4167	0.2083	50%
Player 52	97.92%	0.5334	0.0833	0.0417	50%
Player 54	96.9%	0.9667	0.0333	0.0333	100%
Player 6	66.29%	0.3964	0.2667	0.1333	50%

Table 1 Numerical summaries of MISD algorithm for the leading 14 players exhibiting subsequent injuries. *Diagonal mass represents the percentage (Per.) of matrix probabilities on diagonal. Number (No.) of offspring is the estimated number of injuries triggered from previous injuries (subsequent), the number of offspring in 2 months are the estimated number of subsequent injuries within 2 months of an injury and percentage of offspring in 2 months is the estimated percent of an injury total offspring which occur within the first 2 months after an injury.*

For player 13, the model estimated the expected number of offspring per injury event to be approximately 0.23 injuries with 0.13 of those injuries occurring in the first two months. This then implies that player 13 has 57% of subsequent injuries likely to be within the first two months and the remaining 43% of injuries occurring sometime afterward. Player 16 shows the largest number of offspring events at 0.67 injuries with all subsequent injuries occurring in the first 2 month (see figure (4), player 13 & 16 model and fit). However for the remaining players, the overwhelming majority of injuries are triggered by the background rate.

In reviewing estimated parameters (Fig.3 and Tbl.3), it is noted that players 42 and 54 demonstrate the largest background intensity, μ, while player 6 and 47 have notably high branching ratios ($\int \phi(\tau)d\tau$) indicating endogenous (subsequent) injuries or influence of the kernel on overall conditional intensity function. For reference, n $>$ 1 indicates an injury typically triggers at least one extra injury and n=0 reduces to a Poisson process with constant intensity. Finally, by examining the decay rate of injury β, we see that player 32 exhibits the fastest return to a homogeneous Poisson rate and player 47 has the slowest relative recovery.

We review the MISD performance, particularly for support with small sample sizes. Table (11) compares the conventional maximum likelihood approach to the MISD (EM type) algorithm. Both maximization of the likelihood and MISD algorithm can be influenced by

a poor choice of starting values; therefore presented results are estimates of 100 runs to mitigate this issue. As seen, the likelihood, AIC and BIC reported metrics on all players show superior results with the MISD approach. This is particularly noticeable for players that have fewer observed injuries (e.g Players 64, 32, 20, 16) as the maximum likelihood approach has an increased bias compared to the MISD algorithm, due to the requirement for substantially larger sample sizes in order to attain the asymptotic regime guaranteeing consistent estimation.

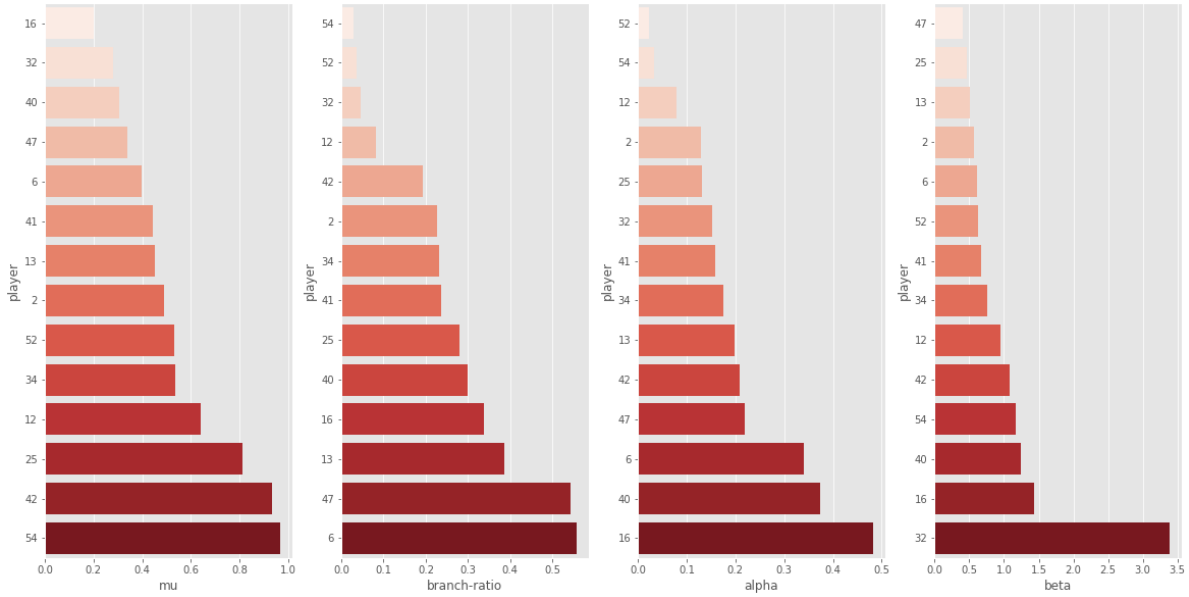

Figure 3 Players rank parameters

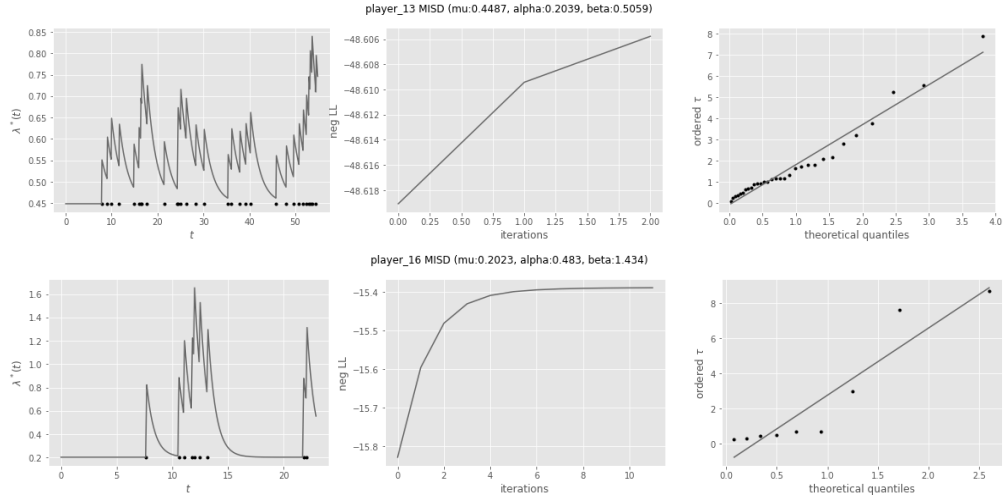

Figure 4 Player 13(above) and 16(below) estimated conditional intensity(left), iteration(middle) and QQ-plot(right)

4.1.1 Summary

The MISD algorithm for the Hawkes process was presented and fitted to data, modelling AFL players subsequent injuries. Stochastic declusting of the Poisson process was applied to the context of self exciting injuries, where a number of highlighted players were shown to exhibit this behaviour, with triggering and background probabilities were reported. In addition, robustness and accuracy of the MISD algorithm was shown and compared against ML, with superior estimates demonstrated specifically where limited data are available and the asymptotic regime is not reached in numerical the maximisation.

4.2 Gibbs

Building on incorporating the latent model formulation shown in the previous stochastic declustering method, the following MCMC scheme aims to further support smaller size samples while providing full posterior parameter distribution. Performance of algorithm is shown with parameter estimates of all players, table(4) and figure(5) report posterior means and standard deviations.

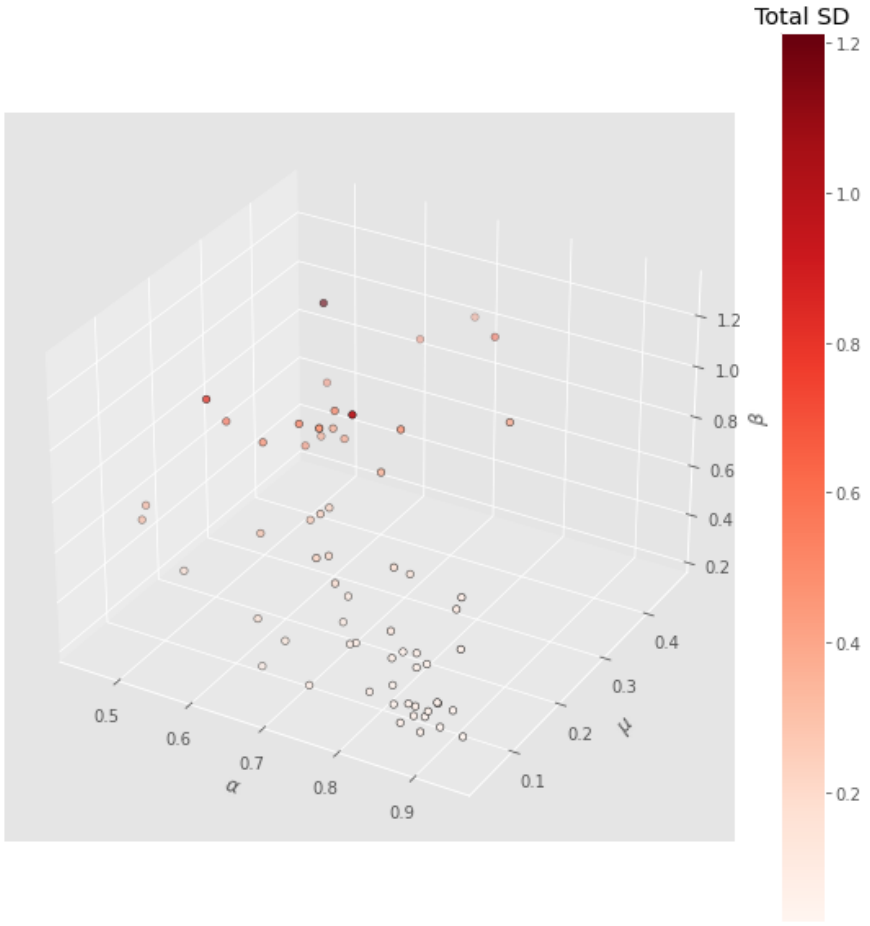

Figure 5 All players(63) parameters, total standard deviation.

Evaluation of model priors are first intended as non-informative, where parameters (μ, α, β) are tried on specification of gamma(1,1). However, due to observed injuries small sample sizes, posterior estimates can be sensitive to specification of the prior distribution. Therefore, we

now specify priors intended to be suitable, shown below are two players (17 and 22) with the priors on θ specified as gamma(1,1) compared against gamma(2,5) and gamma(3,10) for of varying degrees of informality. Players 17 has 28 observed injuries while player 22 has 9 injuries with both over approximately 30 months, highlighting the extremes of players injuries. Player 17 supporting diagnostics are included in the appendix, including MCMC runs(Fig.12), parameters correlation (Fig.13) and forecast (Fig.14). Figure(8) illustrates the full conditional Gibbs sampler, is fitted to injury model for all players. Highlighted below are player 17 and 22 MCMC trace, autocorrelation, parameter spread and forecast.

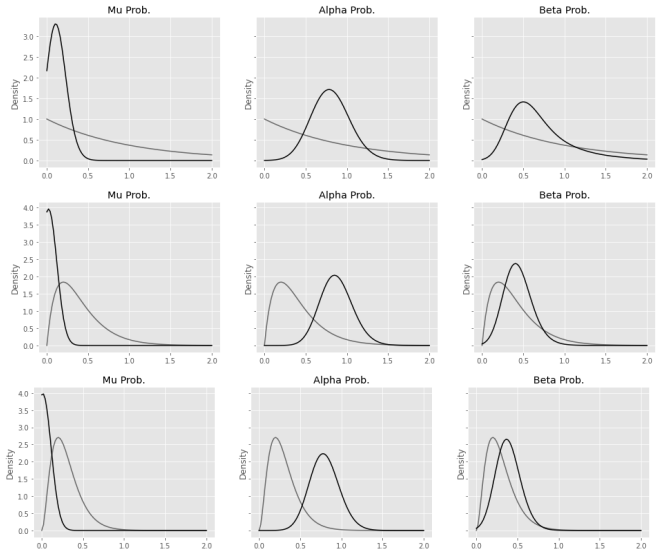

Figure 6 Players 17, posterior(black) distributions (μ, α, β) and prior(grey) gamma (1,1), gamma (2,5), gamma (3,10) rows top, middle and bottom.

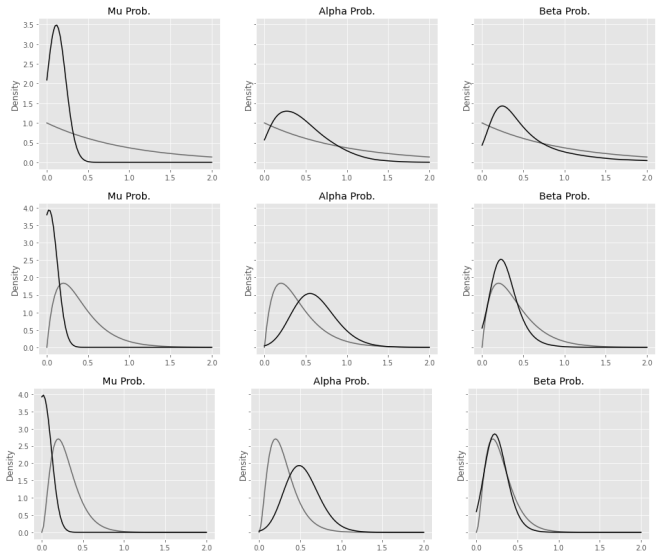

Figure 7 Players 22, posterior(black) distributions (μ, α, β) prior(grey) gamma (1,1), gamma (2,5), gamma (3,10) rows top, middle and bottom.

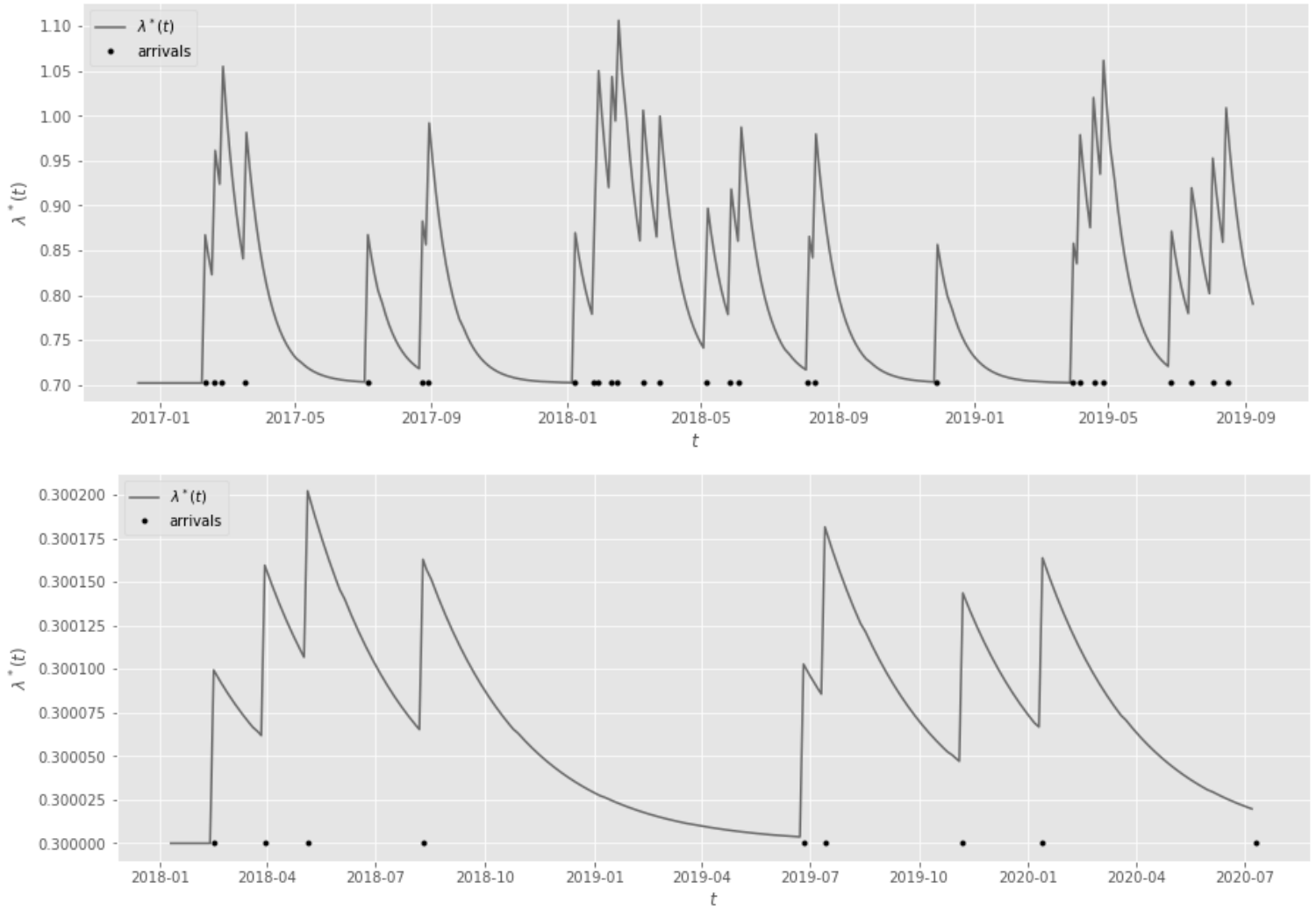

Figure 8 Players 17 and 22 estimated conditional intensities.

4.3 Griddy Gibbs

To demonstrate the improved performance of our APGG sampler, we compare estimates obtained under previous Gibbs scheme with the synthetic dataset, then the parameter estimates with AFL injury records. We evaluate the two sampling approaches for the desirable properties of efficiency and mixing leading to faster convergence. The inefficiency factor metric is defined as

$$1 + 2\sum_{s=1}^{n} \rho(s) \tag{14}$$

where $\rho(s)$ is the sample autocorrelation at lag s from sampled values and is a measure of the ratio of numerical variance of the sample posterior mean to the variance of the sample mean from the hypothetical uncorrelated draw (Chib, 2001). Table(2) compares this inefficiency factor and computational time of MCMC simulations. Posterior inference is empolyed after burn-in on MCMC chains of lengths (10000, 20000, 50000) where adaptive grid are set with local error tolerance of 0.001, spanning ranges [0,1.5].

	Gibbs				APGG			
iter	**time**	**mu**	**alpha**	**beta**	**time**	**mu**	**alpha**	**beta**
n = 30								
10000	39.6	1.23	1.21	1.24	54.85	1.23	1.19	1.03
20000	97.77	1.02	1.22	1.16	124.79	1.01	1.0	0.88
50000	406.6	1.03	1.11	1.13	449.54	1.02	1.05	0.92
n = 50								
10000	94.46	1.07	1.28	1.12	101.91	1.12	1.05	1.05
20000	204.58	1.25	1.29	1.18	223.52	1.0	1.02	1.08
50000	702.04	1.08	0.99	1.1	712.07	0.97	1.02	1.1
n = 60								
10000	24.44	1.14	1.03	0.96	41.76	1.02	1.23	0.93
20000	67.2	1.02	0.91	0.95	98.58	0.93	0.83	0.81
50000	341.23	1.05	0.99	1.08	371.71	1.03	1.04	0.98

Table 2 Inefficiency factors and convergence time.

Reported inefficiency factors show that in almost all cases APGG was superior with some notable results. With n=30 all inefficiency factors obtained for APGG are smaller than Gibbs. Increasing n to 60 results in μ with larger differences although some slightly higher values in α where MCMC chain cycle of 10000 and 50000. As such, we can see APGG general performance improvement overall as run length increases, leading to faster convergences rate.

Runtime results are on CPU Intel Xeon 2, with APGG overall computation time increasing in spite of parallelisation and latent variable approximations. However, numerical experimentation shows in most cases as the iterations increase the computation difference decreases, for example in n=60 chain 20000 compared to 50000 maintains similar differences (appx 30 secs) despite additional computational cost. Summarising where observation and iterations are not small AGGS shows suitability, however further improvement are expected where either are increased.

Table(5) contains the posterior means and standard deviations all players. Low standard error from each of the parameter show favorable estimation results, moreover in comparison to

the previous MCMC scheme similar parameter convergence are obtained despite approximate nature of Griddy Gibbs. Figure(9) highlights player 17 comparable parameter estimates as kernel density plots, indicative of the report values.

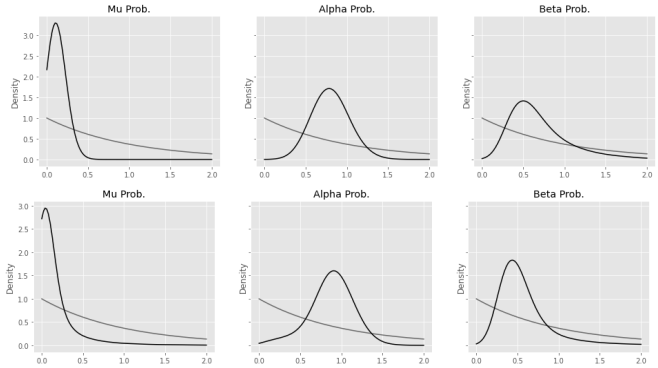

Figure 9 Players 17 posterior(black) μ, α, β and prior(grey) gamma (1,1). Gibbs (above) and APGG (below).

.

4.4 Forecasting

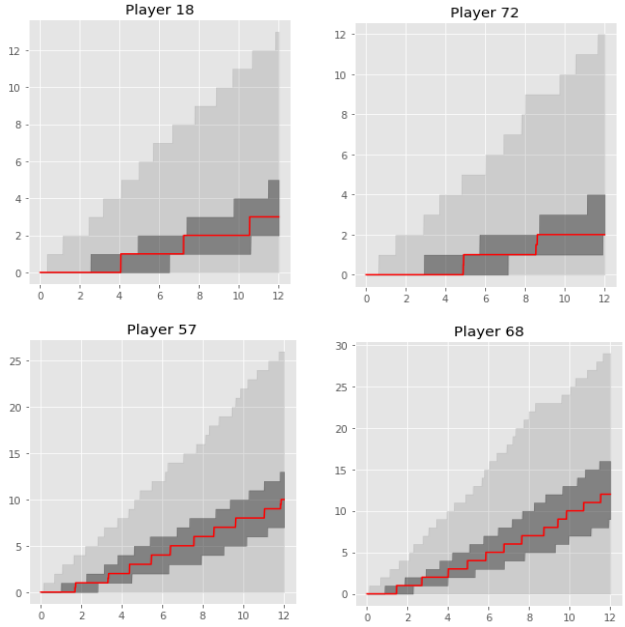

Figure 10 Forecast(12 months) players injuries for 4 players. Shaded areas corresponds to 68 and 95 percent credible regions and solid line represents medians of predictive density.

The fitted model can be used to forecast future player injuries. Under the latent variable formulation, each player's injury profile is inferred up to 12 months with 100 simulations from Gibbs estimated conditional intensities. Figure(17) shows the number of injuries by months with estimates of percentile regions shaded and median(red) plotted for all players. The majority of players have a median of 1-4 injuries in the following season (typical players 18 and 72, top row Fig.10); however some notable players (bottom row Fig.10) have substantially larger numbers of predicted values(10-12).

5 Discussions and Conclusion

In this paper, we presented an MISD algorithm that maximizes the expected complete data log-likelihood function with an introduced branching variable. The advantages of this algorithm compared to conventional ML estimation are substantial. The first advantage is uncovering temporal causal structure to support subsequent events. The proposed estimation method is shown to yield more accurate results in terms of convergence, bias, and robustness to limited observations, given conventional ML asymptotic regime requirements. The MISD latent procedure when converges was shown provide estimate of subsequent latent structure that may be especially attractive for inferring causal structure of injuries and related statistics.8

In contrast to the above point estimate that ignores inherent parameter uncertainty, we apply a Bayesian latent method. This approach does not rely on asymptotics and addresses the hindrance of this property given the small samples in injuries. Performance of efficiency and accuracy of Gibbs scheme is demonstrated and discussed. We investigate the sensitivity of the Bayesian model to different choices of prior distribution. For gamma prior on μ, α and β, we specify non-informative and construct informative to varying degrees, specific for the use of injury modelling.

Finally, an adaptive parallel Griddy Gibbs approach was implemented, showing good mixing of MCMC and efficiency under a parallel environment. With the choice of grid points limits determined by latent variable along with an adaptive parallel adaptive approach provided an accurate efficient characterisation of the conditional distribution. The algorithm in general performed well and addresses the known issues when discretising parameters, including algorithm performance in higher dimensional systems. The latent variable support for parameter μ adaptive weighting and framework was highlighted and was shown to be relatively easy to implement and code.

In summary, this article discusses techniques used in the self-exciting point process to investigate causal structure of injuries and the risk of profession AFL players. In particular, we examine low sample size context and illustrate a number of useful approaches to efficiently modelling both exogenous and endogenous clustering.

Acknowledgement

This research was supported by the ARC Centre of Excellence for Mathematical and Statistical Frontiers (ACEMS) and the Australian Research Council (ARC) Laureate Fellowship Program under the project 'Bayesian Learning for Decision Making in the Big Data Era' (ID: FL150100150).

References

Bacry, E., Mastromatteo, I., & Muzy, J.-F. (2015). Hawkes Processes in Finance. *Market Microstructure and Liquidity*, *01*(01), 1550005.

Bahr, R., & Holme, I. (2003). Risk factors for sports injuries—a methodological approach. *British journal of sports medicine*, *37*(5), 384–392.

Bittencourt, N. F., Meeuwisse, W., Mendonça, L., Nettel-Aguirre, A., Ocarino, J., & Fonseca, S. (2016). Complex systems approach for sports injuries: Moving from risk factor identification to injury pattern recognition—narrative review and new concept. *British journal of sports medicine*, *50*(21), 1309–1314.

Browning, R., Sulem, D., Mengersen, K., Rivoirard, V., & Rousseau, J. (2021). Simple discrete-time self-exciting models can describe complex dynamic processes: A case study of COVID-19. *PLoS ONE*, *16*(4 April), 1–28.

Casals, M., Langohr, K., Carrasco, J. L., & Rönnegård, L. (2015). Parameter estimation of poisson generalized linear mixed models based on three different statistical principles: A simulation study. *SORT: statistics and operations research transactions*, *39*(2), 0281–308.

Chen, J., Hawkes, A. G., & Scalas, E. (2021). A Fractional Hawkes Process. *SEMA SIMAI Springer Series*, *26*, 121–131.

Chiang, W. H., Liu, X., & Mohler, G. (2021). Hawkes process modeling of COVID-19 with mobility leading indicators and spatial covariates. *International Journal of Forecasting*, (40).

Chib, S. (2001). Markov chain monte carlo methods: Computation and inference. *Handbook of econometrics*, *5*, 3569–3649.

Daley, D. J., & Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes*. Springer-Verlag.

Donnet, S., Rivoirard, V., & Rousseau, J. (2020). Nonparametric bayesian estimation for multivariate hawkes processes. *The Annals of statistics*, 2698–2727.

Drew, M. K., Raysmith, B. P., & Charlton, P. C. (2017). Injuries impair the chance of successful performance by sportspeople: A systematic review. *British journal of sports medicine*, *51*(16), 1209–1214.

Hägglund, M., Waldén, M., & Ekstrand, J. (2005). Injury incidence and distribution in elite football—a prospective study of the danish and the swedish top divisions. *Scandinavian journal of medicine & science in sports*, *15*(1), 21–28.

Hall, E. C., & Willett, R. M. (2016). Tracking Dynamic Point Processes on Networks. *IEEE Transactions on Information Theory*, *62*(7), 4327–4346.

Hawkes, A. G. [Alan G.]. (1971). Point Spectra of Some Mutually Exciting Point Processes. *Journal of the Royal Statistical Society: Series B (Methodological)*, *33*(3), 438–443.

Hawkes, A. G. [Alan G], & Oakes, D. (1974). A cluster process representation of a self-exciting process. *Journal of applied probability*, *11*(3), 493–503.

Kalair, K., Connaughton, C., & Di Loro, P. A. (2020). A non-parametric Hawkes process model of primary and secondary accidents on a UK smart motorway.

Kelly, J. D., Park, J., Harrigan, R. J., Hoff, N. A., Lee, S. D., Wannier, R., Selo, B., Mossoko, M., Njoloko, B., Okitolonda-Wemakoy, E., Mbala-Kingebeni, P., Rutherford, G. W., Smith, T. B., Ahuka-Mundeke, S., Muyembe-Tamfum, J. J., Rimoin, A. W., & Schoenberg, F. P. (2019). Real-time predictions of the 2018–2019 Ebola virus disease outbreak in the Democratic Republic of the Congo using Hawkes point process models. *Epidemics*, *28*, 100354.

Lewis, E., & Mohler, G. (2011). A nonparametric em algorithm for multiscale hawkes processes. *Journal of Nonparametric Statistics*, *1*(1), 1–20.

Li, Z., Cui, L., & Chen, J. (2018). Traffic accident modelling via self-exciting point processes. *Reliability Engineering and System Safety*, *180*(July), 312–320.

Markwick, D. (2020). Bayesian Nonparametric Hawkes Processes with Applications.

Marsan, D., & Lengliné, O. (2008). Extending earthquakes' reach through cascading. *Science*, *319*(5866), 1076–1079.

Meeuwisse, W. H., Tyreman, H., Hagel, B., & Emery, C. (2007). A dynamic model of etiology in sport injury: The recursive nature of risk and causation. *Clinical journal of sport medicine*, *17*(3), 215–219.

Mohler, G. O., Short, M. B., Brantingham, P. J., Schoenberg, F. P., & Tita, G. E. (2011). Self-Exciting Point Process Modeling of Crime. *Journal of the American Statistical Association*, *106*, 100–108.

Ogata. (1981). On Lewis' Simulation Method for Point Processes. *IEEE Transactions on Information Theory*, *27*(1), 23–31.

Porter, M. D., & White, G. (2010). Self-exciting hurdle models for terrorist activity. *Annals of Applied Statistics*, *4*(1), 106–124.

Rasmussen. (2013). Bayesian Inference for Hawkes Processes. *Methodology and Computing in Applied Probability*, *15*(3), 623–642.

Ritter, C., & Tanner, M. A. (1992). Facilitating the gibbs sampler: The gibbs stopper and the griddy-gibbs sampler. *Journal of the American Statistical Association*, *87*(419), 861–868.

Ross, G. J. (2021). Bayesian estimation of the etas model for earthquake occurrences. *Bulletin of the Seismological Society of America*, *111*(3), 1473–1480.

Stares, J., Dawson, B., Peeling, P., Heasman, J., Rogalski, B., Drew, M., Colby, M., Dupont, G., & Lester, L. (2018). Identifying high risk loading conditions for in-season injury in elite australian football players. *Journal of science and medicine in sport*, *21*(1), 46–51.

Stares, J. J., Dawson, B., Peeling, P., Heasman, J., Rogalski, B., Fahey-Gilmour, J., Dupont, G., Drew, M. K., Welvaert, M., & Toohey, L. (2019). Subsequent injury risk is elevated above baseline after return to play: A 5-year prospective study in elite australian football. *The American journal of sports medicine*, *47*(9), 2225–2231.

White, G., Porter, M. D., & Mazerolle, L. (2012). Terrorism Risk, Resilience and Volatility: A Comparison of Terrorism Patterns in Three Southeast Asian Countries. *Journal of Quantitative Criminology*, *29*(2), 295–320.

Yang, Y., Etesami, J., He, N., & Kiyavash, N. (2018). Nonparametric Hawkes Processes: Online Estimation and Generalization Bounds. (Nips 2017), 1–39.

Zhang, R., Walder, C., & Rizoiu, M.-A. (2020). Variational inference for sparse gaussian process modulated hawkes process. *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*(04), 6803–6810.

Zhou, F., Li, Z., Fan, X., Wang, Y., Sowmya, A., & Chen, F. (2020). Fast multi-resolution segmentation for nonstationary hawkes process using cumulants. *International Journal of Data Science and Analytics*, *10*(4), 321–330.

Zhuang, J., & Mateu, J. (2019). A semiparametric spatiotemporal Hawkes-type point process model with periodic background for crime data. *Journal of the Royal Statistical Society. Series A: Statistics in Society*, *182*(3), 919–942.

Zou, G. (2004). A modified poisson regression approach to prospective studies with binary data. *American journal of epidemiology*, *159*(7), 702–706.

A MISD

Player	no-injury	mu -mle	alpha -mle	beta -mle	mu - em	alpha -em	beta -em
Player 2	33	0.493	0.129	0.444	0.489	0.128	0.560
Player 3	15	0.474	0.270	0.843	0.478	0.249	0.745
Player 6	16	0.375	0.394	0.609	0.402	0.331	0.606
Player 12	24	0.648	0.072	0.898	0.642	0.079	0.943
Player 15	35	0.531	0.214	0.782	0.525	0.212	0.724
Player 16	10	0.194	0.538	1.213	0.202	0.483	1.434
Player 17	29	0.714	0.161	1.526	0.701	0.174	1.506
Player 20	10	0.438	0.194	1.454	0.469	0.114	1.223
Player 24	33	0.443	0.226	0.493	0.431	0.232	0.513
Player 29	19	0.773	0.195	2.064	0.792	0.164	1.696
Player 30	60	0.809	0.284	0.820	0.764	0.313	0.836
Player 32	6	0.260	0.282	2.768	0.283	0.152	3.393
Player 34	38	0.551	0.161	0.697	0.534	0.177	0.730
Player 35	31	0.579	0.116	1.523	0.569	0.127	1.445
Player 40	22	0.298	0.409	1.237	0.307	0.373	1.224
Player 41	31	0.449	0.157	0.674	0.446	0.153	0.629
Player 42	40	0.977	0.185	0.871	0.947	0.199	1.084
Player 43	10	0.484	0.039	1.035	0.480	0.040	0.960
Player 53	16	0.251	0.120	0.611	0.247	0.127	0.610
Player 54	31	0.956	0.049	0.826	0.967	0.033	1.170
Player 64	7	0.257	0.156	1.598	0.267	0.111	2.054

LL - mle	LL - em	AIC - mle	AIC - em	BIC - mle	BIC -em
-50.362	-50.375	-94.72	-94.75	-96.17	-96.19
-19.986	-19.990	-33.97	-33.98	-36.44	-36.45
-21.818	-21.838	-37.64	-37.68	-40.02	-40.06
-31.262	-31.263	-56.52	-56.53	-58.38	-58.39
-47.363	-47.366	-88.73	-88.73	-90.09	-90.10
-15.350	-15.388	-24.70	-24.78	-27.70	-27.78
-32.365	-32.367	-58.73	-58.73	-60.34	-60.35
-14.499	-14.533	-23.00	-23.07	-26.00	-26.07
-50.107	-50.114	-94.21	-94.23	-95.66	-95.67
-18.643	-18.653	-31.29	-31.31	-33.45	-33.47
-51.848	-51.872	-97.70	-97.74	-98.36	-98.41
-10.109	-10.178	-14.22	-14.36	-17.88	-18.02
-52.755	-52.767	-99.51	-99.53	-100.77	-100.79
-42.613	-42.615	-79.23	-79.23	-80.75	-80.76
-33.257	-33.273	-60.51	-60.55	-62.49	-62.52
-49.053	-49.055	-92.11	-92.11	-93.63	-93.64
-32.248	-32.286	-58.50	-58.57	-59.69	-59.77
-15.235	-15.235	-24.47	-24.47	-27.47	-27.47
-33.800	-33.801	-61.60	-61.60	-63.99	-63.99
-29.989	-29.999	-53.98	-54.00	-55.50	-55.52
-13.091	-13.106	-20.18	-20.21	-23.65	-23.68

Table 3 MLE vs MISD parameters (left) and fit (right) on 100 estimates.

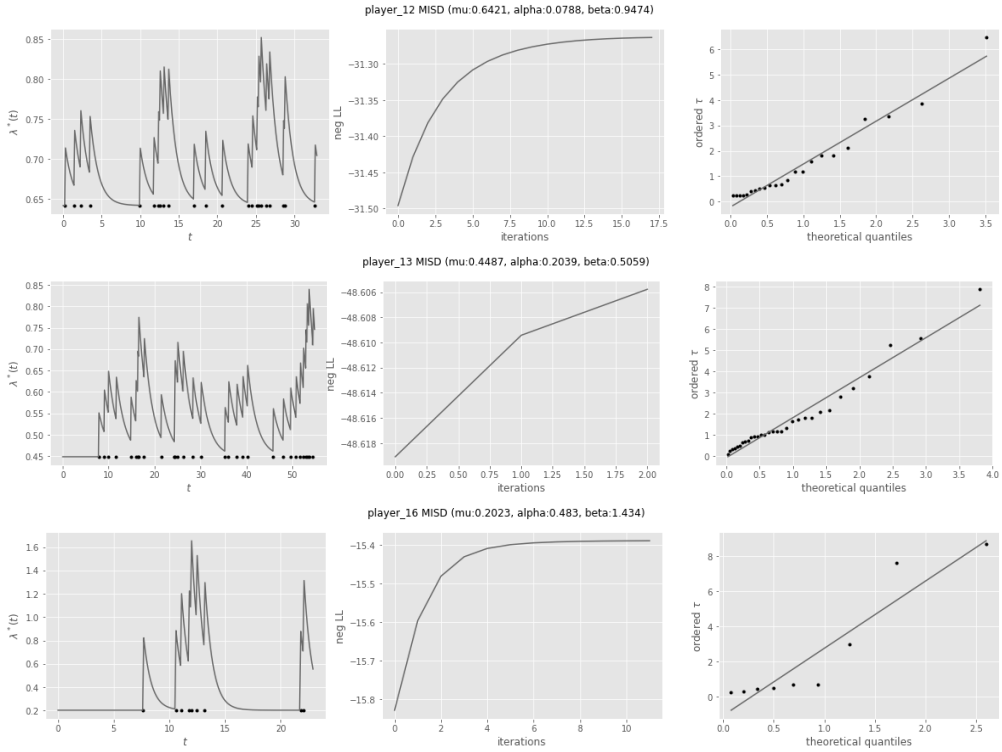

Figure 11 Player 12(above), 13(middle) and 16(bottom) rows, corresponding estimated conditional intensity(left), iteration(middle) and QQplot(right) columns

B Gibbs

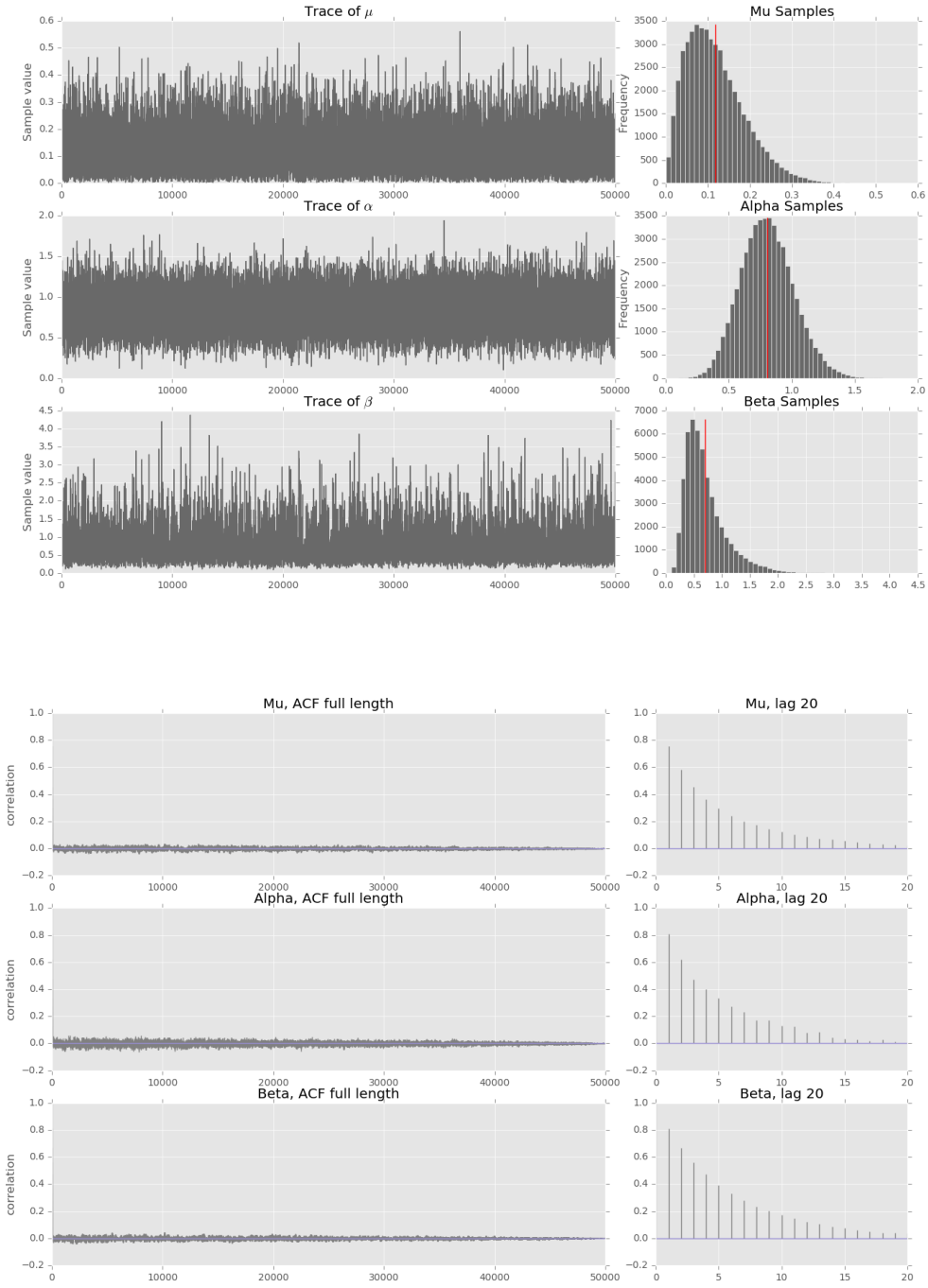

Figure 12 Players 17 parameters trace and auto correlation plot.

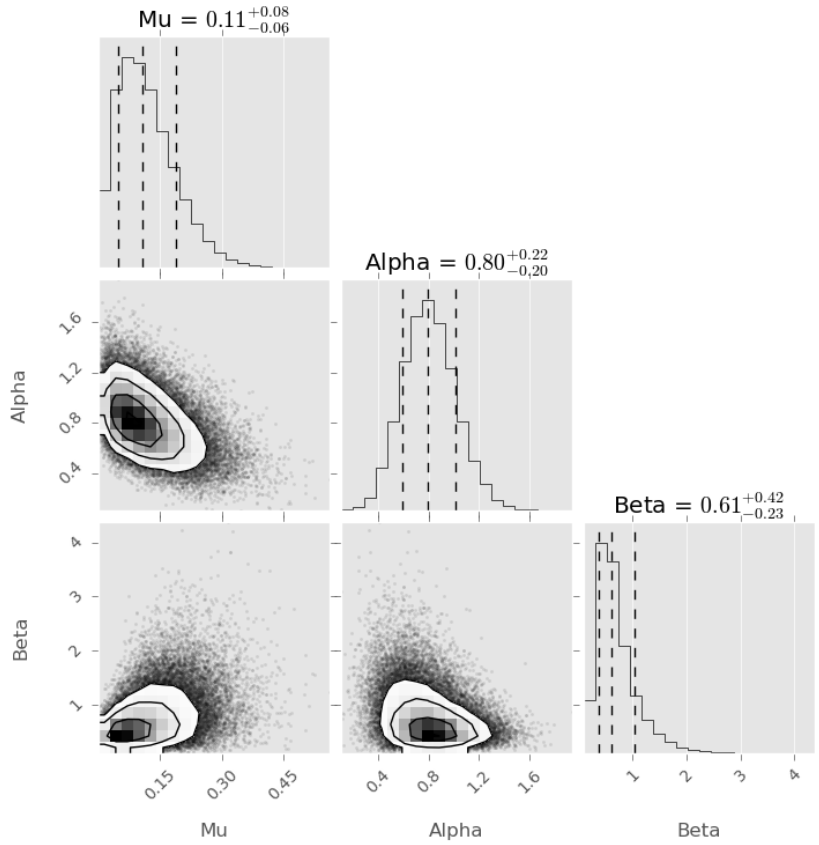

Figure 13 Players 17 chain sampling correlation.

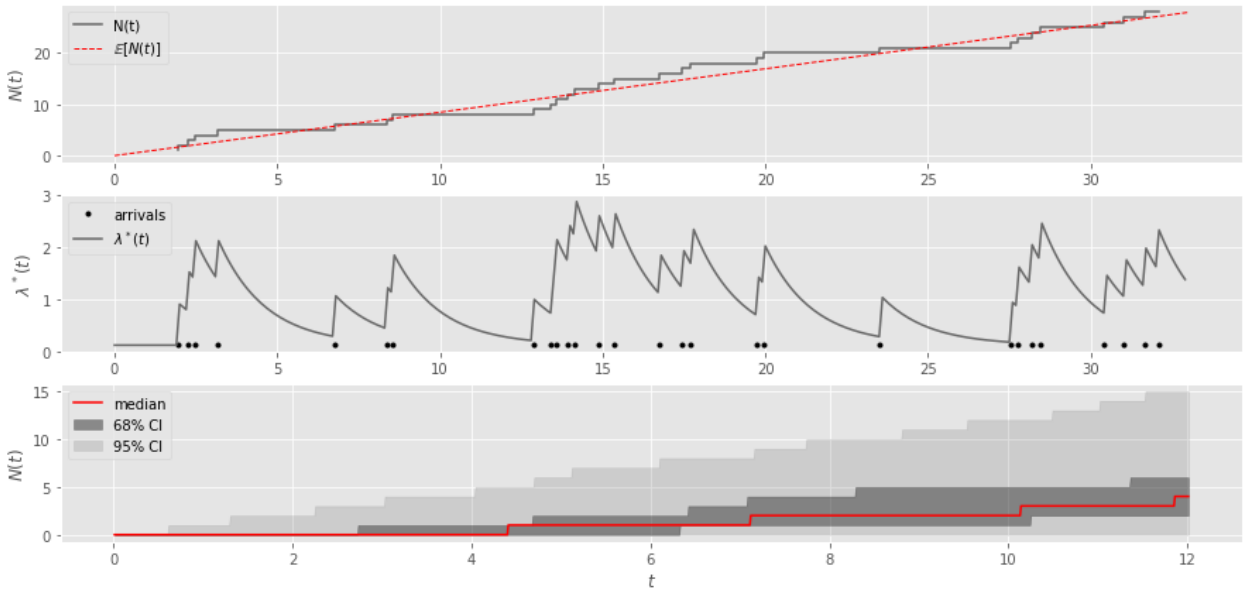

Figure 14 Players 17 chain forecast/HP model.

C APGG

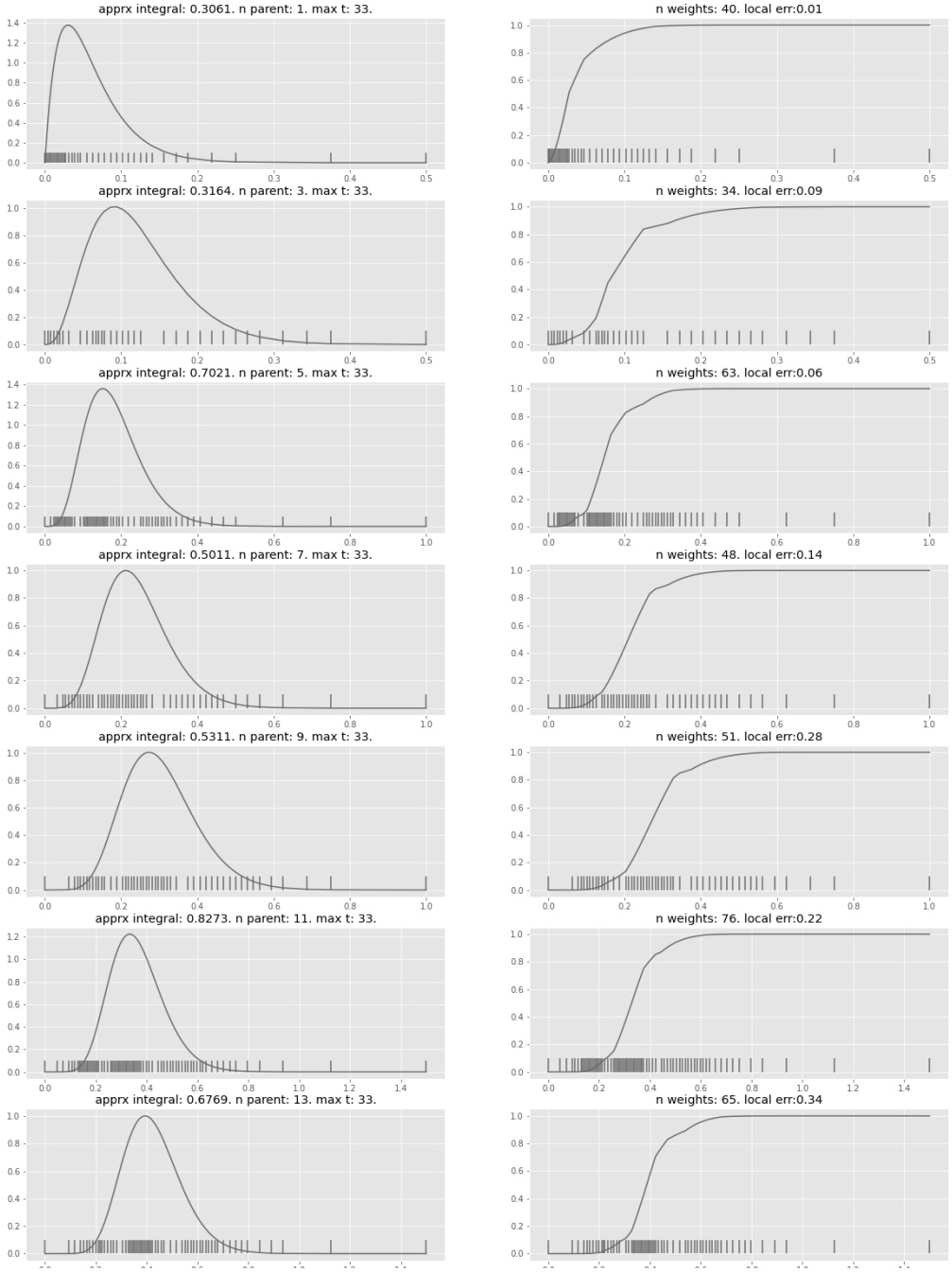

Figure 15 APGG μ, approximate integral(left), ECDF (right) and $B = 1, 3, 5, 7, 9, 11, 13$ (rows).

Players	mu	alpha	beta	Players	mu	alpha	beta
10	0.088 (0.054)	0.677 (0.292)	0.346 (0.293)	45	0.052 (0.035)	0.937 (0.17)	0.277 (0.096)
12	0.12 (0.07)	0.741 (0.236)	0.524 (0.277)	46	0.082 (0.055)	0.824 (0.237)	0.425 (0.193)
13	0.053 (0.036)	0.89 (0.196)	0.331 (0.13)	47	0.047 (0.032)	0.858 (0.218)	0.267 (0.111)
15	0.058 (0.04)	0.9 (0.188)	0.362 (0.165)	48	0.127 (0.08)	0.729 (0.273)	0.398 (0.242)
16	0.099 (0.054)	0.718 (0.311)	1.193 (0.654)	49	0.214 (0.114)	0.66 (0.408)	1.025 (0.724)
17	0.117 (0.071)	0.804 (0.212)	0.7 (0.388)	50	0.074 (0.047)	0.8 (0.258)	0.284 (0.142)
18	0.101 (0.063)	0.845 (0.23)	0.429 (0.212)	52	0.086 (0.051)	0.766 (0.227)	0.42 (0.223)
19	0.159 (0.099)	0.798 (0.28)	0.595 (0.319)	53	0.083 (0.039)	0.543 (0.261)	0.514 (0.408)
2	0.064 (0.041)	0.839 (0.199)	0.289 (0.144)	54	0.113 (0.073)	0.888 (0.211)	0.62 (0.268)
20	0.17 (0.088)	0.642 (0.34)	1.032 (0.725)	56	0.279 (0.148)	0.602 (0.454)	1.297 (1.093)
21	0.079 (0.046)	0.652 (0.28)	0.237 (0.217)	57	0.407 (0.219)	0.651 (0.465)	0.997 (0.763)
22	0.118 (0.056)	0.459 (0.325)	0.584 (0.603)	58	0.332 (0.168)	0.57 (0.441)	0.876 (0.767)
23	0.089 (0.049)	0.639 (0.282)	0.399 (0.376)	59	0.187 (0.094)	0.53 (0.36)	0.922 (0.798)
24	0.059 (0.038)	0.861 (0.196)	0.323 (0.144)	6	0.122 (0.069)	0.714 (0.264)	0.656 (0.421)
25	0.064 (0.043)	0.916 (0.164)	0.338 (0.117)	60	0.26 (0.138)	0.588 (0.482)	0.754 (0.714)
27	0.062 (0.04)	0.728 (0.286)	0.264 (0.208)	61	0.109 (0.059)	0.731 (0.306)	0.587 (0.392)
28	0.099 (0.066)	0.86 (0.202)	0.404 (0.151)	62	0.041 (0.029)	0.888 (0.22)	0.269 (0.113)
29	0.196 (0.108)	0.761 (0.273)	1.065 (0.676)	63	0.219 (0.117)	0.655 (0.342)	0.943 (0.611)
3	0.144 (0.079)	0.688 (0.275)	0.762 (0.467)	64	0.123 (0.063)	0.55 (0.37)	1.128 (0.889)
30	0.07 (0.044)	0.921 (0.14)	0.573 (0.193)	65	0.177 (0.096)	0.678 (0.307)	0.72 (0.456)
31	0.178 (0.097)	0.747 (0.374)	0.916 (0.577)	66	0.056 (0.038)	0.905 (0.163)	0.277 (0.093)
32	0.138 (0.075)	0.735 (0.499)	1.194 (0.976)	67	0.056 (0.037)	0.885 (0.18)	0.299 (0.113)
34	0.058 (0.039)	0.899 (0.186)	0.366 (0.155)	68	0.466 (0.253)	0.689 (0.422)	1.025 (0.731)
35	0.089 (0.056)	0.818 (0.216)	0.514 (0.305)	69	0.256 (0.144)	0.645 (0.44)	0.834 (0.643)
38	0.266 (0.142)	0.606 (0.406)	0.797 (0.645)	71	0.273 (0.178)	0.856 (0.447)	1.045 (0.585)
39	0.075 (0.048)	0.83 (0.2)	0.336 (0.145)	72	0.091 (0.059)	0.851 (0.209)	0.397 (0.181)
4	0.058 (0.037)	0.871 (0.223)	0.324 (0.179)	7	0.099 (0.061)	0.765 (0.251)	0.4 (0.241)
40	0.093 (0.048)	0.717 (0.231)	0.701 (0.456)	70	0.178 (0.101)	0.651 (0.354)	0.648 (0.51)
41	0.056 (0.038)	0.87 (0.207)	0.288 (0.151)	73	0.105 (0.065)	0.824 (0.213)	0.408 (0.178)
42	0.109 (0.068)	0.898 (0.178)	0.683 (0.254)	75	0.148 (0.082)	0.603 (0.36)	0.603 (0.547)
43	0.175 (0.086)	0.589 (0.347)	0.908 (0.712)	8	0.399 (0.236)	0.758 (0.531)	1.101 (0.792)
44	0.135 (0.062)	0.453 (0.321)	0.609 (0.628)				

Table 4 Gibbs, Chain = 50,000 with prior Gamma(1,1). Parameters mean and (standard deviation).

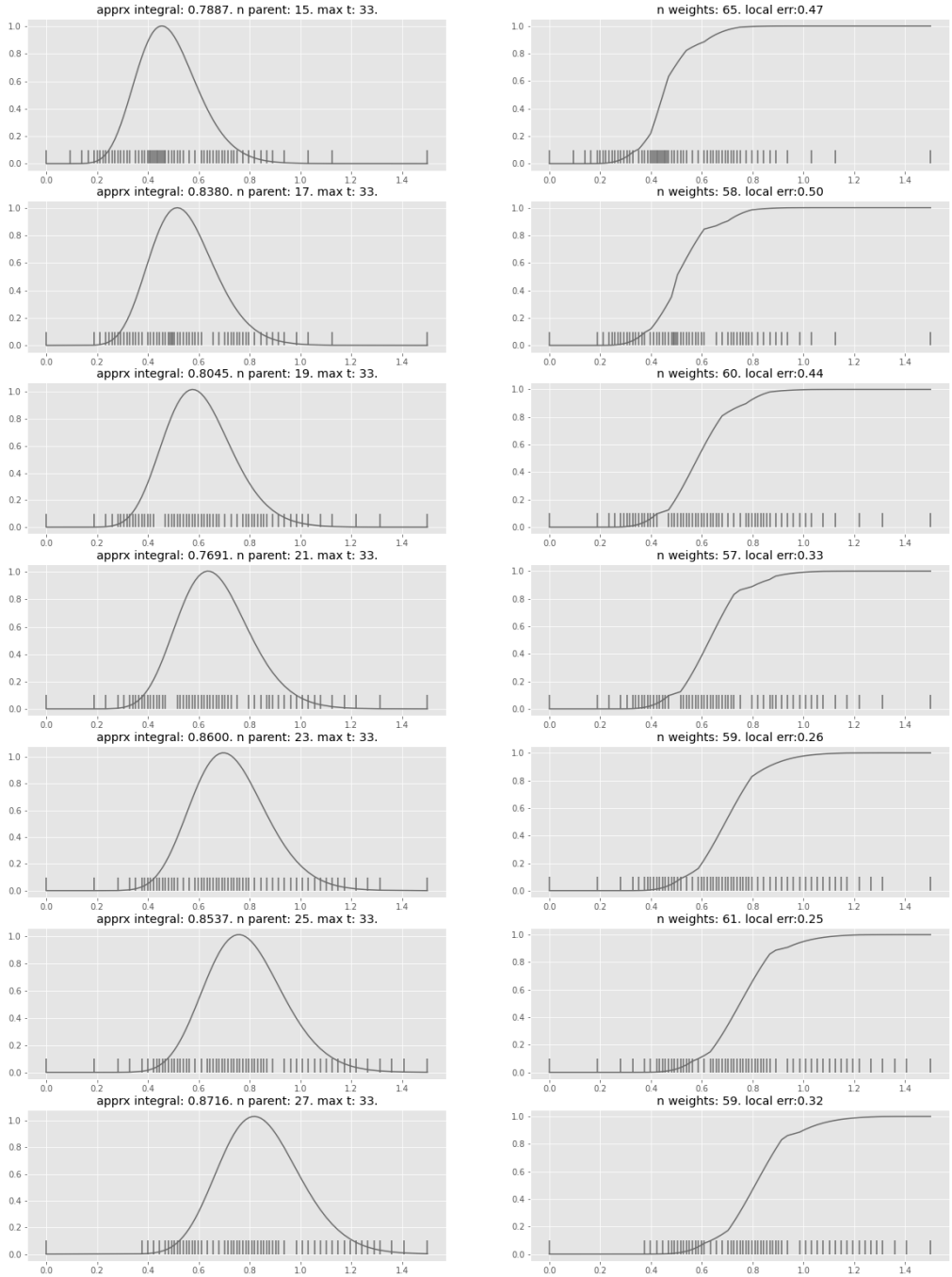

Figure 16 APGG μ, approximate integral(left), ECDF (right) and $B = 15, 17, 19, 21, 23, 25, 27$ (rows).

Players	mu	alpha	beta	Players	mu	alpha	beta
10	0.045 (0.078)	0.944 (0.257)	0.247 (0.119)	45	0.226 (0.456)	0.942 (0.206)	0.279 (0.104)
12	0.098 (0.193)	0.904 (0.247)	0.423 (0.222)	46	0.075 (0.144)	0.924 (0.24)	0.398 (0.173)
13	0.226 (0.467)	0.894 (0.251)	0.336 (0.177)	47	0.225 (0.47)	0.874 (0.256)	0.268 (0.149)
15	0.056 (0.181)	0.977 (0.196)	0.33 (0.143)	48	0.088 (0.171)	0.919 (0.261)	0.35 (0.158)
16	0.159 (0.223)	0.724 (0.356)	1.033 (0.683)	49	0.285 (0.398)	0.771 (0.474)	0.887 (0.668)
17	0.142 (0.285)	0.879 (0.261)	0.577 (0.357)	50	0.039 (0.079)	0.985 (0.237)	0.251 (0.098)
18	0.074 (0.149)	0.987 (0.227)	0.361 (0.16)	52	0.052 (0.108)	0.939 (0.222)	0.31 (0.151)
19	0.173 (0.32)	0.923 (0.305)	0.539 (0.288)	53	0.052 (0.178)	0.905 (0.279)	0.214 (0.176)
2	0.247 (0.523)	0.88 (0.244)	0.272 (0.166)	54	0.141 (0.291)	0.946 (0.249)	0.567 (0.266)
20	0.226 (0.337)	0.762 (0.411)	0.808 (0.655)	56	0.306 (0.398)	0.729 (0.514)	1.057 (0.958)
21	0.036 (0.112)	0.951 (0.239)	0.177 (0.079)	57	0.353 (0.492)	0.819 (0.528)	0.893 (0.621)
22	0.077 (0.128)	0.878 (0.368)	0.28 (0.264)	58	0.312 (0.431)	0.794 (0.502)	0.696 (0.566)
23	0.044 (0.159)	0.969 (0.246)	0.205 (0.114)	59	0.21 (0.305)	0.713 (0.425)	0.688 (0.634)
24	0.254 (0.529)	0.873 (0.253)	0.322 (0.207)	6	0.129 (0.229)	0.839 (0.295)	0.508 (0.36)
25	0.048 (0.088)	0.984 (0.161)	0.323 (0.106)	60	0.245 (0.343)	0.818 (0.535)	0.601 (0.492)
27	0.04 (0.132)	0.949 (0.255)	0.205 (0.093)	61	0.099 (0.18)	0.937 (0.337)	0.423 (0.303)
28	0.076 (0.15)	0.96 (0.2)	0.389 (0.138)	62	0.224 (0.488)	0.886 (0.265)	0.278 (0.159)
29	0.297 (0.497)	0.817 (0.363)	0.882 (0.657)	63	0.288 (0.421)	0.772 (0.407)	0.802 (0.57)
3	0.162 (0.282)	0.816 (0.32)	0.593 (0.402)	64	0.156 (0.225)	0.694 (0.441)	0.795 (0.772)
30	0.131 (0.382)	0.937 (0.192)	0.531 (0.213)	65	0.206 (0.328)	0.812 (0.354)	0.618 (0.408)
31	0.243 (0.357)	0.841 (0.442)	0.804 (0.566)	66	0.046 (0.151)	0.977 (0.161)	0.27 (0.086)
32	0.182 (0.246)	0.837 (0.553)	0.955 (0.875)	67	0.268 (0.578)	0.887 (0.239)	0.307 (0.158)
34	0.26 (0.514)	0.885 (0.252)	0.376 (0.213)	68	0.391 (0.535)	0.865 (0.483)	0.939 (0.61)
35	0.07 (0.148)	0.953 (0.225)	0.386 (0.234)	69	0.262 (0.378)	0.8 (0.477)	0.743 (0.527)
38	0.273 (0.409)	0.796 (0.449)	0.681 (0.502)	71	0.291 (0.436)	0.915 (0.491)	1.033 (0.571)
39	0.053 (0.105)	0.951 (0.194)	0.303 (0.119)	72	0.063 (0.134)	0.971 (0.204)	0.353 (0.14)
4	0.248 (0.504)	0.898 (0.283)	0.311 (0.237)	7	0.069 (0.127)	0.936 (0.243)	0.326 (0.175)
40	0.089 (0.171)	0.881 (0.266)	0.443 (0.342)	70	0.156 (0.263)	0.888 (0.377)	0.511 (0.365)
41	0.25 (0.514)	0.884 (0.261)	0.302 (0.21)	73	0.082 (0.158)	0.945 (0.216)	0.376 (0.167)
42	0.178 (0.359)	0.91 (0.24)	0.651 (0.313)	75	0.127 (0.211)	0.866 (0.382)	0.432 (0.331)
43	0.194 (0.299)	0.812 (0.429)	0.631 (0.588)	8	0.304 (0.41)	0.902 (0.591)	1.056 (0.693)
44	0.096 (0.165)	0.836 (0.371)	0.31 (0.287)				

Table 5 APGG Chain = 50,000 with prior Gamma(1,1). Parameters mean and (standard deviation).

D Forecast

Figure 17 Players number of injury forecast (12 months). The shaded areas corresponds to 68 and 95 percent credible regions and solid line represents medians of predictive density.