ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS

Journal of the Royal Statistical Society
Applied Statistics
Series C

Appl. Statist. (2020)
69, *Part* 5, *pp.* 1307–1336

A Bayesian quest for finding a unified model for predicting volleyball games

Leonardo Egidi

Università degli Studi di Trieste, Italy

and Ioannis Ntzoufras

Athens University of Economics and Business, Greece

[Received September 2019. Final revision July 2020]

Summary. Volleyball is a team sport with unique and specific characteristics. We introduce a new two-level hierarchical Bayesian model which accounts for these volleyball-specific characteristics. In the first level, we model the set outcome with a simple logistic regression model. Conditionally on the winner of the set, in the second level, we use a truncated negative binomial distribution for the points earned by the losing team. An additional Poisson-distributed inflation component is introduced to model the extra points played in the case that the two teams have a point difference less than two points. The number of points of the winner within each set is deterministically specified by the winner of the set and the points of the inflation component. The team-specific abilities and the home effect are used as covariates on all layers of the model (set, point and extra inflated points). The implementation of the proposed model on the Italian SuperLega 2017–2018 data shows exceptional reproducibility of the final league table and satisfactory predictive ability.

Keywords: Bayesian regression model; Losing team point; Predictive ability; Team-specific ability; Truncated negative binomial model; Volleyball results

1. Introduction

Sports analytics and modelling have a long tradition among the statistical community with initial work published back to the 1950s and 1960s. For example, seminal work has been initiated in the literature in the most popular sports like baseball (Mosteller, 1952; Albert, 1992), association football–soccer (Reep and Benjamin, 1968), American football (Mosteller, 1970; Harville, 1977) and basketball (Stefani, 1980; Schwertman *et al.*, 1991). The World Wide Web and recent technologies have given many scientists access to interesting sport-related data which are now widely and freely available (see for example `www.football-data.co.uk` for association football and `http://www.tennis-data.co.uk/` for tennis). Moreover, new interesting problems have been raised due to the big data sets that can be derived by in-play sensor and camera-driven technologies; see for example Metulini *et al.* (2017) and Facchinetti *et al.* (2019) for applications in basketball.

Sports analytics are currently a fashionable and attractive topic of research with a growing community of both academics and professionals. Regarding team sports prediction, there

Address for correspondence: Ioannis Ntzoufras, Sports Analytics Group, Computational and Bayesian Statistics Laboratory, Department of Statistics, Athens University of Economics and Business, 76 Patission Street, Athens, Attica 10434, Greece.
E-mail: ntzoufras@aueb.gr

 0035–9254/20/691307

are two main alternative outcomes of interest, and thus two modelling approaches (using two response variables) may be followed:

(a) the win–draw–loss outcome modelled by logistic or multinomial regression models and other similar approaches (Carpita *et al.*, 2019), and
(b) the goals or points scored by each team, which are modelled by sport-specific models depending on the nature of the game.

In this work, we focus on the second approach, where the response is richer in terms of information that is used for the estimation of team abilities and enables a model with better prediction accuracy to be obtained. The biggest group of team sports is the group where the score is measured with a discrete number of goals (of equal value) such as association football, water polo, handball and hockey (among others). In such cases, Poisson log-linear models and their extensions are the most popular choices of models. The historic time line of models developed for such team sports includes the use of simple double-Poisson models in the work of Maher (1982) and Lee (1997), the extension of Dixon and Coles (1997) with an adjustment for 0–0 and 1–1 draws, the diagonal bivariate Poisson model of Karlis and Ntzoufras (2003), the Poisson difference model (Karlis and Ntzoufras, 2009) and the dynamic models of Rue and Salvesen (2000), Owen (2011) and Koopman and Lit (2015); see Tsokos *et al.* (2019) and references therein for further details and an up-to-date review.

Unlike what happens for other major team sports, modelling volleyball match outcomes has not been thoroughly addressed by statisticians and data scientists. Early attempts to model volleyball data date back to Lee and Chin (2004), who analysed the effect of a team deciding to serve or to receive the service in the fifth set. Concerning prediction models in volleyball, several researchers have considered implementing Markov chain models to estimate the winning probabilities of a set and a game; see Barnett *et al.* (2008) for a first attempt and Ferrante and Fonseca (2014) for a more recent and complete treatment of the problem. Miskin *et al.* (2010) used a Bayesian multinomial formulation based on a Markovian transition matrix to model each point and the effect of specific volleyball skills. A similar approach was used by Drikos *et al.* (2019) to analyse top level international teams in several age categories. Sepulveda *et al.* (2017) used a Markov chain model as a useful tool to analyse players' probability of attack in terms of team rotation. A simpler alternative was based on logistic regression models for the probability of winning a set (Marcelino *et al.*, 2009; Fellingham *et al.*, 2013) or a point (Miskin *et al.*, 2010). Recently, Gabrio (2020) proposed a Bayesian hierarchical model for the prediction of the rankings of volleyball national teams, which also enabled the estimation of the results of each match in the league. At the points level, Gabrio (2020) used a double-Poisson model component. Sonnabend (2020) published an empirical study on the characteristics of beach volleyball, including details on the distribution of points. He used a normal regression model to study the effect of several game characteristics (game heterogeneity, referees, home effect, tournament phase, the fact of winning or losing the previous sets, gender and age) on the difference in points in a set. Finally, concerning research on player performance evaluation, Mendes *et al.* (2018) used Bayesian multilevel models to analyse sports activities of *élite* adult Brazilian players whereas Hass and Craig (2018) implemented a plus–minus approach to obtain volleyball players' evaluation metrics.

Unlike volleyball, in most sports (like in basketball and football, for example) there is a single performance outcome, namely the number of points or goals scored, which is measured cumulatively from the beginning to the end of the game. In these situations, a model with the total goals or points as a response is required. In contrast, in volleyball, the winner is announced in two stages or levels of outcomes: sets and points within each set. Hence, the winner is the team that reaches three sets first. For this reason, the second-level outcome, i.e. the total number of sets, is a random variable which ranges from a minimum of 3 to a maximum of 5. Each set is won by the team that first reaches a prespecified number of points which is 25 for the first four sets and 15 for the final tie-break set. (This points system was adopted in 1998, during the Men's and Women's World Championships held in Japan (source: `https://ncva.com/info/general-info/history-of-volleyball/`).) Nevertheless, the number of points that are required by the team winning the set varies further depending on whether there is a margin of two points. Hence, volleyball outcomes consist of a natural hierarchy of sets and points within sets, with both measurements being random variables.

In this work, we follow the approach of modelling both outcomes of volleyball: sets and points. In this way, the response data are richer in terms of information which enables us to estimate team abilities more accurately and to increase the prediction accuracy of our model.

In our perspective, the task of modelling volleyball match results should follow a top-down strategy, from sets to single points. Thus, defining the probability of winning a set is the first step; building up a generative discrete model for the points realized in each set is the second step. Although following this kind of hierarchy is not mandatory, we maintain it in all our fitted models. Hence, we propose a set-by-set statistical model for the points of the losing team, conditionally on the set result. Another aspect to consider is the difference in strength between the teams: weaker teams are of course not favoured when competing against stronger teams, and a parametric assumption about teams' skills is needed. In the Bayesian approach, teams' abilities are easily incorporated in the model by the use of weakly informative prior distributions (Gelman *et al.*, 2008): similarly to what happens for football models (Karlis and Ntzoufras, 2003), the abilities may be regarding both attack and defence skills, and, moreover, be considered as dynamic over the season (Owen, 2011).

The rest of the paper is organized as follows. The main features of the game are presented in Section 2. In Section 3 we introduce the basic negative binomial model for volleyball outcomes. Model extensions are thoroughly presented in Section 4, whereas model estimation, goodness-of-fit diagnostics and out-of-sample prediction measures are detailed in Section 5. Markov chain Monte Carlo (MCMC) replications for the selected negative binomial model are used in Section 5.3 to assess its plausibility and to reconstruct the final rank of the league. The paper concludes after a detailed discussion.

The programs that were used to analyse the data can be obtained from

`https://rss.onlinelibrary.wiley.com/hub/journal/14679876/series-c-datasets`.

2. The features of the game

Volleyball is different from other team sports of invasion (like football or basketball in which the purpose is to invade the opponent's territory and to score a goal or a point) since the two teams are separated and there is no contact between the players of the two competing teams. It belongs to a category of net and ball sports (footvolley, headis or sepak takraw, tennis, badminton, pickleball and table tennis) and therefore it has some unique characteristics that cannot be modelled by using the approaches that are adopted in other sports such as the Poisson regression models that are commonly used in football.

Here we summarize these characteristics and we address theses issues one by one.

(a) The first and most important characteristic is that the main outcome of the game is split into two levels: the sets and the points inside each set. Roughly speaking, a set is played until one of the two teams first wins 25 points. This team is the winner of the set. The game is played until a team wins three sets. Hence we have two levels of outcome (sets and points) which are interconnected and should be modelled simultaneously.

(b) Moreover, the number of sets in a volleyball game ranges from 3 to 5 and, hence, it is reasonable to test for the assumption of repeated measures of the points which are correlated across different sets. The existence of repeated measurements of points needs to be addressed stochastically and tested within our modelling approach.

(c) The points of the winning team are (almost) fixed by the design and the rules of the game. So, given that we know who won the set, the only outcome variability is reflected by the points of the team that lost the particular set.

(d) An additional rule, that creates further complication, is that the winning team should have at least two points margin of difference to win a set. So conceptually, if two teams are close in terms of their abilities, they could play for infinite time and points until the required difference of two points is achieved.

(e) Finally, the fifth set of the game is terminated at 15 points (and not at 25 points) and it is called a tie-break. The two-points margin of difference is also required for the tie-break.

In this work, we deal with each of unique characteristics of volleyball by adding a corresponding component to the model formulation. The resulting model is a unified approach for the volleyball data and it is unique in the literature. To be more specific, we model the two response outcomes (sets and points) hierarchically, using a binomial model for each set and, conditionally on the winner of the set, we use a negative binomial distribution for the points of the losing team assuming $r=25$ or $r=15$ successes for normal sets and tie-breaks respectively (features (a), (c) and (e)). We further truncate this distribution to deal with the two-points margin of difference that is required in each set (feature (d)), and we model the excess of points due to ties (sets with less than two points difference) by using a zero-inflated Poisson (ZIP) distribution (feature (d)). Furthermore, we consider normal random effects to account for the correlation between sets of the same game (feature (b)). Finally, we take into consideration the connection between sets and points by considering general team abilities in contrast with point- or set-specific team abilities (feature (a)).

As the reader might initially think, our approach is counterintuitive and apparently in contrast with what a usual sports model may consider. But this counterintuitive logic is the main innovation of the model that we propose. By using this approach, our aim is to exploit the fact that we (almost) know the points of the winning team. So if we consider modelling the win or loss of each set in the first level of the model then, conditionally on the set winner, we can specify a sensible distribution for the points of the losing team (whereas the points of the winning team are specified deterministically). In contrast, considering the usual approach, i.e. modelling directly the number of points by using a bivariate distribution is more cumbersome and challenging because of the restrictions that are imposed by the game regulations.

In Section 3 which follows, we formalize the basic structure and assumptions of our proposed model whereas further considerations and extensions of the model are provided in Section 4.

3. The basic model for volleyball

3.1. Truncated negative binomial model

Let Y_s^{A} and Y_s^{B} be the random variables of the points in set $s=1,2,\ldots,S$ of two competing teams A and B playing at the home and away stadium respectively. Furthermore, W_s is a binary indicator denoting the win or loss of sets for the home team. To begin with, assume for the moment that each set finishes at a fixed number of points (25 or 15 depending on the type of set);

then the points of the winning team are fixed and not random. Hence interest lies in the random variable Y_s which denotes the number of points for the team losing the sth set. Concerning the observed realization of the points gained by the losing team, this will be obtained by

$$y_s = w_s y_s^{\mathrm{B}} + (1 - w_s) y_s^{\mathrm{A}}.$$

So, in our data set, we shall eventually model the data for two responses: the binary response W_s and the count variable Y_s. Our model is built hierarchically. For the outcome of each set, we use a simple logistic regression model given by

$$W_s \sim \text{Bernoulli}(\omega_s), \tag{1}$$

$$\text{logit}(\omega_s) = H^{\text{set}} + \alpha_{A(s)} - \alpha_{B(s)}, \tag{2}$$

where α_T is a parameter capturing the ability of team T to win a set (set abilities henceforth), and $A(s)$ and $B(s)$ are the home and away team indices respectively for the teams competing with each other in set s. Now, conditionally on the winner of the set, we then model the points of the losing team for each set by using a negative binomial model (ignoring for the moment that the game may continue if the margin of points' difference is less than two points). Hence, the model formulation will be now given by

$$Y_s | W_s \sim \text{NegBin}(r_s, p_s) \mathscr{I}(Y_s \leqslant r_s - 2), \tag{3}$$

which is the right-truncated negative binomial distribution with parameters r_s and p_s. $\mathscr{I}(A)$ denotes the event indicator, which is equal to 1 if the event A is true and 0 otherwise. The first parameter, r_s, is the number of successes (points here) required to finish the set and it is equal to 25 for sets 1–4, and equal to 15 for the last (fifth) set. Mathematically this can be written as

$$r_s = 25 - 10\, \mathscr{I}(R_s = 5), \tag{4}$$

where R_s is the sequential set number for the specific game $G(s)$. Parameter p_s is the probability of realizing a point for the team winning set s. Equivalently, $q_s = 1 - p_s$ denotes the probability of realizing a point for the team losing set s. The right truncation has been fixed at $r_s - 2$ (23 or 13) points since this is the highest number of points that can be achieved by the losing team (under the assumption of no ties). Moreover, the point success probability will be modelled as

$$\eta_s = \mu + (1 - W_s) H^{\text{points}} + (\beta_{A(s)} - \beta_{B(s)})(1 - 2W_s), \tag{5}$$

$$p_s = \frac{1}{1 + \exp(\eta_s)}, \tag{6}$$

where η_s is the (fixed effects) linear predictor for the points of the losing team given by equation (5). The constant μ is a common baseline parameter, H^{points} is the points home advantage for the host team, and $\beta_{A(s)}$ and $\beta_{B(s)}$ are the points abilities for teams $A(s)$ and $B(s)$ respectively. Consider the first equation: the larger is the difference between the abilities of team A and team B, $\xi_s = \beta_{A(s)} - \beta_{B(s)}$, the higher is the expected number of points that team A will win when losing a set. Equivalently, in this case, the lower will be the number of points of team B when losing a set. Hence the multiplier $1 - W_s$ in equation (5) controls the presence of the home effect, whereas the multiplier $1 - 2W_s$ controls the sign of the difference in the abilities of the two teams (depending on which team is playing at home).

Before we proceed, we focus for a moment on the untruncated negative binomial distribution, for which the average number of points for team A (evaluated if $W_s = 0$) and team B (evaluated if $W_s = 1$) in the sth set are respectively

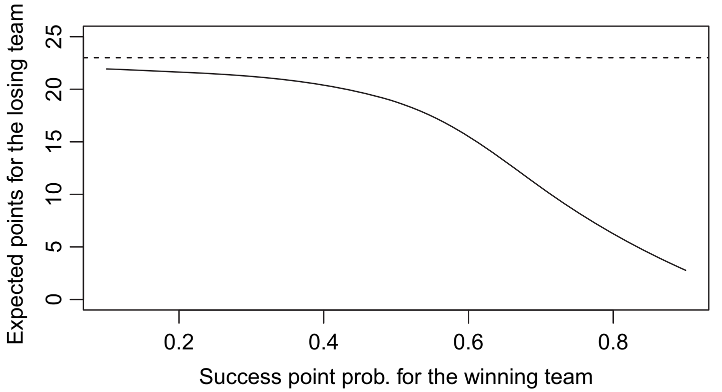


Fig. 1. Expected number of points collected by the team losing set s against the success point probability p_s for the winning team, truncated negative binomial with upper truncation at r_s-2: as the points probability for the team winning the set increases, the expected number of points for the team losing the set decreases

$$\begin{aligned} E[Y_s^{\mathrm{A}}|W_s=0]&=r_s\exp(\mu+H^{\mathrm{points}}+\beta_{A(s)}-\beta_{B(s)})=r_sM\xi_s\exp(H^{\mathrm{points}}),\\ E[Y_s^{\mathrm{B}}|W_s=1]&=r_s\exp(\mu-\beta_{A(s)}+\beta_{B(s)})=r_s\ M/\xi_s, \end{aligned} \tag{7}$$

where $M=\exp(\mu)$ and $\xi_s=\exp(\beta_{A(s)}-\beta_{B(s)})$.

However, in this initial model formulation the losing set team can reach at most r_s-2 points (in the case of no extra points); then we need to reconsider the expected number of points of the losing team (i.e. equation (7)) in the light of the upper truncation. Shonkwiler (2016) reported the mathematical expression for the truncated negative binomial distribution which in our case becomes

$$\left.\begin{aligned} E[Y_s^{\mathrm{A}}|Y_s^{\mathrm{A}}\leqslant r_s-2,W_s=0]E[Y_s^{\mathrm{A}}|W_s=0]-c_s^*/p_s,\\ E[Y_s^{\mathrm{B}}|Y_s^{\mathrm{B}}\leqslant r_s-2,W_s=1]=E[Y_s^{\mathrm{B}}|W_s=1]-c_s^*/p_s,\\ c_s^*=\frac{(r_s-1)f_{\mathrm{NB}}(r_s-1)}{F_{\mathrm{NB}}(r_s-2;r_s,p_s)},\qquad c_s^*>0, \end{aligned}\right\} \tag{8}$$

where f_{NB} and $F_{\mathrm{NB}}(x;r,p)$ are the probability mass function and the cumulative function respectively of the negative binomial distribution with parameters r and p. The interpretation is identical to the untruncated case: the higher is the point ability of a team, the higher will be the number of points when losing a set, given a fixed winning team ability. However, the untruncated mean is subtracted by the positive factor c_s^*/p_s, which forces the mean of the points of the losing team to be lower than or equal to r_s-2. For illustration Fig. 1 displays the expected number of points collected by the team losing set s against the success point probability p_s: as the point probability for the team winning the set increases, the expected number of points for the team losing the set decreases. In Section 3.3 we shall extend the model to allow for extra points after r_s due to the required margin of two points difference.

The random variables of the points of each team, under the model assumed, can be now written as

$$\begin{aligned} Y_s^{\mathrm{A}}&=W_sr_s+(1-W_s)Y_s,\\ Y_s^{\mathrm{B}}&=W_sY_s+(1-W_s)r_s, \end{aligned}$$

whereas the expected number of points of each set are given by

$$E[Y_s^{\mathrm{A}}] = \omega_s r_s + (1-\omega_s) r_s \xi_s M \exp(H^{\mathrm{points}}) - (1-\omega_s)\{1+\xi_s M \exp(H^{\mathrm{points}})\} c_s^*,$$
$$E[Y_s^{\mathrm{B}}] = \omega_s r_s \frac{M}{\xi_s} + (1-\omega_s) r_s - \omega_s \left(1+\frac{M}{\xi_s}\right) c_s^*, \tag{9}$$

where c_s^* is given in expression (8) whereas ξ_s and M are defined in expression (7).

The Bayesian model is completed by assigning some weakly informative priors (Gelman *et al.*, 2008) to the set and points abilities, for each team $T = 1, \ldots, N_T$:

$$\alpha_T^*, \beta_T^* \sim \mathcal{N}(0, 2^2),$$
$$\mu, H^{\mathrm{points}}, H^{\mathrm{set}} \sim \mathcal{N}(0, 10^6), \tag{10}$$

where N_T is the total number of teams in the league. To achieve identifiability, set and points abilities need to be constrained; in such a framework we impose sum-to-zero (STZ) constraints for both α and β by centring the free parameters α_T^* and β_T^* using the equations

$$\alpha_T = \alpha_T^* - \bar{\alpha}^*,$$
$$\beta_T = \beta_T^* - \bar{\beta}^*,$$

for $T = 1, \ldots, N_T$, where $\bar{\alpha}^*$ and $\bar{\beta}^*$ are the means of the unconstrained abilities given by $\bar{\alpha}^* = (1/N_T)\Sigma_{T=1}^{N_T} \alpha_T^*$ and $\bar{\beta}^* = (1/N_T)\Sigma_{T=1}^{N_T} \beta_T^*$. Note that the constrained abilities α_T and β_T are finally used in the model which automatically satisfies the STZ constraint and this centring is applied in every iteration of the MCMC algorithm. In terms of interpretation, the STZ parameterization implies that an average team will have an ability parameter close to 0.

3.2. Using random effects to capture within-game correlation

We further introduce game additive random effects to capture the induced correlation between the set repetition and the fact that we have 3–5 measurements of the points of the losing team. Hence, the points probability in each set given by equation (6) is slightly changed to

$$p_s = \frac{1}{1+\exp(\eta_s + \varepsilon_{G(s)})},$$

where η_s is the (fixed effects) linear predictor as defined in equation (5) and $\varepsilon_{G(s)}$ are the game random effects which are used to capture any potential correlation across the measurements of the points within each game.

To complete the model formulation, we include a hierarchical step to assume exchangeability of the game random effects by

$$\varepsilon_G \sim \mathcal{N}(0, \sigma_\varepsilon^2),$$

for $G = 1, \ldots, N_G$ with N_G denoting the number of games in the data set. Moreover, for the variance of the random effects we consider the hyperprior

$$\sigma_\varepsilon^2 \sim \mathrm{InvGamma}(a_\varepsilon, b_\varepsilon),$$

with fixed hyperparameters a_ϵ and b_ϵ. Small posterior values of σ_ε indicate that there is no need for such game effects, whereas large values indicate the need for unconnected (fixed) game effects (and possibly bad fit of the model without any game effects). Fig. 2 displays the posterior marginal distribution for σ_ε for the Italian SuperLega 2017–2018 data: although a change of 0.1–0.15 in the random effects may have a non-negligible effect in the logit scale of the points probability p_s defined above, there is little evidence of any set effect in terms of the model's goodness of fit, as will be demonstrated in Section 3.4.

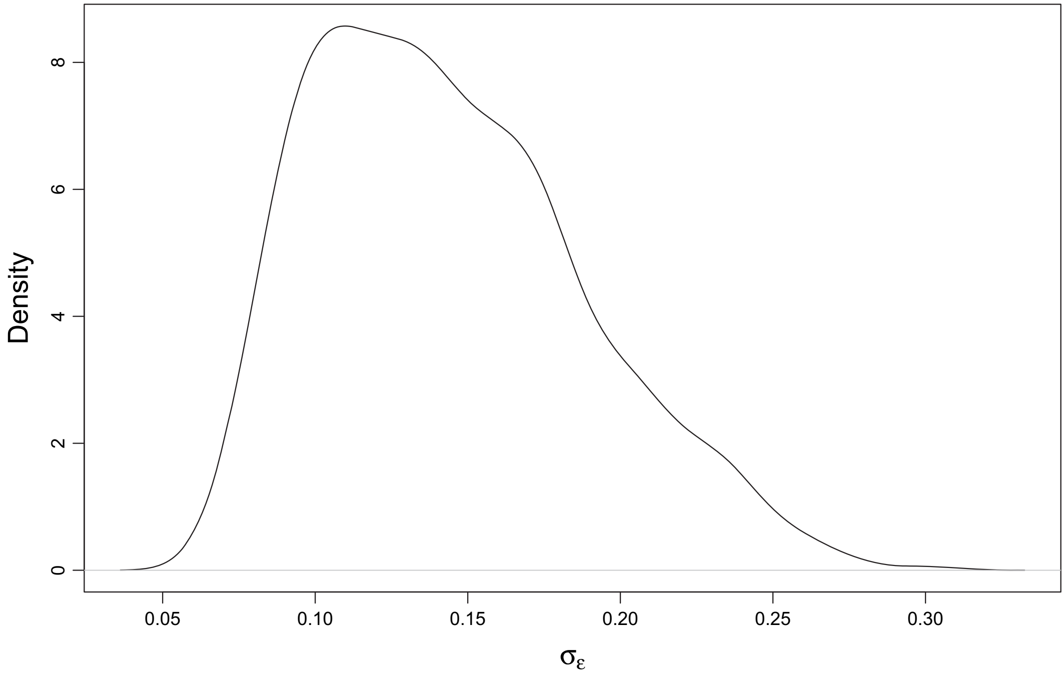


Fig. 2. Estimated posterior marginal distribution of the standard deviation σ_ε of the random effects $\varepsilon_{G(s)}$ for the Italian SuperLega 2017–2018 data

3.3. Zero-inflated Poisson distribution for the extra points

To allow for the extra points arising from the 24-deuce (or 14-deuce) score, the model that was proposed in Section 3.1 is extended by specifying a ZIP latent variable for the extra points that are collected by the losing set team. The number of extra points is 0 if the losing set team does not reach 24 points and greater than 0 otherwise. So the model for the random variable of the points collected by the losing team is now defined as

$$\begin{aligned} Z_s &= Y_s + O_s, \\ Y_s &\sim \text{NegBin}(r_s, p_s)\, \mathscr{I}(Y_s \leqslant r_s - 2), \\ O_s &\sim \text{ZIPoisson}(\pi_s, \lambda). \end{aligned} \tag{11}$$

The ZIP distribution for the number of extra points O_s that are collected by the team losing the sth set is then defined as

$$f_{\text{ZIP}}(o_s) = \pi_s \mathscr{I}(o_s = 0) + (1 - \pi_s) f_{\text{P}}(o_s; \lambda), \tag{12}$$

where π_s describes the proportion of 0s and $f_{\text{P}}(x; \lambda)$ is the probability mass function of a Poisson distribution with rate parameter λ evaluated at x. In this section, we assume a constant inflation probability for all games, but in Section 4.4 we explore the possibility of expressing π_s as a function of the team abilities.

Following expression (9), under this model the random variables of the points are now given by

$$Z_s^{\text{A}} = Y_s^{\text{A}} + O_s = W_s r_s + (1 - W_s) Y_s + O_s$$

and

$$Z_s^{\mathrm{B}} = Y_s^{\mathrm{B}} + O_s = W_s Y_s + (1 - W_s) r_s + O_s$$

for the home and the away team respectively. Now the expected points are adjusted for the extra points, and hence they are given by

$$E[Z_s^{\mathrm{A}}] = E[Y_s^{\mathrm{A}}] + (1 - \pi_s)\lambda$$

and

$$E[Z_s^{\mathrm{B}}] = E[Y_s^{\mathrm{B}}] + (1 - \pi_s)\lambda,$$

where $E[Y_s^{\mathrm{A}}]$ and $E[Y_s^{\mathrm{B}}]$ are given in expression (9) and represent the expected number of points under the truncated model of Section 3.1 which does not consider any extra points for each set.

3.4. Model comparisons for the basic model formulation by using the deviance information criterion

Table 1 reports the deviance information criterion DIC (Spiegelhalter *et al.*, 2002) values and the effective number of parameters on the Italian SuperLega 2017–2018 data for a simple Poisson model and the basic models that were presented in Sections 3.1–3.3, computed by running 3000 iterations obtained by three parallel chains of 1000 iterations of Gibbs sampling via the R package `rjags` (Plummer, 2018). In the Poisson model, the rates have a log-linear specification depending on the points abilities. Models 1 and 2 use unrestricted data (with no explicit modelling of the ties) and both report a higher DIC than does the truncated negative binomial model with extra points (model 3). As far as we can conclude from the DIC, using random effects to capture within-game correlation (model 4) improves the fit only slightly (DIC = 4537.2 *versus* DIC = 4537.7); see also the posterior marginal distribution of σ_ε^2 in Fig. 2 and the considerations in Section 3.2. So we recommend use of the truncated negative binomial model allowing for extra points (model 3 in Table 1; see Section 3.3 for details) since it has similar predictive accuracy (in terms of DIC) to that of the corresponding random-effects model (model 4 in Table 1), whereas the computational burden and its model complexity are considerably lower.

Table 1. Details of the fitted models with different distributional assumptions for the points of the losing team for the Italian SuperLega 2017–2018 season†

Points distribution of the losing team‡	*Equations*	*Additional model details at the points level*	*Number of effective parameters*	*DIC*
1, Poisson	Log-linear model§	No upper limit	29	4557.2
2, truncated negative binomial	(3)–(6)	Upper limit, no extra points	29	4674.2
3, ZIP truncated negative binomial	(4)–(6) and (10), (11)	Upper limit and extra points	133	4537.7
4, ZIP truncated negative binomial	(4), (5), (10), (11) and Section 3.2	Model 3 and game random effects	151	4537.2

†MCMC sampling; 3000 iterations; `rjags` package.
‡In all models we use a logistic regression model for the sets (equations (1) and (2)), and overall disconnected team abilities α_T and β_T for the set and the points level.
§$Y_s|W_s \sim \text{Poisson}(\lambda_s)$ with $\log(\lambda_s/r_s) = \eta_s$ where r_s and η_s are given by equations (4) and (6) respectively.

4. Model extensions concerning team abilities

4.1. Attacking and defensive abilities

A common practice in many team sports (such as football, basketball and hockey) is to model the attacking and the defensive team abilities separately. This is also relevant for coaches and sports scientists because modern sports are highly specialized: estimates of attack and defence abilities give an indication of athletes' or team performance. Following this practice also in our proposed model for volleyball, we can assume the following decomposition of the points abilities of team T, $T=1,\ldots,N_T$:

$$\beta_T = \beta_T^{\text{att}} + \beta_T^{\text{def}}, \tag{13}$$

where the global points abilities β are defined as the sum between the attack and the defence abilities at the points level for each team. It is worth noting that assuming different attacking and defensive abilities at the set level would make the logit model (2) not identifiable.

4.2. Connecting the abilities

In equations (2) and (5), set and points abilities influence separately the set and points probabilities respectively: conditionally on winning or losing a set, points abilities are then estimated from the probability of realizing a point. However, we could combine them by defining a global ability measure. Here we consider a model where the abilities of winning a point also influence the probability of winning a set by a different scaling factor (controlled by the parameter θ). Hence the probability of winning a set is now given by

$$\text{logit}(\omega_s) = H^{\text{set}} + v_1(\alpha_{A(s)} - \alpha_{B(s)}) + v_2\theta(\beta_{A(s)} - \beta_{B(s)}), \tag{14}$$

where v_1 and v_2 are indicator variables, and θ summarizes the effect of the points abilities on winning a set. If $v_1=1$ and $v_2=0$ we obtain the basic model of Section 3.1 with set probability as defined by equation (2); if $v_1=0$ and $v_2=1$ we assume connected points and set abilities where the set ability parameters are simply proportional to points abilities, whereas if $v_1=v_2=1$ we assume connected point and set abilities and extra set-specific abilities. For illustration only, just view everything from the perspective of team A. Now consider the model with connected abilities and extra set abilities ($v_1=v_2=1$). If two teams are almost equally strong in terms of points, then the points abilities difference $\beta_{A(s)} - \beta_{B(s)}$ will be very small, and the set probability will be solely driven by the extra set abilities. Conversely, when two teams are expected to be quite far apart in terms of point performance, then the set winning probability will be mainly affected by the points performance.

In this generalized version of the model (the case $v_1=v_2=1$), the set abilities will capture diversions of teams in the set efficiency in comparison with the points efficiency. For most of the teams, intuitively we do not expect an excess of set abilities and the probability of winning a set will be mainly driven by a unified (set and points) ability. But a limited number of teams is expected to be more or less efficient at the set level than at the points level. Therefore, we have used posterior intervals and DIC to identify which teams behave in a different way in terms of sets and therefore an extra parameter is needed to handle these differences.

In Table 2 the DIC-values and the effective number of parameters for each model are reported with respect to the Italian SuperLega 2017–2018 data. According to this analysis the ZIP truncated negative binomial model with connected abilities and extra set abilities for only Verona and Padova and constant zero-inflated probability is the best fitted model.

Table 2. Details of the fitted logistic–ZIP truncated negative binomial models for the Italian SuperLega 2017–2018 season†

Model‡	*Connected team abilities*	*Additional model features*	*Number of effective parameters*	*DIC*
3	No	—	133	4537.7
5	No	Separate attacking and defensive abilities at the points level	147	4541.1
6	Yes	Connected abilities only ($v_1=0; v_2=1$)	122	4524.3
7	Yes	Plus extra set abilities ($v_1=v_2=1$)	133	4536.3
8	Yes	Plus extra set abilities for Verona	123	4522.2
9	Yes	Plus extra set abilities for Verona and Padova	124	4521.1
10	No	β_t dynamic (points level)	174	4569.5
11	No	α_t dynamic (set level)	132	4530.1

†MCMC sampling; 3000 iterations; `rjags` package.
‡In all models we use a logistic regression model for the sets (equations (1) and (2)), the ZIP model for extra points and the truncated negative binomial model was used for the points, and a constant probability and Poisson rate for the ZIP component of the extra points is assumed.

4.3. Dynamic abilities

The performance of each team is likely to change within a season. Hence, temporal trends may be helpful for modelling the ability of each team within a season. A dynamic structural assumption for the ability parameters is a step forward. A natural choice is an auto-regressive model for the set and points abilities. For each team $T=1,\ldots,N_T$ and game $G=2,\ldots,N_G$ we specify

$$\begin{aligned}\alpha_{T,G} &\sim \mathcal{N}(\alpha_{T,G-1},\sigma_\alpha^2),\\ \beta_{T,G} &\sim \mathcal{N}(\beta_{T,G-1},\sigma_\beta^2),\end{aligned} \tag{15}$$

whereas for the first match we assume that

$$\begin{aligned}\alpha_{T,1} &\sim \mathcal{N}(0,\sigma_\alpha^2),\\ \beta_{T,1} &\sim \mathcal{N}(0,\sigma_\beta^2).\end{aligned} \tag{16}$$

Analogously to Section 3.1, STZ constraints are required for each match day to achieve identifiability. The variance parameters σ_α^2 and σ_β^2 are assigned the following hyperpriors:

$$\sigma_\alpha^2,\sigma_\beta^2 \sim \text{InvGamma}(0.001,0.001). \tag{17}$$

As is evident from Table 2, the assumption of dynamic ability parameters does not improve the fit of the model for the Italian SuperLega data that we consider here. However, modelling dynamic patterns may be very useful in other leagues when considering distinct subsets of a league (such as regular season and play-offs). Fig. 3 displays posterior 95% intervals for the dynamic point abilities for the Italian SuperLega 2017–2018 data, whereas the corresponding marginal posterior distributions for the standard deviations σ_α and σ_β are plotted in Fig. 4: the time variability is negligible in the Italian data that we analyse in this paper. Although the within-team variability of the points abilities may look high, we believe that it is reasonable, since it corresponds only to a small portion of the total variability of the response measurement

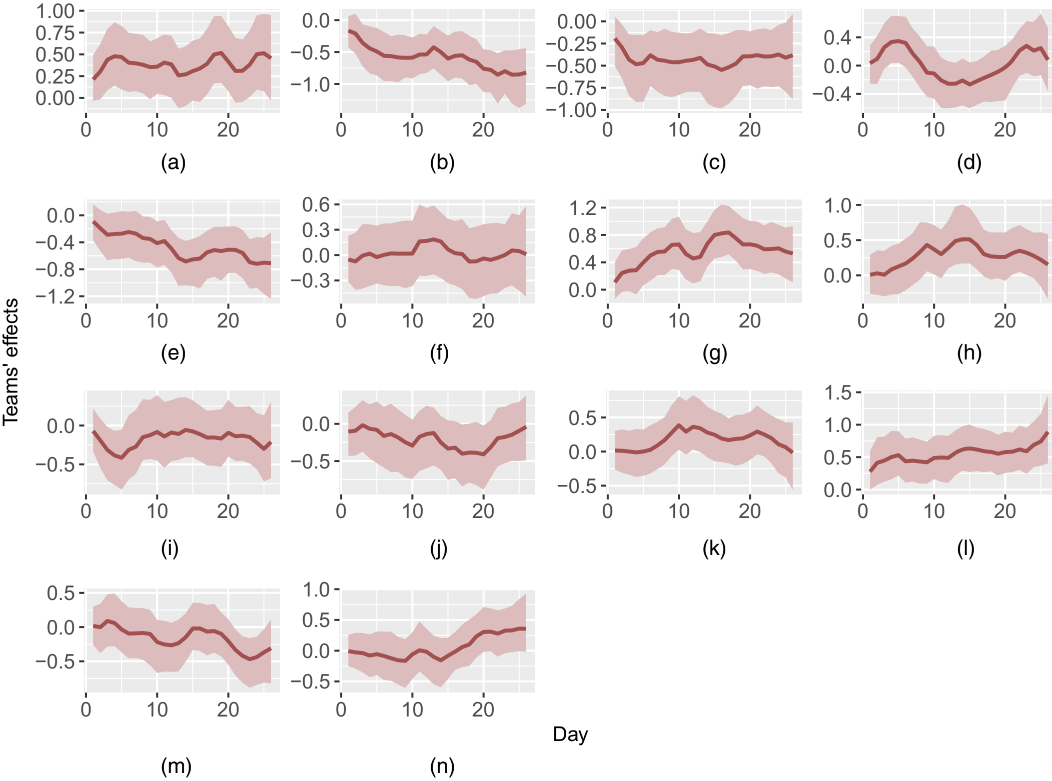


Fig. 3. Posterior medians and 95% density intervals for the dynamic points abilities parameters β for the Italian SuperLega 2017–2018 data: (a) Azimut Modena; (b) BCC Castellana Grotte; (c) Biosì Sora; (d) Bunge Ravenna; (e) Callipo Vibo Valentia; (f) Calzedonia Verona; (g) Cucine Lube Civitanova; (h) Diatec Trentino; (i) Gi Group Monza; (j) Kioene Padova; (k) Revivre Milano; (l) Sir Safety Perugia; (m) Taiwan Exc. Latina; (n) Wixo LPR Piacenza

of this model's component (which is the logit of the proportion of points earned by the losing team after removing the extra points played because of ties). To confirm this, we have calculated the proportion of points won by the losing team (after removing extra points played because of ties) which, on average (plus or minus posterior standard deviation), was found to be equal to 0.795 ± 0.12. The corresponding logits of these proportions were found to be equal to 1.5 on average with standard deviation 0.742. According to our posterior results, the posterior standard deviations of the points parameters were found to be approximately 0.10–0.15, which corresponds only to 12–20% of the total variability of the response measurement (as we stated previously).

4.4. Modelling extra points as a function of team abilities

In this section we explore whether the probability of observing extra points due to ties (i.e. the inflation component probability) can be written as a function of the set and/or points abilities. In principle, we would expect a negative association between the team ability differences and the probability of playing extra points in each set. Therefore, the closer the two teams are, the higher is the probability of being tied in each set and as a result of playing for extra points. For the probability π_s we considered various versions of the linear predictor which can be summarized by

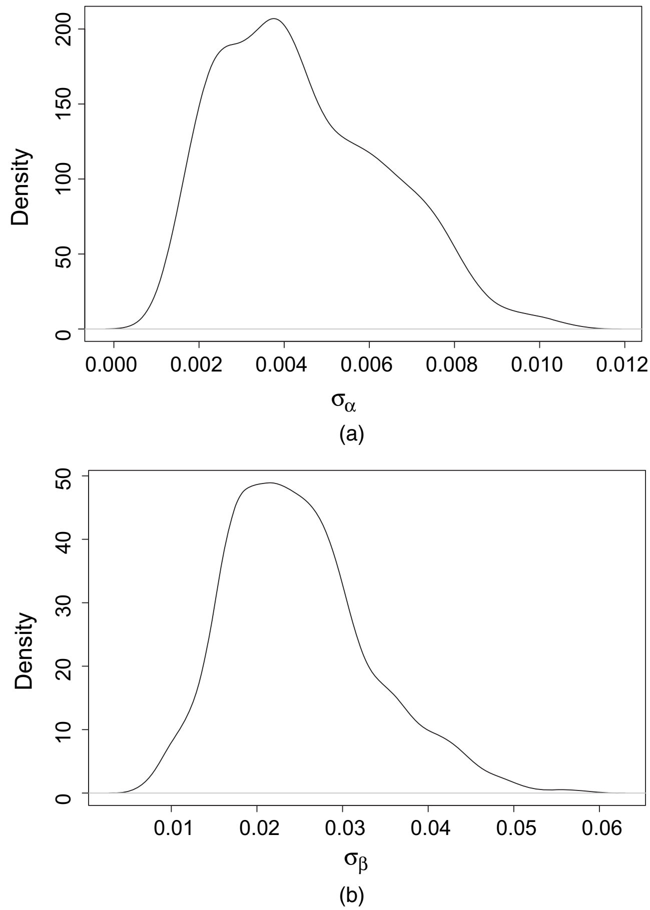


Fig. 4. Posterior marginal distribution for the standard deviations of the dynamic set and points abilities for the Italian SuperLega 2017–2018 data (a) for sets, σ_α, and (b) for points, σ_β

$$\begin{aligned} &\text{logit}(\pi_s) = m + \delta\,\Phi(\alpha_{A(s)} - \alpha_{B(s)}) + \gamma\,\Phi(\beta_{A(s)} - \beta_{B(s)}) \\ &m \sim \mathcal{N}(0,1), \qquad \delta, \gamma \sim \mathcal{N}(0, 10^6), \qquad \lambda \sim \text{LN}(0,1), \end{aligned} \tag{18}$$

where $\text{LN}(\mu, \sigma^2)$ denotes the log-normal distribution with parameters μ and σ^2, m is a constant parameter, δ and γ are the coefficients that are associated with the set and points abilities differences, and $\Phi(\cdot)$ is a specific function for the set or points abilities differences, depending on the model that we consider:

(a) constant probability by using the null function, $\Phi(x) = 0$;
(b) linear effect of ability differences by using the linear function, $\Phi(x) = c_1 x$;
(c) linear effect of absolute ability differences by using the linear absolute function, $\Phi(x) = c_1 |x|$;
(d) quadratic effect of ability differences by using the quadratic function, $\Phi(x) = c_1 x + c_2 x^2$;
(e) quadratic effect of absolute ability differences by using the quadratic function of absolute values of x, $\Phi(x) = c_1 |x| + c_2 x^2$.

Here c_0, c_1 and c_2 are further parameters for estimation. In Table 3 the DIC-values and the effective number of parameters for each model are reported with respect to the Italian SuperLega

Table 3. Details of the fitted logistic–ZIP truncated negative binomial models with different structure on the probability of extra points for the Italian SuperLega 2017–2018 season†

Model‡	*Structural assumption about the probability of extra points*	$\Phi(x)$	*Number of effective parameters*	*DIC*
9	Constant probability	$\Phi(x)=0$	124	4521.1
12	Linear effects	$\Phi(x)=c_1 x$	125	4523.0
13	Linear absolute effects	$\Phi(x)=c_1 \vert x\vert$	125	4525.8
14	Quadratic effects	$\Phi(x)=c_1 x+c_2 x^2$	126	4524.6
15	Quadratic absolute effects	$\Phi(x)=c_1 \vert x\vert+c_2 x^2$	126	4526.6

†MCMC sampling, 3000 iterations, `rjags` package.
‡In all models we use a logistic regression model for the sets (equations (1) and (2)), the ZIP model for extra points and the truncated negative binomial model for points, and connected team abilities and extra set abilities for Verona and Padova.

2017–2018 data. No model improves the fit that we obtain when using the constant Poisson rate for the ZIP component of the extra points (model 9).

5. Analysis and results of the Italian SuperLega 2017–2018 data

5.1. Data and computational details

The data come from the regular season of the Italian SuperLega 2017–2018 and consist of a seasonal sample of 680 set observations, for a total number of 182 matches and 14 teams involved (source, Italian SuperLega web page `https://www.legavolley.it/category/superlega/`). Posterior estimates are obtained with the `rjags` R package (MCMC sampling from the posterior distribution by using Gibbs sampling), for a total of 3000 iterations obtained by three parallel chains of 1000 iterations and a burn-in period of 100. Following the suggestions of Gelman *et al.* (2013), we monitored the convergence of our MCMC algorithms by checking the effective sample size of each chain parameter and by implementing the Gelman–Rubin statistic (Gelman and Rubin, 1992) which resulted in a lower-than-usual threshold of 1.1 for all the parameters (details are given in Table 4).

In 39 matches out of 182 (21.4%) the final winner was determined in the tie-break (i.e. in the fifth set), whereas in 101 out of 680 (14.8%) sets it was required to play for extra points to declare the winner of the corresponding set.

5.2. Interpretation of the model selected

Here we focus on the analysis of the Italian SupeLega 2017–2018 data by using the model that was selected in Sections 3.4 and 4.2 (model 9 in Table 2), i.e. the model with connected abilities for all teams and extra set abilities only for Verona and Padova. The complete model formulation, including likelihood specification, priors and identifiability constraints, is summarized in Table 4. Posterior estimates for the set home advantage H^{set}, the points home advantage H^{points}, the intercept μ and the ZIP parameters λ and m are reported in Table 5: there is a clear indication of home advantage which seems to be smaller for the set level (posterior median 0.16; 95% posterior interval marginally containing zero), and higher at the points level (posterior median 0.20; 95% posterior interval not containing zero). In terms of percentage change, this means that in a game between two teams of equal strength we expect that the home team will have 17%

Table 4. Final model formulation for model 9 of Table 2: likelihood, priors and identifiability constraints (24 parameters in total)

Likelihood	
Total set points	$Z_s = Y_s + O_s$
Losing team points	$Y_s \mid W_s \sim \text{NegBin}(r_s, p_s)\,\mathscr{I}(Y_s < r_s - 2)$
Extra points	$O_s \sim \text{ZIPoisson}(\pi_s, \lambda)$
Home win indicator	$W_s \sim \text{Bernoulli}(\omega_s)$
Logit of set win	$\text{logit}(\omega_s) = H^{\text{set}} + v_1(\alpha_{A(s)} - \alpha_{B(s)}) + v_2\theta(\beta_{A(s)} - \beta_{B(s)})$
ZIP: log-odds of extra points	$\text{logit}(\pi_s) = m + \delta\,\Phi(\alpha_{A(s)} - \alpha_{B(s)}) + \gamma\,\Phi(\beta_{A(s)} - \beta_{B(s)})$
Linear predictor for points	$\eta_s = \mu + (1 - W_s)H^{\text{points}} + (\beta_{A(s)} - \beta_{B(s)})(1 - 2W_s)$
Winning team points probability	$p_s = 1/\{1 + \exp(\eta_s)\}$
Required success points	$r_s = 25 - 10\,\mathscr{I}(R_s = 5)$
Constraints	
Extra set abilities (only for specific teams)†	$\alpha_T = \alpha_T^*$ with $\alpha_T^* \equiv 0,\ T \neq 10, 12$
STZ for points abilities	$\beta_T = \beta_T^* - \bar{\beta}^*;\ \bar{\beta}^* = (1/N_T)\Sigma_{T=1}^{N_T}\beta_T^*$
Connecting abilities set-up	$v_1 = v_2 = 1$
ZIP general probability function	$\Phi(x) = c_0 x + c_1\lvert x\rvert + c_2 x^2$
ZIP finally selected model	$c_0 = c_1 = c_2 = 0$
Priors	
Set abilities (Padova and Verona)	$\alpha_{10}^*, \alpha_{12}^* \sim \mathcal{N}(0, 2^2)$
Points abilities (unconstrained)	$\beta_T^* \sim \mathcal{N}(0, 2^2)$
Constant for points	$\mu \sim \mathcal{N}(0, 10^6)$
Points ability coefficient for sets	$\theta \sim \mathcal{N}(0, 10^6)$
Home effects	$H^{\text{points}}, H^{\text{set}} \sim \mathcal{N}(0, 10^6)$
Constant for extra points	$m \sim \mathcal{N}(0, 1)$
Poisson rate for ZIP distribution	$\lambda \sim \text{LN}(0, 1)$

†This parameterization is used in the final model, where extra abilities are considered only for two specific teams; in the general formulation with all set abilities then the STZ parameterization is recommended where $\alpha_T = \alpha_T^* - \bar{\alpha}^*$, $\bar{\alpha}^* = (1/N_T)\Sigma_{T=1}^{N_T}\alpha_T^*$ and N_T is the number of teams under consideration. Recommended prior when all set abilities are used with STZ: $\alpha_T^* \sim \mathcal{N}(0, 2^2)$.

(posterior 95% interval (0%, 42%)) and 22% (posterior 95% interval (8%, 40%)) higher odds of winning a set and a point respectively.

The scaling factor θ (posterior mean 4.60) shows a very strong positive association between the points abilities and the probability of wining the set, as assumed in equation (14). No evidence was found for the parameters γ and δ in equation (18), describing the influence of set and points abilities differences respectively, on the probability of observing extra points; however, we feel that these parameters could be beneficial for other data sets or other leagues.

The number of effective sample sizes, n_eff, and Gelman–Rubin statistics $\hat{R}$ appear to be quite satisfactory for each parameter.

The 95% posterior intervals for set and points team abilities are displayed in Figs 5 and 6 for the model with connected abilities and extra set abilities for all teams (model 7 in Table 2) and the corresponding model with extra set abilities only for Verona and Padova (model 9 in Table 2) respectively: dots denote the posterior medians, whereas thicker and thinner segments yield the 50% and the 95% posterior density intervals respectively. From Fig. 5(b) we observe that the points abilities of Verona and Padova are slightly misaligned compared with the actual rank. Moreover, from Fig. 5(a) we note that all the 95% posterior intervals of the extra set abilities contain the value zero. However, for Verona and Padova we obtained a marginal effect in terms of set extra abilities. For this reason, we moved to the model with connected abilities and set

Table 5. Posterior summaries for the sets home H^{set}, the points home H^{points}, the connecting abilities scaling factor θ, the intercept μ and ZIP parameters λ and m for the Italian SuperLega 2017–2018 data using the ZIP truncated negative binomial model with connected abilities and extra set abilities for Verona and Padova (model 9 in Table 2)†

Description	*Parameter*	*Mean*	*Median*	*Standard deviation*	*2.5%*	*97.5%*	*n_eff*	$\hat{R}$
Set home advantage	H^{set}	0.16	0.15	0.09	−0.01	0.33	2839	1.00
Points home advantage	H^{points}	0.20	0.20	0.06	0.08	0.34	1664	1.00
Connecting abilities	θ	4.60	4.52	0.80	3.36	6.30	2310	1.00
Intercept	μ	0.36	0.36	0.05	0.27	0.46	2213	1.00
ZIP Poisson rate	λ	3.97	3.97	1.07	3.45	4.52	2115	1.00
Tie probability intercept	m	2.12	2.13	0.13	1.87	2.39	2460	1.00

†Also reported are the effective sample size n_eff and the Gelman–Rubin statistics $\hat{R}$.

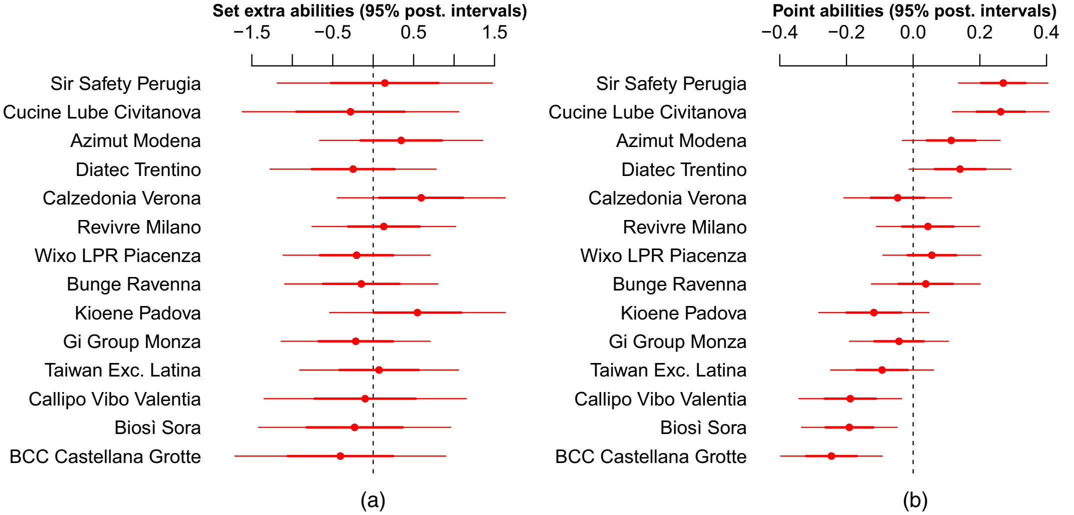
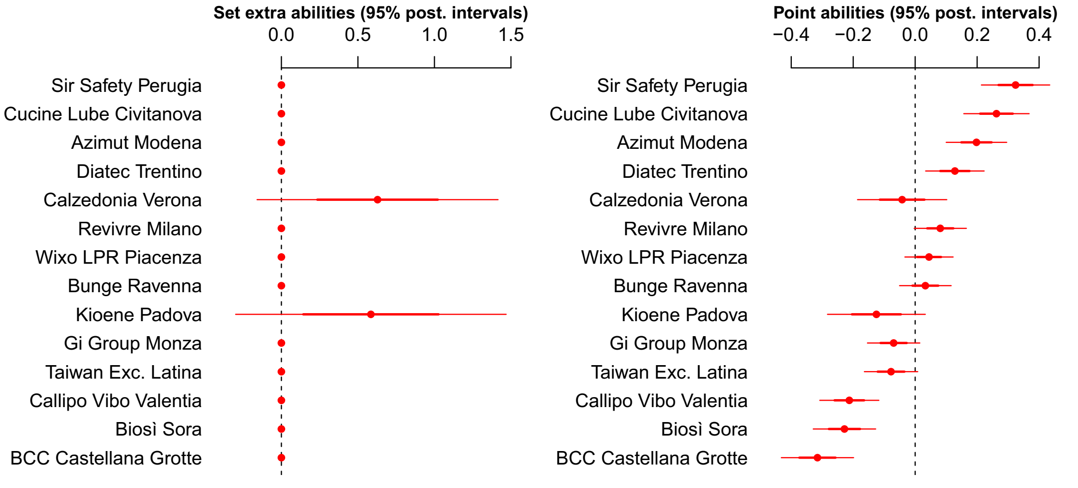


Fig. 5. Forest plots of posterior error bars for set and points team abilities for the Italian SuperLega 2017–2018 data, using the ZIP truncated negative binomial model with connected abilities (model 7 in Table 2) (intervals are ordered by the actual final ranking of each team): ▬, 50% posterior intervals; —, 95% posterior intervals; ●, posterior medians

extra abilities for these two teams only (Fig. 6), by forcing all the remaining extra set abilities to be restricted to 0. In such a way, we reduced the model complexity by 12 parameters whereas we improved our final model in terms of predictive accuracy (see Section 4.2). Finally, for this model we can specify the actual effects (abilities) of each team on the winning probability of a set as

$$\alpha'_T = \alpha_T + \theta\beta_T.$$

These overall set abilities are depicted in Fig. 7 where we can clearly see that the overall abilities depict the actual observed rankings since the extra set abilities for Padova and Verona correct for any inconsistencies between the set and point level concerning the efficiency of the teams.



(a) (b)

Fig. 6. Forest plots of posterior error bars for set and points team abilities for the Italian SuperLega 2017–2018 data, using the ZIP truncated negative binomial model with connected abilities and extra set abilities for Verona and Padova (model 9 in Table 2) (ordered by the actual final ranking of each team): ——, 50% posterior intervals; ——, 95% posterior intervals; ●, posterior medians

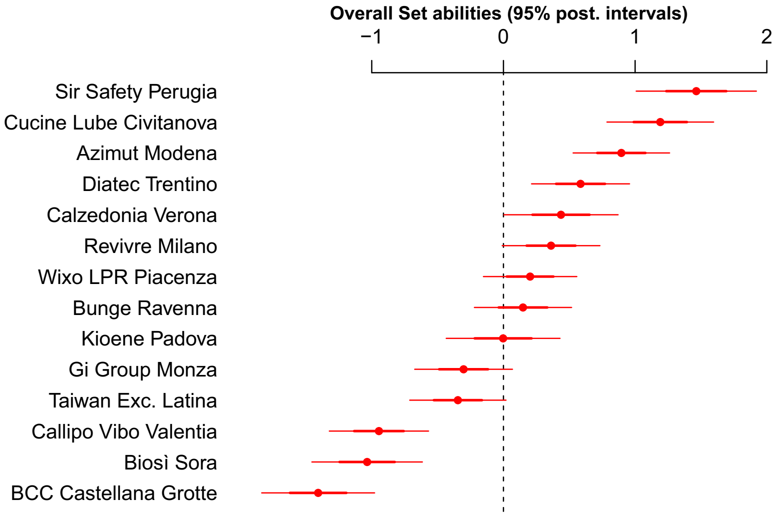


Fig. 7. Forest plots of posterior error bars for overall set team abilities $\alpha'_T = \alpha_T + \theta\beta_T$ for the Italian SuperLega 2017–2018 data, using the ZIP truncated negative binomial model with connected abilities and extra set abilities only for Verona and Padova and constant ZIP probability for extra points (model 9 in Table 2) (intervals are ordered by the actual final rank of each team): ——, 50% posterior intervals; ——, 95% posterior intervals; ●, posterior medians

5.3. League reconstruction and predictive measures of fit

To assess the in-sample predictive accuracy of our final model, we reconstruct the league in terms of final points and rank positions from the predictive distribution of the model. To do so, for each iteration of the MCMC sampling, we draw values from the model's sampling likelihood (see Table 4) for the given set of parameter values generated at each each iteration, resulting in

Table 6. Algorithm 1: volleyball stochastic league reconstruction algorithm

Input: $(\pi_G^{(t)}, p_G^{(t)}, \lambda^{(t)}, \omega_G^{(t)}, \mathrm{HT}_G, \mathrm{AT}_G)$; MCMC values of the parameters of the model for each game G and MCMC iteration t; HT_G and AT_G denote the home and the away teams in each game G

Output: $\mathbf{L}^{(t)} = (P_T^{(t)}, \mathrm{SW}_T^{(t)}, \mathrm{SL}_T^{(t)}, \mathrm{GPW}_T^{(t)}, \mathrm{GPL}_T^{(t)})$, T_{mcmc} leagues with number of league points, total sets won and lost and total number of game points won and lost respectively, for each team T and MCMC iteration t

for $t = 1$ *to* T_{mcmc} *do*

  initialize the league output for iteration t

  $\mathbf{L}^{(t)} = \mathbf{0}$;

  for $G = 1$ *to* N_G *do*

    initialize set and points for each game

    $S_H = S_W = P_H = P_W = 0$;

    while $\max\{S_H, S_W\} < 3$ (number of sets won by each team lower than 3) *do*

      $R = S_H + S_W$ (calculate total number of sets until now)

      $r < -15 + 10\mathcal{I}(R = 5)$ (calculate maximum number of points)

      $W \sim \mathrm{Bernoulli}(\omega_G^{(t)})$ (generate the winner of the set)

      $O \sim \mathrm{ZIPoisson}(\pi_G^{(t)}, \lambda^{(t)})$ (generate the extra points)

      $Y \sim \mathrm{NegBin}(r, p_G^{(t)})\mathcal{I}(Y < r - 2)$ (generate the points of the losing team)

      $S_H = S_H + W$; $S_A = S_A + 1 - W$ (update the winning sets of the teams)

      $P_W = r + O$; $P_L = Y + O$ (points of the winning and losing team)

      $P_H = P_H + WP_W + (1 - W)P_L$ (update the total points of the home team)

      $P_A = P_A + (1 - W)P_W + WP_L$ (update the total points of the away team)

    end

    updating the league parameters for the home team

    $P_{\mathrm{HT}_G}^{(t)} = P_{\mathrm{HT}_G}^{(t)} + 3\mathcal{I}(S_H - S_A > 1) + \mathcal{I}(S_H - S_A = 1)$ (league points)

    $\mathrm{SW}_{\mathrm{HT}_G}^{(t)} = \mathrm{SW}_{\mathrm{HT}_G}^{(t)} + S_H$ (sets won $\mathrm{SL}_{\mathrm{HT}_G}^{(t)} = \mathrm{SW}_{\mathrm{HT}_G}^{(t)} + S_A$; sets lost)

    $\mathrm{PW}_{\mathrm{HT}_G}^{(t)} = \mathrm{PW}_{\mathrm{HT}_G}^{(t)} + P_H$ (game points won)

    $\mathrm{PL}_{\mathrm{HT}_G}^{(t)} = \mathrm{PL}_{\mathrm{HT}_G}^{(t)} + P_A$ (game points lost)

    updating the league parameters for the away team

    $P_{\mathrm{AT}_G}^{(t)} = P_{\mathrm{AT}_G}^{(t)} + 3\mathcal{I}(S_A - S_H > 1) + \mathcal{I}(S_A - S_H = 1)$ (league points)

    $\mathrm{SW}_{\mathrm{AT}_G}^{(t)} = \mathrm{SW}_{\mathrm{AT}_G}^{(t)} + S_A$ (sets won)

    $\mathrm{SL}_{\mathrm{AT}_G}^{(t)} = \mathrm{SW}_{\mathrm{AT}_G}^{(t)} + S_H$ (sets lost)

    $\mathrm{PW}_{\mathrm{AT}_G}^{(t)} = \mathrm{PW}_{\mathrm{AT}_G}^{(t)} + P_A$ (game points won)

    $\mathrm{PL}_{\mathrm{AT}_G}^{(t)} = \mathrm{PL}_{\mathrm{AT}_G}^{(t)} + P_H$ (game points lost)

  end

end

Return league results $\mathbf{L}^{(t)}$ for $t = 1, \ldots, T_{\mathrm{mcmc}}$; indices $t = 1, \ldots, T_{\mathrm{mcmc}}$ (T_{mcmc}, number of MCMC iterations); $G = 1, \ldots, N_G$ (N_G, number of games); $T = 1, \ldots, N_T$ (N_T, number of teams in the league)

a new sample of match results obtained from the posterior predictive distribution of the model. It is worth mentioning that predicting future matches in volleyball is not as easy as in other sports. First, we need to simulate the actual number of sets for each game by using equations (1) and (2); we terminate the sets' simulation when one of the two teams wins three sets first. Then, we calculate the number of points that the teams collected after winning a game at each reconstructed league of each iteration. For each MCMC iteration, each match is simulated in terms of both points and sets from their posterior predictive distribution. This means that a new replicated data set y^{rep} is sampled from

$$p(y^{\text{rep}}|y) = \int_{\Theta} p(y^{\text{rep}}|\theta)\, p(\theta|y) \mathrm{d}\theta, \tag{19}$$

where $p(\theta|y)$ is the posterior distribution for the parameter θ, and $p(y^{\text{rep}}|\theta)$ is the sampling distribution for the hypothetical values. In many applications, the posterior predictive distribution in equation (19) is not available in a closed form, and therefore we sample from it by using MCMC methods. Algorithm 1 (Table 6) presents the entire procedure for the stochastic league reconstruction from the posterior predictive distribution: the entire league is simulated and the final points and rankings are obtained for further analysis.

Table 7 reports the expected final points and the 95% predictive intervals estimated from the MCMC algorithm along with the observed points and the actual team rankings: the points that are reported in Table 7 are obtained by computing the median and the 95% predictive intervals of the posterior predictive distribution of the final points, for each team. The agreement between the actual and the expected number of points is remarkable since the difference is at most equal to four points, and the simulated positions mirror perfectly the final observed rank, with the exception of switch in the expected positions between Sora and Vibo Valentia. Generally speaking, the model's in-sample predictions mirror almost perfectly the observed results in terms of expected points and final rank positions.

Beyond that, it is straightforward to obtain a measure of model goodness of fit at the points level. For each set s, $s=1, \ldots, S$, we denote by d_s the set points difference $Y_s^{\text{A}} - Y_s^{\text{B}}$, and with $\tilde{d}_s^{(t)}$ the corresponding points difference arising from the tth MCMC replications, $y_s^{\text{A, rep}(t)} - y_s^{\text{B, rep}(t)}$. Once we replicate new existing values from our model, it is of interest to assess how far they are if compared with the actual data that we observed. Fig. 8 displays the posterior predictive distribution of each $\tilde{d}_s^{(t)}$ (the shaded areas) plotted against the true observed distribution for d_s: there is quite good agreement between the replicated distributions and the observed distribution, and this is further corroboration of the goodness of fit of our final model (the plot was obtained through the `bayesplot` package (Gabry and Mahr, 2019), which always provides a continuous approximation for discrete dstributions).

5.4. Out-of-sample prediction

Our final task is to assess the out-of-sample predictive ability of our proposed model. As usual, we expect lower predictive accuracy than that obtained for in-sample measures. Nevertheless, it is crucial to ensure that our proposed model has satisfactory predictive performance.

The procedure is similar to the stochastic league regeneration that was described in Section 5.3 and algorithm 1. The main difference here is that a specific number of games is now known and fixed (i.e. data) and only the remaining games are generated from the predictive distribution, whereas in Section 5.3 the data for the whole season were available and the data of a new full league were regenerated assuming that we have exactly the same characteristics and team abilities as those observed. Here we proceed with two scenarios:

(a) the mid-season prediction scenario and
(b) the play-off prediction scenario.

In the first case we assume that we are at the middle of the season where half of the game results are available and we try to predict the final league standings. In the second scenario, we use the full league data to predict the final results in the play-off phase.

With the latter case, a further complication arises which is due to the formation of the *play-off* phase. In this after-season tournament, the best eight teams are competing from the quarter-finals: the team that wins three matches first goes to the next stage. Thus, each game say between

Table 7. Final reconstructed league for the Italian SuperLega 2017–2018 data, using the ZIP truncated negative binomial model with connected abilities and extra set abilities for Verona and Padova (model 9 in Table 2)

Predicted rank†	*Team*	*Points*	
		Expected (actual)	*95% credible interval‡*
1 (–)	Sir Safety Perugia	66 (70)	(57–73)
2 (–)	Cucine Lube Civitanova	61 (64)	(51–69)
3 (–)	Azimut Modena	56 (60)	(45–66)
4 (–)	Diatec Trentino	52 (51)	(40–63)
5 (–)	Calzedonia Verona	49 (50)	(35–60)
6 (–)	Revivre Milano	45 (44)	(32–57)
7 (–)	Wixo LPR Piacenza	42 (42)	(29–54)
8 (–)	Bunge Ravenna	40 (41)	(27–52)
9 (–)	Kioene Padova	35 (35)	(22–48)
10 (–)	Gi Group Monza	28 (28)	(17–42)
11 (–)	Taiwan Exc. Latina	28 (25)	(18–40)
12 (+1)	Biosì Sora	17 (13)	(8–26)
13 (–1)	Callipo Vibo Valentia	16 (13)	(9–27)
14 (–)	BCC Castellana Grotte	13 (10)	(5–21)

†Rank based on the expected predicted points; in parentheses is the change in predictive ranking in comparison with the actual ranking.
‡95% credible interval based on the 2.5% and 97.5% percentiles of the posterior predictive distribution.

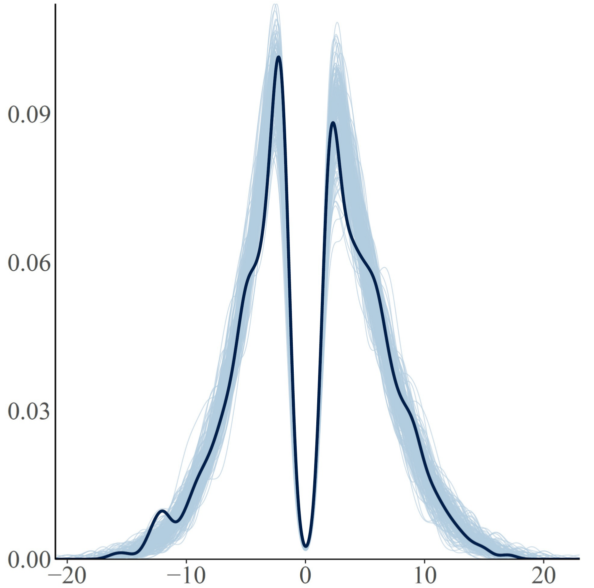


Fig. 8. Distribution of the observed set points differences $d_s = Y_s^{\mathrm{A}} - Y_s^{\mathrm{B}}$ (——) plotted against the posterior predictive distribution of $\tilde{d}_s^{(t)} = y_s^{\mathrm{A},\mathrm{rep}(t)} - y_s^{\mathrm{B},\mathrm{rep}(t)}$ (——) for the Italian SuperLega 2017–2018 data, using the ZIP truncated negative binomial model with connected abilities and extra set abilities for Verona and Padova (model 9 in Table 2)

team A and team B consists of a random number of repeated measurements, ranging from 3 to 5, whereas the set points system is the same as that described in the previous sections.

5.4.1. Mid-season prediction

In this section we predict the second half of the season by using the data for the first half of the season as a training set. This time point is important psychologically for sports fans. For example, in national football (soccer) leagues, there is the informal title of the 'Winter Champion' which is mainly promoted by sports media and newspapers. In terms of data, at this time point, a considerable amount of games is available and all teams have played against all their opponents once. Hence, reliable estimates of the model parameters and team abilities can be obtained, leading to sufficiently accurate predictions about the final league ranking.

To assess the model's predictive accuracy preliminarily we use the percentage of agreement between predicted games or sets from the MCMC sample and the observed games or sets: the posterior distribution of the percentage of agreement of the correctly predicted games for the mid-season prediction scenario is presented in Fig. 9(a). The posterior mean (plus or minus the posterior standard deviation) of the percentage of games predicted correctly in terms of

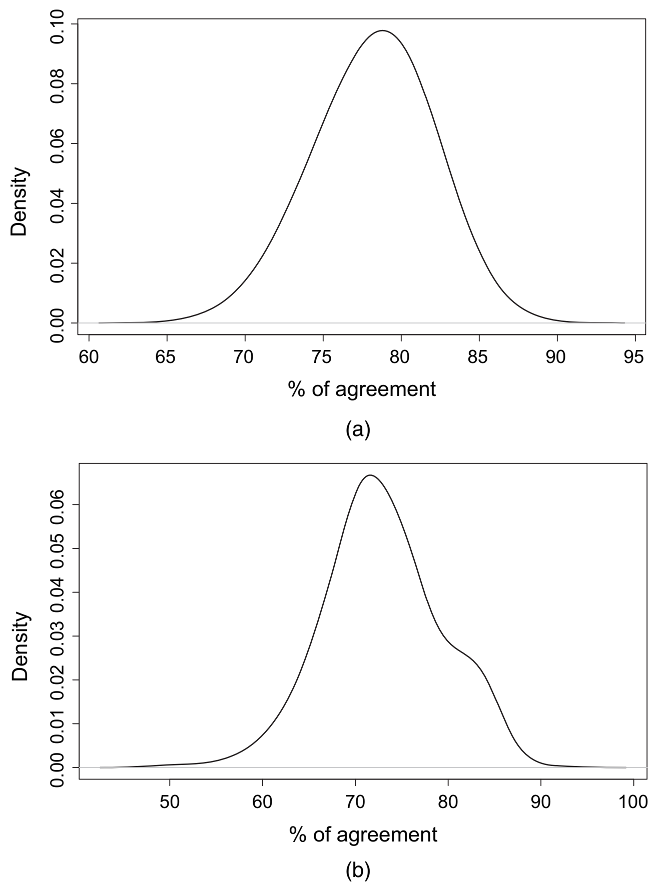


Fig. 9. Out-of-sample predictions: posterior distribution of percentage of correctly predicted games for (a) the mid-season prediction and (b) the play-off prediction phase for the Italian SuperLega 2017–2018 data, using the ZIP truncated negative binomial model with connected abilities and extra set abilities for Verona and Padova (model 9 in Table 2)

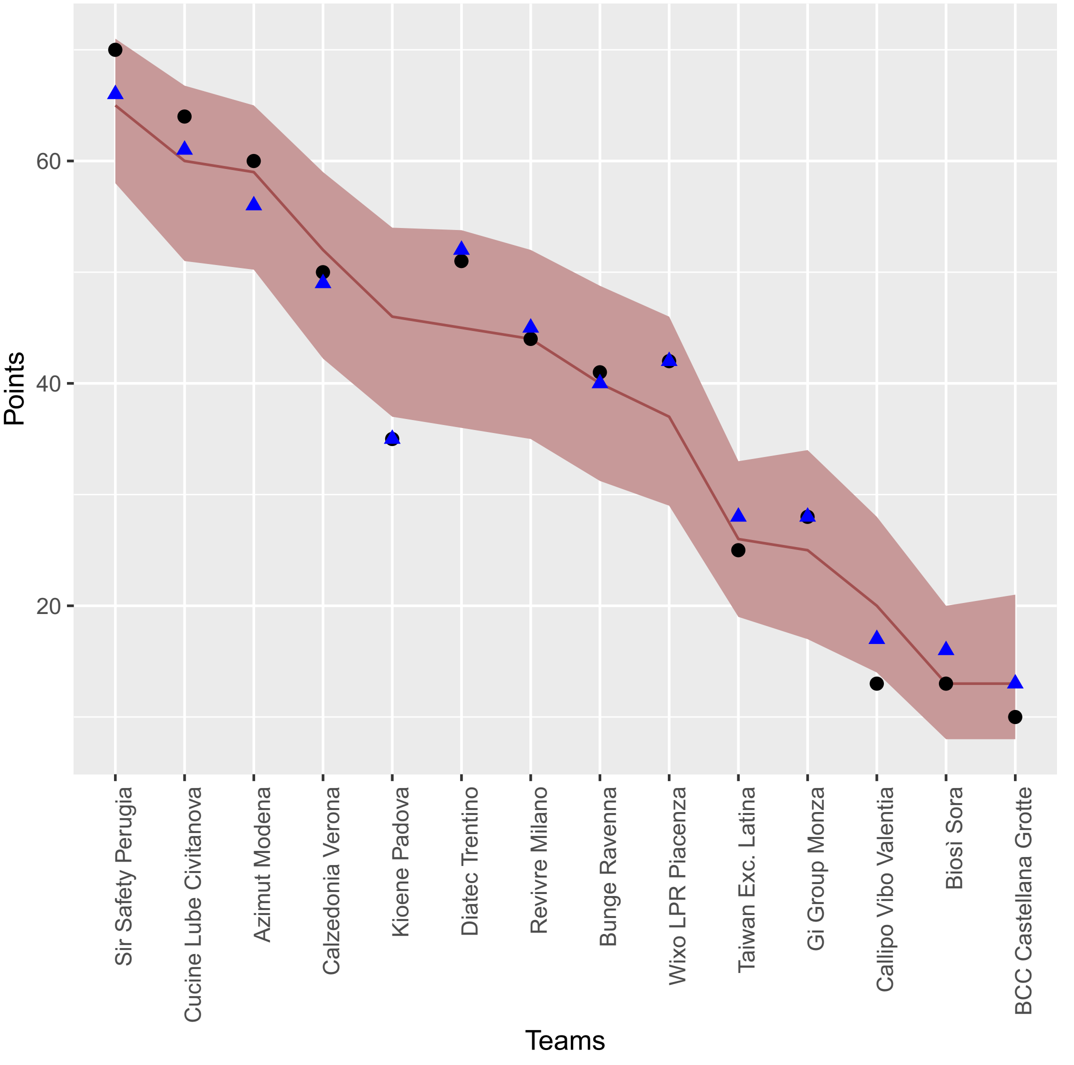


Fig. 10. Mid-season out-of-sample prediction: 95% predictive intervals (▬) from the posterior predictive distribution of the final points collected by each of the 14 teams of the Italian SuperLega 2017–2018 data along with the observed final points (●) and the expected points from the in-sample league reconstruction (▲; see Table 6): ——, median

the final result (win or loss) is found to be equal to 78.26% (±3%) whereas the corresponding percentage of sets predicted correctly (not displayed in the plot) is equal to 69.5% (±1%). Fig. 10 displays 95% predictive intervals (the ribbon) for the predicted achieved points of the 14 teams competing in the Italian SuperLega 2017–2018, using the first half as the training set, along with the observed final points (dots), and the expected points from the in-sample league reconstruction (triangles; see Table 7). At first glance, there is high agreement between the predicted and expected rankings, especially for the top-three teams (Perugia, Civitanova and Modena) and the bottom three teams (Vibo Valentia, Sora and Castellana Grotte). Padova and Trentino are the only teams whose observed points fall outside the 95% predictive intervals. Moreover, for three teams (Piacenza, Padova and Monza) the observed final points (dots)

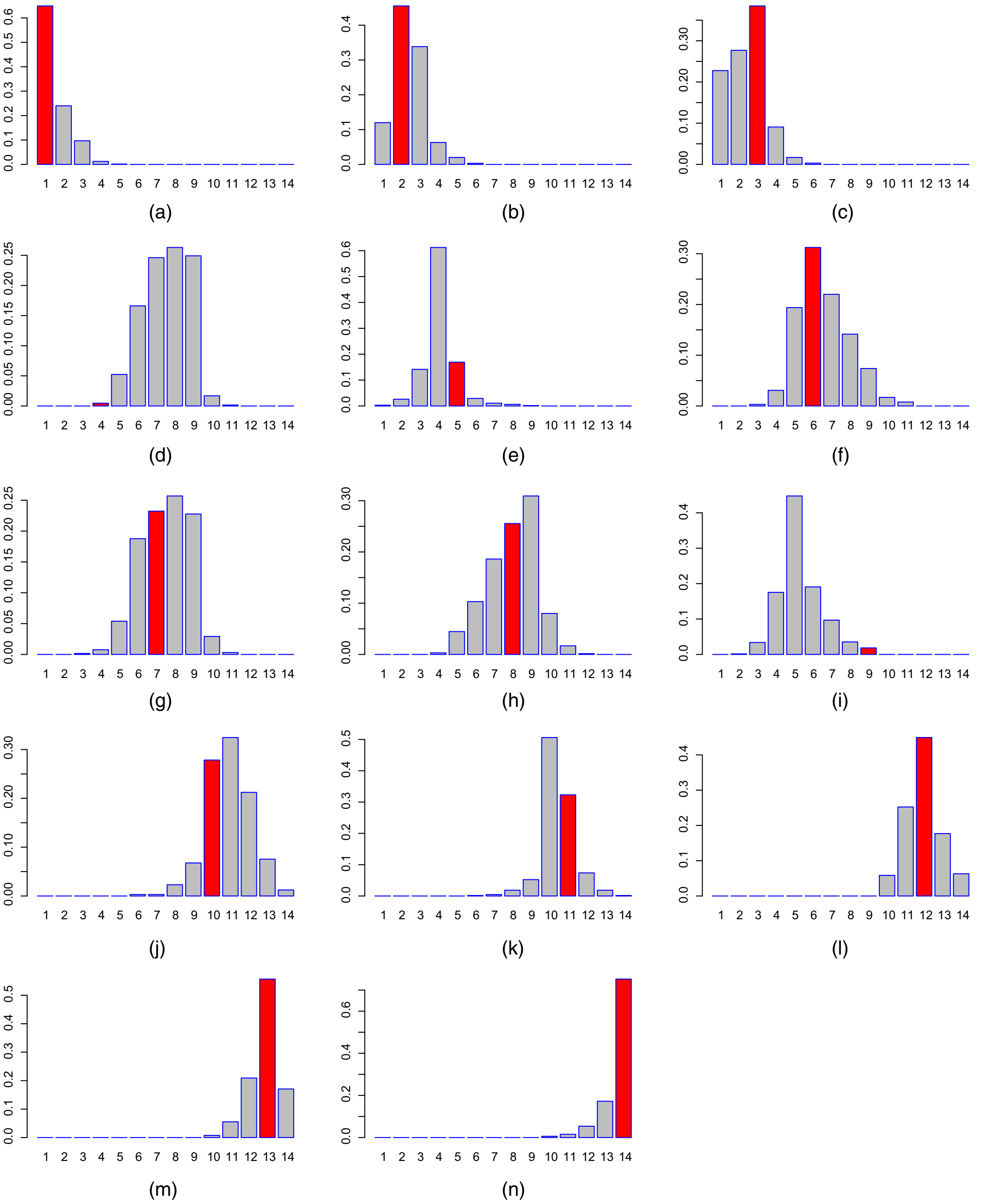


Fig. 11. Mid-season out-of-sample prediction—posterior predictive distribution of the final league ranking of each team in the Italian SuperLega 2017–2018 data (|, actual final rank position) (the final league ranking is given within parentheses after each team name): (a) Sir Safety Perugia (1); (b) Cucine Lube Civitanova (2); (c) Azimut Modena (3); (d) Diatec Trentino (4); (e) Calzedonia Verona (5); (f) Revivre Milano (6); (g) Wixo LPR Piacenza (7); (h) Bunge Ravenna (8); (i) Kioene Padova (9); (j) Gi Group Monza (10); (k) Taiwan Exc. Latina (11); (l) Callipo Vibo Valentia (12); (m) Biosì Sora (13); (n) BCC Castellana Grotte (14)

coincide with the in-sample simulated points (triangles). Fig. 11 shows the posterior predictive distribution of the league ranking of each team for the Italian SuperLega 2017–2018 data. The differently coloured bar, which corresponds to the actual rank, is the highest (i.e. is associated with the highest probability) both for the top three teams and for the bottom three teams, suggesting again good predictive ability for our model.

5.4.2. Play-off prediction using regular season games

Here we predict the games of the play-off phase by using the entire regular season as the training set. Fig. 9(b) displays the posterior distribution of the percentage of agreement of the correctly predicted games: the posterior mean (plus or minus the posterior standard deviation) is 73.06% (±6.05%). The posterior agreement of correctly predicted sets (not displayed in the plot) is 61.5% (±2.54%).

The play-off phase consists of a small knock-out tournament between the best eight teams in the regular season: Sir Safety Perugia, Cucina Lube Civitanova, Azimut Modena, Diatec Trentino, Calzedonia Verona, Revivre Milano, Bunge Ravenna and Wixo LPR Piacenza. Table 8 shows for each team the probability of winning in each play-off stage and progress to the next round until winning the tournament: Perugia is associated with the highest probability (0.75) of winning the play-off phase (Perugia actually won this phase, defeating Civitanova in the final) and, generally, reports the highest probabilities of progressing in each stage. Civitanova, the second-best team during the regular season, is associated with a high probability of entering the semifinals (0.87) and with the second-highest probability of winning the play-off (0.14). Modena and Trentino, who reached the semifinals, report high probabilities of progressing to the semifinals, 0.76 and 0.61 respectively, whereas Piacenza and Ravenna yield 0.13 and 0 probabilities of reaching the semifinals respectively (they were actually eliminated in the quarter-finals). Globally, these probabilities seem to mirror realistically the actual strength of each team in the final stage of the season.

Fig. 12 displays the play-off results of the matches that were actually played along with the posterior probabilities of progressing in each play-off stage. These probabilities have been obtained simply by considering the regular season results. As we can see, Perugia, the play-off winner, is associated with the highest probabilities in each match, especially against Ravenna and Trentino: although it may seem a little unrealistic that Perugia has probability 1 of beating Ravenna, this happens because the model's probability for Perugia is so high that during the MCMC simulation it is never defeated by Ravenna. To give some insight into this issue, when

Table 8. Play-off data: out-of-sample of prediction for the Italian SuperLega 2017–2018: probability of progressing to each stage of the play-off phase along with the actual results

Team	*Semifinal*	*Final*	*Winner*	*Actual*
Sir Safety Perugia	1.00	0.96	0.75	Winner
Cucine Lube Civitanova	0.87	0.58	0.14	Finalist
Azimut Modena	0.76	0.28	0.05	Semifinalist
Diatec Trentino	0.61	0.03	0.02	Semifinalist
Calzedonia Verona	0.39	0.01	0.00	Quarter-finalist
Revivre Milano	0.24	0.07	0.02	Quarter-finalist
Bunge Ravenna	0.00	0.00	0.00	Quarter-finalist
Wixo LPR Piacenza	0.13	0.06	0.01	Quarter-finalist

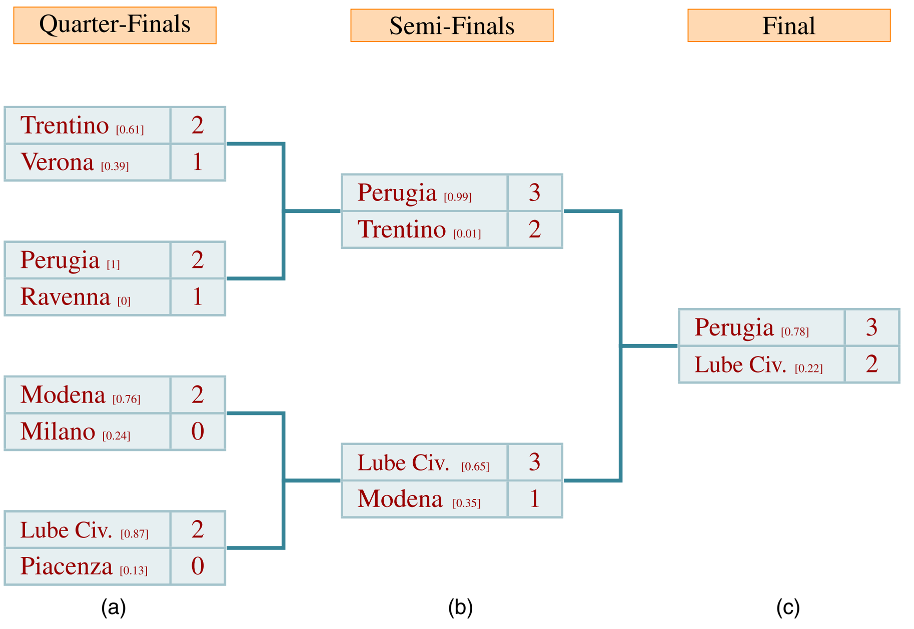


Fig. 12. Play-off out-of-sample prediction by using the full season data of the Italian SuperLega 2017–2018 data (probabilities for each team to progress to each play-off stage are reported in parentheses): (a) quarter-finals; (b) semifinals; (c) finals

playing at home and away Perugia has probabilities of 0.81 and 0.76 of winning a set against Ravenna respectively. In general, the highest probabilities are always associated with the teams that actually won the match and, consequently, progressed to the next stage.

Overall, our model yields good out-of-sample predictive performance, especially for the second mid-season.

6. Discussion

With this work, we propose a unified hierarchical framework for modelling volleyball data by using both outcomes (sets and points) of the game. The model follows a top-down approach (from sets to points) which initially seems counterintuitive but it helps to capture the characteristics of the game itself. The core model structure is based on truncated versions of the negative binomial distribution for the points. Moreover, the two levels of the outcomes (set and points) are connected via common abilities with extra set abilities when needed. Finally, the main characteristics of the game are taken into consideration including that

(a) the winner of the set is the team that first scores a prespecified number of points and
(b) the winner needs at least two points difference to win the set (and the set continues until this is achieved).

The latter characteristic is modelled via an extra latent component which is assumed to follow a ZIP distribution. We have also tested for the existence of correlations between sets (by using random effects), the appropriateness of dynamic set and points abilities and, finally, whether the probability of playing for extra points is influenced by the abilities of the teams (using a variety of functional forms). For the former check, there is some evidence that correlation might be present, as expected; however, the DIC provided similar predictive ability compared with the model without random effects, and therefore we proceeded with the simplified version of our model for computational convenience. Dynamic abilities do not seem to improve the model (although point ability dynamics seem to be more useful than set ability dynamics). Finally, the difference in the abilities between the two competing teams does not seem to alter the probability of playing for extra time (as we might expect). We have concluded our modelling quest by selecting a ZIP truncated model with connected abilities and extra set abilities for Verona and Padova and constant probability of observing extra points of each set. Posterior predictive checks show good agreement between our model and the observed results, and an overall exceptional ability to replicate the final rankings of the league. Concerning future out-of-sample predictions, our proposed model is well behaved with acceptable predictive accuracy for future matches both for the mid-season and for the play-off phase.

6.1. Prior considerations

6.1.1. Prior sensitivity analysis

We have used low information priors following an informal objective Bayes approach. For this reason, all prior distributions used are proper with relatively large variances. To ensure that the prior parameters selected had a minor influence on the inference, we also conducted a sensitivity analysis. Detailed results in the form of 95% posterior error bars can be found in the on-line supplementary material of the paper; see Figs A.1–A.3 in appendix A there.

6.1.2. Prior elicitation

Our modelling approach can be used also to incorporate prior information coming from historical data or from expert opinion. A standard method can be developed by using the power prior approach (Chen and Ibrahim, 2000) where historical data (even with incomplete or different covariate information) can be incorporated in our modelling approach. In this framework, the power parameter controls how reliable we believe this prior information to be and, therefore, how much it will influence the posterior results. Empirical Bayes approaches or fully Bayesian hierarchical approaches can be used to estimate the power parameter from the data; see for example the work of Gravestock and Held (2017, 2019) in medicine. Simpler approaches can be used when proportions of wins for each game are available, which is very common in sports. A simple approach based on generalized linear models can be used to build prior estimates of the team abilities at the game level and further to convert them into priors for the winning probabilities of sets and the associated team abilities. Finally, prior elicitation techniques for information coming from experts can be used to extract winning proportions and team abilities for each game. This can be implemented by following the earlier work of Chen *et al.* (1999) or the more recent and general framework of Albert *et al.* (2012). Nevertheless, experts, such as coaches, have rather limited skills and training on the quantification of their intuition or empirical knowledge. Therefore, the use of a downweight parameter (similar to the power parameter) is recommended as in Drikos *et al.* (2019). In this way, most of the inference will be based on the actual data whereas a small portion of it will come from expert knowledge. The latter information will correct for possible model misspecification or it may increase confidence for some specific parameters. Alternatively, we may incorporate predictions based on more reliable sources of prior information such as bookmakers as in Egidi *et al.* (2018) for football. Generally, prior elicitation in sports, and more specifically in volleyball, is an intriguing topic because of the general availability of historical data, the large amount of data published by betting companies and the easy access to sport experts (betting players, coaches and people working in the sports industry). For any of these cases, a more elaborate treatment is needed which is outside the scope of this work.

6.2. Limitations of the methodology proposed

Naturally, our approach embodies some assumptions which were tested by using the data of the Italian SuperLega 2017–2018 season. For example, for the final model we have assumed independence of sets and points conditionally on the explanatory information. This was tested for the specific data set by using both random effects and dynamic ability components. In both cases, no convincing evidence was found for incorporating either of these components in our final model. Moreover, we have assumed that the extra points follow a zero-inflated distribution. This was found to be sufficient in our implementation but further investigation is needed to validate this result.

6.2.1. Limitation I: team-specific covariates

The main aim of this paper is to validate a general modelling formulation for volleyball data. Therefore we focus on modelling the main characteristics of the game and in developing a simple basic 'vanilla' model by using only the competing teams in each game as covariates without considering any extra explanatory information for the simultaneous modelling and prediction of the two volleyball outcomes (i.e. sets and points). Therefore, a limitation of our approach is that we do not consider any further covariates to improve either the interpretational or the predictive ability of the model. In this direction, Gabrio (2020) used a variety of team-specific covariates: attack and defence types, service, service reception, blocks, passing abilities and roster quality. According to the results of Gabrio (2020), the use of some of these covariates may be beneficial in terms of both game explanation and predictive power. We are currently working on a more enriched volleyball data set to include other characteristics of the game by using two different approaches:

(a) descriptive, using the end-of-game statistics for interpretational reasons, and
(b) predictive, using statistics that are available at the beginning of the game to improve prediction.

Both of these approaches will be applied in combination with Bayesian variable-selection techniques.

6.2.2. Limitation II: separate attacking and defensive team abilities

As a referee pointed out, we did not use separate attacking and defensive team abilities either at the set or the points level. Concerning sets, it was not possible to separate attacking and defensive abilities because of identifiability reasons (also, no related models in the literature consider separate attacking and defensive team abilities). For the points level, since the actual response is represented by only the points of the losing team, conditionally on the winner of the set, we believe that the data will not have enough information to estimate both attacking and defensive abilities accurately. This was confirmed by the empirical results appearing in Table 2; see model 5. Although in this work we illustrate this finding by using a single-season data set, we intuitively believe that this result also holds generally for other leagues and data sets.

6.2.3. Limitation III: not considering the fatigue in the model

Another important athletic characteristic that we have not included in our modelling approach is fatigue. Although this is of interest for every sport, it might be of prominent importance for volleyball since sets are terminal important time points of each volleyball game and these are achieved sequentially. This can be done in several ways: using fixed trend effects or random effects in the modelling of sets. Nevertheless, the exploration of the function that optimally influences the sets might be cumbersome and therefore we believe that it should be treated separately, in other research work which is more focused on empirical results.

6.3. Comparisons with other methods

Early attempts to model volleyball were more focused on winning match and set probabilities through Markovian models (Ferrante and Fonseca, 2014), whereas Bayesian logistic regression to determine how the performance of individual skills affects the probability of scoring a point has been proposed by Miskin *et al.* (2010). However, the majority of previous studies in volleyball is not directed at modelling the entire game and validating the model strategy through league reconstruction and prediction for future matches.

As far as we know from reviewing the literature, the only attempt to implement a generative model for volleyball results was proposed by Gabrio (2020) for the women's volleyball Italian SuperLega 2017–2018 season. Our work presents some similarities with Gabrio (2020) such as the Bayesian framework and the posterior predictive validation of the model in terms of final points and ranking positions. However, distributional assumptions are deeply different: Gabrio's model is an adjustment of the double-Poisson model that has been adopted for football (Maher, 1982; Lee, 1997), whereas we propose a model which takes into consideration the special characteristics of the game itself which is different from goal-scoring team sports for which the double-Poisson model and its extensions were introduced.

6.4. Limitations regarding data-specific results

6.4.1. Considering only one data set

Concerning the empirical implementation, a limitation of our results concerns the use of a single-season data set from a single league. To check the model's adequacy in a wider sense, we should apply our model on a variety of seasons and tournaments. One problem in this direction is that volleyball data sets are not so widely available as in other sports (e.g. basketball and football). We are currently in touch with volleyball experts to obtain richer data sets (including game-specific covariates) and more seasons from the Greek league.

6.4.2. No covariates in the model structure

Moreover, additional covariates were not used here; therefore we have not touched on topics that other researchers have dealt with in the past (with simpler approaches), such as fatigue (Shaw *et al.*, 2008), the effect of service (Papadimitriou *et al.*, 2004; Lopez-Martinez and Palao, 2009) and the effect of specific volleyball skills on final outcomes (Miskin *et al.*, 2010; Drikos *et al.*, 2019; Gabrio, 2020).

6.4.3. Not considering the service advantage

There is an increasing literature which focuses on which team serves first, not only for volleyball (Shaw *et al.*, 2008) but also for tennis (Cohen-Zada *et al.*, 2018). This seems to be an important determinant at the points level (when we model the individual success of a point) but it might be less relevant at the accumulated points level for each set that we consider here. This is reasonable since the two teams are playing by having the advantage of serve in turn, especially when the two competing teams are of high level. Nevertheless, it might be more relevant for a tie-break where every small detail may count on determining the final winner. We believe that this effect will be minor when the game is unbalanced in terms of abilities (i.e. one team is much better than the other) but it might play a role (similar to the home effect) if the two teams are close in terms of abilities. Unfortunately, for our data set this information was not available but it might be of great interest to study this effect in the future.

6.5. Final conclusion

To conclude, we have introduced an alternative model for volleyball data which uses a top-down approach modelling both sets and points and by considering the sport-specific characteristics such as the extra points played due to the required two points margin of win. Our work focuses on the validation of the simple model without considering extra covariates structure or characteristics such as fatigue, serve or specific sport skills. We expect and hope that this work will initiate further quests for finding new methods and models for predicting and understanding volleyball and other sports belonging to the group of net and ball games.

7. Supplementary material

Electronic supplementary material with further plots and a sensitivity check is available from `https://github.com/LeoEgidi/Bayesian-Volleyball-paper`.

Acknowledgements

We thank Dr Sotirios Drikos for motivating us to work with volleyball data and the two referees who improved the quality of the manuscript with their fruitful comments. This research is financed by the Research Centre of Athens University of Economics and Business, in the framework of the project entitled 'Original scientific publications 2019'.

References

Albert, I., Donnet, S., Guihenneuc-Jouyaux, C., Low-Choy, S., Mengersen, K. and Rousseau, J. (2012) Combining expert opinions in prior elicitation. *Baysn Anal.*, **7**, 503–532.

Albert, J. (1992) A Bayesian analysis of a Poisson random effects model for home run hitters. *Am. Statistn*, **46**, 246–253.

Barnett, T. J., Brown, A. and Jackson, K. (2008) Modelling outcomes in volleyball. In *Proc. 9th Australasn Conf. Mathematics and Computers in Sport, Tweed Heads* (ed. J. Hammond), pp. 130–137. London: Chapman and Hall.

Carpita, M., Ciavolino, E. and Pasca, P. (2019) Exploring and modelling team performances of the kaggle European soccer database. *Statist. Modllng*, **19**, 74–101.

Chen, M.-H. and Ibrahim, J. G. (2000) Power prior distributions for regression models. *Statist. Sci.*, **15**, 46–60.

Chen, M.-H., Ibrahim, J. G. and Yiannoutsos, C. (1999) Prior elicitation, variable selection and Bayesian computation for logistic regression models. *J. R. Statist. Soc.* B, **61**, 223–242.

Cohen-Zada, D., Krumer, A. and Shapir, O. M. (2018) Testing the effect of serve order in tennis tiebreak. *J. Econ. Behav. Organizn*, **146**, 106–115.

Dixon, M. J. and Coles, S. G. (1997) Modelling association football scores and inefficiencies in the football betting market. *Appl. Statist.*, **46**, 265–280.

Drikos, S., Ntzoufras, I. and Apostolidis, N. (2019) Bayesian analysis of skills importance in World Champions men's volleyball across ages. *J. Comput. Sci. Sprt*, **18**, 24–44.

Egidi, L., Pauli, F. and Torelli, N. (2018) Combining historical data and bookmakers' odds in modelling football scores. *Statist. Modllng*, **18**, 436–459.

Facchinetti, T., Metulini, R. and Zuccolotto, P. (2019) Detecting and classifying moments in basketball matches using sensor tracked data. In *Smart Statistics for Smart Applications* (eds G. Arbia, S. Peluso, A. Pini and G. Rivellin), pp. 297–304. Milan: Pearson.

Fellingham, G. W., Hinkle, L. J. and Hunter, I. (2013) Importance of attack speed in volleyball. *J. Quant. Anal. Sprts*, **9**, 87–96.

Ferrante, M. and Fonseca, G. (2014) On the winning probabilities and mean durations of volleyball. *J. Quant. Anal. Sprts*, **10**, no. 2, 91–98.

Gabrio, A. (2020) Bayesian hierarchical models for the prediction of volleyball results. *J. Appl. Statist.*, to be published, doi 10.1080/02664763.2020.1723506.

Gabry, J. and Mahr, T. (2019) bayesplot: plotting for Bayesian models. *R Package Version 1.7.0*.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A. and Rubin, D. B. (2013) *Bayesian Data Analysis*. Boca Raton: Chapman and Hall–CRC.

Gelman, A., Jakulin, A., Pittau, M. G. and Su, Y.-S. (2008) A weakly informative default prior distribution for logistic and other regression models. *Ann. Appl. Statist.*, **2**, 1360–1383.

Gelman, A. and Rubin, D. B. (1992) Inference from iterative simulation using multiple sequences. *Statist. Sci.*, **7**, 457–472.
Gravestock, I. and Held, L. (2017) Adaptive power priors with empirical Bayes for clinical trials. *Pharm. Statist.*, **16**, 349–360.
Gravestock, I. and Held, L. (2019) Power priors based on multiple historical studies for binary outcomes. *Biometr. J.*, **61**, 1201–1218.
Harville, D. (1977) The use of linear-model methodology to rate high school or college football teams. *J. Am. Statist. Ass.*, **72**, 278–289.
Hass, Z. and Craig, B. A. (2018) Exploring the potential of the plus/minus in NCAA women's volleyball via the recovery of court presence information. *J. Sprts Analyt.*, **4**, 285–295.
Karlis, D. and Ntzoufras, I. (2003) Analysis of sports data by using bivariate Poisson models. *Statistician*, **52**, 381–393.
Karlis, D. and Ntzoufras, I. (2009) Bayesian modelling of football outcomes: using the Skellam's distribution for the goal difference. *IMA J. Mangmnt Math.*, **20**, 133–146.
Koopman, S. J. and Lit, R. (2015) A dynamic bivariate Poisson model for analysing and forecasting match results in the English Premier League. *J. R. Statist. Soc.* A, **178**, 167–186.
Lee, A. (1997) Modeling scores in the Premier League: is Manchester United really the best? *Chance*, **10,** 15–19.
Lee, K. T. and Chin, S. T. (2004) Strategies to serve or receive the service in volleyball. *Math. Meth. Oper. Res.*, **59**, 53–67.
Lopez-Martinez, A. and Palao, J. (2009) Effect of serve execution on serve efficacy in men's and women's beach volleyball. *Int. J. Appl. Sprts Sci.*, **21**, 1–16.
Maher, M. (1982) Modelling association football scores. *Statist. Neerland.*, **36**, 109–118.
Marcelino, R., Mesquita, I., Palao, J. M. and Sampaio, J. (2009) Home advantage in high-level volleyball varies according to set number. *J. Sprts Sci. Med.*, **8**, 352–356.
Mendes, F. G., Nascimento, J. V., Souza, E. R., Collet, C., Milistetd, M., Côté, J. and Carvalho, H. M. (2018) Retrospective analysis of accumulated structured practice: a bayesian multilevel analysis of elite Brazilian volleyball players. *High Abilty Stud.*, **29**, 255–269.
Metulini, R., Marisera, M. and Zuccolotto, P. (2017) Space-time analysis of movements in basketball using sensor data. In *Statistics and Data Science: New Challenges, New Generations*. Firenze: Firenze University Press.
Miskin, M. A., Fellingham, G. W. and Florence, L. W. (2010) Skill importance in women's volleyball. *J. Quant. Anal. Sprts*, **6**, no. 2, article 5.
Mosteller, F. (1952) The world series competition. *J. Am. Statist. Ass.*, **47**, 355–380.
Mosteller, F. (1970) Collegiate football scores, U.S.A. *J. Am. Statist. Ass.*, **65**, 35–48.
Owen, A. (2011) Dynamic Bayesian forecasting models of football match outcomes with estimation of the evolution variance parameter. *IMA J. Mangmnt Math.*, **22**, 99–113.
Papadimitriou, K., Pashali, E., Sermaki, I., Mellas, S. and Papas, M. (2004) The effect of the opponents' serve on the offensive actions of Greek setters in volleyball games. *Int. J. Perform. Anal. Sprt*, **4**, 23–33.
Plummer, M. (2018) rjags: Bayesian graphical models using MCMC. *R Package Version 4-8*.
Reep, C. and Benjamin, B. (1968) Skill and chance in association football. *J. R. Statist. Soc.* A, **131**, 581–585.
Rue, H. and Salvesen, Ø. (2000) Prediction and retrospective analysis of soccer matches in a league. *Statistician*, **49**, 399–418.
Schwertman, N. C., McCready, T. A. and Howard, L. (1991) Probability models for the NCAA regional basketball tournaments. *Am. Statistn*, **45**, 35–38.
Sepulveda, D., Lleida, U. D. and Alcaraz, A. G. D. (2017) Analysis of volleyball attack from the Markov chain model. In *Proc. Int. Congr. Complex Systems in Sport, Linking Theory and Practice* (eds C. Torrents, P. Passos and F. Cos), pp. 166–168. Frontiers Media.
Shaw, M. Y., Gribble, P. A., Atc, À. and Frye, J. L. (2008) Ankle bracing, fatigue, and time to stabilization in collegiate volleyball athletes. *J. Athl. Train.*, **43**, 164–171.
Shonkwiler, J. (2016) Variance of the truncated negative binomial distribution. *J. Econmetr.*, **195**, 209–210.
Sonnabend, H. (2020) On discouraging environments in team contests: evidence from top-level beach volleyball. *Mangrl Decsn Econ.*, to be published, doi 10.1002/mde.3153.
Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and van der Linde, A. (2002) Bayesian measures of model complexity and fit (with discussion). *J. R. Statist. Soc.* B, **64**, 583–639.
Stefani, R. (1980) Improved least squares football, basketball, and soccer predictions. *IEEE Trans. Syst. Man Cyber.*, **10**, 116–123.
Tsokos, A., Narayanan, S., Kosmidis, I., Baio, G., Cucuringu, M., Whitaker, G. and Király, F. (2019) Modeling outcomes of soccer matches. *Mach. Learn.*, **108**, 77–95.

Supporting information

Additional 'supporting information' may be found in the on-line version of this article:

'Electronic supplementary material for the paper: "A Bayesian quest for finding a unified model for predicting volleyball games"'.