**ORIGINAL PAPER**



# **Ranking handball teams from statistical strength estimation**

**Florian Felice[1](http://orcid.org/0009-0001-0222-6148)**

Received: 16 February 2024 / Accepted: 12 June 2024 © The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2024

# **Abstract**

In this work, we present a methodology to estimate the strength of handball teams. We propose the use of the Conway-Maxwell-Poisson distribution to model the number of goals scored by a team as a fexible discrete distribution which can handle situations of non equi-dispersion. From its parameters, we derive a mathematical formula to determine the strength of a team. We propose a ranking based on the estimated strengths to compare teams across diferent championships. Applied to female handball club data from European competitions over the 2022/2023 season, we show that our new proposed ranking can have an echo in real sports events and is linked to recent results from European competitions.

**Keywords** Handball · Conway-Maxwell-Poisson · Ranking · Strength estimation

# **1 Introduction**

Handball is a popular sport with growing interest across the world. To date, there does not exist any official ranking tool to compare clubs performances. The only quantitative metrics available are coefficient ranks that compare countries (based on championships) provided by the European and International Handball Federations (EHF and IHF respectively) or player's performance index available in some countries (Liqui Moly HBL [2024](#page-10-0)). Handball also sufers from a lack of literature (Saavedra [2018](#page-10-1)), especially in the predictive or analytical felds. In this work, we aim to establish a methodology to estimate the strength of teams.

Estimating the strength of a team has long been discussed in the literature, and is typically seen in sports such as football. Rating methods often assume some

 $\boxtimes$  Florian Felice forian.felice@uni.lu

Disclaimer: This work is not related to Amazon.

<sup>&</sup>lt;sup>1</sup> Department of Mathematics, University of Luxembourg, 6, avenue de la Fonte, Esch-sur-Alzette L-4364, Luxembourg

probability distributions to represent the distribution of the outcome of a match. Some methods are based on the Thurstone-Mosteller model (Thurstone [1927](#page-10-2)) or the Bradley–Terry model (Bradley and Terry, [1952\)](#page-10-3) to model the outcome of a match based on a probability distribution where the location parameter corresponds to the strengths of the modeled teams. These popular techniques, however, assume that the underlying probability distribution is continuous. This is by nature contrary to the structure of the majority of scoring-based team sports data (e.g. football, handball, basketball, rugby, etc).

Another topic of controversy is the choice of the underlying probability distribution representing the number of goals scored by a team. Reep et al.  $(1971)$  $(1971)$  demonstrated that the Negative Binomial distribution is suitable for modeling scores in several ball games. Maher ([1982\)](#page-10-5) however argued that tests for goodness-of-ft plead in favor of the independent Poisson distribution to model football scores. Ley et al. [\(2019](#page-10-6)) further investigated the idea of Poisson distributions and after a broad comparison of models, suggested the bivariate Poisson model (Karlis and Ntzoufras, [2003](#page-10-7)) to best represent the outcome of football games. From the estimated parameter *𝜆* obtained via Maximum Likelihood Estimation, they assume a structure from that parameter for team *i* with opponent team *j* as

$$
\log(\lambda_i) = \beta_0 + (r_i - r_j) + h \cdot \mathbb{1}(\text{team } i \text{ playing at home})
$$
 (1)

where  $\beta_0 \in \mathbb{R}$  is a common intercept and  $h > 0$  is the effect of playing at home. The parameters  $r_i \in \mathbb{R}$  and  $r_j \in \mathbb{R}$  represent the abilities of team *i* and *j* that are used as estimation of team strengths.

In the context of handball, Groll et al. ([2020\)](#page-10-8) analyzed historical international games to determine the best probability distribution to model the number of goals scored in handball matches. Given the level of under-dispersion observed, they concluded that the standard Poisson distribution cannot be used and a Gaussian distribution with low variance is the most appropriate.

In this article, we propose a method to derive a ranking based on handball teams strengths. These strengths are obtained using the estimated parameters of an appropriate discrete probability distribution by means of maximum likelihood. We defne formulae to transform such statistical estimates into sports abilities and observe how mathematical expressions can translate into sports facts. To illustrate our results, we apply our method to historical European female matches from the 2022/2023 season and obtain a ranking which is linked to the end of season standings.

Our work is organized as follows. In Sect. [2](#page-2-0), we compare distributions from the existing literature with the Conway-Maxwell-Poisson distribution. After motivating the use of this fexible discrete probability distribution, we will generate a metric representing the strength of a team. In Sect. [3,](#page-7-0) we will illustrate the results of the proposed methodology on female club data and propose a ranking of the best performing teams based on statistical fndings. Finally, we discuss next steps and future considerations in Sect. [4](#page-8-0) and conclude in Sect. [5](#page-9-0).

### <span id="page-2-0"></span>**2 Methodology**

In this section, we review the methodology used for modeling handball data to represent the strength of a team. First, we explain why the classical Poisson distribution cannot be used as the underlying probability distribution. We then propose the Conway-Maxwell-Poisson distribution as a fexible probability distribution from which we estimate, using its parameters, the strength of a team.

#### **2.1 Non equi‑dispersion of handball data**

When analyzing historical data from female handball matches, one can observe situations with non equi-dispersion. We defne the dispersion index *DI* as the ratio between the expectation  $E(X)$  and the variance  $V(X)$  of a random variable *X*:

$$
DI = \frac{\mathbb{E}(X)}{\mathbb{V}(X)}.\tag{2}
$$

When  $DI < 1$ , we are in the situation of over-dispersion, since the variance is larger than the expectation. When  $DI > 1$ , the variance is lower than the average which indicates under-dispersion. The final situation where  $DI = 1$  leads to equi-dispersion.

To measure such index for handball data, we analyzed games over the 2022/2023 season of the European championships and performed a statistical test to assess whether the data are equi-dispersed or not. De Oliveira ([1963\)](#page-10-9) proposed a dispersion test to compare the mean to the standard deviation of a discrete distribution. Böhning ([1994\)](#page-10-10) later proposed an update of the test, correcting the computation of the asymptotic standard deviation. Under the null hypothesis  $H_0: X \sim \mathcal{P}(\lambda)$ , one can assess whether a variable *X* follows a Poisson distribution with parameter  $\lambda$  given that the distribution fulfills the property  $E(X) = V(X) = \lambda$ . The alternative hypothesis is  $H_1$ :  $\mathbb{E}(X) \neq \mathbb{V}(X)$  highlighting non equi-dispersion. The test statistic is thus defned as

$$
T = \frac{\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2 - \bar{X}}{\sqrt{\frac{2}{n-1} \bar{X}}}
$$
(3)

where  $X_i \in \mathbb{N}$  records the number of goals scored in past matches (e.g. ongoing season). The location parameter is approximated by the empirical mean  $\bar{X} = \sum_{i=1}^{n} X_i$ . Under  $H_0$ , the test statistic *T* follows a  $\chi^2$  distribution with *n* − 1 degrees of freedom (Hoel, [1943\)](#page-10-11).

We present in Table [1](#page-3-0) the results of the tests performed over 819 European clubs (323 being female and 496 male). We could not reject  $H_0$  (at level 5%) for only 47 clubs (5.7%), all others either show over- or under-dispersion. As an illustration, the female team of Metz Handball (France) scored on average 32.25 goals over the 2022/2023 season with a variance of 21.23. Performing the test yields a test statistic

<span id="page-3-0"></span>

<span id="page-3-1"></span>of −1.53 with a *p*-value *<* 0.001, indicating under-dispersion. Therefore, aligned with conclusions from Groll et al. ([2020\)](#page-10-8), we do not determine that equi-dispersed Poisson distribution is suitable to model scored goals during handball matches.

#### <span id="page-3-2"></span>**2.2 Modelling handball games with Conway‑Maxwell‑Poisson**

As an alternative to the standard Poisson distribution, we consider the Conway-Maxwell-Poisson (CMP) distribution (Sellers, [2022](#page-10-12)). It is a generalization of the common Poisson distribution, but with the ability to handle under- and over-dispersion. Its probability mass function is defned by

$$
\mathbb{P}(X = x | \lambda, v) = \frac{\lambda^x}{(x!)^v} \frac{1}{\sum_{j=0}^{\infty} \frac{\lambda^j}{(j!)^v}}.
$$
\n(4)

The parameter  $v \ge 0$  represents the level of dispersion. When  $v = 1$ , this indicates an equi-dispersed Poisson distribution. When  $v < 1$ , we are in the situation of over-dispersion while  $v > 1$  represents under-dispersion. Though it does not have an explicit interpretation,  $\lambda \in \mathbb{R}^+_*$  can be seen as a location parameter whose value gets closer to the mean as  $v \rightarrow 1$ . Other special cases of the Conway-Maxwell-Poisson distribution include the Bernoulli distribution with parameter  $\lambda/(1 + \lambda)$  as  $\nu \to \infty$  and the geometric distribution with probability of success  $1 - \lambda$  when  $\lambda < 1$  and  $\nu = 0$ . The CMP distribution can also be a good alternative to the classical Poisson distribution given its fexibility to handle diferent levels of dispersion.

To evaluate the goodness of ft of the distribution on handball data, we compare the CMP with the Gaussian and Negative Binomial distributions as mentioned in Groll et al. ([2020\)](#page-10-8). In Table [2](#page-3-1), we report the estimated log-likelihood (*L̂* ) and the associated Akaike Information Criterion (AIC) for the club of Metz Handball over the 2022/2023 season. We observe from Table [2](#page-3-1) that, although the three distributions seem to similarly ft the data, the Conway-Maxwell-Poisson distribution exhibits the

<span id="page-4-0"></span>



<span id="page-4-1"></span>**Fig. 1** Relation between CMP parameters  $\lambda$  and  $\nu$  and the empirical mean  $\bar{X}$  (from simulated data)

maximum log-likelihood. However, the AIC aims to penalize complex distributions with numerous parameters to estimate. Given that  $k = 2$  for all three distributions, minimizing the AIC or maximizing the log-likelihood leads to the same conclusion.

Furthermore, when ftting and comparing the three distributions for 819 European male and female clubs (reported in Table [3\)](#page-4-0), the Conway-Maxwell-Poisson distribution is the most suitable in the majority of the cases (382 team out of 819, 46,6%). Although the Gaussian distribution is also appropriate for 36% of the teams, the log-likelihood for the three distributions remain very close to each other. Therefore, considering the fexibility of the distribution, which can handle under-, equiand over-dispersion situations, the Conway-Maxwell-Poisson distribution would be the most appropriate choice for modeling handball data.

We represent in Fig. [1](#page-4-1) the relation between the empirical mean from a CMP distribution and its associated parameters  $\lambda$  and  $\nu$ . We notice a logarithmic relationship between the parameter  $\lambda$  and the empirical mean. This relation will be of particular interest in the next Sect. [2.3](#page-4-2) when defning the team's strength.

### <span id="page-4-2"></span>**2.3 Estimation of team strengths**

As with most competitive sports, the strength of a team can be expressed by its ability to perform both the areas of attack and defense. We thus introduce diferent formulae to represent defense and attack strengths of a team. We then defne the overall strength of a team, creating a combination of attack and defense abilities. In this section, we will refer to a difficulty parameter  $\omega$  defined as

<span id="page-5-2"></span>
$$
\omega = \frac{1}{n} \sum_{i=1}^{n} \omega_i \in [0, 1].
$$
 (5)

This parameter corresponds to the average difficulty of the  $n \in \mathbb{N}$  matches played by a team over a fixed period of time. Each match is assigned a level of difficulty  $\omega_i \in [0, 1]$  which is a function of the competition. It is based on the European Handball Federation's (EHF) place distribution of competitiveness of leagues. Some tuning is performed to fnd a realistic value to represent the competitivity of the leagues. Our experiments suggested to scale the values, using min-max normalization, of the place distribution between 0.9 for the least competitive countries (i.e. Luxembourg or Belgium for women and Moldova for men) and 1 for the EHF Champions League. These results, validated with sport professionals, indicate that a lower bound of 0.9 for  $\omega$  allows to not devalue competitions while still being realistic about their competitivity. Values lower than 0.9 tend to severely penalize some teams and lead to unrealistic rankings, the scaling method allows to keep the amplitude of the place distribution and protect potential gaps between countries.<sup>[1](#page-5-0)</sup> The parameter serves as a penalty to highlight teams playing in more competitive championships.

#### **2.3.1 Defense strength**

Adopting the selected Conway-Maxwell-Poisson distribution, we use its parameters to represent the defensive strength of a team. The distribution of goals conceded by a team, denoted by  $Y_d$ , is assumed to follow a  $CMP(\lambda_d, v_d)$ , where the parameter  $\lambda_d > 0$  can act as a location parameter and  $v_d \ge 0$  as the dispersion parameter. We then defne the defense strength as

<span id="page-5-1"></span>
$$
s_d = \left(\frac{v_d}{\log(\lambda_d)}\right)^{\omega}.\tag{6}
$$

The strength of a team's defense is inversely proportional to the goals they concede. This is refected in Eq. ([6\)](#page-5-1) in the sense that the higher the average number of conceded goals are (i.e. the higher  $\lambda_d$ ) the lower the strength  $s_d$  will be. We notice the logarithmic transformation  $log(\lambda_d)$  to account for the relation with the empirical mean as mentioned and illustrated in Fig. [1](#page-4-1). On the other hand, we want to penalize for inconsistencies of a team, therefore we want the parameter  $v_d$  to be as large as possible corresponding to under-dispersion. The penalty  $\omega \in [0, 1]$  then makes sure that highly competitive matches are put forward in the strength. We can thus interpret formula ([6\)](#page-5-1) as follows: a team is a strong defender if it constantly concedes few goals during matches.

<span id="page-5-0"></span><sup>&</sup>lt;sup>1</sup> E.g. In female competitions, Hungary has the highest weight ( $\omega = 0.999$ ) followed by France  $(\omega = 0.992)$  and Norway ( $\omega = 0.989$ ).

#### **2.3.2 Attack strength**

We also assume that the distribution of scored goals follows a CMP distribution,  $Y_a \sim \text{CMP}(\lambda_a, \nu_a)$ . A team is considered strong in attack if the average number of scored goals is large. The logic can therefore be considered as the inverse from Eq. ([6\)](#page-5-1). We defne the attack strength of a team as

<span id="page-6-0"></span>
$$
s_a = \left(\frac{\log(\lambda_a)}{v_a}\right)^{\omega} \tag{7}
$$

where the location parameter  $\lambda_a$  is used as the numerator to show that a high number of goals scored on average increases the attack strength. The dispersion parameter  $v_a$  is used as the denominator, behaving like a penalty as we expect teams to have consistent performances over the season. We finally include the weight  $\omega \in [0, 1]$ , as defined in Eq.  $(5)$  $(5)$ , to penalize for the difficulty of the matches played.

#### **2.3.3 Global strength**

A team is considered strong when it can perform well in attack and defense. We defne the overall strength of a team as the combination of attack and defense strengths by

<span id="page-6-1"></span>
$$
s = s_a \cdot s_d = \left(\frac{\log(\lambda_a) \cdot v_d}{v_a \cdot \log(\lambda_d)}\right)^{\omega}.
$$
 (8)

We observe that a high score for overall strength can be driven by two factors. On the one hand, the team should have a high average of scored goals while demonstrating consistent defensive performances over time. On the other hand, a team should be able to adapt its attack strategies to teams and be able to take their opponent by surprise by scoring more than expected. They should also be able to prevent conceding too many goals and have a scoring capacity. In other words, the goal diference in the competition's ranking should be as large as possible. This can usually be verifed in diferent competitions where leading teams tend to have a high difference (+229 goals for Metz Handball in the French female championship at end of 2022/2023 season or +257 for Vipers Kristiansand in Norway) while teams at the bottom of the season standings have highly negative goal diferences (-107 for Toulon Métropole Var Handball in France for the same season or -170 for Volda in Norway).

We can now note the importance of the nonlinear transformation for  $\lambda_a$  and  $\lambda_d$ . Given the logarithmic rate of these parameters, a team may have to record a much higher average of scored goals to distinguish itself from other teams. First, the slope of the strength *s* with respect to the scored goals  $\lambda_a$  is

$$
\frac{\partial s}{\partial \lambda_a} = \frac{\omega \log^{2-1}(\lambda_a)}{\lambda_a} \left( \frac{v_d}{v_a \log(\lambda_d)} \right)^{\omega} > 0.
$$
\n(9)

As the team gets stronger,  $\lambda_a$  increases (everything else being equal) and differentiators with other teams become marginal since  $\lim_{\lambda_a \to \infty} (\frac{\partial s}{\partial \lambda_a}) = 0$ .



<span id="page-7-1"></span>**Fig. 2** Histogram of goals scored by Metz Handball over season 2022/2023 vs. theoretical CMP distribution estimated via Maximum Likelihood

Second, the derivative of the strength *s* with respect to conceded goals  $\lambda_d$  is

$$
\frac{\partial s}{\partial \lambda_d} = -\frac{\omega}{\lambda_d \log^{\omega+1}(\lambda_d)} \left( \frac{\log(\lambda_d) v_d}{v_a} \right)^{\omega} < 0.
$$
 (10)

This suggests that any reduction of conceded goals leads to improvements in the overall strength. These statements have an actual echo in sports terms. It is common knowledge for handball players and coaches that the best way to improve a team's performance is to start by improving their defense.

### <span id="page-7-0"></span>**3 Illustrative applications**

As illustrated in Sect. [2.2,](#page-3-2) the CMP distribution seems to be the most appropriate choice to model goals scored during a handball match. We plot in Fig. [2](#page-7-1) the histogram of scored goals over the 2022/2023 season for the female club of Metz Handball (France) and compare with the ftted theoretical CMP distribution. The ftted distribution smooths the empirical histogram which highlights two distinct scoring regimes. The situation when  $x \leq 32$  mostly corresponds to Champions League matches with aggressive defensive teams (average scored goals  $=$  30.4). On the other hand, when  $x > 32$  we mostly have matches from the French championship with teams being substantially weaker (average scored goals  $=$  33.6). To account for these diferences, we will need to use a penalty for the competitiveness as introduced in Sect. [2.3](#page-4-2).

Furthermore, we estimate the strength parameters for all European female clubs and display the ranking in Table [4.](#page-8-1) The estimations are derived from all matches over the 2022/2023 season that were played in female frst division competitions (ranging from friendly games to the regular championships and Champions League).

Team	Avg. scored	Avg. conceded	$\omega$	$S_a$	$S_d$	S
MKS Zaglebie Lubin	28.93	23.41	91.10%	3.13	3.53	11.06
Győri Audi ETO KC	33.71	23.70	99.93%	3.48	3.13	10.92
<b>Vipers Kristiansand</b>	33.88	25.51	98.95%	3.52	3.09	10.89
Team Esbjerg	31.01	25.54	98.82%	3.44	3.15	10.85
Metz handball	31.26	24.79	99.21%	3.43	3.15	10.81
FTC Rail-Cargo Hungaria	31.11	24.97	99.93%	3.44	3.11	10.73
<b>SG BBM Bietigheim</b>	32.43	25.74	95.20%	3.35	3.19	10.71
<b>IK Sävehof</b>	29.17	24.97	90.00%	3.03	3.52	10.69
<b>CSM</b> Bucuresti	29.81	24.57	97.64%	3.37	3.14	10.56
<b>DVSC Schaeffler</b>	29.72	25.32	99.92%	3.37	3.13	10.55

<span id="page-8-1"></span>**Table 4** Top 10 strongest female teams in Europe for the 2022/2023 season

We can observe from the top clubs that Győri Audi ETO KC, Vipers Kristiansand, Team Esbjerg and FTC Rail-Cargo Hungaria were the participants of the EHF fnal four in June 2023, representing the fnal stage of the toughest European competition. Vipers Kristiansand fnally won the competition for the third year in a row. All other clubs are leading their championships in their respective countries and were members of the EHF Champions League in season  $2023/2024$ . As handball does not have any official ranking, we asked feedback from sports professionals as the only way to validate our results. They could confrm that our results make sense and are in line with the performance of European teams.

We also notice in Table [4](#page-8-1) that, even though the ranking is sorted by the overall estimated strength, the average number of scored and conceded goals seem to follow some sort of hierarchy. To be ranked toward the top, teams have to show a high average of scored goals and a relatively lower number of conceded goals. The exceptions (e.g. MKS Zaglebie Lubin) fnd a justifcation in the consistency of their performance. Indeed, such teams exhibit a lower value for  $v_a$  or higher value for  $v_d$  suggesting consistencies in attack or defense and boosting their final strength ranking. This justifies the requirement for the use of formulae  $(6)$  $(6)$  and ([7\)](#page-6-0) instead of purely relying on average scored goals.

# <span id="page-8-0"></span>**4 Discussion**

Our proposal ofers an estimation of attack and defense strengths to rank teams and generate features that can be informative and meaningful in subsequent modelling tasks. Provided that one has access to such data, the presented exercise can be extended to other objectives such as estimation of player abilities or be generalized to other sports.

# **4.1 From team strengths to player abilities**

Using more granular data (not publicly available) on player performances for each game and over several seasons, one can also estimate the attack strength of a player. Considering that the data will most likely also suffer from under- or overdispersion, the CMP distribution seems to be a good choice to ft the number of scored goals by a player. Using formula ([7](#page-6-0)), we can therefore estimate the attack strength of an individual player. Not focusing only on goals scored, playing ability could also include components such as passing ability and combine scoring and passing abilities as a global attack strength. Accessing data such as interception, successful blocks (e.g. faults with no penalty such as yellow card, 2 min penalty, etc.), the defense ability can be modeled in a similar fashion so that one can derive a defensive ability at player level.

Therefore, combining attack and defense abilities as defned by Eq. [\(8\)](#page-6-1), one can estimate the individual abilities and derive a ranking. Such ranking can help subsequent modelling exercises by adding informative variables regarding the strength of the individual players and not only the global strength of a team. Additionally, the individual ranking can be used as a new source of information for team managers to assess the potential abilities of a player when recruiting. Indeed, one can obtain a time dependent ranking and observe the evolution of a player over several seasons. This can further lead to forecasting exercises to identify players with high potential to be added to the squad.

# **4.2 Generalization to other sports from Conway‑Maxwell‑Poisson distribution**

Modelling sports requires the reliance on discrete distributions though the issue of over- or under-dispersion is a recurrent problem (Karlis and Ntzoufras, [2008;](#page-10-13) Van Bommel et al., [2021](#page-11-0)). Given the similar constraints as we have seen in the present work, one can replicate the discussed logic on other sports' data. The methodology from Ley et al. ([2019](#page-10-6)) can be merged with our proposed methodology to obtain football team abilities based on a distribution that can handle the problem of under-dispersion. One can thus defne new rankings and generate new informative features to include in predictive Machine Learning models. Using a similar methodology as Groll et al. ([2019](#page-10-14)), one can include such generated features in the feature set to improve the predictive model.

# <span id="page-9-0"></span>**5 Conclusion**

Handball is a fast-paced sport of which goals cannot be analyzed via standard count distributions due to the problem of under- or over-dispersion. We showed that, using an appropriate probability distribution, one can defne meaningful statistical estimates that approximate the strengths of a team. The choice of the Conway-Maxwell-Poisson distribution can therefore be a suitable option to model not only handball games but any scoring-based team sports.

The proposed methodology allows the generation of very informative context about the performance of a team. In a future work, this will allow to generate new covariates that can be included in predictive models in the spirit of Groll et al.  $(2019)$  $(2019)$ . They can also offer the possibility to consider data-driven analyses of a team's performance to later support team managers in their personal strategies and tactical motivations. With access to more granular data, this methodology can be adapted to the estimation of player abilities and offer tools to allow coaches take data-driven decisions in their recruitment processes.

**Supplementary Information** The online version contains supplementary material available at [https://doi.](https://doi.org/10.1007/s00180-024-01522-0) [org/10.1007/s00180-024-01522-0](https://doi.org/10.1007/s00180-024-01522-0).

**Acknowledgements** The author would like to thank Emmanuel Mayonnade, Head Coach of Metz Handball (France) for his insightful feedback to validate and help improve the methodology.

# **References**

- <span id="page-10-10"></span>Böhning D (1994) A note on a test for Poisson overdispersion. Biometrika 81(2):418–419. [https://doi.org/](https://doi.org/10.1093/biomet/81.2.418) [10.1093/biomet/81.2.418](https://doi.org/10.1093/biomet/81.2.418)
- <span id="page-10-3"></span>Bradley RA, Terry ME (1952) Rank analysis of incomplete block designs: I. The method of paired comparisons. Biometrika 39(3/4):324. <https://doi.org/10.2307/2334029>. (**Accessed 2023-07-01**)
- <span id="page-10-9"></span>De Oliveira T (1963) Some Elementary Tests for Mixtures of Discrete Distribution. Columbia University, New York
- <span id="page-10-14"></span>Groll A, Ley C, Schauberger G, Van Eetvelde H (2019) A hybrid random forest to predict soccer matches in international tournaments. J Quant Anal Sports 15(4):271–287. [https://doi.org/10.1515/](https://doi.org/10.1515/jqas-2018-0060) [jqas-2018-0060](https://doi.org/10.1515/jqas-2018-0060)
- <span id="page-10-8"></span>Groll A, Heiner J, Schauberger G, Uhrmeister J (2020) Prediction of the 2019 IHF World men's handball championship - a sparse gaussian approximation model. J Sports Anal 6(3):187–197. [https://doi.org/](https://doi.org/10.3233/JSA-200384) [10.3233/JSA-200384](https://doi.org/10.3233/JSA-200384)
- <span id="page-10-11"></span>Hoel PG (1943) On indices of dispersion. Ann Math Stat 14(2):155–162. [https://doi.org/10.1214/aoms/](https://doi.org/10.1214/aoms/1177731457) [1177731457](https://doi.org/10.1214/aoms/1177731457)
- <span id="page-10-7"></span>Karlis D, Ntzoufras I (2003) Analysis of sports data by using bivariate Poisson models. J Royal Stat Soc: Series D (The Statistician) 52(3):381–393.<https://doi.org/10.1111/1467-9884.00366>
- <span id="page-10-13"></span>Karlis D, Ntzoufras I (2008) Bayesian modelling of football outcomes: using the Skellam's distribution for the goal diference. IMA J Manag Math 20(2):133–145. <https://doi.org/10.1093/imaman/dpn026>
- <span id="page-10-6"></span>Ley C, Wiele T, Van Eetvelde H (2019) Ranking soccer teams on the basis of their current strength: a comparison of maximum likelihood approaches. Stat Model 19(1):55–73. [https://doi.org/10.1177/](https://doi.org/10.1177/1471082X18817650) [1471082X18817650](https://doi.org/10.1177/1471082X18817650)
- <span id="page-10-0"></span>Liqui Moly HBL - Handball Performance Index: season overview. [https://www.liquimoly-hbl.de/en/s/](https://www.liquimoly-hbl.de/en/s/handball-performance-index/widgets/liqui-moly-hbl-en/) [handball-performance-index/widgets/liqui-moly-hbl-en/](https://www.liquimoly-hbl.de/en/s/handball-performance-index/widgets/liqui-moly-hbl-en/) Accessed 2024-04-20
- <span id="page-10-5"></span>Maher MJ (1982) Modelling association football scores. Stat Neerl 36(3):109–118. [https://doi.org/10.](https://doi.org/10.1111/j.1467-9574.1982.tb00782.x) [1111/j.1467-9574.1982.tb00782.x](https://doi.org/10.1111/j.1467-9574.1982.tb00782.x)
- <span id="page-10-4"></span>Reep C, Pollard R, Benjamin B (1971) Skill and chance in ball games. J Royal Stat Soc. Series A (General) 134(4):623.<https://doi.org/10.2307/2343657>
- <span id="page-10-1"></span>Saavedra JM (2018) Handball research: state of the art. J Hum Kinet 63(1):5–8. [https://doi.org/10.2478/](https://doi.org/10.2478/hukin-2018-0001) [hukin-2018-0001](https://doi.org/10.2478/hukin-2018-0001)
- <span id="page-10-12"></span>Sellers KF (2022) The Conway-Maxwell-Poisson Distribution. Institute of Mathematical Statistics monographs. Cambridge University Press, Cambridge
- <span id="page-10-2"></span>Thurstone LL (1927) Psychophysical analysis. Am J Psychol 38(3):368. <https://doi.org/10.2307/1415006>

<span id="page-11-0"></span>Van Bommel M, Bornn L, Chow-White P, Gao C (2021) Home sweet home: quantifying home court advantages for NCAA basketball statistics. J Sports Anal 7(1):25–36. [https://doi.org/10.3233/](https://doi.org/10.3233/JSA-200450) [JSA-200450](https://doi.org/10.3233/JSA-200450)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional afliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.