

バレーボール各国代表チームのレーティング手法 の提案および結果予測・大会形式評価への応用

小中 英嗣[†]

(受付 2016 年 12 月 20 日；改訂 2017 年 5 月 2 日；採択 8 月 28 日)

要 旨

バレーボールの各国代表チームのランキングである FIVB ランキングは、ランキングポイントの設計に統計的な根拠が無く、さらに、主要な国際大会を独占的に主催する特定の国が優遇されていることなどが原因となり、その値を実力の定量的評価には活用できない。そこで本研究では、過去の試合結果(各セットでの得点)から各チームの実力(レーティング)を算出する手法を提案する。具体的には、レーティング差がロジスティック回帰モデルを通し各セットでの得点率を説明するモデルを仮定する。また、試合結果からレーティングを算出するアルゴリズムをあわせて提案する。

リオデジャネイロオリンピックを含む複数の国際大会での実際の得点率と、提案手法により算出したレーティング差から予測される得点率との間に中程度から強い相関(相関係数がおおよそ 0.65 から 0.80)があり、その相関はおおむね FIVB ランキング(ポイント)差に基づくもの(相関係数がおおよそ 0.45 から 0.70)より強いことが分かった。さらに、提案手法に基づくレーティングにより各大陸間の実力差を定量的に明らかにし、FIVB ランキングが実力を正しく反映していないことの根拠を明示する。この事実に基づき、具体的な事例としてリオデジャネイロオリンピックの予選および本選の大会形式の不備を指摘する。具体的には、敗退行為を誘引しかねない最終予選出場国の決定方法、および実力を反映していない FIVB ランキングにより引き起こされたオリンピック本選での不均衡な組み分けについて指摘する。

キーワード：スポーツ，バレーボール，レーティング，ロジスティックモデル。

1. 研究背景

本稿ではバレーボール各国代表チームのレーティング手法の提案とその検証を目的とする。

スポーツにおいて、試合結果に基づき選手・チーム間の序列を定めるランキング制度は、大会への参加基準やシード対象選手(チーム)の選定基準となるなど、その適切な設計が要求される事項である。

全ての選手(チーム)が等しい回数総当たりする(できる)場合は、勝利数や勝率がランキングに最も適した指標であることは論を待たない。しかし、多くの選手(チーム)が参加し、総当たりが不可能な場合に適切に順序を定めることには困難を伴う。

多くのスポーツでは、過去の一定期間の各大会での成績(順位)を数値化し、その総計(ランキングポイント)に基づいたランキングを採用している(Ray, 2011)。ここではこれを加算方式

[†]名城大学 理工学部：〒468-8502 愛知県名古屋市天白区塩釜口 1-501

と呼ぶ。バレーボールの各国代表チームのランキングも加算方式により算出されている(FIVB, 2016)。この手法では、各チームが自分自身の結果のみと得点算出基準を参照することでランキングポイントを算出でき、この簡便さは利点である。しかし、各選手・チーム間の合計ポイントの差や比が実力差を定量的に表すためには、各大会・各順位に対する付与ポイント、大会参加条件、大会形式(シード権の設定など)を含めた総合的な設計が不可欠である。

一部のスポーツでは、ポイント交換(point exchange)と呼ばれる、各対戦結果に基づき対戦チーム間でランキングポイントを総和が変わらないように交換する手法に基づきランキングポイントを算出している。最も有名な例はチェスのイロレーティング(Elo rating)(Elo, 1979)であり、球技ではラグビーがイロレーティングを若干修正した手法を採用している(World Rugby, 2014)。この手法では、各チームの実力がある期間一定であり、その期間内に十分な対戦数が行われれば、算出されたランキングポイントが実力に相当する値に近づく特徴を持っている。ただし、この算出には対戦時点での両チームのランキングポイントが既知である必要があり、原理的には期間開始時点での全チームのランキングポイントおよび期間内の全対戦の結果を必要とする。また、得点差などによるポイント交換量の調整などを含んでいるスポーツも多く、その量の調整によりランキングが適切に実力を反映できているかどうかは明確でなくなる点は加算方式と同様である。

ここで、ランキング(ranking)とレーティング(rating)を定義し、区別することとする。ランキングは各チーム間の順位を定めることであり、それに対しレーティングは各チームの実力に相当する値を算出することと定義する。レーティングの結果に基づいてランキングを作成することが可能である。

2つのチームが対戦した場合、大きく二つに分けて、

- (1)それぞれのチームが数年程度のある一定の期間にわたり安定的に持っている技能・実力
- (2)当日の調子・運、対戦チーム間の相性

の要因が得点や勝敗に影響すると仮定する。本稿では、バレーボールなどの2チームが互いの得点の多少を競うスポーツに対する狭義のレーティングとして

- ・上記の(1)に対応する、過去の試合結果から推定した、あるチームが得点する確率を説明するパラメータ

と定義する。加算方式のランキングポイントは上記の(1)および(2)の双方を区別せずに反映させる方法であり、これが(1)のみを反映させるようにするためには大会の参加条件、トーナメント形式、および順位ごとに付与されるポイントなど統一的な設計が必要である。それに対してイロレーティングに代表されるポイント交換手法は(2)の影響を排除して(1)のみを推定する、つまり狭義のレーティングを統計的に算出することを意図した手法である。以降本稿で「レーティング」は「狭義のレーティング」を指すこととする。

バレーボールに関する統計的な研究として、各技能(サーブ、レセプション、スパイク、ディグ、ブロック、およびセットなど)や戦略が得点にどのように寄与しているかを調査したものは多数ある(たとえば、Eom and Schutz, 1992; Zetou et al., 2007; Florence et al., 2008; Araújo et al., 2010; Ferrante and Fonseca, 2014; Burton and Powers, 2015など)。しかし、これらの研究には試合動画を利用した各プレーの詳細な分析が必要であり、その調査コストは無視できない。また、状況を細かく分析しすぎることは説明変数を増やしてその次元を上げてしまうが、活用できるデータ数には上限がある。このようにして「次元の呪い」(Indyk and Motwani, 1998)を引き起こしかねないため、推定するパラメータ数が試合数に対して十分少なくなるようなモデルの構築が肝要である。

これら技能の評価に関する研究が非常に多岐にわたるのに対し、バレーボールのチームのレーティングに関する研究は(著者の調査の範囲では)非常に少ない。Masseyの方法(Massey, 1997)に基づき、大学女子バレーボールの勝率を予測したもの(Knapper and McIlwain, 2015)や、ビーチバレーに適用したもの(Glasson et al., 2001)、公式ランキング差をもとにロジスティック重回帰分析を適用したもの(Dziedzic and Hunter, 2015)などが散見されるのみである。

バレーボールはサービスから始まる一連のプレイが必ずどちらかの得点で終わるという特徴を有する競技である。また、サービスから始まる一連のプレイを1試合中に150回から200回繰り返す。得点は試合進行の最小単位、かつ(総得点が多いほうが勝利するとは限らないが)勝敗に対する主要な原因であり、さらに公式記録に必ず記載されて調査が容易である。この観点から、本研究では各試合、各セットの得点を基本データとし、これに基づく統計的な分析を行う。

本研究では、各チームが得点する確率が、それぞれのチームのレーティングの差により説明できると仮定し、国際試合の結果からその統計的な算出を試みる。得点率の説明モデルにはロジスティック回帰モデルを利用する。対象となる全試合に対する最適化を行うことで、最適なレーティングを算出する。これは試合ごとの逐次更新に基づくポイント交換(イロレーティングなど)の欠点を補う明確な利点となる。

提案手法により算出したレーティングに基づき、いくつかの国際大会の結果を予測する。具体的には、ロンドンオリンピック、リオデジャネイロオリンピック世界最終予選1、およびリオデジャネイロオリンピックの各試合の結果および得点率を、その直近1年の主要国際大会の結果から算出したレーティングに基づき予測する。

その結果、

- レーティングから予測された得点率と各セットの実際の得点率には中程度の相関があり(相関係数が0.5強)、
- レーティングから予測された得点率と各試合の実際の得点率には中程度から強い相関がある(相関係数が0.7付近)

ことを示す。

これらの相関、および得点率の大小の予測結果をFIVBランキングおよびランキングポイントに基づく手法と比較し、提案するレーティングが優位であることを示す。さらに、定量的なレーティングに基づく大会形式の評価の応用として

- 世界最終予選が敗退行為を助長しかねない不適切な設計となっていたこと
- FIVBランキングが実力を反映していないことが原因でリオデジャネイロオリンピックでのプールわけに不均衡が生じ、実力の高いチームが早期敗退を強いられたこと

の2点を指摘する。

2. レーティングの定義と算出方法

本節ではまず対象とするバレーボールの世界ランキングの特性について議論し、その後レーティング算出アルゴリズムを提案する。

2.1 FIVB ランキング

国際バレーボール連盟(FIVB)はオリンピック、世界選手権などの主要国際大会の順位をランキングポイントに換算し、一定期間の和をそのチームのランキングポイントとし、その値で

表 1. FIVB Ranking Point System.

Standing	Competition name				
	Olympic	World Cup	World Championship		Continental Championship
			Men	Women	
1	100	100	100	100	30
2	90	90	90	90	26
3	80	80	80	80	22
4	70	70	70	70	18
5	50	50	62	58	10
6	—	40	56	—	14
7	—	30	50	50	10
8	—	25	—	—	5
9	30	5	45	45	—
10	—	5	—	—	3
11	20	5	40	40	—
12	—	5	—	—	—
13 Tie			36	36	2
15 Tie			33	33	—
17 Tie			30	30	—
21 Tie			25	25	—

世界ランキングを定めている (FIVB, 2016)。

順位に対しどの程度のランキングポイントを付与するか設計により、どのようなチームがランキング上位となるかが異なる。表 1 に、主要国際大会で得られるランキングポイントを示す。

順位決定にトーナメント形式を採用している大会(オリンピック、世界選手権)では順位決定戦を行わない場合があり(たとえば、オリンピックでは準々決勝敗退の 4 チーム間での順位決定戦を行わない)、その場合は同順位として同じランキングポイントが付与される。

この表から、世界大会での 1 位から 4 位までの順位ひとつの差は 10 点で統一されているが、5 位以下でのランキングポイントの設計法には一貫性がないことが分かる。具体的には、オリンピック、ワールドカップでは 4 位と 5 位の差が 20 ポイントであるのに対し、世界選手権では 8(男子)、12(女子)ポイントである。それ以下の順位となると順位間でのポイント差はまた縮まる傾向にある。大陸選手権は 5 大陸(アジア、アフリカ、ヨーロッパ、北中アメリカ、南アメリカ)それぞれで 2 年ごとに開催されるが、大陸間の実力差はポイントには反映されていない。また、特定の大会を継続して同一国で開催し、その開催国が常にランキングポイントを獲得している、などの問題点も指摘できる(第 4 章で詳述する)。

これに対し、順位による加算方式を採用している中でも設計に一貫性のあるテニス(ATP ランキングポイント)(ATP World Tour, 2016)と比較する。一例としてグランドスラム(主要四大大会)でのランキングポイントを示す。

優勝：2000、準優勝：1200、ベスト 4：720、ベスト 8：360、ベスト 16：180、
 ベスト 32：90、ベスト 64：45、ベスト 128：10、予選通過：25、予選決勝：16、
 予選 2 回戦：8

ベスト 64 からベスト 4 までは 1 回勝ち抜くごとに付与ポイントが 2 倍、それ以降は 5/3 倍となっている。グランドスラム以下の大会(マスターズ 1000、500 シリーズ、250 シリーズ)では優勝者に付与されるランキングポイントは大会名に冠されている値(1000、500、250)であり

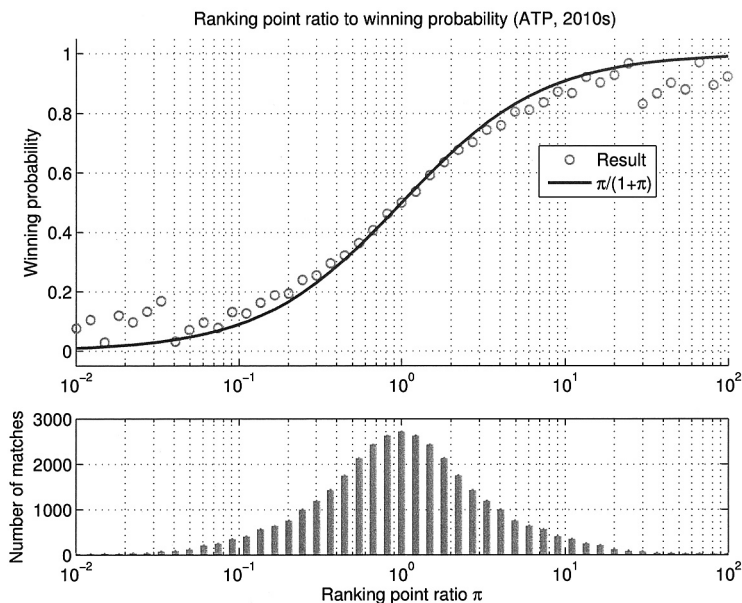


図 1. ATP ranking point ratio and winning probability.

それらは均等な比率を保っている。また優勝と準優勝以降の比率も保持されている。テニスのこのランキングポイントシステムは、どこまで勝ち抜いたかを二進数で表現していることとほぼ等価である。

このような一貫した付与ポイントの設計に加え、上位選手に対して主要大会への出場義務があり、ランキングによりシードを決定するトーナメント方式が安定して運営されている。したがって、その選手がどの規模の大会でどの程度まで勝ち抜いているかと合計ランキングポイントには強い関係ができ上がる。その結果、試合直前の対戦選手間のランキングポイント比が予測される勝率を説明することになる。図 1 に、試合直前の ATP ランキングポイント比と試合結果(勝率)の関係を示す。試合結果は 2010 年 1 月から 2016 年 1 月までの 18197 試合のもの (Sackmann, 2017) を利用した。ランキングポイント比を π としたとき、勝率が $\pi/(1+\pi)$ で近似できることが示されている。

2.2 レーティング算出方法

前述の通り、FIVB ランキングポイントの設計には数理的根拠が無く、明確な意図を持って設計されている ATP ランキングポイントと比較してもそれは明らかである。したがって、FIVB ランキングポイントによりそのチームの実力を推定し、さらに今後の対戦の勝敗予測などに利用するのは困難である(第 3 章および第 4 章で具体例を示す)。

そこで本稿では、試合結果として必ず公式に記載される各セット毎の得点に基づき、各チームの得点率を説明するパラメータを推定する手法を提案する。具体的には、各チーム i の実力を示すパラメータ(今後、これをチーム i のレーティングと呼ぶ)を r_i で示し、チーム i と j が対戦した場合、ひとつのプレイで i が得点する確率 $p_{i,j}$ が次式に従うと仮定する。

$$(2.1) \quad p_{i,j} = \frac{1}{1 + e^{-(r_i - r_j)}}.$$

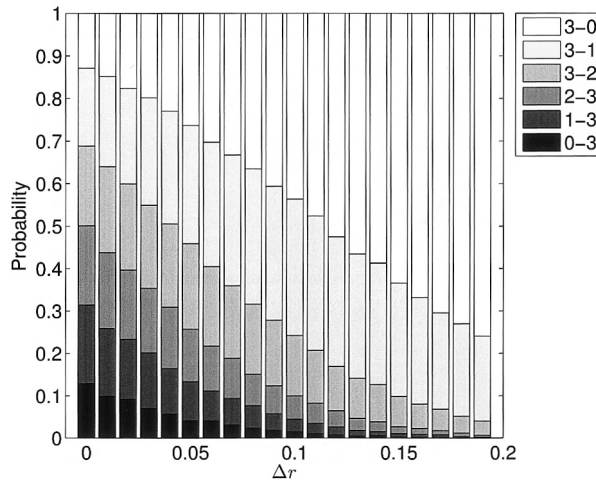


図 2. Won-lost sets probability.

この仮定はロジスティック回帰モデルと呼ばれ、チェスのレーティングであるイロレーティング (Elo, 1979) での勝率の仮定や、教育学における項目応答理論 (Hambleton, 1991; de Ayala, 2008) での問題の正答率の仮定など、幅広い分野で利用されているものと同じ形式である。大まかな目安として、パラメータ差 0.1 は、パラメータが大きいチームの 1 プレイでの得点率が 0.525 であることに対応する。この場合、パラメータの大きいチームが 1 セット取る (25 点得点する) までに小さいチームは平均で 22.6 点得点していることになる。

2 チームのパラメータ差 $\Delta r = r_i - r_j$ を 0, 0.01, 0.02, ..., 0.20 と選び、各プレイ間での得点確率が一定で独立であると仮定し、それぞれ 10^4 回の対戦をシミュレーションした結果を図 2 に示す。実力差が小さいところではセット数 3-2, 2-3 が多く、 $\Delta r = 0.05$ では実力が高いチームが 2-3 で敗れる割合が約 12.4% である。 Δr が 0.1 を超えると番狂わせ (レーティングが低いチームが勝利する) の確率は 5% 程度となり、0.2 を超えるとほぼ勝敗は明確であることが示唆されている。

2.2.1 ポイント交換法の特徴

イロレーティング (Elo, 1979) は、式 (2.1) の線形変換である

$$(2.2) \quad p_{i,j} = \frac{1}{1 + 10^{\frac{-(r_i - r_j)}{400}}}$$

により選手 i, j 間の勝率が説明できると仮定している。したがって、レーティングが正しく計算されていれば、その差を利用して試合結果を予測することが可能な方法である。ラグビーでは上式を折れ線で近似したものを利用している (World Rugby, 2014)。このようにしてランキングポイント差を実力差と関連付けているため、「日本初戦が番狂わせランキングトップ」 (World Rugby, 2015) のように、ある試合の結果がどの程度の番狂わせであったのかを定量的に評価することが可能となっている。

また、選手 (チーム) i, j 間の試合結果が

$$(2.3) \quad s_{i,j} = \begin{cases} 1 & (i \text{ wins}), \\ 0 & (i \text{ loses}). \end{cases}$$

で表されるとしたとき、試合後のレーティングが

$$(2.4) \quad r'_i = r_i + K(s_{i,j} - p_{i,j})$$

と更新される。 r_j の更新規則も同様である。式(2.4)は以下のように解釈できる：勝利したときはレーティングが増加する。これは、「勝利したのだから試合前のレーティングは過小評価であり、レーティングはより高く修正されるべきである」ということに他ならない。増加量は予測される勝率が高いほど小さい。これは「試合前のレーティングから予想される結果が起きているので、レーティングを大きく修正する必要は無い」ということである。敗北したときはこの逆である。

このような方式であるため、レーティングの計算には試合前の両チームのレーティングが必要であり、1チームのレーティングを計算するために、関与する全員(全チーム)の対戦結果が必要である。

式(2.4)の定数 K はレーティングの更新量を定める定数であり、チェスのプロレベルでは16が利用されている。 K が大きければ直近の対戦の影響が大きく、逆に小さければ過去の影響が大きい。この値の設定により、レーティングの性質が異なる。

r_i の真値が既知かつ一定であり、対戦が十分な回数実施されれば、式(2.4)の更新規則により r_i は真値に近づく。しかし、同程度の実力である二者が対戦した場合でも、対戦後に必ずレーティングが変動してしまうため、収束は保障されない。

2.2.2 提案手法

前述の、試合ごとのポイント交換手法の欠点は、1試合ごとの結果に対するレーティングの更新量(式(2.4)内のパラメータ K により定まる)の適切な設定が困難であること、およびレーティングの更新が試合ごとであり、レーティングが収束に至るまでには更新回数が少ないこと、の二点にまとめられる。本稿ではこれらの欠点を克服する手法を提案する。

提案手法の概要を以下に述べる。

- (1) 対象とする期間の全試合のデータをランダムにソートした後、順に選択し、その得点率に対して式(2.4)でレーティングを更新する。
- (2) 試合ごとの結果に対するレーティングの更新量(式(2.4)内の K)を動的に調整する。
レーティングの更新量が直前よりも大きくなることは K が大きすぎることを意味するので、その値を小さくする。
- (3) 上記の更新を、全チームのレーティングが収束するまで十分な回数繰り返す。

以下にアルゴリズムを示す。表2に記号の意味を示す。

レーティング算出アルゴリズム。

- (1) $\mathbf{r}^{(0)} = \mathbf{0}$ とし、収束判定のための $\epsilon_{th} > 0$ を小さな値に設定する。 $K > 0$ を定める。繰り返し回数を $k = 0$ とする。 N_S セットの結果を保持したデータベースを用意する。各セットの結果を $\langle i, j, s_i, s_j \rangle$ で示す。
- (2) データベースの各セットをランダムにソートする。
- (3) ソートされたデータベースからセットを順に取り出し、これを $\langle i, j, s_i, s_j \rangle$ とする。
- (4) レーティング r_i および r_j を以下の規則で更新する。

$$(2.5) \quad p_{i,j} = \frac{1}{1 + e^{-(r_i^{(k)} - r_j^{(k)})}}, \quad s_{i,j} = \frac{s_i}{s_i + s_j},$$

$$(2.6) \quad r_i^{(k+1)} = r_i^{(k)} + K(s_{i,j} - p_{i,j}), \quad r_j^{(k+1)} = r_j^{(k)} + K((1 - s_{i,j}) - (1 - p_{i,j})).$$

表 2. Notations.

N_T	Number of teams
$\mathbf{r} = (r_1, \dots, r_{N_T})^T$	Rating vector
N_S	Number of sets
$\langle i, j, s_i, s_j \rangle$	Result of one set. Team i and j scored s_i and s_j points in a set. N_S tuples are stored in database.
ϵ_{th}	Threshold value
K	Parameter used in rating update
k	Iteration index
$\mathbf{0}, \mathbf{1}$	Column vector composed of zeros and ones with suitable dimensions
$\ \mathbf{x}\ $	Euclidean norm of vector x

データベース内の全セットに対して上記の更新を行う。

- (5) $\|\mathbf{r}^{(k+1)} - \mathbf{r}^{(k)}\| > \|\mathbf{r}^{(k)} - \mathbf{r}^{(k-1)}\|$ の場合, $K \leftarrow 0.5K$ とする。
 (6) $\|\mathbf{r}^{(k+1)} - \mathbf{r}^{(k)}\| < \epsilon_{th}$ が成り立つ場合, $\mathbf{r}^{(k+1)}$ を出力としてアルゴリズムを終了する。そうでない場合 $k \leftarrow k + 1$ としてステップ(2)へ戻る。

このレーティングは 2 チーム間の差のみに意味がある(間隔尺度である)ため, 分かりやすさのために全体に定数を加減算してもかまわない。提案手法ではレーティングの和が保存されるので, $\mathbf{r}^{(0)} = \mathbf{0}$ から計算を開始すると $r_i = 0$ は全チーム内で中央の実力を持つことを意味する。また,

$$(2.7) \quad \mathbf{r} \leftarrow \mathbf{r} - (\max \mathbf{r}) \cdot \mathbf{1}$$

とすれば, レーティングの最上位を必ず 0 とし, そこからの差を表すことができる。

2.3 他の手法との関係および提案手法の拡張性

対の比較結果から全体の順位付けを行う他の手法として, Bradley-Terry モデル(BT モデル)(Rao and Kupper, 1967; Davidson, 1970)がよく知られており, スポーツのレーティングに関する研究も多い。BT モデルでは, チーム i とチーム j が対戦した場合, i が j に勝利する確率 $p_{i,j}$ が各チームのパラメータ $\pi_i, \pi_j > 0$ を用いて

$$(2.8) \quad p_{i,j} = \frac{\pi_i}{\pi_i + \pi_j}$$

で表すことができると仮定する。この仮定は,

$$(2.9) \quad p_{i,j} = \frac{\pi_i}{\pi_i + \pi_j} = \frac{1}{1 + \frac{\pi_j}{\pi_i}} = \frac{1}{1 + \exp(\log \frac{\pi_j}{\pi_i})} = \frac{1}{1 + \exp(\log \pi_j - \log \pi_i)}$$

と変形できるため,

$$(2.10) \quad r_i = \log \pi_i$$

の非線形変換により提案手法(式(2.1))と等価である。

これら二つの手法は等価であるが, 文献(Ray, 2011)の調査では, BT モデルに基づくレーティング手法を採用している競技は無く, それに対し提案手法の線形変換であるイロレーティングはいくつかの競技団体で採用されている。

また, 項目応答理論(Hambleton, 1991; de Ayala, 2008)では, 項目弁別度 a_i , 項目難易度 b_i , 個人能力パラメータ θ_j を用いて, 個人 j が項目 i に正解する確率 $p_{i,j}$ を

$$(2.11) \quad p_{i,j} = \frac{1}{1 + \exp(-a_i(\theta_j - b_i))}$$

により説明するモデルが採用されることがあり、そのパラメータ推定手法なども広く研究されている。

これをスポーツに拡張すると、大会 k でチーム i とチーム j が対戦した場合、 i が j に勝利する確率 $p_{i,j}$ が

$$(2.12) \quad p_{i,j} = \frac{1}{1 + \exp(-a_k(r_i - r_j))}$$

で説明されると仮定するモデルも構築することができる。 $a_k > 0$ は大会ごとの番狂わせの起こりやすさ・起こりにくさを意味するパラメータと解釈できる。

これらの、「既存競技での採用状況」、および「モデルの拡張性とその解釈」の観点から、本研究では(提案手法と等価ではあるが)BTモデルを採用せず、提案手法(式(2.1))を採用する。

3. レーティング算出および試合結果予測への適用

本章では、前章で提案したレーティング算出手法を実装し、各国代表チームのレーティングを算出する。

3.1 使用したデータ

以下のデータを使用した。オリンピック本選2大会(ロンドン, リオデジャネイロ)およびオリンピック世界最終予選, 男女計6大会を予測対象とし、それぞれの直近1年以内を目安とし、FIVB主催大会および各大陸のオリンピック予選の結果に基づきレーティングを算出した。このことにより、各チームが1年程度の期間を通じて安定的に保持している技能・実力をレーティングとして抽出できる。

データセット W1

- 2011年ワールドカップバレーボール(女子)[2011年11月]
- ロンドンオリンピック ヨーロッパ予選(女子)[2012年5月], 南アメリカ予選(女子)[2012年5月], 北アメリカ予選(女子)[2012年5月], アフリカ予選(女子)[2012年2月], 世界最終予選(女子)[2012年5月]
- 2012年ワールドグランプリ[2012年6月-7月]

予測対象: ロンドンオリンピック(女子)[2012年7月-8月]

データセット W2

- 2015年ワールドカップバレーボール(女子)[2015年8月-9月]
- リオデジャネイロオリンピック ヨーロッパ予選(女子)[2016年1月], 南アメリカ予選(女子)[2016年1月], 北アメリカ予選(女子)[2016年1月], アフリカ予選(女子)[2016年2月]

予測対象: リオデジャネイロオリンピック 世界最終予選1(女子), 2016年5月

データセット W3

W2に

- リオデジャネイロオリンピック 世界最終予選1(女子)[2016年5月]
- 2016年ワールドグランプリ[2016年6月-7月]

を追加したもの。

予測対象: リオデジャネイロオリンピック(女子)[2016年8月]

表 3. Root mean squares (RMSs).

	by Set	by Game		by Set	by Game
Dataset W1	0.0697	0.0429	Dataset M1	0.0594	0.0381
Dataset W2	0.0701	0.0369	Dataset M2	0.0612	0.0305
Dataset W3	0.0710	0.0390	Dataset M3	0.0611	0.0321

データセット M1

- 2011 年ワールドカップバレーボール(男子) [2011 年 11 月–12 月]
- ロンドンオリンピック ヨーロッパ予選(男子) [2012 年 5 月], 南アメリカ予選(男子) [2012 年 5 月], 北アメリカ予選(男子) [2012 年 5 月], アフリカ予選(男子) [2012 年 1 月], 世界最終予選(男子) [2012 年 6 月]
- 2012 年ワールドリーグ [2012 年 5 月–7 月]

予測対象：ロンドンオリンピック(男子) [2012 年 7 月–8 月]

データセット M2

- 2015 年ワールドリーグ [2015 年 5 月–7 月]
- 2015 年ワールドカップバレーボール(男子) [2015 年 9 月]
- リオデジャネイロオリンピック ヨーロッパ予選(男子) [2016 年 1 月], 南アメリカ予選(男子) [2015 年 10 月], 北アメリカ予選(男子) [2016 年 1 月], アフリカ予選(男子) [2016 年 1 月]

予測対象：リオデジャネイロオリンピック 世界最終予選 1(男子) [2016 年 5 月]

データセット M3

M2 に

- リオデジャネイロオリンピック 世界最終予選 1(男子) [2016 年 5 月]
- 2016 年ワールドリーグ [2016 年 6 月–7 月]

を追加したもの。

予測対象：リオデジャネイロオリンピック(男子) [2016 年 8 月]

3.2 結果

3.2.1 数値的な安定性

アルゴリズムの数値的な安定性を確認するため、以下の数値実験を行った。データセット W3 に対し、初期レーティングベクトルの要素を正規乱数で与えた後和が 0 となるよう定数を加え、提案アルゴリズムを 25 回実行した。パラメータは $\epsilon_{th} = 10^{-5}$, $K = 0.1$ とした。得られた各国のレーティングの最大値と最小値の差の上限は 2.2×10^{-5} 、標準偏差の上限は 4.7×10^{-6} であり、以降の分析に対し十分な精度の値が安定して得られた。

表 3 に、それぞれのデータセットのレーティング算出対象の試合について、レーティングから予測される得点率 p_k と実際の得点率 s_k (k はセットまたは試合番号) との残差の二乗平均平方根 RMS (次式)を示す。

$$(3.1) \quad RMS = \sqrt{\frac{1}{N_S} \sum_{k=1}^{N_S} (p_k - s_k)^2}$$

データセット W3 を例とすると、RMS はセットごとで 0.0710、試合ごとで 0.0390 であった。セットごとの残差 $p_k - s_k$ を調べたところ、95% のセットが $|p_k - s_k| < 0.139$ を満たしていた。残差 0.139 は平均で 1 セットあたり約 6 点程度の予測誤差に相当する。試合ごとでは 95% の試

表 4. Correlation coefficients based on proposed rating method.

	Correlation coefficient		Correlation coefficient	
	by Set	by Game	by Set	by Game
Dataset W1	0.5110	0.7909	Dataset M1	0.5313 0.6775
Dataset W2	0.5549	0.7244	Dataset M2	0.4764 0.6413
Dataset W3	0.5605	0.7384	Dataset M3	0.5607 0.7381

合が $|p_k - s_k| < 0.076$ を満たしている. 3, 4, 5 セットの試合ではそれぞれ約 9.8, 13.1, 15.1 点の予測誤差に相当する.

3.2.2 主結果

表 4 に説明変数をレーティング差から予測される得点率(式(2.1)), 決定変数を予測対象の試合での実際の得点率(セットごと, および試合ごと)としたときの相関係数を示す. 図 3 および図 4 に, W3 および M3 で算出したレーティング差とリオデジャネイロオリンピックで行われた実際の試合ごとの得点率を示す. これらの図中の実線は式(2.1)で予測される得点率を示している.

3.3 考察

表 4 の結果より,

- レーティングから予測された得点率と各セットの実際の得点率には中程度の相関があり(相関係数が 0.5 強),
- レーティングから予測された得点率と各試合の実際の得点率には中程度から強い相関がある(相関係数が 0.7 付近)

ことが確認できた.

また図 3 および図 4 より, 番狂わせはレーティング差が小さい対戦で若干起こっており, 一定以上のレーティング差では番狂わせが起きていないことが示されている.

これらの結果より, 提案手法に基づくレーティングは各チームの実力を適切に評価し, その後の試合結果をある程度予測できることが示された.

3.4 FIVB ランキングとの比較

本節では提案手法に基づくレーティングと FIVB ランキングのどちらが試合結果をより正確に予測できるかについて議論する.

全てのデータセットに対し, 予測対象とした大会での各セットおよび各試合の得点率と, 直前の FIVB ランキング差および FIVB ランキングポイント差との相関係数を求めたものを表 5 に示す. ロンドンオリンピック(データセット W1 および M1)については, 開催国枠で出場したイギリスの世界ランキングが極端に低く(女子 69 位, 男子 92 位), 外れ値となる恐れがあったため, イギリスの試合結果を除外した結果もあわせて示す.

ランキング差の算出にはランキングの順位をそのまま使用しているため, 相関係数はマイナスになる(ランキング差がマイナスになるほど自分のランキングが高い). 一部強い相関を示している大会もあるが, おおむね提案手法よりも弱い相関を示している. データセット M1 とデータセット M1*(イギリスを除外したもの)を比較すると, 極端にランキング差およびランキングポイント差の絶対値が大きいデータにより擬似的な相関が導入されていることが分かる.

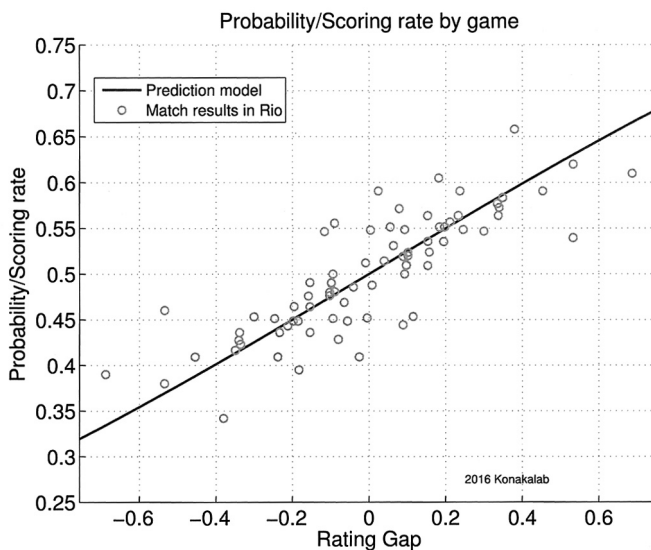


図 3. Rate difference and scoring rate in each game (Rio Olympic Games 2016, Women).

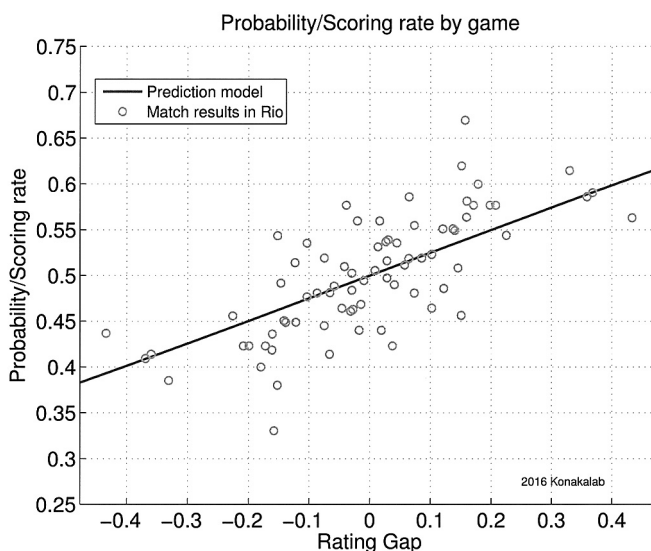


図 4. Rate difference and scoring rate in each game (Rio Olympic Games 2016, Men).

相関が低くなる要因は、ランキングおよびランキングポイントが間隔尺度となっていないことである。ランキングまたはランキングポイントに何らかの非線形な変換を施すことにより相関が見かけ上高くなることも考えられるが、相関を高くするように選択された変換が定性的な意味を持つとは限らない。この点において、レーティング差は得点確率を説明している、つまりレーティングは間隔尺度である、という統計的意味が明確である提案手法が優れている。

図 5 に W3 に基づき算出したレーティングとその時点での FIVB ランキングポイントとの関係を示す。

表 5. Correlation coefficients: FIVB ranking and scoring rate.

	Correlation coefficient			
	FIVB ranking gap		FIVB ranking point gap	
	by Set	by Game	by Set	by Game
Dataset W1	-0.3074	-0.5491	0.4179	0.5983
Dataset W1*	-0.3768	-0.5417	0.3763	0.5267
Dataset W2	-0.5078	-0.6578	0.3658	0.4901
Dataset W3	-0.4948	-0.6629	0.4713	0.6446
Dataset M1	-0.4789	-0.6105	0.5362	0.7016
Dataset M1*	-0.4587	-0.6331	0.4731	0.6512
Dataset M2	-0.3624	-0.5380	0.3171	0.4641
Dataset M3	-0.5992	-0.7899	0.5342	0.7166

*: except GBR.

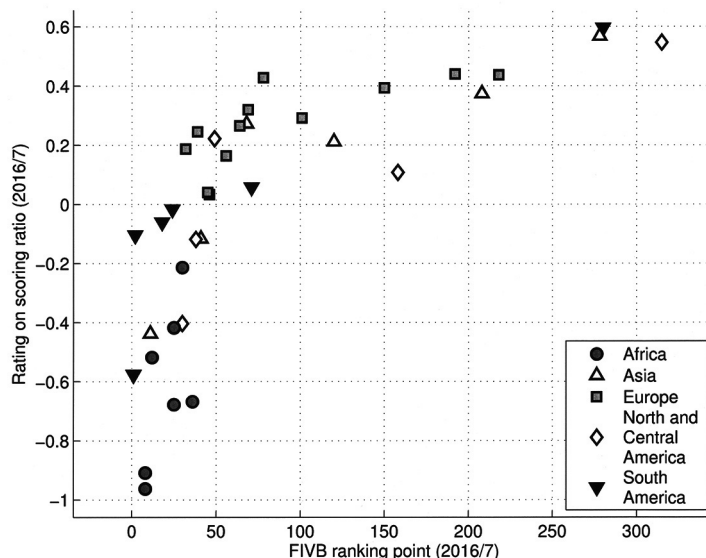


図 5. FIVB ranking point and proposed rating (2016/7, Women).

図 5 より，レーティングが 0.4 前後のヨーロッパ各国間でのランキングポイントの差が大きいことや，ランキングポイントが 40 前後の各国間でのレーティング差が大きいことが分かる．この結果からも，FIVB ランキングポイントがチームの実力を正しく反映していないことが分かる．この原因については次節で議論する．

4. 大会形式の評価への応用

本章では提案手法に基づくレーティングに基づく大会形式の評価(主に不適切な設計の指摘)について述べる．

4.1 世界最終予選の不適切な設計

リオデジャネイロオリンピック世界最終予選の参加国を以下に示す．

- 世界最終予選 1: アジア (FIVB ランキングアジア内上位 3 カ国および開催国(日本)), ヨーロッパ(予選 2・3 位), 北アメリカ(予選 2 位), 南アメリカ(予選 2 位). このうち 4 カ国(アジア最上位国とそれを除く上位 3 カ国)がオリンピック出場.
- 世界最終予選 2: アフリカ(予選 2・3 位), 北アメリカ(予選 3 位), 南アメリカ(予選 3 位). このうち最上位 1 カ国がオリンピック出場.

いずれの最終予選も 1 回戦総当りで, 順位は勝利数が多い順で決められた. この方式では北アメリカ予選で世界最終予選 2 を「選択」することでオリンピックへの出場権を獲得しやすくなる可能性がある. この時点で敗退行為を誘発しうる不適切な制度設計であり, 実際に最終予選 1 の出場国が疑問を呈している(奥村, 2016).

表 6 に世界最終予選直前のレーティング, 世界最終予選前最新の FIVB ランキングポイント(2015 年 10 月発表), 最終的なオリンピック出場権, および各予選での順位を示す. 国名は IOC の 3 文字コード(International Olympic Committee, 2016)で示す. BRA は開催国, CHN および SRB は 2015 年ワールドカップでオリンピック出場権を得ているため, 各大陸予選には参加していない.

前節での考察と同様に, 表 6 は, FIVB ランキングポイントが実力を適切に評価できていないことを示している. 各大陸選手権も含めて, 大会の順位に基づいてポイントが付与されている(表 1)ので, チーム数の多いヨーロッパには高レーティングにも関わらずポイントが少ないチームがあり(特に NED, GER), それに対してそれ以外の大陸の上位チーム(JPN, KOR(アジア), ARG(南アメリカ), KEN(アフリカ)など)はレーティングに対して付与ポイントが多い. FIVB ランキングが「世界」ランキングとして実力を正しく表せていないことが分かる.

レーティング以外にも, 大規模な国際大会での得点率を利用することでも FIVB ランキングの不適切な設計を指摘することができる. バレーボールでは各大陸から代表チームが出場する国際大会(ワールドグランプリなど)が毎年開催され, 大陸間の実力差は定性的には明確である. 表 7 に 2016 年ワールドグランプリでの大陸間対戦での得点率を示す. アフリカを除き, 各大陸で最上位の成績であったチーム(BRA, CHN, RUS, USA)を除外し, 中堅チーム間対戦のみの得点率もあわせて示す.

この結果からもアフリカ各国チームの実力が劣っていることは明確であり, また北アメリカは大陸 1 位(USA)とそれ以下との実力差が大きいことも読み取れる.

これらの統計に基づく, 北アメリカ予選に参加する USA は明確に実力が高く, 1 位での通過がほぼ確実である. 北アメリカの 2 位は世界最終予選 1 に回ることとなるが, ここにはヨーロッパの 2 位および 3 位, およびアジアの上位国(JPN, THA, KOR)が参加する. つまり, レーティングが 0.40 から 0.25 程度の 4 チームの出場が確定しており, レーティングが 0.2 付近の北アメリカ 2, 3 番手から見ると勝利が難しい相手ばかりである. したがって, この中で 4 位以上となることも困難が予想される.

それに対し, 世界最終予選 2 の対戦国は南アメリカ, アフリカからの 3 カ国であり, いずれもレーティングは 0 以下である. 北アメリカの 3 位になり世界最終予選 2 に出場すれば, そこで 1 位を目指す方がハードルが低いように思われる.

北アメリカ大陸予選は 4 チームの 1 回戦総当りであり, かつ USA が 3 勝することがほぼ確実視されている. したがって, 2 位と 3 位を争う 2 チームの順位は直接対決 1 試合の結果のみで決まる可能性は低くはない.

この状況で行われた北アメリカ予選の初戦で DOM-PUR の対戦が行われ, セット数 3-1 で DOM が勝利した. USA は実力どおり 3 連勝し, DOM が北アメリカ 2 位で最終予選 1 へ, PUR が北アメリカ 3 位で最終予選 2 へ進出した. DOM は最終予選 1 でヨーロッパ, アジアの強豪

表 6. Rating just before Rio WOQT (2016/5) and qualifying rankings.

IOC Code	FIVB Ranking		Qualifiers(**)						
	Rating	Points (*)	OL	AF	EU	NA	SA	WOQT1	WOQT2
'BRA'	0.5951	302	✓						
'CHN'	0.5583	310	✓						
'USA'	0.4839	350	✓			1			
'SRB'	0.4308	212	✓						
'RUS'	0.4014	253	✓		1				
'NED'	0.3913	82	✓		2			2	
'GER'	0.3713	93			5				
'JPN'	0.3440	225	✓					3	
'ITA'	0.3330	163	✓		3			1	
'THA'	0.2770	83						5	
'TUR'	0.2472	117			4				
'BEL'	0.2115	76			6				
'CZE'	0.1962	38							
'KOR'	0.1927	146	✓					4	
'POL'	0.1875	44			7				
'PUR'	0.1613	70	✓			3			1
'BUL'	0.0951	57							
'DOM'	0.0685	186				2		6	
'CAN'	0.0494	64				4			
'ARG'	0.0386	87	✓				1		
'CRO'	-0.0427	51			8				
'CUB'	-0.0627	52							
'PER'	-0.0659	51					2	7	
'COL'	-0.1147	41					3		3
'KAZ'	-0.1430	42						8	
'VEN'	-0.1467	16					4		
'KEN'	-0.2257	58			3				2
'CMR'	-0.4291	51	✓	1					
'EGY'	-0.5280	34		2					
'CHL'	-0.6191	11					5		
'ALG'	-0.6747	51		4					4
'TUN'	-0.6906	25		6					
'BOT'	-0.9197	22		5					
'UGA'	-0.9715	8		7					

(*) FIVB Senior World Ranking - Women. As per 6th October 2015 (FIVB, 2015). It is the latest FIVB ranking before WOQTs.

(**) OL: Olympic games, AF: Africa, EU: Europe, NA: North and Central America, SA: South America, WOQT: World Olympic Qualifying Tournament

国と対戦し、8チーム中6位でオリンピック出場を逃している。それに対しPURは実力の劣る南アメリカ、アフリカの3チームと対戦し、1セットも落とすことなく3戦全勝の1位でオリンピック出場権を得ている。

この状況下での世界最終予選の設計は上記の理由から明らかに不適切であり、該当するチームに対し不要な疑いが向けられかねない。なおかつ、前述の通り、各大陸の上位チームには実力(実際の試合で得点を取る能力)に見合わないランキングポイントが多く付与される傾向にあ

表 7. Scoring ratio in inter-continental matches in World Grand Prix 2016.

(a) All teams							(b) Excluding the strongest team in four continents						
(*)	Teams	Opponents					(*)	Teams	Opponents				
		AF	AS	EU	NA	SA			AF	AS	EU	NA	SA
AF	2	—	0.318	0.377	0.422	0.381	AF	2	—	0.318	0.377	0.421	0.381
AS	5	0.682	—	0.494	0.494	0.444	AS	4	0.682	—	0.487	0.516	0.427
EU	11	0.623	0.506	—	0.494	0.479	EU	10	0.623	0.513	—	0.523	0.529
NA	6	0.579	0.506	0.506	—	0.482	NA	5	0.579	0.482	0.477	—	0.482
SA	4	0.619	0.556	0.525	0.518	—	SA	3	0.619	0.571	0.471	0.518	—

(*) continent names; AF: Africa, AS: Asia, EU: Europe, NA: North and central America, SA: South America

り、公式に発表されるランキングポイントを信用すると却って大陸間の実力差が見えにくくなってしまふこともあわせて問題である。

敗退行為・談合試合が起こるかどうかは大会方式を設計する上で非常に重要な点である。サッカーで談合試合の疑惑をもたれている有名な事例が1982年ワールドカップスペイン大会で起こっている(Smyth, 2014)。一次リーグ最終戦で対戦した2チームが、お互いが一次リーグを突破できるように試合を操作したのではないかという疑惑である。原因は一次リーグ最終節の試合時間が異なり、当該の2チームがともに一次リーグを勝ち抜く条件が試合前に確定していたことである。この事例を受け、1986年大会から一次リーグ最終節の2試合は同時刻に開始するよう変更されている。

他のスポーツでも、ロンドンオリンピックのバドミントンでは準々決勝の組み合わせを有利にするために4組のペアが故意の失点を繰り返したことが問題となった(Gilmour, 2012)。これは故意に負けることがその先で明らかに有利になるよう、不適切に一次リーグ戦を設計したことが大きな要因である(これ以前の大会のようにトーナメント戦であれば敗退行為は起こりえない)。これを受け、続くリオデジャネイロオリンピックでは大会形式が変更されている。

Jリーグで2シーズンのみ実施された2ステージ+チャンピオンシップ制についても、敗退行為の可能性を排除するよう制度設計が行われたことが公式に発表されている(J League, 2013; 泉・小中, 2016)。

上記の事例に鑑み、バレーボールでも最終予選制度の適切な再設計が望ましい。

4.2 リオデジャネイロオリンピックの予選プールの不均衡

表8にオリンピック本選における予選プールを示す。各プールは直前のFIVBランキングに基づいて決定されたが、ランキングとレーティングによる実力の評価には大きな隔たりがあり、Pool Bに実力国が集まっていることが分かる。ITAはPool Bで6チーム中5位となり準々決勝に進出できなかったものの、レーティングでは全体の7番目であり、Pool Bの上位4カ国が全て準々決勝を通過していることから、適切に組み分けが設計されていれば準々決勝進出が十分期待できる実力であったと思われる。

Pool Aではアジアおよび南アメリカの3カ国が実力以上のランキングを得ていることが分かる。特にJPNは主要国際大会であるワールドカップバレーボールの1977年以降全ての大会に開催国として参加し、ランキングポイントを獲得している影響が大きい。ワールドカップの出場国数は12カ国のみ、かつ大陸ごとの参加枠があるため、JPNは必ず実力の劣る数チームと対戦でき、ランキングポイントを獲得できる。それに対しヨーロッパからワールドカップへの出場枠は2つのみであり、出場権を獲得できないチーム(2015年大会ではITA, NEDなど)は

表 8. Rio 2016 women's volleyball pools and results.

Pool A					Pool B				
IOC code	FIVB ranking	Rating	Result (Pool)	Result (Final)	IOC code	FIVB ranking	Rating	Result (Pool)	Result (Final)
BRA	3 (Host)	0.5963	1	5T	USA	1	0.5464	1	3
RUS	4	0.4365	2	5T	CHN	2	0.5678	4	1
JPN	5	0.3737	4	5T	SRB	6	0.4396	3	2
KOR	9	0.2115	3	5T	ITA	8	0.3934	5	9T
ARG	12	0.0574	5	9T	NED	14	0.4276	2	4
CMR	21	-0.4174	6	11T	PUR	16	0.2222	6	11T

当然ランキングポイントも 0 であり、これが世界ランキングでの不当な低評価の原因である。

ランキングに基づいた大会制度設計の大前提は実力を適切に反映したランキングであり、その実現のためには、大会の参加権の与え方、および各大会・各順位におけるランキングポイントの抜本的な修正が必要である。

5. 結論

本論文では、バレーボールの各国代表チームの実力の定量的な評価方法として、各セットでの得点率を説明するモデルを構築し、そのパラメータをレーティングとして定義した。またレーティングの具体的な計算アルゴリズムを示し、実際の国際大会の試合結果に基づき値を算出した。

算出したレーティングは公式のランキングである FIVB ランキングよりも試合結果を適切に予測できることを、オリンピックなど複数の大会の結果により確認した。レーティング差が世界ランキング(ポイント)差よりも適切に実力差を説明できることに基づいて、世界最終予選やオリンピックでの大会形式の不適切な設計、およびその原因である FIVB ランキングポイントの不適切な設計を具体的な事例に基づき指摘した。

参 考 文 献

- Araújo, R.M., Castro, J., Marcelino, R. and Mesquita, I.R. (2010). Relationship between the opponent block and the hitter in elite male volleyball, *Journal of Quantitative Analysis in Sports*, **6**(4), 1–12.
- ATP World Tour (2016). Rankings FAQ, <http://www.atpworldtour.com/en/rankings/rankings-faq>, accessed 2016/6/14.
- Burton, T. and Powers, S. (2015). A linear model for estimating optimal service error fraction in volleyball, *Journal of Quantitative Analysis in Sports*, **11**(2), 117–129.
- Davidson, R.R. (1970). On extending the Bradley-Terry model to accommodate ties in paired comparison experiments, *Journal of the American Statistical Association*, **65**(329), 317–328, <http://www.jstor.org/stable/2283595>.
- de Ayala, R.J. (2008). *The Theory and Practice of Item Response Theory (Methodology in the Social Sciences)*, Guilford Press, New York.
- Dziedzic, E. and Hunter, G. (2015). Predicting the results of tennis and volleyball matches using regression models, and applications to gambling strategies, *Proceedings of the 5th International Conference on Mathematics in Sport*, 32–40.
- Elo, A.E. (1979). *Ratings of Chess Players Past and Present*, Arco Publishing, New York.

- Eom, H.J. and Schutz, R.W. (1992). Statistical analyses of volleyball team performance, *Research Quarterly for Exercise and Sport*, **63**(1), 11–18.
- Ferrante, M. and Fonseca, G. (2014). On the winning probabilities and mean durations of volleyball, *Journal of Quantitative Analysis in Sports*, **10**(2), 91–98.
- FIVB (2015). FIVB Senior World Ranking - Women. As per 6th October 2015, http://www.fivb.org/en/volleyball/VB_Ranking_W_2015-10.asp, accessed 2017/2/10.
- FIVB (2016). FIVB volleyball world rankings, <http://www.fivb.org/en/volleyball/Rankings.asp>, accessed 2016/6/14.
- Florence, L.W., Fellingham, G.W., Vehrs, P.R. and Mortensen, N.P. (2008). Skill evaluation in women's volleyball, *Journal of Quantitative Analysis in Sports*, **4**(2).
- Gilmour, R. (2012). Badminton match-fixing scandal: How and why the four pairs were disqualified from the London 2012 Olympics, <http://www.telegraph.co.uk/sport/olympics/badminton/9444025/Badminton-match-fixing-scandal-how-and-why-the-four-pairs-were-disqualified-from-the-London-2012-Olympics.html>, accessed 2015/3/2.
- Glasson, S., Jeremiejczyk, B. and Clarke, S.R. (2001). Simulation of women's beach volleyball tournaments, *Australian Society for Operations Research*, **20**(2), 2–7.
- Hambleton, R. (1991). *Fundamentals of Item Response Theory (Measurement Methods for the Social Science)*, new edition, Sage Publications, California.
- Indyk, P. and Motwani, R. (1998). Approximate nearest neighbors: Towards removing the curse of dimensionality, *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*, 604–613.
- International Olympic Committee (2016). List of all National Olympic Committees in IOC protocol order, <https://www.olympic.org/national-olympic-committees>, accessed 2016/6/15.
- 泉武志, 小中英嗣 (2016). J1 リーグ 2 ステージ + ポストシーズン制度の統計的分析, *日本オペレーションズ・リサーチ学会論文誌*, **59**, 21–37.
- J League (2013). 10 月 30 日開催 大会方式変更に関する意見交換会についてのチェアマンコメント, <http://www.jleague.jp/release/article-00005433/>, accessed 2017/2/14.
- Knapper, E. and McIlwain, H. (2015). Predicting wins and losses: A volleyball case study, *The College Mathematics Journal*, **46**(5), 352–358, <http://www.jstor.org/stable/10.4169/college.math.j.46.5.352>.
- Massey, K. (1997). Massey Rating, <http://www.masseyratings.com/>, accessed 2016/6/14.
- 奥村信哉 (2016). 不公平感漂う「もう一つ」の最終予選! 日本大会出場の監督ら相次ぎ疑問の声 知られざるプエルトリコ大会, <http://www.sankei.com/sports/news/160524/spo1605240002-n1.html>, accessed 2016/5/25.
- Rao, P.V. and Kupper, L.L. (1967). Ties in paired-comparison experiments: A generalization of the Bradley-Terry model, *Journal of the American Statistical Association*, **62**(317), 194–204.
- Ray, S. (2011). The methodology of officially recognized international sports rating systems, *Journal of Quantitative Analysis in Sports*, **7**(4).
- Sackmann, J. (2017). <https://github.com/JeffSackmann>, accessed 2017/4/4.
- Smyth, R. (2014). World Cup: 25 stunning moments ... No3: West Germany 1-0 Austria in 1982, <http://www.theguardian.com/football/blog/2014/feb/25/world-cup-25-stunning-moments-no3-germany-austria-1982-rob-smyth>, accessed 2015/3/3.
- World Rugby (2014). Rankings Explanation, <http://www.worldrugby.org/rankings/explanation>, accessed 2016/6/14.
- World Rugby (2015). World rankings confirm Japan's victory as biggest shock, <http://www.rugbyworldcup.com/news/111746>, accessed 2016/6/14.
- Zetou, E., Moustakidis, A., Tsigilis, N. and Komninakidou, A. (2007). Does effectiveness of skill in complex i predict win in men's Olympic volleyball games?, *Journal of Quantitative Analysis in Sports*, **3**(4).

Statistical Rating Method of Volleyball National Teams to Predict Results and Determine Competition Format Design

Eiji Konaka

Faculty of Science and Technology, Meijo University

The Fédération Internationale de Volleyball (FIVB), the world's governing body for volleyball, regularly ranks its member nations' teams. The FIVB Board of Administration has designed a point system for select FIVB world and other official competitions. However, the point system does not have a clear mathematical or statistical background. Consequently, this system cannot be used as a quantitative measure of a team's skill.

This paper proposes a novel mathematics-based rating and ranking system of national volleyball teams. The rating, which is a parameter reflecting the skill of the team, is calculated based on the scoring ratio of teams in each major international competition. A logistic regression model is employed to explain the scoring ratio with respect to the rating difference between two teams. Additionally, an iterative rating calculation method is proposed. Numerical experiments demonstrate the stability of the proposed method. The correlation coefficient between the proposed rating difference and the results of several major international competitions (e.g., Rio Olympic Games) is about 0.7. This value shows a strong correlation that is higher than that of the FIVB ranking (point) difference. The proposed rating is used to highlight the improper design of the ranking point attribution system and the competition format of the World Olympic Qualifying Tournament and Rio Olympic Games.