

大学野球におけるデータ活用システムの構築と提案

河崎 達成[†] 時井 真紀[‡]

筑波大学 情報学群 知識情報・図書館学類[†]

筑波大学 図書館情報メディア系[‡]

1. はじめに

近年、大学野球現場において体組成やトレーニング計測、試合時スカウティングデータの他、投手が投球する球質を即時的に数値化するトラッキングデータが普及し始めた。そこで多くのデータを活用した指導が求められている。そして、選手視点におけるデータ活用が個人的なパフォーマンス向上につながる^[1]。そのためにデータから得られる知見をアナリストが分析し、選手がその分析を理解することが必要であると考えられる。また、大学野球において普及してきたトラッキングデータが試合時のデータと関連性があることが明らかになっている^[2]。練習時におけるトラッキングデータが試合にどのような影響があるのかをチーム内で比較できれば、選手が数値的指標を獲得することができ、個人の数値的な特徴が明確になると考えられる。

そこで、本システムの構築により、大学野球における多くのデータを一元化し、チーム内での比較やクラスタリングからアナリストが選手個人の数値的実態を理解し、選手に対して容易に分析を伝えることが出来るようになることを目的とする。

2. システム

本システムでは「個人カルテ」、「全体比較」と「分析」を3つの指針とし、それぞれのページを作成した。「個人カルテ」では体組成計測データ、リバウンドジャンプ計測データ、パフォーマンス計測データ、トラッキングデータ、スカウティングデータにおいて選手一人分のデータを一元化し表示した。「個人カルテ」の一部を図1に示す。

箱ひげ図や折れ線グラフで過去のデータの推移を示し、レーダーチャートで練習と実戦練習の比較を明らかにした。



図1 個人カルテの一部

また、トラッキングデータにおける変化量項目に関しては図2に示すように散布図を各球種における散らばりをマハラノビス楕円で囲み、各球種の平均値での点の散らばり方をポリゴン図で作成した。年月ボタンを作成し、ユーザーの選択により月別の散布図を透過させ、重ね合わせることでそれぞれの散布図を時系列での比較ができるようにした。

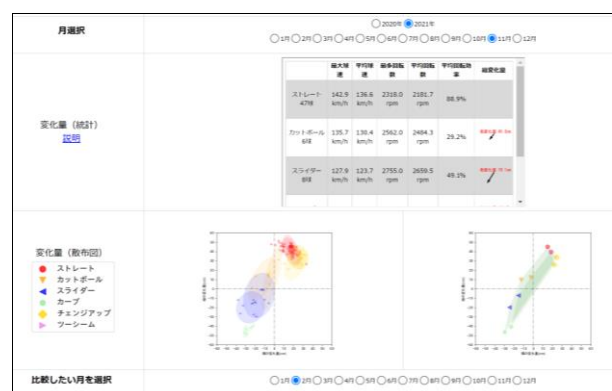


図2 変化量データ

Development of a Data Visualization and Analysis System for College Baseball

[†]Tatsunari Kawasaki

College of Knowledge and Library Sciences, School of Informatics, University of Tsukuba

[‡]Maki Tokii

Faculty of Library, Information and Media Science, University of Tsukuba

「全体比較」では「個人カルテ」で表示したデータ5項目（体組成計測データ、リバウンドジャンプ計測データ、パフォーマンス計測データ、ストレート球質データ、変化量データ）を項目ごとにチーム内の選手全員分を一覧表示し

た。さらに、時系列での数値変化に応じて色分けし、チーム内で全体比較をできるようにした。

「分析」では投手が投げるストレートの実戦練習時の4か月分(2021年8月～11月)の計測値(打者を相手にした時のトラッキングデータ)における球速(km/h)、回転数(rpm)、回転効率(%)、横の変化量(cm)、縦の変化量(cm)でグループ分けを行い、試合での打球傾向について明らかにすることを目的とした。球質5項目のそれぞれの平均値を正規化し、その数値をもとに階層クラスタリングを行い、図3に示すように樹形図から5クラスに分けた。

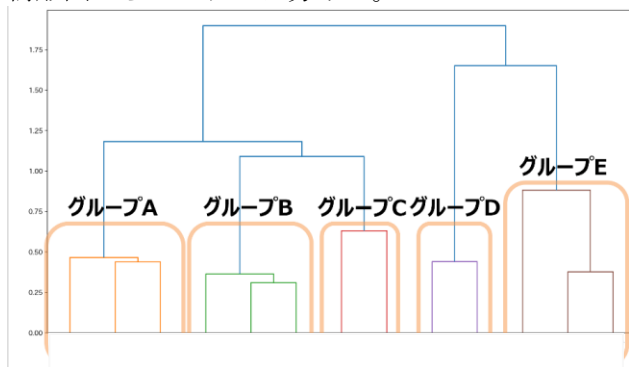


図3 グループ分け結果

クラスタリングの結果からグループ別の統計数値を表示した。この統計数値から、グループごとの球質の特徴を早見表や、理論も含めた文章にして表示をした。グループ分けによる球質の特徴が実際の試合ではどのような影響が及ぼされているのかを比較するために、実際の試合での打球傾向(空振り、ファール、ゴロ、フライ、ライナー)をグループごとにまとめた。使用したデータは4か月間(2021年8月～11月)のスカウティングデータでグループごとに打者がスイングした数に対しての空振り率、ファール率、ゴロ率、フライ率、ライナー率を積み上げ棒グラフにして表示した。

3. 結果

アナリスト経験の異なる大学生2名、大学院生1名と硬式野球部に所属する投手の大学生5名を対象に、アナリスト1名が本システムの利用を通じて投手1名に対して分析を報告する実験を行った。アナリストに対してはシステムに関して、投手に対してはデータ分析の理解の観点で事前事後アンケートを行った。

アナリスト経験の違いによってシステム利用する際に選手に分析を行う順序や説明するために使用するページが異なっていた。また、共通して投手に対して感覚と数値をつなぎ合わせる

ような問いかけが見られた。そして、アナリスト経験が長く、データ理解度が高いほど「システムをスムーズに利用できる」という結果が得られた。さらに、アナリスト、投手両方に対して、事前のデータ理解度は関係なく、「システム利用後にデータへの興味関心」、「データ理解度」が上昇したという結果が得られた。

4. 考察

多くのデータを可視化し、他人比較や時系列の比較を行うことでアナリストが選手個人の数値の実態を理解でき、興味や関心を選手に対して引き付けることが出来ると考えられる。しかし、アナリストのデータへの理解度がデータ分析をスムーズに行うことに影響するため、データ理解度にかかわらず、容易にシステムから分析が可能となるような補足説明の工夫は必要である。補足説明により、経験の浅いアナリストだけに関わらず、データ理解の低い選手に対しても理解が深まるものになることが考える。

また、体組成データやパフォーマンスデータといった身体的な計測データが投球のデータに対してどのような影響があるかわからないといった意見が得られた。よって、一元化して表示するだけでなく、関連性がある項目を強調して比較することが今後の発展として必要である。

5. おわりに

本研究では多くのデータを一元化し、可視化することでアナリストが分析することが容易になり、選手が理解しやすくなることの実現を目指した。そこで、データの一元化と他人や過去のデータと比較することで選手の課題を明確化できるシステムの構築とその評価を行った。

データを理解することでスムーズに分析が出来るようになるだけでなく、データの理解が出来ていなくてもデータに関する説明を充実させることで経験に左右されずに容易に分析できるシステムになると考えられる。

参考文献

- [1] 徐広孝. 高等学校体育におけるスポーツアナリティクス教材の有効性: バレーボールのゲームパフォーマンス分析を経験した生徒の質問紙調査報告. スポーツと人間. 2021, vol. 5, no. 1, p. 37-45.
- [2] 松木拓弥, 鈴木秀男. トラッキングデータとセイバーメトリクス指標による MLB 投手の能力評価モデルの構築. 電子情報通信学会通信ソサイエティマガジン. 2018, vol. 12, no. 2, p. 117-125.