

2022年度 3月修了 修士論文

二塁走者がフライ性の単打によって
本塁に生還する可能性を規定する要因の分析

早稲田大学 大学院スポーツ科学研究科
スポーツ科学専攻 身体運動科学研究領域

5021A039-0

福本 健人

研究指導教員： 矢内 利政 教授

目次

緒言	1
第一章 deterministic model の構築	3
方法	3
結果	8
考察	10
第二章 統計モデルによる予測	13
方法	13
データ収集	13
モデリングに用いる変数	14
モデリング手法	26
結果	29
考察	38
リミテーション	45
実践における活用	47
結論	49
謝辞	49
参考文献	50

緒言

プロ野球界では高性能弾道測定器「TrackMan」や、画像解析トラッキングシステム「Hawk-Eye」が多数の球場に常設されたことにより、試合中の投球や打球の物理的特徴、野手のスプリント速度など、あらゆる運動学的データが収集可能となっている。これらを集約したビッグデータは統計手法を用いて客観的に分析され、投球における最適な変化球の特徴の分析(e.g. Foster,2020)¹⁾や打撃成績向上(e.g. Kato and Yanai,2022)²⁾などに数多く貢献している。しかし、ビッグデータを用いて、得点を増加させるために重要な走塁のパフォーマンス向上に着目した研究は少ない。

走塁のパフォーマンスが重要になる場面として、走者が二塁に出塁しており、外野手が第一捕球者となった単打が発生した場面が挙げられる。この場面について、三塁コーチ経験者は「一番難しいのが、走者が二塁にいるときにワンヒットで、本塁まで行かせるか、三塁で止めるかであろう。」と述べている(澤宮,2018)³⁾。三塁コーチは、二塁走者や三塁走者の走塁を指示する役割を持つ。プロ野球選手として活躍した村松有人は、二塁走者の走塁において、単打で本塁へ進塁するかどうかの判断について問われた際、「これはコーチャーに任せます。コーチャーを見ながら全力でサードに行きます。もちろんホームへ行くつもりで走ります。」と述べている(成美堂出版,2004)⁴⁾。このように、二塁走者の走塁判断の決定権は、三塁コーチに与えられているといえる。そのため、この場面で二塁走者が生還する難しさを、三塁コーチは誰よりも知っているといえる。これほど重要で難しいプレーであるのにも関わらず、この場面の走塁パ

パフォーマンスに関する先行研究は、私の知る限りない。

そこで本研究では、走者が二塁におり、外野手が第一捕球者であるフライ単打が発生した場面において、二塁走者の生還成否を規定する走者・守備者及び打球特徴の要因分析を、2段階に分けて行った。研究の目的は、以下の2つとなる。1つ目は、二塁走者が本塁生還する可能性を規定する力学的要因を系統的に示す deterministic model(Hay, 1988)⁵⁾を構築することである。deterministic modelsとは、力学的要因から構成され、かつ入力指定されたときに出力が決定されるモデルのことである。このモデルの末端要因の内、本研究では取得不可能並びに選手やコーチが変更できない要因除き、残った要因を決定要因として扱う。2つ目は、deterministic modelにおいて定義された決定要因を用いた統計モデルを作り、予測精度を高める影響が大きい要因を導出することである。

第一章 deterministic model の構築

方法

本研究では、二塁走者の生還率を決定する力学的要因を系統的に導くために、deterministic model を構築した。deterministic model の第一段は、二塁走者の生還率である。第一段の二塁走者の生還率は、「インパクトから送球が本塁に到達するまでの時間」と「インパクトから走者が本塁に到達するまでの時間」によって決定される。「インパクトから送球が本塁に到達するまでの時間」は、打球のインパクトから外野手が捕球をし、送球が本塁に到達するまでの守備側の時間を示している。「インパクトから走者が本塁に到達するまでの時間」は、打球のインパクトから二塁走者及び三塁コーチが走塁判断をし、二塁走者が走り始めて本塁に到達するまでの攻撃側の時間を示している。二塁走者の生還率は、守備側の時間と攻撃側の時間の差によって決定される。守備側の時間が攻撃側の時間より長い場合、二塁走者が送球より早く本塁に到達するため、二塁走者は本塁へ生還することができる。守備側の時間が攻撃側の時間より短い場合、送球が二塁走者より早く本塁に到達するため、二塁走者は本塁への生還を阻止される。これらのことから、守備側の時間が大きくなるほど生還率が高くなり、攻撃側の時間は短くなるほど生還率が高くなる。第二段の「インパクトから送球が本塁に到達するまでの時間」は、「インパクトから捕球までの時間」と「捕球から送球が本塁に到達するまでの時間」によって決定される。「インパクトから捕球までの時間」は打球の物理的特徴に起因する時間であり、「捕球

から送球が本塁に到達するまでの時間」は外野手の守備能力に起因する時間となる。第二段の「インパクトから走者が本塁に到達するまでの時間」は、「インパクトから走塁開始までの時間」と「走塁開始から走者が本塁に到達するまでの時間」によって決定される。「インパクトから走塁開始までの時間」はコーチが走塁判断・指示をし、走者が本塁へ走り始めるまでの時間であり、「走塁開始から走者が本塁に到達するまでの時間」は二塁走者の走力と本塁までの残り距離に起因する時間となる。第四段以降は、第三段それぞれの要因でその段数が異なる。

第三段の「インパクトから捕球までの時間」は、「打球の飛翔時間」、「打球の落下地点と捕球者の初期位置の成す角度」、「打球の落下地点と捕球者の初期位置間の距離」、「バウンド後の打球速度」によって決定される。インパクトから打球が落下するまでの時間は「打球の飛翔時間」となり、落下してから捕球するまでの時間は、「打球の落下地点と捕球者の初期位置の成す角度」、「打球の落下地点と捕球者の初期位置間の距離」、「バウンド後の打球速度」の値の組み合わせで変化する。第四段の「打球の飛翔時間」は、「打球のインパクト位置」、「打球初速度」、「鉛直打球角度」、「重力加速度」によって決定される。これらは、飛翔時間を力学的に決定する要因となっている。第四段の「打球落下地点と捕球者の初期位置の成す角度」と「打球落下地点と捕球者の初期位置間の距離」は、「捕球者の初期位置」と「打球落下地点」によって決定される。図1に示したように、捕球者の初期位置と打球落下地点を結んだ線分と、捕球者の初期位置と本塁中心を結んだ線分が成す角度が「打球落下地点と捕球者の初期位置の成す角度」となる。捕球者の初期位置と打球落

下地点を結んだ線分の長さが、「打球落下地点と捕球者の初期位置間の距離」となる。第五段の「打球落下地点」は、「水平打球角度」と「打球飛距離」によって決定される。

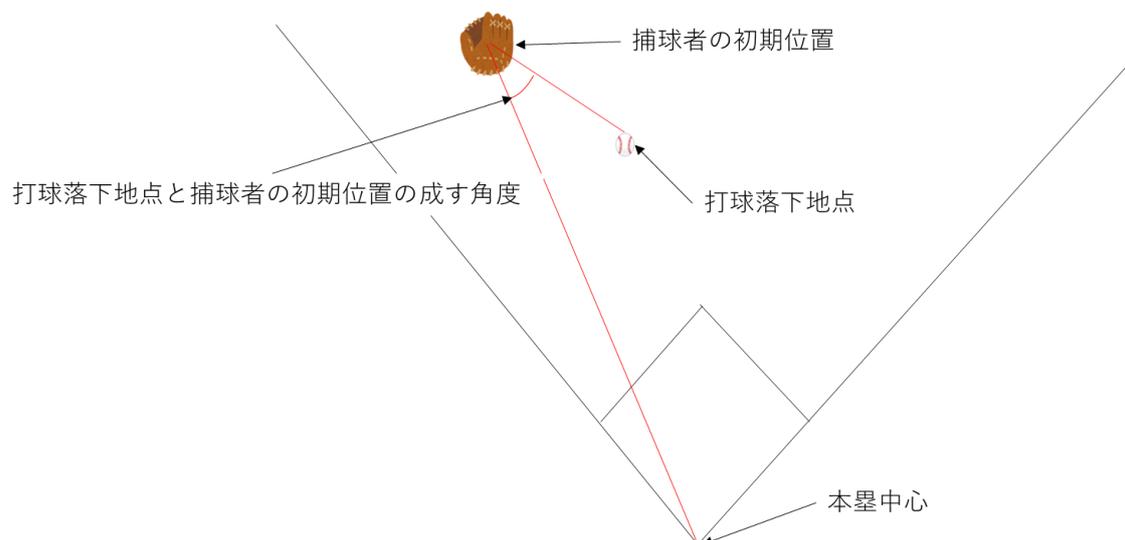


図 1 捕球者の初期位置と打球落下地点と本塁中心の関係

水平打球角度と打球飛距離から三角関数を用いて「打球落下地点」の座標を算出することが可能となる。第四段の「バウンド後の打球速度」は、「打球初速度」と「外野の芝に起因する摩擦係数」によって決定される。打球初速度から外野の芝の質や深さに応じた抵抗を受け減速した打球速度が、「バウンド後の打球速度」となる。第六段が最後の段となっており、「打球のインパクト位置」、「打球初速度」、「鉛直打球角度」、「重力加速度」、「捕球者の初期位置」、「水平打球角度」、「打球飛距離」、「バウンド後の打球速度」、「外野の芝の深さ」が末端要因となった。

第三段の「捕球から送球が本塁に到達するまでの時間」は、「外野手の送球動作」と「外野手の送球精度」によって決定される。「捕球から送球が本塁に到達するまでの時間」は、捕球から送球リリース

までの時間である「送球動作時間」と、送球リリースから本塁到達までの時間である「送球が本塁に到達するまでの時間」の和となる。第四段の「送球が本塁に到達するまでの時間」は、「送球の速度」と「送球精度」によって決定される。送球の速度によりおおまかな時間が決定され、送球の上下左右方向の精度によって捕手が捕球をしてから走者をタッチするまでの時間が決定される。第五段が最後の段となっており、「送球動作時間」、「送球の速度」、「送球精度」が末端要因となった。

第三段の「インパクトから走塁開始までの時間」は、「走者と三塁コーチの走塁判断に要する時間」によって決定される。第四段の「走者と三塁コーチの走塁判断に要する時間」は、「走者と三塁コーチの打球判断に要する時間」、「外野の芝に起因する摩擦係数」、「天候」、「イニング」、「アウトカウント」、「塁状況」、「点差」、「二塁走者の走力」、「二塁走者のリードの大きさ」によって決定される。走者とコーチの走塁判断・指示の際、これらの要因を複合的に考慮し、その値の組み合わせにより「走者と三塁コーチの走塁判断に要する時間」が決定される。質的変数である「外野の芝に起因する摩擦係数」、「天候」、「イニング」、「アウトカウント」、「塁状況」、「点差」は、直接的には力学的変数ではない。しかし、これらは「時間」という力学的変数を規定する質的変数であるため、本研究では生還率の規定因子として扱う。第四段の「走者と三塁コーチの打球判断に要する時間」は、「水平打球角度」、「打球飛距離」、「打球の飛翔時間」によって決定される。「走者と三塁コーチの打球判断に要する時間」は打球がヒットになるかアウトになるかの判断のことであり、「水平打球角度」、「打球飛距離」、「打球の飛翔時間」から判断さ

れる。「打球の飛翔時間」は前述の通りであり、「打球のインパクト位置」、「打球初速度」、「鉛直打球角度」、「重力加速度」から決定される。第七段が最後の段となっており、「水平打球角度」、「打球飛距離」、「打球のインパクト位置」、「打球初速度」、「鉛直打球角度」、「重力加速度」、「外野の芝に起因する摩擦係数」、「天候」、「イニング」、「アウトカウント」、「塁状況」、「点差」、「二塁走者の走力」、「二塁走者のリードの大きさ」が末端要因となった。

第三段の「走塁開始から走者が本塁に到達するまでの時間」は、「二塁走者の走力」と「本塁までの距離」によって決定される。「二塁走者の走力」と「本塁までの距離」が速度と距離の関係となっており、「走塁開始から走者が本塁に到達するまでの時間」を表す。第四段の「本塁までの距離」は、「二塁走者のリードの大きさ」によって決定される。二塁から本塁までの距離から「二塁走者のリードの大きさ」を引いた値が「本塁までの距離」となる。第五段が最後の段となっており、「二塁走者の走力」と「二塁走者のリードの大きさ」が末端要因となった。

結果

構築した deterministic model を図 2 に示す。二塁走者の生還率を出力とした deterministic model は、最大で第七段まで構築された。

「打球のインパクト位置」、「打球初速度」、「鉛直打球角度」、「重力加速度」、「捕球者の初期位置」、「水平打球角度」、「打球飛距離」、「バウンド後の打球速度」、「外野の芝に起因する摩擦係数」、「送球の速度」、「送球動作時間」、「外野手の送球精度」、「天候」、「イニング」、「アウトカウント」、「塁状況」、「点差」、「二塁走者の走力」、「二塁走者のリードの大きさ」が末端要因となった。これらの末端要因の内、本研究では取得不可能な要因並びに選手やコーチが変更できない要因として、「打球のインパクト位置」、「重力加速度」、「捕球者の初期位置」、「バウンド後の打球速度」、「外野の芝に起因する摩擦係数」、「送球動作時間」、「天候」、「二塁走者のリードの大きさ」を除外した。その結果、「打球初速度」、「鉛直打球角度」、「水平打球角度」、「打球飛距離」、「送球精度」、「イニング」、「アウトカウント」、「塁状況」、「点差」、「二塁走者の走力」が、二塁走者の生還する可能性を規定する力学的な決定要因となった。

考察

決定要因となった変数が、どのように値が変化すると生還率にどのような影響を与えるかについての仮説を述べる。外野手の送球精度は守備側の時間を決定する。外野手の送球精度は、値が小さくなるほど送球の質が悪くなるため、送球が本塁に到達するまでの時間が長くなり守備側の時間が長くなるため、生還率が高くなると考えられる。

イニングは攻撃側の時間を決定する。「7」、「8」、「9」の終盤を示す値の時、三塁コーチは目先の1点を狙うことを優先して、通常より早いタイミングで二塁走者に本塁突入の指示を与えると考えられる。そのため、判断に要する時間が短くなるため、攻撃側の時間が短くなり生還率が高くなると考えられる。アウトカウントは攻撃側の時間を決定する。2アウト時には、打球がヒットになるかアウトになるかを見分ける打球判断の問題がなくなり、二塁走者が打球のインパクトと同時に本塁へスタートを切ることができる。そのため走塁判断に要する時間が短くなるため、攻撃側の時間が短くなり生還率が高くなると考えられる。塁状況は攻撃側の時間を決定する。先の塁に走者がいない状況「二塁のみ」、「一・二塁」では、二塁走者が先の走者を気にせず早めにスタートを切ることができ、走塁判断に要する時間が短くなるため、生還率が高くなると考えられる。点差は攻撃側の時間を決定する。接戦である±3点以内の値であれば、三塁コーチは目先の1点を狙うことを優先して、通常より早いタイミングで二塁走者に本塁突入の指示を与えると考えられる。そのため、判断に要する時間が短くなるため、攻撃側の時間が短くなり生

還率が高くなると考えられる。二塁走者の走力は攻撃側の時間を決定する。値が大きくなるほど、走塁開始から走者が本塁に到達するまでの時間が短くなるため、攻撃側の時間が短くなり生還率が高くなると考えられる。

打球飛距離は守備側の時間と攻撃側の時間双方を決定する。打球飛距離の値は、外野手の定位置にあたる95m以下では、値が小さくなるほど外野手が捕球するまでの距離が大きくなるため、守備側の時間が長くなり生還率が高くなると考えられる。外野手の定位置にあたる95mを境に値が大きくなるほど、外野手が捕球するまでの距離が大きくなるため、守備側の時間が長くなり生還率が高くなると考えられる。同様に外野手の定位置を境に値が大きくなるほど、打球がヒットになるかアウトになるかの判断が容易になり判断に要する時間が短くなるため、攻撃側の時間が短くなり生還率が高くなると考えられる。打球初速度は守備側の時間と攻撃側の時間双方を決定する。打球初速度の値が小さくなるほど、外野手が捕球するまでの時間が長くなるため、守備側の時間が長くなり生還率が高くなると考えられる。値が大きくなるほど、打球がヒットになるかアウトになるかの判断が容易になり判断に要する時間が短くなるため、攻撃側の時間が短くなり生還率が高くなると考えられる。水平打球角度は守備側の時間と攻撃側の時間双方を決定する。水平打球角度の値は、左翼線、右翼線にあたる $\pm 45^\circ$ の値や、左中間、右中間にあたる $\pm 15^\circ$ の値で、外野手が捕球するまでの距離が大きくなり時間が長くなるため、守備側の時間が長くなり生還率が高くなると考えられる。同様に左翼線、右翼線にあたる $\pm 45^\circ$ の値や、左中間、右中間にあたる $\pm 15^\circ$ の値で、打球がヒットになるかアウトになるか

の判断が容易になり判断に要する時間が短くなるため、攻撃側の時間が短くなり生還率が高くなると考えられる。鉛直打球角度は守備側の時間と攻撃側の時間双方を決定する。鉛直打球角度は、打球の飛行時間を決定する要因となっており、値が大きくなるほど、打球の飛行時間が長くなり外野手が捕球するまでの時間が長くなるため、生還率が高くなると考えられる。10°前後の値でライナー性の打球となるため、打球がヒットになるかアウトになるかの判断が容易になり判断に要する時間が短くなるため、攻撃側の時間が短くなり生還率が高くなると考えられる。

第二章 統計モデルによる予測

方法

データ収集

本研究におけるデータの収集、フィルタリング及びモデリングには、統計ソフト R(R Core Team, 2020)⁶⁾を用いた。Rにはパッケージという関数やデータセットを集めた拡張機能がある。パッケージには集計やモデル構築などの分析が簡単にできる関数を備えたものもあり、本研究では `baseballr`(Petti et al.,2022)⁷⁾、`remotes`(Csárdi et al.,2021)⁸⁾、`tidyverse`(Wickham et al.,2021)⁹⁾、`psych`(Revelle et al.,2022)¹⁰⁾、`openxlsx`(Schauberger et al.,2021)¹¹⁾、`mgcv`(Wood,2022)¹²⁾、`mgcViz`(Fasiolo et al.,2021)¹³⁾、`MuMIn`(Bartoń,2022)¹⁴⁾を用いた。本研究で用いるデータは2016年から2021年において、MLB公式ウェブサイト「Baseball Savant」(MLB,online)¹⁵⁾と、MLB公認で FanGraphs 社が公開しているウェブサイト「FanGraphs」(FanGraphs,online)¹⁶⁾から取得した。どちらのサイトのデータも個人的、非商業的な家庭での利用が可能とされている。また、U.S. Copyright Office の Circ.21 に述べられている、学術または研究などの目的のための著作物の公正使用(Fair Use)に本研究は該当すると考えられるため、取得したデータの利用は著作権侵害に該当しないと想定される(U.S. Copyright Office,2017)¹⁷⁾。これらの理由により、本研究は早稲田大学より倫理申請不要の研究として認証されている。2016年以降のデータのみを

使用する理由は、2015年以前は記録されていなかった変数を研究に用いるためである。

Baseball Savantからは、アウトカウント(outs_when_up)、イニング(inning)、打球飛距離(hit_distance_sc)、打球初速度(launch_speed)、鉛直打球角度(launch_angle)、二塁走者の走力(sprint_speed)、打球位置の座標(hc_x, hc_y)を取得した。FanGraphsからは、外野手の送球による得点抑止力(arm)を取得した。二塁走者の生還有無(scored_second)、塁状況(runner_situation)、点差(run_difference)、水平打球角度(spray_angle)については、Baseball Savantのデータを基に生成した変数となっている。各変数の詳細については後述する。

モデリングに用いる変数

二塁走者の生還有無に関する直接的なデータは、Baseball Savantからは取得できない。そのため、本研究では、塁状況と得点数から二塁走者の生還有無を推測した。表1は二塁走者の生還有無判別結果を示した表となっている。

表1 二塁走者の生還有無判別方法

得点	出塁状況				
	二塁のみ	一・二塁	二・三塁	満塁	
1点	●	●	×	×	● : 二塁走者生還有り × : 二塁走者生還無し
2点以上	●	●	●	●	

安打が放たれた結果としての得点数が 0 点の場合、必然的に二塁走者の生還は「無し」と判定した。得点数が 1 点で三塁走者が出塁していない場合、二塁走者が最も先の塁にいる走者となるため、二塁走者の生還は「有り」とみなした。得点数が 1 点で三塁走者が出塁している場合、二塁走者は三塁走者の後の塁にいる走者となり、三塁走者が二塁走者よりも先に生還すると考えられるため、二塁走者の生還は「無し」とみなした。得点数が 2 点以上の場合、二塁走者が三塁走者の後の塁にいる走者の場合(二・三塁、満塁)でも、二塁走者が最も先の塁にいる走者の場合(二塁のみ、一・二塁)でも、二塁走者の生還は「有り」とみなした。守備側のエラーによりこの判別方法が当てはまらない場合もあるが、希少な例であり予測に影響を及ぼさないと判断した。二塁走者の生還有無は「生還 = 1」、「非生還 = 0」の二値変数として表される。

アウトカウントは、「0 アウト」、「1 アウト」、「2 アウト」の 3 クラスの値のいずれかを持つ。

イニングは、「1 回」から「9 回」の値のいずれかを持つカテゴリカル変数である。本研究では 10 回以降、つまり延長戦のデータを除外している。これは 10 回以降の各イニングにおいて、回帰分析を行うに十分なサンプル数が得られなかったための処置である。

塁状況は、3 桁の数字で表す。一塁走者がいる場合、3 桁目には「1」の値が割り振られる。同様に、二塁走者がいる場合、2 桁目には「2」の値が、三塁走者がいる場合、1 桁目には「3」の値が割り振られる。それぞれの塁に走者がいない場合、その桁には「0」の値が割り振られる。具体的には、走者二・三塁の場合、塁状況として「023」という値が与えられる。

点差は、Baseball Savant から取得できる攻撃側の得点 (bat_score) と守備側の得点 (fld_score) から算出する。式①が算出式である。

$$run_difference = bat_score - fld_score \quad \text{①}$$

正の値となる場合は、リードしている状態であり、負の値となる場合は、ビハインドの状態を表すこととなる。また、回帰分析を行うに十分なサンプル数が得られなかったため、点差は 10 点差以内に限定している。

二塁走者の走力は、Baseball Savant から取得できる sprint speed を用いる。2017 年シーズンに導入された sprint speed は、選手が 1 秒間に何フィート走ったかを測定することで、より正確に走力を数値化することを目的とした指標だ。sprint speed は、2017 年までは二塁から長打で生還した場合を除く、本塁打以外の 2 つ以上進塁した走塁だけで測定されていた。その後 2018 年に改善がなされ、先述の走塁に加えて、弱い打球における本塁から一塁への走塁も測定の対象となった。これらの測定のうち、ある選手の上位約 3 分の 2 の速度を記録した走塁を抽出し、その選手のシーズン平均スプリント速度として算出する。このシーズン平均スプリント速度が、sprint speed として用いられている

打球飛距離は、Baseball Savant から取得した段階では、単位が [ft] になっているので、[m] へと単位変換を行っている。ここで打球飛距離は、打球を守備者が捕球した位置ではなく、打球がファーストバウンドした地点を示している。

打球初速度は、Baseball Savant から取得した段階では、単位が [mile/h] になっているので、[km/h] へと単位変換を行っている。

鉛直打球角度は、地面と水平な打球角度を 0° とし、鉛直上向きの

角度が正の値、鉛直下向きの角度が負の値となっている。

水平打球角度に関する直接的なデータは、Baseball Savant からは取得できない。そのため、本研究では先行研究を参考に以下の式②で算出した (Petti,2017)¹⁸⁾。

$$\tan^{-1}\{(hc_x - 125.42) \div (198.27 - hc_y)\} \times 180 \div \pi \quad ②$$

ここで、hc_x と hc_y は Baseball Savant から取得でき、打球位置の座標を示す変数である。また、hc_x が 125.42 の値、hc_y が 198.27 の値の点がホームベースの座標として定義されている。

外野手の送球精度を直接的に表す変数は、Baseball Savant からは取得できない。そのため、本研究では、FanGraphs から取得できる外野手の送球による得点抑止力 (ARM) を、外野手の送球精度を表す変数として採用する。ARM の算出にはまず、ある外野手に対して図 3 の 4 場面において余分な進塁を「抑止」したのか「抑止できず」だったのかを分類する。

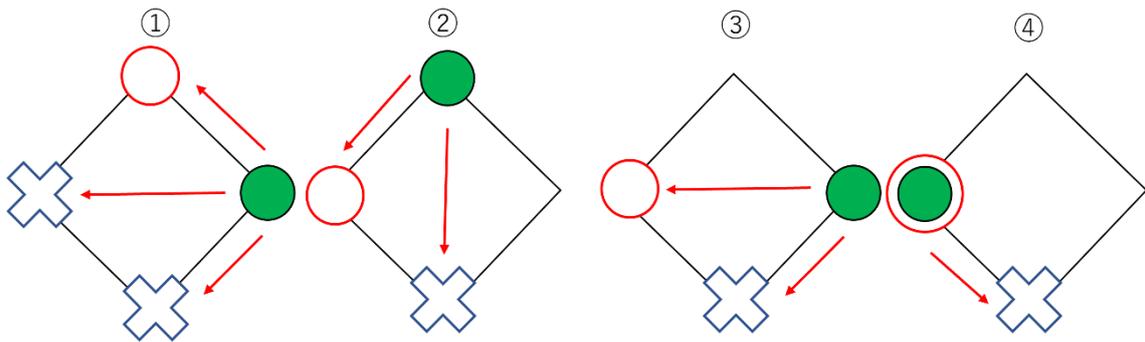


図 3 ARM の算出方法

緑丸は分類対象となる走者の置かれている塁を示している。赤丸は余分な進塁を「抑止」したと分類される進塁状況を示している。一方、青バツは余分な進塁を「抑止できず」と分類される進塁状況を示している。

①は一塁走者がいて単打を打たれた場面を表す。ただし、二塁走者がいる場合は計算対象から除外される。この時、一塁走者を二塁進塁のみに止めた場合、外野手は余分な進塁を「抑止」したと分類される。一方、一塁走者を三塁や本塁に進塁させてしまった場合、外野手は余分な進塁を「抑止できず」だったと分類される。②は二塁走者がいて単打を打たれた場面を表す。この時、二塁走者を三塁進塁のみに止めた場合、外野手は余分な進塁を「抑止」したと分類される。一方、二塁走者を本塁に進塁させてしまった場合、外野手は余分な進塁を「抑止できず」だったと分類される。③は一塁走者がいて二塁打を打たれた場面を表す。この時、一塁走者を三塁進塁のみに止めた場合、外野手は余分な進塁を「抑止」したと分類される。一方、一塁走者を本塁に進塁させてしまった場合、外野手は余分な進塁を「抑止できず」だったと分類される。④は三塁走者がいて外野フライを打ち上げられた場面を表す。この時、三塁走者を進塁なしに止めた場合、外野手は余分な進塁を「抑止」したと分類される。一方、三塁走者を本塁に進塁させてしまった場合、外野手は余分な進塁を「抑止できず」だったと分類される。これら4場面において「抑止」と「抑止できず」の8パターンそれぞれの発生数×変化した得点期待値を合計した値がARMとして算出される。外野手の送球による得点抑止力の値は、負に大きいほど抑止力が小さく、正に大きいほど抑止力が大きくなる。ここで、算出に用いられる得点期待値とは、表2のようなアウトカウントとランナーの組み合わせごとに、イニングの残りで得点され得る平均得点数を計算したものと定義されている(Weinberg,2014)¹⁹⁾。

表 2 得点期待値表 (Weinberg(2014)¹⁹⁾を改変)

runner_situation	0outs	1out	2outs
000	0.461	0.243	0.095
100	0.831	0.489	0.214
020	1.068	0.644	0.305
120	1.373	0.908	0.343
003	1.426	0.865	0.413
103	1.798	1.140	0.471
023	1.920	1.352	0.570
123	2.282	1.520	0.736

Note: runner_situation は塁状況を示す。「1」は一塁走者、「2」は二塁走者、「3」は三塁走者が出塁していることを示し、走者がいない塁は

ARM の他に外野手の送球精度を評価できる変数として、Baseball Savant から取得可能な Arm Strength も存在する。この変数は、その選手がそのシーズンの送球における最高球速を取得することができる。この変数を採用しなかった理由の一つとして、測定が開始されたのが 2020 年以降であり、本研究において 2 年分のデータしか取得できないことが挙げられる。また、球速だけでは送球の精度を十分に評価できないと考えたため、ARM を採用した。

研究対象となるシチュエーションは、走者が二塁に出塁している状況で、外野手が第一捕球者となるフライ打球の単打が放たれた場面に限定する。他の塁に走者が出塁しているかどうかは問わない。打席結果は単打のみを分析対象とする。これは、二塁打以上の結果であれば二塁走者はほぼ確実に生還でき、一方凡打であれば二塁走者はほぼ確実に生還できない為である。また、打球性質はフライに

限定する。本研究では、フライ打球は打球飛距離が 45m 以上の打球と定義する。45m という距離は、内外野の芝と土の分け目とほぼ同じラインとして設定した。研究対象となるシチュエーションでの二塁走者生還率は、64.1% となった。研究対象シチュエーションに限定した場合における、各変数のヒストグラムを図 4 から図 13 に示す。

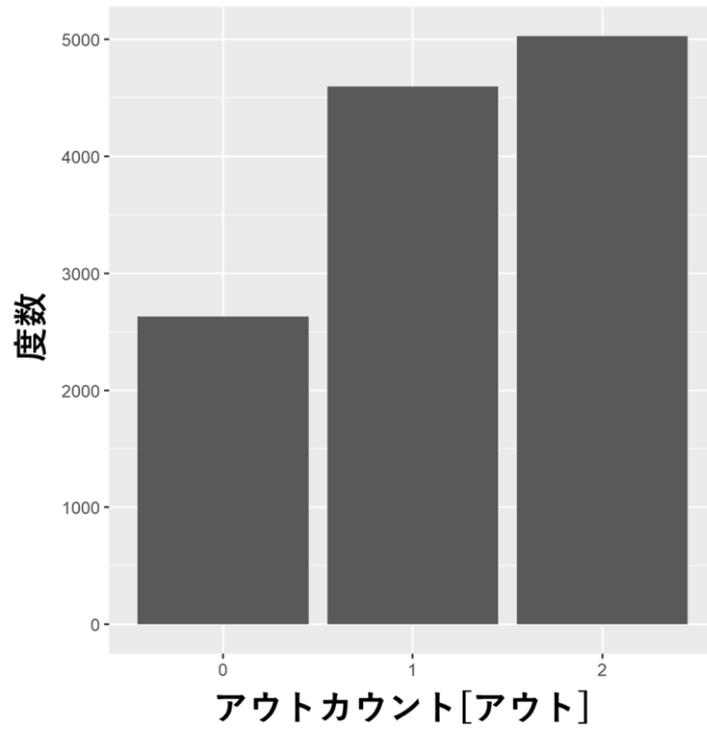


図 4 アウトカウントのヒストグラム

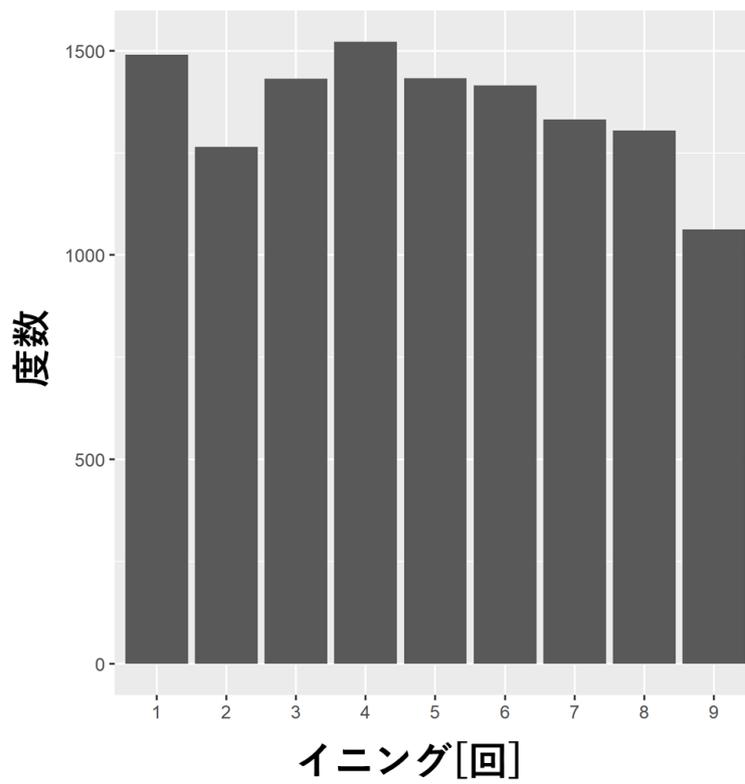


図 5 イニングのヒストグラム

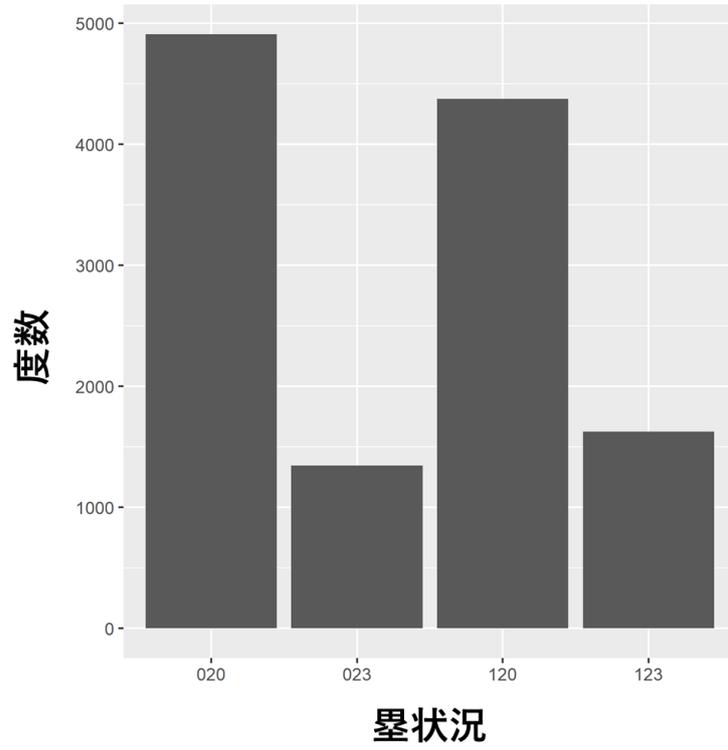


図 6 塁状況のヒストグラム

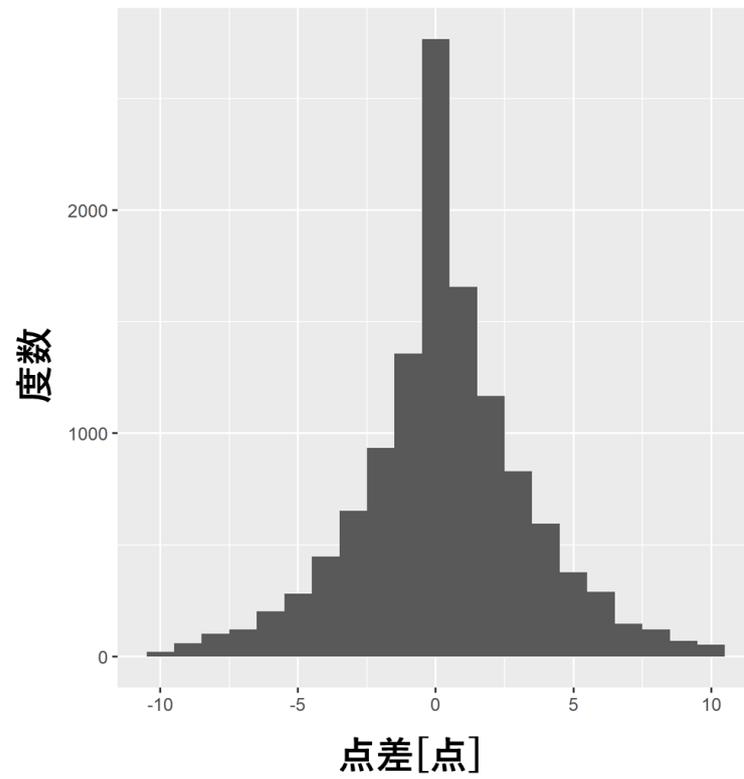


図 7 点差のヒストグラム

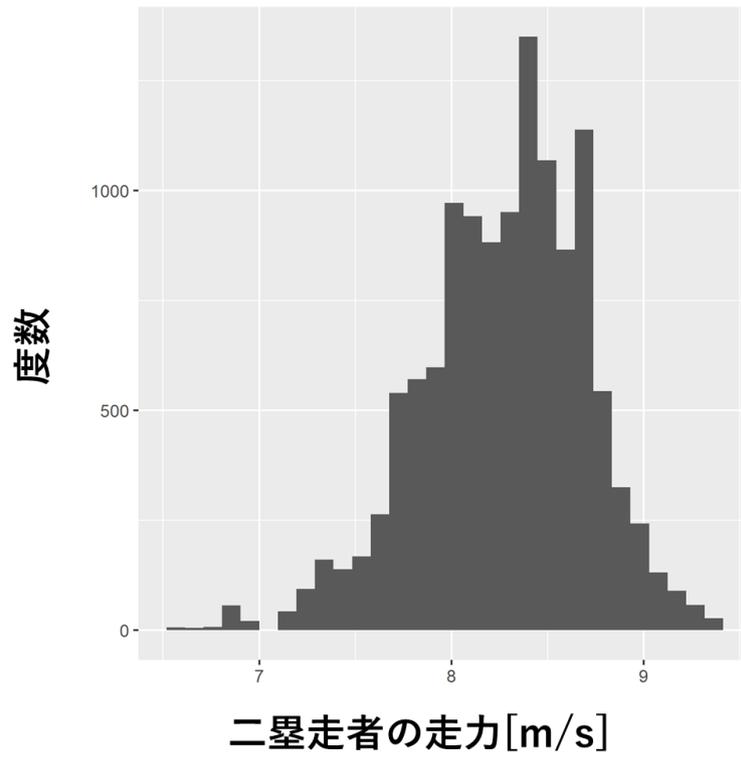


図 8 二塁走者の走力のヒストグラム

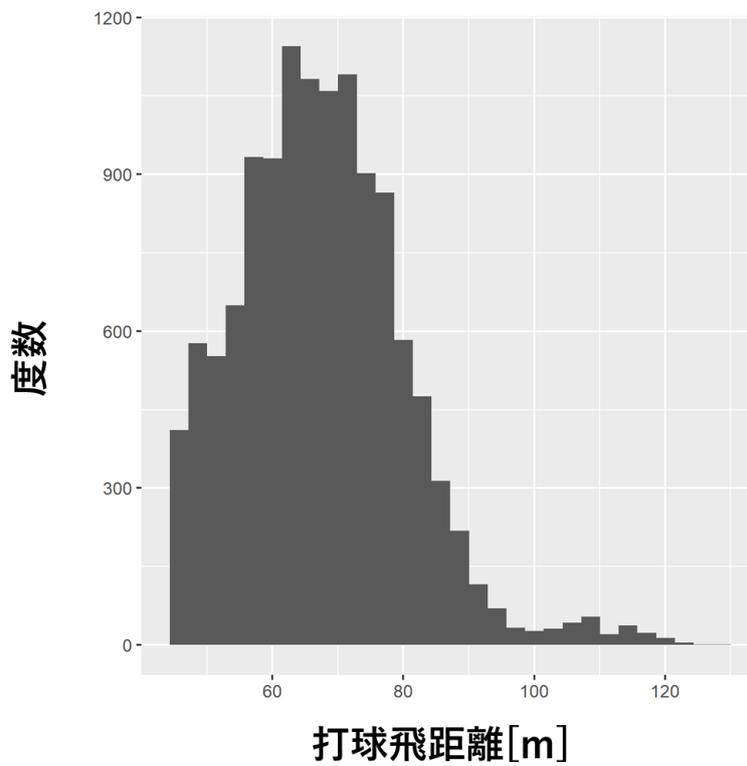


図 9 打球飛距離のヒストグラム

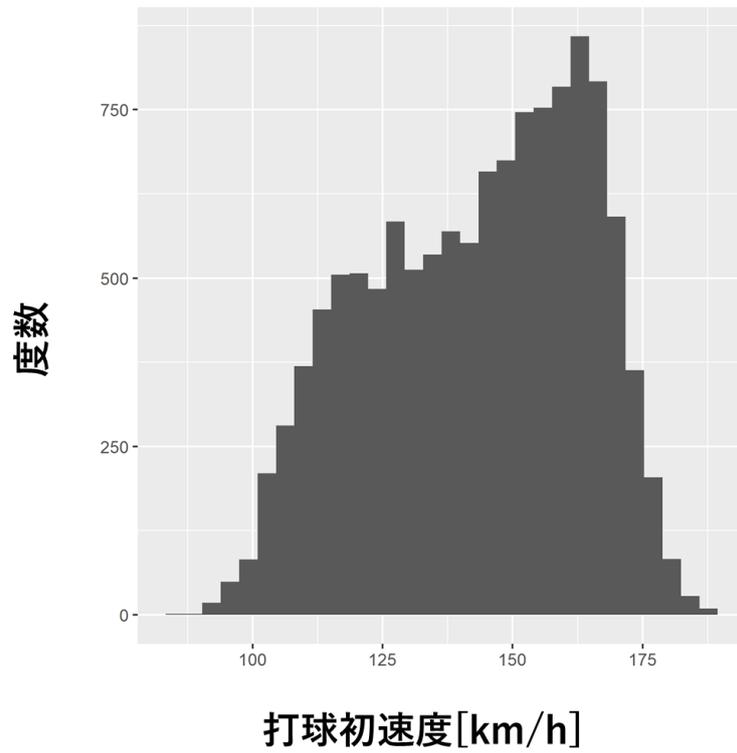


図 10 打球初速度のヒストグラム

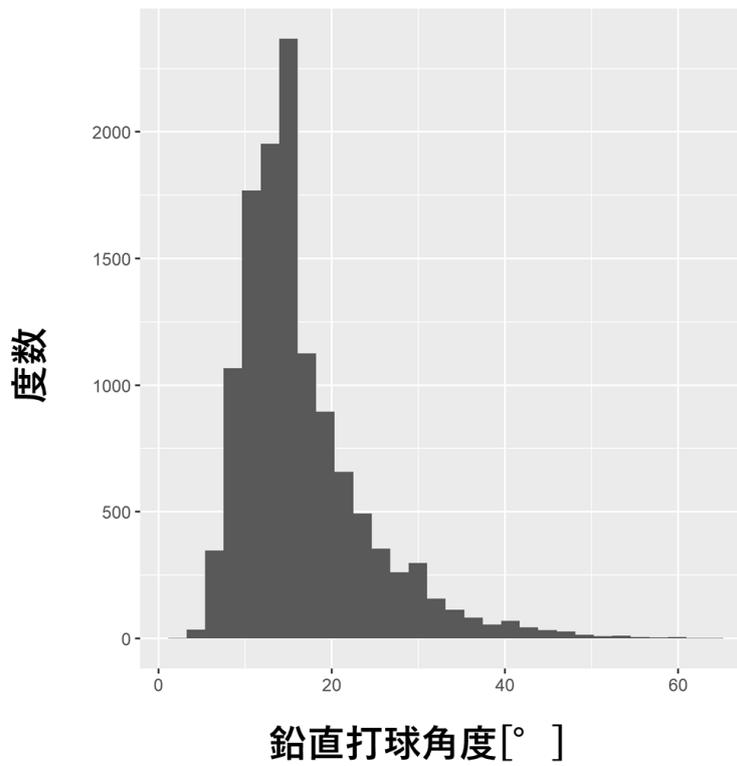


図 11 鉛直打球角度のヒストグラム

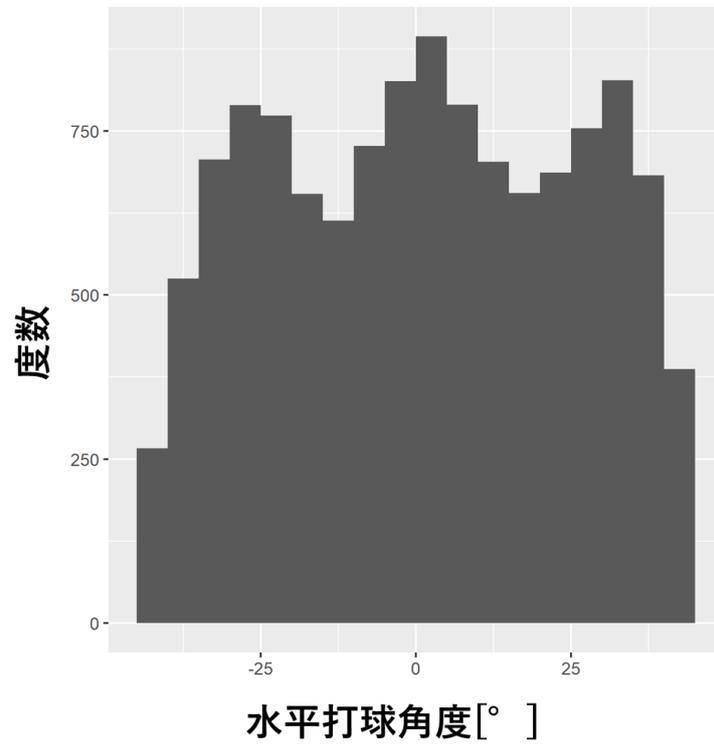


図 12 水平打球角度のヒストグラム

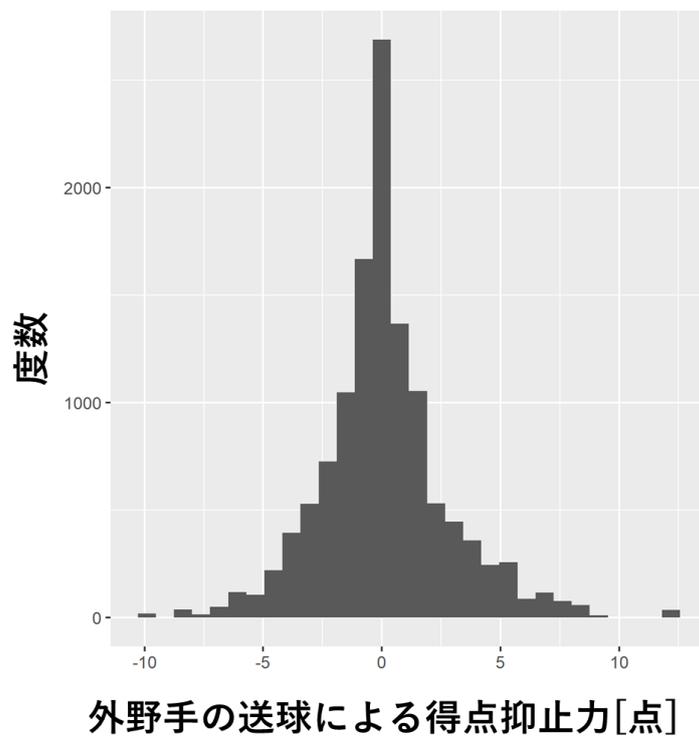


図 13 外野手の送球による得点抑止力のヒストグラム

モデリング手法

二塁走者が生還したか否かを目的変数として、12個の説明変数を用いた統計モデルを構築する。統計モデルの構築には、一般化加法モデル(GAM)(Hastie and Tibshirani,1986)²⁰⁾を用いた。GAMとは、説明変数ごとに目的変数との関係を推定し、それぞれの関係性を足し合わせ、全説明変数と目的変数の関係とする手法である。GAMを用いるのは、モデルの解釈性が高いことに加え、非正規分布の説明変数と目的変数にも対応できるためである。

予測に用いるデータ(n=12257)は、学習データ8割とテストデータ2割に分割した。モデリングに用いた変数は、表3に記したものになる。本研究ではmgcvパッケージのgam関数を用いて学習データに対するモデリングを行い、MuMInパッケージのMuMIn関数を用いて変数選択を行い、Rの基本関数であるpredict関数を用いてテストデータに対する予測を行った。gam関数において、目的変数は二項分布に、リンク関数はlogitに設定し、表3に記したように説明変数の回帰方法を設定した。ここで平滑化スプラインとは、平滑化パラメータを用いて離散的な説明変数と予測値の関係を連続的に補完する手法である。平滑化パラメータは、GCV(一般化交差検証)を用いて過学習を防いだ最適な値を推定している。MuMIn関数において、最も予測精度が高い説明変数の組み合わせを導き出すために、赤池情報量規準(AIC)(Akaike,1973)²¹⁾を用いた変数選択を行った。ここでAICとは、統計モデルの予測精度の高さを評価する指標のことである。AICは以下の式④によって定義される。

$$AIC = -2\log L + 2k \quad \text{④}$$

ここで L はモデルの最大尤度、 k はモデルに含まれる説明変数の数を示す。式④の第一項が小さくなるほど予測精度が高くなり、第二項の罰則項が小さくなるほど説明変数が少ないシンプルなモデルとなる。つまり AIC の値が小さくなるほど、予測精度が高く、かつ説明変数の数が少ないシンプルなモデルとなる。この変数選択において AIC が最も低くなったモデルをベストフィットモデルとみなし、本研究では分析した。

`predict` 関数においては、引数 `type` に `response` を指定した。`response` は 0 から 1 の値を返す。なお二塁走者生還成否予測正解率の計算は、`predict` 関数による予測値が 0.5 以上であれば「生還」を予測したとみなし、`predict` 関数による予測値が 0.5 未満であれば「非生還」を予測したとみなした。

決定要因となった変数の予測精度を比較するため、説明変数の AIC 寄与の大きさ (`AIC_dif`) を計算した。AIC 寄与の大きさは、先行研究 (Evans and Blackburn, 2020)²²⁾ を参考に以下の式⑤で計算した。

$$AIC_dif_i = AIC_{subi} - AIC_{best} \quad \text{⑤}$$

ここで AIC_dif_i は i 個目の説明変数の `AIC_dif` の値、 AIC_{subi} はベストフィットモデルから i 個目の説明変数を除いた AIC の値、 AIC_{best} はベストフィットモデルの AIC の値を示す。

表 3 モデリングに用いた変数表

変数名	単位	取得方法
二塁走者の生還有無(scoring_second)	点	算出
アウトカウント(outs_when_up)	アウト	Baseball Savant
イニング(inning)	回	Baseball Savant
塁状況(runner_situation)	—	算出
点差(run_difference)	点	算出
二塁走者の走力(sprint_speed)	m/s	Baseball Savant
打球飛距離(hit_distance_sc)	m	Baseball Savant
打球初速度(launch_speed)	km/h	Baseball Savant
鉛直打球角度(launch_angle)	°	Baseball Savant
水平打球角度(spray_angle)	°	算出
外野手の送球による得点抑止力(arm)	点	FanGraphs

[Note]変数名のカッコ内は変数名の略称となっている。二塁走者の生還有無はモデリングにおいて目的変数となるため、回帰方法は設定されていない。

結果

表 4 はベストフィットモデルに含まれた変数と、それぞれの変数の AIC 寄与の大きさを示している。ベストフィットモデルに含まれた変数はアウトカウント、打球飛距離、鉛直打球角度、二塁走者の走力、外野手の送球による得点抑止力、塁状況、点差、水平打球角度、打球の落下地点と捕球者の初期位置間の距離、打球の落下地点と捕球者の初期位置の成す角度で、含まれなかった変数はイニングのみだった。ベストフィットモデルの二塁走者生還成否予測正解率は 76.3% であった。また二塁走者生還予測正解率は 80.2% で、二塁走者非生還予測正解率は 68.7% となった。

表 4 ベストフィットモデルに含まれたそれぞれの説明変数の AIC 寄与の大きさ

変数名	AIC寄与
アウトカウント(outs_when_up)	2254
塁状況(runner_situation)	29
点差(run_difference)	51
二塁走者の走力(sprint_speed)	259
打球飛距離(hit_distance_sc)	73
打球初速度(launch_speed)	50
鉛直打球角度(launch_angle)	139
水平打球角度(spray_angle)	968
外野手の送球による得点抑止力(arm)	157

ベストフィットモデルに含まれた各説明変数と予測値の関係性を、図 14 から図 22 に示す。横軸は各説明変数の値、縦軸は予測値となっている。縦軸のラベルに記されている $s()$ は平滑化スプラインによる予測を行ったことを示しており、 $f()$ はロジスティック回帰が行われたことを示している。また、 $s()$ 内に記されている数字は平滑化スプラインの次数を決める有効自由度の値となっている。有効自由度の値が大きいほど、平滑化スプラインによる予測値は複雑な曲線となる。横軸上内向きに存在する線は、説明変数の値の密度を表していて、色が濃いほど密度が高いという関係になっている。グラフ内の実線は予測値の平均値(この論文内では単に予測値と呼ぶ)の推移となっており、破線はその値の 95% 信頼区間を表す。予測値が正の値だと生還率を高める影響をその変数を与えており、負の値だと生還率を低下させる影響をその変数を与えていると解釈できる。また、正の値・負の値それぞれの絶対値が大きくなるほど、与える影響も大きくなる。

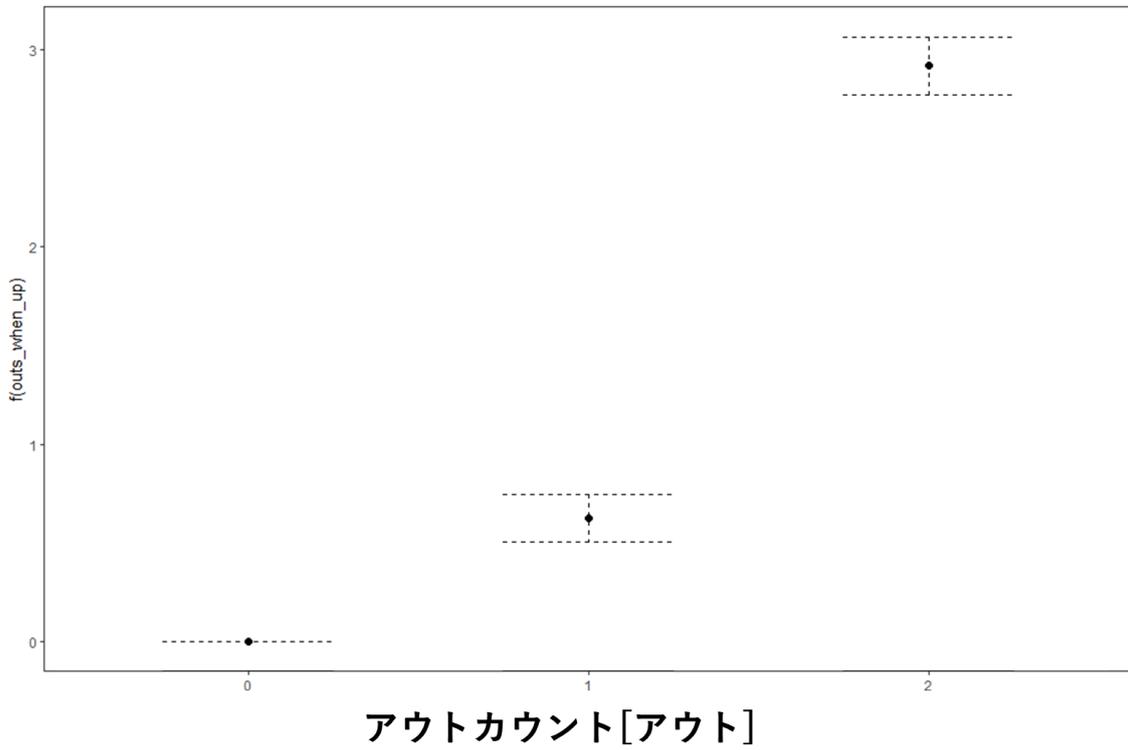


図 14 アウトカウントの値と予測値の関係

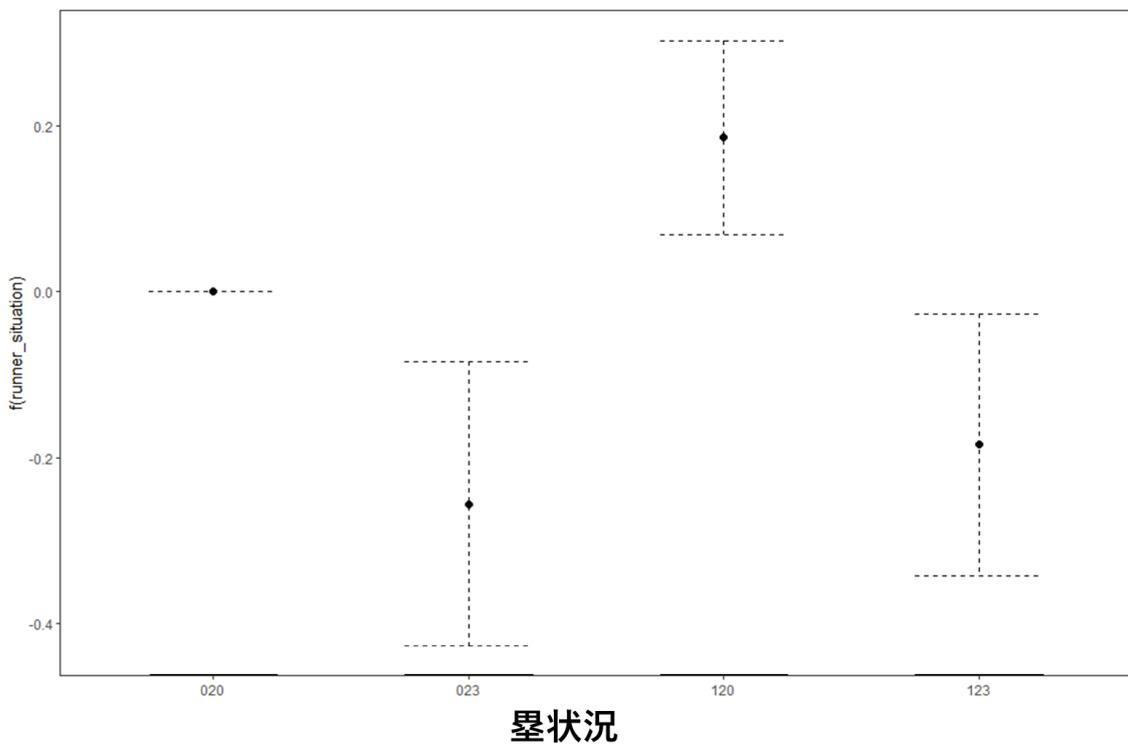


図 15 塁状況の値と予測値の関係

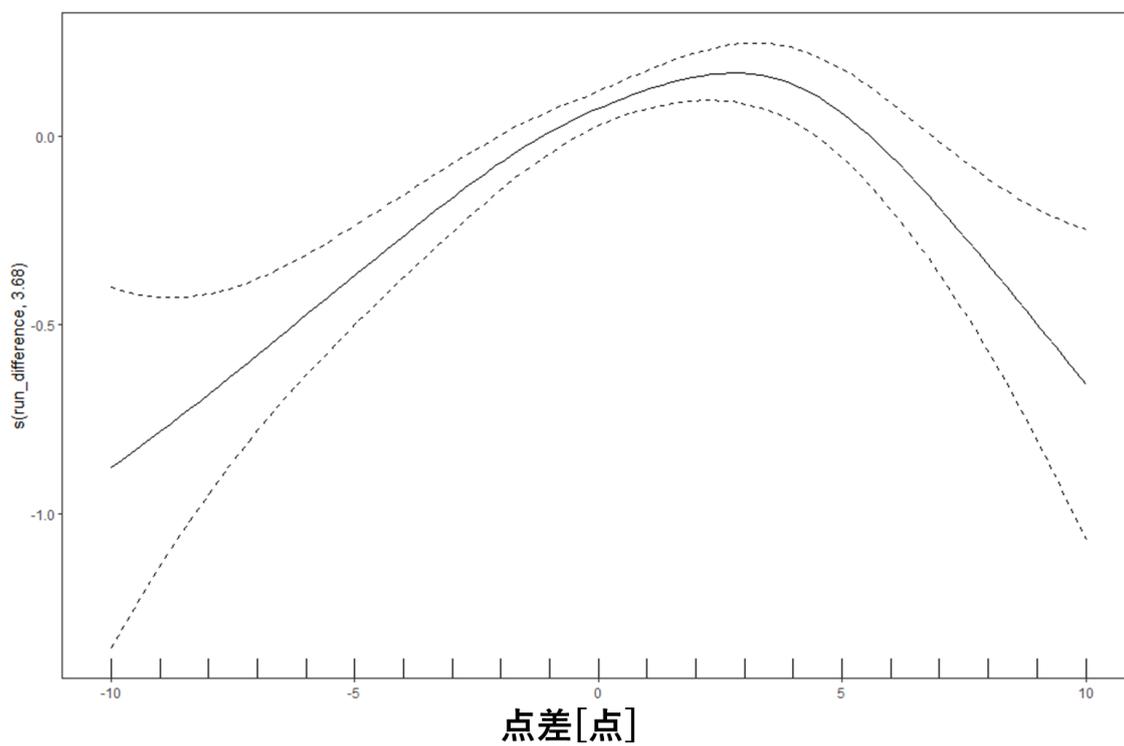


図 16 点差の値と予測値の関係

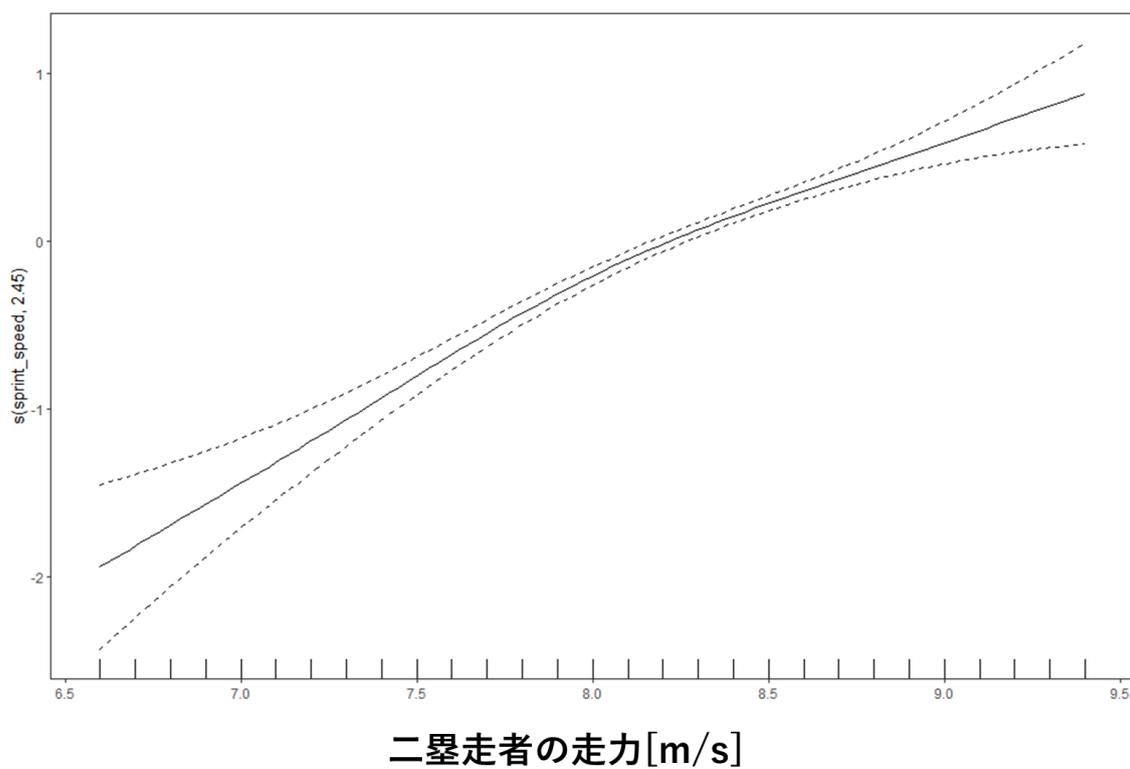


図 17 二塁走者の走力の値と予測値の関係

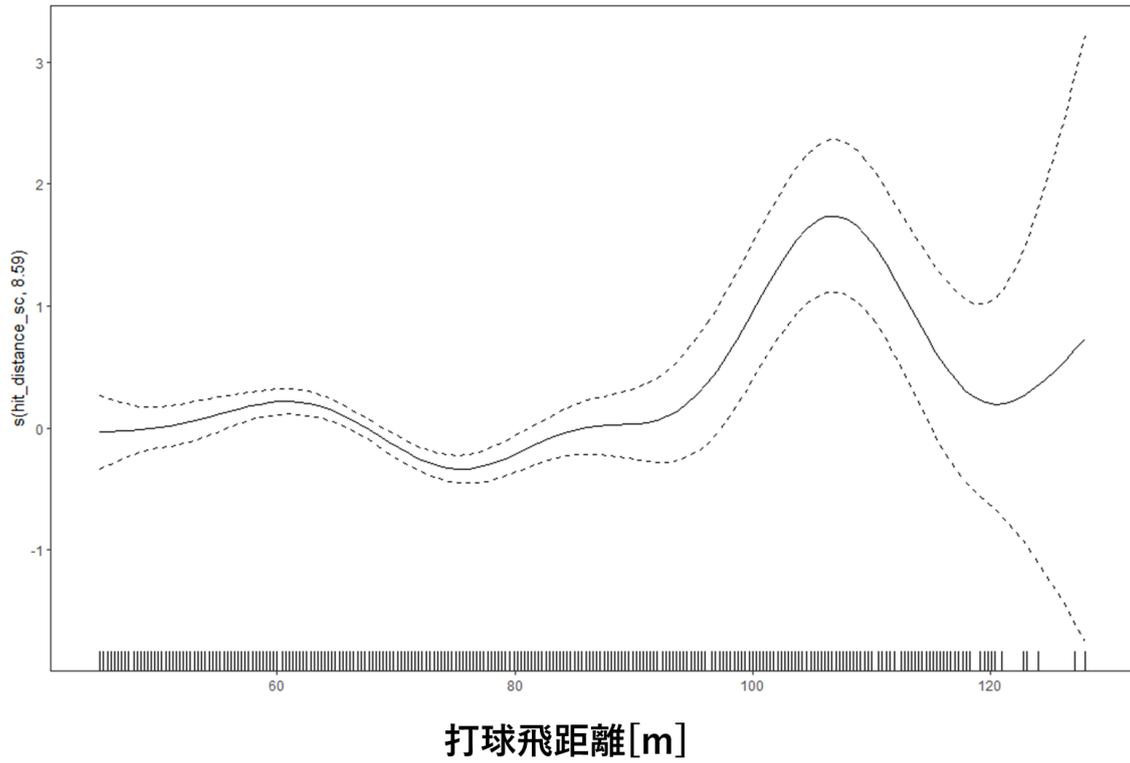


図 18 打球飛距離の値と予測値の関係

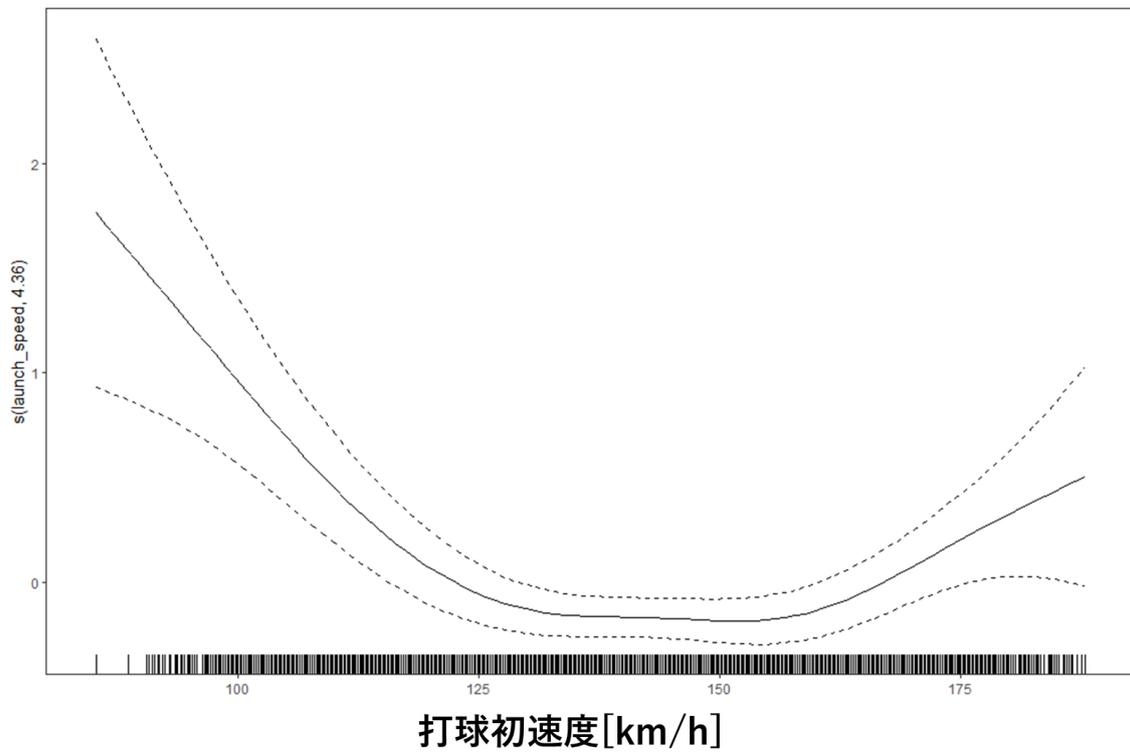


図 19 打球初速度の値と予測値の関係

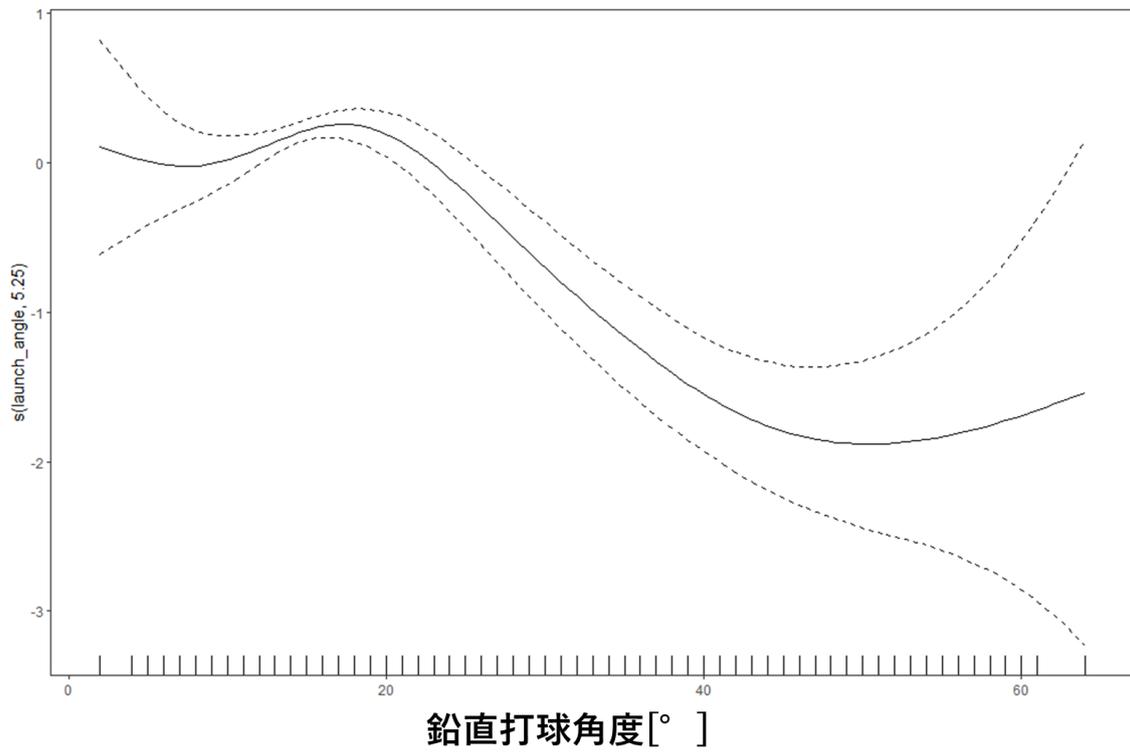


図 20 鉛直打球角度の値と予測値の関係

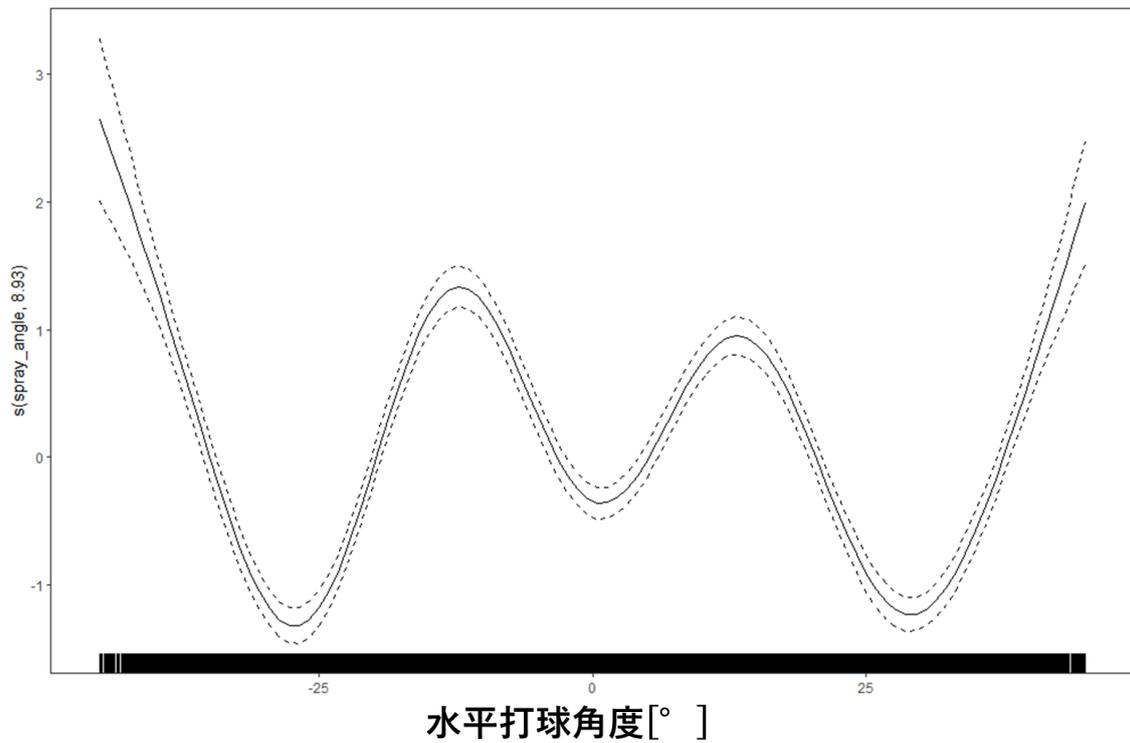


図 21 水平打球角度の値と予測値の関係

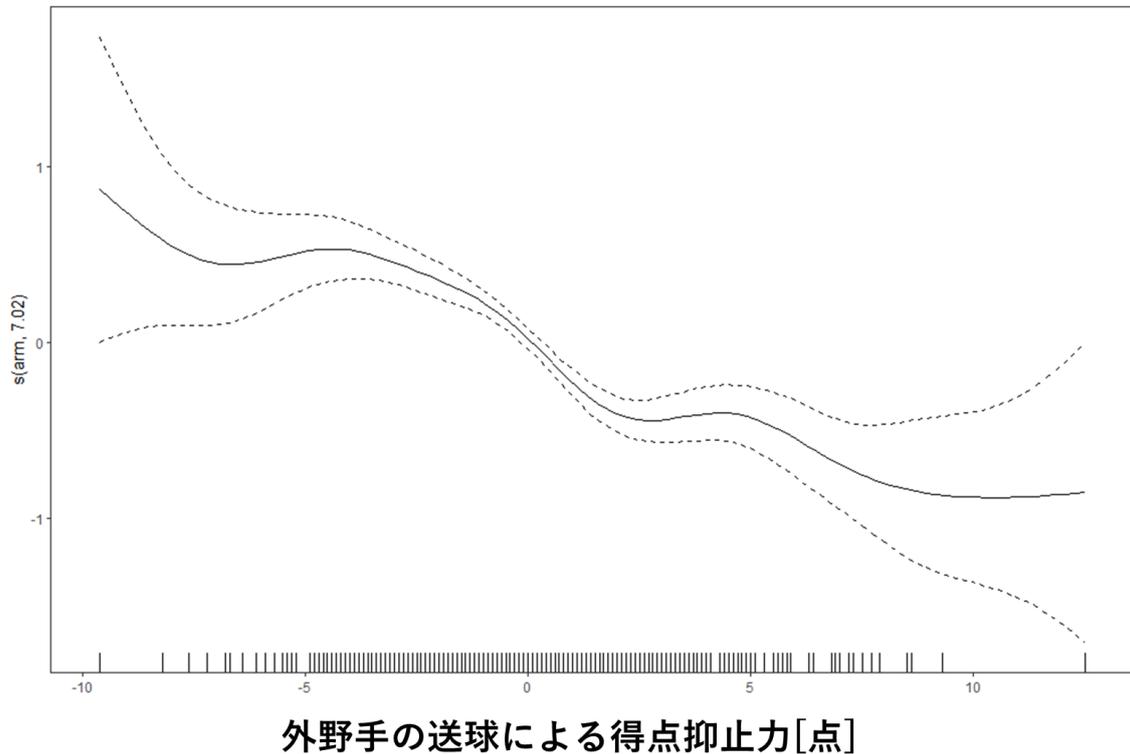


図 22 外野手の送球による得点抑止力の値と予測値の関係

図 14 はアウトカウントの値と予測値の関係性について示したグラフである。0 アウトを基準の 0 としている。1 アウトで予測値が 1 付近の正の値を取っている。また、2 アウトでは予測値が 3 付近の正の値を取っている。この変数の AIC 寄与の大きさの値は 2254 で、ベストフィットモデルに含まれた変数の内 1 番目の大きさだった。

図 15 は塁状況の値と予測値の関係性について示したグラフである。走者二塁のみの塁状況を基準の 0 としている。走者二・三塁と満塁で予測値が負の値を取っている。一方、走者一・二塁では予測値が正の値を取っている。しかし、どの塁状況においても予測値は 0 付近の小さい値を取っている。この変数の AIC 寄与の大きさの値は 29 で、ベストフィットモデルに含まれた変数の内 9 番目の大きさだった。

図 16 は点差の値と予測値の関係性について示したグラフである。値が 3 点付近の時、予測値が 0 付近で頂点となる。そして、値が変動するにつれて予測値も低くなっている。この変数の AIC 寄与の大きさの値は 51 で、ベストフィットモデルに含まれた変数の内 7 番目の大きさだった。

図 17 は二塁走者の走力の値と予測値の関係性について示したグラフである。値が大きくなるにつれて、ほぼ線形に予測値が高くなっている。値が 8.3m/s までは負の区間で予測値が高くなっており、8.3m/s 以上は正の区間で予測値が高くなっている。この変数の AIC 寄与の大きさの値は 259 で、ベストフィットモデルに含まれた変数の内 3 番目の大きさだった。

図 18 は打球飛距離の値と予測値の関係性について示したグラフである。値が 65m 以下では、予測値は負の値から 0 付近まで高くなっている。値が 65m から 110m の間では、予測値は 0 付近でほぼ水平に推移している。値が 110m から 120m へと変化すると、-2 まで予測値が低くなっている。この変数の AIC 寄与の大きさの値は 73 で、ベストフィットモデルに含まれた変数の内 6 番目の大きさだった。

図 19 は打球初速度の値と予測値の関係性について示したグラフである。値が 125km/h 以下では、正の値の区間で 0 付近まで予測値が低くなっている。値が 125km/h から 160km/h の間では、予測値は 0 付近でほぼ水平に推移している。値が 160km/h 以上になると、予測値が高くなっている。この変数の AIC 寄与の大きさの値は 50 で、ベストフィットモデルに含まれた変数の内 8 番目の大きさだった。

図 20 は鉛直打球角度の値と予測値の関係性について示したグラフである。値が 20° までの区間では、予測値は 0 付近でほぼ一定とな

っている。値が 20° から 50° の間では、予測値は負の区間で低くなっている。値が 50° 以降では、負の区間で予測値が高くなっている。この変数の AIC 寄与の大きさの値は 139 で、ベストフィットモデルに含まれた変数の内 5 番目の大きさだった。

図 21 は水平打球角度の値と予測値の関係性について示したグラフである。値が 0° 付近では予測値も 0 近くになり、予測値の変動は 0° を中心に左右対称となっている。値が $\pm 15^{\circ}$ と $\pm 45^{\circ}$ 付近の時、予測値は正の区間で高くなっている。値が $\pm 30^{\circ}$ 付近の時、予測値は負の区間で低くなっている。この変数の AIC 寄与の大きさの値は 968 で、ベストフィットモデルに含まれた変数の内 2 番目の大きさだった。

図 22 は外野手の送球による得点抑止力の値と予測値の関係性について示したグラフである。値が 0 点以下の時、正の区間で概ね右肩下がりに予測値が低くなっている。そして、値が 0 点付近で予測値も 0 近くになる。値が 0 点以上の時、今度は負の区間で右肩下がりに予測値が低くなっている。この変数の AIC 寄与の大きさの値は 157 で、ベストフィットモデルに含まれた変数の内 4 番目の大きさだった。

考察

二塁走者が本塁生還する可能性を規定する力学的要因を系統的に示す deterministic model において定義された決定要因を用いた統計モデルを作り、予測精度を高める影響が大きい要因を導出することを、本研究の目的とした。結果として、deterministic model において定義された決定要因の内、イニングを除いた 9 つがベストフィットモデルに含まれた。そして、予測精度を高める影響が大きい順にアウトカウント、水平打球角度、二塁走者の走力、外野手の送球による得点抑止力、鉛直打球角度、打球飛距離、点差、打球初速度、塁状況となった。

ベストフィットモデルには 9 つの変数が選択され、二塁走者生還成否予測正解率は 76.3% となった。4 回に 1 回の割合で誤った予測が起こる予測モデルとなった。二塁走者生還予測正解率は 80.2% で、二塁走者非生還予測正解率は 68.7% となった。本研究で選択されたベストフィットモデルは、走者を生還させる予測は正解率が高く得意であり、走者が生還できない予測は正解率が低く苦手という傾向が見られた。図 23 は AIC 寄与の大きさの総和における各変数の割合を示した円グラフとなっている。AIC 寄与の大きさの総和の内、アウトカウントが 50% 以上を占めており、アウトカウント以外の変数全てを足し合わせた数値より大きいことがわかった。水平打球角度が全体の 4 分の 1 を占めており、アウトカウントと水平打球角度だけで予測精度を高める要因の 8 割を占めていた。走者が二塁に出塁しており、外野手が第一捕球者となった単打が発生した場面において、三塁コーチは図 23 の割合で走塁判断基準の意識づけをす

ることで、より正確な走塁判断・指示が可能になると考えられる。

AIC寄与の大きさの総和における各変数の割合

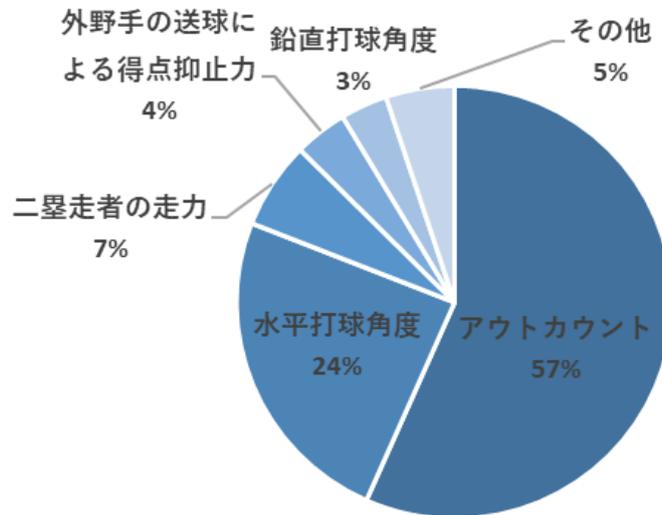


図 23 AIC 寄与の大きさの総和における各変数の割合

アウトカウントは、生還率に最も大きな影響を及ぼす決定要因となった。1アウトの場合生還率を高める影響を与えている。これはチェンジになるまでの攻撃の機会が1アウト分少ないため、0アウト時より積極的に本塁へ突入するからだと考えられる。裏を返せば、0アウト時は本来なら際どいタイミングだが生還が可能なプレーにおいて、次打者にもチャンスの場면을継続させるために、二塁走者を本塁へ突入させていないと考えられる。そのため、必要以上に慎重な走塁を指示した結果、0アウト時に取れた1点をみすみす逃してしまう場面が多くなったのではないかと思われる。また、2アウトの場合さらに生還率を高める影響が大きくなる。これは放たれたフライ打球が、ヒットになるかアウトになるかを見分ける打球判断の問題がなくなるためであると考えられる。アウトカウントが0アウトもしくは1アウトの時、フライ打球が守備側に直接捕球されアウト

となった場合、二塁走者は守備側が二塁へ送球するより先に二塁へ帰塁しなければならない。このプレーを怠ると、二塁走者が二塁に戻れずアウトになってしまい、打者のアウトに加え無駄なアウトを献上してしまうこととなる。そのため、打球がヒットになる確信がなければ、二塁走者は本塁へスタートすることができない。しかし2アウトであれば、フライ打球がアウトになってもそこで攻守が入れ替わる。そのため、二塁走者が二塁へ帰塁する必要がなくなり、打球が放たれた瞬間に本塁へのスタートを切れるのである。

塁状況は、生還率に10番目に大きな影響を及ぼす決定要因となった。塁状況が「走者二・三塁」、「満塁」といった走者が三塁にいる場合、生還率を低下させる影響を与えていた。これは、先の塁に走者がいることで早めのスタートが切りにくいからだと考えられる。

「走者一・二塁」の場合、生還率を高める影響を与えていた。これは、一塁に走者がいることで二塁走者がアウトになっても得点圏に走者を置き続けられるため、二塁走者が積極的に本塁を狙いやすいからだと考えられる。

点差は、生還率に6番目に大きな影響を及ぼす決定要因となった。点差は、3点リードから値が変動するにつれて生還率を低下させる影響を与えている。これは、点差が開くにつれて走者を貯めて大量点を狙う傾向があり、ビハインドだとその傾向がより強まるのではないかと考えられる。

二塁走者の走力は、生還率に2番目に大きな影響を及ぼす決定要因となった。値が大きくなるにつれてほぼ線形に生還率を高める影響を与えていた。また、予測値の正負は8.3m/sで反転している。ここで8.3m/sという速度は、MLB2016～2021年までにおける全選手

のスプリント速度の上位 56.6%にあたる。そのため、二塁走者が MLB の平均以上の走力を持つだけで、生還に有利に働くと示唆された。

打球飛距離は、生還率に 11 番目に大きな影響を及ぼす決定要因となった。値が 65m 以下までは、生還率を低下させる影響を与えており、値が大きくなるほど低下させる影響が小さくなった。65m から 110m の間では、生還率への影響はほとんどなかった。また値が 110m から 120m へと変化すると、生還率を低下させる影響を再び与えた。ホームから 110m 以上離れた位置は外野フェンス際の打球となる。しかし、外野フェンス付近への単打は、サンプル数が少なく、本研究で生還率が低下している理由として、過学習となっている可能性が考えられる。

打球初速度は、生還率に 6 番目に大きな影響を及ぼす決定要因となった。値が 125km/h 以下では、生還率を高める影響を与えており。その影響は値が大きくなるにつれて小さくなっていった。値が 125km/h から 160km/h の間では、生還率への影響はほとんどなかった。値が 160km/h 以上になると、生還率を高める影響を与えていた。打球初速度が小さければ、外野手が捕球するまでの時間が長くなり、二塁走者はその時間で本塁への距離を縮めることができるため生還が容易となる。打球初速度が大きくなるにつれ、その関係は弱まるが、160km/h 以上になると、再び生還が容易となる。打球初速度が大きすぎる打球は、外野手が捕球するまでは速いが、その分外野手の初期位置と捕球位置が近くなり、捕球位置と本塁までの距離が大きくなるため、送球距離が長くなり生還しやすくなるのだと推察される。

鉛直打球角度は、生還率に 5 番目に大きな影響を及ぼす決定要因となった。値が 20° 以下では、生還率への影響はほとんどなかった。値が 20° から 50° の区間では、値が大きくなるにつれて生還率を低下させる影響を与えていた。鉛直打球角度が大きいほど、打球の飛翔時間も大きくなる。打球の飛翔時間が長ければ二塁走者が本塁到達するまでの猶予時間が伸びるため、生還率が高くなると考えるほうが自然な発想に思われる。この一見不可解な関係性には、アウトカウントと同様に、フライ打球におけるヒットかどうかを見分ける打球判断の問題が関わっていると考えられる。そして打球の飛翔時間が長くなるにつれて、外野手が打球を直接捕球してアウトにする可能性も高くなるため、打球判断はさらに難化していく。そのため、鉛直打球角度が大きくなるにつれて生還率を低下させる効果を与えているのだと考えられる。

水平打球角度は、生還率に 3 番目に大きな影響を及ぼす決定要因となった。値が $\pm 30^\circ$ 付近の時、生還率を低下させる影響を与えていた。また値が $\pm 15^\circ$ と $\pm 45^\circ$ 付近の時、生還率を高める影響を与えていた。ここで、図 24 は MLB の 2021 年における各ポジションの守備位置平均を示した図である。外野手の守備位置を見ると、センターの選手はほぼ 0° 、レフトの選手はほぼ -30° 、ライトの選手はほぼ 30° となっている。この図と照らし合わせると、 0° と $\pm 30^\circ$ 付近は外野手の正面の打球、 $\pm 15^\circ$ 付近は左中間、右中間の打球、 $\pm 45^\circ$ 付近は左翼線、右翼線の打球となる。 $\pm 30^\circ$ 付近は外野手の正面の打球となり、打球を捕球するまでの時間が短く、助走をつけた送球体勢も作りやすいため、生還が困難になるのだと考えられる。 0° 付近も外野手の正面の打球となるが、生還率を低くする影響は与えて

いなかった。これは、ライトレフトよりセンターは定位置が後方にあり、捕球までの時間が長くなる分、正面の打球というメリットを相殺しているのだと考えられる。±15°付近は左中間、右中間の打球、±45°付近は左翼線、右翼線の打球となり、打球を捕球するまでの時間が長く、助走をつけた送球体勢も作りにくいいため、生還が容易になるのだと考えられる。

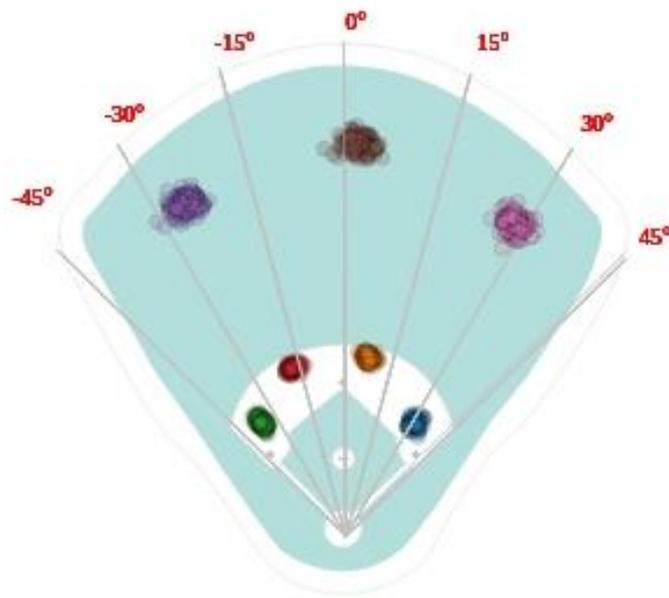


図 24 MLB2021 年における野手のポジショニングデータ

ポジショニングデータは Baseball Savant から取得した。このデータは、走者無し・走者一塁のみ以外の塁状況における、右打者に対する守備シフトを敷いていないときのポジショニングデータとなっている。

外野手の送球による得点抑止力は、生還率に 4 番目に大きな影響を及ぼす決定要因となった。値が負の区間では生還率を高める影響を与えており、正の区間では生還率を低下させる影響を与えている。このことから、捕球した外野手の送球による得点抑止力が高くなるほど、二塁走者の生還は困難になると考えられる。

リミテーション

本研究では、打球性質をフライ性の打球に限定したため、本研究の成果はゴロ性の打球に対して活用した場合の予測精度は、本研究結果で提示された76.3%より低くなる可能性がある。フライ性の打球に限定した理由として、本研究で用いた打球飛距離は、打球がファーストバウンドした位置を示しており、ゴロ性の打球の打球飛距離が解釈しにくい点が挙げられる。ゴロ性の打球だとその多くでホームベース付近が打球の落下地点となっており、各ゴロ性の打球飛距離の差を認識することが難しいと考えられる。ゴロ性の打球の生還率は73.5%となっており、フライ性の打球における生還率64.1%よりも高い。ゴロ性の打球を加える、もしくはゴロ性の打球に限定することで、予測精度が高いモデルが作成できるものと思われる。

二塁走者生還予測に関わるが、本研究では取得できなかったために考慮できなかった変数がいくつかある。文献において、三塁コーチが考慮する要因として、「よくあるケースが、走者二塁で、シングルヒットが出たケースである。打球の速さ、外野手の肩の強さ、捕球体勢、走者の足の速さ、そして球場の風の向き、強さ、グラウンドの土が乾いているか、湿っているか、芝生の向き、これらを瞬時に判断材料に入れて、走者を三塁に止めるか、本塁まで突入させるかの指示を出さなければならない。」と述べられている(澤宮,2018)³⁾。本研究では、外野手の捕球体勢、球場の風の向き、強さ、グラウンドの土が乾いているか、湿っているか、芝生の向きといったデータを取得することができなかった。また、二塁走者のリードの大きさに関しては、かつてはBaseball Savantから取得するこ

とができただが、現在は取得することができないようになっている。
将来的に球場におけるデータ収集がさらに進化し、それらのデータ
が誰にでも取得可能となれば、より予測精度の高いモデリングが可
能になると考えられる。

実践における活用

本研究における成果は、二塁走者の走塁を指示する三塁コーチによって活用されることが特に望まれ、勝負の分かれ目となる1点を、奪い取る可能性を高めるための方策を提示できると考えられる。二塁走者が単打によって本塁へ突入する場合、二塁走者が打球方向へよそ見をせず走塁するほうが速いことは自明である。そのため、本研究における成果を最大限活用する場合は、三塁コーチに二塁走者の走塁を全権委任すべきだと考える。

アウトカウント、二塁走者の走力や外野手の送球による得点抑止力は、プレー開始までに予め把握しておくことが、三塁コーチが精度の高い走塁指示をするために必要となる。また、水平打球角度や鉛直打球角度の値の変化を迅速かつ正確に察知することが、三塁コーチが精度の高い走塁指示をするために重要だと示唆された。これらの値は実践において、三塁コーチがコーチボックスの位置から瞬時に把握することは難しい。また、打球がヒットになるかアウトになるかの打球判断を二塁走者が行う必要がない2アウトの場面では、生還率が大きく高まることが示された。これは、2アウトであるために走者が打球のインパクトの直後に本塁へのスタートを躊躇なく切れるためと考えられる。アウトカウントが0や1の場合においても二塁走者が早いタイミングでスタートを切り、本塁生還率を高めるためには、選手及び三塁コーチは普段の練習から、水平打球角度、鉛直打球角度や打球がヒットになるかアウトになるかどうかを、迅速かつ正確に判断する練習を行う必要があると考えられる。8.3m/s以上の走力を持つ選手が二塁走者であれば、生還率を高める

ことが示されたことから、三塁コーチはチームの選手に対し、最高速度 8.3m/s 以上の走力を身に付けるよう指導すべきと考えられる。

三塁コーチは「走者が、無事に生還すれば、好走した選手の手柄にされ、アウトになれば、『壊れた信号機』と呼ばれ、三塁ベースコーチが非難の矢面に立たされる。」精神的に過酷な役割であり、その過酷さから走塁を指示する手が回らなくなる、いわゆるイップスに陥るコーチも珍しくない(澤宮,2018)³⁾。様々な判断材料がある中、ある程度の優先順位をつけて走塁判断を行えることは、正常な判断を続けるうえで非常に重要になると考えられる。本研究は三塁コーチを精神的に支える端緒となるだろう。

結論

deterministic model を構築し、二塁走者の生還率の決定要因として 12 の変数を定義した。実測値を用いたベストフィットモデルにはその内の 11 つの変数が選択され、その予測正解率は 76.3% となった。生還率に大きな影響を及ぼす要因として、アウトカウント、二塁走者の走力、水平打球角度、外野手の送球による得点抑止力、鉛直打球角度があげられた。この結果は、二塁走者の生還率を高めるためには水平打球角度と鉛直打球角度を元に打球がヒットになるか否かを、迅速かつ正確に判断する能力が、三塁コーチには必要であることを示唆するものである。

謝辞

本研究にあたり、指導教官として終始多大なご指導を賜った、矢内利政教授に深く感謝いたします。また、松井泰二教授、並びに深見英一郎准教授には、本論文の作成にあたり、副査として適切なご助言を賜りました。心より御礼申し上げます。最後に、矢内研究室の皆様には、本研究にあたり多大なご助言とご協力をいただきました。深く感謝いたします。

参考文献

- 1) Foster, J. (2020). What Actually Makes a Curveball Effective?
[URL] <https://community.fangraphs.com/what-actually-makes-a-curveball-effective/>. (アクセス日: 2023年1月10日)
- 2) Kato, M., Yanai, T. (2022). Launch fly balls for better batting statistics: Applicability of “fly-ball revolution” to Japan’s professional baseball league. *International Journal of Performance Analysis in Sport.*, 22(3): 437-453.
- 3) 澤宮 優 (2018). 三塁ベースコーチが野球を変える. 河出文庫. 東京
- 4) 成美堂出版 (2004). バッティング 守備・走塁 トッププレイヤーの技術. 東京
- 5) Hay, J.G., & Reid, J.G. (1988). *Anatomy, mechanics, and human motion*. Prentice Hall.
- 6) R Core Team (2020). *R: A Language and Environment for Statistical Computing* [URL] <https://www.R-project.org/>. (アクセス日: 2023年1月10日)
- 7) Petti, B. (2021). *baseballr: Functions for acquiring and analyzing*

baseball data. R package version 0.8.4.

8) Csárdi,G., Hester,J., Wickham,H., Chang,W., Rstudio,
Morgan,M., Tenenbaum,D., Mango Solutions(2021). remotes: R
Package Installation from Remote Repositories, Including 'GitHub'.
R package version 2.3.0. [URL] [https://CRAN.R-
project.org/package=remotes](https://CRAN.R-project.org/package=remotes) (アクセス日:2023年1月10日)

9) Wickham,H., RStudio(2019). Welcome to the tidyverse. Journal
of Open Source Software, 4(43), 1686, [URL]
<https://doi.org/10.21105/joss.01686> (アクセス日:2023年1月10
日)

10) Revelle,W. (2022) psych: Procedures for Personality and
Psychological Research, Northwestern University, Evanston,
Illinois, USA, [URL] <https://CRAN.R-project.org/package=psych>
Version = 2.2.5. (アクセス日:2023年1月10日)

11) Schauberger,P., Walker,A., Braglia,L., Sturm,J.,
Garbuszus,J.M., & Barbone,J.M. (2021). openxlsx: Read, Write and
Edit xlsx Files. R package version 4.2.5. [URL] [https://CRAN.R-
project.org/package=openxlsx](https://CRAN.R-project.org/package=openxlsx) (アクセス日:2023年1月10日)

12) Wood,S.N. (2011) Fast stable restricted maximum likelihood
and marginal likelihood estimation of semiparametric generalized

linear models. *Journal of the Royal Statistical Society (B)* 73(1):3-36

13) Fasiolo, M., Nedellec, R., Goude, Y., Capezza, C., & Wood, S.N. (2021). *mgcViz: Visualisations for Generalized Additive Models* [URL] <https://cran.r-project.org/web/packages/mgcViz/index.html> (アクセス日:2023年1月10日)

14) Bartoń, K. (2022). *MuMIn: Multi-Model Inference*. R package version 1.46.0. [URL] <https://CRAN.R-project.org/package=MuMIn> (アクセス日:2023年1月10日)

15) MLB(online). *Baseball Savant* [URL] <https://baseballsavant.mlb.com/> (アクセス日:2023年1月10日)

16) FanGraphs(online). *FanGraphs* [URL] <https://www.fangraphs.com/> (アクセス日:2023年1月10日)

17) U.S. Copyright Office(2017). *Reproduction of Copyrighted Works by Educators and Librarians* [URL] <https://www.copyright.gov/circs/> (アクセス日:2023年1月10日)

18) Petti, B. (2017). *Research Notebook: New Format for Statcast*

Data Export at Baseball Savant [URL]

<https://tht.fangraphs.com/research-notebook-new-format-for-statcast-data-export-at-baseball-savant/> (アクセス日:2023年1月10日)

19) Weinberg,N.(2014). RE24 [URL]

<https://library.fangraphs.com/misc/re24/> (アクセス日:2023年1月10日)

20) Hastie, T., & Tibshirani, R. (1986). Additive Models Generalized. Stat. Sci, 1: 297-310.

21) Akaike, H.(1973). Information theory and an extension of the maximum likelihood principle, Proceedings of the 2nd International Symposium on Information Theory, Petrov, B. N., and Caski, F. (eds.), Akadimiai Kiado, Budapest

22) Evans, T., & Blackburn, T.M.(2020). Global variation in the availability of data on the environmental impacts of alien birds. Biol Invasions., 22: 1027-1036