

Explanation of Runs Lost Using Combined Fielding Indices in Korean Professional Baseball

Hyuk Joo Kim^{a,1} · Yea Hyoung Kim^a

^aDivision of Mathematics & Informational Statistics, Wonkwang University

(Received August 18, 2015; Revised September 23, 2015; Accepted October 12, 2015)

Abstract

We studied indices to explain runs lost for Korean professional baseball teams. Kim and Kim (2014) studied batting indices to explain run productivity of teams; subsequently, we studied fielding indices to explain runs lost. We considered several combined indices made by combining fielding indices closely connected with the runs lost of teams. Data analysis from all games in the regular seasons of 1982~2014 show that weighted WPH (defined as weighted average of WHIP and number of home runs allowed per game) best explain runs lost. Weighted WPH consisting of WHIP (with weight 81%) and number of home runs allowed per game (with weight 19%) was found optimal weighted WPH having correlation coefficient 0.95033 with average runs lost per game. Analysis by chronological periods gave results not much different.

Keywords: WHIP, number of home runs allowed per game, Weighted WPH, runs lost, correlation coefficient

1. 서론

한국 프로야구팀의 득점과 타격 능력에 관한 연구가 최근 국내의 학자들에 의해 활발히 이루어지고 있다. Cho와 Cho (2005a)는 OPS(On-base percentage Plus Slugging average)와 득점에 관한 연구를 통하여 OPS가 득점에 관한 매우 좋은 추정 도구가 된다는 점을 밝혔다. Lee와 Kim (2005)은 타자들의 공격 능력의 측정 도구로 제안되는 여러 가지의 득점 측도를 비교하였다. Cho와 Cho (2005b)는 득점과 실점을 이용하여 승률을 추정하는 문제를 연구하였으며, Chea 등 (2010)은 정규리그 자료를 이용하여 포스트시즌 진출 여부를 예측할 수 있는 여러 가지 통계모형의 분석 결과를 비교하였다. Hong과 Park (2008)은 한국 프로야구에서 팀들이 기록한 안타의 방향을 분석하고 이를 시각적으로 표현하는 방법을 개발하였다. Lee (2014a)는 타자의 능력을 측정하는 지표를 제안하였으며, Lee (2014b)는 출루율 계수를 추정하는 문제를 연구하였다. 한편 Kim (2012)은 가중OPS를 제안하여 팀의 득점력을 설명하였으며, Kim과 Kim (2014)은 가중OPS를 수정한 가중수정OPS를 이용하면 팀의 득점력을 좀 더 잘 설명할 수 있음을 보였다.

다른 경기도 마찬가지이지만, 야구경기는 득점이 실점보다 많아야 승리하는 경기이다. 따라서 득점 못 지않게 관심을 가질 필요가 있는 것이 실점이다. 실점과 관련이 많은 투수의 자책점에 관련된 연구로서

This paper was supported by Wonkwang University in 2013.

¹Corresponding author: Division of Mathematics & Informational Statistics and Institute of Basic Natural Sciences, Wonkwang University, 460, Iksan-daero, Iksan, Jeonbuk 54538, Korea.

E-mail: hjkim@wonkwang.ac.kr

Table 2.1. Correlation coefficient and rank correlation coefficient between each fielding index and average runs lost per game

	ABG	AVG	BHG	OBP	WHIP	SAG
corr. coeff.	0.72220	0.87061	0.59885	0.88226	0.90711	0.10604
rank corr. coeff.	0.67395	0.85192	0.60635	0.87653	0.89854	0.10596
	HR	HRG	KG	WPG	ERG	DEF
corr. coeff.	0.72107	0.75471	0.15987	0.52142	-0.02571	-0.04582
rank corr. coeff.	0.70678	0.72457	0.19570	0.50619	-0.10510	0.02438

Cho와 Cho (2004)는 2003년 시즌의 한국 프로야구 자료를 바탕으로 WHIP(Walks plus Hits divided by Innings Pitched)가 방어율에 미치는 영향을 연구하였다. 이로부터 WHIP가 팀의 실점에도 영향을 미칠 것으로 추측할 수 있다. 하지만 출루율과 장타율이 결합된 OPS를 사용하면 출루율과 장타율 중 하나만을 사용하는 것보다 훨씬 팀 득점에 대한 설명력이 높아지듯이, 실점의 경우에도 두 개의 지표로 결합한 적절한 수비지표를 사용하면 하나의 지표만을 사용하는 경우보다 설명력이 높아질 것으로 기대된다. 본 논문에서는 Kim과 Kim (2014)이 팀들의 득점력을 설명하기 위한 공격지표를 연구한 것과 유사하게 팀들의 실점을 설명하기 위한 수비지표를 연구하고자 한다.

2. 몇 가지의 수비지표와 실점 사이의 관계

본 논문에서 사용된 자료는 한국 프로야구 원년인 1982년부터 2014년까지 33년에 걸친 정규리그 전 경기(15,660경기)의 수비자료이다. 이 자료는 Korea Baseball Organization (2009)의 한국프로야구 기록대백과와 Korea Baseball Organization (2010, 2011, 2012, 2013, 2014, 2015)의 한국프로야구 연감에서 구한 것이다. 본 논문의 통계분석과 그래프 작성은 모두 Minitab Release 14.12.1을 이용하여 수행하였다.

1982년부터 2014년까지 연 253개 팀의 1년 단위 기록들을 근거로 하여 여러 가지의 수비지표와 실점 사이의 관계를 알아보려 한다. 팀들의 피안타율과 피출루율, (상대팀)경기당타수 등 여러 지표와 경기당 평균실점 간의 상관계수를 구하여 Table 2.1에 나타냈다. 또한 각각의 지표를 기준으로 하여 팀들의 순위를 구한 뒤 경기당 평균실점을 기준으로 한 순위와의 순위상관계수를 구하여 표에 함께 나타냈다. 이 표에서 ABG는 경기당타수, AVG는 피안타율, BHG는 경기당4사구, OBP는 피출루율, SAG는 경기당 희생타, HR은 피홈런, HRG는 경기당피홈런, KG는 경기당탈삼진, WPG는 경기당폭투, ERG는 경기당실책, DEF는 수비율을 각각 나타낸다.

Table 2.1에서 WHIP는 다음 식으로 정의되는 것으로서, 최근 몇 년 사이에 투수의 능력을 평가하는 중요한 지표로 사용되고 있다.

$$WHIP = \frac{\text{볼넷} + \text{피안타}}{\text{투구이닝}}$$

Table 2.1에서 볼 수 있듯이 상관계수와 순위상관계수는 거의 모든 지표에서 매우 유사한 값을 갖는다. 여러 지표들 중 평균실점과 높은 상관관계를 갖는 것을 순서대로 써보면 WHIP(상관계수 0.90711), 피출루율(상관계수 0.88226), 피안타율(상관계수 0.87061), 경기당피홈런(상관계수 0.75471), 경기당타수(상관계수 0.72220)이다. 그리고 피홈런수 자체보다는 피홈런수를 경기수로 나눈 경기당피홈런이 평균실점과 좀 더 높은 상관관계를 가지며, 탈삼진과 실책 및 수비율은 평균실점과의 상관관계가 약한 것으로 나타났다. 투구이닝과 보크 등에 관해서도 조사했는데, 역시 평균실점과의 상관관계가 미미한 것으로 나타났기 때문에 표에 포함하지 않았다.

3. 실점을 설명하기 위한 수비지표

공격지표로 중요하게 사용되고 있는 OPS는 출루율과 장타율의 단순합계인데, 이것이 널리 쓰이는 것은 단지 계산이 간편하다는 이유 때문이다. 출루율과 장타율 중 어느 한 쪽에 더 큰 비중을 둔 가중평균을 사용하여 팀의 득점력을 좀 더 잘 설명할 수 있다면 단순한 OPS보다 더욱 합리적인 타격지표가 될 것이라는 의미에서 Kim (2012)은 다음의 식으로 정의되는 가중OPS를 제안하였다.

$$\text{가중OPS} = a \times \text{출루율} + (1 - a) \times \text{장타율},$$

여기서 a 는 0과 1 사이의 값이다. Kim (2012)은 2007년부터 2011년까지의 한국 프로야구 정규리그 전 경기 자료를 분석하였는데, 그 결과 출루율에 57%의 비중을 두고 장타율에 43%의 비중을 두는 가중OPS가 팀의 득점력을 가장 잘 설명하는 것으로 밝혀졌다. 한편 Kim과 Kim (2014)은 출루율을 약간 수정하여 정의되는 수정출루율과 장타율의 가중평균인 가중수정OPS를 이용하면 팀의 득점력을 좀 더 잘 설명할 수 있음을 보였다.

본 논문에서는 중요한 공격지표 두 개(출루율과 장타율)의 가중평균으로 정의되어 팀의 득점력을 잘 설명해주는 가중OPS처럼, 중요한 수비지표 두 개의 가중평균으로 정의되어 팀의 실점을 잘 설명해주는 수비지표를 제시하고자 한다. 2절에서 평균실점과 높은 상관관계를 갖는 지표들이 WHIP, 피출루율, 피안타율, 경기당피홈런, 경기당타수로 밝혀졌으므로, 이것들을 두 개씩 짝지어 결합하여 다음과 같이 10가지의 결합 수비지표를 고려한다.

$$\begin{aligned} \text{가중BPA} &= a \times \text{경기당타수} + (1 - a) \times \text{피안타율}, \\ \text{가중BPO} &= a \times \text{경기당타수} + (1 - a) \times \text{피출루율}, \\ \text{가중BPW} &= a \times \text{경기당타수} + (1 - a) \times \text{WHIP}, \\ \text{가중BPH} &= a \times \text{경기당타수} + (1 - a) \times \text{경기당피홈런}, \\ \text{가중APO} &= a \times \text{피안타율} + (1 - a) \times \text{피출루율}, \\ \text{가중APW} &= a \times \text{피안타율} + (1 - a) \times \text{WHIP}, \\ \text{가중APH} &= a \times \text{피안타율} + (1 - a) \times \text{경기당피홈런}, \\ \text{가중OPW} &= a \times \text{피출루율} + (1 - a) \times \text{WHIP}, \\ \text{가중OPH} &= a \times \text{피출루율} + (1 - a) \times \text{경기당피홈런}, \\ \text{가중WPH} &= a \times \text{WHIP} + (1 - a) \times \text{경기당피홈런}. \end{aligned}$$

위의 식에서 a 는 0과 1 사이(0과 1 포함)의 값이다. a 의 값을 0부터 1까지 0.01 간격으로 변화시켜 가며 팀들의 수비지표들을 계산하고, 이 수비지표와 팀의 경기당 평균실점 간의 상관계수를 구하여, 각 수비지표에 대해 상관계수가 최대가 되는 값과 그때의 상관계수를 구한 결과가 Table 3.1에 나와 있다. 이 표에서 각 지표의 이름 중 맨 앞의 W는 ‘가중’을 의미한다.

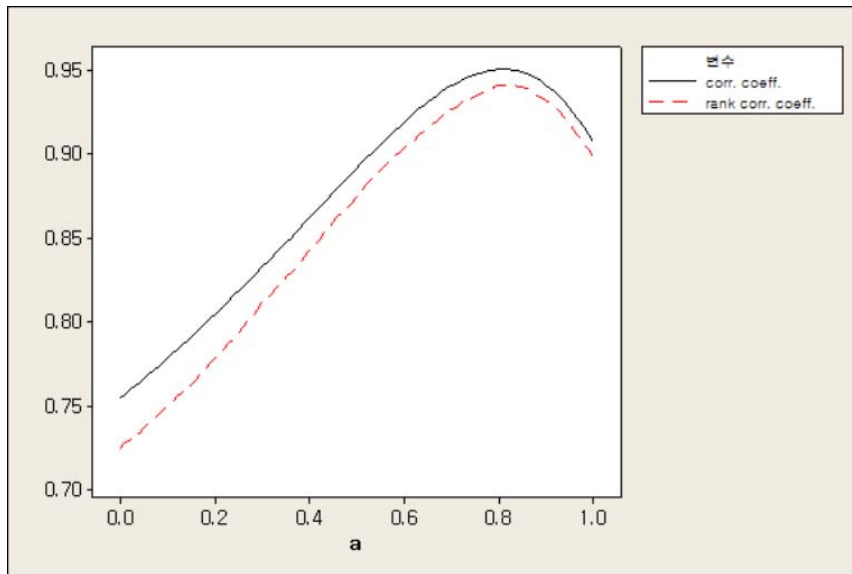
Table 3.1에 있는 값들 중 가장 큰 값은 $a = 0.81$ 일 때의 가중WPH와 평균실점 간의 상관계수인 0.95033이다. 이것은 가중평균을 이용한 지표들 중에서 다음 식과 같이 WHIP에 81%, 경기당피홈런에 19%의 비중을 두는 가중WPH를 사용할 때 팀들의 실점이 가장 잘 설명된다는 것을 의미한다.

$$\text{가중WPH}(0.81) = 0.81 \times \text{WHIP} + 0.19 \times \text{경기당피홈런}.$$

가중WPH와 평균실점 간의 상관계수와 순위상관계수가 a 값에 따라 변화하는 내용을 그래프로 작성한 것이 Figure 3.1이다. 이 그림에서 실선은 상관계수를 나타내고 점선은 순위상관계수를 나타낸다. 순위

Table 3.1. Maximum correlation coefficient (with a value) between each combined fielding index and average runs lost per game

fielding index	WAPO	WAPW	WAPH	WOPW	WOPH
a	0.48	0.75	0.96	0	0.96
max. corr. coeff.	0.90129	0.91525	0.92905	0.90711	0.93971
fielding index	WWPH	WBPA	WBPO	WBPW	WBPH
a	0.81	0.01	0.01	0.05	0.23
max. corr. coeff.	0.95033	0.87339	0.91775	0.92416	0.83031

**Figure 3.1.** Plot of correlation coefficient and rank correlation coefficient between weighted WPH and average runs lost per game, against a .

상관계수는 $a = 0.80$ 일 때 최댓값 0.94126을 갖는 것으로 밝혀졌고, $a = 0.81$ 일때는 0.94093으로서 최댓값과 거의 차이가 없었다. 상관계수를 최대로 하는 a 값과 순위상관계수를 최대로 하는 a 값에 0.01이라는 미세한 차이가 있으나, 지표의 값들로부터 직접 계산된 상관계수가 우선이므로 위의 식으로 정의된 가중WPH(0.81)을 사용하는 것이 타당할 것이다.

가중WPH(0.81)이 팀의 경기당 평균실점에 관한 좋은 지표가 되려면 가중WPH(0.81)과 경기당 평균실점 사이에 직선에 가까운 유의한 관계가 존재해야 할 것이다. 이를 확인하기 위하여 이 두 변수 간의 산점도를 최소제곱회귀직선과 함께 나타낸 것이 Figure 3.2이다. 이 그림에서 RLG는 경기당 평균실점을 나타낸다. 가중WPH(0.81)을 x , 경기당 평균실점을 y 라 할 때, 최소제곱회귀직선의 방정식은

$$\hat{y} = -2.436 + 5.466x$$

였다. 결정계수는 90.3%(이 값은 물론 앞에서 얻은 상관계수 0.95033의 제곱과 같다)로 매우 높았으며, 회귀분석의 F 통계량(자유도 1과 251)의 값은 2,340.09였고 p 값은 논하는 것이 무의미할 정도로 작았다. Table 3.2는 위의 회귀방정식에 근거를 두고 표준화잔차의 절댓값이 큰 비정상적인 관측값들을 찾아낸 것이다. 예컨대 1982년에 삼미의 실제 평균실점이 7.1750점으로 가중WPH(0.81)에 의한 적합값

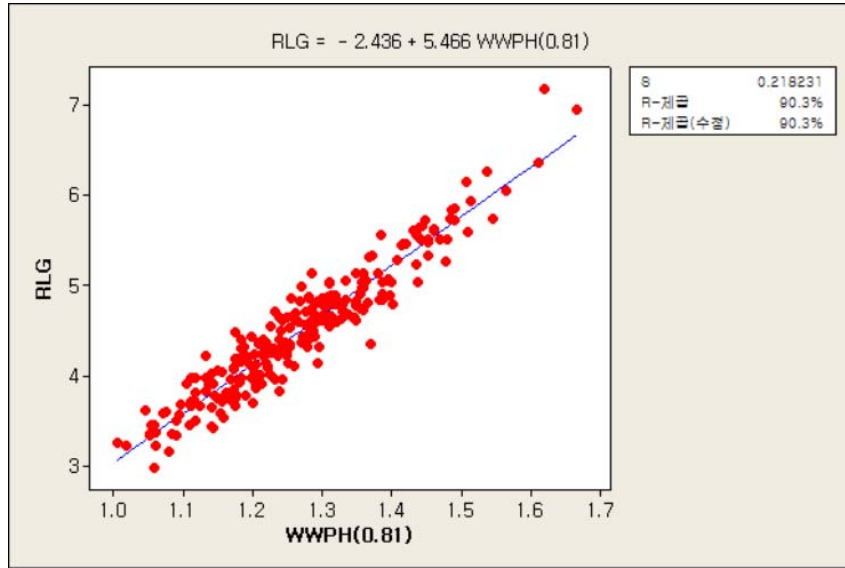


Figure 3.2. Plot of average runs lost per game against WWP(0.81) (1982~2014).

Table 3.2. Unusual observations based on the regression equation

Year	Team	WWP(0.81)	RLG	\widehat{RLG}	$\widehat{SE}(\widehat{RLG})$	Residual	Standardized residual
1982	Sammi	1.62	7.1750	6.4233	0.0424	0.7517	3.51
1998	Lotte	1.29	5.1270	4.5926	0.0139	0.5344	2.45
1982	MBC	1.17	4.4875	3.9859	0.0171	0.5016	2.31
1988	Pacific	1.27	4.9815	4.5040	0.0137	0.4775	2.19
1996	Hanwha	1.13	4.2222	3.7503	0.0204	0.4719	2.17
2003	LG	1.37	4.3534	5.0570	0.0181	-0.7036	-3.24
2007	Hanwha	1.24	3.8175	4.3297	0.0141	-0.5122	-2.35
2009	SK	1.29	4.1353	4.6380	0.0141	-0.5027	-2.31
2004	LG	1.40	4.7895	5.2307	0.0207	-0.4412	-2.03

6.4233점보다 비정상적으로 많이 나왔다는 것을 의미하며, 1998년의 롯데, 1982년의 MBC, 1988년의 태평양, 1996년의 한화도 유사한 경우이다. 반면, 2003년의 LG, 2007년의 한화, 2009년의 SK, 2004년의 LG는 실제 평균실점이 가중WPH(0.81)에 의한 적합값보다 비정상적으로 적게 나온 경우이다. 해당 연도에 이 팀들의 실제 평균실점이 어떠한 이유로 비정상적인 양상을 보였는지는 야구 전문가들의 기술적 관점에서 연구해볼 문제라고 생각된다.

4. 시대별 분석

이 절에서는 한국 프로야구를 시대별(1980년대, 1990년대, 2000년대, 2010년대)로 구분하여 실점과 수비지표들의 관계를 살펴보고자 한다. Table 4.1에는 여러 수비지표와 경기당 평균실점 사이의 상관계수를 시대별로 구한 것이 나와 있다.

Table 4.1에서 볼 수 있듯이 평균실점과 가장 높은 상관관계를 갖는 수비지표는 모든 시대에서 공통적으

Table 4.1. Correlation coefficient between each fielding index and average runs lost per game (by chronological periods)

	ABG	AVG	BHG	OBP	WHIP	SAG
1982~1989	0.70619	0.87688	0.36260	0.87124	0.89842	0.20310
1990~1999	0.65512	0.88402	0.62770	0.87928	0.90235	0.25969
2000~2009	0.68019	0.84650	0.58490	0.84841	0.87590	-0.02903
2010~2014	0.76562	0.89872	0.50185	0.90930	0.94120	-0.03412
	HR	HRG	KG	WPG	ERG	DEF
1982~1989	0.66858	0.75701	-0.43932	0.33787	0.47588	-0.51191
1990~1999	0.75868	0.76170	-0.11219	0.50933	0.05043	-0.09587
2000~2009	0.72567	0.72935	-0.16094	0.31793	0.27088	-0.33502
2010~2014	0.76718	0.78588	-0.13592	0.54077	0.45582	-0.52131

로 WHIP이며, 다음으로 피출루율과 피안타율인데 1980년대와 1990년대에는 피안타율의 상관계수가 피출루율의 상관계수보다 근소하게 높았고 2000년대와 2010년대에는 그 반대였다. 그 다음으로 상관계수가 높은 지표는 경기당피홈런인데, 모든 시대에 피홈런수 자체보다 경기당피홈런이 평균실점과 좀 더 높은 상관관계를 갖는 것으로 나타났다.

이제 시대별로 WHIP, 피출루율, 피안타율, 경기당피홈런, 경기당타수를 두 개씩 짝지어 결합하여 만들어지는, 3절에서 고려했던 10가지의 결합 수비지표를 생각해 보자. 시대별로 10가지의 결합 수비지표와 팀의 경기당 평균실점 간의 상관계수를 $a = 0, 0.01, \dots, 1$ 에 대해 구하여, 각 수비지표에 대해 상관계수가 최대가 되는 a 값과 그때의 상관계수를 구한 결과가 Table 4.2에 나와 있다. 이 표에서 볼 수 있듯이, 평균실점과의 상관계수가 가장 높은 경우는 모든 시대에 대하여 공통적으로 WHIP와 경기당피홈런의 가중평균인 가중WPH에서 나왔으며, 최대상관계수의 값은 0.95 내외로서 Table 3.1에서 1982년부터 2014년까지의 전 기간에 걸쳐서 얻어진 결과인 0.95033과 매우 비슷하고, 그때의 a 값은 1980년대에 0.75, 1990년대에 0.80, 2000년대에 0.80, 2010년대에 0.82로서 역시 Table 3.1에서 얻어진 $a = 0.81$ 과 큰 차이가 없다. 이 결과로부터 알 수 있는 사실을 다음과 같이 정리할 수 있겠다. 가중WPH는 팀들의 경기당 평균실점을 매우 잘 설명하는 결합 수비지표이며, 1980년대에는 최고의 설명력을 보이는 경우의 WHIP의 비중이 75% 정도였는데, 1990년대 이후 2010년대(현재)까지는 80%~82% 정도로 WHIP의 비중이 다소 증가했다.

5. 결론

본 논문에서는 WHIP와 경기당피홈런의 가중평균으로 정의되어 팀의 실점을 설명하기 위한 수비지표인 가중WPH를 제시하였다. 팀의 경기당 평균실점과 높은 상관관계를 갖는 다섯 개의 지표인 WHIP, 피출루율, 피안타율, 경기당피홈런, 경기당타수를 두 개씩 짝지어 생각할 수 있는 열 가지의 결합 수비지표 중 평균실점을 가장 잘 설명해주는 수비지표인 가중WPH를 얻었다. 한국 프로야구 원년인 1982년부터 2014년까지의 정규리그 전 경기 자료를 분석한 결과 WHIP에 81%, 경기당피홈런에 19%의 비중을 두는 가중WPH가 팀의 실점을 가장 잘 설명하는 것으로 나타났으며, 경기당 평균실점과의 상관계수는 0.95033이었다. 시대별(1980년대, 1990년대, 2000년대, 2010년대)로 구분하여 살펴본 결과도 이와 크게 다르지 않았다.

이러한 수비지표를 연구한 의도는, 중요한 공격지표 두 개(출루율과 장타율)의 합으로 정의되어 유용하게 사용되고 있고 또한 팀의 득점력을 잘 설명해주는 것으로 밝혀진 OPS처럼, 적은 수(두 개 정도)의 중요한 수비지표의 가중평균으로 정의되어 팀의 실점을 잘 설명해주는 결합 수비지표를 찾아내고자 한

Table 4.2. Maximum correlation coefficient (with a value) between each combined fielding index and average runs lost per game (by chronological periods)

1982~1989					
fielding index	WAPO	WAPW	WAPH	WOPW	WOPH
a	0.53	0.80	0.95	0	0.94
max. corr. coeff.	0.89916	0.91497	0.90509	0.89842	0.92949
fielding index	WWPH	WBPA	WBPO	WBPH	WBPW
a	0.75	0	0.01	0.18	0.05
max. corr. coeff.	0.94771	0.87688	0.90461	0.84125	0.91862
1990~1999					
fielding index	WAPO	WAPW	WAPH	WOPW	WOPH
a	0.58	0.83	0.97	0	0.95
max. corr. coeff.	0.90628	0.91646	0.93297	0.90235	0.94666
fielding index	WWPH	WBPA	WBPO	WBPH	WBPW
a	0.80	0	0.01	0.18	0.06
max. corr. coeff.	0.95346	0.88402	0.92390	0.81027	0.92816
2000~2014					
fielding index	WAPO	WAPW	WAPH	WOPW	WOPH
a	0.53	0.79	0.96	0	0.96
max. corr. coeff.	0.87122	0.88490	0.92851	0.87590	0.93411
fielding index	WWPH	WBPA	WBPO	WBPH	WBPW
a	0.80	0.01	0.02	0.25	0.07
max. corr. coeff.	0.94617	0.87110	0.91719	0.80155	0.91720
2010~2014					
fielding index	WAPO	WAPW	WAPH	WOPW	WOPH
a	0.47	0.71	0.96	0	0.96
max. corr. coeff.	0.93348	0.94735	0.93706	0.94120	0.95429
fielding index	WWPH	WBPA	WBPO	WBPH	WBPW
a	0.82	0	0.01	0.19	0.03
max. corr. coeff.	0.96563	0.89872	0.94143	0.86217	0.94931

것이다. 그 결과로 얻어진 가중WPH는 경기당 평균실점을 설명하는 데에 있어서 90%가 넘는 결정계수를 보였다.

특점에 관한 좋은 설명력을 가지고 있는 OPS와 가중OPS는 개인기록의 측면에서도 유용하게 활용될 수 있고, 실점과 관련된 가중WPH의 주요 성분인 WHIP도 마찬가지이다. 그러나 가중WPH 자체는 WHIP와 경기당피홈런이 결합된 지표이기 때문에 개인기록으로는 사용될 수 없다는 한계를 갖는다. 투수가 경기에 나와서 던지는 투구횟수가 일정하지 않으므로 단순히 경기당 피홈런 수로 투수를 평가할 수 없기 때문이다. 따라서 가중WPH는 투수 개인 평가의 목적이 아니고, 팀의 차원에서 실점을 줄이기 위해서는 우선적으로 어디에 주력할 것인지를 파악하는 용도로 활용될 수 있을 것이다.

물론 야구에는 실점과 관련하여 생각할 수 있는 여러 가지의 수비지표가 있다. 회귀분석의 변수선택 기법을 사용하면 실점에 대한 설명력을 좀 더 끌어올리는 회귀모형을 얻을 수도 있겠으나, 이렇게 되면 너무 많은 수의 설명변수가 사용되어 본 논문의 취지인 지표의 간편성이 훼손되므로 후후의 연구로 미루고자 한다.

References

- Chea, J. S., Cho, E. H. and Eom, H. J. (2010). Comparisons of the outcomes of statistical models applied to the prediction of post-season entry in Korean professional baseball, *The Korean Journal of Measurement and Evaluation in Physical Education and Sport Science*, **12**, 33–48.
- Cho, Y. S. and Cho, Y. J. (2004). Study about the influence that WHIP has on ERA in 2003 season Korean professional baseball, *Journal of the Korean Data Analysis Society*, **6**, 1415–1424.
- Cho, Y. S. and Cho, Y. J. (2005a). A study on OPS and runs from Korean baseball league, *Journal of the Korean Data Analysis Society*, **7**, 221–231.
- Cho, Y. S. and Cho, Y. J. (2005b). A study on winning percentage using batter's runs and pitcher's runs in Korean professional baseball league, *Journal of the Korean Data Analysis Society*, **7**, 2303–2312.
- Hong, C. S. and Park, H. (2008). Visual representation and applications of hitting direction in Korean baseball records, *Journal of the Korean Data & Information Science Society*, **19**, 539–549.
- Kim, H. J. (2012). Effects of on-base and slugging ability on run productivity in Korean professional baseball, *Journal of the Korean Data & Information Science Society*, **23**, 1065–1074.
- Kim, H. J. and Kim, Y. H. (2014). Explanation of run productivity using weighted adjusted OPS in Korean professional baseball, *The Korean Journal of Applied Statistics*, **27**, 731–741.
- Korea Baseball Organization (2009). *The Official Baseball Encyclopedia 2009*, Korea Baseball Organization, Seoul.
- Korea Baseball Organization (2010). *2010 Official Baseball Guide*, Korea Baseball Organization, Seoul.
- Korea Baseball Organization (2011). *2011 Official Baseball Guide*, Korea Baseball Organization, Seoul.
- Korea Baseball Organization (2012). *2012 Official Baseball Guide*, Korea Baseball Organization, Seoul.
- Korea Baseball Organization (2013). *2013 Official Baseball Guide*, Korea Baseball Organization, Seoul.
- Korea Baseball Organization (2014). *2014 Official Baseball Guide*, Korea Baseball Organization, Seoul.
- Korea Baseball Organization (2015). *2015 Official Baseball Guide*, Korea Baseball Organization, Seoul.
- Lee, J. T. (2014a). Measurements for hitting ability in the Korean pro-baseball, *Journal of the Korean Data & Information Science Society*, **25**, 349–356.
- Lee, J. T. (2014b). Estimation of OBP coefficient in Korean professional baseball, *Journal of the Korean Data & Information Science Society*, **25**, 357–363.
- Lee, J. T. and Kim, Y. T. (2005). A study on runs evaluation measure for Korean pro-baseball players, *Journal of the Korean Data Analysis Society*, **7**, 2289–2302.

결합된 수비지표들을 이용한 한국 프로야구의 실점 설명

김혁주^{a,1} · 김예형^a

^a원광대학교 수학·정보통계학부

(2015년 8월 18일 접수, 2015년 9월 23일 수정, 2015년 10월 12일 채택)

요약

한국 프로야구에서 팀들의 실점을 설명하기 위한 지표를 연구하였다. Kim과 Kim (2014)이 팀들의 득점력을 설명하기 위한 공격지표를 연구한 것과 유사하게 본 논문에서는 팀들의 실점을 설명하기 위한 수비지표를 연구하였다. 여러 가지의 수비지표 중 팀의 실점과 관련이 큰 것들을 결합하여 만든 몇 가지의 결합지표들을 고려하였다. 프로야구 원년인 1982년부터 2014년까지의 정규리그 전 경기 자료를 분석한 결과 WHIP와 경기당피홈런의 가중평균으로 정의되는 가중WPH가 실점을 가장 잘 설명해줬다. 구체적으로 WHIP에 81%, 경기당피홈런에 19%의 가중값을 주는 가중WPH가 팀의 평균실점과 0.95033의 상관계수를 갖는 최적의 가중WPH인 것으로 나타났다. 시대별 분석에서도 크게 다르지 않은 결과를 얻었다.

주요용어: WHIP, 경기당피홈런, 가중WPH, 실점, 상관계수

논문은 2013학년도 원광대학교의 교비 지원에 의해서 수행됨.

¹교신저자: (54538) 전북 익산시 익산대로 460, 원광대학교 수학·정보통계학부 및 기초자연과학연구소.

E-mail: hjkim@wonkwang.ac.kr