The Impact of Cloud Cover on Major League Baseball

WES P. KENT

Medina County Highway Engineers, Medina, and Department of Geography, Kent State University, Kent, Ohio

SCOTT C. SHERIDAN

Department of Geography, Kent State University, Kent, Ohio

(Manuscript received 1 September 2010, in final form 26 January 2011)

ABSTRACT

Although it is often suggested that direct sunlight may affect a player's vision, no published studies have analyzed this interaction. In this research, a variety of statistical tests were utilized to study how baseball variables respond to different cloud cover conditions. Data from more than 35 000 Major League Baseball games, spanning the seasons from 1987 through 2002, were studied. Eleven baseball variables covering batting, pitching, and fielding performance were included. Overall responses were analyzed, as well as individual responses at 21 different stadiums. Home and away team performances were evaluated separately. This study then synthesized the synergistic differences in offensive production, pitching performance, and fielding performance into changes in the "home field advantage."

Offensive production generally declines during clearer-sky daytime games compared to cloudy-sky daytime games, while pitching performance increases as conditions become clearer. Strikeouts show the strongest response in the study, increasing from 5.95 per game during cloudy-sky conditions to 6.40 per game during clear-sky conditions. The number of errors per game increases during clear-sky daytime games compared to cloudy-sky daytime games, while fly outs increase and ground outs decrease between daytime and nighttime games, regardless of the amount of cloud cover. Results at individual stadiums vary, with some stadiums displaying a very strong association between baseball performance and changes in cloud cover, while others display a weak association. All of these impacts affect the home field advantage, with the home team winning 56% of the games played under clear skies compared to 52.3% of the games played under cloudy skies.

1. Introduction

Weather is a force of nature that can impact many facets of our everyday lives. For some humans, the weather can have a direct impact on their profession. From construction workers (Moselhi et al. 1997) to emergency responders battling wildfires (Sun et al. 2009), different meteorological occurrences can elicit different responses. Another profession that can be greatly influenced by the weather is a Major League Baseball player. Although both baseball and climate have been individually analyzed by many researchers, baseball has been studied much less frequently, most often in the light of economic impacts (e.g., Jewell et al. 2004; Kinnard et al. 1997) or analysis of the correlation between team success and attendance levels (Davis 2008). Quantitative analysis of baseball statistics and trends has also been conducted (Albert 2008); however, at present, literature relating baseball to climate is surprisingly limited. The common notion on how weather affects Major League Baseball often deals with inclement conditions. If it is moderately raining, then a game is typically postponed or cancelled. Several studies have assessed how various weather variables can impact Major League Baseball games. An earlier study by Kingsley (1980) examined the effects of temperature and humidity on home run frequencies in Atlanta. Skeeter (1988) analyzed wind and its impact on individual stadiums from 1965 to 1974 while theorizing how those wind patterns might favor left- or right-handed hitters and pitchers. A larger study analyzing a greater number of weather variables was completed by Kraft and Skeeter (1995), focusing on fly ball distances in Major League Baseball games and how weather variables, such as temperature, humidity, wind, air pressure, and altitude, can act to

Corresponding author address: Wes Kent, Engineering Center, Medina County, 791 W. Smith Road, P.O. Box 825, Medina, OH 44256.
E-mail: wpk611@hotmail.com

DOI: 10.1175/2011WCAS1093.1

TABLE 1. Mean values of each baseball variable studied in this research by level of cloud cover.

	Clear	Partly cloudy	Cloudy	Closed/night
Batting average (by away batters)	0.251	0.256	0.256	0.253
Batting average (by home batters)	0.259	0.264	0.266	0.263
Home runs (by away batters)	0.95	1.02	1.01	0.97
Home runs (by home batters)	0.98	1.04	0.96	0.95
Slugging % (by away batters)	0.386	0.398	0.398	0.391
Slugging % (by home batters)	0.405	0.418	0.414	0.410
ERA (allowed by away pitchers)	4.50	4.72	4.68	4.55
ERA (allowed by home pitchers)	3.93	4.21	4.26	4.07
WHIP (allowed by away pitchers)	1.44	1.48	1.47	1.45
WHIP (allowed by home pitchers)	1.34	1.37	1.38	1.35
Walks (allowed by away pitchers)	3.56	3.62	3.50	3.44
Walks (allowed by home pitchers)	3.37	3.39	3.43	3.30
Strikeouts (recorded by away pitchers)	6.14	6.07	5.67	5.83
Strikeouts (recorded by home pitchers)	6.65	6.58	6.22	6.40
Fly ball outs (allowed by away pitchers)	9.69	9.77	9.80	9.45
Fly ball outs (allowed by home pitchers)	10.00	10.09	10.08	9.83
Ground ball outs (allowed by away pitchers)	11.42	11.66	11.85	12.20
Ground ball outs (allowed by home pitchers)	12.10	12.27	12.45	12.75
Errors (committed by away team)	0.80	0.74	0.73	0.73
Errors (committed by home team)	0.77	0.76	0.75	0.72
Winning % (away)	0.440	0.470	0.477	0.463
Winning % (home)	0.560	0.530	0.523	0.537

increase or decrease the distance traveled by a baseball. In addition, a more site-specific study dealing with the impacts of altitude and air density on fly ball distance was conducted for the Colorado Rockies team by Chambers et al. (2003).

These baseball and climate studies have all contributed greatly to the discipline. However, no study analyzing the impact of sunlight or cloud cover changes on Major League Baseball has been published. On an international level, analysis of the impact of light levels on cricket matches has been researched, where it has become accepted practice to alter the order of batsmen based on the cloud cover conditions present during the match (Norman and Clarke 2006). Therefore, similar to cricket, knowledge of the baseball–cloud cover relationship could provide Major League Baseball teams and managers with valuable information to assist them with game-day preparations. The goal of this study is thus to analyze how variability in cloud cover can impact Major League Baseball player productivity. To achieve this goal, a variety of statistical tests were utilized to study how baseball variables respond to different cloud cover conditions. Eleven baseball performance measures from batting, pitching, and fielding were included. Overall responses were analyzed, as well as individual responses at 21 different stadiums, with home and away team performances evaluated separately. This study then examined whether the synergistic differences in offensive production, pitching performance, and fielding performance led to changes in the ''home field advantage.''

2. Data and methodology

a. Data

This research required two different sets of data: cloud cover and baseball statistics. Baseball data were obtained from STATS Inc. for all Major League Baseball games played during the period from 1987 to 2002, a total of 35 101 games. The 11 baseball variables used in this study are shown in Table 1 and were compared across three defined levels of cloud cover. All variables were calculated as away and home team totals on a per game basis; that is, for every game during the study period, the team batting average for the home team and the away team was separately calculated.

The cloud cover data used for this study were obtained from the National Climatic Data Center (NCDC). These data were obtained for all cities where a Major League Baseball team resided during the length of the study period. One of the difficulties in obtaining meteorological data for individual stadiums is that the stadiums themselves do not record standardized weather observations. Therefore, cloud cover conditions at 11:00 a.m., 5:00 p.m., and 11:00 p.m. EDT from the nearest airport to the stadium were used. Although there may be small microclimatic differences, cloud cover is more consistent

TABLE 2. The number of games per stadium by cloud cover group.

Stadium	Clear	Partly cloudy	Cloudy	Daytime total	Night/closed
Chicago (NL): Wrigley Field	189	448	375	**1012**	256
Oakland: Oakland Coliseum	243	284	79	**606**	650
San Francisco: Candlestick Park	272	205	81	**558**	466
New York (AL): Yankee Stadium	80	216	145	**441**	816
New York (NL): Shea Stadium	80	216	132	**428**	831
Boston: Fenway Park	62	191	170	**423**	846
Cincinnati: Cinergy Field	95	172	137	**404**	863
Milwaukee: County Stadium	95	167	131	**393**	704
St. Louis: Busch Stadium	102	163	118	**383**	885
Detroit: Tiger Stadium	85	166	110	**361**	659
San Diego: Qualcomm Stadium	122	171	61	**354**	902
Philadelphia: Veterans Stadium	55	152	89	**296**	971
Kansas City: Kauffman Stadium	69	113	99	**281**	979
Pittsburgh: Three Rivers Stadium	44	120	108	**272**	835
Chicago (AL): U.S. Cellular Field	46	133	91	**270**	667
Los Angeles (NL): Dodger Stadium	108	125	35	**268**	993
Baltimore: Oriole Park at Camden Yards	36	147	66	**249**	607
Denver: Coors Field	22	183	28	**233**	334
Los Angeles (AL): Angel Stadium	88	117	27	**232**	1038
Cleveland: Progressive Field	59	100	59	**218**	470
Toronto: Rogers Centre	61	98	52	**211**	866

on a mesoscale level than other meteorological variables (e.g., temperature) between two close places.

Three groups, based on the local start time of a game, were defined. The *daytime* group includes those games starting between 11:00 a.m. and 3:59 p.m., while the *nighttime* group includes games starting at 7:00 p.m. and later. In addition, all games played in a dome, or under a closed roof, were included in the nighttime category. A third category, *in-between*, encompasses games starting between 4:00 and 6:59 p.m. Because of the late-afternoon start times, portions of these games may be played during the day, other portions at night, which made it difficult to clearly classify in-between games as either daytime or nighttime. Since only about 5% of the total number of games in the study fell into this category, these games were removed from analysis. Overall, the daytime game total over the 16-yr period is 10 758, being outnumbered by nighttime games by more than 2 to 1.

A classification procedure was next applied to all games to determine those games that may be impacted by cloud cover and those games where the amount of cloud cover was irrelevant. The cloud cover conditions at 11:00 a.m., 5:00 p.m., and 11:00 p.m. EDT were utilized in this process. For the eastern and central time zones, the earlier two cloud cover conditions were averaged, while the latter two observations were averaged for the mountain and Pacific time zones. These measurements were chosen because they occur closer to the average starting and ending times for day games played in these time zones. After determining the average amount of cloud cover present for a daytime game, each game was assigned to one of three cloud cover groups. A game was placed into the *clear* group if the mean cloud cover was less than 30%. The *partly cloudy* group included those games that were played under a mean cloud cover greater than 30% but less than 80%. Games played with 80% or greater mean cloud cover were assigned to the *cloudy* group. The fourth group contained those games that fall into the nighttime category. These four cloud cover classification groups were used in analyzing the baseball variables in the study. The analysis was performed on a stadium-level basis, with home team and visiting team performances evaluated separately.

With numerous new stadiums being built over the course of the study period, the sample size for some of these stadiums is quite small. To provide more robust statistical testing, only stadiums for which at least 200 daytime games were available were analyzed, as this threshold represented a distinct break in day-game sample size. The stadiums studied in this research are listed in Table 2.

b. *Methodology*

A variety of statistical testing methods were used in this study. Samples from clear and cloudy daytime games were analyzed via a two-sample difference of means test using the Wilcoxon rank-sum test (WRT). Of all of the sky conditions a daytime baseball game is played under, clear and cloudy are the most dissimilar. Therefore, a higher probability exists for a significant difference in performance between these two conditions.

Samples from games played under partly cloudy skies were not tested because the amount of cloud cover may not be as consistent under this sky condition.

Least squares regression analysis was used to determine the strength of the association between cloud cover and individual baseball variables. The first regression test employed the use of two binary ''dummy'' variables (clear daytime game and cloudy daytime game), whose values were set to 0 when analyzing games in the nighttime category. Similar to the WRT, partly cloudy daytime games were not analyzed. Each of the baseball variables were then regressed against the two cloud cover dummy variables, and equations were produced showing the association each of the dummy variables has with the particular baseball variable being tested. For example, the results of the binary regression test on home team batting average in Wrigley Field are calculated as follows:

$$\mathrm{AVG}_{\mathrm{home}} = .264 - .016_{\mathrm{clear}} + .006_{\mathrm{cloudy}}.$$

The results show that when cloud cover is irrelevant (during night games), the home team batting average in Wrigley Field was 0.264. Under clear-sky conditions, the batting average dropped by 0.016. Under cloudy-sky conditions, however, the batting average rose to 0.270. This form of regression was performed for the other baseball variables as well, and all coefficients were tested for statistical significance.

The second regression test had only one independent variable, average cloud cover. Like the previous test, this regression analyzed how baseball variables responded to changes in cloud cover. While the first test utilized the defined categories of cloud cover, this second test directly used the average cloud cover amount observed during a particular game, for example, ''80% cloud cover'' as opposed to a ''partly-cloudy-sky game.'' Because of the irrelevance of cloud cover conditions during nighttime games and games played under a closed roof, only daytime games were included in this regression analysis.

Changes in winning percentage due to cloud cover were tested using a parametric two-sample difference of proportions test. As in the WRT, clear- and cloudy-sky daytime games were analyzed against each other. The testing methods reviewed above were performed for all variables at the 21 stadiums selected for the study. Unless otherwise noted, all tests in this study are regarded statistically significant at the 95% level ($\alpha = 0.05$).

As all of the statistical tests were initially performed at stadium level, the collective significance across all stadiums was determined by a binomial test. That is, for each baseball variable, a value of 1 (0) was assigned if an increase (decrease) in performance occurred at a particular stadium. For example, when studying batting average on cloudy days, those instances when the batting average increased above the control received a value of 1. Since the testing methods employed in this study analyzed different variables, the control varies: for the WRT, clear daytime game values are the control, while nighttime game values represent the control for the binary variable regression tests. If 14 out of 21 stadiums exhibited a trend in the same direction, then this was considered statistically significant at a 5% level.

3. Results

The following sections examine the results of the statistical tests, broken down by baseball variable. The results for all 11 variables used in the study are shown in Table 1; only the strongest and most relevant results are discussed in detail below.

a. Batting average

Batting average may be the most popular way to measure offensive proficiency in baseball. Overall, across all stadiums, both away and home team batting averages show a slight increase between games played under clear- and cloudy-sky conditions, with the home team batting average increasing from 0.259 to 0.266 and the away team batting average rising from 0.251 to 0.256. When comparing daytime games to nighttime games, the batting average for both the away and home teams decreases during clear-sky games, while an increase in batting average is seen for cloudy-sky games. An opposite trend, however, is observed when analyzing batting average using the average cloud cover regression test, with a statistically significant 16 stadiums ($p < 0.01$) showing a decrease in away and home team batting averages as sky conditions become cloudier.

Examining stadiums individually, the home team batting average increases at 14 out of 21 stadiums between clear- and cloudy-sky games, with two of these increases statistically significant. The largest increase is seen at Shea Stadium, where the home team batting average rises from 0.231 during clear-sky games to 0.264 during cloudy-sky games (Table 3). Statistically significant increases are also present for the away team between clear- and cloudy-sky daytime games, with the largest change (+.029) occurring at Oakland Coliseum (Table 4). Across all tests, several stadiums exhibit consistent statistically significant results for batting average. Oakland Coliseum shows statistically significant increases in batting average with increasing cloud cover in three out of the four away team tests, while Shea Stadium and Wrigley Field have significant changes in batting average in three out of the four home team tests.

TABLE 3. The numerical change in baseball variables from clear-sky games to cloudy-sky games for the home team at each stadium. **Bold** values indicate statistically significant results ($p \leq 0.05$), *italicized* values are near significant ($0.05 < p \leq 0.10$).

Stadium	Batting average	Home runs	ERA	Strikeouts	Errors	Fly outs	Win %
Baltimore	−0.024	+0.07	+0.31	−0.34	+0.06	+0.44	*−.156*
Boston	−0.005	−0.10	−0.18	+0.47	−0.02	+0.27	−0.022
Chicago (AL)	+.002	+0.06	*−0.22*	**−0.72**	*+0.20*	+0.54	+.082
Chicago (NL)	**+.022**	−0.03	+0.54	−0.58	−0.14	+0.34	−0.055
Cincinnati	*+.024*	−0.15	−0.10	−0.69	−0.14	*−0.11*	+.036
Cleveland	+.003	*+0.22*	+0.33	−0.33	+0.15	+0.04	+.034
Denver	+.011	−0.61	−0.63	+0.41	−0.16	−1.50	+.068
Detroit	−0.009	−0.03	−0.25	**−0.87**	+0.01	−0.28	−0.009
Kansas City	+.010	−0.01	+0.01	+0.27	−0.01	+0.49	−0.002
Los Angeles (AL)	+.009	+0.08	−0.46	−0.58	−0.34	+0.92	+.019
Los Angeles (NL)	+.010	−0.21	−0.57	−0.63	−0.10	−0.54	+.025
Milwaukee	−0.002	−0.06	+0.19	−0.39	−0.16	−0.03	−0.064
New York (AL)	+.009	+0.22	−0.12	+0.11	+0.11	−0.02	−0.002
New York (NL)	**+.033**	+0.17	*+0.65*	**−0.67**	−0.01	+0.73	+.035
Oakland	−0.001	+0.14	**+1.14**	*−0.63*	−0.17	**+1.03**	**−.145**
Philadelphia	+.010	+0.10	+0.29	−0.08	+0.05	**+1.20**	+.012
Pittsburgh	+.011	+0.32	+0.57	−0.38	+0.15	*+0.86*	+.060
San Diego	+.010	+0.13	+0.68	−0.32	−0.01	−0.25	−0.065
San Francisco	+.001	−0.01	*+0.41*	−0.10	−0.06	−0.33	−0.060
St. Louis	−0.003	−0.07	*+0.48*	−0.50	−0.08	+0.38	−0.029
Toronto	**−.041**	−0.15	+0.06	−0.31	+0.09	+0.53	−0.121

b. Home runs

Though home runs are also a measure of offensive ability, the results for this variable are not as consistent as those for batting average. Although the majority of stadiums show fewer home team home runs during cloudy-sky games than clear-sky games, away team home runs increase at most stadiums as sky conditions become cloudier. These results are opposite of those for batting average, where the majority of stadiums show an

TABLE 4. The numerical change in baseball variables from clear-sky games to cloudy-sky games for the away team at each stadium. **Bold** values indicate statistically significant results ($p \leq 0.05$), *italicized* values are near significant ($0.05 < p \leq 0.10$).

Stadium	Batting average	Home runs	ERA	Strikeouts	Errors	Fly outs	Win %
Baltimore	+.019	+0.30	+0.07	**−1.11**	+0.09	+0.60	*+.156*
Boston	−0.009	+0.03	−0.56	+0.01	−0.03	+0.70	+.022
Chicago (AL)	−0.004	*−0.13*	−0.34	**−0.99**	−0.18	+0.26	−0.082
Chicago (NL)	+.014	+0.13	*+0.46*	**−0.65**	+0.03	+0.34	+.055
Cincinnati	+.007	−0.10	**+0.12**	−0.36	−0.09	−0.96	−0.036
Cleveland	0.000	0.00	+0.70	+0.25	**−0.34**	+1.17	−0.034
Denver	−0.018	−0.25	+0.20	*−1.37*	+0.07	−1.00	−0.068
Detroit	−0.011	−0.30	−0.12	−0.22	+0.08	−0.19	+.009
Kansas City	−0.007	−0.19	+0.36	**−1.01**	−0.07	+0.19	+.002
Los Angeles (AL)	**−.032**	+0.12	−0.01	*−0.72*	−0.08	+0.08	−0.019
Los Angeles (NL)	*−.020*	−0.08	−0.25	**−1.21**	−0.15	−0.19	−0.025
Milwaukee	+.005	+0.07	+0.07	−0.25	+0.18	+0.43	+.064
New York (AL)	−0.013	−0.16	+0.21	−0.01	−0.02	+0.43	+.002
New York (NL)	+.008	*+0.23*	**+0.77**	*−0.57*	*−0.27*	+0.20	−0.035
Oakland	**+.029**	+0.19	−0.05	−0.40	+0.05	+0.02	**+.145**
Philadelphia	−0.002	+0.14	+0.68	+0.07	−0.02	+0.07	−0.012
Pittsburgh	+.011	+0.08	**+1.36**	−0.11	+0.08	+0.98	−0.060
San Diego	+.006	+0.11	+0.42	−0.55	−0.11	*+0.77*	+.065
San Francisco	+.014	+0.12	−0.06	−0.15	−0.15	+0.63	+.060
St. Louis	+.010	+0.18	+0.09	−0.31	−0.08	+0.20	+.029
Toronto	−0.008	−0.06	**−1.64**	+0.20	−0.07	−0.56	+.121

increase in home team batting average and a decrease in away team batting average as cloud cover increases.

The cloudy-sky binary variable test shows home team home runs increasing by 0.01 HR per game, although the majority of stadiums show a decrease in home runs between nighttime games and cloudy-sky games. Across the levels of cloudiness during daytime games, home team home runs do not exhibit any statistically significant trend, as they are highest during games under partly cloudy skies. For the away team, a 2% decrease in home runs occurs when comparing nighttime games to clear-sky daytime games, with a slight increase in home runs during cloudy-sky games.

c. *Earned run average*

Earned run average (ERA) may be the most popular way to measure pitching performance. Overall, a decrease in ERA occurs as conditions become clearer, with the away team ERA. (i.e., the ERA allowed by away team pitchers) decreasing from 4.68 to 4.50 between cloudy- and clear-sky daytime games. A larger change is seen with home team ERA, which drops to 3.93 during clear-sky games from 4.26 during cloudy-sky games. Strong responses are also observed between cloudy-sky daytime games and nighttime games, with a statistically significant amount of stadiums displaying an increase in ERA for both teams under cloudy-sky conditions. These results coincide with the overall decreases seen in the majority of the offensive variables as sky conditions become clearer, which should be expected, as a decline in offensive production should lead to an improved pitching performance.

Several stadiums show statistically significant changes in ERA between different levels of cloud cover. When comparing cloudy-sky daytime games to nighttime games, Shea Stadium (+0.69) shows a statistically significant increase in home team ERA, along with Busch Stadium (+0.62). Three Rivers Stadium has the largest increase in away team ERA, rising from 4.31 during nighttime games to 5.04 during those games played under cloudy-sky conditions. The largest away team ERA increase between clear and cloudy day games also occurs at Three Rivers Stadium, with a rise of 1.36. Lastly, Oakland Coliseum has the largest home team response, with ERA rising from 3.72 under clear skies to 4.86 under cloudy skies. Using the average cloud cover regression test, a rise in home team ERA of 0.08 for every 10% increase in cloud cover is also uncovered at Oakland Coliseum, although the goodness of fit is minimal ($r = 0.10$).

d. *Strikeouts*

A strikeout is perhaps the most valuable out in baseball from the defensive standpoint, since it prevents the ball from being put in play. In most cases, as cloud cover decreases, strikeouts recorded by the pitcher increase, with five out of the eight strikeout tests displaying a statistically significant number of stadiums with similar changes in strikeouts. Some of the results of the individual tests are among the strongest in the study, with statistically significant differences in strikeouts occurring for away and home teams at many stadiums.

The statistical tests performed for the strikeout variable show consistency when comparing the results across all levels of cloud cover (Table 1). During nighttime games, the away team pitchers strike out 5.83 batters per game, with that number decreasing to 5.67 for cloudy-sky daytime games. During clear-sky daytime games, however, the number of strikeouts per game for the away team pitchers rises to 6.14. The results are similar for home team pitchers, with the highest level of strikeouts occurring during clear-sky games (6.65) and the lowest during cloudy-sky games (6.22).

The strong results of the statistical tests on the strikeout variable are consistent at a stadium level as well, with at least one occurrence of statistical significance in 15 stadiums. Individually, Coors Field has a 20% decrease in strikeouts by the away team pitchers, from 6.73 per clear-sky game to 5.36 per cloudy-sky game. Other notable statistically significant decreases occur at Dodger Stadium (−1.21) and Oriole Park (−1.11). Busch Stadium and Wrigley Field also have statistically significant decreases in both away and home team pitcher strikeouts with increasing cloud cover. Average cloud cover regression analyses at Wrigley Field suggest strikeouts recorded by the away team pitchers drop from 6.60 per game with 0% cloud cover to 5.80 per game with 100% cloud cover ($r = -0.086$).

e. *Fly outs and ground outs*

In baseball, in addition to the strikeout, a batter can also make an out by fly and ground out. The results of the binary variable regression tests show that fly outs increase in all games played during the day, when compared to nighttime games. Fly outs increase for both the away and home team when comparing nighttime games to clear-sky daytime games (+2%) as well as cloudy-sky daytime games (+3%); for cloudy-sky daytime games, 20 of the 21 individual stadiums show an increase for the away team. The greatest increase in fly outs allowed by away team pitchers occurs at Progressive Field, where they rise from 9.42 per nighttime game to 10.97 per cloudy-sky daytime game. Oakland Coliseum has the highest statistically significant increase in fly outs allowed by the home team, rising by 1.75 per game.

Fly outs also increase with increasing cloud cover when comparing clear-sky daytime games to cloudy-sky daytime games. Overall, there is not a large mean difference between these two scenarios, but an increase in both away and home team fly outs occurs at a statistically significant number of stadiums. Qualcomm Stadium has the greatest statistically significant increase in away team fly outs allowed, rising by 9% to 8.38 per game under cloudy-sky conditions. An 11% increase in home team fly outs occurs at Veterans Stadium, where a clear-sky game value of 9.44 fly outs per game rises to 10.64 per cloudy-sky game.

As would be expected, ground ball outs show an opposite response to fly outs with changes in cloud cover. An overall decrease in ground ball outs is observed in all daytime games compared with nighttime games, with this decrease becoming more pronounced as conditions become clearer.

f. Errors

Errors are the only variable in the study that measures fielding proficiency. Despite not being directly linked to either batting or pitching, the results of the statistical tests on errors do show some similarities to those seen for the variables discussed above. For example, errors show an increase from nighttime games to clear-sky daytime games, rising by 10% for the away team and 6% for the home team. Comparing the daytime games to each other, errors increase under clearer skies for both teams; however, home team errors only show a modest rise of 3% between cloudy- and clear-sky games, while away team errors increase 10%.

This impact on away team errors manifests itself at a stadium level as well. Cinergy Field shows a 24% jump in away team errors between nighttime games and cloudy-sky daytime games, rising from 0.67 to 0.88. Statistically significant increases in away team errors between cloudy- and clear-sky games also occur at Progressive Field and Shea Stadium, where errors rise by 0.34 and 0.27 per game, respectively. Wrigley Field is the only stadium where a statistically significant increase in the number of home team errors per game is observed, with a rise from 0.73 per cloudy-sky game to 0.87 per clear-sky game.

g. Winning percentage

An overall increase in away team winning percentage is seen between clear- and cloudy-sky games when data from all stadiums are used. This difference is also statistically significant, with the away team's winning percentage almost 4% points higher under cloudy-sky conditions. A slight majority of the stadiums (12 out of 21) exhibit this trend as well. Only Oakland Coliseum displays a statistically significant difference between clear- and cloudy-sky games, with the home team winning percentage decreasing from 0.613 under clear skies to 0.468 under cloudy skies. Camden Yards displays the largest difference in winning percentage in the study (0.156), although this result is only statistically significant at a 10% level. Finally, the home team winning percentage at the Skydome drops to 0.404 under cloudy-sky conditions from 0.525 during clear skies; however, this result does not show statistical significance ($p = .101$).

4. Discussion

Baseball, like most sports, is a highly visual game. The main interaction in the sport comes in the form of the pitcher versus the batter. Thus, the analysis in this study focuses on this interaction and how the performance of one side or the other may be affected by changes in cloud cover.

One of the most crucial aspects to the pitcher–batter interaction is how well the batter can see the pitch. Changing cloud cover presents different playing conditions, with some playing conditions potentially helping a batter see a pitch, whereas others may make reading a pitch more difficult. For example, brighter conditions may result in increased eye strain for a batter and a higher level of glare in a ballpark. These factors could contribute to less than favorable conditions for a batter trying to focus on a pitch, impacting performance in a number of areas.

The majority of the offensive variables display a decrease in performance under clear-sky daytime conditions compared to both cloudy-sky daytime conditions and those games played under a closed roof or at night. These decreases are consistent for both away and home teams, and could be linked to the visualization problems discussed above.

With cloudier sky conditions benefiting the batter, it should be expected that clearer conditions benefit the pitcher. The ERA and walks/hits per inning pitched (WHIP) do reflect increased pitching performance as sky conditions become clearer. Walks issued by pitchers do not show any significant difference relative to cloud cover, but cloud-cover-related variability in strikeouts is stronger than any other variable in the study. This may be due to the simplicity of a strikeout compared to some of the other variables studied, as the two main players involved are the batter and the pitcher. Conversely, batting average and slugging percentage are dependent not only on how well a batter hits a ball but how it is fielded or how the batter runs the bases.

Fly outs and ground outs show complementary results when comparing daytime and nighttime games. Fly outs increase during daytime games and ground outs decrease, regardless of cloud cover. This may be due to higher temperatures during clear-sky daytime games, as Kraft and Skeeter (1995) suggest. The number of errors per game shows a slight increase as conditions get clearer and during all daytime games, perhaps because of fielders losing fly balls in the sun or misjudging a ball because of increased eye strain with brighter conditions. However, these increases are most likely limited because an error is not always given when a fielder loses a fly ball in the sun. If the ball drops and the fielder never touches it, then the official scorer may not credit him with an error.

Despite having similar trends with many of the variables, the level of response differs between the away and home teams. Across all games in this study, the home team wins approximately 54% of the time, similar to historical values (*Sports Illustrated*, 3 February 2003). The home field advantage becomes more pronounced during clear-sky daytime games, when 56% of games are won by the home team—a winning percentage that is 3.7 percentage points higher than during cloudy-sky games. While both teams are impacted by visual disadvantages caused by increased sunlight, the home team may compensate better for the adverse playing conditions because of its familiarity with the ballpark. This familiarity with daytime light conditions in a stadium may work in conjunction with other meteorological variables to strengthen the home field advantage. For example, Borghesi (2007) found that game-day temperature in National Football League games significantly affects team performance, especially when the difference in recent average temperature between the cities where the competing teams are located is higher.

Certain stadiums show more consistent results in this research. The strongest responses occurred at Oakland Coliseum, Shea Stadium, and Wrigley Field. Sample size may play a role in explaining the significance of these results, as Oakland Coliseum and Wrigley Field have the two largest numbers of daytime games in the study. Additionally, the players on these teams may be more experienced with playing in brighter daytime conditions, which could provide them with a stronger home-field advantage when they compete against players who have not participated in as many daytime contests. Moreover, all three of these stadiums have a more open architectural style, which may allow sunlight and cloud cover changes to have more of an impact on the daytime games played in these stadiums. The remainder of the stadiums in the study all show statistical significance in some of their results but nothing comparable to the stadiums discussed above, likely at least partially attributable to smaller sample sizes.

The relatively low goodness-of-fit statistics observed in this research may be explained by a myriad of other factors that confound the relationship between baseball statistics and cloud cover. For instance, the location of shadows can vary between stadiums, due to differences in stadium orientation (measured by the direction the home plate faces) and architectural style. Similarly, the height and type of cloud conditions present during a game may also impact baseball performance. High and thin cirrus clouds could create a brighter playing environment, while lower stratus clouds should result in less sunlight-induced glare, potentially benefiting the batter. Cloud type and amount can also vary over the course of a game; therefore, a single number representing cloud cover (i.e., 80%) may not be an appropriate metric for all games. Concurrently, a change in the amount of cloud cover present in the atmosphere may also alter other weather variables, such as humidity and temperature, which could also impact player performance (Kraft and Skeeter 1995). Ultimately, the interaction between cloud cover and baseball performance is at the individual level. Previous research has shown that an individual player on a given team can influence team averages significantly (Davenport 2001); hence, the impact of cloud cover changes at a player level should be further explored.

5. Conclusions

In this research, more than 35 000 Major League Baseball games were analyzed. Eleven baseball variables were compared across varying levels of cloud cover, using a variety of statistical methods. The results show that offensive production generally declines during clearer-sky daytime games compared to cloudy-sky games. Pitching performance displays complementary results to the offensive variables, with performance increasing as conditions became clearer. Strikeouts show the strongest response in the study, increasing from 5.95 per game during cloudy-sky conditions to 6.40 per game during clear-sky conditions. An overall increase in the number of errors occurs during clear-sky daytime games compared to cloudy-sky daytime games, with an increase observed during all daytime games when compared with nighttime games. Lastly, fly outs increase and ground outs decrease in all daytime games compared to nighttime games, with an increase in both variables also occurring between clear- and cloudy-sky daytime games.

Results at individual stadiums differ in their magnitude, with some stadiums showing a very strong association between baseball performance and changes in cloud cover, while others display little association. Some of this difference can be attributed to sample size, as stadiums where the most daytime games were played

(Oakland Coliseum and Wrigley Field) display some of the most significant results. Although all of the impacts due to cloud cover discussed above influence both away and home teams, the home team winning percentage increases during clear-sky conditions compared to cloudy-sky conditions, possibly indicating that players on the home team are not affected as much as away team players when more sunlight is present.

Although this study uncovered many significant results, clearly new avenues for research are present. Whereas overall impacts on batters, pitchers, and fielders have been discussed, more specific player analyses can be performed. For example, it is not known if left- or right-handed batters are more impacted by changes in cloud cover at particular stadiums, along with which defensive positions may be most affected. The role of sunlight/shadow differences could be more quantitatively studied by obtaining stadium architecture measurements, as well as sun elevation angle and azimuth at different times of the baseball season. Intraseasonal studies or trend analyses could also be completed to determine if baseball performance or cloud cover patterns change over time. Last, further analysis can examine whether teams that participate in a greater number of daytime games experience less disadvantage while on the road than teams that participate in fewer daytime games.

Acknowledgments. We would like to express our gratitude to Milton Harvey and Chad Kauffman, as well as the editor and reviewers, for their work in improving this research and manuscript.

REFERENCES

Albert, J. H., 2008: Streaky hitting in baseball. *J. Quant. Anal. Sports,* **4,** doi:10.2202/1559-0410.1085.

Borghesi, R., 2007: The home team weather advantage and biases in the NFL betting market. *J. Econ. Bus.,* **59,** 340–354.

Chambers, F., B. Page, and C. Zaidins, 2003: Atmosphere, weather, and baseball: How much farther do baseballs really fly at Denver's Coors Field? *Prof. Geogr.,* **55,** 491–504.

Davenport, C., 2001: Park factors. *Baseball Prospectus: 2001 Edition,* J. Bower, Ed., Brassey's Inc., 517–518.

Davis, M. C., 2008: The interaction between baseball attendance and winning percentage: A VAR analysis. *Int. J. Sport Finance,* **1,** 58–73.

Jewell, R. T., M. A. McPherson, and D. J. Molina, 2004: Testing the determinants of income distribution in Major League Baseball. *Econ. Inquiry,* **42,** 469–482.

Kingsley, R. H., 1980: Lots of home runs at Atlanta? *Baseball J.,* **2,** 66–71.

Kinnard, W. N., M. B. Geckler, and J. DeLottie, 1997: Team performance, attendance and risk for Major League Baseball stadiums: 1970-1994. *Real Estate Issues,* **1,** 8.

Kraft, M. D., and B. R. Skeeter, 1995: The effect of meteorological conditions on fly ball distances in North American Major League Baseball games. *Geogr. Bull.,* **37,** 40–48.

Moselhi, O., D. Gong, and K. El-Rayes, 1997: Estimating weather impact on the duration of construction activities. *Can. J. Civ. Eng.,* **24,** 359–366.

Norman, J. M., and S. R. Clarke, 2006: Dynamic programming in cricket: Optimizing batting order for a sticky wicket. *J. Oper. Res. Soc.,* **58,** 1678–1682.

Skeeter, B. R., 1988: Winds and the baseball stadium: General wind characteristics in Major League Baseball stadiums. *Sport Place,* **2,** 37–44.

Sun, R., S. K. Krueger, M. A. Jenkins, M. A. Zulauf, and J. J. Charney, 2009: The importance of fire–atmosphere coupling and boundary-layer turbulence to wildfire spread. *Int. J. Wildland Fire,* **18,** 50–60.