The New York Times

April 7, 2009

Answering Baseball's What-Ifs

By ALAN SCHWARZ

You can learn a lot during a major league baseball game. Like Ukrainian, if it is a particularly slow nine innings.

As for the science of baseball strategy, one game teaches precious little. A well-timed sacrifice bunt can backfire and lose the game; a foolish steal can appear brilliant. The vagaries of randomness — the way Sandy Koufax got battered occasionally and a pipsqueak named Bucky Dent hit one of the most famous home runs ever — camouflage the game's inner forces, which for 150 years have operated somewhere between fact and fable.

One game has little meaning. A thousand seasons can take a while. Thank goodness for quad-core processors.

"Computer simulations work pretty well in baseball for two reasons," said Carl Morris, a professor of statistics at Harvard University who has written several papers that commingled baseball and formal statistical theory. "In general, they allow you to study fairly complicated processes that you can't really get at with pure mathematics. But also, sports are great for simulations — you can play 10,000 seasons overnight."

No one can afford to wait less than major league teams, which crave every extra run or victory they can wring from their $100 million rosters. John Abbamondi, the assistant general manager for the St. Louis Cardinals, says his team and about 10 others use simulations to evaluate potential trades and how they might affect the pennant race.

"It's all part of the statistical analysis that complements the more traditional scouting we do," he said.

Using computer simulations to explore in-game and other baseball strategies is by no means new. As early as 1958, a professor at the Massachusetts Institute of Technology programmed a behemoth I.B.M. 704 mainframe to investigate whether the sacrifice bunt was a smart play. (More on that later.) Simulators have since grown so complex that the most sophisticated one available to the public, called Diamond Mind, not only runs lickety-split on laptops but even considers minutiae like the effects of wind in individual ballparks.

Under what conditions is bunting advantageous? When does trying to steal make sense, and when does it decrease the chances of scoring? Questions like these turn out to be ideally suited to computer programs through which millions of iterations can smooth out the peaks and valleys of randomness, and converge toward a reliable approximation.

Known among formal statisticians as the Monte Carlo method, this approach takes spectacularly complex phenomena like weather patterns and stock performance and allows their behavior to be approximated, if not determined.

What are the chances of winning a game of solitaire? Rather than writing an equation that tries to take into account the trillions of trillions of possible hands and moves, a statistician can run a computer program that simply plays the game a few million times in minutes to see how often it wins. Dr. Morris says he has seen the Monte Carlo method used to improve computer graphics and explore gene sequences.

Like such competitors as Strat-O-Matic — which made its debut in 1961 with at-bats determined by cards and dice, and remains popular on the personal computer — Diamond Mind is designed to allow fans to play fictional games and seasons, exploring what-if scenarios that real life would be too slow and controversial to allow.

Take the age-old question of how much difference a team's lineup order makes. This issue so vexed the former manager Billy Martin that he once literally picked his Detroit Tigers batting order out of a hat.

Luke Kraemer of Imagine Sports, which owns Diamond Mind, programmed the simulator to force the 2008 Yankees to bat their best hitter and cleanup man, Alex Rodriguez, ninth — to see how scoring was affected. Mr. Kraemer got the run total not for just one season, which can fluctuate as much as 80 runs in each direction from simple randomness, but for 100 seasons — more than 16,000 Yankees games in all.

The result? The Yankees scored 747 runs per season, 40 fewer than their real-life 787. (Diamond Mind was so accurate that 100 seasons with A-Rod batting fourth averaged 789, almost dead-on.) Most research suggests that those 40 runs would mean only about four fewer victories, for a strategy no manager would ever consider; so the difference with Rodriguez batting third or fifth would be insignificant, and nowhere near worth the forests of trees that would give their lives to the ensuing sports-page debate.

Diamond Mind took its cuts at several other baseball knucklers, running 100 full seasons of games for each:

•

The intentional walk. This frequently used defensive strategy avoids dangerous hitters and can set up a double play, but it also awards a free base, and even the best hitters usually make an out. So is it smart in the long run? Diamond Mind found that it was not, though the difference was only about five runs per team per season.

•

The stolen base. Advancing from first to second puts the runner in scoring position, but he — and the rest of your hitters — will have a hard time scoring if he gets thrown out. Mr. Kraemer looked at a recent team that ran wild (the 2008 Tampa Bay Rays) and one that barely stole at all (the 2005 Oakland A's) and switched their mind-sets to see what happened. The A's scored 20 runs fewer, which probably says more about their players' inability to run in the first place. But when the speedy Rays stole sparingly, they increased their scoring by 47 runs per season — suggesting that perhaps the Rays were running too often in real life.

•

The sacrifice bunt. Is it worth making an out intentionally to move a runner from first to second? Forcing a team that hated that maneuver (the 2005 Boston Red Sox) to do it a lot cost them 19 runs per season. But making a bunting team (the 2008 New York Mets) avoid it also cost them — by 15 runs on average — suggesting that the Mets' managers, Willie Randolph and Jerry Manuel, used it quite intelligently. (The 1958 M.I.T. statisticians found that the sacrifice was rarely a good move; major league managers paid little attention.)

One problem with computer simulations is that no matter how realistically they might be programmed, they can say more about the programmer than baseball itself. A computer, after all, cannot feel human emotions like pressure or the will to hit in the clutch.

"We can run the experiment in the simulation environment and think we're measuring the effect of a great defense on a pitching staff, but it might tell us more about how we modeled defense," said Tom Tippett, who wrote the original Diamond Mind code in the early 1980s. "The simulation is real close to real-life baseball, but in the end it isn't real-life baseball."

After developing Diamond Mind into the industry standard, Mr. Tippett was hired a few years ago by the Boston Red Sox — a sign of how much some teams have come to value simulation research. While none will discuss exactly what they model and how, Mr. Abbamondi, of the Cardinals, said they could provide objective insight into how an offense might be affected by trading for a hitter in midseason; how many games that might improve the team; and how that hitter might improve or deteriorate as he ages. Many of these measurements come in the form of scenarios of increasing uncertainty, not unlike the projection of hurricane paths.

As Mr. Tippett suggested, however, simulations have inherent limits, and probably will not ever model baseball's vicissitudes of fate — how scrubs morph into all-stars and some teams just collapse. (Indeed, fans of the recent New York Mets would be relieved that some things defy re-creation.) Tony La Russa, the Cardinals' manager, who is a sure bet for the Hall of Fame, said the value of computer simulations in baseball tended to stop at the dugout entrance.

"There's way too much importance given to what you can produce from a machine," he said. "These are human beings, and I don't think any computer is going to model that close to what we deal with at this level."

That can be as true now as it was 25 years ago, when a Tank McNamara cartoon captured it best. A downtrodden manager peered over his computer. He asked plaintively, "But will it take the blame?"