MASTERSBALL

2010 Projection Methodology Guide – The Translation Process

Introduction

The following series of essays will be an explanation of the process used to create projections at Mastersball.com. Our process is continually updated based on new data, testing, and evolving thought processes. While we would be the first to suggest that there is no "right way" to do projections, we believe that our method is as rigorous and defensible as any out there today.

What Are We Trying To Create Here?

When we undertook the effort of creating this projection system, the first issue we had to deal with was exactly what we wanted to create. We started with the theory that there are two ways to approach a projecting a Major League baseball player.

1. A projection represents the statistical set that the forecaster expects each player to put up for the upcoming season, taking into account the numbers, splits, scouting, gut feeling, and whatever else the forecaster wishes to include in their thought process. This may involve consistent or inconsistent treatment of players at the forecaster's discretion.

2. A projection set should involve a thorough but consistent methodology for evaluating each player that takes into account their historical performance and is as free of speculation and bias as possible. Any qualitative commentary should be saved for accompanying material such as profiles or scouting reports.

It is important to realize that the majority of projections at the Major League level are for players with a fairly stable three-to-five year track record. There is unlikely to be significant variation in statistical sets. For the most part, these projections will represent a three-year average with some tweaking. There is not a tremendous amount of added value that any forecaster is going to add to the process. To be fair, there could be some, but it's not significant, and most of it is going to be purely their hunch.

Where the forecaster really earns their stripes is in players clearly on the rise, or on the decline. Should the forecaster regress everything towards a mean or project further rise or decline?

The most significant reason to use a hunch-driven product is if you trust the forecaster's judgment and you understand exactly what they're doing. After all, you're paying good money for their insights, and ultimately you want your projections to reflect their evaluation. This of course is quite reasonable, and no one could argue with it. OK, well I'm going to – and here's why.

Theo Epstein gave Julio Lugo four-years and $36 million and Brian Sabean gave Barry Zito seven-years and $126 million. These are two very stat-heavy organizations with vibrant,

smart decision-makers, and one assumes they have every resource they possibly could available to them such as scouts and statisticians. Yet these two teams completely swung and missed with a significant amount of money. What is the lesson of the story? It's that player forecasting is freaking hard. If anyone, anywhere offered you a real competitive advantage with their methodology, John Henry or Hank Steinbrenner would have them locked in a room with computers all around, paying them millions to guide their decision-making. The Yankees ended up spending nearly $250-$300 million in player contracts in the 2009 offseason. Do you think they wouldn't spend a small percentage of that money on a statistical whiz if they existed? Hell, the Red Sox got Bill James, the godfather of all of this, and while he's a valued part of their organization they trust their internal scouting and hunches as much or more than anything he puts in a spreadsheet.

Now back to the projection set again. So I've posited the theory that no one who does this is any good at it and if they were their work would be more classified than Iraq war plans. So what do we do about it? Is my goal to completely disavow you of any confidence in what we're doing? And the answer is precisely the opposite. The Mastersball team shares a common goal, and that is to be the best information provider in this business by clearly outlining a methodology, utilizing it, testing it thoroughly and often, and then focusing on the best analysis we can provide to go with it, completely with the knowledge of where the risk lies in everything we do. And even better, we combine this projection methodology with a complementary valuation and strategic methodologies which also captures all of the risks/rewards/issues we'll cover in these essays. The final result is not the world's most accurate projections necessarily (not that we'd know what those were if we saw them until after the season) or the perfect valuation system, but a process for managing your fantasy baseball team that day in and day out will give you confidence in what you're doing.

So, to quickly conclude on Part 1 of this Series, I'll walk you through what we are going to be doing this year to accomplish all that we've addressed above.

The first thing we've done is create our own proprietary method for translating each individual statistic from A-ball to the majors into neutral statistics – those being statistics each player would put up in a neutral major league park based on their major league and minor league output over the last 3 years. This has been (and continues to be) rigorously back-tested against actual player performance.

The second thing we've done is back tested over 50 years of major league seasons to come up with an aging mechanism to apply to players so that we take into account growth/peak/decline phases in players. As you might expect, it's not as simple as "players peak at 27 and decline by 31" – each statistic has its own way of developing, and our methodology takes it into account.

Third, we apply a weighting mechanism to statistics for each level to weight more heavily statistics earned at higher levels and then perform a three-year average based on weighting more recent years more heavily.

The end result is a system which takes into account every reasonable factor we can think of, allows us to project literally thousands of players (they won't all be in the majors immediately, but we know many of you out there are in deep leagues with minor league farm systems and we want to take care of you, too), and allow us to immediately pick up on changes in player performance via update statistics in-season.

We intend as part of our product to provide to you to continually test our assumptions, provide the results of that testing to our readers, and where appropriate update our methodology. This is the real benefit to us of being as transparent about our system as is possible – we'll get the feedback of smart people like you. Now, I'll be honest, we're not going to give away every last secret piece of information, but you'll know what we did, how we did it, and what it meant.

And lastly, what we'll do with the next many pieces of this series is take you behind the scenes of our projection methodology, no secrets, with a clear outline of the strengths and weaknesses of our approach and more importantly, what that means for you at the draft table and during the season, so that not only will you show up for your draft more prepared and armed with material than you've ever been, but with a greater understanding of the process and what it means for your fantasy team.

And as always, we'll welcome critique and comment, and will be more than happy to debate any and all ideas that come out of this.

The Translation Process

So to start, the #1 thing you are going to get with our projection set is that we are going to utilize a player's full 3 year performance history in their projections regardless of whether those stats were accumulated in the majors or minors.

League Translation

Multiple things must be done in order to accomplish this. The first is that we have to take into account the level of competition that the player's statistics were put up at. Obviously 15 HR in AA ball is different than 15 HR in MLB, for example. So, we went and back-tested several years worth of minor league and major league results in order to calculate how each statistic would translate to the major league level.

While all of the details of the testing are not particularly important for this purpose, the most important thing to take into account is that we only evaluated players who had accumulated enough PA or IP at multiple levels in the SAME season – the idea is to see how

their performance at one level translated to the next level when they moved up. Obviously, very few AA hitters are promoted directly to MLB, so rather than use a translation of AA to MLB, what we did was translate AA stats into AAA stats and then translate that result into MLB stats.

The findings were similar to those found elsewhere on the internet (you can Google Major League Equivalents if you want more background on the subject), though our testing did find some variances, and we use our translations for the purposes of these projections.

Park Translation

The second thing we need to do is take into account the park that statistics were put up in – clearly offensive production in Colorado in the 90's is different than San Diego in 2010, and in the minors, things like lighting and field conditions can be even more of a factor. Park Factors provide a multiplier to adjust statistics based on the parks they've been produced in. They are basically calculated taking the statistics put up in a park and then comparing those statistics to those put up on the road by the same teams to see how much the home park aids/hurts statistical output.

The Mastersball team has discussed the idea of park factors for years, and generally have been very hesitant to utilize it. Fundamentally, we've embraced the idea, but it has always seemed to be a very imperfect way to manipulate statistics. For example, lefty and righty hitters in Fenway Park and Yankee Stadium have very different experiences – to try and encapsulate this effect in one number is something we've stayed away from.

And to be very open, there are many reasons to take issue with the process – the data is rocked with sample size constraints, schedules are very unbalanced in the minors (especially for teams that might be strong in the early season and then have their players all called up later), and factors like travel might unfairly make certain parks look more friendly than they are.

However, after careful consideration of the following points, we have included these factors in our projection process.

For the most part, the adjustment is fairly minor – for example, say a park is 10% friendly to HR. Well, a player plays half their games on the road and half at home, so that means their HR are only affected by 5% (the road games played on aggregate in neutral parks). So for a 20 HR hitter than would be 1 HR. You're not making a buying decision based on 19 HR verses 20. And 10% is a significant park effect seen in very few parks.

We have the information. It's there, it exists, and it's quantifiable and makes sense based on what our eyes see – San Diego is a pitchers park, Philly a hitter's park.

The impact of park factors only affect players changing parks – a player putting up 35 HR in Fenway in 2009 and then returning to Boston in 2010 would expect 35 HR again. So the sample of affected players at the major league level is not large.

However, minor leaguers do move around from team to team and this measurement would aid in the understanding of performance changes from team to team.

Lastly, and most importantly we asked the question: Would our projections be more accurate, appropriate, and practical using the information, or not using it?

It was the answer to #5 that won the day. We have the information, its imperfect, but it's better in our opinion than ignoring it.

Other Translation Considerations

The following discussion is specific to our analysis of BABIP but really serve as proxy for an entire discussion of the projection process as a whole.

The BABIP Quandary – Part I

Up to this point, there is nothing particularly noteworthy about what we've done to the translations. But then we get into the subject of how to translate hits.

Our general translation of hits shows that in the aggregate, the translation percentage for hits is roughly 80% from AA to MLB and about 88% from AAA to MLB. This is primarily due to the fact that strikeouts increase significantly from the minors to MLB and as such, balls on play decrease and therefore hits decrease. Further, at each level we see a decrease of 5-10% of BABIP from the previous one.

However, there is an additional issue to take into account – there exists in essence a "survivor's bias" – that is, that the players who played in multiple levels in a given season likely were 1) promoted because they had put up a high BABIP which aided their AVG or 2) demoted because of a low BABIP which hurt their AVG, and as such, the regression towards the mean is due at least in part to regression that was going to happen anyways, regardless of level change. In fact, even when looking at major league hitters, we see them regressing towards the mean in a majority of cases. So at the end of the day, we may or may not know anything about the player's BABIP ability looking at their minor league numbers, because we are uncertain how much of the change from level to level is due to expected regression and how much is due to differences in talent level.

Put another way, I'll show this in a chart, with all minor leaguers who moved a level being sorted into 4 groups: those above .330 BABIP, those between .300 and .330, those between .270 and .300, and those below .270. *(Note, this is only an example, we are not evaluating players in aggregate but based on each minor league level)* The group as a whole saw a 8.0%

decrease moving up a level. We will present what happened to their BABIP at the next level in the first column and compare it to the average of the group in the second.

Group	Based on Group	Straight-Lined
>.330	-12.0%	-8.0%
>.300	-5.5%	-8.0%
>.270	3.0%	-8.0%
<.270	9.0%	-8.0%
Total	-8.0%	-8.0%

So one has two options here, the first is to take players at each level and "translate" their BABIP based on their BABIP at the lower level, thus adjusting players differently. The second would be to adjust everyone on a straight-lined basis, just adjusting everyone based on the average decrease. A coherent argument can be made for both.

The group adjuster makes the argument that we simply cannot accept that a .350 BABIP and .250 BABIP players are going to react the same way that we're aware of the regression inherent in the projection and should modify for it. Further, they point out that the adjustment makes the focus of the translation a batter's underlying skills and less about a more elusive factor like BABIP. Further, the group adjuster argues that this will smooth out outlier performances which generally make projections more difficult.

The Straight-Lined adjuster makes the point that modifying their actual performance to some sort of theoretical adjustment basically renders their actual performance meaningless in the big picture of actually creating the projection. Further, the straight-liner points out that we're using multiple years for the purpose of producing the projection, so any regression will be caught up once multiple years are averaged into set. And, to counter the group adjuster's final point, the Straight-Liner argues that those spikes are a part of the record, no matter whether they are an outlier or not, and that pretending they don't exist is no way to properly assess a player going forward.

The BABIP Quandary – Part II

The second part of the issue is dealing with the different levels of the minors and the BABIP translation. Basically, if a batter puts up a .280 BABIP in A-ball, the truth is that we know very little about what they would look like as a major leaguer. Once you translated them through several levels, you'd predict them to be a .300 (average) major league BABIP hitter.

But if a batter puts up a .280 in AAA ball, by that point we do have a more established translation to the majors, and you'd expect his BABIP to be around .290 or so (again, expectation of significant regression to the mean). And herein lies the problem – it seems entirely counterintuitive to take a .280 BABIP season in A-ball and a .280 BABIP season in

AAA ball, and actually translate the A-ball season to a HIGHER MLB BABIP because of the fact that we are completely in the dark about the A player, and we know a bit more about the AAA player, but that is what the numbers suggest we should do.

The real question is whether or not we prefer a common-sense approach even if it means it might be less supported by the research, or whether once we've done the research and identified the issues at hand whether we should include the methodology even if it doesn't feel good. (Note, numerically it makes perfect sense, it just….well, has that feel to it….)

The BABIP Quandary – Part III

Then, the question even further moves to major leaguers. We also know that similar to the minor leaguers listed above those major leaguers in the groupings listed above also regress to the mean in a similar fashion (this research will be provided as part of site content). The question again lies in whether or not we adjust their BABIP for the regression likely to occur or whether we just let their performance speak for itself and allow the process of evaluating multiple years to do its thing.

So now let's back away from all of the technical talk for a minute and just focus back on the overall exercise here. We are trying to show a translation of a player's statistics to create a MLB Neutral stat line which can then be included within a three-year average to create an appropriate projection for each player.

We understand that most players will regress to THEIR individual mean in every category during their three year period under consideration, and as such the debate between manipulating statistics significantly based on statistical study verses properly adjusting them but otherwise leaving them to stand on their own is one that has merits on both sides of it.

Having tested this both ways and having reviewed the 2009 projections in depth in preparation for doing the 2010 projections, we have come to the conclusion that we should regress BABIP to the mean for minor leaguers and leave the BABIP in place for major leaguers, issues raised above notwithstanding. This might be seen as inconsistent treatment. Actually, it is. However, our take on the matter is that we should have a very good reason before we discount actual MLB performance and try to re-create it. There is regression to the mean, no doubt, but a player's history over time in the majors should stand on its own rather than create some sort of hypothetical world. Whereas for minor leaguers we are trying to estimate what their minor league performance would mean as a major leaguer, and as such the regression methodology more accurately reflects what MLB results would have looked like.

We've run this all three ways (regress no one, regress everyone, regress MiLB) – and this is the process that "feels" best. Over the coming months and years we will keep messing with this and try and improve it if there is a better method.

Applying Proper Aging Theory

One of the things we set out to do in producing our projection system was to have it properly address the affects of a player's age in the year they produced the statistics so that stats produced at younger and older ages could reflect different things. Typical projection systems force the user to qualitatively evaluate a player's age, we wanted to take it a step further and do the legwork ourselves.

Aging Translation

So what we did was take 75 years of major league baseball history and evaluated the growth and/or decline in each category we address from playing time to caught stealing that players experienced at each age. So we took every 24 year old and said "What happened to their statistics going from age 24 to age 25". And we did this for every age covered by MLB players during the period, and the results were striking.

What we found was almost every category had a growth curve that was very closely fit by a logarithmic curve and that there was a very clear, evident aging pattern that player's were following. Now, it isn't as easy as "players peak at 27 years old" – each statistic had a different aging pattern, and as such we utilized that aging curve for the purposes of determining how the player would age. For example, the curve might show that the average 28 year old will have a HR rate 99.5% as much as a 27 year old (fictitious example).

Using Aging Methodology with the Three Year History

So what we wanted to do was take a slightly different approach to the process than other methods do while retaining the core values of the 3-year average process.

The first thing we do is take the translated MLB neutral statistics for each of the last three years. What we then want to do is evaluate what they tell us. For example, we have a player who was 25 years old in 2007, 26 in 2008, and 27 in 2009. Each season's statistics would be translated a bit differently as it related to the year they were produced in.

But what do we care about? We care about 2010!! So what we do is take the stats put up in each year and translate them based on our aging curve to what they represent the player would put up in 2009. So we take that 25 year old in 2007 and say, "these statistics put up in 2006 on average represent X in 2010", same for 2008 and 2009. So we're not interested in the raw statistics, we're interested in what those statistics would represent aged forward to the season we're actually projecting.

And what we find is that most players do indeed age in line with our aging curves. Not perfect, of course, but the skills really do seem to translate as we'd expect. This allows us great perspective in looking at player trends and seeing which are real changes in skill and

which are the typical improvement/decline we'd expect from a player as they go through their career.

Significant Translation Considerations

The most important thing to address as a weakness/concern/consideration when looking at our evaluation process is something that has come up in other sections of this guide, and that is the common happening of regression to the mean and in pitchers, the risk of breakdown due to ineffectiveness or injury that certainly has aging elements to it but is somewhat a hazard of the trade.

For offensive players, this ended up not appearing to be much of a concern. The type of regression we typically see in these players is more to their individual mean than to a player-wide mean, and as such when we did the three year average, any regression that occurred washed out through the averaging process.

For pitchers, however, when we first ran this, we saw that basically every age group showed some level of decline, because of how often pitchers get hurt, lose it, etc. And there's no doubt that if we wanted to show that level of decline across the board, we could justify it from a statistical standpoint. But the problem with that is that projections should present a pitcher's skill set, not their skill set taking into account what might happen if they blow up. We already take that risk into account by the way we allocate dollars to pitching and through our overall strategies. For 2009 we tried to ballpark the adjustment because we felt it was important to include an aging adjustment. Our internal review for 2010 didn't really bear this out in any meaningful way. It is true that pitchers on either end of the spectrum (extremely young or extremely old) likely require an adjustment – the problem was that trying to get these 10% of the population correct was hurting the accuracy on the middle 90%, and we still struggled with the 10%, because as a whole pitching projection accuracy is a very inexact science. We would rather point out the upside or downside of a player in a profile and allow the skill indicators to control the projection for pitchers. You still should see the affects of age in the three year history. This is also an illustration of where some regression for pitchers can make a lot of sense (we will cover regression methodology elsewhere in this series).

For hitters this simply doesn't manifest itself in the same way and the aging approach is entirely appropriate.

Other Aging Considerations

I think the most significant issue one could raise with this approach is to point out that it's silly to expect that each player will age the same way. And I will tell you that we felt the exact same way until we looked at the three year histories to see how many players really do seem to follow their aging path. Given the amount of translation a player has gone through by the time we're evaluating their histories (park, league, year, age), the results are striking.

But we will concede the point that a 32 year old catcher likely ages more severely than a 32 year old first baseman from a wear and tear standpoint. Ideally you could look at many issues – games played, body type, position played, drafted out of college, high school, etc. And maybe someday enough data will exist to really be able to implement all of those differences. But we are very comfortable that even though this isn't perfect, it encapsulates the factors we are trying to include very well.

Player Evaluation Considerations

The second significant consideration is that this has changed the way we look at all players now. For example, instead of saying, "The player hit X HR in Y AB in 2007, we are saying, X HR in Y AB means he'd hit __ HR in __ AB in 2010", and that 2010 translation is what we're evaluating when we look at his three year history. So we're no longer evaluating a player based on the incline/decline of his raw statistics, we're judging whether or not he is following the aging curve that his previous stats projected he would.

This is actually a very cool thing because in essence, every player presents three different career aging paths in their three year histories, and we fit their 2010 projection to the path that makes the most sense based on their history.

When we discuss the projection/profile process we'll address this once more.

At-Bats and Innings Pitched

Most of the attention with respect to projections focuses on historical performance rates and how future performance can be derived using that as a foundation. While that is obviously a very important aspect of projection methodology, somewhat overlooked is the other element, the amount of anticipated playing time, usually expressed as at bats and innings pitched. The following is a review outlining our process with respect to assigning playing time.

Hitters

Though the metric that most fantasy enthusiasts look at is at bats, we actually base playing time on plate appearances, and then calculate how many at bats each player garners after backing out their expected walks.

Some prognosticators like to set up a positional grid and assign playing time based on the percentage of the time each player is expected to accrue at each position. While this system is easy to automate, there are some shortcomings such as needing to adjust for position in the batting order and it is human nature to assign percentages in intervals of 5%. While it may seem trivial, as a means to illustrate its importance, if you assign Alex Rodriguez 20 additional at bats, and prorate the rest of the stats accordingly, his value increases by $2.

As such, we decided to use a method a bit less automated, but admittedly every bit as subjective as estimating the percentage of playing time at each position.

The first step in the process is coming up with a target number of plate appearances per team, based on recent history. Next, the plate appearances for the anticipated starters are projected for each position. This is done by reviewing recent year's playing time, considering injury risk and quality of backups. The remainder of the playing time is assigned to the reserves, again considering their historical level in concert with the available time left by the starters. Care is taken to distribute the plate appearances logically with respect to insuring each position is covered by a player able to play the position. We have the model automated to adjust all the stats dynamically based on the number of plate appearances. As teams sign or release players, or it is apparent an un-projected player will make the team we will adjust the at bats globally, making sure the overall total remains logical as well as keeping adequate coverage at each position.

Pitchers

Projecting innings is not a significantly different than plate appearances. The target number of innings is simpler, just 162 x 9 or 1458. Obviously, road teams do not have to work the 9th inning when they are behind and there are extra inning games, but these come close enough to washing out that aiming for 1458 is fine. Not to diminish their relevance, but these last few innings most often are assigned to the relievers that will not be drafted, even in the deepest of leagues.

Innings for the projected starting-5 are done first. This is based on history and current situation. On the average, most team's starters work about 1000 innings, so anything short of that based on the first five is assigned to the most likely spot starters and swingman. Next, the innings for the top relievers are estimated, based on history and current bullpen makeup. The rest are distributed amongst the remaining relievers, once again using history as a basis.

Passing the Smell Test

Perhaps one of the more under-discussed aspects of a projection set is how logical it is in terms of representing the eventual composition of the statistical pool. Some believe the perfect projection set predicts as many hits and runs scored as it does hits runs allowed. They feel each team's W-L should add up to 162 and globally, there should be as many wins as there are losses. And while intuitively this makes sense as year to year, the cumulative stats are fairly close, the recent drop in power notwithstanding. However, there is a big problem with forcing these constraints at the expense of looking at player in a vacuum and accepting how the result ends up globally.

The problem can best be described as baseball's version of natural selection, survival of the fittest. At the beginning of each season, the poorer players will be replaced by better

players and better players whose playing time was lessened by injury will play more. The end result is a predicted global statistical pool greater than the previous year. By the end of the season, some better players will get hurt and will be substituted for by inferior players, driving down the accumulated stats to levels below originally predicted. If this lower level is forced upon the projected player pool, players will be short-changed at the expense of mythical global accuracy.

That said, there are still some logical parameters that can be met in a global sense. Early on in the projection process, some common sense assumptions can be made, if it is obvious a team will still be making a move to fill a hole in their lineup or pitching staff so the cumulative at bats and innings do not have to be within reason of the eventual target. But as the spring progresses, playing time issues should clarify, and the total plate appearances and innings should be assigned in a sensible manner.

It is fairly difficult to use a strict Pythagorean method to compute wins and losses for as explained above, the number of runs scored is not going to be the same as runs allowed, which is the basis for Bill James' estimation of wins and losses. In addition, there is some bullpen bias as the starters generally enjoy the benefit of having the better relievers follow them. However, we do employ a modified Pythag method that also takes a player's history into account. The end result is a team record very close to having 162 decisions, but we don't sweat it if we fall a little short of or exceed 162 games. More importantly, we don't adjust to have the global number of wins equal the global number of losses. Here, we look at how each pitcher "should" do based on their expected runs allowed and offensive support and not worry about how they fit into the big picture.

Saves is the only projected stat that is basically all subjective, with the guideline that all teams will average between 45% and 55% of their wins being saved. To facilitate, we just assign a save to 50% of the team's projected wins. So a 70 win team will be projected to have 35 saves, with 30 or 31 going to the closer assuming they have a viable candidate. One reason teams only win 70 games is a weak bullpen, so we don't need to spread around more than a few more saves. Teams projected to win 90 games will get 45 saves, with 37 or 38 going to the closer and the others distributed to the setup guys. Teams projected to win close to 100 games might have their closer assigned 40 or 41 saves. Based on this, the implication is closers on better teams are more likely to get more saves which makes intuitive sense. However, many preach saves are a crapshoot and good closers on bad team will also get saves, sometimes more than closers on good teams. The reason for this lies in the 45%-55% range above. A team winning 70 games with 55% getting saves results in the closer getting about 35 saves. A team projected to win 90 games with only 45% resulting in a save will cost their closer a handful, knocking him down to the 35 range. But the problem is the 45%-55% is happenstance and not predictable, so all we can do is split the difference and assign saves based on opportunity accordingly. If we are not confident that one closer will get the vast majority of saves, we will distribute them amongst the viable candidates in proportion to the likelihood we feel they will fall.

The only other logical test we subject to the projections is looking at the percentage of RBI versus runs scored – teams usually average between 94% and 96%. If the result is extreme, we try to identify the reason. It is usually that a power hitter that had a large number of RBI is being replaced by a speedster that knocked in fewer runs but scored more or vice versa. This is one of the few instances an adjustment is made by hand, and only in extreme cases.

Final Weightings

The first thing we wanted to do was make sure that we valued statistics put up at higher levels more than at lower levels. So we weight MLB statistics at 1.0, AAA stats at .75, AA at .5, A/A+ stats at .25. We consider Japanese stats to be the equivalent of AAA performance.

Then for the three year weighting, we weight 2009 1.0, 2008 0.7, 2007 0.5. This is in line with most weighting systems of its kind. Remember, we're not weighting the raw statistics; we're weighting the translated-forward-to-2010 projections for each year.

The Artwork

Then we go back and we evaluate whether the yearly weightings are appropriate. A player may have played hurt for much of a season. Probably not all that proper to weight that season as fully as we otherwise would. Or a player experienced a significant skills growth or decline that we feel is permanent. We'd make the adjustment to depress the weights of earlier seasons. We'll try and address this through our content and on the message board, but we're not simply feeding these to a computer and forgetting about it. There's thorough evaluation of every player and their track record before the projection goes out. And if we don't like the projection at that point, we simply modify it until we do.

After all that is done, and we have an MLB-neutral projection taking into account everything we know, we apply a projected 2010 park factor for each team (this is the three year process described above for every team except the New York teams, Washington, and Minnesota). For the teams without a three-year history we regressed them somewhat to the mean but took into account all of the information we had.

And then we have a projection.

A final note about the profiles. You will see them constantly referring to skill declines or inclines. Please note that these reflect the major league and minor league translated stats of a players record, and they reflect the aging history we expect a player to follow. So it is not as easy as looking at the major league stats. Please post on the board or send an email if the projections or profiles don't seem to make sense, we'd be delighted to explain anything that doesn't follow. We hope that in including so much in our analysis that our projections and profiles really give our users an added edge.