

The Topsy-turvy Factor in College Football

Discussing and disputing the games of the season against the games of the past is a time-honored tradition of sports fans. Passions are born and beer consumed when reminiscing about a particular season being “wild” or “topsy-turvy.” Sportscasters become increasingly apoplectic each week in their attempts to describe this situation. Maybe this increases ratings. But what is it, exactly, that makes one season a ho-hummer (except if your team wins, of course) and another season turns into a chaos of triumph and tragedy. In these seasons of upsets and missed field goals, we are mesmerized by watching the mighty fall, over and over again, to the underdogs of Backwater U.

Since early man (and we know it was a man, for reasons not covered in this article), threw a rock and challenged his buddy to throw one farther, athletic accomplishments have always been lovingly recorded by those of a more analytical nature, the statisticians. Joe the Gladiator would have been just a smear of graffiti on the Coliseum wall if the scribes had not kept track of the number of lions, fellow gladiators and assorted Christians he dispatched in the prior season. Sports cannot survive without statistics.

So what kind of statistics do we have that describe not wins and losses, but the nature of a single week or season? How to quantify and define a “wild” season over another, more predictable season? Thanks to Dr. Autar Kaw, professor of Mechanical Engineering, and Dr. Ali Yalcin, associate professor of Industrial & Management Systems Engineering, we now have a metric, a number (always positive!) that will settle arguments and preserve friendships for seasons to come.

“Imagine a bunch of football fans sitting around the couch on a Sunday and talking about what a wild week it was,” says Dr. Yalcin. “One claims ‘It was the wildest he had ever seen in college football!’ Another fan counters ‘It wasn’t all that wild, this team lost to that team that was totally predictable.’ So our math ensures that instead of getting into a brawl and hurting each other, they can actually calculate it, put a number on it.”

Dr. Kaw explains how he approached the definition of topsy-turvy, and we are not talking about the Middle English *terven*, to overturn. He wanted to quantify, or assign a value to a season that would immediately indicate its volatility or predictability. “I thought the more the guy talked about a wild week or a topsy-turvy week, the more people would tune in to see what the heck was going on with the college report. So, it was interesting figuring out how we can quantify this. That’s how it all started in 2007,” says Dr. Kaw.

Weeks make up a season, so it was necessary to assign a number to a week. Dr. Kaw calls this number Week TT Factor. This weekly number is handy when our couch-loving fans get together. Instead of one-upping each other in volume, eye-rolling and potential violence, they can discuss the Week TT Factor, a number between 0 and 200.

The resulting Season TT Factor is calculated after each week to gauge how topsy-turvy the season has been so far. At the end, the Season TT Factor is the defining value of the volatility, disarray, chaos or whatever mathematical term you choose for the entire season. According to this methodology, season 2007 came in with a whopping Season TT of 50, and Season 2004 with a snoozer of 33.

The complete Week TT Factor and Season TT Factor figures are available on the website at: <http://www.eng.usf.edu/~kaw/ttfactor/index.html>. This is one of the most popular websites at USF, and has produced some topsy-turvy bandwidth issues of its own.

Dr. Kaw explains how the weekly topsy-turvy figure is arrived at. “At the end of each college football week, the Associated Press (AP) poll rankings are calculated by polling 65 sportswriters and broadcasters across the nation. Each voter supplies his or her ranking of the top 25 teams. The individual votes are added by giving 25 points to the first place vote, 24 points to the second place vote, etc. The addition of the points then produces the list of the AP top 25 teams of the week.” Those team rankings are compared to the previous week’s rankings. The difference between the current and previous rankings is calculated and then squared. Why squared? “It allocates proportionately higher importance on bigger week-to-week changes in rankings for a given team,” explains Dr. Kaw. These squared numbers are added together and then normalized by the average of the lowest and highest possible value of the sum and multiplied by 100. This produces a Week TT factor between 0 and 200.

The other teams could make trouble for us if they win
—Ugoi Berra

So before you launch your spreadsheet program, be aware there are some serious pitfalls. What about the unfortunate teams who fall out of Week 3 rankings? And those lucky underdogs who were unranked in Weeks 1 and 2, but have now joined the, ah, ranks of ranked teams? This involves some serious drawbacks concerning the fact that a paltry one or two votes is not really a true ranking, not all teams get unranked, and a host of other issues best covered in the articles on the website. Normalization numbers, formulae and other measures of disarray are clearly defined; the numerically-challenged among us can actually get a grip on how the professors worked out these dilemmas.

In reality, the topsy-turvy factors are a measure of disarray. It does not measure individual teams’ chances of winning. Because the base data is actually a set of opinions (albeit expert opinions), as opposed to win-loss statistics, the formula describes the changes between what is predicted and what actually happens. And if graphs showing parallel trends make your heart stop, be sure and check out the comparison on the website between the AP poll and the USA Today poll. Using the same formula, Drs. Kaw and Yalcin used the rankings by 63 head coaches in Division I-A, the USA Today experts. Using the same 25 votes for first place, 24 for second, the top twenty-five teams were created. The difference between using AP and USA Today for calculation of TT factors was less than 5%.



If you are still a little fuzzy on this concept, try Dr. Yalcin’s simplified explanation. Five teams, A, B, C, D, and E are ranked by experts. To make it easy, A gets the most votes, and E gets almost none. Teams C and D have just a couple of votes difference. It is a far bigger catastrophe if A gets thrashed by E, than if C gets beat by D. The disarray of A choking is measurably greater than C losing to D. And to prove all of this, if the rankings remain exactly the same, and the projected teams win, then the TT factor is a big fat zero. While the chances of this are practically nil, it proves the methodology. A number above 54 is considered “very topsy-turvy” and under 30 is “predictable.”

The individual weeks and seasons since 2002 are all available on the website. This methodology can be applied to anything that is “ranked” such as baseball, tennis and basketball. The bad news is that the TT factors, both weekly and seasonally, are random - maybe not as random as roulette, but still, nothing here to take to the bank. But, if you are willing to bet that in 2007, week 5 beat the daylights out of week 5 in 2004, you can take that bet to the bank.

by Janet Dawald