Predicting the NCAA Men's Postseason Basketball Poll More Accurately

John A. Trono and Philip Yates
Saint Michael's College,
Colchester, VT

Tech report: SMC-2015-CS-001
(Presented at the 27th European Conference on Operational Research, also known as EURO 27, July 12-15, 2015, Glasgow, Scotland)

Abstract

A previous study investigated how well a linear model could predict where teams would be ranked in the final NCAA coaches' poll (for men's basketball) which is announced right after the post season, single elimination, championship tournament (known as *March Madness*) has concluded. Monte Carlo techniques were able to improve upon those results, which were obtained via a weighted, linear regression model. This Monte Carlo approach produced a model whose Spearman correlation coefficients were roughly equal to 0.85 for the top 15, top 25 and top 35 teams, respectively, with regards to said final poll. This article will describe a non-linear model that is approximately 10% more accurate than the previous model, and incorporates Zipf's law – and a quantity known as the Tournament Selection Ratio.

Introduction

"It's tough to make predictions, especially about the future." – Yogi Berra

The men's National Collegiate Athletic Association (NCAA) month-long basketball tournament, also referred to as 'March Madness', continues to increase in popularity, especially with the appearance of Warren Buffet's billion dollar bracket challenge in 2014. With this additional monetary incentive, everyone now has a vested interest in trying to predict the winners throughout the entire tournament. However, once that tournament's champion has been crowned, the task of ranking the top 25 teams, for the final ESPN/*USA Today* poll, falls to a relatively small group of coaches, who have been asked to cast their ballots to generate this subjective ordering.

No one, to our knowledge, had ever reported on trying to predict how that final ranking would turn out, which prompted us to investigate several statistical models that might accurately forecast where teams would be ranked. These results were published in a recent issue of *Chance* [4], and even though we were fairly satisfied that our model did successfully replicate how the coaches vote, several unexpected insights led to the creation of several new approaches – the most accurate of those is roughly 10% closer to the final polls than our previous work. Therefore, after providing a quick overview regarding the content of our previous paper, we will describe these insights and how they influenced our latest models. (The results from the National Invitation Tournament (NIT) are also incorporated into both our previous and our newest models; more details on the specifics of how those results are included will be described later.)

Recap

In our earlier study, we relied on several objective team quantities, besides how many wins were earned in the NCAA postseason tournament, when designing our linear regression models – where the dependent variable was the total number of votes a team was awarded in the final coaches' poll. The most accurate model utilized a team's tournament victory count plus one – so that a zero count would represent teams that were not invited to the tournament – along with each team's winning percentage and power rating [1]. Using a weighted, least squares regression model (iterating until the estimated regression coefficients became stable), multipliers were determined for these three independent variables. Finally, the Spearman Correlation Coefficients (SCC), for the top 15, top 25 and top 35 teams, were computed to determine how closely our linear regression models matched the actual poll.

During this study, we realized that we were actually more interested in accurately matching the poll positions than the specific vote total for each team appearing in the final poll. We took advantage of the information we had gathered so far, and then initiated a slightly different approach. By employing Monte Carlo simulation techniques, we began to randomly generate weights (for the three independent variables) that were relatively close to the values selected for our most accurate regression model (referred to as Model-2 in our earlier article [4]). A separate computer program was then used to generate a large number of these random weights, and to evaluate the results – with the program also recording the weights that maximized the sum of the three aforementioned SCC values.

Much to our surprise, that program discovered 26 different sets of weights which produced the same SCC sum of 2.57253 – as did the average of those sets of weights. After examining the sets, we also noted that the particular weights weren't as important as the relative sizes between each of these weights since our linear prediction equation was no longer attempting to match the actual poll's vote total (for each team). In fact, as long as the weight for tournament wins was roughly 5 times larger than the power rating, which itself was roughly 2.64 times larger than the winning percentage weight, the SCC sum was the same. We called this set of average weights MC-Best (since it represented the best results from our Monte Carlo search). These results were produced by comparing how well our models corresponded to the actual polls in our training data set (1993 to 2007). We then evaluated how our most accurate model (MC-Best) performed on the subsequent, final coaches' polls (2008-2012).

The MC-Best model was very accurate in the first three (of the five) years that followed the training set years. However, the 2013 NCAA tournament, which completed just after our previous article had been accepted for publication, illustrated some limitations of the MC-Best model – as did the final polls following the NCAA tournament in 2011 and 2012.

Improved Prediction Models

The combination of two crucial, independent insights led to an initial, exploratory model that surpassed the prediction accuracy of the MC-Best model. The first of these insights occurred when one of us came across ‘Zipf’s Law’ while reading a work of fiction (by Robert J. Sawyer) close to when the NCAA tournament’s Final Four matchups were about to be played in 2013. In the story, this ‘law’ was explained in the context of the relative number of occurrences of words in a large body of text where typically the most popular word appears twice as many times as the second most popular word, three times more than the third most popular word, and so on.

The most direct way to represent these relative frequencies would be the following set of values: 1, 1/2, 1/3, 1/4, 1/5, etc. These values could be summed, and normalized, to generate each word’s probability of appearance. However, these specific values do embody Zipf’s law, and seeing this particular set of values made us wonder: what if the function related to the number of NCAA tournament wins was non-linear? Given the patterns observed in previous final polls, one conjecture was that perhaps each NCAA tournament win was incrementally worth a little more, to the coaches who vote in the final poll, than the previous one. If this was the case, then the fractions listed above would follow that progression.

The new model was then to reward a team that was invited to the tournament with a bonus of 1/7, a team with one win would be awarded a bonus of 1/7 + 1/6, and so on up to the tournament runner-up’s bonus (1/7 + 1/6 + 1/5 + 1/4 + 1/3 + 1/2); the NCAA champion’s bonus would be 1 larger than the runner-up’s bonus. (And, no invitation implies a bonus value of zero.) This bonus would be added to some quantitative team attribute, presumably normalized to be in the range zero to one. This seemed to be a reasonable range for adding said bonus values to – to effectively predict the final polls being investigated here. The additional +1 bonus, above the runner-up’s bonus, would also guarantee that the champion would sit atop the final poll, using this approach – if this team attribute was constrained to remain in the zero to one range – when making a prediction concerning how the final poll might turn out.

The second insight was observed from recent final polls: most teams with X wins typically remain ranked above other such teams, as ordered in the coaches’ poll that was taken **before** the NCAA tournament began (which will now be referred to as the ‘penultimate poll’). Table 1 illustrates this behavior for those teams who have won two or three tournament games, from 2007 to 2014, ignoring any ‘play-in’ games that reduce the field to 64 teams. (This behavior is harder to validate for 1 win, since roughly half of those teams are not ranked in the penultimate poll. However, teams who appear in the penultimate poll, and who earn 1 tournament win, typically appear in the final poll as well.)

This observed pattern of tournament win/voting behavior held true to form perfectly in all eight years (in Table 1), for those teams who won three games, except for Duke leapfrogging over Ohio State in 2013, Arizona doing likewise (to Florida) in 2011, and similarly for Louisville (overtaking Xavier) in 2008. Perhaps the specific tournament game scores, or opponents they defeated, impacted the coaches' votes in those cases. (Teams with four wins, i.e. those that lost their Final Four contest, behaved exactly as this second insight would forecast – for this eight year span.) Teams with two tournament wins followed this pattern as well in 2011, 2012 and 2014: only one team (Kansas) near the top of this group in 2013, two teams (Xavier and Cornell) closer to the bottom of this list of teams in 2010, two teams near the top in 2009 (Kansas and Syracuse) and just Tennessee (or Wisconsin) in 2008, were slightly 'out of order', while in 2007, Butler slipped past UNLV – as did Tennessee (moving ahead of Vanderbilt). In years where two teams earned two tournament wins, and both were unranked in the penultimate poll, these teams ended up being only separated by 1 vote in the 2012 final poll, by 6 votes in 2013, by 7 votes in 2011, and by 18 votes in 2014.

	NCAA wins	Final Poll	Pen. Poll		NCAA Wins	Final Poll	Pen. Poll
2014				2010			
Arizona	3	5	5	Kentucky	3	5	2
Michigan	3	6	8	Kansas State	3	7	9
Michigan State	3	8	12	Tennessee	3	9	13
Dayton	3	18T	NR	Baylor	3	10	21
Louisville	2	9	3	Syracuse	2	8	4
Virginia	2	10	4	Ohio State	2	11	6
Iowa State	2	11	9	Purdue	2	12	11
San Diego State	2	12	11	Northern Iowa	2	13	24
UCLA	2	15	23	Xavier	2	14	33
Baylor	2	18T	28	Cornell	2	17	29
Tennessee	2	22	NR	St. Mary's (CA)	2	19	26
Stanford	2	24	NR	Washington	2	21	30
2013				2009			
Duke	3	5	7	Louisville	3	5	1
Ohio State	3	6	6	Pittsburgh	3	6	4
Florida	3	9	12	Oklahoma	3	7	8
Marquette	3	11	16	Missouri	3	8	9
Indiana	2	7	4T	Memphis	2	9	2
Kansas	2	8	3	Kansas	2	10	13
Miami(F)	2	10	4T	Duke	2	11	5
Michigan State	2	13	9	Syracuse	2	12	15
Arizona	2	14	20	Gonzaga	2	13	10
Oregon	2	15	24	Purdue	2	14	18

LaSalle	2	24	NR	Xavier	2	15	22
Florida Gulf Coast	2	25	NR	Arizona	2	24	NR
2012				2008			
Syracuse	3	5	2	Texas	3	5	7
North Carolina	3	6	5	Louisville	3	6	13
Baylor	3	8	10	Xavier	3	8	12
Florida	3	9	21	Davidson	3	9	23
Michigan State	2	7	4	Tennessee	2	7	6
Marquette	2	10	11	Wisconsin	2	10	5
Wisconsin	2	12	13	Stanford	2	11	11
Indiana	2	13	17	Michigan State	2	13	20
Cincinnati	2	18	29	Washington State	2	15	21
NC State	2	20	36T	West Virginia	2	17	36T
Xavier	2	24	NR	W Kentucky	2	20	36T
Ohio Univ.	2	25	NR	Villanova	2	24	NR
2011				2007			
Kansas	3	4	2	Kansas	3	5T	2
North Carolina	3	8	7	North Carolina	3	5T	4
Arizona	3	9	15	Memphis	3	7	5
Florida	3	10	12	Oregon	3	8	12
Ohio State	2	5	1	Texas A&M	2	9	9
Duke	2	7	5	Pittsburgh	2	10	11
San Diego State	2	11	6	Southern Illinois	2	11T	15
BYU	2	13	8	Butler	2	13	19
Wisconsin	2	15	13	UNLV	2	14	18
Florida State	2	19	36	USC	2	15	25
Marquette	2	20	NR	Tennessee	2	20	32
Richmond	2	21	NR	Vanderbilt	2	21	31

Table 1 – Comparison of team ranks in the final two polls.

When these two insights were combined into the ZPF model – by adding said bonus value, as derived from Zip's Law, to the normalized, penultimate poll's vote total – all three SCC values were close to 0.9 (or higher), when applied to the same training set (1993-2007). This is quite an improvement over the three SCC values (all roughly equal to 0.85) that were observed with the MC-Best model – and these results (for the ZPF model) were produced before any additional improvements were incorporated to account for NIT tournament wins. (The calculated SCC-15 value would remain unchanged since the NIT champion hasn't been ranked higher than the #20-#25 range in any final poll after 1980, and more recently, said champion now typically appears in the #30-#35 final rank range.)

Generating Other Possible Bonus Values

Since this new approach, as employed in the ZPF model, produced a significant improvement over the MC-Best model, it seemed reasonable to consider that perhaps other bonus values should be evaluated because there is no inherent reason why the values motivated by Zip's Law are necessarily the most accurate predictors for this new strategy. Table 2 contains 12 sets of bonus values, and we will briefly explain how these dozen came about. (Three other sets of bonus values were generated, all involving more convoluted generation strategies, but since none of them produced results close to the best in Table 2, they were omitted.)

Name	0 Wins	1 Win	2 Wins	3 Wins	4 Wins	5 Wins	6 Wins
ZPF	0.142857	0.309524	0.509524	0.759524	1.092857	1.592857	2.692857
ZP2	0.125000	0.267857	0.434524	0.634524	0.884524	1.217857	1.717857
PRI	0.058824	0.135746	0.226656	0.369512	0.565913	0.902846	1.402846
PR2	0.117647	0.271493	0.453311	0.739026	1.139026	1.805692	2.805692
LIN	0.100000	0.200000	0.400000	0.700000	1.100000	1.500000	2.100000
LN2	0.100000	0.300000	0.600000	1.000000	1.500000	2.100000	2.800000
FIB	0.100000	0.200000	0.300000	0.500000	0.800000	1.300000	2.100000
FB2	0.200000	0.300000	0.500000	0.800000	1.300000	2.100000	3.400000
BAS	1.000000	2.000000	3.000000	4.000000	5.000000	6.000000	7.000000
DBL	0.100000	0.300000	0.700000	1.300000	2.100000	3.100000	4.300000
50T	0.240000	0.360000	0.540000	0.810000	1.210000	1.810000	2.710000
33T	0.270000	0.360000	0.480000	0.640000	0.850000	1.130000	1.510000

Table 2 – Sets of bonus values studied.

Because the relative differences between each pair of bonus values for ZPF, moving left to right in Table 2, increases slightly in magnitude, we considered as many different sequences that we could think of that would mimic this pattern. We attempted to produce these sets of bonus values, using relatively straightforward approaches that would still produce somewhat different sets of values – while also striving to maintain that the difference between the largest two bonus values would remain roughly in the range of 0.5 to 1.

ZP2 was generated by shifting the weights used in ZPF to the right, and then, inserting 1/8 as the bonus for uninvited teams. PRI and PR2 are similar to ZPF except that the denominators used in the fractional increment are the smallest seven prime numbers (17, 13, 11, 7, 3, 2) – instead of 7 down to 1; PR2 uses the same denominators as PRI but replaces the current numerator (1) with 2. The first bonus value was 'established' at 0.1 for LIN (other initial values might be more accurate), and then each subsequent bonus increment is increased by 0.1, starting with 0.1, i.e., add 0.1 to the first value (0.1), then add 0.2 to the second value (0.2), then 0.3 to the third value (0.4), etc. LN2 is a variation on LIN, adding 0.1 to zero to produce the first bonus value, and then adding 0.2, 0.3, and so on to the following bonus values. DBL starts with 0.1 and then simply doubles the increments used in LIN, so 0.2, then 0.4, 0.6, 0.8, 1.0 and finally 1.2 are added to the previous bonus values.

FIB and FB2 rely on the rule used to enumerate the Fibonacci sequence, i.e. each bonus value is the sum of the previous two; FIB begins with 0.1 and 0.2 while FB2 uses 0.2 and 0.3 as its first two bonus values. (FB2 sort of 'violates' a previously stated tenet since the relative increase from 0 – for uninvited teams – to 0.2 is larger than the 0.1 increase between the next pair of bonus values: 0.2 and 0.3. This also occurs with the first two bonus values in 50T and 33T described next.) The idea behind the 50T strategy was to use 50% of the previous bonus value as the increment, and 0.24 was chosen as the initial bonus value because it could be repetitively increased by 50% without needing 3 significant digits (until the last three bonus values), and the final increase was 0.9, which is close to 1 as in ZPF. Likewise with 33T – the increases were one third of the previous bonus value (though the final increase, 0.38, was much smaller). Lastly, BAS represents a baseline set of bonus values, where the integer values from 1 to 7 are used. These values essentially order teams by their NCAA tournament win total, breaking ties by the preexisting order that exists in the penultimate coaches' poll, i.e. the embodiment of the second insight mentioned previously.

As indicated in the Pen. Poll column in Table 3, the predictions of eight of these twelve sets of bonus values averaged close to 0.9 (per SCC), when applied to the training data set (1993-2007), and this is higher than the 2.57253 previously reported for the MC-Best model. The next five years (after 2007) were also predicted, on average, more accurately by ZPF (than the MC-Best model). However, the final poll in 2013 illustrated a potential shortcoming in this approach – when using the normalized, penultimate poll vote totals as the quantitative measure that would be added to each team's earned bonus value.

Abbrev.	Pen. Poll	TSR
ZPF	2.763298	2.767448
ZP2	2.692076	2.808283
LIN	2.742170	2.777374
LN2	2.785163	2.658235
FIB	2.568073	2.769227
FB2	2.735494	2.766658
BAS	1.801038	2.036637
DBL	2.746272	2.534972
50T	2.719305	2.782339
33T	2.539936	2.776811
PRI	2.431091	2.711560
PR2	2.739301	2.772363

Table 3 – Model results (using the training data set).

Given that the ninth seeded, Final Four participant Wichita State played eventual 2013 champion Louisville competitively throughout their contest, even leading by a significant margin in the second half of that game, most interested observers would've guessed that the Shockers would end up as the #4 team in the final poll that year, based upon the recognized voting patterns in other recent, final polls. MC-Best predicted the Shockers to be #8, but the ZPF model said they would be #13 primarily because Wichita State received zero votes in the penultimate poll. Since only 35 to 50 teams typically receive one or more votes in that penultimate poll, before the NCAA tournament begins, every one of the 300+ teams that are not present in said pool are essentially ‘lumped together’. This seemed unfair to strong teams (e.g. Wichita State) since they would then start out ‘equal’ to teams who might have only won a few games that year, because all of these teams received zero votes in the penultimate poll.

Thankfully, we were aware of a quantitative measure that assigns every team a value between 0 and 1, and this measure was designed to reflect for each team, how good the season was (that was just completed). The Tournament Selection Ratio (TSR) [3] was initially designed to evaluate how fairly teams were bracketed in the early years of the NCAA tournament, i.e. before teams were seeded (which began in 1979), when teams were assigned more by geography than by any attempt to evenly distribute quality across all four regions (whose victors would reach the Final Four). Like Coleman and Lynch’s Dance Card model [2], the TSR also reliably predicted which teams would garner the ‘non-automatic qualifier’ invitations to the NCAA tournament. TSR has correctly predicted roughly 90% of these invitations, from 1985 to the present, which is slightly below the Dance Card’s accuracy of 93%. The Dance Card model was trained using data from 1994-1999, and it has maintained this high level of accuracy from 2000 to the present.

The TSR metric is a deterministic formula that has many similarities to the formula used in the Bowl Championship Series (BCS). The BCS formula was used to select the teams who would compete for the NCAA‘s championship in football. The TSR utilizes two human polls as well: each normalized vote total contributes 25% to the final TSR value, and the other 50% comes from eight computer rating/ranking systems. Rating systems rely on the full margin of victory while ranking systems typically only use win/loss outcomes for its calculations. Four of each type of system are included in the TSR, and two of them are based on the power rating system [1] that was previously mentioned (and incorporated into) the MC-Best model. This power rating system is included both as a ranking and a rating system in the TSR, where the margin of victory is capped at one point in the former. (The final Sagarin ratings, as they appear in the *USA Today*, are also included as one of the eight systems.) The trimmed Borda count adjustment is administered, removing the highest and lowest rank. The normalized average rank, which is computed from the remaining six computer models, is multiplied by 50% and added to the poll’s contribution to this measure, to produce the TSR value for each team.

Table 4 illustrates who the TSR thought were the top eight teams in 2014, along with the three other teams who reached the Final Four that year. Each row lists the TSR rank (and value), the team’s record, its seed number, the normalized AP and coaches’ penultimate poll vote totals, along with the rank where that team was placed by each of the eight systems, with 351 representing the highest rating, according to that system. (Pw1, RPI, Rew and Mod are the ranking systems.)

```
 #     TSR  W  L S#    AP    COA Pow Pw1 RPI Rew Exp Mod  SD Sag Team Name
 1 0.99494 32  2  1 0.991  0.994 348 351 351 351 349 351 350 349 Florida
 2 0.97787 34  0  1 0.967  0.968 333 347 348 350 348 350 351 339 WichitaSt
 3 0.93227 29  5  4 0.869  0.896 351 335 334 344 351 340 348 350 Louisville
 4 0.92906 30  4  1 0.875  0.849 350 350 350 349 350 349 349 351 Arizona
 5 0.92729 28  6  1 0.880  0.870 336 342 345 345 343 343 345 347 Virginia
 6 0.87267 28  4  2 0.758  0.755 346 349 346 348 346 348 347 348 Villanova
 7 0.84972 26  8  3 0.690  0.761 347 338 344 338 347 339 332 345 Duke
 8 0.83666 25  8  2 0.715  0.704 338 340 342 339 334 338 322 341 Michigan
13 0.74606 26  7  2 0.549  0.484 339 345 347 343 335 347 338 343 Wisconsin
18 0.61157 26  8  7 0.310  0.269 328 323 331 332 327 330 316 328 Connecticut
23 0.51649 24 10  8 0.047  0.129 335 332 336 315 337 326 327 334 Kentucky
```

Table 4 – Full TSR ratings for its eight top teams, and other teams who reached the Final Four.

If the values in the two columns in Table 3 are compared, nine of the twelve possible bonus values approaches performed better when added to the full TSR value than when added to just the normalized vote total in the penultimate coaches' poll. Table 5 provides a complete breakdown for the top two performers in each category – along with MC-Best. (A year by year breakdown, of where each system ranked the top 35 teams, can be found at http://academics.smcvt.edu/jtrono/FPModels.html.) Table 6 does likewise for the years after the training set.

Abbrev.	SCC-15	SCC-25	SCC-35	Sum
ZP2 (TSR)	0.924510	0.953089	0.930684	2.808283
50T (TSR)	0.904523	0.948560	0.929256	2.782339
LN2 (Pen.)	0.905958	0.941282	0.937923	2.785163
ZPF (Pen.)	0.890953	0.934244	0.938101	2.763298
MCB (OLR)	0.850241	0.861102	0.861191	2.572534
RND (TSR)	0.937348	0.956744	0.936786	2.830878

Table 5 – Results when using the training data set: 1993-2007.

Abbrev.	SCC-15	SCC-25	SCC-35	Sum
ZP2 (TSR)	0.948979	0.958654	0.949890	2.757523
50T (TSR)	0.959440	0.960904	0.953456	2.873830
LN2 (Pen.)	0.930229	0.952280	0.952101	2.834610
ZPF (Pen.)	0.869459	0.932499	0.942463	2.735421
MCB (OLR)	0.857907	0.812381	0.855306	2.525594
RND (TSR)	0.955357	0.964121	0.955642	2.875120

Table 6 – Predictive results after training: 2008-2014.

It is interesting to note, when comparing the results contained in Tables 5 and 6, that even though RND (described below) was the best in all categories (excluding the SCC-35 value) when evaluated against the training data, it just barely surpassed the SCC sum for the 50T model when applied to the non-training data; 50T also had the highest SCC-15 value of all models in Table 6.

ZP2's performance seems to have slipped some, regarding how well it has performed on the more recent polls, whereas LN2 has joined 50T in more accurately predicting where teams would be placed these last seven years. (By the way, 50T also did the best job predicting where Wichita State would be ranked after the tournament in 2013 – at #6.)

Our previous article also outlined that the NIT tournament invites teams after the NCAA tournament field is selected, and so an NIT win is not as impressive as one in the NCAA tournament. We discovered that our MC-Best model performed most accurately when the value of an NIT win was one quarter of the value of an NCAA win. (The number of NIT wins earned was also incremented since the number of invited teams in the NIT is 32 – not 64+ – so, the champion must win 5, not 6, games to earn the NIT crown.) For each of the strategies (and each component added to the bonus values – TSR and Pen. poll) in Table 3, a different divisor was determined, using the training set, which maximized the sum of the three different SCC metrics. For ZP2 (TSR), 50T(TSR), LN2(Pen.) and ZPF(Pen.), these were 6, 6, 15, and 8, respectively.

We also applied the Monte Carlo technique to our models that utilize the two new insights, as described in this article, generating random bonus values that increased slightly (on average) as more tournament wins were earned by a team. Of the 10 million sets of bonus values generated, eight had an SCC sum >= 2.82 (where the highest computed value in Table 3, for ZP2, was roughly 2.81.) These eight sets of bonus values were all quite different from each other. For instance, the initial bonus value awarded to invited teams who lost their opening round NCAA contest ranged from 0.1195 to 0.1985, and the increase to the next bonus value ranged from 0.0612 to 0.1234. However, the difference between the second and third bonus values only varied from 0.2025 to 0.2235 for all eight of those sets, and only from 0.1345 and 0.1566 for the third and fourth bonus values. This range increased to be from 0.3195 to 0.4064 for the next two bonus values, with much larger variations between the last two pairs of bonus values. (The randomly generated set of bonus values that produced the highest SCC sum is: 0.198471, 0.289600, 0.492093, 0.641069, 0.985658, 1.640018, and 2.160332. The accuracy of these bonus values appear in Tables 5 and 6 under the name RND. Six was chosen as the NCAA win divisor in this Monte Carlo approach, and other divisors were not explored. All of the bonus value pair differences in RND, like FB2, 50T and 33T, did not progressively increase – as illustrated by the smaller difference between the first two, and the third and fourth bonus values.)

Final Thoughts

We are thankful for the inspirations that led to the confluence of these new ideas which fueled our further investigation into the non-linear models described in this article. Some critics may object to the inclusion of any subjective attributes in these models, i.e. the votes offered by the coaches in the penultimate poll and/or both polls in the TSR. Given all the bonus value generation strategies that we examined, it is somewhat surprising that ZPF(Pen.), with the highest SCC-35 value in the training set, and ZP2(TSR), with the highest SCC sum, and the highest SCC-15 and SCC-25 values in the same data set, appear in the top two in both variations (when excluding RND). Perhaps there is some subconscious reason behind this inherent adherence (to Zip's Law) as observed in the voting behavior of the coaches. It does appear that the coaches are fairly consistent in their voting as they continue to favor those teams that they previously believed to be high achieving teams, given the new information gained by the tournament outcomes. (We may retroactively add 'predictions' of how the coaches might have voted, in the years prior to 1993, to our web site which displays the results produced by our models. However, there are several caveats to consider when viewing such speculative results; these 'special accommodations' for final poll predictions, especially before 1975, were listed in our previous article, and have been omitted here.)

Perhaps randomly tweaking the RND bonus values will lead to even higher SCC sums, but there is more satisfaction for us to use a model for making predictions, that produces very accurate results, which relies solely on bonus values that follow a particular pattern, as in ZP2, 50T, LN2, etc., than those that incorporate randomly generated bonus values (as when applying Monte Carlo techniques). The average difference in rank observed for the ZP2, 50T and LN2 models has been between 2.1972 and 2.2519 for the top 35 teams, for the polls spanning 1993-2014. These prediction models appear to match the actual polls quite accurately, and even the other, less accurate models listed in Table 5 do fairly well, with an average difference for ZPF coming in at 2.3705 and even MC-Best was only off by 3.4071 rank positions per team on average.

Finally, there will always be some shortcomings of our predictive models that attempt to accurately approximate this ranking (as produced by the coaches for this final poll); subtle details, that aren't captured by the quantitative measurements included in our models, can influence the voting coaches. Therefore, the types of models included here will typically incorrectly predict where some teams will be ranked because of this inability to 'consider everything', e.g. final scores of tournament games, in the same manner that the coaches seem to do.

References

[1] Carroll, B., Palmer, P., Thorn, J. 1988. *The Hidden Game of Football.* Warner Books.

[2] Coleman, B. J., Lynch, A. K. 2001. Identifying the NCAA Tournament 'Dance Card'. *Interfaces* 31(3):76-86.

[3] Trono, J. A. 2013. Evaluating Regional Balance in the NCAA Men's Basketball Tournament using the Tournament Selection Ratio. Proceedings of the 4th International Conference on Mathematics in Sport.

[4] Trono, J. A. Yates, P. 2014. How Predictable is the Overall Voting Pattern in the NCAA Men's Basketball Post Tournament Poll? Chance, 27(2):4-12.