A comparison of preseason collegiate football polls with corresponding end-of-the-year rankings

Brent Hancock, Matt Hancock, Marc Schmidt

Dr. David Strong, Advisor

PEPPERDINE UNIVERSITY

Out-of-order rankings and inversion numbers

The number of "out-of-order rankings" is the sum of the total number of other ranks each rank is out of order with. For example, in the right most column below, the 5 *incorrectly* comes before the 4, 3, 2 and 1, the 4 incorrectly comes before the 3, 2 and 1, and so on, for 4 + 3 + 2 + 1 = 10 total out-of-order rankings. It turns out that this is always simply the inversion number of the permutation.

The inversion number of a permutation is the minimal number of interchanges of consecutive elements necessary to rearrange them in their natural order. The inversion number of a particular permutation is unique, although the steps that take you from that permutation to the natural ordering are not unique . [2]

At right are progressively more out-of-order permutations and their inversion numbers for permutations of the numbers 1 to 5. In ordering teams, a lower inversion is better, as it means the predicted order is closer to the actual outcome.

Inversion number										
0	1	2	3	4	5	6	7	8	9	10
1	1	1	3	3	3	3	3	3	5	5
2	3	3	1	4	4	4	4	5	3	4
3	2	4	4	1	1	5	5	4	4	3
4	4	2	2	2	5	1	2	2	2	2
5	5	5	5	5	2	2	1	1	1	1

Computing probability distributions:

There are *n*! permutations of *n* items (in this case, teams to order). The maximum inversion number for a permutation of *n* items (teams) is *n*(*n*–1)/2. The minimal inversion number, of course, is 0. For example, a permutation of 5 items has a maximum inversion number of 10, as seen in the right most column of the table above. A permutation with this maximal inversion number would correspond to a list of teams whose predicted order was exactly opposite of the actual order. This happened three times in our study. In these cases, the preseason poll gave the worst possible prediction for the final outcome: the teams were in the exact opposite order.

For each permutation of size *n*, an explicit formula for computing the number of occurrences $I_n(k)$ of each inversion number *k*, *k* = 1 to *n*(*n*–1)/2, is at right. [3]

We use this to compute the probability (relative frequency) distribution for *n* = 5, 10, 15, 20 and 25, the corresponding cumulative probability distributions, and the corresponding percentiles for what fraction of the *n*! permutations each inversion number is better than. A percentile of 0 corresponds to the maximal inversion number, and means the preseason ordering was better than none of the *n*! possible orderings of the teams. This happened for the two preseason polls mentioned above. A percentile near 100% means the preseason prediction was better than nearly all of the other possible team orderings.

$$I_n(k)=\binom{n+k+1}{k}+\sum_{j=1}^{\infty}\binom{n+k-u_j-j-1}{k-u_j-j}+\sum_{j=1}^{\infty}\binom{n+k-u_j-1}{k-u_j}$$

where $u_j=\frac{j(3j-1)}{2}$ and $\binom{a}{b}\equiv 0$ for $b<0$.

Note that $\sum_{k=0}^{n(n-1)/2} I_n(k)=n!$

Summary

We looked at how well preseason polls predict the *relative* end-of-the-year rankings in NCAA football. That is, rather than address the issue of predicting the actual top 25 teams and their order, we looked at how each poll's preseason top 25 (and top 5, 10, 15, and 20) teams ended the season ranked *relative to each other*.

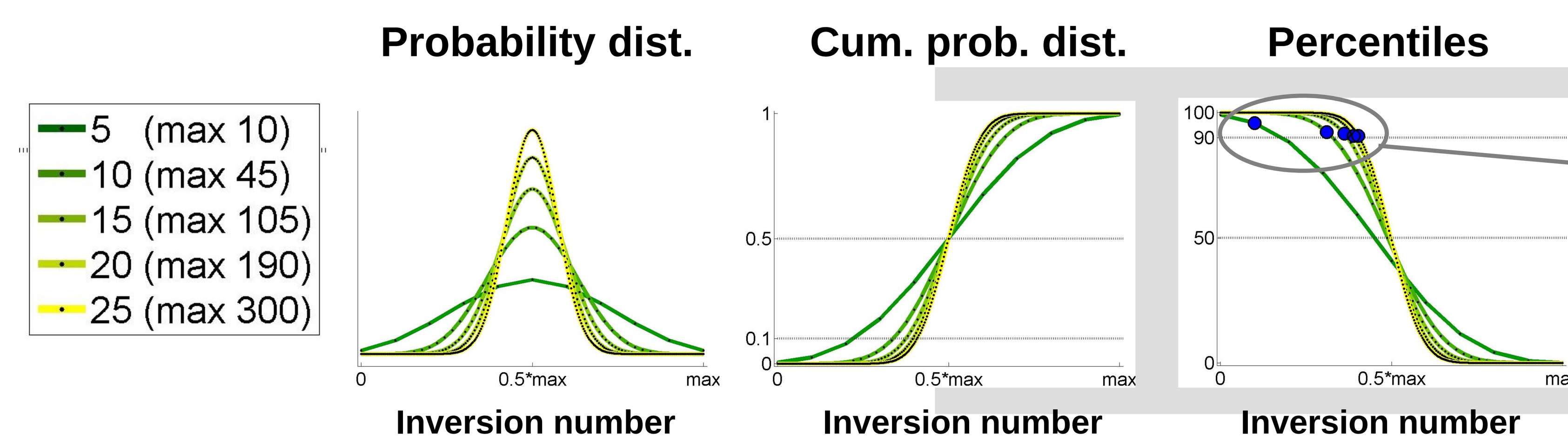
We measured how well the ordering of a poll's teams predicts the final poll simply by looking at how many pairwise comparisons of teams are out of order (the more out of order, the worse the poll's prediction). It turns out this measure is simply the *inversion number* of that poll relative to the final poll. Thus we simply found the inversion number for each poll's ordering of teams relative to the final end-of-the-season rankings, for which we simply used the average of the computer polls used in computing the BCS rankings for each year. We considered polls compiled at [1] for which we had data from all six years of 2001 – 2006.

For each poll, the order of its top 25 (and 5, 10, 15 and 20) teams is compared to the order of those same teams in the final poll. The results below, for Street and Smith's 2004 preseason poll, had the highest percentile of any poll for any year. This poll had an inversion number of 43 (out of 300 max), and a percentile of 99.99999839.

S & S preseason		Final computer		
1	USC	1	USC	1
2	Georgia	2	Oklahoma	3
3	Oklahoma	3	Georgia	2
4	LSU	4	LSU	4
5	Florida State	5	Iowa	11
6	Texas	6	Texas	6
7	Miami	7	Miami	7
8	Michigan	8	Michigan	8
9	Florida	9	Florida State	5
10	Tennessee	11	Ohio State	12
11	Iowa	12	Utah	21
12	Ohio State	13.5	Tennessee	10
13	Auburn	13.5	Auburn	13
14	Virginia	15	California	17
15	West Virginia	16	Florida	9
16	Clemson	17	Maryland	18
17	California	18	Purdue	22
18	Maryland	21	Clemson	16
19	Missouri	25	Virginia	14
20	Wisconsin	26	Wisconsin	20
21	Utah	29	Minnesota	24
22	Purdue	31	West Virginia	15
23	Memphis	38	Missouri	19
24	Minnesota	40	Memphis	23
25	Toledo	42	Toledo	25

Original order

Permuted order



Percentiles of inversion numbers for permuations of size *n*.

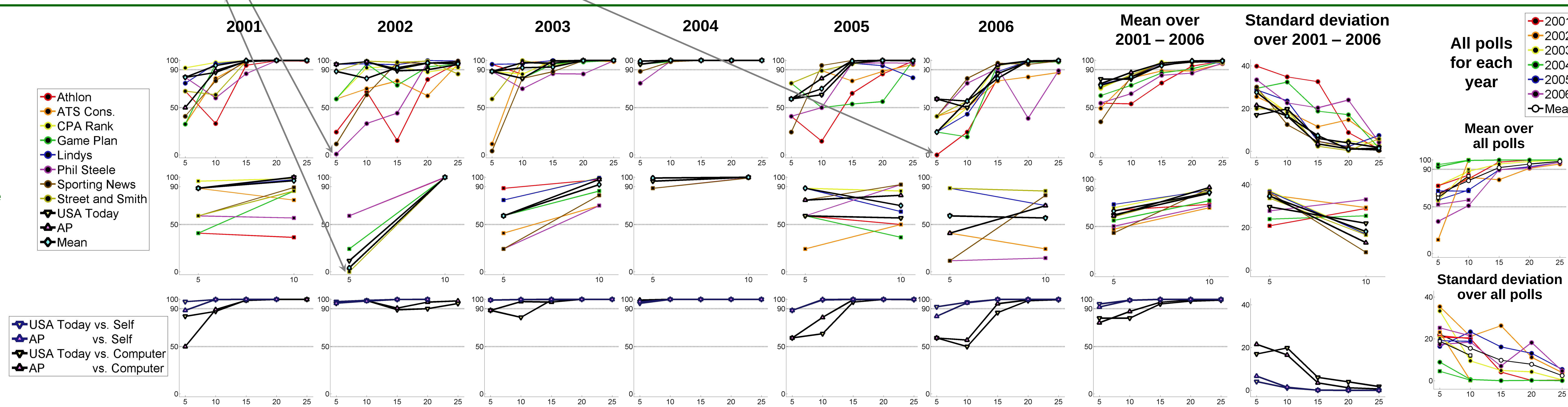
		Actual		Relative to max	
n	max	90	99	90	99
5	10	1	0	0.100	0.000
10	45	14	8	0.311	0.178
15	105	38	28	0.362	0.267
20	190	74	59	0.389	0.311
25	300	121	100	0.403	0.333

Three comparisons for the 2001 – 2006 seasons

For each poll, compare its own top 5, 10, 15, 20 and 25 to final computer polls.

For each poll, compare its ordering of the top 5 and the top 10 teams present in *all* preseason polls to the final computer polls.

For AP (sports writers) and USA Today (coaches) polls, compare its preseason poll to its *own* final poll.



Observations and conclusions:

In general, the polls do a good job (better than we expected) at predicting the final outcomes of relative rankings of teams. Most polls were above the 90th percentile in most cases.

Accurately ordering just a few teams proved more difficult than ordering several teams. This is clearly seen in the percentile plots for each year, as well as the *Mean* and *Standard deviation* plots: percentiles increase and standard deviations decrease as the number of teams *n* being ordered increases. This increase in percentiles as the number of teams *n* being ranked increases is explained by the plots of *Probability distribution*, *Cumulative probability distribution*, and *Percentiles*: for increasing *n*, an increasing proportion of inversion numbers is concentrated in the center. In short, while it was a bit difficult to accurately predict the relative rankings of just a few teams, it is even more difficult to not do well in ordering a large number of teams. One extreme example of this is the 2004 Street and Smith's preseason top 25 (see far upper right), which has a seemingly unexceptional inversion number of 43 (maximum possible 300), with a remarkable corresponding percentile of 99.99999839.

The polls with the best results are those with larger and more diverse group of voters. In particular, the USA Today and AP polls generally seemed to perform better than any of the others. Moreover, while they did well relative to final *computer* rankings, they did even better when comparing their preseason polls with their own final polls. This suggests one (or both) of two things: where a team ends up ranked in a poll is to some degree dependent on where it started in that poll; or whatever bias those who vote in a particular poll have in their preseason voting is still evident in their final voting.

Overall Athlon, Phil Steele, and Game Plan seemed to be the worst performers, as seen in both the mean and standard deviation over 2001 – 2006, as well as most of the individual years. Two polls (ATS Cons. and Sporting News) who were at or below the 50th percentile in ranking their own top 5 over all of 2001 – 2006; a purely *random* ordering of their top 5 would have been better.

Some years don't go at all as predicted. As seen in both the *Mean over all polls* and the *Standard deviation over all polls,* as well as the individual years, 2002, 2005 and 2006 were especially hard to predict. In particular, many of the 2006 preseason polls for top 5 and top 10 were below the 50th percentile. On the other hand, 2004 went quite as predicted, as seen in the various plots.

Further questions to consider

We had planned to do most of these (we had actually started on some of them!) as part of this project, but ran out of time.

For each ranked team, how does variance amongst the various polls correlate with the agreement between the average rank for that team and its final poll position?

For each preseason poll, how does the average or cummulative error/difference between that poll and the average of all polls correlate with how accurately that poll predicts the final poll?

What weighting (e.g. by finding least squares solutions) of the polls best predicts the final poll?

How much does when each preseason poll is released in the summer affect that poll's reliability?

References

[1] College Football Preseason Magazines, compiled at http://preseason.stassen.com.

[2] T. Muir, *A Treatise on the Theory of Determinants*, New York: Dover, 1960.

[3] D. E. Knuth, *The Art of Computer Programming*, Addison-Wesley, Reading, MA, Vol. 3, p. 15.