
Copyright © 1988 Operational Research Society.

This is the author’s version of the work. It is posted here with the permission of the publisher for your personal use. No further distribution is permitted. If your library has a subscription to this journal, you may also be able to access the published version via the library catalogue.
Providing football predictions for a major daily newspaper proved to be a demanding but instructive experience for Stephen Clarke. Computer output was subject to misinterpretation, while the quality of forecasts was often judged harshly on the basis of odd way-off predictions.

During the latter part of 1981, a computer prediction model was developed for VFL (Australia Rules) football. The program used as data only the draw for the season and previous results. The computer had a rating for each team, and the result of a match was predicted using the difference in the teams' ratings plus any home ground advantage. The ratings were updated by a simple exponential smoothing technique. Thus, if the prediction was in error by 25 points, one team's rating was increased by some fraction of 25, and the other team's rating decreased by the same amount. Stephani (1987) gives technical details of such methods and their application to American football. Although the program was originally written for interest only, I subsequently wrote to a major daily newspaper in Melbourne, suggesting they include the computer tip in their football coverage by setting up a competition with their flamboyant writer, Lou Richards. A consulting arrangement was agreed, and for six years, early in each week of the football season, the paper was supplied with a computer printout giving various predictions for the round of matches to be played on the following Saturday and the remainder of the year. The printout listed predicted winning teams, with predicted point margins and probabilities of winning, and a predicted final ladder. In 1983, simulation output estimating the chance of each team making the finals was also included. On the day of the matches the paper published extracts from the printout. Occasionally, extra material submitted separately would be published.

The computer was quickly labelled 'Tinhead' by the chief writer, and the name caught on. Unfortunately, my friends and colleagues confused the computer with its programmer, and made frequent use of the nickname. I generally had to put up with much (good natured?) ribbing on the computer's (my?) performance when it had a bad day. I often found myself in discussion with friends, neighbours, etc., on the results of my research - answering such questions as "How does the computer do it?", "How can it possibly pick Hawthorn this week?", "Does it take player injuries into account?", and so on.

Lies, damned lies...

In publishing for a lay audience, one must compromise to their level of expertise. Sometimes this required minor adjustments, such as referring to probabilities as percentage chances. Thus, the printout might show Hawthorn to win with an 80% chance and a predicted point margin of 30 points. Hawthorn to win by 30 points meant that Hawthorn was just as likely to win by more than 30 points as to win by less than 30 points or lose. In a sense, the probability is never wrong (if Hawthorn lose, well it did say they had a 20% chance of losing!), whereas the point margin is virtually never right. Unfortunately, in six years the newspaper
always managed to publish the margins but never the probabilities. In tuning the program at the start of the season, the value of the parameters could be chosen to optimize the number of wins, or the accuracy of the point margins, or the final ladder predictions. This, however, had to be done before it was known which sections were to be published. To make matters worse, most readers interpret the margin as being a prediction of the margin that will actually occur. This resulted in harsh judgements on the computer’s performance. For example, in predicting Melbourne to win by 62 points when most commentators were suggesting a closer game, the computer had performed well when the margin was 129 points. However, to readers this was just a whopping error of 67 points. It was simply not possible to resort to confidence intervals. A prediction that Melbourne will probably either lose by less than 20 points or win by less than 120 points would hardly be exciting reading. This misinterpretation of margins also necessitated altering the printout. Hawthorn to win by 0 points was a perfectly reasonable margin statement, but the average reader would think the computer had blown a fuse; a margin of 0 points would be interpreted as a draw. In these cases the printout had to be altered to show a win by 1 point.

Sometimes information had to be suppressed from the printout to prevent the publication of seemingly contradictory material. To predict the final ladder, the computer calculated for each team a probability of winning each match (classing the match as a win if the probability exceeded 0.5), and accumulated the probabilities for each team to get the expected number of wins for the rest of the season. The following printout might occur:

<table>
<thead>
<tr>
<th>Team</th>
<th>Number of wins so far</th>
<th>Results in remaining 4 matches</th>
<th>Predicted number of wins for the year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Melbourne</td>
<td>12</td>
<td>LLLL</td>
<td>13</td>
</tr>
</tbody>
</table>

Thus, the printout showed Melbourne was expected to lose each of its remaining four matches, while also showing they were expected to win one match for the rest of the season (in much the same way as for each of 10,000 houses, the most likely result is that the house will not burn down, but for the 10,000 houses as a whole, the most likely result may be that two will burn down). While this is perfectly clear to a statistician, the general public, on seeing the above table, would think the computer had made an additional mistake. This necessitated excluding the individual match predictions for the remainder of the year from the printout.

**Rapid responses**

The popular press has a much shorter lead-time than academic publishing. The daily press, in particular, lives from hand to mouth. At the beginning of each season, there was usually only a few days’ notice that the paper wanted to use the computer tips. This left little time for detailed research work. Periodically, letters or comments on the computer’s performance appeared, from the sports writers or the public, and any comments or rebuttals had to be ready for the following day’s paper. For the weekly tips, while it was refreshing sometimes to submit material at 9.00 p.m. Friday and see it in print at 7.00 a.m. Saturday, such haste left no time for checking or for reflection. Inevitably, misprints sometimes occurred. Subsequent retractions were out of the question. Last week’s sports news is no news at all and readers are not interested in reading on Monday that the Saturday prediction was meant to be Hawthorn to win, not lose. We just had to live with any mistakes. In fact, in six years and nearly 800 tips,
only twice was the incorrect predicted winner published. One of these was in the computer's favour, the other against.

Surprisingly, even with such a short lead-time, information was sometimes out of date. With the introduction of Friday night matches, the paper published the predictions for these matches on the Friday. However, the final ladder prediction was held over until Saturday and could therefore be based on out-of-date information. Thus, a team having a particularly bad loss on a given Friday, with little hope now of making the finals, might on the following Saturday still be shown high up in the final ladder prediction.

The immediacy of the forecast also made life difficult. In forecasting energy demands in the year 2000, a researcher is judged on his methods, not (at least in the short term) on the accuracy of his predictions. My situation was akin to a weather forecaster, with many people even reading the forecast after the event. In this case the computer, like the much-maligned weather forecaster, was often judged harshly. People remembered the odd way-off prediction, and forgot the majority of correct or near-miss ones. Over a season, in predicting winners of 138 matches with an average 70% success rate (about the success rate achieved by the computer and most commentators), the standard deviation of the number of correct winners is about 5. Thus for a commentator to 'beat' the computer by 2 over a year was insignificant. Of course, to the average football follower, this was damning proof of the computer's inferiority. I was then placed in a no-win situation if I preserved academic integrity and did not claim superiority with a similar margin in the computer's favour. In fact, in six years, the major football writer of the paper never once beat the computer - although he did manage to draw one year.

"NO FOOTBALL PREDICTIONS UNTIL YOUVE FINISHED YOUR LINEAR PROGRAMMING!"

Was it worth it?

I suspect the paper judged the computer not by its accuracy but by the interest it created. As it was possible to select parameters to give conservative or controversial forecasts, I had the problem common to most projects - Do I give the clients the answer I think they want? Any tendency by the paper to print the more controversial forecasts (and newspapers thrive on controversy) increased interest in the predictions but gave a biased view of the computer's performance. Was it worth it? I confess there was a pecuniary element to my endeavours, but there were other motives too.

It is my belief that much operational research should be published not only in OR and mathematical journals, but also in journals relevant to the application areas. If results are applied, and we are interested in implementation, then the results should be directed at the users of the results, as well as those who merely happen to use the same techniques for other application areas. Such publishing also has the side-effect of educating the general public in the possible uses of OR. With my particular interest, applications in sport, this means publishing in the popular press - newspapers or sports magazines.
All operational researchers, having had a paper accepted by a traditional journal, should seriously consider submitting an article based on their work to a popular form of publishing. Obviously, not all research is suitable, but I believe a greater effort should be made to popularize research results. It increases awareness in the general public of the applications of OR, mathematics and statistics, as well as publicizing the writer's institution. In an educational context, where there is competition for funds and students, the Importance of this sort of publicity should not be underestimated. The reputation gained through such publishing may be just as important as more traditional forms of publishing which contribute to one's standing in the academic world. On a personal level, your work does become known to a wider group of people. In my case, I felt part of the Australian football scene for several Interesting years. I hope this article will encourage other people to attempt alternative forms of publishing.

Reference

About the Author
STEPHEN CLARKE is a senior lecturer in the Mathematics Department at the Swinburne Institute of Technology, Hawthorne, Australia. He is currently responsible for the OR component of the Applied Science degree at Swinburne. His research interests lie in the mathematical applications of OR in sport, with papers on squash, badminton, football and cricket to his credit.