

Comments on *Flexible marked spatio-temporal point processes with applications to event sequences from association football* by Narayanan et al.

Paul A. Smith, University of Southampton

It was a pleasure (at a time when the World Cup is being televised) to hear some analytically based evidence on football rather than the usual pundit-based discussions. I enjoyed reading the paper, and both the paper and the proposer and seconder prompted some thoughts.

Splitting the field into three parts is clearly a pragmatic approach, but this is done deterministically and taken as a fixed part of the analysis. On the other hand, we can imagine that different areas of the field are related to different events (or chains of events). Is it possible to split the field into areas based on the observed data, and then to make inferences about the importance of different parts of the field?

I was intrigued that events in football are more regular than a Poisson process, and it seems to me that this is most likely because the shortest inter-event times are missing. Could you therefore model the inter-event times more effectively with a truncated Poisson process?

The paper indicates that predicting future events is challenging, but shows some encouraging results on prediction of shots at goal. Can you extend the idea of simulating passages of play to simulate whole games? And if you do, is the scoreline for a simulated game plausible (eg 2-0), or implausible (eg 14-3)? This would give some evidence for the usefulness of the fitted parameters in defining the important features of the game.

Finally, I read somewhere that statistical modelling for football is difficult because a lot of activity important to the outcome takes place off the ball. But the method in the paper focuses exclusively on ball-related events. Could you incorporate other information about the positions or qualities of players or the team in order to improve your model as a game descriptor?