Time-varying Associations in Multivariate Mixed Models: Applications in Sports Analytics

Supervisor: Dr Lisa McFetridge

Soccer is one of the most popular sports in the world, with over 3.5 billion fans across the globe. In 2021, it was estimated to be worth $3.02 billion globally and predicted to reach a value of $3.87 billion by 2027. With huge financial rewards for success in this sport, in recent years, all major football clubs have begun to substantially invest in the analysis of big data and the utilisation of statistical and machine learning methods applied to sports analytics.

Such investments have been aided with the development of wearable technology that allows real-time monitoring of health and performance metrics, such as players' heart rates, movements and location both during matches and in training scenarios. The dynamic changes in players' biomarkers, such as heart rate variability, act as pseudo variables to measure their adoption to training programmes, in-match performance and uncover the onset of fatigue. If utilised correctly, such information can uncover trends and patterns over time to reduce injury risks, enhance recovery and ultimately aid players' health and performance (Viegas 2024).

Wearable technology typically collects multiple observations a second and accumulates to vast amounts of data over the course of a season. To model such intensive longitudinal data, multivariate mixed effects models may be utilised (Hickey 2016). These approaches allow the modelling of multiple biomarkers or metrics over time, capturing different aspects that impact players' internal loads simultaneously and thus providing key insights into not only how each biomarker uniquely impacts players' health but also how different biomarkers relate to each other.

Whilst such insights are valuable, multivariate mixed models don't currently capitalise fully on the time-varying nature of the data. Current methods assume that the relationship between biomarkers doesn't change over time. This assumption of time-invariant correlations may be restrictive and likely unrealistic when analysing information that naturally evolves over time. To better capture the time-varying nature of these relationships, this project will develop new multivariate mixed model methodology that incorporates time-varying correlations.

This project will feed into a collection of work in sports analytics lead by the Predictive Sports Analytics (PSA) group. To maximise the potential for impact of this work, the PhD student would work closely with other PSA researchers alongside local elite football clubs, including those playing in the top tier of their national leagues, to embed cutting-edge data-driven decision making into their processes and implement the desired real-world impacts.

For further details about the project, please contact the primary supervisor l.mcfetridge@qub.ac.uk.

Viegas, J.M., Dores, H., Freitas, A., Cavigli, L. & D'Ascenzi, F., Developments in sports cardiology: The way to a brighter future, *Revista Portuguesa de Cardiologia*, 2024; **43**(2), 87-89.

Hickey, G. L., Philipson, P., Jorgensen, A., & Kolamunnage-Dona, R., Joint modelling of time-to-event and multivariate longitudinal outcomes: recent developments and issues, *BMC Medical Research Methodology*, 2016; **16**(1), 1-15.