COMPETITIVE SWIMMING REIMAGINED: using data to help get the win

KEY CONTACT Dr Paul Wu
p.wu@qut.edu.au

Imagine this: It's the 2024 Olympic games, and the Australian swimming team is optimised to win—not just through dedicated training and exceptional skill, but algorithmic prediction models to refine the team's makeup into a composition with the best chance of getting the gold on the big day.

With sports becoming ever more competitive, a misstep could be the difference between gold and going home empty-handed. Data analytics and predictive models can assist coaches and athletes to better understand what is needed for a win, and help better prepare for elite competition.

Dr Paul Wu at QUT's Centre for Data Science is working alongside industry partners from Swimming Australia, the Australian Institute of Sport and the Queensland Academy of Sport to develop predictive models for relay team composition, qualifying swimming times and pacing.

New time trials

"Predicting winning times is a lot like goal setting—it's about understanding how fast you need to swim to achieve first place, or third place, or qualify for the Olympics or world championships," Wu explained.

A more traditional model might see coaches looking through recent elite swimming times as a guide and setting a goal to beat the previous year's record.

"Our model is more sophisticated, and is 19% more accurate compared to just using the previous winning time as the standard to beat," Wu said.

The model can't tell a swimmer whether they'll win or not based on their swimming times, but instead works in the realm of probability.

"We build up a probabilistic model based on trends over time. This means we can predict not just winning times, but also estimate the probability of events such as new world records, or the probability a certain time has of winning," Wu said.

The accurate model has a mean error rate of 0.5% - this means that if the winning time for a race is 100 seconds, the model's mean prediction will be within 0.5 seconds of the actual winning time on average.

"Trends over time have shown a steep decrease in winning times over the middle part of the twentieth century, before levelling out in recent decades as we inch toward the peak of human ability," Wu said.

"Once you can characterise the patterns of winning times, and especially the uncertainty and variation in it, you can get useful insights that allow for risk-informed ways to train and prepare athletes."

Fine-tuning the relay

Choosing who makes the cut and how they are run in a relay race is a complex decision for coaches and national selectors —one with a lot of variables and a lot of potential outcomes.

"If you had the four top-ranked swimmers in the world on one team, you'd obviously put them in your relay team," Wu said.

"But that's rarely the case—so there are a lot of different variables and statistics to consider when looking at relay team make-up.

"Our machine learning models use real data from elite competitions to explore trends in relay team performance, which can then help predict the probability of different race outcomes depending on swimmer calibre and racing order."

Imagine you have iconic Australian swimmers Shane Gould and Susie O'Neill on a relay team. If Gould swims first and O'Neill swims third, your probability of winning a silver medal may be 0.175, and your probability of winning a bronze medal might be 0.345 – not bad if you're in the business of winning bronze. If Gould swims second and O'Neill swims first, your probability of winning a silver medal may go up to 0.299, but your probability of winning no medal at all may jump to 0.320.

"It's about risk verses reward," Wu said.

"Is it worth risking a higher probability of not medalling in order to have a higher probability of a better medal? Or should you play it safe and aim for the bronze?

"Our model can help coaches and decision makers virtually explore team selection and ordering strategies, and they can then trial these strategies in the water to see how their teams perform."

Fresh ideas from young minds

Wu and his team worked with undergraduate mathematics students from QUT who were eager to explore research opportunities as part of ARC Centre of Excellence for Mathematical and Statistical Frontiers internships and QUT's Vacation Research Experience Scheme.

"The undergraduate students are always eager and learn quickly, and it's a great opportunity for them to see the impact of their work as it applies to end users," Wu said.

"When you work with industry, just being good at maths isn't good enough—you have to be able to communicate, to be organised, and to create a human connection.

"That's the most valuable skill that these students can take out of a project like this: how to deal with industry partners, and how to put the things they learn into practice in a real environment."