Swimming in Data

Katherine Douglass, Augustus Lamb, Jerry Lu, Ken Ono, and William Tenpas

Ladies and Gentlemen, *this is the Olympic Final*," thunders from the loudspeakers, yet it is barely audible over the deafening roar of a crowd ten thousand strong. One long blast of the referee's whistle summons eight of the world's fastest swimmers to the starting blocks. The air is thick with anticipation as the noise of the crowd gradually fades and tunnel vision sets in. To compete in the Olympics is to have arrived at the pinnacle of one's sport. It is the realization of a dream shared by the youngest summer league swimmer and the most outstanding professional. It is the realization of years of sacrifice—grueling workouts, lost sleep, missing out on adolescent rites of passage. And now, in front of millions of fans, the chance to bring home a medal, to stand on the podium draped in the national flag, to feel that all of the hard work and sacrifice was worth it. The chance to make history is at hand.

"*Take your mark*."

Olympic Dreams

This summer in Nanterre, a western suburb of Paris, sports fans worldwide will have the opportunity to see this thrilling scene repeated dozens of times in the Paris La Défense Arena. For the competitors—the Olympians—the chance to compete in these games will be a dream come true. For the American Olympians, the journey probably began when they were children in summer swim leagues, for in the warm summer months, when kids are out of school, swimming is a popular pastime supported by community swim clubs across the country.

Much like the music box ice cream truck or an afternoon spent running through the sprinkler, a summer league meet, with its handcrafted prize ribbons and competitive but friendly energy, is for many a staple of summer. It is estimated that more than three million children in America participate in summer swim leagues each year [11].

Turning Olympic dreams into reality is a monumental challenge, one that requires years of total commitment. Predawn practices are the norm, and as athletes rise through the ranks, the workouts become more demanding. For those that reach the state and national levels, there are multiple workouts each day. In addition to school, the motto "eat, sleep, swim" controls their lives. Practices bookend classes, and weekends that might have been spent socializing are instead devoted to competition and sleep.

How difficult is the challenge? In the United States, becoming a competitor in collegiate swimming is a major accomplishment. According to the National Collegiate Athletic Association (NCAA) [8], only about seven percent of high-school swimmers are recruited to a Division 1 college team. An invitation to the United States Olympic trials, arguably the most competitive swim meet in the world, is an honor that only the top 100 to 120 athletes can boast of in each event. Furthermore, outside of fellow swimming powerhouses Australia and China, the depth of American competition is unmatched. For example, in the 2023 Speedo world rankings [12], four of the top five women in the 100 meter backstroke event are from the United States. And for a further dose of reality, only the two best men and women at the Olympic trials will make the Olympic team. To add to these tall odds, races are often decided by mere hundredths of a second. Indeed, the difference between a trip to the Olympics and retirement has been as little as one hundredth of a second. Can the stakes be any higher?

How should coaches prepare Olympic hopefuls? Should they instruct their athletes to swim just like Katie Ledecky and Michael Phelps, with dreams of replicating their success? Should they train swimmers to imitate those icons, on the theory that their technique represents absolute perfection? Definitely not! Athletes come in different shapes and sizes. Athletes have different strengths and weaknesses. The stuff of Olympians is nothing like the "do it yourself" movement of home repair, where a simple internet search reveals the one correct answer. Coaches aren't hanging drywall; they are forging potential Olympians.

This is where the introduction of mathematics, physics, and technology has revolutionized swimming in recent years. Waves of new devices and methods have raised the level of world-class swimming, and it is now possible to "precision train" Olympic hopefuls with a little help from math and physics. The results have been eye-popping.

The Evolution of Sport

To place recent innovations in swim training in proper perspective, it is worthwhile to reflect on the evolution of the Olympics and Olympic sports. The 2024 Paris Olympics marks the one hundredth anniversary of the 1924 Paris

https://doi.org/10.1007/s00283-024-10339-0

Olympics. Although some venues will be reused, much has changed. Apart from the iconic Olympic flame and Olympic rings, which symbolize the union of the five inhabited continents,[1] the 2024 Olympics will bear little resemblance to their Parisian predecessor. Forty-four nations were represented in 1924, but this summer, that number will be eclipsed when athletes from over two hundred nations will arrive, illustrating remarkable growth in spite of significant world events that have disrupted the Olympic movement. Indeed, the Olympics have survived the tumult of two world wars and the Cold War.

Much like world politics, sports have also changed significantly. The first standard-length Olympic marathon[2] was held in 1924. The men's race was won by the Finnish runner Albin Stenroos in 2 hours, 41 minutes. The mighty Kenyan Eliud Kipchoge won the 2021 Tokyo Olympic marathon in a winning time of 2 hours, 8 minutes, representing a 20% time improvement. Stenroos's gold-medal-winning time wouldn't even earn him a spot on the starting line of the 2024 United States' women's Olympic trials, for his performance time misses the qualifying standard by four minutes.

Also consider the team pursuit event in track cycling, in which two teams of four athletes start at opposite sides of an oval track and chase each other in a desperate attempt to overtake the other team. The riders on each team take turns in the lead, allowing the unit to maintain a higher speed due to the protection from air resistance, known as sharing the draft. The pursuit is usually futile, for a team is rarely caught. Instead, the winning team is the first to have three of its riders cross the four-kilometer mark. A quartet of Italian men won in 1924 with a time of 5 minutes, 15 seconds. Illustrating their country's prolonged prowess in cycling, the men of Team Italy won gold again in 2021 with a time of 3 minutes, 42 seconds, a 30% improvement. What is remarkable is that in the individual pursuit, in which two cyclists chase each other over four kilometers without the benefit of teammates to share the draft, British cyclist Bradley Wiggins won gold in 2008 in 4 minutes, 15 seconds, beating the 1924 Italian team score by one minute! Furthermore, his English compatriot Rebecca Romero won the women's individual pursuit with an average speed that is four miles per hour faster than that of the 1924 Italian men!

What explains such a vast improvement in these performances? There are some obvious answers. Partly, it is technology. Bicycles of the past were primitive steel contraptions compared to today's sleek and lightweight machines fashioned from carbon fiber and fitted with aerodynamic disc wheels. Innovations like wind tunnels, "teardrop" helmets, and wind-defying skinsuits play further significant roles in reducing air resistance. Such advances in technology are common across events that require sporting equipment.

Advances in nutrition and improved diet have also greatly enhanced athletic performance. Indeed, today's marathoners understand the importance of pre-raceday meals that include high carbs combined with protein bars and energy gels for in-competition nutrition. For contrast, it is amusing to compare the modern regimen with that of the American runner Thomas Hicks, the 1904 Olympic marathon champion. He was refused water by his trainers and instead was fed strychnine and egg whites twice over the course of the race,[3] the second time with a dose of brandy [1].

Are there further reasons for leaps in athletic performance? For human-driven sports events, a better understanding of biomechanics has lifted sports to new heights. Indeed, there have been complete paradigm changes. For example, consider the high jump, in which competitors jump over a bar that is continually raised until only one competitor can clear the height. For decades, the world's best high jumpers leapt face forward and attempted to clear the bar with a "straddle technique." That changed essentially overnight at the 1968 Mexico City Olympics, when American Dick Fosbury won gold with a radical new technique, the "Fosbury flop." Instead of an awkward straddle, the high-jumper runs toward the bar along an arc and then gracefully leans into the turn with the goal of launching over the bar by arching backward over it. Scientists quickly understood the superiority of this strategy, the antithesis of the Caribbean limbo party game, where the aim is to pass under a bar by bending over backward. In the final steps before the jump, an athlete's center of gravity is lowered enough to offer significantly more time to generate liftoff.

As these examples show, Olympic performances have improved dramatically over the last hundred years, driven to a great extent by the emergence of sports science. Our aim here is to describe a new aspect of sports science, the role of mathematics and physics in the precision training of 2024 Olympic swimming hopefuls.

One Hundred Years of Swimming

To properly appreciate present-day training methods, we first comment on the evolution of the sport of swimming. How much has it changed? You might guess that swimming in Paris 2024, like marathon running, will look much as it did in 1924. After all, a swimmer needs only a swimsuit and a pool, just as a marathoner needs only a running suit and a pair of shoes. And of course, the swimmers will be churning through the same inorganic H_2O, while the marathoners will scamper over the same cobbled and paved streets. There do not seem to be many variables to tinker with.

If that's what you supposed, you would be wrong! Much about swimming has changed over the past hundred years, and the sport today would be unrecognizable to the 1924 Olympians. For starters, the suits are far more hydrodynamically efficient, a nod to the need to combat drag. Compare the bathing costumes of the 1920s with the form-fitting

---

[1] Why only five? The two American continents count as one in this iconic logo.

[2] The official distance for the marathon is 26 miles and 385 yards, or 42.195 kilometers.

[3] Strychnine is a toxic chemical that is often used in pesticides. At the time, it was considered a performance-enhancing drug.

Figure 1. Johnny Weismuller. (Credit: Library of Congress.)

Lycra suits of the 1980s, the now illegal "supersuits" of the 2000s, and the present-day carbon-fiber "tech suits." A cursory glance reveals a sport that has gotten much faster, with each garment change a reflection of advances made in materials science. Even more surprising is the fact that goggles were prohibited before the 1976 Olympics. Goggles offer protection, and importantly, they allow athletes to see in the water. Don't you need to see where you are going in order to swim fast?

How do the winning times from the 1924 games compare to those in the recent Tokyo 2020 games?[4] In 1924, the Olympic champion in the men's hundred meter freestyle was the American Johnny Weissmuller (Figure 1), who is perhaps better known for portraying Tarzan on the silver screen in the 1930s and 1940s. He won with a time of 59.0 seconds. We regret to say that the mighty vine-swinging Tarzan would find no success today as a swimmer. Indeed, the American Caeleb Dressel (Figure 2) won the event in 2021 with a time of 47.0 seconds. And the Australian Emma McKeon, the women's gold medalist, would have beaten the Hollywood icon by over seven seconds!

Unlike sports that rely critically on equipment, like cycling, some of the most significant advances in swimming have come from a better understanding of biomechanics. This means stroke technique. How should one execute the backstroke, breaststroke, butterfly, and freestyle? Each stroke involves many variables that have to be taken into account.

Technique has evolved significantly since 1924. For example, swimmers of the twentieth century were constantly pushing the boundaries of what was considered legal. Indeed, as athletes and coaches learned more about biomechanics, they introduced changes in execution in search of faster and faster times. Some of those changes were downright bizarre, and the high priests and priestesses of FINA (Fédération internationale de natation, today known as World Aquatics), the sport's governing body, have felt the need to respond for the good of the sport.

Consider the case of 1956 Olympian Masaru Furukawa. This bold Japanese swimmer employed a novel approach to win the 200 meter breaststroke event, four laps of the pool swimming on his stomach, upper body bobbing up and down, with arms performing sweeping semicircular movements, combined with a frog kick. He swam most of the event underwater! This stirred up quite a controversy, since he won the breaststroke event without actually swimming the breaststroke as generally understood! FINA responded by forbidding the practice in breaststroke races. For déjà vu, the American swimmer David Berkoff set a world record in the 100 meter backstroke at the 1988 Olympics again by essentially swimming the entire race underwater. FINA responded by outlawing this practice in all of its sanctioned events. Berkoff championed an underwater technique known as the dolphin kick (Figure 3), whereby the swimmer generates speed with aggressive full-body kicks that emulate darting dolphins. For the lucky few, this underwater technique in which the athlete resembles a human torpedo with a powerful fluke tail is much faster than surface swimming.

What is remarkable is that neither Furukawa's nor Berkoff's time would be competitive in 2024. Their times don't stack up! Furukawa would need to cut a whopping 30 seconds from his time to be in contention, an eternity for a sport generally decided by fractions of a second. In Berkoff's case, he would still have five or six meters left to swim when the 2024 medal contenders had finished their races.

Advances in training, nutrition, and recovery strategy have propelled the sport to new heights. But they are only part of the story, and this is where the math and physics come in. The idea is that the minutiae of biomechanical and hydrodynamic constraints, adding to the monumentally difficult challenge of developing athletic talent, are variables of a complex physical and mathematical problem whose optimization can result in an individual's perfect race. For 2024 Olympic hopefuls, the advent of sensor technology has turned this idea into a reality in which mathematics and physics produce actionable items that can help athletes as they strive to reach the limit of their potential.

---

[4] Due to Covid, the 2020 games were held in 2021.

Figure 2. Caeleb Dressel. (Credit: Speedo.)

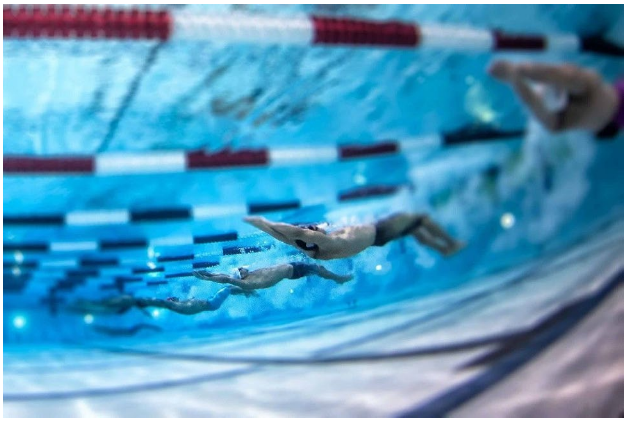

Figure 3. Underwater dolphin kicks. (Credit: Mike Lewis.)

Mathematics and the Physics of Swimming

What Is Swimming?

Swimming involves applying forces to move and control the body through water. Competitive swimmers strive to cut through water as fast as possible. However, the aquatic environment poses unique challenges. Indeed, humans evolved as land creatures, and so it comes as no surprise that our biology is not well suited to the sport of swimming. It's a good thing that the "undersea world" is not invited to the Olympics. Humans would stand no chance against the whip-fast sailfish, which can reach speeds up to 80 miles an hour. Our Olympians are giving it their all to briefly reach five or six miles per hour.

On land, our main foe is gravity, which impedes our forward momentum significantly, leaving air resistance a less-relevant factor. The pumping of a runner's legs at the limits of human performance as they fight gravity creates searing pain. Swimming, however, takes place in water, a dense and viscous fluid. Thrashing one's limbs through water can feel like wading through waist-high mud, but near-neutral human buoyancy minimizes the effect of gravitational forces. The primary obstacle swimmers have to overcome is drag—the frictional forces that push back against the forward motion.

To demonstrate, wave your hand rapidly back and forth through the air and then repeat the motion underwater. You will find it much more difficult to push through water. While water supports the swimmer, it also hinders movement. The density and viscosity that provide buoyancy also create drag. Therefore, swimmers require both power to overcome drag and technique to glide through water with maximum efficiency. The best swimmers find a balance between the two.

Newton's Laws of Motion

While the World Aquatics rulebook [13] that governs Olympic swimming is over 360 pages long, far more fundamental to the way swimming works are Newton's laws of motion [9].

The universal nature of Newton's laws of kinematics governs not just our solar system, but also the minute movements of a swimmer. When a swimmer dives into a pool and begins undulating to propel themselves forward, Newton's laws govern the connection between the propulsive forces generated and the resulting acceleration of the swimmer's body. Whether one is trying to launch a satellite into orbit or analyzing the breaststroke, Newton's discovery underpins our quantitative understanding of dynamics in the skies, on the ground, and in the water.

Newton's first law: the principle of inertia. Translated from his *Principia Mathematica* [9, p. 62], Newton's first law of motion is:

> Every body preserves in its state of being at rest or of moving uniformly straight forward, except insofar as it is compelled to change its state by forces impressed.

This law is the one that drives the need for strokes in swimming. The body is at rest when the race begins, with only the athletes' minds racing. But within tenths of a second, the nervous system is firing, and legs and arms drive the competitors' bodies off the blocks and toward athletic glory. The instant that the body contacts the water after the dive, it immediately undergoes a drag force that would return it to a state of rest from its newly found state of forward motion. Thus, the swimmer must undulate their hips, rotate their arms, and pump their legs to create forces that keep them moving forward.

Let us place this law into the context of the Olympic 50 meter freestyle final, in which eight athletes thrash their limbs with the goal of completing one lap of the pool first. This event is not a competition between athletes. It just looks like that when viewed in person or on screen. Instead, it is eight individual swimmers, each in an individual battle against the physics of inertia, resisting the force of drag, and creating forces that move the body to the finish in the hopes of a gold medal. Indeed, in the world of elite swimming, the only body that may be more compelled to stay at rest is the swimmer's body when their alarm goes off at 4:30 a.m., urging them to get to their predawn practice.

Newton's second law ($\mathbf{F} = m\mathbf{a}$). Newton's second law is burned into the brains of all high-school physics students [9, p. 62]:

> A change in motion is proportional to the motive force impressed and takes place along the straight line in which that force is impressed.

Newton first defined the quantity of motion called momentum as the product of velocity and mass:

$$\mathbf{p} = m\mathbf{v}.$$

Newton called a force applied to a mass over a time interval the impulse. The applied force produces a change in momentum:

$$\mathbf{I} = \mathbf{F}\Delta t = \Delta\mathbf{p}.$$

Taking the limit as the time interval Δt tends to zero, force can therefore be written as

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = m\frac{d\mathbf{v}}{dt} = m\mathbf{a}.$$

In swimming, the second law dictates how forces affect the mass of interest: the human body. Forces of various magnitudes and directions are creating accelerations of various magnitudes on the body in their respective directions. The drag force is accelerating the swimmer against their intended motion, while the violent kick of their leg against the starting wedge creates a force accelerating them. Maximizing forces that accelerate them to the finish is the goal behind explosive and complex weight-room exercises, building up a world-class aerobic system that can power muscles to pull the body through the water and developing a start that flings the body into the water as quickly as possible.

It is thanks to the second law that we first see the opportunity for the use of technology in the training of elite swimmers. The forces swimmers apply can be studied and measured by an inertial measurement unit (IMU). These sensors are known as accelerometers, and they are designed to capture acceleration data. Because a swimmer's mass is constant in a race, the acceleration data, combined with gyroscopic measurements, can accurately represent the thrust a swimmer is generating and also the drag experienced during deceleration. Indeed, a swimmer wants to maximize periods of acceleration while minimizing all sources of unnecessary drag.

Newton's third law: equal and opposite reaction. Newton's third law reads [9, p. 63] as follows:

> To any action there is always an opposite and equal reaction; in other words, the actions of two bodies upon each other are always equal and always opposite in direction.

"Equal and opposite" is how a swimmer turns the movement of the body into a win. As a swimmer's hands push backward against the water, the water pushes the hand, and the body connected to it, forward with equal force. Similarly, when the legs kick back during flutter and dolphin kicks, the reactive force from the water pushes the body forward. Timed improperly, this reactive force is small. But when large muscles pull through the water at carefully calibrated speeds, the water feels like a solid block of concrete, and the body accelerates over the limbs creating these forces. Elite swimmers seek perfection in this area, and talk about having the "feel for the water," while despising the sensation of "slipping strokes." Mastering this level of execution requires years of practice and training. Furthermore, this level of mastery is so delicate that even the world's top swimmers complain about losing the "feel" after a few days out of the water. This is further indication that we are land animals. Swimming is unnatural for the human body.

Applied Fluid Dynamics

Finally, we turn to the issue of drag—a swimmer's worst enemy. Determined by Bernoulli's equation for pressure in a fluid and the definition of pressure P as force F per unit area A, that is, $P = F/A$, the drag force for an object in motion in a fluid can be written as "dynamic pressure" over an area:

$$F_{\mathrm{d}} = \frac{1}{2}\rho v^2 C_{\mathrm{d}} A.$$

Here, ρ is the density of the fluid, A is the area of the object in motion, v is its velocity, and C_{d} is its drag coefficient. Reducing this drag force is a cornerstone of Olympic glory, and swimmers turn to everything from shaving the hair off their bodies to compressive suits that minimize the coefficient of drag. Our 2024 Olympians will sport carbon-fiber-reinforced tech suits that are crafted to *just barely* accommodate their chiseled bodies. For women, getting "suited" is a Herculean task that resembles a contorted wrestling match with the sharkskin garment. In some cases, especially for women, it can take half an hour or more to get into these suits. However, these steps are all worth the precious fractions of seconds saved in the heat of competition.

Swimmers carefully build muscle in regions of their bodies that won't affect the quantity $C_{\mathrm{d}}A$, and the best possess the famous "swimmer's build" (the V-shaped torso of broad shoulders and narrow waist) to slice through the water. Many develop the ability to contort their bodies with flexibility drills, and they apply their yoga-enhanced skills throughout their training. The goal is to reduce their frontal area to a fraction of its resting position, achieved with awkward head positions, hips that ride high on the surface of the water, and so on. The fact is that races are often won and lost based on the quality of an athlete's streamline, a position in which an athlete aims to glide sleekly underwater before commencing with surface swim strokes. Indeed, races can be lost in the segments of a swim that involve no actual swimming!

Applying Science and Mathematics to Swimming

Since 2015 [2, 4, 5, 7], teams of researchers at Emory University and the University of Virginia, led by the fourth author, have been combining the physics of Newton's laws with mathematical modeling and optimization with the goal of enhancing the training of elite swimmers. The idea is to make use of sensitive accelerometers fitted with internal gyroscopes and accurate directional force meters. Athletes perform a battery of tests with these sensors, and the collected data are used to fashion "digital twins" of the athletes. The granular data captures much more information than ordinary digital video, which generally records an image at 24 frames per second. In contrast, our sensors capture movements and (directional) generated force 512 times per second.

IMUs and Force Sensor Bands

We have been using inertial measurement units (IMUs) to capture acceleration data in each of the three traditional coordinate axis directions. These sensors can be placed on swimmers' wrists, ankles, or back (see Figure 4) to quantify precisely how the swimmer is accelerating. The effect from every rotation, splash, pull, and kick can be quantified in each direction and analyzed.

Recently, we have begun employing advanced sensors (see Figure 5) that measure force generated by an athlete's hands. These high-tech bands are placed on the athletes' hands to measure the pressure differential between the palm and the side of the hand. These sensors generate nuanced force field data that quantify the otherwise incredibly complex fluid dynamics that govern a hand propelling through the water. What was previously evaluated purely by looking at the swimmer above the water can now be distilled into a sequence of vector fields that show the distribution of force in all three axial directions. Force applied in any direction other than forward is not helping an athlete achieve their dream of Olympic gold.

What Do We Do with All This Data?

We use these streams of numbers to assemble an athlete's digital twin. The twin captures an athlete's movements down to the millisecond. We have assembled a massive database of digital twins from over one hundred of the best swimmers in America. Thanks to the cooperation of USA Swimming and leading collegiate coaches from across America, we have tested and assembled twins of numerous NCAA champions, Olympic medalists, and world champions.

Figure 4. An IMU on athlete Sebastien Sergile. (Credit: Jerry Lu.)

Figure 5. A force sensor band. (Credit: eo SwimBETTER.)

What do we do with these digital twins? They allow us to make recommendations that immediately improve technique, offer suggestions for race strategy, and point

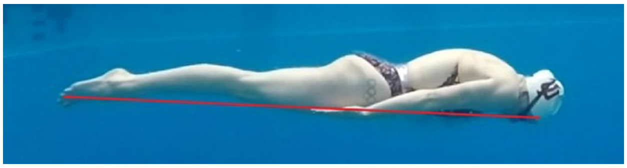

Figure 6. Lilly King in streamline. (Credit: Ken Ono.)

Figure 7. Kate Douglass in streamline in November 2020. (Credit: Ken Ono.)

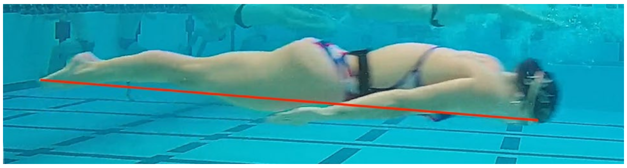

Figure 8. Kate Douglass in streamline in October 2023. (Credit: Ken Ono.)

to long-term aspirational goals—all in pursuit of the optimal race.

In terms of technique, we are able to digitally identify an athlete's comparative strengths and weaknesses without having to hold an actual live race. Obviously, the identification of a technical flaw helps a coach offer immediate precision training that leads to improved performance. Furthermore, the digital twin quantifies the severity of a flaw. Indeed, thanks to Newton's equations and the acceleration data, we are able to accurately predict the time savings that an athlete can expect to achieve with a given change. It boils down to the numerical integration of the acceleration data, since these values are the derivatives of velocity. Thank you, Newton!

Flaws come in many forms; examples include poor head position, anchoring of the legs, imbalance in body rotation, and breathing inefficiencies, to name but a few. To give one concrete example, consider the execution of streamline in breaststroke, which is the underwater glide phase in which one isn't even really swimming. The goal is to preserve as much speed as possible off the opening dive and after pushing powerfully off walls in exiting turns. One might think that there is little opportunity for improvement in these phases of breaststroke, since the swimmer seems to be doing nothing at all. On the contrary, races can be won or lost, and records set, during this innocuous-seeming phase.

Figure 6 shows the textbook streamline of 2016 Olympic gold medalist Lilly King. On the other hand, take a glance at Figure 7 from a November 2020 test with Kate Douglass, the first author and former University of Virginia collegian. By comparison, and without the need for expertise, one can guess that her head position introduces extra turbulence and drag. Her digital twin allowed us to quantify the significance of this flaw. Using numerical integration, which takes into account the dynamics of her personal active A and C_d while in motion, we predicted that she stood to gain 0.10 to 0.15 seconds per streamline glide by making suitable modifications. In the 200 meter breaststroke event, an athlete performs four of these streamline glides, and so we predicted that this one recommendation could amount to a 0.4 to 0.6 second time savings.

Douglass meticulously worked to fine-tune her streamline glide to near perfection. To illustrate, consider Figure 8, from October 2023. After 36 months of work, Douglass's improved technique resulted in a 0.11 second savings on average per streamline glide, which amounts to 0.44 seconds in the 200 meter breaststroke event. Is that savings significant? Twelve weeks after this test, Douglass broke the twelve-year-old American record in the event with a time of 2 minutes, 19.3 seconds, dipping under the previous mark by 0.29 seconds.

This quantitative analysis of Douglass's streamline is just one of countless examples that confirm the utility of digital twinning. We have been applying this type of analysis to the myriad body movements and positions that impact speed and drag. The resulting recommendations, with their predictive power, have helped coaches and athletes prioritize immediate technical targets of opportunity.

To offer another concrete example, let's revisit Douglass's American record in breaststroke. Obviously, the record lends credence to the idea that "big data" can enhance training. However, what is more remarkable is her meteoric rise in the event. At the time of her first test, in November 2020, the 200 meter breaststroke was not on her event list. She didn't race the event at all in 2020, since her personal best before college was a good, but not stellar, 2 minutes, 30.4 seconds. Within hours of compiling her digital twin, we knew that she had both the physical ability and aerobic capacity to compete at the world championship level. We then ran the simulations, and in turn, supplied a list of targets of opportunity if she chose to pursue breaststroke.

Clearly, Douglass made the decision to add breaststroke to her event list, and after 36 months of hard work, she became the best 200 meter breaststroker in American history. This painstaking process is difficult for both the athlete and the coach. The world's best coaches are remarkable individuals; they have a gift for helping their athletes reach their potential through a delicate balance of courage, dedication, patience, and virtue. Douglass honed her technique and unusual timing under the watchful eye of mastermind Todd DeSorbo, University of Virginia and 2024 U.S. Women's Olympic Head Coach.

Digital twins also play an important role in devising race strategy. Analysis of a twin can lead to suggested changes in tempo, modifications of the timing of body movements, adjustments to the number of kicks taken in various

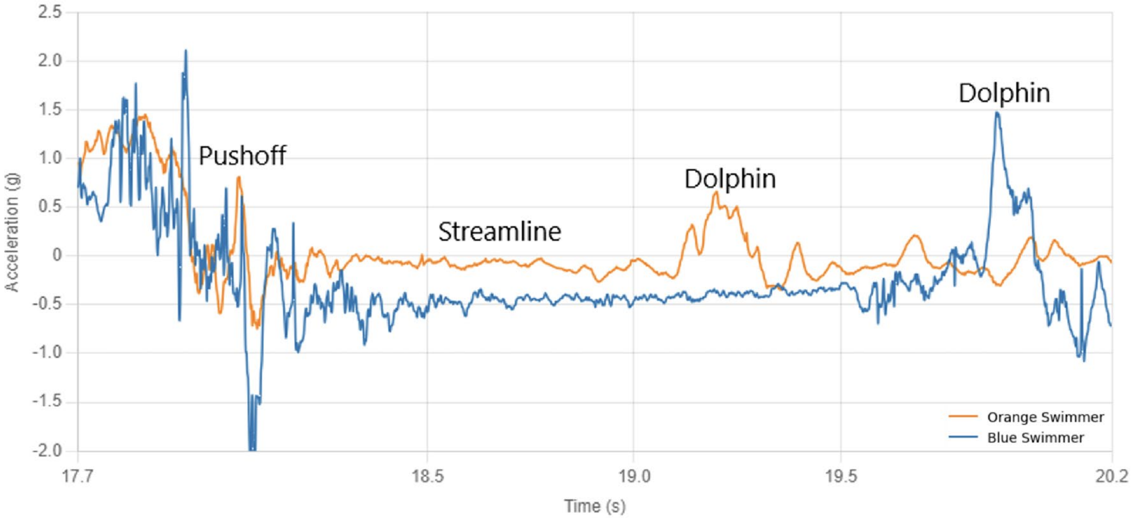

Figure 9. Comparing two breaststrokers. (Credit: Jerry Lu.)

phases, or recommended breathing patterns. Should an athlete breathe on both sides in freestyle? How many breaths should be taken in a 100 meter sprint?

Since there are so many variables and little time for experimentation with athletes in training, the digital twin is a godsend. The luxury of having an athlete's digital doppelgänger is easily understood through our ability to run different race scenarios. This isn't the stuff of Nintendo's Mario Brothers, where you can race the Mario Kart ghosts for fun. This is the stuff of Olympic dreams! The goal is to determine an athlete's optimal "formula" for execution, which can be updated with minimal retesting.

What does this look like? In Figure 9, we compare the digital twins of two elite breaststrokers executing the first phase of a "pullout," which fans cannot see, since it takes place underwater. The pullout phase consists of a powerful push off the wall followed by a streamline glide, and it ends with a single dolphin kick. The graph in the figure overlays the acceleration in the direction of the swim measured in g's, gravitational acceleration. One can see that the orange swimmer has an extraordinary streamline, since her graph sits slightly below $0g$, reflecting almost no deceleration. On the other hand, the blue swimmer decelerates significantly in glide. The orange breaststroker also has a weaker dolphin kick, which she executes almost one second earlier. In terms of strategy, the orange swimmer might consider delaying the execution of the dolphin kick due to her superior streamline and weak kick, while the other breaststroker might want to execute her more powerful kick earlier to mitigate the inferiority of her glide. By running different simulations, we are able confirm these speculations, offer optimal timing of execution with confidence, and also provide the expected time savings to boot. Why guess?

Finally, this quantitative approach to swim analysis can be used to formulate aspirational goals that can become reality after months and years of extensive training. Indeed, some of the desired simulations are not realistic, given an athlete's current aerobic capacity. After all, the digital twin doesn't feel the pain of burning muscles and oxygen-starved lungs. Over time, however, a coach might be able to help an athlete increase their aerobic capacity, transforming an unrealistic simulation into a genuine race strategy.

For example, consider a sprint butterfly specialist who has world-class underwater dolphin kicks that shoot her forward powerfully off the dive and out of turns. A simulation might recommend that this athlete take ten or eleven dolphin kicks off the dive and the same number after the turn in the 100 meter butterfly sprint. At the world-class level, such a recommendation would have been considered loony a few years ago, since these kicks are expensive in terms of oxygen consumption. Indeed, the 2016 Olympics was won by the Swede Sarah Sjöström, who took ten kicks off the dive and only seven off the turn.

For the world's best dolphin kickers, the ability to squeeze in one extra kick might cut 0.1 or 0.12 seconds off their time. For them, the long-term training required to develop the necessary aerobic capacity might be worth it. How are these athletes identified? The underwater tests that are part of the crafting of a digital twin play a central role. To underscore this point, we note that the Canadian swimmer Maggie Mac Neil won the 2019 World Championships by taking nine and ten kicks, and the following year, the American Claire Curzan won the 2020 U.S. Open by taking even more kicks: ten and eleven.

Some More Mathematics

There is a lot of math that goes into this analysis. Most important are the calculations related to Newton's three laws of motion as described above. But there is even more.

Readers familiar with accelerometer data will know that massive streams of high-frequency sensor data are noisy. To deal with this, modern data visualization techniques help us make sense of the time series performance data at hand. Smoothing splines [3] is a way to fit discrete data points into a continuous curve, which provides meaningful visualizations. A smoothing spline is usually constructed as a function f with a smoothing parameter λ that minimizes

$$\sum_{i=1}^{n}(y_i - f(x_i))^2 + \lambda \int f''(x)^2 \, dx \, ,$$

where n is the number of data points, (x_i, y_i) is the acceleration time series data, f is our fitted function, and λ is our regularizer. In this minimization function, the first term,

$$\sum_{i=1}^{n}(y_i - f(x_i))^2 \, ,$$

is the residual sum of squares of the model, and

$$\lambda \int f''(x)^2 \, dx$$

is the penalty term for the roughness of the function, where $f''(x)$ is the rate of change of the slope of f at x regularized by λ. When we fit our function f to data, we control the smoothness by penalizing its integrated squared second derivative.

For our acceleration data, we define λ as $10n - 1$ to control smoothing. The choice of an n-dependent regularizer is an empirical one based on our desired smoothing level and characteristics of the data. For context, we note that choosing $\lambda = 0$ leads to interpolating between the data points.

High-Definition Video

As much as we trust the numbers and our calculations, we have to say that we don't ignore what the eyes can see. Indeed, we record all of our sensor tests with an extraordinary underwater camera system. This high-definition video allows us to corroborate the integrity of the data. Furthermore, "a picture is worth a thousand words." We are able to include video evidence that speaks to our recommendations. Although the video does not allow us to analyze stroke mechanics in extreme detail at all angles, the footage generally offers glimpses of technical flaws that stand out in the data. This video evidence, combined with a mathematical explanation, assists coaches in prescribing drills to correct technical inefficiencies and flaws.

Does It Help?

The fourth author, Ken Ono, began this research at Emory University, a powerhouse in the nonscholarship realm of collegiate sports. Andrew Wilson, a walk-on member of the school's swim and dive team, serendipitously enrolled in Ken's fall 2014 course on number theory, marking the first step of this unlikely story. Although Wilson would end up writing an honors thesis in the abstract theory of elliptic curves, earning him graduate fellowship offers from MIT, Oxford, and the University of Texas,[5] the work described here began as an academic curiosity. It began with amateur experiments involving surgical gloves, Saran Wrap, and accelerometers designed for sharks.

Wilson's work ethic is the stuff of legend. He left no stone unturned. He took advantage of every resource available to him, from the Emory coaches, to USA Swimming staff, to coaching icons Jack Bauerle and Eddie Reese, and ultimately to the math and science described here. He made history [10] as the first NCAA Division 3 swimmer to be named a U.S. Olympian. He won a coveted gold medal to boot.[6]

Head coach Todd DeSorbo and his world-class coaching staff at the University of Virginia (UVA) embraced the math and science when Ken joined UVA's Department of Mathematics in fall 2019. He began his work as an external consultant to the team with the 2020–2021 NCAA season. The crafting of digital twins involves four or five testing sessions each season, each followed by hours of work involving computer simulations and numerical calculations. The math and science then provides each athlete a list of personalized recommendations and targets of opportunity. It must be stressed that the athletes and coaches perform all of the actual hard physical and mental preparation. The coaches work their magic, using these recommendations as guidance, and the athletes swim thousands of laps fine-tuning their bodies in pursuit of perfection.

But does it work? The UVA women won their first team national championship that season, and they have not lost since. From their first team title in 2021 to the present day, the women's team has rewritten the record books. UVA women currently own the NCAA and American records in four of the five relay events, and five individual NCAA and American records. The UVA men's team owns the American record in the 200 yard freestyle relay. All told, UVA owns more records than any other collegiate team in America. In terms of the Olympic Games, four 2021 U.S. Olympians had the benefit of digital twin modeling. Each of them returned from Tokyo with an Olympic medal.

And at the recently completed 2024 world championships in Doha, the American Claire Curzan (Figure 10, left), who transferred from Stanford to UVA to work on this project, swept the 50, 100, and 200 meter backstroke events for gold medals after only a few months with us, obtaining her best personal times in all three events.

---

[5] Wilson is presently a doctoral student in applied mathematics at Oxford University.

[6] Wilson swam the breaststroke leg of the 2021 men's medley relay in preliminaries.

Figure 10. World champions Claire Curzan and Kate Douglass. (Credit: Mike Lewis.)

We should also add that our coauthor Kate Douglass (Figure 10, right) won gold in the 200 meter individual medley, swimming the sixth fastest time ever recorded. And she won silver in the 50 meter freestyle, setting an American record of 23.91 seconds, the fourth fastest time ever recorded. In all, United States swimmers took home twenty medals. Seven of those were won by women in individual events. In fact, Kate Douglass and Claire Curzan won all seven of them. GoMath! And to ice the cake, Claire was named Female Swimmer of the Championships.

Paris 2024

The return of the Olympics to Paris after a 100-year hiatus offers an elegant opportunity to reflect on a central pillar of the Olympics—its offering of a consistent tradition to a constantly evolving world. The Seine will still flow throughout the competition, and the Eiffel Tower will still preside over many of the same events as it did a century ago. However, among these stoic landmarks are a city, and games, that have been unquestionably altered by the modern era in which they take place, one full of science, electronics, and an abundant supply of resources. These leaps in technology will result in equally magnificent leaps in performance, with athletic feats that would have been quite literally unimaginable 100 years earlier. These athletes, armed with troves of data, refined training techniques, and complex analytics, demonstrate the beauty of the games, as both a driver and display of what humans and technology can achieve, redefining our common limits.

Millions will watch the swimming events unfold in Paris's La Défense Arena. Many American swimmers will make Olympic history with medals and records, and for some, hidden on computers out of sight, will be their digital doppelgängers that were somehow also part of the team.

Acknowledgments

The fourth author thanks Jane Street Capital, the National Science Foundation (DMS-2002265 and DMS-2055118), the National Security Agency (H98230-23-1-0016), the Templeton World Charity Foundation, and the Thomas Jefferson Fund for their support. The authors thank Blaire Bachman, Jack Bauerle, Sean Conway, Claire Curzan, Maddie Dawsey, Todd DeSorbo, Tyler Fenwick, Cindy Fontana, Kirk Grand, Curtis Goss, Bobby Hackett, Coleman Hodges, Jon Howell, Taylor Knibb, Mike Lewis, Abi Liu, Paige Madden, Russell Mark, Sven Mehisovic, Fiona Muir, Sage Ono, John Petroff, Keenan Robinson, Larry Rolen, Megan Santaniello, Sebastien Sergile, Gary Taylor, Andrew Wilson, and Jaycee Yegher for their help along the way. They also thank the folks at Sideline Scout and eo SwimBETTER for their technical assistance.

Open Access

References

[1] K. Abbott. The 1904 Olympic marathon may have been the strangest ever. *Smithsonian Magazine*, August 7, 2012. Available at https://www.smithsonianmag.com/history/the-1904-olympic-marathon-may-have-been-the-strangest-ever-14910747/.

[2] K. Barnes. The sorcery and science behind Virginia's swimming dynasty. *ESPN Magazine*, March 14, 2023.

[3] C. de Boor. *A Practical Guide to Splines*. Springer, 1978.

[4] M. Dawsey and K. Ono. Seeking speed. *Splash*, Summer 2019, 38–39.

[5] M. De George. The math behind the medals. *Swimming World*, January 31, 2023.

[6] International Olympic Committee (2024). Olympic Games. Available at https://olympics.com/en/.

[7] National Broadcasting Company. How data has UVA swimming's Olympians on top. Available at https://www.nbcsports.com/watch/how-ono-uses-data-to-refine-olympians-swimming, November 21, 2023.

[8] National Collegiate Athletic Association (2024). Estimating probability of competing in college athletics. Available at https://www.ncaa.org/sports/2015/3/2/estimated-probability-of-competing-in-college-athletics.aspx.

[9] I. Newton. *The Principia: The Authoritative Translation and Guide: Mathematical Principals of Natural Philosophy*, edited by I. Bernard Cohen. University of California Press, 2016.

[10] K. Suigura. Emory grad Andrew Wilson's spot on U.S. Olympic team is "mind-blowing." *Atlanta Journal Constitution*, June 25, 2021.

[11] Summer Leagues Swimming (2024). Available at https://www.summerleagueswimming.com.

[12] Swimswam (2024). Speedo world swimming rankings. Available at https://swimswam.com/ranking/.

[13] World Aquatics (2024). Swimming rules. Available at https://www.worldaquatics.com/swimming/rules.

Katherine Douglass, Department of Statistics, University of Virginia, Charlottesville, VA 22904, USA. E-mail: kcd6znz@virginia.edu

Augustus Lamb, School of Data Science, University of Virginia, Charlottesville, VA 22904, USA. E-mail: abl6ywp@virginia.edu

Jerry Lu, Sloan School of Management, MIT, Cambridge, MA 02139, USA. E-mail: jerry_lu@mit.edu

Ken Ono, Department of Mathematics, University of Virginia, Charlottesville, VA 22904, USA. E-mail: ken.ono691@virginia.edu

William Tenpas, School of Data Science, University of Virginia, Charlottesville, VA 22904, USA. E-mail: wet9me@virginia.edu