

Article

Mathematical Modeling of Stroke Rate and Stroke Length for Swimming Speed and Energy Efficiency

Haochen Luo^{1,*}

¹ Shanghai Starriver Bilingual School, Shanghai, China

* Correspondence: Haochen Luo, Shanghai Starriver Bilingual School, Shanghai, China

Abstract: This study develops a theoretical framework for identifying stroke rate and stroke length combinations that balance swimming speed and energy expenditure. Cycle averaged speed is expressed as their product, while a simplified hydrodynamic model links metabolic power to drag and effective efficiency. Energy expenditure is evaluated per unit distance and over a fixed event distance. The analysis shows that constant efficiency determines an energy efficient speed but cannot identify a unique stroke combination. A dimensionless calculation demonstrates why swimming more slowly does not necessarily minimize total energy expenditure. An assumed efficiency function with quadratic penalties in logarithmic stroke variables then yields an explicit optimal rate and length allocation at a given speed. Physiological and technical limits restrict these solutions, and an event specific speed requirement or energy budget selects among feasible tradeoffs. A calibration and validation protocol is provided for practical application. The results are conditional analytical deductions; individualized parameter estimates and empirical performance improvements remain to be established.

Keywords: swimming biomechanics; stroke rate; stroke length; energy expenditure; multiobjective optimization

1. Introduction

Swimming speed depends on how frequently a swimmer completes a stroke cycle and how far the swimmer travels during that cycle. Increasing either quantity can increase speed if the other is maintained, but neither can generally be adjusted independently of technique, fatigue, or force production. A review by Zamparo, Cortesi, and Gatta identifies hydrodynamic resistance and propulsion efficiency as major determinants of swimming energy cost [1]. The practical question is therefore how to choose a feasible stroke rate and stroke length combination that delivers sufficient speed with an acceptable energy demand.

The research question addressed here is: How can mathematical modeling identify the optimal combination of stroke rate and stroke length to maximize swimming speed while minimizing energy expenditure? These objectives require a defined comparison. Metabolic power measures energy use per second, whereas total energy over a fixed distance also depends on completion time. Minimizing power alone can favor very slow swimming without minimizing the energy needed to finish an event. This paper uses energy cost per distance and a fixed event distance to make the speed and energy objectives comparable.

The central hypothesis is that technique dependent efficiency and physiological limits determine which combinations are advantageous, and that the preferred solution depends on the swimmer and the event. The study develops a theoretical model, derives an analytical benchmark, and specifies how individual measurements could support calibration. Its contribution is a transparent formulation of the decision problem and its identifiability limits. The results are mathematical consequences of stated assumptions, rather than evidence from an athlete intervention.

Received: 12 August 2026

Revised: 19 September 2026

Accepted: 01 October 2026

Published: 08 October 2026



Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Related Research and Study Scope: Experimental evidence supports distinguishing swimming styles and intensities. Barbosa and colleagues examined 18 swimmers assigned to groups for four competitive strokes and related energy cost to stroke frequency, stroke length, and speed. At a given speed, higher stroke frequency was associated with higher energy cost in backstroke, breaststroke, and butterfly; the corresponding relationship was not significant in freestyle. Longer stroke length was associated with lower energy cost in breaststroke [2]. A separate study of 26 international swimmers in stroke specific groups found that the energy costs of the four strokes differed, with front crawl showing the lowest cost over the investigated speed range [3]. These findings do not establish a universal instruction to reduce rate or increase length.

A nonlinear response surface study of elite freestyle swimming investigates the joint relationship of stroke rate and stroke length with speed [4]. Such modeling is relevant to the search for favorable combinations, but a speed model alone does not resolve energy expenditure. Moreover, when stroke length is calculated from speed and rate, their relationship partly follows from the definition of stroke length. Predictive accuracy for speed must therefore be interpreted in relation to the measurement procedure.

A mathematical study of swimming stroke patterns also illustrates that changing the temporal coordination of stroke rate and stroke length can change finishing time and modeled mechanical work [5]. Mechanical work is not a direct measurement of metabolic expenditure, and the simulations were not experimentally validated. They provide a conceptual comparison rather than individual calibration data. Building on these strands, the present framework connects a cycle averaged kinematic identity with a metabolic energy model and an explicit feasible set. It focuses on steady clean swimming within a single stroke style; starts, turns, and underwater phases require additional components.

2. Mathematical Model

The model separates observable stroke variables from physiological parameters and event objectives. All comparisons refer to the same swimmer, stroke style, environment, and event distance unless stated otherwise. Table 1 summarizes the quantities used in the equations.

Table 1. Model Variables and Parameters

Symbol	Definition	Unit
r ; R	Stroke rate in cycles per second; cycles per minute	cycle s ⁻¹ ; cycle min ⁻¹
L	Distance traveled per complete stroke cycle	m cycle ⁻¹
v	Cycle averaged clean swimming speed	m s ⁻¹
d ; T	Modeled steady distance; duration	m; s
P ; P_0	Metabolic power; baseline nonpropulsive power	W
K	Effective quadratic drag coefficient	kg m ⁻¹
η	Effective conversion efficiency	dimensionless
C ; E	Metabolic energy cost per distance; total energy	J m ⁻¹ ; J
X	Feasible set of rate and length combinations	not applicable

2.1. Stroke Definitions and Measurement

Stroke rate r is measured in complete cycles per second, and stroke length L is the distance traveled per complete cycle. For front crawl and backstroke, a cycle includes both arm actions; counting individual arm actions would require a corresponding change in the length definition. If R is reported in cycles per minute, the conversion is included explicitly:

$$v = rL = \frac{RL}{60} \quad (1)$$

Measurements should use a clean swimming segment that excludes acceleration after the start and the effects of a turn. Rate can be estimated from the time for several complete cycles. Speed can be estimated from the distance and time of the same segment. If L is then calculated as v/r , Equation (1) is an identity and reproducing it is not an independent

validation of the model. The optimization concerns achievable stroke combinations and their metabolic consequences, rather than proving the identity.

2.2. Drag and Metabolic Energy Expenditure

For a cycle averaged, approximately steady condition, hydrodynamic drag is represented by $F_D = Kv^2$. K summarizes water density, frontal area, and an effective drag coefficient over the calibrated range. The mechanical power needed to overcome drag is $F_D v$. Effective efficiency η combines the conversion of metabolic energy into mechanical work and useful propulsion; it is not identical to propelling efficiency alone. A simplified metabolic power model is:

$$P(r, L) = P_0 + \frac{Kv^3}{\eta(r, L)} \quad (2)$$

P_0 represents approximately speed independent nonpropulsive metabolic power, with $P_0 > 0$, $K > 0$, and $0 < \eta \leq 1$. The drag and efficiency interpretation follows the energetic framework reviewed in [1], while the additive baseline is a simplifying modeling assumption. Cycle fluctuations, stroke dependent changes in drag, and anaerobic contributions can require a richer model. For a fixed modeled clean swimming distance d and constant speed, energy cost C , energy E for that distance, and its duration T are:

$$C = \frac{P}{v} = \frac{P_0}{v} + \frac{Kv^2}{\eta}, \quad E = dC, \quad T = \frac{d}{v} \quad (3)$$

C and E are proportional when d is fixed, so minimizing either produces the same optimum. P is measured in watts, C in joules per meter, and E in joules. The baseline term accounts for the energy expended during a longer completion time. Its magnitude depends on whether gross or net metabolic expenditure is measured; these definitions must be kept consistent during calibration. Whole race predictions require separate start, turn, and underwater phase contributions; $T = d/v$ describes only the modeled steady portion.

2.3. Feasible Stroke Combinations

The feasible set X contains positive rate and length pairs that a swimmer can maintain for the target duration. Its constraints include measured rate bounds, a rate dependent interval of achievable stroke lengths, and a duration dependent metabolic power ceiling. A joint interval $L_{min}(r) \leq L \leq L_{max}(r)$ is preferable to treating every point in a rectangular rate and length range as attainable. Event distance, technique, and fatigue determine the limits. Heart rate can support monitoring but does not, by itself, determine total metabolic power or energy expenditure.

The model assumes X is nonempty and compact and that the objective functions are continuous, with speed and efficiency bounded away from zero on X . These conditions ensure that constrained optima exist. They do not guarantee a unique pair. Anthropometry, strength, flexibility, and skill enter through the measured feasible set or calibrated parameters; their effects should not be assigned numerical values without supporting measurements.

2.4. Optimization Objectives and Selection

The two objectives are to maximize speed and minimize energy cost over X :

$$\text{maximize}_{(r, L) \in X} v(r, L), \quad \text{minimize}_{(r, L) \in X} C(r, L) \quad (4)$$

A pair is Pareto optimal if no other feasible pair is at least as fast and at most as costly, with a strict improvement in one objective. The Pareto set describes the available tradeoffs. It need not be a smooth curve, and several stroke pairs can map to the same objective values. Choosing a race or training objective then selects a particular point or a small region.

For an endurance session, a clear decision rule is to minimize energy cost subject to a required speed v_{req} :

$$\text{minimize}_{(r, L) \in X} C(r, L) \quad \text{subject to} \quad rL \geq v_{req} \quad (5)$$

For a fixed event with an available energy budget, an alternative is to maximize speed subject to $dC \leq E_{max}$. An infeasible requirement should be reported as infeasible. Weighted sums may also be used after speed and energy are normalized, but their weights

express a preference and cannot be inferred from mechanics alone. An energy allowance, target time, or training purpose is therefore needed to turn the research question into a specific recommendation.

3. Analytical Results and Solution Methods

The constant efficiency case provides an exact benchmark for interpreting the objectives and checking a numerical implementation. The technique dependent case then explains what additional information is needed to select a stroke combination.

3.1. Constant Efficiency Benchmark

Let $\eta = \eta_0$ be constant and retain a constant K . Differentiating Equation (3) gives the unconstrained energy efficient speed v_e :

$$v_e = \left(\frac{P_0 \eta_0}{2K} \right)^{\frac{1}{3}} \quad (6)$$

The second derivative is $2P_0/v^3 + 2K/\eta_0 > 0$ for positive v , so this stationary point is the unique unconstrained minimum of C as a function of speed. If feasible speeds form an interval, the constrained minimum is v_e when it lies inside that interval and the appropriate endpoint otherwise. With a feasible minimum speed requirement, the minimum cost speed is the projection of v_e onto the remaining feasible interval. This interval argument must not be applied to a disconnected feasible set without checking each component. This interior optimum depends on including a positive baseline in expenditure. If a net expenditure model sets $P_0 = 0$, cost is Kv^2/η_0 and increases with speed, so its minimum occurs at the lowest feasible speed.

Define $u = v/v_e$. Using the stationary condition produces a dimensionless result that requires no assumed athlete parameters:

$$u = \frac{v}{v_e}, \quad \frac{C(v)}{C(v_e)} = \frac{\frac{2}{u} + u^2}{3} \quad (7)$$

Table 2 evaluates Equation (7). These values are exact model calculations rounded to four decimals, rather than experimental observations. Swimming at half the energy efficient speed costs approximately 41.7% more energy per meter because the longer duration increases baseline expenditure. At 1.5 times that speed, cost is approximately 19.4% higher because drag related expenditure increases. Both results depend on the stated constant efficiency and baseline assumptions.

Table 2. Dimensionless Energy Cost in the Constant Efficiency Model

Speed ratio u	Cost ratio $C(v)/C(v_e)$	Comparison if $u = 1$ is feasible
0.50	1.4167	Slower and more costly
0.75	1.0764	Slower and more costly
1.00	1.0000	Minimum energy cost
1.25	1.0542	Faster with higher energy cost
1.50	1.1944	Faster with higher energy cost

If v_e is feasible, every feasible speed below v_e is dominated by v_e in the constant efficiency benchmark. Above v_e , increasing speed increases cost, producing a Pareto tradeoff over the feasible speed interval. However, all feasible pairs with $rL = v$ have the same cost when efficiency is constant. The benchmark therefore identifies a speed, not a unique rate and length. This underdetermination is central to the research question.

3.2. Technique Dependent Efficiency

When η varies with rate and length, different pairs producing the same speed may have different metabolic costs. With K and P_0 fixed, the minimum cost pair at a target speed is the feasible pair with the highest effective efficiency:

$$r^*(v) \in \arg \max_{r \in R(v)} \eta \left(r, \frac{v}{r} \right), \quad L^*(v) = \frac{v}{r^*(v)} \quad (8)$$

Here $R(v)$ is the set of rates for which $(r, v/r)$ lies in X . If it is empty, that speed is unattainable. If the efficiency maximum is unique, a unique pair can be selected; otherwise several pairs remain equivalent. Evaluating this inner problem over attainable speeds

constructs a minimum cost envelope. Nondominated points on that envelope answer the speed and energy tradeoff. A longer stroke is advantageous at a fixed speed only when its associated lower rate improves this envelope without violating technical constraints.

To obtain a solvable technique dependent example, assume that effective efficiency is reduced by departures from a reference rate r_0 and length L_0 through quadratic penalties in logarithmic variables:

$$\eta(r, L) = \eta_{max} \exp \left(-a \ln \left(\frac{r}{r_0} \right)^2 - b \ln \left(\frac{L}{L_0} \right)^2 \right) \quad (9)$$

Here a and b are positive dimensionless penalty coefficients, r_0 and L_0 are positive reference values, and $0 < \eta_{max} \leq 1$. The reference speed is $v_0 = r_0 L_0$. This function is an explicit local modeling assumption, rather than a measured efficiency law. It provides a conditional solution that can be tested and calibrated. With $x = \ln(r/r_0)$ and $s = \ln(v/v_0)$, the fixed speed constraint gives $\ln(L/L_0) = s - x$. Maximizing efficiency is equivalent to minimizing $ax^2 + b(s - x)^2$. Its second derivative is $2(a + b) > 0$, so the unconstrained optimum is unique:

$$r^*(v) = r_0 \left(\frac{v}{v_0} \right)^{\frac{b}{a+b}}, \quad L^*(v) = L_0 \left(\frac{v}{v_0} \right)^{\frac{a}{a+b}} \quad (10)$$

The resulting product is exactly v . When a is large relative to b , the rate stays closer to its reference value and more of the speed change is assigned to length; when b is larger, more is assigned to rate. If the feasible rates at the target speed form an interval, the optimum is obtained by clipping the unconstrained rate to that interval and setting $L = v/r$. For a disconnected feasible set, each component must be examined. The formula therefore identifies a specific conditional combination without asserting a universal stroke prescription.

Substituting the optimal pair into Equation (9) gives the maximum efficiency at each speed whenever that pair is feasible. Equation (3) then evaluates its energy cost, and nondominance filtering selects the speed and energy tradeoffs. If constraints force a different pair, its actual efficiency must be used instead. Metabolic measurements alone generally identify K/η_{max} rather than its components separately; independent drag measurements are needed to distinguish hydrodynamic change from efficiency change.

3.3. Numerical Solution and Sensitivity

Because the decision space has two variables, a measured feasible grid combined with the fixed speed inner problem provides an interpretable initial solution. Grid refinement and local constrained searches can check whether apparent optima persist. For a more complex or disconnected feasible set, the nondominated sorting genetic algorithm II, NSGA-II, offers an established method for approximating a Pareto set [6]. It should be compared with the grid result, repeated with different random seeds, and checked for constraint violations. Its use does not establish optimality or validate the energy model.

In the constant efficiency benchmark, the logarithmic sensitivities of v_e to P_0 , η_0 , and K are respectively $1/3$, $1/3$, and $-1/3$. Thus a small relative change in each parameter produces one third as large a relative change in v_e , with greater drag lowering the minimum cost speed. For the calibrated technique dependent model, uncertainty should be propagated by resampling trials or varying parameter estimates within supported ranges, then recomputing feasible nondominated solutions. A narrow optimum that disappears under small plausible changes is unsuitable as a precise training target.

4. Calibration and Practical Application

Application requires measurements that connect feasible stroke combinations to metabolic expenditure for the intended swimmer and event. The following protocol specifies the evidence needed to move from theoretical results to an individualized recommendation.

4.1. Measurement and Parameter Estimation

Repeated swimming trials should cover several speeds and several feasible rates at each speed, using a pacing cue where appropriate. Pool conditions, stroke style,

equipment, and the clean swimming segment should be recorded consistently. Sufficient recovery and counterbalanced trial order can reduce fatigue and order effects. Synchronized video or validated sensors should record cycles, segment time, and within trial changes in technique. Anthropometric and training information can help interpret differences between swimmers.

Metabolic power should be estimated using a measurement method appropriate to trial intensity and duration. Oxygen uptake supports assessment of aerobic expenditure, while high intensity trials also require a defensible treatment of anaerobic contributions. The experimental approaches in [2] and [3] show why speed, oxygen uptake, and lactate related information are relevant, but their published participant counts do not constitute a dataset for this study. Stopwatch records or elite race times alone cannot estimate energy expenditure.

Fit the energy model to measured metabolic expenditure, with parameter constraints that preserve positivity and feasible predictions. Establish whether expenditure is gross or net and define P_0 accordingly. Independent evidence is needed to separate drag from efficiency; otherwise fit the effective coefficient K/η and avoid attributing changes uniquely to either mechanism. The rate and length domain should be restricted to combinations supported by measurement, because extrapolated optima may reflect the fitted surface rather than attainable technique.

4.2. Independent Validation

Separate complete trials, sessions, or swimmers into calibration and validation sets according to the intended application. Splitting individual cycles from the same trial between sets can exaggerate performance. Compare predicted and measured energy cost using absolute error, systematic bias, and uncertainty intervals across speed and rate conditions. Benchmark the fitted model against a constant efficiency model and a simpler empirical energy model. When L is calculated as v/r , reproducing speed with rL should not be presented as an independent prediction.

A prospective test would compare a selected feasible combination with the swimmer's usual technique at the same speed and distance, using randomized or counterbalanced conditions. The principal outcome would be a change in measured energy cost; a speed focused comparison would instead hold the energy allowance or effort protocol consistent. Conclusions about superior performance, fatigue reduction, or long term training effects require such observations. Model calibration and a simulation alone cannot establish them.

4.3. Training and Event Selection

For a steady endurance session, the model can compare nearby feasible combinations at the prescribed speed and identify those with lower predicted cost. A sprint objective places greater weight on attainable speed and event specific power limits. Rate and length should be adjusted together, and the recommendation should specify a tolerable range rather than an unsupported exact target. Each stroke style requires separate calibration because its coordination and energetic characteristics differ.

Coaches could use the minimum cost envelope to compare a technical change with an established baseline. A recommendation becomes actionable when it gives the target speed, feasible rate range, corresponding length range, expected energy cost, and uncertainty. Wearable feedback may support monitoring after sensor accuracy is checked. Adaptation over weeks or months requires repeated measurements and cannot be predicted from a single steady state fit.

5. Discussion and Limitations

The analysis explains two distinct optimization tasks. First, the event requirement determines an acceptable balance between speed and expenditure. Second, technique dependent measurements identify efficient rate and length combinations within that balance. This agrees with prior evidence that stroke mechanics and energetic cost interact, while avoiding a universal optimum inferred from group averages. The analytical

benchmark also shows why lower metabolic power and lower total energy over a distance are different objectives.

The main limitations concern the model's steady, cycle averaged representation. Actual swimming has fluctuating velocity and force, and averaging a cubic power relation can underestimate the contribution of fluctuations. K may change with body position, speed, and stroke coordination. Effective efficiency may also vary with intensity and fatigue. Starts, turns, and underwater phases alter event time without following the same stroke identity. Temperature, equipment, and pool conditions can limit transfer between sessions.

A positive P_0 creates an interior minimum in the benchmark, but net expenditure definitions or speed dependent baseline costs may change that result. A time dependent event model would require rate, length, fatigue, and energy supply to evolve during the swim; the steady model cannot infer when their peaks should occur. Separate calibrations are needed before quantitative comparisons across strokes or athletes. The present study supplies analytical results and a testable protocol, without establishing empirical parameter values, algorithm performance, or physiological improvement.

6. Conclusion

Mathematical modeling can identify favorable stroke rate and stroke length combinations by coupling their kinematic relationship with an energy cost model and measured feasibility limits. For a fixed event distance, minimizing energy cost per meter is equivalent to minimizing total modeled energy, while maximizing speed defines a competing objective. The constant efficiency model has a unique minimum cost speed but leaves the stroke combination undetermined. Technique dependent efficiency, or a calibrated effective energy coefficient, is necessary to distinguish pairs at the same speed.

The practical solution is a feasible Pareto set followed by a selection based on a required speed, target time, or energy allowance. Compared with speed only or descriptive stroke analyses, this formulation states the energy objective explicitly and exposes the measurements needed for personalization. Future work should calibrate the model with repeated physiological trials, validate selected combinations independently, and extend the framework to fatigue and event phases. Individual training benefits must be established through those tests.

References

1. P. Zamparo, M. Cortesi, and G. Gatta, "The energy cost of swimming and its determinants," *European Journal of Applied Physiology*, vol. 120, no. 1, pp. 41–66, 2020, doi: 10.1007/s00421-019-04270-y.
2. T. M. Barbosa, R. J. Fernandes, K. L. Keskinen, and J. P. Vilas-Boas, "The influence of stroke mechanics into energy cost of elite swimmers," *European Journal of Applied Physiology*, vol. 103, no. 2, pp. 139–149, 2008, doi: 10.1007/s00421-008-0676-z.
3. T. M. Barbosa, R. Fernandes, K. L. Keskinen, P. Colaço, C. Cardoso, J. Silva, and J. P. Vilas-Boas, "Evaluation of the energy expenditure in competitive swimming strokes," *International Journal of Sports Medicine*, vol. 27, no. 11, pp. 894–899, 2006, doi: 10.1055/s-2006-923776.
4. S. M. Jameel, "Nonlinear interaction between stroke length and stroke rate in elite swimming," *Frontiers in Sports and Active Living*, vol. 8, p. 1830689, 2026, doi: 10.3389/fspor.2026.1830689.
5. A. Ding, "Modeling for optimizing the patterns of swimming strokes," *The National High School Journal of Science, NHSJS Reports*, pp. 1–13, 2025.
6. K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002, doi: 10.1109/4235.996017.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of Publisher and/or the editor(s). Publisher and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.