Diving into a simple anguilliform swimmer's sensitivity

Nicholas A. Battista
Dept. of Mathematics and Statistics, 2000 Pennington Road, The College of New Jersey, Ewing Township, NJ 08628, USA
battistn@tcnj.edu

July 8, 2021

Abstract

Computational models of aquatic locomotion range from modest individual simple swimmers in $2D$ to sophisticated $3D$ multi-swimmer models that attempt to parse collective behavioral dynamics. Each of these models contain a multitude of model input parameters to which its outputs are inherently dependent, i.e., various performance metrics. In this work, the swimming performance's sensitivity to parameters is investigated for an idealized, simple anguilliform swimming model in $2D$. The swimmer considered here propagates forward by dynamically varying its body curvature, similar to motion of a *C. elegans*. The parameter sensitivities were explored with respect to the fluid scale (Reynolds number), stroke (undulation) frequency, as well as a kinematic parameter controlling the velocity and acceleration of each upstroke and downstroke. The input Reynolds number and stroke frequencies sampled were from [450,2200] and [1,3] Hz, respectively. In total, 5000 fluid-structure interaction simulations were performed, each with a unique parameter combination selected via a Sobol sequence, in order to conduct global sensitivity analysis. Results indicate that the swimmer's performance is most sensitive to variations in its stroke frequency. Trends in swimming performance were discovered by projecting the performance data onto particular $2D$ subspaces. Pareto-like optimal fronts were identified. This work is a natural extension of the parameter explorations of the same model from [1].

Keywords: aquatic locomotion; anguilliform motion; fluid-structure interaction; immersed boundary method; sensitivity analysis; Sobol sensitivity

1 Introduction

Mathematical models inherently depend on parameter estimation and selection. In the study of disease [2] or drug epidemics [3], transmission and recovery rates are estimated based on available epidemiological data. Similarly, scientists investigating cellular processes, immunology, or oncology may attempt to fine tune model parameters based off experimental data available to investigate novel therapies [4, 5, 6]. Parameter explorations have provided insight into a number of organismal systems, including odor-capture by crabs [7], soaring strategies of birds [8], or tubular heart pumping [9]. A chief focus in each of these studies involved not only explorations within a particular input parameter space, but the *sensitivity* of the model's output to the model's input parameters. A system is *sensitive* to an input parameter, if slight variations in the parameter result in significant changes in the system's output. Global sensitivity analyses quantify the impact of parameter uncertainty (inputs) on the overall model prediction uncertainty (outputs) in a holistic fashion [10]. These analyses attempt to determine which parameters, or their combinations, are most sensitive for the system output across a given parameter space.

In organismal biology contexts, uncertain parameters, or fluctuations in parameters, could be viewed slightly differently than those in some population epidemiology or cellular contexts. Rather than only focusing on fine tuning model parameters to match experimental data, variations in parameters could be interpreted as biodiversity, even extending towards evolutionary contexts by exploring robustness of a performance metric to parameter sensitivities [11, 12, 9]. In particular, biodiversity has highlighted elegant "many-to-one mapping" solutions, where different morphologies (form) can lead to similar performance (function) [13, 14]. However, not all combinations of traits produce similar performance. To that extent, slight variations may to lead non-linear consequences in functional performance [15]. Upon exploring regions of higher performance within performance landscapes, global sensitivity analysis can determine what variation (or combinations thereof) induce the greatest changes in performance. Thus one is able to rank the importance of specific traits for maintaining a particular performance level (within a specific region on the landscape) while assessing its robustness. Furthermore, quantifying such sensitivities could further provide insight into the evolution rate of a given mechanical system [16, 17].

In the spirit of interdisciplinary collaboration from the 2020 SICB Symposium *Melding Modeling and Morphology: integrating approaches to understand the evolution of form and function*, this paper highlights a formal mathematical approach for assessing the sensitivity of a model's output due its input parameters. This paper explores the global sensitivity analysis of a mathematical model of a biological system - an idealized swimmer that resembles an aquatic nematode. The selection of a fluid-structure interaction model was two fold: (1) to illustrate that such global sensitivity analysis can be performed on

a complex model and (2) to highlight the challenges (necessary computational resources) to perform such an analysis on a computational fluid dynamics model. The swimmer itself, in resembling a nematode, propagates itself forward through its fluid environment by varying its body curvature, i.e., bending [18, 19, 20, 21], see Figure 1.

Previous experimental work directly focused on *C. elegans* has highlighted that variations in bending frequency impact locomotion performance at low Reynolds numbers (Re < 1) [21]. The Reynolds number is a dimensionless number that is given as a combination of four parameters: two parameters describing the physical properties of the fluid - its density and viscosity, ρ and μ, respectively, and two system parameters - a characteristic length scale, L and characteristic velocity scale, V. Mathematically, it is given as

$$Re = \frac{\rho L V}{\mu}. \tag{1}$$

Low Re situations arise when viscous forces are much greater than inertial forces. Many computational models of nematodes have only considered the low Reynolds limit, i.e., the Stokes flow Regime (Re $= 0$) [22, 23, 24]. However, anguilliform swimming gaits are observed in other organisms, such as fly larva, leeches, eels, or lamprey [25, 26, 27, 28]. These organisms are much larger and live at intermediate Re (Re $\sim 10s$ or $100s$) or higher Re (Re $\gtrsim 1000$) ranges. A recent computational study sought out to explore the effectiveness of this anguilliform gait across a variety of fluid scales, f, and intrinsic curvature dynamics [1].

This work continues that of [1], which explored the same idealized anguilliform swimmer's performance through thorough extensive parameter subspace investigations. The parameter space considered was composed a fluid scaling parameter, given by an input Reynolds number (Re_{in}), the stroke (undulation) frequency (f), and a kinematic control parameter (p). The kinematic parameter controls the velocity and acceleration of the swimmer's body curvature as it undulates between a concave up and concave down state. That is, it controls each stroke's acceleration from rest in its current curvature (concavity) state to its maximal velocity and back to rest in its next curvature (concavity) state. Previous experimental work suggests that a nematode's undulatory amplitude does not change when places in increasing viscosity environments [29], i.e., decreasing Re settings; however, this work will show that does not appear be the case as Re increases.

While [1] performed extensive parameter explorations using the same anguilliform swimmer model, specific 2D slices through the overall 3D parameter space were chosen in which to investigate swimming performance. The 2D slices were chosen in a rectangular fashion. For example, as the entire parameter space considered was $\mathrm{Re}_{in} \times f \times p = [0.3, 4500] \times [1, 2.5] \times [0.075, 0.425]$, one 2D slice could be all (Re, f) combinations for $p = 0.25$. The range of Re_{in} given above is based off the definition of input Reynolds number provided in Eq. 2. Although this systematic approach allows one to identify a parameter

subspace that results in desired swimming performance, i.e., determine threshold forward swimming speeds or cost of transports, it has a few significant drawbacks. First, it is unclear whether the resulting performance data could be interpolated with any accuracy between parallel 2D slices due to nonlinear relationships that may exist. Second, it is restrictive in that it does not allow one to perform a formal global sensitivity analyses on 3D parameter space, due to the manner in which the 3D space is sampled. Thus you cannot determine which parameter (or combinations thereof) most significantly affect the model's output across the input parameter ranges considered.

The anguilliform swimmer model explored here has substantial forward swimming speeds when $\mathrm{Re}_{in} \gtrsim 450$ [1]. Therefore, with the goal of performing global sensitivity analyses, the selected parameter subspace to perform the analyses was $\mathrm{Re} \in [450, 2200]$, $f \in [1, 3]$ Hz, and $p \in [0.05, 0.45]$. In the remainder of this paper, Sobol sensitivity analysis [30] is performed on a variety of swimming performance metrics over the parameter ranges above. By comparing multiple output metrics sensitivity to parameters it is possible to determine how different performance metrics may have differing parameter sensitivities, both in determining which parameter it is most sensitive to as well as its relative importance (magnitude). Discernible patterns are also uncovered in swimming performance by projecting the $3D$ parameter space sampled onto specific $2D$ subspaces. Furthermore, Pareto-like optimal fronts in swimming performance are identified across various combinations of the input parameters [31, 32, 33]. Lastly, the parameter sensitivity suggested by these Pareto-like fronts is consistent with that of the formal Sobol sensitivity analysis. Both Sobol sensitivity analysis and Pareto-like optimal fronts will be described in further detail in Sections 2 and 3, respectively.

2 Methods

In this work, I continued the investigation of the anguilliform swimming model first presented in [34] and further explored in [35, 1]. To propagate forward, i.e., swim, the swimmer changes the curvature along its $1D$ body, a motion that resembles an aquatic nematode [36, 20], see Figure 1. Its total stroke (undulation) cycle is compromised of both an upstroke and downstroke, which end when the swimmer's body is either concave down and concave up, respectively. As the swimmer propagates forward, it leaves behind a vortex wake. This observation suggests that hydrodynamic force generation by dynamically changing curvature is sufficient for locomotion [37]. The swimmer given in Figure 1 corresponds to the case of $(\mathrm{Re}_{in}, f, p) = (1011.75, 1.6875, 0.4125)$. As these parameters are varied, the swimmer's performance may substantially change, as can be seen in Figures 6, 7, and 8 (and as shown in [1]).

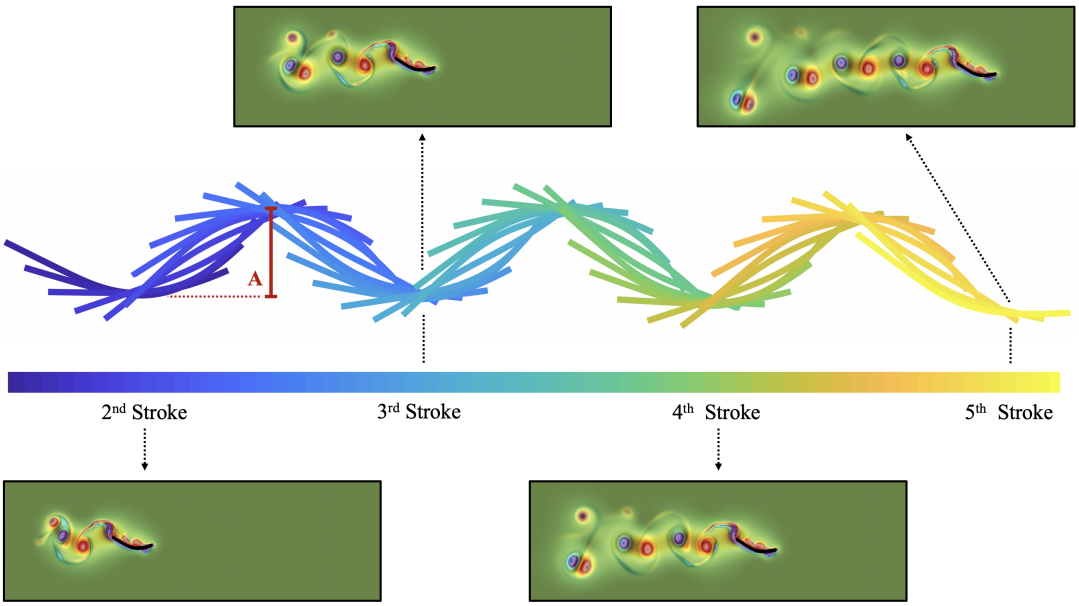

Figure 1: An illustration of how the swimmer changes its preferred curvature to propagate forward during the 2^{nd} through 5^{th} stroke cycles. Snapshots of the swimmer and the associated fluid vorticity (colormap) are given at the beginning of each successive stroke cycle. The definition of the peak-to-peak amplitude, A, is given. This amplitude is an output of the model.

I used the open-source fluid-structure interaction software, *IB2d* [38, 39, 34], based on the immersed boundary method (IB) [40] to perform the computational fluid dynamics simulations. This framework allows one to study this particular anguilliform swimming mode in an incompressible, viscous fluid, thus allowing for studies across numerous fluid scales i.e., Reynolds numbers. This anguilliform model is one of the built-in examples contained in the software offered to the scientific community. It can be found at `github.com/nickabattista/IB2d` in the sub-directory:

`IB2d/matIB2d/Examples/Examples_Education/Interpolation/Swimmer`.

A background introduction to IB and the swimmer's implementation are provided in the Supplemental Materials; greater detail regarding the mathematical framework, formulation, and implementation can be found in [34, 35, 1]. The computational grid, temporal, and material property parameters are identical to those used in [1]; they are given in Table S1 in the Supplemental Materials. The model parameters varied in this study are an *input* Reynolds Number, Re_{in}, stroke (undulation) frequency, f, and a kinematic control parameter, p. The kinematic parameter, p, helps govern the kinematic profile of each stroke. Although, each upstroke (or downstroke) always performs an undulatory motion, varying p is akin to changing the acceleration and velocity through stroke pattern itself. That is, it controls how quickly each stroke accelerates from rest in the starting curvature (concavity) state, to its maximal velocity, and back to rest in the other curvature (concavity) state.

In other words, in this study I varied one parameter that encompasses size variations (the input Reynolds number, Re_{in}), another that details how often swimming movement patterns are performed (the frequency, f), and a third that changes the intrinsic kinematics within each stroke (kinematic parameter, p). The model parameter ranges were slightly different than those in [1]. Here a $3D$ parameter subspace was selected based on the previous explorations in which there was significant forward swimming speed: $\mathrm{Re}_{in} \times f \times p = [450, 2200] \times [1, 3]$ Hz $\times [0.05, 0.45]$. The range of f was selected to center about the undulation frequency that *C. elegans* display while swimming in water, roughly 2 Hz [21, 41]. Although, in wet granular media *C. elegans* swim with frequencies around $\sim$ 1 Hz as well as exhibit faster swimming behavior [36, 42]. On the other hand, *C. elegans* undergoing dietary restrictions undulate at higher frequencies of $\sim$ 3 Hz while also swimming faster than the control [21]. Other anguilliform swimmers such as eels and lampreys swim at higher frequencies. For example, eels have been observed to steadily swim with an undulation frequency of $\sim$ 3 Hz [43]. Thus, the range of frequencies was chosen to be between [1,3] Hz. The values of the kinematic parameter p were chosen to span the range in which lead to a symmetric interpolation function, see Figure S1 in the Supplemental Materials.

To achieve the desired input Reynolds number, Re_{in}, first a tuple was selected, i.e., $\mathrm{Re}_{in}, f, p)$, followed by finding the appropriate dynamic viscosity, μ, to give the appropriate Re_{in}, i.e.,

$$\mathrm{Re}_{in} = \frac{\rho L(fL)}{\mu}, \tag{2}$$

where L is the swimmer's bodylength (characteristic length) and the product fL gives an input frequency-based velocity scale. In anguilliform studies it is common to use the product of f and a peak-to-peak undulation amplitude, A, as the characteristic velocity when calculating Re. However, since undulation amplitude is an output of the model and cannot be known *a priori*, I elected to use a characteristic velocity seen in fish literature as an input velocity scale, fL [44, 45, 46]. Illustrations of the peak-to-peak amplitude, A are given in Figures 1 and 2. The output Reynolds number is defined as $\mathrm{Re}_{out} = \rho \cdot L \cdot fA/\mu$, whose frequency based velocity scale is given as fA. Comparisons of the input and output Reynolds numbers, Re_{in} and Re_{out}, respectfully, are provided in the Supplemental Materials.

A variety of swimming performance metrics were computed in each simulation. The motivation behind calculating multiple metrics was to identify whether different performance metrics were most sensitive to different input parameters. Rather than report the dimensional swimming speed, V_{dim}, for each simulation, swimming speeds are provided in dimensionless form, given by the inverse of the Strouhal number. The Strouhal number is defined as

$$\mathrm{St} = \frac{fA}{V_{dim}}. \tag{3}$$

where A is the peak-to-peak undulation amplitude, see Figures 1 and 2. Note that A is an output of the model itself, whose changing values due to variations input parameter changes was also analyzed.

The Strouhal number has been used previously to assess efficient swimming [47, 48]. It compares the wavespeed to the average speed of the swimmer. Three other performance metrics were computed: a power-based cost of transport [49, 28], COT, a distance effectiveness ratio, d_{eff}, and an angular trajectory metric, θ. The COT provides the energetic cost per unit speed of the swimmer, d_{eff} compares the forward distance swam to the total linear distance swam overall, and θ indicates how horizontal the overall swimming movement stayed. All of these quantities are provided in non-dimensional units, i.e.,

$$COT = \frac{1}{N}\ \frac{1}{V_{dim}}\ \frac{1}{\rho(fA)^2L^2}\sum_{j=1}^{N}|F_j||U_j| \tag{4}$$

$$d_{eff} = \frac{D_S}{D_{Tot}} \tag{5}$$

$$\theta = \frac{1}{N}\sum_{j=1}^{N}\tan^{-1}\left(\frac{Y_H(t) - Y_{M0}}{X_H(t) - x_{M0}}\right) \tag{6}$$

where D_S and D_{Tot} are the forward distance swam (horizontal distance) and total linear distance moved by the swimmer during a specific period of time encompassing N time-points. F_j and U_{r_j} are the applied force and tangential body velocity of the swimmer, at the j^{th} time step, respectively. The applied force was computed as the force perpendicular to the original direction of motion of the swimmer. $X_H(t)$ and $Y_H(t)$ are the time-dependent Lagrangian position of the swimmer's head and (X_{M0}, Y_{M0}) is the initial position of the last point along the straight section of the swimmer's body (see Figure 2). The non-dimensional COT is similar the energy-consumption coefficient in [49].

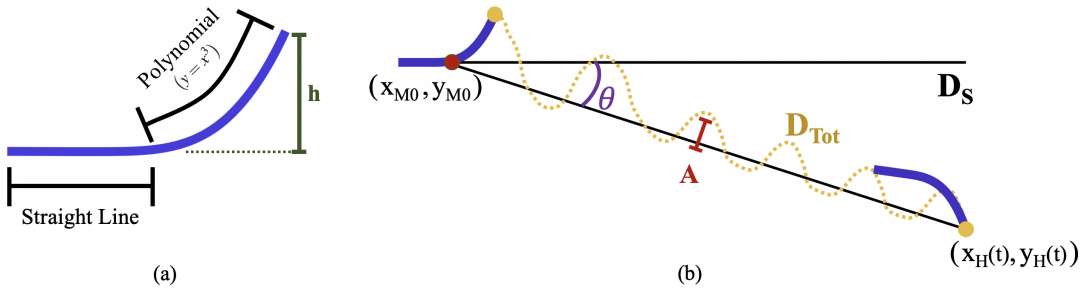

Figure 2: (a) The swimmer's geometry is composed of a straight line segment and a curved portion, given by a cubic polynomial. The swimmer's height, h, is the height of the curved portion off the straight segment. It remained uniform across all simulations. (b) The definitions for the angular trajectory off the horizontal (θ), as measured off the swimmer's head and a fixed point along the horizontal (last point along the straight portion of the swimmer's body), the horizontal distance swam (D_S), and the total linear distance moved by the swimmer (D_{Tot}). The swimmer in (b) corresponds to the case of $(Re_{in}, f, p) = (270, 1.25, 0.25)$. Figure adapted from [1].

To assess this swimming model's overall sensitivity to its parameters, I sampled the overall parameter space using Saltelli's extension of the Sobol sequence [10, 50]. A total of 5000 combinations were selected, $\{(\mathrm{Re}_{in}, f, p)_j\}_{j=1}^{5}000$, in which to simulate the model. Once the simulations were performed, they were analyzed to produce a peak-to-peak undulation amplitude A and a Strouhal number, St, and thereby a time-averaged forward swimming speed (1/St), d_{eff}, COT, and θ. Once the metrics were computed, Sobol sensitivity analysis could then be conducted. Sobol sensitivity is a variance-based sensitivity analysis that can provide *global* sensitivity to parameters, rather than only local sensitivity [51, 52]. By quantifying global sensitivity, one can determine which parameter, when varied within a particular range, results in the most significant changes in the model's output, even with respect to other parameters being varied [50, 52]. More detail regarding global vs local sensitivity analyses is provided in the Supplemental Materials.

The workflow for the entire simulation and sensitivity analysis process is illustrated in Figure 3.

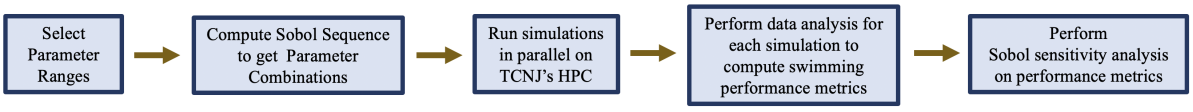

Figure 3: Workflow of the entire simulation and global sensitivity process

3 Results

A Sobol sequence was generated using $N = 1000$ and $d = 3$ (the dimension of the parameter space) using Saltelli's extension of the Sobol sequence [10, 50] for a total of $N(d + 2) = 5000$ different parameter combinations to be simulated in order to find the

Sobol sensitivity indices of each performance metric. The simulations were performed on The College of New Jersey's high performance computing cluster [53]. As each simulation took approximately $\sim$ 24 hours to run, the entire effort of this study took required approximately a total of 120,000 computational hours.

Once the swimming performance metrics were computed, the Sobol sensitivity indices were found [30, 50]. Figure 4 provides the indices for both the first-order and total-order parameter interactions. Over the input parameter ranges considered, the performance metrics were most sensitive to variations in stroke (undulation) frequency, f. Moreover, higher-order interactions appear to exist, as the first-order and total-order indices are not equivalent. However, both indicate that system is most sensitive to f. This remained true for the dimensional swimming speed and cost of transport data, see Figure S7 in the Supplemental Materials. The degree of sensitivity to each parameter did vary among the output metrics. While this study only varied Re_{in}, f, and p, it is possible that the performance metrics could be more sensitive to other parameters not explored here, such as variations in the swimmer's morphology [54, 55, 56, 57].

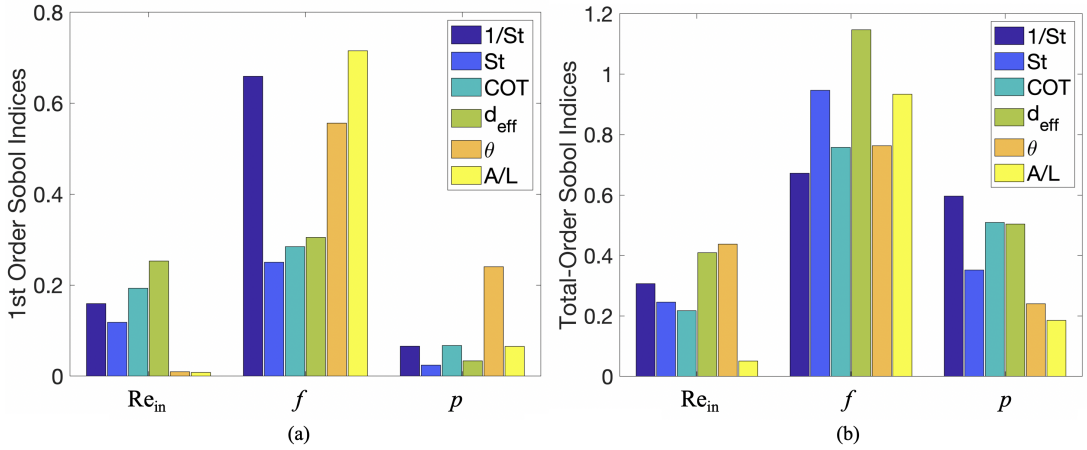

Figure 4: (a) First-order and (b) Total-order Sobol indices of the three varied parameters Re_{in}, f, and p for 5 swimming performance metrics: non-dimensional swimming speed (1/St), Strouhal number (St), non-dimensional cost of transport (COT), a distance effectiveness ratio (d_{eff}), and the average angular trajectory from horizontal (θ).

With only the Sobol indices in hand, it is unclear how the performance metrics are sensitive to f, e.g., if f is increased, it is unclear whether swimming speeds would increase or decrease. More so, it isn't clear how to determine a finer parameter subspace where higher swimming performance resides directly from the Sobol indices alone. If another parameter subspace were tested, as the input ranges would be different in that subspace, it would inherently change the sensitivity indices, and in possibly substantial ways as well.

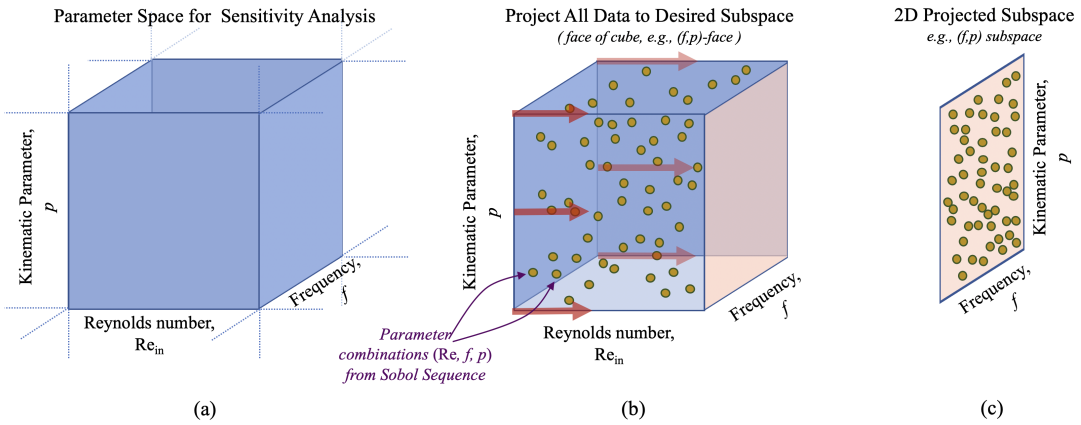

Figure 5: (a) The overall $3D$ parameter space being sampled by Sobol sequences. (b) Illustration of sampling points via Sobol sequences being projected into a $2D$ subspace, here the (f, p) subspace (c) Visualization of where data is projected from a higher dimension space into a $2D$ subspace.

From the Sobol sensitivity analysis performed in this paper, $2D$ subspace colormaps are not able to be produced in the same manner as in [1]. Since the $3D$ parameter space was sampled using Sobol sequences, coplanar $2D$ subspaces, like those in [1], are not highly resolved. The sampled points simply do not heavily lie on any of the same 2D subspaces. Therefore, the sampled data here was projected onto three distinct subspaces: (Re_{in}, f), (Re_{in}, p), and (f, p), as in the process illustrated by Figures 5b-c. Having performed 5000 simulations, these projected subspaces appear "filled" and can be parsed for general performance trends.

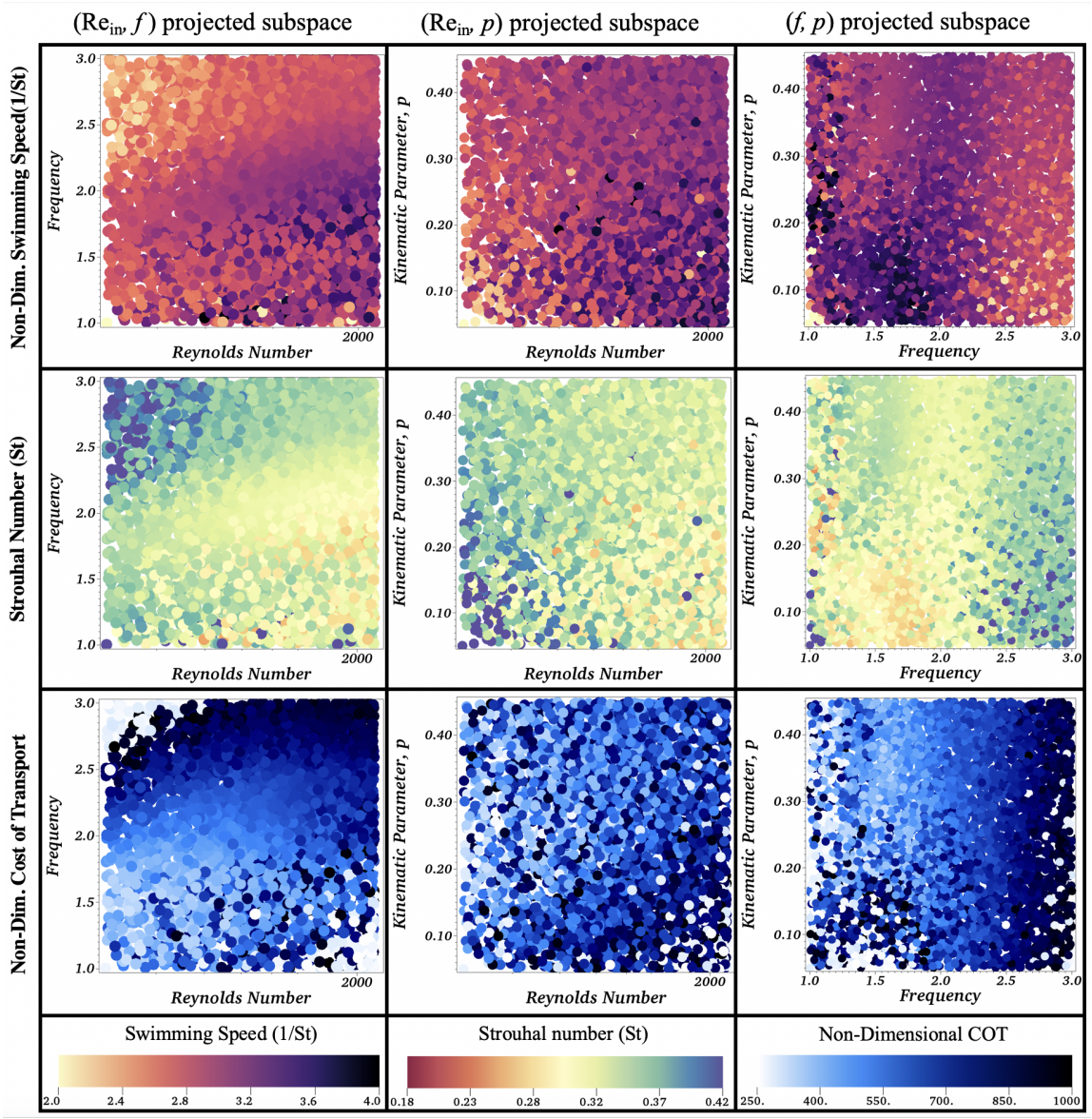

Figure 6: Colormaps corresponding to the non-dimensional forward swimming speeds (1/St), Strouhal numbers (St), and cost of transports (COT) for all the data sampled from Sobol sequences when projected onto either the (Re_{in}, f), (Re_{in}, p), or (f, p) subspaces.

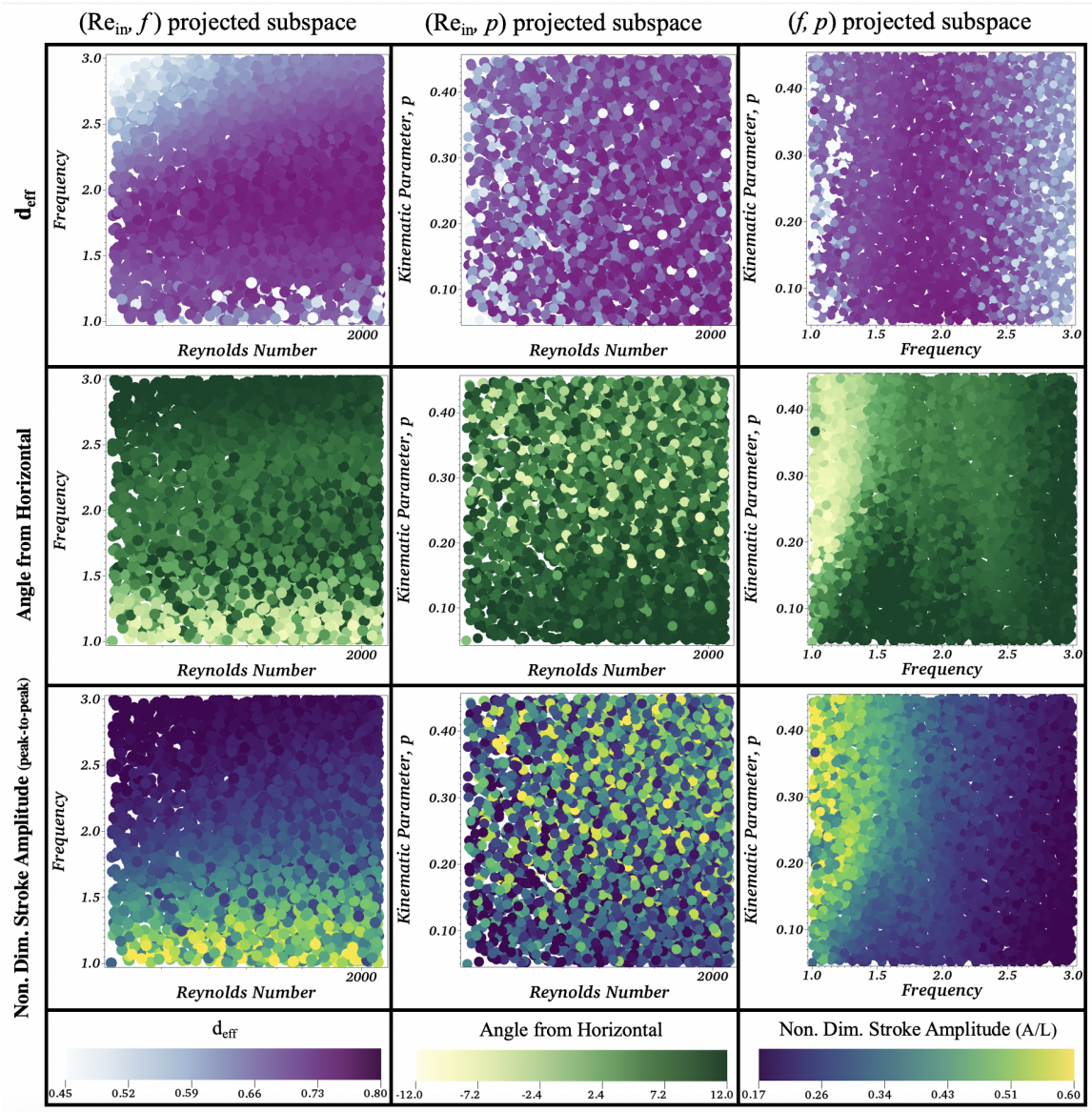

Figure 7: Colormaps corresponding to d_{eff}, θ, and A/BL for all the data sampled from Sobol sequences when projected onto either the (Re_{in}, f), (Re_{in}, p), or (f, p) subspaces.

Figures 6 and 7 provide the projected data onto each respective subspace: (Re_{in}, f), (Re_{in}, p), or (f, p). From a quick glance, qualitative patterns emerge within regions of the (Re_{in}, f) and (f, p) subspaces for all performance metrics. The data in the (Re_{in}, p) subspace looks more like random noise. For example, the COT, d_{eff}, θ, or A panels corresponding to the (Re_{in}, p) subspace do not show much, if any, immediate discernible qualitative pattern(s). However, it does appear that lower f corresponds to more noisy data within almost every subspace. Furthermore, the subspaces in which patterns arose had f as one of its component axes. The patterns and trends that emerged within the performance data aligned with the f-axis in the (Re_{in}, f) and (f, p) subspaces. While slight variations do occur in the other component's direction, either Re_{in} or p, in these two subspaces, respectively, the most substantial change appears to be in the direction of

f. This agrees with the Sobol sensitivity analysis above, see Figure 4. Although, as these plots suggest, f is not the only significant parameter contributing to the model output.

The patterns that emerged in Figures 6 and 7 show similar general trends as those seen in [1]. For example, the high swimming speeds arose from high Re_{in} and lower f. Interestingly, the minimal value regions for COT fell for either high Re_{in} and low f or low Re_{in} and high f. Although, low Re_{in} and high f correspond to a region of higher St, where St leaves the biological range of $0.2 < \mathrm{St} < 0.4$ [47]. High values of d_{eff} in the (Re_{in}, f) subspace also corresponded to lower values in COT, moderately high swimming speeds, and emergent peak-to-peak stroke amplitudes of approximately 1/3 of their bodylength. These occurred around $f \sim 2$ Hz. *C. elegans* have been observed swimming at roughly $f \sim 2$ Hz with peak-to-peak amplitudes that were roughly 25% of their bodylength, but at $\mathrm{Re} \approx 0.5$ [42]. It appeared overall that low frequency swimming resulted in higher peak-to-peak stroke amplitudes and downward swimming trajectories. Moreover, the performance metrics show non-linear dependence on f and p, as inferred from the (f, p) panels. The observable trends in the (Re_{in}, f) and (Re_{in}, p) projected subspaces are consistent with those involving the output Reynolds number, Re_{out}, i.e., (Re_{out}, f) and (Re_{out}, p). This data is provided in Figures S5 and S6 in the Supplemental Materials.

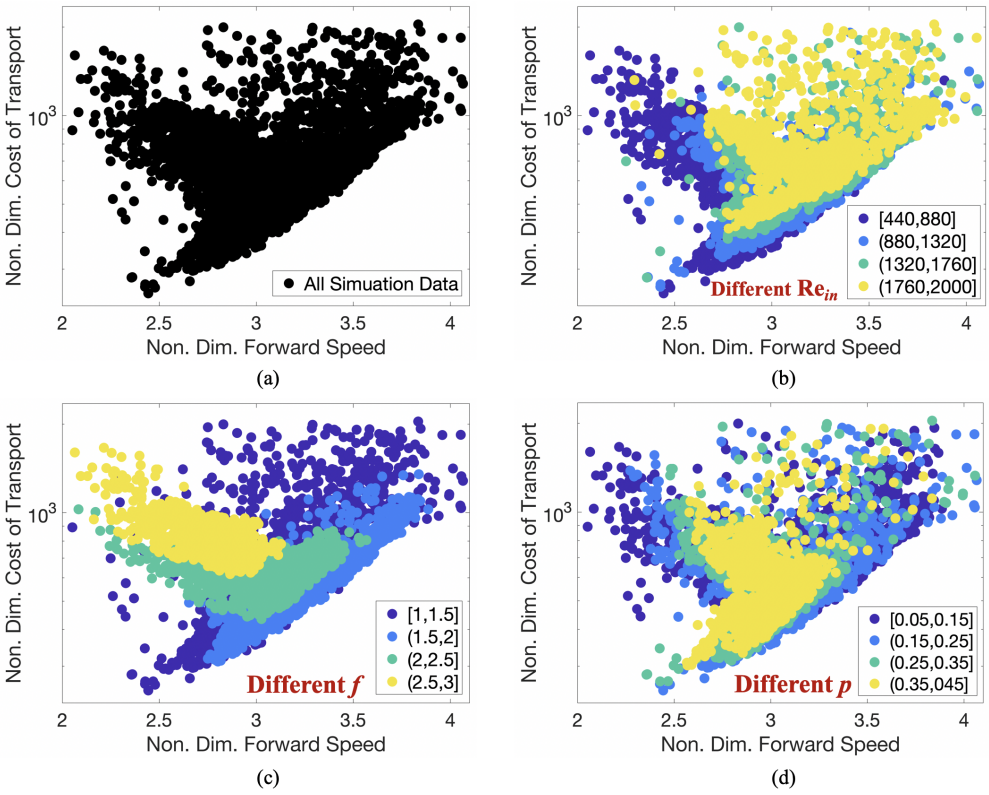

Figure 8: The non-dimensional COT and forward swimming speeds (1/St) plotted against each other for (a) every simulation performed, as well as colors indicating different ranges of the input parameters: (b) different Re_{in} (c) different f, and (d) different p.

Lastly, a Pareto-like optimal front is identified when plotting the non-dimensional cost of transport (COT) against the non-dimensional swimming speed (1/St) [31, 32, 58, 33, 59], see Figure 8a. This is referred to as *Pareto-like* simply because minimal COT is desired for maximal 1/St, thus both metrics are not being maximized as is standard in traditional Pareto optimization. The COT appears to increase exponentially (note the logarithmic axis) with increasing swimming speed. Figures 8b-d, depict where different input parameter ranges lie within the performance landscape across all simulations performed. (b) and (d) illustrate that for a given Re_{in} or p within the input parameter space that combinations of the other two parameters, either (f, p) or $\mathrm{Re}_{in}, f)$, respectively, could result in a swimmer whose performance could almost be anywhere in the entire landscape. For example, given $\mathrm{Re}_{in} = 500$, one could create a swimmer whose swimming speed was anywhere between $[2, 4]$ by selecting the appropriate f and p combination. On the other hand, distinct clusters emerge in (c) for different f. As f increases, swimming speed generally decreases while COT increases. Undulation frequencies within the range of $[1, 2]$ extend the length of the Pareto-like front. However, frequencies within the $[1, 1.5]$ range span the entire performance space, while frequencies within $[1.5, 2]$ are more restricted towards the Pareto-like front. This further supports that the model output is most sensitive to undulation frequency, f, within the aforementioned parameter subspace considered in this study. Where different parameter combination swimmers lie within the performance space is given in Figure 9. This figure also gives some swimmer's position and vortex wakes after their 5^{th} stroke cycle. There do not appear to be any qualitative patterns among the vortex wakes that suggest higher or lower cost of transport. This performance space data is also given in *dimensional* form in Figures S9 and S10 in the Supplemental Materials. A similar trend is observed in the dimensional data, in which distinct clusters form for different f, as in Figure 8c, although higher f corresponds to higher dimensional cost of transport. Moreover, there appears to be more clustering for different Re_{out} than in Figure 8b, where higher Re_{out} appear to correspond to faster swimming and slightly lower cost of transport in dimensional units.

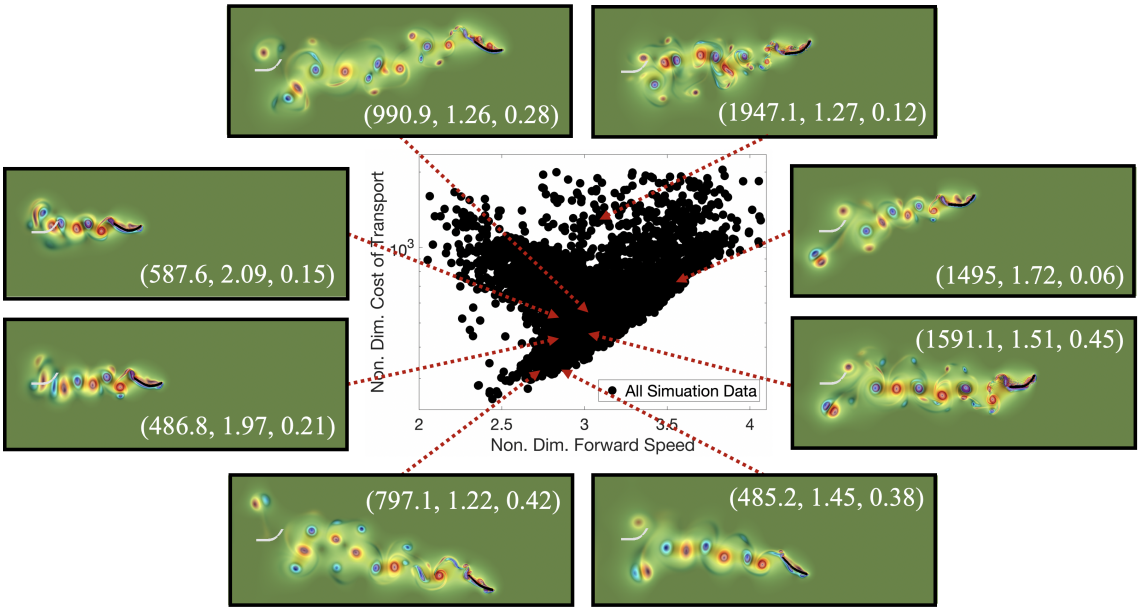

Figure 9: Numerous swimmer's vortex wake and position after their 5^{th} full stroke cycle and indicating where they fall onto the non-dimensional cost of transport vs. swimming speed performance space. The colormap illustrates vorticity and the greyed out swimmer shows the starting position of the swimmer in each case. Different parameter combinations (Re_{in}, f, p) lead to different swimming behavior, as indicated by its placement within the performance space and vortex wake left behind.

4 Discussion

Previous experimental and computational studies have highlighted the importance of size (scale) [60, 61], stroke/undulation frequency [62, 63, 64], bending curvature and stiffness [61, 65], and intrinsic kinematics [66, 67] on swimming performance. In this work, 3 input parameters were varied: the scale (Re_{in}), stoke (undulation) frequency (f), and a kinematic control parameter (p) for an idealized, simple anguilliform fluid-structure interaction model. The input parameter space was composed entirely of a region in which there is substantial forward swimming performance [1]. Rather than focusing on the net effect that variations of a single parameter has on swimming performance, the model's global sensitivity to its input parameters was uncovered using Sobol sensitivity analysis. The sensitivity analysis indicated that swimming performance was most sensitive to variations in frequency (Figures 4 and S7).

While it is already well-known that frequency affects forward speed and performance [62, 63, 64], this study indicates that varying the frequency (within the parameter space considered) will more significantly affect the resulting model's performance than changes in the other two input parameters and combinations thereof. Although, all of the performance metrics considered were most sensitive to frequency, the degree of sensitivity to each parameter varied between different metrics. For example, after frequency, the

total-order Sobol indices in Figure 4b illustrates that 1/St, St, COT, and d_{eff} appear more sensitive to p than Re_{out}, while θ is more sensitive to Re_{out} than p. Therefore when discussing a model's sensitivity to a particular parameter, it can only be with respect to a specific performance metric. Moreover, the Sobol index rankings after f among first-order indices are different for every output metric except θ, see Figure 4a. Since the first-order and total-order indices are not equivalent, it suggests that higher-order interactions among parameters are important in the model output.

The morphology and kinematics of the swimmer studied here resemble that of a *C. elegans* [36]. This model has previously predicted swimming speeds ($1/\mathrm{St} \sim 0.4 - 0.7$) in agreement at lower biologically relevant Re ($\mathrm{Re}_{out} \sim 0.5$) and frequencies ($f \sim 1.8 - 2.2$ Hz) [1] with the organism itself ($1/\mathrm{St} \sim 0.53 - 0.75$) [42]. However, in this work the fluid scale investigated was approximately two or more orders of magnitude greater than that of a *C. elegans* [42]. At higher Re, the emergent peak-to-peak stroke amplitude no longer appeared conserved, in contrast to experimental studies of *C. elegans* [29]. The choice of Re in this work was deliberate to investigate this simple anguilliform mode over a range of Re that boasted higher forward swimming speeds [1]. This selection of input Re (and resulting output Re) across the intermediate Re regime is a particularly interesting fluid scale to explore due to the importance and intricate balance of both viscous and inertial forces [68, 69, 70]. Many of the seminal anguilliform studies have assumed either low Re [71, 18] or high Re [72] settings. Much of the work since has also focused on one these two regimes, although a few studies have focused their efforts on intermediate Re [69, 73].

The C-start escape mechanism in larval fish falls within the intermediate Re regime ($\mathrm{Re} \sim 100s$) [74, 73]. However, it is not energetically efficient for the swimmer to exclusively use this mode. They only use it to evade a predator before continuing with a less energy intensive anguilliform mode [75]. During larval stages, some insects may use an anguilliform mode at $\mathrm{Re} \sim 100s$, such as Ceratopogonid larva, which have been observed swimming at an estimated $\mathrm{Re} \sim 160$ at 2.17 bodylength/s [25]. Other anguilliform swimmers, such as eels or lampreys, swim at higher Re in the 1000s or 10,000s [43, 27, 76]. However, their anguilliform modes are not equivalent to the simple swimmer studied here. A CFD model of lamprey swimming at $\mathrm{Re} \sim 1000s$ and $f = 1$ Hz resulted in swimming speeds between $0.25 - 0.5$ bodylengths/s [76]. The model explored here produced swimmers within the range of $0.7 - 2.3$ bodylengths/s for swimmers operating at $\mathrm{Re}_{out} \sim 100s$ (Figures S8 and S9 in the Supplemental Materials). Anguilliform swimmers like American eels (*Anguilla rostrata*), larval sea lampreys (*Petromyzon marinus*, 5-7 years old), and medicinal leeches (*Hirudo medicinalis*) have been observed to swim at steady state speeds of $0.5 - 2$ bodylengths/s [27, 77], $1.6 - 1.75$ bodylengths/s [75], and $1.8 - 2$ bodylengths/s [26], respectively. These organisms all swim at high Re. Varying their Re across orders of magnitude could substantially change their hydrodynamics, as

suggested by differences in the flow fields of lamprey models [78, 79].

Recently, much emphasis has been placed on designing optimal undulatory swimmers [80, 73, 81, 31, 56, 66, 82, 59]. In general these studies attempted to optimize either shape, kinematics, or both in regards to swimming speed and either Froude efficiency or cost of transport. These ideas could help inform the design of faster, more efficient, or more maneuverable underwater vehicles [83, 84, 85]. The Froude efficiency is the ratio between the useful propulsive power generated to the rate of energy input. Optimal morphologies and kinematics were different depending on whether maximal speed, maximal Froude efficiency, or lowest cost of transport were desired [80, 56, 66, 81, 59]. Optimization studies by Tokić and Yue [81, 59] suggest that swimming speed and cost of transport are the main drivers of evolution, rather than Froude efficiency. However, the studies above assumed high Re settings; only the studies by Gazzola et al. 2012 [73], Van Rees et al. 2013 [56], and Van Rees et al. 2015 [66] explored designs within the intermediate Re space - in particular, one specific Re, Re = 550. None of these studies varied Re across an intermediate regime nor did they vary undulation frequency. The global sensitivity results here indicated that varying frequency within an Re regime of $50 < \mathrm{Re}_{out} < 1220$ will most significantly affect swimming performance. On the other hand, this study did not alter the swimmer's morphology, such as its length, curvature, or thickness, nor stroke dynamics, such as burst and glide modes or asymmetric stroke patterns.

Furthermore, the parameter sensitivity indices may change if other subspaces of the whole (Re_{in}, f, p) space are considered, as suggested by the data [1]. Changing the input parameter space may lead to significant changes in the sensitivity results, e.g., rather than sampling $\mathrm{Re}_{in} \in [450, 2200]$, selecting $\mathrm{Re}_{in} \in [10, 100]$. For example, investigating the performance of the anguilliform swimming mode at different phases of development when Re may be smaller than 100, might suggest a higher degree of sensitivity to Re rather than f, or other parameters entirely. Thus, to perform sensitivity analysis properly, insights from experimental data or other parameter explorations must first be carefully analyzed to make sure the analysis is performed over the appropriate input space [86, 87].

Moreover, projections of the Sobol sampled $3D$ parameter space, $\{(\mathrm{Re}_{in_j}, f_j, p_j)\}_{j=1}^{5000}$, were projected onto two-parameter subspaces, i.e., (Re_{in}, f), (Re_{in}, p), and (f, p). Similar trends within the performance metrics could be recognized through these projections (Figures 6, 7, and S8) and the Pareto-like front identified (Figures 8 and S9 to those observed in [1]. However, there are distinct differences between the two analyses.

First, in [1], the parameter combinations were selectively sampled in a rectangular fashion. Each subspace considered variations in two parameters at a time while holding the third parameter constant, rather than varying all three at once. The latter is what was done in the work here, via Sobol sequence sampling. The $3D$ data here had to be projected from a higher dimensional subspace ($3D$) to a lower dimensional subspace ($2D$) to investigate trends. Second, the projected colormaps here (Figures 6 and 7)

divulged that the performance metrics appear most sensitive to changes in frequency - most discernible qualitative patterns align with changes in f. However, each subspace also contained regions in which looked more like noise. These noisy regions do not suggest a dominate parameter or parameter combinations. Third, the $3D$ Sobol sampling approach allowed for performing a global sensitivity analysis. The broad parameter explorations in [1] did not allow this. From enough $2D$ subspace explorations, naive hypotheses could be formulated, but only in a strict qualitative sense [1]. Fourth, the parameter space sampled in this paper was only realized from the thorough explorations in [1]. That is, a $3D$ subspace (with high swimming performance) was first identified from [1], to which was then sub-sampled for the sensitivity analysis performed here. Hence, in order to perform this analysis, such a region must be identified *a priori*. Thus, parameter explorations are still necessary before performing sensitivity analysis if one is interested in the sensitivity to parameters within regions of desired performance, unless experimental data is readily available.

Both the study presented here and [1] only considered a $3D$ parameter space in which to analyze swimming performance. It took nearly 5000 simulations for the Sobol sensitivity indices to converge when studying a 3D input space. Therefore studying a higher dimensional parameter space (> 3) using this formulation will require exponentially more simulations [88]. Contemporary methods involving polynomial chaos have become popular to reduce the computational burden, i.e., reduce the amount of simulations required for accurate indices [89, 90, 91].

Finally, when using mathematical modeling for biological inquiry, sensitivity analyses can help further knowledge of a biological system. It provides insight into the importance of parameters for specific output metrics. For example, if one wished to decipher which parameter (phenotype) was most important to the success of an active predation strategy, they would have to restrict the parameter space for a given model to a biologically relevant one for that particular organism. However, sensitivity insights alone aren't sufficient to fully understand any possible limitations of a particular predation strategy. Other parameter explorations would complement the sensitivity results to possibly inform why that organism has evolved a particular way, e.g., maybe such a strategy would not work (or be nearly as successful) if it grew to 10x its size (thus increasing the Re by a factor of 10). Both parameter explorations as well as sensitivity analyses can be used in-conjunction with one another to provide greater understanding a biological system. Symbiosis [1].

[1] only until you run out of computing time (environmental factors)

References

[1] Battista, N. A. Swimming through parameter subspaces of a simple anguilliform swimmer. Integr. Comp. Biol. **60(5)**, 1221–1235 (2020).

[2] Wu, J., Dhingra, R., Gambhir, M. & Remais, J. V. Sensitivity analysis of infectious disease models: methods, advances and their application. J. R. Soc. Interface **10(86)**, 1020121018 (2013).

[3] Battista, N. A., Pearcy, L. B. & Strickland, W. C. Modeling the prescription opioid epidemic. Bull. Math. Biol. **81(7)**, 2258–2289 (2019).

[4] Lodhi, H. & Gilbert, D. Bootstrapping parameter estimation in dynamic systems. In Elomaa, T., J, H. & Mannila, H. (eds.) Discovery Science, chap. 17, 194–208 (Springer, Berlin, Germany, 2011).

[5] Barish, S., Ochs, M. F., Sontag, E. D. & Gevertz, J. L. Evaluating optimal therapy robustness by virtual expansion of a sample population, with a case study in cancer immunotherapy. Proc. Natl. Acad. Sci. **114**, E6277–E6286 (2017).

[6] Eriksson, O. et al. Uncertainty quantification, propagation and characterization by bayesian analysis combined with global sensitivity analysis applied to dynamical intracellular pathway models. Bioinformatics **35(2)**, 284–292 (2019).

[7] Waldrop, L. D., He, Y. & Khatri, S. What can computational modeling tell us about the diversity of odor-capture structures in the pancrustacea? J. Chem. Ecol. **44**, 1084–1100 (2018).

[8] Waldrop, L. D., He, Y., Hedrick, T. L. & Rader, J. Functional morphology of gliding flight I. Modeling reveals distinct performance landscapes based on soaring strategies. Int. Comp. Biol. (2020).

[9] Waldrop, L. D., He, Y., Battista, N. A., Neary Peterman, T. & Miller, L. A. Uncertainty quantification reveals the physical constraints on pumping by peristaltic hearts. J. R. Soc. Inter. **17**, 2020032 (2020).

[10] Saltelli, A. Making best use of model evaluations to compute sensitivity indices. Comp. Phys. Comm. **145**, 280–297 (2002).

[11] Anderson, P. S. L. & Patek, S. N. Mechanical sensitivity reveals evolutionary dynamics of mechanical systems. Proc. R. Soc. B: Biol. Sci. **282**, 20143088 (2015).

[12] Muñoz, M. M., Anderson, P. S. L. & Patek, S. N. Mechanical sensitivity and the dynamics of evolutionary rate shifts in biomechanical systems. Proc. R. Soc. B: Biol. Sci. **284**, 20162325 (2017).

[13] Wainwright, P. C., Alfaro, M. E., Bolnick, D. I. & Hulsey, C. D. Many-to-One Mapping of Form to Function: A General Principle in Organismal Design? Int. Comp. Biol. **45**, 256–262 (2005).

[14] Wainwright, P. C. Functional versus morphological diversity in macroevolution. Ann. Rev. Ecol., Evo., and Syst. **38**, 381–401 (2007).

[15] Arnold, S. J. Performance Surfaces and Adaptive Landscapes1. Int. Comp. Biol. **43**, 367–375 (2003).

[16] Muñoz, M. M., Hu, Y., Anderson, P. S. L. & Patek, S. Strong biomechanical relationships bias the tempo and mode of morphological evolution. eLife **7**, e37621 (2018).

[17] Muñoz, M. M. The Evolutionary Dynamics of Mechanically Complex Systems. Int. Comp. Biol. **59**, 705–715 (2019).

[18] Gray, J. & Lissmann, H. W. The locomotion of nematodes. J. Exp. Biol. **41**, 135–154 (1964).

[19] Ghosh, R. & Emmons, S. W. Episodic swimming behavior in the nematode c. elegans. J. Exp. Biol. **211**, 3703–3711 (2008).

[20] Majmudar, T., Keaveny, E. E., Zhang, J. & Shelley, M. J. Experiments and theory of undulatory locomotion in a simple structured medium. J. R. Soc. Inter. **9**, 1809–1823 (2012).

[21] Lüersen, K., Faust, U., Gottschling, D.-C. & Döring, F. Gait-specific adaptation of locomotor activity in response to dietary restriction in caenorhabditis elegans. J. Exp. Biol. **217**, 2480–2488 (2014).

[22] Berman, R. S., Kenneth, O., Sznitman, J. & Leshansky, A. M. Undulatory locomotion of finite filaments: lessons from *Caenorhabditis elegans*. New J. Phys. **15**, 075022 (2013).

[23] Gutierrez, J., Sorenson, M. & Strawbridge, E. Modeling fluid flow induced by *c. elegans* swimming at low reynolds number. In Dediu, A. H., M, M. L. & Martín-Vide, C. (eds.) Theory and Practice of Natural Computing, chap. 7, 71–82 (Springer, Cham, Switzerland, 2014).

[24] Montenegro-Johnson, T. D., Gagnon, D. A., Arratia, P. E. & Lauga, E. Flow analysis of the low reynolds number swimmer c. elegans. Phys. Rev. Fluids **1**, 053202 (2016).

[25] Taylor, G. I. Analysis of the swimming of long and narrow animals. Proc. R. Soc. Lond. Ser. A **214**, 158–183 (1952).

[26] Jordan, C. E. Scale effects in the kinematics and dynamics of swimming leeches. Can. J. Zool. **76**, 1869–1877 (1998).

[27] Tytell, E. D. The hydrodynamics of eel swimming ii. effect of swimming speed. J. Exp. Biol. **207**, 3265–3279 (2004).

[28] Hamlet, C., Fauci, L. J. & Tytell, E. D. The effect of intrinsic muscular nonlinearities on the energetics of locomotion in a computational model of an anguilliform swimmer. J. Theor. Biol. **385**, 119–129 (2015).

[29] Korta, J., Clark, D. A., Gabel, C. V., Mahadevan, L. & Samuel, A. D. T. Mechanosensation and mechanical load modulate the locomotory gait of swimming *C. elegans*. J. Exp. Biol. **210**, 2383–2389 (2007).

[30] Sobol, I. M. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Math. Comput. Simul. **55**, 271–280 (2001).

[31] Eloy, C. On the best design for undulatory swimming. J. Fluid Mech. **717**, 48–89 (2013).

[32] Verma, S., Hadjidoukas, P., Wirth, P., Rossinelli, D. & Koumoutsakos, P. Pareto optimal swimmers. In Proc. of the Platform for Advanced Scientific Computing Conference (Assoc. for Comp. Mach., NY, USA, 2017).

[33] Smits, A. J. Undulatory and oscillatory swimming. J. Fluid Mech. **874**, P1 (2019).

[34] Battista, N. A., Strickland, W. C., Barrett, A. & Miller, L. A. IB2d Reloaded: a more powerful Python and MATLAB implementation of the immersed boundary method. Math. Meth. Appl. Sci **41**, 8455–8480 (2018).

[35] Battista, N. A. Fluid-structure interaction for the classroom: Interpolation, hearts, and swimming! (accepted, in production) SIAM Review (2020).

[36] Jung, S., Lee, S. & Samuel, A. Swimming c. elegans in a wet granular medium. Chaos **18**, 041106 (2008).

[37] Rayner, J. M. Dynamics of the vortex wakes of flying and swimming vertebrates. Symp. Soc. Exp, Biol. **49**, 131–155 (1995).

[38] Battista, N. A., Baird, A. J. & Miller, L. A. A mathematical model and matlab code for muscle-fluid-structure simulations. Integr. Comp. Biol. **55(5)**, 901–911 (2015).

[39] Battista, N. A., Strickland, W. C. & Miller, L. A. IB2d: a Python and MATLAB implementation of the immersed boundary method. Bioinspir. Biomim. **12(3)**, 036003 (2017).

[40] Peskin, C. S. The immersed boundary method. Acta Numerica **11**, 479–517 (2002).

[41] Backholm, M., Kasper, A. K. S., Schulman, R. D., Ryu, W. S. & Dalnoki-Veress, K. The effects of viscosity on the undulatory swimming dynamics of *C. elegans*. Phys. Fluids **27**, 091901 (2015).

[42] Jung, S. Caenorhabditis elegans swimming in a saturated particulate system. Phys. Fluids **22**, 031903 (2010).

[43] Tytell, E. D. & Lauder, G. V. The hydrodynamics of eel swimming. J. Exp. Biol. **207**, 1825–1841 (2004).

[44] Cui, Z., Gu, X., Li, K. & Jiang, H. Cfd studies of the effects of waveform on swimming performance of carangiform fish. Appl. Sci. **7**, 149 (2017).

[45] Dai, L., He, G., Zhang, X. & Zhang, X. Stable formations of self-propelled fish-like swimmers induced by hydrodynamic interactions. J. R. Soc. Inter. **15**, 20180490 (2018).

[46] Cui, Z., Yang, Z. & Jiang, H. Sharp interface immersed boundary method for simulating three-dimensional swimming fish. Enging. Appl. Comp. Fluid Mech. **14**, 534–544 (2020).

[47] Taylor, G. K., Nudds, R. L. & Thomas, A. L. Flying and swimming animals cruise at a strouhal number tuned for high power efficiency. Nature **425**, 707–711 (2003).

[48] Eloy, C. Optimal strouhal number for swimming animals. J. Fluids and Structures **30**, 205 – 218 (2012).

[49] Bale, R., Hao, M., Bhalla, A. & Patankar, N. A. Energy efficiency and allometry of movement of swimming and flying animals. Proc. Natl. Acad. Sci. **111(21)**, 7517–7521 (2014).

[50] Saltelli, A. et al. Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index. Comp. Phys. Comm. **18**, 259–270 (2010).

[51] Link, K. G. et al. A local and global sensitivity analysis of a mathematical model of coagulation and platelet deposition under flow. PLOS ONE **13(7)**, e0200917 (2018).

[52] Saltelli, A. et al. Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. Environ. Model. & Software **114**, 29 – 39 (2019).

[53] The College of New Jersey. Electronic laboratory for science & analysis (elsa) (2020). URL `https://docs.hpc.tcnj.edu/`. Accessed Online; accessed 24 January 2020.

[54] Gillis, G. B. Undulatory locomotion in elongate aquatic vertebrates: Anguilliform swimming since sir james gray. American Zoologist **36**, 656–665 (1996).

[55] Padmanabhan, V. et al. Locomotion of c. elegans: A piecewise-harmonic curvature representation of nematode behavior. PLOS ONE **7**, 1–11 (2012). URL `https://doi.org/10.1371/journal.pone.0040121`.

[56] van Rees, W. M., Gazzola, M. & Koumoutsakos, P. Optimal shapes for anguilliform swimmers at intermediate reynolds numbers. J. of Fluid Mech. **722**, R3 (2013).

[57] Williams, T. L. & McMillen, T. Strategies for swimming: explorations of the behaviour of a neuro-musculo-mechanical model of the lamprey. Biology Open **4**, 253–258 (2015).

[58] Schuech, R., Hoehfurtner, T., Smith, D. J. & Humphries, S. Motile curved bacteria are pareto-optimal. PNAS **116(29)**, 14440–14447 (2019).

[59] Tokić, G. & Yue, D. Energetics of optimal undulatory swimming organisms. PLOS Comp. Biol. **15**, 1–25 (2019).

[60] Alben, S., Witt, C., Baker, T. V., Anderson, E. & Lauder, G. V. Dynamics of freely swimming flexible foils. Phys. Fluids **24**, 051901 (2012).

[61] Shelton, R. M., Thornycroft, P. J. M. & Lauder, G. V. Undulatory locomotion of flexible foils as biomimetic models for understanding fish propulsion. J. Exp. Biol. **217**, 2110–2120 (2014).

[62] Bainbridge, R. The speed of swimming of fish as related to size and to the frequency and amplitude of the tail beat. J. Exp. Biol. **35**, 109–133 (1958).

[63] Gray, J. Animal Locomotion (World Naturalist) (Weidenfeld and Nicolson, London, UK, 1968).

[64] Steinhausen, M. F., Steffensen, J. F. & Andersen, N. G. Tail beat frequency as a predictor of swimming speed and oxygen consumption of saithe (*Pollachius virens*) and whiting (*Merlangius merlangus*) during forced swimming. Marine Biology **148**, 197–204 (2005).

[65] Hoover, A. P., Cortez, R., Tytell, E. & Fauci, L. Swimming performance, resonance and shape evolution in heaving flexible panels. J. Fluid. Mech. **847**, 386–416 (2018).

[66] van Rees, W. M., Gazzola, M. & Koumoutsakos, P. Optimal morphokinematics for undulatory swimmers at intermediate reynolds numbers. J. of Fluid Mech. **775**, 178–188 (2015).

[67] Hamlet, C. L., Hoffman, K. A., Tytell, E. D. & Fauci, L. J. The role of curvature feedback in the energetics and dynamics of lamprey swimming: A closed-loop model. PLOS Comp. Biol. **14**, 1–29 (2018). URL `https://doi.org/10.1371/journal.pcbi.1006324`.

[68] Jordan, C. E. A model of rapid-start swimming at intermediate reynolds number: undulatory locomotion in the chaetognath *Sagitta elegans*. J. Exp. Biol. **163**, 119–137 (1992).

[69] Tyson, R., Jordan, C. E. & Hebert, J. Modelling anguilliform swimming at intermediate reynolds number: A review and a novel extension of immersed boundary method applications. Comp. Meth. Appl. Mech. Enging. **197**, 2105 – 2118 (2008).

[70] Klotsa, D. As above, so below, and also in between: mesoscale active matter in fluids. Soft Matter **15**, 8946–8950 (2019).

[71] Gray, J. & Hancock, G. J. The propulsion of sea-urchin spermatozoa. J. Exp. Biol. **32**, 802–814 (1955).

[72] Lighthill, M. J. Hydromechanics of aquatic animal propulsion. Ann. Rev. Fluid Mech. **1**, 413–446 (1969).

[73] Gazzola, M., van Rees, W. M. & Koumoutsakos, P. C-start: optimal start of larval fish. J. of Fluid Mech. **698**, 5–18 (2012).

[74] Müller, U. K., van den Boogaart, J. G. M. & van Leeuwen, J. L. Flow patterns of larval fish: undulatory swimming in the intermediate flow regime. J. Exp. Biol. **211**, 196–205 (2008).

[75] Du Clos, K. T. et al. Thrust generation during steady swimming and acceleration from rest in anguilliform swimmers. J. Exp. Biol. **222** (2019).

[76] Tytell, E., Hsu, C., Williams, T., Cohen, A. & Fauci, L. Interactions between internal forces, body stiffness, and fluid environment in a neuromechanical model of lamprey swimming. Proc. Natl. Acad. Sci. **107**, 19832–19837 (2010).

[77] Tytell, E. D. Kinematics and hydrodynamics of linear acceleration in eels, *Anguilla rostrata*. Proc.: Biol. Sci. **271**, 2535–2540 (2004).

[78] Borazjani, I. & Sotiropoulos, F. Numerical investigation of the hydrodynamics of anguilliform swimming in the transitional and inertial flow regimes. J. Exp. Biol. **212**, 576–592 (2009).

[79] Tytell, E. D. et al. Disentangling the functional roles of morphology and motion in the swimming of fish. Int. Comp. Biol. **50(6)**, 1140–1154 (2010).

[80] Kern, S. & Koumoutsakos, P. Simulations of optimized anguilliform swimming. J. Exp. Biol. **209**, 4841–4857 (2006).

[81] Tokić, G. & Yue, D. Optimal shape and motion of undulatory swimming organisms. Proc. R. Soc. B: Biol. Sci. **279**, 3065–3074 (2012).

[82] Gazzola, M., Argentina, M. & Mahadevan, L. Gait and speed selection in slender inertial swimmers. PNAS **112**, 3874–3879 (2015).

[83] Low, K. H., Prabu, S. & Pattathil, A. P. Initial prototype design and development of hybrid modular underwater vehicles. In 2006 IEEE Int. Conf. on Robotics and Biomimetics, 311–316 (2006).

[84] Niu, X. & Xu, J. Modeling, control and locomotion planning of an anguilliform robotic fish. Unmanned Systems **02**, 295–321 (2014).

[85] Feng, H., Sun, Y., Todd, P. A. & Lee, H. P. Body wave generation for anguilliform locomotion using a fiber-reinforced soft fluidic elastomer actuator array toward the development of the eel-inspired underwater soft robot. Soft Robotics **7**, 233–250 (2020).

[86] Gan, Y. et al. A comprehensive evaluation of various sensitivity analysis methods: A case study with a hydrological model. Envir. Model. Soft. **51**, 269–285 (2014).

[87] Zhang, X. Y., Trame, M., Lesko, L. & Schmidt, S. Sobol sensitivity analysis: A tool to guide the development and evaluation of systems pharmacology models. CPT Pharmacometrics Syst. Pharmacol. **4**, 69–79 (2014).

[88] Nossent, J., Elsen, P. & Bauwens, W. Sobol' sensitivity analysis of a complex environmental model. Envir. Model. Soft. **26(12)**, 1515–1525 (2011).

[89] Xiu, D. & Karniadakis, G. E. Modeling uncertainty in flow simulations via generalized polynomial chaos. J. Comp. Phys. **187**, 137–167 (2003).

[90] Xiu, D., Lucor, D., Su, C. H. & Karniadakis, G. E. Performance evaluation of generalized polynomial chaos. In Sloot, P. et al. (eds.) Int. Conf. on Comp. Sci., chap. 36, 346–354 (Springer, Berlin, Heidelberg, Germany, 2003).

[91] Sudret, B. Global sensitivity analysis using polynomial chaos expansions. Reliab. Enging. & Sys. Safety **93(7)**, 964–979 (2008).

Acknowledgements

The author would like to thank Lindsay Waldrop and Jonathan Rader for the invitation to participate in the SICB Symposium *Melding Modeling and Morphology: integrating approaches to understand the evolution of form and function symposium* at the 2020 annual meeting. The author would also like to thank Christina Battista, Karen Clark, Jana Gevertz, Laura Miller, Matthew Mizuhara, Emily Slesinger, Edward Voskanian, and Lindsay Waldrop for comments and discussion. The author also wishes to thank the anonymous reviewers for their careful reading of the manuscript and their very insightful, constructive feedback that significantly strengthened the manuscript. Computational resources were provided by the NSF OAC #1826915 and the NSF OAC #1828163. Support for N.A.B. was provided by the TCNJ Support of Scholarly Activity Grant, the TCNJ Department of Mathematics and Statistics, and the TCNJ School of Science.

Supplemental Materials

Computational Parameters and Geometry

Table S1 offers the computational parameters used in the study. They are identical to those in [1] with the exception of the parameter ranges for Re_{in}, f, and p.

Parameter	Variable	Units	Value
Domain Size	$[L_x, L_y]$	m	$[6, 16]$
Spatial Grid Size	$dx = dy$	m	$L_x/1024 = L_y/384$
Lagrangian Grid Size	ds	m	$dx/2$
Time Step Size	dt	s	2.5×10^{-5}
Total Simulation Time	T	*stroke cycles*	6
Fluid Density	ρ	kg/m^3	1000
Fluid Dynamic Viscosity	μ	$kg/(ms)$	[0.475,17750]
Swimmer Length	L	m	1.5
Swimmer Height	h	m	0.5
Stroke Frequency	f	s^{-1}	[1,3.0]
Reynolds number (input)	Re_{in}	-	[450,2200]
Kinematic Parameter	p	–	[0.05,0.45]
Spring Stiffness	k_{spr}	$kg \cdot m/s^2$	9.5625×10^9
Non-invariant Beam Stiffness	k_{beam}	$kg \cdot m/s^2$	2.03634×10^{12}

Table S1: Numerical parameters used in the two-dimensional immersed boundary simulations of the idealized, simple anguilliform swimmer

The undulations of the swimmer's body that produce forward swimming are governed by interpolating between two curvature states. The function that governs the interpolation is given by a cubic spline [35, 1]. This interpolant satisfies the following criteria:

1. Continuity of the interpolation polynomial, its velocity, and its acceleration between successive half-strokes and stroke cycles
2. No instantaneous acceleration or deceleration
3. Symmetry within each half-stroke cycle

The resulting interpolant can be seen in Figure S1 for a variety of kinematic parameter (p) values. This figure was modified from [1]. As p increases, the interpolant deviates from trivial sinusoidal curvature interpolation functions. Some anguilliform studies investigating optimal morphokinematics use a mix of cubic splines and sinusoidal functions to prescribe the swimmer's body kinematics [80, 66]; however, these constructions result

in many more free parameters that govern the body's shape and kinematics, rather than only one (p) in the simplified model explored here.

The curvature interpolation scheme used here provides functionality to subtly change properties of the stroke, i.e., the maximum velocity and accelerations of the changing body curvature during each upstroke or downstroke. Given two (or more) curvature (geometric) states, one could design a custom interpolation function(s) based on an animal's kinematic data. This could be done by tracking points along the moving body and in addition to recording the position over time, using such data to approximate the points' associated velocity and acceleration between different positions. The curvature interpolant here used a cubic spline interpolant [35, 1], which allowed for one free parameter (p) that could be varied to change the body curvature's maximal velocity and acceleration. Increasing p would continue to deviate away from pure sinusoidal behavior, see Figure S1. Moreover, this framework allows one to enforce that no instantaneous accelerations occur along the swimmer's body, see Figure S1c. Introducing a higher-order interpolation function would give rise to more free parameters which would allow one to more closely replicate the exact kinematics of a moving body.

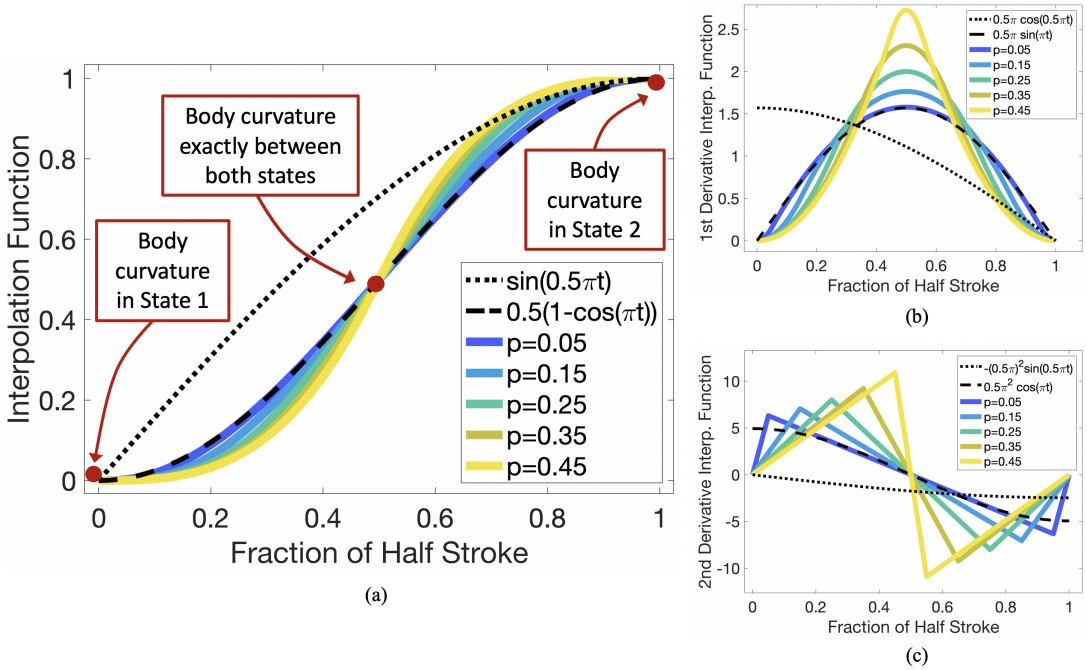

Figure S1: (a) Examples of how the interpolation function deviates away from a trivial sinusoidal interpolation between body curvature states as p increases. The associated velocity and acceleration profiles are given in (b) and (c) respectively. Figure modified and courtesy of [1].

Since it was desired to have a fluid scaling parameter as an input to the model, an input Reynolds number was defined in Eq. 2, based on a characteristic length that was the swimmer's body (L) and a frequency based characteristic velocity that was the product of the undulation frequency and the swimmer's body length, fL [44, 45]. Many

anguilliform swimming studies use the characteristic velocity as fA, where A is the peak-to-peak undulation amplitude. However, since A is an output of the model, using a characteristic velocity scale of fA was impossible for the input Re. Figure S2 provides plots of the input Re vs. output Re, as organized by the input parameters f and p in Figures S2b and c, respectively. The output Re, Re_{out} depends on model output A, as a pattern emerges, where higher f corresponds to lower Re_{out}, given a particular Re_{in}. This occurs as f and A are inversely related, i.e., higher f results in a smaller A, see Figure 7). Figure S3 provides a deeper look into the relationship among the three input parameters and Re_{out}, while Figure S4 confirms that the output Reynolds number is most significantly affected by changes to the input Reynolds number.

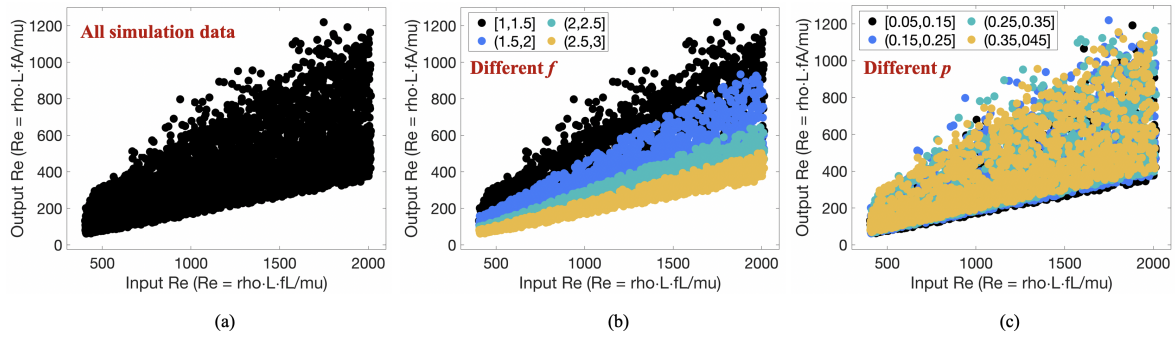

Figure S2: The input Reynolds number, $\mathrm{Re}_{in} = \rho L(fL)/\mu$, plotted against the output Reynolds number, $\mathrm{Re}_{out} = \rho L(fA)/\mu$, whose frequency based velocity scale (fA) is based off the undulation peak-to-peak amplitude, A, which is an output of the model.

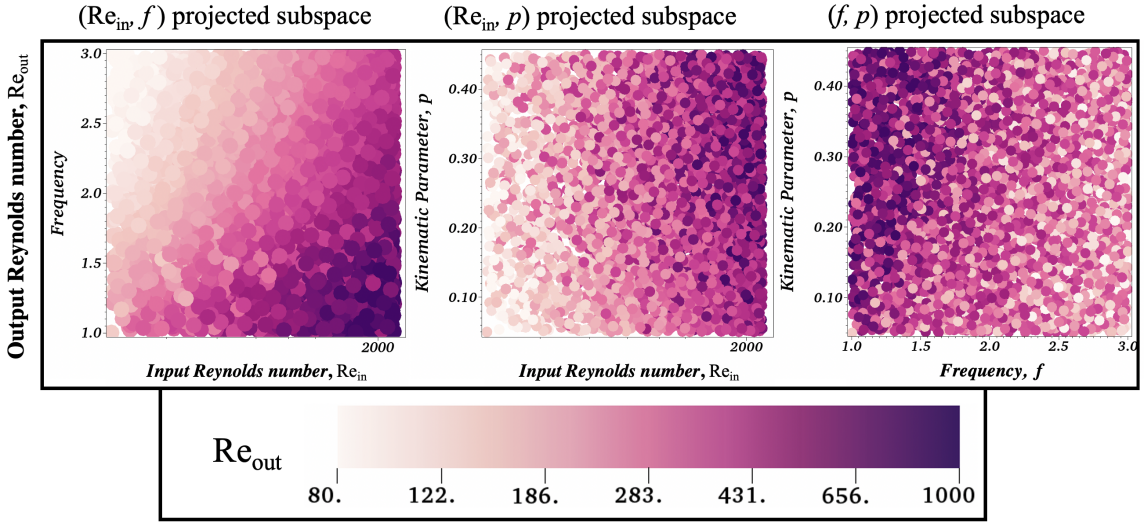

Figure S3: Colormaps corresponding to the output Reynolds number, Re_{out}, for all the data sampled from Sobol sequences when it is projected onto either the (Re_{in}, f) and (Re_{in}, p), or (f, p) subspaces.

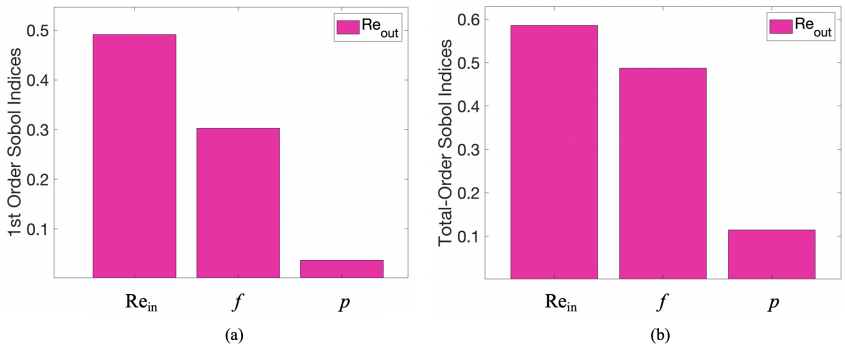

Figure S4: (a) First-order and (b) Total-order Sobol indices of the three varied parameters Re_{in}, f, and p for the output Reynolds number, Re_{out}.

Global vs. local sensitivity analyses

Sobol sensitivity is a variance-based sensitivity analysis that can provide *global* sensitivity to parameters, rather than only local sensitivity [51, 52]. By quantifying global sensitivity, one can determine which parameter, when varied within a particular range, results in the most significant changes in the model's output, even with respect to other parameters being varied. Moreover, Sobol analysis is able to efficiently calculate first-order parameter sensitivity indices, i.e., perturbations of one parameter at a time, but also higher-order indices, i.e., those corresponding to perturbations of two or more parameters at a time, and total-order indices, i.e., all combinations of other parameters [50, 52]. Furthermore, due to this, the importance of higher-order interactions can be inferred by comparing first-order and total-order sensitivity indices. If there are significant differences between these indices, it suggests the presence of higher-order interactions. Higher-order interactions occur when two or more parameters are changed and it causes a greater variation in the output than when varying each of those inputs alone.

If one changes the input parameter space sampled for sensitivity, the sensitivity indices, like those in Figure 4 or S7, could significantly change. Thus, while Sobol sensitivity analysis gives global sensitivities of model output metrics to model parameters, the sensitivity results are highly dependent on the sampled parameter space [86, 87]. Thus, careful consideration must be made to choose the appropriate input parameter ranges from either experimental data or previously performed parameter studies.

In comparison to *global* sensitivity analyses, *local* sensitivity analyses, on the other hand, have the misfortune of necessitating that only one parameter can be varied at a time, which considerably restricts the parameter space that is able to be explored and analyzed. Under-resolving the input parameter space could easily lead to inaccurate sensitivities, unless the model is linear [52]. The model examined here was previously seen to exhibit non-linear behavior [1]. Therefore it was deemed necessary to use global approaches for sensitivity.

Furthermore, higher-order interactions are difficult to accurately parse out of the data

using only local sensitivity methods. However, with enough work, some local sensitivity estimates might be able to arrive at the same conclusion that a model is most sensitive to a particular parameter [51]. Unfortunately, if the goal of the sensitivity analysis is for model reduction, i.e., eliminating parameters or dynamics from a model, it might incorrectly suggest some parameters are not important, as it is difficult to predict the importance of parameter interactions using local sensitivity methods [52].

Additional Data

Non-dimensional Data

Figures S5 and S6 provide colormaps of the performance data in terms of the output Reynolds number, $\mathrm{Re}_{out} = \rho \cdot L \cdot fA/\mu$, based on the output peak-to-peak stroke amplitude, across the (Re_{in}, f) and (Re_{in}, p) projected subspaces, respectively. The resulting output Reynolds numbers spanned $[50, 1220]$. Patterns emerge across both projected subspaces as either Re_{out} and either f or p vary in each subspace. Although, there are regions within the subspaces that for particular performance metrics look qualitatively like noise, e.g., the colormap for the (Re_{out}, p) subspace for COT when $p \gtrsim 0.25$ for all Re_{out}. Such noisy regions might suggest that the two varying subspace parameters are not the main parameters driving that performance metric in that region. The third free parameter (and possible interactions with the others) may have a more significant impact on that performance metric there.

Interestingly, both subspaces suggest that the largest regions of highest swimming speed are located within regions that may encompass both high and low values in COT, near $f \sim 2.0$ Hz. These regions also correspond to higher d_{eff}. Moreover, the Strouhal numbers almost everywhere within the subspaces fall within the optimal region of $0.2 < St < 0.4$ [47], except near the lower end of the Re_{out} spectrum, where they begin to rise greater than 0.4 as Re_{out} decreases.

Furthermore, the (Re, f) projected parameter subspace data is qualitatively very similar when organized by either Re_{in} (Figures 6 and 7) or Re_{out} (Figure S5). These consistencies suggest that the Sobol sensitivity analysis would remain consistent for either Re_{in} or Re_{out}. However, there are clear differences in the case of (Re, p) among Figures 6 and 7) when compared to Figure S6. That is, the data appears noisy in the former, while patterns form in the latter. The patterns that emerge in Figure S6 indicate a strong dependence on fA, as it organized by Re_{out}. However, this is actually to be expected since Re_{out} depends on the output A and it is observed that f and A are inversely related (see the A panel for the (Re, f) and (f, p) subspaces in Figure 7). So the patterns that emerge are a result that although the input frequency f is not one of the axes, its presence is felt by the dependence of A in Re_{out}.

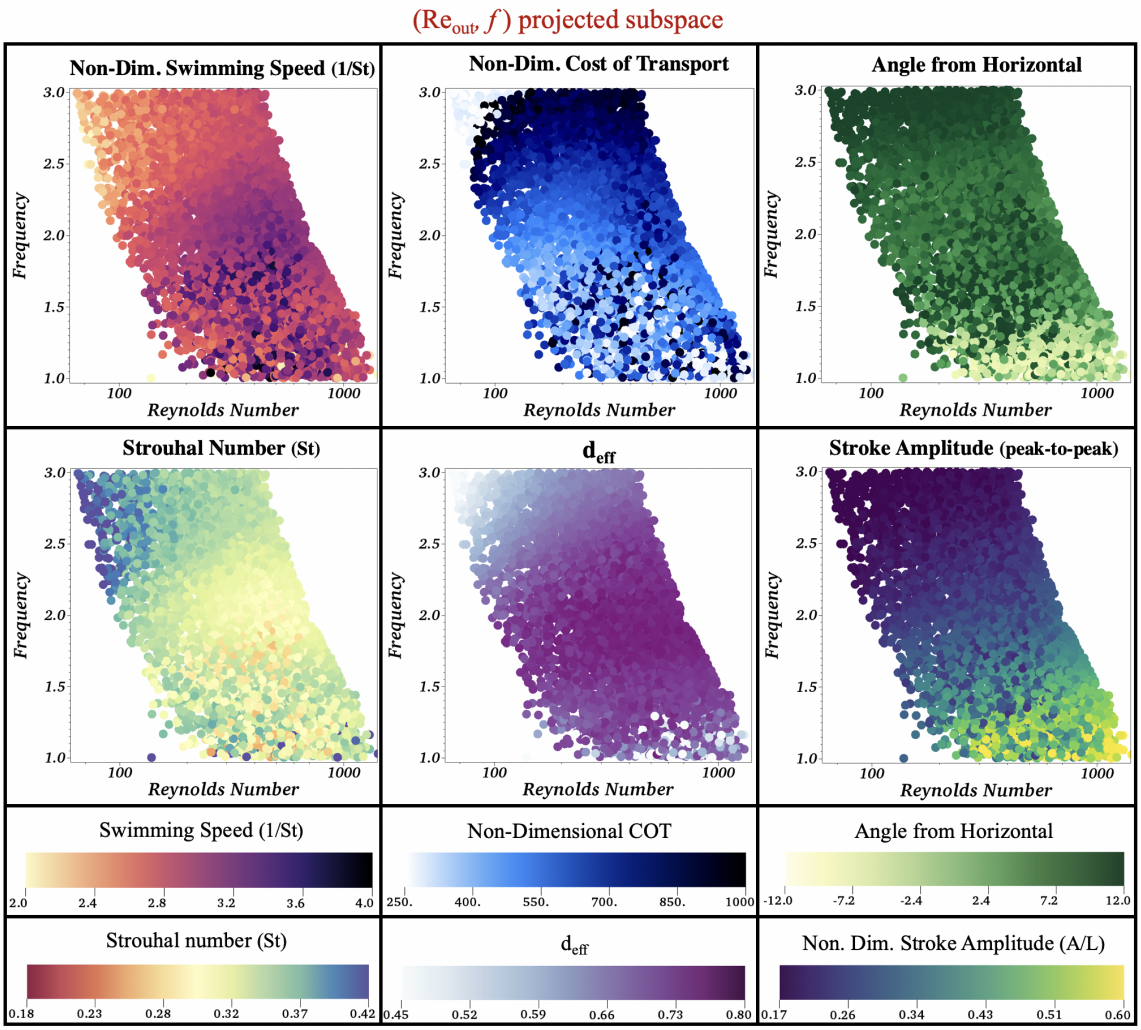

Figure S5: Colormaps corresponding to the (Re_{out}, f) projected subspace, providing the non-dimensional forward swimming speeds (1/St), cost of transports (COT), and stroke amplitudes (A/L), as well as the Strouhal numbers (St), the distance effectiveness ratios (d_{eff}), and the angular trajectories off the horizontal (θ) for all the data sampled from Sobol sequences. Note that the Reynolds numbers shown here are the Re_{out}, whose frequency based velocity scale (fA) is given in terms of the output stroke amplitude, A.

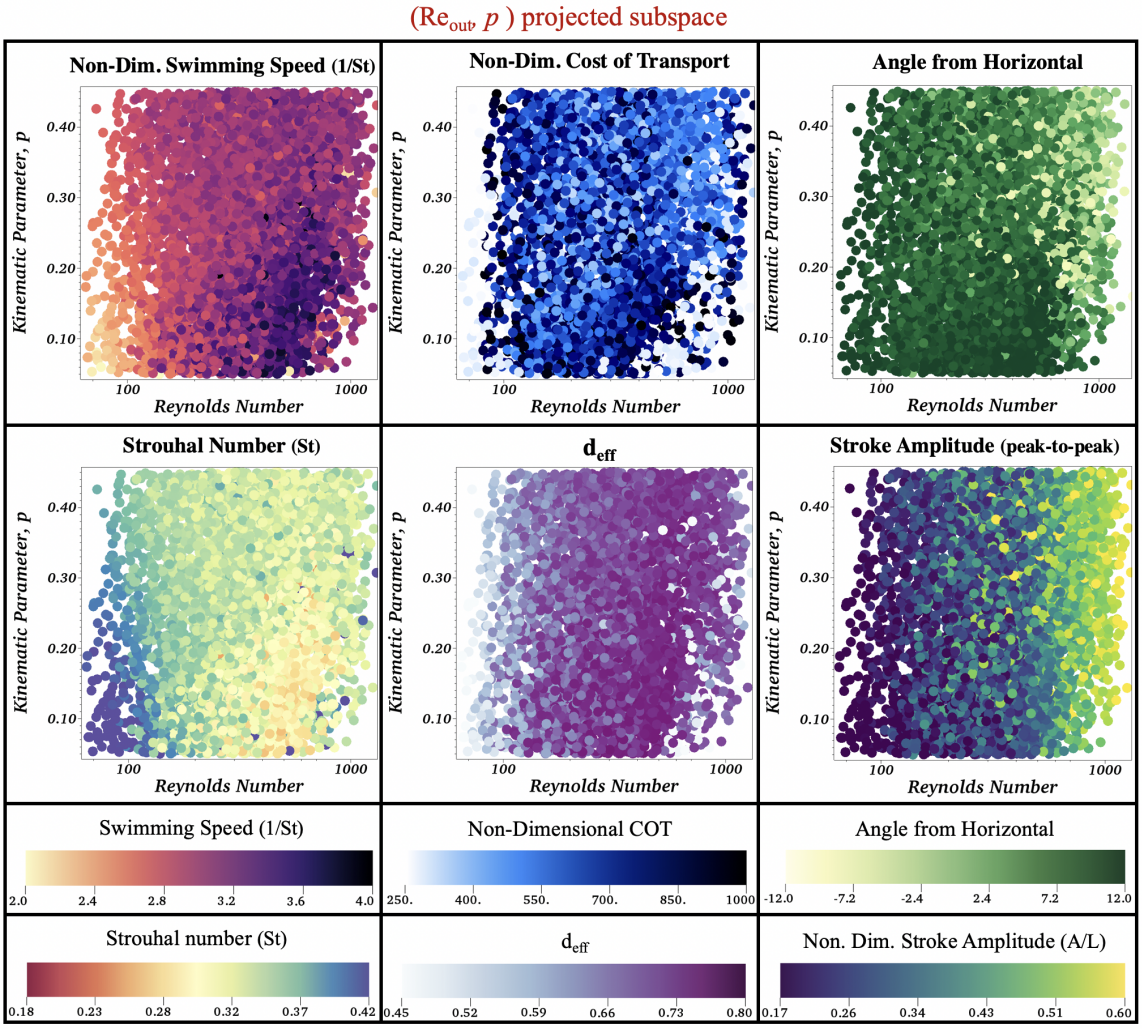

Figure S6: Colormaps corresponding to the (Re_{out}, p) projected subspace, providing the non-dimensional forward swimming speeds (1/St), cost of transports (COT), and stroke amplitudes (A/L), as well as the Strouhal numbers (St), the distance effectiveness ratios (d_{eff}), and the angular trajectories off the horizontal (θ) for all the data sampled from Sobol sequences. Note that the Reynolds numbers shown here are the Re_{out}, whose frequency based velocity scale (fA) is given in terms of the output stroke amplitude, A.

Dimensional Data

A comparison of the global parameter sensitivities for swimming speed and cost of transport between their dimensional and non-dimensional form is provided in Figure S7. The dimensional output metrics for swimming speed (bodylength/s) and COT (N/kg) are still most sensitive to variations in the stroke (undulation) frequency within the $\mathrm{Re}_{in} \times f \times p = [450, 2200] \times [1, 3] \times [0.05, 0.45]$ parameter subspace. However, both of the dimensional swimming speed and cost of transport are slightly more sensitive to Re_{in} than p.

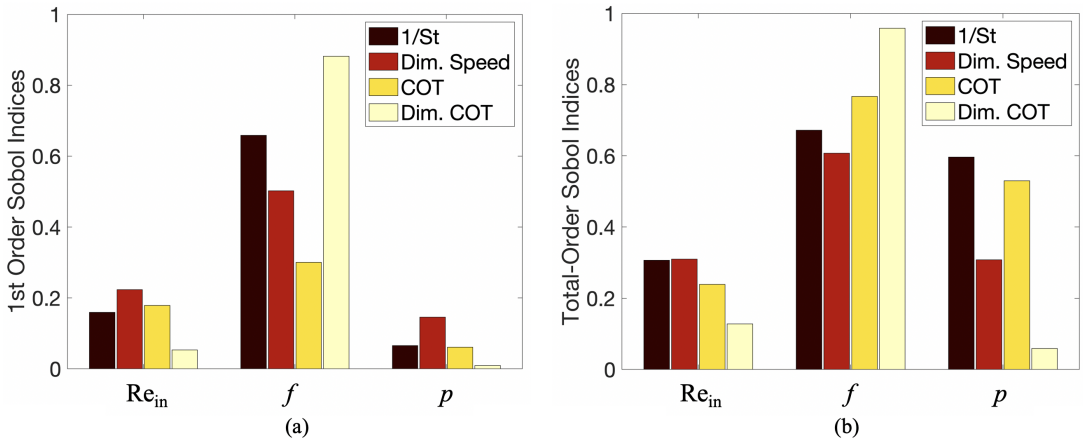

Figure S7: (a) First-order and (b) Total-order Sobol indices of the three varied parameters Re_{in}, f, and p to compare global parameter sensitivities of dimensional to non-dimensional quantities for swimming speed and cost of transport.

Colormaps showing the *dimensional* forward swimming speeds (bodylength/s) and cost of transports (N/kg) over each projected parameter subspace are provided in Figure S8. Similar to Figures S5 and S6, patterns emerge within the subspaces as different parameters are varied. Note that these Figures give the Reynolds number as Re_{out}, unlike Figures 6 and 7 in the main manuscript. Beyond similar trends in f, dimensional swimming speed appears to be highly correlated with higher Re_{out} and higher p. As Figure S1 suggests, higher p correspond to slower initial accelerations, but higher maximal velocities and accelerations of the changing body curvature throughout each half-stroke.

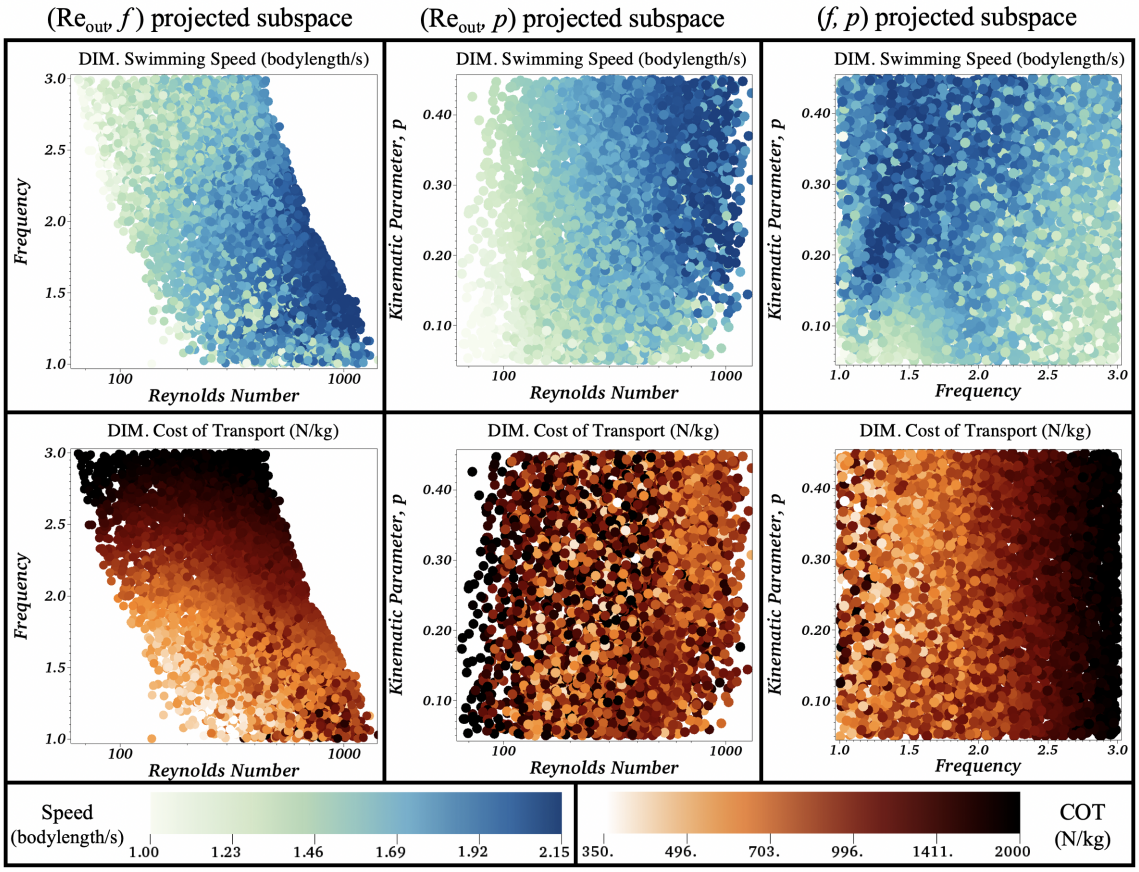

Figure S8: Colormaps corresponding to the *dimensional* forward swimming speeds (bodylength/s) and cost of transports (N/kg) for all the data sampled from Sobol sequences when projected onto either the (Re_{out}, f) and (Re_{out}, p) subspaces.

Pareto-like optimal fronts were identified by plotting the *dimensional* cost of transport (N/kg) against swimming speed (bodylength/s) for each simulation performed. The data is presented in Figure S9. From Figure S9d, given a p, depending on values of the other two parameters, one could construct a swimmer that falls almost anywhere within the performance space. However, Figure S9c shows distinct clusters where different frequency ranges reside and Figure S9b suggests higher Re_{out} result in faster swimming speeds (bodylength/s) and lower cost of transport (N/kg). This complemented the Sobol sensitivity results for the dimensional data analyzed; choosing a particular f would most significantly determine where in the performance space a swimmer may reside.

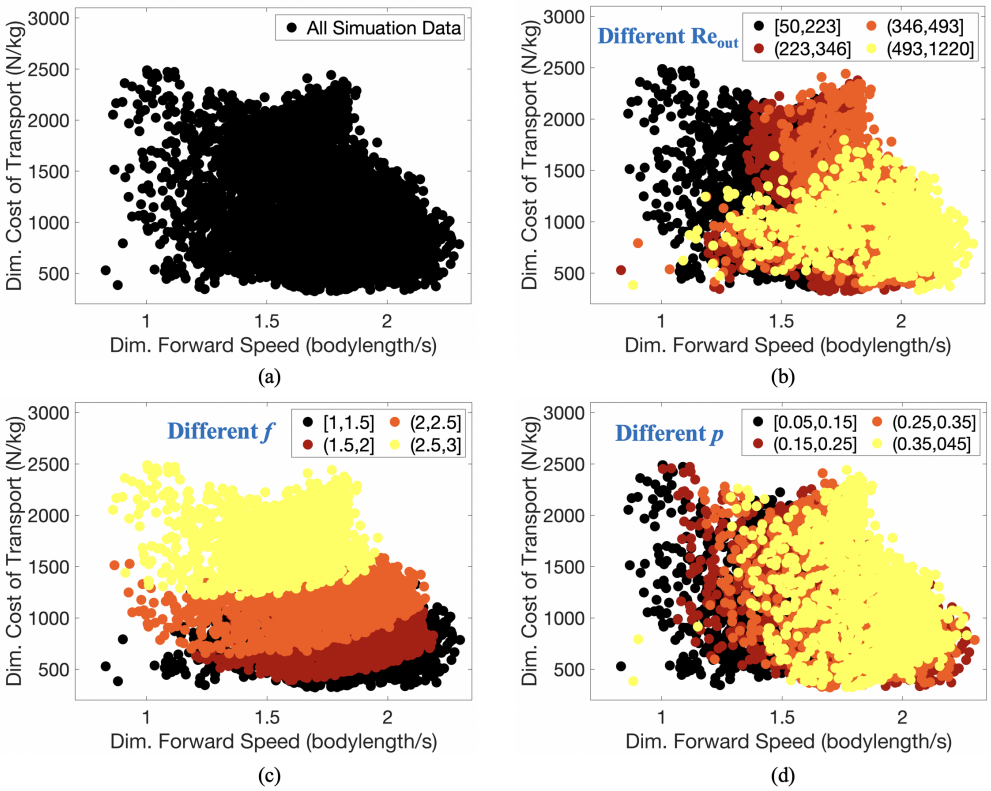

Figure S9: The *dimensional* cost of transport (N/kg) and forward swimming speeds (bodylength/s) plotted against each other for (a) every simulation performed, as well as colors indicating different ranges of the input parameters: (b) different Re_{in} (c) different f, and (d) different p.

Figure S10 shows where different parameter combination swimmers lie within the performance space in Figure S10. This figure also illustrates those swimmer's position and vortex wakes after their 5^{th} stroke cycle. Again, the Reynolds number indicated here is the output Re, $\mathrm{Re}_{out} = \rho L(fA)/\mu$, where A is the peak-to-peak amplitude, computed as output from the model. These are the same swimmers illustrated in Figure 9, but in the context of the dimensional performance space.

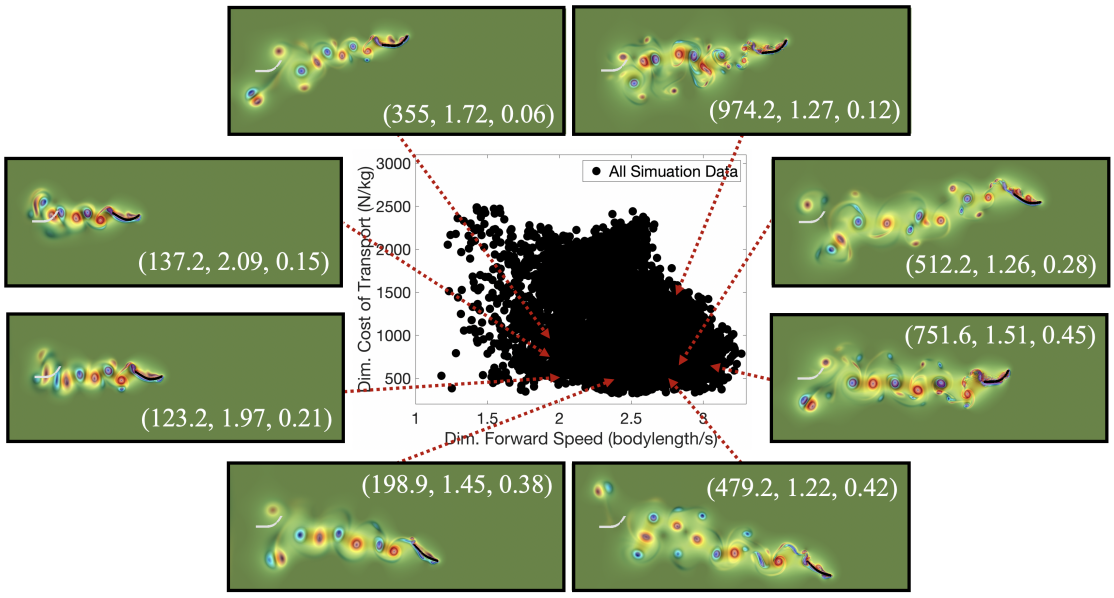

Figure S10: Numerous swimmer's vortex wake and position after its 5^{th} full stroke cycle as well as indicating where they fall onto the *dimensional* cost of transport (N/kg) vs. swimming speed (bodylength/s) performance space. The colormap illustrates vorticity and the greyed out swimmer shows the starting position of the swimmer in each case. Different parameter combinations $(\mathrm{Re}_{out}, f, p)$ lead to different swimming behavior, as indicated by its placement within the performance space and vortex wake left behind.

Details regarding the Immersed Boundary Method (IB)

Here the *immersed boundary method* (IB) [40] will be briefly introduced. The IB was fluid-structure interaction method that solved the equations that coupled the angulliform swimmer and the fluid to which it was immersed. The open-source IB software `IB2d` [38, 39, 34] was used for all the simulations presented in this work.

The full viscous, incompressible Navier-Stokes equations were used to model the fluid since $Re \in [250, 750]$, i.e.,

$$\rho\Big[\frac{\partial \mathbf{u}}{\partial t}(\mathbf{x},t) + \mathbf{u}(\mathbf{x},t)\cdot\nabla\mathbf{u}(\mathbf{x},t)\Big] = -\nabla p(\mathbf{x},t) + \mu\Delta\mathbf{u}(\mathbf{x},t) + \mathbf{F}(\mathbf{x},t) \tag{7}$$

$$\nabla\cdot\mathbf{u}(\mathbf{x},t) = 0 \tag{8}$$

where $\mathbf{u}(\mathbf{x},t)$ and $p(\mathbf{x},t)$are the fluid's velocity and pressure, respectively, at spatial location $\mathbf{x}$ at time t. $\mathbf{F}(\mathbf{x},t)$ is the force per unit area applied to the fluid by the immersed boundary, i.e., the swimmer. These three quantities are modeled in an Eulerian framework on a fixed rectangular mesh. ρ and μ are the fluid's density and dynamic viscosity, respectively.

All *interactions* between the swimmer and fluid are governed by integral equations with delta function kernels. As the swimmer bends, deformation forces are spread from its body to the nearest fluid mesh points. Similarly, the fluid velocity is interpolated back

onto the swimmer to ensure the no-slip condition is satisfied. The integral equations that govern these dynamics are given as:

$$\mathbf{F}(\mathbf{x},t) = \int \mathbf{f}(s,t)\delta\left(\mathbf{x} - \mathbf{X}(s,t)\right) ds \tag{9}$$

$$\frac{\partial \mathbf{X}}{\partial t}(s,t) = \int \mathbf{u}(\mathbf{x},t)\delta\left(\mathbf{x} - \mathbf{X}(s,t)\right) d\mathbf{x}. \tag{10}$$

$\mathbf{X}(s,t)$ and $\mathbf{f}(s,t)$ give the Cartesian coordinates and deformation forces along the immersed boundary (the swimmer) for each point denoted by Lagrangian parameter, s, and at time, t, respectively. Eqns (9)-(9) essentially transform Lagrangian variables to Eulerian variables and vice versa. Here $\delta(\mathbf{x})$ is a 2D delta function. These delta functions help ensure that forces from the immersed body are spread only to the nearest fluid mesh points to the immersed boundary, and vice-versa for when the velocity field is interpolated back to immersed boundary.

To construct a discretized 1D swimmer with physical meaning, throughout the entire swimmer's body linear springs and beams are used to connect adjacent Lagrangian points. The deformation force equations for springs and beams are given as the following,

$$\mathbf{F}_{spr} = -k_{spr}\left(1 - \frac{R_L}{||\mathbf{X}_F - \mathbf{X}_L||}\right) \cdot \left(\mathbf{X}_L - \mathbf{X}_F\right). \tag{11}$$

$$\mathbf{F}_{beam} = -k_{beam}\frac{\partial^4}{\partial s^4}\Big(\mathbf{X}(s,t) - \mathbf{X}_B(s,t)\Big), \tag{12}$$

where k_{spr} and k_{beam} are the spring stiffness and beam stiffness coefficients, respectively. In (11), the terms X_L and X_F represent the Cartesian positions of two Lagrangian nodes to which are connected by a spring - a leader (L) and a follower (F) node, at time, t. R_L is that spring's corresponding resting length. In (12), $\mathbf{X}_B(s,t)$ represents the preferred curvature (shape) of the swimmer's body at time, t. This model dynamically changes $\mathbf{X}_B(s,t)$ over time, interpolating changing between preferred curvature states. Therefore this self-propelled swimmer propagates forward only due to time-varying body curvature and not through explicit prescribed motion of its Lagrangian points.

`IB2d` discretizes each beam using 3 Lagrangian points; define those three points as

$$\begin{aligned} \mathbf{X}(s,t) &= (X_q, Y_q), \\ \mathbf{X}(s+1,t) &= (X_r, Y_r), \\ \mathbf{X}(s-1,t) &= (X_p, Y_p). \end{aligned} \tag{13}$$

As described in [34], since Newton's Second Law of Motion relates forces to an overall acceleration, (12) is further discretized by a second derivative with respect to time. Thereby, all deformation forces arises from beams can be calculated in the following

manner

$$
\begin{aligned}
F_{beam}(s-1,1) &= -k_{beam}\begin{pmatrix} X_r - 2X_q + X_p - C_x(t) \\ Y_r - 2Y_q + Y_p - C_y(t) \end{pmatrix}, \\
F_{beam}(s,1) &= 2k_{beam}\begin{pmatrix} X_r - 2X_q + X_p - C_x(t) \\ Y_r - 2Y_q + Y_p - C_y(t) \end{pmatrix}, \\
F_{beam}(s+1,1) &= -k_{beam}\begin{pmatrix} X_r - 2X_q + X_p - C_x(t) \\ Y_r - 2Y_q + Y_p - C_y(t) \end{pmatrix}.
\end{aligned} \tag{14}
$$

Note that $C_x(t)$ and $C_y(t)$ are the preferred curvatures at time t. In these discretizations they are defined as

$$
\mathbf{C} = \begin{pmatrix} C_x \\ C_y \end{pmatrix} = \begin{pmatrix} X_{r_{pref}} - 2X_{q_{pref}} + X_{p_{pref}} \\ Y_{r_{pref}} - 2Y_{q_{pref}} + Y_{p_{pref}} \end{pmatrix}, \tag{15}
$$

where the $pref$ subscript denotes the preferred geometric configuration. These quantities all may be time-dependent.

To discretize (9) and (10) regularized delta functions from [40] were used , i.e., $\delta_h(\mathbf{x})$,

$$
\delta_h(\mathbf{x}) = \frac{1}{h^3}\phi\left(\frac{x}{h}\right)\phi\left(\frac{y}{h}\right)\phi\left(\frac{z}{h}\right), \tag{16}
$$

where $\phi(r)$ is defined as

$$
\phi(r) = \begin{cases} \frac{1}{8}(3 - 2|r| + \sqrt{1 + 4|r| - 4r^2}), & 0 \le |r| < 1 \\ \frac{1}{8}(5 - 2|r| + \sqrt{-7 + 12|r| - 4r^2}), & 1 \le |r| < 2 \\ 0 & 2 \le |r|. \end{cases} \tag{17}
$$