





Actuation response model from sparse data for wall turbulence drag reduction

Daniel Fernex ^{1,*}, Richard Semaan ¹, Marian Albers ², Pascal S. Meysonnat,²
Wolfgang Schröder,^{2,3} and Bernd R. Noack ^{4,5,1}

¹*Institut für Strömungsmechanik, Technische Universität Braunschweig,
Hermann-Blenk-Str. 37, 38108 Braunschweig, Germany*

²*Institute of Aerodynamics, RWTH Aachen University, Willnerstr. 5a, 52062 Aachen, Germany*

³*JARA - High-Performance Computing, Forschungszentrum Jülich, 52425 Jülich, Germany*

⁴*Institut für Strömungsmechanik und Technische Akustik (ISTA), Technische Universität Berlin,
Müller-Breslau-Straße 8, 10623 Berlin, Germany*

⁵*Institute for Turbulence-Noise-Vibration Interactions and Control, Harbin Institute of Technology,
58800 Shenzhen, China*



(Received 17 September 2019; accepted 28 May 2020; published 2 July 2020)

We compute, model, and predict drag reduction of an actuated turbulent boundary layer at a momentum-thickness-based Reynolds number of $Re_\theta = 1000$. The actuation is performed using spanwise traveling transversal surface waves parametrized by wavelength, amplitude, and period. The drag reduction for the set of actuation parameters is modeled using 71 large-eddy simulations (LESs). This drag model allows us to extrapolate outside the actuation domain for larger wavelengths and amplitudes. The modeling novelty is based on combining support vector regression for interpolation, a parametrized ridgeline leading out of the data domain, a scaling for the drag reduction, and a discovered self-similar structure of the actuation effect. The model yields high prediction accuracy outside the training data range.

DOI: [10.1103/PhysRevFluids.5.073901](https://doi.org/10.1103/PhysRevFluids.5.073901)

I. INTRODUCTION

The skin friction associated with a turbulent boundary layer constitutes about 50% of the total drag of an airplane. Owing to its importance, passive or active skin-friction reduction means have widely been investigated [1]. Promising strategies include riblets [2], compliant surfaces [3], spanwise wall oscillations and similar variations [4,5], and spanwise traveling waves of spanwise forcing [6] or wall-normal deflection [7–9]. To determine the optimal actuation settings, the parameter space is typically scanned by performing a large number of numerical simulations, which is very costly and sometimes untractable, e.g., for high Reynolds number or large actuation wavelength. This reliance on numerical simulations is partially due to the fact that experiments for many of these actuation concepts are either currently unrealizable or are limited by design to a small actuation range. In this study, a methodology is developed to model sparse flow response data to spanwise traveling surface waves using a machine learning regression algorithm for interpolation and a ridgeline modeling for extrapolation and optimization, which reduces the necessity for a large parametric study. Investigations on the boundary layer response sensitivities show a self-similar response behavior starting at a certain wavelength.

*d.fernex@tu-braunschweig.de

Actuation employing a spanwise pressure gradient has been shown to attenuate the boundary layer low-speed streaks and reduce turbulence production. This principle was first put into practice by a spanwise wall oscillation for turbulent channel flows [4,10] and for turbulent boundary layers [11]. The actuation generates a thin spanwise Stokes layer and reduces the wall-shear stress. A more efficient variant of the spanwise actuation is the streamwise traveling wave of spanwise forcing [5], where a maximum drag reduction of 48% can be achieved for turbulent channel flows. Using the same actuation technique Gatti and Quadrio [12] performed a comprehensive study of 4020 direct numerical simulations of a channel flow with varying wave number, amplitude, frequency, and Reynolds number. Such a large parameter study using high-fidelity simulations is unusual and computationally very expensive. The results showed that, for channel flows, the drag reduction at higher Reynolds numbers can be estimated using the vertical shift of the logarithmic velocity profile.

Another actuation variant employs the transversal traveling wave. A first implementation was conducted for a channel flow by Du and Karniadakis [6], where the wave effect was generated with a Lorentz force. To enable real-life applications, more recent studies proposed a similar traveling-wave effect by means of surface deformation. This approach has been experimentally tested for a turbulent boundary layer by Li *et al.* [13], where the surface was deflected using electromagnetic actuators. They achieved a drag reduction of 4.5%. Spanwise traveling transversal surface waves have also been numerically simulated for a turbulent boundary layer over a flat plate [9], with a maximum drag reduction of 26%, and over a wing section [8], where the pressure varies in the streamwise direction. For the wing flow, the total drag was reduced by 7.5 % and a slight lift increase was also achieved. It is worth noting that all these aforementioned surface actuation techniques, including the currently used one, are still investigated at a low technology readiness level. Several conceptual, as well as practical implementation hurdles, remain to be overcome. One important aspect is the net energy savings, which was addressed by Albers *et al.* [9], where they have shown a positive net power saving for several actuation settings.

Drag reduction optimization in a rich actuation space constitutes a challenge. In experimental setups, many degrees of freedom are fixed by design, whereas high-fidelity high Reynolds number simulations are able to explore a rich spectrum of actuation settings. However, high-fidelity numerical computations are costly and thus limited to a small number of control laws. Surrogate models are computationally cheap models that approximate the behavior of complex systems based on a limited number of data. Surrogate models are typically used for optimization [14,15] and for visualization and design space analysis [16]. There exists many approaches and algorithms. The response surface methodology (RSM) is one of the earliest approaches [17]. The models from RSM are often polynomials up to second order [18–21], which cannot represent highly nonlinear or multimodal design landscapes. Radial basis functions (RBFs) is an interpolation technique based on a weighted sum of radial basis functions [22]. Various types of basis functions can be used to accommodate the response distribution complexity [14]. RBF surrogate models have successfully been used to optimize groundwater remediation design [23]. Another widely used technique is kriging, which is a kernel-based probabilistic approach [24]. Kriging models can yield high predictive power [14] and have been used, for instance, to optimize a two-dimensional airfoil geometry [25].

Other computationally more expensive algorithms include support vector regression (SVR) [26] and artificial neural networks (ANNs), first developed by McCulloch and Pitts [27]. SVR is a kernel-based regression technique, which tolerates prediction errors within a user-defined interval. More details about SVR are given in Sec. III B. SVR has been shown to outperform RSM, kriging, and RBF for a test bed of 26 complex engineering functions [28], and has successfully been used to optimize railway wind barriers [29]. ANNs are nonlinear regression models inspired by biological neural networks. They have been used to accurately predict the drag reduction in oil pipelines [30].

A common shortcoming of the aforementioned as well as most data-driven surrogate modeling algorithms is their rapidly diminishing accuracy outside the training parameter range. This limitation means strong disadvantages for the investigated boundary layer application. Initial analyses have

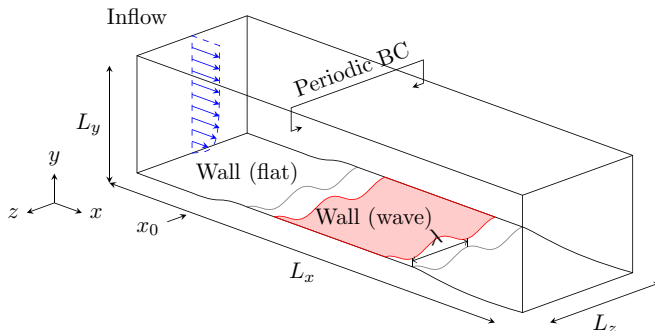


FIG. 1. Overview of the physical domain of the actuated turbulent boundary layer flow, where L_x , L_y , and L_z are the domain dimensions in the Cartesian directions, λ is the wavelength of the spanwise traveling wave, and x_0 marks the actuation onset. The shaded red surface A_{surf} marks the integration area of the wall-shear stress τ_w .

shown a higher drag reduction trend leading beyond the training parameter space, where simulations become increasingly less affordable.

In this study, we present a modeling methodology capable of extrapolating drag reduction beyond the parameter range. The starting point is a sparse set of 71 large-eddy simulations (LESs) of a turbulent boundary layer actuated by spanwise traveling transversal waves. Our approach consists of two steps: First, a surrogate model is built using SVR to interpolate the drag reduction in the training parameter space. Then, extrapolation is enabled through the identification of a ridgeline in the drag reduction response. The model is used to analyze the actuation sensitivities and to infer higher drag reduction and the corresponding actuation settings.

The paper is structured as follows: The numerical method of the high-fidelity simulations and the computational setup of the flat plate undergoing transversal spanwise traveling waves are defined in Sec. II. The modeling approach is described in Sec. III for a simple problem, before being applied to the actuated boundary layer data in Sec. IV. Finally, conclusions are presented in Sec. V.

II. NUMERICAL SETUP

In this section, the open-loop actuation study of wall turbulence drag reduction is recapitulated. In particular, the investigated actuation parameters are enumerated. Section II A describes the configuration, while Sec. II B details the employed large-eddy simulation (LES) solver.

A. Configuration

The fluid flow is described in a Cartesian frame of reference where the streamwise, wall-normal, and spanwise coordinates are denoted by $\mathbf{x} = (x, y, z)^T$ and the velocity components by $\mathbf{u} = (u, v, w)^T$. The Mach number is set to $M = 0.1$ such that nearly incompressible flow is considered. An illustration of the rectangular physical domain is shown in Fig. 1. A momentum thickness of $\theta = 1$ at $x_0 = 0.0$ is achieved such that the momentum-thickness-based Reynolds number $Re_\theta = 1000$ at x_0 . The distance between the inflow plane and x_0 is about $\Delta x = 5.06\delta_{99}$ where δ_{99} is the boundary layer thickness at x_0 . The domain length and height in the streamwise and wall-normal direction are $L_x = 190\theta$ and $L_y = 105\theta$. In the spanwise direction, different domain widths $L_z \in [21.65\theta, 108.25\theta]$ are used to simulate different actuation wavelengths.

At the domain inlet, a synthetic turbulence generation method is applied to generate a natural turbulent boundary layer flow after a transition length of 2–4 boundary layer thicknesses [31]. That is, at x_0 the properties of the turbulent boundary layer are comparable to that of natural boundary layer flow or direct numerical simulation (DNS) of turbulent boundary layer flow in much

larger domains. A comparison between the boundary layer statistics of the nonactuated reference simulation and a DNS from the literature [32] can be found in Albers *et al.* [9]. Characteristic boundary conditions are used at the domain exit and a no-slip wall boundary condition is enforced at the lower domain boundary for the unactuated and actuated wall. The wall actuation is prescribed by the space- and time-dependent function

$$y_{\text{wall}}^+(z^+, t^+) = A^+ \cos\left(\frac{2\pi}{\lambda^+} z^+ - \frac{2\pi}{T^+} t^+\right) \quad (1)$$

in the interval $-5 \leq x/\theta \leq 140$. The quantities λ^+ , T^+ , and A^+ denote the wavelength, period, and amplitude in inner coordinates, i.e., the parameters are scaled by the viscosity ν and the friction velocity of the unactuated reference case u_τ^n . In the area just upstream and downstream of the wave-actuation region, a spatial transition is used from a flat plate to an actuated plate and vice versa [9]. In total, 71 actuation configurations with wavelength $\lambda^+ \in [500, 3000]$, period $T^+ \in [20, 120]$, and amplitude $A^+ \in [10, 78]$ are simulated. Two additional validation simulations are performed for $\lambda^+ = 5000$, $T^+ = 44$, and $A^+ = 92$ and for $\lambda^+ = 5000$, $T^+ = 44$, and $A^+ = 99$. The parameter range was chosen based on our own experience and on that of others from the literature. It covers a reasonable span of what is experimentally and numerically feasible. For example, very large wavelengths ($\lambda^+ > 5000$) are not considered because of their high numerical cost. Cases that change the fundamental principle of the employed actuation (i.e., $A^+ = 0$ or $T^+ = 0$) are excluded. All operating conditions and the corresponding drag reductions are listed in Table II. The current setup is identical to that in Ishar *et al.* [33]. However, a considerably larger parameter set is computed in this study, containing also larger wavelengths.

The physical domain is discretized by a structured block-type mesh with a resolution of $\Delta x^+ = 12.0$ in the streamwise and $\Delta z^+ = 4.0$ in the spanwise direction. In the wall-normal direction, a resolution of $\Delta y^+|_{\text{wall}} = 1.0$ at the wall is used with gradual coarsening away from the wall. This yields a DNS-like resolution near the wall. Away from the wall, the resolution requirements are lower such that, overall, a LES resolution is achieved. Depending on the domain width, the meshes consist of 24 to 120 million cells.

The simulation procedure is as follows: First, the reference simulations for all domain widths are run for $tu_\infty/\theta = 650$ convective time units. Then, the actuated simulations are initialized by the solution from the unactuated reference case and the temporal transition from the flat plate to the actuated wall is initiated. When a converged state of the friction drag is obtained, statistics are collected for $tu_\infty/\theta = 1250$ convective times.

The drag coefficient c_d is computed by integrating the wall-shear stress over the streamwise interval $50 \leq x/\theta \leq 100$ and over the entire domain spanwise width, i.e., the colored surface in Fig. 1:

$$c_d = \frac{2}{\rho_\infty u_\infty^2 A_{\text{ref}}} \mathbf{e}_x \cdot \int_{A_{\text{surf}}} \bar{\bar{\tau}} \cdot \mathbf{n} dA.$$

The quantities \mathbf{n} , \mathbf{e}_x denote the unit normal vector of the surface and the unit vector in the x direction; the reference surface is $A_{\text{ref}} = 1$. The drag reduction is defined as

$$J = \Delta c_d = \frac{c_d^u - c_d^a}{c_d^u},$$

where the superscripts u and a refer to the unactuated reference and actuated cases.

B. Numerical method

The actuated flat plate turbulent boundary layer flow is governed by the unsteady compressible Navier-Stokes equations in the arbitrary Lagrangian-Eulerian formulation for time-dependent domains. A second-order accurate finite-volume approximation of the governing equations is used in which the convective fluxes are computed by the advection upstream splitting method (AUSM) and

time integration is performed via a five-stage Runge-Kutta scheme. The smallest dissipative scales are implicitly modeled through the numerical dissipation of the AUSM scheme. This monotonically integrated large-eddy simulation approach [34] is capable of accurately capturing all the physics of the resolved scales [35]. Further details on the numerical method can be found in Albers *et al.* [8] and Ishar *et al.* [33].

III. METHODOLOGY

In this section, we propose a data-driven response surface methodology for interpolation and extrapolation. The methodology is developed to handle the observed relative drag reduction sensitivities $J = \Delta c_D$. We note that J is *positive* for *reduced* drag. Initial analyses indicate that, for the spanwise traveling wave, in every λ^+ plane, the drag reduction Δc_d features a single global maximum (T_r^+, A_r^+) with respect to the actuation period T^+ and amplitude A^+ of the current spanwise traveling-wave type. The curve of (λ^+, T^+, A^+) connecting all these λ^+ -dependent Δc_d maxima is the *ridgeline*, denoted by the subscript r . These optimal amplitude and period (T_r^+, A_r^+) increase with λ^+ beyond the currently simulated parameter range. Such a response behavior is challenging for optimization and requires specially developed tools.

Section III A introduces an analytical example which features similar topology to the drag reduction response distribution. The machine learning algorithm used to interpolate the response within the parameter range is detailed in Sec. III B. In Sec. III C, a data-driven modeling approach is proposed and exemplified for the analytical example.

A. Analytical response surface

To sharpen our data-driven tools, we start with an analytic response function $J(\mathbf{p})$ that we construct to behave qualitatively similar to the drag reduction problem. From the parameter vector $\mathbf{p} = (p, q, s)$, p mimics the wavelength, q the period, and s the amplitude. The analytical example

$$J(\mathbf{p}) = \underbrace{\tanh(1+p)}_{=:G(p)} \underbrace{\exp[-(1-2q/\sqrt{p})^2 - (1-2s/\sqrt{p})^2]}_{=:F_p(q,s)} \quad (2)$$

is investigated in the domain

$$\Omega := [0, 1] \times [0, 1] \times [0, 1]. \quad (3)$$

The function J factors into one monotonically increasing term $G(p)$ and one p -independent monomodal term $F_p(q, s)$ with a single maximum in (q, s) . At a given p , J assumes the maximum $J_p := \tanh(1+p)$ on the ridgeline $q_r = \frac{1}{2}\sqrt{p}$, $s_r = \frac{1}{2}\sqrt{p}$. The ridgeline marks the maxima of q and s for constant p . The global maximum $J_{\max} = \tanh(2)$ in the domain Ω is at the boundary $\mathbf{p}_{\max} = (1, 0.5, 0.5)$. These trends are observed in Fig. 2, which illustrates the analytical response surface as isosurfaces, the ridgeline (black), and the lines of steepest ascent (red). The lines of steepest ascent provide a direct indication to the response sensitivities and point in the direction of the global optima. The lines of steepest ascent are simply streamlines of the gradient field $\nabla J = (\frac{\partial J}{\partial p}, \frac{\partial J}{\partial q}, \frac{\partial J}{\partial s})$ seeded from various points.

Evidently, larger J values are obtained outside the domain Ω on the ridgeline. Following this curve is a good extrapolation strategy for testing new and better parameters. The extrapolation to suboptimal parameters outside the domain $p > p_{\max}$ is facilitated by the self-similar structure of this particular response function. The response J can be parametrized by a p -dependent function multiplying a properly scaled (q, s) -dependent function:

$$J(p, q, s) = J_r(p)F(q^*, s^*) = J_r(p)F\left(\frac{q}{q_r}, \frac{s}{s_r}\right), \quad (4)$$

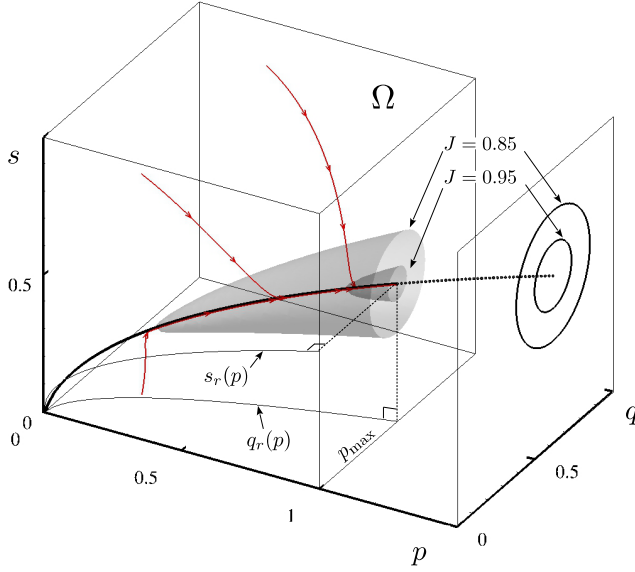


FIG. 2. Analytical response surface of Eq. (2). Also shown is the ridgeline (black) in the interpolated (solid line) and the extrapolated (dotted line) regimes. The red lines denote lines of steepest ascent seeded at various domain locations.

where

$$J_r(p) = \tanh(1 + p), \quad q_r = \frac{1}{2}\sqrt{p}, \quad s_r = \frac{1}{2}\sqrt{p}, \quad (5a)$$

$$F(q^*, s^*) = \exp[-(1 - q^*)^2 - (1 - s^*)^2]. \quad (5b)$$

Thus, knowing $J(p, q, s)$ in a plane $p = \text{const} \leq p_{\max}$ allows us to extrapolate all response functions J for $p > p_{\max}$ via Eq. (4).

B. Support vector regression

In this study, the analytical response formula from sparse data points is obtained by support vector regression (SVR) [26,36]. SVR belongs to the family of supervised-learning algorithms that trains from M observations to find a mapping between N features or inputs $\mathbf{x}_m = [x_m^1, x_m^2, \dots, x_m^N]$, and the corresponding response y_m , $m = 1, \dots, M$. In the application presented in Sec. IV, the features are the wavelength λ^+ , period T^+ , and amplitude A^+ and the output is the relative drag reduction Δc_d .

Following good practices of machine learning [37], the inputs for the response formula are centered features which are normalized to unit variance. This normalization gives every feature a similar weight in interpolation. In this study, the normalization is particularly important as the ranges of investigated wavelengths and periods differ by more than one order of magnitude.

SVR yields a regression model $\hat{J}(\mathbf{x})$ smoothly interpolating from data points (\mathbf{x}_m, y_m) , $m = 1, \dots, M$, employing a Gaussian kernel $K(\mathbf{x}, \mathbf{x}_m)$ and optimized weights ω_m :

$$\hat{J}(\mathbf{x}) = \mu + \sum_{m=1}^M \omega_m K(\mathbf{x}, \mathbf{x}_m) = \mu + \boldsymbol{\omega}^T \mathbf{K}(\mathbf{x}). \quad (6)$$

Here, μ is a constant to which \hat{J} converges far away from the data points, $\boldsymbol{\omega}^T = [\omega_1, \omega_2, \dots, \omega_M]$ denotes the weight vector, and $\mathbf{K}^T = [K(\mathbf{x}, \mathbf{x}_1), K(\mathbf{x}, \mathbf{x}_2), \dots, K(\mathbf{x}, \mathbf{x}_M)]$ comprises the Gaussian kernel functions.

Calibrating the response model (6) for $\hat{J}(\mathbf{x}_m) = y_m$, $m = 1, \dots, M$ leads to m linear equations for m weights ω_m . Under generic conditions, such a linear system can be solved and the formula will exactly reproduce the input data. Yet, this vanishing in-sample error may come at the price of overfitting. Noise may be incorporated as a data feature, thus leading to an unphysical model complexity. The over-fitted model may amplify noise outside the training data, implying a large generalization error or, equivalently, a large out-of-sample error.

To account for noise and new data points, an error of ε is tolerated, i.e., a prediction $|\hat{J}(\mathbf{x}_m) - y_m| < \varepsilon$ is accepted. Complexity is characterized and penalized by the vector norm $\|\boldsymbol{\omega}\|^2$. This leads to the regularized optimization problem

$$\min \frac{1}{2} \|\boldsymbol{\omega}\|^2, \quad \text{subject to } |y_m - \boldsymbol{\omega}^T \mathbf{K} - \mu| \leq \varepsilon. \quad (7)$$

However, weights $\boldsymbol{\omega}$ which satisfy the ε constraint at all points (\mathbf{x}_m, y_m) might not exist, particularly for the validation data. This constraint is relaxed by introducing so-called *slack variables* ξ_m^+ in case $\hat{J}(\mathbf{x}_m) - y_m > \varepsilon$ and ξ_m^- if $y_m - \hat{J}(\mathbf{x}_m) < \varepsilon$. The slack variables extend the permissible ε interval for $\hat{J}(\mathbf{x}_m)$ to $[y_m - \varepsilon - \xi_m^-, y_m + \varepsilon + \xi_m^+]$. Now, the relaxed regularized optimization problem becomes

$$\begin{aligned} \min \quad & \frac{1}{2} \|\boldsymbol{\omega}\|^2 + C \frac{1}{M} \sum_{m=1}^M (\xi_m^+ + \xi_m^-) \quad \text{subject to } y_m - \boldsymbol{\omega}^T \mathbf{K} - \mu \leq \varepsilon + \xi_m^+, \\ & \boldsymbol{\omega}^T \mathbf{K} + \mu - y_m \leq \varepsilon + \xi_m^-, \quad \xi_m^+, \xi_m^- \geq 0. \end{aligned} \quad (8)$$

The trade-off between model complexity and errors beyond the ε limit is controlled by the penalty parameter C . The extreme choice $C = 0$ leads to unpenalized, arbitrarily large slack variables and minimal complexity since the minimization solely focuses on $\frac{1}{2} \|\boldsymbol{\omega}\|^2$. In other words, $\hat{J} \equiv \mu$ is a constant function. For sufficiently large C , the accuracy of the response model is optimized tolerating maximum complexity. The value of ε is set to the data-noise level, if available. Note that too large a value will decrease the prediction accuracy.

The interpolation is performed with the radial basis function

$$K(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{|\mathbf{x} - \mathbf{x}'|^2}{\sigma^2}\right). \quad (9)$$

The reader can refer to Ref. [14] for more details about the SVR formulation and the solution of the constrained optimization problem (8).

C. Data-driven response surface

The analytical example preludes our data-driven approach for the actuated turbulent boundary layer. The approach consists of the following steps:

Step 1: We consider M computed response function values J_m for parameter points \mathbf{p}_m , $m = 1, \dots, M$, covering well the parameter range of interest Ω . Each parameter may need to be centered and scaled to unit variance for the regression problem.

Step 2: Interpolate all function values in the domain with an accurate and smooth machine learning regression.

Step 3: Apply a gradient search technique for several initial conditions. If the corresponding steepest ascent curves converge to a point inside the domain, the purpose of a response surface model is served.

Step 4: Identify the ridgeline coordinates $(q_r(p), s_r(p))$ and response $J_r(p)$ leading out of the domain Ω , and model them using simple functions. This simple model can now be used to

extrapolate the ridgeline outside Ω towards the global response optimum. Note the choice of the parametrizing ridgeline parameter (here, p) is problem dependent.

Step 5: In some cases, like in the example (2), the response function J exhibits self-similar behavior and can be expressed as a p -dependent function multiplying a scaled (q, s) -dependent function as

$$\hat{J}(p, q, s) = J_r(p)F\left(\frac{q}{q_r(p)}, \frac{s}{s_r(p)}\right),$$

where J_r is the ridgeline response, and F is a shape function with the maximum $F(1, 1) = 1$.

In this case, the shape function and extrapolated ridgeline can be used to predict the response \hat{J} to parameter inputs away from the ridgeline.

Note that the parameters p , q , and s used in this analytical example correspond to the wavelength λ^+ , period T^+ , and amplitude A^+ for the boundary layer application.

IV. RESULTS

The previous section discussed the modeling methodology of response functions, which assumes the optimal value at the boundary of the explored parameter space. This section applies the approach to drag reduction for an actuated boundary layer with spanwise traveling surface waves. We follow the steps outlined in Sec. III.

The process begins with the interpolation of the sparse parameter space using support vector regression (SVR). As presented in Sec. II, the investigated parameter space spanned by λ^+ , T^+ , and A^+ is large, and a dense coverage is beyond the reach of feasibility.

The SVR algorithm is chosen for its prediction accuracy and its smooth response distribution (see Appendix A for details). The algorithm is trained on a subset of 80% of the dataset, while the remaining 20% is used for testing the prediction performance. This separation of training and testing data reduces the risk of overfitting. The algorithm hyperparameters are tuned using three-fold cross-validation. In this study, the SVR model yields $R^2 = 0.93$, which indicates excellent prediction accuracy. The definitions of k -fold cross-validation and R^2 are given in Appendix A. Using the SVR model, we interpolate the parameter space with drag coefficient predictions.

With the parameter space densely populated, it is now possible to compute and to visualize the streamlines of the gradient field and the ridgeline. It is shown in Fig. 3(d) that the streamlines (red) and the ridgeline (solid black line) terminate at the domain boundary in the $T^+ - \lambda^+$ plane at the exit point A ($\lambda^+ = 1875$, $T^+ = 44$, and $A^+ = 78$). This indicates that the optimal drag reduction lies outside the current range. Along the ridgeline, the relative drag reduction increases from $\Delta c_d = 7.0\%$ at $\lambda^+ = 500$ to $\Delta c_d = 22.5\%$ at the ridgeline exit point A.

Figure 4 shows the projection of the ridgeline onto the $\lambda^+ - T^+$ and $\lambda^+ - A^+$ planes. Starting at $\lambda^+ \approx 1000$, T_r^+ asymptotes rapidly toward 44. In other words, the optimum wave period remains constant at $T^+ = 44$, even with increasing wavelength and amplitude. Similarly, A_r^+ shows an asymptotic behavior with higher λ^+ , albeit at a slower rate.

This asymptotic behavior of the ridgeline starting at $\lambda^+ = 1000$ is easily modeled as

$$\begin{aligned} T_r^+ &= 44 - 46721 \exp(-0.0128\lambda^+), \\ A_r^+ &= 100 - 113 \exp(-0.0009\lambda^+). \end{aligned} \quad (10)$$

The fitted curves are presented in Fig. 4 by dotted lines and show good agreement with the reference lines over the common range ($1000 \leq \lambda^+ \leq 1875$).

Having established the T_r^+ and A_r^+ sole dependence on λ^+ along the ridgeline, we turn our attention to drag reduction. Similarly to T^+ and A^+ , J_r also shows the sole dependence on λ^+ or equivalently on T_r^+ and A_r^+ . This is best expressed using a scaling proposed by Tomiyama and Fukagata [38], defined as $A_r^+ \sqrt{2\pi/T_r^+}$, which is the product of the velocity amplitude of the actuation $2\pi A_r^+/T_r^+$ and the thickness of the Stokes layer $\sqrt{T_r^+/(2\pi)}$ along the ridge. Note that

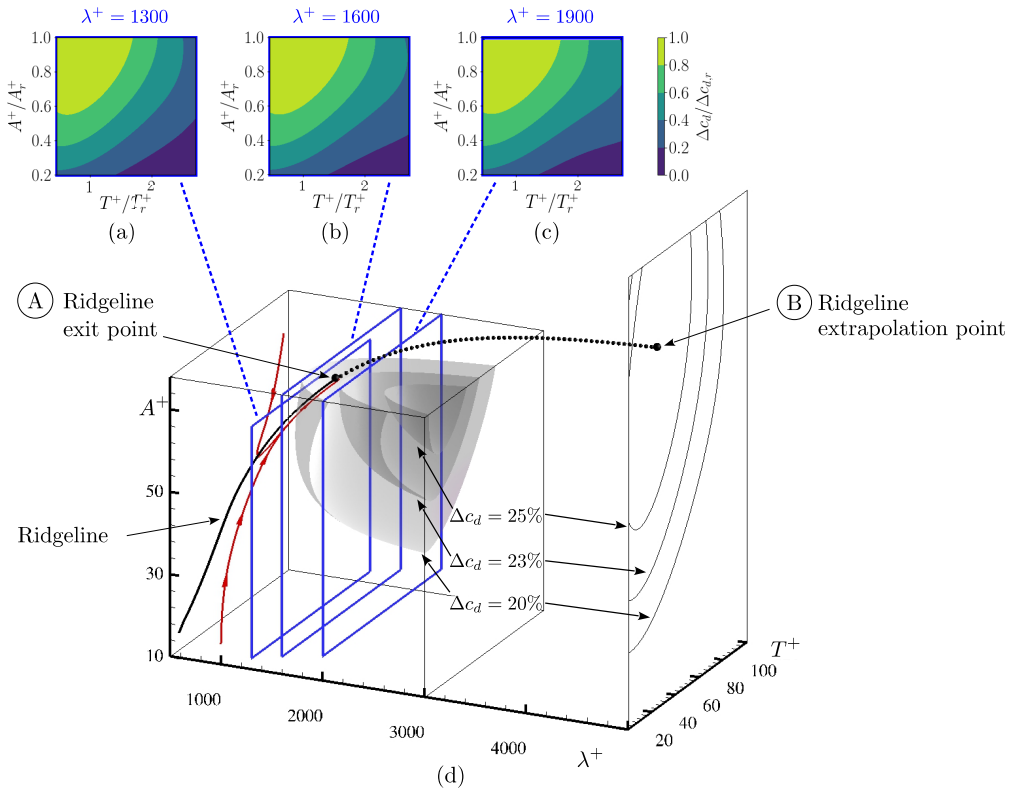


FIG. 3. (d) The drag reduction model $\widehat{\Delta c_D}(\lambda^+, T^+, A^+)$. The gray surfaces represent three drag reduction levels: 20%, 23%, and 25%. The ridgeline (black) is displayed in the interpolated (solid line) and the extrapolated (dotted line) regimes. The ridgeline leaves the investigated domain at point A and predicts the optimal drag reduction for $\lambda^+ \leq 5000$ at point B ($\lambda^+ = 5000$, $T^+ = 44$, and $A^+ = 99$). The red curves denote lines of steepest ascent seeded at various domain locations. The contour distributions at the top [panels (a)–(c)] represent scaled-drag reductions. These reductions are the shape functions F defined in Eq. (13) and modeled in Eq. (14). The three slices correspond to the blue-framed $T^+ - A^+$ rectangles in the three-dimensional plot.

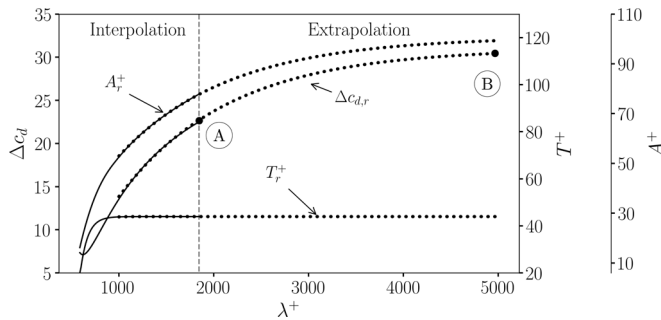


FIG. 4. Projection of the ridgeline onto the $\lambda^+ - T^+$ and $\lambda^+ - A^+$ planes, as well as the drag reduction along the ridge as function of λ^+ . The solid lines are interpolated with SVR, whereas the dotted lines are obtained by Eqs. (10) (for T_r^+ and A_r^+), and (11) (for $\Delta c_{d,r}$). Points A and B are the same as those in Fig. 3. The vertical gray dashed line separates the interpolation and extrapolation regions.

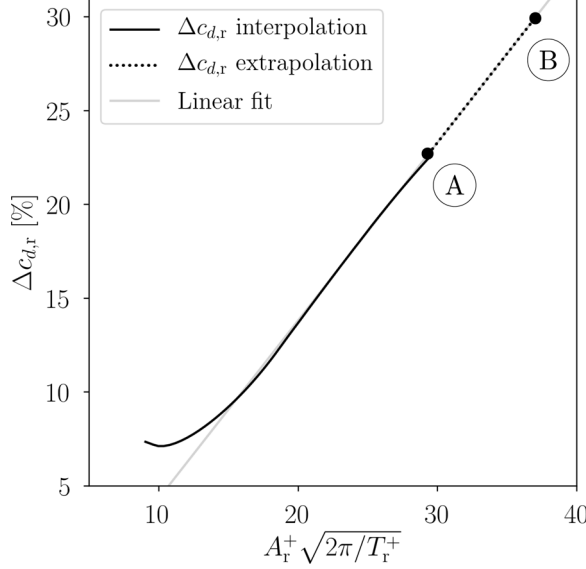


FIG. 5. Drag reduction along the ridgeline as function of the Tomiyama and Fukagata scaling. The figure shows a linear behavior starting at $\lambda^+ \approx 1000$. The solid line is obtained from data interpolated with SVR. The dotted line is obtained with Eq. (11). Points A and B are the same as those in Fig. 3.

this scaling is originally defined for the skin-friction coefficient, but for the considered cases, the amount of added wetted surface is negligible and the scaling holds [9]. The wetted surface is the area that interacts with the working fluid. The data in Appendix B show the change of the wetted surface ΔA_{surf} with different actuation settings. The evolution of $\Delta c_{d,r}$ towards a linear behavior is illustrated in Fig. 5. As the figure shows, the drag reduction along the ridge starts exhibiting linearity around $\lambda^+ \approx 1000$ corresponding approximately to $A^+ \sqrt{2\pi/T_r^+} = 19$. It is worth to note that this almost perfect linear Tomiyama and Fukagata scaling only holds along the ridgeline. Away from the ridgeline, the scaling shows scatter. It is now straightforward to model the relative drag reduction in the linear range, i.e., $\lambda^+ \geq 1000$, as

$$\widehat{\Delta c_{d,r}} = 0.95 A_r^+ \sqrt{\frac{2\pi}{T_r^+}} - 5.16. \quad (11)$$

This linear fit is shown with a dotted line in Fig. 5 as function of the Tomiyama and Fukagata scaling, and in Fig. 4 as a function of λ^+ . Note that we assume that the linear behavior continues for a finite range beyond $\lambda^+ = 1875$.

Based on the optimal drag reduction behavior being only dependent on λ^+ , which is consistent with a self-similar behavior, we assume a response of the form

$$\hat{J} = \widehat{\Delta c_d}(\lambda^+, T^+, A^*) = J_r(\lambda^+) F(T^*, A^*), \quad (12)$$

where $\hat{J}_r = \widehat{\Delta c_{d,r}}$ is the constant-linear model from substituting Eq. 10 into Eq. 11, and T^* and A^* are properly scaled actuation parameters. The natural scaling choice is the maximum relative drag reduction along the ridgeline, which yields

$$F(T^*, A^*) = F\left(\frac{T^+}{T_r^+}, \frac{A^+}{A_r^+}\right) = \frac{J(\lambda^+, T^+, A^+)}{J_r(\lambda^+)}. \quad (13)$$

Hence, self-similarity is validated when F becomes independent of λ^+ . This is confirmed in Figs. 3(a)–3(c), where the F distributions collapse starting at $\lambda^+ \approx 1000$. The smooth distribution

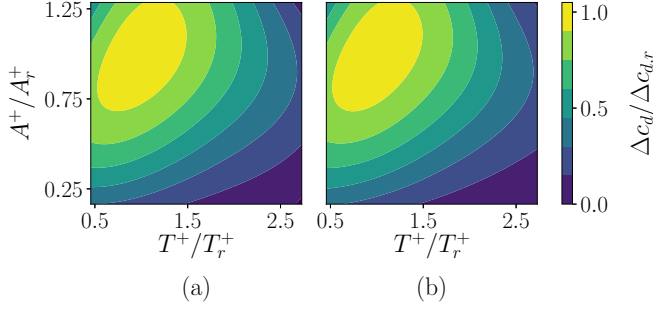


FIG. 6. (a) The reference F distribution, and (b) the nonlinear model \hat{F} . The model faithfully represents the reference response.

F can also be analytically modeled. In this study, we use a sparsified multivariate regression method with a predefined library of functions to identify a model. As detailed in Appendix C, the method yields a set of possible candidate nonlinear models with increasing accuracy and complexity. The selected model \hat{F} reads

$$\begin{aligned} \hat{F} = & 1.82A^* - 0.7T^*A^* + -4.34A^{*2} + 2.0T^*A^{*2} - 0.79T^*A^{*3} + 0.84A^{*4} \\ & + (0.36 - 0.52T^* + 5.39A^*f - 1.59T^*A^* - 4.79A^{*2} + 0.79T^{*2}A^* + 2.53A^{*3})P(T^*, A^*), \end{aligned} \quad (14)$$

where $T^* = T^+ / T_r^+$, $A^* = A^+ / A_r^+$, and

$$P(T^*, A^*) = \frac{1}{\sigma_T \sigma_A 2\pi} \exp \left\{ -0.5 \left[\left(\frac{T^* - 1}{\sigma_T} \right)^2 + \left(\frac{A^* - 1}{\sigma_A} \right)^2 \right] \right\},$$

with $\sigma_T = 0.9$ and $\sigma_A = 0.5$. Figure 6(a) shows the reference F distribution along with the nonlinear model \hat{F} [Fig. 6(b)]. As can be clearly seen, the model duplicates the reference distribution very accurately. Besides yielding an analytical expression, the model \hat{F} offers the possibility to extrapolate (within reason) beyond the current $A^+ - T^+$ range.

Note that the preceding analysis not only examined the sensitivities of the flow response and its self-similar behavior but also yielded a simple powerful model of the relative drag reductions.

This self-similar drag reduction model for $\lambda^+ \geq 1000$ proceeds as follows:

- (i) For given actuation setting λ^+ , T^+ and A^+ , compute T_r and A_r from Eq. (10).
- (ii) Determine the drag reduction along the ridgeline $J_r = \Delta c_{d,r}(\lambda^+)$ using Eq. (11).
- (iii) Compute $F(\frac{T^+}{T_r^+}, \frac{A^+}{A_r^+})$ from Eq. (14) or read it from the distributions in Fig. 3.
- (iv) Deduce the relative drag reduction from $J = \Delta c_d = J_r F$.

In the interpolation regime, and for $1000 \leq \lambda^+ \leq 3000$, this simple model has a coefficient of determination of $R^2 = 0.92$, which is very close to that of the SVR model. In the extrapolation regime, the model is validated with two points at $\lambda^+ = 5000$, which is well beyond the training range. The first validation point B is situated on the ridgeline (cf. Fig. 3), whereas the second point B' is off the ridgeline at coordinates $\lambda^+ = 5000$, $T^+ = 44$, and $A^+ = 92$. For these two operating conditions, the relative drag reductions predicted by the model are $\widehat{\Delta c_d} = 30.45\%$ and $\Delta c_d = 30.23\%$, which compare favorably with those of the reference LES data of $\Delta c_d = 31.09\%$ and $\Delta c_d = 30.03\%$. These predictions yield relative errors of 2.1% and 0.7% for B and B' , respectively.

The prediction accuracy for the extrapolation at $\lambda^+ = 5000$, i.e., 67% beyond the maximal investigated values $\lambda^+ = 3000$ is impressive. Yet, the model (12) should not be assumed to hold at much larger wavelengths ($\lambda^+ \rightarrow \infty$). In this limit, the actuation approaches that of a flat plate

moving up and down without height variations in the spanwise direction. In this scenario, the boundary layer remains unchanged and no drag reductions can be expected.

V. CONCLUSIONS

We target improved drag reduction of an actuated turbulent boundary layer with spanwise traveling surface waves at $Re_\theta = 1000$. Seventy-one large-eddy simulations are used to determine a machine learned model to predict drag reduction as a function of the actuation parameters: amplitude, period, and spanwise wavelength. The first enabler for this formula is the support vector regression (SVR) for smooth interpolation. For this dataset, SVR is found to be distinctly superior to many other common regression solvers. The second enabler is a ridgeline pointing outside the computed domain indicating further drag reduction potential at unexplored higher wavelengths. This ridgeline is then modeled and used for extrapolation. The results indicate a potential around 31% drag reduction with increasing wavelength, which is denoted as point B in Fig. 3. This result is confirmed by an additional LES. The corresponding period seems to asymptote against 44 plus units while the amplitude slowly increases with wave number. The ridgeline parameters are consistent with the Tomiyama and Fukagata scaling. More precisely, at wavelengths above 1000 plus units within the analyzed range, the drag reduction linearly increases with the Tomiyama and Fukagata parameter on the ridgeline.

Surprisingly, the drag reduction formula exhibits a self-similar behavior starting at $\lambda^+ \approx 1000$. As such, drag reduction can be expressed as the product of a factor depending only on the wavelength and a shape factor depending on amplitude and period normalized with their ridgeline values. The ridgeline parameters and the drag reduction values in a plane with constant wavelength allow us to extrapolate drag values for amplitudes and periods for wavelengths above 1000 plus units. The self-similar drag reduction formula beautifully parametrizes all investigated simulations and allow us to predict further performance potential at unexplored larger wavelengths.

The proposed machine learning method for the drag reduction formula can easily be applied to other performance parametrics from sparse data. This includes more challenging problems with larger parameter space, albeit with increased difficulty. The strategy is (1) to interpolate the sparse parameter space using an accurate machine learning algorithm; (2) to compute several steepest ascent lines and ridgelines; (3) to search for the global optimum inside the domain; (4) if the steepest ascent lines terminate at the boundary to extrapolate the ridgeline out of the domain. We note the possibility of having multiple ridgelines exiting the domain. For such a scenario, we recommend tracking the most promising one, which is typically the ridgeline with the steepest ascent or descent; (5) to test for self-similarity based on this ridgeline. Self-similarity opens the possibility to extrapolate the performance away from the ridgeline.

The drag reduction formula may guide future simulations in search of larger drag reduction. In addition, the observed self-similarity guides and constrains future physics-based models. The authors are actively exploring these avenues.

ACKNOWLEDGMENTS

The research was funded by the Deutsche Forschungsgemeinschaft (DFG) in the framework of the research projects SE 2504/2-1, SCHR 309/52, and SCHR 309/68. The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time on the GCS Supercomputers Hazelhen at HLRS Stuttgart and JURECA at Jülich Supercomputing Centre (JSC). B.R.N. acknowledges support from the Agence Nationale de la Recherche (ANR) under Grant No. ANR-17-ASTR-0022 (FlowCon).

APPENDIX A: MACHINE LEARNING REGRESSION MODEL

Drag reduction modeling as a function of the actuation parameters for the actuated boundary layer is a challenging problem. The complexity of the response topology led to the utilization of

TABLE I. Comparison of the prediction accuracy R^2 and smoothness TV of the three tested machine learning algorithms.

Algorithm	R^2	TV
k NN	0.76	1.62×10^6
RF	0.97	1.60×10^6
SVR	0.93	1.31×10^6

machine learning (ML) approaches. For this application, ML is used to model the drag reduction Δc_d under varying actuation conditions (λ^+ , T^+ , and A^+). ML algorithms are evaluated based on their prediction accuracy, given by the coefficient of determination R^2 , defined as

$$R^2 = 1 - \frac{\sum_i^N (\Delta c_{d,i} - \widehat{\Delta c}_{d,i})^2}{\sum_i^N (\Delta c_{d,i} - \overline{\Delta c_d})^2}, \quad (\text{A1})$$

where $\Delta c_{d,i}$ are the reference computed data points, $\widehat{\Delta c}_{d,i}$ are the predicted ones, $\overline{\Delta c_d}$ is the mean of $\Delta c_{d,i}$, and N is the number of samples in the test set. A value of $R^2 = 1$ denotes perfect prediction score. Besides accuracy, the model smoothness is the second criterion for the model selection. The ML algorithm smoothness was quantified with the total variation (TV) defined as

$$y = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K \{(\widehat{\Delta c}_{d,i,j,k} - \widehat{\Delta c}_{d,i-1,j,k})^2 + (\widehat{\Delta c}_{d,i,j,k} - \widehat{\Delta c}_{d,i,j-1,k})^2 + (\widehat{\Delta c}_{d,i,j,k} - \widehat{\Delta c}_{d,i,j,k-1})^2\}^{1/2}, \quad (\text{A2})$$

where I , J , and K are the number of discretized points in the λ^+ , T^+ , and A^+ directions. A smooth response is indicated by a lower TV value. Three machine learning algorithms were benchmarked: the k -nearest neighbors (k NN), random forest (RF), and support vector regression (SVR). The hyperparameters of each algorithm were optimized using cross-validation, yielding five neighbors for k NN and 300 trees for RF. Cross-validation is a resampling procedure used to evaluate machine learning models. It involves randomly dividing the set of observations into k groups, or folds, of approximately equal size. The first fold is used for validation, whereas the model is fit on the remaining $k - 1$ folds. Radial basis functions are used for SVR. The R^2 and TV values for the three algorithms are summarized in Table I. Based on the results, SVR offers the best compromise between smoothness and accuracy; it is smoother than RF and more accurate than k NN. Therefore, it is selected for this study.

APPENDIX B: OPERATING CONDITIONS OF THE LARGE-EDDY SIMULATIONS

Table II lists the operation condition of the large-eddy simulation.

APPENDIX C: NONLINEAR MULTIVARIATE REGRESSION

This section describes the methodology to identify a sparse nonlinear model of the distribution F . The motivation behind promoting model sparsity is to reduce complexity and avoid overfitting. The current approach is comparable to the sparse identification of nonlinear dynamical systems (SINDY) method [39], but for a response surface. Practically, it is a nonlinear multivariate regression with a preselected library of functions. In the current application, we are concerned with identifying a sparse nonlinear model as a function of T^+ and A^+ . In other words, we seek to identify the following model

$$\mathbf{F} = \mathbf{F}(\mathbf{X}), \quad (\text{C1})$$

TABLE II. Actuation parameters of the turbulent boundary layer simulations, where each setup is denoted by a case number N . The quantity λ^+ is the spanwise wavelength of the traveling wave, T^+ is the period, and A^+ is the amplitude, all given in inner units, i.e., nondimensionalized by the kinematic viscosity ν and the friction velocity u_τ . Each block includes setups with varying period and amplitude for a constant wavelength. The list includes the values of the averaged relative drag reduction Δc_d , the averaged relative skin friction reduction $\Delta c_f = \frac{c_f^u - c_f^d}{c_f^u}$, and the relative increase of the wetted surface, ΔA_{surf} .

N	L_z^+	λ^+	T^+	A^+	Δc_d [%]	Δc_f [%]	ΔA_{surf} [%]
1	1000	500	20	30	0	4	3.5
2	1000	500	30	22	9	10	1.9
3	1000	500	40	21	8	9	1.7
4	1000	500	40	30	8	11	3.5
5	1000	500	60	30	5	8	3.5
6	1000	500	70	36	3	8	4.9
7	1000	500	70	64	-10	4	14.6
8	1000	500	100	48	-3	5	8.6
9	1000	1000	20	10	5	5	0.1
10	1000	1000	20	30	13	13	0.9
11	1000	1000	20	50	0	3	2.4
12	1000	1000	40	10	3	3	0.1
13	1000	1000	40	20	7	8	0.4
14	1000	1000	40	30	12	13	0.9
15	1000	1000	40	40	15	16	1.6
16	1000	1000	40	50	15	17	2.4
17	1000	1000	40	60	13	16	3.5
18	1000	1000	80	10	1	1	0.1
19	1000	1000	80	20	3	4	0.4
20	1000	1000	80	30	6	6	0.9
21	1000	1000	80	40	9	10	1.6
22	1000	1000	80	50	9	11	2.4
23	1000	1000	80	60	9	12	3.5
24	1000	1000	120	10	1	1	0.1
25	1000	1000	120	20	0	1	0.4
26	1000	1000	120	30	3	4	0.9
27	1000	1000	120	40	3	5	1.6
28	1000	1000	120	50	2	5	2.4
29	1000	1000	120	60	2	6	3.5
30	1200	600	30	44	2	7	5.1
31	1200	600	40	59	-4	5	8.9
32	1200	600	50	36	9	12	3.5
33	1200	600	60	21	5	6	1.2
34	1200	600	70	29	6	8	2.3
35	1200	600	80	66	-5	6	11.0
36	1200	600	90	51	-1	6	6.8
37	1200	600	100	14	2	2	0.5
38	1600	1600	20	22	11	11	0.2
39	1600	1600	40	34	14	14	0.4
40	1600	1600	40	48	19	19	0.9
41	1600	1600	50	60	19	20	1.4
42	1600	1600	50	73	21	22	2.0
43	1600	1600	60	27	8	8	0.3
44	1600	1600	70	71	17	19	1.9

TABLE II. (Continued.)

N	L_z^+	λ^+	T^+	A^+	Δc_d [%]	Δc_f [%]	ΔA_{surf} [%]
45	1600	1600	80	17	2	2	0.1
46	1600	1600	90	65	13	14	1.6
47	1600	1600	100	40	8	8	0.6
48	1800	900	30	49	10	12	2.9
49	1800	900	40	63	7	12	4.7
50	1800	900	50	22	7	7	0.6
51	1800	900	50	44	12	14	2.3
52	1800	900	70	28	7	8	0.9
53	1800	900	80	17	3	4	0.4
54	1800	900	80	60	6	9	4.3
55	1800	900	90	39	6	7	1.8
56	1800	1800	30	14	5	5	0.1
57	1800	1800	40	51	19	20	0.8
58	1800	1800	40	70	22	23	1.5
59	1800	1800	50	59	20	21	1.1
60	1800	1800	60	44	15	15	0.6
61	1800	1800	60	75	21	22	1.7
62	1800	1800	70	29	7	7	0.3
63	1800	1800	80	36	9	9	0.4
64	1800	1800	90	66	13	14	1.3
65	1800	1800	100	21	3	3	0.1
66	3000	3000	40	51	21	21	0.3
67	3000	3000	50	78	26	26	0.7
68	3000	3000	60	26	7	7	0.1
69	3000	3000	70	64	19	19	0.4
70	3000	3000	80	11	1	1	0.0
71	3000	3000	90	66	16	16	0.5
B	5000	5000	44	99	31	31	0.0
B'	5000	5000	44	92	30	30	0.0

where $\mathbf{F} = [F_1, F_2, \dots, F_m]^T$ are the m discrete points that constitute the distribution and

$$\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} T_1^* & A_1^* \\ T_2^* & A_2^* \\ \vdots & \vdots \\ T_m^* & A_m^* \end{bmatrix}, \quad (\text{C2})$$

are the normalized actuation settings, $T^* = T^+/T_r^+$, $A^* = A^+/A_r^+$. Modeling begins by constructing an augmented library $\Theta(\mathbf{X})$ consisting of candidate functions of the columns of \mathbf{X} . For example, $\Theta(\mathbf{X})$ may consist of constant, polynomial or trigonometric functions, such as

$$\Theta(\mathbf{X}) = \begin{bmatrix} | & | & | & | & | & \dots \\ 1 & \mathbf{X} & \mathbf{X}^{P_2} & \mathbf{X}^{P_3} & \sin \mathbf{X} & \dots \\ | & | & | & | & | & \dots \end{bmatrix}. \quad (\text{C3})$$

Here, \mathbf{X}^{P_i} denote higher polynomials. For example, \mathbf{X}^{P_2} denotes the quadratic nonlinearities in \mathbf{X} . For the current study, the highest polynomial order is set to four. We have also supplemented the

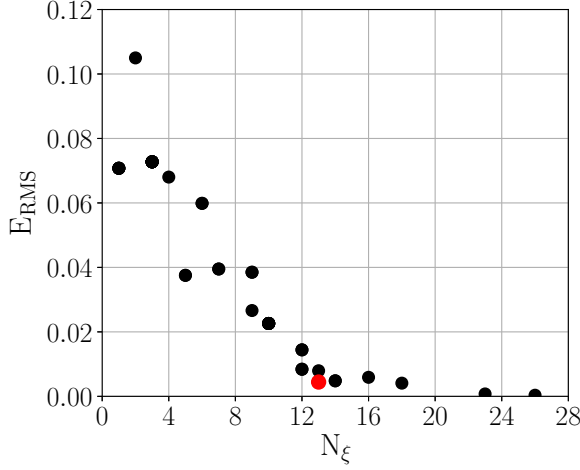


FIG. 7. The cross-validated mean square error for a range of models with increasing complexity N_ξ .

library of functions $\Theta(\mathbf{X})$ with the product of the polynomials with the exponential function

$$\mathbf{P} = P(T_i^*, A_i^*) = \frac{1}{\sigma_T \sigma_A 2\pi} \exp \left\{ -0.5 \left[\left(\frac{T_i^* - 1}{\sigma_T} \right)^2 + \left(\frac{A_i^* - 1}{\sigma_A} \right)^2 \right] \right\},$$

where $\sigma_T = 0.9$ and $\sigma_A = 0.5$, yielding

$$\Theta(\mathbf{X}) = \begin{bmatrix} | & | & | & | & | & | & | & | & | & | \\ 1 & \mathbf{X} & \mathbf{X}^{P_2} & \mathbf{X}^{P_3} & \mathbf{X}^{P_4} & \mathbf{P} & \mathbf{P}\mathbf{X} & \mathbf{P}\mathbf{X}^{P_2} & \mathbf{P}\mathbf{X}^{P_3} & \mathbf{P}\mathbf{X}^{P_4} \\ | & | & | & | & | & | & | & | & | & | \end{bmatrix}. \quad (\text{C4})$$

Each column of $\Theta(\mathbf{X})$ represents a candidate function for the model in Eq. (C1). There is a large number of possible entries in this matrix of nonlinearities. Since we are seeking a parsimonious model, only a few of these nonlinearities are active. Moreover, the simulated response contains random uncertainties. The solution to this overdetermined system with noise is obtained by solving

$$\mathbf{F} = \Theta(\mathbf{X})\boldsymbol{\Xi} + \eta\mathbf{Z}, \quad (\text{C5})$$

which yields the sparse vectors of coefficients $\boldsymbol{\Xi} = [\xi_1, \xi_2, \dots, \xi_n]$. Here, \mathbf{Z} represents white noise with vanishing mean and unit variance while η denotes noise gain. System (C5) can be easily solved using the LASSO [40], which is an ℓ_1 regression that promotes sparsity.

We emphasize that we seek the simplest and most accurate model. We define complexity N_ξ as the number of nonzero coefficients ξ_i , whereas accuracy is quantified as the root mean square error (E_{RMS}) between the model prediction and the reference data. As Fig. 7 shows, the root mean square error reduces with higher model complexity. In other words, as the model becomes more complex it becomes more accurate. However, higher complexity comes at the cost of overfitting and less interpretability. We select a model on the Pareto front as a compromise between complexity and accuracy. The red circle in Fig. 7 indicates the selected model, which possesses thirteen nonzero coefficients and a root mean square error of ≈ 0.005 . It reads

$$\begin{aligned} \hat{F} = & 1.82A^* - 0.77T^*A^* - 4.34A^{*2} + 2.0T^*A^{*2} - 0.79T^*A^{*3} + 0.84A^{*4} \\ & + (0.36 - 0.52T^* + 5.39A^*f - 1.59T^*A^* - 4.79A^{*2} + 0.79T^{*2}A^* + 2.53A^{*3})P(T^*, A^*). \end{aligned} \quad (\text{C6})$$

- [1] B. Fan and G. Dong, *Principles of Turbulence Control* (National Defense Industry Press with Wiley, Singapore, 2016).
- [2] M. J. Walsh and M. Lindemann, Optimization and application of riblets for turbulent drag reduction, in *AIAA 22nd Aerospace Sciences Meeting* (AIAA, Reston, VA, 1984).
- [3] M. Luhar, A. S. Sharma, and B. J. McKeon, On the design of optimal compliant walls for turbulence control, *J. Turbul.* **17**, 787 (2016).
- [4] W. J. Jung, N. Mangiavacchi, and R. Akhavan, Suppression of turbulence in wall-bounded flows by high-frequency spanwise oscillations, *Phys. Fluids A* **4**, 1605 (1992).
- [5] M. Quadrio, P. Ricco, and C. Viotti, Streamwise-travelling waves of spanwise wall velocity for turbulent drag reduction, *J. Fluid Mech.* **627**, 161 (2009).
- [6] Y. Du and G. Karniadakis, Suppressing wall turbulence by means of a transverse traveling wave, *Science* **288**, 1230 (2000).
- [7] S. Klumpp, M. Meinke, and W. Schröder, Friction drag variation via spanwise transversal surface waves, *Flow, Turbul. Combust.* **87**, 33 (2011).
- [8] M. Albers, P. S. Meysonnat, and W. Schröder, Actively reduced airfoil drag by transversal surface waves, *Flow, Turbul. Combust.* **102**, 865 (2019).
- [9] M. Albers, P. S. Meysonnat, D. Fernex, R. Semaan, B. R. Noack, and W. Schröder, Drag reduction and energy saving by spanwise traveling transversal surface waves for flat plate flow, *Flow, Turbul. Combust.* **105**, 125 (2020).
- [10] E. Toubert and M. A. Leschziner, Near-wall streak modification by spanwise oscillatory wall motion and drag-reduction mechanisms, *J. Fluid Mech.* **693**, 150 (2012).
- [11] S. Lardeau and M. A. Leschziner, The streamwise drag-reduction response of a boundary layer subjected to a sudden imposition of transverse oscillatory wall motion, *Phys. Fluids* **25**, 075109 (2013).
- [12] D. Gatti and M. Quadrio, Reynolds-number dependence of turbulent skin-friction drag reduction induced by spanwise forcing, *J. Fluid Mech.* **802**, 553 (2016).
- [13] W. Li, D. Roggenkamp, T. Hecken, W. Jessen, M. Klaas, and W. Schröder, Parametric investigation of friction drag reduction in turbulent flow over a flexible wall undergoing spanwise transversal traveling waves, *Exp. Fluids* **59**, 105 (2018).
- [14] A. I. J. Forrester and A. J. Keane, Recent advances in surrogate-based optimization, *Prog. Aerosp. Sci.* **45**, 50 (2009).
- [15] R. Yondo, E. Andrés, and E. Valero, A review on design of experiments and surrogate models in aircraft real-time and many-query aerodynamic analyses, *Prog. Aerosp. Sci.* **96**, 23 (2018).
- [16] C. M. E. Holden and A. J. Keane, Visualization methodologies in aircraft design, in *Proceedings of the 10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, 30 August–1 September 2004, Albany, New York*, AIAA Paper 2004-4449 (AIAA, Reston, VA, 2004).
- [17] G. E. P. Box and K. B. Wilson, On the experimental attainment of optimum conditions, *J. Roy. Stat. Soc. B Met.* **13**, 1 (1951).
- [18] N. E. Sevant, M. I. G. Bloor, and M. J. Wilson, Aerodynamic design of a flying wing using response surface methodology, *J. Aircr.* **37**, 562 (2000).
- [19] J. I. Madsen, W. Shyy, and R. T. Haftka, Response surface techniques for diffuser shape optimization, *AIAA J.* **38**, 1512 (2000).
- [20] J. Ahn, H.-J. Kim, O. Lee, and D.-H. Rho, Response surface method for airfoil design in transonic flow, *J. Aircr.* **38**, 231 (2001).
- [21] H. R. Karami, M. Keyhani, and D. Mowla, Experimental analysis of drag reduction in the pipelines with response surface methodology, *J. Pet. Sci. Eng.* **138**, 104 (2016).
- [22] D. S. Broomhead and D. Lowe, *Radial Basis Functions, Multi-Variable Functional Interpolation and Adaptive Networks*, Tech. Rep. RSRE-MEMO-4148 (Royal Signals and Radar Establishment Malvern, United Kingdom, 1988).
- [23] T. Akhtar and C. A. Shoemaker, Multi objective optimization of computationally expensive multi-modal functions with RBF surrogates and multi-rule selection, *J. Global Optim.* **64**, 17 (2016).
- [24] G. Matheron, Principles of geostatistics, *Econ. Geol. Bull. Soc. Econ. Geol.* **58**, 1246 (1963).

- [25] S. Jeong, M. Murayama, and K. Yamamoto, Efficient optimization design method using Kriging model, *J. Aircr.* **42**, 413 (2005).
- [26] H. Drucker, C. J. C. Burges, L. Kaufman, A. J. Smola, and V. Vapnik, Support vector regression machines, in *Advances in Neural Information Processing Systems* (MIT Press, Cambridge, 1997), pp. 155–161.
- [27] W. S. McCulloch and W. Pitts, A logical calculus of the ideas immanent in nervous activity, *Bull. Math. Biophys.* **5**, 115 (1943).
- [28] S. M. Clarke, J. H. Griebisch, and T. W. Simpson, Analysis of support vector regression for approximation of complex engineering analyses, *J. Mech. Des.* **127**, 1077 (2005).
- [29] H. Xiang, Y. Li, H. Liao, and C. Li, An adaptive surrogate model based on support vector regression and its application to the optimization of railway wind barriers, *Structural and Multidisciplinary Optimization* **55**, 701 (2017).
- [30] R. Zabihi, D. Mowla, and H. R. Karami, Artificial intelligence approach to predict drag reduction in crude oil pipelines, *J. Pet. Sci. Eng.* **178**, 586 (2019).
- [31] B. Roidl, M. Meinke, and W. Schröder, A reformulated synthetic turbulence generation method for a zonal RANS–LES method and its application to zero-pressure gradient boundary layers, *Int. J. Heat Fluid Flow* **44**, 28 (2013).
- [32] P. Schlatter and R. Örlü, Turbulent boundary layers at moderate Reynolds numbers: Inflow length and tripping effects, *J. Fluid Mech.* **710**, 5 (2012).
- [33] R. Ishar, E. Kaiser, M. Morzyński, D. Fernex, R. Semaan, M. Albers, P. S. Meysonnat, W. Schröder, and B. R. Noack, Metric for attractor overlap, *J. Fluid Mech.* **874**, 720 (2019).
- [34] J. P. Boris, F. F. Grinstein, E. S. Oran, and R. L. Kolbe, New insights into large eddy simulation, *Fluid Dyn. Res.* **10**, 199 (1992).
- [35] M. Meinke, W. Schröder, E. Krause, and T. Rister, A comparison of second- and sixth-order methods for large-eddy simulations, *Comput. Fluids* **31**, 695 (2002).
- [36] C. Cortes and V. Vapnik, Support-vector networks, *Machine Learning* **20**, 273 (1995).
- [37] A. Burkov, *The Hundred-Page Machine Learning Book* (Andriy Burkov, 2019).
- [38] N. Tomiyama and K. Fukagata, Direct numerical simulation of drag reduction in a turbulent channel flow using spanwise traveling wave-like wall deformation, *Phys. Fluids* **25**, 105115 (2013).
- [39] S. L. Brunton, J. L. Proctor, and N. J. Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems, *Proc. Natl. Acad. Sci. USA* **113**, 3932 (2016).
- [40] R. Tibshirani, Regression shrinkage and selection via the lasso, *J. Roy. Stat. Soc. B Met.* **58**, 267 (1996).