Ocean Engineering 210 (2020) 107566

Contents lists available at ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng

# Comparison of meta-modeling methodologies through the statistical-empirical prediction modeling of hydrodynamic bodies

Christopher S. Thurman, J. Ryan Somero *

*Newport News Shipbuilding, A Division of Huntington Ingalls Industries, 4101 Washington Ave., Newport News, VA 23607, United States of America*

ARTICLE INFO

*Keywords:*
Hydrodynamics
Maneuvering prediction
Design of experiments
Machine learning
Neural network

ABSTRACT

A methodology for directly establishing the forces and moments on a maneuvering body based on the instantaneous state of the vehicle is demonstrated. Both classical and modern Design of Experiments were used to develop experimental design regions consisting of different combinations of inflow conditions. Data was generated using computational fluid dynamic simulations which were performed on DARPA SUBOFF and the Virginia Tech ellipsoid model at the parameters prescribed by the designed experiments. Nonlinear polynomial regression modeling and artificial neural network modeling were used to develop prediction models for the normal and side force coefficients of both geometries with functional dependence on the inflow conditions. The prediction models were compared to published experimental data which showed excellent agreement within the experimental design region. The classical and modern approaches were also compared to each other and various strengths and applicability were illustrated. Nonlinear polynomial regression modeling and classical Design of Experiments proved to be very insightful by elucidating which inflow parameters and inflow parameter interactions were significant to the prediction model. Artificial neural network modeling coupled with modern Design of Experiments was shown to be more accurate, however, gave no insight to the underlying physical flow phenomena.

## 1. Introduction

Prediction modeling of aerodynamic and hydrodynamic bodies has always been of utmost interest in the commercial and defense sectors. Accurate prediction modeling can lead to computationally inexpensive meta-models which fully characterize the geometric body in question. Ultimately, these meta-models can be used for geometric or other forms of goal-oriented optimization, such as drag reduction, at minimal to no cost.

Prediction modeling can also be applied to determining the maneuvering characteristics of aero/hydrodynamic vehicles. Traditional maneuvering models use a series of coefficients based on Taylor series expansion of the forces and moments to describe the motion of the vehicle. The accuracy of these models is based on the number of coefficients used to describe the motion. Designers tend to limit motion to a single plane which reduces the number of coefficients required, but these techniques fail to capture nonlinear effects due to motion in multiple planes. In order to more accurately capture these nonlinear effects, additional coefficients are required, which significantly increases the number of experiments that must be run to establish these coefficients. Racine and Paterson (2005) explored methods to reduce the number of coefficients required to accurately model the motion of a vehicle by determining the relative importance of each term. Terms that were considered less significant were dropped from the equations of motion, however the final model produced mixed results when compared with experimental data.

Simonsen et al. (2012) also explored the use of simulated Planar Motion Mechanism (PMM) tests where the minimum number of PMM tests required to provide sufficient accuracy was determined. It was found that the number of PMM tests could be reduced by 30% without significant loss of accuracy. The test matrix reduction, however, could not be accomplished in a systematic manner as the maneuvering model would easily become unstable and diverge. A significant understanding of the vehicle's maneuvering characteristics was required prior to reduction of the test matrix to ensure stability would be maintained, reducing this method's attractiveness for concept design.

Hess (2002) demonstrated the power of Recursive Neural Networks (RNN) for application to maneuvering prediction during the Office of Naval Research Maneuvering Challenge, where the RNN out performed models using lifting line theory, modified potential flow, and Unsteady Reynolds-Averaged Navier–Stokes (URANS). The challenge in the application of this RNN is it used 78 time-accurate experimental data sets to train. Generation of this data would be cost and time prohibitive for

* Corresponding author.
*E-mail address:* john.r.somero@hii-nns.com (J.R. Somero).

https://doi.org/10.1016/j.oceaneng.2020.107566
Received 18 November 2019; Received in revised form 8 April 2020; Accepted 23 May 2020
Available online 29 May 2020

**Nomenclature**

| | |
|---|---|
| $C_Z$ | normal force coefficient |
| $C_Y$ | side force coefficient |
| $C_n$ | yaw moment coefficient |
| $\alpha$ | angle of attack |
| $\beta$ | side slip angle |
| $MS_E$ | mean square of the error |
| $\vec{u}$ | local velocity in three-dimensional space |
| $\rho$ | fluid density |
| $p$ | local fluid pressure |
| $\nu_{eff}$ | effective viscosity |
| $\nu$ | kinematic viscosity |
| $\nu_t$ | turbulent eddy viscosity |
| $\Delta x$ | cell size (dimensional) |
| $\Delta s$ | cell size (non-dimensionalized by L) |
| L | body length |
| y+ | dimensionless wall distance |
| $Re$ | Reynolds number ($UL/\nu$) |
| $\vec{U}_b$ | velocity, body-frame |
| $\vec{U}_I$ | velocity, inertial-frame |
| $\psi$ | yaw angle |
| $\theta$ | pitch angle |
| $\phi$ | roll angle |
| $k$ | turbulent kinetic energy |
| $\omega$ | specific dissipation rate |
| $\xi_0$ | regression model intercept |
| $\xi_i$ | regression coefficients |
| $\varepsilon_i$ | sum of squares of the error |
| E(y) | overall mean |
| $\bar{y}(D_n)$ | sample mean |

ship design applications using physical experiments. Even considering the significant advances in computational fluid dynamics (CFD) and available computational resources, generation of this type of data set is currently computationally prohibitive. The question, therefor, remains of how to optimally gain insightful information with regard to characterization while keeping the computational cost relatively low. The solution to this conundrum is through the use of Design of Experiments (DoE).

This paper aims to explore various statistical sampling methods inherent to DoE which are somewhat specific to two different modeling techniques to be compared, polynomial regression and artificial neural network modeling. These two modeling techniques will be used to develop nonlinear relationships between normal force coefficient, $C_Z$, and side force coefficient, $C_Y$, as respective functions of the inflow conditions: angle of attack, $\alpha$, and side-slip angle, $\beta$. This method allows the forces and moments to be established for a 6-Degree of Freedom (6-DOF) simulation directly based on the state of the vehicle, in lieu of small angle limited linear derivatives.

The method is demonstrated here using two different geometries: the DARPA SUBOFF model and the Virginia Tech scalene ellipsoid model. The two different modeling techniques will be validated using experimental data for the two geometries published by Roddy (1990) and Granlund (2009) for the DARPA SUBOFF model and the ellipsoid model, respectively.

## 2. Design of experiments

DoE can be thought of as a process for planning an experiment so that appropriate data can be collected and analyzed by statistical methods, resulting in valid and objective conclusions (Montgomery, 2017; Box et al., 1978). DoE is typically used in systems comprised of multiple inputs when a nonlinear functional relationship between the quantities of interest, such as force and moment coefficients, and the regressors, or inputs, and their interactions is required. DoE can be broken up into two categories: classical and modern. A brief comparison of which is given in Table 1.

### 2.1. Classical DoE

Classical DoE sampling methods are generally coupled with preliminary least-squares regression modeling, the Analysis of Variance (ANOVA) significance testing, and regression model reduction. Regressors, such as the main effects from the inputs, the effect of input interactions, and the effect of various higher order terms are included in preliminary least-squares regression models before significance testing is performed to determine which effects are inconsequential to the regression model. Insignificant terms are removed from the model and added to the error estimate of the model. In this fashion, nonlinear polynomial regression models can be created and refined to yield sought after functional relationships between outputs and regressors.

Classical DoE sampling methods were originally developed for use in physical experiments (Yondo et al., 2018) and have been used with great success for complex aerospace and motorsport vehicle characterization applications (Murphy and Landman, 2015; Doane and Landman, 2012; Landman et al., 2002), as well for suspension system design (Mohsen et al., 2015) and instrumentation calibration (Reinholtz et al., 2012).

These sampling methods assume stochasticism, or the presence of random pure error, inherent to the physical experiment which is normally distributed, N(0, $\sigma^2$). The presence of this pure error and its normality are key main assumptions used in ANOVA which decouples the total variability in the response, or functional output from least-squares regression, and tests the significance of the regressors to the meta-model by comparing a measure of their variability: their individual mean squares, to the mean square error ($MS_E$) of the regression model which serves as a measure of the stochasticism inherent to the experiment. With the presence of this inherent error, design points are chosen in a combinatorial fashion, a priori, which exploit aggressive regressor spacing, allowing for a decreased number of samples.

Typically, factorial type designs and central-composite designs (CCD) are used in classical DoE with many alterations and augmentations possible which allow for the accommodation of testing limitations or the optimization of the experimental design.

Augmentation usually concerns the addition of samples to the design space in order to enhance the predictive capabilities of the resultant regression model or to fulfill some alphabetic optimality criteria. D-optimality, which is often used, focuses on good model parameter estimation, allowing for superior regressor screening capabilities (Myers et al., 2016). I-optimality works to reduce the average scaled prediction variance, allowing for optimal predictive capabilities within the design space. There are other alphabetic optimality criteria as well that will not be addressed in this paper, but if interested, can be found in Montgomery (2017) and Myers et al. (2016).

### 2.2. Modern DoE

Contrary to classical DoE, modern DoE deals almost exclusively with computer-based experiments. These types of experiments are deterministic, meaning that there is no measure of the inherent stochasticism. The error is typically associated with grid discretization error, round off error, numerical convergence error, etc., meaning that for a fixed number of iterations, a replicated computer experiment should yield the same results. Because of this, there is no reference quantity with which to perform significance testing, so other methods for statistical-empirical modeling must be used which require different experimental sampling techniques.

**Table 1**
Comparison between classical and modern DoE.

| Classical DoE | Modern DoE |
|---|---|
| Application to physical experiments | Application to computer-based simulations |
| Fewer experimental points required | More experimental points required |
| Based upon statistical methodology | No statistical basis; more of an interpolation approach |
| Fewer available modeling procedures | More available modeling procedures |
| Factor screening and model reduction | N/A |
| Limited order of polynomial model (typically $< (xy)^3$) | Unlimited nonlinear modeling capability |

Various alphabetically optimal designs may be used, like with classical DoE, however, space-filling designs are predominantly used. There are several different space-filling designs such as the Uniform Design (UD), Latin Hyper-Cube Design, Sphere-Packing Design, etc., which, in general, spread the design points out nearly evenly or uniformly throughout the region of experimentation, adhering to some measure of uniformity determined by the design being used (Myers et al., 2016). The purpose of this uniform sampling is to acquire data in a manner that will represent the entire experimental domain while minimizing the difference of the overall mean between the meta-model and the experimental data (Fang et al., 2005).

These space-filling designs often require more runs than typical classical DoE designs, such as factorial experiments, since the experimental design points are placed at interior points rather than at the outer limits of the experimental domain. Alphabetic optimal designs are somewhat of a hybrid between classical and modern DoE designs since they place points both on the exterior of the design space in a factorial fashion, as well as on the interior to adhere to some optimality criterion.

There are many modeling techniques used with modern DoE. Polynomial regression is often used, however, contrary to its use with classical DoE, ANOVA may not be used due to the lack of an inherent stochasticism estimate and other forms of significance testing are needed for factor screening. Since the modeling procedure for modern DoE does not involve ANOVA, it is more of an interpolation problem (Fang et al., 2005). This broader interpolation problem lends itself to the application of machine learning (ML) algorithms such as Bayesian estimation, decision trees, artificial neural networks (ANN), etc. All of which approach the general data interpolation problem somewhat abstractly in that they define the functional relationship between input and output in a "black-box" mannerism without giving much insight as to which input factors are significant to the problem. The ML modeling technique of particular interest to this research is the ANN which will be discussed in subsequent sections. The reader is referred to Da Ronch et al. (2017), Cleaver et al. (2016), Montevechi et al. (2007) and Nazalla et al. (2006) for various modern DoE and modeling techniques, as applied to the simulation of aerospace vehicles and manufacturing processes, respectively.

## 3. Geometric configurations

This research not only looks to compare different modeling techniques coupled with their respective sampling methodologies, but it also aims to determine if there are any limitations imposed on the experimental design region by the CFD processes used. To validate both the capabilities of the CFD solver and the meta-models, results are compared with the experimental data of the DARPA SUBOFF and the Virginia Tech scalene ellipsoid model. The SUBOFF geometry (Roddy, 1990) was used in a bare-hull configuration for simplicity and will generalize typical submersible bodies-of-revolution (BOR). The scalene ellipsoid generalizes non-bodies-of-revolution (NBOR) and its asymmetric geometry will induce flow physics of greater complexity. Its geometric parameters were taken from Granlund (2009) and Somero (2010).

The coordinate system used for both geometries was located at each geometry's center of gravity and was such that the positive X-direction faced upwind, the positive Y-direction faced starboard, and the positive Z-direction pointed downward.

## 4. CFD solver

HPCMP CREATE$^{TM}$SH NavyFOAM was used throughout the scope of this work. NavyFOAM is a CFD solver that was developed under the DoD High Performance Computing Modernization Program's (HPCMP) Computational Research and Engineering Acquisition Tools and Environments (CREATE$^{TM}$) Ship's Hydrodynamics Project. NavyFOAM builds upon the computational framework of the open source CFD toolkit, OpenFOAM, by implementing a large number of new features and capabilities including new top-level solvers for single- and multi-phase flows, libraries of new numerical schemes and algorithms, advanced physical models, and custom postprocessing utilities. Here, the single phase incompressible flow solver, sRansFoam, is used to solve the equations for the conservation of mass and momentum given as:

$$\nabla \cdot \vec{u} = 0 \tag{1}$$

$$\frac{\partial \vec{u}}{\partial t} + \vec{u} \cdot \nabla \vec{u} = -\frac{1}{\rho}\nabla p + \nabla \cdot \left[\nu_{eff}\left(\nabla \vec{u} + \nabla^T \vec{u}\right)\right] \tag{2}$$

where $\nu_{eff} = \nu + \nu_t$, $\nu$ is the kinematic viscosity, $\nu_t$ is the turbulent eddy viscosity, and the superscript $T$ in $\nabla^T$ represents the matrix transpose. The Navier–Stokes solvers in NavyFOAM employ a cell-centered finite-volume method based on a multidimensional linear reconstruction scheme that permits use of arbitrary polyhedrals (Kim et al., 2017). One of the features in NavyFOAM which was taken advantage of in this work is the inclusion of Wilcox's updated 2006 $k$ - $\omega$ turbulence model. This updated, two-equation turbulence closure model used in Reynolds-Averaged simulations has been proven to resolve separated flows more efficiently than its preceding rendition (Wilcox et al., 2006; Wilcox, 2008).

## 5. Technical procedure

### 5.1. Mesh generation

#### 5.1.1. Surface mesh generation

The surface mesh for the DARPA SUBOFF model and the ellipsoid were both created such that adequate cell coverage spanned the entirety of each model. Both surface meshes were triangulated and unstructured. The DARPA SUBOFF model had variable surface mesh density whereas the ellipsoid had a constant density. The DARPA SUBOFF surface mesh had diminishing cell sizes with a maximum cell size, $\Delta x$, of 0.01 m along the parallel mid-body, or non-dimensionalized by body length, $\Delta s = 0.0024$.

The stern was separated into three sections with the approximate separation located at geometric slope changes along the surface in the downwind direction. Each of these sections decreased in cell size by a factor of 1.5 sequentially, starting from the parallel mid-body, while maintaining one-to-one cell connectivity between sections. The surface mesh on the bow section was 1.5 times smaller than the parallel mid-body. The ellipsoid geometry's surface mesh had a constant mesh density, with an approximate $\Delta s = 0.0044$.

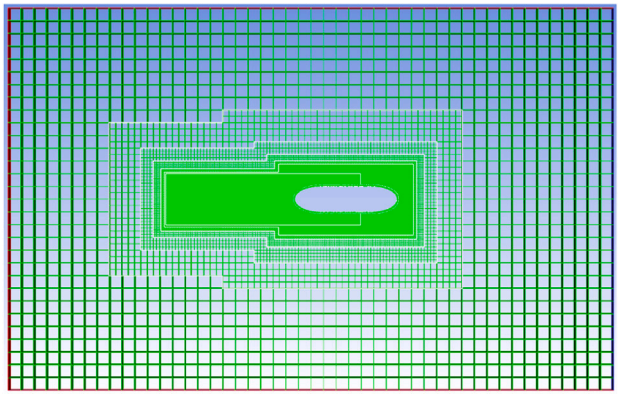

**Fig. 1.** Ellipsoid volume mesh.

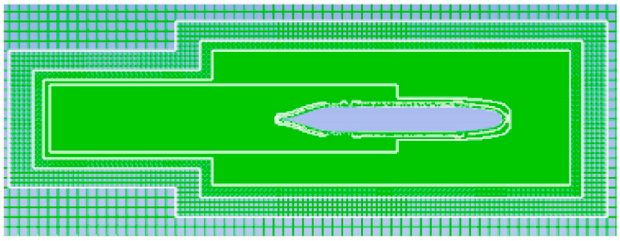

**Fig. 2.** SUBOFF refinement regions.

#### 5.1.2. Volume mesh generation

The volume meshes for both the DARPA SUBOFF model and the ellipsoid, shown in Figs. 1 and 2, were created as hex-dominant volume meshes with several levels of refinement from the far field. The outer limits imposed on the volume mesh for both geometries were as follows: 3 body lengths in the downwind direction, 2.5 body lengths in the upwind direction, and 2 body lengths port/starboard of both geometries. The maximum cell length along the outer region was approximately $\Delta s = 0.2$.

Refinement regions were added around the anticipated region of the wake for at least 1 full body length as well as around both geometries. The wake refinement regions were 7 levels (8× volumetrically each level) of refinement greater than the outer region $\Delta s$ and the geometry refinement regions were 6 levels of refinement greater. The addition of these two refinement regions served the purpose of increasing the flow field resolution in regions of particular interest to this study. Since forces integrated over the geometries are of utmost importance, accurate modeling of the flow field around the hull and in the wake region is necessary. Fig. 2 depicts the two refinement regions around the DARPA SUBOFF model.

The first layer of cells adjacent to the geometries was sized to accommodate a y+ distance of approximately 60 for the Reynolds number, Re, specific to each geometry. The Reynolds numbers were set to match those given in the published experimental data and will be addressed in subsequent sections (Roddy, 1990; Granlund, 2009).

### 5.2. CFD implementation & validation

Prior to comparing DoE sampling methodologies and modeling techniques, CFD validation was necessary to determine the accuracy of the numerical techniques implemented and their ability to properly simulate the physical flow phenomena surrounding each geometry. This was done by running simulations where only one of the two inflow variables was modified in an effort to replicate the experimental results. Since the DARPA SUBOFF model is axisymmetric, only one of the input parameters will be used, whereas the NBOR nature of the scalene ellipsoid will require that both inflow parameters be investigated individually.

For the incompressible flow about our two geometries, a steady-state solution was contrived. The Laplacian terms associated with the diffusion of momentum, turbulent kinetic energy, and turbulent specific dissipation were discretized spatially using second-order linear interpolation schemes. First-order spatial upwind discretization was applied to the divergence schemes associated with the convection of the turbulent kinetic energy and turbulent specific dissipation. The convective momentum, on the other hand, was spatially discretized using a blend between a second-order linear interpolation scheme and a first-order spatial upwind scheme.

The matrix representation of the discretized set of linear algebraic equations for the momentum, turbulent kinetic energy, and turbulent specific dissipation is preconditioned using a symmetric, diagonal incomplete-Cholesky method and solved with a preconditioned bi-conjugate gradient scheme (PBiCG). The Poisson's equations for the pressure and pressure correction, derived from the momentum equations and utilized by the Semi-Implicit Method for Pressure Linked Equation (SIMPLE) algorithm, were preconditioned using an asymmetric, diagonal incomplete-Cholesky method and solved with a preconditioned conjugate gradient scheme (PCG).

Solution tolerances were specified such that iteration would ensue until the residuals associated with the turbulence closure equations were reduced by 6 orders of magnitude and the residuals associated with the pressure, pressure correction, and momentum equations were reduced by 7 orders of magnitude. These convergence criteria were satisfied before the SIMPLE method's outer loop began the next iteration. Under-relaxation was also specified to promote the stability of the iterative process and two non-orthogonal correctors were utilized due to the non-orthogonality associated with the unstructured meshes used.

#### 5.2.1. Prescribed simulation parameters

To accurately replicate the incompressible flow around the geometries in question, dynamic similarity was maintained by prescribing values of Reynolds number that matched those of the published experiments. For DARPA SUBOFF, $Re = 1447 \times 10^6$ and for the scalene ellipsoid, $Re = 4.5 \times 10^6$ (Roddy, 1990; Granlund, 2009). Inflow velocity magnitudes and kinematic viscosities were modified to accommodate these Reynolds numbers for all inflow conditions throughout the simulations conducted in this research.

To determine what the proper inflow conditions should be in the body-fixed reference frame coordinate system described in Section 3, it was necessary to transform the velocity components from the inertial reference frame, using the Euler rotation angles given in Eq. (3) (Bossert et al., 2003).

$$U_B = \begin{bmatrix} \cos(\psi) & \sin(\psi) & 0 \\ -\sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} \cos(\theta) & 0 & -\sin(\theta) \\ 0 & 1 & 0 \\ \sin(\theta) & 0 & \cos(\theta) \end{bmatrix} * \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi) & \sin(\phi) \\ 0 & -\sin(\phi) & \cos(\phi) \end{bmatrix} * U_I \tag{3}$$

Where $\psi$ is the yaw angle, $\theta$ is the pitch angle, and $\phi$ is the roll angle.

Utilizing the coordinate transformation given in Eq. (3), velocity conditions were calculated and imposed at the inlet and the far field sides while a zero gradient condition was given at the outlet. The geometric surfaces were given a no-slip velocity condition. The pressure was given a gauge value of zero at the outlet. A zero gradient pressure condition was imposed along all other domain boundaries, as well as along the geometric surfaces. Since the $k$ - $\omega$ turbulence model was used, the turbulent kinetic energy, $k$, and the turbulent specific dissipation, $\omega$, were prescribed to $3.75 \times 10^{-5}$ $m^2/s^2$ and 1000 $s^{-1}$,

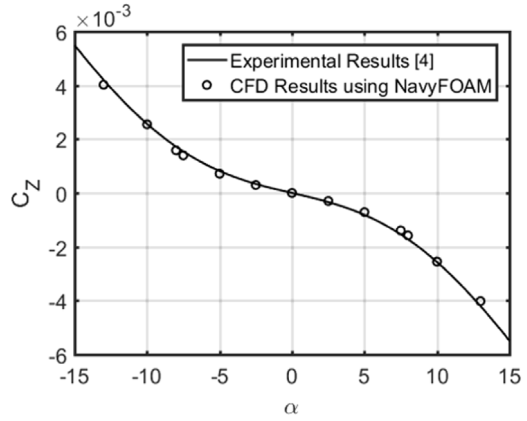


**Fig. 3.** DARPA SUBOFF $C_Z$ Comparison ($\beta = 0°$).

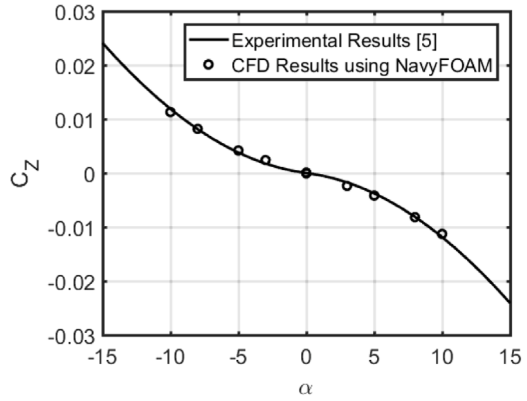


**Fig. 4.** Scalene ellipsoid $C_Z$ comparison ($\beta = 0°$).

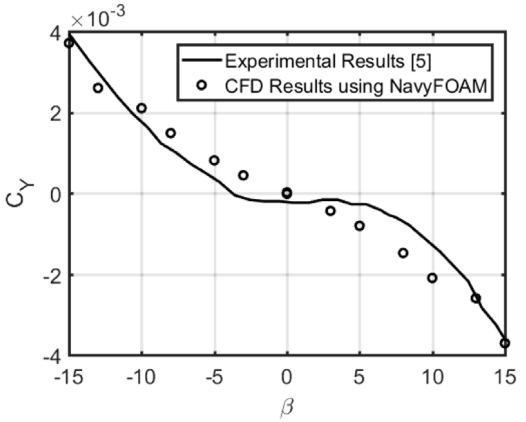


**Fig. 5.** Scalene ellipsoid $C_Y$ comparison ($\alpha = 0°$).

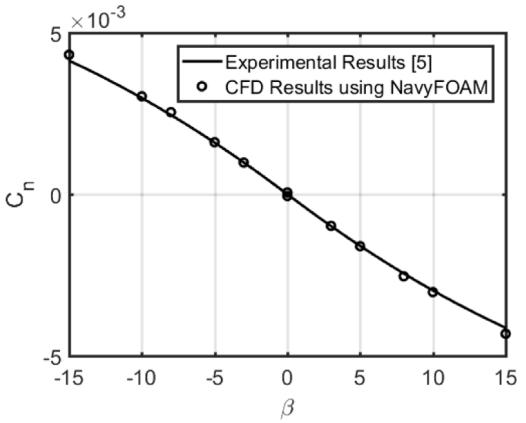


**Fig. 6.** Scalene ellipsoid $C_n$ comparison ($\alpha = 0°$).

respectively, at the inlet. Zero gradient turbulence conditions were given at the far field sides and the outlet.

For the CFD validation, $\alpha$ values of 0°, 3°, 5°, 7.5°, 8°, 10°, and 13° were simulated and compared with the experimental results for the DARPA SUBOFF geometry. For the scalene ellipsoid model $\alpha$ and $\beta$ were each treated individually, while holding the other parameter constant. Values of 0° 3°, 5°, 8°, and 10° for $\alpha$ and 0°, 3°, 5°, 8°, 10°, 13°, and 15° for $\beta$ were simulated and compared with experimental results. The inflow velocities corresponding to these angles and for combinations of angles in subsequent sections were all transformed from the inertial frame of reference to the body-fixed frame using Eq. (3).

#### 5.2.2. CFD validation

The numerical methods outlined previously were executed in NavyFOAM for each geometry which produced good agreement with the published experimental results (Roddy, 1990; Granlund, 2009) and will be compared in this section.

The DARPA SUBOFF model was simulated at $\alpha$ values ranging from 0° to 13° and the results for $C_Z$ are shown in Fig. 3 where force is non-dimensionalized by $L^2$ as

$$C_z = \frac{Z}{\frac{1}{2}\rho U^2 L^2}. \tag{4}$$

It should be noted that the force coefficient data for all experiments conducted in the scope of this research were averaged over the last 200 iterations. Since the geometry was axisymmetric, the results were extended to negative $\alpha$ values by mirroring the results about the X and Y axes to aid in the visualization of the result comparison. Analysis of $C_Y$ was not needed due to the symmetry. The percentage errors between the fine mesh CFD results and the published experimental results were calculated with their averaged value being 5.18%. The same comparison was done with the Virginia Tech scalene ellipsoid results, however, since it is a NBOR, the lateral direction was also analyzed. The results are plotted in Figs. 4, 5, and 6.

Since there were discrepancies in the $C_Y$ data, the yaw moment coefficient, $C_n$, has also been plotted in Fig. 6. The percentage differences averaged over the aforementioned discrete data points for $C_Z$ and $C_n$ were 5.67% and 2.17%, respectively.

### 5.3. DoE and modeling procedure

Accurate results have been shown achievable for angles up to at least to 13°–15° in both $\alpha$ and $\beta$, respectively. For the purpose of this research, the region of interest was chosen to be between −8° and 8° for both $\alpha$ and $\beta$ to provide the ability to evaluate model performance in prediction beyond the range of data included.

#### 5.3.1. Classical DoE implementation

Typically, there is a sequential nature involved with classical DoE methodology. After the region of experimentation has been defined, a two-level factorial experiment is first generated to allow for factor screening, or significance testing. This involves the generation of a preliminary least-squares regression fit to the data. This regression model takes the generic form of Eq. (5) which is defined as a multiple linear regression model (Myers et al., 2016).

$$y_i = \xi_0 + \sum_{j=1}^{k} \xi_j x_{ij} + \varepsilon_i, \qquad for \;\; i = 1, 2, \ldots, n \tag{5}$$

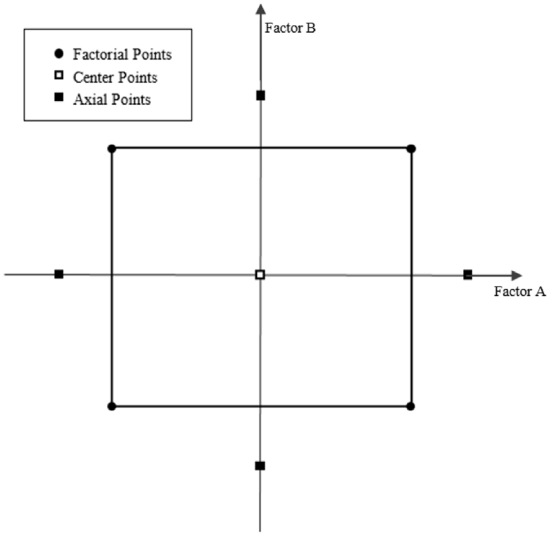


**Fig. 7.** Typical central composite design experiment.

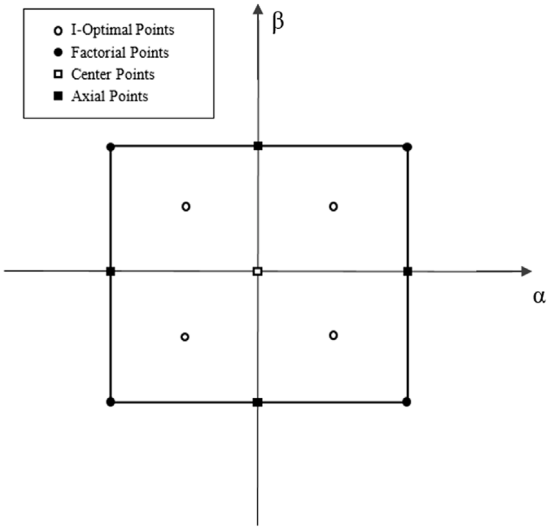


**Fig. 8.** FCD experiment with I-optimal points (Design space ranging from $\alpha = \pm 8°$ and $\beta = \pm 8°$).

**Table 2**
FCD experimental design points.

| RUN | $\alpha$ | $\beta$ |
|---|---|---|
| 1 | −8° | 0° |
| 2 | 8° | 0° |
| 3 | 0° | 0° |
| 4 | −8° | 8° |
| 5 | 8° | −8° |
| 6 | 8° | 8° |
| 7 | −8° | −8° |
| 8 | 0° | −8° |
| 9 | 0° | 0° |
| 10 | 0° | 8° |
| 11 | 0° | 0° |
| 12 | 4° | −4° |
| 13 | 4° | 4° |
| 14 | −4° | −4° |
| 15 | −4° | 4° |
| 16 | 0° | 0° |

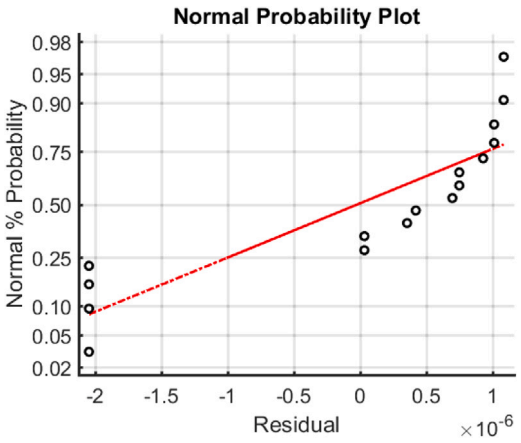


**Fig. 9.** DARPA SUBOFF — normal probability plot of $C_Z$ (line represents normal distribution).

where $\xi_0$ is the model intercept, $\xi_j$ are regression coefficients, and $\varepsilon_i$ are sum of squares of the error.

The standard procedure decomposes the error in the regression model into two parts: a contribution from the variability in each individual regressor, and the contribution from the stochasticism inherent to the experiment. These contributions are namely the mean square of each regressor and the $MS_E$, respectively. The ANOVA process then determines the significance of inputs and multi-input interactions to the regression model by performing an F-test which compares the error estimate of the regressors to the error estimate of the stochasticism. Insignificant terms are dropped from the model and added to the estimate of the pure error. Once it is known which terms are insignificant, sequential experimentation can be done without these terms, occasionally reducing the number of experimental runs to be conducted if a factor is eliminated completely from the experiment.

Most sequential experimentation with a two-level factorial experiment involve the augmentation of the experimental design space with the addition of axial points and center points, as shown in Fig. 7, allowing for the estimation of quadratic curvature. A design of this form is typically called a Central Composite Design (CCD). Axial points are generally spaced such that certain properties of the experimental design space are achieved, such as rotatability. Significance testing is performed again with this new design space to determine which higher order terms and interactions are significant to the new, second-order regression model which takes the general form of Eq. (6) (Myers et al., 2016). The final polynomial regression model will be a direct consequence of the significance testing of regressors to Eq. (6).

$$y = \xi_0 + \sum_{j=1}^{k} \xi_j x_j + \sum_{i<j=2}^{k} \sum \xi_{ij} x_i x_j + \sum_{j=1}^{k} \xi_{jj} x_j^2 + \varepsilon \tag{6}$$

To perform the standard approach outlined above, certain assumptions must be satisfied which define characteristics of the experimental data. The first assumption concerns the normality of the experimental data and of the error, requiring that they both conform to a Gaussian distribution. The second assumption is that the data have no time dependence. This ensures that the stochasticism in the experiment is not due to a process error such as hysteresis, wind tunnel warm-up error, or other time dependent forms of error. The last assumption is that the data and the stochastic error both have constant variance. With these three assumptions met, the residuals, or discrete differences between the experimental data and the regression model prediction of the experimental data should be completely structure-less, allowing for the $MS_E$ to be used as a measurement statistic for ANOVA.

Since the form of the polynomial regression can be assumed to be cubic for both the DARPA SUBOFF model and the scalene ellipsoid, as suggested by Roddy (1990) and Granlund (2009), an experimental design was chosen which included axial and center points to determine

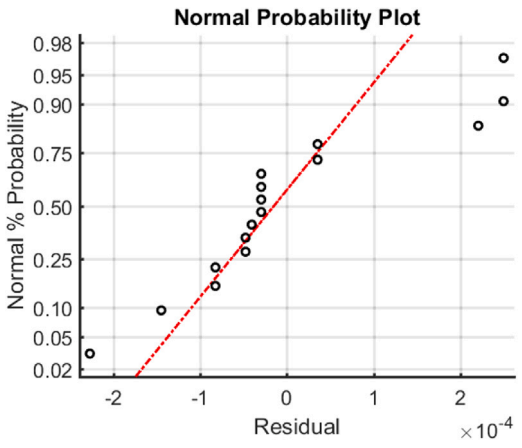


**Fig. 10.** Ellipsoid — normal probability plot of $C_Z$ (line represents normal distribution).

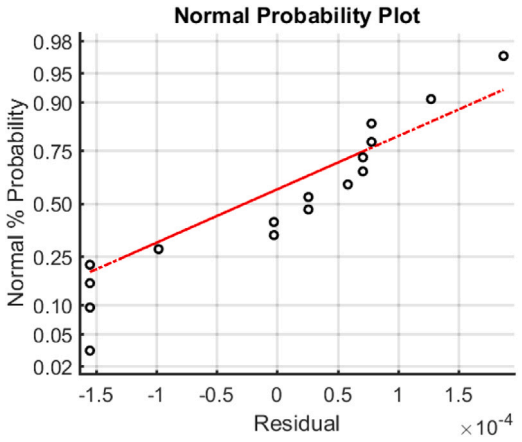


**Fig. 11.** Ellipsoid — normal probability plot of $C_Y$ (line represents normal distribution).

the curvature of the responses, $C_Z$ and $C_Y$ which were functions of $\alpha$ and $\beta$. A Face Centered Design (FCD) was used with axial points located along the faces of the experimental region. The general form of a cubic regression model in two factors has 10 model terms which is equivalent to the number of experimental design points, indicating saturation. To address this, six additional model points were chosen in an I-optimal fashion to reduce the average scaled prediction variance

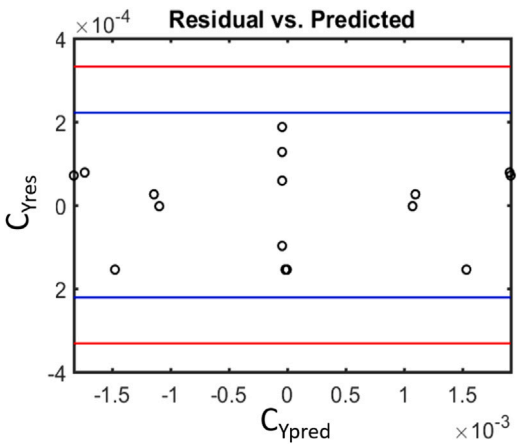


**Fig. 13.** Ellipsoid — residuals vs. predicted for $C_Y$ (lines are 2 and 3 residual standard deviations).

throughout the region of experimentation. Two of the points involved with this augmentation were center runs to help stabilize the prediction variance and provide additional degrees of freedom for pure error estimation (Myers et al., 2016). This 16 point design space in two factors produced the randomly ordered experimental design depicted in Fig. 8 and tabulated in Table 2.

Since the regression model form has already been determined to be cubic, the preliminary model with which to perform significance testing takes the form of Eq. (7) for both $C_Z$ and $C_Y$ for both geometries.

$$y = \xi_0 + \xi_1\alpha + \xi_2\beta + \xi_{12}\alpha\beta + \xi_{11}\alpha^2 + \xi_{22}\beta^2 + \xi_{112}\alpha^2\beta + \xi_{122}\alpha\beta^2 + \xi_{111}\alpha^3 + \xi_{222}\beta^3 \tag{7}$$

The normal probability plots for $C_Z$ of the DARPA SUBOFF model and the scalene ellipsoid and $C_Y$ of the scalene ellipsoid are shown in Figs. 9–11.

The results of which, show that the normality assumption is violated for the $C_Z$ of SUBOFF. This means that ANOVA cannot be performed without data manipulation due to the deterministic nature of the experiment. The residuals of the scalene ellipsoid, aside from a few outliers, seem to trend towards normality without any significant defects in behavior. It should be noted that these normal probability plots are preliminary and correspond to residuals from the full regression model, Eq. (7). They are used simply to determine the general trend in data and outliers will be addressed in subsequent sections. Once ANOVA is performed, the normal probability plots of the reduced models will be analyzed and detailed in Section 6.

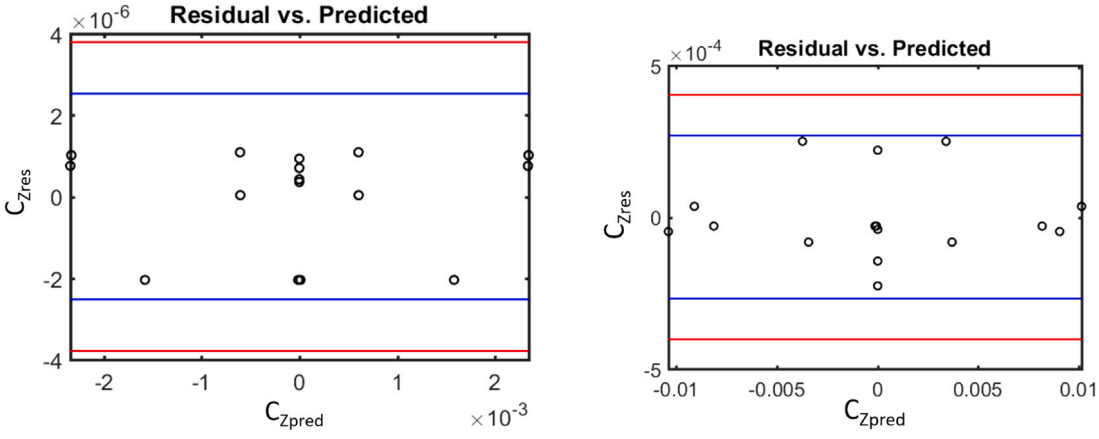


**Fig. 12.** Residuals vs. predicted for $C_Z$ (lines are 2 and 3 residual standard deviations). Left — DARPA SUBOFF, Right — ellipsoid.

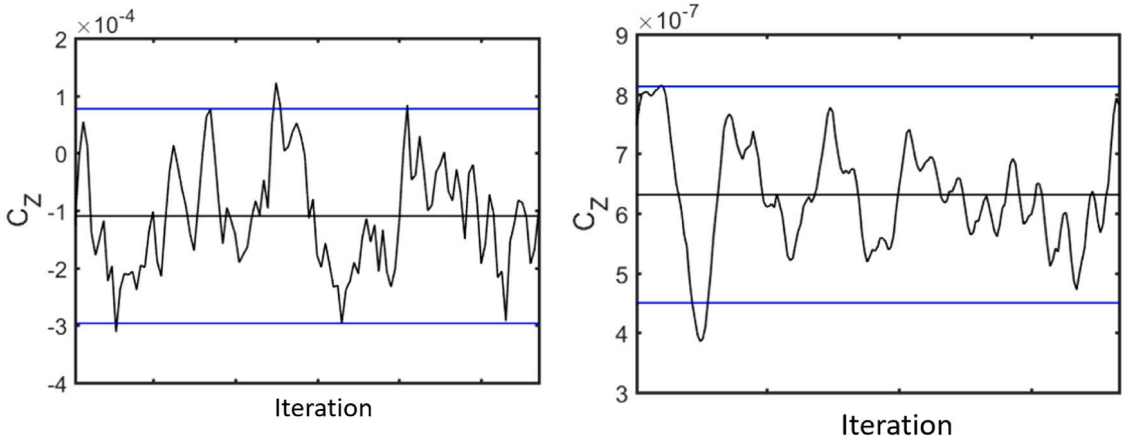


**Fig. 14.** $C_Z$ pure error at center point (Run 3) (lines are 2 residual standard deviations). Left — Ellipsoid, Right — DARPA SUBOFF.

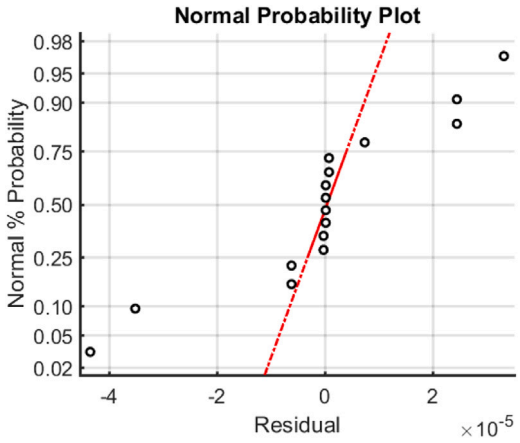


**Fig. 15.** DARPA SUBOFF — normal probability plot of $C_Z$ after adding wind tunnel error (line reflects normal distribution).

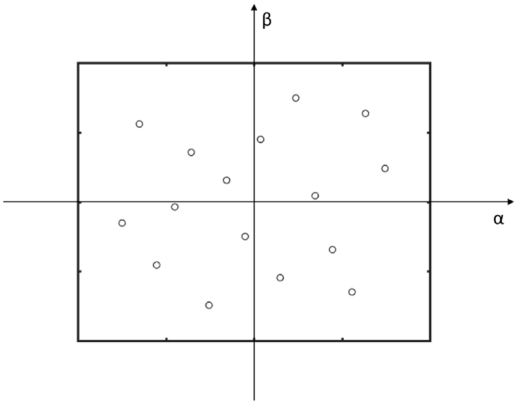


**Fig. 16.** Uniform design in two factors (Design space ranging from $\alpha = \pm 8°$ and $\beta = \pm 8°$).

The residuals vs. the predicted responses for the $C_Z$ of the DARPA SUBOFF model and the scalene ellipsoid and the $C_Y$ of the ellipsoid are plotted in Figs. 12 and 13 to determine if there is any structure to the residuals. Aside from the obvious symmetry of the residuals, there appears to be no real structure to the residuals. The lines in Figs. 12 and 13 are indicative of 2 and 3 residual standard deviations, respectively. Since none of the residual values lie outside of these standard deviation boundaries, it is safe to assume that the residuals conform to a Gaussian distribution and are unstructured, with a mean of 0.

An analysis of the time history of the residuals is not necessary since each simulation was run independently of one another. What this preliminary residual analysis suggests is that the CFD data for the DARPA SUBOFF model adheres to two of our three assumptions necessary for ANOVA, and the data for the scalene ellipsoid adheres to all three of our assumptions.

It is believed that the unsteadiness associated with the bluff-bodied nature of the scalene ellipsoid is the proprietor of the variance in the CFD data. Not only is the ellipsoid a complex geometry over which to resolve the flow field, but the Reynolds number is low, reducing the nearfield flow stability. All of these indicate that the problem is truly an unsteady one and yet it was solved using a steady-state solution. A look at the CFD force or moment coefficients at an experimental center point shows the pure error associated with noise caused by unsteadiness. These values plotted against the iteration number are shown in Fig. 14 and indicate that the steady-state simulation of a truly unsteady flow problem not only has very similar behavioral trends as a wind tunnel test, but also has a similar variance (Stringer, 2017).

The lines plotted in Fig. 14 take on the same meaning as in Figs. 12 and 13, indicating that the pure error of an individual run also has constant variance and is structure-less. A look at the $C_Z$ residuals of the DARPA SUBOFF model in Fig. 14 suggests that there is not enough error present for significance testing.

The flow field of the DARPA SUBOFF was less complex than the ellipsoid because it was axisymmetric with a tapered stern. It was also run at a higher Reynolds number promoting flow field stability and numerical convergence, becoming more deterministic in nature than the ellipsoid simulations. Because of this and because of the lack of normality of the residuals (Fig. 9), a normally distributed pure error, $N(0, \sigma^2 = 5 \times 10^{-5})$, calculated from wind tunnel testing data at $\alpha = 0°$, tabulated by Stringer (2017), was added to the 16 experimental runs. This pure error, typically associated with wind tunnel tests, was added to the data in an effort to provide stochasticism as a measurement statistic for significance testing. Another look at the normal probability plot of the residuals, using the DARPA SUBOFF data with injected error, Fig. 15, shows that the normality assumption of ANOVA has now been met.

Therefore, ANOVA will be done for both the DARPA SUBOFF model and the scalene ellipsoid, with the inherent stochasticism being solely a measurement of flow unsteadiness, either from the numerical simulation for the ellipsoid, or from typical wind tunnel testing error added to the SUBOFF data. If a regressor is shown to be significant by comparing it to the unsteadiness, it will not be excluded from the regression model. Taking the mean over 200 iterations will be thought of as equivalent to time averaging wind-tunnel test data for the ellipsoid. The same was done to the DARPA SUBOFF test data prior to adding actual wind

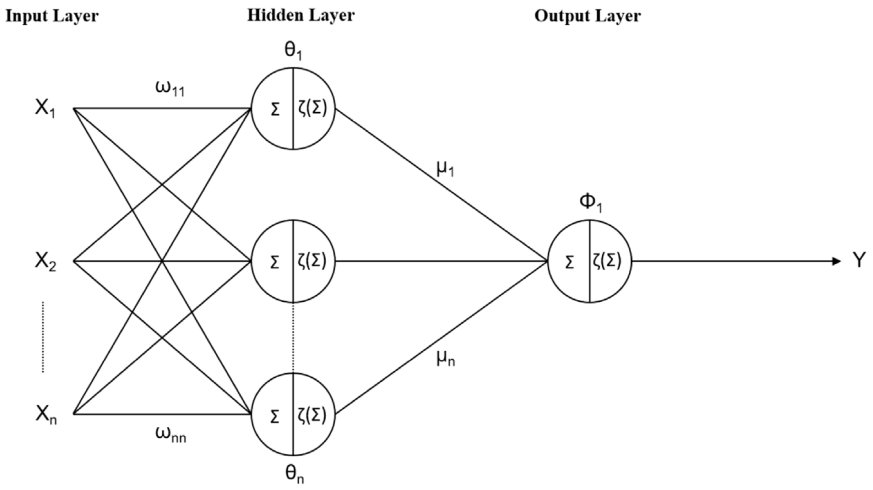


**Fig. 17.** Generic ANN representation (Yondo et al., 2018).

tunnel testing error. Significance testing and model reduction were performed on the $C_Z$ data for both the DARPA SUBOFF model and the scalene ellipsoid, as well as for the $C_Y$ data for the ellipsoid. The results are detailed in Section 6.

### 5.4. Modern DoE implementation

A modern DoE approach has also been explored since these types of experimental designs are tailored towards computer-based experiments. CFD was used throughout the entirety of this experiment, which is typically deterministic in nature, and was shown in the previous section. The same number of experimental design points was used for the modern DoE approach as for the classical approach to more directly compare the two design spaces, and a uniform design space was chosen based upon the limited number of runs, the meta-model selection (ANN), and the desired outcome: optimal prediction capability. The uniform design (UD) places points in a fashion that minimizes some measure of non-uniformity, in this case the centered $L_2$ discrepancy, between the design points and a theoretical uniform distribution (Hickernell, 1998). The design is shown in Fig. 16 and tabulated in Table 3. The uniformly placed points are meant to minimize the difference between the overall mean, Eq. (8), and the sample mean, Eq. (9), throughout the entirety of the design space (Fang et al., 2005).

$$E(y) = \int f(x)dx \tag{8}$$

$$\bar{y}(D_n) = \frac{1}{n}\sum_{i=1}^{n} f(x_i) \tag{9}$$

Many of the modeling procedures used with modern DoE techniques work in an iterative fashion, adjusting, or "training", model parameters by looping over the collected data until some goodness of fit measurement is sufficiently minimized. Typically, these goodness of fit measures are called penalty functions, the most common of which being the quadratic loss function, or the sum of the squared difference between the collected data and the data predicted by an arbitrary model. The minimization of this function is a problem in convex optimization and typically machine learning (ML) algorithms are implemented to solve this minimization problem. The overall goal is to not only minimize this function and provide the best fit of the meta-model to the collected data, but also to be able to extend the meta-model to arbitrary data within the design region not collected in the experiment. The ability of the meta-model to conform to both the "training" data and the "validation" data is the over-arching problem, especially for prediction modeling.

An Artificial Neural Network (ANN) was chosen as the meta-model to fit the data; more specifically, a single layer perception (SLP). The SLP aims at replicating the architecture of the neurons in the human brain, set up in layers as shown in Fig. 17. Each layer consists of a number of activation functions, or neurons, aligned in parallel. All of the neurons of a particular layer are activated in unison, with different multiplicative weights along the connections between neurons, inputs, and outputs.

**Table 3**
UD experimental design points.

| RUN | $\alpha$ | $\beta$ |
|---|---|---|
| 1 | −6.50 | 5.59 |
| 2 | −2.54 | −7.46 |
| 3 | −3.55 | 3.55 |
| 4 | −5.51 | −4.56 |
| 5 | −0.485 | −2.50 |
| 6 | −4.48 | −0.379 |
| 7 | 3.49 | 0.427 |
| 8 | 1.50 | −5.47 |
| 9 | 5.59 | −6.50 |
| 10 | 6.35 | 6.35 |
| 11 | −1.54 | 1.54 |
| 12 | 7.46 | 2.39 |
| 13 | 2.39 | 7.46 |
| 14 | 4.47 | −3.45 |
| 15 | 0.391 | 4.48 |
| 16 | −7.47 | −1.54 |

The SLP architecture used throughout the modern DoE modeling procedure contained one hidden layer consisting of 5 hyperbolic tangent activation functions. Each activation function can be thought of as a hyperplane with values between −1 and 1, following the nonlinear form of the hyperbolic tangent function (Alpaydin, 2014). These activation functions are necessary to introduce nonlinearity to the ANN, enabling it to effectively model complex, nonlinear relationships between inputs and outputs (Sharma, 2017; Walia, 2017). The number and type of activation functions used are heuristically determined and it should be noted that ANNs of different architectures and consisting of different activation functions were investigated, but showed significant bias, or over-fitting, to the simulated results. The chosen SLP architecture was determined to be optimal for this research.

The input layer contains the values for $\alpha$ and $\beta$ at each design point, which are passed through the SLP randomly and the multiplicative weights along the connections are modified until the quadratic loss function is minimized in a feed-forward, gradient descent training manner. Due to the limited number of design points, a k-fold method was used for validating the SLP where the data was divided into 3 subsets. Each subset was used to validate the model fit after training and the kth model with the best fit was selected as the final model.

Statistical analysis of the collected data is not necessary for a ML approach because input significance is inconsequential to most ML modeling procedures. Therefore, no insight is gained about which inputs, or input interactions, are significant to the ANN. Rather than excluding insignificant terms, the ANN works with combinations of highly nonlinear mixtures of inputs, determined by the number of hidden layers, the number of activation functions in each layer, and the form of the activation function specified by the engineer or researcher. Based on these specifications, an ANN effectively takes on a "black-box" form to accurately map the inputs to the outputs throughout the experimental design region in a way that generalizes to arbitrary data inside this design region. The specific ANNs created to model the previously defined outputs for the DARPA SUBOFF and the scalene ellipsoid models will be analyzed in Section 6.

## 6. Results and discussion

### 6.1. Classical DoE results

Significance testing was carried out against the inherent stochasticism due to unsteadiness for the scalene ellipsoid and against the injected error for the DARPA SUBOFF model. This involved an F-test on each factor against the error in the preliminary regression models at a 6.5% level of significance. Insignificant terms were dropped from the preliminary models and newer, more accurate regression models for each output quantity of interest were fashioned. The final form of the regression model for $C_Z$ of the DARPA SUBOFF model as a cubic function of $\alpha$, $\beta$, and various interaction terms is shown by Eq. (10). Likewise, the final regression models of $C_Z$ and $C_Y$ for the scalene ellipsoid are shown by Eqs. (11) and (12), respectively.

$$\begin{aligned} C_{Z_{SO}} &= (1.0025\times10^{-5}) - (9.70794\times10^{-5}\alpha) \\ &\quad - (4.23273\times10^{-7}\beta) + (4.74999\times10^{-8}\alpha\beta) + \\ &\quad (7.51526\times10^{-8}\alpha^2) - (4.42186\times10^{-7}\beta^2) \\ &\quad - (1.49511\times10^{-6}\alpha\beta^2) - (1.63054\times10^{-6}\alpha^3) \end{aligned} \tag{10}$$

$$\begin{aligned} C_{Z_{EL}} &= (6.56441\times10^{-5}) - (7.89135\times10^{-4}\alpha) \\ &\quad + (2.53750\times10^{-6}\beta) - (9.27757\times10^{-6}\alpha\beta) - \\ &\quad (1.28967\times10^{-6}\alpha^2) - (1.55608\times10^{-6}\beta^2) \\ &\quad - (2.88818\times10^{-6}\alpha\beta^2) - (3.68148\times10^{-6}\alpha^3) \end{aligned} \tag{11}$$

$$\begin{aligned} C_{Y_{EL}} &= (6.75017\times10^{-5}) + (1.75589\times10^{-6}\alpha) \\ &\quad - (2.93189\times10^{-4}\beta) + (5.0057\times10^{-7}\alpha\beta) - \\ &\quad (1.25388\times10^{-6}\alpha^2) + (1.31612\times10^{-6}\beta^2) \\ &\quad - (6.9502\times10^{-7}\alpha^2\beta) + (1.68044\times10^{-6}\beta^3) \end{aligned} \tag{12}$$

The pure quadratic terms, $\alpha^2$ and $\beta^2$, as well as the second-order interaction term, $\alpha\beta$, for both $C_Z$ and $C_Y$ were found to be insignificant to the models, however, were included to maintain term hierarchy. This is a common practice associated with ANOVA and was done because higher order terms, $\alpha^3$ and $\alpha\beta^2$ for $C_Z$ and $\beta^3$ and $\alpha^2\beta$ for $C_Y$ were found to be significant to the models. A residual analysis was performed on these new regression models to determine if our three assumptions of ANOVA were maintained through the significance testing. The normal probability plot of $C_Z$ for both the SUBOFF model and the scalene ellipsoid are shown in Figs. 18 and 19 and the normal probability plot of $C_Y$ for the ellipsoid is shown in Fig. 20, all of which validate the normality assumption.

Figs. 21 and 22 show the residuals plotted against the predicted values to determine whether the variance is constant and structureless. Based off these plots, it can be ascertained that all assumptions of ANOVA have been validated. The third assumption is inherent to the experimental procedure since all CFD runs were conducted separately, it is intuitive that the run order has no effect on the residuals. There were a few points on the two standard deviation limit in Figs. 21 and

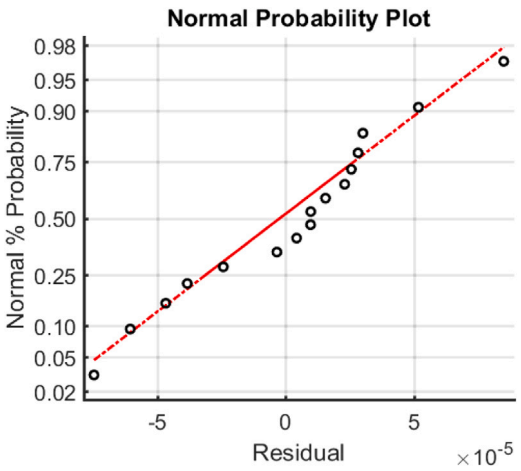


**Fig. 18.** DARPA SUBOFF — normal probability plot of $C_Z$, Final (line reflects normal distribution).

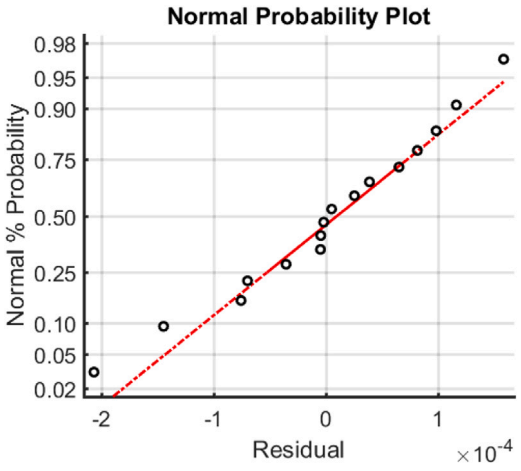


**Fig. 19.** Ellipsoid — normal probability plot of $C_Z$, final (line reflects normal distribution).

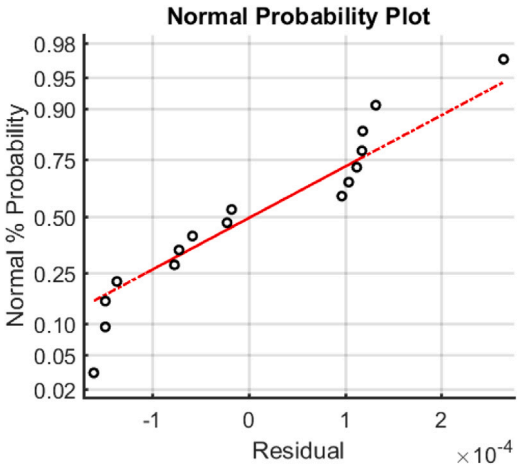


**Fig. 20.** Ellipsoid — normal probability plot of $C_Y$, final. Line reflects normal distribution.

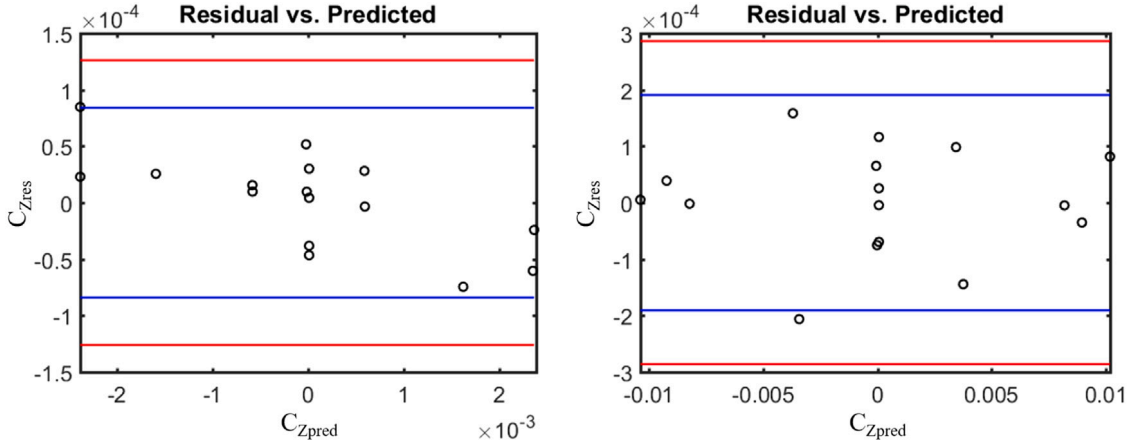


**Fig. 21.** Ellipsoid — residuals vs. predicted for $C_Z$, final (lines are 2 and 3 residual standard deviations). Left — DARPA SUBOFF, right — ellipsoid.

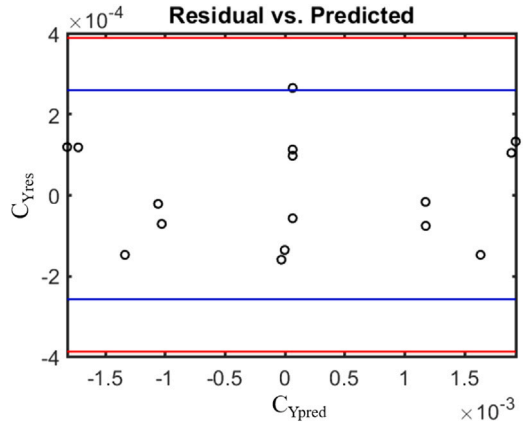


**Fig. 22.** Ellipsoid — residuals vs. predicted for $C_Y$, final. Lines are 2 and 3 residual standard deviations.

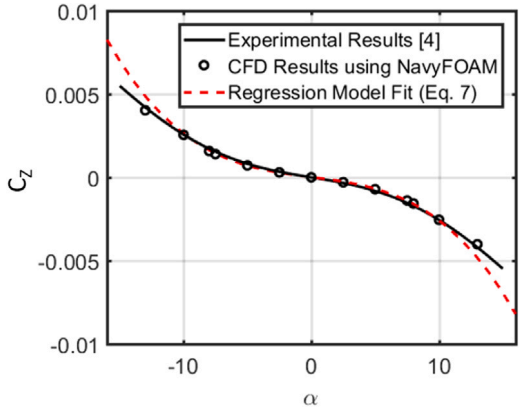


**Fig. 23.** DARPA SUBOFF — $C_Z$ regression model fit.

**Table 4**
Regression model fit summary.

| Model | $R^2_{ADJ}$ | $R^2_{PRED}$ |
|---|---|---|
| $C_{Z_{SO}}$ | 0.9996 | 0.9979 |
| $C_{Z_{EL}}$ | 0.9995 | 0.9987 |
| $C_{Y_{EL}}$ | 0.9791 | 0.7979 |

22 but are inconsequential to this analysis and are not considered outliers. With the residual analysis producing satisfactory statistical results which validated that the ANOVA assumptions were not violated, the regression models can confidently be assumed to adequately represent the data in a statistical sense.

Regression model fit statistics were calculated for all three regression models and are shown in Table 4. The adjusted $R^2$ was used to show the amount of reduction in the variability of the experimental data points by using the regression models (Montgomery, 2017). Higher values indicate a better representation of the experimental data by the regression model. It is scaled to the number of regressors making it an unbiased estimate which will not inflate with the number of model terms, contrary to the unadjusted $R^2$.

As it can be seen from Table 4, all three regression models represent the experimental data quite well. A look at the predicted $R^2$ values indicate the predictive capabilities of the regression model in the experimental design region, accounting for the variability in predicting new data points within the experimental domain (Montgomery, 2017). The predicted $R^2$ values are calculated using a leave-one-out method where a data point is excluded from the data set used to fit the model. This model is then used to predict the data point left out and the error is determined. This is done for all possible subsets and these errors are added together. The values for $C_Z$ for both the DARPA SUBOFF model and the scalene ellipsoid model indicate that the regression models, Eqs. (10) and (11), account for over 99% of the variability in predicting new observations. Eq. (12), however, only accounts for approximately 80% of the variability in predicting new observations as opposed to the 98% of the variability in the original experimental data. The $C_Z$ regression models for both geometries, Eqs. (10) and (11), were used to predict responses at $\alpha$ values both in the experimental domain and outside of the experimental domain to determine the regression model's extrapolation capabilities. The parameter $\beta$ was held constant at 0° to perform a direct comparison to the published data and to the CFD validation points from Section 5.2.2. The range of $\alpha$ was chosen to be from −16° to 16° with 1° increments and the results of this analysis have been plotted in Figs. 23 and 24. The same was done for Eq. (12), however, $\alpha$ was held constant and $\beta$ ranged from −16° to 16° in increments of 1°, as shown in Fig. 25.

It should be reiterated that the experimental domain ranges from −8° to 8°, meaning that Eqs. (10) and (11) not only show exemplary prediction performance within the experimental domain, but these regression models can also be used outside of the experimental domain and will extrapolate well up to approximately 13° for DARPA SUBOFF and 16° for the scalene ellipsoid. Fig. 25 not only shows ambiguity between the regression model fit and the CFD data, but also between the regression model fit and the data published by Granlund (2009) which was noted in Section 5.2.2. The regression model was not made to fit the published data, however, it was made to fit the CFD data obtained from simulating the design points in Table 2.

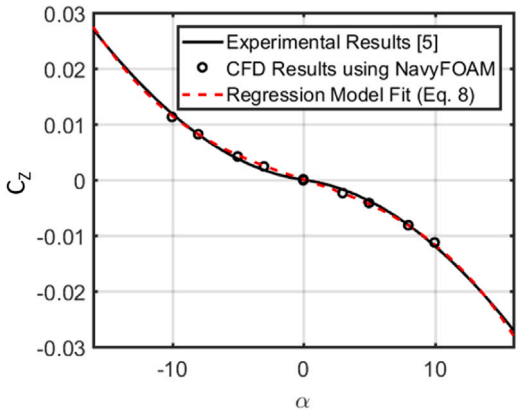


**Fig. 24.** Ellipsoid — $C_Z$ regression model fit.

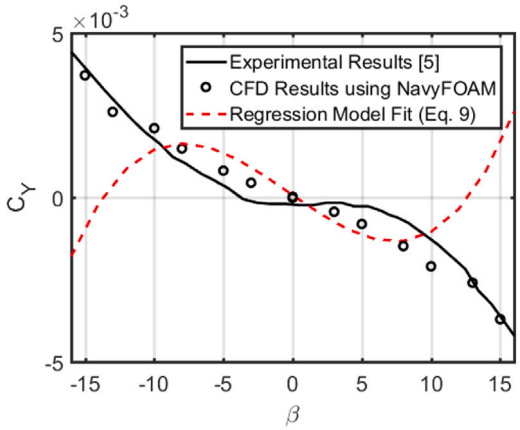


**Fig. 25.** Ellipsoid — $C_Y$ regression model fit.

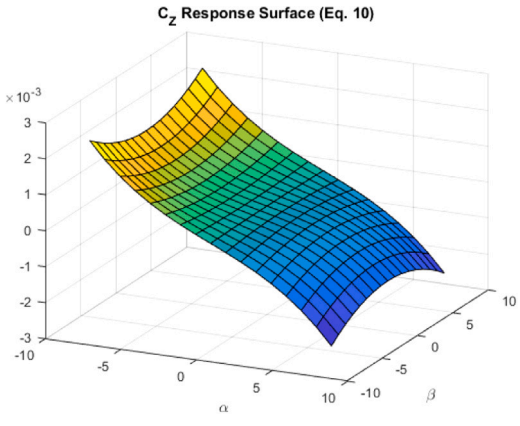


**Fig. 26.** DARPA SUBOFF — $C_Z$ response surface.

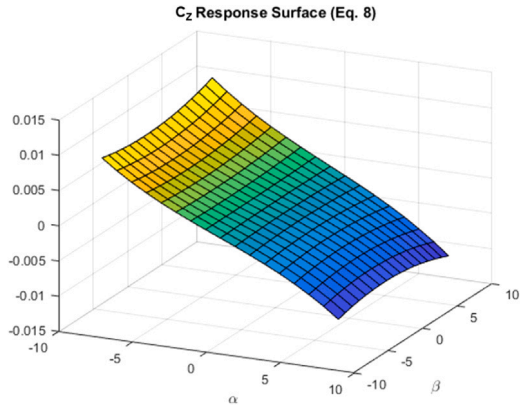


**Fig. 27.** Ellipsoid — $C_Z$ response surface.

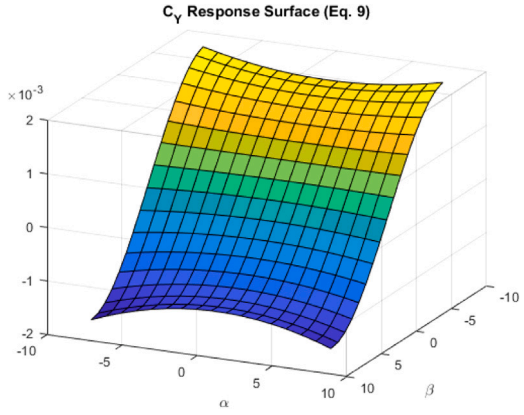


**Fig. 28.** Ellipsoid — $C_Y$ response surface.

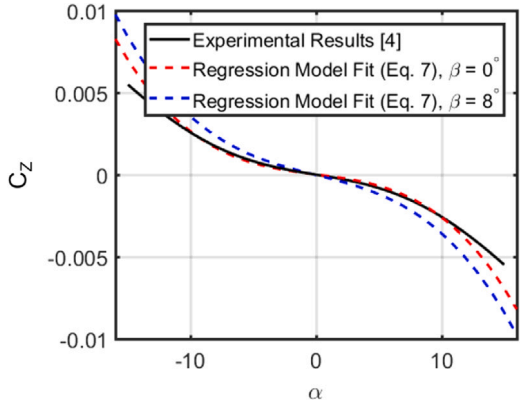


**Fig. 29.** DARPA SUBOFF — $C_Z$ Regression Model Fit at Different Values of $\beta$.

With that, it can be said that Eq. (12) does represent the CFD data to some level of accuracy within the experimental domain but does not extrapolate well and has opposite curvature to that of the CFD validation points and to the published data beyond the extents of the experimental domain. Response surfaces within the experimental region for the three regression models, Eqs. (10)–(12), were also created and are shown in Figs. 26–28.

These response surfaces aid in the visualization of the effects on the response from varying both $\alpha$ and $\beta$. The main purpose of this research was to analyze and compare modeling procedures with their respective experimental design spaces and not necessarily to perform a proper fluid dynamic study so the results of the response surface will only briefly be mentioned. Figs. 26 and 27 show that for both geometries, $\beta$ acts to parabolically augment the normal force at $\alpha$ angles greater or less than zero as the total angle of incidence is increased. The side force of the ellipsoid is shown to exhibit a similar behavior in Fig. 28. This effect can be seen more in depth for $C_Z$ of the SUBOFF model in Fig. 29.

Comparing Figs. 26 and 27 shows that the normal force augmentation from $\beta$ is less for the ellipsoid than it is for SUBOFF. Fig. 28 shows that $\alpha$ has a very similar effect on the side force as $\beta$ does on the normal force. Further, for the ellipsoid, the rolling moment coefficient should also be investigated for a proper fluid dynamic study.

**Table 5**
Artificial neural network model fit summary.

| Model | $R^2_{ADJ}$ | $R^2_{PRED}$ |
|---|---|---|
| $C_{Z_{SO}}$ | 0.9936 | 0.9932 |
| $C_{Z_{EL}}$ | 0.9989 | 0.9992 |
| $C_{Y_{EL}}$ | 0.9523 | 0.9553 |

### 6.2. Modern DoE results

A similar study was conducted on the ANNs that were fit to the experimental data using the procedure outlined in Section 4.D. These ANN's are shown by Eqs. (13)–(15). Since the ANN structure depends on activation functions, significance testing was not necessary for model formulation.

$$C_{Z_{SO}} = 6.6140 - 34.240V_1 - 15.480V_2 + 2.6978V_3 + 10.571V_4 - 70.852V_5 \tag{13}$$

$$V_1 = \tanh(0.047591 + 0.003124\alpha + 0.0032646\beta)$$
$$V_2 = \tanh(-0.057567 + 0.0060278\alpha + 0.007707\beta)$$
$$V_3 = \tanh(-0.0006128 + 0.0098289\alpha - 0.10423\beta)$$
$$V_4 = \tanh(-0.011212 + 0.0081894\alpha + 0.0090236\beta)$$
$$V_5 = \tanh(0.0816 - 0.0012298\alpha - 0.002114\beta)$$

$$C_{Z_{EL}} = 0.00941 + 0.000204V_1 + 0.0045V_2 - 0.0064V_3 - 0.00114V_4 + 0.0412V_5 \tag{14}$$

$$V_1 = \tanh(0.13663 + 0.26006\alpha - 0.21853\beta)$$
$$V_2 = \tanh(-0.47072 - 0.068779\alpha - 0.12504\beta)$$
$$V_3 = \tanh(-0.29733 - 0.1275\alpha - 0.10423\beta)$$
$$V_4 = \tanh(-0.12726 - 0.191737\alpha + 0.29288\beta)$$
$$V_5 = \tanh(-0.23350 - 0.03803\alpha + 0.0035824\beta)$$

$$C_{Y_{EL}} = 0.002681 - 0.01517V_1 - 0.000114V_2 + 0.0053V_3 - 0.01572V_4 - 0.00101V_5 \tag{15}$$

$$V_1 = \tanh(-0.44783 + 0.10746\alpha - 0.060789\beta)$$
$$V_2 = \tanh(-0.56462 - 0.14723\alpha - 0.15626\beta)$$
$$V_3 = \tanh(-0.82391 + 0.14583\alpha - 0.045996\beta)$$
$$V_4 = \tanh(0.32526 - 0.057428\alpha + 0.054880\beta)$$
$$V_5 = \tanh(0.78819 - 0.024872\alpha + 0.14822\beta)$$

Equations (13)–(15) can be seen to have highly nonlinear combinations of inflow conditions without giving much insight into which effects are actually significant to the response. Looking at the fit statistics for each ANN in Table 5 not only indicates that Eqs. (13)–(15) represent the experimental data quite well, but it also indicates that these ANN's have good prediction capabilities within the experimental region.

These results are further validated by Figs. 30–32 which use the ANNs, Eqs. (13)–(15), to predict responses for the same input parameters as for Figs. 23–25. Fig. 30 verifies the model fit of Eq. (13) and also shows that it can extrapolate outside of the experimental design region quite well.

This ANN is very comparable in accuracy to its regression model counterpart, Eq. (10). Figs. 31 and 32, however, show that although the ANN's fit with good accuracy compared with the experimental data points, they cannot be extrapolated outside of the experimental domain. Fig. 32 shows that the ANN for $C_Y$ actually has better accuracy than its regression model counterpart within the experimental domain. The ANN was modeled to fit the CFD simulation data at the design points in Table 3, which it does quite well. The normal force augmenting effects of $\beta$ for SUBOFF can again be seen in the response surface shown in Fig. 33. In comparison to the regression model counterpart in Fig. 26, the ANN seems to over-exaggerate these effects for negative angles. Fig. 34 shows that the normal force augmenting effects of $\beta$ are actually skewed towards one side which can be thought of as a slight misprediction since the geometry is symmetric.

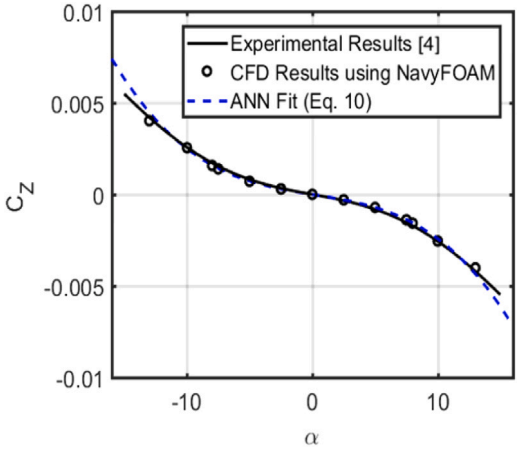


**Fig. 30.** DARPA SUBOFF — $C_Z$ ANN model fit.

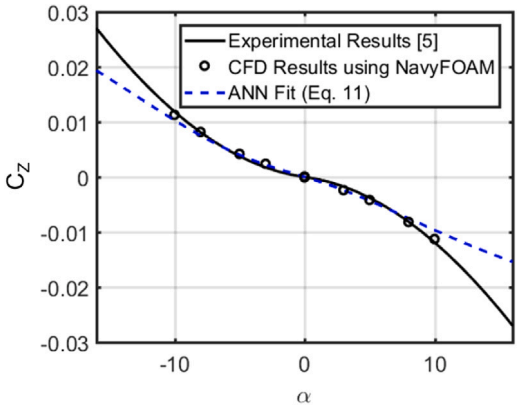


**Fig. 31.** Ellipsoid — $C_Z$ ANN model fit.

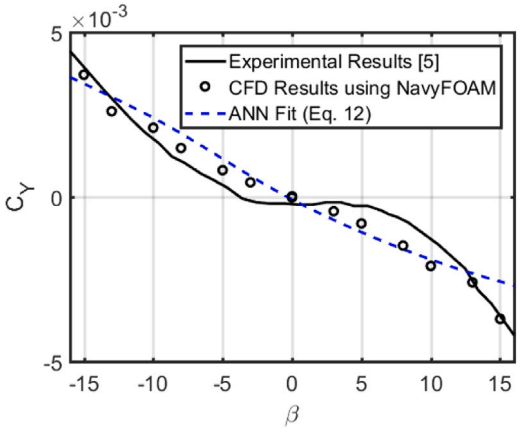


**Fig. 32.** Ellipsoid — $C_Y$ ANN model fit.

The response surfaces for Eqs. (14) and (15) have also been plotted in Figs. 35 and 36 for comparison purposes.

The response surface for $C_Z$ of the ellipsoid shows that there is also an over exaggeration of the normal force augmentation due to $\beta$, skewed towards negative $\alpha$ values. The response surface for $C_Y$

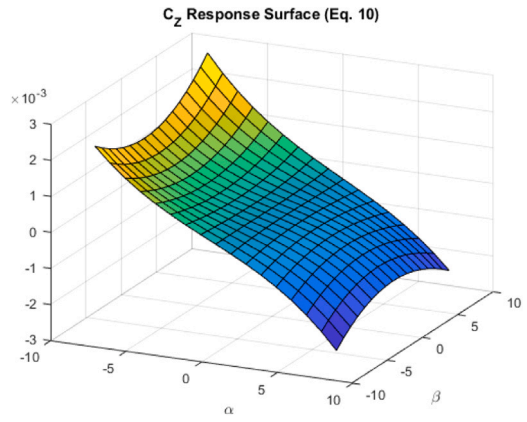


**Fig. 33.** DARPA SUBOFF — $C_Z$ response surface, ANN.

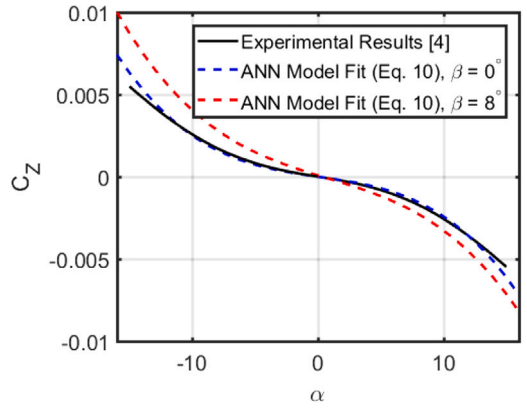


**Fig. 34.** DARPA SUBOFF — $C_Z$ ANN model fit at different values of $\beta$.

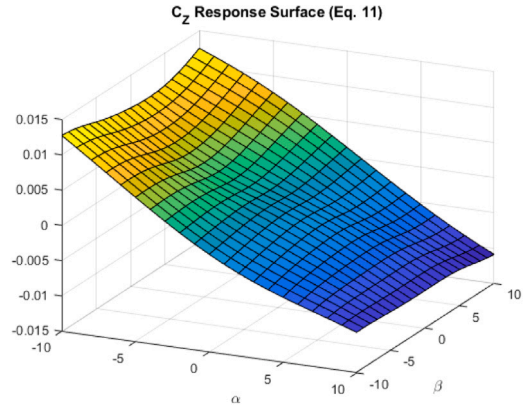


**Fig. 35.** Ellipsoid — $C_Z$ response surface, ANN.

shows that the ANN lacks symmetric behavior in both $\alpha$ and $\beta$ which is incorrect.

It is difficult to definitively make a statement pertaining to predictions of $C_Y$ for both the regression model and the ANN, Eqs. (12) and (15), respectively. This is due to the issue of small numbers, with the response, $C_Y$, being an order of magnitude less than $C_Z$ values and only one order of magnitude larger than the residuals. It is believed that there is not enough variance in the data for general trends to be predicted which is somewhat alluded to by the lower predicted $R^2$ value of the regression model, Eq. (12), in Table 4. Although the ANN, Eq. (15), does a better job fitting the data and predicting the experimental design points, noted by a high predicted $R^2$, it essentially

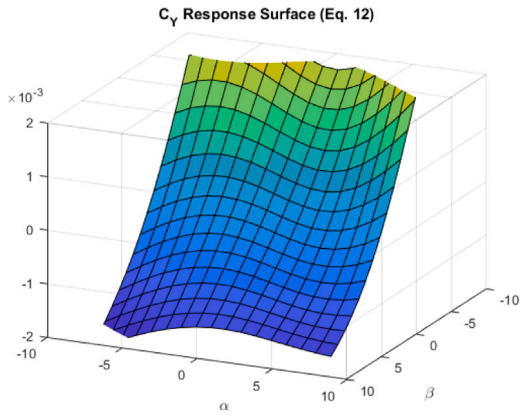


**Fig. 36.** Ellipsoid — $C_Y$ response surface, ANN.

interpolates the experimental design points and does not follow any particular trend inherent to the underlying physics, as noted by Fig. 36. The regression model, Eq. (12), on the other hand, exhibits symmetric side force augmentation from $\alpha$, as expected.

## 7. Conclusions

Throughout this work, two modeling techniques, nonlinear polynomial regression modeling and artificial neural network modeling, coupled with DoE techniques commonly used by each respective modeling procedure, have been analyzed and compared. Both classical and modern DoE were used to develop experimental design regions consisting of different combinations of inflow conditions, namely $\alpha$ and $\beta$. Nonlinear polynomial regression modeling and artificial neural network modeling were employed to map these inputs to the normal force and side force coefficients, $C_Z$ and $C_Y$, of the DARPA SUBOFF model and the Virginia Tech scalene ellipsoid model. NavyFOAM was used to perform CFD simulations on both geometries at the combinatorial inflow conditions prescribed by each experimental design region.

Significance testing and model reduction produced the final nonlinear polynomial regression model forms based upon the experimental design region created using classical DoE. An ANN was also fit to the CFD data produced by simulating points in the design space created using modern DoE. Fit statistics were calculated which showed that both the regression models and the ANNs fit the experimental data quite well, however, the ANN outperformed the regression model for $C_Y$ of the ellipsoid. Response surfaces were also generated to visualize general response trends at combinatorial inflow conditions which indicated that $\beta$ had an augmenting effect on the normal force coefficient for both geometries and $\alpha$ had an augmenting effect on the side force coefficient of the ellipsoid.

Both modeling techniques were shown to each have their own advantages. Classical DoE coupled with ANOVA and regression modeling adequately modeled the data with more rigorously spaced experimental design points. Regression modeling also gave much more insight into the physical flow phenomena, signifying which inflow parameters and parameter interactions were insignificant to the regression model.

The ANN modeling technique was shown to be more accurate with respect to the experimental design points. This is, in part, due to the uniform distribution of the design points inherent to the UD and the highly nonlinear structure of the ANN. These modern DoE methods and modeling procedures associated with them are generally considered interpolation methods, giving no insight to the underlying physics. For this reason, they are typically used for purely deterministic problems and more data points enhance their prediction capabilities.

In general, Classical DoE methodology coupled with ANOVA and regression modeling is preferred by the author in cases where the number of design points is limited but this methodology may not be used

unless the assumptions inherent to ANOVA are validated. Steps may be taken to inject stochasticism in a typically deterministic procedure, however, must be done so with care as to not over inflate the variance. The added error must also be consistent with the physical phenomena being simulated. In this case, wind tunnel testing error was added to the DARPA SUBOFF data since the CFD is simulating and being compared with wind/water tunnel testing data.

The artificial neural network has proven its capabilities in this paper by exemplifying excellent accuracy with respect to the experimental design points. Only 16 points were used to provide a more direct comparison to classic DoE methodology, however, more points can be added to the UD which would decrease the overall error between the meta-model and the experimental data, thus increasing the prediction accuracy over the experimental design space. If the number of experimental design points is not limited, ANN's may be used to produce better accuracy of the design region through interpolation.

Application of these methods to incorporate the affects of appendages and rotational velocities are ongoing and will be the focus of future publications.

## CRediT authorship contribution statement

**Christopher S. Thurman:** Methodology, Software, Validation, Formal analysis, Writing - original draft. **J. Ryan Somero:** Conceptualization, Formal analysis, Writing - review & editing, Supervision.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

This work was supported by funding from Newport News Shipbuilding's internal research and development program as well as support from the Department of Defense High Performance Computing Modernization Office (HPCMO) Computational Research and Engineering Acquisition Tools and Environments (CREATE) Program.

## References

Alpaydin, E., 2014. Introduction to Machine Learning. MIT press.

Bossert, D.E., Morris, S.L., Hallgren, W.F., Yechout, T.R., 2003. Introduction to Aircraft Flight Mechanics: Performance, Static Stability, Dynamic Stability, and Classical Feedback Control. American Institute of Aeronautics and Astronautics.

Box, G.E., Hunter, W.G., Hunter, J.S., et al., 1978. John Wiley and sons New York.

Cleaver, T.A., Gutman, A.J., Martin, C.L., Reeder, M.F., Hill, R.R., 2016. Using design of experiments methods for applied computational fluid dynamics: A case study. Qual. Eng. 28 (3), 280–292.

Da Ronch, A., Panzeri, M., Abd Bari, M., d'Ippolito, R., Franciolini, M., 2017. Adaptive design of experiments for efficient and accurate estimation of aerodynamic loads. Aircr. Eng. Aerosp. Technol. 89 (4), 558–569.

Doane, S.R., Landman, D., 2012. A wind tunnel investigation of ship airwake/rotor downwash coupling using design of experiments methodologies. In: 50th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition. Nashville, Tennessee. January 2012.

Fang, K.-T., Li, R., Sudjianto, A., 2005. Design and Modeling for Computer Experiments. Chapman and Hall/CRC.

Granlund, K.O., 2009. Steady and Unsteady Maneuvering Forces and Moments on Slender Bodies (Ph.D. thesis). Virginia Tech.

Hess, D., 2002. Using recursive neural networks for blind predictions of submarine manoeuvres. In: 24th Symposium on Naval Hydrodynamics. Fukuoka, Japan. July 2002.

Hickernell, F., 1998. A generalized discrepancy and quadrature error bound. Math. Comp. Amer. Math. Soc. 67 (221), 299–322.

Kim, S.-E., Shan, H., Miller, R., Rhee, B., Vargas, A., Aram, S., Gorski, J., 2017. A scalable and extensible computational fluid dynamics software framework for ship hydrodynamics applications: NavyFOAM. Comput. Sci. Eng. 19 (6), 33–39.

Landman, D., Simpson, J., Brian, H., Sumner, T., 2002. Use of Design of Experiments in Wind Tunnel Testing of Performance Automobiles. SAE Technical Paper 2002-01-3313.

Mohsen, M., Eltaher, H.M., Sharaf, A.M., El-demerdash, S.M., 2015. A systematic approach for the study and analysis of vehicle dynamics using design of experiments. In: 16th International Conference on Aerospace Sciences & Aviation Technology. ASAT. Cairo, Egypt. May 2015.

Montevechi, J.A.B., Pinho, A., Leal, F., Marins, F., 2007. Application of design of experiments on the simulation of a process in an automotive industry. In: Proceedings of the 2007 Winter Simulation Conference. pp. 1601–1609.

Montgomery, D.C., 2017. Design and Analysis of Experiments. John wiley & sons.

Murphy, P.C., Landman, D., 2015. Experimental design for complex VTOL aircraft with distributed propulsion and tilt wing. In: 2015 SciTech Forum: AIAA Atmospheric Flight Mechanics Conference. Kissimmee, Florida. January 2015.

Myers, R.H., Montgomery, D.C., Anderson-Cook, C.M., 2016. Response Surface Methodology: Process and Product Optimization using Designed Experiments. John Wiley & Sons.

Nazalla, D., Mollaghasemi, M., Anderson, C.D., 2006. A simulation-based evaluation of the cost of cycle time reduction in agere systems wafer fabrication facility - a case study. Int. J. Prod. Econ. 100 (2), 300–313.

Racine, B., Paterson, E., 2005. CFD-based method for simulation of marine-vehicle maneuvering. In: 35th AIAA Fluid Dynamics Conference and Exhibit. pp. 4904.

Reinholtz, C., Yoder, D., Landman, D., Simpson, J., 2012. A design of experiments approach to wind tunnel balance calibration at arnold engineering development complex. In: 51st AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition. Grapevine, Texas. January 2013.

Roddy, R.F., 1990. Investigation of the Stability and Control Characteristics of Several Configurations of the DARPA SUBOFF Model (DTRC Model 5470) from Captive-Model Experiments. Tech. rep.. David Taylor Research Center Bethesda MD Ship Hydromechanics Dept.

Sharma, S., 2017. Activation functions in neural networks. https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6.

Simonsen, C.D., Otzen, J.F., Klimt, C., Larsen, N.L., Stern, F., 2012. Maneuvering predictions in the early design phase using CFD generated PMM data. In: 29th Symposium on Naval Hydrodynamics. pp. 26–31.

Somero, J.R., 2010. Computational Simulations of a Non-Body of Revolution Ellipsoidal Model Utilizing RANS (M.S. thesis). Virginia Tech.

Stringer, D., 2017. Airfoil 360:s1210 Raw Wind Tunnel Data -2017. Mendeley Data, https://data.mendeley.com/datasets/yn4hxc8m8y/1. Retrieved Nov. 2018..

Walia, A.S., 2017. Activation functions and it's types - which is better? https://towardsdatascience.com/activation-functions-and-its-types-which-is-better-a9a5310cc8f.

Wilcox, D.C., 2008. Formulation of the $k-\omega$ turbulence model revisited. AIAA J. 46 (11), 2823–2838.

Wilcox, D.C., et al., 2006. Turbulence Modeling for CFD, Vol. 3. DCW industries La Canada, CA.

Yondo, R., Andrés, E., Valero, E., 2018. A review on design of experiments and surrogate models in aircraft real-time and many-query aerodynamic analyses. Prog. Aerosp. Sci. 96, 23–61.