EasyChair Preprint

№ 14537

Enhanced Prediction of Airfoil's Drag Coefficient Using Curve Fitting and Artificial Neural Network Preprint Version Information

Mohssen Elshaar and Naef Qasem

August 26, 2024

Enhanced prediction of airfoil's drag coefficient using curve fitting and artificial neural network

Mohssen E. Elshaar, Naef A.A. Qasem

Abstract

This study explores the application of Artificial neural networks (ANNs) for predicting the aerodynamic coefficients of airfoils, with a focus on the drag coefficient (C_D), as the literature has not predicted it as precisely as other aerodynamic coefficients. A novel quadratic fitting function is introduced to improve the accuracy of C_D predictions. Two datasets, DI and DII, with varying ranges of Mach numbers, were prepared, and the performance of the ANN was evaluated. Model I was trained on Dataset I (Mach 0.1 to 0.3), while Model II was trained on Dataset II (Mach 0.1 to 0.8). The results indicate that a larger and more diverse dataset significantly enhances the predictive capabilities of the model. Additionally, the model's ability to generalize to airfoils and flight conditions outside the training data was tested, revealing the generalization power of the model.

Keywords: Airfoil optimization; Drag coefficient; Artificial neural networks (ANNs); Polynomial regression; Aerodynamic prediction; Data-driven models

Nomenclature

C_L	Lift coefficient
C_M	Moment coefficient
C_D	Drag coefficient
C_{D0}	Drag coefficient at zero Lift and zero Angle of Attack
$C_{D_{\text{approximate}}}$	Approximate drag
k	Lift coefficient weight
α	Angle of attack
ζ	Angle of attack fitting coefficient

1. Introduction

Airfoil design is essential for aircraft aerodynamics. An airfoil-shaped object in motion produces a lift perpendicular to its direction and drag parallel to it. Aerospace engineers use airfoil geometry to estimate lift, drag, and other parameters like the center of pressure. These calculations help evaluate a wing's aerodynamic performance, emphasizing the importance of finding the right airfoil shape (Narendra and Parthasarathy (1990)). Advances in computational engineering offer efficient ways to compute aerodynamic coefficients and optimize airfoil shapes. The rise of computer science and AI has led to reliable data-driven models for aerodynamics, capable of accurately predicting solutions to complex problems (Hunt et al. (1992)). While Computational fluid dynamics (CFD) is powerful, it's often time-consuming and relies on physical laws, limiting its use in real-time applications. In contrast, data-driven models like Artificial neural networks (ANNs) can rapidly map inputs to outputs without needing detailed system knowledge (Calise and Rysdyk (1998)). Thus, ANNs are ideal for flight control and dynamic system identification (Narendra and Parthasarathy (1990)).

Numerous studies have explored using deep learning techniques, particularly Convolutional neural networks (CNNs), to predict flow fields and aerodynamic force coefficients of airfoils. Multiple CNNs were employed to learn lift coefficients across various airfoil shapes under different Angles of Attack (AoA), Mach numbers, and Reynolds

numbers, estimating the flow field over an airfoil without directly solving the Navier-Stokes equations (Zhang et al. (2018); Guo et al. (2016)). Moreover, CNNs were used to predict velocity and pressure fields over an airfoil and to anticipate non-uniform steady laminar flow in both 2D and 3D domains (Bhatnagar et al. (2019); Chen et al. (2020)).

Conversely, simpler ANN architectures were frequently employed for inverse airfoil design problems. Pressure distribution (Rai and Madavan (2001)) and 15 other design variables (Rai and Madavan (2000)) were fed as inputs for ANNs to design turbomachinery airfoils. The Eppler method (Huang et al. (1994)), which is a multipoint inverse airfoil design approach that predicts airfoil performance under various conditions by specifying its velocity distribution, was designed and assessed using ANNs. To optimize their shapes, a swarm-based method was combined with an Artificial Neural Network (ANN) on PARSEC airfoils (Khurana et al. (2008)). Additionally, neural networks were used to enhance the high lift performance of a multi-element airfoil (Greenman and Roth (1999)). A genetic algorithm was used along with a trained ANN to expedite the search in airfoil design space (Hacioglu (2007)) and (Xu et al. (2019)) optimized airfoil design in the presence of buffeting phenomena using neural networks. A unique approach for predicting steady turbulent aerodynamic fields was developed in (Sun et al. (2015)) using neural networks. Finally, ANNs were used to reconstruct models for improved prediction of lift coefficients and flow separation over airfoils (Li et al. (2020)).

In some studies, CNNs were combined with ANNs to optimize airfoil shape. For example, (Sekar et al. (2019)) introduced a novel sampling method for airfoils and wings using a deep Convolutional-Generative adversarial network (DC-GAN).

In (Moin et al. (2022)), ANNs are used to estimate the aerodynamic coefficients of various airfoils by training on normalized 2D coordinates of the airfoil geometry rather than traditional airfoil design parameters. Although the predicted lift and moment coefficients were of acceptable accuracy, the trained model could not accurately represent the drag coefficient trend. Moreover, the model was limited to low-speed flight conditions, i.e., (0.1 0.3).

This work is considered an extension to the literature work (Moin et al. (2022)) for predicting the aerodynamic performance of airfoils under subsonic conditions, where two main contributions are proposed:

- A quadratic fitting function that approximates the drag coefficient as a function of the lift coefficient and the angle of attack. This attempts to overcome the drag coefficient's poor prediction accuracy due to the usual irregularities with the drag coefficient trend.
- Extending the prediction model to work on a wider range of Mach numbers, i.e., (0.1 - 0.8).

2. Dataset preparation

2.1. Data generation

The NACA classification system provides a structured approach to describe these parameters. The National Advisory Committee for Aeronautics (NACA), later part of NASA, gathered extensive airfoil data, classifying them into NACA 4- and 5-digit series. A combination of both series is used as the data pool. To generate the coordinates of a wide range of airfoils, along with their respective experimental aerodynamic coefficients, Lift Coefficient C_L, Drag Coefficient C_D, and Moment Coefficient C_M, JavaFoil software was used. 560 NACA 4-digit and 1120 NACA 5-digit series Airfoil data were generated. These airfoils spanned a wide range of airfoil parameters. It was proved that combining both 4-digit and 5-digit data during training yielded a better network performance (Moin et al. (2022)).

The airfoils' coordinate points are interpolated at consistent intervals along the x-axis using cosine spacing, focusing on denser point distributions near the leading and trailing edges to accurately capture their shapes. The upper and lower surface points are represented as $y_{U,k}$ and $y_{L,k}$ respectively for all k within the range [1, N]. While the leading and trailing edges are set at (0, 0) and (1, 0), respectively, these fixed points are omitted from the dataset as they remain unchanged across all training examples. According to the literature (Moin et al. (2022)), setting N to 10 data points was found to give a better average performance.

Two datasets with similar operating conditions were prepared, differing only in the range of Mach numbers. The first dataset included combinations of angles of attack (AoAs) ranging from -10 to 10 degrees in steps of 1, Reynolds numbers ranging from 100,000 to 500,000 in steps of 100,000, and Mach numbers ranging from 0.1 to 0.3 in steps of 0.1. The second dataset had the same ranges for AoAs and Reynolds numbers but included Mach numbers ranging

Table 1: Sample fitting results for drag coefficient approximation, $C_{D_0} = 0.0124$, $k = -0.03218$, and $\zeta = 4.9931$

α (°)	C_L	C_D	$C_{D_{\text{estimate}}}$
-10	-0.334	0.1614	0.1609
-8	-0.442	0.10163	0.10345
-6	-0.459	0.05855	0.06037
-4	-0.349	0.03876	0.032809
-2	-0.228	0.00996	0.016804
0	0	0.00924	0.012393
2	0.228	0.00996	0.016804
4	0.349	0.03876	0.032809
6	0.459	0.05855	0.06037
8	0.442	0.10163	0.10345
10	0.334	0.1614	0.16090

from 0.1 to 0.8 in steps of 0.1. Consequently, the first dataset D_I consisted of 171,431 samples from the 4-digit airfoil data and 283,244 samples from the 5-digit airfoil data, while the second dataset D_{II} comprised 312,891 samples from the 4-digit airfoil data and 557,404 samples from the 5-digit airfoil data.

2.2. Drag coefficient approximation

Based on the observation that the drag coefficient in previous work followed a roughly parabolic trend, the proposed drag approximation incorporates three fitting parameters: a constant (C_{D_0}), a second-order lift coefficient gain (k), and a coefficient ζ that is multiplied by the squared angle of attack in radians. The parameters k and ζ act as weights, reflecting the contributions of the lift coefficient and the angle of attack to the approximation. The constant C_{D_0}, which is the Drag coefficient at zero Lift and zero AoA, serves as a relaxation factor to prevent the approximation from strictly following a parabolic curve, thereby reducing the risk of overfitting.

$$C_{D_{\text{approximate}}} = C_{D_0} + kC_L^2 + \zeta\alpha^2$$

The actual drag coefficient C_D is fitted using polynomial regression. The approximated drag average error over all samples was found to be 10% for the first dataset and 18% for the second dataset. Table 1 shows sample C_D vs $C_{D_{\text{approximate}}}$ fitting. Since at higher speeds the drag coefficient deviates slightly from the parabolic trend, the fitting model accuracy is deprecated, as shown in the contrast between Figs. A.3 and A.4 in Appendix A.

2.3. Model inputs and outputs

Along with the 10 coordinate points of the upper and lower airfoil surfaces, the Reynold's number, AoA, and Mach number are considered the features the model should train upon. This makes the size of the input layer $2N + 3 = 23$. The outputs of the model are the three aerodynamic coefficients in addition to the three fitting parameters proposed in this work. Tables 2 and 3 show a sample data instance of inputs (features) and outputs (required predictions).

Table 2: Sample Input/Features

$y_{U,1}$	..	$y_{U,10}$	$y_{L,1}$	..	$y_{L,10}$	Re	M	α
0.009814	..	0.001179	-0.009810	..	-0.001180	100000	0.1	-10

Table 3: Sample Outputs/Predictions

C_L	C_D	C_m	C_{D_0}	k	ζ
-0.334	0.1614	0.001	0.01239	-0.03218	4.9931

3. ANN model

3.1. Architecture

A neural network with three hidden layers and neuron configurations of $[512, 256, 128]$ is chosen for the model. Literature showed that both shallower and deeper models tend to perform worse. The input layer consists of 23 inputs, while the output layer has 6 outputs. The Rectified Linear Unit (ReLU) activation function was used across all layers for its efficiency and effectiveness in addressing vanishing gradients. Only the output layer used a linear activation function.

3.2. Training and validation

Two models are trained: Model I and Model II, using D_I and D_{II}, respectively. The training process followed the procedures in (Moin et al. (2022)), using the Adam optimizer with an initial learning rate of 0.0005, first and second moments at 0.9 and 0.999, and Mean Squared Error (MSE) as the loss function. Performance was evaluated using Root Mean Squared Error (RMSE) and R^2 values. The learning rate was reduced by 10% if the validation loss did not improve for 5 epochs. Each network was trained for 50 epochs with a batch size of 128 samples.

4. Results and discussion

Two primary tests were performed to assess the model's performance. The first involved applying the model to a sample test subset drawn from the main dataset, which the model had not encountered during training. The second test evaluated the model's performance on sample airfoils beyond the design domain it was trained on. The approximated predictions of C_D were evaluated versus the actual ground truth $C_{D_{GT}}$, and the approximated ground truth $C_{D_{\text{approximate\_GT}}}$.

4.1. Performance on test data

Table 4 illustrates the performance metrics for Model I. The R^2 values indicate strong correlations between the predicted and actual values for C_L and C_M, and a weak correlation when it comes to the direct prediction of C_D (0.7680), as expected. The fitting parameters C_{D_0}, k, ζ prediction was excellent, and thus, the prediction of the $C_{D_{\text{approximate}}}$ scored well against $C_{D_{\text{approximate\_GT}}}$ (0.8977). Notably, the predicted $C_{D_{\text{approximate}}}$ scored better against the actual $C_{D_{GT}}$ as well (0.8148), indicating that the approximated predictions are even closer to the actual values than the direct predictions.

Table 4: Model I Performance on Test Data

Metric	C_L	C_D	C_m	C_{D_0}	k	ζ	$C_{D_{\textbf{approximate}}}$	$C_{D_{\textbf{approximate/Actual}}}$
R^2	0.9985	0.7680	0.9903	0.8968	0.9253	0.9963	0.8977	0.8148
RMSE	0.0321	0.0125	0.0081	0.0040	0.0294	0.0861	0.0077	0.0111

When it came to evaluating the fitting solution on Model II, new insights came to prevail. Table 5 shows that the model performance on C_D predictions enhanced significantly (0.8125), while the scores of $C_{D_{\text{approximate}}}$ stayed the same – more or less.

Table 5: Model II Performance on Test Data

Metric	C_L	C_D	C_m	C_{D_0}	k	ζ	$C_{D_{\textbf{approximate}}}$	$C_{D_{\textbf{approximate/Actual}}}$
R^2	0.9963	0.8125	0.9490	0.8955	0.8640	0.9981	0.8714	0.7994
RMSE	0.0531	0.0169	0.0153	0.0072	0.0319	0.1145	0.0134	0.0175

The improved performance observed in Model II compared to Model I could be attributed to the larger dataset size. With approximately twice the number of samples, dataset DII provides a more comprehensive representation of flight conditions. The increased data granularity allows the model to capture more nuanced relationships between input variables and output predictions, leading to more accurate predictions of C_D. The consistent scores of $C_{D_{\text{approximate}}}$ (0.8714 and 0.7994) despite the improvements in Model II suggest that further training epochs might have enhanced the performance of estimations. Increasing the number of epochs could allow the model to further refine its learned representations and optimize its predictive capabilities.

4.2. Performance beyond the design domain

In (Moin et al. (2022)), four airfoils—NACA 0045, 2412, 6408, and 136138—were selected to evaluate the model's performance on airfoils that were not included in the training dataset. Figure A.5 reveals that NACA 0045 and 136138 are slightly outside the designated design space, whereas NACA 2412 and 6408 are entirely within it. These airfoils were tested under flight conditions that were also new to the network, specifically at Mach 0.25 in Model I, and Mach 0.25 and 0.65 in Model II, with a Reynolds number of 250,000.

Fig. 1 reveals the deprecation of Model I performance on the beyond design Airfoils, i.e. NACA 0045, and 136138. Neither the direct prediction of C_D, nor the estimated prediction $C_{D_{\text{approximate}}}$ follow the ground truth $C_{D_{\text{GT}}}$ trend. On NACA 2412 and 6406 the model seems to perform a little better on estimations, whereas the direct prediction of C_D remains poor. This suggests that Model I struggles to generalize effectively to new airfoils and flight conditions not represented in the training dataset. Table 6 gives an oversight of the performance of the Model I on the unseen Airfoils.

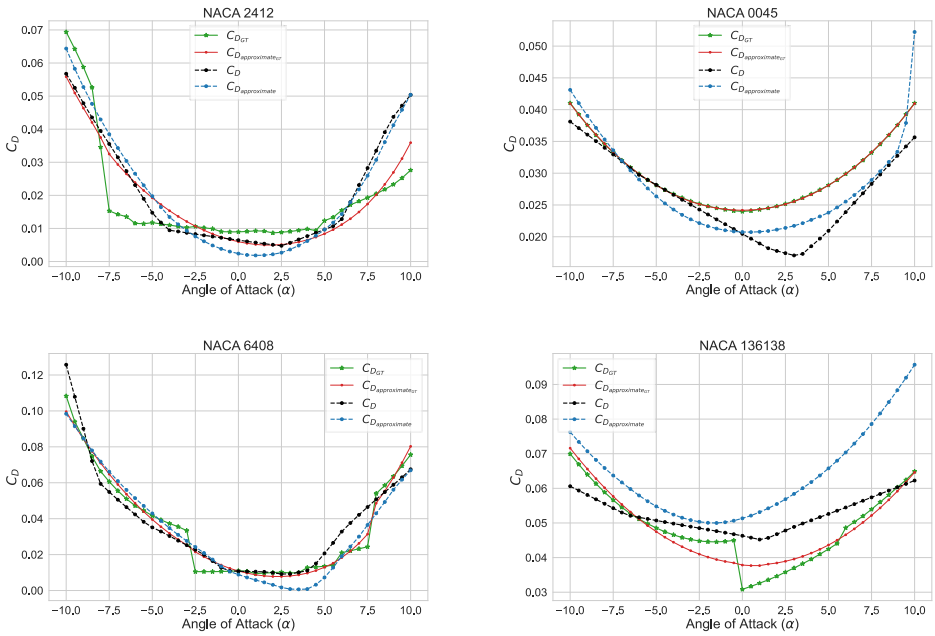
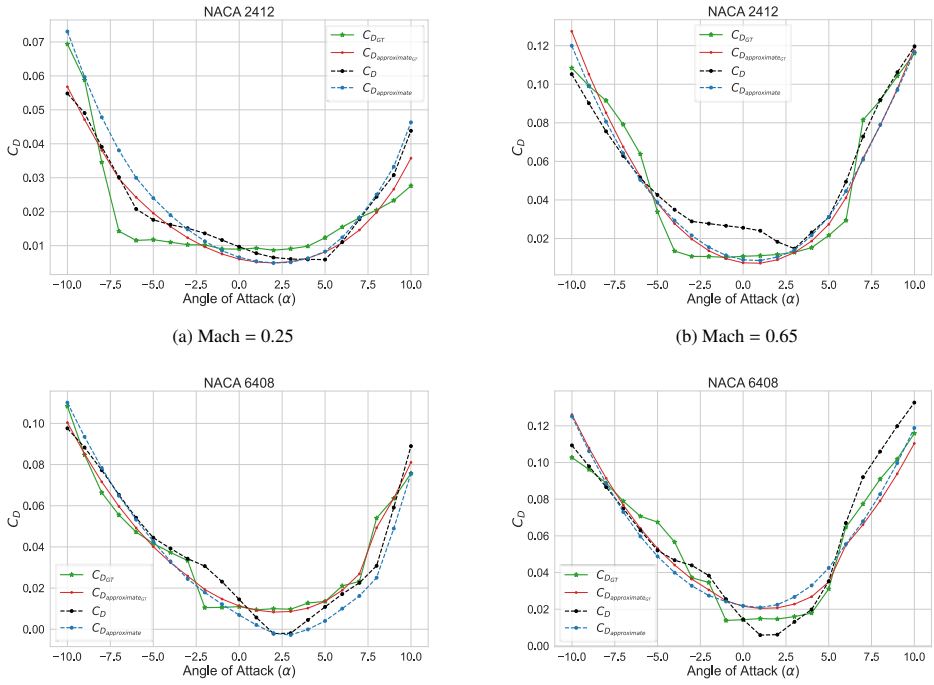


Fig. 1: $C_{D_{\text{GT}}}$, $C_{D_{\text{approximate\_GT}}}$, C_D, and $C_{D_{\text{approximate}}}$ versus AoA on NACA 2412, 6408, 0045, and 136138 for Model I. (Mach = 0.25 and Re = 250,000)



(a) Mach = 0.25 (b) Mach = 0.65

(c) Mach = 0.25 (d) Mach = 0.65

Fig. 2: $C_{D_{\text{GT}}}$, $C_{D_{\text{approximate\_GT}}}$, C_D, and $C_{D_{\text{approximate}}}$ versus AoA on NACA 2412 and 6408 for Model II. (Mach = 0.25, 0.65 and Re = 250,000)

Table 6: Oversight of Model I performance on NACA 2412, 6408, 136138, and 0045 for Model I. Mach = 0.25, Re = 250,000

NACA	Metric	C_L	C_D	C_m	$C_{D_{\text{approximate}}}$	$C_{D_{\text{approximate/Actual}}}$
0045	R^2	0.9990	0.3539	0.8647	0.5150	0.5147
	RMSE	0.0260	0.0045	0.0049	0.0037	0.0037
2412	R^2	0.9977	0.5536	Poor	0.7612	0.5613
	RMSE	0.0313	0.0107	0.0110	0.0069	0.0106
6408	R^2	0.9879	0.4233	0.6223	0.5031	0.4929
	RMSE	0.0618	0.0111	0.0293	0.0183	0.0187
136138	R^2	0.9938	0.4922	0.1567	Poor	Poor
	RMSE	0.0611	0.0073	0.0220	0.0148	0.015

For Model II, emphasis was placed on NACA 2412 and 6406, as Model I demonstrated promising performance on these airfoils. The extensive dataset D_{II} revealed its advantage once more, with Model II achieving high accuracy in both direct and estimated predictions. The rich data provided by D_{II} significantly enhanced the model's predictive capabilities, resulting in more precise approximations (Table 7 and Fig. 2).

Table 7: Oversight of Model II performance on NACA 2412 and 6408 for Model II. Mach = 0.25, 0.6 and Re = 250,000

MACH	NACA	Metric	C_L	C_D	C_m	$C_{D_{\text{approximate}}}$	$C_{D_{\text{approximate/Actual}}}$
0.25	**2412**	R^2	0.9987	0.7806	0.4632	0.8071	0.6528
		RMSE	0.0239	0.0075	0.0068	0.0064	0.0095
0.25	**6408**	R^2	0.9864	0.8770	0.5536	0.8922	0.8622
		RMSE	0.0665	0.0100	0.0321	0.0093	0.0106
0.65	**2412**	R^2	0.9933	0.9015	0.5291	0.9940	0.9372
		RMSE	0.0522	0.0125	0.0110	0.0030	0.0100
0.65	**6408**	R^2	0.9909	0.9230	0.4611	0.9850	0.9003
		RMSE	0.0615	0.0095	0.0358	0.0151	0.0148

5. Conclusion

This paper presents a comprehensive evaluation of using ANNs to predict aerodynamic coefficients, with a particular focus on the drag coefficient C_D. By introducing a quadratic fitting function, the study aimed to address the challenges in accurately predicting C_D. The performance of the models was assessed using two distinct datasets, differing in their range of Mach numbers. The findings revealed that the model trained on the larger and more diverse Dataset D_{II} exhibited enhanced predictive capabilities. Despite this, the models struggled to generalize effectively to airfoils and flight conditions outside the training data, particularly for direct C_D predictions. The results suggest that further training epochs and even larger datasets could improve the models' performance. Additionally, the finer sampling of Reynolds number, Mach number, and angle of attack (AoA) ranges will likely enhance the network's performance. The study underscores the promise of data-driven approaches in aerodynamic predictions and sets the stage for future research to refine these models and extend their applicability.

Acknowledgements

The authors thank the Department of Aerospace Engineering, KFUPM, for supporting this research.

Appendix A. Additional figures

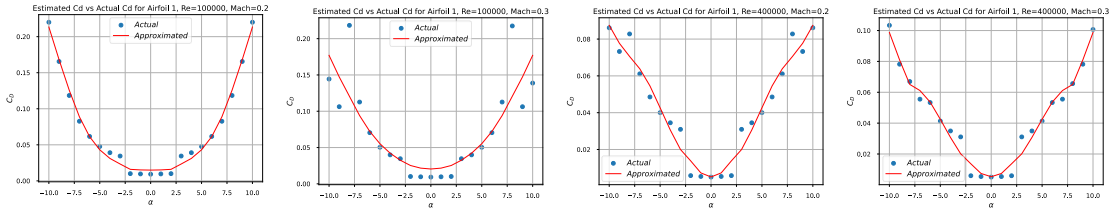


Fig. A.3: Dataset D_I Fitting samples showing excellent fitting on the drag coefficient trend

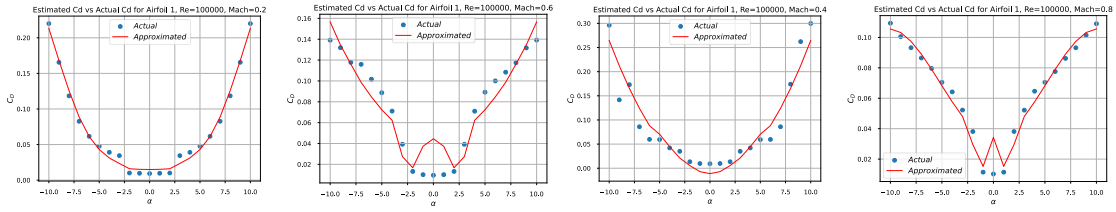


Fig. A.4: Dataset D_{II} Fitting samples exhibiting excellent fit for the drag coefficient trend at low Mach numbers, but reduced accuracy at higher Mach numbers. The drag coefficient trend diverges from the parabolic pattern, causing the approximating fitting function to be less accurate.

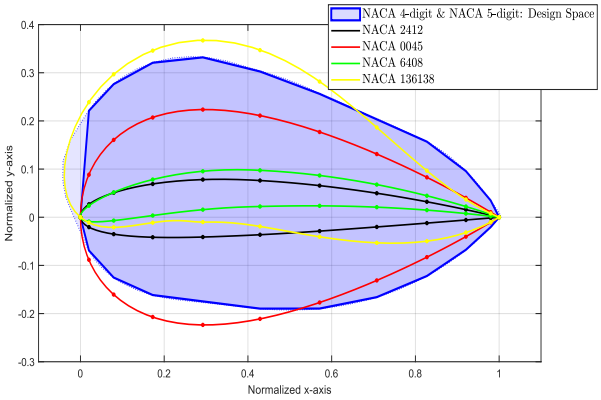


Fig. A.5: Test airfoils vs the training design space (borrowed from (Moin et al. (2022)))

References

Bhatnagar, S., Afshar, Y., Pan, S., Duraisamy, K., Kaushik, S., 2019. Prediction of aerodynamic flow fields using convolutional neural networks. Computational Mechanics 64(2), 525–545.

Calise, A. J., Rysdyk, R. T., 1998. Nonlinear adaptive flight control using neural networks. IEEE Control Systems Magazine 18(6), 14–25.

Chen, H., He, L., Qian, W., Wang, S., 2020. Multiple aerodynamic coefficient prediction of airfoils using a convolutional neural network. Symmetry 12(4), p. 544.

Greenman, R. M., Roth, K. R., 1999. High-Lift Optimization Design Using Neural Networks on a Multi-Element Airfoil. Journal of Fluids Engineering 121(2), 434–440.

Guo, X., Li, W., Iorio, F., 2016. Convolutional neural networks for steady flow approximation. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 481–490.

Hacioglu, A., 2007. Fast evolutionary algorithm for airfoil design via neural network. AIAA Journal 45(9), 2196–2203.

Hassoun, M. H., 1995. Fundamentals of Artificial Neural Networks. MIT Press.

Huang, S., Miller, L., Steck, J., 1994. An exploratory application of neural networks to airfoil design. 32nd Aerospace Sciences Meeting and Exhibit, p. 501.

Hunt, K., Sbarbaro, D., Zbikowski, R., Gawthrop, P., 1992. Neural networks for control systems—a survey. Automatica 28(6), 1083–1112.

Khurana, M., Winarto, H., Sinha, A., 2008. Application of swarm approach and artificial neural networks for airfoil shape optimization. 12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, p. 5954.

Li, J., Zhang, M., Martins, J. R. R. A., Shu, C., 2020. Efficient aerodynamic shape optimization with deep-learning-based geometric filtering. AIAA Journal 58(10), 4243–4259. [Online]. Available: https://doi.org/10.2514/1.J059254

Moin, H., Khan, H. Z. I., Mobeen, S., Riaz, J., 2022. Airfoil's Aerodynamic Coefficients Prediction using Artificial Neural Network. 2022 19th International Bhurban Conference on Applied Sciences and Technology (IBCAST), 175–182. doi: 10.1109/IBCAST54850.2022.9990112.

Narendra, K. S., Parthasarathy, K., 1990. Identification and control of dynamical systems using neural networks. IEEE Transactions on Neural Networks 1(1), 4–27.

Rai, M. M., Madavan, N. K., 2000. Aerodynamic design using neural networks. AIAA Journal 38(1), 173–182.

Rai, M. M., Madavan, N. K., 2001. Application of artificial neural networks to the design of turbomachinery airfoils. Journal of Propulsion and Power 17(1), 176–183.

Sekar, V., Zhang, M., Shu, C., Khoo, B. C., 2019. Inverse design of airfoil using a deep convolutional neural network. AIAA Journal 57(3), 993–1003.

Sun, G., Sun, Y., Wang, S., 2015. Artificial neural network based inverse design: Airfoils and wings. Aerospace Science and Technology 42, 415–428.

Xu, Z., Saleh, J. H., Yang, V., 2019. Optimization of supercritical airfoil design with buffet effect. AIAA Journal 57(10), 4343–4353. [Online]. Available: https://doi.org/10.2514/1.J057573

Zhang, Y., Sung, W. J., Mavris, D. N., 2018. Application of convolutional neural network to predict airfoil lift coefficient. AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, p. 1903.