Journal of Physics: Conference Series **2767** (2024) 052055 doi:10.1088/1742-6596/2767/5/052055

Opensource machine learning metamodels for assessing blade performance impairment due to general leading edge degradation

A. Castorrini[1,2], A. Ortolani[2], E. Minisci[3], M. S. Campobasso[2]

[1] Dept. of Mechanical and Aerospace Engineering, Sapienza University of Rome, Italy.
[2] School of Engineering, Lancaster University, Lancaster LA1 4YW, United Kingdom.
[3] University of Strathclyde, Department of Mechanical and Aerospace Engineering, Glasgow G1 1XQ, United Kingdom.

E-mail: alessio.castorrini@uniroma1.it

Abstract. Blades leading edge erosion can significantly reduce annual energy production of wind turbines. Accurate estimates of the resulting blade performance impairment are paramount to predict the resulting energy losses and enable cost-informed decisions on optimal maintenance and operational strategies, maximizing energy production and reducing maintenance costs. Computational Fluid Dynamics (CFD) is a robust approach for predicting the performance losses due to LEE. However, the impact of the damage on blade aerodynamics varies depending on damage pattern, extent and location. Therefore, direct CFD simulation of a sufficiently general set of damaged blades is computationally not viable in industrial applications, since the energy loss assessment needs to be performed for hundreds of turbines at many times of the wind farm operation. To address this issue, previous studies showed how CFD can be used to train machine learning metamodels of the perfomance of damaged blade sections, enabling the definition of multi-fidelity energy loss prediction systems. This study presents improved metamodels, using validated CFD to generate training datasets that cover a more general and wider range of erosion patterns, from low-amplitude roughness to severe grooves. In order to provide the industry with additional erosion geometry-linked tools for estimating energy yield losses, and foster further research and development in this area, the developed meta-models have been made available online with unrestricted access.

1. Introduction

Airborne particles such as rain droplets, hail stones and sand grains can significantly alter the external shape and roughness of wind turbine (WT) blades, yielding a significant loss of annual energy production (AEP) of the wind farms. Recent studies estimated the AEP loss and its uncertainty using probabilistic frameworks [1, 2], showing that an average AEP loss from 1.5% to 3% can occur, depending on the site (onshore or offshore), turbine model and operating conditions. In those studies, the considered uncertainty was that associated with a particular LE erosion (LEE) pattern along the blade. In general, the impact of LEE on blade aerodynamics abd turbine performance depends on the shape of damage, its depth and distribution and location along the radius [3, 4]. In [5], AEP losses were found to vary from 1% for LEE made of pits and gauges to 6% for LEE due to extended loss of coating. In [6], a field recorded damage geometry was used to define the CFD model for different possible damage severity states. The study

shows that up to 1% of AEP can be lost when LEE trips the boundary layer (BL) laminar-to-turbulent transition at the LE along a significant portion of the blade, whereas the presence of severe damages can further spoil the blade performance, from 1 to 3.5%, depending on the damage geometry and location.

All the aforementioned studies calculate the AEP loss making use of the blade element momentum theory (BEMT) for rotor aerodynamics embedded in multi-disciplinary wind turbine codes. A fundamental information required to generate a reliable BEMT model of the damaged blade is the accurate estimate of its airfoils' lift (c_l) and drag (c_d) coefficients for each LEE damage considered. The effects of LEE on c_l and c_d have been extensively investigated through wind tunnel tests [3, 7] and numerical simulation [8, 6], showing that both small-scale surface roughness and large-scale surface alterations have a significant effect on the blade performance.

CFD offers a robust approach to predict the performance losses due to LEE, but the variety of possible damage geometries makes this approach impractical for AEP loss estimates from an O&M perspective. In [9], CFD was used to define a suitable dataset of damaged airfoil force coefficients to train c_l and c_d metamodels to evaluate the aerodynamic force of each damaged blade strip. In that preliminary work, the focus was on large grooves extending from the pressure to the suction side. A similar approach to metamodel generation was used in [10], which focused on uncertainty affecting the overall geometry of the blade sections, such as that due to manufacturing errors.

In this study, we present extended and improved metamodels for predicting c_l and c_d of damaged blade strips, based on more accurate CFD models and setups, and a broader variety of LEE patterns, examining damages ranging from low-amplitude distributed roughness to severe grooves.

The metamodels are feed-forward artificial neural networks (ANNs), trained with two databases of c_l and c_d, each associated with two damage types. The database of the first damage type covers the range of damage patterns corresponding to small-scale roughness, the database of the second damage type is that associated woth large grooves over the LE. The damage scale separation is accounted for in the definition of the CFD model. Similarly to [6], a fully turbulent approach and damage resolving CFD mesh is used for airfoils with grooves, assuming the BL laminar-to-turbulent transition to always occur at the LE. For small-amplitude roughness, the BL transition location is crucial for accurately predicting the force coefficients; hence, a transitional RANS model is used to simulate the nominal airfoil geometry with rough wall. Small-amplitude roughness is accounted for by using a two-parameter model, the sand grain roughness K_s for the rough-wall boundary conditions, and a measure of the actual roughness height K for the rough-wall transition model.

Defining metamodels with K and K_s as separate inputs is an important aspect of novelty associated with this study. Indeed, the determination of the correct combination of roughness parameters associated with a given LEE damage is paramount for reducing the uncertainty of AEP loss prediction, but this is still an outstanding research question. More specifically, the estimate of K may be derived from the geometric characteristics of the measured roughness profile, as reported in [11]. The K_s parameter, on the other hand, depends on how the roughness profile interacts with the BL. Studies based on direct Navier-Stokes simulations [12], and experimental studies [13, 14] show that the relationship between K and K_s can differ significantly for the same measured value of K, depending on the detailed geometric features of the roughness geometry, e.g. density, shape, and orientation with respect to the aerodynamic flow of the roughness elements. In order to account for this uncertainty, the metamodels here is are structured so as to provide c_l and c_d estimates for a fairly wide range of values and combinations of K and K_s.

The paper structure is as follows. Section 2 presents the parametrization used to define the LEE damage types and properties. Section 3 describes the CFD modelling used for the

computation of damaged airfoils c_l and c_d. Section 4 describes the algorithm for defining and generating the force coefficient databases used to train the machine learning metamodels defined in Section 5. Finally, Section 6 presents the analysis of the databases and assesses the performance of the new metamodels.

In order to provide the industry with additional erosion geometry-linked tools for estimating energy yield losses, and foster further research and development in this area, the developed metamodels have been made available online with unrestricted access at github.com/LANCASTER-CFD/Leading-Edge-Erosion/tree/main/TORQUE-2024. The website will also publish new data generated by follow-on work of the authors.

2. LEE damage parametrization

Two parametric models of LEE damages are defined: one for cases with small-amplitude roughness, and a second one for cases with resolved severe damages approximated with constant depth chordwise grooves. Figure 1 shows the parameters definition used to define the damages. The first three sub-figures refer to the grooves, and next three to distributed roughness. The parameters used to define the severe damage grooves are: s_u and s_l which are the curvilinear extension of the groove on the suction side and pressure side respectively, and d which is the groove depth. For the case with surface roughness, the curvilinear extension of the rough wall on the airfoil is defined by the same s_u and s_l parameters, while the roughness is defined through the roughness height K and the sand grain roughness K_s.

Both damages can be positioned on the suction side, on the pressure side, or on both sides of the airfoil by a proper combination of positive or negative values of s_u and s_l.

3. Computational aerodynamics model

The force coefficients (cl and c_d) of the damaged airfoil are computed with 2D computational fluid dynamics simulations. For all the simulations, the Reynolds Averaged Navier Stokes (RANS) equations for incompressible and turbulent air flow are solved, coupled with the $k-\omega$ Shear Stress Transport (SST) eddy viscosity model [15]. A fully turbulent approach is used when a groove damage is resolved by the computational mesh. For nominal airfoil geometry simulations, laminar-to-turbulent transition of the BLs is simulated by using the Langtry-Menter four-equation γ–Re_θ SST transition model [16, 17, 18]. The model couples two additional equations to the SST model, one for the momentum thickness Reynolds number Re_θ and the other for the turbulence intermittency γ.

For cases where the damage pattern scales cannot be resolved by the computational grid, small amplitude roughness is accounted for by using a two-parameter model. The model requires defining the sand grain roughness K_s value for the rough-wall boundary conditions, and a measure of the actual roughness height K for the rough-wall transition model. The sand grain roughness is used in an the automatic wall treatment based on rough wall functions (details in [19, 8, 20, 6]). The roughness height parameter K is used to define a modification term in the γ–Re_θ SST transition model. In particular, the Re_θ field is modified to account for the increase of the BL momentum thickness due to the unresolved roughness [19, 8, 6].

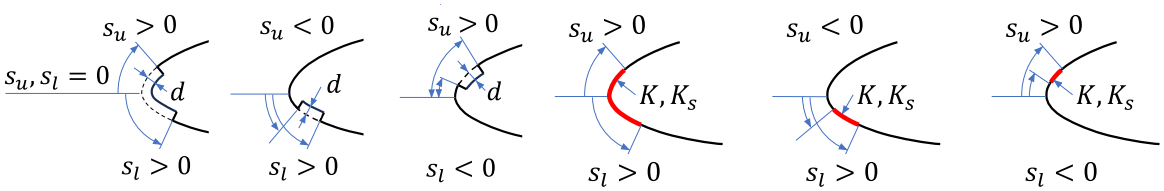
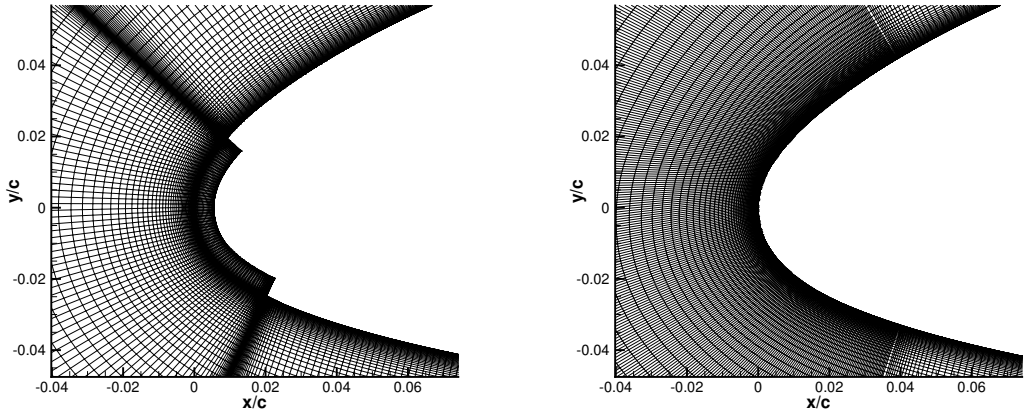


Figure 1: LEE damages and parameters.



(a) (b)

Figure 2: Computational mesh. Left: geometry with positive s_u/c and s_l/c groove. Right: mesh of the nominal airfoil geometry

Figure 2a shows the LE mesh for a case with groove damage of positive curvilinear extension on both suction side and pressure side. Figure 2b show the computational mesh for all the cases with surface roughness. In the latter case, 350 elements are used to discretize the LE region covering the span where roughness is applied. The rest of airfoil line is divided in 350 elements on both the suction side and pressure side, with 2.5×10^{-6} used as minimum wall distance to assure $y^+ < 1$ on all no-slip wall boundary. The number of elements normal to the airfoil and from the TE to the outlet of the domain are 150 and 110 respectively. The mesh for the resolved groove case is different only in the LE zone, where an additional structured block of mesh is used for the flow inside the constant depth groove. A grid sensitivity study was performed to find the optimal density of cells per unit length to set in the discretization of the groove curvilinear extension.

Both steady and time-dependent solvers are used in the generation of the c_l and c_d computations. As in previous works [20, 21], the constant a_1 in the expression of the SST eddy viscosity is changed from 0.31 to 0.29 to improve the agreement with measured force coefficients of the nominal and clean NACA64_3-618 airfoil.

This CFD modelling has been validated against experiments for a NACA$63_3$418 airfoil in two cases: distributed roughness on the leading edge (LE) in a wind tunnel experiment at Re=3×10^6 [3], and a constant groove depth of 0.3% of the chord at Re=5×10^6 [4]. Figure 3 shows the comparison of CFD and measurements for the drag polar as reported also in [6, 21].

All the simulations are performed with Ansys FLUENT v.21 R2 on the Lancaster University HEC cluster [22].

4. Databases of force coefficients

Table 1 lists the value of the parameters defining the two damaged airfoil CFD databases used for training the force coefficient metamodels, with the selected airfoil being the NACA64_3-618, which covers the outer 30% of the reference blade. In all cases, CFD simulations are performed at angles of attack (AoAs) from -10° to 16° with a 1° step.

The database generation is automatic and scripted. The automatic framework builds the geometry of the airfoil (with or without a groove), generates the CFD mesh and writes the simulation journal file listing all the settings for the aerodynamics modelling and boundary conditions for the 2D simulations. An automatic control on the damage curvilinear extension is

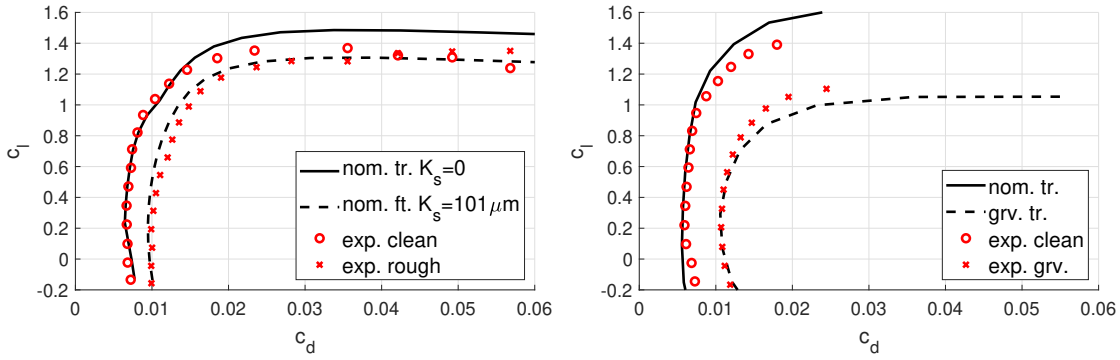


Figure 3: Comparison of CFD vs. experiments for a NACA63$_3$418 with distributed roughness of $K_s = 101$ μm (left) and 3 mm LE groove (right).

Table 1: Database definition and bounds.

Parameter	Database 1	Database 2
$s_u/c\times100$	-2 : 0.3$_\Delta$: 4	-2 : 0.6$_\Delta$: 4
$s_l/c\times100$	-2 : 0.3$_\Delta$: 4	-2 : 0.6$_\Delta$: 4
$d/c\times100$	0.05 : 0.1$_\Delta$: 0.75	-
$K/c\times10^6$	-	(50, 100, 150, 200, 300)
K_s/K	-	(0.5, 1, 2, 5, 10)

applied in the automatic generation of the groove geometries and rough distribution, to avoid combinations such that $s_u/c+s_l/c < 0.003$. This condition avoid negative curvilinear extensions (which are not possible), but it also avoid cases where the curvilinear extension of the damage is too small to produce significant effects on the alteration of the aerodynamic force coefficients.

Automated convergence tests are defined and implemented to assess convergence of residual and force coefficient in the scripted database generation. The scripted system starts with steady state simulations, and automatically switches to time-dependent simulations for non-converging solutions, running until time-averaged convergence of the aerodynamic force coefficients is achieved.

The steady RANS simulation run for a maximum of 2500 iterations, while 3,000 time-steps is the limit of the time-dependent simulations. The time dependent simulation starts from the final steady RANS solution. Each time-step is solved by performing 25 sub-iterations, and the time-step is obtained dividing by 35 the time the air flow takes to cover one chord length. A Reynolds number of 9×10^6 is imposed for all the simulations. The far field TI and turbulent length scale are set to 0.1% and 0.2 chords, respectively.

5. Metamodels of force coefficients

All the metamodels use a multi-layer perception feed-forward Artificial Neural Networks (ANNs) with one hidden layer of 80 neurons. The number of neurons is chosen to give a good trade-off of fitting and computational cost. Three ANNs are generated for each database, one for c_l and one for c_d. The learning process consists of a gradient-based optimization problem, solved with the Levenberg–Marquardt (LM) method. The dataset splitting is 80% for training 15% for validation and 5% for testing.

The effectiveness of ANNs can significantly increase by augmenting the input set, especially when used for regression tasks. In a standard ANN used for regression, the input layer receives the raw features $X = \{x_1, x_2, ..., x_d\}$ (where d is the number of the original inputs). These are processed through one or more hidden layers that apply weights (w) and biases (b), followed by

activation functions like ReLU or sigmoid. The output layer then produces the continuous value $\hat{y}$. The learning process involves adjusting w and b to minimize the difference between $\hat{y}$ and the actual output y, often using a mean squared error (MSE) loss function: $MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$ (where N is the number of samples).

Augmenting input features in ANNs for regression means adding new inputs derived from the original set. The input augmentation can have different benefits. Augmented features allow the ANN to model complex relationships without needing additional layers or neurons, which can be computationally more efficient. By including interactions and nonlinearities explicitly, the model can often achieve higher accuracy in regression tasks. In some cases, augmented features can lead to a simpler model that generalises better, reducing the risk of overfitting. Although ANNs are often seen as black boxes, augmented features derived from known interactions can offer some interpretability regarding how certain variables influence the output.

In this preliminary attempt to use the augmented input approach, quadratic and simple interactions (such as $x_{12} = x_1 \times x_2$) terms have been used. The results look promising, but future work and tests will be needed to understand the effect of the augmented inputs and then optimise the set.

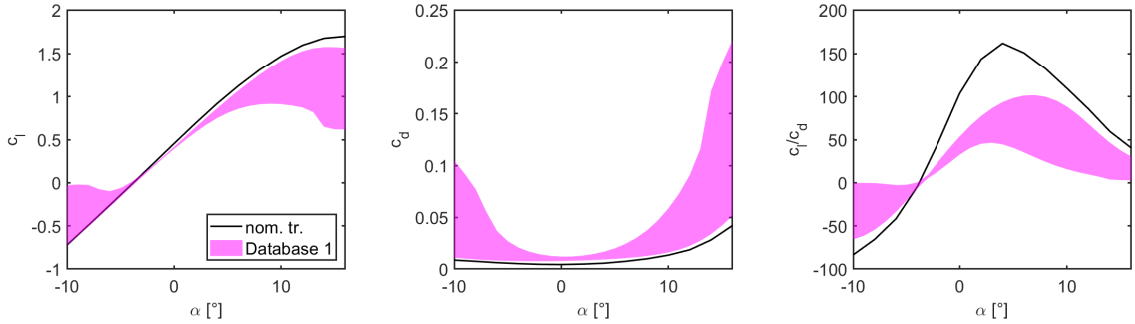
6. Results and discussion

Figure 4 shows the ranges of force coefficients (filled areas) spanned by the damage cases in the two databases, as well as the force coefficients of the nominal clean NACA64$_3$-618 airfoil (black curves). Figure 4a presents the values of c_l, c_d, and c_l/c_d versus AoA for Database 1, which includes LEE grooves. This type of damage significantly affects both lift and drag, reducing lift at high AoA and increasing drag across the entire AoA range. Figure 4b refers to Database 2 of LEE cases with small amplitude roughness. Here, the span of the filled area is narrower than in the erosion grooves case, and it approaches the nominal clean case for scenarios with the smallest roughness height and extent.

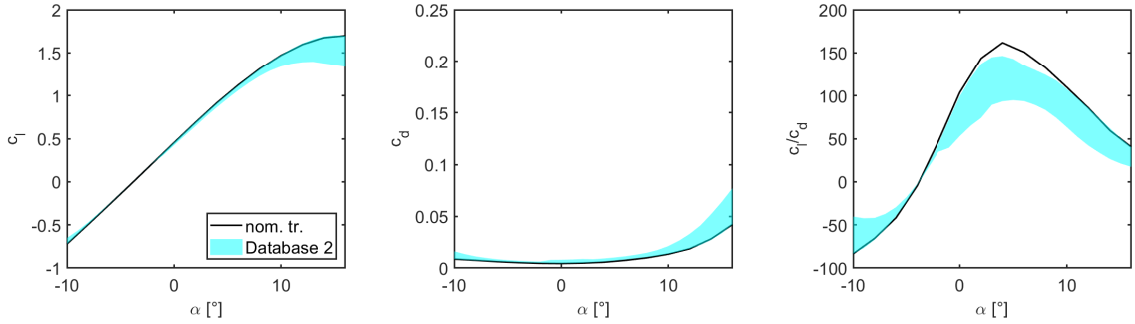
Merging the plots from Figures 4a and 4b would demonstrate that the combination of the two databases covers a continuous spectrum from the clean and nominal case to the most severe groove damage case. There is an overlapping area of conditions where roughness and groove damage cause similar performance impairments.

Figures 5 and 6 illustrate the isolate effects of individual damage parameters on the output of force coefficients for Databases 1 and 2, respectively. Figure 5 show how the c_l and c_d curves vary with changes in s_u/c (as shown in Figure 5a) or d/c (as shown in Figure 5b), when other parameters are held constant at a reference point of (s_u/c, s_l/c, d/c) = (0.022, 0.022, 0.0025). Increasing the damage depth d consistently leads to a monotonic decrease in performance, evidenced by a reduction in c_l at positive AoAs and an increase in c_d across all AoAs. Conversely, the impact of varying the suction side curvilinear extension s_u/c is not monotonic for both c_l and c_d. The set of cases in Figure 5a indicates that the worst performance occurs with an intermediate value of s_u/c. This suggests that positioning the forward-facing step farther from the LE can mitigate the impact on the c_l and c_d of the damaged airfoil.

Figure 6 shows the change of c_l and c_d curves for a damage case defined by (s_u/c, s_l/c, K/c, K_s/K) = (0.022, 0.022, 150×10^{-6}, 2), analyzing the variation of s_u/c (as shown in Figure 6a) or K_s/K (as shown in Figure 6b). In this scenario, the effects of both parameters on c_l and c_d are monotonic. Indeed, the primary influence comes from the roughness effects, particularly in terms of its interaction with the BL transition and shear friction. As the extent of the roughness over the airfoil line increases, so does its impact on airfoil performance. Specifically, an increase in s_u/c leads to a decrease in c_l at high AoAs, and alters the AoA at which the BL transition occurs at the LE of the airfoil, which is reflected in a discontinuity in the c_d vs. AoA curve. Figure 6b shows that an increase in K_s/K produces a qualitatively similar effect. This result demonstrate that even with a reliable measurement of K, the performance of the damaged airfoil can vary



(a) Database 1



(b) Database 2

Figure 4: Databases of c_l and c_d.

significantly depending on the assumed equivalent sand grain roughness K_s. This underscores the importance of accurately estimating K_s/K to reliably assess the impact of roughness on blade performance.

The accuracy of the metamodels is verified on two separate sets of data that were not used for training. One set (on-grid) is defined by extracting 128 cases from the database detailed in Table 1. The second set of data for testing (off-grid) is created by running additional CFD simulations for a set of parameter values that fall within the bounds but do not correspond exactly to any of the grid points defined in Table 1. In constructing the off-grid test cases, CFD simulations were performed over a finer range of AoA, from -10° to 16° in steps of 0.5°. Figure 7 provides an overview of the performance of the ANNs for Database 1. The left sub-plots show a histogram collecting the results of the tests, which involve comparing the ANN c_l predictions to the CFD results for the same damaged geometry. The metric employed to verify prediction accuracy is the Root Mean Squared error (RMSe) of the difference between ANN and CFD results across the entire range of AoAs. The distributions indicate that the total c_l RMSe is below 0.02 for the on-grid cases (Figure 7a), and below 0.01 for the off-grid cases (Figure 7b), with averages of 0.0045 and 0.0046, respectively. A similar pattern is observed for the c_d ANN test, where the average RMSe for on-grid cases is 0.00071 and for off-grid cases is 0.00068. The middle and right sub-plots in Figure 7 present the CFD and ANN c_l and c_d curves for the tested geometry with RMSe closest to the modal value of the error distribution in the left sub-plot. For most geometries, the ANNs accurately predict c_l and c_d, both in conditions of a fully attached BL and near the maximum c_l

Figure 8 employs the same visualization approach to present an overview of the performance of the ANNs for Database 2. For on-grid cases (Figure 8a), the average RMSe is 0.0029 for c_l and 0.0002 for c_d. For off-grid cases (Figure 8b), the average RMSe is 0.0087 for c_l and 0.001 for c_d. From the right plots in Figure 8, it is apparent how the metamodels handle the discontinuity

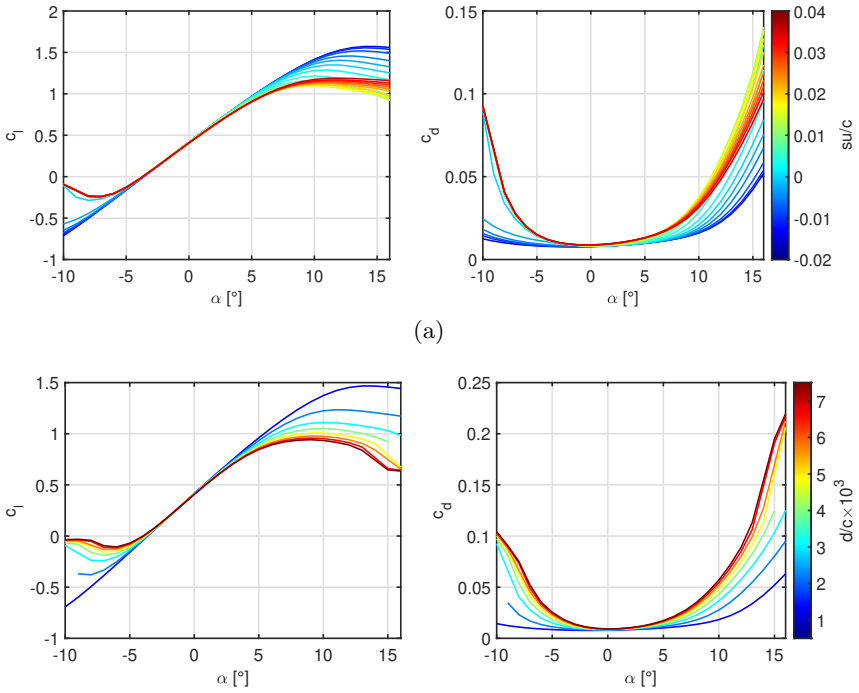


Figure 5: c_l and c_d vs. AoA for cases of Database 2 with $s_l/c = 0.022$. Top: cases with fixed $d/c = 0.0025$ and variable su/c; bottom: cases with fixed $su/c = 0.022$ and variable d/c.

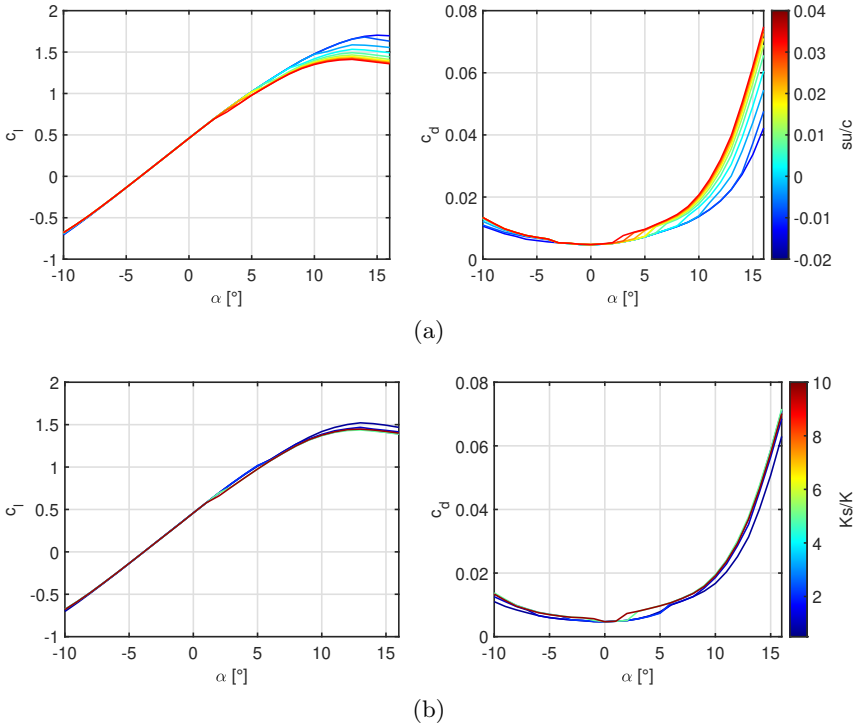
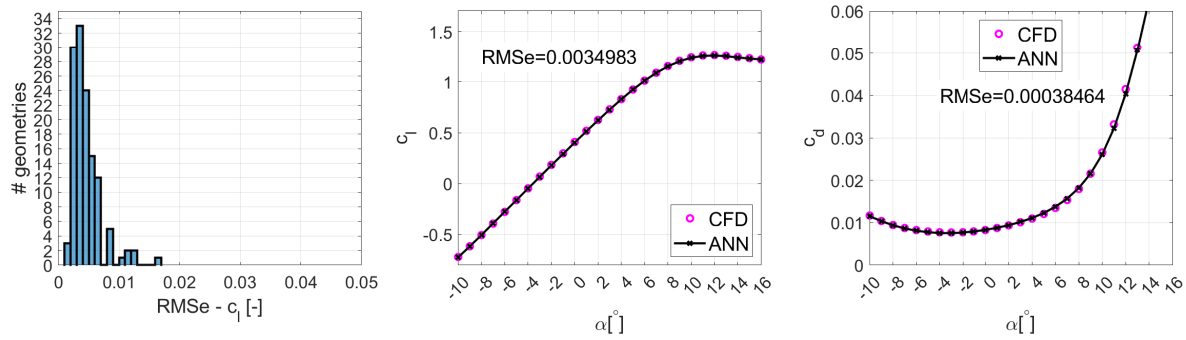
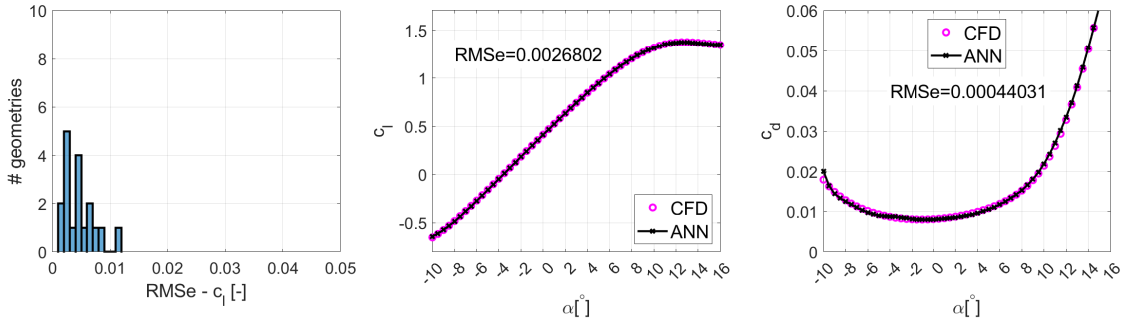


Figure 6: c_l and c_d vs. AoA for cases of Database 2 with $s_l/c = 0.022$. Top: cases with fixed $K_s/K = 2$ and variable su/c; bottom: cases with fixed $su/c = 0.022$ and variable K_s/K.



(a) On-grid tests



(b) Off-grid tests

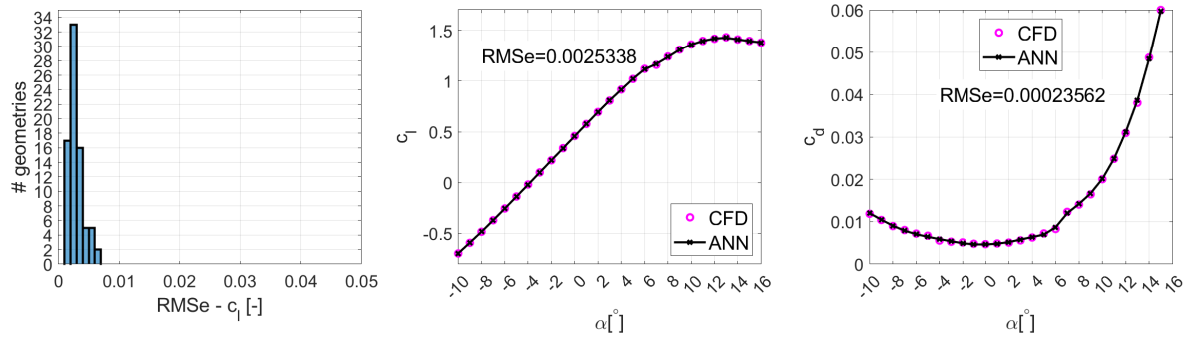
Figure 7: ANNs tests for Database 1. Left: histogram of RMSe in c_l prediction. Center and right plots shows prediction in the case closest to the median RMSe for the c_l ANNs.

in the c_d curve, which is due to the onset of BL transition at the LE of the airfoil. In cases where the prediction is requested for an AoA that is included in the training dataset, the prediction perfectly matches. However, when the resolution in terms of AoA is increased, the ANN is unable to capture the discontinuity and provides an interpolated value for the points that are off-grid with respect to the AoA.

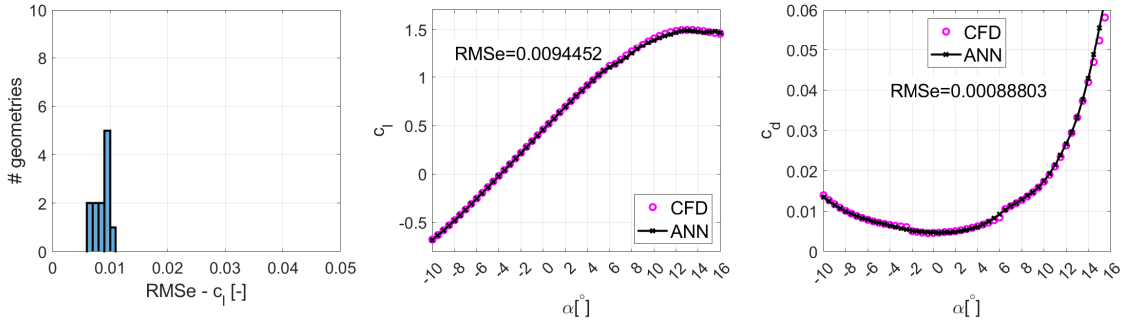
7. Conclusions

This article presented new metamodels, based on ANNs, for the rapid estimate of the aerodynamic performance of the NACA64$_3$-618 airfoil subjected to LEE. Compared to previous studies, the metamodels are designed to cover a wider and more comprehensive range of damage cases, ranging from small-amplitude distributed roughness to cases with significant material removal from the LE. Two databases were generated using 2D RANS simulations for the training of the ANNs. The first database is generated with fully turbulent simulations of geometrically resolved erosion groove, whereas the second database is generated using a transitional laminar-to-turbulent RANS model combined with a two-parameter roughness modelling applied to the nominal geometry. This latter case represents a novel aspect enabling the ANNs to cover different type of roughness profiles of the blade surface. The new metamodels demonstrated very good accuracy compared with CFD results of eroded geometries not included in the training dataset.

Making the developed metamodels openly available at github.com/LANCASTER-CFD/Leading-Edge-Erosion/tree/main/TORQUE-2024, will provide new support to wind farm operators and wind energy researchers, contributing to further advances in cost analyses and blade predictive maintenance. The website will also publish new data generated by follow-on work of the authors.



(a) On-grid tests



(b) Off-grid tests

Figure 8: ANNs tests for Database 2 . Left: histogram of RMSe in c_l prediction. Center and right plots shows prediction in the case closest to the median RMSe for the c_l ANNs.

Acknowledgments

The project was partly supported by the UK Engineering and Physical Sciences Research Council, Grant No. EP/R511560/1, Lancaster University's Impact Fund and the Ministry of Enterprises and Made in Italy under Grant Agreement "RdS PTR 2022-2024 - Energia elettrica dal mare". All CFD simulations were performed on Lancaster University's HEC cluster.

References

[1] Papi F, Balduzzi F, Ferrara G and Bianchini A 2021 *Renewable Energy* **165** 701–715 ISSN 0960-1481
[2] Campobasso M S, Castorrini A, Ortolani A and Minisci E 2023 *Renew. Sustain. Energy Rev.* **178**
[3] Langel C, Chow M, Van Dam C and Maniaci D 2017 *SAND2017-11289*
[4] Krog Kruse E, Bak C and Olsen A S 2021 *Wind Energy* **24** 1263–1274
[5] Saenz E, Mendez B and Muñoz A 2022 *Journal of Physics: Conference Series* **2265** 032059
[6] Castorrini A, Ortolani A and Campobasso M S 2023 *Renewable Energy* 119256
[7] Sareen A, Sapre C A and Selig M S 2014 *Wind Energy* **17** 1531–1542
[8] Campobasso M S, Castorrini A, Cappugi L and Bonfiglioli A 2022 *Wind Energy* **25** 168–189
[9] Campobasso M S, Cavazzini A and Minisci E 2020 *Journal of Turbomachinery* **142** 071002
[10] Marepally K, Jung Y S, Baeder J and Vijayakumar G 2022 *Journal of Physics: Conference Series* **2265**
[11] Bhushan B 2000 *Modern tribology handbook, two volume set* (CRC press) pp 79–150
[12] Busse A and Zhdanov O 2022 *Flow, Turbulence and Combustion* **109** 1195–1213
[13] Schlichting H 1937 *Experimental investigation of the problem of surface roughness* 823 (NACA)
[14] Jiménez J 2004 *Annual Review of Fluid Mechanics* **36** 173–196
[15] Menter F 1994 *AIAA Journal* **32** 1598–1605
[16] Menter F, Langtry R, Likki S, Suzen Y, Huang P and Völker S 2006 *J. Turbomach.* **128** 413–422
[17] Langtry R, Menter F, , Likki S, Suzen Y, Huang P and Völker S 2006 *J. Turbomach.* **128** 423–434
[18] Langtry R and Menter F 2009 *AIAA Journal* **47** 2894–2906
[19] Inc A 2021 *Fluent Theory Guide* 2021st ed accessed on December 1, 2022
[20] Ortolani A, Castorrini A and Campobasso M S 2022 *Journal of Physics: Conference Series* **2265** 032102
[21] Cappugi L, Castorrini A, Bonfiglioli A, Minisci E and Campobasso M S 2021 *Energy Conv. Manag.* **245**
[22] Lancaster University High End Computing (HEC) `www.lancaster.ac.uk/iss/hec` accessed on 20 April 2021