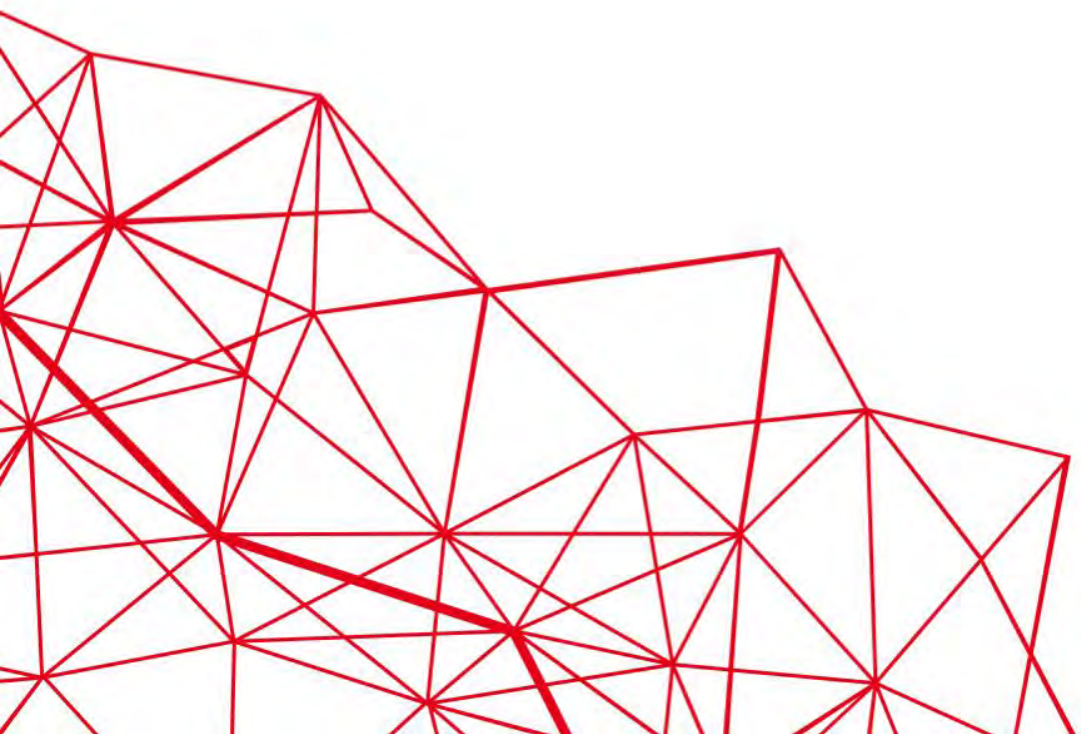
SHAPING TOMORROW

ISC High Performance

ISC 2020 DIGITAL

JUNE 22-25

#ISC20



Data-Driven Techniques to Enhance and Supplement Computational Fluid Dynamics Prediction Capabilities

Philipp Bekemeyer, Florian Jäckel and Cornelia Grabe

German Aerospace Center (DLR), Institute of Aerodynamics and Flow Technology, Center for Computer Applications in AeroSpace Science and Engineering

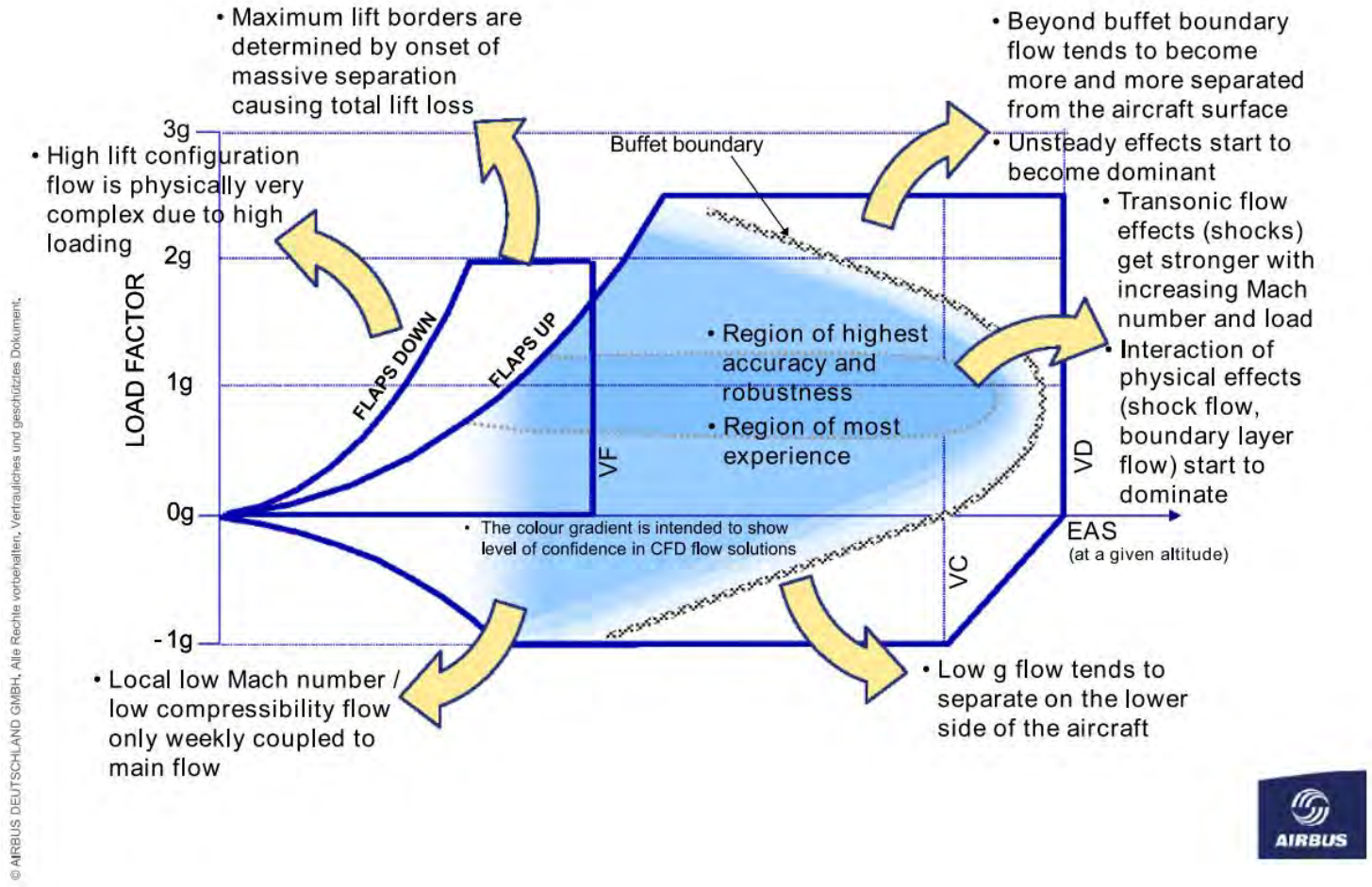
philipp.bekemeyer@dlr.de

Outline

- Motivation
- SMARTy
- Methods to Supplement Computational Fluid Dynamics
- Methods to Enhance Computational Fluid Dynamics
- Conclusions & Future Work

Motivation

- CFD simulations – in particular RANS simulations – play an important role in modern aircraft design
- High-quality aerodynamic data hopefully enables better design
- Region of high accuracy and robustness is centered around the cruise point
- Computational cost is still a bottleneck for multi-query scenarios (e.g. optimization or loads analysis)



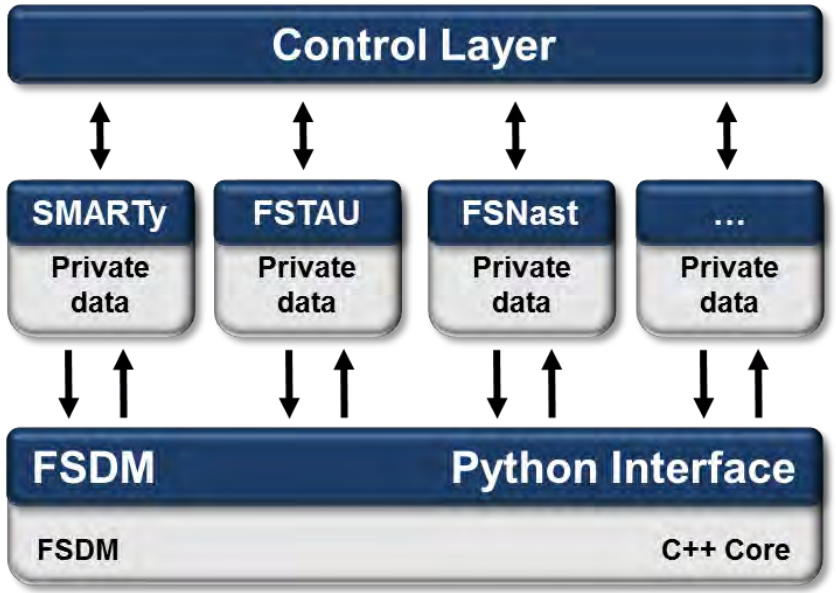
Estimation of CFD capabilities across the flight envelope [1]

[1] Abbas-Bayoumi, A., Becker, K. **An industrial view on numerical simulation for aircraft aerodynamic design**. J.Math.Industry 1, 10 (2011). https://doi.org/10.1186/2190-5983-1-10

Surrogate Modeling for AeRo Data Toolbox (SMARTy)

Key features

- Design of experiments
- Adaptive sampling
- Surrogate Modeling
- Hyperparameter optimization
- Variable-Fidelity Modeling
- Dimensionality reduction
- Reduced Order Modeling
- FlowSimulator “plug-in”, using FSDM & FSTAU functionalities
- Parallelized, partly differentiated, MKL, interfaces to Scikit-learn
- standalone version with use-case specific C-interface*



Surrogate Modeling for AeRo Data Toolbox (SMARTy)

Key features

- Design of experiments
- Adaptive sampling
- Surrogate Modeling
- Hyperparameter optimization
- Variable-Fidelity Modeling
- Dimensionality reduction
- Reduced Order Modeling
- FlowSimulator “plug-in”, using FSDM & FSTAU functionalities
- Parallelized, partly differentiated, MKL, interfaces to Scikit-learn
- standalone version with use-case specific C-interface*

Applications

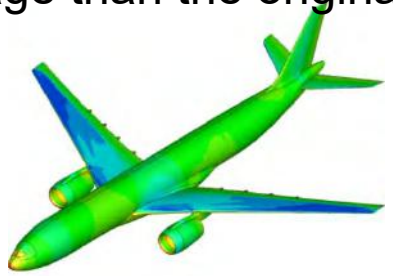
- Surrogate-based shape optimization
- Uncertainty quantification
- Robust & inverse design

- Fusing experimental + CFD data
- Optimal sensor placement
- Wind-tunnel corrections and support

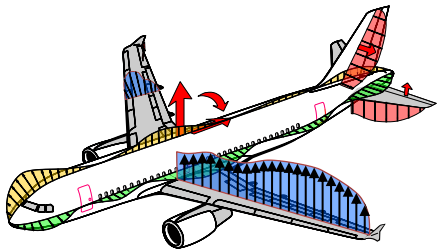
- Accelerating CFD computations
- Virtual flight testing
- Rapid aero-data & loads prediction across the flight envelope (and structural sizing in MDO context)

Motivation to Supplement Computational Fluid Dynamics

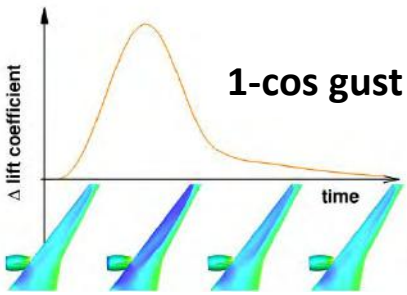
- From aircraft design to certification many **aerodynamic data** are needed for the entire flight envelope → pressure & shear stress distributions, global coefficients, derivatives from steady and unsteady simulations
- **Goal:** provide predictions of the aerodynamics based on **high-fidelity CFD** but with lower evaluation time and storage than the original CFD model → **ROMs**



LOADS



PERFORMANCES



HANDLING QUALITIES

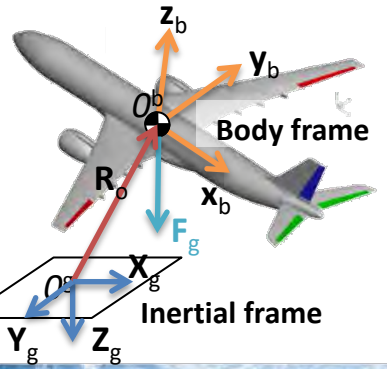
50 flight conditions (A)

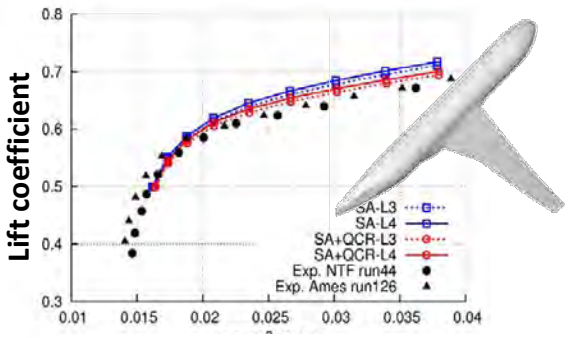
C: 4 control laws

M: 100 mass configs.

S: 5 maneuvers + 25 gust lengths

$\mathcal{O}(10^6)$ simulations





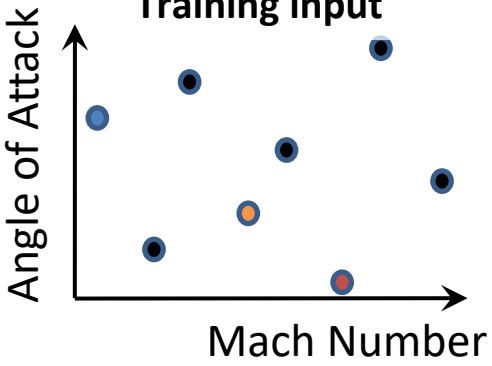
Idealized drag coefficient

DLR

Methodology

1

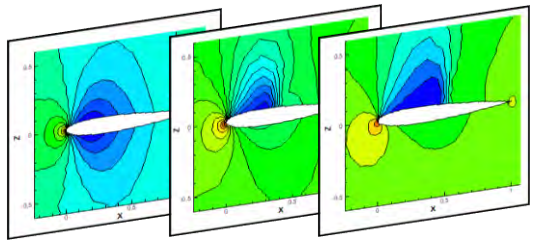
Training input



Define training points

2

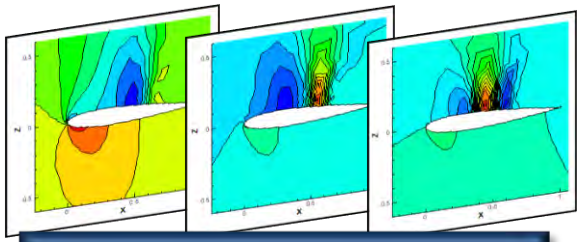
Training output



Compute training outputs with CFD

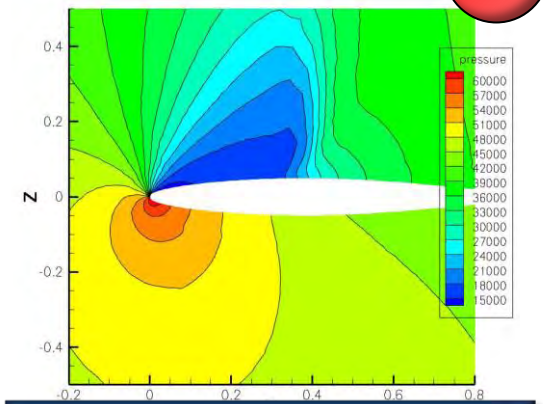
3

a) POD + Interpolation
b) Isomap + Residul Min
c) Neural Networks



Build a data-driven Model

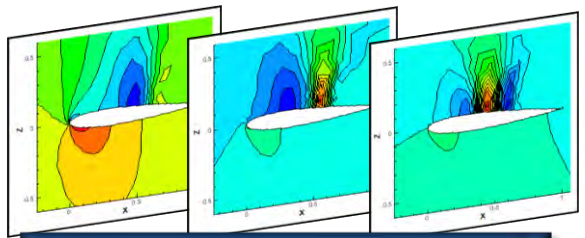
4



Flow prediction at untried condition

Methodology

a) POD + Interpolation
b) Isomap + Residual Min.
c) Neural Networks



Build data-driven Model

a) POD + Interpolation

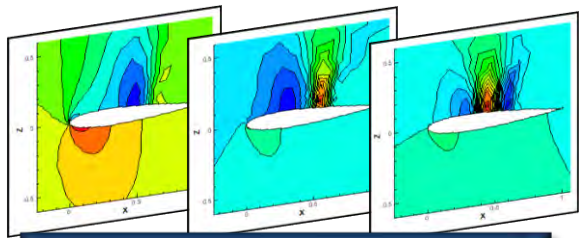
- Proper Orthogonal Decomposition for dimensionality reduction
- Gaussian Processes to interpolate on the latent space
- Non-intrusive, data-driven approach

\+ Works on surface data only
\+ Fast training and predictions
\+ Easy to implement and parallelize
\- Does not take any physical information into account
\- POD is inherently linear but CFD problems are typically not

POD + Interpolation can be seen as the standard baseline approach which is widely used in academia and industry

Methodology

a) POD + Interpolation
b) Isomap + Residual Min.
c) Neural Networks



Build data-driven Model

3

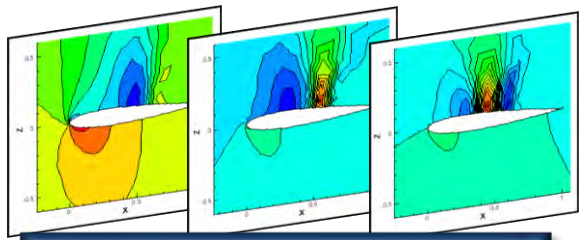
b) Isomap + Residual Minimization

- Isomap as Manifold Learning Technique is used for dim. reduction
  - Finds the closest solutions by searching on the manifold
- Solution are determined by minimizing the full or parts of the discretized residual
- Intrusive, physics-informed approach

\+ Takes physical knowledge into account
\+ Explores the solution space by searching the manifold for closest points
\+ Faster predictions compared to full CFD runs
\- Needs the full flow-state (at least in its current implementation)
\- Needs access to the underlying discretized residual

Methodology

a) POD + Interpolation
b) Isomap + Residual Min.
c) Neural Networks



Build data-driven Model

3

c) Neural Networks

- Building a neural network based on flow-state and surface location
  - Predicting one pressure coefficient for one point at the time
  - Greatly boosts the number of training points compared to predicting the full surface at once
- Non-intrusive, data-driven approach

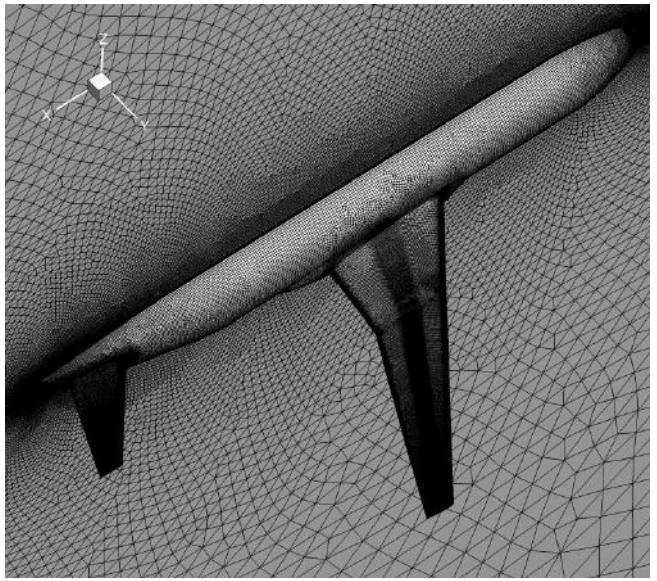
\+ Works on surface data only
\+ Capable of predicting also complex behavior
\- Large amount of architectures and parameters to be determined
\- High computational costs for training needed

In particular a fully connect MLP network is applied herein

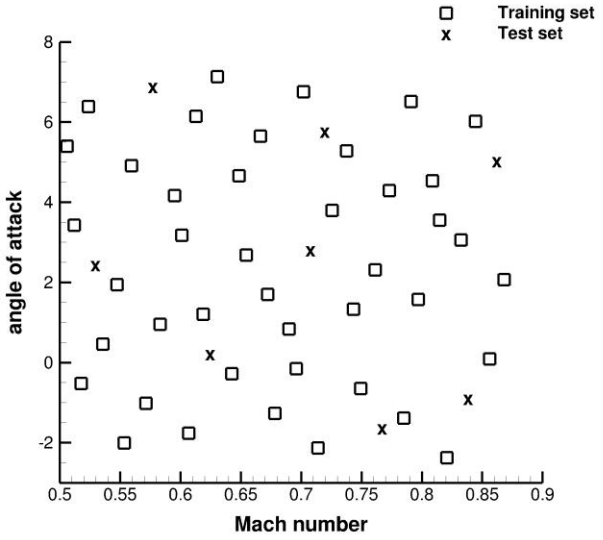
- 3 hidden layers with 500 neurons each
- ReLU as activation function
- Adam optimizer with roughly 600 epochs

Test Case NASA CRM

- The NASA CRM is a generic aircraft model for research
- Scaled to full scale and elastically deformed to obtain flight shape
- DLR-TAU Code solving the RANS equations with SA turbulence model
- Roughly 7 million grid nodes with 225 000 on the surface
- Varying angle of attack and Mach number
  - 42 snapshots as training dataset
  - 8 snapshots as test dataset

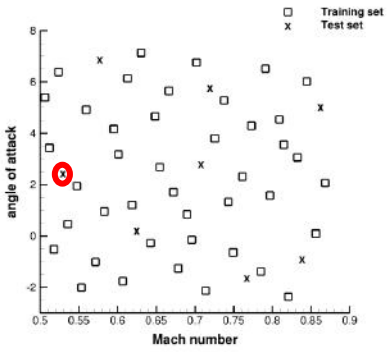


Computational grid around the aircraft

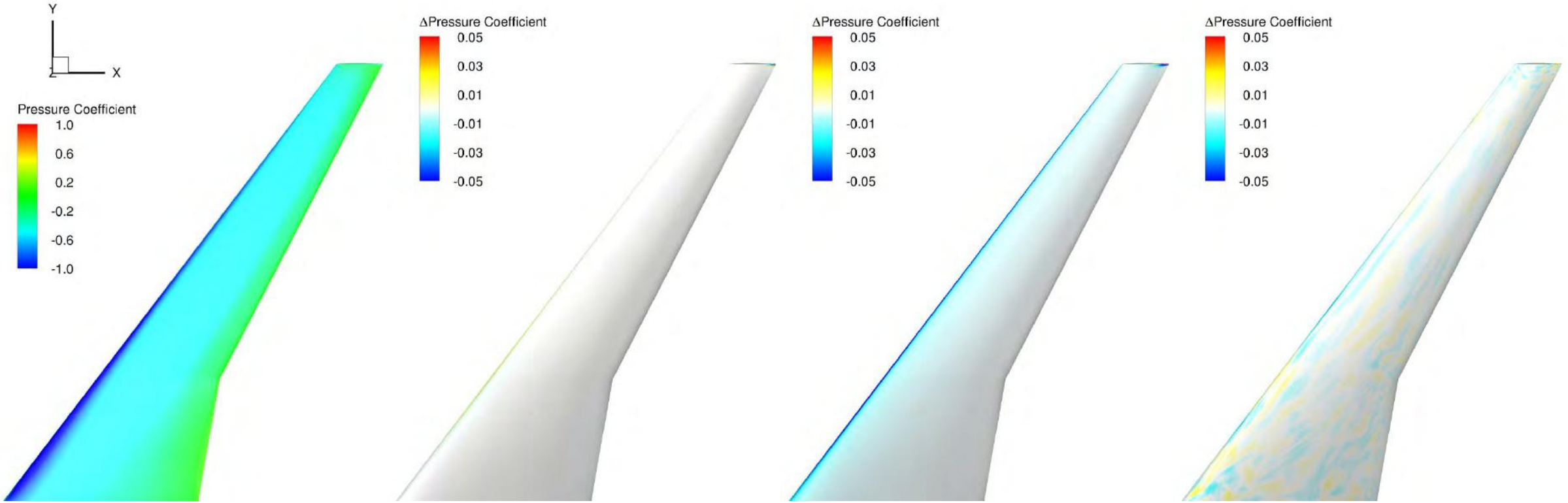


Training and Test data set

Results

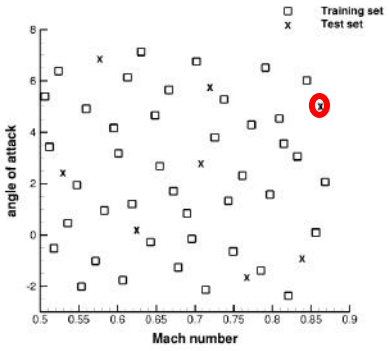


- Prediction at Mach = 0.52969 and AoA = 2.43827

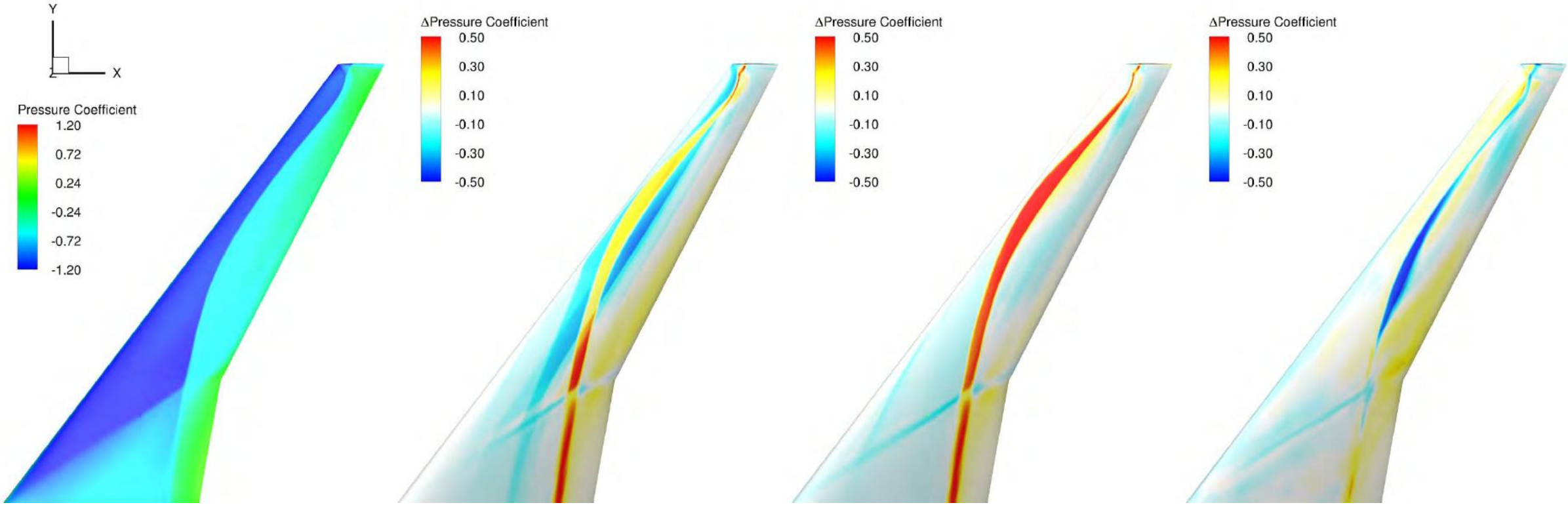


RANS Reference Solution | POD + I Solution | Isomap + Res. Min. Prediction | NN Prediction

Results



- Prediction at Mach = 0.86219 and AoA = 5.03086



RANS Reference Solution | POD + I Solution | Isomap + Res. Min. Prediction | NN Prediction

Motivation to Enhance Computational Fluid Dynamics

- DNS, LES are computationally expensive since all turbulent scales need to be resolved, experiments are expensive as well
- RANS are computationally cheaper since influence of turbulence on the mean flow is modeled and not all scales need to be resolved
- Turbulence models are often less precise, especially at the borders of the flight envelope

Methodology

How can high-fidelity data gathered from DNS, LES and experiments together with Machine Learning be used to improve existing Turbulence Models?

Idea:

- Augment the existing turbulence model (here: Spalart-Allmaras) by introducing a correction factor β:

$$\frac{D\tilde{\nu}}{Dt} = \beta P(\tilde{\nu}, U) - D(\tilde{\nu}) + T(\tilde{\nu}, U)$$

- For multiple cases, compute the optimal distributions for β so that the results from the RANS model match the high-fidelity reference data (Field Inversion)
- Using these β-distributions, find a function for β using Machine Learning
- Insert the ML-model in the augmented turbulence model and apply it to new cases

- Proposed by K. Duraisamy and colleagues, around 2015

Methodology

- The goal of Field Inversion is to modify β so that the results d_β from the RANS calculations with the augmented turbulence model match the results d_{ref} from the high-fidelity methods
- Thus, minimize the cost function

$$\mathcal{I}(\beta) = \frac{1}{2}\left(\sum_j^{N_{\text{ref}}}(d_{\text{ref},j} - d_{\beta,j})^2 + \lambda \sum_i^{N}(\beta_i - 1)^2\right)$$

- The second term of the cost function is a regularization since we have many DOFs (= number of cells)
- This is a constrained optimization problem, since the RANS-equations need to stay satisfied (R=0)
- For efficient optimization, a gradient-descent based method can be applied, since the gradient dI/dβ can be calculated efficiently using the adjoint method

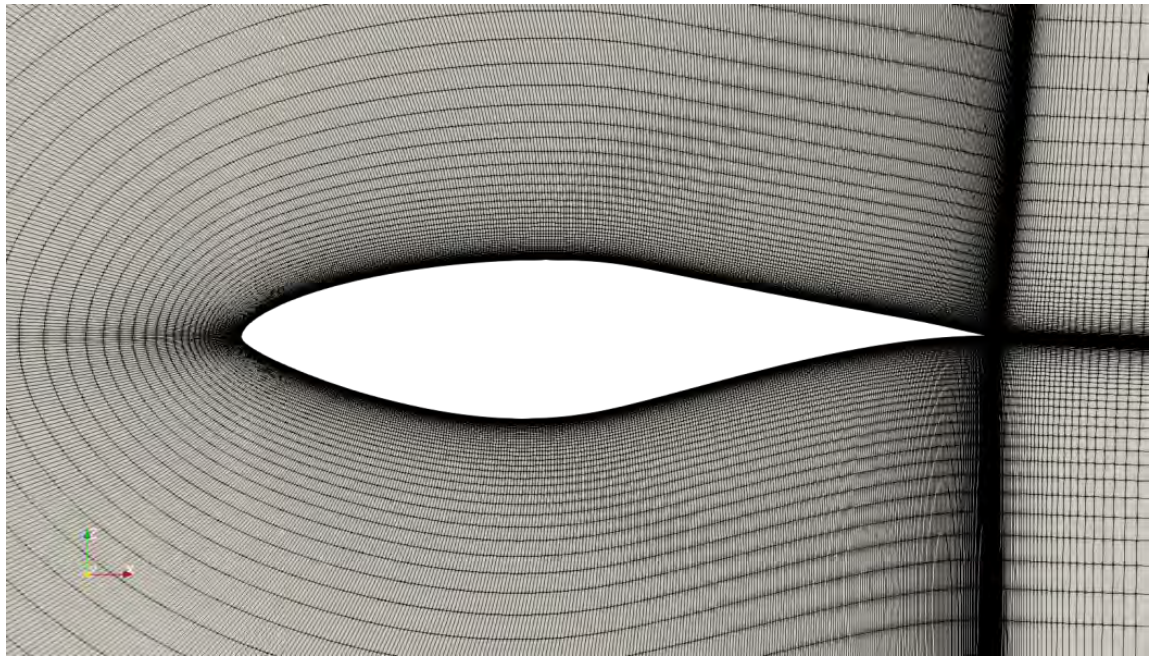
Methodology

- Field Inversion is only the first step to improve the turbulence model,
  - β-distributions is only valid for the respective cases

- A function $\beta(\eta_1, \eta_2, ...)$ depending on flow features $\eta_1, \eta_2, \ldots$ needs to be found
  - Use Machine Learning to determine such a function
  - Neural Networks are especially promising, since they can *approximate any limited function* (Universal Approximation Theorem)

Results: The S809 airfoil

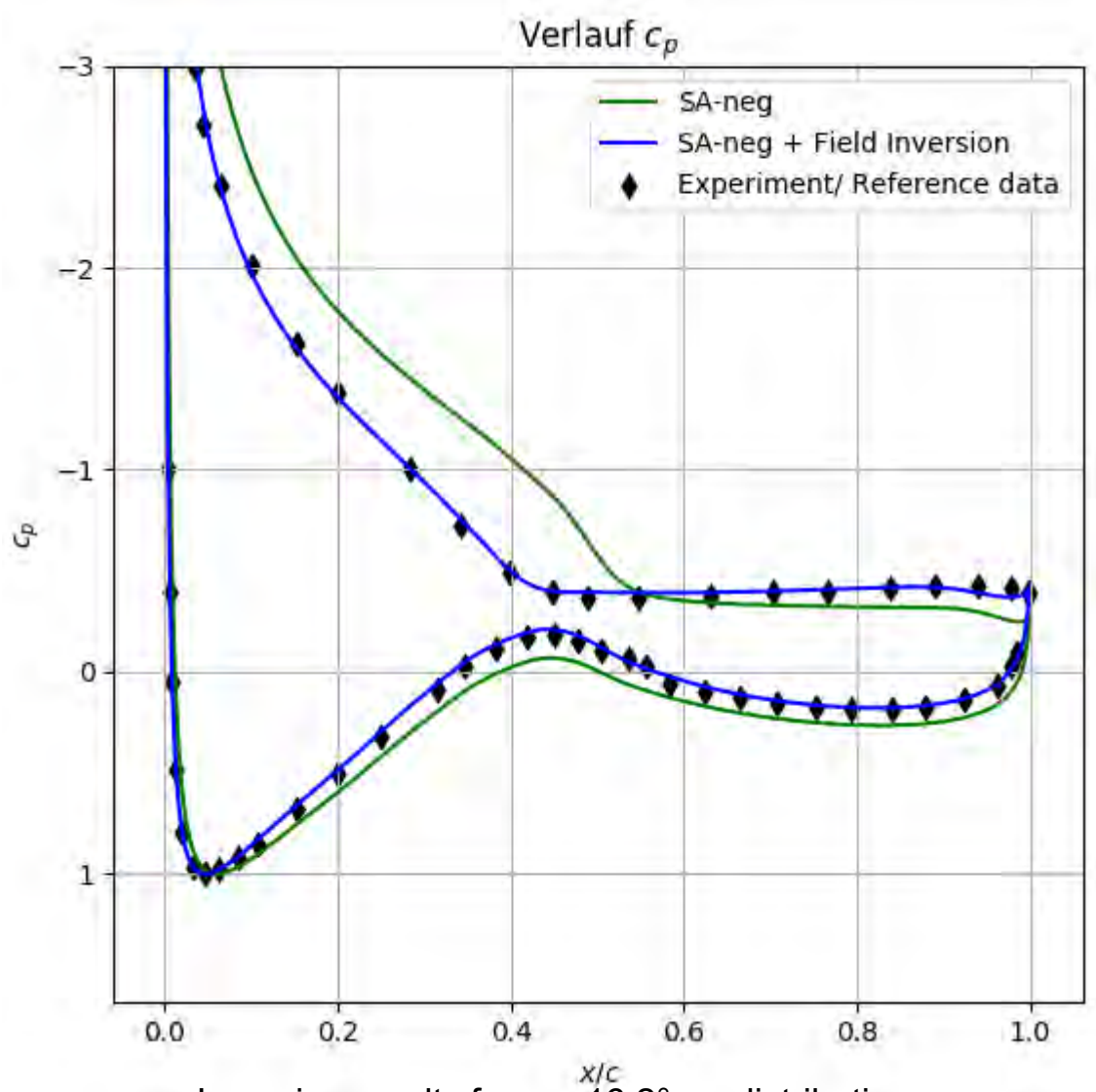
- The S809 is an airfoil for wind energy researched by NREL
- Reliable experimental data available for c_p and c_L at a wide range of angles of attack [3]
- Roughly 300 000 grid nodes
- Singh, Holland have used the S809 airfoil previously



Computational grid around the airfoil surface

Results: The S809 airfoil

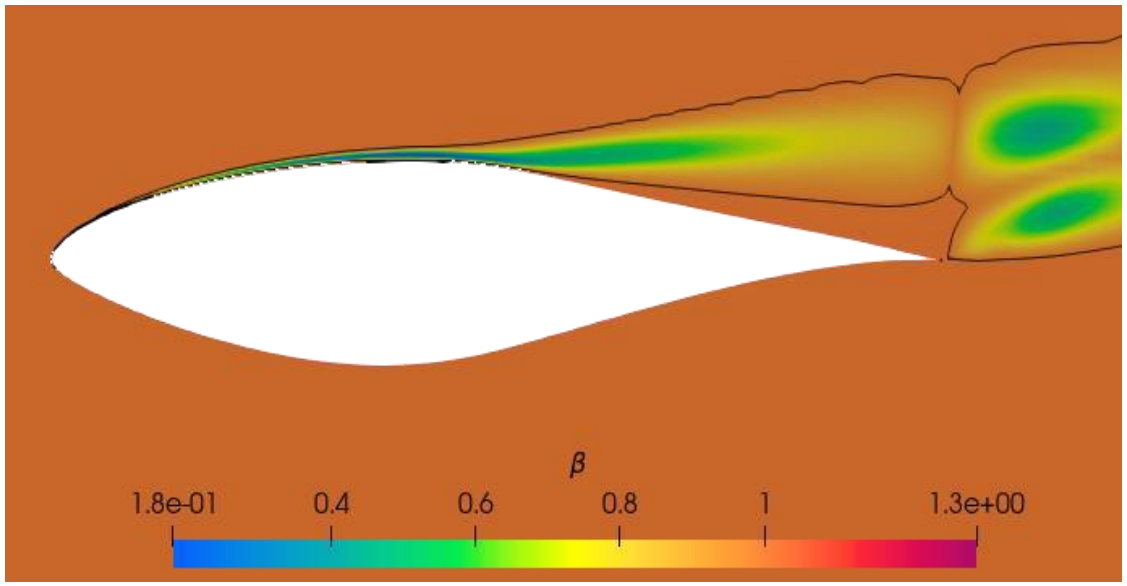
- Field Inversions were performed for α = 12.2°, 16.2° and 18.2°
- minimizing the difference of the c_p distributions between experiment and RANS simulation
- The experimental data can be matched almost perfectly
- Matching c_p also reduces the deviation in c_l e.g. from ~35% to ~4% for α = 16.2°



Inversion results for α = 16.2°, c_p-distribution

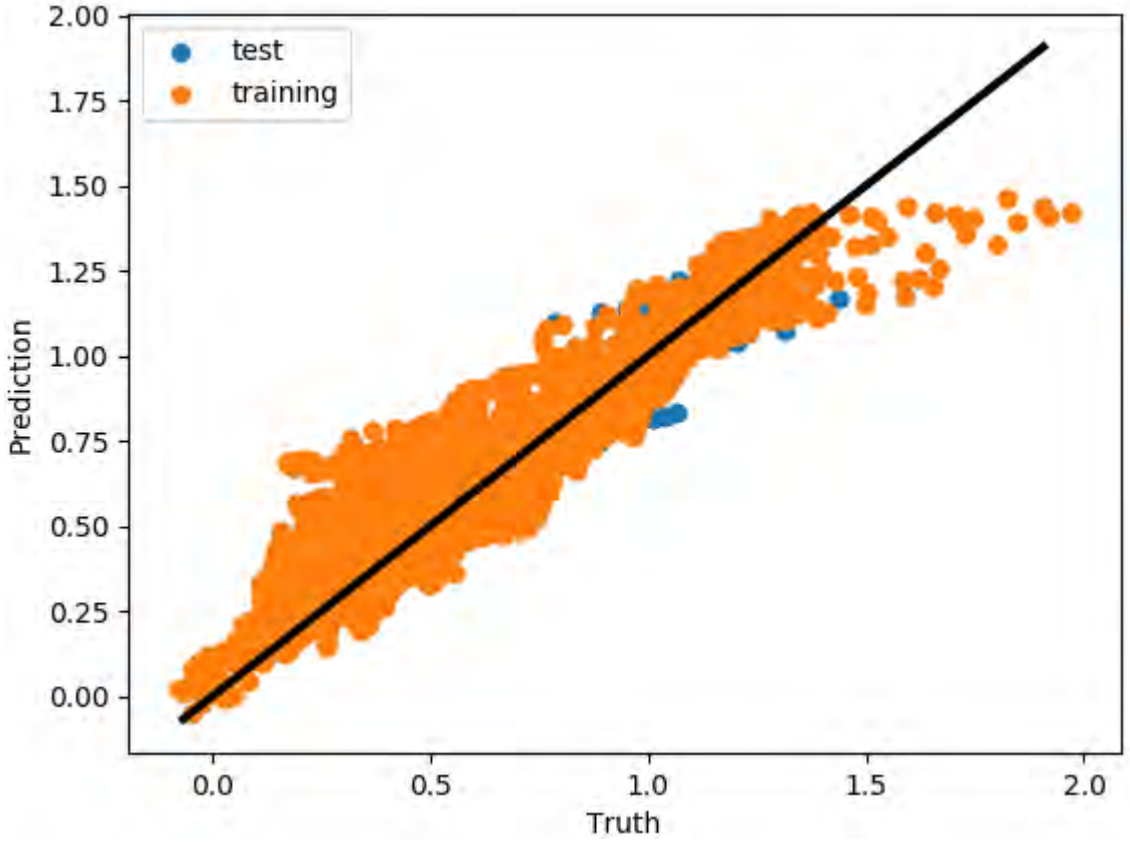
Results: The S809 airfoil

- Using the three datasets from the Inversion (α = 12.2°, 16.2°, 18.2°) but neglecting point with β = 1
- Training a Neural Network
- Neural Network
  - Fully connected, 2 layers à 200 neurons
  - ReLU and dropout
  - Adam – 200 epochs
  - Input features [2]: $\{\frac{P}{D}, \chi, \frac{|S|}{|\Omega|}, \frac{\mu_t|S_{ij}|}{1.5\tau_w}\}$
  - Target: β

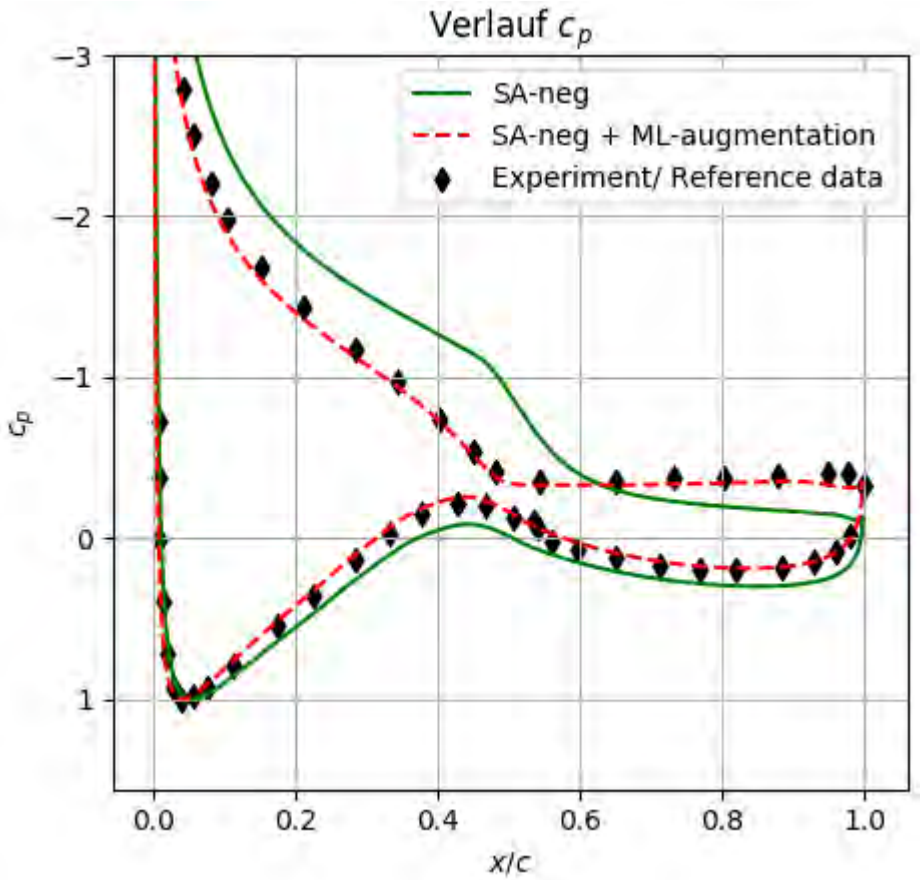


Inversion results for α = 16.2°, β-field

Results: The S809 airfoil



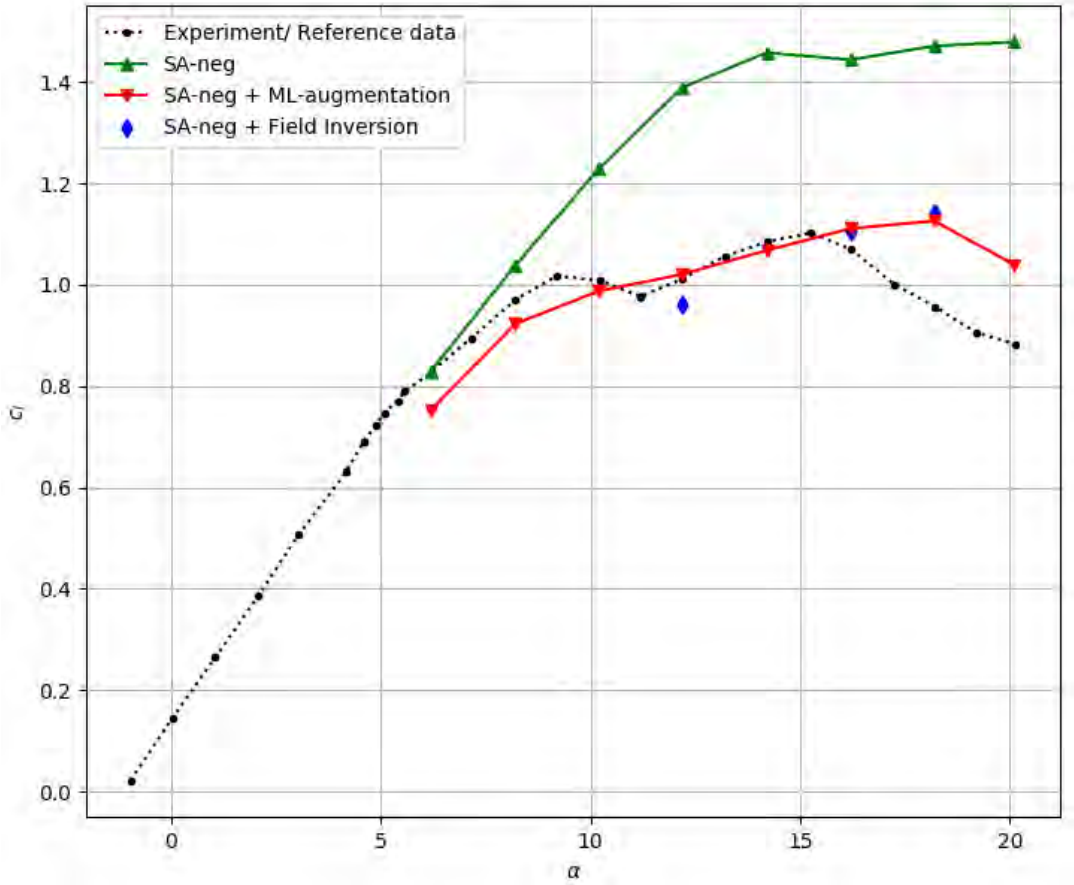
Ground truth vs Prediction, ML-model



RANS calculation with the ML-augmented Turbulence model, α = 14.2°

Results: The S809 airfoil

- The ML-augmented turbulence model greatly improves the RANS results in areas where c_l vs α is non-linear, i.e. where flow separation is dominant
- This is not only true for the datasets the ML-model was trained on (α = 12.2°, 16.2°, 18.2°), but also over a wider range (6.2° < α < 20.1°)
- Differences between match in c_p and c_l as a result of experimental data



Lift Coefficient over angle of attack for different modeling approaches

Conclusions

- Provided insight into DLRs work on data-driven techniques

1) Supplementing CFD capabilities by ML-based models for multi-query problems
    - Compared 3 different techniques
        a) POD + Interpolation
        b) Isomap + Residual Minimization
        c) Regression based on Neural Networks
    - Presented results for NASA CRM

2) Enhancing CFD capabilities by ML-based turbulence modeling
    - Revisited field inversion and ML step
    - Presented results for S809 airfoil

Future Work

- Next steps for the data-driven reduced order modeling part
  - Enabling residual minimization techniques on subset of points
  - Exploring different network architectures
  - Combining physical insights with neural networks

- Next steps for the data-driven turbulence modelling part
  - Investigations on the influence of the field inversion, i.e. the β-distribution due to
    - Regularization, type of cost function, amount of reference data available
  - Application of the FIML-approach to new test cases
  - Extensive data exploration and feature selection/ engineering