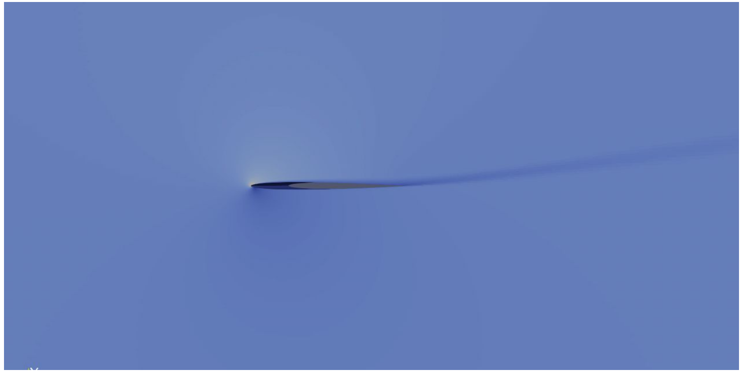
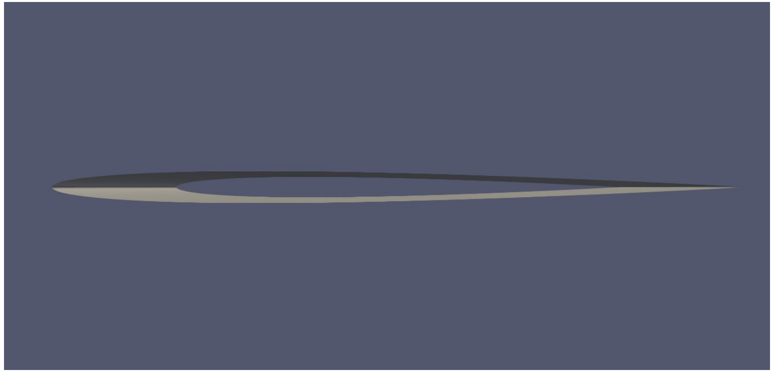
Inferring Functional Properties from CFD

Francesco Montanaro

Aim of the study



CFD-computed flow fields around the airfoil



shape of the airfoil

The aim of the study is to derive the geometrical features of the airfoils from the CFD outcomes by means of their **NACA** numbers:

- **Maximum Camber.**
- **Maximum Camber position.**
- **Thickness.**

Since an analytical link between the CFD-computed flow fields and the required information is not available, we refer to the use of **Machine Learning**.

Dataset

2600 .*vtk* files containing the values of the CFD-computed flow fields within the space.

Target values: **NACA** numbers.

1. Maximum camber → *[0 - 9] in units of c/100*
2. Maximum camber position → *[0 - 9] in units of c/10*
3. Maximum thickness → *[05 - 50] in units of c/100*

Features extraction

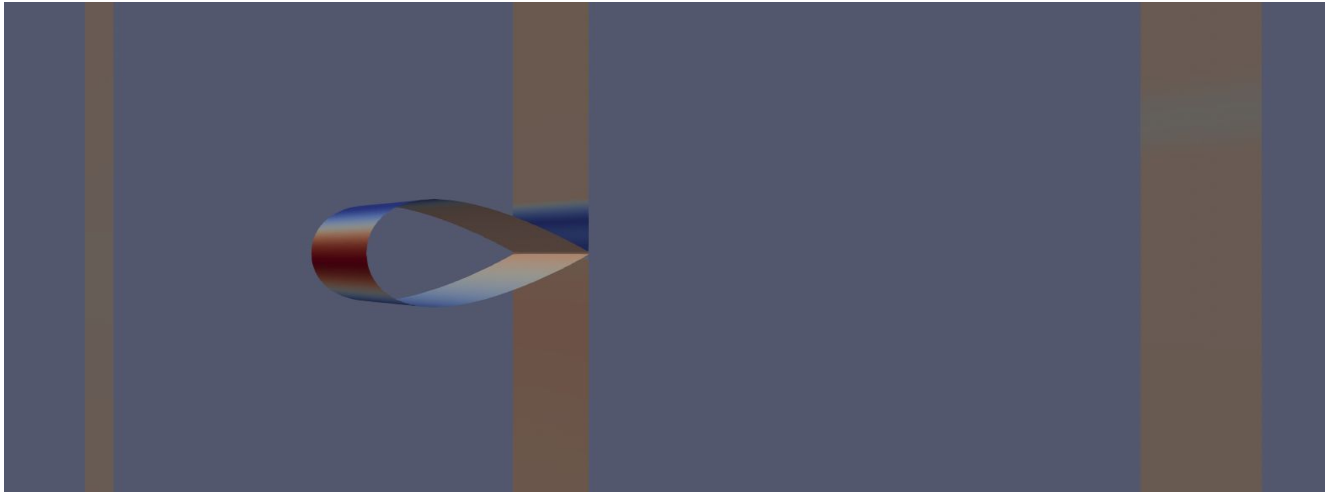
1. **Regional Averages** of the flow fields.
2. **Arrival times** of the streamlines.
3. **Regional arrival times** of the streamlines.

→ **Machine Learning**

4. **Signals** of the flow fields.
5. **Signals** of the streamlines.

→ **Deep Learning**

1. Regional Averages

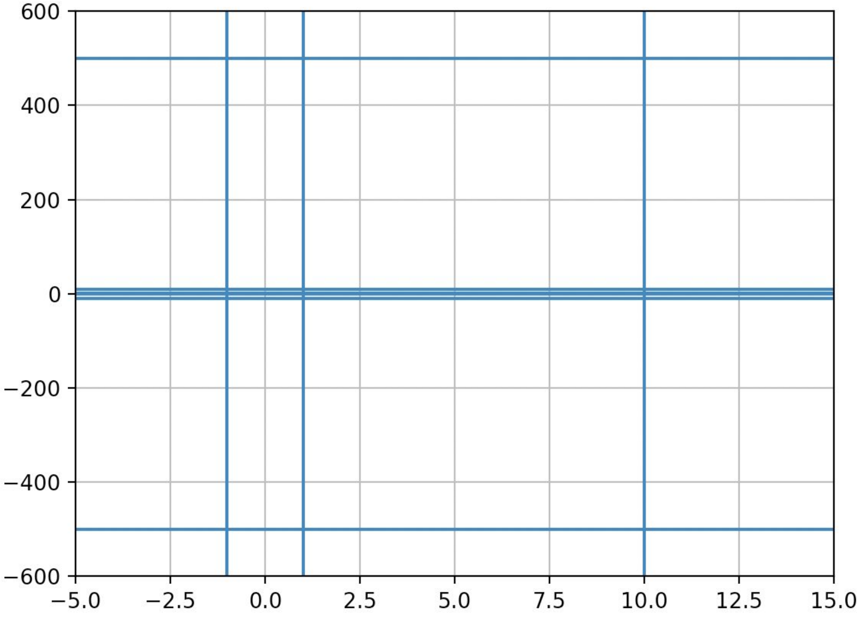


Sections extracted from the space

Goal:

- Extract the regions of interest.
- Compute the average of the flow fields for each region.

1. Regional Averages: cutting the space



- 3 sections considered.
  X = [-c, c, 10c]
- 8 regions per section.
  Y = [500, 10, 1, 0.1, 0, -0.1, -1, -10, -500]
- 24 regions in total.

c: chord length

1. Regional Averages: Features extraction

$$\overline{P}_k = \frac{\sum_{i}^{n} A_i p_i}{\sum_{i}^{n} A_i}$$

Regional average of the **pressure** field.

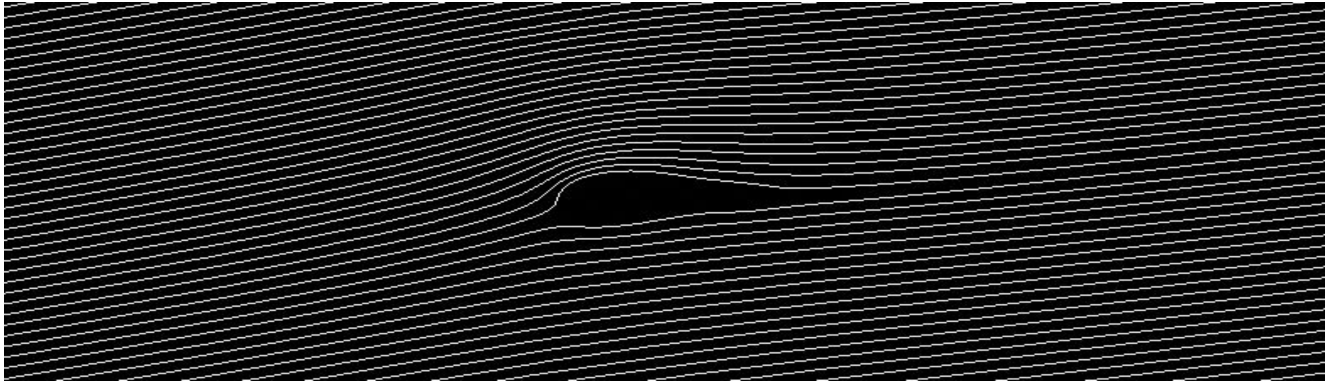
n: *number of cells belonging to the k-th region.*
pi: *pressure of the i-th cell.*
Ai: *surface of the i-th cell.*

$$\overline{U}_k = \frac{\sum_{i}^{n} A_i U_i}{\sum_{i}^{n} A_i}$$

Regional average of the **velocity** field.

n: *number of cells belonging to the k-th region.*
Ui: *velocity of the i-th cell.*
Ai: *surface of the i-th cell.*

2. Arrival times

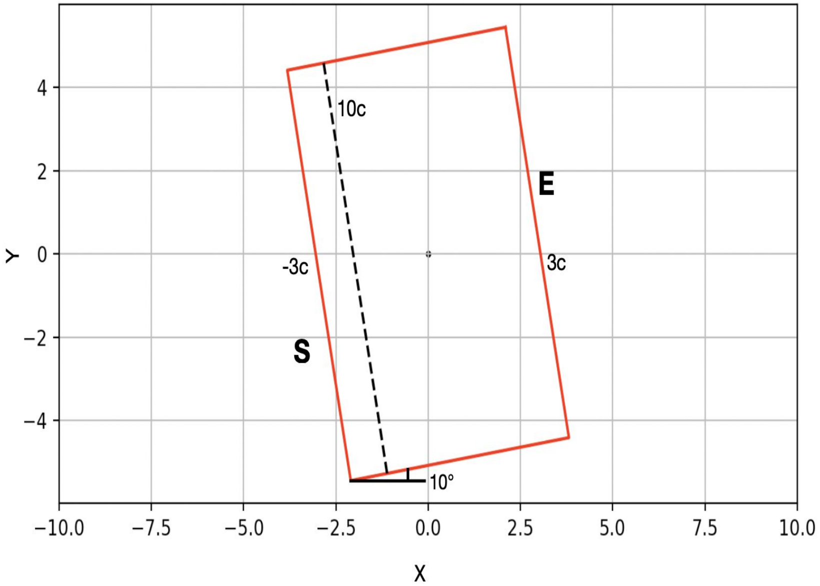


Streamlines affecting an airfoil

Goal:

- Extract streamlines.
- Compute the arrival time of each streamline from a starting section *S* to an arrival section *E*.

2. Arrival times: cutting the space - 1/2

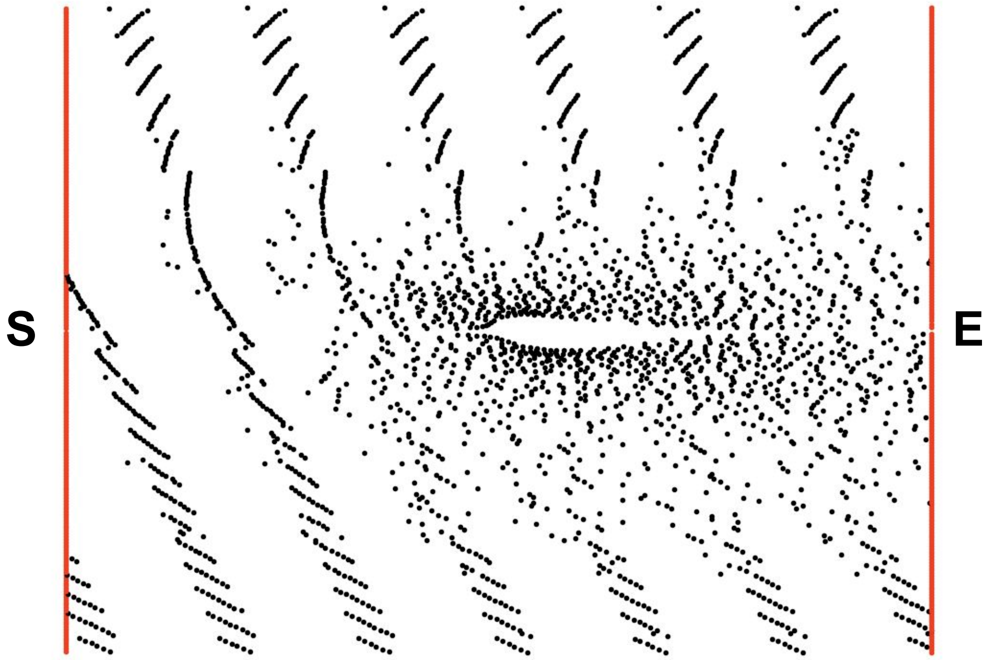


Starting section **S** and arrival section **E** both **orthogonal** to the free stream direction (**10°**):

- Sections length: **10c**
- Sections distances from the origin: **3c**

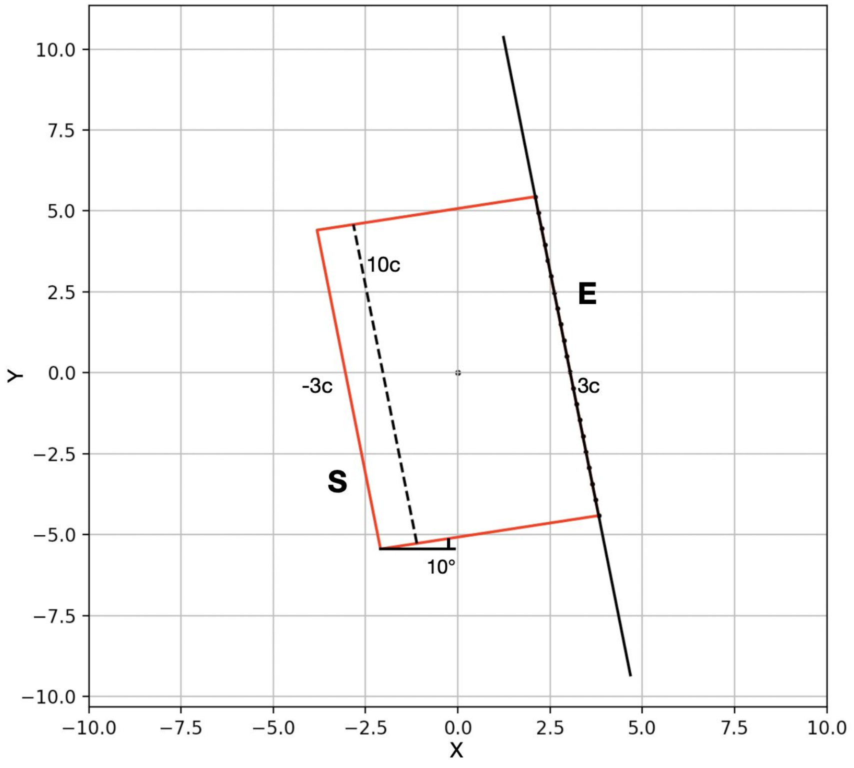
c: chord length

2. Arrival times: cutting the space - 2/2



- Extracting streamlines belonging to the cutting plane.
- Streamlines computed by using an adaptive integrator (**RungeKutta45**) with a variable step length.

2. Arrival times: Binning operation



- *1* segment considered at the arrival section *E*.
  section length = 10c
- Segment divided into *64* bins of equal length.

c: chord length

2. Arrival times: features extraction

$$\Delta t_{i,j} = \frac{2\sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}}{|U_j| + |U_i|}$$

Time distance of two consecutive points.

Ui*: velocity field of the i-th point.*

xi*,* ***yi****: spatial coordinates of the i-th point*

$$T_k = \sum_{i=1}^{n-1} \Delta t_{i,i+1}$$

Arrival time of the k-th streamline.

n*: number of points, of the k-th streamline, belonging to the cutting plane.*

$$T_{b_i} = \frac{1}{n_i} \sum_{h=1}^{n_i} T_h$$

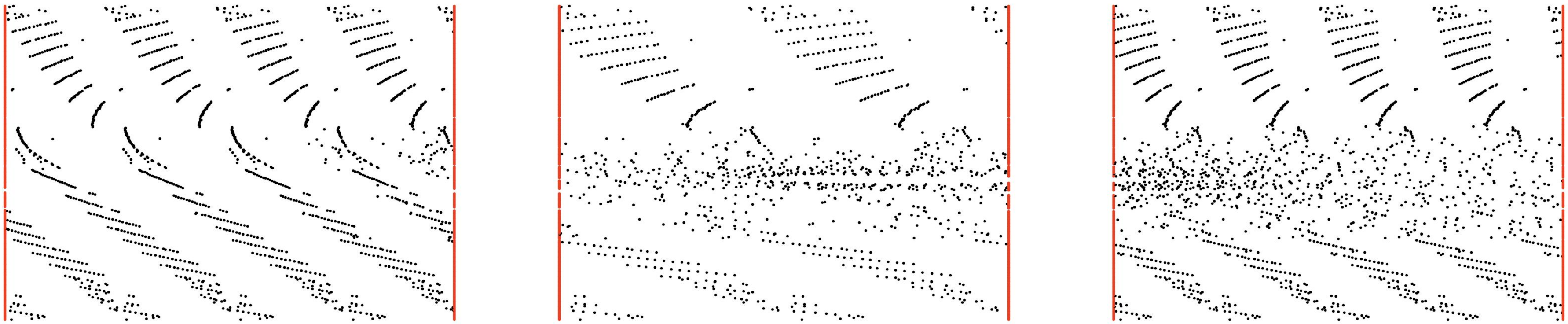
Arrival time of the i-th bin.

bi*: i-th bin.*

ni*: number of streamlines belonging to the i-th bin.*

Th*: arrival time of the h-th streamline.*

3. Regional arrival times

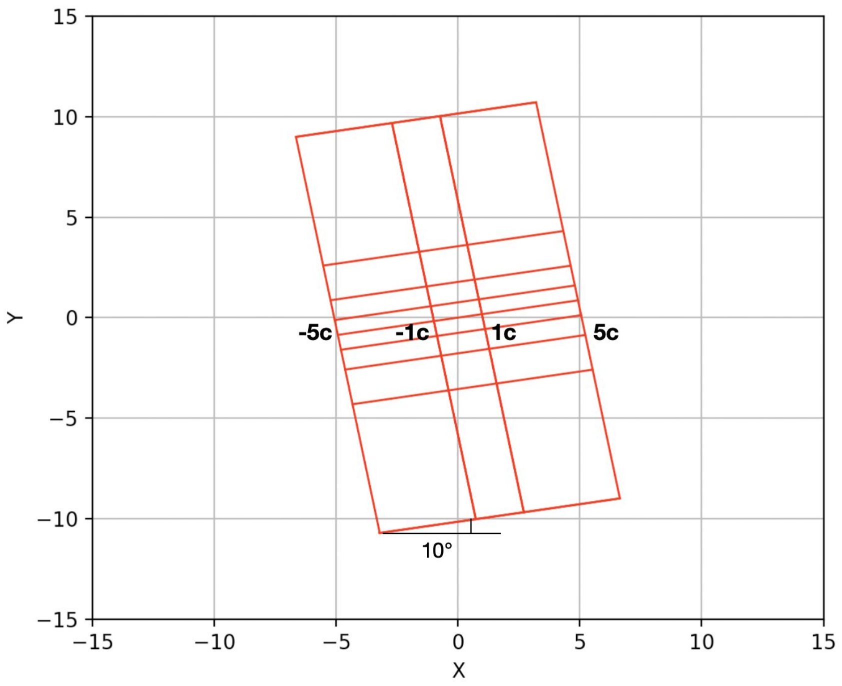


Streamlines extracted from different sections

Goal:

- Extract streamlines.
- Compute the arrival time of each streamline for multiple cutting sections.

3. Regional arrival times: cutting the space



- 3 regions on the x axis.
  X = [[-5c, -1c], [-1c, 1c], [1c, 5c]]
- 8 regions on the y axis.
  Y = [-10, -3.5, -1.75, -0.75, 0, 0.75, 1.75, 3.5, 10]
- 24 regions in total.

c: chord length

3. Regional arrival times: features extraction

$$\Delta t_{i,j} = \frac{2\sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}}{|U_j| + |U_i|}$$

Time distance of two consecutive points.

Ui: *velocity field of the i-th point.*

xi, ***yi***: *spatial coordinates of the i-th point*

$$T_k = \sum_{i=1}^{n-1} \Delta t_{i,i+1}$$

Arrival time of the k-th streamline.

n: *number of points, of the k-th streamline, belonging to the cutting plane.*

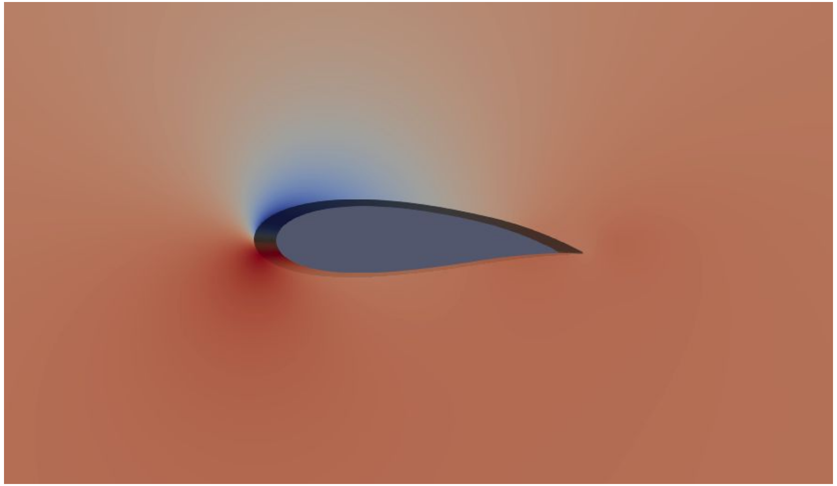
$$R_{T_i} = \frac{1}{n_j} \sum_{j=1}^{n_j} T_j$$

Regional average of the arrival i-th time.

nj: *number of streamlines belonging to the i-th region.*

Tj: *arrival time of the j-th streamline.*

4. Signals of the flow fields: features extraction

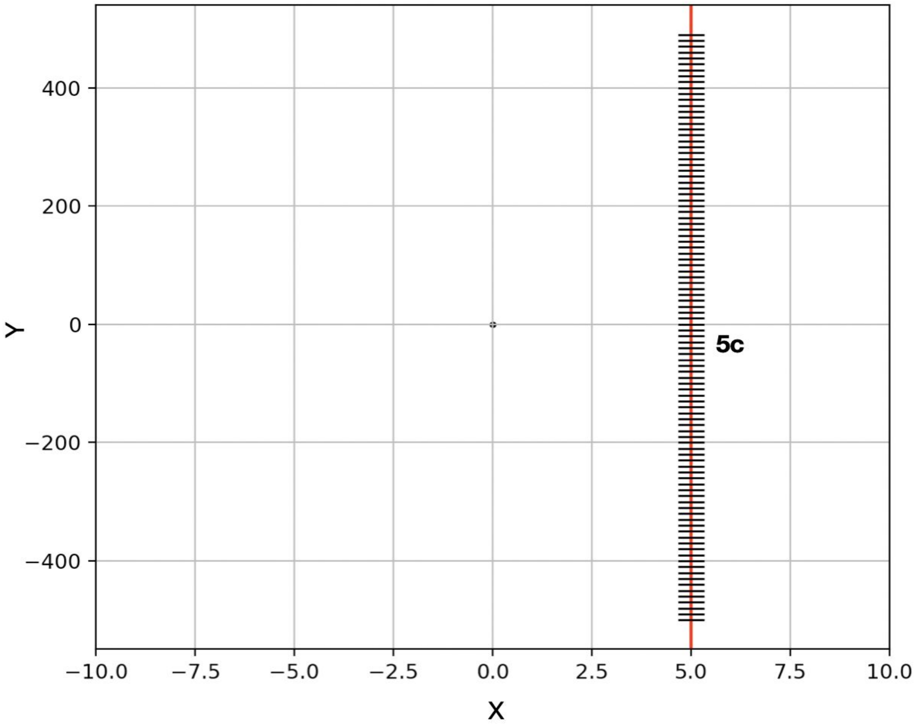


CFD-computed flow fields around the airfoil

Goal:

- Extract the values of the CFD-computed flow fields for a specific segment of the space.
- Generate spatial signals from the obtained values.

4. Signals of the flow fields: Binning operation



Binning operation

- *1* segment considered.
  x = 5c
  y = [-500, 500]
  z = 0.5

- Section divided into *1024* bins of equal length.

c: chord length

4. Signals of the flow fields: features extraction - 1/2

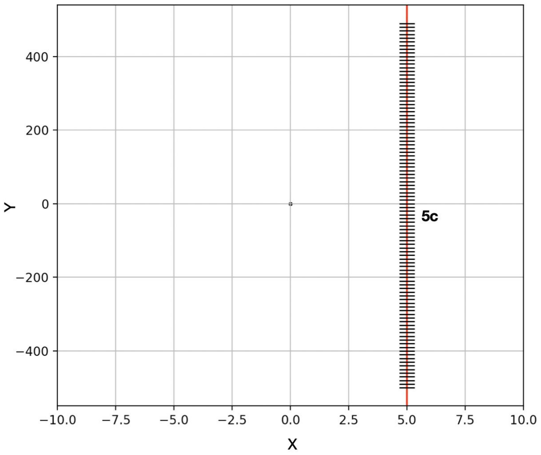
$$P_{b_i} = \frac{1}{n_i}\sum_{h=1}^{n_i} p_h$$

Pressure of the i-th bin.
bi: *i-th bin.*
ni: *number of points belonging to the i-th bin.*
ph: *pressure of the h-th point of the i-th bin.*

$$|U|_{b_i} = \frac{1}{n_i}\sum_{h=1}^{n_i} \left|u_h\right|$$

Magnitude of the **Velocity** of the i-th bin.
bi: *i-th bin.*
ni: *number of points belonging to the i-th bin.*
uh: *magnitude of the velocity of the h-th point of the i-th bin.*

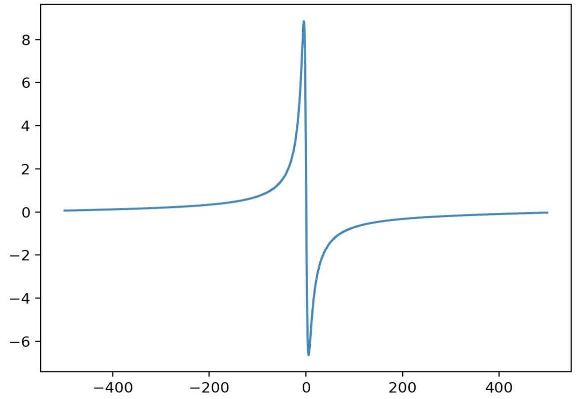
features extraction - 2/2



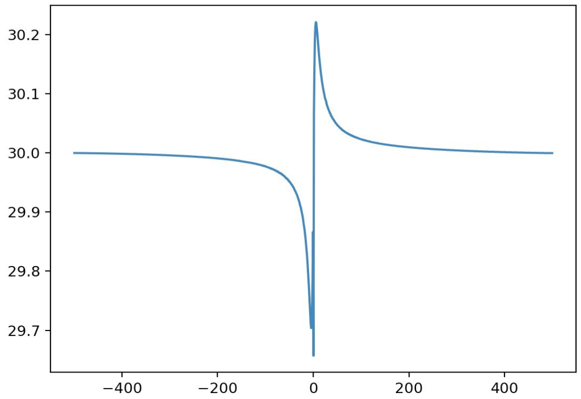
Binning operation

$$P_{b_i} = \frac{1}{n_i}\sum_{h=1}^{n_i} p_h$$

$$|U|_{b_i} = \frac{1}{n_i}\sum_{h=1}^{n_i} \left|u_h\right|$$

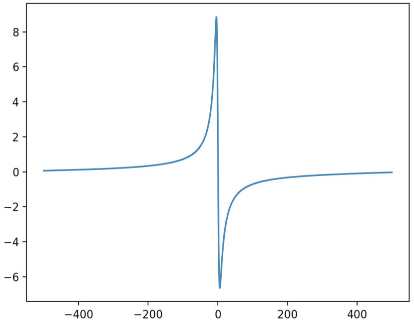


Pressure signal

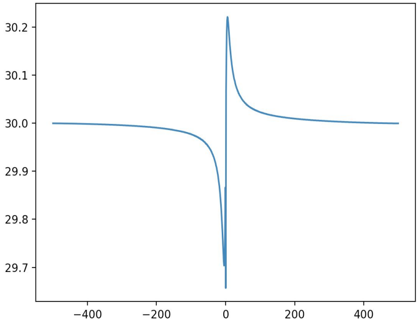


Velocity signal

NACA: **0005**

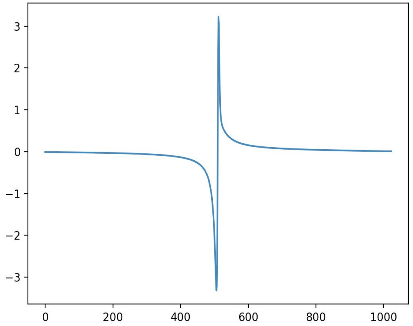


Pressure

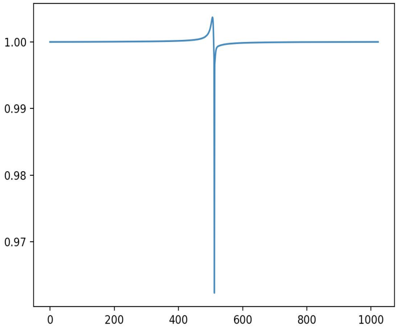


Velocity

NACA: **0050**

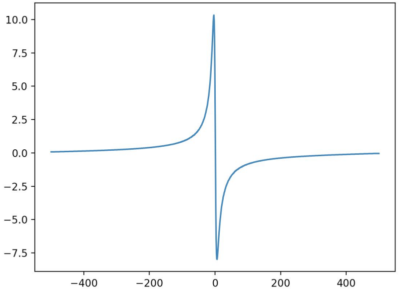


Pressure

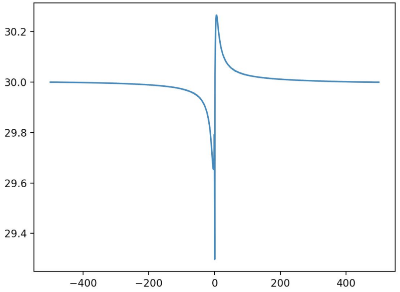


Velocity

NACA: **5732**

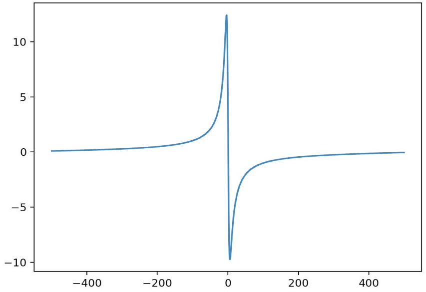


Pressure

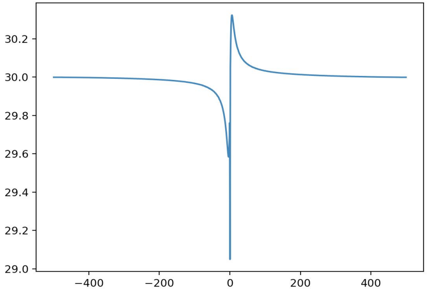


Velocity

NACA: **8732**

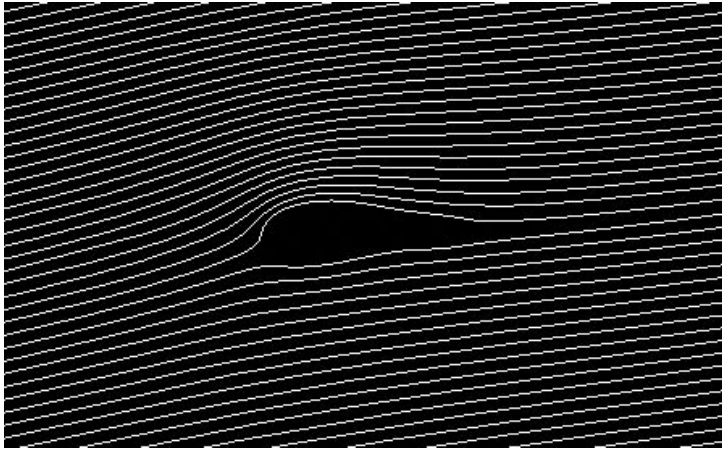


Pressure

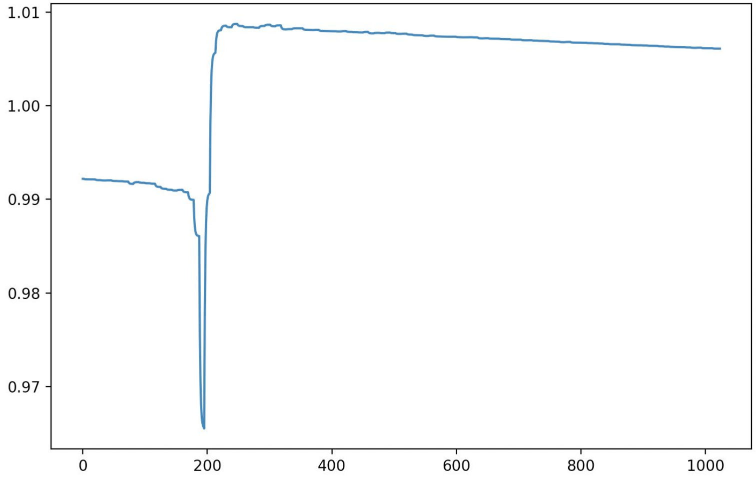


Velocity

5. Signals of the streamlines



Streamlines affecting an airfoil

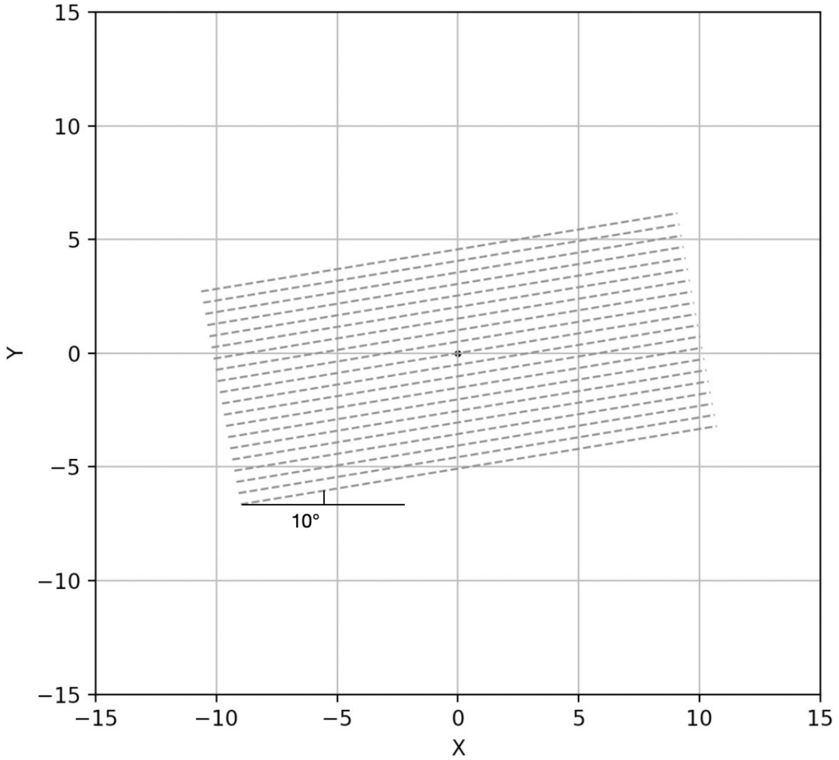


Signal of the streamline

Goal:

- Extract streamlines within the space.
- Generate spatial signals of the velocity field.

5. Signals of the streamlines: Binning operation



- Space divided into *1024* bins of equal length.
- Bins orthogonal to the free stream direction.

c: chord length

5. Signals of the streamlines: features extraction - 1/2

$$|u|_h = \frac{1}{n_h} \sum_{j=1}^{n_h} |u|_j$$

Mean velocity magnitude of the points belonging the h-th streamline.

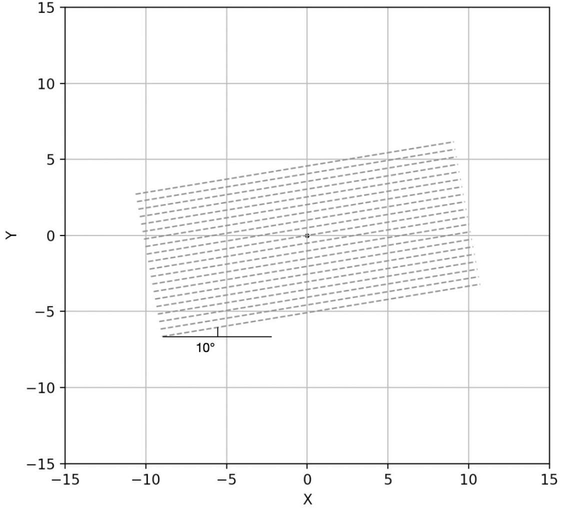
ui: *velocity magnitude of the i-th point.*
nh: *number of points of the i-th streamline.*

$$|U|_{b_i} = \frac{1}{n_i} \sum_{j=1}^{n_i} |U|_j$$

Velocity magnitude of the i-th bin.

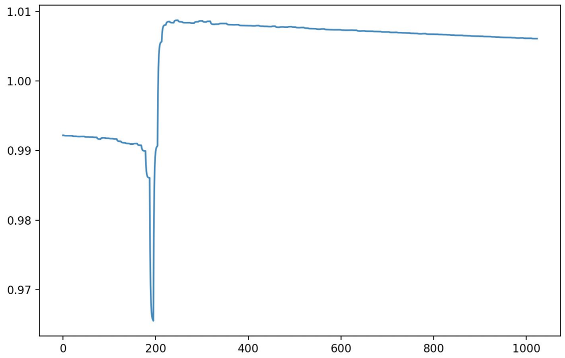
bi: *i-th bin.*
ni: *number of streamlines belonging to the i-th bin.*
Ui: *mean velocity magnitude of the j-th streamline.*

5. Signals of the streamlines: Features extraction - 2/2



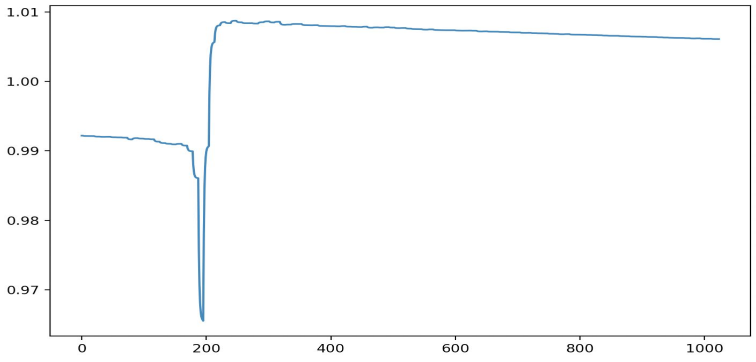
Binning operation

$$|U|_{b_i} = \frac{1}{n_i} \sum_{j=1}^{n_i} |U|_j$$

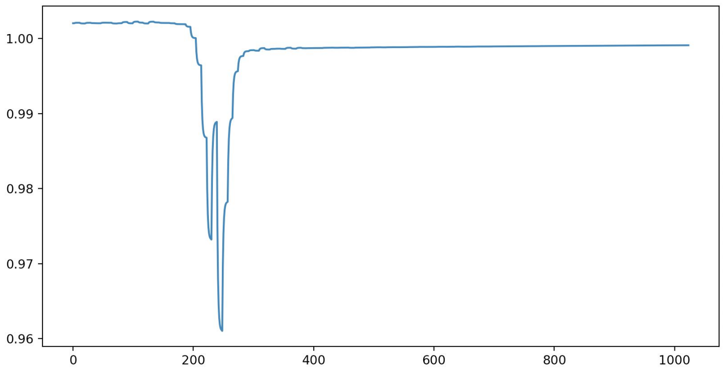


Signal of the streamlines

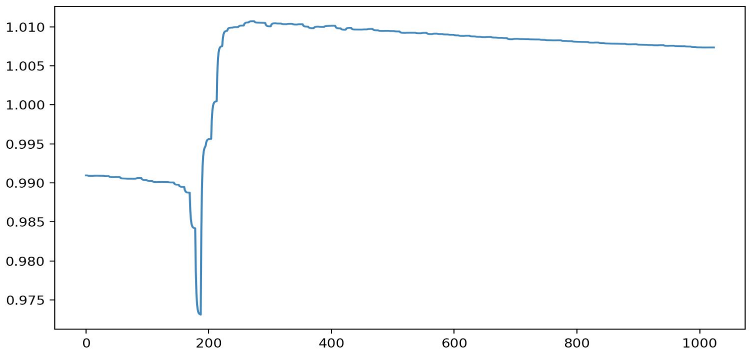
NACA: **0005**



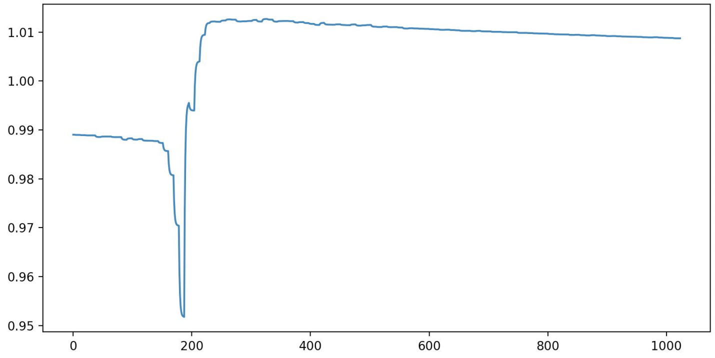
NACA: **0050**



NACA: **5732**



NACA: **8732**



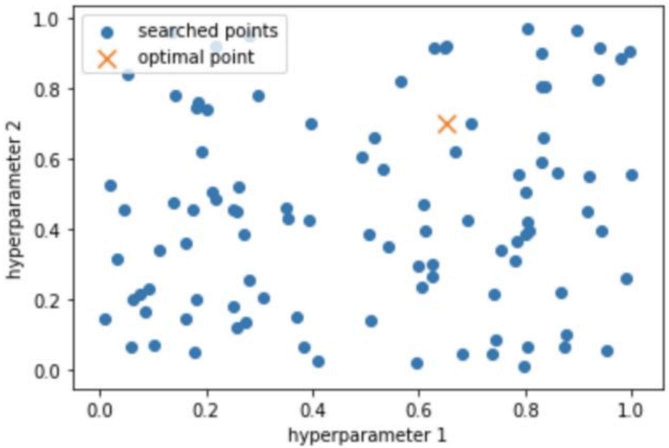
Prediction of Geometrical Features

TensorFlow → Building the models

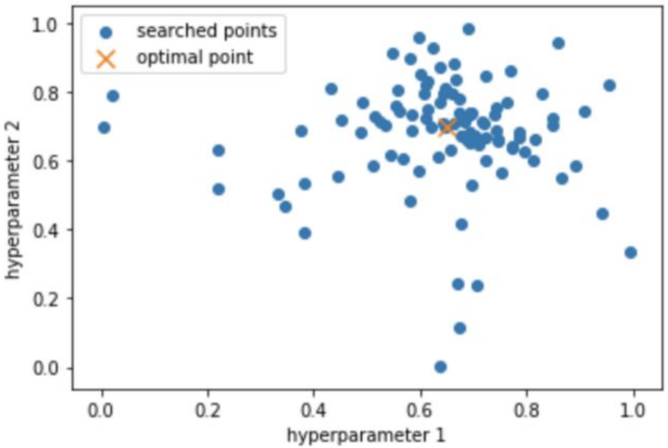
+

OPTUNA → Optimizing the models

Optuna for hyperparameters tuning - 1/2



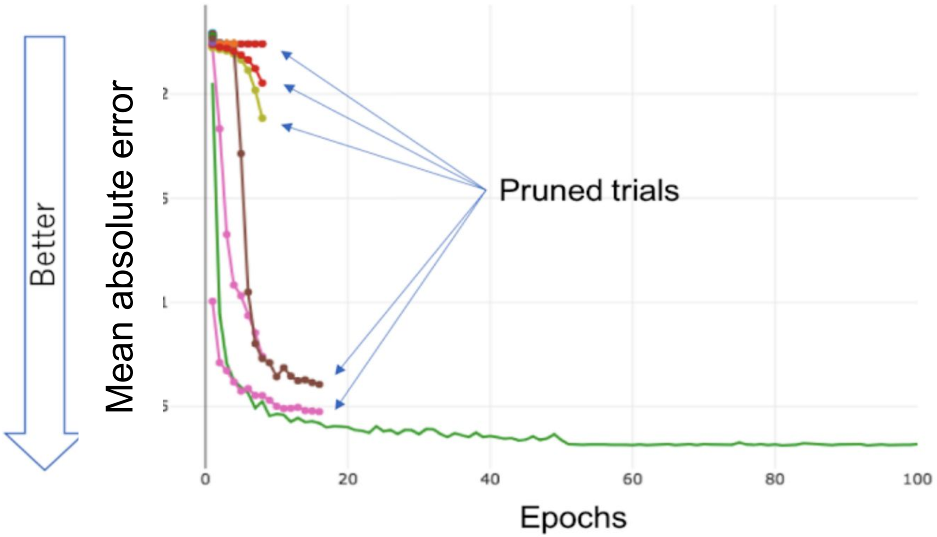
Random search



Sequential Model-based Optimization (TPE)

Optuna uses Tree-structured Parzen Estimator (TPE) to **search** more efficiently than a random search, by iteratively choosing points closer to previous good results.

Optuna for hyperparameters tuning - 2/2



- More than **150 architectures tested** for each class of features.

- More than **750 architectures tested** in total.

Optuna uses a **Pruning** optimization technique to automatically stop unpromising trials at the early stages of the training, so that computing time can be used for trials that show more potential.

Prediction of Geometrical Features

- Two classes of models considered:
  1. Fully Connected Neural Network (**FCNN**):
     - Regional Averages.
     - Arrival times.
     - Regional arrival times.
  2. Convolutional Neural Network 1D (**CNN-1D**):
     - Signals of the flow fields.
     - Signals of the streamlines.

- Loss function: Mean Square Error (**MSE**) computed on the NACA numbers.

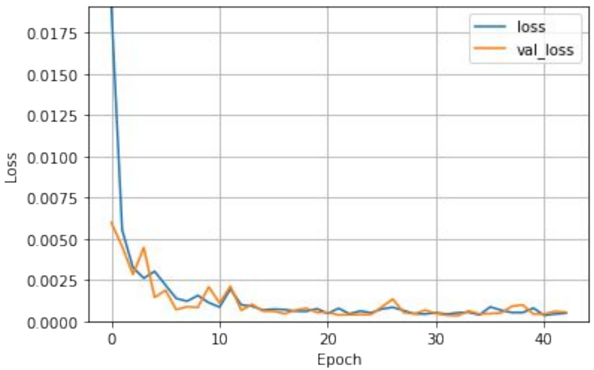
- Performance evaluation metric: Mean Absolute Error (**MAE**) computed on the NACA numbers.

Experimental setup

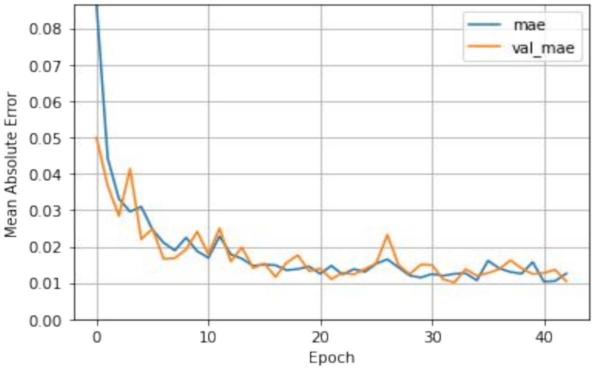
First Approach: all samples used in the training phase.

- Data **sampled randomly**.
- Data splitting:
    - **Training set**: *80%* of the available samples.
    - **Validation set**: *20%* of the training samples.
    - **Test set**: *20%* of the available samples.

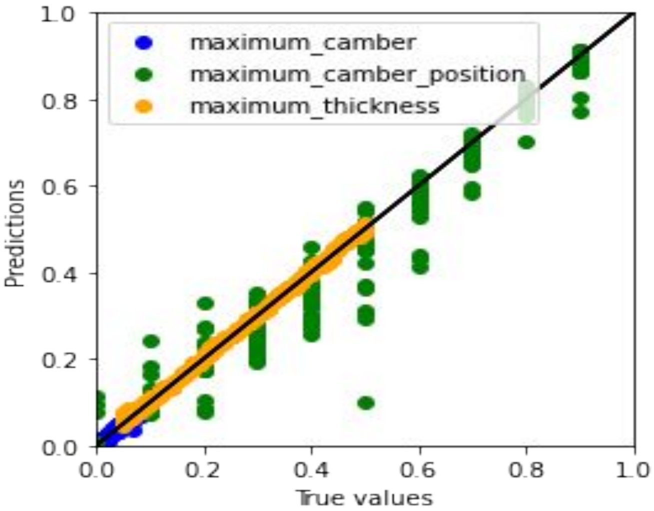
Regional Averages



Training and Validation loss trends across the epochs



Training and Validation MAE trends across the epochs



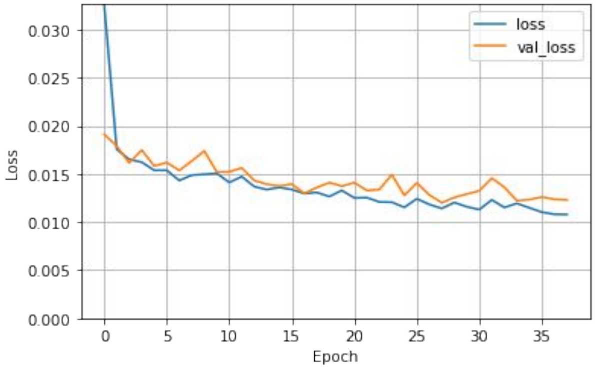
Prediction of the geometrical features

Metrics on the test set:

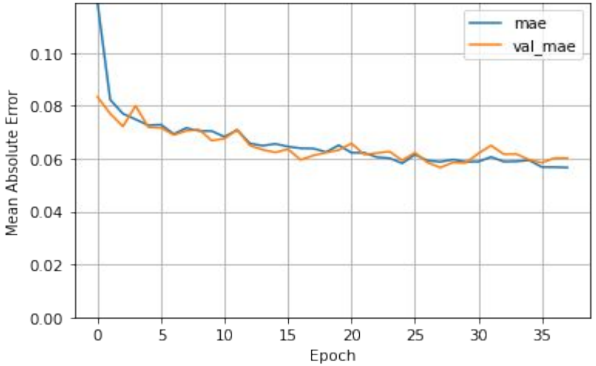
Loss (MSE): **0.00066293**

Mean Absolute Error (MAE): **0.011075435**

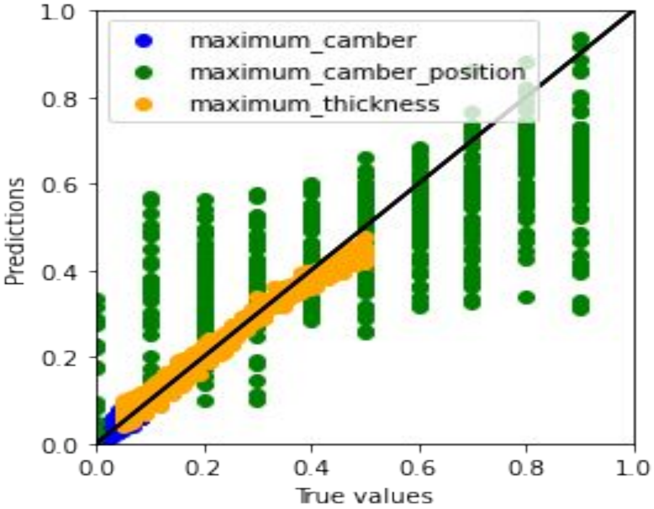
Arrival times



Training and Validation loss trends across the epochs



Training and Validation MAE trends across the epochs



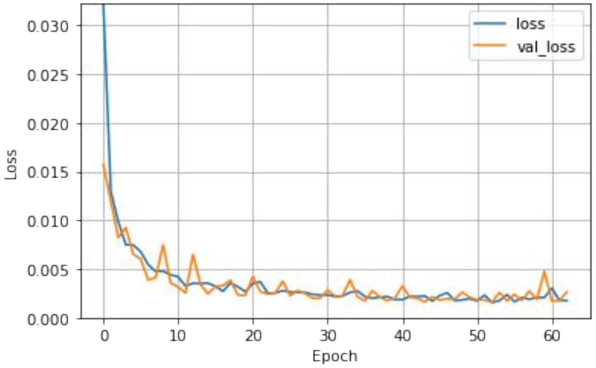
Prediction of the geometrical features

Metrics on the test set:

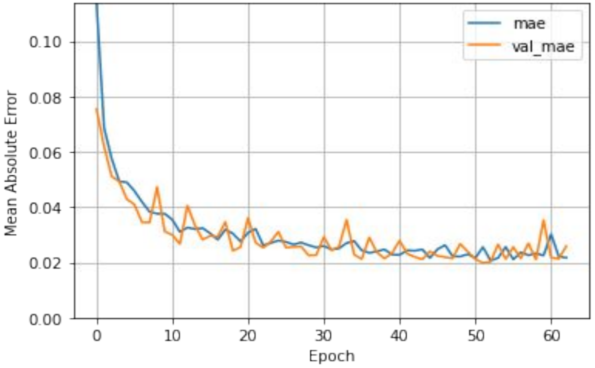
Loss (MSE): **0.012493741**

Mean Absolute Error (MAE): **0.060213528**

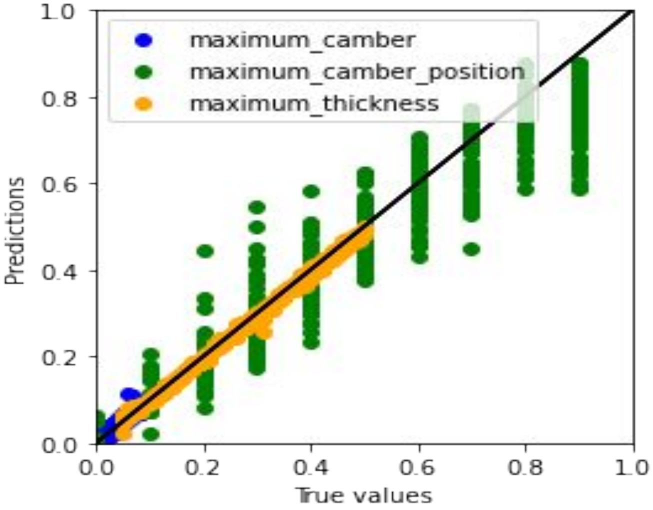
Regional arrival times



Training and Validation loss trends across the epochs



Training and Validation MAE trends across the epochs



Prediction of the geometrical features

Metrics on the test set:

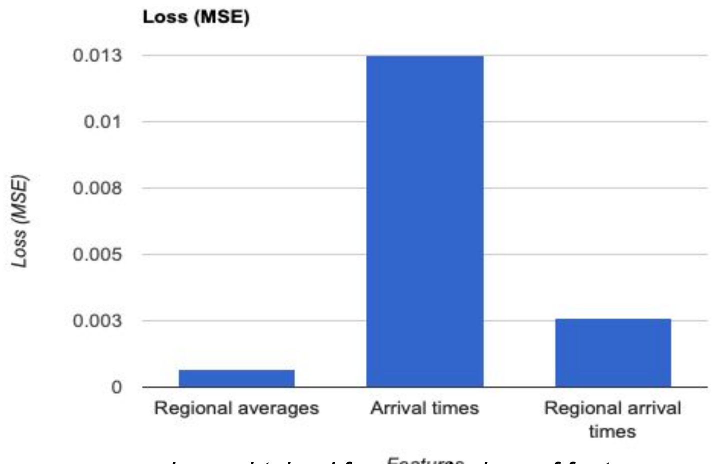
Loss (MSE): **0.002624211**

Mean Absolute Error (MAE): **0.025725699**

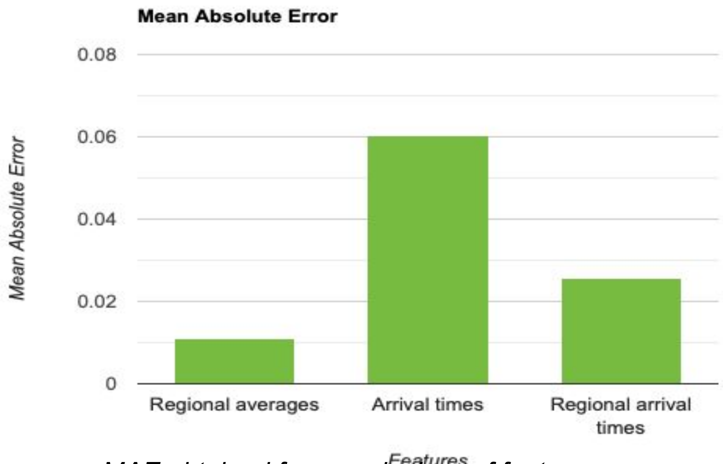
Comparison of metrics - FCNN

	Regional Averages	Arrival times	Regional arrival times
Loss (MSE)	0.00066293	0.012493741	0.002624211
Mean Absolute Error (MAE)	0.011075435	0.060213528	0.025725699

Metrics obtained from each class of features

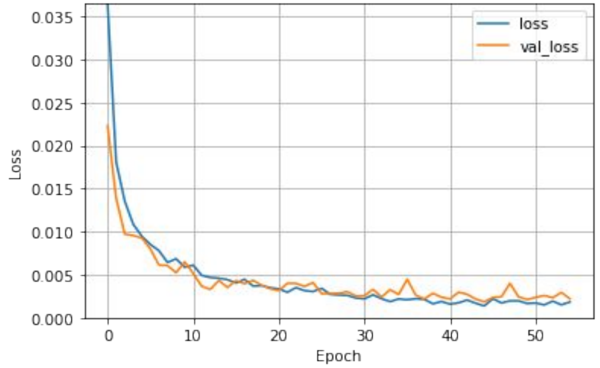


Loss obtained from each class of features

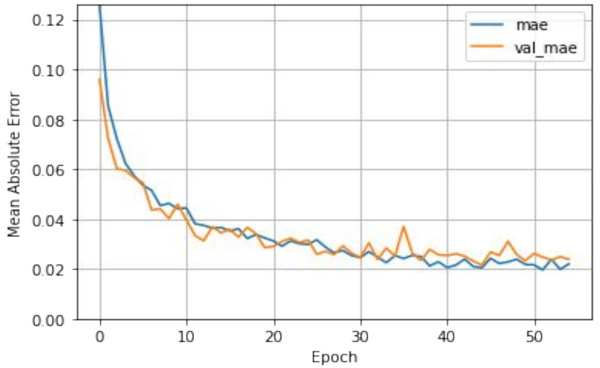


MAE obtained from each class of features

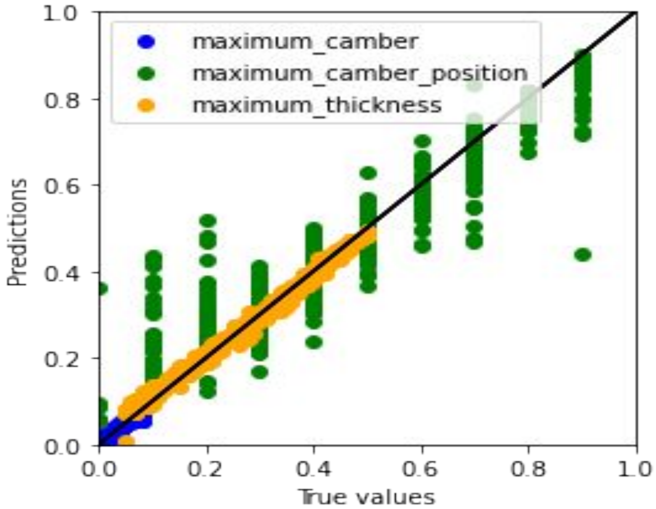
Signals of the flow fields



Training and Validation loss trends across the epochs



Training and Validation MAE trends across the epochs



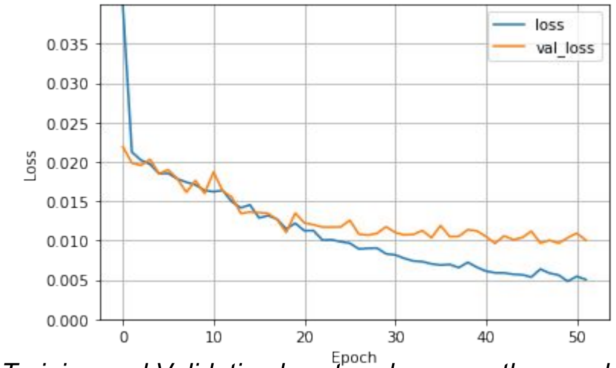
Prediction of the geometrical features

Metrics on the test set:

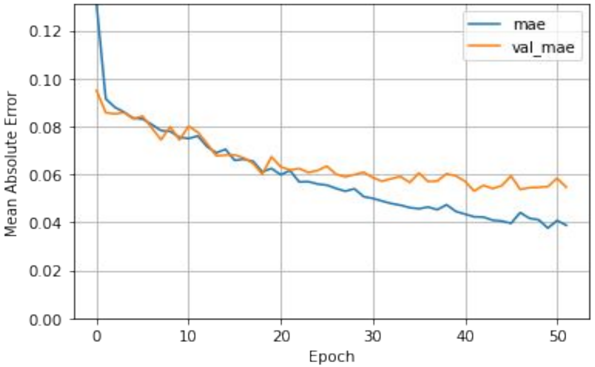
Loss (MSE): **0.002248327**

Mean Absolute Error (MAE): **0.024343444**

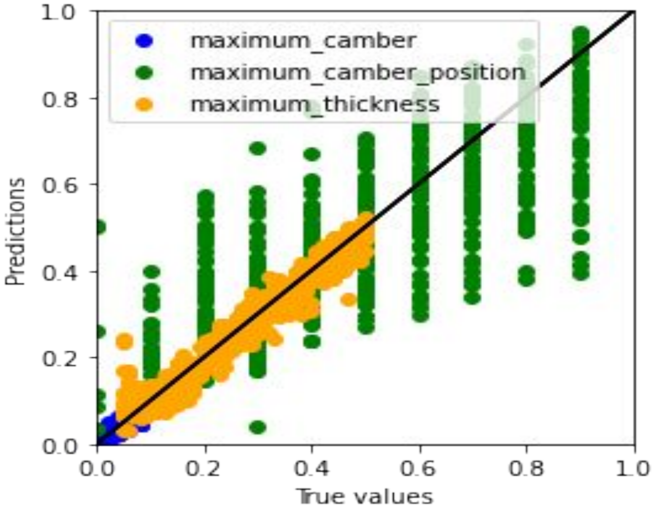
Signals of the streamlines



Training and Validation loss trends across the epochs



Training and Validation MAE trends across the epochs



Prediction of the geometrical features

Metrics on the test set:

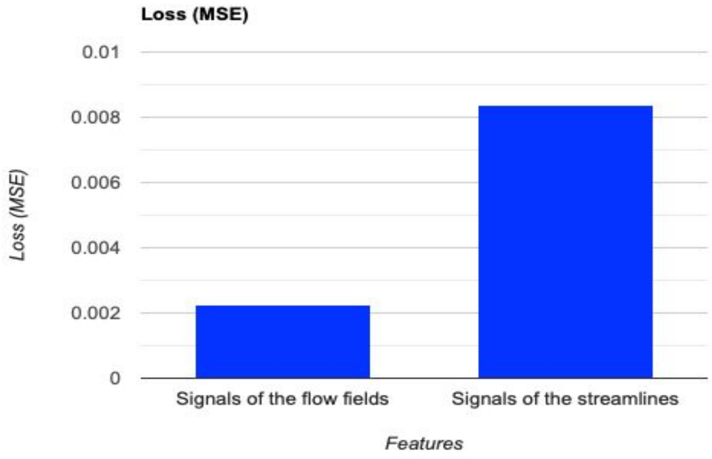
Loss (MSE): **0.008360184**

Mean Absolute Error (MAE): **0.049679368**

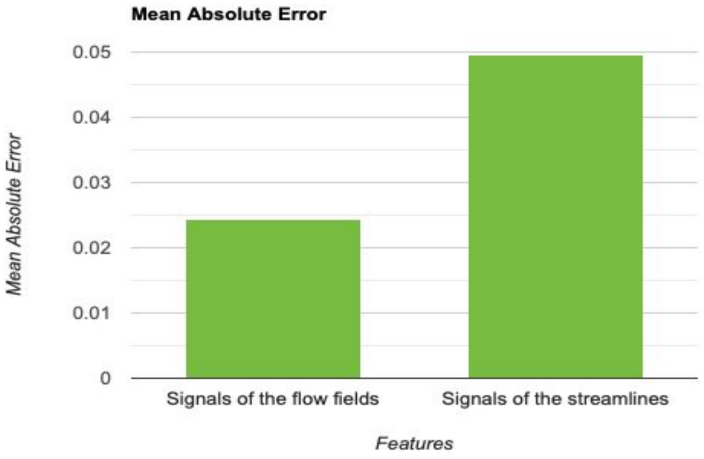
Comparison of metrics - CNN

	Signals of the flow fields	Signals of the streamlines
Loss (MSE)	0.002248327	0.008360184
Mean Absolute Error (MAE)	0.008360184	0.049679368

Metrics obtained from each class of features

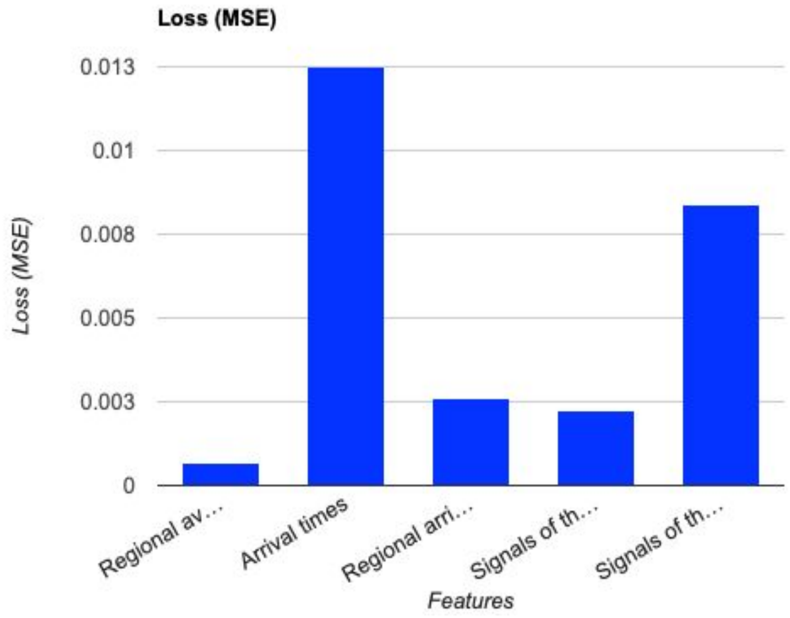


Loss obtained from each class of features

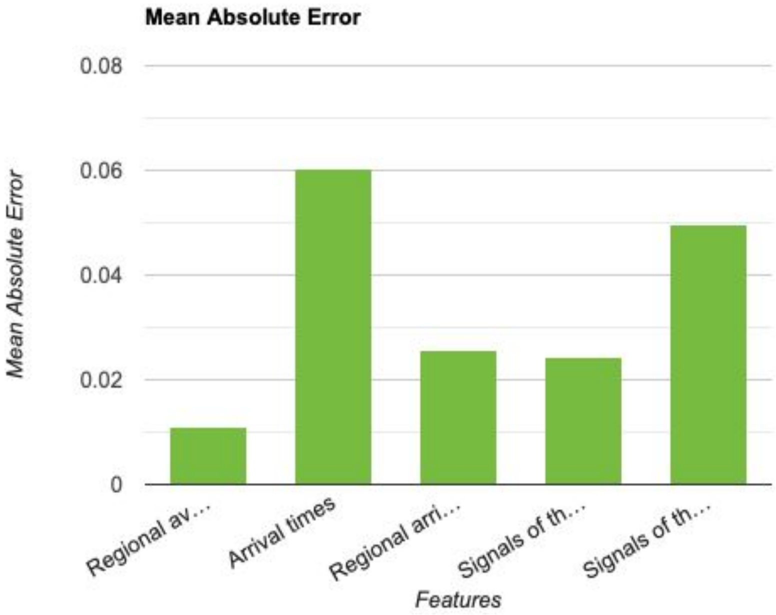


MAE obtained from each class of features

Comparison of metrics



Loss obtained from each class of features



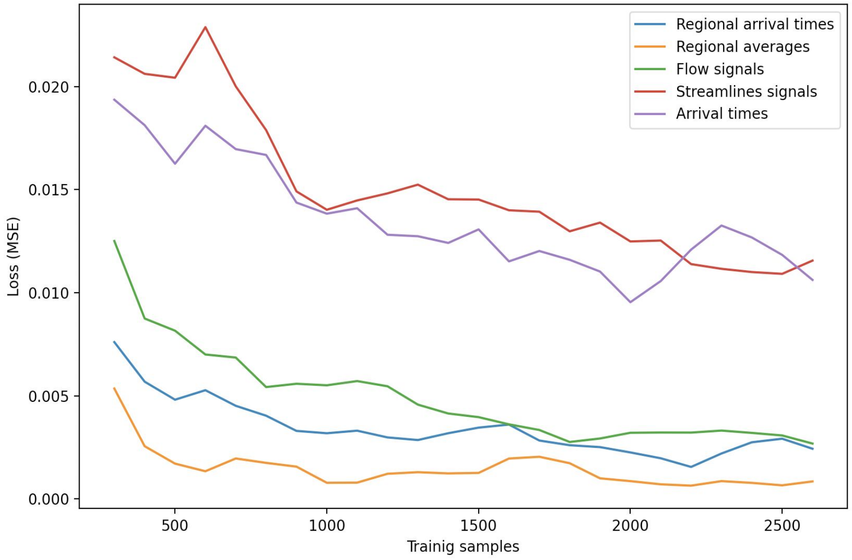
MAE obtained from each class of features

Experimental setup

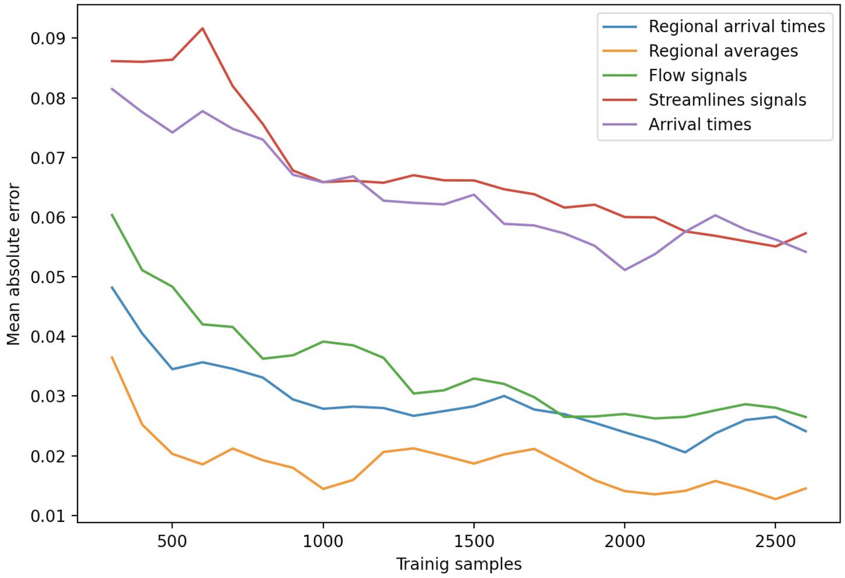
Second Approach: variation of the training set size.

- Incremental **step size** of *100* samples per experiment.
- Data **sampled randomly**.
- Data splitting:
    - **Training set**: *80%* of the available samples for training.
    - **Validation set**: *20%* of the training samples.
    - **Test set**: *20%* of the available samples.

Comparison of metrics



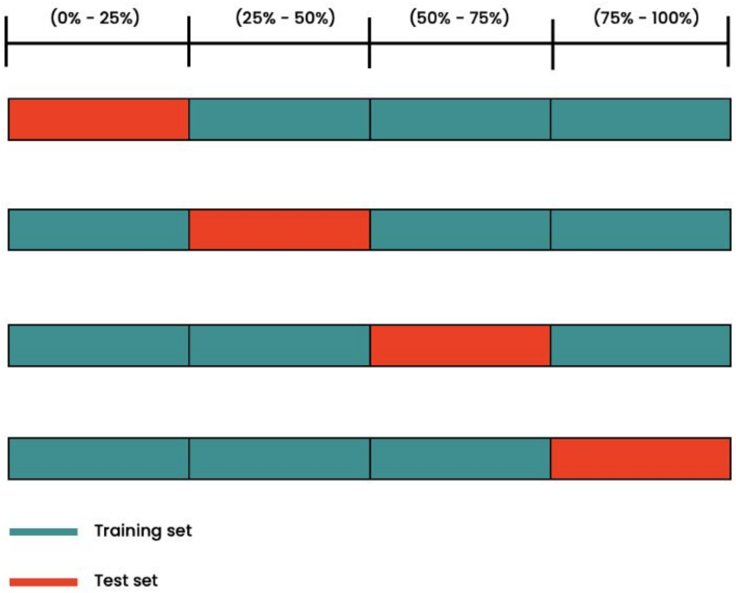
Loss trend w.r.t. the training set size for each class of features



MEA trend w.r.t. the training set size for each class of features

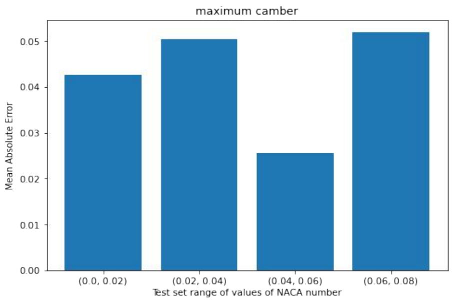
Extrapolation

Studying how the composition of the training set can affect the performances.

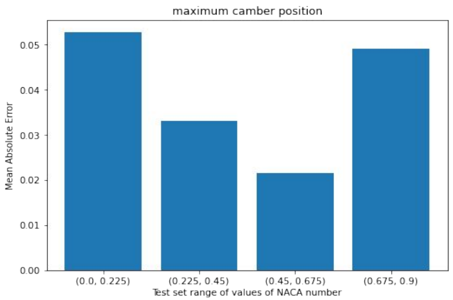


- *4* **not overlapping** subsets extracted for each NACA number by excluding specific range of values from the training set.
- *12* subsets in total.

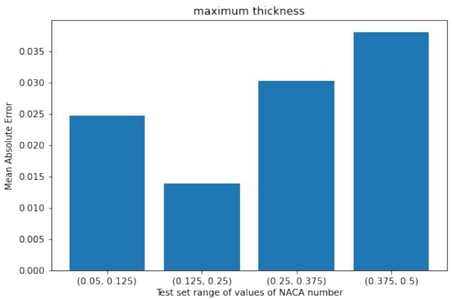
Regional Averages



Maximum camber

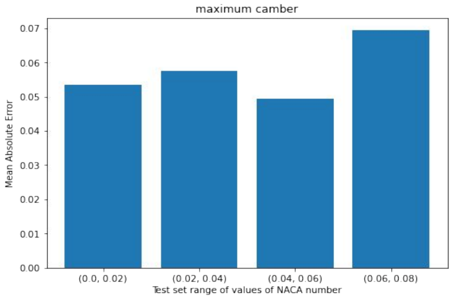


Maximum camber position

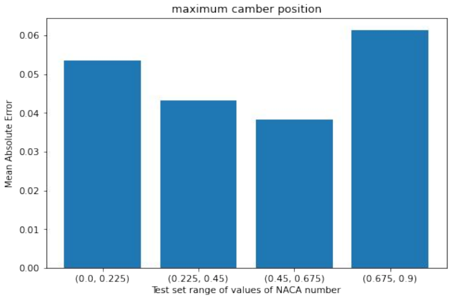


Maximum thickness

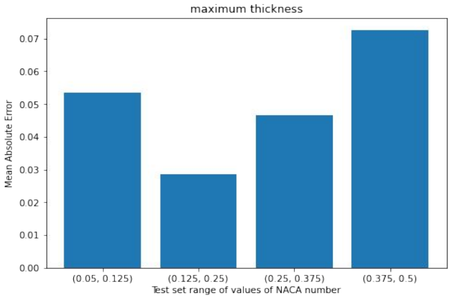
Regional arrival times



Maximum camber

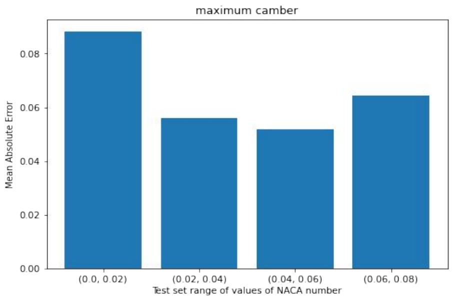


Maximum camber position

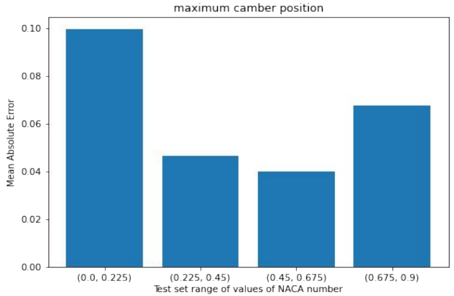


Maximum thickness

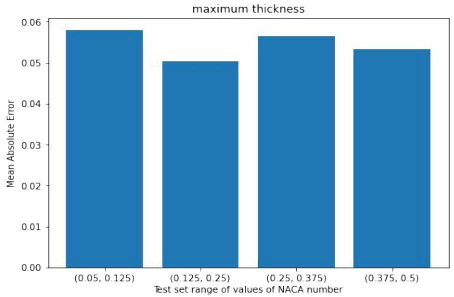
Signals of the flow fields



Maximum camber

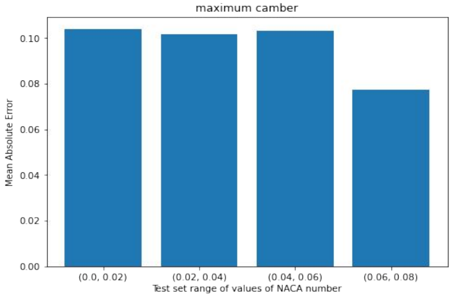


Maximum camber position

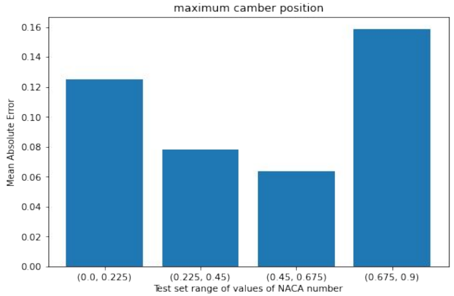


Maximum thickness

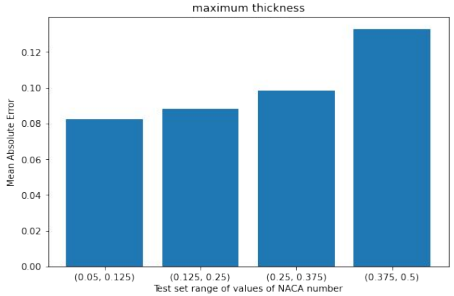
Signals of the streamlines



Maximum camber



Maximum camber position



Maximum thickness