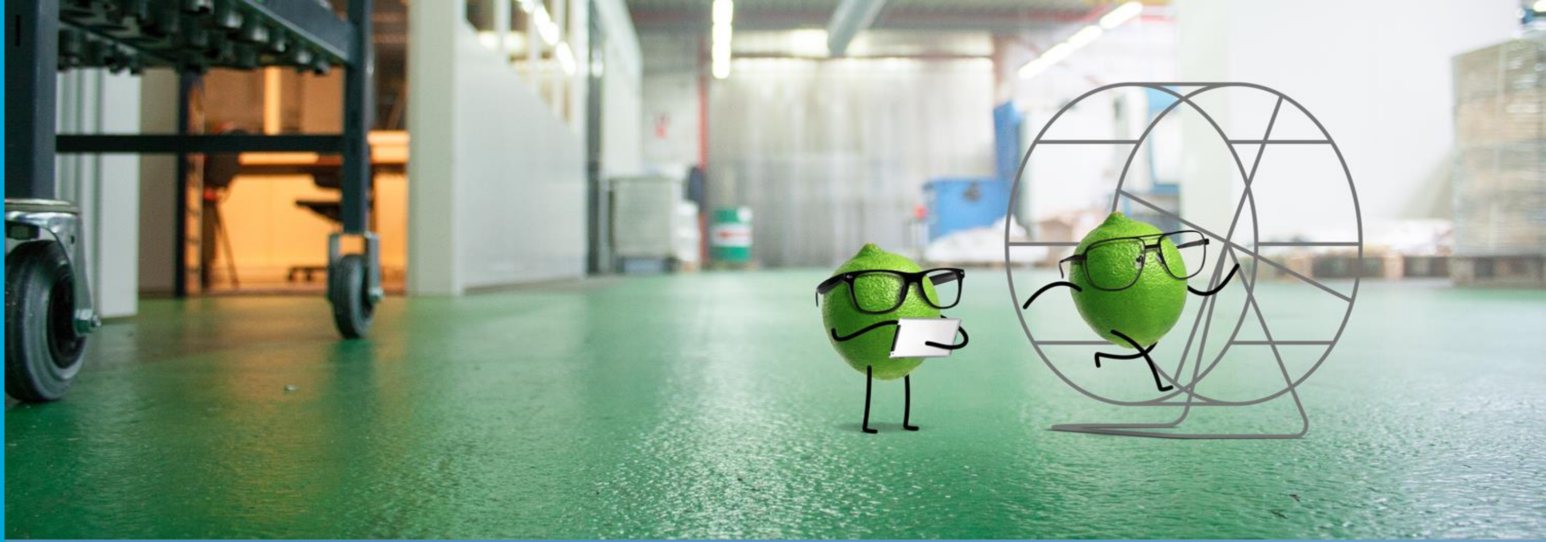
Data Driven Turbulence Modeling

Midterm Presentation:

Elske van Leeuwen

TUDelft

SIOUX TECHNOLOGIES



Turbulence and its Modeling

TU Delft

SIOUX TECHNOLOGIES

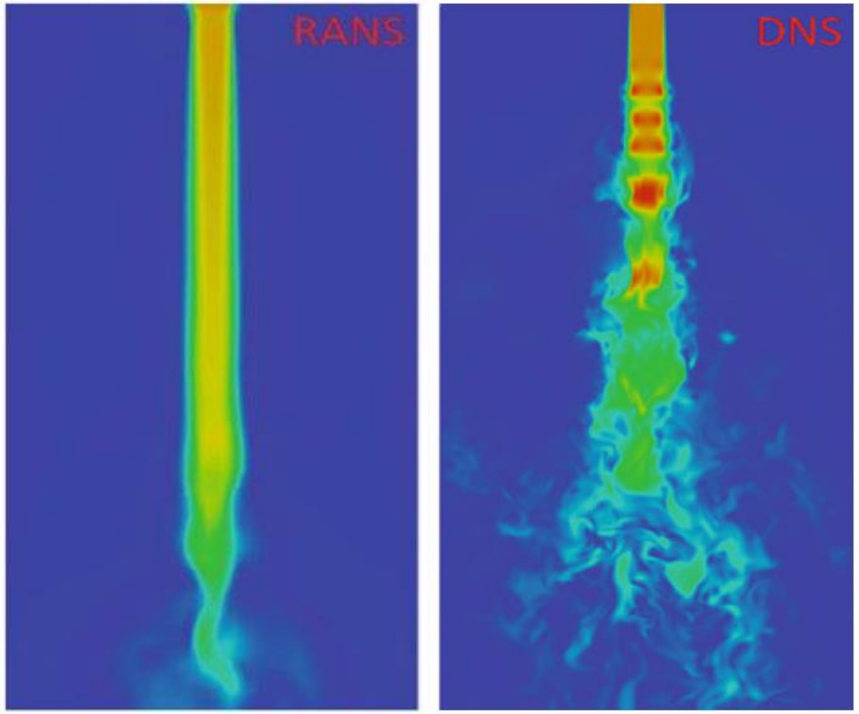
Turbulence

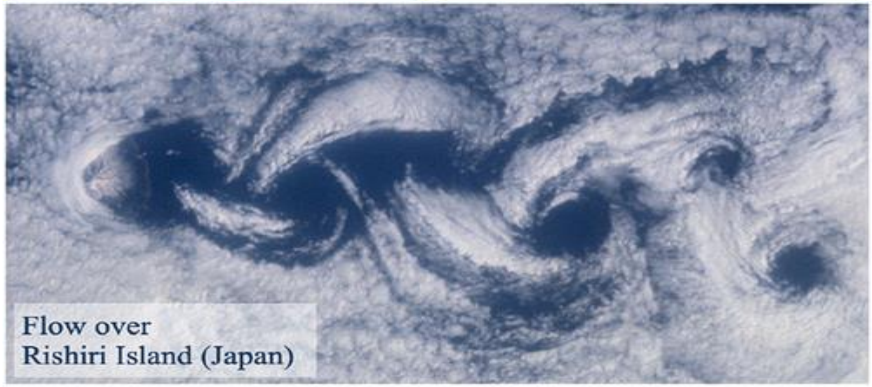
- Chaotic behavior of fluid flow
- Modeling improves the disign of technological applications

- Navier-Stokes equations describe the motion of a flow
  - $\rho\left(\frac{\partial u_i}{\partial t}+u_j\frac{\partial u_i}{\partial x_j}\right)=\frac{\partial p}{\partial x_i}+\mu\frac{\partial^2 u_i}{\partial x_j^2}$

Turbulence Modeling

- Direct Numerical Simulations (DNS)
  - High accuracy
  - High computational complexity
- Reynolds Average Navier Stokes (RANS) models
  - Low accuracy
  - Low computation
- Interest in Machine Learning increased:
  - ML is able to find patterns
  - More data available since the of computational power has increased

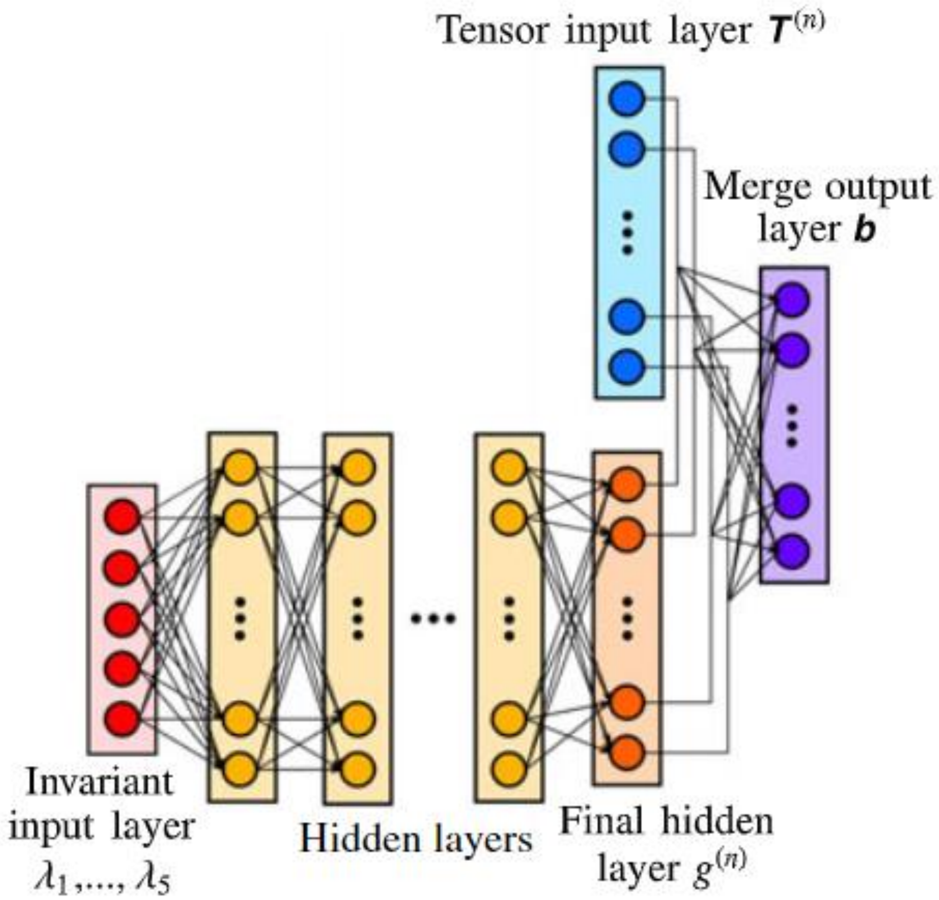



Flow over
Rishiri Island (Japan)


Flow over a cylinder
($Re = 100$)

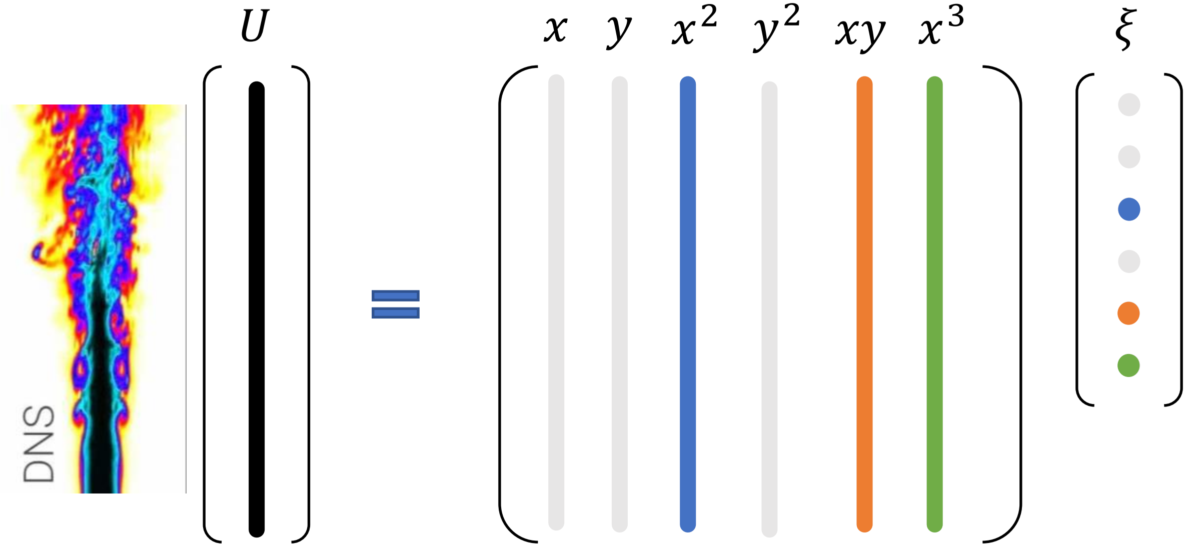
Challenges ML in Turbulence Modeling

- Interpretability
- Generalizability
- Simplicity
- Physically Informed
- Integration with CFD Solvers



Sparse Symbolic Regression

- Finds a mathematical expression for a quantity of interest
- Requires a library of candidate functions
- Selects function with sparse regression techniques



Target

Library of candidate functions $\mathbb{C}$

Model identification

$$\xi = \min \|U - \mathbb{C}\xi\|_2^2 + \lambda\|\xi\|_1$$

$$U = \mathbb{C}\xi$$

Sparse Symbolic Regression

- Interpretable
- Physics informed
- Efficient
- Simple
- Integration with CFD solvers (OpenFoam) possible and “easy”

Objectives and Questions

Objective

- Develop a data-driven turbulence model using sparse symbolic regression to improve a RANS turbulence model, which contains physical knowledge and should be interpretable, generalizable and robust

Questions

- Which features are the most relevant?
- How can physical knowledge be included?
- Which sparse symbolic regression technique finds the best performing algebraic model?
- How does the obtained turbulence model perform in terms of robustness, generalizability and interpretability?

Methodology

RANS Turbulence Modeling

- Filling in $U = \bar{U} + U'$ in the NS equation (Reynolds' decomposition):
  - NS: $\frac{DU_i}{Dt} = -\frac{1}{\rho}\frac{\partial P}{\partial x_i} + \nu \frac{\partial^2 U_i}{\partial x_j^2}$
  - RANS: $\frac{D\bar{U}_i}{Dt} = -\frac{1}{\rho}\frac{\partial P}{\partial x_i} + \nu \frac{\partial^2 \bar{U}_i}{\partial x_j^2} - \frac{\partial}{\partial x_j}\left(\overline{U_i' U_j'}\right)$
- Reynolds stress $= \overline{U_i' U_j'}$
  - Requires modeling
  - Boussinesq: $= \overline{U_i' U_j'} = -2\nu_t S_{ij} + \frac{2}{3} k \delta_{ij}$

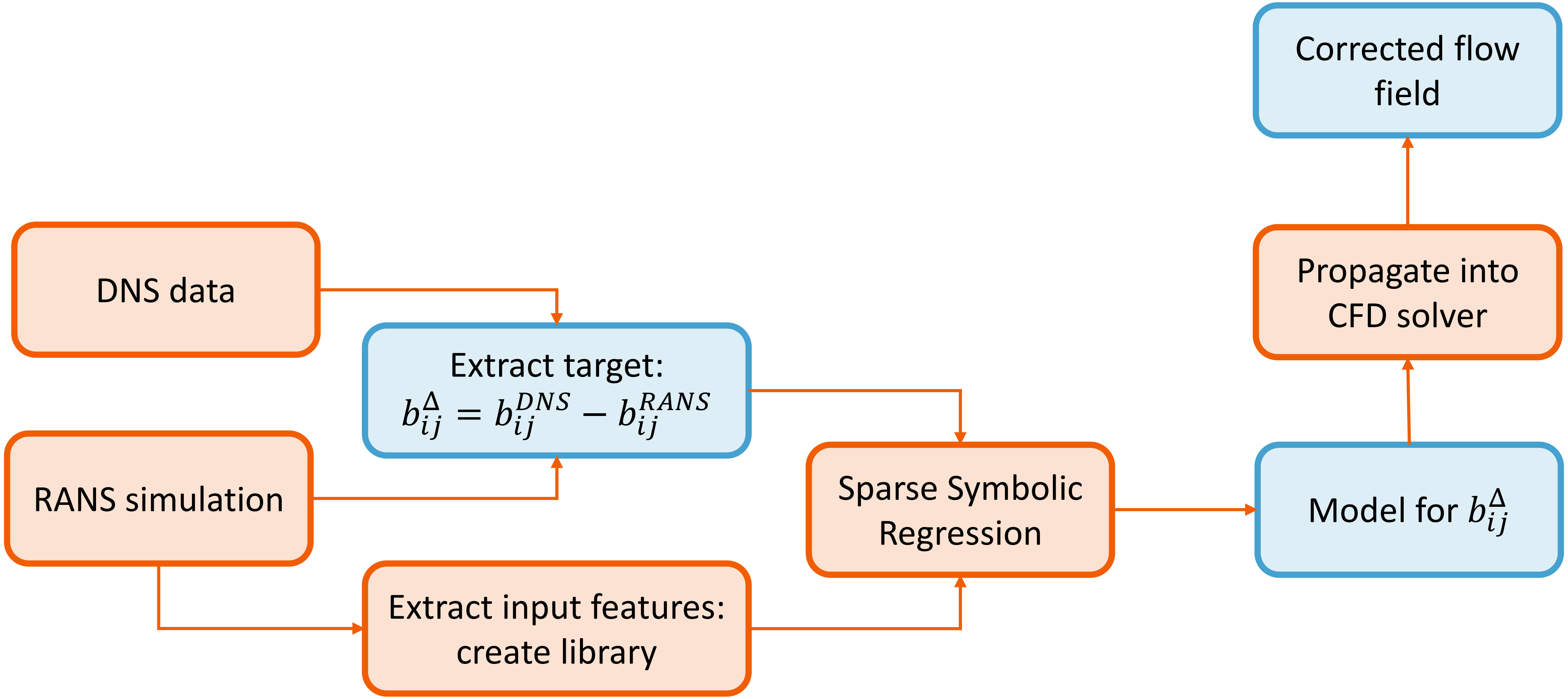
$$S_{ij} = \frac{1}{2}\left(\frac{\partial \bar{U}_i}{\partial x_j} + \frac{\partial \bar{U}_j}{\partial x_i}\right)$$

Subject of Modeling

$T_{ij}^{(1)} = S_{ij}$	$T_{ij}^{(6)} = R_{ik}R_{kl}S_{lj} + S_{ik}R_{kl}R_{lj} - \frac{2}{3}S_{pk}R_{kl}R_{lp}\delta_{ij}$
$T_{ij}^{(2)} = S_{ik}R_{kj} - R_{ik}Skj$	$T_{ij}^{(7)} = R_{ik}S_{kl}R_{lp}R_{pj} - R_{ik}R_{kl}S_{lp}R_{pj}$
$T_{ij}^{(3)} = S_{ik}S_{kj} - \frac{1}{3}S_{lk}S_{kl}\delta_{ij}$	$T_{ij}^{(8)} = S_{ik}R_{kl}S_{lp}S_{pj} - S_{ik}S_{kl}R_{lp}S_{pj}$
$T_{ij}^{(4)} = R_{ik}R_{kj} - \frac{1}{3}R_{lk}R_{kl}\delta_{ij}$	$T_{ij}^{(9)} = R_{ik}R_{kl}S_{lp}S_{pj} + S_{ik}S_{kl}R_{lp}R_{pj} - \frac{2}{3}S_{qk}S_{kl}R_{lp}R_{pq}$
$T_{ij}^{(5)} = R_{ik}S_{kl}S_{lj} - S_{ik}S_{kl}R_{lj}$	$T_{ij}^{(10)} = R_{ik}S_{kl}S_{lp}R_{pq}R_{qj} - R_{ik}R_{kl}S_{lp}S_{pq}R_{qj}$

- $-\overline{U_i'U_j'} = 2k\left(b_{ij} + \frac{1}{3}\delta_{ij}\right)$
- Here, b_{ij} is the anisotropic stress tensor:
  - Boussinesq: $b_{ij} = -\frac{\nu_t}{k}S_{ij}$
- Pope: $b_{ij}\left(\hat{S}_{ij}, \hat{R}_{ij}\right) = \sum_{n=1}^{10} G^{(n)}(\lambda_1, \ldots, \lambda_2)T_{ij}^{(n)}$
  - $S_{ij} = \frac{1}{2}\left(\frac{\partial \bar{U}_i}{\partial x_j} + \frac{\partial \bar{U}_j}{\partial x_i}\right)$
  - $R_{ij} = \frac{1}{2}\left(\frac{\partial \bar{U}_i}{\partial x_j} - \frac{\partial \bar{U}_j}{\partial x_i}\right)$
- Split into linear and nonlinear portions: $b_{ij} = \overbrace{-\frac{\nu_t}{k}\hat{S}_{ij}}^{\text{linear}} + \overbrace{b_{ij}^{\Delta}}^{\text{non-linear}}$

Modeling Overview



Library of Candidate Functions

- Select appropriate features
- Create more complex functions by multiplying the features with each other
- Multiply each function with the tensor bases $T_{ij}^{(n)}$

$$\begin{pmatrix} \lambda_1 & \lambda_2 & q_1 & q_2 & \cdots & q_n \end{pmatrix}$$

$$
\begin{aligned}
T_{ij}^{(1)} &= S_{ij} & T_{ij}^{(6)} &= R_{ik}R_{kl}S_{lj} + S_{ik}R_{kl}R_{lj} - \tfrac{2}{3}S_{pk}R_{kl}R_{lp}\delta_{ij} \\
T_{ij}^{(2)} &= S_{ik}R_{kj} - R_{ik}Skj & T_{ij}^{(7)} &= R_{ik}S_{kl}R_{lp}R_{pj} - R_{ik}R_{kl}S_{lp}R_{pj} \\
T_{ij}^{(3)} &= S_{ik}S_{kj} - \tfrac{1}{3}S_{lk}S_{kl}\delta_{ij} & T_{ij}^{(8)} &= S_{ik}R_{kl}S_{lp}S_{pj} - S_{ik}S_{kl}R_{lp}S_{pj} \\
T_{ij}^{(4)} &= R_{ik}R_{kj} - \tfrac{1}{3}R_{lk}R_{kl}\delta_{ij} & T_{ij}^{(9)} &= R_{ik}R_{kl}S_{lp}S_{pj} + S_{ik}S_{kl}R_{lp}R_{pj} - \tfrac{2}{3}S_{qk}S_{kl}R_{lp}R_{pq} \\
T_{ij}^{(5)} &= R_{ik}S_{kl}S_{lj} - S_{ik}S_{kl}R_{lj} & T_{ij}^{(10)} &= R_{ik}S_{kl}S_{lp}R_{pq}R_{qj} - R_{ik}R_{kl}S_{lp}S_{pq}R_{qj}
\end{aligned}
$$

Sparse Symbolic Regression

Two step modeling

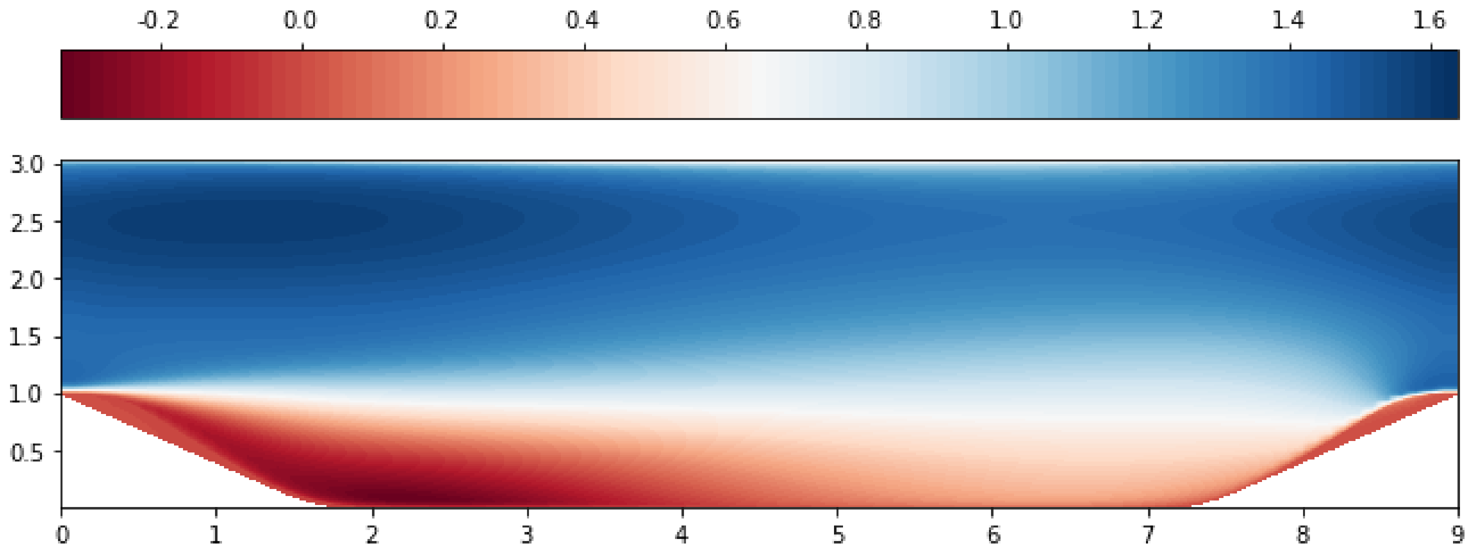
1. Model Discovery
   - Standardization of the features
   - Lasso Regression: $\boldsymbol{\xi} = \underset{\xi}{\operatorname{argmin}} \|U - \mathbb{C}\boldsymbol{\xi}\|_2^2 + \lambda \|\boldsymbol{\xi}\|_1$
   - Selecting the 'active' features
2. Model Inference
   - No standardization of the features
   - Only using the selected features
   - Ridge Regression: $\boldsymbol{\xi} = \underset{\xi}{\operatorname{argmin}} \|U - \mathbb{C}\boldsymbol{\xi}\|_2^2 + \lambda \|\boldsymbol{\xi}\|_2^2$
   - Calibration of the coefficient

Case Studies

- Channel Flow
- Periodic Hill
- Square Duct
- Taylor Couette flow
- Industrial Problem: Vortex Gripper

Some first results

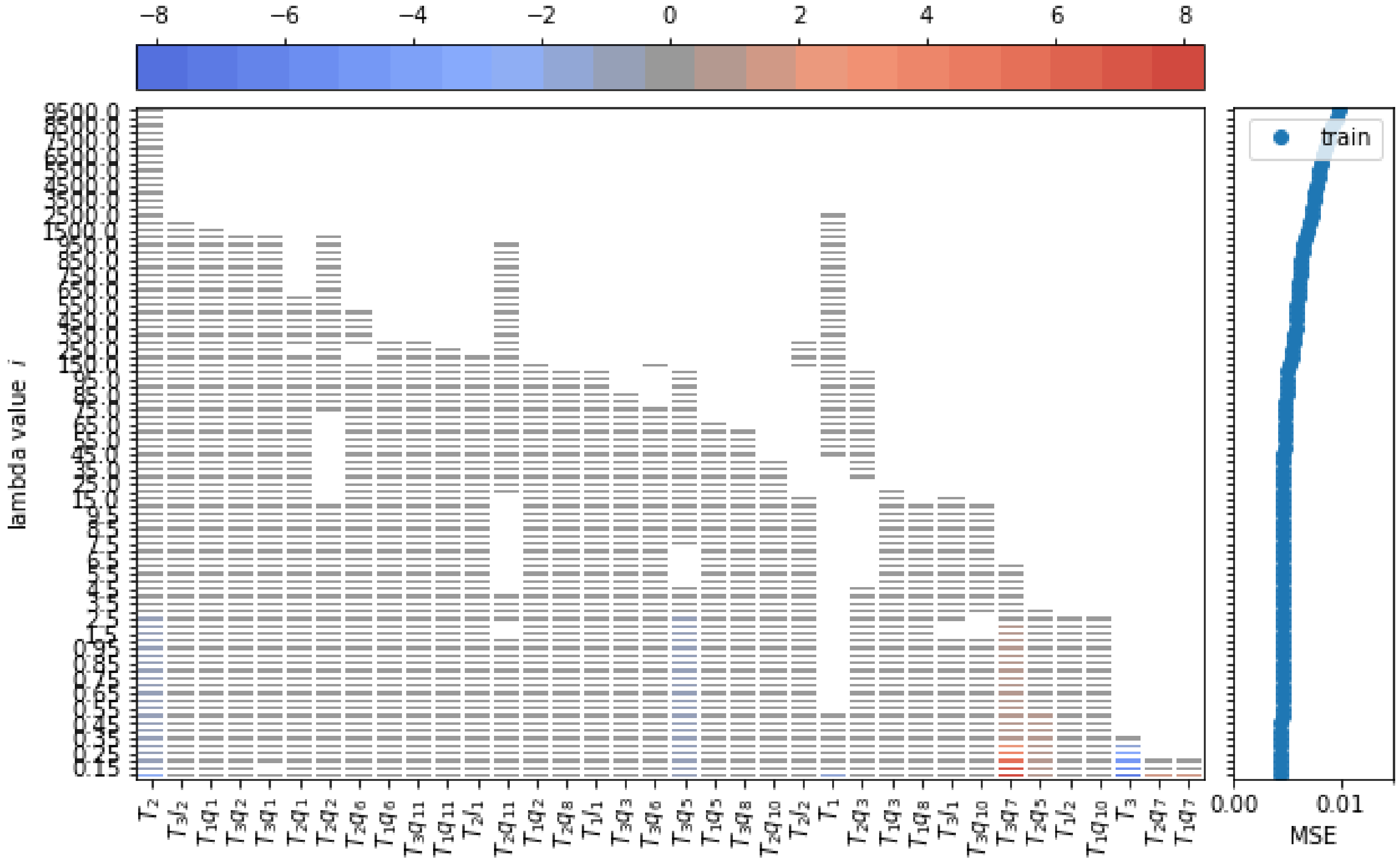
Periodic Hill



Model Discovery

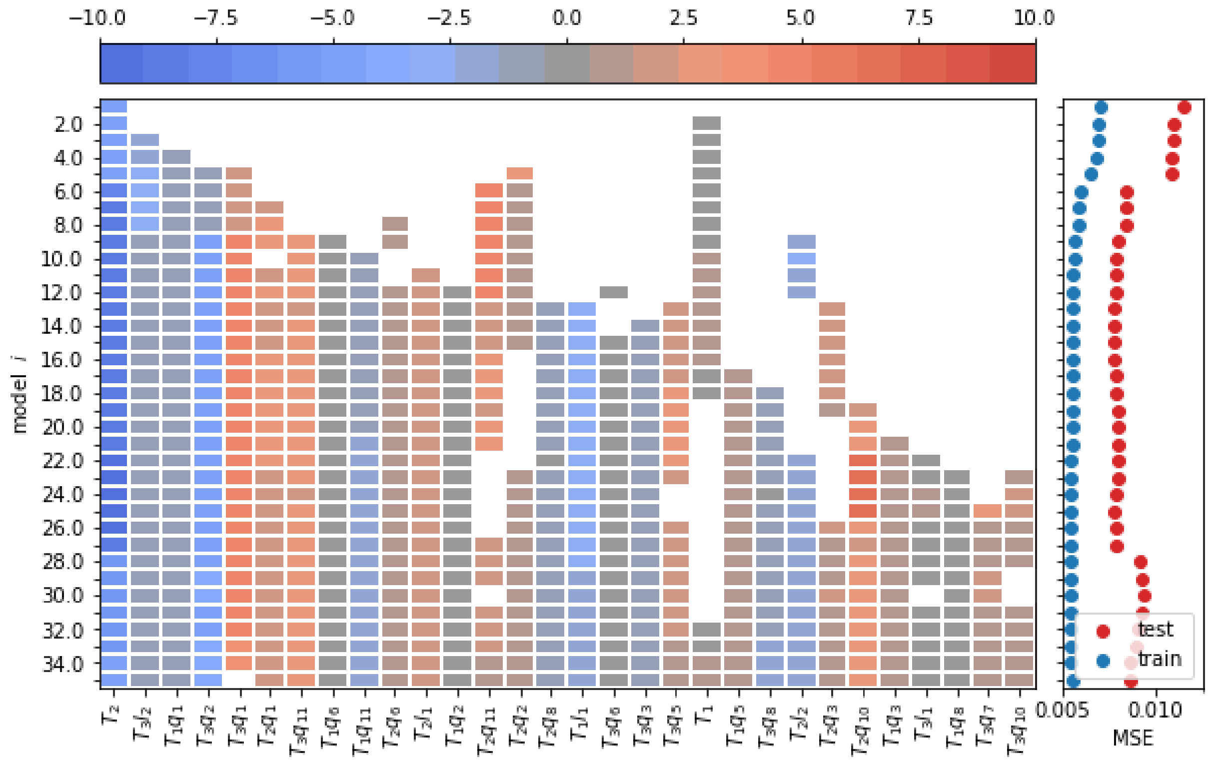
- Lasso Regression: $\xi = \min\|b^* - \mathbb{C}\xi\|_2^2 + \lambda\|\xi\|_1$

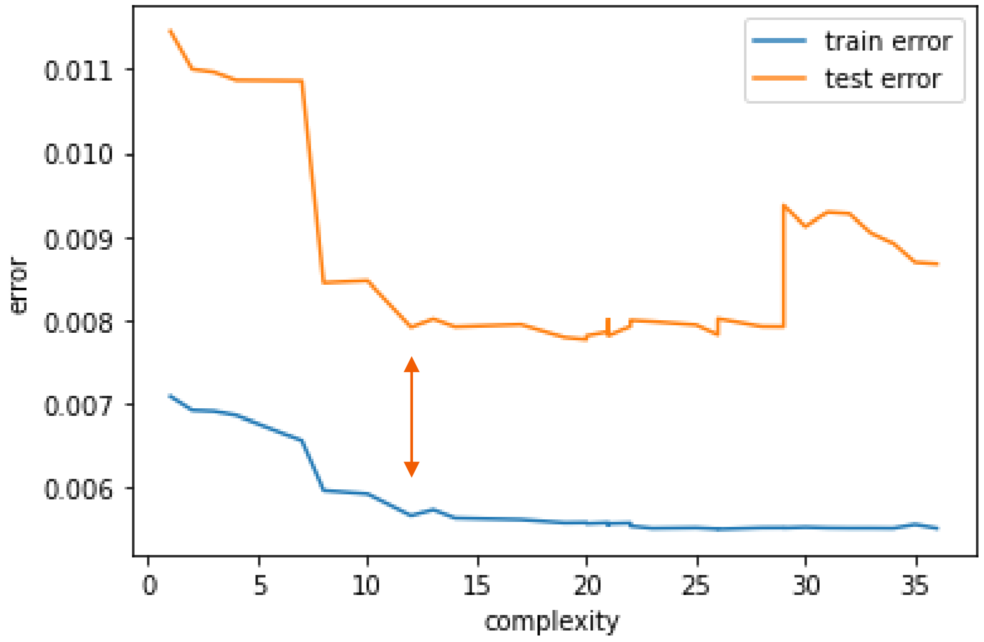
index	Features
q_1	$\frac{1}{2}(\|\|R\|\|^2 - \|\|S\|\|^2)$
q_2	k
q_3	$\min(\frac{\sqrt{k}d}{50\nu}, 2)$
q_4	$\overline{u}_i \frac{\partial p}{\partial x_k}$
q_5	k/ϵ
q_6	$\overline{u}_i \frac{\partial k}{\partial x_i}$
q_7	$\|\|\overline{u_i' u_j'}\|\|$
q_8	$\sqrt{\frac{\partial p}{\partial x_i}\frac{\partial p}{\partial x_i}}$
q_9	$\|\overline{u}_i \overline{u}_j \frac{\partial \overline{u}_i}{\partial x_j}\|$
q_{10}	f_1
q_{11}	f_1^2
l_1	trace(S^2)
l_2	trace(Ω^2)



Model Inference

- Ridge regression: $\boldsymbol{\xi} = \min \|b^* - \mathbb{C}\boldsymbol{\xi}\|_2^2 + \lambda\|\boldsymbol{\xi}\|_2^2$,
  - $\lambda = 0.1$
- Selecting models:
  - Trade-off between complexity and accuracy



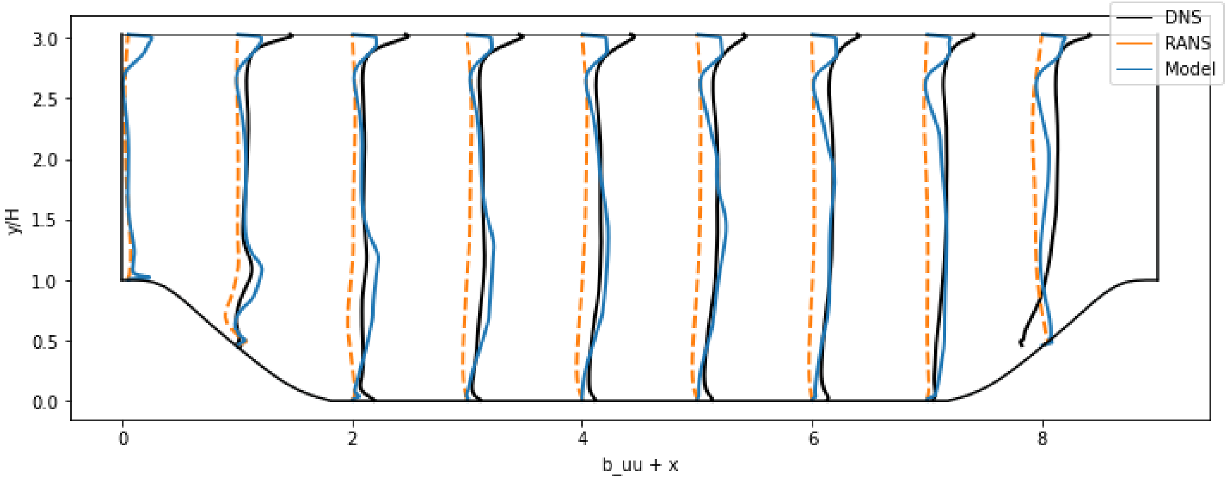


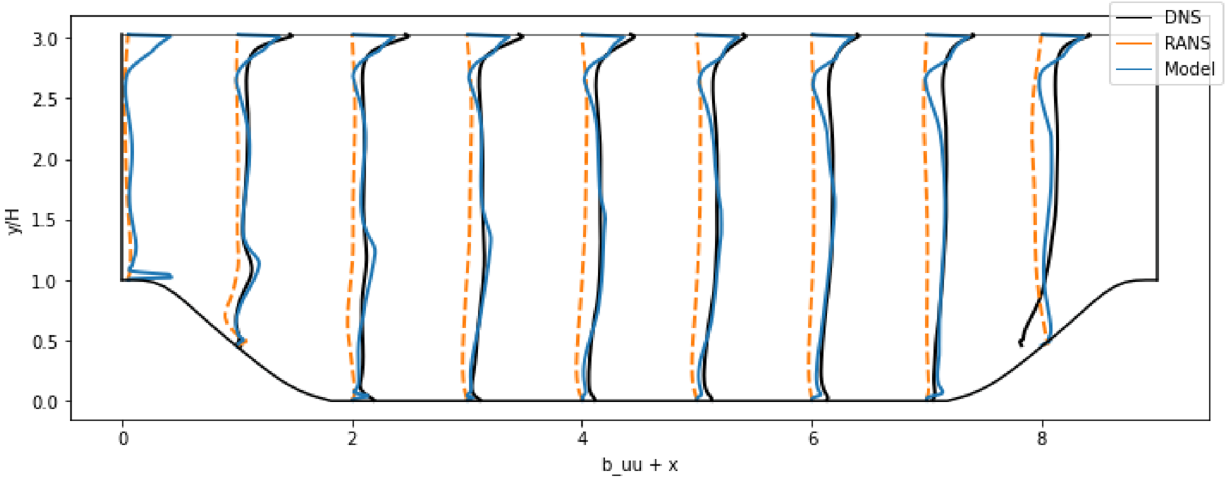
Result of selected models

Anisotropic stress: $b_{ij} = \frac{\nu_t}{k} S_{ij} + b^*_{ij}$

- Simple: $b^*_{ij} = -4.3 * T_2$

More complex: $b^*_{ij} = (1.72q_1 - 0.71q_{11} + 0.08q_6 + 0.80)T_1 + (-2.55l_2 + 4.71q_{11} + 0.72q_2 - 8.20)T_2 + (-1.24l_2 + 4.51q_1 + 3.23q_{11} - 4.56q_2)T_3$



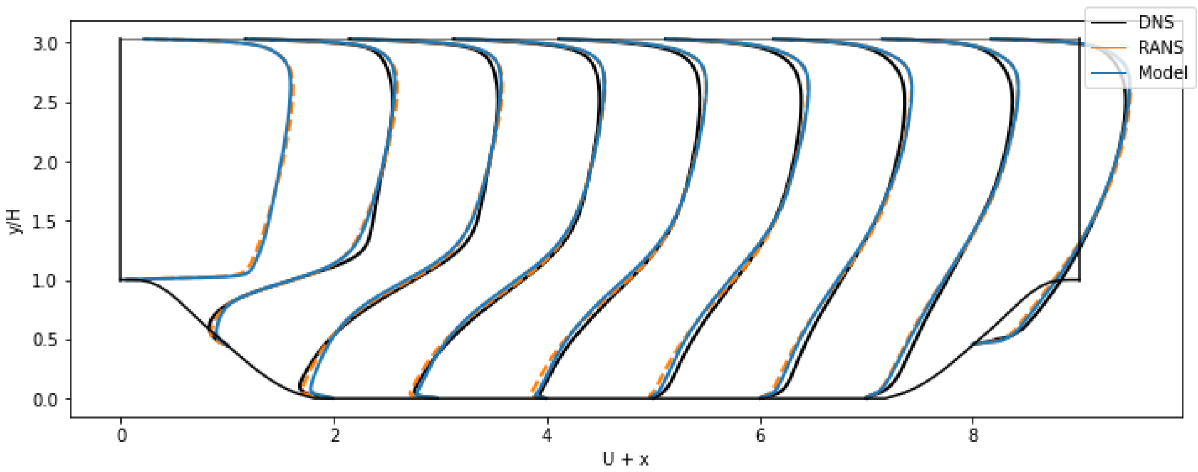


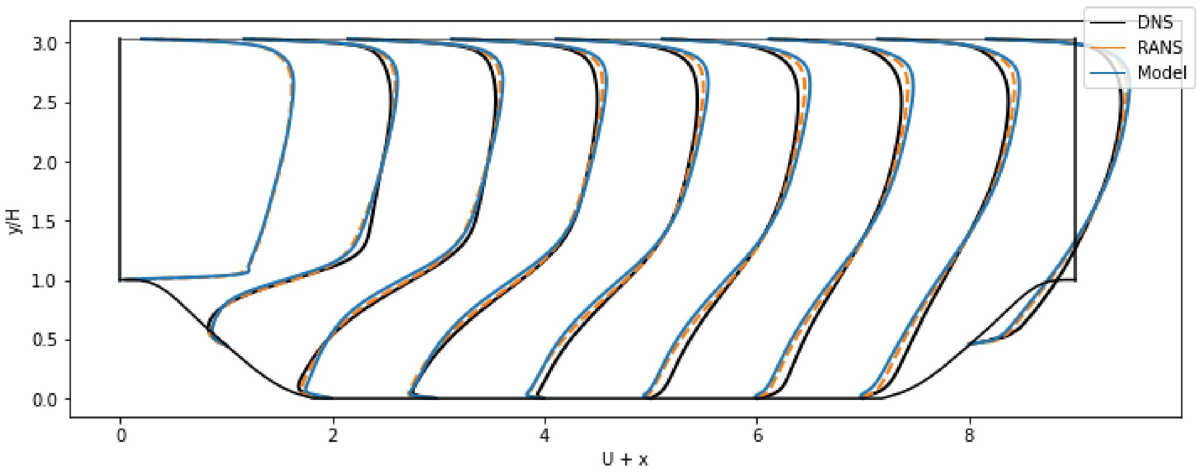
Model Propagation

Streamwise velocity profiles

- Simple: $b_{ij}^* = -4.3 * T_2$

More complex: $b_{ij}^* = (1.72q_1 - 0.71q_{11} + 0.08q_6 + 0.80)T_1 + (-2.55l_2 + 4.71q_{11} + 0.72q_2 - 8.20)T_2 + (-1.24l_2 + 4.51q_1 + 3.23q_{11} - 4.56q_2)T_3$





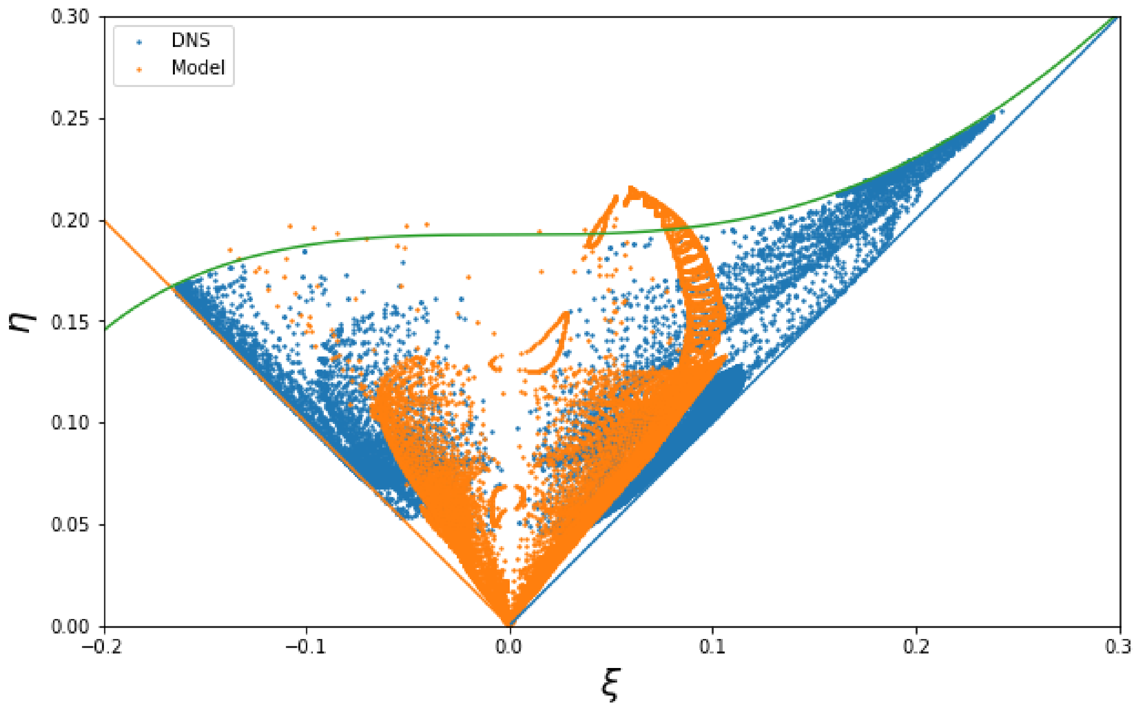
Constraints

For realizable flows: $b_{ii} \in \left[-\frac{1}{3}, \frac{2}{3}\right]$ and $b_{ij} \in \left[-\frac{1}{2}, \frac{1}{2}\right]$

Creating a Lumley triangle using: $\zeta^3 = b_{ij}b_{in}b_{jn}/2$ and $\eta^2 = -\frac{b_{ij}b_{ji}}{2}$

Values of ζ and η should be inside the triangle

Adding constraints in regression function could enforce the realizability



Intermediate Conclusions

Intermediate conclusions

- Frame-work to discover models with sparse symbolic regression
- Possible to add constraints
- No improvement of the velocity profiles yet

What's left?

To do:

- Finding a good model that improves the velocity
- Research different regression functions
- Research the effect of adding constraints to the regression functions
- Perform the other test cases

To be continued…

Questions?