Applied Machine Learning to Estimate Length of Separation and Reattachment Flows as Parameter Active Flow Control in Backward Facing Step

Ahmad Fakhri Giyats, Mohamad Yamin[*], Cokorda Prapti Mahandari
Department of Mechanical Engineering, Faculty of Industrial Technology,
Gunadarma University, Depok 16424, INDONESIA
[*]mohay@staff.gunadarma.ac.id

ABSTRACT

Recently, large amounts of data from experimental measurements and simulations with high fidelity have extensively accelerated fluid mechanics advancement. Machine learning (ML) offers a wealth of techniques to extract data that can be translated into knowledge about the underlying fluid mechanics. Backward-Facing Step (BFS) is well-known for its application to fluid mechanics, particularly flow turbulence. Typically, a numerical approach can be used to understand the flow phenomena on BFS. In some instances, numerical investigations have a computational time limitation. This paper examines the application of ML to predict reattachment length on BFS flow. The procedure begins with a simulated meshing sensitivity of 1.27 cm in step height. This numerical analysis was conducted in the turbulent zone with a Reynolds number between 35587 and 40422. OpenFOAM® was used to perform numerical simulations using the turbulence model of k-omega shear stress transport. ML employed information in the form of Velocity and Pressure at every node to represent the type of turbulence. Using Recurrent Neural Networks (RNNs) as the most effective model to predict reattachment length values, the reattachment length was predicted with a Root Mean Square Error of 0.013.

Keywords: *Data-Driven; Machine-Learning; Fluids Mechanics; Backward-Facing Step*

ISSN 1823-5514, eISSN 2550-164X

https://doi.org/10.24191/jmeche.v20i3.23904

Received for review: 2022-10-13
Accepted for publication: 2023-05-09
Published: 2023-09-15

Introduction

In recent years, fluid mechanics has been a field full of data and complex problems. Many conventional research data, flow field observations, and extensive numerical simulations are accessible, accompanied by the development of high computing. The advantages of developing high computing-based programming architecture, current experimental capabilities, measurement methodologies, and "Big Data" have become indispensable to the advancement of fluid mechanics. In the meantime, a suitable and efficient method for processing enormous data quantities, such as the "cluster database," has been developed by Perlman et al. [1] and Giovanni [2] for data analysis and disclosure, which can be carried out.

The rapid development of data has spread across various disciplines, so obtaining potential information quickly and accurately is the focus of research in this decade. Considering some hardware architecture improvements, storage, more efficient data transmission, rapid algorithm development, and development of open source-based frameworks, as well as data-driven research methods, have received much attention from academics and commercial opportunities. Deep Learning (DL), one with a neural network technique, has distinct benefits when confronted with nonlinear high dimensional problems, which are rapidly integrated into fluid mechanic research.

Machine learning classes are supervised, semi-supervised, and unsupervised learning [3]. With many advantages, machine learning has been gradually applied to reduced-order models, prediction, reconstruction, closed turbulence models, and active flow control systems in fluid mechanics data analysis [3]-[4]. Rowley et al. [5] used Proper Orthogonal Decomposition (POD) to examine near-wall flow features at varying Reynolds numbers in a turbulent channel flow.

The integration of machine learning with fluid dynamics has a history. Teo et al. [6] developed a neural network (neural network) to create particles in the photo to measure the velocity. The same is done by adding how many neural layers (multi-layer). The development continues, and the application of neural networks for adaptive controllers is carried out to reduce turbulence barriers [7]. This study describes a simple control network that employs suction and blowing based on the shear stress of the wall in the span direction to reduce up to 20 percent less wall friction. At this period, neural network applications are still difficult to comprehend, and neural network development is still in its infancy.

The machine learning algorithm will simulate some fluid features, such as the lift profile given a specific airfoil geometry, and provide a proxy that can be optimized. Utilizing machine learning (ML) to tackle the fluid optimization problem directly is also possible by developing a machine learning model to influence the fluid's behaviour towards some engineering

aim via active control. Besides, many successful advances in using DL techniques to accelerate topology optimization [8].

In 2021, Usman et al. [9] were concerned with Fluid-Structure Interaction (FSI) in machine component design. Computational Fluid Dynamics (CFD) accuracy strongly comes from mesh size; hence, the computational cost is proportional to the resolution of tiny features. Multiple physics and scales exponentially increase the computing complexity, extending the process's duration. ML has demonstrated a highly promising ability to predict solutions for differential equations. It has provided excellent approximations in a fraction of the time required by standard simulation techniques [9].

In 2022, Vinuesa et al. [10] had a perspective on ML. They highlighted some areas of the highest potential impact, including accelerating direct numerical simulations, improving turbulence closure modelling, and developing enhanced reduced-order models.

Based on the literature study above, DL is still being developed to estimate reattachment lengths in flow separation [11] with installability in the separation flow [12]. Reattachment length refers to the distance the separated flow returns to the surface. The reattachment length can be calculated using numerical models such as Unsteady Reynolds-Averaged Navier-Stokes (URANS) or Large Eddy Simulation (LES). In fluid flow analysis, reattachment length is commonly employed to determine how the flow separates and returns to the surface [11].

Determining the reattachment length is difficult to estimate, so this research's estimation is intended to validate changes due to changing parameters (pressure, velocity, *Cf*). A robust system has been developed in this case to determine the reattachment length accurately. Where the reattachment length is one of the sequences to minimize bubble separation, if it is reduced, it can be avoided to a minimum (drag reduction for external flow, pressure drop for internal flow) with the aim that it can be applied to active flow control or fluid power systems.

The following is the structure of this paper: The section "Methodology" demonstrates our proposed technique and the experimental details. The following section provides the results and discussion. The last section summarizes our research.

Methodology

Selection turbulence models

Consider the flow of fluid across a thin plate from the horizontal direction. When a fluid constructed velocity to a plate's leading edge, the layer of laminar boundary form begins. The flow is highly anticipated in the locality. A small distance from the leading edge, the transition zone evolves into a fully-fledged

turbulent region, as depicted schematically in Figure 1(a) [13]. The Reynolds number characterizes the flow transition between the three regions over the flat plate. The velocity and pressure fields can be predicted using the Navier-Stokes equations in a steady-state laminar regime. The flow is assumed to be steady and uniform to predict the flow behaviour accurately. The Reynolds number remains constant in the laminar regime, so no averaging is required. Figure 1(b) depicts the fundamental physical and geometric models of the Backward-Facing Step (BFS) (under the 2-D scheme). The BFS methodology characterized a uniform velocity inflow, which may be either turbulent or laminar, originating from a channel of height (H). A step of height (*h*) is also present on the lower or upper sides. As depicted in Figure 1(b), The BFS flow field consists of several regions: the separated shear layer, the recirculation shear layer, the attached or recovery region, and the reattachment region.

Figure 1(c) depicts a fundamental model encompassing the three essential characteristics of a separated flow: reattachment length, vortex evolution, and free shear flow separation.

Solving the equations of Navier-Stokes is challenging because the Reynolds number fluctuates with time and space, leading to minor flow eddies and oscillations within an insufficiently short period. Under such conditions, the Reynolds Averaged Navier Stokes formulation is preferable.

The turbulent flow close to a flat wall can be divided into four categories: "buffer layer," "laminar or viscous sublayer," "free stream," and "transition zone." Eventually, the flow turns completely turbulent at the transition zone, and the mean fluid velocity is related to the distance from a flat plate or stationary wall. Consequently, this region is sometimes referred to as the log law region. The laminar sublayer is too thin in the turbulence model. Hence, it is beneficial to approximate this region.

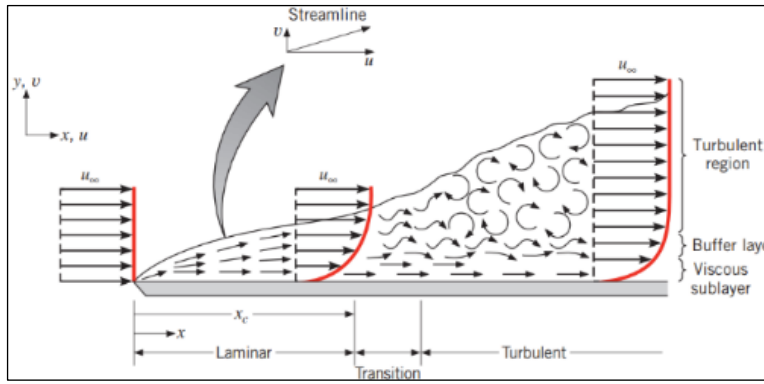
A more detailed review of problems and successes of turbulent computing flow and the appropriate sources of turbulence modelling was explained by Argyropoulos et al. [16].

RANS turbulence models

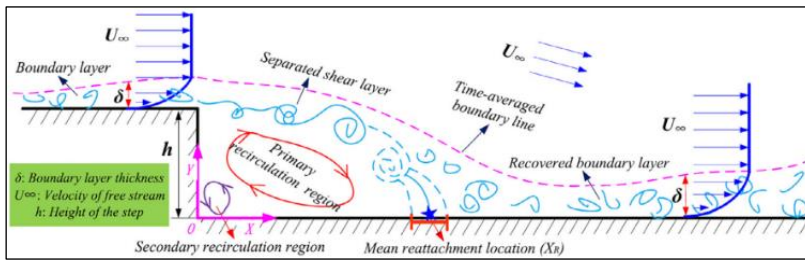
First, eddy viscosity with zero equations is the simplest turbulence model. The first model is zero-equation, embracing the mixing-length concept developed by Prandtl (the equivalent of a gas's mean free paths) hypothesis [17].

Zero-equation models are insufficient for simulating every flow type because they disregard diverse physical characteristics, such as non-local effects on turbulent eddy viscosity. In addition, flow history is not considered to overcome these deficiencies.

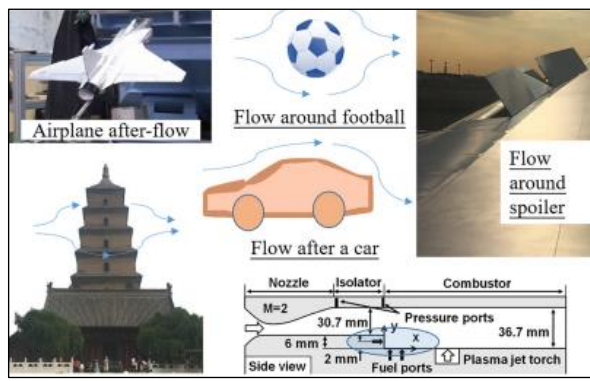
Second, similar to the turbulent model with zero-equation, a length scale must be specified for turbulence in this model. Results indicate that the One-equation eddy-viscosity model is unsuitable for indoor environments due to its inability to replicate turbulent flow at corners in the presence of flow barriers. This model demonstrates promising findings for zero flow separation.



(a)



(b)



(c)

Figure 1: Generalization theory of BFS and their applications; (a) fluid flow over thin horizontal plate [13], (b) schematic BFS flow evolutions [14], and (c) various BFS model applications [15]

Third, two equations of the eddy-viscosity answer the kinetic energy dissipation rate and kinetic energy from the Reynolds transport equation. This model includes a partial differential transport equation and the expression for turbulent kinetic energy 'k' K-epsilon model (*k-ε*). Launder and Spalding's [18] one of the well-known models for modelling indoor and outdoor airflow is the *k-ε* model [2], [18].

The RNG k-epsilon (RNG *k-Ɛ*) employs statistical techniques to eliminate motion characteristics with low scales methodically. The RNG k–epsilon is obtained using a statistical technique. The fluid flow regulating equation is adjusted concerning significant motion elements. The RNG k-model cannot predict the effect of corners due to restrictions on using a coarse grid near walls and corners [19]-[20].

In the k-omega model (*k-ω*), Omega-ω is the ratio of (dissipation rate of energy from turbulent kinetic) over k (turbulent flow's kinetic energy). Omega-ω is the conversion rate of turbulent kinetic energy (*k*) to internal energy. Omega-ω represents the turbulence scale. Comparing the *k-ε* model to the *k-ω* model, The *k- ω* model is more accurate in predicting places with unfavourable pressure flow conditions [21]. Several *k-ω* models have been developed to address specific flow problems that the traditional model cannot adequately evaluate. These models include the shear stress transport (SST) *k-ω* models and customized *k-ω* models. The (SST) k-ω model is commonly used when accurate flow predictions near wall boundaries are required, owing to its high efficacy. The (SST) *k-ω* is similar to the *k-ε* model and yields virtually identical results for various flow circumstances.

Large Eddy Simulation (LES) and Detached Eddy Simulation (DES)

Higher computer capacity and user skills are required for LES. LES uses sub-grid scale eddies filtering to solve the numerical simulation equation for large eddies. Comparing LES, *K-ω*, and RNG *k-ω* for air circulation analysis in a room environment, the study conducted by Tian and colleagues provides evidence that utilizing all three turbulence models results in a precise prediction of the experimental configuration [14]. In addition, he concluded that LES provides more accurate results than RANS models and that the results are close to the actual conditions. The LES technique can validate the K-ε model, given its capacity to address complex flow phenomena and consider the wall function.

Some researchers use the DES model for complex enclosed airflow analysis, but it does not play a significant role in cold storage airflow analysis. RANS models are not good at predicting massive separation in free shear flows, whereas the DES model is suitable for very high flow instability.

The performance and cost of DES are situated between those of the LES and RANS models. In BFS analysis, achieving steady-state conditions within the chamber is crucial. Once the flow has been properly established, there are

minimal perturbations. Utilizing LES or DES for such needs is an expensive and time-consuming endeavour.

Recurrent Neural Networks (RNNs)

RNNs are neural network data sequences where each value depends on previous values. This RNN is a feed-forward network with feedback loops [22]. RNNs are superior at simulating temporal dynamic behaviour to traditional feed-forward neural networks because they bring the concept of time to them [23]. Some RNNs units maintain a previous time step's internal memory state, denoting a context window of indefinite size. Numerous RNNs applications have been proposed and researched [24]. Basic RNNs and long short-term memory are the two most common units described in the following section.

Basic recurrent neural networks

As shown in Figure 2(a), The term of input time series with T the model of the sample is $\{x_t\}_{t=1}^{T}$. And the term of the output of models containing T samples of the specified time series is $\{s_t\}_{t=1}^{T}$. At time t, the input of models (x_t) and it produces the result (prediction), S_t. The following equations define a fundamental RNNs unit:

$$S_t = tanh\,(x_t\,u + S_{t-1}\omega + b) \tag{1}$$

The tangent function of hyperbolic, $\tanh(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$. Moreover, the model's parameters u, ω, and b are given. In addition to the current input x_t, the model at time t also receives its output from the prior period (s_{t-1}). The hyperbolic tangent activation function's argument is the linear combination of $\mathrm{x}_t u + S_{t-1}\omega + b$, which enables the unit to simulate nonlinear input-output relationships. Additional activation functions, such as logistic functions, rectified linear units (ReLU), or sigmoid functions, may be used in various implementations [25].

Long short-term memory networks

RNNs are known to exhibit the issue of limited "short-term memory": historical data is utilized to generate forecasts if a sequence is of adequate length. The inability to effectively transmit vital information from preceding eras to subsequent ones, such as significant patterns from the same month in previous years. LSTM is a neural network that addresses short-term memory problems by utilizing gates to preserve and combine significant long-term memory with the latest input [26]. LSTM paved the path for substantial advancements in various domains, including speech recognition and natural language processing [27].

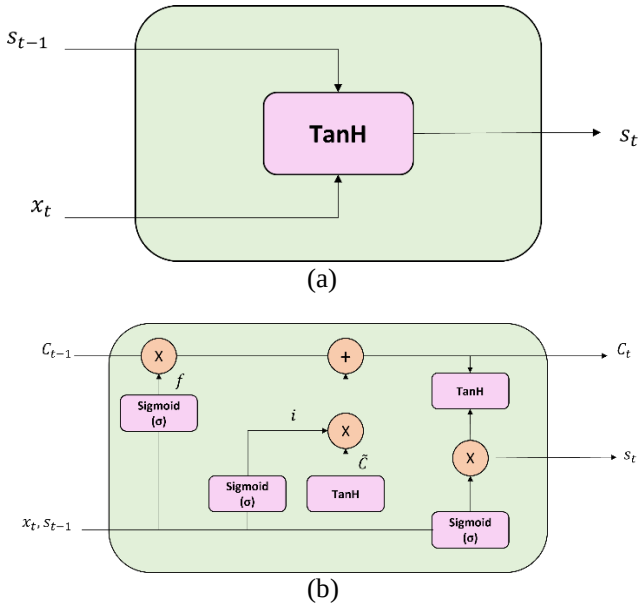


Figure 2: Illustration of recurrent neural networks; (a) illustration of the basic-RNNs unit, and (b) illustration of LSTM unit

Following Figure 2(b), it can be observed that each line facilitates the transmission of an entire vector value from the output of a given node to serve as the input for subsequent nodes. The circular symbols in orange signify point-wise operations, while the rectangular shapes in pink represent the layers of a trained neural network. When lines merge, they become one, whereas when lines fork, their material is copied and sent to separate locations.

The LSTM unit can selectively "remember" or "forget" information through the precise control of three gates, namely the input, forget, and output, which a specific memory cell state facilitates. The gates control the flow of data into and out of the state of memory cells. The following equations define an LSTM unit:

$$i = \sigma\left(x_t u^i + S_{t-1}\omega^i + b^i\right) \tag{2}$$

$$f = \sigma\left(x_t u^f + S_{t-1}\omega^f + b^f\right) \tag{3}$$

$$o = \sigma\left(x_t u^o + S_{t-1}\omega^o + b^o\right) \tag{4}$$

$$\tilde{c} = tanh\left(x_t u^c + S_{t-1}\omega^c + b^c\right) \tag{5}$$

$$c_t = f \times c_{t-1} + i \times \tilde{c} \tag{6}$$

$$s_t = o \times tanh(c_t) \tag{7}$$

The logistic activation function or sigmoid is $\sigma(x) = \frac{1}{1+e^{-x}}$. The parameters learned to regulate the input gate i are u^i, ω^i, and b^i. The parameters learned that govern the forget gate f are u^f, ω^f, and b^f. The parameters learned to control the output gate o are u^o, ω^o, and b^o , and c is the newly identified candidate activation for the condition of the cell u^c, ω^c, and b^c. The cell state c_t is utilizing a linear combination update $c_t = f \times c_{t-1} + i \times \tilde{c}$, where the previous cell state value is c_{t-1}. The input gate i identifies the aspects of the candidate c will be utilized to alter the status of a memory cell, while the forget gate f decides which elements of the previous memory are retrieved (c_{t-1}) will be deleted. The output gate o then determines which portions of the newly updated cell state (c_{t-1}) will be displayed in the output s_t.

Evaluation metrics

Under Aparicio et al. [28], we publish our findings on evaluation metrics with Root Mean Squared Error (RMSE). The RMSE is calculated using:

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(x_t - \hat{x}_t)^2} \tag{8}$$

where x_t is the Reynolds change rate for t, and $\hat{x}_t$ is the corresponding prediction.

Model geometry

This study uses numerical simulation to develop and solve the model, meshing, numerical equation, and boundary condition setup. In addition, this work utilizes Python with the RNNs Method to support the neural networks. The main method in this study is divided into three major groups: training cases, machine learning models, and test cases. This study's flow methodology is explained in Figure 3.

In the training case, it is the stage for generating datasets that will be used in the ML model stage; at this stage, the resulting dataset is in the form of fluid characteristics at each coordinate spread over the slicing plane and image capture where each data is obtained at each speed interval. The ML model is the next stage of research, namely creating a model that can estimate the reattachment distance. In constructing this model, the LSTM architecture is used.

In the final stage, the Test case is used to prove and evaluate the model that has been made.

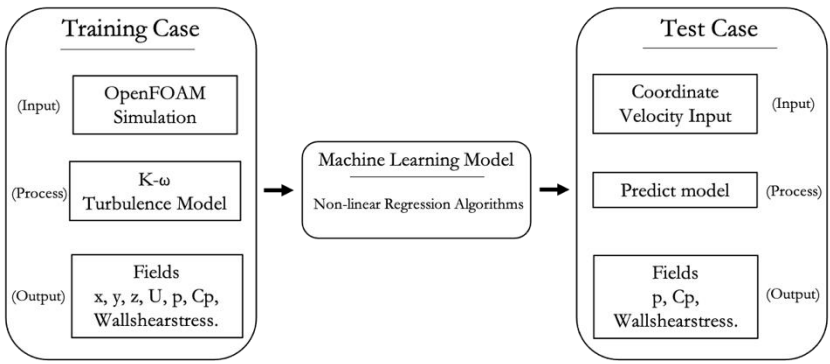


Figure 3: The current research framework

Before starting the numerical simulations with various configurations of multiple components and levels, the CFD simulation program must be validated as suitable software for subsequent step activities. The validation technique aims to validate that the computational fluid dynamics with several parameter settings accurately depict the actual conditions and are consistent with experimental findings. In addition, after completing the validation procedure, the following step is to simulate a model. The estimation procedure is concluded by utilizing RNNs and analysing the evaluation metrics.

BFS is one of the essential separation-flow models for theoretical and technical advancement. Airfoils at extreme attack angles, spoiler flows, inlet tunnel flow of an engine or inside a condenser/combustor, flow separation behind a vehicle, and flow around a boat or a complex building are examples of daily applications for backward-facing step flow [15].

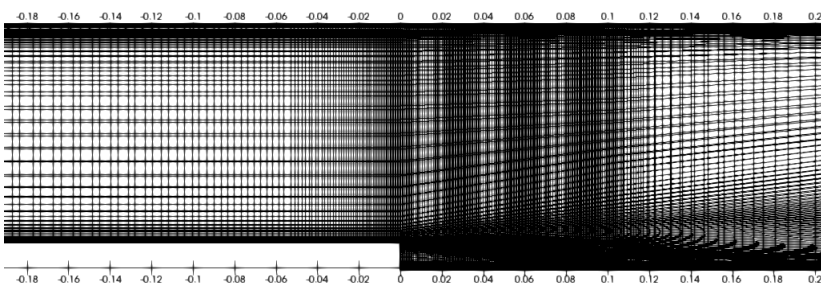


Figure 4: Cross-section of BFS fine mesh

BFS flows known as "backward flow," "sudden expansion flows," "back-step flow," "circular expanding flow," or "diverging channel" involve the fundamental characteristics of general separation flows. It is ideal for two-dimensional topics discussed under certain flow conditions or with simple step geometries and has a three-dimensional nature. It is essential for separation flows over a wide Reynolds number range [15].

From a flow dynamics perspective, the BFS flow is characterized by enormous separation vortices in the backward step zone and a few minor vortices at the corner. Occasionally, complex vortex series can occur under certain physical and geometric conditions. Many researchers have compared predictions with validated turbulent models against existing data. BFS in turbulent flow at boundary walls is a common occurrence. The investigation of flow structure in the BFS has been rigorously examined and investigated in the literature, shown in Table 1.

Table 1: Recent study representative numerical studies of BFS

Title	*Re*	Method	*Xi/h*	Comment
McQueen et al. [29]	5000 - 6700	OpenFOAM	5 – 6.7	Incompressible
Sazhin et al. [30]	10, 389, 648	Direct Monte Carlo	0.7-2	Flowrate
Talib et al. [31]	5000 - 20000	CFD (fluent)	2-3	Heat transfer
Loksupapaiboon et al. [32]	15500	OpenFOAM	2-5	Incompressible

A model of the BFS is created using blockMesh OpenFOAM. Figure 4 shows the cross-section of the BFS. Additionally, the air is used as a working fluid in this task due to its simplicity and ease of operation. The model adapts the 2D Backward-Facing Step research, Langley Research Centre Turbulence Modelling Resource [33].

A mesh independence check was performed to make the solution more independent. In order to achieve the convergence criteria, the residual error was set to 10E-4, and the mesh must be refined globally to obtain more acceptable mesh cells [34]. The mesh must be refined until the error is reduced. At 767,500 cells, mesh independence is achieved, as shown in Table 2.

Table 2: Number of cells in each case and mesh-dependent analysis *Re*=36000

Case	Coarse	Medium	Fine	Rumsey [33]
Number of cells	22,165	363,090	767,500	Experimental
Reattachment length (mm)	0.0835	0.0797	0.0786	0.0790

Coarse, medium, and fine are the three types of two-dimensional meshing. As depicted, three two-dimensional scenarios were executed to undertake a grid independence analysis. The cell number in each scenario is displayed in Table 2.

Table 3: Number of cells in each case and mesh-dependent analysis
Re=36000

Indicator	Value
Points	942312
Internal Points	599800
Faces	2473755
Internal faces	2131245
Cells	767500

The distance from the "Step" to the point at which the direction of the Wallshear value receives a change called the reattachment length. Table 3 show compares the experimental results and the recorded reattachment length values.

Numerical setup

It is intended to obtain flow characteristics at this stage as it passes the BFS. Numerical computing is performed with a custom-built open-source CFD package. OpenFOAM® "blockMesh" creates geometries along the *x*, *y*, and *z* axes. As shown in Figure 4, the model is given a unit width for the two-dimensional condition [34].

For the meshing arrangement, the geometry is divided into six blocks with edge names: inlet, outlet, front and back, top wall, and bottom wall. Meshing is performed with a greater concentration of cells in the centre and along the walls. This method captures the turbulent flow in this region more accurately. A grid independence test was carried out [34]. In Table 3, the number of cells was shown in each case. The mesh arrangement for numerical computation is carried out on an acceptable mesh type with a 767500-cell number in the two-dimensional case.

Based on the selection in the section "RANS Turbulence Models," it was considered that k-omega turbulence is a good model for understanding the flow separation process, where the turbulence model will be used in this study. The Reynolds Averaged Navier Stokes equation, represented by equation [35], determines the flow in the Backward-Facing Step.

$$\bar{u}_j \frac{\partial \bar{u}_i}{x_j} - \frac{\partial}{\partial x_j}\left[v_{eff}\left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i}\right)\right] = -\frac{\partial p}{\partial x_i} \quad (9)$$

In this numerical computation, several equations are used in iterations. *K-ω* is built from the two-equation model for the turbulence kinetic energy (*k*) and turbulence-specific dissipation rate (*ω*) base model. The equation for the turbulence-specific dissipation rate is [35]:

$$\frac{\partial \epsilon}{\partial t} + \bar{u}_j \frac{\partial \epsilon}{\partial x_j} - \frac{\partial}{\partial x_j}\left[\left(v + \frac{v_T}{\sigma\epsilon}\right)\frac{\partial \epsilon}{\partial x_j}\right] = C_1 \frac{\epsilon}{k} v_T \frac{\partial \bar{u}i}{\partial x_j}\left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i}\right) - C_2 \frac{\epsilon^2}{k} \quad (10)$$

And the kinetic energy of turbulence:

$$\frac{\partial k}{\partial t} + \bar{u}_j \frac{\partial k}{\partial x_j} - \frac{\partial}{\partial x_j}\left[\left(v_{eff}\right)\frac{\partial k}{\partial x_j}\right] = v_T \frac{\partial \bar{u}i}{\partial x_j}\left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i}\right) - \epsilon \quad (11)$$

By defining the specific dissipation $\omega = \frac{\epsilon}{k}$ as the second transported variable, we have the *k-ω* model. The equation used for *k* is the same implemented for the *k-ϵ* model, while the equation for *ω* becomes:

$$\frac{\partial \omega}{\partial t} + \bar{u}_j \frac{\partial \omega}{\partial x_j} - \frac{\partial}{\partial x_j}\left[\left(v + \alpha_\omega v_T\right)\frac{\partial \omega}{\partial x_j}\right] = \alpha \frac{\omega}{k} v_T \frac{\partial \bar{u}i}{\partial x_j}\left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i}\right) - \beta\omega^2 \quad (12)$$

The turbulence viscosity SST is obtained using [36]:

$$v_t = \frac{\alpha_1 k}{max(\alpha_1 \omega, b_1 F_{23})} \quad (13)$$

In isotropic turbulence, the kinetic energy of turbulence can be approximated by [36]:

$$k = \frac{3}{2}\left(I|u_{ref}|\right)^2 \quad (14)$$

where I represents intensity, and u_{ref} reference velocity. The turbulence-specific dissipation rate (ω) can be determined as [36]:

$$\omega = \frac{k^{0.5}}{C_\mu^{0.25} L} \quad (15)$$

where $C\mu$ is equal to 0.09 (constant) and L as the reference length scale, the algorithm used is Semi-Implicit Method for Pressure-Linked Problems (SIMPLE) to solve model equations in OpenFOAM®.

The boundary conditions given are "constant velocity profile" for inlet faces, "zero gradients" for outlet faces, and "wall (no slip)" for lower and upper wall faces in the 2D case. Only the front and rear faces are "empty" in 2D [36].

$$Re = \frac{u\,L}{\nu} \tag{16}$$

where ν: kinematic viscosity of a fluid ($\nu = \mu/\rho$); μ : dynamic viscosity of a fluid; ρ: fluid density; L: characteristic length; u: inlet velocity (m/s) in temperature of air 25 ℃.

This numerical method is used as a dataset which is carried out on variations in the entry velocity at the entry velocity interval, V=44.2 – 50 m/s with a training frequency of 0.01 m/s so that the data set obtained has a total of 600 datasets *.csv in the form of data reattachment length along the x-axis from step.

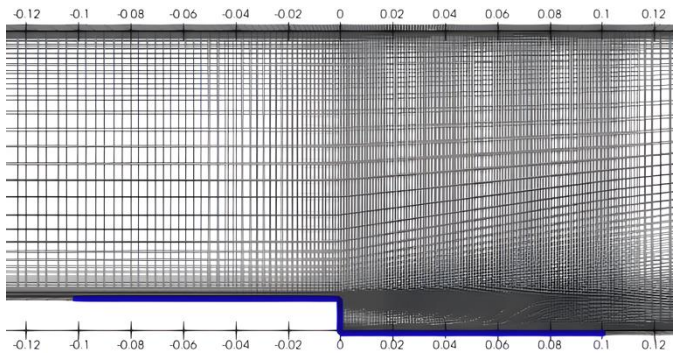


Figure 5: Sample area data of BFS (blue line)

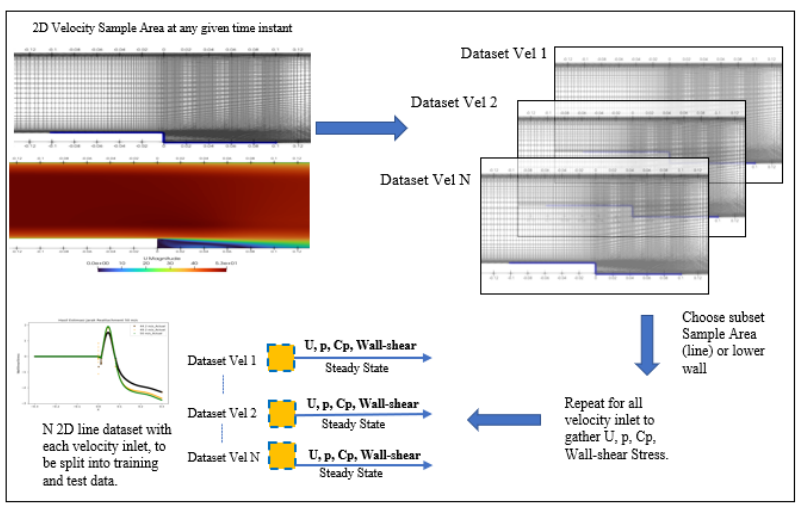


Figure 6: Extracting 2-D datasets process from a line of the flow field

Data from numerical simulations include the coefficient of pressure, pressure, wall shear, and velocity in the area of 1.0 < x > -1.0 on the surface of the lower wall of the model, see Figure 5. The number of samples taken at each velocity is 3000 nodes.

Long short-term memory setup

The data obtained from the simulation results, namely the reattachment length (change in the value of the x-axis wall shear direction), is stored at each velocity increase in the simulation iteration process. Data is retrieved at 600 different velocities and extracted with a mechanism, as shown in Figure 6.

The data are combined in one *.csv document, which is then estimated by creating a model using the Neural Networks LSTM approach in subsection "long short-term memory networks." The estimation results are evaluated using Equation 8. This evaluation is a parameter of the model's feasibility to make estimates of the next Reynolds number.

Results and Discussion

Numerical simulation results

According to the specific model dimensions and operational conditions, the validation procedure is executed by evaluating the numerical simulation results with the actual research data from prior research [33]. The validation step must confirm that the CFD software depicts the actual state appropriately.

Based on the graph in Figure 7, the numerical data from prior research and numerical results from the initial setup for BFS with a modified Reynolds Number do not differ significantly. The disparity arises since the numerical simulation is conducted assuming the system conditions. Nonetheless, numerous numerical settings were challenging to manage. However, since both plots reflect the same trendline, there is no visible variation in value.

As shown in Figure 5, the sampling area for data from the coefficient of pressure (*Cp*), pressure (*p*), and wall shear (*Wss_x*). From the simulation results, the data is processed to get the attachment length by measuring the step distance to the change in the direction of the wall shear value, which shows the reattachment point. Paraview is used for post-processing data to observe the structure as a contour graph.

Figure 8(a) is a flowing contour depicting a velocity contour plot for the case at 50 m/s as a mean velocity. From the results of numerical studies, each cell element is produced as coordinates with velocity, pressure, pressure coefficient, and shear stress (wall shear stress). The reattachment distance value is obtained from the selection of data on changes in the direction value of the wall shear stress in the lower wall (sample) area [34].

As shown in Figure 8(b), the reattachment point occurs between the distances 0.07-0.08 as a sign of the limit of changes in backward and forward

flow. Table 4 shows the reattachment distance with the inlet condition at 44.2 m/s, having a length of 0.07837 mm.

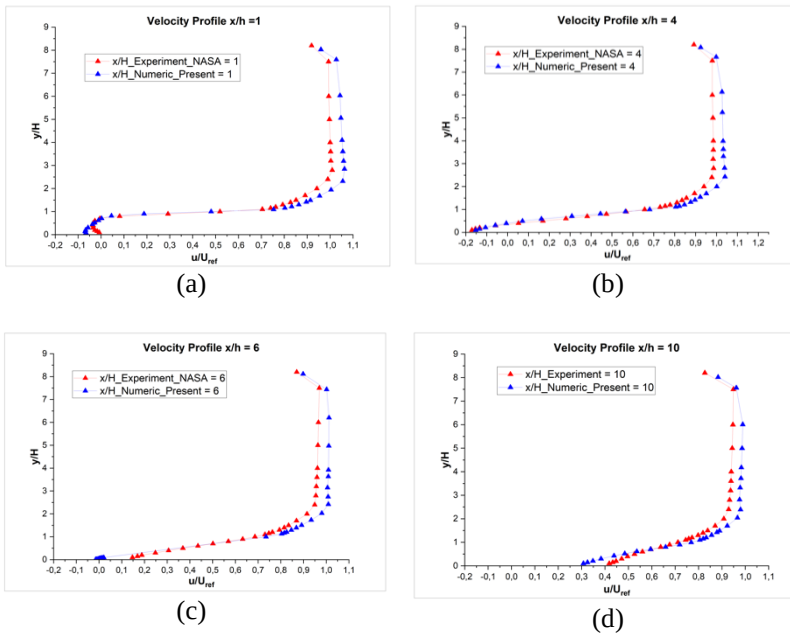


Figure 7: Comparison between the experimental [33] and the present numerical value with variation range "*x*" from step "*h*" based on the setup; (a) 1 x/h, (b) 4 x/h, (c) 6 x/h, and (d) 10 x/h

While in Tables 5 and 6 show the reattachment distance with inlet-velocity conditions of 48.2 m/s and 50.0 m/s. The result of reattachment length is 0.07747 mm and 0.774775 mm, respectively. This result has a slightly varied reattachment length value. Compared with the results of [33] with the same configuration, it also has a reattachment value of 6.26+-0.1 *x/h* or 0.0790 mm. This difference is due to the instability phenomenon caused by a strong adverse pressure gradient. An adverse pressure gradient study has been carried out by Driver and Seegmeler [37].

From picking random data 44.2, 48.2, and 50.0 m/s. Observably, the distance of the reattachment points from the step decreases with continuing to increase velocity. Likewise, it can be seen in Figure 8(c), a wall shear graph from various velocities. From Figure 8(c), the green line is the wall shear line at 50 m/s as the mean velocity, which is relatively lower than the other two wall shears.

The results of each reattachment length from velocity variations are processed in a pre-processing stage. Processing is done by separating the overall data until there are only Reynolds number, velocity, and Wallshear_x values. The data changes as the velocity increases or the Reynolds number changes.

Figure 9(a) illustrates the results of separating the reattachment length data for each Reynolds number or velocity variation. Due to the small Reynolds range, the data was insufficient because of awake vorticities and different Reynold numbers.

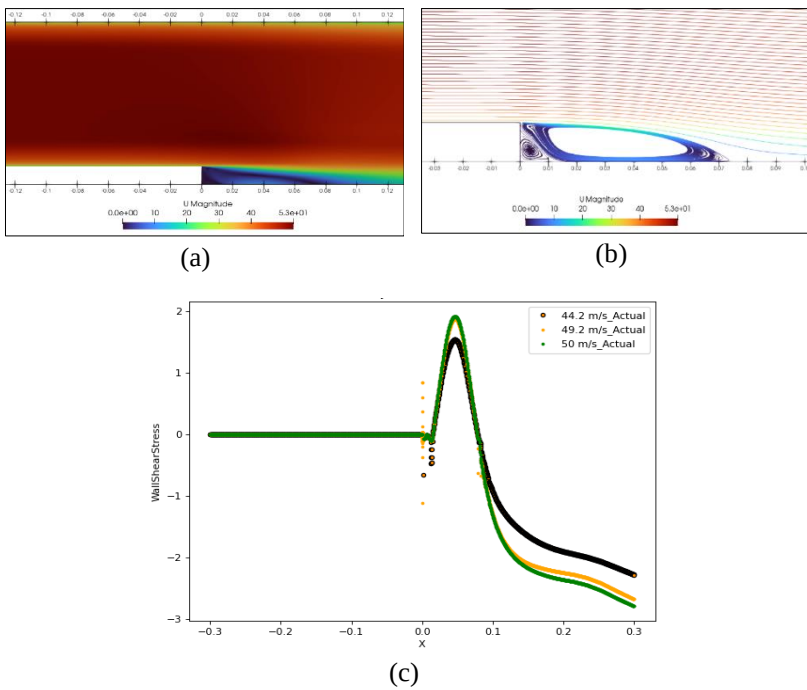


Figure 8: Numerical results of the current study; (a) freestream contour on Backward-Facing Step (50 m/s), (b) streamline contour on Backward-Facing Step (50 m/s), and (c) compared wall shear 44.2 m/s, 48.2 m/s, and 50 m/s

Table 4: Results of reattachment point simulation in velocity 44.2 m/s

Coor_x	*Cp*	*p*	Wss_x
0.0780781	-0.00306	-389.648	0.02241340
0.0783784	-0.00214	-209.273	0.00374021
0.0786787	-0.00124	-121.148	-0.0147147
0.0786787	-0.00036	-0.36035	-0.0328505

Table 5: Results of reattachment point simulation in velocity 48.2 m/s

Coor_x	*Cp*	*p*	Wss_x
0.0771772	-0.00208	-389.648	0.0423038
0.0774775	-0.00101	-209.273	0.0172398
0.0777778	-0.00009	-121.148	-0.0073336
0.0780782	0.00100	-0.36035	-0.0318490

Table 6: Results of reattachment point simulation in velocity 50.0 m/s

Coor_x	*Cp*	*p*	Wss_x
0.0771772	-0.00306	-389.648	0.02241340
0.0774775	-0.00214	-209.273	0.00374021
0.0777778	-0.00124	-121.148	-0.0147147
0.0780781	-0.00036	-0.36035	-0.0328505

Long Short-Term Memory (LSTM) result

The results of training simulations of numerical data have been sorted and tidied up based on variations in velocity. In this case, the data for training and testing is separated with a ratio of 80:20. As shown in Figure 9(b), the orange line represents the test data, while the line with blue colour represents the data used for testing.

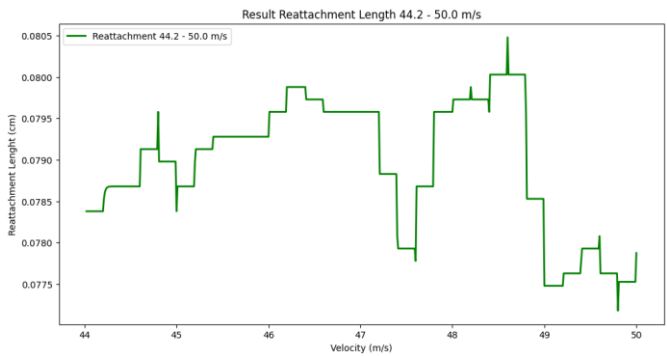
Data training is carried out from the model that has been created and modified. The "data loss" training process from every data was carried out. From Figure 10(a), it can be seen that the loss value continues to fall, at least when doing training. From the estimation results using the LSTM approach, the reattachment length value is obtained with a relatively more significant value (0.079745) than the numerical simulation results.

Figure 10(b) shows that the orange line is described as the prediction line, and the blue line is numerical data. Based on that figure, the nonlinear line was predicted from recent data, and the result was significantly close to the actual data value. The neural network employed is developing a base model in some previous studies with satisfactory results. After performing hypertuning to optimize parameters such as epoch, batch, iteration, and activation function, select the optimal model that can be used to estimate reattachment length. As evaluation metrics, the model's results are evaluated with Root Mean Square Error (RMSE) [28] metric: 0.013.

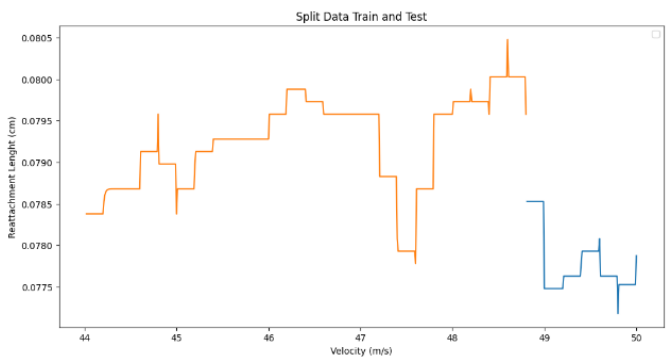
Conclusion

This study succeeded in carrying out a numerical study approach to describe the fluid structure and implement ML using the LSTM method. From the simulation process, the data change length of the reattachment based on the

velocity increase is used as a dataset to estimate and reduce the iteration time of the numerical simulation. The estimation results using the LSTM algorithm resulted in reasonably good training. Reattachment length can be estimated with parameter data in geometric configuration and suitable conditions in this study. This model can estimate the reattachment length. Using the machine learning model, it gets the predicted reattachment length from *Re* 35587 - 40422 variation with a step height of 1.27 cm. The numerical simulation process in modelling the flow structure is intended to optimize a system. Besides that, the simulation is used as data from the results of physical interactions. The machine learning methods can be applied by one of the trainings using a data processing approach to streamline time. The estimation process is expected to Velocity increase the process of consideration for optimizing the system.

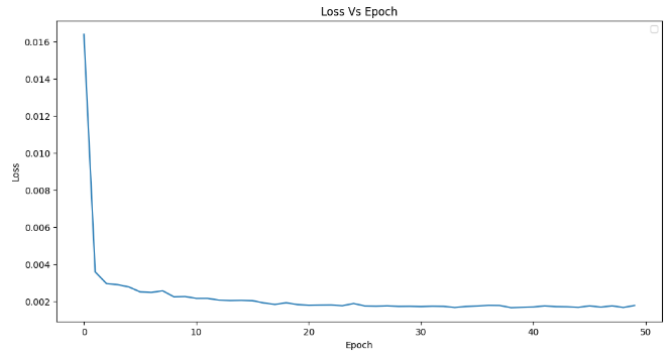


(a)

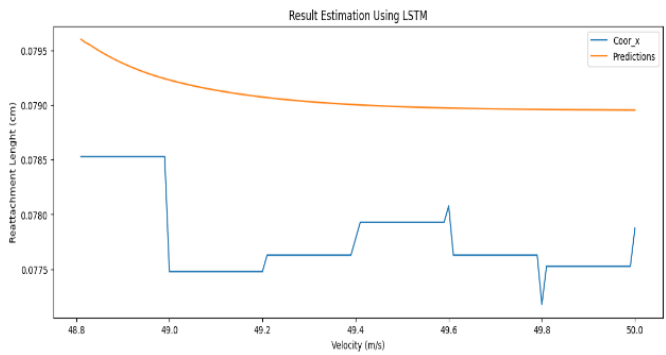


(b)

Figure 9: Results of the present numerical study and split data; (a) reattachment length vs. velocity, and (b) split data training and testing from the present setup



(a)



(b)

Figure 10: Results of long short-term memory; (a) Epoch vs. loss training process, and (b) result estimation prediction vs. simulation

Contributions of Authors

The authors confirm the equal contribution in each part of this work. All authors reviewed and approved the final version of this work.

Funding

This work was supported by the Ministry of Education and Culture, Research and Technology, Higher Education of the Republic of Indonesia [grant number 155/E5/PG.02.00.PT/2022].

Conflict of Interests

All authors declare that they have no conflicts of interest.

Acknowledgment

We thank the High Processing Computer Gunadarma University team for assistance in providing resource computation.

References

[1] E. Perlman, R. Burns, Y. Li, and C. Meneveau, "Data exploration of turbulence simulations using a database cluster", *Proceedings of the 2007 ACM/IEEE conference on Supercomputing - SC '07, Reno, Nevada: ACM Press*, p. 1, 2007. https://doi.org/10.1145/1362622.1362654

[2] Giovanni Calzolari and W. Liu, "Deep learning to replace, improve, or aid CFD analysis in built environment applications: A review", *Building and Environment*, vol. 206, p. 108315, 2021. https://doi.org/10.1016/j.buildenv.2021.108315

[3] S. L. Brunton, B. R. Noack, and P. Koumoutsakos, "Machine Learning for Fluid Mechanics", *Annual Review of Fluid Mechanics*, vol. 52, no. 1, pp. 477–508, 2020. https://doi.org/10.1146/annurev-fluid-010719-060214

[4] B. Podvin, Y. Fraigneau, J. Jouanguy, and J.-P. Laval, "On Self-Similarity in the Inner Wall Layer of a Turbulent Channel Flow", *Journal of Fluids Engineering*, vol. 132, no. 4, p. 041202, 2010. https://doi.org/10.1115/1.4001385

[5] C. W. Rowley and S. T. M. Dawson, "Model Reduction for Flow Analysis and Control", *Annual Review of Fluid Mechanics*, vol. 49, no. 1, pp. 387–417, 2017. https://doi.org/10.1146/annurev-fluid-010816-060042

[6] C. L. Teo, K. B. Lim, G. S. Hong, and M. H. T. Yeo, "A neural net approach in analyzing photograph in PIV", in *Conference Proceedings 1991 IEEE International Conference on Systems, Man, and Cybernetics, Charlottesville, VA, USA: IEEE*, pp. 1535–1538, 1991. https://doi.org/10.1109/ICSMC.1991.169906

[7] C. Lee, J. Kim, D. Babcock, and R. Goodman, "Application of neural networks to turbulence control for drag reduction", *Physics of Fluids*, vol. 9, no. 6, pp. 1740–1747, 1997. https://doi.org/10.1063/1.869290

[8] J. Luo, Y. Li, W. Zhou, X. Chen, and W. Yao, "A Novel Method to Accelerate the Solution of Compliance Using Deep Learning for

Topology Optimization", *Advances in Mechanical Design*, vol. 111, pp. 1781–1792, 2022. https://doi.org/10.1007/978-981-16-7381-8_111

[9] A. Usman, M. Rafiq, M. Saeed, A. Nauman, A. Almqvist, and M. Liwicki, "Machine Learning Computational Fluid Dynamics", *2021 Swedish Artificial Intelligence Society Workshop (SAIS), Sweden: IEEE*, pp. 1–4, 2021. https://doi.org/10.1109/SAIS53221.2021.9483997

[10] R. Vinuesa and S. L. Brunton, "Enhancing computational fluid dynamics with machine learning", *Nature Computational Science*, vol. 2, no. 6, pp. 358–366, 2022. https://doi.org/10.1038/s43588-022-00264-7

[11] D. Teso-Fz-Betoño, M. Juica, K. Portal-Porras, U. Fernandez-Gamiz, and E. Zulueta, "Estimating the Reattachment Length by Realizing a Comparison between URANS k-Omega SST and LES WALE Models on a Symmetric Geometry", *Symmetry*, vol. 13, no. 9, p. 1555, 2021. https://doi.org/10.3390/sym13091555

[12] H. Nowruzi, S. S. Nourazar, and H. Ghassemi, "On the Instability of Two Dimensional Backward-Facing Step Flow using Energy Gradient Method", *Journal of Applied Fluid Mechanics*, vol. 11, no. 1, pp. 241–256, 2018. https://doi.org/10.29252/jafm.11.01.28235

[13] "1000+ COMSOL Multiphysics® Modeling Examples for Download", 2022. [Online] https://www.comsol.com/models (Accessed Sep 05, 2022).

[14] Z. F. Tian, J. Y. Tu, G. H. Yeoh, and R. K. K. Yuen, "On the numerical study of contaminant particle concentration in indoor airflow", *Building and Environment*, vol. 41, no. 11, pp. 1504–1514, 2006. https://doi.org/10.1016/j.buildenv.2005.06.006

[15] L. Chen, K. Asai, T. Nonomura, G. Xi, and T. Liu, "A review of Backward-Facing Step (BFS) flow mechanisms, heat transfer and control", *Thermal Science and Engineering Progress*, vol. 6, pp. 194–216, 2018. https://doi.org/10.1016/j.tsep.2018.04.004.

[16] C. D. Argyropoulos and N. C. Markatos, "Recent advances on the numerical modelling of turbulent flows", *Applied Mathematical Modelling*, vol. 39, no. 2, pp. 693–732, 2015. https://doi.org/10.1016/j.apm.2014.07.001

[17] W. Tollmien, H. Schlichting, and R. W. Riegels, "Uber die ausgebildete Turbulenz," in Ludwig Prandtl Gesammelte Abhandlungen, Springer, Berlinm Heidelberg, pp. 736-751, 1961.

[18] B. E. Launder and D. B. Spalding, "The numerical computation of turbulent flows", *Computer Methods in Applied Mechanics and Engineering*, vol. 3, no. 2, pp. 269–289, 1974. https://doi.org/10.1016/0045-7825(74)90029-2

[19] V. Yakhot, S. A. Orszag, S. Thangam, T. B. Gatski, and C. G. Speziale, "Development of turbulence models for shear flows by a double expansion technique", *Physics of Fluids A: Fluid Dynamics*, vol. 4, no. 7, pp. 1510–1520, 1992. https://doi.org/10.1063/1.858424

[20] V. Yakhot and S. A. Orszag, "Renormalization group analysis of turbulence. I. Basic theory", *Journal of Scientific Computing*, vol. 1, no. 1, pp. 3–51, 1986. https://doi.org/10.1007/BF01061452

[21] P. G. Huang, P. Bradshaw, and T. J. Coakley, "Assessment of Closure Coefficients for Compressible-Flow Turbulence Models", *NASA Technical Memorandum 103882*, p. 18, 1992.

[22] D. P. Mandic and J. A. Chambers, "Recurrent Neural Networks for Prediction in Wiley Series", in Adaptive and Learning Systems for Signal Processing, Communications, and Control. Chichester, UK, 2001. https://doi.org/10.1002/047084535X

[23] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling", *arXiv*, 2014. Accessed: Sep. 05, 2022. [Online]. Available: http://arxiv.org/abs/1412.3555

[24] Z. C. Lipton, J. Berkowitz, and C. Elkan, "A Critical Review of Recurrent Neural Networks for Sequence Learning", *arXiv*, 2015. https://doi.org/10.48550/ARXIV.1506.00019

[25] P. Ramachandran, B. Zoph, and Q. V. Le, "Searching for Activation Functions," *arXiv*, 2017. https://doi.org/ 10.48550/ARXIV.1710.05941.

[26] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory", *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997. https://doi.org/10.1162/neco.1997.9.8.1735

[27] Y. Yu, X. Si, C. Hu, and J. Zhang, "A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures", *Neural Computation*, vol. 31, no. 7, pp. 1235–1270, 2019. https://doi.org/10.1162/neco_a_01199

[28] D. Aparicio and M. I. Bertolotto, "Forecasting inflation with online prices", *International Journal of Forecasting*, vol. 36, no. 2, pp. 232–247, 2020. https://doi.org/10.1016/j.ijforecast.2019.04.018

[29] T. McQueen, D. Burton, J. Sheridan, and M. C. Thompson, "Active control of flow over a backward-facing step at high Reynolds numbers", *International Journal of Heat and Fluid Flow*, vol. 93, p. 108891, 2022. https://doi.org/10.1016/j.ijheatfluidflow.2021.108891

[30] O. Sazhin, "Gas outflow into vacuum over a forward- and backward-facing step in a wide range of rarefaction", *International Journal of Heat and Mass Transfer*, vol. 179, p. 121666, 2021. https://doi.org/10.1016/j.ijheatmasstransfer.2021.121666

[31] A. R. Abu Talib and A. K. Hilo, "Fluid flow and heat transfer over corrugated backward-facing step channel", *Case Studies in Thermal Engineering*, vol. 24, p. 100862, 2021. https://doi.org/10.1016/j.csite.2021.100862

[32] K. Loksupapaiboon and C. Suvanjumrat, "Numerical simulation of flow over a passive disturbance and backward-facing step", *IOP Conference*

Series: Materials Science and Engineering, vol. 1137, no. 1, p. 012046, 2021. https://doi.org/10.1088/1757-899X/1137/1/012046.

[33] C. Rumsey, "2DBFS: 2D Backward Facing Step", Turbulence Modeling Resource, Langley Research Center, Nov 18, 2021. [Online]. Available https://turbmodels.larc.nasa.gov/backstep_val.html/ (Accessed Aug 7, 2023)

[34] A. Satheesh Kumar, A. Singh, and K. B. Thiagarajan, "Simulation of backward facing step flow using OpenFOAM®", *AIP Conference Proceedings*, vol. 1, p. 2204, 2020. https://doi.org/10.1063/1.5141565

[35] D. Cappelli and N. N. Mansour, "Performance of Reynolds Averaged Navier-Stokes Models in Predicting Separated Flows: Study of the Hump Flow Model Problem", *31st AIAA Applied Aerodynamics Conference*, 2013. https://doi.org/10.2514/6.2013-3154

[36] F. R. Menter, M. Kuntz, and R. Langtry, "Ten years of industrial experience with the SST turbulence model. Begell", *Turbulence, Heat and Mass Transfer*, pp. 625-632, 2003.

[37] D. M. Driver and H. L. Seegmiller, "Features of a reattaching turbulent shear layer in divergent channel flow", *AIAA Journal*, vol. 23, no. 2, pp. 163–171, 1985. https://doi.org/10.2514/3.8890