Contemporary Mathematics
https://ojs.wiserpub.com/index.php/CM/

Research Article

ML-Based Approach to Predict Carotid Arterial Blood Flow Dynamics

T Raja Rani [1], Abdullah Al Shibli [2], Mohamed Siraj [3], Woshan Srimal [4], Nooh Zayid Suwaid Al Bakri [5] and T S L Radhika [6]

[1,4] Foundation Programme Department, Military Technological College, 111, Muscat, Oman
[2] Applied & Research Department, Military Technological College, 111, Muscat, Oman
[3] Systems Engineering Department, Military Technological College, 111, Muscat, Oman
[5] MTC Clinic, Military Technological College, 111, Muscat, Oman
[6] Department of Mathematics, BITS Pilani, Hyderabad Campus, Hyderabad, 500078, India

E-mail: rani.t@mtc.edu.om (https://orcid.org/0000-0002-7050-2191), abdullah.alshibli@mtc.edu.om, mohamed.siraj@mtc.edu.om, woshan.srimal@mtc.edu.om, nooh.albakri@mtc.edu.om, radhikatsl@hyderabad.bits-pilani.ac.in

Received: 13 June 2023; **Revised:** 18 July 2023; **Accepted:** 31 July 2023

Abstract: In the current study, a numerical model has been developed to simulate the blood flow characteristics in the human carotid artery. The data thus generated is analyzed to understand the blood flow variations and predict the flow characteristics using Machine Learning techniques. In developing the numerical model, the key features of the system, namely, the blood, is modeled as an incompressible Newtonian fluid, and the artery is an elastic pipe. This model is simulated using COMSOL software by varying the material properties of the artery. Univariate analysis was performed to gain insight into the features' behaviour and target variables. Subsequently, machine-learning regression models were trained using the data generated from the idealized human carotid artery. Furthermore, the validity of the data was ensured by comparing it with flow division ratios available in the literature. The evaluation of these models was conducted by calculating the Mean Absolute Error values for the test dataset, resulting in the following values: polynomial regressor (0.0106), hyper-tuned support vector regressor (0.0487), decision tree regressor (0.000), random forest regressor (0.0156), Adaboost (0.0508), gradient-boosting (0.0044), and XGboost (0.0043). A quantile loss function was employed to assess the prediction uncertainty. According to the theory of loss function, models with low loss values are considered good predictors. The prediction uncertainty was measured by applying quantile loss function, and it identified that the random forest regressor as the best predictor model for the data, followed by the polynomial regression of degree 3. Prediction intervals for the target variable were computed by leveraging the random forest quantile regressor model. Moreover, the developed polynomial model was utilized to investigate the presence of stenosis in the artery.

Keywords: carotid artery; blood flow; mathematical model; COMSOL multiphysics; ML algorithms; ensemble models; quantile loss function.

MSC: 76Z05, 92C10, 92C35

DOI: https://doi.org/10.37256/cm.5120243224

Nomenclature

Term	Description
AvgVel	Average blood velocity (in each segment)
CCA	Common Carotid Artery
D	Blood Density
ECA	External Carotid Artery
InPr	Inlet Pressure (Blood pressure at the CCA entrance)
InVFR	Inlet Volumetric Flow Rate (at the CCA entrance)
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
Out1Vel	Blood velocity at outlet1, i.e., ICA
Out1VFR	Volumetric flow rate at outlet1, i.e., ICA
Out2Vel	Blood velocity at outlet2, i.e., ECA
Out2VFR	Volumetric flow rate at outlet2, i.e., ECA
PARDISO	Parallel Direct Sparse Solver
RBF	Radial Basis Function
RFR	Random Forest Regressor
RMSE	Root Mean Square Error
SVR	Support Vector Regressor
VIF	Variance Inflation Factor
Vis	Blood viscosity

1. Introduction

Every organ of the human body has a purpose to serve. All these purposes fall under one umbrella: "To maintain and sustain the functioning of the biological system." Also, every artery has the ultimate function of supplying oxygenated blood to the target organ and boosting its functioning.

It is known that carotid arteries supply blood to the face, neck, and brain. There are two carotid arteries, one on the right and the other on the left side of the human neck. These pairs of arteries are not just for supplying blood to the face, neck, and brain but also indicate a person's breathing by pulsating on palpation of the artery. Moreover, the carotid artery is one of the vital points of the human body, meaning they are extremely sensitive: any trauma to this artery is proven fatal; the reason that even a brief pause in the blood supply can cause severe problems as brain cells start to die within no time after this pause. Thus, this paper proposes to work on understanding blood flow in the carotid artery bifurcation under different flow conditions.

Another reason for the current study is that stroke is the major contributor to deaths worldwide. Furthermore, this study is believed to be essential, particularly in the current situation where COVID-19 has affected most of the global population, and recent findings have revealed that COVID-19-affected persons are more at risk for artery and vein occlusions [1-3]. Secondly, studying flow characteristics in such a crucial artery is beneficial as it would aid medical practitioners in predicting the reason for the malfunctioning of its targeted organs with some degree of accuracy. Though this study is not the first of its kind, the novelty lies in computing numerics related to diverse blood flow and blood vessel conditions.

Briefly, this section brings a new look to the existing literature in the following areas: (I) Experimental, (II) Modeling approach. While the experimental approach involved invasive and non-invasive methods to understand the blood flow dynamics, Mathematical, Engineering, and Numerical models are the three modeling approaches. Over the past few decades, much insightful work has been done in the current field of study. A few works on the experimental methods are: Uematsu et al. used over 8000 measurements of the blood flow in 120 normal subjects and 550 patients with neurological diseases and assessed cerebral blood flow using an ultrasonic volume flow meter [4]. Nigel et al. predicted stenosis in the ICA (Internal Carotid Artery) by relating it to the blood flow in the CCA (Common Carotid Artery) measured with a pulsed Doppler duplex scanner [5]. Oktar et al. obtained the values of BFV (blood-flow volume) in the ICA of healthy individuals using color Doppler, power Doppler, and B-flow ultrasound techniques. They compared the results with those from phase-contrast MR imaging [6].

Srinivasacharya and Madhava Rao developed a mathematical model for the bifurcated artery and studied the blood flow by modeling blood as a couplestress fluid [7]. They observed an increase in the flow rate with an increase in the values of the model parameters. However, impedance has decreased with an increase in these

values. Chakravarty and Sen developed a mathematical model to understand the flow of blood in a catheterized bifurcated artery under stenotic conditions [8]. They considered the geometry of a bifurcated artery with stenosis in the daughter segment. A two-dimensional mathematical model with blood as a Newtonian fluid has been developed to study the flow, and they solved it using the finite difference scheme. Their findings revealed that arterial catheterization remarkably influences viscous dissipation in the parent segment more than in the daughter segments. Ruchi Agarwal et al. developed a mathematical model coupling the effect between the fluid flow and the elastic deformation of the artery. They calculated the volumetric flow rate and the impedance due to the pulsatile flow in the carotid artery. They analyzed flow at various locations in CCA and ICA [9]. Muraca et al. developed a numerical model in COMSOL software that aids in assessing the wall shear stress in carotid artery bifurcation [10]. Perktold and Rappitsch quantitatively analyzed the flow dynamics due to atherogenesis in the carotid artery [11]. Onwuzu et al. compared the wall shear stress values computed using a CFD (Computational Fluid Dynamics) model with those measured using Ultrasound Doppler Velocimetry [12]. Their study revealed that the computed values are comparable with Doppler velocimetry measurements, thus promising the reliability of numerical or CFD models in the blood flow study.

Abdi et al. developed a lumped model for the cardiovascular system to study the effect of stenosis in ICA [13, 14]. Onaizah et al. constructed a lumped parameter model for blood flow in the carotid artery [15]. Their study showed that the reduction in blood flow in ICA was significant only when stenosis was above 70%. Moreover, vascular stiffening altered the carotid flow or flow waveform.

In the present study, a numerical model has been developed to simulate the blood flow characteristics in the human carotid artery. The system's key features, namely, the blood, is modeled as an incompressible Newtonian fluid, and the artery is an elastic pipe. This model is simulated using the COMSOL software by varying the material properties of the artery, the blood, and the flow characteristics and conditions. The data thus generated is analyzed to understand the blood flow variations and predict the flow characteristics using suitable ML models. The novelty of this work is in assessing the possible prevalence of stenosis in the artery using vitals such as blood pressure and CCA velocity and computed prediction intervals for the target variable.

This article is organized as follows. Section 2 presents the problem identification and research objectives. Section 3 contains a detailed description of the adopted methodology. Section 4 presents the results of experimental research along with their interpretation. Section 5 presents the final conclusions of the research, indicating their limitations, practical application, and future research directions in this field.

2. Problem Identification

Most of the studies mentioned in the literature survey are on understanding the blood flow in ICA or CCA under stenotic conditions in ICA. However, the current study considers the cases where stenotic or atherosclerotic conditions prevail in either or both the bifurcated segments, ICA and ECA. Thus, the scientific aim of this work is to discuss the effect of the presence of these conditions in one segment on the other. Further, this study aims to extend the existing research in carotid artery studies, wherein we intend to fill critical knowledge gaps and provide a comprehensive foundation for future research endeavors.

The dimension of the present study is translational, wherein we propose to predict blood flow characteristics to provide accurate information without using invasive or expensive methods like MRI and Ultrasound. Thus, this study would help solve the need for non-invasive techniques in assessing vascular changes in the carotid artery that Navneet Singh et al. pointed out in their work [16].

Further, we propose a quantitative and experimental (using computer-based simulations) study with a descriptive research question defined as:

Can ML models on simulated data help identify the presence of stenosis in the artery?

Research Objectives:

1. To develop and validate a numerical model for the human carotid artery.
- Simulate the model for healthy and unhealthy conditions.
- Compute appropriate physical quantities and compare their values with the clinical results.

2. To estimate the average blood velocity in the carotid artery using ML algorithms.
- Develop ML models that best fit the data.
- Compute the metrics on the train and test data and identify the most reliable ML prediction model.

3. To use the ML model to assess the blood flow variations in the artery.

- Plot the average blood velocity in CCA as a function of the blood flow characteristics in the carotid artery using MATHEMATICA software.

Exposures and Outcomes:

Fluid (blood) related parameters such as density and viscosity, flow-related parameters such as the inlet pressure and outlets' blood velocities, and material parameters such as the artery's density and Poisson ratio are the independent variables. The outcomes are the average blood velocity and volumetric flow rate in all three arterial segments.

3. Methodology

The step-by-step procedure from data collection to data analysis to achieve the proposed objectives is presented in this section.

Step:1 Prepare data collection instrument:

As described in the introduction, the carotid artery supplies blood to the face, neck, and brain. The main vessel, CCA, is bifurcated into the ECA and ICA, as shown in Figure 1. ECA primarily supplies blood to the face and neck, while ICA supplies blood to the brain.

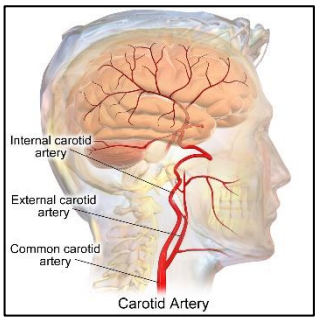

Figure 1. Anatomy of Carotid Artery [Source: Internet]

In this paper, we developed a numerical model to capture the blood flow dynamics in the carotid artery. The key features of the blood circulatory system, namely, the blood, have been modeled as a Newtonian fluid, and the artery is an elastic circular pipe with bifurcation. This model was simulated using clinical data on the anatomy of the artery and the physiological behavior of the blood flow. The COMSOL Multiphysics software is used to model this blood flow problem in an idealized artery, as shown in Figure 2.

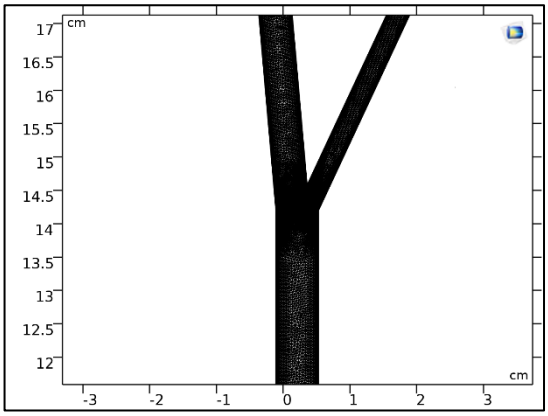

Figure 2. Geometry of the Artery in COMSOL

The stationary Navier Stokes equation was considered and implemented using the laminar flow interface in this study. The viscosity and density values were determined based on available data for human blood. The parameters describing the elastic behavior of the artery were obtained from anatomical data of the carotid artery. For the discretization of both velocity and pressure variables, the P1-P1 linear finite element method was employed. A physics-based mesh consisting of 40637 triangular elements and 4192 quadrilateral elements was

utilized, resulting in a total of 76758 degrees of freedom to determine the velocity (51172) and pressure (25586) variables. The mesh parameters can be found in Table 1.

Table 1. Parameters of the extremely fine mesh

Parameter	Amount
Number of elements	44829
Number of vertex elements	17
Number of edge elements	2466
Average element quality	0.8004
Minimum element quality	0.08677
Mesh area	15.76cm^2

The resulting nonlinear flow problem is solved using Newton's method. Zero initial conditions and no-slip boundary conditions were imposed on the boundary walls, ensuring the absence of backflow. To solve the system, the PARDISO solver was employed.

Step:2 Data Collection (by performing simulations):

The data on the different input parameters to the COMSOL model builder for modelling carotid arterial blood flow in humans are displayed in Table 2 [17–20]. The model parameters, density, and viscosity assume values of human blood since blood is modelled as a Newtonian fluid. According to Table 2, the parameters indicating the carotid artery's elastic nature, density, bulk modulus, and Poisson ratio are given values linked to the human carotid artery. The artery length, diameter, and dimensions are in accordance with the information provided in the literature. Additionally, the ranges listed in Tables 2 & 3 are used to determine the intake pressure and output velocities needed to create blood flow.

Table 2. Data on characteristics of blood and carotid artery [17- 20]

Fluid (blood) Properties		**Artery Properties**	
Density (kg/m^3)	1060	Density	960
Viscosity (Pa.s)	0.004	Bulk Modulus (N/m^2)	1.2×10^8
		Poisson Ratio	0.45
ARTERY	**CCA**	**ICA**	**ECA**
Diameter (mm)	6.10 ± 0.8	4.8 ± 0.3	3.0 ± 0.6
Length (cm)	13.6 ± 1.2	8.6 ± 1.4	8.6 ± 1.4
Velocity (m/sec)	-	0.187- 0.295	0.121-0.185
		Bifurcation Angle (36 ±11)$^\circ$	

The data for carrying out simulations are shown in Tables 3(a) &(b). The inlet pressure has taken 100-, 120-, and 130-mm Hg values. The blood velocities at outlets 1 (ICA) and 2 (ECA) are assigned values from the prescribed range (normal or healthy), as shown in the eighth row (third and fourth columns) in Table 2.

Table 3(a). Dimensions of the simulated artery

ARTERY	CCA	ICA	ECA
Diameter (mm)	6.2	4.8	3.0
Length (cm)	14.2	9.0	9.0

Table 3(b). Data for running simulations.

InPr (mmHg)	Out1vel (ICA) (m/s)	Out2vel (ECA) (m/s)	Bifurcation Angle (degrees)
100	0.241	0.153	
120	0.193	0.122	30°
130	0.1205	0.077	

❖ **Review the collected data for quality and completeness.**

Table 4 presents the information on data types and counts of the numerical model parameters in the segment CCA. This table shows there is no missing data. Similar observations were noted in the other two segments, ECA and ICA.

Table 4. Information on numerical model parameters from python code

#	Column	Non-null Count
0	Vis	90non-null
1	D	90non-null
2	InPr	90non-null
3	Out1vel	90non-null
4	Out2vel	90non-null
5	AvgVel	90non-null
6	InVFR	90non-null
7	Out1 VFR	90non-null
8	Out2 VFR	90non-null

dtypes: float64(9), int64(3)

Exposures:

Vis: Blood viscosity (Pa.s)
D: Blood Density (kg/m3)
InPr: Blood pressure at the CCA entrance (mm Hg)
Out1Vel: Blood velocity at exit 1, i.e., ICA (mm Hg)
Out2Vel: Blood velocity at exit 2, i.e., ECA (mm Hg)

Outcomes:

AvgVel: Average blood velocity in each segment (m/s)
InVFR: Volumetric flow rate at the CCA entrance (m^3/s)
Out1VFR: Volumetric flow rate at exit 1, i.e., ICA (m^3/s)
Out2VFR: Volumetric flow rate at exit 2, i.e., ECA (m^3/s)

Step:3 Data Management:

Data from the numerical model has been collected as per the format shown in Table 5. For each set of exposures, the outcomes are evaluated using the COMSOL software and noted in the designated columns in the table. The flow division ratios, the ratio of volumetric flow rate in ICA to CCA, the ratio of the flow rate in ECA and CCA, and that of ECA and ICA are computed and noted.

Table 5. Data collection format

	Exposures					Outcomes				Flow Division Ratios		
Sl. No	Vis	D	InPr	Out1vel	Out2vel	AvgVel	InVFR	Out1VFR	Out2VFR	Out1VFR/ InVFR	Out2VFR/ InVFR	Out2VFR/ Out1VFR
1												

Step 4: Data Analysis:

The first step is to compute the statistics of exposures to get some basic information on the exposures, as shown in Table 6.

Table 6. Descriptive statistics of exposure

Statistics	Vis	D	InPr	Out1vel	Out2vel
count	90	90	90	90	90
mean	0	1061.67	112.22	0.17	0.11
std	0	4.74	11.39	0.07	0.04
min	0	1060	100	0.07	0.05
25%	0	1060	100	0.12	0.08
50%	0	1060	120	0.19	0.12
75%	0	1060	120	0.24	0.15
max	0.01	1075	130	0.24	0.15

The next step in the process is data validation, where the reliability of the developed numerical model as a source for generating data is assessed.

❖ Data Validation

Table 7 indicates what percentage of fluid (blood) in CCA flows into ICA and ECA at various inlet pressures. Generally, around 72% of the blood from CCA enters the ICA under normal conditions. However, atherosclerosis or stenotic conditions in ICA has reduced this percentage, as shown in Table 7.

The computed flow division ratios of ICA/CCA, ECA/ICA, and ECA/CCA were found to be in the range as reported by Marshall [17], and the details are in Table 7. Further, it has been observed that stenotic and other unhealthy conditions in one or both segments had reduced the blood flow in CCA, as shown in Table 8. The current study considers the situation where the flow in ICA or ECA got reduced due to conditions like atherosclerosis in these arteries. Hence, the states with a 20%, 50%, or 70% reduction in the blood velocity at the outlets have also been discussed and presented in Tables 7 & 8.

Table 7. Percentage of blood flow from CCA to ICA & ECA

InPr	InPr		Deviation of blood velocity from the normal range	
	ICA	ECA	ICA	ECA
	×	×	71.68	28.32
	✓	×	66.96	33.04
	✓✓	×	55.86	44.14
	✓✓✓	×	43.16	56.84
100	×	✓	76.04	23.96
	×	✓✓	83.41	16.59
	×	✓✓✓	89.38	10.62
	✓	✓	71.68	28.32
	✓✓	✓✓	66.68	33.1
	✓✓✓	✓✓✓	55.86	44.14
	×	×	71.75	28.25
	✓	×	67.04	32.96
	✓✓	×	55.95	44.05
	✓✓✓	×	43.25	56.75
	×	✓	76.11	23.89
120	×	✓✓	83.46	16.54
	×	✓✓✓	89.42	10.58
	✓	✓	71.84	28.16
	✓✓	✓✓	71.62	23.38
	✓✓✓	✓✓✓	71.71	28.29
	×	×	71.68	28.32
	✓	×	66.97	33.03
	✓✓	×	55.86	44.14
	✓✓✓	×	43.16	56.84
	×	✓	76.05	23.95
130	×	✓✓	83.42	16.56
	×	✓✓✓	89.38	10.62
	✓	✓	71.77	28.23
	✓✓	✓✓	71.55	28.45
	✓✓✓	✓✓✓	71.64	28.36

Table 8. Percentage decrease in blood flow from CCA to ICA & ECA

Inlet Pressure in (mmHg)	Deviation of blood velocity from the normal range		% Decrease in blood flow in CCA	% Decrease in blood flow in ICA	% Decrease in blood flow in ECA
	ICA	ECA	-	-	-
	×	×	-	-	-
	✓	×	14.27	19.92	0
	✓✓	×	35.84	50	0
	✓✓✓	×	50.18	70	0
100	×	✓	5.74	0	20.26
	×	✓✓	14.07	0	49.67
	×	✓✓✓	19.81	0	69.94
	✓	✓	20.02	19.92	20.26
	✓✓	✓✓	49.91	50	49.67
	✓✓✓	✓✓✓	69.98	70	69.94

× indicates no deviation from the normal range of blood velocity. ✓ indicates a 20% decrease from the normal velocity. ✓✓ indicates a 50% decrease from the normal velocity. ✓✓✓ indicates a 70% decrease from the normal velocity.
Here blood viscosity is 0.004 Pa.s and density = 1060 kg/m^3

4. Results & Discussions

This section presents plots of velocity and pressure distribution in the carotid artery for one set of parameters, followed by data analysis using ML algorithms.

Figures 3(a) & 3(b) depict velocity and pressure contours when InPr is 100, out1vel =0.241, out2vel is 0.077, D =1060 and Vis=0.0045.

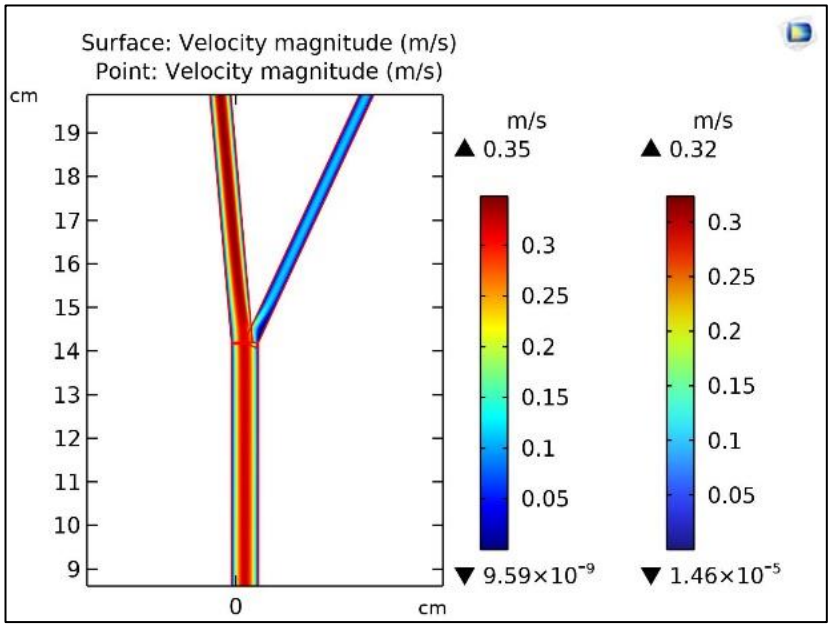

Figure 3(a). Velocity Contour

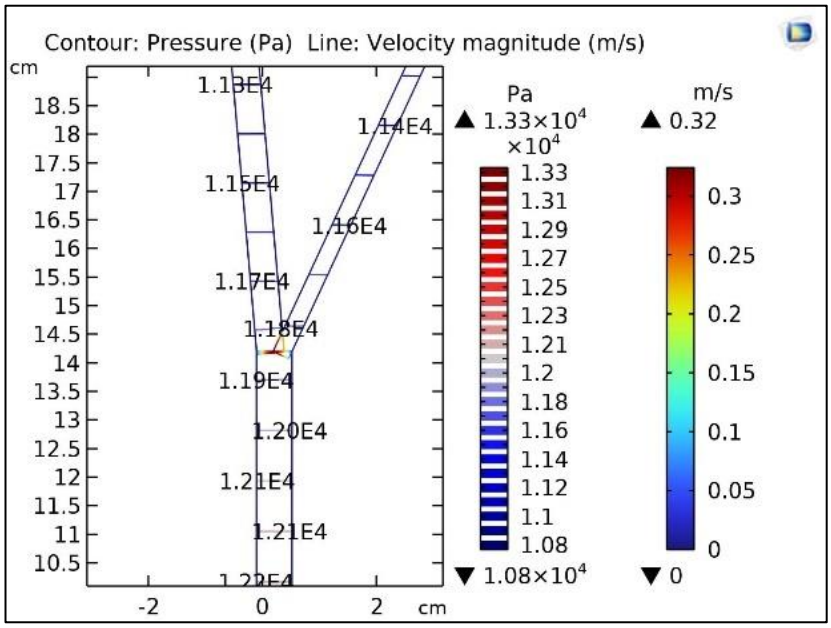

Figure 3(b). Pressure Contour

Further analysis of the collected (simulated) data using ML algorithms is now presented, with a step-by-step process being depicted in the flow chart in Figure 4.

This section provides the detailed description of the data analysis process: To begin with, univariate analysis was performed to explore the features in the data set, where the range and central tendency were examined. The descriptive statistics of features were also analyzed for the presence of outliers. It was observed that the maximum and 75% exposure (features) values were not significantly different, indicating no outliers in the data.

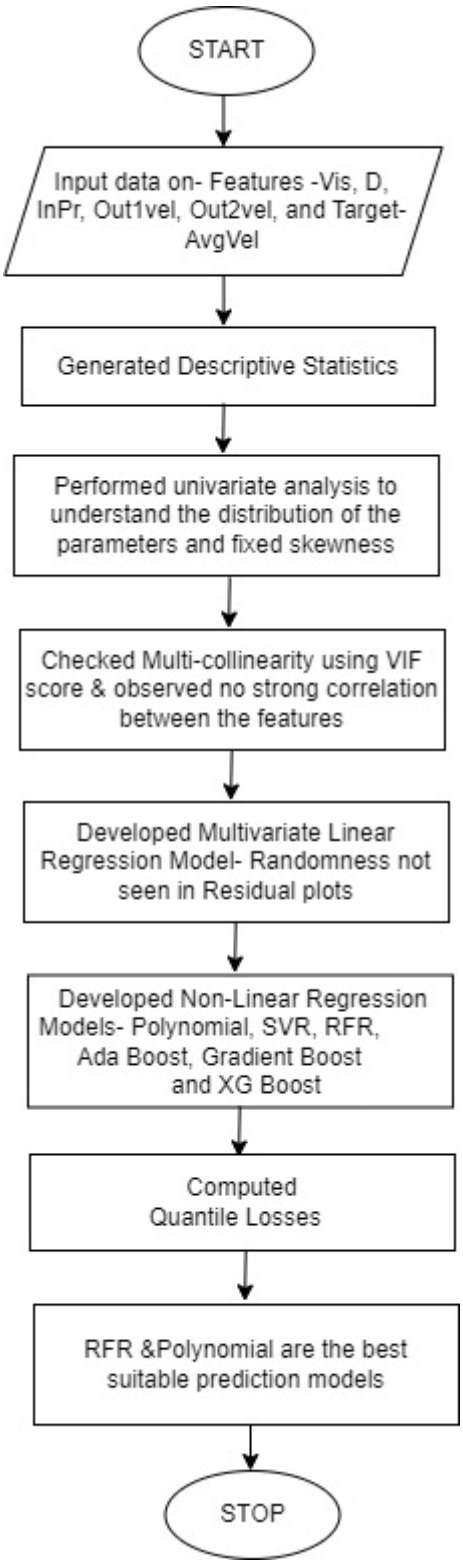

Figure 4. Flow Chart

Then, a univariate analysis on the target variables was performed through bar charts, and no skewness in the data was observed. Subsequently, the next step of performing a multivariate analysis was undertaken, which involved multiple dependent variables or features resulting in one outcome or target variable.

The VIF score was computed for each feature to determine the extent of correlation among them. It was observed that they are not correlated, as the VIF score of each feature was less than 5. Consequently, the first model, namely, the linear regression model with Vis, D, InPr, Out1vel, and Out2vel as the features and AvgVel as the target, was developed. The metrics, namely, MAE (Mean Absolute Error), RMSE (Root Mean Square Error), R2, and adj-R2 scores were calculated for the train and test data. It was noted that the RMSE and MAE

for the train and test data set are almost similar (as per Table 9), indicating a good fit of the model to the data. However, patterns were observed in the residual plots, leading to the adoption of nonlinear models, namely, Polynomial, SVR, Decision Tree, RFR, Adaboost, Gradient Boosting, and XGBoost. An overview of these models is provided below:

A polynomial model is a basic linear regression model with a higher degree that can capture the complex relationship between the features and the targets. For instance, a two-feature quadratic (polynomial of degree 2) model has 6 terms as shown;$z = ax^2 + bxy + cy^2 + dx + ey + f$. Here x and y are the features, and z is the target variable. In general, a polynomial model of degree n with k features has $\binom{n+k}{k}$ terms. Thus, a polynomial model of a higher degree is usually very complex but gives us more flexibility and higher accuracy as it utilizes multiple variables and their combinations in the same equation. However, the choice of the degree of the polynomial is crucial, as it is a hyperparameter that needs to be selected wisely. A higher-degree polynomial may result in overfitting the data, whereas lower-degree models may lead to underfitting.

The decision tree is a classifier that adopts a tree-like structure, where features of a dataset are represented by internal nodes, decision rules are depicted by branches, and each leaf node corresponds to an outcome. It serves as a graphical representation to derive possible solutions or decisions based on provided conditions. The key challenge lies in determining the attribute for the root node at each level, a task known as attribute selection. Two widely used measures for attribute selection are Information Gain and Gini Index.

Support Vector Regressor (SVR) is an ML algorithm used for regression analysis. Unlike the Support Vector Machine (SVM) algorithm used for classification, SVR seeks a hyperplane that fits the data in a continuous space. This is attained by mapping the input variables to a high-dimensional feature space and finding the hyperplane that maximizes the distance between the hyperplane and the data points closest to it, while minimizing the prediction error. Nonlinear relationships between the features and targets can be addressed by SVR through the use of a kernel function that maps the data into a higher-dimensional space. model, the hyperparameters include the hyperplane, C, which controls the error, and the choice of the kernel.

Random Forest Regressor (RFR) is an ensemble learning method that enhances predictive accuracy by combining predictions from multiple ML algorithms to create a more robust and accurate predictive model. This algorithm uses multiple decision trees and a technique called Bootstrap and Aggregation, commonly called bagging. Bootstrap refers to the random sampling of a small subset of data from the data set with replacement, which allows to understand bias and variance within the data. Bagging is a procedure used to reduce variance, and it makes each decision tree model run independently and then aggregates the outputs at the end without preference for any model. The important hyperparameters are n_estimators representing the number of trees to be built, max_features for the maximum number of features that the algorithm considers in splitting a node; mini_sample_leaf for determining the minimum number of leaves for splitting an internal node; criterion for how to split the node in each tree, entropy, Gini, impurity or Log Loss and max_leaf_nodes for fixing the maximum leaf nodes in each tree.

Adaboost (short for Adaptive Boosting) is a boosting algorithm combining multiple weak classifiers to create a more robust classifier. It works by training multiple weak classifiers on different subsets of the training data and then combining their predictions to make a final prediction.

Gradient Boosting is another boosting algorithm that combines multiple weak learners to create a strong learner. It works by sequentially adding weak learners to the model and updating the weights of the training examples based on the residual errors of the previous models.

XGBoost (short for Extreme Gradient Boosting) is a highly optimized implementation of the Gradient Boosting algorithm. This algorithm is similar to Gradient Boosting but includes additional features such as regularization (both L1 and L2) and tree pruning to improve performance.

Returning to the data analysis using these nonlinear models, the polynomial model will be explored first. A Python code has been developed to determine the metrics of the polynomial regressors with different degrees. In order to identify the best-fit polynomial, the graph for MAE vs. the degree of the polynomial was depicted for training and testing data, as shown in Figure 5. It can be inferred from these graphs that a third-degree polynomial serves as a superior nonlinear prediction model for this data. Subsequently, the model was constructed, and it is presented in the Appendix.

Moving forward, the SVR, RFR, Adaboost, Gradient boosting, and XG boost algorithms were employed, and the corresponding metrics were computed as presented in Table 9. According to the table, the metrics computed on the train and test data sets revealed a high MAE for SVR, although R2 and adj-R2 were closer to unity. Hence, it can be concluded that SVR is not a good predictor model.

However, the hyper-tuned model, built using the RBF (Radial Basis Function) kernel with the parameters C=1000 and gamma=0.001, demonstrated a significant reduction in MAE. Moreover, upon reviewing the MAE

values from Table 9, it was evident that the Decision Tree, RFR, and boosting algorithms exhibited lower MAE values on our dataset. While MAE serves as a useful metric for identifying acceptable regressor models, it is essential to consider the quantile loss function to make a more comprehensive evaluation and determine the most suitable regressor models [23]. This allows us to account for prediction uncertainty and obtain a more robust understanding of the models' performance.

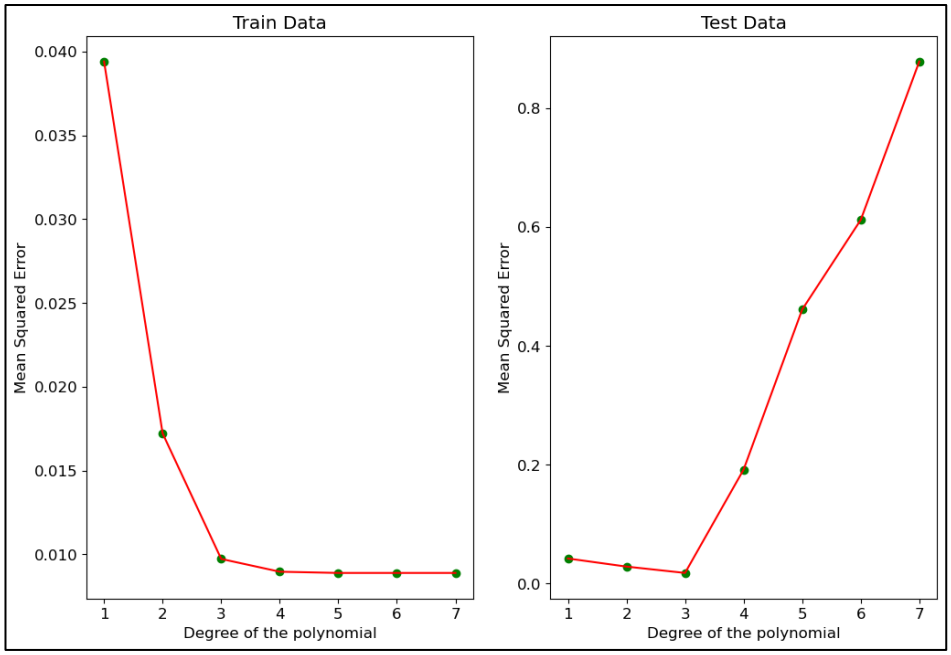

Figure 5. Graphs depicting MSE vs. degree of the polynomial

Table 9. Metrics of regressor models

Regressor Models	Data Set	RMSE	MAE	R^2	Adj-R^2	MAPE
Linear	**Train**	0.0394	0.0314	0.9984	0.9983	-
	Test	0.0422	0.0341	0.9984	0.9977	-
Polynomial degree 3	**Train**	0.0097	0.0055	0.9999	0.9995	-
	Test	0.0179	0.0106	0.9997	1.0001	-
Support Vector	**Train**	0.1396	0.1032	0.9780	0.9760	4.0573
	Test	0.3154	0.2249	0.9177	0.8981	2.5274
Support Vector Tuned	**Train**	0.0582	0.0480	0.9962	0.9958	2.6915
	Test	0.0593	0.0487	0.9971	0.9964	0.2094
Decision Tree	**Train**	0.0000	0.0000	1.0000	1.0000	0.0000
	Test	0.0000	0.0000	1.0000	1.0000	0.0000
Random Forest	**Train**	0.0068	0.0051	0.9999	0.9999	0.1703
	Test	0.0226	0.0156	0.9995	0.9992	0.4461
Ada Boost	**Train**	0.0445	0.0360	0.9980	0.9978	2.9066
	Test	0.0626	0.0508	0.9962	0.9941	3.5796
Gradient Boosting	**Train**	0.0005	0.0003	1.0000	1.0000	0.0205
	Test	0.0065	0.0044	0.9999	0.9999	0.4254
XG Boost	**Train**	0.0008	0.0006	0.9999	0.9999	0.0211
	Test	0.0064	0.0043	0.9999	0.9999	0.4406

In the last and most crucial step, the focus was directed towards quantifying the prediction uncertainty associated with these models. The Quantile loss function was employed to provide insights into the level of uncertainty surrounding point estimation. Figure 6 illustrates the results for the 0.1, 0.5, and 0.9 quantiles, offering a comprehensive view of the models' performance in terms of uncertainty. Analysis of the plot reveals that RFR exhibits the minimum quantile loss, indicating its ability to capture the prediction uncertainty effectively. Following closely behind, the polynomial regressor model emerges as the second-best option in terms of quantile loss. A Python code was developed for the quantile random forest regressor model, enabling the computation of interval estimates for the target variable. The results of these interval estimates were presented in Table 10.

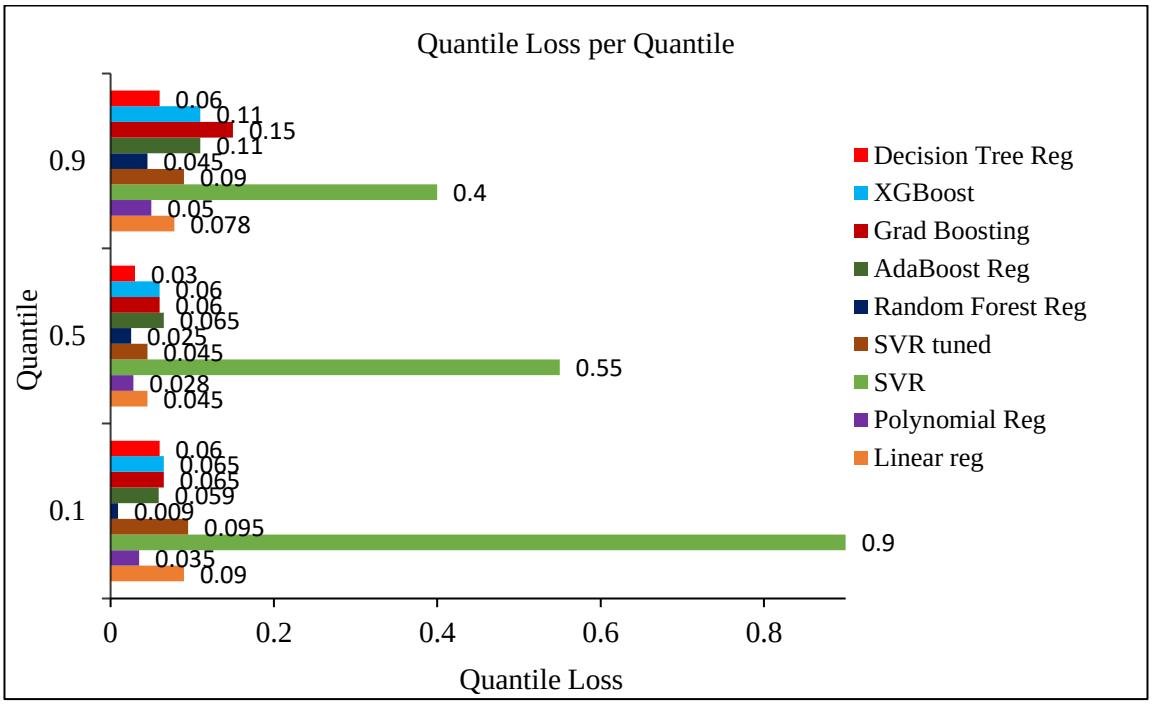

Figure 6. Quantile Loss per Quantile for the ML models

Table 10. Interval estimates for the target variable, avgvel

Features					Target (True value)	Interval Estimate of Target Variable		
Vis	**D**	**InPr**	**Out1vel**	**Out2vel**	**AvgVel**	**0.1 Quantile**	**0.5 Quantile**	**0.9 Quantile**
0.0045	1075	120	0.1205	0.153	0.232483	0.17894559 0	.23229798 0	0.23873813
0.0045	1075	120	0.0723	0.153	0.179428	0.17894559	0.17894559	0.1829619
0.0045	1075	120	0.241	0.122	0.342363	0.29123726	0.34279391	0.35336791
0.0045	1075	120	0.241	0.077	0.31443	0.30300203	0.31470395	0.32619867
0.0045	1075	120	0.241	0.046	0.295083	0.29524845	0.29524845	0.31470395
0.0045	1075	120	0.193	0.122	0.291094	0.29123726	0.29123726	0.32041498
0.0045	1075	120	0.1205	0.077	0.184646	0.18419325	0.18419325	0.23229798
0.0045	1075	120	0.0723	0.046	0.1116	0.11073822	0.11073822	0.11275982
0.004	1060	130	0.241	0.153	0.372377	0.32396078	0.37297629	0.37297629
0.004	1060	130	0.193	0.153	0.320109	0.30056721	0.32041498	0.37297629
0.004	1060	130	0.1205	0.153	0.238887	0.18407012	0.23873813	0.23873813
0.004	1060	130	0.0723	0.153	0.183422	0.17232702	0.1829619	0.23873813
0.004	1060	130	0.241	0.122	0.352878	0.32396078	0.35336791	0.35336791
0.004	1060	130	0.241	0.077	0.324027	0.30386354	0.32435475	0.32725607
0.004	1060	130	0.241	0.046	0.30365	0.29524845	0.30386354	0.32435475
0.004	1060	130	0.193	0.122	0.300372	0.29963422	0.30056721	0.32435475
0.004	1060	130	0.1205	0.077	0.189463	0.18419325	0.1890368	0.23873813
0.004	1060	130	0.0723	0.046	0.113611	0.11073822	0.11275982	0.11275982

The developed polynomial degree 3 regressor model is now being deployed to investigate the presence of stenosis in the artery. The model presented in the Appendix was utilized to predict variations in blood flow in the two carotid artery segments based on a given set of features (exposures). Contour plots of AvgVel (CCA velocity) as a function of Out1vel (ICA) and Out2vel (ECA) were generated using MATHEMATICA software, as shown in Figures 7 to 9. These graphs indicate that the maximum velocity, which represents that of a healthy individual, falls within the range of 30 – 40 cm/s. These findings align with the values reported by the authors in reference [21].

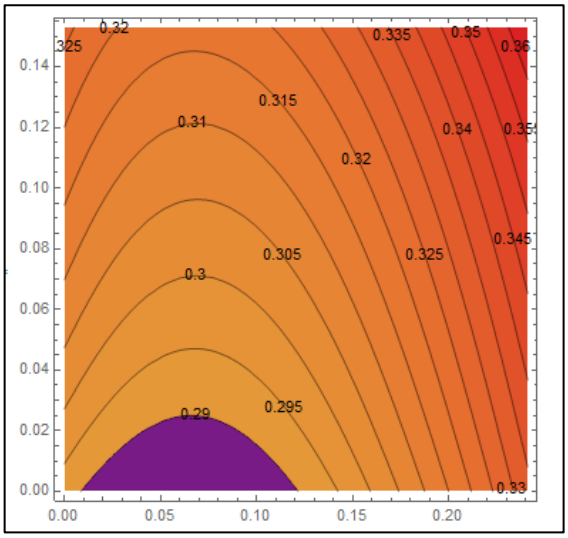

Figure 7. Contour plot for AvgVel (m/s) as a function of Out1vel & Out2vel when Vis=0.004, D=1050, and InPr=120

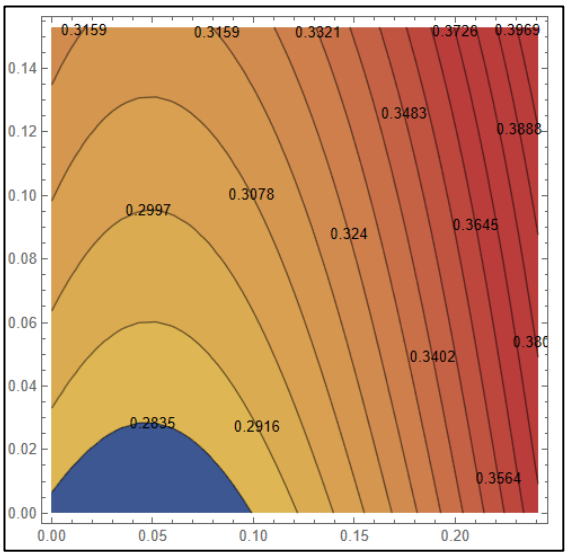

Figure 8. Contour plot for AvgVel (m/s) as a function of Out1vel & Out2vel when Vis=0.004, D=1050, and InPr=125

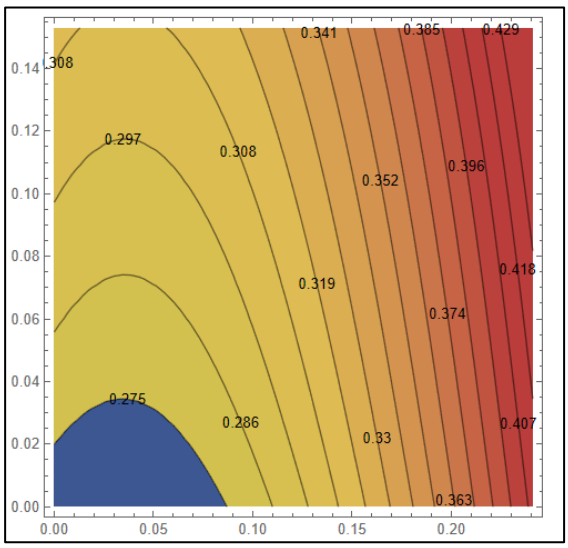

Figure 9. Contour plot for AvgVel (m/s) as a function of Out1vel & Out2vel when Vis=0.004, D=1050, and InPr=130

Based on Figure 7, further observations can be noted. When the blood pressure is 120mm Hg, the normal (healthy) CCA velocity range predicted by the ML model falls between 33.5 and 36 cm/s. If the blood velocity in CCA is 33 cm/s, there is a 2.5% decrease in the velocity of ICA, which reduces from 0.241 m/s to 0.235 m/s. Additionally, a decrease in CCA velocity results in reduced blood velocity in both ICA and ECA. For instance, if the CCA velocity is 31cm/s, the ML model predicts a reduction in blood velocity of up to 69% in ICA and 80% in ECA, suggesting the potential presence of stenosis in both segments. Similar interpretations can be drawn from Figures 8 & 9. However, it is important to note that these observations are based on a numerical model of an idealized carotid artery developed with certain assumptions, as mentioned in the methodology section of this paper. Therefore, further investigation is required to validate the results. Nonetheless, this work introduces a novel approach to assessing the potential prevalence of stenosis in the human carotid artery using vital indicators such as blood pressure and CCA velocity. This approach offers a less expensive and non-invasive alternative compared to other procedures mentioned in reference [5].

Limitations of this study:

The model presented in this paper specifically pertains to an idealized carotid artery. In order to achieve numerical tractability, certain assumptions have been made, including the assumption that the blood behaves as a Newtonian fluid, the arterial segments are circular elastic pipes with a constant radius, and the blood flow is steady.

5. Conclusions

The present work focused on developing a numerical model for the blood flow in the carotid artery and simulating it under different medical conditions arising from stenotic or atherosclerotic situations. Appropriate ML models were developed for the simulated data and used to predict the blood flow characteristics in the artery. By utilizing the quantile loss function, the prediction uncertainty was measured, revealing that the Random Forest Regressor performed as the best model, followed by the polynomial regression of degree 3 as the second best. Through the implementation of the random forest quantile regressor model, prediction intervals were computed for the average blood velocity in CCA. A novel approach was discussed to identify the prevalence of stenosis in ICA and CCA, utilizing vital indicators such as blood pressure and CCA velocity. It should be noted that certain limitations exist in this study concerning the assumptions made to ensure numerical tractability for the problem. Therefore, the future scope of our work involves extending it to a patient-specific model to enhance its reliability as a reference source for the medical fraternity.

Acknowledgments

MTC authors would like to thank the Ministry of Defence, Muscat, Oman, for providing financial support for executing this project MTC/F/M1/MoD/01/23.

Authors would like to express their sincere gratitude to the reviewers for their valuable insights, constructive comments, and thorough evaluation of this research manuscript. Their expertise and feedback have significantly contributed to improving the quality and clarity of the study.

Conflict of interest

Declare conflicts of interest or state “There is no conflict of interest for this study”.

Appendix

```
AvgVel[Vis, Den, InPr, Out1vel, Out2vel]
```

```
 0.265409` + 0.078564` (-0.0022750283` + 0.00022381376682196116` (-1061.666667` + Den)
- 0.0001739537517079976` (-1061.666667` + Den)^2 - 0.0002993633491530194` (-1061.666667` + Den)^3
+ 0.00004733892608231857` (-112.22222` + InPr) +
0.0010605483376810376` (-1061.666667` + Den) (-112.22222` + InPr) -
0.00002926853427527918` (-1061.666667` + Den)^2 (-112.22222` + InPr) -
0.00011584581222704803` (-112.22222` + InPr)^2 -
2.65533629192115` *10^-6 (-1061.666667` + Den) (-112.22222` + InPr)^2 +
6.75610566985723` *10^-6 (-112.22222` + InPr)^3 -
1.8627033751462465` (-0.17356` + Out1vel) +
0.004576007361036121` (-1061.666667` + Den) (-0.17356` + Out1vel) +
0.007451128321580791` (-1061.666667` + Den)^2 (-0.17356` + Out1vel) +
0.0016609074228502178` (-112.22222` + InPr) (-0.17356` + Out1vel) -
0.03905508812992131` (-1061.666667` + Den) (-112.22222` + InPr) (-0.17356` + Out1vel) +
0.00485949913843544` (-112.22222` + InPr)^2 (-0.17356` + Out1vel) -
0.11547011867167774` (-0.17356` + Out1vel)^2 + 0.02821675144718525` (-1061.666667` + Den) (-0.17356` + Out1vel)^2 +
2.0686448767661743` (-112.22222` + InPr) (-0.17356` + Out1vel)^2 -
31.337060662417937` (-0.17356` + Out1vel)^3 +
2.732267903950571` (-0.1102` + Out2vel) + 0.03404275083163098` (-1061.666667` + Den) (-0.1102` + Out2vel) +
0.0027749030173769623` (-1061.666667` + Den)^2 (-0.1102` + Out2vel) +
0.04328394858196776` (-112.22222` + InPr) (-0.1102` + Out2vel) -
20.490957022893234` (-112.22222` + InPr) (-0.004222` + Vis) +
18.15090721567112` (-1061.666667` + Den) (-112.22222` + InPr) (-0.004222` + Vis) -
0.03625887596798143` (-112.22222` + InPr)^2 (-0.004222` + Vis) +
82.87927764043116` (-0.17356` + Out1vel) (-0.004222` + Vis) +
0.840759008506112` (-1061.666667` + Den) (-0.17356` + Out1vel) (-0.004222` + Vis) -
755.0195150460552` (-112.22222` + InPr) (-0.17356` + Out1vel) (-0.004222` + Vis) +
759.2899113664167` (-0.17356` + Out1vel)^2 (-0.004222` + Vis) +
15.551147953584664` (-0.1102` + Out2vel) (-0.004222` + Vis) +
88.6694445323135` (-1061.666667` + Den) (-0.1102` + Out2vel) (-0.004222` + Vis) -
32.184121934536016` (-112.22222` + InPr) (-0.1102` + Out2vel) (-0.004222` + Vis) +
648.1344731201417` (-0.17356` + Out1vel) (-0.1102` + Out2vel) (-0.004222` + Vis) -
14756.086901041232` (-0.1102` + Out2vel)^2 (-0.004222` + Vis) +
8246.650374304616` (-0.004222` + Vis)^2 + 6823.624803736039` (-1061.666667` + Den) (-0.004222` + Vis)^2 +
350868.2941984279` (-112.22222` + InPr) (-0.004222` + Vis)^2 +
318305.1010880145` (-0.17356` + Out1vel) (-0.004222` + Vis)^2 -
1.6730246957202835`*^7 (-0.1102` + Out2vel) (-0.004222` + Vis)^2 +
0.00047500789762314384` (-0.004222` + Vis)^3)
```

References

[1] Zakeri, A.; Jadhav, A.P.; Sullenger, B. A.; et al. Ischemic stroke in COVID-19-positive patients: an overview of SARS-CoV-2 and thrombotic mechanisms for the neurointerventional. *Journal of Neuro Interventional Surgery*. 2021, 13, 202-206.

[2] Qureshi, A.I.; Baskett, W.I.; Huang, W.; et al. Acute Ischemic Stroke and COVID-19: An Analysis of 27,676 Patients. *Stroke*. 2021, 52(3).

[3] Belani, P.; Schefflein, J.; Kihira, S.; et al. COVID-19 Is an Independent Risk Factor for Acute Ischemic Stroke. *AJNR Am J Neuroradiol*. 2020, 41(8), 1361-1364.

[4] Uematsu, S.; Yang, A.; Preziosi, T.J.; Toung, T. J. Measurement of carotid blood flow in man and its clinical application. *Stroke*. 1983,14(2), 2.

[5] Ackroyd, N.; Gill, R.; Griffiths, K.; et al. Appleberg, M. Quantitative common carotid artery blood flow: Prediction of internal carotid artery stenosis. *Journal of Vascular Surgery*. 1986, 3(6),846-853.

[6] Oktar, S.O.; Yücel, C.; Karaosmanoglu, D.; et al. Blood-Flow Volume Quantification in Internal Carotid and Vertebral Arteries: Comparison of 3 Different Ultrasound Techniques with Phase-Contrast MR Imaging. *American Journal of Neuroradiology*. 2006, 27(2), 363-369.

[7] Srinivasacharya, D.; Madhava Rao, G. Mathematical Model for Blood Flow Through a Bifurcated Artery Using Couple Stress Fluid. *Mathematical Biosciences*. 2016, 278, 37-47.

[8] Chakravarty; Sen, S. A Mathematical Model of Blood flow in a Catheterized Artery with Stenosis. *Journal of Mechanics in Medicine and Biology*. 2009, 09(3), 377-410.

[9] Ruchi Agarwal, V. K.; Katiyar; Prabhakar Pradhan. A Mathematical Modeling of Pulsatile Flow in Carotid Artery Bifurcation. *International Journal of Engineering Science*. 2008, 46 (11),1147-1156.

[10] Muraca, E.; Gramigna; Vera Fragomeni; et al. Mathematical Model of Blood Flow in Carotid Bifurcation. *Excerpt from the Proceedings of the COMSOL Conference*.2009, Milan.

[11] Perktold, K.; Rappitsch, G. Mathematical modeling of arterial blood flow and correlation to atherosclerosis. *Technol Health Care*. 1995, 3(3), 139-151.

[12] Onwuzu.; Sobechukwu.; Ugwu, A.; Mbah; Godwin; Onwuzu; Ifunanya. Measuring Wall Shear Stress Distribution in the Carotid Artery in an African Population: Computational Fluid Dynamics versus Ultrasound Doppler Velocimetry. *Radiography*. 2020, DOI: 10.1016/j.radi.2020.11.018.

[13] Abdi, M.; Navidbakhsh, M.; Razmkon, A. A Lumped Parameter Method to Calculate the Effect of Internal Carotid Artery Occlusion on Anterior Cerebral Artery Pressure Waveform. *J Biomed Phys Eng*. 2016, 6(1), 33-40.

[14] Sadraie, S.; Abdi, M.; Navidbakhsh, M.; et al. Modeling the heartbeat, circle of Willis, and related cerebral stenosis using an equivalent electronic circuit. *Biomedical Engineering: Applications, Basis and Communications*. 2014, 26.

[15] Onaizah, O.; Poepping, T. L.; Zamir, M. A model of blood supply to the brain via the carotid arteries: Effects of obstructive vs. sclerotic changes. *Medical Engineering & Physics*. 2017, 49, 121-130.

[16] Singh, N.; Moody, A. R.; Roifman, I.; et al. Advanced MRI for carotid plaque imaging. *Int J Cardiovasc Imaging*. 2016, 32(1), 83-89.

[17] Marshall, I.; Papathanasopoulou, P.; Wartolowska, K. Carotid flow rates and flow division at the bifurcation in healthy volunteers. *Physiol Meas*. 2004, 25(3), 691-697.

[18] Fojas; Jhalique Jane; de Leon; et al. Carotid Artery Modeling Using the Navier-Stokes Equations for an Incompressible, Newtonian and Axisymmetric Flow. *APCBEE Procedia*. 2013, 7, 86–92.

[19] Shigehiko Ogoh; Takuro Washio; Julian, F.R.; et al. Gravitational effects on intracranial pressure and blood flow regulation in young men: a potential shunting role for the external carotid artery. *Journal of Applied Physiology*. 2020, 129(4), 901-908.

[20] Kamenskiy, A. V.; MacTaggart, J. N.; Pipinos II; et al. Three-dimensional geometry of the human carotid artery. *J Biomech Eng*. 2012, 134(6).

[21] Lee, W. General principles of carotid Doppler ultrasonography. *Ultrasonography*. 2014, 33(1), 11-17.

[22] Marcel Hamda Soulouknga.; Hasan Huseyin Coban.; Ruben Zieba Falama.; et al. Noel Djongyang. Comparison of Different Models to Estimate Global Solar Irradiation in the Sudanese Zone of Chad. *Jurnal Elektronika dan Telekomunikasi*. 2022, 22 (2).

[23] Xi Chen; Weidong Liu; Xiaojun Mao; Zhuoyi Yang. Distributed High-dimensional Regression Under a Quantile Loss Function. *Journal of Machine Learning Research*. 2020, 21, 1-43.