*Article*

Machine Learning Based Extraction of Boundary Conditions from Doppler Echo Images for Patient Specific Coarctation of the Aorta: Computational Fluid Dynamics Study

Vincent Milimo Masilokwa Punabantu [1], Malebogo Ngoepe [1,*], Amit Kumar Mishra [2,*], Thomas Aldersley [3], John Lawrenson [3] and Liesl Zühlke [4]

1 Mechanical Engineering, University of Cape Town, Cape Town 7700, South Africa; pnbvin001@myuct.ac.za
2 National Spectrum Centre, Aberystwyth University, Aberystwyth SY23 3AH, UK
3 The Children's Heart Disease Research Unit, Red Cross Children's Hospital, Cape Town 7700, South Africa
4 Extramural Research and Internal Portfolio, South African Medical Research Council, Cape Town 7505, South Africa
* Correspondence: malebogo.ngoepe@uct.ac.za (M.N.); akmishra@ieee.org (A.K.M.)

Abstract: Patient-specific computational fluid dynamics (CFD) studies on coarctation of the aorta (CoA) in resource-constrained settings are limited by the available imaging modalities for geometry and velocity data acquisition. Doppler echocardiography is considered a suitable velocity acquisition modality due to its low cost and safety. This study aims to investigate the application of classical machine learning (ML) methods to create an adequate and robust approach to obtain boundary conditions (BCs) from Doppler echocardiography images for haemodynamic modelling using CFD. Our proposed approach combines ML and CFD to model haemodynamic flow within the region of interest. The key feature of the approach is the use of ML models to calibrate the inlet and outlet BCs of the CFD model. In the ML model, patient heart rate served as the crucial input variable due to its temporal variation across the measured vessels. ANSYS Fluent was used for the CFD component of the study, whilst the Scikit-learn Python library was used for the ML component. We validated our approach against a real clinical case of severe CoA before intervention. The maximum coarctation velocity of our simulations was compared to the measured maximum coarctation velocity obtained from the patient whose geometry was used within the study. Of the 5 ML models used to obtain BCs, the top model was within 5% of the maximum measured coarctation velocity. The framework demonstrated that it was capable of taking into account variations in the patient's heart rate between measurements. Therefore, it allowed for the calculation of BCs that were physiologically realistic when the measurements across each vessel were scaled to the same heart rate while providing a reasonably accurate solution.

Keywords: machine learning; computational fluid dynamics; coarctation of the aorta; boundary conditions

1. Introduction

Congenital heart disease (CHD) is the most common birth defect, with a global prevalence of approximately 9 per 1000 births. Coarctation of the aorta (CoA) is one of the most common forms of CHD, comprising approximately 7% of all CHDs [1]. CoA is defined as the narrowing of a point within the aorta, usually at the isthmus, which may be discrete or elongated and occurs on a spectrum of varying severity. This narrowing leads to increased resistance across the vessel, thus causing upper body hypertension and reduced blood supply to downstream vessels and body parts. The resultant pressure difference across the coarctation is used by clinicians as a means of diagnosis and is referred to as the pressure gradient. A peak systolic pressure difference or drop of greater than 20 mmHg or

imaging diagnosis after hypertension warrants intervention [2,3]. Common treatments for CoA are balloon angioplasty, end-to-end anastomosis (REEA) resection, and stenting in older patients [2]. The general objective of treatment is to alleviate the pressure difference and restore normal flow throughout the coarctation.

Recoarctation can occur after intervention, and the haemodynamic environment after surgery is a strong driver of this occurrence [4]. Computational fluid dynamics (CFD) has been identified as a potential avenue for studying patient haemodynamics and identifying key flow metrics such as pressure, velocity, and wall shear stress with high temporal (time) and spatial resolution [5]. Another key benefit is the non-invasive nature of the technique. Furthermore, with the design of patient-specific CFD pipelines, flow modelling can be specific to an individual patient. However, the incorporation of CFD into clinical workflows has been limited due to challenges in patient data acquisition for modelling and the skill set required to perform the modelling [6]. Thus, most applications of the technique have resided in research.

In April 2017, the Partnerships in Congenital Heart Disease (PROTEA) project was started, which aims to establish a densely phenotyped and genotyped congenital heart disease (CHD) cohort for southern Africa [7]. The project has four core aims, with the third being to study repaired tetralogy of Fallot and coarctation of the aorta using computational fluid dynamics in order to demonstrate its potential to assist in the clinical assessment of CHD, including long-term prediction of growth and remodeling from local blood flow [7]. Swanson et al. [5] conducted a proof-of-concept study in response to the third aim. They developed an open-source, patient-specific CFD modelling pipeline using data obtained from a single CoA patient from the PROTEA project cohort. They used CTA for geometry acquisition and Doppler TTE for velocity acquisition [5]. These two modalities were used, as they were both technically feasible and easily available. CTA and other modalities such as 4D-MRI and contrast-enhanced CT have been used for geometry acquisition in previous studies [8–11], while boundary conditions (BCs) have been prescribed by either direct measurement from imaging modalities such as MRI-based techniques and Doppler echocardiography or by values obtained from the literature [9–11]. Through their study, the authors highlighted the suitability of Doppler echocardiography-derived BCs and presented the case for further study of its implementation. PC-MRI has been favoured within the literature [12–17]; however, it can be costly and difficult to implement for patients within the pediatric population.

Swanson et al. [18] highlighted the current challenges of using Doppler echocardiography to prescribe BCs in a CFD study. One of the main concerns was variations in patient heart rate across different measurements, which is an inherent limitation of the modality. This discrepancy arises because measurements at each spatial point must be taken one at a time. As a result, the velocity profiles at the inlet and outlets do not have the same heart rate. Furthermore, mass conservation was not conserved, as the sum of flow at the outlets was greater than at the inlet. To remedy the different heart rates across each measured vessel, the velocity profiles were scaled in time to a common period using a heart rate of 120 BPM. It was then observed that there was a phase mismatch between peak flow at the supra-aortic branches, since peak flow at the left common carotid artery and left subclavian artery led when compared to the ascending aorta. This seemed unrealistic due to the capacitive effect of artery distensibility. This was thought to be a result of the variation in heart rate between the measured vessels [18]. Four different adjustment schemes were tested to prescribe BCs by correcting for either mass conservation, phase mismatch at peak flow, or both [18].

Machine learning (ML) has been used as a tool to improve BCs for CFD studies [19,20]. In their study, Lassila et al. use ML to account for the variability in flow BCs between subjects to model intracranial aneurysms. They combine a log-linear mixed effects model with a Gaussian process model. The model was trained on carotid ultrasound measurements and demographic data from a cohort of 103 elderly volunteers [19]. In Véras et al. [20], ML was used to define inlet BCs in order to simulate draft tube flows. In their approach, they

used an artificial neural network (ANN) to predict the inlet BCs from downstream flow results and trained the model using data generated by numerical simulation.

Our work develops a framework that uses an ML-based approach to improve the fidelity of Doppler TTE-derived BCs. This is applied to a CFD model for CoA using the data collected by the PROTEA project. By accounting for changes in the velocity profiles due to changes in heart rate, we hypothesize that more-accurate inlet and outlet BCs can be computed. An ML approach using nearest neighbour- and tree-based methods was selected, as our dataset was relatively small [21]. Our approach focuses on the translation potential of patient-specific CFD pipelines into real clinical settings by using available imaging modalities.

Our work has two major novelties. First, we use ML to account for changes in heart rate in Doppler TTE measurements in order to prescribe inlet and outlet BCs that result in an accurate flow solution. Second, this is achieved using a small and incomplete dataset with a cohort size of five patients.

The paper is organised as follows: It begins by describing the development of the ML-CFD pipeline. The ML subsection describes model training, and the subsequent CFD section focuses on the modelling approach and validation process. The results, discussion, and conclusion are then presented.

2. Materials and Methods

In this section, we outline the methods used to develop our approach. Figure 1 provides an overview of the full pipeline. Beginning with data pre-processing, key features are extracted from the Doppler TTE images to build the ML training dataset. This dataset is split into the respective inputs and outputs, as shown in the ML Model Training box, which then undergo a feature engineering process. This converts the data to a format that is compatible with the ML model. Once the regression model is trained, the evaluation inputs are specified, and BCs are obtained. These are then input into the CFD solver along with a meshed patient geometry. The CFD solver then outputs the velocity and pressure results.
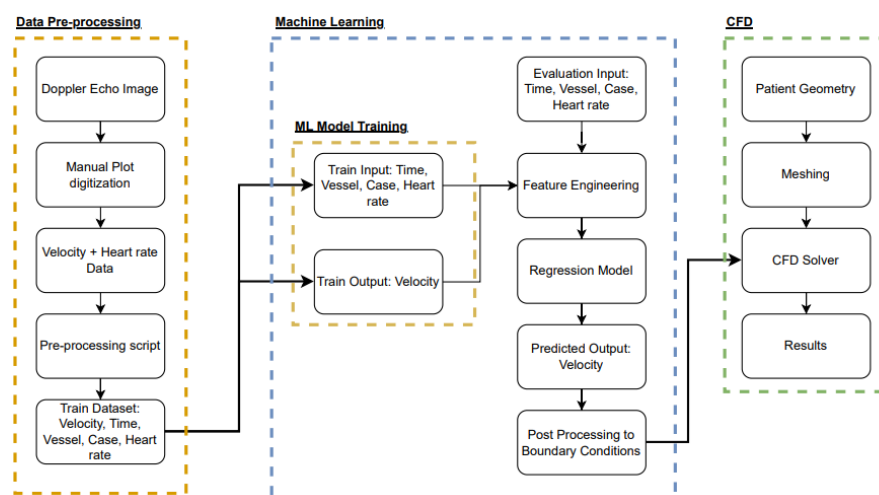

Figure 1. Flow diagram representation of the full framework used to develop the coupled ML-CFD pipeline, for modelling patient-specific flow using Doppler TTE as the velocity data acquisition modality. The pipeline starts with the Data Pre-processing box in yellow, is where key features are extracted from the Doppler TTE images to build the ML training dataset. The dataset is split into the respective inputs and outputs, shown in the ML Model Training sub-box in gold. The Machine Learning box, in blue, then shows the process through which the ML models are trained and evaluated. Inputs are converted to a format that can be input to the models via a feature engineering process, and the predicted output velocity is then processed to obtain the boundary conditions for CFD modelling. The CFD box, in green, shows the process used to obtain the the CFD result, which, in this study, is the patient-specific velocity field.

2.1. Data Pipeline and Clinical Data

In this study, Doppler TTE measurements were obtained from 5 unique patients. However, the data for all patients were incomplete to varying degrees, which posed a significant challenge. A total of 4 out of the 5 patients only had Doppler TTE measurements across the coarctation. Additionally, information on when the scans were taken with respect to one another and the severity of the stenosis was not available. Only 1 out of the 5 had a more complete dataset: Doppler TTE measurements across the ascending aorta, supra aortic branches (innominate artery, left common carotid artery, and left subclavian artery), coarctation and descending aorta, and a 3D patient-specific geometry prior to intervention. Although, for post-intervention only Doppler TTE measurements were available after the intervention across the ascending aorta, the innominate artery, and the coarctation; these data were previously used in Swanson et al. [5]. Therefore, the pre-intervention geometry and Doppler TTE measurements for this patient were used for subsequent CFD modelling for the study, as the dataset was more complete. All clinical data used in the study were provided by Red Cross War Memorial Hospital (RXH).

A data pre-processing pipeline was developed to process the Doppler TTE images and build a dataset that could be used to train the ML models. The velocities were extracted from the Doppler TTE images using a free online plot digitiser [22]. This was a manual process carried out by the researcher. Points were first set along the axes of the velocity graph in order to convert the image to a plot. The values at the designated points were then entered manually, which scaled the plot accordingly. Points were then traced along the velocity profile within the graph so as to obtain the velocity readings. The tracings were taken for velocity profiles that were adequately developed; that is, the distinct shape of the profile was visible and there was minimal noise across the profile. Tracings during diastole were set to zero. The coordinates of the tracings were then saved in a csv file. This process was repeated for all Doppler TTE images in the dataset from [5] and for the 4 additional coarctation images from RXH.

The same plot digitiser was used to determine the number of beats per second, allowing the heart rate of each profile to be determined. The clinicians noted that measuring heart rate directly from the image by measuring the peak-to-peak difference of the cardiac waveforms was a more reliable approach compared to using the value provided in the bottom right-hand corner of the image, as shown in Figure 2. Thus, in images in which the electrocardiogram (ECG) was present—for example, Figure 2a–c—the heart rate was calculated by measuring the time interval between successive peaks of the QRS complex. However, in images such as Figure 2d, in which the ECG was not present, the time interval between the successive peaks of the velocity profiles was used to calculate the heart rate. In both cases, the time intervals between 4–5 successive peaks were measured, and the average was calculated, which was then used as the associated heart rate. We note that this approach introduces human error at the plot digitization stage; however, calculation of the heart rates was then conducted via an in-house Python script. The sampling rate at which the velocity profiles were measured was not available, as this metric did not fall within the clinical protocol when the data were collected. As a result, methods such as using a Fourier transform were not an option. As such, it would not have been possible to find the right scaling for the y-axis.

The csv files from the plot digitiser were processed using an in-house Python script to build the dataset for ML model training. First, the script zeroed the velocity profiles so that they started from 0 and inverted the coarctation velocity values to the positive y-axis. Second, any non-zero values during diastole were set to 0, and the velocity profile was interpolated for 200 or 350 steps within the time interval, depending on the length of the time interval and number of cycles. Last, the associated vessel name (ascending aorta, innominate artery, left common carotid artery, left subclavian artery, or coarctation and descending aorta), case (pre- or post-intervention), and heart rate were captured. A training dataset was then generated from this information in the form of a table.
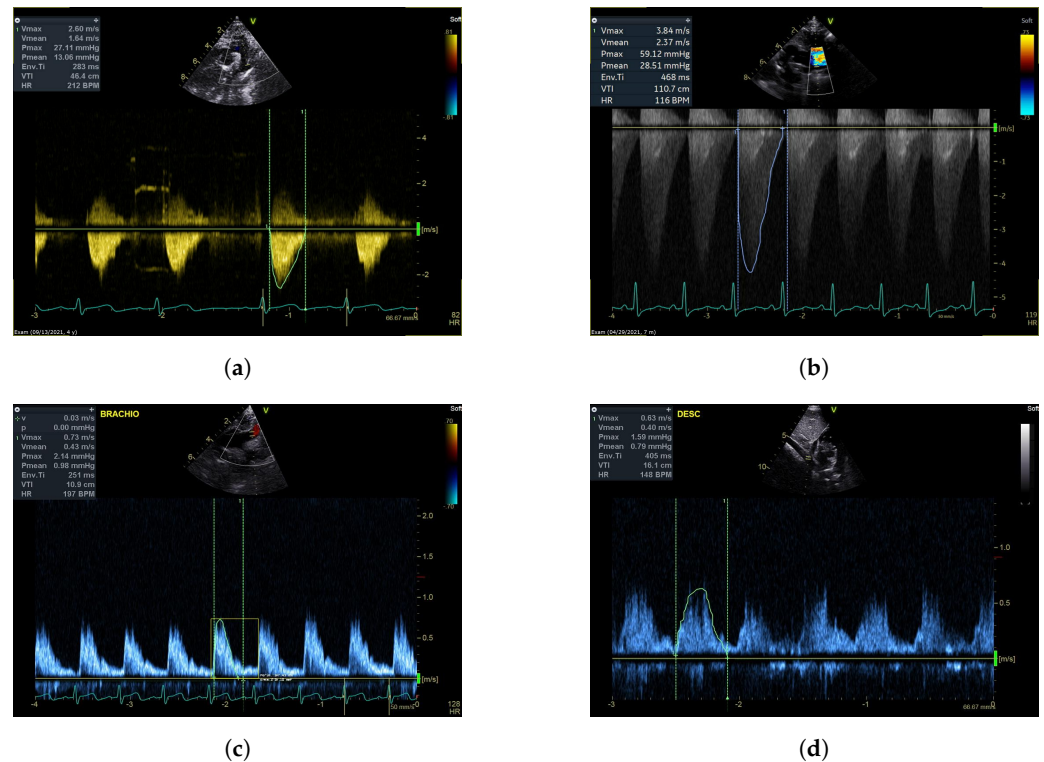
Figure 2. Sample Doppler TTE images that were used within the study. The images in (**a**–**d**) show a coarctation measurement post intervention, coarctation measurement pre-intervention, innominate artery measurement post intervention, and descending aorta measurement pre-intervention.

2.2. Machine Learning Pipeline

Our training dataset consisted of 3 numerical values: time in seconds, area-averaged velocity across the vessel cross section in metres per second, and heart rate in beats per second. The case and vessel features/columns are comprised of categorical or non-numeric entries. Additional pre-processing steps were employed before training the ML model. First, the categorical entries case and Vessel were converted to numerical entries using the Scikit-learn label encoding function [23], as the models within the Scikit-learns library require numerical inputs. For the numeric entries (particularly the velocity values), it was found that the distribution of the data was skewed to the right and contained several outliers. Data points that were set to 0 during diastole were partially responsible for the skewness of the distribution, resulting in several outliers in a box-and-whisker plot. High velocity values at the tail end were due to the coarctation. This would contribute to poor performance for models that are sensitive to outliers, which could lead to higher errors in ML model predictions. However, the non-zero outliers were biologically plausible and not the result of noise or measurement errors. Numpy's log1p function [24], a Log transform, was applied to the velocity entries in order to reduce the effect of the outliers during training and improve ML model performance while preserving the outlier values within training.

The class of ML problem posed is a regression problem, as the objective of the ML model is to predict a continuous value given a set of discrete inputs. Recalling that the main challenge faced was the discontinuity in the time measured between vessels, the heart rate at the current measured vessel varied from the previous vessel. As noted earlier, changes in heart rate affect the velocity profile across the measured vessel, which indicates a present relationship between heart rate and velocity. For the CFD model, a mass flow rate is prescribed at the inlets and outlets of the region of interest and is a function of velocity. Therefore, the patient's heart rate was considered the key input that could relate the time-dependent in vivo flow conditions with the velocity, which was then used to calculate the BCs. Thus, time (time interval over which the velocity was measured), velocity

(area-averaged velocity across the measured vessel site), and heart rate (average heart rate across 3–4 cardiac cycles) were the identified key numerical inputs. However, these values changed depending on the vessel and the case (whether an intervention had occurred or not). Therefore, these two parameters were included as categorical variables. It is noted that although the name of the vessel, which represents the anatomical position, can be better represented by geometric parameters that could convey more insightful information, in our case, this was not possible. Geometric information such as the vessel diameter or length were inconsistent within the given data, whilst the vessel name was consistently given.

For this study, the inputs to the ML models were time in seconds, the name of the vessel (ascending aorta, innominate artery, left common carotid artery, left subclavian artery, or coarctation and descending aorta), case (pre- or post-intervention), and heart rate in beats per second. Each of these inputs was scaled using the Scikit-learn standard scaler function [23]. This was to limit error in models sensitive to scaling, such as linear, K-nearest neighbour, and support vector machine regressors. The evaluation metric used to assess model performance was the root mean square error (*RMSE*), shown in Equation (1), a common metric for regression problems that calculates the Euclidian norm between the predicted and actual value [21].

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{1}$$

where n is the number of data points and y_i is the actual (true) output value of the ith data point, and $\hat{y}_i$ is the predicted value for the ith data point.

In addition, the coefficient of determination (R^2), shown in Equation (2), was used as the second evaluation metric due to its use in other similar studies [25,26]. R^2 is the proportion of the variation in the output that is explained by the input of the predictive model. A score of 1.0 indicates a perfect prediction, and it may be negative because a model can be arbitrarily worse, according to Scikit-learn's r2_score documentation [23].

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{2}$$

where $\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$ is the mean of the actual output value across all data points.

The models trained in this study were, namely, the linear regression model, which provided the baseline, support vector machine regressor, random forest regressor, gradient boosted regressor, K-nearest neighbour regressor (k-NN), and a voting ensemble method. The training dataset was divided into train, validation, and test with a 60% (2190 × 5), 20% (730 × 5), and 20% (730 × 5) split, respectively. The parameters for the K-nearest neighbour regressor, random forest regressor, gradient-boosted regressor, LightGBM regressor [27], and support vector machine regressor were optimised in the train set and validated against the validation set [23]. Finally, all models were fit to the combined training and validation set and were evaluated on the test set.

2.3. CFD Pipeline

ANSYS Fluent (ANSYS Academic Research [Fluent], release 2020R2) was the commercial software package used for the CFD study. The mesh for the unrepaired geometry from Swanson et al. [5] was used for this portion of the study. Figure 3 shows the unrepaired geometry. The mesh constituted 2 million tetrahedral elements and five prism boundary layers meshed using ANSYS ICEM-CFD. The number of elements used was greater than the minimum number of elements required for grid independence. For further details on the grid independence study, please see [5]. The incompressible Navier-Stokes equations were solved using the finite volume method. While pressure was calculated using the PRESTO discretization scheme, momentum was determined using a second-order upwind scheme. Pressure and velocity were coupled using the PISO scheme. For temporal discretization, a bounded second-order implicit scheme with a time step of 0.05 ms was used. Flow was

modelled as laminar, and the blood was assumed to be Newtonian, with a density of 1060 kg/m^3 and viscosity of 0.004 kg/(m s). The vessel wall was modelled as rigid, and a no-slip boundary condition was applied. Convergence criteria of 10^{-3} were used (ANSYS Academic Research [Fluent], release 2020R2, 37.21.1, Judging Convergence, ANSYS, Inc., Canonsburg, PA, USA). Simulations were run as time-dependent with constant boundary conditions for 1 s since flow after the coarctation was unsteady, as was previously done in [5]. This allowed start-up pressure wave reflection and propagation effects to dampen out [18]. All simulations were run on a single cluster node with 40 Intel CPU cores and 386 GB of RAM, on the University of Cape Town's High Performance Computing Cluster (HPC). Only peak systole was simulated and not the full cardiac cycle, as each simulation, on average, took approximately 50 h to complete and simulating the entire cardiac cycle for 24 simulations was deemed too computationally expensive.
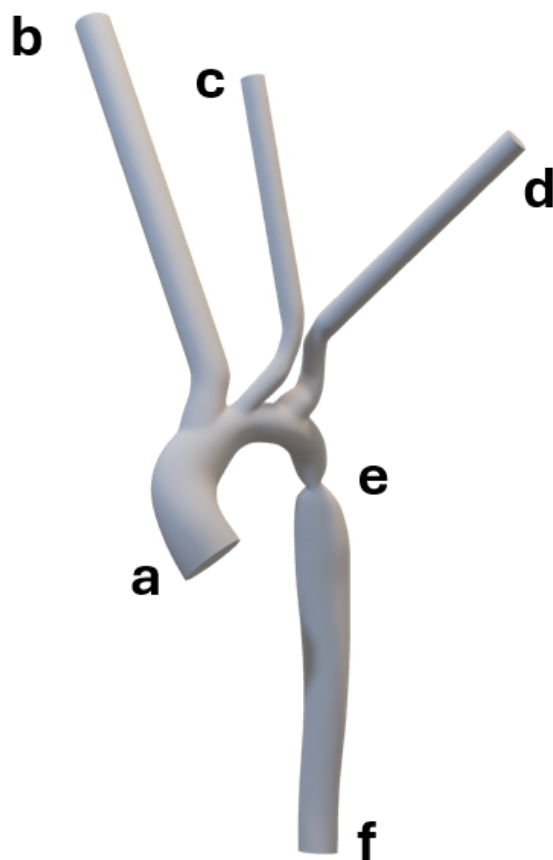

Figure 3. The image shows the CoA geometry for patient 1, and the labels a to f are the ascending aorta, innominate artery, left common carotid artery, left subclavian artery, coarctation, and descending aorta, respectively. The inlet was prescribed at a, and the outlets at b, c, d, and f, in that order.

Five ML models with the least *RMSE* and highest R^2 values were selected to generate BCs for the CFD modelling process. A time interval of 0.442 s was used, which corresponds to 135.6 BPM. This was the calculated heart rate at which the measurement across the coarctation was taken. The vessel names and case (pre-intervention) were also input to each ML model. The CFD mass flow rate input profiles were calculated from the velocity profiles output by each ML model using the stated density of blood and the vessel area. The set of BCs was adapted from [5,28,29] in order to assess BCs that were commonly used in the literature whilst also taking into account the limited data constraint due to the use of Doppler TTE and the absence of pressure data. Each set of BCs had to be simulated for the not-adjusted case (max velocity values obtained from Doppler TTE measurements with no additional processing) and

for the outputs of the five selected ML models. Therefore, the computational cost of running the simulations was an important factor.

Table 1 shows the set of BCs specified at the inlet and outlets of the patient geometry for each ML model and the not-adjusted case. The outlets consisted of the arch branches (AoB) (innominate artery (outlet 1); left common carotid artery, (LCCA, outlet 2); left subclavian artery, (LSA, outlet 3)), and descending aorta (DAo, outlet 4). The inlet was defined at the ascending aorta. Four different combinations of BCs were explored. For these combinations, the inlet BC was kept constant, while the outlet BCs varied, as these BCs had a greater impact on the flow within the coarctation [29]. BC 1, which specified zero pressure at all outlets, has been shown to result in the least-accurate flow solution compared to other types of outlet BCs [28]. Thus, BC 1 was primarily included as a benchmark case to provide the baseline performance of the ML-CFD pipeline. The zero-pressure BC with target mass flow adds a specified mass flow rate as a constraint that must be met when calculating pressure at the outlet. This provided an alternate way for the pressure to be prescribed at the outlet, as the mass flow rate can be seen as an alternate form of prescribing velocity at the outlets. This can be seen in BC 3 and BC 4.

Given that Madhavan et al. showed that the effects of the inlet boundary condition are only significant up to 2 inlet diameters distal to the inlet patch [29], a plug flow velocity profile was assumed for the inlet. The inlet is sufficiently far from the coarctation, which is the main region of interest. BCs at the outlets are registered to affect up to 5 diameters distal to the outlet [29]. Due to the limited data availability imposed by using Doppler TTE and the absence of pressure data, pressure-based BCs using the three-element Windkessel model (3-EWM) proved difficult to implement. Instead, outlet BCs, namely zero pressure and mass flow rate, were used, as this was more suitable to implement. BCs derived from the ML model were compared with not-adjusted BCs, which were BCs obtained from Doppler TTE in [5].

Table 1. Boundary conditions: This table contains the 4 sets of boundary conditions (BCs) that were used within the study, namely BC 1–4. Outlets are located at the arch branches (AoB) and the descending aorta (DAo), with the inlet at the ascending aorta. Note that the AoB are considered separate from the DAo. The types of BCs used comprised of mass flow rate [kg/s], zero pressure (static pressure [Pa] set to zero), and zero pressure target mass flow rate (static pressure [Pa] set to 0 initially but constrained to meet a specified mass flow [kg/s] value in subsequent iterations).

BC	Inlet BC	AoB BC	DAo
BC 1	mass flow rate	zero pressure	zero pressure
BC 2	mass flow rate	mass flow rate	zero pressure
BC 3	mass flow rate	zero pressure, target mass flow rate	zero pressure, target mass flow rate
BC 4	mass flow rate	mass flow rate	zero pressure, target mass flow rate

The maximum coarctation velocity obtained from the numerical simulation was compared with the measured value from Doppler TTE. This was achieved by calculating the relative error between the two values, as shown in Equation (3).

$$Relative\ Error = \frac{|v_s - v_m|}{|v_m|} \times 100\% \qquad (3)$$

where v_s is the simulated maximum coarctation velocity from CFD, and v_m is the measured maximum coarctation velocity from Doppler TTE.

As noted in [5], Doppler TTE is capable of reliably measuring blood velocity within the body, and the measurement was taken by an experienced sonographer. Therefore, this measurement was used to validate the numerical solution. Although pressure is the clinically relevant flow parameter, the simplified Bernoulli method used by the device has been shown to be inaccurate. Therefore, velocity was a more reliable flow parameter to validate the framework.

3. Results

This section begins with a brief description of the data processing results followed by the results of the ML model and, finally, the CFD results.

3.1. Data Pipeline and Clinical Data Results

The output of the data processing pipeline is the dataset used to train the ML models of size (3650 × 5). The five columns are titled time in seconds, velocity in metres per second, case (pre- or post intervention), vessel, and heart rate in beats per second. The 3650 data points were obtained from plot digitisation and post-processing of the 11 Doppler TTE images using the in-house Python script. The data points were concatenated into a single tabular dataset. This was then divided into three consisting of sizes (2190 × 5), (730 × 5), and (730 × 5) for training, validation, and testing, respectively.

Table 2 contains the heart rate values in BPM across the indicated vessels for each patient both pre- and post intervention. It can be seen that the heart rate varies both between patients and within patients across each measured vessel. For patient 1, who was unique in having measurements taken across multiple vessels, the heart rate had a standard deviation of 5.38 BPM and 6.68 BPM pre- and post intervention, respectively. The standard deviation of the heart rates across the coarctation between the recorded patients both pre- and post intervention was 12.73 BPM and 44.92 BPM, respectively.

Table 2. The table shows the heart rate values in beats per minute [BPM] across the measured vessels for each patient both pre- and post intervention.

Patient No	Case	Vessel	Heart Rate [BPM]
Patient 1	Pre-intervention	Ascending aorta	131.4
		Innominate Artery	123.6
		Left common carotid artery	138.6
		Left subclavian artery	127.8
		Coarctation	135.6
		Descending Aorta	130.2
Patient 1	Post intervention	Ascending aorta	114.6
		Innominate artery	127.8
		Coarctation	123.0
Patient 2	Pre-intervention	Coarctation	117.6
Patient 3	Post intervention	Coarctation	85.8
Patient 4	Post intervention	Coarctation	85.2
Patient 5	Post intervention	Coarctation	180.6

3.2. ML Model Training Results

Among the models, the k-NN regressor was found to have fitted the data the best, as it had the lowest $RMSE$ and highest R^2. The voting regressor, which is an ensemble of k-NN and random forest regressors, performed second best. Table 3 shows the $RMSE$ and R^2 values for the ML models that were trained and evaluated on the test set. The $RMSE$ was calculated by comparing the ML model-predicted velocity to the measured velocity from the Doppler TTE images for the inputs within the test set. The k-NN, random forest, support vector machine, and LightGBM regressors are the models that underwent hyperparameter tuning. However, it was only the k-NN and random forest regressor that demonstrated some improvement in performance. Minimal improvement in the other models resulted in the use of default settings. The number of nearest neighbours for the k-NN was set to 2. This was determined by determining the $RMSE$ and R^2 values on the validation set for a range of nearest neighbours between 1 and 40. The remaining parameters were set to their default values. For the random forest regressor, an exhaustive search was conducted over a set parameter space using the Scikit-learn GridSearchCV function [23]. The number of

estimators, maximum tree depth, minimum sample split per internal node, and minimum sample split per leaf node were tuned, and cross-validation was set to 5. It should be noted that the *RMSE* and R^2 were not the final metrics with which the ML model performance was evaluated. Instead, the comparison of maximum velocity across the coarctation, obtained from ANSYS Fluent, compared to the ground truth value obtained by Doppler TTE was also used. The support vector machine regressor and linear regression had the poorest results, as they had the highest *RMSE* and lowest R^2 values.

Table 3. ML Model results: This table presents the root mean square error (*RSME*) and coefficient of determination (R^2) values for each machine learning model when evaluated against the test or hold out set. The test set constituted 20% of the original dataset built for the study.

Model	*RMSE* [m/s]	R^2
Linear Regression	0.44076	0.08963
Support Vector Machine Regressor	0.39030	0.28614
Gradient Boosted Regressor	0.28556	0.61789
LightGBM Regressor	0.12076	0.93167
Random Forest Regressor	0.05499	0.98579
Random Forest Regressor Optimized	0.05497	0.98594
Voting Regressor	0.04261	0.99181
K-Nearest Neighbours Regressor	0.03671	0.99369

3.3. CFD Boundary Condition Results

From Figure 4 in the outlet mass flow rate distribution graph, on the left-hand side, we observe that the ML models underestimate the total mass flow at the outlets in comparison to the not-adjusted case. The differences in flow are mainly seen at the innominate artery and the descending aorta, whilst there is agreement at the left common carotid artery and left subclavian artery. The graph on the right-hand side of Figure 4 displays the mass flow rate splits across each outlet. From this graph, we observe that the k-NN regressor distributes more flow to the aortic branches in comparison to the other ML models and the not-adjusted case. Meanwhile, the voting ensemble, random forest, and LightGBM regressors distribute less flow to the aortic branches and more to the descending aorta when compared to the not-adjusted case.
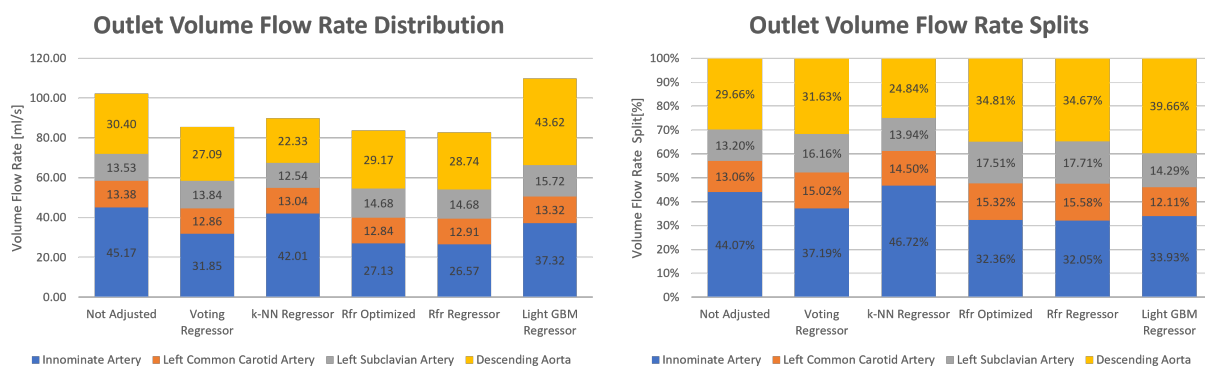

Figure 4. Comparison of the magnitude of flow distributed for the not-adjusted and machine learning (ML) model cases (**left**) and comparison of the mass flow rate splits (**right**). From the graph on the left we observe that the ML models underestimate the total amount of flow at the outlets in comparison to the not-adjusted case, with the exception of the LightGBM regressor. The differences in flow are mainly seen at the innominate artery and the descending aorta, whilst there is agreement at the left common carotid artery and left subclavian artery. This distinction is better seen in the graph on the right, where the flow splits for each case are shown and in which the K-nearest neighbours regressor distributes more flow the innominate artery and less to the descending aorta, which is in contrast to the other ML models.

3.4. Coarctation Velocity Results

Figure 5 presents the CFD model results for each BC type and model used, with Figure 5a illustrating the maximum coarctation velocities and Figure 5b showing the relative error (the relative error between the ground truth maximum coarctation velocity 3.49 [m/s] and the output of the simulation). At first glance, the performance of the ML models varied among BCs 1–4, which are defined in Table 1. Of the ML-derived BCs, we observe that the k-NN model BCs were found to have the least relative error, with BC 2 having the lowest value of 4.88%. Similarly, the k-NN regressor using BC 4 produced the second most accurate result, with a relative error of 7.12%. The remaining ML models generally performed poorly in comparison across all BCs. BC 1 (zero pressure outlet BCs) resulted in the least accurate results, and only the LightGBM showed marginal improvement as compared to the not-adjusted case. From BC 2–4 the voting ensemble, random forest (hyperparameter tuned and default), and LightGBM regressor's as well as the not-adjusted case's performance did not vary significantly. However, the k-NN regressor's performance varied more significantly in comparison.
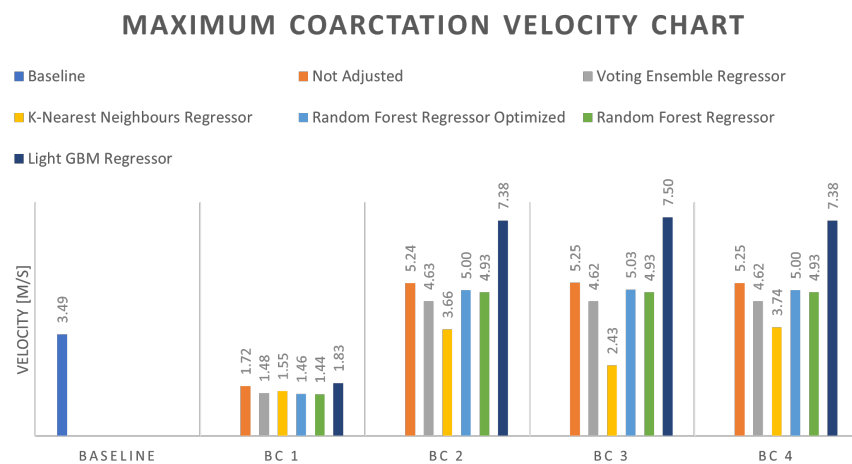
(**a**) CFD Maximum Coarctation Velocity Results
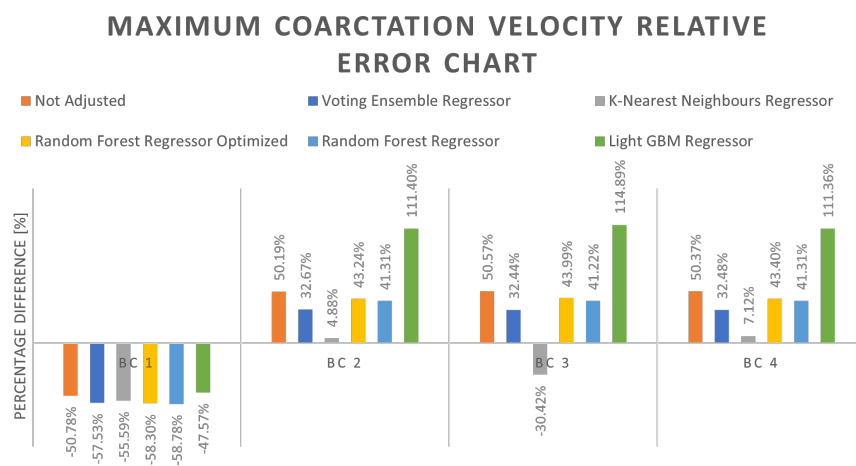
(**b**) CFD Maximum Coarctation Compared to Measured Result

Figure 5. The graph in (**a**) shows the maximum simulated coarctation velocity from the CFD model, arranged according the type of BC and ML model used; (**b**) shows the relative error when the simulated maximum coarctation velocity is compared to the measured result from Doppler echocardiography.

Table 4 presents the CFD pressure drop results for the ML models and not-adjusted case for BC2. BC2 was used, as it resulted in the best results for the maximum coarctation

velocity. The different approaches were evaluated against the measured value from Doppler TTE. We observed that the voting regressor had the best agreement with the baseline, with a relative error of 5.7%. The k-NN was the only model that underestimated the pressure drop; the others overestimated it. However, since the pressure drop from the Doppler TTE was not reliable, these results were not used to quantitatively evaluate the pipeline's performance but provided a clinical indication of the CFD results. When compared with the established standard of a pressure drop greater than 20 mmHg warranting intervention, each model and the not-adjusted case indicated that intervention was required, which was the clinical decision that was determined for patient 1.

Table 4. ΔP Results: This table presents the pressure drop values for each model and the not-adjusted case for BC2. BC2 was used, as it returned in the best results for the maximum coarctation velocity. The different approaches were evaluated against the measured value from Doppler TTE. Equation (3) was used to calculate the error.

	CFD Coarctation ΔP (mmHg)	ΔP Error (%)
Doppler TTE Baseline	**48.65**	-
LightGBM Regressor	128.19	163.5
Not Adjusted	64.36	32.3
Random Forest Regressor Optimized	58.62	20.5
Random Forest Regressor	56.90	17.0
Voting Regressor	51.42	5.7
k-Nearest Neighbours Regressor	35.31	−27.4

Figure 6 shows the pressure and velocity distributions of the flow for the k-NN BC2 case. From the pressure distribution (on the left), we see that the region before the coarctation is at a higher pressure, while the region after the coarctation is at a lower pressure, illustrating a drop in pressure across the coarctation. This result is similar to those found in [5,11]. In addition, the pressure distribution qualitatively shows hypertension in the upper extremity vessels by the high pressure flow through the supra-aortic branches, which is a common symptom of CoA. With regards to the velocity distribution (on the right), we observe that the peak occurs at the coarctation region. Furthermore, within the region after the coarctation, we can observe the helical nature of the flow marked by the areas with higher velocity values. It is also worth noting that recirculation zones are present as well, mainly in the low-velocity areas of the distribution. The velocity flow field also has good agreement with the results found in [5].
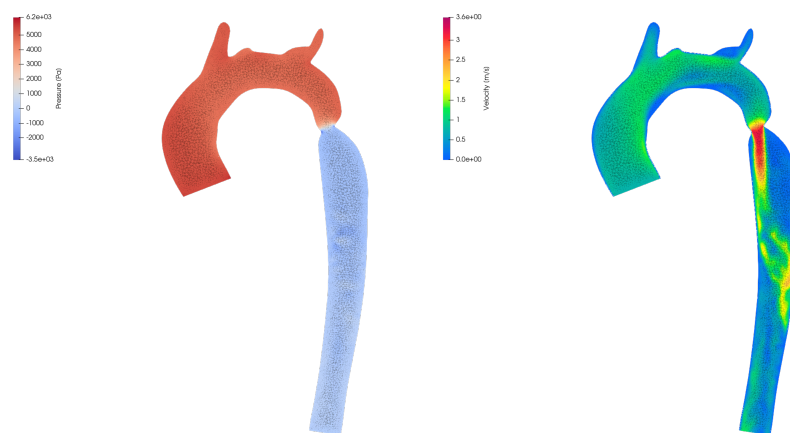

Figure 6. The figure shows the pressure distribution on the (**left**) and velocity distribution on the (**right**) along a plane across the fluid domain for the k-NN BC2 boundary conditions case.

4. Discussion

Numerical modelling offers a means of improving the analysis and understanding of the cardiovascular haemodynamics of CoA through the simulation of flow in complex geometries. However, the inputs required for these methods can prove challenging to obtain. This study aimed to add to the body of work at the intersection of empirically derived patient-specific BCs in CoA, Doppler echocardiography, CFD, and ML [5,19,30].

Our framework is capable of producing velocity results that are accurate while accounting for variations in the patient's heart rate. This is evidenced by the relative errors within 5% for BC 2 and 7% for BC 4 using the k-NN regressor. Furthermore, for BCs 2–4, the ML model-derived BCs have better agreement with the measured maximum coarctation velocity in comparison to the not-adjusted case, except for the LightGBM regressor. This indicates that there is a potential benefit to accounting for the effect of heart rate on the velocity profile; however, further work will be required to conclusively determine this. The CFD simulations were limited to a single case, and only flow at peak systole was simulated. Therefore, it will be key to see if this increased accuracy is consistent across a full cardiac cycle.

The heart rates for each vessel were shown in Table 2. These varied both within and between patients as well as between pre- and post-intervention cases. For patient 1, the standard deviation was 12.73 BPM and 44.92 BPM pre- and post-intervention, respectively. Patient restlessness contributed to a higher BPM during post-intervention echocardiographic imaging and prevented measurement of the three remaining vessels. Regardless, this instance demonstrates that variability in patient flow conditions can be experienced during the measuring process. Lassila et al. also noted this in their study and highlighted the inadequacy of using population-averaged or one-shot measurements as a true indication of a specific patient's flow regime [19]. To account for intra-patient flow variability, they used a lumped parameter model, with changes in heart rate and blood pressure being the key inputs [19]. Their work also demonstrated that the patient's heart rate can be used to account for flow variability. In addition, carotid ultrasound was used for patient velocity data acquisition. Similar to Lassila et al, our work demonstrates the applicability of heart rate and ultrasound to account for flow variability, the difference being that in our case, we make use of ML models for within-patient variability as opposed to a physics-based model. Moreover, our approach makes use of a smaller dataset with a cohort size of 5 as opposed to 103.

Of the ML models used within this study, the k-NN regressor fit the data the best, as it had the lowest $RMSE$ and highest R^2 value of 0.03671 m/s and 0.99369, respectively, when evaluated on the test set as shown in Table 3. In addition, the BCs derived using this model produced the most-accurate CFD result when compared to the measured value of 3.49 m/s from Doppler TTE. However, of the models used, the k-NN model is the simplest, with the exception of the linear regression model, which was used as a base line. This was a surprise, for one would have expected a more complex model such as the support vector machine, random forest, or LightGBM regressor to provide better results. However, it is the case that ML models will perform differently due to factors such as the size of the data, the data's structure, and the tuning of model hyper-parameters. For example, Cai et al. [25] in their study compared k-NN, XGBoost, and multi-layer perceptron (MLP) regressors as finite element method (FEM) surrogate models to estimate material properties from clinical data. They found the k-NN model to have the worst performance, which is in contrast to our case. Note that their dataset was larger compared to this study, with 10,000 samples, and that their data were generated from FEM simulations rather than measurement instruments. Similarly, Cilla et al. [31] and Mehdi et al. [26] used ML models as surrogates for FEM simulations to predict plaque rupture and estimate myocardial stiffness, respectively. Although the ML models were applied to FEM-based biomechanics studies, their performance varied between studies. For example, Mehdi et al. [26] found that the MLP provided the best results followed by the support vector machine and random forest, while Cilla et al. [31] found that the support vector machine, the simpler

model, outperformed the MLP (only these two models were evaluated in Cilla et al. [31]). Meanwhile, in Cai et al. [25], the XGBoost performed best. Although we cannot compare like for like as the nature of the problems between these studies are not equivalent, we still observe that model performance differs across different applications even in the same domain, building FEM surrogates. Therefore, the need to select the model on a case-by-case basis and evaluate a sizable set is supported. As per the law of parsimony, there are cases where the simpler model or approach tends to yield better results.

From Figure 5, we observe that the choice of BCs strongly influences the accuracy of the CFD result, highlighted by the varied performance of the CFD-simulated results for different ML model-derived BC values. This observation is in agreement with similar studies [28,32,33]. Additionally, Figure 4 shows that the ML models have similar flow splits but differ mainly on the volume of flow that exits the descending aorta and the innominate artery. The bulk of the flow runs through the supra-aortic branches, which is realistic due to the presence of the coarctation, as it restricts flow from exiting the descending aorta. This is in contrast to a healthy case, in which the bulk of flow leaves the descending aorta, with an 80% to 30% split between it and the supra-aortic branches [28].

The results suggest that BC2 and BC4 are potential alternatives to pressure-based 3-EWM, which is traditionally preferred in the literature [3,30,32]. This could be beneficial in cases where reliable pressure data are absent. The 3-EWM has been proven capable of prescribing physiologically accurate BCs that result in highly accurate flow solutions. However, acquiring accurate patient pressure data and calibrating model parameters such as resistance and capacitance can pose a challenge to its implementation. Although there are approaches such as the ones presented in Pant et al. and Romarowaski et al. [30,34] which are capable of calibrating 3-EWM parameters in cases where limited data are available, pressure data are still required. Therefore, such an approach would have been difficult to implement in this study. It would be possible to obtain parameters from the literature with similar studies; however, this would still introduce errors due to variability between patients, as highlighted by Lassila et al. [19]. Therefore, our approach opens the door for haemodynamic modelling not only in cases where Doppler TTE is used as the velocity data acquisition modality, but also in those where pressure data may not be available. This allows the approach to be used with data collected from current clinical protocols for CoA where only Doppler TTE measurements are taken without the need for brachial cuff pressure measurements. Although a similar approach can be extended to other large arteries, additional considerations specific to those cases will be required. Noting the limitation of pressure data, future work would be to extend the framework to include 3-EWM, given the strong motivation within the literature for the approach.

Limitations

The dataset built for this study was limited, as it was restricted to one patient with a more complete dataset and the four additional patients with only Doppler TTE images across the coarctation. Therefore, the extension of the ML models to cases beyond those used within the study without retraining may result in a performance drop of the ML-CFD pipeline. Furthermore, the study was limited in the type of cases modelled. Modelling for intra-patient variability, for example, when the patient is in a stressed state, was not possible. Therefore, modelling was limited to a range of heart rates that captured a relaxed state. Both Lassila et al. [19] and Pant et al. [30] note that there is a significant difference between flow parameters for the two different states, with Pant et al. [30] further noting increased difficulty in accurately modelling flow in stressed states. Similarly, the effects of different patient geometries on the performance of the framework were not evaluated. Among the inputs used to train the ML models, the relative importance between vessel names was not taken into account when label encoding was used. Lastly, the maximum coarctation velocities results show that the choice of ML model used (and the type of BC prescribed at the outlet) had an impact on the outcomes. This motivates the need to

evaluate the performance of other models, such as artificial neural networks (ANNs), that may be able to determine more complex relationships within the data [35–37].

Assumptions made in the CFD modelling also contributed to the limitations of the study. These include the use of fixed BCs and the rigid wall assumption. However, accounting for these considerations comes at increased computational cost; hence, their inclusion should be balanced against improvements in accuracy [30]. A suitable compromise could be including these factors in the ML model component. In their paper, Raissi et al. [37,38] demonstrated the applicability of physics-informed neural networks, which can be seen as a hybrid approach that combines physics-based and ML models. Thus, future work could include the development of a hybrid ML and physics-based model that accounts for the noted limitations.

The evaluation of the framework was limited to comparing the simulated maximum coarctation velocity with the measured value from Doppler TTE at peak systole. However, it is necessary to simulate a full cardiac cycle in order to verify whether the improved accuracy in the velocity field is maintained. Additionally, the absence of reliable pressure data from the patient prevented the pressure field from being evaluated. This would ideally need to be evaluated, as velocity-based BCs have a strong influence on the pressure field, as highlighted by [34], and can result in non-physiological pressure ranges (see [28]) even though the velocity field is accurate. However, lack of access to a catheterisation lab and the additional risk to the patient of conducting an unnecessary catheterisation procedure to obtain pressure measurements make this unfeasible. Therefore, this is an open challenge within our framework.

5. Conclusions

Our novel contribution in this paper is the use of ML models in combination with echocardiography for application in a physics-based CFD model. Echocardiography, which is widely available across different resource contexts, is used to derive patient-specific BCs based on heart rate rather than pressure. The ML model then bridges the heart rate variability arising from the echo measurements being taken at different time points, which is inevitable given the measurement process. This process reduces the influence of the time difference, thereby improving the BCs and resulting in a more accurate CFD solution which better matches the measured data.

Author Contributions: V.M.M.P. contributed to the design of the pipeline, methodology, analysis, and writing of the manuscript. M.N. and A.K.M. contributed to the conceptualization of the work, analysis, methodology, and critical revision of the manuscript. T.A., J.L. and L.Z. contributed to data collection and critical revision of the manuscript. All authors have read and agreed to the published version of the manuscript.

Informed Consent Statement: The study used secondary data from the Red Cross War Memorial Children's Hospital through the PROTEA project. Patient consent was obtained.

Data Availability Statement: The code used within this study is available at the given link: https://figshare.com/s/a86bf491ef0a5b1e38cb accessed on 26 July 2022. The data is available upon reasonable request to the corresponding authors.

Conflicts of Interest: The authors declare that they have no conflicts of interest.

References

1. Liu, Y.; Chen, S.; Zühlke, L.; Black, G.C.; Choy, M.K.; Li, N.; Keavney, B.D. Global birth prevalence of congenital heart defects 1970–2017: Updated systematic review and meta-analysis of 260 studies. *Int. J. Epidemiol.* **2019**, *48*, 455–463. [CrossRef]
2. Zühlke, L.; Lawrenson, J.; Comitis, G.; De Decker, R.; Brooks, A.; Fourie, B.; Swanson, L.; Hugo-Hamman, C. Congenital heart disease in low-and lower-middle–income countries: Current status and new opportunities. *Curr. Cardiol. Rep.* **2019**, *21*, 163. [CrossRef]
3. Obaid, D.R.; Smith, D.; Gilbert, M.; Ashraf, S.; Chase, A. Computer simulated "Virtual TAVR" to guide TAVR in the presence of a previous Starr-Edwards mitral prosthesis. *J. Cardiovasc. Comput. Tomogr.* **2019**, *13*, 38–40. [CrossRef]
4. Cecchi, E.; Giglioli, C.; Valente, S.; Lazzeri, C.; Gensini, G.F.; Abbate, R.; Mannini, L. Role of hemodynamic shear stress in cardiovascular disease. *Atherosclerosis* **2011**, *214*, 249–256. [CrossRef]
5. Swanson, L.; Owen, B.; Keshmiri, A.; Deyranlou, A.; Aldersley, T.; Lawrenson, J.; Human, P.; De Decker, R.; Fourie, B.; Comitis, G.; et al. A patient-specific CFD pipeline using doppler echocardiography for application in coarctation of the aorta in a limited resource clinical context. *Front. Bioeng. Biotechnol.* **2020**, *8*, 409. [CrossRef]
6. Capelli, C.; Sauvage, E.; Giusti, G.; Bosi, G.M.; Ntsinjana, H.; Carminati, M.; Derrick, G.; Marek, J.; Khambadkone, S.; Taylor, A.M.; et al. Patient-specific simulations for planning treatment in congenital heart disease. *Interface Focus* **2018**, *8*, 20170021. [CrossRef]
7. Aldersley, T.; Lawrenson, J.; Human, P.; Shaboodien, G.; Cupido, B.; Comitis, G.; De Decker, R.; Fourie, B.; Swanson, L.; Joachim, A.; et al. PROTEA, a Southern African multicenter congenital heart disease registry and biorepository: Rationale, design, and initial results. *Front. Pediatr.* **2021**, *9*, 763060. [CrossRef]
8. Nita, C.I.; Puiu, A.; Bunescu, D.; Mihai Itu, L.; Mihalef, V.; Chintalapani, G.; Armstrong, A.; Zampi, J.; Benson, L.; Sharma, P.; et al. Personalized Pre-and Post-Operative Hemodynamic Assessment of Aortic Coarctation from 3D Rotational Angiography. *Cardiovasc. Eng. Technol.* **2022**, *13*, 14–40. [CrossRef]
9. Zhu, Y.; Chen, R.; Juan, Y.H.; Li, H.; Wang, J.; Yu, Z.; Liu, H. Clinical validation and assessment of aortic hemodynamics using computational fluid dynamics simulations from computed tomography angiography. *Biomed. Eng. Online* **2018**, *17*, 53. [CrossRef] [PubMed]
10. Goubergrits, L.; Riesenkampff, E.; Yevtushenko, P.; Schaller, J.; Kertzscher, U.; Hennemuth, A.; Berger, F.; Schubert, S.; Kuehne, T. MRI-based computational fluid dynamics for diagnosis and treatment prediction: Clinical validation study in patients with coarctation of aorta. *J. Magn. Reson. Imaging* **2015**, *41*, 909–916. [CrossRef] [PubMed]
11. Rinaudo, A.; D'Ancona, G.; Baglini, R.; Amaducci, A.; Follis, F.; Pilato, M.; Pasta, S. Computational fluid dynamics simulation to evaluate aortic coarctation gradient with contrast-enhanced CT. *Comput. Methods Biomech. Biomed. Eng.* **2015**, *18*, 1066–1071. [CrossRef]
12. Pirola, S.; Jarral, O.; O'Regan, D.; Asimakopoulos, G.; Anderson, J.; Pepper, J.; Athanasiou, T.; Xu, X. Computational study of aortic hemodynamics for patients with an abnormal aortic valve: The importance of secondary flow at the ascending aorta inlet. *APL Bioeng.* **2018**, *2*, 026101. [CrossRef]
13. Antonuccio, M.N.; Mariotti, A.; Fanni, B.M.; Capellini, K.; Capelli, C.; Sauvage, E.; Celi, S. Effects of Uncertainty of Outlet Boundary Conditions in a Patient-Specific Case of Aortic Coarctation. *Ann. Biomed. Eng.* **2021**, *49*, 3494–3507. [CrossRef]
14. Brüning, J.; Hellmeier, F.; Yevtushenko, P.; Kühne, T.; Goubergrits, L. Uncertainty quantification for non-invasive assessment of pressure drop across a coarctation of the aorta using CFD. *Cardiovasc. Eng. Technol.* **2018**, *9*, 582–596. [CrossRef]
15. Campbell, I.C.; Ries, J.; Dhawan, S.S.; Quyyumi, A.A.; Taylor, W.R.; Oshinski, J.N. Effect of inlet velocity profiles on patient-specific computational fluid dynamics simulations of the carotid bifurcation. *J. Biomech. Eng.* **2012**, *134*, 051001. [CrossRef]
16. Conti, M.; Long, C.; Marconi, M.; Berchiolli, R.; Bazilevs, Y.; Reali, A. Carotid artery hemodynamics before and after stenting: A patient specific CFD study. *Comput. Fluids* **2016**, *141*, 62–74. [CrossRef]
17. Tan, F.; Borghi, A.; Mohiaddin, R.; Wood, N.; Thom, S.; Xu, X. Analysis of flow patterns in a patient-specific thoracic aortic aneurysm model. *Comput. Struct.* **2009**, *87*, 680–690. [CrossRef]
18. Liam, S. The Development of a Patient-Specific, Open Source Computational Fluid Dynamics Tool to Comprehensively and Innovatively Study Coarctation of the Aorta in a Limited Resource Clinical Context. Master's Thesis, University of Cape Town, Cape Town, Western Cape, South Africa, 2020. Available online: http://hdl.handle.net/11427/33033 (accessed on 1 March 2021).
19. Lassila, T.; Sarrami-Foroushani, A.; Hejazi, S.; Frangi, A.F. Population-specific modelling of between/within-subject flow variability in the carotid arteries of the elderly. *Int. J. Numer. Methods Biomed. Eng.* **2020**, *36*, e3271. [CrossRef]
20. Véras, P.; Métais, O.; Balarac, G.; Georges, D.; Bombenger, A.; Ségoufin, C. Reconstruction of proper numerical inlet boundary conditions for draft tube flow simulations using machine learning. *Comput. Fluids* **2023**, *254*, 105792. [CrossRef]
21. Gèron, A. *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*; O'Reilly Media: Sebastopol, CA, USA, 2017.
22. *PlotDigitizer*, Version 2.2; PlotDigitizer: Phoenix, AZ, USA, 2022.
23. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
24. Harris, C.R.; Millman, K.J.; van der Walt, S.J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N.J.; et al. Array programming with NumPy. *Nature* **2020**, *585*, 357–362. [CrossRef]
25. Cai, L.; Ren, L.; Wang, Y.; Xie, W.; Zhu, G.; Gao, H. Surrogate models based on machine learning methods for parameter estimation of left ventricular myocardium. *R. Soc. Open Sci.* **2021**, *8*, 201121. [CrossRef] [PubMed]

26. Mehdi, R.R.; Mendiola, E.A.; Sears, A.; Choudhary, G.; Ohayon, J.; Pettigrew, R.; Avazmohammadi, R. *Chapter 19—Comparison of Three Machine Learning Methods to Estimate Myocardial Stiffness*; Academic Press: Cambridge, MA, USA, 2023. [CrossRef]

27. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.Y. Lightgbm: A highly efficient gradient boosting decision tree. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017; Volume 30.

28. Pirola, S.; Cheng, Z.; Jarral, O.; O'Regan, D.; Pepper, J.; Athanasiou, T.; Xu, X. On the choice of outlet boundary conditions for patient-specific analysis of aortic flow using computational fluid dynamics. *J. Biomech.* **2017**, *60*, 15–21. [CrossRef] [PubMed]

29. Madhavan, S.; Kemmerling, E.M.C. The effect of inlet and outlet boundary conditions in image-based CFD modeling of aortic flow. *Biomed. Eng. Online* **2018**, *17*, 66. [CrossRef]

30. Pant, S.; Fabrèges, B.; Gerbeau, J.F.; Vignon-Clementel, I. A methodological paradigm for patient-specific multi-scale CFD simulations: From clinical measurements to parameter estimates for individual analysis. *Int. J. Numer. Methods Biomed. Eng.* **2014**, *30*, 1614–1648. [CrossRef]

31. Cilla, M.; Martinez, J.; Pena, E.; Martínez, M.Á. Machine learning techniques as a helpful tool toward determination of plaque vulnerability. *IEEE Trans. Biomed. Eng.* **2012**, *59*, 1155–1161. [CrossRef]

32. Morbiducci, U.; Ponzini, R.; Gallo, D.; Bignardi, C.; Rizzo, G. Inflow boundary conditions for image-based computational hemodynamics: Impact of idealized versus measured velocity profiles in the human aorta. *J. Biomech.* **2013**, *46*, 102–109. [CrossRef] [PubMed]

33. Gallo, D.; De Santis, G.; Negri, F.; Tresoldi, D.; Ponzini, R.; Massai, D.; Deriu, M.; Segers, P.; Verhegghe, B.; Rizzo, G.; et al. On the use of in vivo measured flow rates as boundary conditions for image-based hemodynamic models of the human aorta: Implications for indicators of abnormal flow. *Ann. Biomed. Eng.* **2012**, *40*, 729–741. [CrossRef]

34. Romarowski, R.M.; Lefieux, A.; Morganti, S.; Veneziani, A.; Auricchio, F. Patient-specific CFD modelling in the thoracic aorta with PC-MRI–based boundary conditions: A least-square three-element Windkessel approach. *Int. J. Numer. Methods Biomed. Eng.* **2018**, *34*, e3134. [CrossRef]

35. Feiger, B.; Gounley, J.; Adler, D.; Leopold, J.A.; Draeger, E.W.; Chaudhury, R.; Ryan, J.; Pathangey, G.; Winarta, K.; Frakes, D.; et al. Accelerating massively parallel hemodynamic models of coarctation of the aorta using neural networks. *Sci. Rep.* **2020**, *10*, 9508. [CrossRef]

36. Yevtushenko, P.; Goubergrits, L.; Gundelwein, L.; Setio, A.; Heimann, T.; Ramm, H.; Lamecker, H.; Kuehne, T.; Meyer, A.; Schafstedde, M. Deep Learning Based Centerline-Aggregated Aortic Hemodynamics: An Efficient Alternative to Numerical Modelling of Hemodynamics. *IEEE J. Biomed. Health Inform.* **2021**, *26*, 1815–1825. [CrossRef]

37. Raissi, M.; Perdikaris, P.; Karniadakis, G. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **2019**, *378*, 686–707. [CrossRef]

38. Raissi, M.; Yazdani, A.; Karniadakis, G.E. Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations. *Science* **2020**, *367*, 1026–1030. [CrossRef]