# City Research Online

## City, University of London Institutional Repository

**Citation:** Bradley, G. R. E. & Kyriacou, P. A. (2024). Evaluating the effectiveness of non-invasive intracranial pressure monitoring via near-infrared photoplethysmography using classical machine learning methods. Biomedical Signal Processing and Control, 96(Part B), 106517. doi: 10.1016/j.bspc.2024.106517

This is the published version of the paper.

This version of the publication may differ from the final published version.

**Permanent repository link:** https://openaccess.city.ac.uk/id/eprint/33134/

**Link to published version:** https://doi.org/10.1016/j.bspc.2024.106517

**Copyright:** City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

**Reuse:** Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

City Research Online: http://openaccess.city.ac.uk/ publications@city.ac.uk

Biomedical Signal Processing and Control 96 (2024) 106517

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

# Evaluating the effectiveness of non-invasive intracranial pressure monitoring via near-infrared photoplethysmography using classical machine learning methods

George R.E. Bradley [*], Panayiotis A. Kyriacou

*Research Centre for Biomedical Engineering, City, University of London, United Kingdom*

## ARTICLE INFO

*Keywords:*
Traumatic brain injury
Photoplethysmography
Machine learning
Signal processing

## ABSTRACT

This study investigates the feasibility of utilising photoplethysmography signals to estimate continuous intracranial pressure (ICP) values in patients with traumatic brain injury. A clinical dataset was compiled, comprising synchronised data from a non-invasive optical sensor and an invasive gold standard ICP monitor from 27 patients. Two datasets, derived from short and long-distance NIRS, were generated from this data. For each dataset, 141 features were extracted for every one-minute window of non-invasive data. A total of 5 regression models were assessed. The study aimed to evaluate the models' performance for the continuous, non-invasive monitoring of ICP using a leave-one-patient out cross validation approach. The 5 models were trained on both the long and short distance NIRS data. The lowest mean absolute error (MAE) and root mean squared error (RMSE) were obtained using features derived from long-distance NIRS. A Random Forest (RF) model achieved the lowest MAE and RMSE of 5.030 and 4.067 mmHg respectively. The RF exhibited wide limits of agreement with the reference method. This was reflected in the 95% Bland–Altman limits of agreement, ranging from 8.782 to -8.487 mmHg.

## 1. Introduction

Traumatic brain injury (TBI) is a condition marked by the impairment of brain function resulting from sudden trauma, inflicting damage on the brain. The severity of symptoms varies, ranging from mild to moderate, or severe, depending upon the extent of damage to the brain [1]. Annually, the global incidence of TBI is estimated at 69 million cases, 5.48 million of which are classified as severe [2]. This condition imposes a considerable burden on affected individuals and their families, through health loss and disability. Moreover, TBI poses a large and growing burden on healthcare systems and nations, due to the complex and expensive medical care that the condition necessitates, coupled with substantial productivity loss, with an estimated annual cost reaching USD 400 billion [3]. The incidence rate of TBI is increasing, attributed to various factors such as population density, population aging, and the growing use of motor vehicles, motorcycles, and bicycles, particularly in low to middle-income countries (LMICs) where road traffic collisions are the primary cause [4–6].

Intracranial pressure (ICP) monitoring plays a crucial role in neurocritical care for the timely identification of intracranial hypertension (IH). Elevated ICP is associated with poor neuropsychological performance and functional outcomes in TBI patients. ICP refers to the pressure inside the cranial vault, and changes in ICP result from alterations in brain volume, cerebral blood volume, and the production or clearance of cerebrospinal fluid (CSF) [7,8]. Pathological processes, often related to TBI, can cause changes in ICP through conditions such as mass lesions, venous sinus obstruction, and cerebral edema. If these changes lead to volumetric increases within the cranial vault, compensatory mechanisms are activated to maintain ICP within normal ranges, typically 10 to 15 millimetres of mercury (mmHg) for adults and 3 to 7 mmHg for young children [9]. A critical threshold is reached when space-occupying lesions can no longer expand, which without efficacious monitoring and intervention can potentially lead to secondary brain injury, herniation, and brain death [10–12].

ICP monitoring is used either as a guide to treatment or as a diagnostic modality in a number of pathological conditions resulting in neurological injury. The gold standard for ICP monitoring is intraventricular pressure monitoring using an extraventricular drain (EVD) [13,14]. Intraventricular-based monitoring involves inserting a catheter into a ventricle of the brain through a burr hole in the skull. Despite being the gold standard, EVD monitoring is expensive and may lead to a series of complications. The most common complications

---

[*] Corresponding author.
*E-mail address:* george.bradley@city.ac.uk (G.R.E. Bradley).

https://doi.org/10.1016/j.bspc.2024.106517
Received 30 November 2023; Received in revised form 8 April 2024; Accepted 28 May 2024
Available online 14 June 2024
1746-8094/© 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

include intracranial hemorrhage (reported in 21.6% of patients), infection (8.3% to 23.5% cumulative incidence rates), and ventricle tap failure or misplacement (malplacement rate of 12.3%) [15–18]. The estimated diagnostic, procedural, and material cost of EVD placement has been suggested to be in the range of USD $1300–$3200 [19,20].

### 1.1. Current state of the art

Non-invasive ICP monitoring approaches have been explored but there are no existing continuous and non-invasive ICP monitoring techniques that are accurate across the pressure range in relation to invasive monitors. Transcranial Doppler (TCD), has been shown to not be accurate across the pressure range, particularly in elevated ICP values [21]. Additionally, TCD is susceptible to intra and inter-observer variability [22–25]. Moreover, approximately 10%–15% of patients may not benefit from TCD due to the inability of ultrasound waves to penetrate the skull [26]. Similarly, the optic nerve sheath diameter (ONSD) technique suffers from intra-observer variance [27–29]. Additionally, there are a number of conditions that can affect the optical nerve sheath diameter rendering the ONSD approach un-efficacious [30].

There has been a growing body of work investigating approaches utilising cardiac waveforms, including diffuse correlation spectroscopy and near-infrared spectroscopy (NIRS) to estimate intracranial pressure non-invasively. A pilot study in 1997 demonstrated that NIRS values when ICP was > 25 mmHg were significantly lower than those when ICP < 25 mmHg. These results suggested that NIRS could be of value in the evaluation of increased ICP [31]. Ruesch et al.'s study investigated the estimation of ICP non-invasively in non-human primates using near-infrared spectroscopy [32]. NIRS data was recorded whilst varying ICP levels were induced. Hemodynamic responses to ICP changes were measured and compared to changes detected by a conventional intraparenchymal ICP probe. The results demonstrated that there was a high correlation with haemoglobin concentrations and induced ICP suggesting that NIRS has the potential for non-invasive ICP monitoring. Ruesch et al.'s later study utilised morphological and time series features extracted from diffuse correlation spectroscopy (DCS) data in addition to mean arterial pressure for the prediction of ICP non invasively [33]. This work demonstrated the potential to use morphological and time series features derived from DCS pulsations to monitor ICP non-invasively resulting in a Coefficient of determination ($R^2$) of 0.92 and an mean squared error (MSE) of 3.3 mmHg. The same group extended this research to the use of NIRS derived photoplethysmography (PPG) features in combination to MAP to estimate ICP [34]. This study returned an increased $R^2$ of 0.937 and a MSE of 2.703 mmHg. More recently there has been research demonstrating the use of PPG waveform features and their relationship to ICP levels, reporting a $R^2$ of 0.66 suggesting that these features may provide a method for measuring ICP non-invasively [35].

This field of research aims to contribute knowledge towards the pioneering of continuous, non-invasive ICP monitoring helping reduce the barrier to entry through the reduction in cost, complexity, and risk. This specific study builds upon the research of Roldan et al. [36,37]. Its contribution is an evaluation of the feasibility of continuous monitoring of ICP non-invasively using morphological and time-series features derived from PPG signal data from the largest and clinically collected labelled dataset.

## 2. Materials and methods

### 2.1. Non-invasive sensor

This study is based upon data produced by an in-house near-infrared spectroscopy (NIRS) reflectance non-invasive optical ICP sensor hereinafter referred to as the "nICP" sensor [38]. The nICP sensor consists of four LEDs at four different wavelengths (770, 810, 855 and 880 nm) and two photodiodes "*proximal*" and "*distal*" arranged as shown in Fig. 1. Pulsatile signals were acquired from both the *proximal* and *distal* photodiodes of the nICP sensor positioned at 10 mm and 35 mm from the light source we refer to this data as "NIR-PPG".

The reflected light from extracerebral tissue reached the proximal photodetector placed 10 mm from the LEDs, and the non-absorbed light from deeper tissues travelled back to the distal photodiode placed 35 mm from the LEDs. It is proposed that the data from the proximal photodiode corresponds to extracerebral data, while the data from the distal photodiode represents a combination of extracerebral and cerebral data [39]. The nICP monitor was calibrated for each patient through the adjustment of the LED intensity and amplification gain according to patients' characteristics and ambient light. Calibration was performed before recording started. The data was sampled at a frequency of 100 Hz.

### 2.2. Data

The primary dataset consists of data from 40 patients with severe TBI recruited from the intensive therapy unit of the Royal London Hospital. The non-randomised data collection was performed between January 2020 and July 2021 (ClinicalTrials No. NCT05632302). Each patient had an implanted invasive ICP probe (Raumedic® Neurovent-P intra-parenchymal pressure probe) and the nICP sensor was affixed to the patient's forehead below the hairline. The non-invasive and invasive data were collected synchronously. The patients' average age was 43.92 years, with a male-to-female ratio of 13:2. Data collection lasted for an average of 42.16 h. None of the patients received an EVD.

This research focuses on the data collected from the 810 nm wavelength collected from the proximal and distal photodiodes. The selection of the 810 nm wavelength is based on the property that oxyhemoglobin and deoxyhemoglobin exhibit the same absorption characteristics at this wavelength, enabling the extraction of an optical signal independent of blood oxygenation and eliminating it as a confounding factor [40]. Fig. 2 depicts traces of the signal for the same patient at elevated and low ICP values from both the proximal and distal photodiodes.

### 2.3. Data preprocessing

Initially, the dataset consisted of data from 40 patients. Of these 40, the data of 6 patients were identified as either being corrupted or missing information resulting in a dataset of 34 patients. Of these, the data of 7 patients were removed due to data integrity, resulting in a final dataset from 27 patients.

Within the dataset, two main sources of noise are hypothesised to be present (i) motion artifacts, which are identified by irregular signal morphology or high amplitude variance, and (ii) photodetector saturation, characterised by areas with little or no amplitude variance referred to as "flat lines". Fig. 3 shows an example of the noise within the data. The non-invasive data were denoised utilising the *Envelope PPG Denoising Algorithm* [41]. Any instances of data identified as anomalous were removed from both the nICP data and the reference ICP data to ensure temporal synchronicity between the datasets. Additionally, the assumption was made that values of ICP $\geq$ 60 mmHg or $<$ 0 mmHg, could be erroneous or biologically implausible. Consequently, these instances were removed from both the reference and nICP data. Values above 60 mmHg may indicate potential measurement errors, while negative values are physiologically implausible.

### 2.4. Feature engineering

The NIR-PPG dataset was segmented into 1-minute windows, with the median ICP value over each window being used as a label. This 1-minute window size was chosen considering the possible clinical relevance and practicality of the prediction frequency. Longer window durations might be less favourable in a clinical setting, and our goal

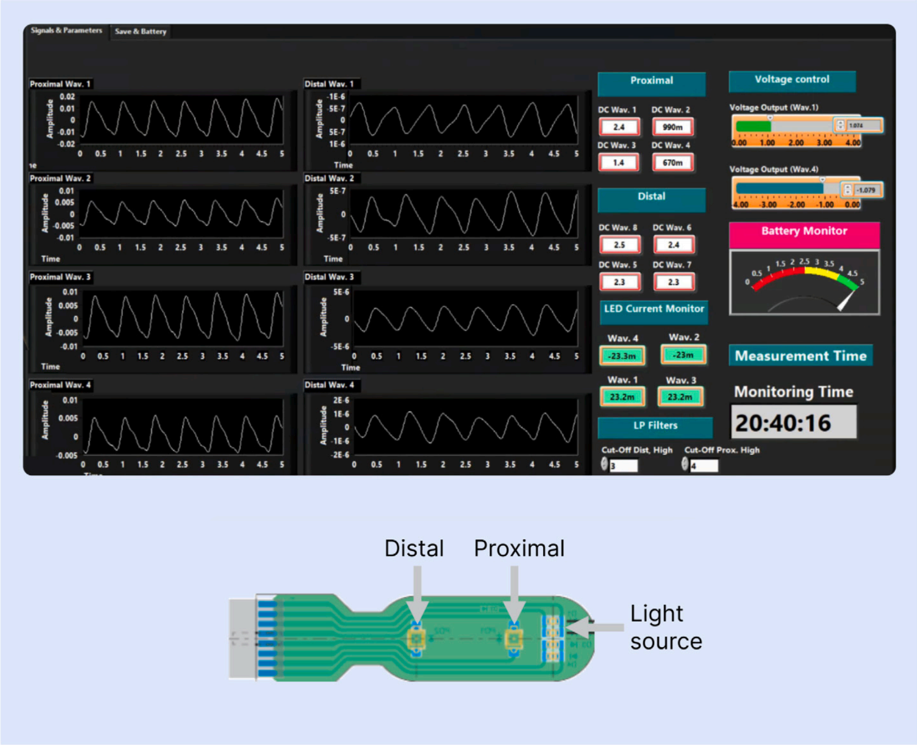
Signals & Parameters
Save & Battery
Proximal Wav. 1
Proximal Wav. 2
Proximal Wav. 3
Proximal Wav. 4
Distal Wav. 1
Distal Wav. 2
Distal Wav. 3
Distal Wav. 4
Amplitude
Time
Proximal
DC Wav. 1 2.4
DC Wav. 2 990m
DC Wav. 3 1.4
DC Wav. 4 670m
Distal
DC Wav. 8 2.5
DC Wav. 6 2.4
DC Wav. 5 2.3
DC Wav. 7 2.3
LED Current Monitor
Wav. 4 -23.3m
Wav. 2 -23m
Wav. 1 23.2m
Wav. 3 23.2m
LP Filters
Cut-Off Dist. High
Cut-Off Prox. High
Voltage control
Voltage Output (Wav.1)
Voltage Output (Wav.4)
Battery Monitor
Measurement Time
Monitoring Time
20:40:16
Distal
Proximal
Light source

**Fig. 1.** nICP sensor's design and interface.

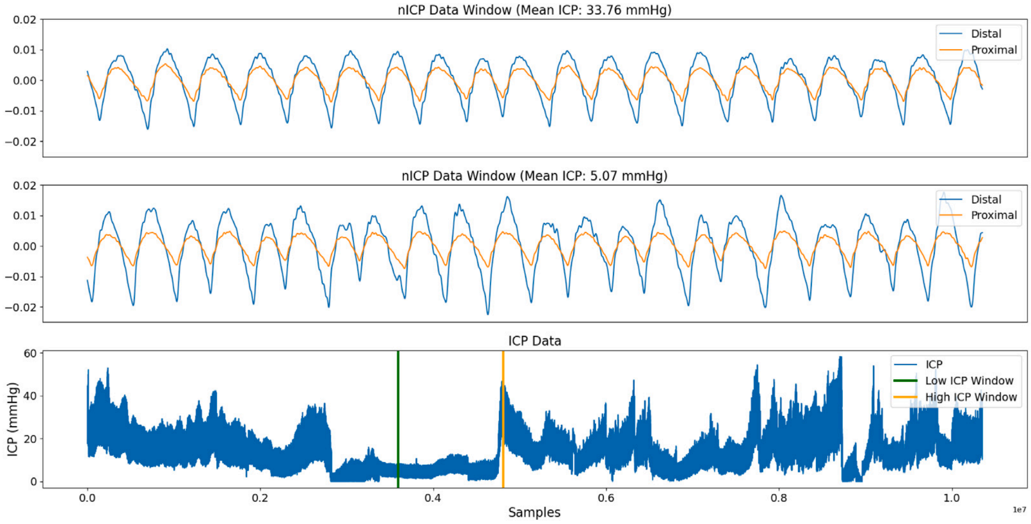

**Fig. 2.** A three-subplot figure illustrating non-invasive ICP signal data. Two subplots showcase a 30-second window each, with one capturing data from a period of elevated ICP and the other from a period of low ICP. The third subplot presents the entire ICP dataset, emphasising the sampled areas.

was to use a window size that maximises the amount of relevant data available for prediction while still being clinically manageable. The extracted features are derived from the AC, pulsatile component of the signal. The pulsatility of the signal is as a consequence of the absorption of light resulting from blood volume changes during the cardiac cycle. During systole there is an increase in blood volume resulting in increased absorbance of light compared to the diastolic state. This change in light absorbance creates the pulsatile waveform synchronous with each heartbeat. In order to capture as much of the morphology of the signal which is linked to haemodynamic changes, for each window of data 141 morphological and time-series features were extracted across the original signal and its first and second derivative as shown in Table 5 in the appendix. Figs. 4 and 5 depict the primary time series and morphological features extracted from the cardiac pulses. The feature representation for each window was determined by computing the median of each feature across all pulses within that window. The derivatives of the signal were calculated using a Savitzky–Golay filter using a polynomial of 7 and a window size of 101.

In order to attempt to diminish the possible affect of calibration and baseline wandering on the data collected from multiple patients the alternating current (AC) and direct current (DC) components of each patient's signal data were isolated using a bandpass and lowpass filter respectively. The bandpass frequency band ranged from 0.4 to 10 Hz, the low pass cutoff frequency was set to 0.4 Hz. By dividing the AC component of the signal by the DC component we aim to minimise the effects of the patient level calibration and provide a normalised

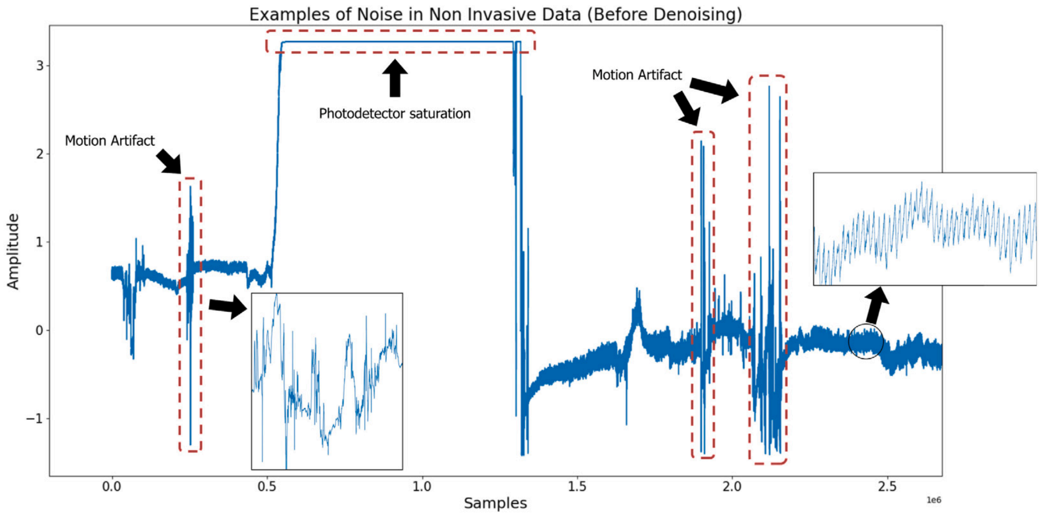

**Fig. 3.** Distal signal data with highlighted noise attributed to motion artifact and photodetector saturation. Visual aids emphasise the presence of noise within the window.

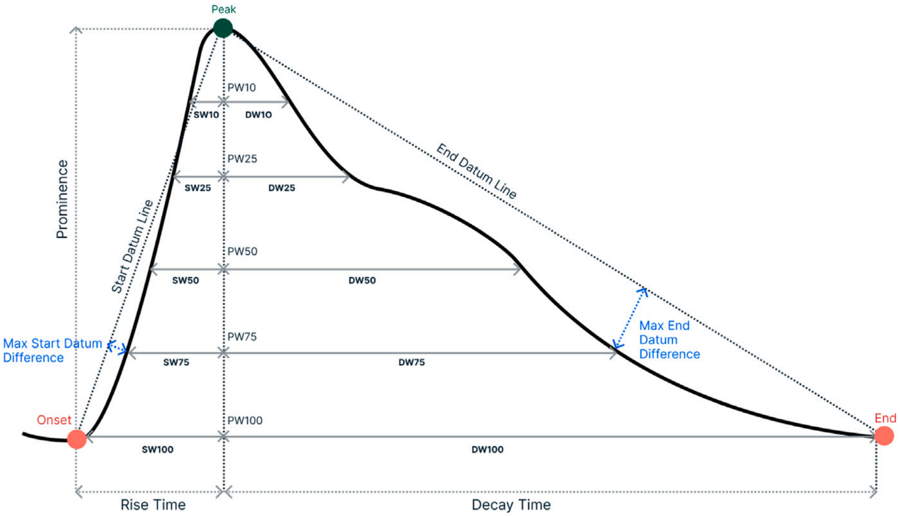
Peak
Prominence
PW10
SW10
DW10
PW25
SW25
DW25
End Datum Line
Start Datum Line
PW50
SW50
DW50
Max End Datum Difference
Max Start Datum Difference
PW75
SW75
DW75
Onset
End
PW100
SW100
DW100
Rise Time
Decay Time

**Fig. 4.** A figure illustrating the key time-series and morphological features extracted from individual cardiac pulses. Within the figure, SW, DW, and PW represent systolic width, diastolic width, and pulse width, respectively. The numerical value associated with SW, DW, or PW indicates the position along the pulse prominence where the corresponding measurement is taken.

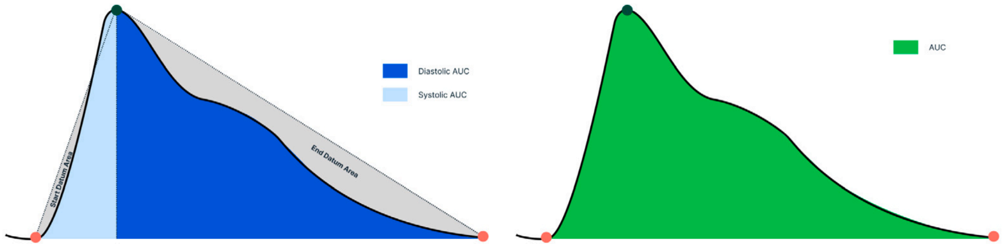
Diastolic AUC
Systolic AUC
Start Datum Area
End Datum Area
AUC

**Fig. 5.** A figure depicting the Area Under the Curve (AUC), along with the Diastolic AUC and Systolic AUC. Additionally, the start and end datum areas are identified.

representation of the signal that is less influenced by calibration variations and allows for better inter-patient comparison. Following this, each window's data were normalised to a range of (0,1) to account for inter-patient amplitudinal variance and to aid in the creation of a shared distribution across patients in the feature space.

The extracted features were predicated upon the detection of cardiac cycles within the signal. In order to effectively extract features a pulse detection algorithm was developed. The detection algorithm is predicated upon the crossing points of the signal and the moving average of the signal. The general concept of the algorithm is, if the mean of the data between crossing points is greater than the mean of the moving average between crossing points the location of the maximum value between crossing points is classified as a peak. Vice versa if the mean of the data between points is less than the mean of the moving average between crossing points the index of data which has the lowest value between points is classified as a trough. The identified

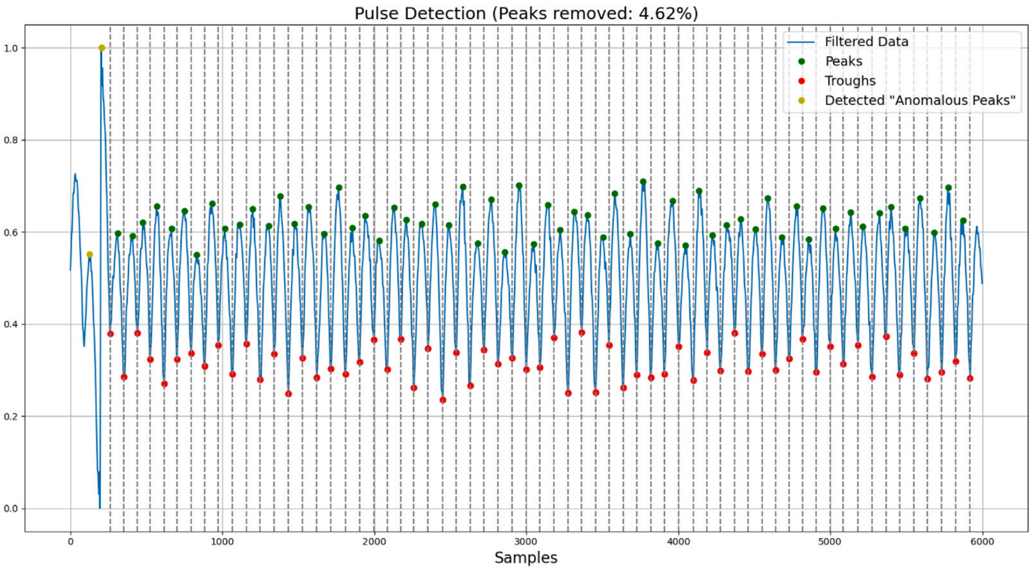

**Fig. 6.** 1-minute window of NIR-PPG signal data with detected peaks, troughs and anomalous peaks.

peaks and associated troughs are filtered to ensure the correct peak–trough association. Fig. 6 shows an example of the identified pulses within a window of data.

In order to handle the presence of possible residual anomalous peaks within each window of data a z-score outlier detection approach was applied using three pulse characteristics: (i) difference between consecutive peaks, (ii) pulse onset-peak difference and (iii) peak-end difference. A z-score threshold of 3 was used. No features were extracted from pulses which were detected as anomalous. During model training and testing the aggregate feature data across patients was normalised using MinMaxScaler from the Python package Scikit-learn [42]. MinMaxScaler was fit to the training data and the transformation was applied to the testing dataset.

### 2.5. Feature selection

A significant amount of the features within the feature sets are colinearly related due to being derived from similar or the same morphological elements of the pulse. In order to perform feature selection whilst dealing with the presence of colinearly related features a three step feature selection process was performed. Initially the maximum information coefficient (MIC) was calculated for each feature against the label. MIC captures linear, non-linear and non-monotonic relationships offering a more precise representation of complex dependencies between features and the label.

To address the issue of collinearity among variables, we identified pairs of features exhibiting a Pearson correlation coefficient $\geq 0.9$. Within each of these pairs, we then removed the feature with the lower MIC score, prioritising the retention of features with higher potential predictive power. Following this, Recursive Feature Elimination with Cross-Validation (RFECV), utilising a 10-fold cross-validation scheme, was implemented to further refine the selection of predictive features. When conducting the RFECV we employed a group-wise holdout method in the cross-validation, thereby preventing any patient's data from simultaneously appearing in both training and validation sets.

### 2.6. Model optimisation and evaluation

The *proximal* and *distal* datasets were both segmented into two distinct datasets: (i) *evaluation* and (ii) *optimisation*. In order to create the evaluation and optimisation datasets, the aggregate dataset of both the *proximal* and *distal* data was split by patients in a ratio of 70:30. Patients were randomly sampled, the patient data of 19 patients and 8 patients made up the evaluation and optimisation sets respectively. In order to effectively evaluate the models' ability for continuous monitoring the temporal order of the windows was maintained. The data of the same 19 and 8 patients comprised the evaluation and optimisation datasets for both the proximal and distal data.

Five regression models were evaluated: (i) K-Nearest Neighbours (KNN), (ii) Light Gradient Boosting Machine (LGBM), (iii) Random Forest (RF) (iv) Extreme Gradient Boosting (XGB) and Epsilon-Support Vector Regression (SVR). The hyperparameters of the models were optimised using the optimisation dataset exclusively.

The hyperparameters were optimised using the aggregate root mean squared error (RMSE) across leave-one-patient-out cross-validation (LOPOCV). For each model 50 hyperparameter optimisation "trials" were performed. To mitigate the risk of overfitting to the training dataset during model optimisation, the hyperparameter search spaces were tuned. Initially, the hyperparameter search spaces encompassed broader ranges, but as the optimisation progressed if any instances of overfitting to the training data were seen the search space was refined. This involved narrowing the range of specific hyperparameters, implementing stronger regularisation penalties, and adjusting learning rates to encourage improved generalisation. Our primary objective was to strike a balance, achieving good performance on the training data while ensuring the model's effectiveness in generalising to the validation data. The hyperparameter search spaces for each model are outlined in Table 1. The hyperparameter optimisation was carried out twice, once for both the *proximal* and *distal* optimisation data.

Following optimisation the hyperparameters which returned the lowest loss during optimisation were used when evaluating each model using the evaluation dataset. The resultant optimised hyperparameters are summarised in Table 1. In order to produce a robust evaluation of the models ability to predict continuous ICP values each model was evaluated using LOPOCV across the 19 patients which made up the evaluation dataset.

The evaluation metrics MSE, Root mean squared error, Mean absolute error (MAE), $R^2$, the upper and lower Bland–Altman limits of agreement as well as the correlation coefficient were calculated for the testing and training datasets to provide an overview of model performance.

**Table 1**
Table containing the defined search spaces for each regression model and the resultant optimised hyperparameters.

| Model | Hyperparameter search space | Optimised hyperparameters (Proximal) | Optimised hyperparameters (Distal) |
|---|---|---|---|
| RF | max depth = 15 to 25<br>number of estimators = 100 to 500<br>min samples split = 10 to 20<br>min samples leaf = 10 to 15 | max depth = 18<br>number of estimators = 101<br>min samples split = 13<br>min samples leaf = 11 | max depth = 17<br>number of estimators = 196<br>min samples split = 10<br>min samples leaf = 11 |
| KNN | number of neighbours = 3 to 10<br>algorithm = [auto, ball tree, kd tree, brute]<br>metric = [euclidean, manhattan, chebyshev, minkowski] | number of neighbours = 10<br>algorithm = auto<br>metric = chebyshev | number of neighbours = 10<br>algorithm = auto<br>metric = chebyshev |
| LGBM | max depth = 5 to 10<br>number of iterations = 100 to 250<br>learning rate = 0.1 to 0.5<br>number of leaves = 25 to 35<br>extra tress = True or False | max depth = 10<br>number of iterations = 102<br>learning rate = 0.179<br>number of leaves = 28<br>extra tress = False | max depth = 8<br>number of iterations = 241<br>learning rate = 0.160<br>number of leaves = 35<br>extra tress = True |
| XGB | number of estimators = 10 to 20<br>max depth = 10 to 15<br>learning rate = 0.3 to 0.7<br>min child weight = 45 to 55 | number of estimators = 19<br>max depth = 10<br>learning rate = 0.351<br>min child weight = 53 | number of estimators = 18<br>max depth = 14<br>learning rate = 0.382<br>min child weight = 55 |
| SVR | C = 10 to 50<br>gamma (DISTAL) = 5 to 10<br>gamma (PROXIMAL) = 1 to 6<br>epsilon = 0.0001 to 0.01 | C = 10.855<br>gamma = 5.986<br>epsilon = 0.000164 | C = 10<br>gamma = 9.972<br>epsilon = 0.002 |

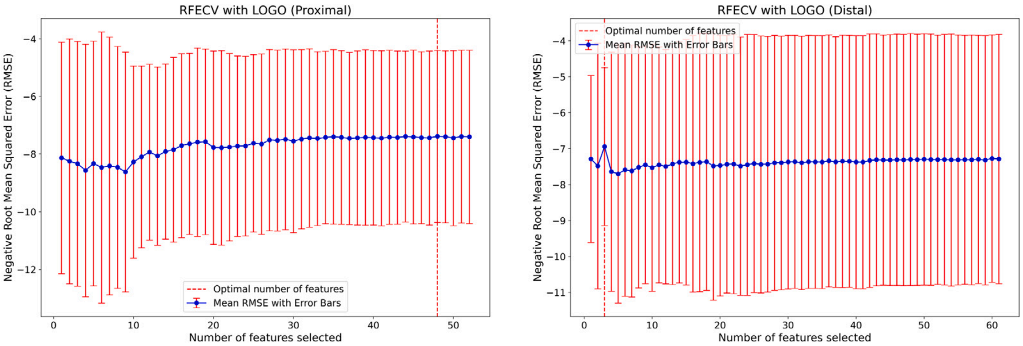

**Fig. 7.** A figure with two subplots illustrating recursive feature elimination plots applied to a subset of features. The plots demonstrate the variation in negative Root Mean Squared Error as the number of features changes. A dashed red line represents the number of features with support returned by recursive feature elimination. The left subplot corresponds to the proximal dataset, while the right subplot corresponds to the distal dataset.

## 3. Results

### 3.1. Feature selection

To address the issue of collinearity among variables, we first identified pairs of features exhibiting a Pearson correlation coefficient ≥ 0.9. Within the proximal and distal sets, 377 and 162 correlated feature pairs were identified. Once the features with the lower MIC score were removed and a unique set of the remaining features was taken, the resultant feature sets had a length of 52 and 61 features for the proximal and distal sets respectively. With the resulting features, RFECV was applied to the feature subsets to determine the most predictive features. Fig. 7 depicts plots for both the proximal and distal datasets, demonstrating the number of features selected and subsequent optimal number of features determined by RFECV.

The feature selection approach returned 48 features with support for the proximal dataset and 3 features with support for the distal dataset. The features with support for the distal dataset led to underfitting, to counteract this, the number of features which returned the second lowest loss for the distal dataset was selected (n=18). Table 2 contains the selected features for both the proximal and distal datasets.

### 3.2. Model evaluation

Five regression models where evaluated. The MSE, RMSE, MAE, $R^2$, the upper and lower Bland–Altman limits of agreement as well as the correlation coefficient were calculated. The training and testing results are presented in Tables 3 and 4.

Extrapolating from the results presented in Table 3, an examination of predictive performance is conducted by focusing on four key metrics: RMSE, MAE, correlation coefficient and Bland–Altman limits of agreement. MAE assesses the average absolute errors between predicted and actual values. Unlike MSE, MAE does not square errors meaning it does not amplify the influence of outliers or large errors. By not penalising large errors, MAE offers a balanced view of model performance.

On the other hand, RMSE also gauges the discrepancy between predicted and actual values. In contrast to MAE, RMSE calculates the square root of the average of squared errors. Squaring the errors gives more weight to larger errors, meaning they have a proportionally greater impact on the RMSE. This makes RMSE sensitive to outliers and large errors, emphasising their importance. We posit that in the context of estimating ICP, the clinical implications of errors can vary significantly. Larger errors in ICP prediction can have critical consequences for patient care and safety. For instance, underestimating ICP might lead to inadequate intervention, risking patient health. Therefore, prioritising the reduction of larger errors is crucial to enhance the clinical utility and safety of the model. RMSE aligns well with this objective. Given this, the feature selection process using RFECV and model hyperparameter optimisation were centred around minimising squared errors as the primary cost function. RMSE, being the square root of MSE, allows for a more intuitive interpretation of the model's performance as it brings the evaluation metric back to the original units

**Table 2**

Table containing the selected features for both the proximal and distal datasets.

| Proximal | Distal |
|---|---|
| onset-end slope (deriv 1) | pulse width 100 |
| ds ratio 25 | kurt (deriv 2) |
| std ibi (deriv 2) | pulse width 10 |
| rise time - decay time ratio (deriv 1) | diastolic width 50 (deriv 1) |
| systolic width 100 | systolic width 75 |
| upslope | rise time - decay time ratio |
| kurt | rise time - decay time ratio (deriv 1) |
| rise time - decay time ratio | max end datum difference |
| pulse length - height ratio | ibi (deriv 2) |
| std ibi | skew |
| ds ratio 25 (deriv 1) | ds ratio 50 (deriv 2) |
| kurt (deriv 1) | start datum difference |
| systolic width 10 | kurt (deriv 1) |
| systolic width 75 (deriv 2) | zcr |
| max start datum difference (deriv 2) | skew (deriv 1) |
| zcr | ds ratio 50 (deriv 1) |
| ds ratio 75 (deriv 2) | skew (deriv 2) |
| skew | systolic width 25 (deriv 2) |
| end datum difference (deriv 1) | |
| downslope length (deriv 2) | |
| onset-end slope (deriv 2) | |
| pulse length - height ratio (deriv 2) | |
| end datum difference (deriv 2) | |
| max start datum difference | |
| start datum area (deriv 1) | |
| end datum area (deriv 1) | |
| systolic width 100 (deriv 1) | |
| max end datum difference | |
| ds ratio 10 | |
| downslope | |
| downslope (deriv 1) | |
| std ibi (deriv 1) | |
| systolic width 25 (deriv 2) | |
| end datum difference | |
| onset-end slope | |
| zcr (deriv 2) | |
| ds ratio 25 (deriv 2) | |
| start datum difference (deriv 2) | |
| skew (deriv 1) | |
| ds ratio 75 | |
| max end datum difference (deriv 1) | |
| start datum difference (deriv 1) | |
| diastolic width 10 (deriv 1) | |
| end datum area (deriv 2) | |
| diastolic width 75 (deriv 2) | |
| pulse width 10 (deriv 2) | |
| systolic width 10 (deriv 1) | |
| diastolic width 10 | |

**Table 3**

Table containing the results (test) for the 5 regression models, estimating intracranial pressure. The table contains the mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), coefficient of determination ($R^2$), the upper and lower Bland–Altman limits of agreement as well as the correlation coefficient (corr coef) values for each model.

| Distal results (Evaluation) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Model | RMSE test | MSE test | MAE test | R2 test | Bland–Altman upper test | Bland–Altman lower test | Bland–Altman bias test | Test corr coef |
| KNN | 5.529 | 32.735 | 4.445 | −1.496 | 10.057 | −9.202 | 0.427 | 0.021 |
| LGBM | 5.253 | 29.902 | 4.233 | −1.217 | 9.220 | −8.862 | 0.179 | −0.016 |
| RF | 5.030 | 27.364 | 4.067 | −1.046 | 8.782 | −8.487 | 0.147 | −0.007 |
| SVR | 5.473 | 32.148 | 4.404 | −1.370 | 9.266 | −9.348 | −0.041 | 0.027 |
| XGB | 5.379 | 30.986 | 4.338 | −1.354 | 9.372 | −9.154 | 0.109 | 0.001 |
| Proximal results (Evaluation) | | | | | | | | |
| Model | RMSE test | MSE test | MAE test | R2 test | Bland–Altman upper test | Bland–Altman lower test | Bland–Altman bias test | Test corr coef |
| KNN | 5.734 | 35.212 | 4.650 | −1.766 | 10.055 | −8.995 | 0.530 | −0.052 |
| LGBM | 5.627 | 33.680 | 4.634 | −1.677 | 9.414 | −8.661 | 0.376 | −0.022 |
| RF | 5.538 | 32.969 | 4.538 | −1.555 | 9.423 | −8.598 | 0.413 | −0.026 |
| SVR | 5.640 | 33.268 | 4.608 | −1.701 | 9.289 | −9.159 | 0.065 | −0.020 |
| XGB | 5.689 | 34.444 | 4.663 | −1.729 | 9.625 | −9.025 | 0.300 | −0.013 |

**Table 4**
Table containing the results (train) for the 5 regression models, estimating intracranial pressure. The table contains the mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), coefficient of determination ($R^2$), the upper and lower Bland–Altman limits of agreement as well as the correlation coefficient (corr coef) values for each model.

| Model | RMSE train | MSE train | MAE train | R2 train | Bland–Altman upper train | Bland–Altman lower train | Bland–Altman bias train | Train corr coef |
|---|---|---|---|---|---|---|---|---|
| Distal results (Evaluation) | | | | | | | | |
| KNN | 3.063 | 9.387 | 2.011 | 0.625 | 6.021 | −5.986 | 0.017 | 0.792 |
| LGBM | 2.993 | 8.961 | 2.121 | 0.642 | 5.866 | −5.866 | 0.000 | 0.807 |
| RF | 2.709 | 7.341 | 1.854 | 0.707 | 5.316 | −5.303 | 0.006 | 0.861 |
| SVR | 3.203 | 10.267 | 1.986 | 0.590 | 5.965 | −6.542 | −0.288 | 0.772 |
| XGB | 2.573 | 6.625 | 1.816 | 0.735 | 5.041 | −5.045 | −0.002 | 0.862 |
| Proximal results (Evaluation) | | | | | | | | |
| Model | RMSE train | MSE train | MAE train | R2 train | Bland–Altman upper train | Bland–Altman lower train | Bland–Altman bias train | Train corr coef |
| KNN | 3.613 | 13.058 | 2.586 | 0.478 | 7.128 | −7.034 | 0.047 | 0.696 |
| LGBM | 3.052 | 9.316 | 2.198 | 0.628 | 5.981 | −5.981 | 0.000 | 0.798 |
| RF | 2.556 | 6.540 | 1.738 | 0.739 | 5.019 | −5.002 | 0.008 | 0.877 |
| SVR | 2.427 | 5.892 | 1.199 | 0.765 | 4.474 | −4.985 | −0.256 | 0.880 |
| XGB | 2.643 | 6.991 | 1.876 | 0.721 | 5.177 | −5.184 | −0.003 | 0.855 |

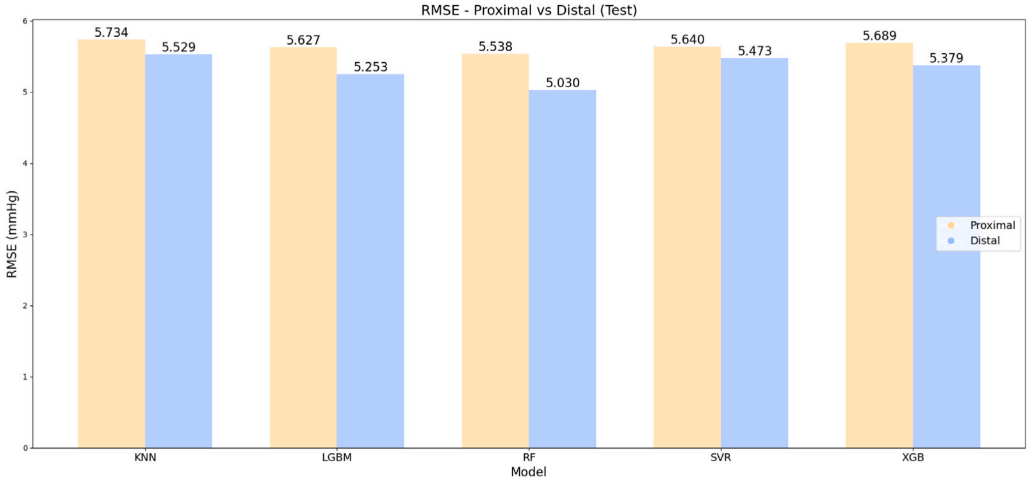

**Fig. 8.** A bar chart presenting the root mean squared error of each model for both the short and long distance near-infrared spectroscopy data.

of the target variable. Both MAE and RMSE are presented in the same units as the target variable (mmHg), facilitating easier comprehension and comparison (see Fig. 8).

Overall the models developed using distal features performed better than model developed using proximal features. The mean MAE and RMSE across models developed using distal features were 4.297 and 5.333 mmHg, approximately 5.7% and 7.2% lower than the mean MAE and RMSE across models developed using proximal features. Of the models developed using distal data, the best performing model with reference to RMSE and MAE was a RF model which returned a mean RMSE and MAE across the testing folds of LOPOCV of 5.030 and 4.067 mmHg respectively. The worst performing distal model with reference to RMSE and MAE was a KNN model which produced an RMSE and MAE of 5.529 and 4.445 mmHg.

To complement the RMSE and MAE we investigate the correlation coefficient and Bland–Altman limits of agreement of the developed models. Fig. 9 depicts the Bland–Altman limits of agreement for each developed model. Bland–Altman is a statistical technique used to assess the agreement between two quantitative measurements. The Bland–Altman plot provides insights into the extent of agreement and any potential biases between the predicted and actual values. The limits of agreement, marked in the plot, illustrate the range within which most of the differences between predicted and actual values lie. Specifically, they represent the mean difference between the predicted and actual values, with 95% confidence intervals meaning that 95% of the differences fall within this range, offering a valuable indication of the agreement's reliability.

The mean upper and lower limits of agreement of the distal and proximal derived models are 9.334 and −9.011 and 9.561 and −8.888 mmHg accordingly. The limits of agreement between the models trained using distal and proximal data are similar. Both groups of models returned negligible mean biases of 0.337 and 0.164 mmHg suggests a minimal systematic difference between the predicted and actual ICP values. This implies that, on average, the ICP predictions from the models are within approximately 8 mmHg of the actual values.

In the case of ICP prediction a preferable correlation coefficient would be a strong positive correlation with label. The mean correlation coefficient returned by both the distal and proximal derived models are very low. The distal derived models returned a mean correlation coefficient of approximately 0.005, 118.5% greater than the mean correlation coefficient returned by the proximal derived models. The higher and positive mean correlation coefficient of the distal derived models apposed to the mean negative correlation coefficient of the proximal derived models suggests that distal features may be more useful for the prediction of ICP. Despite this, the correlation coefficients remain low.

## 4. Discussion

This study has investigated the effectiveness of non-invasive ICP monitoring using NIR-PPG derived features. Five classical machine learning models were optimised and evaluated on both short distance and long distance NIR-PPG data.

From the models developed, the distal models returned lower RMSE and MAE errors on average than the models developed using proximal

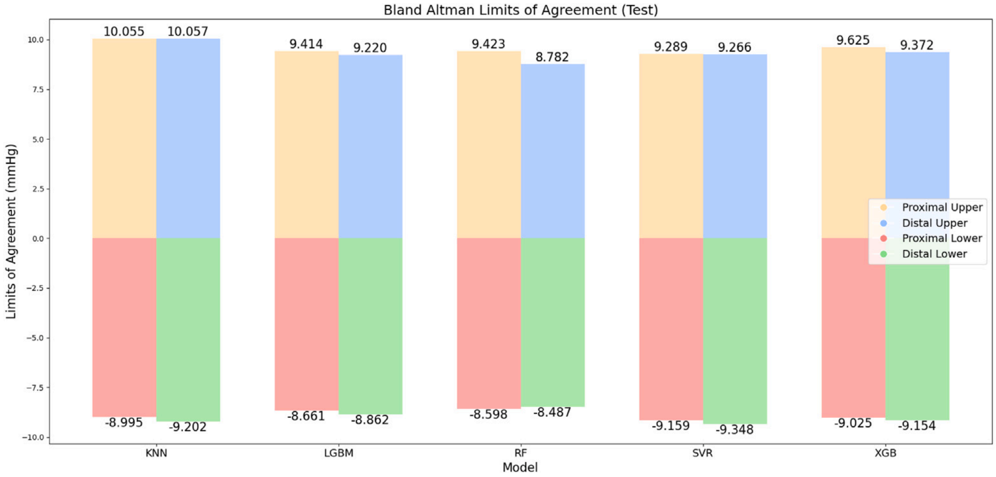

**Fig. 9.** A bar chart representing the Bland–Altman upper and lower limits of agreement for each model, calculated at the 95% confidence level. The bars in the chart indicate the upper and lower limits of agreement, representing the range within which most of the differences between predicted and actual values lie with 95% confidence.

derived features. This supports our hypothesis that long distance NIR-PPG data is associated with changes in ICP. Additionally although both groups of models returned low correlation coefficients the distal models returned a higher mean correlation coefficient in comparison to the proximal models which returned a mean negative correlation coefficient.

The low correlation coefficients and wide limits of agreement observed in our analysis suggest that the models struggled to generalise to new, unseen data. We hypothesise that this challenge is primarily attributable to noise within the data and inter-patient data heterogeneity. Specifically, variances in the distribution of ICP among patients may impact model performance.

Moreover, the presence of inter-patient feature data heterogeneity and the consequential challenge of creating a shared distribution exacerbates the difficulty in achieving optimal model generalisation. This heterogeneity can stem from diverse patient characteristics, disease severity, and treatment responses among others. This coupled with varying distributions of ICP levels among the patients, can impair the model's ability to learn underlying patterns effectively thereby resulting in wider limits of agreement.

Addressing these challenges could involve increasing the dataset, particularly with more data collected at the high and low ends of the ICP distribution as well as increasing the diversity of the patient population which comprise the dataset. A larger and more diverse dataset could help in reducing the observed wide limits of agreement by enabling the model to better learn and predict the outcomes for patients with less common characteristics.

Animal studies have investigated the estimation of ICP non-invasively using NIRS, features extracted from DCS pulsations and NIR-PPG cardiac waveforms [32–34]. These approaches within these studies have demonstrated significant agreement with ICP returning $R^2$ of up to 0.937 and MSE lower than < 3 mmHg. Despite this, these studies use MAP as an additional feature for model development. MAP was found to be the strongest predictive feature across both studies [33,34]. This study aimed to investigate the efficacy of predicting ICP non-invasively using NIR-PPG morphological and time-series features exclusively.

The evaluation and optimisation methodology could be optimised further. The overall patient data was divided into two distinct datasets: (i) the *evaluation* set and (ii) the *optimisation* set. Random sampling was employed, with 19 patients (70%) contributing to the evaluation set and 7 patients (30%) forming the optimisation set. The feature selection and model optimisation were performed using the optimisation dataset exclusively to minimise potential data leakage. Model optimisation utilised the aggregate RMSE across LOPOCV. Following hyperparameter optimisation, model evaluation occurred through LOPOCV on the evaluation dataset.

This methodology was chosen over nested cross-validation to strike a balance between performance and computational efficiency. Despite this choice, there remains a possibility of bias in the models towards the optimisation set's data. The 30% of patients' data constituting the optimisation dataset was deemed a reasonably representative sample. Even if the models exhibit bias towards the optimisation dataset, they do not positively bias the evaluation results; if anything, they could negatively impact them. This approach was considered reasonable and serves to prevent data leakage and provide a robust evaluation of model performance.

This research brings novel insights to non-invasive ICP monitoring through the investigation of the largest clinically collected NIR-PPG dataset from patients with severe TBI and invasive ICP monitoring to date. The outcomes of this study provide solid groundwork for future research within this domain especially the investigation of the causes of possible inter-patient heterogeneity.

The search for non-invasive alternatives to ICP monitoring is significant. The potential of NIR-PPG technology to change TBI diagnosis and ICP monitoring arises as a promising prospect. NIR-PPG based monitoring offers the potential for an inexpensive, easy-to-use, safe, and real-time alternative. Moreover, a NIR-PPG device as a consequence of its non-invasive nature and simplicity of use could make an ideal point-of-care monitor in various settings, from emergency rooms and ambulances to sports events and remote locations. The cost-effectiveness of NIR-PPG further improves its appeal, as it offers a more affordable approach to TBI diagnosis and ICP monitoring.

## 5. Conclusion

This study aimed to assess the potential of non-invasive ICP monitoring using NIR-PPG derived features. Five classical machine learning models were optimised and evaluated on both short distance *proximal* and long distance *distal* NIR-PPG data. The models results indicated that distal models returned better results than models developed using proximal features returning a higher mean RMSE, MAE and correlation coefficient. The best performing model was a RF model which returned a mean RMSE and MAE across the testing folds of LOPOCV of 5.030 and 4.067 mmHg respectively with Bland–Altman limits of agreement of approximately 8.5 mmHg and a low correlation coefficient of −0.007.

This research brings novel insights to non-invasive ICP monitoring through the investigation of the largest clinically collected NIR-PPG dataset from patients with severe TBI and invasive ICP monitoring and the investigation into the possibility of predicting ICP non-invasively using exclusively morphological and time series features from NIRS-PPG. This research area is nascent and evolving and future work is required. The outcomes of this study provide a solid groundwork for future research within this domain.

**Table 5**
Table containing the list of features extracted from the original, first derivative and second derivative of the signal.

| Features |
|---|
| AUC |
| Systolic AUC |
| Diastolic AUC |
| Ratio between the systolic and diastolic AUC |
| Rise time (samples between the pulse onset and peak) |
| Decay time (samples between the pulse peak and end) |
| Ratio between the rise time and decay time |
| Number of beats |
| Inter-beat interval |
| Standard deviation of the inter-beat interval |
| Prominence |
| Upslope (slope between pulse onset to peak) |
| Downslope (slope between peak to pulse end) |
| Onset-end slope (slope between pulse onset and end) |
| Ratio between the upslope and downslope |
| Ratio between the pulse length and height |
| Start datum area<br>(area between a straight line between the pulse onset and peak and pulse data between those points) |
| End datum area<br>(area between a straight line between the pulse peak and end and pulse data between those points) |
| Ratio between the start datum area and end datum area |
| Max start datum difference<br>(maximum element-wise difference between a straight line between the pulse onset and peak and pulse data between those points) |
| Max end datum difference<br>(maximum element-wise difference between a straight line between the pulse peak and end and pulse data between those points) |
| The median of the element-wise difference between a straight line between the pulse onset and peak and pulse data between those points |
| The median of the element-wise difference between a straight line between the pulse peak and end and pulse data between those points |
| Pulse width at 10% of the pulse prominence (in samples) |
| Pulse width at 25% of the pulse prominence (in samples) |
| Pulse width at 50% of the pulse prominence (in samples) |
| Pulse width at 75% of the pulse prominence (in samples) |
| Pulse width at 100% of the pulse prominence (in samples) |
| Systolic width at 10% of the pulse prominence (in samples) |
| Systolic width at 25% of the pulse prominence (in samples) |
| Systolic width at 50% of the pulse prominence (in samples) |
| Systolic width at 75% of the pulse prominence (in samples) |
| Systolic width at 100% of the pulse prominence (in samples) |
| Diastolic width at 10% of the pulse prominence (in samples) |
| Diastolic width at 25% of the pulse prominence (in samples) |
| Diastolic width at 50% of the pulse prominence (in samples) |
| Diastolic width at 75% of the pulse prominence (in samples) |
| Diastolic width at 100% of the pulse prominence (in samples) |
| Ratio between the diastolic and systolic pulse width at 10% of the pulse prominence |
| Ratio between the diastolic and systolic pulse width at 25% of the pulse prominence |
| Ratio between the diastolic and systolic pulse width at 50% of the pulse prominence |
| Ratio between the diastolic and systolic pulse width at 75% of the pulse prominence |
| Ratio between the diastolic and systolic pulse width at 100% of the pulse prominence |
| Variance of the pulse data |
| Skew of the pulse data |
| Kurtosis of the pulse data |
| Zero-crossing rate of the pulse data |

## CRediT authorship contribution statement

**George R.E. Bradley:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Panayiotis A. Kyriacou:** Writing – review & editing, Supervision, Resources, Project administration, Data curation, Conceptualization.

## Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: George Bradley reports financial support was provided by The George Daniels Educational Trust. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

The data that has been used is confidential.

## Acknowledgment

The George Daniels Educational Trust - George Daniels Doctoral Studentship.

## Appendix

See Table 5.

## References

[1] Alexandra Brazinova, et al., Epidemiology of traumatic brain injury in Europe: A living systematic review, J. Neurotrauma 38 (10) (2018) 1411–1440.

[2] Michael C Dewan, et al., Estimating the global incidence of traumatic brain injury, J. Neurosurg. (2018) 1–18.

[3] Jeroen T J M van Dijck, et al., In-hospital costs after severe traumatic brain injury: A systematic review and quality assessment, PLoS One 14 (5) (2019) e0216743.

[4] Spencer, et al., Global, regional, and national burden of traumatic brain injury and spinal cord injury, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016, Lancet Neurol. 18 (1) (2018) 56–87.

[5] Nelofar Kureshi, et al., Long-Term trends in the epidemiology of major traumatic brain injury, J. Community Health 46 (6) (2021) 1197–1203.

[6] Renee Y. Hsia, et al., Ten-year trends in traumatic brain injury: a retrospective cohort study of California emergency department and hospital revisits and readmissions, BMJ Open 8 (12) (2018) e022297.

[7] B. Mokri, The Monro-Kellie hypothesis: applications in CSF volume depletion, Neurology 56 (12) (2001) 1746–1748.

[8] J. Ghajar, Traumatic brain injury, Lancet 356 (9233) (2000) 923–929.

[9] Leonardo Rangel-Castilla, Shankar Gopinath, Claudia S. Robertson, Management of intracranial hypertension, Neurol Clin. 26 (2) (2008) 521–541.

[10] Robert D. Stevens, Michael Shoykhet, Rhonda Cadena, Emergency neurological life support: Intracranial hypertension and herniation, Neurocrit. Care 23 Suppl 2 (Suppl 2) (2015) S76–82.

[11] Charles A. Adams, et al., Does intracranial pressure management hurt more than it helps in traumatic brain injury? Trauma Surg. Acute Care Open 3 (1) (2018) e000142.

[12] D.C. Engel, et al., Pre-hospital and in-hospital parameters and outcomes in patients with traumatic brain injury: a comparison between German and Australian trauma registries, Injury 41 (9) (2010) 901–906.

[13] J. Guillaume, P. Janny, [Continuous intracranial manometry; importance of the method and first results], Rev. Neurol. (Paris) 84 (2) (1951) 131–142.

[14] N. Lundberg, Continuous recording and control of ventricular fluid pressure in neurosurgical practice, Acta Psychiatr. Scand. Suppl. 36 (149) (1960) 1–193.

[15] Catherine Miller, Ramachandra P. Tummala, Risk factors for hemorrhage associated with external ventricular drain placement and removal, J. Neurosurg. 126 (1) (2016) 289–297.

[16] Andrea Saladino, et al., Malplacement of ventricular catheters by neurosurgeons: A single institution experience, Neurocrit. Care 10 (2008) 248–252, http://dx.doi.org/10.1007/s12028-008-9154-z.

[17] S Hagel, et al., External ventricular drain infections: risk factors and outcome, Interdiscip. Perspect. Infect. Dis. 2014 (2014) 708531.

[18] Kelly Wright, et al., Rates and determinants of ventriculostomy-related infections during a hospital transition to use of antibiotic-coated external ventricular drains, Neurosurg. Focus 34 (5) (2013) E12.

[19] Quentin Aten, et al., Causes, complications, and costs associated with external ventricular drainage catheter obstruction, World Neurosurg. (ISSN: 1878-8750) 134 (2020) 501–506, http://dx.doi.org/10.1016/j.wneu.2019.10.105, URL https://www.sciencedirect.com/science/article/pii/S1878875019327251.

[20] Peter Le Roux, Intracranial pressure monitoring and management, in: Translational Research in Traumatic Brain Injury, CRC Press/Taylor and Francis Group, Boca Raton (FL), 2016.

[21] Johan Bellner, et al., Transcranial Doppler sonography pulsatility index (PI) reflects intracranial pressure (ICP), Surg. Neurol. 62 (1) (2004) 45–51; discussion 51.

[22] H Maeda, et al., A validation study on the reproducibility of transcranial Doppler velocimetry, Ultrasound Med. Biol. 16 (1) (1990) 9–14.

[23] R.W. Baumgartner, et al., A validation study on the intraobserver reproducibility of transcranial color-coded duplex sonography velocity measurements, Ultrasound Med. Biol. 20 (3) (1994) 233–237.

[24] C.J. McMahon, et al., The reproducibility of transcranial Doppler middle cerebral artery velocity measurements: implications for clinical practice, Br. J. Neurosurg. 21 (1) (2007) 21–27.

[25] Q. Shen, et al., Inter observer variability of the transcranial Doppler ultrasound technique: impact of lack of practice on the accuracy of measurement, J. Clin. Monit. Comput. 15 (3–4) (1999) 179–184.

[26] Georgios Tsivgoulis, Andrei V. Alexandrov, Michael A. Sloan, Advances in transcranial Doppler ultrasonography, Curr. Neurol. Neurosci. Rep. 9 (1) (2009) 46–54.

[27] Tuba Cimilli Ozturk, et al., Assessment of intra-interobserver reliability of the sonographic optic nerve sheath diameter measurement, Kaohsiung J. Med. Sci. 31 (8) (2015) 432–436.

[28] Olufunso Simisola Aduayi, et al., Optic nerve sonography: A noninvasive means of detecting raised intracranial pressure in a resource-limited setting, J. Neurosci. Rural Pract. 6 (4) (2015) 563–567.

[29] S.A Ballantyne, et al., Observer variation in the sonographic measurement of optic nerve sheath diameter in normal adults, Eur. J. Ultrasound (ISSN: 0929-8266) 15 (3) (2002) 145–149, http://dx.doi.org/10.1016/S0929-8266(02)00036-8, URL https://www.sciencedirect.com/science/article/pii/S0929826602000368.

[30] P H Raboel, et al., Intracranial pressure monitoring: Invasive versus Non-Invasive Methods-A review, Crit. Care Res. Pract. 2012 (2012) 950393.

[31] A. Kampfl, et al., Near infrared spectroscopy (NIRS) in patients with severe brain injury and elevated intracranial pressure, in: Hector E. James, Lawrence F. Marshall, Hans J. Raulen, Alexander Baethmann, Anthony Marmarou, Umeo Ito, Julian T. Hoff, Toshihiko Kuroiwa, Zbigniew Czernicki (Eds.), Brain Edema X, Springer Vienna, Vienna, ISBN: 978-3-7091-6837-0, 1997, pp. 112–114.

[32] Alexander Ruesch, et al., Fluctuations in intracranial pressure can be estimated non-invasively using near-infrared spectroscopy in non-human primates, J. Cereb. Blood Flow Metab. 40 (11) (2020) 2304–2314, http://dx.doi.org/10.1177/0271678X19891359, PMID: 31775565.

[33] Alexander Ruesch, et al., Estimating intracranial pressure using pulsatile cerebral blood flow measured with diffuse correlation spectroscopy, Biomed. Opt. Express 11 (3) (2020) 1462–1476.

[34] Filip A.J. Relander, et al., Using near-infrared spectroscopy and a random forest regressor to estimate intracranial pressure, Neurophotonics 9 (4) (2022) 045001.

[35] Barry Dixon, et al., Assessment of a Non-Invasive brain pulse monitor to measure Intra-Cranial pressure following acute brain injury, Med. Devices (Auckl) 16 (2023) 15–26.

[36] Maria Roldan, Panicos A. Kyriacou, Head phantom for the acquisition of pulsatile optical signals for traumatic brain injury monitoring, Photonics (ISSN: 2304-6732) 10 (5) (2023) http://dx.doi.org/10.3390/photonics10050504, URL https://www.mdpi.com/2304-6732/10/5/504.

[37] Maria Roldan, et al., Non-invasive monitoring of intracranial pressure changes: healthy volunteers study, Front. Physiol. (ISSN: 1664-042X) 14 (2023) http://dx.doi.org/10.3389/fphys.2023.1208010, URL https://www.frontiersin.org/articles/10.3389/fphys.2023.1208010.

[38] Maria Roldan, Panicos A. Kyriacou, A non-invasive optical multimodal photoplethysmography-near infrared spectroscopy sensor for measuring intracranial pressure and cerebral oxygenation in traumatic brain injury, Appl. Sci. (ISSN: 2076-3417) 13 (8) (2023) http://dx.doi.org/10.3390/app13085211, URL https://www.mdpi.com/2076-3417/13/8/5211.

[39] M. Roldan, S. Chatterjee, P.A. Kyriacou, Brain Light-Tissue interaction modelling: Towards a non-invasive sensor for traumatic brain injury, in: Annu Int Conf IEEE Eng Med Biol Soc, Vol. 2021, United States, 2021, pp. 1292–1296.

[40] J.M. Murkin, M. Arango, Near-infrared spectroscopy as an index of brain and tissue oxygenation, Br. J. Anaesth. 103 (2009) i3–i13.

[41] George R.E. Bradley, Panayiotis A. Kyriacou, Opening the envelope: Efficient envelope-based PPG denoising algorithm, Biomed. Signal Process. Control (ISSN: 1746-8094) 88 (2024) 105693, http://dx.doi.org/10.1016/j.bspc.2023.105693, https://www.sciencedirect.com/science/article/pii/S1746809423011266.

[42] F. Pedregosa, et al., Scikit-learn: Machine learning in Python, J. Mach. Learn. Res. 12 (2011) 2825–2830.