

# Using near-infrared spectroscopy and a random forest regressor to estimate intracranial pressure

Filip A. J. Relander,<sup>a</sup> Alexander Ruesch,<sup>a</sup> Jason Yang<sup>ORCID</sup>,<sup>a</sup>  
Deepshikha Acharya<sup>ORCID</sup>,<sup>a</sup> Bradley Scammon,<sup>a</sup> Samantha Schmitt,<sup>a,b</sup>  
Emily C. Crane<sup>ORCID</sup>,<sup>a</sup> Matthew A. Smith<sup>ORCID</sup>,<sup>a,b</sup> and Jana M. Kainerstorfer<sup>ORCID</sup>,<sup>a,b,\*</sup>

<sup>a</sup>Carnegie Mellon University, Department of Biomedical Engineering, Pittsburgh, Pennsylvania, United States

<sup>b</sup>Carnegie Mellon University, Neuroscience Institute, Pittsburgh, Pennsylvania, United States

## Abstract

**Significance:** Intracranial pressure (ICP) measurements are important for patient treatment but are invasive and prone to complications. Noninvasive ICP monitoring methods exist, but they suffer from poor accuracy, lack of generalizability, or high cost.

**Aim:** We previously showed that cerebral blood flow (CBF) cardiac waveforms measured with diffuse correlation spectroscopy can be used for noninvasive ICP monitoring. Here we extend the approach to cardiac waveforms measured with near-infrared spectroscopy (NIRS).

**Approach:** Changes in hemoglobin concentrations were measured in eight nonhuman primates, in addition to invasive ICP, arterial blood pressure, and CBF changes. Features of average cardiac waveforms in hemoglobin and CBF signals were used to train a random forest (RF) regressor.

**Results:** The RF regressor achieves a cross-validated ICP estimation of  $0.937r^2$ , 2.703-mmHg<sup>2</sup> mean squared error (MSE), and 95% confidence interval (CI) of  $[-3.064\ 3.160]$  mmHg on oxyhemoglobin concentration changes;  $0.946r^2$ , 2.301-mmHg<sup>2</sup> MSE, and 95% CI of  $[-2.841\ 2.866]$  mmHg on total hemoglobin concentration changes; and  $0.963r^2$ , 1.688 mmHg<sup>2</sup> MSE, and 95% CI of  $[-2.450\ 2.397]$  mmHg on CBF changes.

**Conclusions:** This study provides a proof of concept for the use of NIRS in noninvasive ICP estimation.

© The Authors. Published by SPIE under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.NPh.9.4.045001](https://doi.org/10.1117/1.NPh.9.4.045001)]

**Keywords:** biophotonics; near-infrared spectroscopy; machine learning; intracranial pressure.

Paper 22043GR received Jun. 16, 2022; accepted for publication Sep. 23, 2022; published online Oct. 11, 2022.

## 1 Introduction

Intracranial pressure (ICP) is the pressure inside the skull.<sup>1</sup> Healthy ICP is maintained through a balance of cerebral blood volume, cerebrospinal fluid (CSF), and brain tissue.<sup>2</sup> Abnormal ICP occurs when an imbalance in one compartment outweighs the compensatory limits of the other two compartments, as explained by the Monro–Kellie doctrine.<sup>2</sup> An increase in ICP can stem from many issues, including brain bleeds, cerebral edema, a mass lesion such as a brain tumor, and others.<sup>1</sup>

The monitoring of ICP is widely used to guide the treatment of various diseases and illnesses, such as traumatic brain injury (TBI) and hydrocephalus.<sup>3,4</sup> TBI, for instance, accounted for over 223,000 hospitalizations in the United States in 2018.<sup>5</sup> Meanwhile, 1 in every 770 babies born in the United States develop congenital hydrocephalus.<sup>6</sup> Accurate ICP monitoring is also pertinent to the estimation of cerebral perfusion pressure (CPP), calculated as the difference between mean arterial pressure (MAP) and ICP. CPP, in turn, is used in the gauging of cerebral autoregulation

\*Address all correspondence to Jana M. Kainerstorfer, [jkainers@andrew.cmu.edu](mailto:jkainers@andrew.cmu.edu)

(CA), which defines the brain's ability to maintain a near constant blood flow when subject to slow changes in blood pressure,<sup>7-9</sup> and changes in CPP have been linked to altered neuronal function and neurovascular coupling.<sup>10,11</sup> The "gold standard" of ICP measurement is done with an invasive external ventricular drain (EVD). Although the accuracy of an EVD is sensitive to its placement, EVDs have the added benefit of enabling the drainage of excessive CSF.<sup>12</sup>

Although invasive ICP monitoring methods, such as EVDs, microtransducers, and lumbar puncture (LP) manometries, can be precise in their assessment of ICP in patients, they are invasive and not without risks. EVDs, for example, require a coronal burr hole to be drilled and a catheter to be placed near the third ventricle. This can be difficult in some patients, especially in pediatrics for which direct intraparenchymal ICP monitoring has become the norm instead.<sup>13</sup> EVDs, and microtransducers specifically, have the potential to cause hemorrhage or infection-related complications.<sup>14</sup> The LP manometries, as a less invasive alternative, measure ICP through CSF pressure in the spinal cord.<sup>15</sup> LP-based ICP sensing has been shown to agree with EVD-based ICP sensing, but appropriate LP-based sensing depends on 30-min recording periods.<sup>15-17</sup> In addition, although LPs are less invasive than EVDs or intraparenchymal probes, LPs not only carry the risk of infection observed with EVDs, but they are also prone to a higher risk of brain herniation and a larger set of contraindications compared with EVDs.<sup>18-20</sup>

Several noninvasive ICP monitoring methods have been proposed to mitigate the medical risks involved with invasive procedures. Transcranial Doppler (TCD) measures cerebral blood flow (CBF) velocity, from which peak systolic and diastolic flow rates can be extracted. Although some studies have shown that these flow rates correlate with ICP levels, other studies suggest that TCD lacks generalizability.<sup>21-24</sup> Likewise, the observation of tympanic membrane displacement coincides with changes in ICP, but the method suffers from poor negative predictive results.<sup>25,26</sup> Other noninvasive methods using ultrasound, computed tomography (CT), or magnetic resonance imaging (MRI) tend to either be imprecise, fail to work on some patients, only perform binary ICP-level classification, or not be applicable for continuous ICP monitoring applications. CT-based methods, specifically, suffer from a lack of supporting research that would indicate a predictive use for abnormal CT scans in ICP estimation.<sup>14,25</sup> Ultrasound methods, meanwhile, are occasionally afflicted by unwanted artifacts and require a standardized technique across patients for reliable ICP sensing.<sup>14,25</sup>

Hemodynamic-based methods that use diffuse optical devices, such as near-infrared spectroscopy (NIRS) or diffuse correlation spectroscopy (DCS), have also been used for noninvasive ICP monitoring.<sup>27-29</sup> Previously, we have successfully estimated ICP from cardiac waveform features extracted from CBF, measured using DCS.<sup>27</sup> We have also demonstrated that relative changes in oxyhemoglobin concentrations correlate with relative changes in ICP.<sup>28</sup> NIRS-based ICP estimation is favorable to other noninvasive ICP monitoring methods, including TCD, CT, MRI, ultrasound, and DCS, due to its potential low cost, user independence, and bedside compatibility for long-term monitoring.

In this work, we explore noninvasive ICP estimation using cardiac waveform features extracted from relative hemoglobin concentration changes, measured using a NIRS device. We compare regression performance on hemoglobin concentration-based signals with regression performance on CBF-based signals taken over approximately the same time periods and across the same set of subjects.

## 2 Materials and Methods

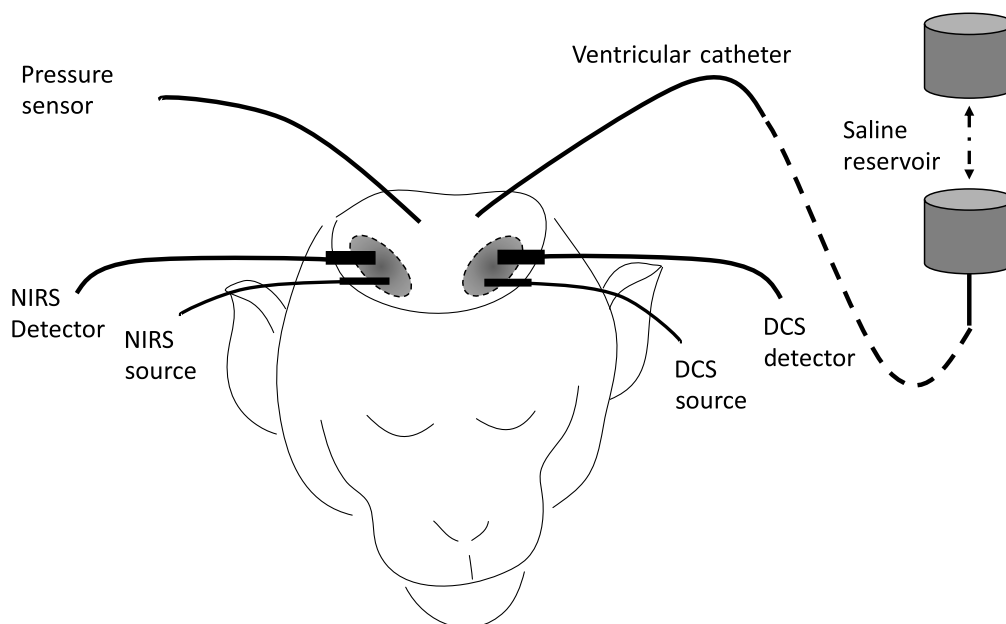
Oxyhemoglobin and deoxyhemoglobin concentration changes along with CBF changes were recorded at various ICP levels with NIRS and DCS across eight nonhuman primates (NHPs). Cardiac pulse waveforms were extracted and processed. Their features were used to train a machine learning algorithm for noninvasive ICP estimation.

### 2.1 Experimental Setup and Signal Extraction

Experimental procedures were approved by the University of Pittsburgh and Carnegie Mellon University's Institutional Animal Care and Use Committees and complied with the guidelines in the National Institute of Health's Guide for the Care and Use of Laboratory Animals.

Details of the experimental protocol and data collection have been reported previously.<sup>28,30</sup> Briefly, eight anesthetized NHPs (NHPs 1 to 4 *Macaca mulatta* aged  $7.9 \pm 1.5$  years, weighing  $9.4 \pm 0.7$  kg, and NHPs 5 to 8 *Macaca fascicularis* aged  $4.2 \pm 0.9$  years, weighing  $5.1 \pm 2.1$  kg) were used for the experiments. Each NHP was sedated throughout experimentation. Prior to their transport into the experimentation room, the NHPs were sedated with 20-mg/kg ketamine. In some animals, an additional 0.04-mg/kg of atropine and 1-mg/kg diazepam were administered. During the experiment, the monkeys were anesthetized with a combination of 0.6% to 1.5% isoflurane (ISO) and 10- to 25.6- $\mu\text{g}/\text{kg}/\text{h}$  fentanyl. Additionally, 0.1-mg/kg/h of vecuronium bromide paralytic was administered intravenously. NHPs were ventilated at 0.18 to 0.4 Hz. The ventilation frequency was kept constant for the duration of the experiment for each NHP.

A frequency-domain NIRS system (OxiplexTS, ISS Inc., United States), as shown in Fig. 1, operating at 690 and 830 nm, was used to measure cerebral hemoglobin concentration changes. The differential pathlength factor (DPF) ratio was held constant between all animals at  $\text{DPF}_{690}/\text{DPF}_{830} = 1.1$ . Although DPF differences are likely between animals, the cardiac waveform shape is not expected to be influenced significantly. To overcome magnitude differences between animals, signals were normalized in steps described later. Optical probes were placed directly on the skull of the animals, above the right prefrontal cortex. For NHPs 1 to 3, the source-detector distance was 2.2 cm, and for NHPs 4 to 8, it was 1.5 cm. The data for this manuscript were based on retrospective analysis. Differences in source-detector distances were due to variations in overall experimental designs between animals. Light intensity changes were recorded at a sampling frequency of 50 Hz and converted to changes in oxygenated ( $\Delta\text{HbO}$ ), deoxygenated ( $\Delta\text{Hb}$ ), and total ( $\Delta\text{HbT} = \Delta\text{HbO} + \Delta\text{Hb}$ ) hemoglobin concentrations using the modified Beer-Lambert's law.<sup>31</sup> In addition, a DCS probe was placed on the left prefrontal cortex in NHPs 1 to 4 and the right temporal cortex for NHPs 5 to 8, with source-detector distances of 2.2 and 1.9 cm, respectively. DCS data from NHP 4 were excluded from further processing due to instrument instabilities. DCS sampling frequency was 100 Hz. Further details of the device are found in Ref. 27. Relative CBF was estimated using a scaled diffusion coefficient measured from the autocorrelation of light intensity changes.<sup>10</sup> The scaled diffusion coefficient has been shown to correlate with CBF.<sup>32</sup> CBF was then used for further processing.



**Fig. 1** Experimental setup with NIRS and DCS placement on the skull above the right and left prefrontal cortex, respectively. The NIRS setup shows a defined source-detector distance. ICP was measured using a parenchymal pressure sensor. Saline was administered using a ventricular catheter. ICP was regulated by adjusting the height of the saline reservoir.

Arterial blood pressure (ABP) was recorded with an MPR1 Datalogger (Raumedic Helmbrechts, Germany) using an arterial line placed in the carotid artery. ICP was altered using a catheter (Lumbar catheter, Medtronic, Minneapolis, Minnesota, United States) placed in the lateral ventricle of the brain, with the other end connected to a saline reservoir, as shown in Fig. 1. A change in the height of the reservoir resulted in a pressure change in the head (simulating hydrocephalus). ICP was measured using a parenchymal pressure sensor recorded by the MPR1 Datalogger (Raumedic Helmbrechts, Germany). Both ICP and ABP were recorded at 100 Hz.

Each NHP experiment, lasting  $\sim 22.6$  h (std 2.3 h), was split between 7 and 10 separate trials. Each trial was  $\sim 90$  min long and corresponded to a particular ICP level, elevated using saline infusion. Induced ICP ranged between 5 and 60 mmHg, with natural ICP fluctuation during experimentation deviating slightly beyond these limits. The distribution of both induced and naturally fluctuating ICP levels was observed to be predominantly between 5 and 30 mmHg. Before ICP elevation, a recording at baseline ICP was also conducted.

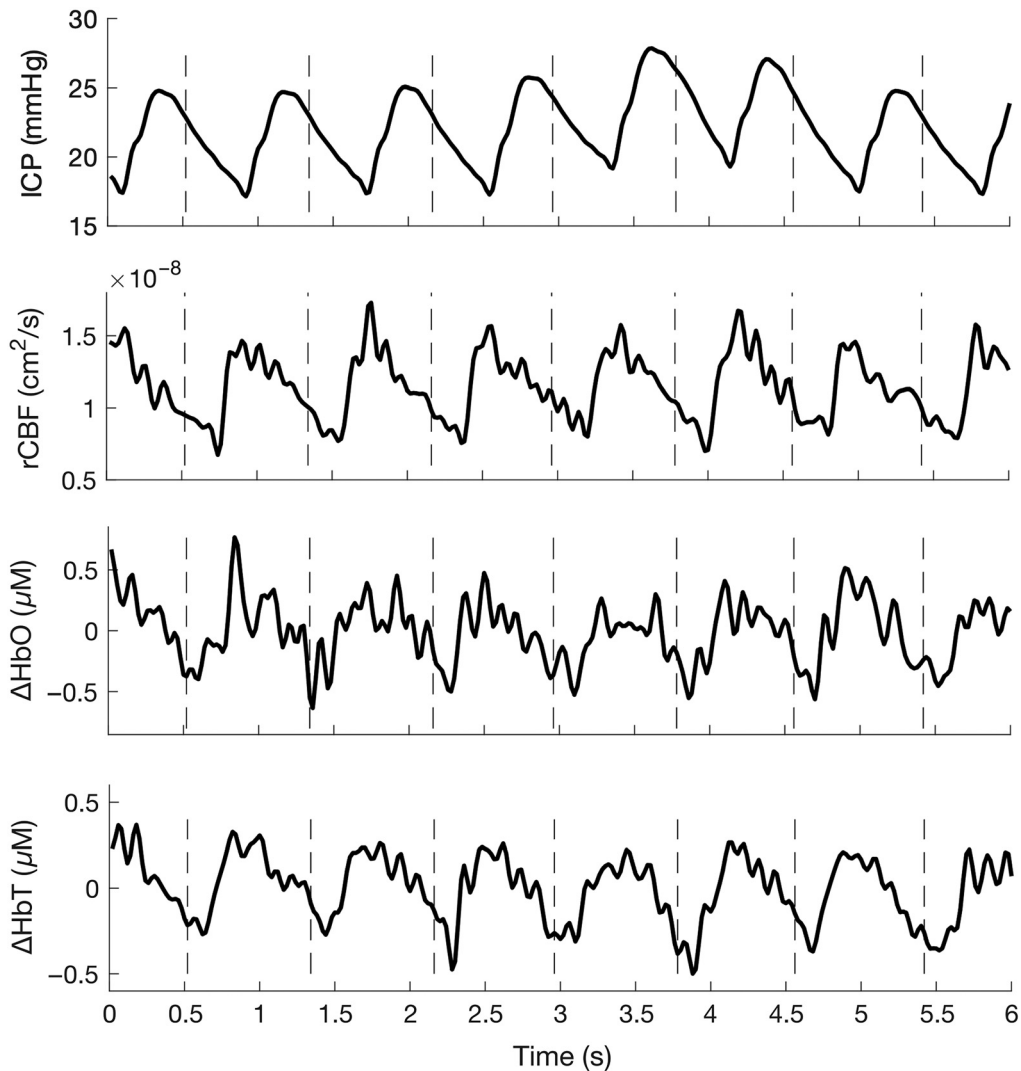
## 2.2 Signal Processing

Following the initial data collection, the signal processing procedures closely followed those outlined in Ref. 27. Analog markers in the form of voltage spikes were sent to the auxiliary ports of the DCS, NIRS, and MPR1 Datalogger devices for data alignment. In NHPs 1 to 3, an amplifier circuit was set to register a cardiac pulse when the electrocardiogram (EKG) signal exceeded an empirically determined threshold that defined the R peak of the QRS complex. When an R peak was detected, a synchronization pulse was sent to the MPR1 Datalogger. This signal was sampled at 100 Hz. For NHPs 4 to 8, EKG was measured through a separate device at 1000 Hz. A maximal overlap discrete wavelet transform was run across the EKG signal to enhance the QRS complex features of the signal (the QRS complex represents ventricular polarization). MATLAB's (MATLAB R2020b, The MathWorks Inc., Natick, Massachusetts, United States) `findpeaks` function was used to index the time point of the peak of each enhanced QRS complex. The respective indices represent the R peak of the QRS complex. During data collection, occasional laser instabilities were observed. Time points in the NIRS- and DCS-based signals, for which instabilities were visually observed, were removed. For NHP 4, the DCS-based CBF signal showed consistent instability across all trials. As a result, NHP 4 was not included in the ICP estimation analysis of CBF. All signals not originally sampled at 50 Hz (ABP, ICP, CBF, and QRS complex peak indices) were downsampled to 50 Hz to match the sampling frequency of the NIRS signal.

To reduce noise in the NIRS and DCS signals, 120 consecutive cardiac pulses were averaged, and an average cardiac waveform (ACPW) was extracted. Because the heart rate of the animals varied, 120 pulses corresponded to between 39 and 78 s across all animals. The ICP values during the 120 pulses were also averaged. The 120-pulse averaging window was moved 20 pulses at a time (resulting in an 83.3% window-to-window overlap) along the entire signal of each trial. All ACPWs were normalized in time and amplitude. With the help of spline interpolation, the length of each ACPW was normalized to 66 data points, corresponding to 1.32 s. ACPW length was measured between two consecutive diastoles. The height of each ACPW (representing the amplitude of  $\Delta\text{HbO}$ ,  $\Delta\text{HbT}$ , and CBF) was normalized to between 0 and 1. Normalization of the  $x$ - and  $y$ -axis removed ACPW length as a feature and thus removed heart rate as a feature of the waveform. We removed heart rate as a feature because heart rate changes can be independent of ICP.<sup>33–35</sup>

To remove outliers, a  $z$ -score rejection method was applied to all ACPWs. For each trial, a  $z$ -score was calculated across all averaged pulses in the trial. This averaged  $z$ -score was then compared against each ACPW. Any  $\Delta\text{HbO}$ ,  $\Delta\text{HbT}$ , or CBF ACPW with a  $z$ -score greater than 3 was rejected.

A Kalman filter was then used to further improve signal quality. The adaptive filter was applied to the ACPWs of each signal for each trial. Each ACPW was compared with the ideal trial pulse produced by the Kalman filter. ACPWs were corrected based on their error from the ideal pulse. In the correction method, the Kalman filter's parameters defined the weight given to the Kalman-produced ideal pulse and the weight given to the calculated ACPW-to-ideal-pulse



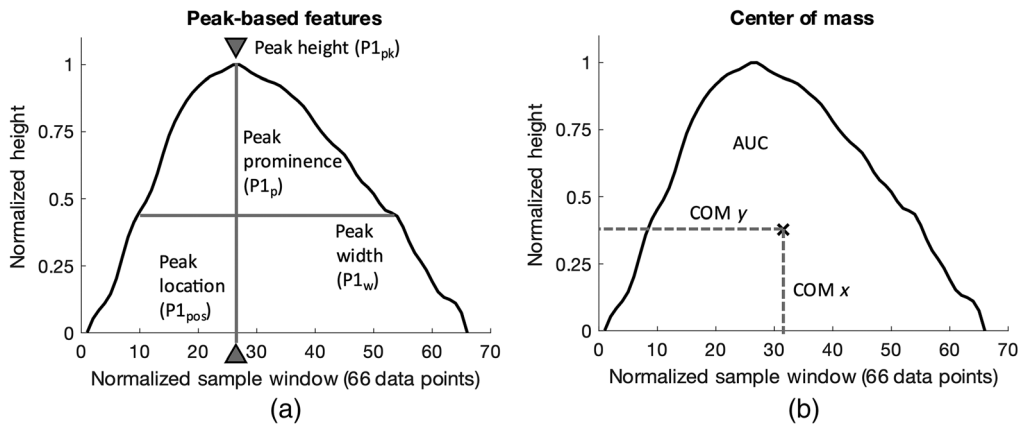
**Fig. 2** Filtered time trace examples of ICP, CBF,  $\Delta\text{HbO}$ , and  $\Delta\text{HbT}$  signals. Dashed lines indicate QRS complex peaks.

error. These parameters were set empirically. The output of the Kalman filter procedure was a set of high signal-to-noise ratio (SNR) ACPWs with feature morphology reflecting changes in ICP. ACPWs reflecting ICP values above 30 mmHg were removed due to their scarcity. Figure 2 shows examples of 50 Hz  $\Delta\text{HbO}$ ,  $\Delta\text{HbT}$ , ICP, and CBF signals. A total of 19,000  $\Delta\text{HbO}$  and 19,258  $\Delta\text{HbT}$  ACPWs sampled from 8 NHPs, along with 17,322 CBF ACPWs sampled from 7 of the same 8 NHPs were used for feature extraction. All ACPWs represented an ICP range of 0 to 30 mmHg.

### 2.3 Feature Extraction

Defining and extracting physiologically relevant features from the processed ACPWs provides the basis for estimating ICP. In our previous work, features of the CBF waveform included the percussion, tidal, and dirotic peaks of the ACPWs.<sup>27</sup> These features were not generally discernable from the NIRS-based hemoglobin concentration signals. As such, the feature set and their extraction method were modified, as shown in Fig. 3. Figure 3 also shows an example of a  $\Delta\text{HbO}$  ACPW.

MATLAB's findpeaks function was used to obtain the peak height ( $P1_{pk}$ : normalized to 1 and used as a measure for model noise), position ( $P1_{pos}$ ), prominence ( $P1_p$ ), and width ( $P1_w$ ) of the ACPWs.



**Fig. 3** (a) Peak-based features, including peak height ( $P1_{pk}$ ) (normalized to 1), peak width ( $P1_w$ ), peak prominence ( $P1_p$ ), and peak location ( $P1_{pos}$ ), for a typical  $\Delta\text{HbO}$  ACPW. (b)  $x$ - and  $y$ -coordinates of the center of mass of the same ACPW. MAP was used as an additional feature, not shown here.

The  $x$ - and  $y$ -coordinates of the center of mass (centroid) of the waveform, as shown in Fig. 3(b), defined as  $\text{COM}_x$  and  $\text{COM}_y$ , were also extracted from individual ACPWs and used as features. This was done by turning each ACPW into a polygon and calculating its centroid using the `polyshape` and `centroid` functions in MATLAB. We hypothesized that  $\text{COM}_x$ , which describes waveform skewness, is related to blood pressure and ICP. A similar reasoning motivated  $\text{COM}_y$ . The area under the curve (AUC) of the waveform was also incorporated as a feature, as was MAP. Similar to ICP, MAP was calculated over each 120-pulse window. The feature engineering resulted in eight human-interpretable and observable features for each processed  $\Delta\text{HbO}$ ,  $\Delta\text{HbT}$ , and CBF ACPW. If a feature was undetected, its value was set to 0, but it was still used.

## 2.4 Machine Learning

Each of the three ACPW datasets was randomly sampled into five cross validation (CV) sets of 80% training and 20% testing. For each CV set, all animals were included. Random sampling ensured that learning became NHP and trial independent, while CV allowed us to alleviate overfitting. Python's `scikit-learn` toolbox's random forest (RF) regression algorithm was used as the ICP estimator.<sup>36</sup> The RF algorithm, explained in detail by Ref. 37, learns a set number of decision trees on a randomly sampled subset of the features using a randomly sampled subset of the training data with replacement (bootstrapping). A total of 100 trees, or estimators, were learned using this bootstrapping method. Every tree received 50% of the features and 80% of the dataset for training. All other hyperparameters were kept as default, defined in the `scikit-learn` documentation.<sup>36</sup> These hyperparameters were chosen empirically to maximize performance while mitigating overfitting. The hyperparameter decision-making process used a random search to gauge approximate hyperparameter ranges, after which a per-parameter and joint-parameter optimization procedure followed. Random search hyperparameter tuning is a common method used for optimizing machine learning models.<sup>38</sup> One important parameter was the number of estimators, or trees. Each of the 100 trees splits its subset of data until all leaves are pure—to max depth. Each tree receives four randomly sampled features for learning. Gini impurity was used as the measure of node split quality.<sup>39</sup> For each CV split (fold), the RF algorithm was trained on the training split and tested on the testing split. During testing, estimated ICP values ( $\text{ICP}_{\text{est}}$ ) were compared with invasively measured ICP values ( $\text{ICP}_{\text{inv}}$ ).  $\text{ICP}_{\text{inv}}$  values were used as ground truth labels. The performance of the model was quantified using the coefficient of determination ( $r^2$ ), mean squared error (MSE), and 95% confidence interval (CI).

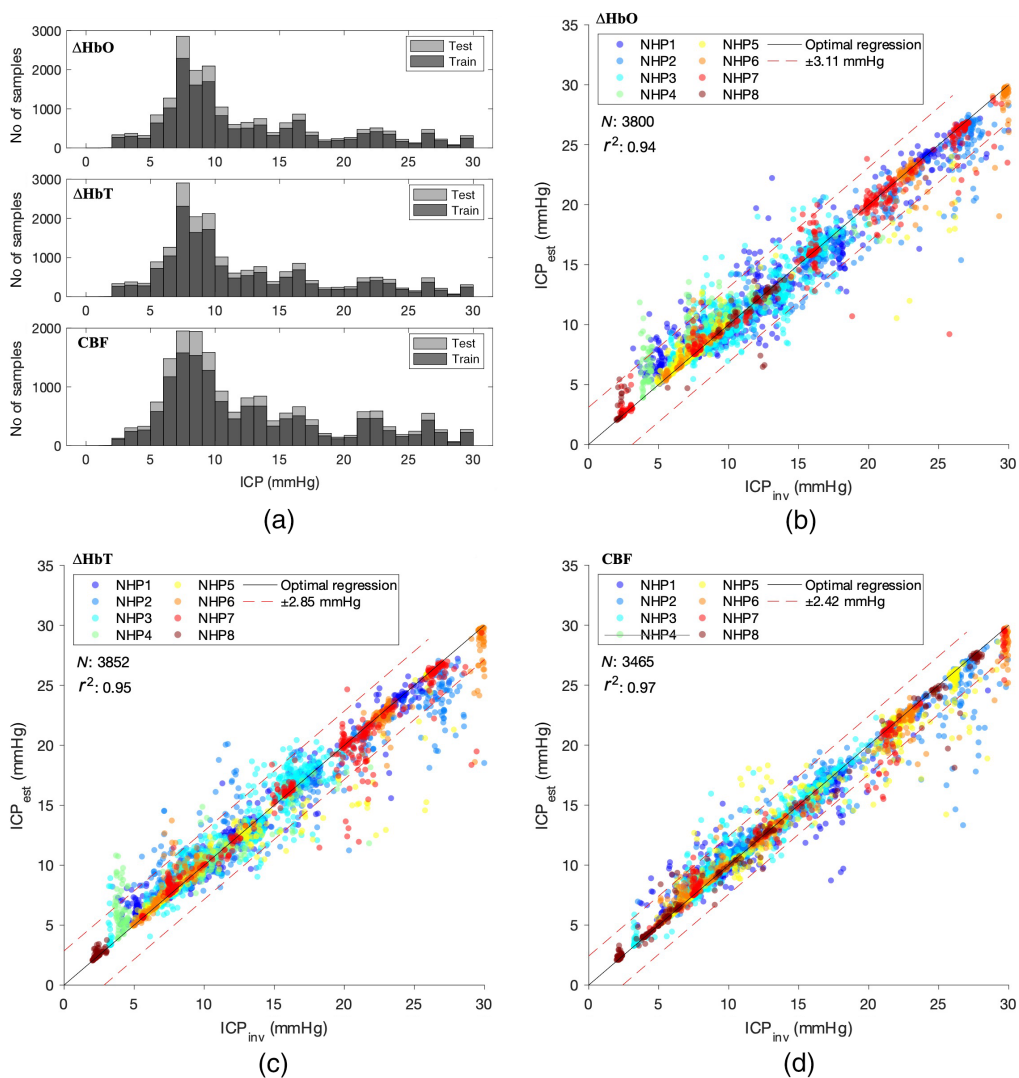
### 3 Results

#### 3.1 Waveform Features and RF Regressors

We have previously found that the CBF ACPW broadens and skews to the right when ICP increases.<sup>27</sup> We qualitatively observed similar trends in the  $\Delta\text{HbO}$  and  $\Delta\text{HbT}$  ACPWs.

Using an RF regressor with modified hyperparameters, we evaluated the ICP estimator's performance on NIRS-derived ACPWs ( $\Delta\text{HbO}$  and  $\Delta\text{HbT}$ ) against its performance on DCS-based ACPWs (CBF). Although the CBF-based waveforms showed more morphological features compared with the  $\Delta\text{HbO}$ - and  $\Delta\text{HbT}$ -based waveforms, the features chosen were the same between all three modalities used. Within the 0 to 30 mmHg range of ICP values, available training and testing data were skewed toward lower ICP values, as shown in Fig. 4(a).

To compare performance differences between the three different types of signals, the coefficient of determination ( $r^2$ ) and MSE were used as evaluation metrics, along with a 95% CI. A fivefold CV was performed individually on the  $\Delta\text{HbO}$ ,  $\Delta\text{HbT}$ , and CBF datasets. We found a mean fold  $r^2 = 0.937$  (averaged over five folds),  $r^2 = 0.946$ , and  $r^2 = 0.963$  for the ACPW

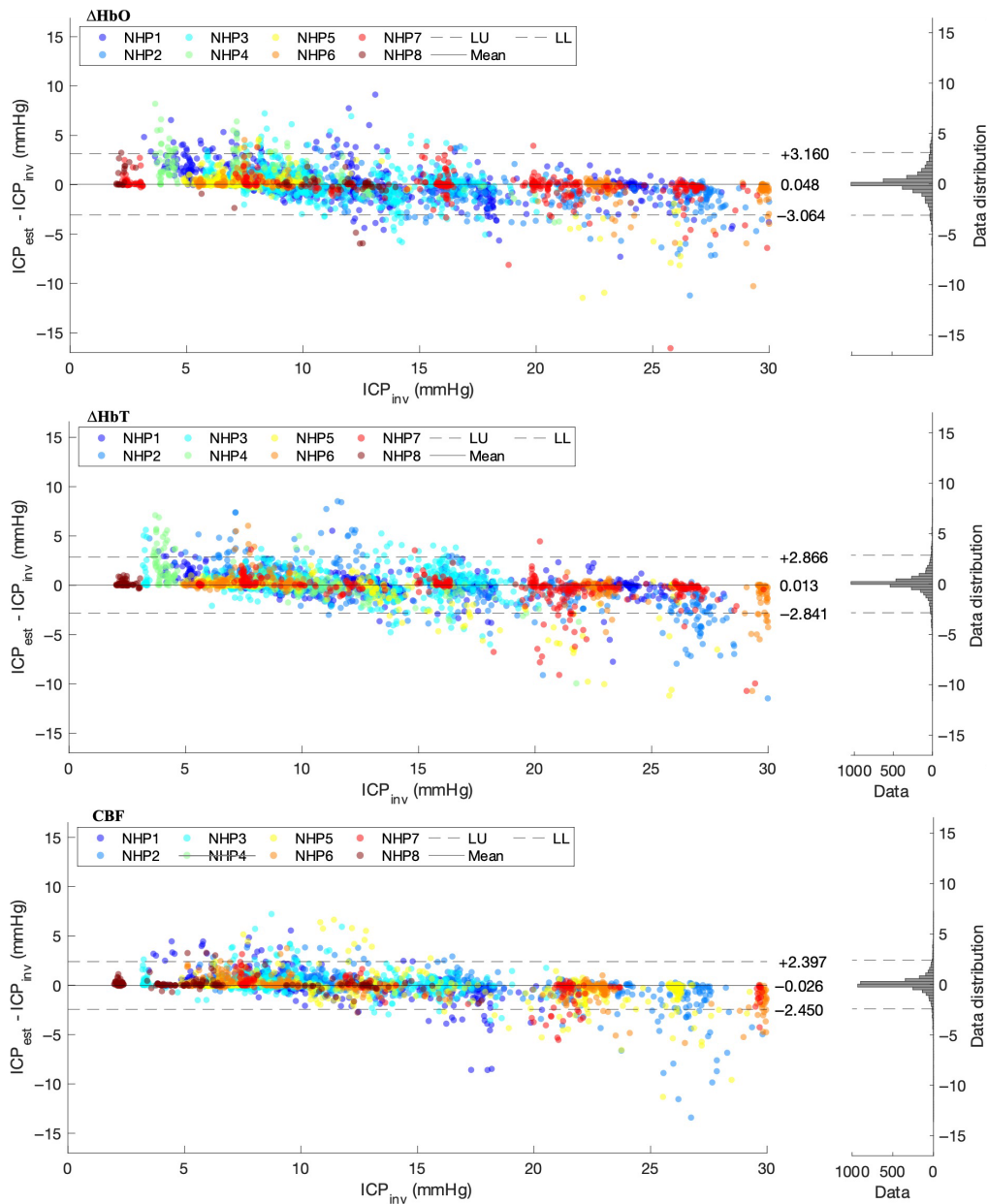


**Fig. 4** (a) Histogram of ICP distribution for both train and test sets across all three modalities. More data were available at lower ICP values, especially between 5 and 10 mmHg. (b)–(d) Correlation plots for  $\Delta\text{HbO}$ ,  $\Delta\text{HbT}$ , and CBF. Strong  $r^2$  for all methods suggests that the model performs well on ICP estimation. Estimation performance drops for higher ICP values across all three modalities due to a lower availability of high ICP data for training.

types  $\Delta\text{HbO}$ ,  $\Delta\text{HbT}$ , and CBF, respectively, with a fivefold  $r^2_{\text{std}}$  standard deviation of  $r^2_{\text{std}} = 0.003$  ( $\Delta\text{HbO}$ ),  $r^2_{\text{std}} = 0.004$  ( $\Delta\text{HbT}$ ), and  $r^2_{\text{std}} = 0.002$  (CBF). We further found a mean fold MSE = 2.703, 2.301, and 1.688  $\text{mmHg}^2$  with a fivefold standard deviation of  $\text{MSE}_{\text{std}} = 0.133$ , 0.163, and 0.131  $\text{mmHg}^2$ , respectively. Our results indicate that CBF waveforms perform better than NIRS-derived parameters in terms of ICP extraction, but ICP can still be quantified within an MSE of  $<3 \text{ mmHg}^2$  when using hemoglobin-based waveforms.

The correlation plots between  $\text{ICP}_{\text{est}}$  and  $\text{ICP}_{\text{inv}}$  for all three signals are shown in Figs. 4(b)–4(d). All three ICP extractions show a good correlation between estimated and invasively measured ICP. Outliers were more common at higher ICP values for which less data were available for training, as shown in the histograms of Fig. 4(a).

A set of three Bland–Altman plots, indicating the level of agreement between  $\text{ICP}_{\text{est}}$  and  $\text{ICP}_{\text{inv}}$ , are shown in Fig. 5 and further confirm the clear fit of our regression model to the invasively measured ICP. The plots show a 95% CI of agreement between  $\text{ICP}_{\text{est}}$  and  $\text{ICP}_{\text{inv}}$  of



**Fig. 5** Bland–Altman plots for  $\Delta\text{HbO}$ ,  $\Delta\text{HbT}$ , and CBF. Histograms show the distribution of data-points across the Bland–Altman plot.



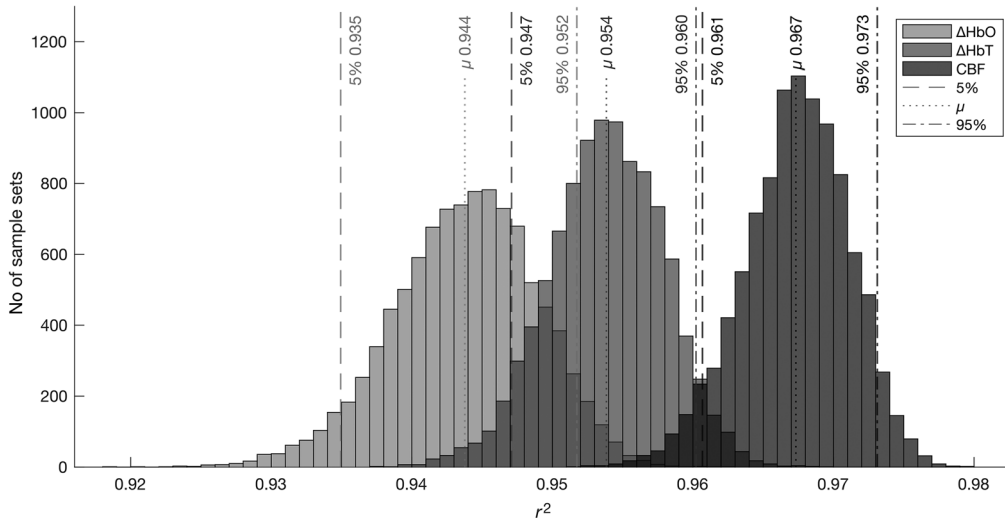
**Table 1** Error metrics across the three modalities.

Modality	$r^2$	MSE	$r^2_{std}$	MSE <sub>std</sub>	95% CI	Mean
CBF	0.963	1.688	0.002	0.131	[-2.450 2.397]	-0.026
$\Delta$ HbT	0.946	2.301	0.004	0.163	[-2.841 2.866]	0.013
$\Delta$ HbO	0.937	2.703	0.003	0.133	[-3.064 3.160]	0.048

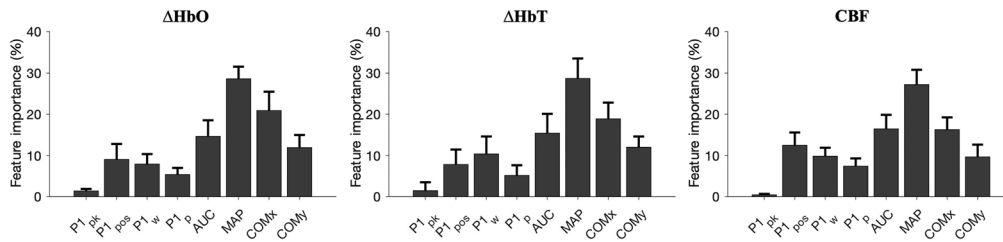
[-3.064 3.160] mmHg with a mean of 0.048 mmHg for  $\Delta$ HbO, a 95% CI of [-2.841 2.866] mmHg with a mean of 0.013 mmHg for  $\Delta$ HbT, and a 95% CI of [-2.450 2.397] mmHg with a mean of -0.026 mmHg for CBF.

Table 1 summarizes the results across all three modalities. When considering the performance of  $\Delta$ HbO against  $\Delta$ HbT and CBF, we find that  $\Delta$ HbT performs slightly better than  $\Delta$ HbO.  $\Delta$ HbO and  $\Delta$ HbT have a 0.009 difference in fivefold mean  $r^2$  and a 0.402 mmHg<sup>2</sup> difference in fivefold mean MSE. CBF performs better than the two NIRS-based methods with an increase of 0.025 fivefold mean  $r^2$  and a 0.402 mmHg<sup>2</sup> decrease in fivefold mean MSE against  $\Delta$ HbO. The 95% CI metric across the modalities echoes this difference in model fit performance. Using bootstrap analysis in which 50% of test samples were randomly drawn 10,000 times with replacement and a respective  $r^2$  was calculated for each set of samples, we found that CBF performs statistically better than either of the NIRS methods (significance level ( $\alpha$ )  $\leq 0.05$ ). Despite this, NIRS performs comparatively well against DCS, along with being cheaper and easier to use. Distributions of  $r^2$  scores with 5%, mean ( $\mu$ ), and 95% thresholds are shown in Fig. 6.

For all three modalities, the feature relevance contributing to ICP extraction was evaluated. The relative magnitude of feature importance is shown in Fig. 7. We found that MAP and COM<sub>x</sub> were the most relevant features used in the training of each estimator tree. AUC, COM<sub>x</sub>, and MAP accounted for ~15%, ~20%, and ~28% of node splits per tree, respectively. Meanwhile, COM<sub>y</sub> represented ~11% of node splits, whereas waveform peak position represented ~9%. The remaining three features represented <7% of node splits per tree. The standard deviation of features used between trees is shown by the error bars of each feature bar. AUC importance for



**Fig. 6** Histograms of bootstrapped  $r^2$  scores for all three modalities. Bootstrapping used 50% of test samples drawn 10,000 times with replacement. The test data from the best performing CV split, per modality, are used. This results in sample mean  $r^2$  slightly above the mean CV  $r^2$  reported.  $r^2$  is calculated for each set of samples across all three modalities ( $\Delta$ HbO,  $\Delta$ HbT, and CBF). A significance level  $\leq 0.05$  is used. CBF statistically outperforms both NIRS methods. NIRS methods are still comparable to CBF in overall performance.



**Fig. 7** Comparison of feature importance across  $\Delta\text{HbO}$ ,  $\Delta\text{HbT}$ , and CBF. MAP is strongest across all three modalities. AUC and COM<sub>x</sub> are relevant for estimation, and peak-based features ( $P1_{\text{pos}}$ ,  $P1_{\text{w}}$ , and  $P1_{\text{p}}$ ) are more important in CBF-based ICP estimation compared with the hemoglobin-based methods. It is also important to note that  $P1_{\text{pk}}$  is low for all methods, especially for CBF. This suggests that the model does not learn irrelevant features.

$\Delta\text{HbO}$ ,  $\Delta\text{HbT}$ , and CBF was approximately the same at around 15% to 18%, whereas peak-related features (peak location, width, and prominence) were more important for CBF compared with the NIRS methods. COM<sub>x</sub> and COM<sub>y</sub> were less important for  $\Delta\text{HbT}$  and CBF compared with  $\Delta\text{HbO}$ . MAP was the most important feature across all modalities, with the level of importance being approximately the same. For  $\Delta\text{HbT}$ , MAP represented  $\sim 29\%$  of splits per tree compared with  $\sim 28\%$  for  $\Delta\text{HbO}$  and  $\sim 27\%$  for CBF. The importance of the height of the peak, which was normalized, is used as a proxy to measure the importance of an unrelated peak. The apparent relevance of this unrelated peak is highest for  $\Delta\text{HbO}$  and lowest for CBF, suggesting that the model learns slightly more uncorrelated patterns on the  $\Delta\text{HbO}$  dataset compared with the datasets of the other modalities.

A more evenly balanced dataset was also tested for  $\Delta\text{HbO}$ . Balancing was performed by randomly removing half of all ICP values between 5 and 10 mmHg. This balanced dataset had  $\sim 23\%$  less data than the skewed dataset. The balanced dataset led to a marginally worse performance compared with the skewed dataset and was not studied further for either  $\Delta\text{HbT}$  or CBF. For  $\Delta\text{HbO}$ ,  $r^2$  dropped by 0.6% from 0.937 to 0.932, MSE increased by 22.1% from 2.703 to 3.299 mmHg<sup>2</sup>, and mean 95% CI increased by [12.3% 11.5%] from  $[-3.064\ 3.160]$  to  $[-3.441\ 3.523]$  mmHg with a 14.6% change in mean from 0.048 to 0.041 mmHg. Because MAP is a dominant feature for all three signals used, we evaluated whether MAP can be removed to make DCS and NIRS standalone methods for ICP extraction. When MAP was removed as a training feature, the importance of the four peak features was increased, and the importance of AUC and COM<sub>x</sub> was maintained. With MAP removed, performance fell to  $r^2 = 0.838$ , MSE = 6.971 mmHg<sup>2</sup>, and 95% CI of  $[-4.960\ 4.984]$  mmHg with a mean of 0.012 mmHg for  $\Delta\text{HbO}$ ;  $r^2 = 0.865$ , MSE = 5.791 mmHg<sup>2</sup>, and 95% CI of  $[-4.556\ 4.700]$  mmHg with a mean of 0.072 mmHg for  $\Delta\text{HbT}$ ; and  $r^2 = 0.916$ , MSE = 3.778 mmHg<sup>2</sup>, and 95% CI of  $[-3.785\ 3.756]$  mmHg with a mean of  $-0.014$  mmHg for CBF. These results indicate that the performance is improved with MAP as an added feature, but ICP can still be quantified to  $< \sim 5$  mmHg for hemoglobin-based estimators and  $< \sim 4$  mmHg for CBF-based estimators.

## 4 Discussion

The goal of this work was to demonstrate that noninvasive ICP measurements are possible when using hemoglobin concentration changes as measured with NIRS. Here we showed that, although the NIRS-based estimate is worse than one obtained using DCS, ICP could still be estimated within  $\sim 3$  mmHg in eight NHPs for an ICP range of between 0 and 30 mmHg. The performance of the estimator used in this manuscript improves on that in our previous work with DCS devices<sup>27</sup> due to the larger dataset and the refined regression model. The results suggest that NIRS could be an alternative to DCS in terms of noninvasive ICP.

### 4.1 Dataset

In building an NHP-based dataset of ICP-dependent ACPW features, it was observed that  $\Delta\text{HbO}$  and  $\Delta\text{HbT}$  cardiac pulse waveforms followed a canonical cardiac pulse arch shape. Using

120-pulse averaging and Kalman filtering improved SNR to the extent that meaningful feature extraction could be performed. The number of pulses averaged over was set to 120 to include enough pulses for appropriate waveform and ICP averaging. Averaging too many pulses would likely lead to lower ICP-to-waveform-feature correlation, whereas not averaging enough pulses would not remove enough noise in the signal waveforms.

Despite selecting only ICP values up to 30 mmHg, the distribution of data was not equal for all ICP values, as shown in the histogram in Fig. 4(a). We found that balancing the dataset did not significantly improve ICP estimation. There appeared to be a more even distribution of outliers in ICP estimation between low and high ICP (0 to 30 mmHg) in the balanced dataset, compared with the skewed dataset (data not shown). Meanwhile, feature importance was not significantly different between the skewed and balanced data. It can be argued that the unbalanced dataset was more likely to cause overfitting to lower ICP values and to artificially inflate performance. Although possible, we do not have enough data points at higher (>20 mmHg) ICP values to quantify this.

## 4.2 Preprocessing and Filtering

To appropriately leverage the raw signals extracted from each NHP, signal preprocessing and filtering were required. The goal was to provide the feature extraction method with low-noise signals while not overfiltering or losing valuable information. Clean signals lead to relevant feature-value-to-ICP mappings for the machine learning model to learn.

Tuning of the signal preprocessing and filtering parameters was done empirically across the cardiac waveform averaging, waveform normalization,  $z$ -score rejection, and Kalman filtering methods. 120 consecutive pulses were found to be a suitable averaging amount to reduce signal to noise, whereas a 20-pulse averaging window struck a balance between obtaining ample amounts of data for training and testing while reducing data replication. Waveform normalization,  $z$ -score rejection, and Kalman filtering all worked to remove data outliers while maintaining a feature distribution that supported model generalization.

## 4.3 Feature Engineering

An investigation into feature importance and performance, as shown in Fig. 7, suggests that MAP on its own, though relevant, is not the only feature driving decision tree learning. This gives weight to our other chosen features, such as AUC and COM $x$ . We also found poor correlation between the raw values of MAP, AUC, and COM $x$  with ICP. This suggests that a feature on its own is not a good indicator of ICP, giving strength to our approach of combining several features in a machine learning method to accurately estimate ICP. Interpretable feature engineering is an important part of clinically relevant and understandable machine learning tools. Explainable models ensure that conclusions made by the model are based on appropriate decisions.<sup>40</sup>

Extracting specific features from data may improve model generalizability for small sample sizes. The FDA's *Good Machine Learning Practice for Medical Device Development: Guiding Principles*<sup>41</sup> highlights the importance of human interpretability of models and their outputs in clinical settings. As such, the features that we engineered are more interpretable when compared with machine-engineered features sometimes used in complex machine learning and neural network-based methods.<sup>29</sup> We also found that our engineered features expressed themselves relatively predictably across modalities.

## 4.4 RF Regression

In our previous work on DCS-based ICP extraction, we showed a cross-validated similarity between estimated ICP and invasively measured ICP of  $r^2 = 0.91$ , an MSE of 3.3 mmHg, and a 95% CI of  $[-3.7 \ 3.7]$  mmHg. Here we compared NIRS-derived waveforms to DCS waveforms for ICP extraction and achieved a cross-validated similarity between DCS-based ICP estimation and invasively measured ICP of  $r^2 = 0.96$ , MSE of 1.7 mmHg, and 95% CI of  $[-2.450 \ 2.397]$  mmHg. This is an improvement to our previous work. The performance increase

was likely due to using more data and a revised machine learning method. Compared with Ruesch et al.'s 14,121 ACPWs across 5 NHPs,<sup>27</sup> our CBF dataset contained 17,322 ACPWs sampled from 7 NHPs. The ICP distributions were similar. Regarding the ICP estimators, Ruesch et al.<sup>27</sup> used a bagged ensemble of trees with subsampling of data. Unlike our RF model, which was chosen for its versatility, robustness, resistance to noise, and ability to manage unbalanced datasets,<sup>42–45</sup> the bagged tree regressor of Ruesch et al.<sup>27</sup> did not randomly subsample features for each tree in the ensemble. In an RF regression, the random sampling of features ensures that no one tree receives the full set of features for learning. This combats overfitting and improves performance.

RF also have a relatively low risk of overfitting with an increase in estimators due to their use of multiple weak and unpruned learners.<sup>46</sup> Each tree in the RF ensemble does overfit to the data and the features that it receives, but the averaged result of the ensembled trees produces a regression prediction that is low in variance and bias. Ultimately, this means that the more trees that are used in the ensemble, the less likely the model is to overfit. More trees, however, increase computational cost. Compared with Ruesch et al.'s 1000 trees,<sup>27</sup> we achieved comparable results with only 100 trees. This means that our approach is more computationally efficient with comparable results.

It is important to note that the signal processing, feature extraction, and training methods used on the three signal types explored in this paper ( $\Delta\text{HbO}$ ,  $\Delta\text{HbT}$ , and CBF) were the same and that the features were specifically engineered with NIRS-based signals in mind. DCS-based (CBF) ICP estimation with these specific features was relatively more dependent on peak-based features than the  $\Delta\text{HbO}$  and  $\Delta\text{HbT}$  ICP estimation, but not to the same extent as observed in Ruesch et al.<sup>27</sup> Meanwhile, features such as AUC,  $\text{COM}_x$ , and  $\text{COM}_y$  were important across the modalities. MAP was the defining feature for all signal types, with varying degrees of relative importance compared with other features. This suggests that the chosen features are potentially useful across signal and sensor modalities. CBF waveforms, however, do carry more waveform-related information compared with waveforms extracted using blood volume changes. Thus, waveform-based features are more relevant for CBF compared with  $\Delta\text{HbO}$  and  $\Delta\text{HbT}$ . The features chosen for this work may therefore not be optimized for CBF, and further improvement in ICP extraction may be possible.

#### 4.5 Feature Differences Across NHPs

Occasionally observed differences in the morphology of the same waveform feature across subjects are likely due in part to influences on the NHPs' natural autoregulation and their abilities to manage varying ICPs. Because ISO gas, commonly used during anesthesia, is known to suppress CA,<sup>47,48</sup> we used <1% to avoid anesthesia induced autoregulation impairment.<sup>30</sup> However, changes in ICP led to CPP levels outside the autoregulatory range.<sup>8,9,30</sup> The Windkessel effect—the elasticity of arteries that moderates blood pressure changes during a cardiac cycle—may also be a factor in observed morphological differences. Studies show that conduit arteries harden with age.<sup>49</sup> The stiffening of cerebral arteries results in an increased pulse pressure per stroke volume and a higher systolic pressure. This effect has primarily been studied in humans, but the effect may be similar in the NHPs used and could therefore contribute to inconsistent morphology of volume-dependent cardiac pulse waveforms between NHPs of different ages. Our youngest NHP was 3 years old, and our oldest was 10 years old. Macaques reach sexual maturity at 3 to 5 years (REFS) and can live to 25 years or more in captivity.<sup>50,51</sup>

Small differences in probe and catheter placement between NHPs were also likely contributors to observed feature dissimilarity and variance between  $\text{ICP}_{\text{inv}}$  and  $\text{ICP}_{\text{est}}$ . Although care was taken in the placement of the NIRS device on the skull of the NHPs, small variations did exist between subjects. In addition, ICP sensor and ventricular catheter placement varied slightly due in part to distinctive subject physiologies.

#### 4.6 Future Work

One avenue of future study along model selection could investigate probabilistic RF (PRF). Reis et al.<sup>52</sup> showed that, compared with normal RF, PRFs improved classification by 10% on noisy features and 30% on noisy labels. This model may also be useful for regression. Natural

continuations of this study would include large-scale human validation. Due to the relatively small sample size, we did not attempt a leave-one-out analysis. All animals were represented in both the training and the testing sets. A meaningful assessment of the generalizability of our ICP estimator to unseen data would be facilitated by a larger dataset.

An additional future consideration could evaluate the importance of the location of MAP estimation. In this manuscript, we measured MAP from the carotid artery. An alternative approach may be the use of a blood pressure cuff, which hypothetically provides a sufficient MAP reading to be used with our approach. This, however, has not yet been evaluated. If successful, an MAP reading using a blood pressure cuff would ease clinical translation of our noninvasive ICP estimation method.

## 5 Conclusions

This work demonstrates that ICP can be accurately predicted using an RF regressor trained on cardiac pulse waveform features extracted from hemoglobin-based signals using NIRS or DCS. It also shows that ICP-dependent cardiac pulse waveform morphology is discernible in hemoglobin-based signals sampled at 50 Hz. We compared the performance of our algorithm on  $\Delta\text{HbO}_2$ ,  $\Delta\text{HbT}$ , and CBF-based data and found comparable accuracy when using features optimized for hemoglobin concentration changes. Our approach is EKG and MAP dependent, both of which are regularly used and measured in clinical settings. An increased NHP sample size, human trials and a balanced representation of ICP values are important to validating our findings for the clinical setting.

## Disclosures

No conflicts of interest are declared.

## Acknowledgments

The authors acknowledge the financial support for the research, authorship, and publication through the Center for Machine Learning and Health (CMLH), Pennsylvania Infrastructure Technology Alliance (PITA), American Heart Association (AHA) 17SDG33700047, and National Institutes of Health (NIH) R21-EB024675. The authors would like to thank ZOLL Medical Corporation (Chelmsford, Massachusetts, United States) for lending the programmable ventilator.

## Code, Data, and Materials Availability

Code and data are available upon request.

## References

1. D. Roytowski and A. Figaji, "Raised intracranial pressure: what it is and how to recognise it," *Contin. Med. Educ.* **31**(11), 390–395 (2013).
2. M. H. Wilson, "Monro-Kellie 2.0: the dynamic vascular and venous pathophysiological components of intracranial pressure," *J. Cereb. Blood Flow Metab.* **36**(8), 1338–1350 (2016).
3. M. Smith, "Monitoring intracranial pressure in traumatic brain injury," *Anesth. Analg.* **106**(1), 240–248 (2008).
4. A. Chari et al., "Intraparenchymal intracranial pressure monitoring for hydrocephalus and cerebrospinal fluid disorders," *Acta Neurochirurg.* **159**(10), 1967–1978 (2017).
5. A. B. Peterson and K. E. Thomas, "Incidence of nonfatal traumatic brain injury-related hospitalizations—United States, 2018," *Morb. Mortal. Wkly. Rep.* **70**(48), 1664–1668 (2021).

6. Hydrocephalus Association, “Congenital hydrocephalus,” <https://www.hydroassoc.org/congenital-hydrocephalus/> (accessed 15 December 2021).
7. P. M. Lewis et al., “Dynamic cerebral autoregulation: should intracranial pressure be taken into account?” *Acta Neurochir.* **149**(6), 549–555 (2007).
8. N. A. Lassen, “Cerebral blood flow and oxygen consumption in man,” *Physiol. Rev.* **39**(2), 183–238 (1959).
9. N. A. Lassen and M. S. Christensen, “Physiology of cerebral blood flow,” *Br. J. Anaesth.* **48**, 719–734 (1976).
10. D. Acharya et al., “Changes in neurovascular coupling with cerebral perfusion pressure indicate a link to cerebral autoregulation,” *J. Cereb. Blood Flow Metab.* **42**(7), 1247–1258 (2022).
11. S. T. Fantini et al., “Cerebral blood flow and autoregulation: current measurement techniques and prospects for noninvasive optical methods,” *Neurophotonics* **3**(3), 031411 (2016).
12. L. C. Padayachy, A. A. Figaji, and M. R. Bullock, “Intracranial pressure monitoring for traumatic brain injury in the modern era,” *Child’s Nerv. Syst.* **26**(4), 441–452 (2010).
13. C. Wiegand and P. Richards, “Measurement of intracranial pressure in children: a critical review of current methods,” *Dev. Med. Child Neurol.* **49**(12), 935–941 (2007).
14. P. H. Raboel et al., “Intracranial pressure monitoring: invasive versus non-invasive methods—a review,” *Crit. Care Res. Pract.* **2012**, 950393 (2012).
15. N. Lenfeldt et al., “CSF pressure assessed by lumbar puncture agrees with intracranial pressure,” *Neurology* **68**(2), 155–158 (2007).
16. M. Czosnyka, “Monitoring and interpretation of intracranial pressure,” *J. Neurol. Neurosurg. Psychiatry* **75**(6), 813–821 (2004).
17. T. W. Langfitt et al., “Transmission of increased intracranial pressure. I. Within the cranio-spinal axis,” *J. Neurosurg.* **21**, 989–997 (1964).
18. P. J. D. Andrews and G. Citerio, “Intracranial pressure: part one: historical overview and basic concepts,” *Intensive Care Med.* **30**(9), 1730–1733 (2004).
19. T. Heldt et al., “Intracranial pressure and intracranial elastance monitoring in neurocritical care,” *Annu. Rev. Biomed. Eng.* **21**, 523–549 (2019).
20. A. Holdgate and K. Cuthbert, “Perils and pitfalls of lumbar puncture in the emergency department,” *Emerg. Med.* **13**(3), 351–358 (2001).
21. J. Klingelhöfer et al., “Evaluation of intracranial pressure from transcranial Doppler studies in cerebral disease,” *J. Neurol.* **235**(3), 159–162 (1988).
22. N. G. Rainov, J. B. Weise, and W. Burkert, “Transcranial Doppler sonography in adult hydrocephalic patients,” *Neurosurg. Rev.* **23**(1), 34–38 (2000).
23. P. W. Hanlo et al., “Value of transcranial Doppler indices in predicting raised ICP in infantile hydrocephalus,” *Child’s Nerv. Syst.* **11**(10), 595–603 (1995).
24. D. Cardim et al., “Non-invasive monitoring of intracranial pressure using transcranial Doppler ultrasonography: is it possible?” *Neurocrit. Care* **25**(3), 473–491 (2016).
25. J. B. Rosenberg et al., “Non-invasive methods of estimating intracranial pressure,” *Neurocrit. Care* **15**(3), 599–608 (2011).
26. A. Reid et al., “The relationship between intracranial pressure and tympanic membrane displacement,” *Br. J. Audiol.* **24**(2), 123–129 (2009).
27. A. Ruesch et al., “Estimating intracranial pressure using pulsatile cerebral blood flow measured with diffuse correlation spectroscopy,” *Biomed. Opt. Express* **11**(3), 1462 (2020).
28. A. 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), 2304–2314 (2020).
29. J. B. Fischer et al., “Non-invasive estimation of intracranial pressure by diffuse optics: a proof-of-concept study,” *J. Neurotrauma.* **37**(23), 2569–2579 (2020).
30. A. Ruesch et al., “Comparison of static and dynamic cerebral autoregulation under anesthesia influence in a controlled animal model,” *PLoS One* **16**(1), e0245291 (2021).
31. A. Sassaroli and S. Fantini, “Comment on the modified Beer-Lambert law for scattering media,” *Phys. Med. Biol.* **49**(14), N255 (2004).

32. T. Durduran and A. G. Yodh, "Diffuse correlation spectroscopy for non-invasive, micro-vascular cerebral blood flow measurement," *Neuroimage* **85**, 51–63 (2014).
33. X. Hu et al., "Characterization of interdependency between intracranial pressure and heart variability signals: a causal spectral measure and a generalized synchronization measure," <http://ieeexplore.ieee.org>
34. S. Dinallo and M. Waseem, "Cushing reflex," *StatPearls* (2021), <https://www.ncbi.nlm.nih.gov/books/NBK549801/> (accessed 8 March 2022).
35. G. Edholm, "The relation of heart rate to intracranial pressure," *J. Physiol.* **98**, 442–445 (1940).
36. F. Pedregosa et al., "Scikit-learn: machine learning in Python," *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
37. L. Breiman, "Random forests," *Mach. Learn.* **45**(1), 5–32 (2001).
38. F. Hutter, L. Kotthoff, and J. Vanschoren, "The Springer series on challenges in machine learning automated machine learning methods, systems, challenges," <http://www.springer.com/series/15602> (accessed 30 March 2022).
39. B. H. Menze et al., "A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data," *BMC Bioinf.* **10**, 213 (2009).
40. M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you? Explaining the predictions of any classifier," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, pp. 1135–1144 (2016).
41. U.S. Food & Drug Administration, "Artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD) action plan," (2021), [www.fda.gov](http://www.fda.gov) (accessed 13 May 2021).
42. G. Biau and E. Scornet, "A random forest guided tour," *Test* **25**(2), 197–227 (2016).
43. M. Khalilia, S. Chakraborty, and M. Popescu, "Predicting disease risks from highly imbalanced data using random forest," *BMC Med. Inf. Decision Making* **11**, 51 (2011).
44. C. Strobl et al., "Bias in random forest variable importance measures: illustrations, sources and a solution," *BMC Bioinf.* **8**, 25 (2007).
45. T. Hengl et al., "Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables," *PeerJ* **6**(8), e5518 (2018).
46. Y. Qi, *Random Forest for Bioinformatics* (2012).
47. S. Yakan and V. Duzguner, "The effects of anesthetic agents on oxidative stress," *AIP Conf. Proc.* **1726**, 171–172 (2016).
48. A. Dagal and A. M. Lam, "Cerebral autoregulation and anesthesia," *Curr. Opin. Anaesthesiol.* **22**(5), 547–552 (2009).
49. S. E. Greenwald, "Ageing of the conduit arteries," *J. Pathol.* **211**(2), 157–172 (2007).
50. L. Kulik et al., "Sex differences in the development of social relationships in Rhesus Macaques (*Macaca mulatta*)," *Int. J. Primatol.* **36**(2), 353–376 (2015).
51. C. M. Luetjens and G. F. Weinbauer, "Functional assessment of sexual maturity in male macaques (*Macaca fascicularis*)," *Regul. Toxicol. Pharmacol.* **63**(3), 391–400 (2012).
52. I. Reis, D. Baron, and S. Shahaf, "Probabilistic random forest: a machine learning algorithm for noisy data sets," *Astron. J.* **157**(1), 16 (2019).

**Filip A. J. Relander** received his bachelor's degree in physics from Cornell University. Currently, he is pursuing a master's degree in biomedical engineering in the Biophotonics Lab at Carnegie Mellon University. Prior to his studies, he was a medical AI research assistant for 18 months at Inception Institute of Artificial Intelligence in Abu Dhabi, UAE. His work there combined novel deep learning models with automated medical image diagnosis systems. His research focuses on NIRS-based intracranial pressure estimation using machine learning.

**Alexander Ruesch** is a postdoctoral research associate at the Neuroscience Institute, Carnegie Mellon University. His research interests include the noninvasive measurement of physiological and functional biological signals. His doctoral work derived intracranial pressure noninvasively from cerebral blood flow measurements using diffuse optical systems. His postdoctoral work

includes translating diffuse optical methods to marine mammals, for animal care and preservation, as well as research, to benefit from noninvasive imaging in their natural habitat.

**Jason Yang** received his bachelor's and master's degrees in electrical engineering and computer science from the Massachusetts Institute of Technology. Currently, he is pursuing a PhD in the Department of Biomedical Engineering at Carnegie Mellon University. His current interests include the intersection of medical devices and optical instrumentation design with a focus on investigating cerebral hemodynamics and autoregulation, noninvasively. He is a member of SPIE.

**Deepshikha Acharya** received her bachelor's degree in instrumentation and control engineering from Delhi University, India, and her master's degree in biomedical engineering from Carnegie Mellon University. Currently, she is pursuing a PhD in the Department of Biomedical Engineering at Carnegie Mellon University, Pittsburgh, Pennsylvania, United States. Her research interests include studying changes in neural and vascular functions with cerebral autoregulatory impairment in preclinical models and clinical patient population.

**Bradley Scammon** received his bachelor's degree in bioelectrical engineering from Clemson University. Currently, he is pursuing a PhD in biomedical engineering at Carnegie Mellon University, Pittsburgh, Pennsylvania, United States. His research interests include the design, development, and testing of optical and electrical hardware that enables the noninvasive investigation of physiological processes, with a particular focus on hemodynamics and neurovascular coupling.

**Samantha Schmitt** received her bachelor's degree in genetics and cell biology from Washington State University. She has worked with Dr. Smith since 2014 in the field of neurophysiology, studying the brain and behavior. Currently, she is a research technician in biomedical engineering in the Smith Lab at Carnegie Mellon University. Her research interests include neurovascular coupling and how anesthesia affects different types of data collection.

**Emily C. Crane** received her bachelor's degree in biomedical engineering from George Washington University and her master's degree in biomedical engineering from Carnegie Mellon University. Previously, she worked for four years in precision medicine at the Food & Drug Administration. Currently, she is pursuing a PhD in the Department of Biomedical Engineering at Carnegie Mellon University and works in the Smith Lab to characterize signal quality changes over time in intracortical microarrays.

**Matthew A. Smith** received his PhD in neural science from New York University and performed postdoctoral work at both Carnegie Mellon University and the University of Pittsburgh. Currently, he is an associate professor in the Department of Biomedical Engineering and the Neuroscience Institute at Carnegie Mellon University. His research focuses on using neurophysiological approaches to reveal the links between brain and behavior, with a focus on understanding the neural circuitry of perception, cognition, and action.

**Jana M. Kainerstorfer** received her PhD in physics from the University of Vienna, Austria, in partnership with the National Institutes of Health and was a postdoctoral fellow at Tufts University. Currently, she is an associate professor of biomedical engineering at Carnegie Mellon University. Her research focuses on optical imaging technology development for monitoring cerebral and tissue health. She serves on program committees at national and international conferences. She is an associate editor for the *SPIE Journal of Biomedical Optics* and is a senior member of Optica (previously OSA).