A Robust Noninvasive Estimator of Intracranial Pressure

C. Marzban[1,2*], W. Gu[2], P. D. Mourad[1,3,4,5]

[1]Applied Physics Laboratory, University of Washington, Seattle, WA 98195

[2]Department of Statistics, University of Washington, Seattle, WA 98195

[3]Department of Neurological Surgery, University of Washington, Seattle, WA 98195

[4]Department of Bioengineering, University of Washington, Seattle, WA 98195

[5]Division of Engineering and Mathematics, University of Washington, Bothell, WA 98011

*Corresponding Author. Email: marzban@stat.washington.edu

Background: Most methods for the noninvasive prediction of Intracranial Pressure (ICP) from Arterial Blood Pressure (ABP) and maximum Flow Velocity (FV) measured by Doppler ultrasound rely on morphology (i.e., shape of the time series across one cardiac cycle). Estimation of the morphology is a complex task requiring sophisticated statistical methods, which in turn jeopardizes model robustness. The quality of the predictions from these models varies across a wide range, with reported typical precision ranging from 3 to 20 mmHg. There exists one method for ICP prediction which does not rely on morphology, and employs only standard statistical models. This model employs a quantity called Zero-Flow-Pressure (ZFP) to predict ICP. Here several ZFP-based models are developed, and it is shown that when between-patient and within-patient variability are taken into account, then the performance the models is within the range of contending model performance values.

Methods: Standard linear regression models are developed using data on 104 patients, using leave-one-out cross-validation. The regression models do not involve the wave form morphology. The accuracy and precision of the predictions is gauged with bias B and standard deviation of the errors S_e, respectively. To properly assess the quality of the predictions, the variability of these measures is decomposed into between-patient and within-patient components.

Results: It is shown that large between-patient variability of B and S_e renders all of the ZFP-based models developed here statistically indistinguishable. However, the within-patient variability of S_e is sufficiently small (i.e., the model is sufficiently robust) to allow for the identification of one of the models as being "better." The between-patient standard deviation of error (i.e., typical precision) for that model is 10.7 mmHg.

Conclusions: It is possible to develop simple - and therefore, robust - regression models to predict ICP from ABP and FV, without the difficult task of estimating the shape of the ABP or FV time series. Although the quality of the predictions is too poor to be clinically useful, all performance numbers are well within the range of those found in the literature.

Keywords: blood flow, brain, head injury, intracranial pressure, noninvasive.

The importance of noninvasively predicting (or estimating) Intracranial Pressure (ICP) has been thoroughly documented[1-5]. There exist several broad categories of methods. In one category, ICP is predicted from Arterial Blood Pressure (ABP) and maximum Flow Velocity (FV), with the latter estimated by Doppler ultrasound[6-8]. The majority of these models employ *features* of the ABP and FV time series as predictors of ICP. For example, Kim et al.[7,8] construct complex features from the peaks and troughs of the wave form across a cardiac cycle. Other complex features are inferred from examining the time series of ABP and FV across multiple cardiac cycles, and in turn mapping the results to ICP in Schmidt et al.[6] All of these models involve the identification of the morphology (i.e., shape or wave-form) of the time series across one or multiple cardiac cycles of ABP and/or FV.

Inferring shape or wave-form generally involves complex and sophisticated statistical methods. For example, Kim et al.[7] develop a feature-extraction method called the Morphological Clustering and Analysis of Intracranial Pressure Pulses (MOCAIP). Schmidt et al.[6] use Systems Analysis[9] to extract features of ABP and FV relevant for ICP prediction. The complexity of these models introduces undesirable variability in the predictions. For instance, consider the spectrogram shown in Figure 1; the time series of the FV for this patient is quite typical, and yet, there are no identifiable peaks or troughs necessary for MOCAIP. Consequently, the MOCAIP algorithm would likely fail on such data. Systems Analysis relies on a high-dimensional parameterization of the morphology of the ABP, FV, and ICP time series, and so, such a method is likely to suffer from large variability of errors.

By contrast there is one ICP-prediction method which does not involve the shape of the time series of ABP and FV. It is based on the Zero-Flow Pressure (ZFP), also known as the critical closing pressure.[10-17] The ZFP, defined as the ABP at which FV is zero, has been shown to approximate ICP.[16,17] In practice, FV in a living patient is never zero, and so, ZFP is normally estimated by examining the scatterplot of ABP versus FV, and then extrapolating to the point where FV is zero. The fact that ICP is correlated with ZFP, and that ZFP can be estimated from ABP and FV, in turn, suggests that one can

develop a model for predicting ICP directly from ABP and FV. Here, several ZFP-motivated models are developed. As seen below, these models are extremely simple with no more than eleven parameters to be estimated by hundreds of observations, thereby significantly reducing the variability of the predictions. (The paucity of the number of parameters also tames overfitting. Overfitting occurs when a model has a sufficiently large number of parameters to allow the model to fit the noise component of data, resulting in poor predictions on future/independent data.)

Although the ultimate aim of all of the above-mentioned methods is to produce accurate and precise predictions of ICP, it is currently impossible to compare the various methods in terms of their performance. The main reason is that they have not been applied/tested on the same data set. The above studies differ in 1) the number of patients, varying from 8 to 98, 2) the medical condition of the patients, including closed and open TBI, presence or absence of vasospam, and ranging from healthy to near-death, 3) the hospitals and their clinical management practices, and 4) the devices with which ICP is measured invasively, e.g., intraventricular catheters and fiber optic pressure sensors.

The absence of such control factors across studies extends into the manner in which prediction errors are assessed. For example, in many of the studies it is not clear if the prediction errors are estimated from an independent data set (e.g., from cross-validation), or from the training set (i.e., the data set on which the models are developed). Prediction errors from the latter are smaller than one might expect on future/independent data. Furthermore, the separation of errors[18] into between-patient and within-patient is often not performed. Consequently, it is unclear if the typical errors reported are expected for a random patient, across time, or across patients at a given time. The measures of performance are equally varied, including Root-Mean-Squared-Error (RMSE), Mean-Absolute-Distance (MAD), and standard deviation of errors. Although all of these quantities are estimators of typical deviation of error (i.e., typical precision), they are not identical. As such, it is difficult to definitively compare the various models in the form they have been reported in the

literature.

These uncontrolled factors explain the wide range of prediction errors reported. Based on a study of 11 patients Schmidt et al.[6] report an MAD of 4 ± 1.8 mmHg. The same method but applied to 17 patients[19] led to an MAD of 8 ± 5 mmHg. On a much larger group of patients, a ZFP-based method[20] yielded a standard deviation (SD) of error of 12.7 mmHg, while Thees et al.[21] reports SD = 8 mmHg based on 70 patients, and Buhre et al.[22] found SD = 15 mmHg in a study of 20 patients. More recently, Behrens et al.[23] finds SD = 10 mmHg in a study that included 8 patients. On the lower end of the range of reported error values, Chacón et al.[24] find SD = 1.7 mmHg, which is in fact smaller than the instrumental error standard deviation of 3.3 mmHg.[25] Most recently, Xu et al.[26] study 23 patients and report a measure of typical error in the 6.0-6.7 mmHg range. In another study of 57 patients Kim et al.[27] report a typical error of 4-6 mmHg. As mentioned previously, in many of these studies it is not clear whether the reported errors are from cross-validation or from a training set, or whether they pertain to between-patient or total variability.

It is worth noting that all of the measures reported in the above studies are measures of precision, and that none of these studies reports any measure of accuracy. A complete assessment of the performance of the predictions must assess both. Accuracy is important because it gauges the agreement between the average of the predictions with the average of the observations. Precision, on the other hand, measures the average of the *spread* (or deviations) of the predictions about the observation. Here, both of these facets of performance are examined; they are measured by the bias and the variance (or standard deviation) of the predictions, respectively, both defined below. Moreover, each of these measures (bias and variance) has two sources of variability (between-patient and within-patient) which must be taken into account because they capture different types of errors. None of the aforementioned studies address both sources of variability. For clinical purposes the most important of the metrics is the between-patient precision, in terms of which the models developed here

will be compared to those in the literature.

The wide range of values of precision reported in the literature - from 1.7 to 15 mmHg - can be partially attributed to the complexity of the underlying models, because complex models are more likely to have higher variability; see, for example, "Bias versus Variance".[28] Here, several very simple, ZFP-based regression models are developed using leave-one-out cross-validation on 104 patients. It is shown that all of the models are statistically equivalent in terms of all of the metrics considered here, except when they are compared in terms of within-patient standard deviation of errors, in terms of which one of the models emerges as "better" than the other models. The between-patient precision (i.e., the clinically most relevant metric) of that model is found to be comparable to the corresponding quantities reported in the literature. The advantage of the models developed here is that they do not involve the morphology of the time series, and as such are relatively robust.

Material and Methods

A total of 104 patients were examined in this study. 59 patients were selected from hospitals within the United States, specifically Harborview Medical Center (Seattle, Washington), Columbia University Medical College (New York City, New York) and the University of Texas Medical School (Houston, Texas). The remaining 45 patients were from Cambridge, England. Schmidt and co-workers[29] describe that data. Further details of the entire data set can be found in Marzban et al.[17] and Marzban et al.[30] In accordance with the institutional review board for each hospital, informed consent was obtained from all patients or their families.

The data contains Doppler ultrasound measurements of blood Flow Velocity (FV) in the middle cerebral artery using clinically approved Transcranial Doppler units. Additionally, ABP was measured by radially-placed arterial blood pressure sensor, and ICP was measured by a Codman® fiber optic pressure sensor. For each patient, the duration of collected data varied from 5 to 30 minutes. For the

present study, only 5 minutes of the time series of ABP, FV, and ICP was used. Retrospective data processing and analysis was performed at the Applied Physics Laboratory, University of Washington. The Doppler time series for all patients were sampled at 40 Hz. This resolution was sufficient to resolve each patient's systolic rise, diachrotic notch, and diastolic minimum, although none of these features were used in the analysis.

Figure 2 shows a scatterplot of the mean (across 120 time-steps, i.e., about 3 seconds in duration) of ICP vs. ABP (top panel), and ICP vs. FV (bottom panel). Different colors/clusters correspond to different patients, and each cluster consists of 100 dots corresponding to 100 non-overlapping time intervals across the patients time series. A wide range of linear associations can be seen, but only within each patient's cluster. There appears to be no between-patient association, linear or otherwise. Therefore, any useful model for the prediction of ICP must involve ABP and FV, jointly.

Data Processing and Statistical Analysis

Statistical Models

As mentioned above, several statistical models are developed for predicting mean ICP (denoted ICP) from mean ABP and mean FV, where the mean is taken over some specified duration of time, here, about 3 seconds. This duration - henceforth, called *interval* - is sufficiently long to contain several cardiac cycles. A longer duration would be desirable from a statistical point of view, for it would lead to a larger sample size for estimating the true mean; but a longer time interval complicates the estimation process because a patient's mean ICP may significantly change during that interval - a change that will not be detectable, if the interval is too long. For these reasons, the interval over which the mean of ICP is to be predicted is set to about 3 seconds (i.e., 120 time-steps).

The structure of the models is motivated by the definition of ZFP. The estimator of ZFP proposed by Weyland et al.[16] is based on a least-squares fit of ABP versus FV when the two have been

aligned to have the same phase. Marzban et al.[17] proposed a revision wherein the least-squares fit is replaced by another - called the SD line. The advantage of the SD line, for estimating ZFP, is that it does not require the time series of ABP and FV to be in phase. As shown in Marzban et al.[17], the original and the revised ZFP estimators are given by

$$ZFP = x_1 - \frac{r\, x_2\, y_1}{y_2}, \qquad ZF\,P' = x_1 - \frac{x_2\, y_1}{y_2} \qquad (1)$$

respectively, where x_1 = ABP, x_2 = FV, y_1 = sd(ABP), y_2 = sd(FV), and r is the correlation coefficient between ABP and FV. mean() and sd() denote sample mean and sample standard deviation, respectively. Note that the only difference between the two estimators is the r. As shown in Marzban et al.[17], of the two estimators (ZFP, ZFP'), the latter is better correlated with ICP.

Table 1 lists the various models examined in this work. Model 1 corresponds to the situation where neither ABP nor FV are used for prediction. As such, the only available quantity for the prediction of future ICP values is the ICP values in the current data set. In order to assess the quality of the predictions, leave-one-out cross-validation is employed.[28] Briefly, one patient is selected (left out), and the sample mean of ICP across the remaining patients is used as the prediction for the ICP of the selected patient. This procedure is repeated 104 times, each time leaving out one of the 104 patients. The prediction error (i.e., predicted ICP - observed ICP) for each of the left-out patients is computed. This model serves as a benchmark in that it should be outperformed by any alternative model which employs ABP and FV for the prediction of ICP. For this reason, it is referred to as the "Null model."

Model 2 employs ZFP' as a predictor of ICP. In other words, it uses the specific combination of ABP and FV given in Eq. (1). Model 3 is simply an algebraic generalization of the definition of ZFP'. Model 4 is, in turn, a generalization of model 3 in that an additional linear term (in x_2) is introduced. Model 5 additionally includes linear terms in y_1 and y_2, and model 6 additionally includes all interaction terms. Note that each successive model in Table 1 has more parameters than the preceding model.

More complex models are possible, but are not considered here in order to avoid overfitting.

The prediction errors for the models 2-8 are assessed in the same manner as model 1, i.e., with leave-one-out cross-validation. Specifically, a patient is left out, while the model is developed on the remaining patients. The model is then used to predict the ICP for the left-out patient, and the prediction error is computed.

Performance Assessment

The aforementioned methodology leads to one value of prediction error, per model, per patient, per interval, where prediction error is given by (predicted ICP - observed ICP), and ICP is a mean over the interval. This design allows one to separate the sources of variability in the errors into two components: between-patient, and within-patient. Consideration of each of these components is important, because the *total* variability of the errors is sufficiently large so as to preclude any useful comparison of the various models.

The quality of the predictions is assessed in terms of the accuracy and precision of the predictions. Here, these two facets are measured with the bias (B) and standard deviation of the errors (S_e), respectively, both defined in the Appendix. The variability of these quantities is assessed through the distribution (technically, histogram) of the prediction errors. To allow easy comparison of these histograms across the various models, each histogram is summarized by a boxplot. The Appendix provides further details of the performance metrics and their between-patient and within-patient decomposition.

Results

The top panels in Figure 3 show the boxplots of bias B for all models, and the boxplots of the standard deviation of the errors (S_e) is shown in the bottom panels. The boxplots in left panels show the variability between patients, and those in the right panels convey the within-patient variability. In

other words, the left panels assess how well the models predict ICP across patients, for a random interval, while the right panels gauge the quality of the predictions across time intervals, for a random patient. In comparing two models, a significant overlap between their boxplots is indication that there is no statistically significant difference between the models. Generally, a model with the lowest overall bias and standard deviation is desirable. Numerical values for the "center" (mean) and "width" (standard deviation) of these boxplots are presented in Table 2.

Examination of the top/left panel in Figure 3 suggests that in terms of bias, the variability between-patients is sufficiently large to preclude any useful comparison of the models in terms of *B* (between patients). All models have near-zero bias between-patients and are statistically indistinguishable. The same conclusion follows when standard deviation of the errors between-patients is considered (bottom/left panel); all of the models are generally comparable.

Consideration of the distribution of bias within-patients (top/right panel) is more informative. For example, whereas models 2 and 3 are negatively biased, models 4 and 5 are positively biased. Models 6 and 7 display a sufficiently large variability so as to preclude any conclusion regarding their bias. Although such comparisons are useful for model comparison and model selection, it is important to point out that all of these bias values are extremely small; as seen in Table 2, the mean of all of these boxplots is in the -0.06 mmHg to +0.06 mmHg range. As such, in terms of within-patient bias, all of these models can be considered adequate.

Examination of the precision of the predictions within-patient (bottom/right panel) leads to more discriminating information. It can be seen that model 5 has the lowest values of S_e. The significant overlap of the boxplots for models 5, 6, and 7, suggests that they are statistically indistinguishable in terms of S_e. The slight increase in S_e across models 5, 6, and 7, may be due to overfitting, because the number of parameters for these models is successively larger. Following the principle of parsimony, it is reasonable to declare model 5 as "the best" model. Having selected the

best model, one can turn attention to the most clinically-relevant metric, namely between-patient standard deviation of errors (i.e., the between-patient standard deviation of *B*). According to the values in Table 2, that quantity for model 5 is 10.7 mmHg.

It is important to reiterate that model 5 emerges as "the best" because of the lower variability of the prediction errors across time intervals (for a random patient). Consideration of between-patient variability alone does not allow one to identify the best model(s). It is only when variability across time is assessed separately that one can distinguish between the various models.

Having selected the best model (i.e. model 5), it is possible to diagnose its performance in a variety of ways. One useful tool is the scatterplot of the predictions versus the observed values of ICP (Figure 4, top). Each cluster/color represents a different patient, and each dot represents the mean across a time interval. The dashed line is a diagonal line of slope one and intercept zero. Ideally, all of the dots would be along the diagonal line if the predictions were exactly equal to the observed values. However, in practice that is not the case. The correlation coefficient - as a measure of the scatter - between the centers of the clusters, i.e., between-patient correlation[31] is 0.60. It is reassuring that all of the dots are randomly and symmetrically distributed about the diagonal. This suggests that the prediction errors are randomly distributed - a desirable feature for a model. Another useful method for visually evaluating the errors is the Bland-Altman plot (Figure 4, bottom). Again, it can be seen that the errors (predicted ICP - observed ICP) are randomly scattered across the diagram. The horizontal lines mark the overall *B* (near zero), and the one- and two-standard-deviation lines.

Discussion

In this paper it is shown that relatively simple multiple regression models, using ABP and FV as predictors, can be developed for predicting ICP noninvasively. The prediction error of the best of the models is not sufficiently low to make the model clinically useful. However, the errors are within the

range of those produced by more complex models reported in the literature. It is difficult to compare the errors of the model developed here to those reported in the literature more closely, because access to the data and the models is limited; however, see below.

In addition to the simplicity of the models developed here, another novel feature of this work is the way the errors are decomposed into between- and within-patient components. It is found that the between-patient variability of the prediction errors is so large that one cannot distinguish between the models developed here. Indeed, it is possible that the same is true when more complex models are included in the comparison. Here, it is the within-patient component of the variability of the errors that allows for the selection of "the best" model.

In many ICP-related articles between-patient RMSE is used for both model selection and for assessing the precision of the predictions. That practice has two flaws: 1) in using a *between-patient* measure one ignores within-patient variability. As seen here, to distinguish between the models one requires a consideration of both between- and within-patient variability. 2) Although RMSE is an adequate measure for model selection, it is inadequate as a measure or precision, because RMSE is related to the sum of two different measures - one measure of precision (S_e^2) and a measure of accuracy (B). The exact relationship is $\mathrm{RMSE}^2 = ((n-1)/n)\, S_e^2 + B^2$, where n denotes sample size; for the between-patient component, n=104 , i.e., the number of patients, while n=100 for the within-patient component. Said differently, RMSE is not a measure of precision, when bias is non-zero. Although in some special circumstances (e.g., if the errors are residuals from a least-squares fit), B=0 by construction, in other cases especially those involving cross-validation, B is non-zero. Given that bias and variance are both important facets of quality, here they are assessed separately. Given that the values of B are generally small, RMSE for these models is approximately equal to S_e.

In order to better assess the utility of the models developed here, an attempt was made to implement some of the methods mentioned in the Introduction, specifically the systems method of

Schmidt et al.[6,19,20,29] and the MOCAIP method of Kim et al.[7,8] The latter has proven difficult to implement mostly because many of the details of the approach have not been completely expressed in the published literature. However, the systems method of Schmidt et al. was implemented and evaluated as in the models above. The resulting model was found to be far inferior to any of the models developed above in terms of B and S_e; see "model 8" in Table 2. It is especially disconcerting that the performance of model 8 is even poorer than that of the Null model, i.e., where neither ABP nor FV are used for prediction. Such a poor level of performance may be attributed to errors in the present implementation of the algorithm proposed by Schmidt et al.[6,19,20,29]

The critical quantity that decides whether a model has clinical utility is the between-patient precision. For the best model developed here, it is ±10.7 mmHg. Smaller values have been reported in the literature: Schmidt et al.[19] report ±8 mmHg, and Kim et al.[27] find precision values in the 4 to 6 mmHg. However, as explained previously, such comparisons are non-informative because the patients in the corresponding groups are different. It is important to point out that the device used here for making invasive ICP measurements itself has an error in the ±3.3 mmHg range, and so, it is possible (in theory) to increase the between-patient precision. More sophisticated techniques are being developed to that end, but all of them appear to have lower within-patient precision.

Another avenue of research currently under investigation is the incorporation of autoregulation in the analysis. Schmidt et al.[29] have proposed a non-invasive measure of the status of autoregulation, and have shown that ICP predictions can be improved when the predictions are analyzed separately, for those with intact and impaired autoregulation. That work is underway.

Conclusion

It is possible to develop relatively simple regression models for predicting ICP from ABP and FV, without the need to perform the generally difficult task of extracting the shape of the time series of

ABP and/or FV. The quality of the predictions is comparable to those reported in the literature from more sophisticated models. The main advantage of the model developed here is its robustness, which is manifested in the small within-patient standard deviation of errors.

Appendix

This appendix provides details of how the prediction error is decomposed into bias (B) and standard deviation of errors (S_e), respectively assessing the accuracy and precision of the predictions. Each of these two components is then further decomposed into between-patient and within-patient components, which respectively gauge how the errors vary across patients and across time.

Recall that the prediction error is defined as the mean of the difference (predicted - observed) of ICP, where the mean is computed across an interval of length 120 points (approximately 3 seconds). For this study 100 non-overlapping time intervals across the entire time series of each patient are used. Also recall that the predicted ICP is arrived at through leave-one-out cross-validation. As such, for each of the 104 patients there are 100 prediction errors.

Let ϵ_{ij} denote the leave-one-out prediction error for the i^{th} interval of the j^{th} patient. Here, i varies from 1 to 100, and j varies from 1 to 104. The marginal means and standard deviations of this error matrix represent all of quantities of interest. Specifically, let $\epsilon_{\cdot j}$ and $s_{\cdot j}$ respectively denote the sample mean and standard deviation of the errors for the j^{th} patient. Similarly, let $\epsilon_{i\cdot}$ and $s_{i\cdot}$ respectively denote the sample mean and standard deviation of the errors for the i^{th} interval. In short,

$$\epsilon_{\cdot j} = \frac{1}{100}\sum_{i}^{100} \epsilon_{ij}, \qquad s^2_{\cdot j} = \frac{1}{100-1}\sum_{i}^{100} \left(\epsilon_{ij} - \epsilon_{\cdot j}\right)^2, \qquad (2a)$$

$$\epsilon_{i\cdot} = \frac{1}{104}\sum_{j}^{104} \epsilon_{ij}, \qquad s^2_{i\cdot} = \frac{1}{100-1}\sum_{j}^{104} \left(\epsilon_{ij} - \epsilon_{i\cdot}\right)^2. \qquad (2b)$$

The boxplot of the 104 values $\epsilon_{\cdot j}$ (top/left panel in Figure 4) conveys information regarding the

distribution of B across patients (i.e., between-patient), while a boxplot of the 100 values $\epsilon_{i.}$ (top/right panel in Figure 4) reflects the distribution of B across the time intervals (i.e., within-patient). Similarly, the boxplot of the 104 values $s_{.j}$ (bottom/left panel in Figure 4) conveys information regarding the distribution of between-patient S_e, while a boxplot of the 100 values $s_{i.}$ (bottom/right panel in Figure 4) reflects the distribution of within-patient S_e. In this paper, model selection is performed in terms of the latter, while comparison with other methods in the literature is performed in terms of the between-patient standard deviation of B.

Acknowledgments

This work has received support from National Institutes of Health (Grant R43NS46824-01A1); National Space Biomedical Research Institute (Grant SMS00701-2009-513); and PhysioSonics Incorporated. P. D. Mourad has a financial interest in PhysioSonics. The authors are grateful to P. R. Illian and D. Morison for contributions made during an early phase of this project.

References

1 Aaslid R, Nornes H. Musical murmurs in human cerebral arteries after subarachnoid hemorrhage. *J Neurosurg* 1984; **60**:32-6.

2 Aaslid R, Huber P, Nornes H. Evaluation of cerebrovascular spasm with transcranial Doppler ultrasound. *J Neurosurg* 1984; **60**:37-41.

3 Aaslid R, Huber P, Nornes H. A transcranial Doppler method in the evaluation of cerebrovascular spasm. *Neuroradiology* 1985; **28**:11-16.

4 Aaslid R. Transcranial Doppler assessment of cerebral vasospasm. *European Journal of Ultrasound* 2002; **16**:3-10.

5 Armonda RA, Bell RS, Vo AH, Ling G, DeGraba TJ, Crandall B, Ecklund J, Campbell WW. Wartime

traumatic cerebral vasospasm: recent review of combat casualties. *Neurosurgery* 2006; **59**(6):1215-1225.

6 Schmidt B, Bernhard J, et al. Noninvasive prediction of intracranial pressure curves using transcranial Doppler ultrasonography and blood pressure curves. *Stroke* 1997; **28**.12: 2465-2472.

7 Kim S, Hu X, et al. Inter-subject correlation exists between morphological metrics of cerebral blood flow velocity and intracranial pressure pulses. *Neurocritical care* 2011 **14**.2: 229-237.

8 Kim S, Hamilton R, Pineles S, Bergsneider M, Hu X. Noninvasive Intracranial Hypertension Detection Utilizing Semisupervised Learning. *Biomedical Engineering, IEEE Transactions on* 2013; **60**(4), 1126-1133.

9 Marmarelis, P. and Marmarelis V. *Analysis of Physiological Systems*. New York, NY, 1978, Plenum Oress, 221 pp.

10 Burton AC. On the physical equilibrium of the small blood vessels. *Am J Physiol*. 1951; **164**: 319-329.

11 Aaslid R, et al. Dynamic pressure-flow velocity relationships in the human cerebral circulation. *Stroke*. 2003; **34**: 1645-1649.

12 Aaslid R, et al. Assessment of cerebral autoregulation dynamics from simultaneous arterial and venous transcranial Doppler recordings in humans. *Stroke*. 1991; **22**:1148-1154.

13 Michel M, et al. Frequency dependence of cerebrovascular impedance in preterm neonates: A different view on critical closing pressure. *J Cereb Blood Flow Metab*. 1997; **17**:1127-1131.

14 Kottenberg-Assenmacher E, et al. Critical closing pressure as the arterial downstream pressure with the heart beating and during circulatory arrest. *Anaesthesiology*. 2009; **110**:370-379.

15 Sherman RW, et al. Cerebral haemodynamics in pregnancy and pre-eclampsia as assessed by transcranial Doppler ultrasonography. *Br J Anaesth*. 2002; **89**:687–92.

16 Weyland A, et al. Cerebrovascular tone rather than intracranial pressure determines the effective

downstream pressure of the cerebral circulation in the absence of intracranial hypertension. *J Neurosurg Anesthesiol*. 2000;**12**:210–6.

17 Marzban C, Illian R, Morison D, Moore A, Kliot M, Czosnyka M, Mourad PD. A method for estimating zero-flow pressure and intracranial pressure. *J Neurosurg Anesthesiol* 2013; **25**(1):25-32.

18 Montgomery D. *Design and Analysis of Experiment*s, 6th ed. New York: John Wiley & Sons, 2005.

19 Schmidt B, Czosnyka M, et al. Cerebral Vasodilatation Causing Acute Intracranial Hypertension & colon; A Method for Noninvasive Assessment. *Journal of Cerebral Blood Flow & Metabolism* **19**.9 (1999): 990-996.

20 Czosnyka M, et al. Critical closing pressure in cerebrovascular circulation. *J Neurol Neurosurg Psych*. 1999; **66**:606-11.

21 Thees C, et al. Relationship between intracranial pressure and critical closing pressure in patients with neurotrauma. *Anesthesiology*. 2002; **96**:595–9.

22 Buhre W, et al. Extrapolation to zero-flow pressure in cerebral arteries to estimate intracranial pressure. *Br J Anaesth.* 2003;**90**:291-295.

23 Behrens A, Lenfeldt N, Ambarki K, Malm J, Eklund A, Koskinen LO. Transcranial doppler pulsatility index: not an accurate method to assess intracranial pressure. *Neurosurgery*. Jun. 2010 p. 1050-1057.[Online]. Available: http://dx.doi.org/10.1227/01.NEU.0000369519.35932.F2

24 Chacón M, Pardo C, et al., Noninvasive intracranial pressure estimation using support vector machine. *32nd Annual International Conference of the IEEE Bioengineering in Medicine and Biology Society (EMBS)*, Buenos Aires, Argentina, August 31 - September 4, 2010, pp. 996-999.

25 Koskinen L-O D, and Olivecrona M. Clinical experience with the intraparenchymal intrcranial pressure monitoring Codman microsensor system. *Neurosurgery*. 2005; **56**(4). 693-698.

26 Xu P, Kasprowicz M, Bergsneider M, Hu X. Improved noninvasive intracranial pressure assessment with nonlinear kernel regression. *Information Technology in Biomedicine, IEEE Transactions on*,

2010; **14**(4), 971-978.

27 Kim S, Scalzo F, Bergsneider M. Vespa P, Martin N, Hu X. (2012). Noninvasive intracranial pressure assessment based on a data-mining approach using a nonlinear mapping function. *Biomedical Engineering, IEEE Transactions on*, **59**(3), 619-626.

28 Hastie T, Tibshirani R, Friedman J, Hastie T, Friedman J, Tibshirani R. *The elements of statistical learning*. New York: Springer, 2009.

29 Schmidt B, et al. Adaptive Noninvasive Assessment of Intracranial Pressure and Cerebral Autoregulation. *Stroke.* 2003;**34**:84-89.

30 Marzban C, Illian R, Morison D, Mourad PD. A double-gaussian, percentile-based method for estimating maximum blood flow velocity. *Journal of Ultrasound in Medicine*. 2013; **32**(11): 1913-20.

31 Bland JM and Altman DG. Calculating correlation coefficient with repeated observations: Part I, correlation within subjects. Statistics. Notes in T*he British Medical Journal*, Vol. **310**, pp. 446, 1995.

Table 1. Models for the prediction of mean ICP. x_1 and x_2 are the mean of ABP and FV, respectively; and y_1 and y_2 denote the standard deviation of ABP and FV, respectively. The notation used here suppresses all regression coefficients; for example, Model 3 as defined below is

$ICP = a_1 + a_2 x_1 + a_3 x_2 y_1 + a_4 x_2 y_2$, where a_i are parameters to be estimated from data.

Model	
1	$ICP = Null\ Model\left(i.e. sample\ mean\ of\ ICP\right)$
2	$ICP = 1 + ZF\ P^{'}$
3	$ICP = 1 + x_1 + x_2 y_1 + x_2 y_2$
4	$ICP = 1 + x_1 + x_2 + x_2 y_1 + x_2 y_2$
5	$ICP = 1 + x_1 + x_2 + y_1 + y_2 + x_2 y_1 + x_2 y_2$
6	$ICP = 1 + x_1 + x_2 + y_1 + y_2 + x_2 y_1 + x_2 y_2 + x_1 y_1 + x_1 y_2$
7	$ICP = 1 + x_1 + x_2 + y_1 + y_2 + x_2 y_1 + x_2 y_2 + x_1 y_1 + x_1 y_2 + x_1 x_2 + y_1 y_2$

Table 2. The performance of the various models developed here in terms of the between-patient and within-patient components of Bias *B* and the standard deviation of errors S_e. Models 1-7 are defined in Table 1; model 8 is the Systems Analysis method of Schmidt et al.[6,19] The entries in the table are (mean ± standard deviation).

Model	Between-Patient B	Within-Patient B	Between-Patient S_e	Within-Patient S_e
1	0.00 ± 13.69	0.00 ± 0.00	1.55 ± 1.35	13.84 ± 0.22
2	-0.04 ± 12.33	-0.04 ± 0.01	2.00 ± 1.41	12.57 ± 0.24
3	-0.03 ± 12.39	-0.03 ± 0.02	1.79 ± 1.19	12.57 ± 0.15
4	0.05 ± 11.52	0.05 ± 0.03	1.95 ± 1.33	11.76 ± 0.22
5	0.05 ± **10.70**	0.05 ± 0.05	2.04 ± 1.33	**10.97** ± 0.29
6	0.06 ± 10.89	0.06 ± 0.10	2.14 ± 1.65	11.21 ± 0.33
7	-0.06 ± 10.99	-0.06 ± 0.08	2.22 ± 1.79	11.35 ± 0.31
8	-2.80 ± 16.53	-2.80 ± 0.21	2.66 ± 1.76	16.82 ± 0.30

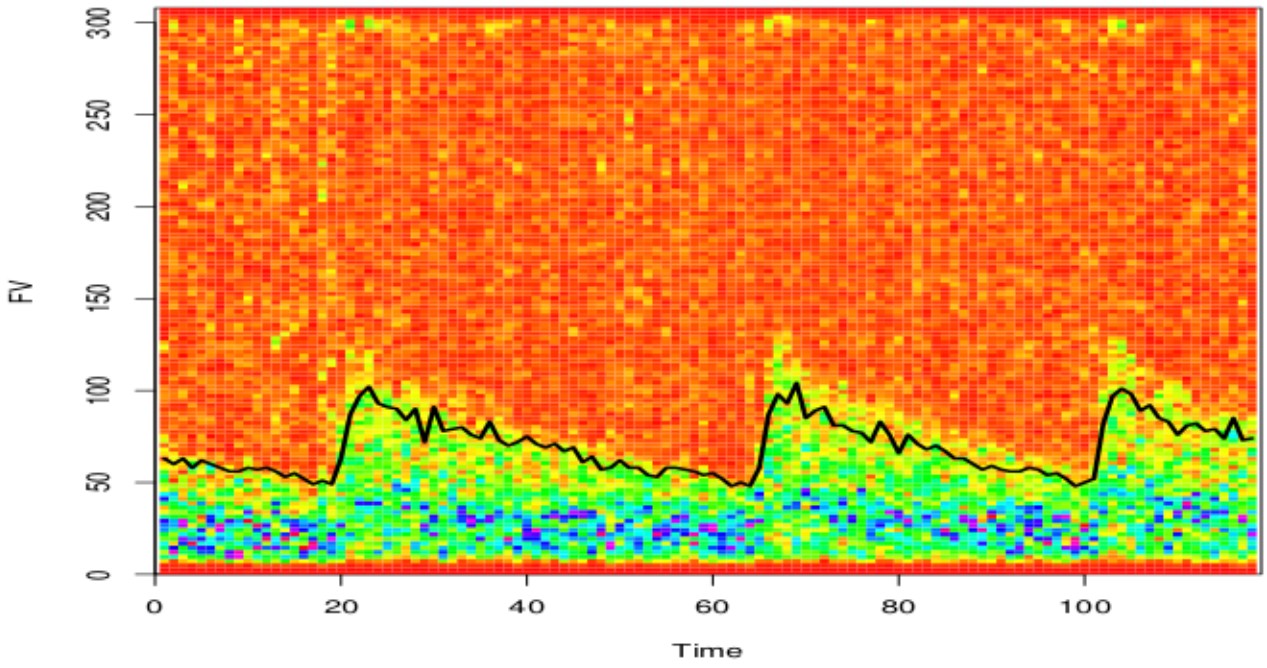

Figure 1. An example of a spectrogram (in color), and the maximum FV envelope (black curve). It can be seen that the shape of this patient envelope has no bumps and troughs, quantities which are necessary for the MOCAIP method.[7,8]

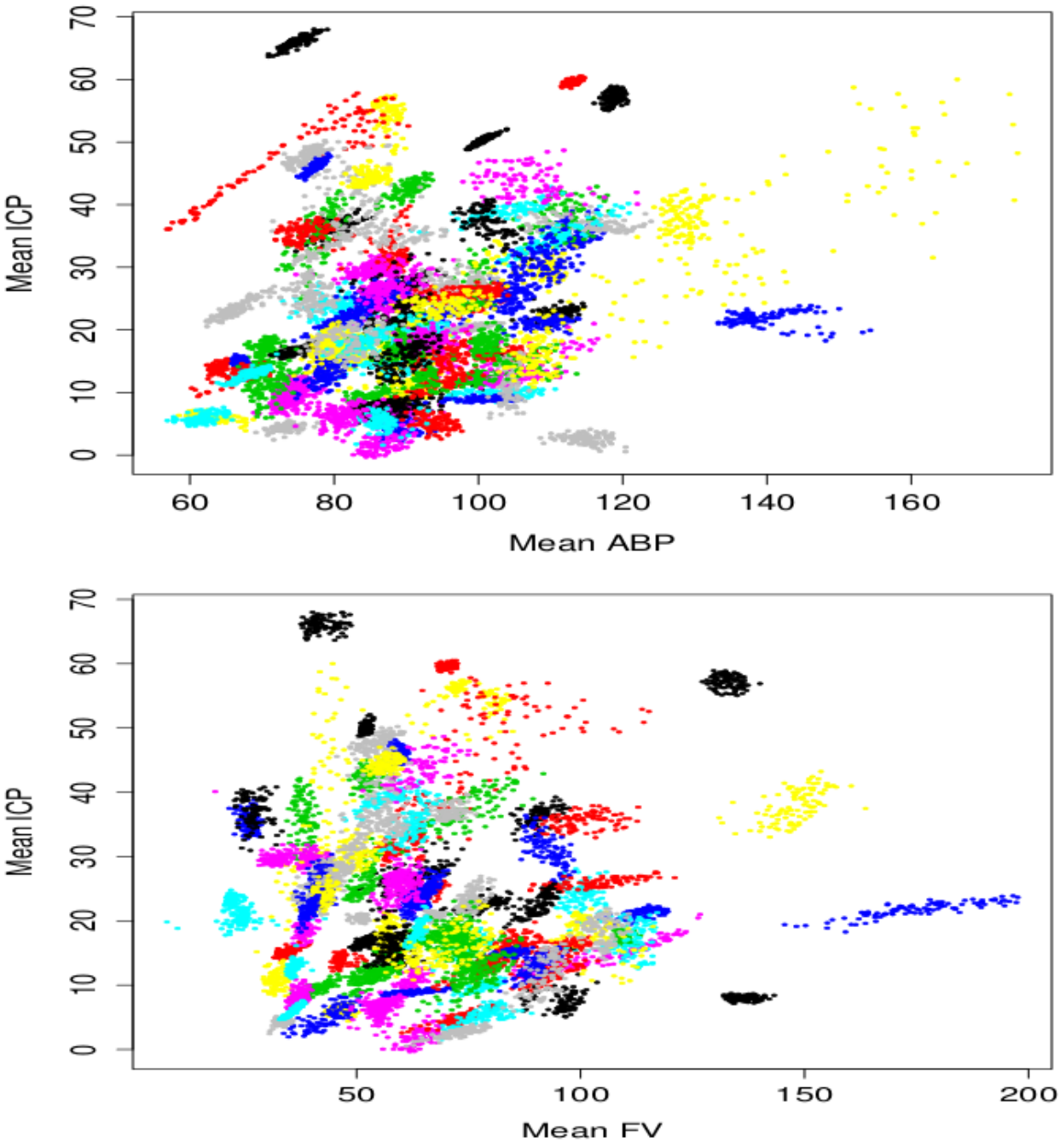

Figure 2. A scatterplot of the mean of ICP versus the mean of ABP, for 104 patients (different colors), and 100 intervals (dots), where the mean is computed across approximately 3 seconds (120 points).

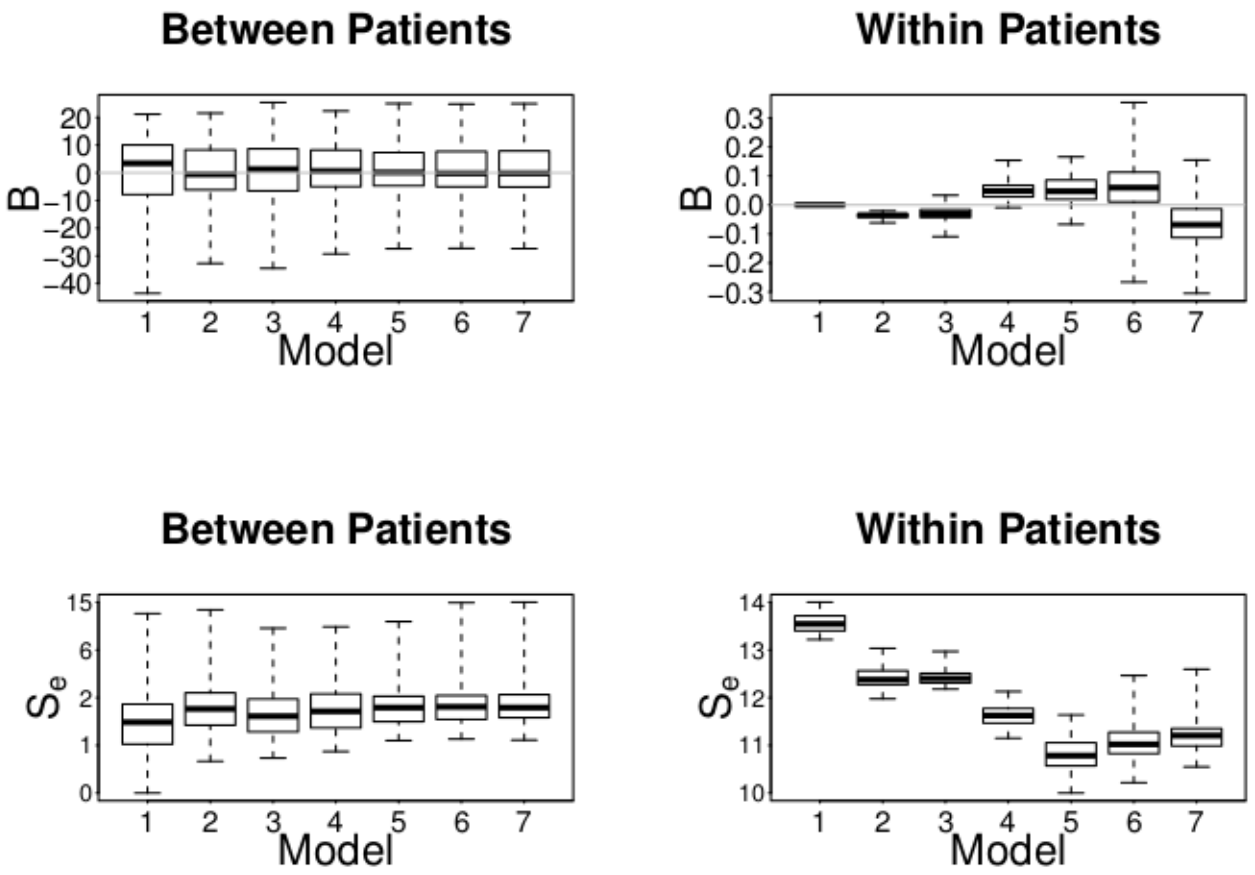

Figure 3. Assessment of prediction error for the various models in Table 1 (denoted along the x-axis). The left panels show the distribution of B (top panel) and S_e (bottom panel) between patients, and the right panels display the distribution of errors within patients. The y-axis on the bottom panels is logarithmic in order to enhance visual clarity.

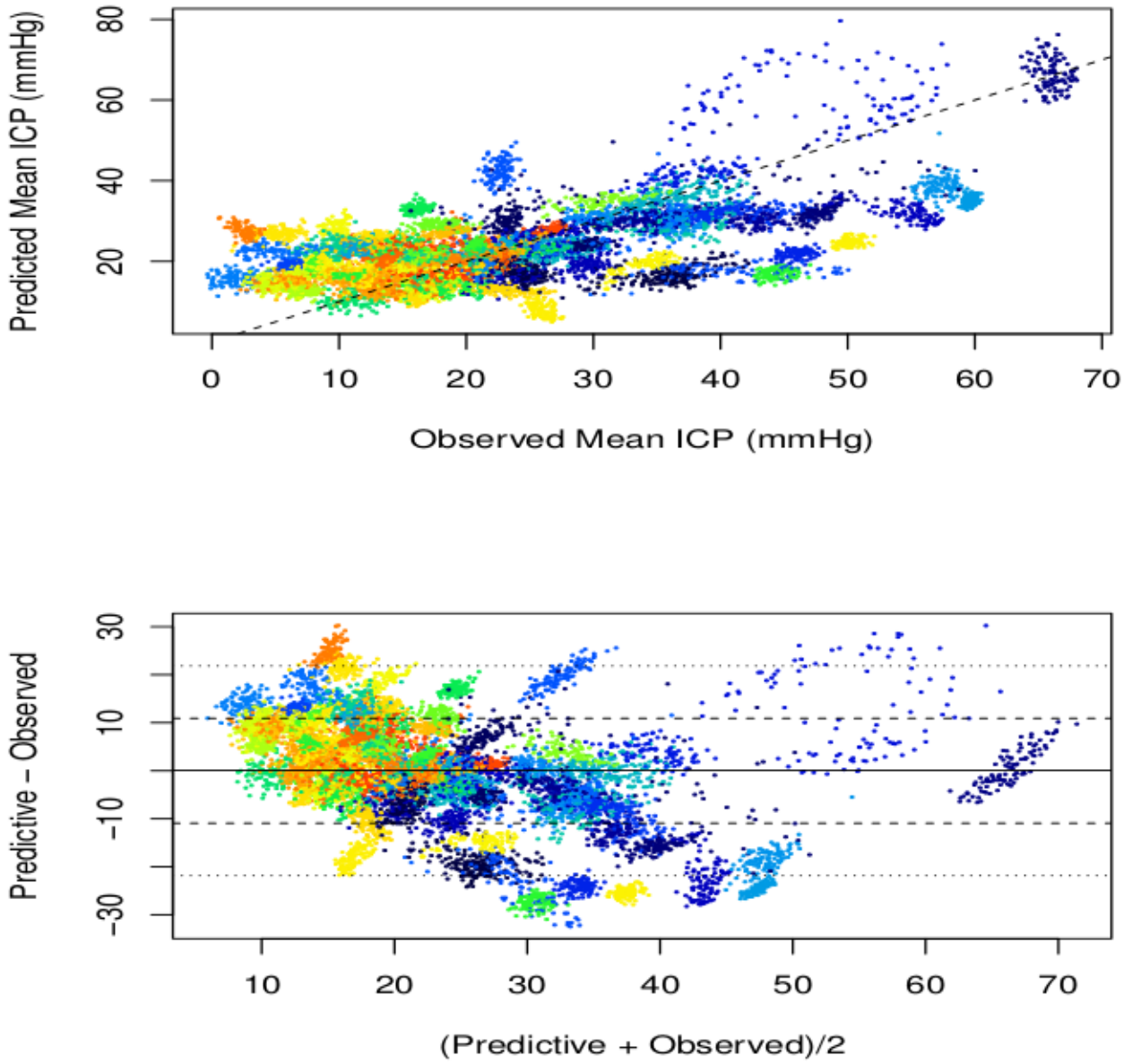

Figure 4. The scatterplot (top) of the predicted ICP versus the observed ICP, for all 104 patients (in color) across 100 time intervals (dots) for model 5. The Bland-Altman plot (bottom) provides an alternative view of the quality of the predictions.