Forecasting Aortic Pressure Cross-Cohort with Deep Sequence Models

Alan Li[1], Zihao Zhou[1], Elise Jortberg[2], Rose Yu[1]

[1] Department of Computer Science and Engineering, UC San Diego, USA
[2] Abiomed, Danvers, MA, USA

Abstract

Mean aortic pressure (MAP) is a primary measurement for monitoring blood and O_2 delivery to major organs. Prolonged periods of hypotension, low MAP, lead to low tissue perfusion and subsequent end organ damage. Patients on mechanical circulatory support (MCS) devices, such as the Impella CP, are managed to maintain sufficient MAP for end-organ perfusion. Forecasting MAP is important for early warning of clinically concerning events, including hypotension and instability as well as device weaning. Patients presenting with cardiogenic shock as a result of acute myocardial infarction (AMI/CGS) have increased hemodynamic instability when compared to patients undergoing high-risk percutaneous coronary interventions (HRPCI). Existing deep sequence models for forecasting often focus on the same patient cohort and cannot generalize across cohorts. In this paper, we examine how deep sequence models respond to the distribution shift of the MAP across the MCS patient cohorts during forecasting. We propose conditional RNN, a deep sequence model that learns to adapt to a different cohort by conditioning on time-invariant cohort features. Our proposed model improves the forecasting performance, achieving a 5.2 mmHg - 6.1 mmHg RMSE for cross-cohort patients.

1. Introduction

Maintenance of a constant mean aortic pressure (MAP) is vital to ensure adequate organ perfusion [1]. Studies show that increase in the duration of time spent below MAP threshold of 65 mmHg is associated with worse patient outcomes such as risk of mortality and organ dysfunction [2, 3]. Patients are monitored and treated to keep MAP above 65 mmHg to avoid low end-organ perfusion. Patients with severe multi-vessel coronary artery disease (CAD), unprotected left main coronary artery stenosis, last remaining patent vessel, and severely reduced left ventricular (LV) ejection fraction (EF) are often turned down from cardiac surgery and are increasingly referred for high-risk percutaneous coronary intervention (HRPCI) [4]. Patients with Cardiogenic Shock resulting from Acute Myocardial Infarction (AMI/CGS) who are dependent on vasopressors are recommended to be escalated to additional hemodynamic support due to unstable, low blood pressures [5]. Many patients at risk for hemodynamic instability, such as patients with AMI/CGS and those undergoing HRPCIs, are treated with mechanical circulatory support (MCS). In both the ICU and cardiac catheter lab (CCL), providing an accurate, continuous MAP forecast would be significant when monitoring hemodynamically unstable patients at risk for decompensation.

For patients treated with MCS, MAP is a function of native heart blood output and the support level of the pump. The challenges include the capability to predict MAP trends across patients with diverse medical presentations whose hemodynamics change throughout the care timeline in response to clinical intervention. Deep sequence models such as Recurrent Neural Networks (RNNs) have demonstrated capability for time series forecasting [6], including blood pressure predictions [7, 8]. Existing studies on MAP forecasting in critical care setting are mostly limited to a specific patient population, hence their conclusions have limited generality across patient cohorts.

The Impella CP is an implantable (percutaneously) catheter-based left ventricular assist device (LVAD) that provides hemodynamic support to the heart. In this paper, we examine the generalization of deep sequence modeling for MAP forecasting across two Impella patient cohorts with differing indications: HRPCI, AMI/CGS. We propose conditional RNN, an extension of RNN that can adapt to the MAP distribution shift across patient cohort. Using a few examples from the new patient cohort, we can compute certain time-invariant cohort features. By conditioning on these cohort-specific features, we can improve the generalization performance of the deep sequence model. We show that conditional RNN improves the forecasting performance for several RNN models including LSTM [9] and LMU [10].

2. Dataset

The left sided Impella devices pump blood from the left ventricle (LV) into the ascending aorta across the aortic valve. Impella is indicated to provide MCS to patients undergoing HRPCI, and patients with Cardiogenic Shock.

Impella device performance signals are collected from the Automated Impella Controller (AIC) at 25 Hz. The signals monitor aortic pressure (in mmHg) from an optical sensor at the Impella outlet, motor current (mA) which modulates with respect to dynamic pressure environment over cardiac cycle to maintain constant speed, and motor speed (rpm) of the pump impeller, and pump flow of blood through the Impella cannula (L/min) as shown in Figure 1.

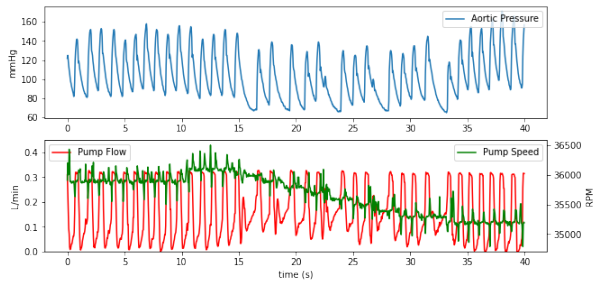

Figure 1. Sample Pump Performance (25 Hz) Data

HRPCI: Coronary artery disease (CAD) results from damage in the arteries that supply blood to the heart. Patients with severe or complex disease who are not eligible for open-heart surgery may undergo a HRPCI with hemodynamic support from an Impella. During the procedure, metrics including MAP and Cardiac Output (CO) are monitored to track hemodynamic stablity. HRPCI procedure restores blood flow to the heart for revascularization.

AMI/CGS: Cardiogenic shock occurs when a severe impairment causes the heart to fail to pump enough blood to sufficiently supply oxygen to critical organs. Many patients who experience cardiogenic shock do so after an acute myocardial infarction [11]. An AMI results from a blockage of the coronary arteries that supply blood to the heart muscle, resulting in damaged tissue. This can lead to cardiogenic shock where the heart enters a weakened state and can no longer pump enough blood to maintain organ perfusion. When compared to patients undergoing HRPCI, AMI/CGS patients' heart disease is more severe and have increased hemodynamic instability.

Table 2 summarizes descriptive parameters of the Impella cases per cohort used. The distribution of the cases is shown in Figure 2.The cases occurred between December 2017-April 2020. The case samples were from 290 hospitals across 32 regions of the USA, Germany, and Canada. The HRPCI median case duration was 1.7 hours [1.2, 2.0] and the AMI/CGS median case duration was 28.0 hours [13.9, 52.1]. We split 50 AMI/CGS cases and 40 HRPCI cases of the 463 cases for out of sample testing.

3. Methods

Previously, [8] examined the problem of MAP forecasting; they concluded that the LMU sequence to sequence

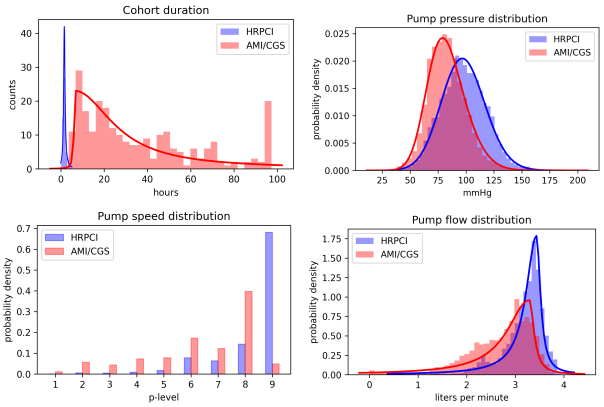

Figure 2. HRPIC and AMI/CGS distributions. Top left: duration of cases; top right: pump (aortic) pressure; bottom left: pump speed; bottom right: pump flow.

Table 1. Demographics of Patients in Cohorts

	HRPCI	**AMI/CGS**
Cases	213	250
Age (years)	72 [64, 79]	65 [58, 71]
Gender (M/F/U)	154/59/7	168/70/8
EF (%)	30 [20, 40]	20 [15, 30]
ECMO	1	15

model worked the best within a patient cohort (HRPCI). Their predictions in an out-of-sample test set in N = 10 patients using RNN-LMU reported a RMSE of 9.27 mmHg. We expand on this work by addressing generalization techniques to better learn patient cohort representations.

Features used were pump speed, pump pressure, pump flow, as well as LV pressure. The input signal at a given timestep is a rolling average over 10 seconds. Our sequence lengths are broken into 10 minutes. Each sequence is then broken into half to form the input and ground truth, respectively. The model takes an input signal of 5 minutes and outputs a sequence of 5 minutes. We categorized the MAP sequences into 3 types based on trends: increasing (I) sequences, decreasing (D) sequences, and stationary (S). In the clinical context, lower ranges in MAP are concerning. For Impella patients, the ranges of MAP can be binned as normal, managed, and critical. Normal MAP ranges are 80 and above, managed is between 60-80 mmHg, and critical would be below 60 mmHg. Changes in MAP in the managed and critical range are significant and would warrant closer observation or intervention. To bin prediction sequences as I, D, or S we perform a linear regression to get the slope and the midpoint of the MAP time series. A slope of $\pm$ 10 mmHg defines an I or D if the midpoint of the linear projection is above 80 mmHg (normal range). If below 80 mmHg (managed, critical), then a slope of $\pm$ 5 mmHg is used. For HRPCI, the number of I, D, and S samples is 7,260 , 7,534 , 67,569. For AMI/CGS,

Table 2. Cross-cohort performance comparison. 5 minute forecasting RMSE (mmHg) of different training/testing cohort combination with AMI/CGS(A), HRPCI(H) and both cohorts (B). Results are averaged over 3 random runs.

Train/Test	A/H	A/A	A/B	B/H	B/A	B/B	H/H	H/A	H/B
CLMU-S2S	**6.110**	**3.362**	**3.449**	**6.380**	4.422	4.529	**6.332**	**5.200**	**5.283**
LMU-S2S	6.235	3.448	3.536	6.986	4.956	5.074	6.604	5.459	5.540
CLSTM-S2S	6.180	3.364	3.454	6.490	**4.357**	**4.462**	6.531	5.467	5.529
LSTM-S2S	6.188	3.408	3.497	6.679	4.675	4.782	6.560	5.387	5.453
LSTM-MLP	6.202	3.491	3.576	6.536	4.712	4.828	6.450	5.369	5.457

Table 3. Cross-cohort generalization performance comparison for each trend category. 5 minute Forecasting RMSE (mmHg) for increasing (I), decreasing (D) and stationary (S) time series. Results are averaged over 3 random runs.

Train/Test	A/H			H/A		
Category	I	D	S	I	D	S
CLMU-S2S	**10.689**	10.585	**4.815**	**9.267**	12.099	**4.686**
LMU-S2S	10.914	10.649	4.940	9.771	11.997	4.954
CLSTM-S2S	10.979	10.585	4.865	9.403	12.388	4.965
LSTM-S2S	10.724	10.723	4.880	9.548	12.575	4.862
LSTM-MLP	11.501	**10.223**	4.929	9.911	**11.788**	4.856

the number of samples in each trend category is 109,522, 112,924, 2,578,005.

To generalize across patient cohorts, we introduce conditioning of deep sequence models based on time invariant cohort data. We take a cohort specific condition, and condition the starting state of a RNN cell based on this for the first timestep. An RNN cell generates the prediction by propagating the hidden state over time. Given input x_t at time step t, an RNN uses the following equations to predict the output y_t. In the case of forecasting, $y_t = x_{t+1}$.

$$h_{t+1} = \sigma(Wh_t + Ux_t + b), \quad y_t = \sigma(Wh_{t+1})$$

Here W, U and b are the trainable parameters, and σ represents the activation function. We then encode our condition variables with a multi-layer perceptron (MLP) to cohort-specific features, which are learned. The encoding module is simply a series of affine transformations followed by $\tanh$ activation and batch normalization, as the magnitude between condition features varies greatly. One such example can be described as follows:

$$[h_1, s_1] = \tanh(Wc + b)$$

where c is the cohort-specific feature. We found that the best conditioning parameters were the standard deviation of each feature according to cohort. A sequence-to-sequence model [12] consists of an encoder and a decoder, both RNNs. The final hidden state of the encoder is used as the starting state of the decoder. To condition on a cohort, we take a similar approach to learn the cohort-specific features and initialize the state of both the encoder and the decoder with the encoded conditional features. We can use different type of RNN cells in the conditional RNN framework, we used two variations: LSTM and LMU [10].

We evaluated the performance of deep sequence models on MAP forecasting across the HRPCI (H) cohort and the AMI/CGS (A) cohort. We rotate training and testing amongst the HRPCI, AMI/CGS, and a mixture of both patients cohorts (B). For the mixture of both, we train on an equal amount of each cohort. Few shot domain adaptation is the ability of a model to generalize to new cohorts despite having only seen a few examples of a new cohort, and having been primarily trained on a different population. For example, by training on the entire HRPCI set, with randomly selected AMI/CGS samples totaling 10% of the HRPCI set.

4. Results

In cross-cohort testing, for example, we train on mainly HRPCI samples, with only a few AMI/CGS samples, and then testing on AMI/CGS samples, as described by Table 2. The conditional RNN is able to discern between cohorts based on the condition encoding. The conditional LMU outperformed the conditional LSTM by a significant RMSE across most of the categories, including the few-shot domain adaptation category. However, it is also apparent that the conditional RNN models outperformed all baselines in all categories consistently.

Table 3 indicates the performance of each model across increasing/decreasing/stationary sequences cross-cohort. We refer to the few-shot domain adaptation experiment for each model. The conditional RNN is consistently the best for stationary and increasing sequences, but does not perform the best across decreasing sequences. The best model actually is the LSTM-MLP, for decreasing sequences on both A/H and H/A. However, the CLMU still achieves decent performance here, with the CLSTM slightly worse. There is possibly a trade off in performance between do-

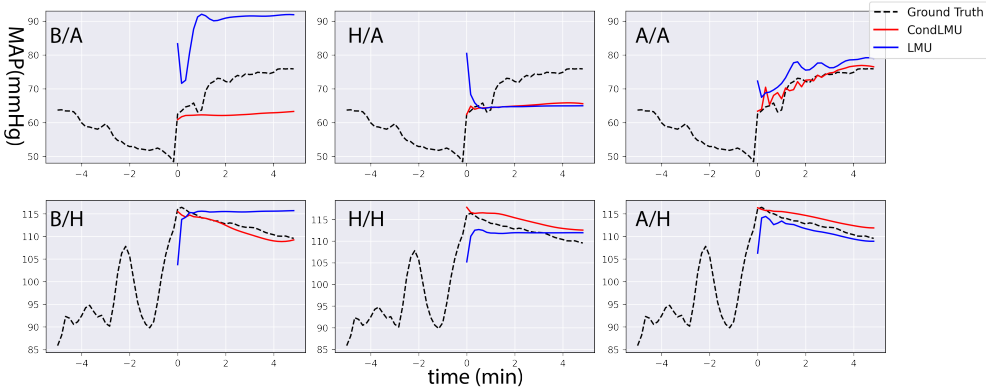

Figure 3. Visualization of predictions using models trained on different cohorts

ing better on stationary sequences and doing worse on increasing/decreasing sequences. As observed, the RMSE is much higher in increasing/decreasing cases on average, because these are less predictable/have more stochasticity.

Figure 3 visualizes the predictions of our proposed model Conditional LMU-Seq2Seq (red) with one of the baselines, LMU (blue) for different training and testing combinations, to examine the advantage conditioning grants over the corresponding ordinary RNN architecture. We observe that CLMU-S2S achieves noticeable improvement over the baseline in forecasting increasing and stationary trend time series. The baseline LMU tends to predict a relatively flat trend, in the examples observed above. For the conditional LMU, the visualizations demonstrate that the conditional LMU provides relatively accurate predictions within and across cohort.

5. Conclusion

We examined the forecasting performance out to 5 minutes across patient cohorts with clinically different presentations and interventions (HRPCI and AMI-CGS cohorts). Through the conditional RNN, we introduced a method of adapting model predictions and applying learning across cohorts with few samples. This indicates that at least for MAP Forecasting, there are learnings and features that we can apply cross cohorts that may have been neglected so far. In doing so, we improve performance overall and improve forecasting, to help patient monitoring to anticipate changes in MAP.

Acknowledgments

This work was sponsored by a research grant from Abiomed, Inc. (Danvers, MA).

References

[1] Chemla D, Antony I, et. al. Mean aortic pressure is the geometric mean of systolic and diastolic aortic pressure in resting humans. Journal of Applied Physiology 996 2278 2284 2005;.

[2] Dunser M, et al. Arterial blood pressure during early sepsis and outcome. Intensive Care Med 351225–1233 2009;.

[3] Dunser M, et al. Association of arterial blood pressure and vasopressor load with septic shock mortality: a post hoc analysis of a multicenter trial. Crit Care Lond Engl 13R181 2009;.

[4] Russo G, et. al. Hemodynamics and its predictors during impella-protected pci in high risk patients with reduced ejection fraction. Int J Cardiol 274221 225 2019;.

[5] S. Pitta RS. Management of acute myocardial infarction with cardiogenic shock: Systematic care and a stepwise intraprocedural guide. Society for Cardiovascular Angiography and Interventions 2018;.

[6] Connor, et. al. Recurrent neural networks and robust time series prediction. IEEE transactions on neural networks 1994;5(2):240–254.

[7] Su R, et. al. Long-term blood pressure prediction with deep recurrent neural networks. In 2018 IEEE EMBS International conference on biomedical & health informatics (BHI). IEEE, 2018; 323–328.

[8] Huang E, Wang R, et. al. Aortic pressure forecasting with deep learning. Computing in Cardiology Vol47 2020;.

[9] Hochreiter S, Schmidhuber J. Long short-term memory. Neural computation 1997;9(8):1735–1780.

[10] Voelker A, Kajic I, Eliasmith C. Legendre memory units: continuous-time representation in recurrent neural networks. Conference on Neural Information Processing Systems 2019;.

[11] Harjola VP, Lassus J, et. al. Clinical picture and risk prediction of short-term mortality in cardiogenic shock: clinical picture and outcome of cardiogenic shock. Eur J Heart Fail March 2015;17:501–509.

[12] Sutskever I, Vinyals O, Le Q. Sequence to sequence learning with neural networks. Conference on Neural Information Processing Systems 2014;.

Address for correspondence:

Elise Jortberg
22 Cherry Hill Dr, Danvers, MA USA 01923
ejortberg@abiomed.com