Master Thesis

Cardiac Output Estimation from Arterial Blood Pressure Waveforms using Echostate Networks

CNSystems Medizintechnik AG
Institute for Theoretical Computer Science, TU Graz

Author:
Martin Hammerschmied

Advisors:
Dr. DI Jürgen Fortin
Prof. Dr. Wolfgang Maass

October 15, 2012

STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

....

date (signature)

EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommene Stellen als solche kenntlich gemacht habe.

Graz, am ....

(Unterschrift)

Abstract

Cardiac output is an important parameter in intensive care medicine and anesthesia. Up to date there is no medically acceptable method to measure the true cardiac output directly. Moreover all commercially available devices capable of measuring cardiac output indirectly are either very complicated to apply and/or highly invasive and therefore risky.

In the past decade devices appeared that are able to estimate cardiac output from continuous blood pressure waveforms taken from peripheral arteries (CO-from-ABP methods). It is well established that this kind of blood pressure curves can be used for high quality estimation of cardiac output. All of these available devices depend on invasively measured input signals using an arterial catheter.

In this thesis I will develop a machine learning approach to estimate cardiac output from arterial blood pressure waveforms based on echo state networks (ESN). A novelty in the present approach to estimate cardiac output is the application of recurrent neural networks (the ESN).

The CNAP Monitor (CNAP) by CNSystems Medizintechnik AG provides a method of continuously measuring peripheral blood pressure in a non-invasive manner. In contrast to the vast majority of the literature the model is based on these non-invasively measured blood pressure curves. Additionally the photoplethysmographic CNAP signal is used.

The present work is based on the data of two different databases. The first database contains patient records including arterial blood pressure measured with the CNAP sensor from trials conducted by CNSystems Medizintechnik AG . The second database is the publicly available MIMIC II medical database including invasively measured arterial blood pressure. Both databases also include cardiac output measurements from different reference methods.

Zusammenfassung

Herzzeitvolumen ist eine wesentliche Messgröße in der Intensivmedizin und in der Anästhesie. Bis heute gibt es keine Methode, mit klinisch akzeptablen Risiko, um Herzzeitvolumen direkt zu messen. Weiters sind alle kommerziell erhältlichen Lösungen Herzzeitvolumen indirekt zu messen, schwierig in der Anwendung und/oder hoch invasiv und daher riskant.

Im vergangenen Jahrzehnt sind Apparate auf den Markt gedrungen welche imstande sind Herzzeitvolumen aus kontinuierlich gemessenen, arteriellen Blutdruckkurven zu schätzen (CO-from-ABP methods). Es ist bekannt, dass sich aus diese Art von Blutdrucksignalen eine gute Schätzung des Herzzeitvolumens ableiten lässt. Im Allgemeinen ist das jedoch eine schwierige Aufgabe. Alle verfügbaren Geräte sind von einer invasiven Blutdruckmessung mittels eines Katheters abhängig.

In dieser Arbeit werde ich eine eine Methode entwickeln, welche auf maschinelles Lernen mithile von "echo state networks"(ESN) basiert. Neu in dieser Herangehensweise ist die Verwendung rekurrenter neuronaler Netze (die ESN).

Der CNAP Monitor (CNAP) von CNSystems Medizintechnik AG stellt eine bequeme Methode dar um kontinuierlichen arteriel Blutdruck nichtinvasiv zu messen. Im Gegensatz zum Großteil der einschlägigen Literatur basiert das Model auf diesen nichtinvasiv gemessenen Blutdruck Kurven. Weiters findet ein CNAP-spezifisches photoplethysmografisches Signal Anwendung.

Die vorliegende Arbeit beruht auf Daten zweier verschiedenen Datenbanken. Die erste Datenbank besteht aus Patientendaten von einer Testreihe, durchgeführt von CNSystems Medizintechnik AG . Diese Daten enthalten Blutdrucksignale aufgezeichnet mit dem CNAP sensor. Die zweite Datenbank ist die frei verfügbare "MIMIC II medical database". Sie enthält invasiv gemessene, arterielle Blutdrucksignale. Beide Datenbanken beinhalten außerdem Referenzmessungen des Herzzeitvolumens.

Acknowledgement

This thesis has been performed at CNSystems Medizintechnik AG . First I want to thank my advisor DI Dr. Jürgen Fortin for inspiration, valuable input and guidance. My deepest gratitude to Prof. Wolfgang Maass for pointing me into the right directions and making this thesis possible and DI Dr. Gregor Hörzer for introducing me to the world of echo state networks. I also want to thank my colleagues at CNSystems for a very comfortable atmosphere and countless talks. My gratitude to Katja Maier for advice and supply of information, Thomas Brunner for explaining the world to me and Julian Grond for support and feedback for my thesis.

Further I want to thank my girlfriend Bettina for encouraging and supporting me throughout the whole time.

List of Acronyms

ABP Arterial Blood Pressure. Often used in conjunction with the terms *central* or *peripheral* to indicate the type of artery.

BA-diagram Bland-Altman diagram. A diagram depicting an error measure versus the reference's absolute value

BIBO Bound input bound output. A term to describe the stability of a dynamic system.

BP Blood pressure.

CNAP Continuous non-invasive arterial pressure. A measurement method developed by CNSystems Medizintechnik AG

CO Cardiac output. The average amount of blood ejected into the systemic circulatory system by the heart.

DD Descriptive Data. Additional information about a patient like weight or hight.

DFT Discrete Fourier transform.

EKG (ECG) Electrocardiogram.

ESN Echo state network. A kind of a recurrent neural network.

HR Heart-rate.

HTTP Hypertext transfer protocol.

ICU Intensive care unit.

MAP Mean arterial pressure.

MIMIC Multiparameter intelligent monitoring in intensive care. An open medical database used for scientific research.

MIT Massachusetts Institute of Technology.

PE Percentage error.

PPV Pulse pressure variation. Typically caused by raspiration cycles.

RMSE Root-mean-square error.

RNN Recurrent neural network.

SFBS Sequential forward backward feature selection. A method to approximate the optimal feature subset for a learning problem.

SSF Slope-sum-function. A function to emphazise the systolic ascent in a blood pressure pulse wave.

TPR Total peripheral resistance The lumped flow resistance in a human body's circulatory system.

VUT Vascular unloading technique. A method to measure arterial blood pressure waveforms non-invasively.

WFDB Waveform database format.

Contents

1 Introduction

The measurement of arterial blood pressure (ABP), i.e. recordings of continuous blood pressure waveforms, is standard procedure in critical care. These waveforms are typically measured at proximal limbs and are periodic pulses caused by heart strokes with an oscillating tail. An example is depicted in figure 1.3. The contour of these waveforms help physicians to assess a patients cardiovascular condition and identify cardiovascular diseases.

The ABP signal also carries potentially very valuable information which is not visible without further ado. Pulse pressure variation (PPV) for example is the magnitude of the pulse pressure oscillation related to the respiratory cycle of patients with mechanical ventilation. PPV is an excellent indicator for the body's fluid balance, to be calculated from the waveform.

Since arterial blood pressure is a direct consequence of the heart beats and the blood propagation through the circulatory system it can also be used to derive parameters of the circulatory system. One of these parameters is cardiac output, the amount of blood the heart is injecting into the circulatory system every minute, i.e. liters per minute. The idea to derive cardiac output from arterial blood pressure is almost a century old and greatly gained popularity in the last decade.

The question of which model is the best for this task is yet unsolved. Variations in the physiology and the condition of each patient make it very difficult to find a generic model that will work with everyone. In this work I will show that a generic model based on echo state networks is a sensible choice with promising results.

1.1 The Human Circulatory System

The cardiovascular (circulatory) system is a highly complex network of blood vessels. It is the main transportation system of body fluids within the human body. Among other tasks it supplies the body with components for metabolism (gases, nutrition matter), regulates body temperatures, is the central component of the immune system and carries messenger substances like hormones. An ill condition of the cardiovascular system disrupts the transportation mechanism and effects all systems within the human body. With the circulatory system failing other body functions start to fail within seconds resulting in irreversible damage within minutes [6].

The cardiovascular system is also a hydrodynamic mechanical system of parallel and serial connected arteries and veins. The schematic of such a cardiovascular system is

depicted in Figure 1.1. The central organ is the heart connecting two circulations. The *systemic circulation* transports blood away from the left half of the heart in order to supply organs with oxygenated blood. The blood vessels containing oxygenated blood are called arteries. As the arteries approach the organs they branch out becoming smaller, arrive at the arterioles and finally at a dense web of capillaries. This is where the gas exchange takes place. Oxygen is exchanged with carbon dioxide. The oxygen depleted blood runs back through the venules and venous to the right half of the heart. From there the blood runs through the *pulmonary circulation* in a similar fashion from the heart through the lungs in order to exchange the carbon dioxide with oxygen, finally arriving at the left half of the heart which closes the circle.

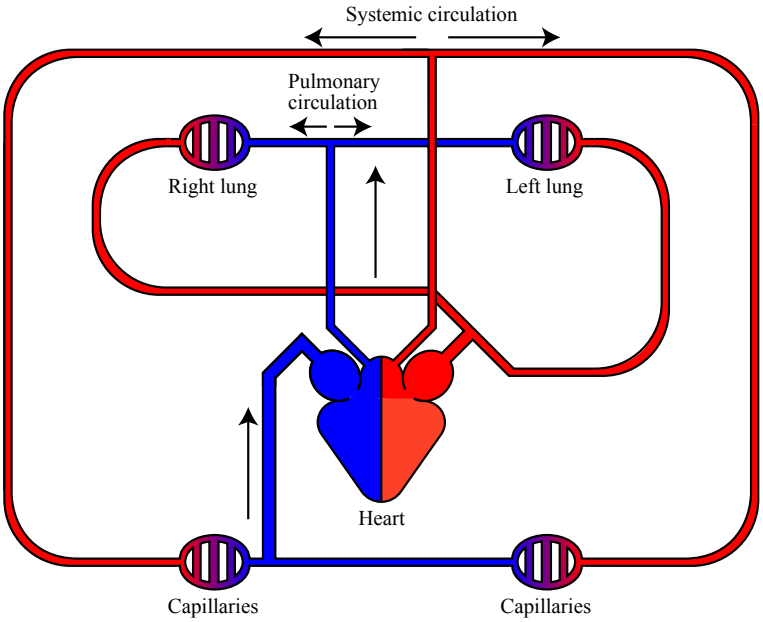

Figure 1.1: Human cardiovascular system. Red indicates oxygen rich blood, blue indicates oxygen poor blood.

Big arteries like the aorta have thick walls and are rather elastic. Therefore they are called *elastic arteries*. The walls of the smaller arteries contain more muscle tissue and are usually less elastic. They are called *muscular arteries*. Over a lifetime the tissue of blood vessels usually become stiffer. [6]

During a heart stroke the pressure in the arteries can almost double which causes them to expand. Depending on the size and elasticity of an artery it can be seen as a capacitor storing blood as it expands and simultaneously increasing the pressure. It has to be noted that unlike a normal capacitor the relation between pressure and volume within an artery is non-linear especially as arterial stiffness increases. The artery stiffness varies significantly on a time scale of months or years.

Blood has a certain viscosity and inertia. Together with the characteristics of blood vessels this results in a specific propagation velocity of blood pressure waves. During a heart contraction a pulse wave propagates forward through the arterial system. The

forward wave is reflected backwards at major and minor branches and interferes with forward pressure waves. This mechanism states another important source of non-linearity. Efforts have been made[12, 29] to determine pulse wave velocity and cancel out reflected waves in mathematical models.

When blood is running through vessels it is naturally exposed to friction. This causes a certain flow resistance depending on the blood's viscosity and the size of the blood vessels. The combined resistance of the circulatory system is called *total peripheral resistance (TPR)*. This TPR is subject to variation. Blood vessels are densely interspersed with muscle tissue controlling the vessels size and therefore its flow resistance. Various events like medication or shock can cause blood vessels to contract or dilate. The resistance can vary significantly within seconds. Also medication like anticoagulant that changes the blood's viscosity causes a change in flow resistance.

1.2 Peripheral Arterial Blood Pressure

Abnormal arterial blood pressure can be an indicator for a variety of diseases. The most common abnormalities are hypertension (high pressure) and hypotension (low pressure). The mean arterial pressure usually shifts upwards with advanced age due to arterial stiffness. Those abnormalities can be easily identified by measuring the maximum (systolic pressure) and the minimum (diastolic pressure) of arterial blood pressure pulse waves. This is usually done using an ordinary sphygmomanometer in combination with a stethoscope as depicted in figure 1.2 or an electronic beat sensor.

Figure 1.2: *Left*: A sphygmomanometer attached to an arm cuff. *Right*: A stethoscope

A lot of other abnormalities like extra systoles, valvular defects or cardiomyopathy diseases are more difficult to identify and require, among other examination, continuous

measurement of arterial blood pressure. Also for monitoring purposes in intensive care, continuous blood pressure measurement is crucial. The standard way to do this is through a catheter inserted into the radial artery (or some other suitable artery) holding a pressure sensor at it's end. Figure 1.3 shows an example of an arterial blood pressure curve taken from the radial artery. Such an arterial line carries the risk of infection and requires a medical doctor to place the catheter.

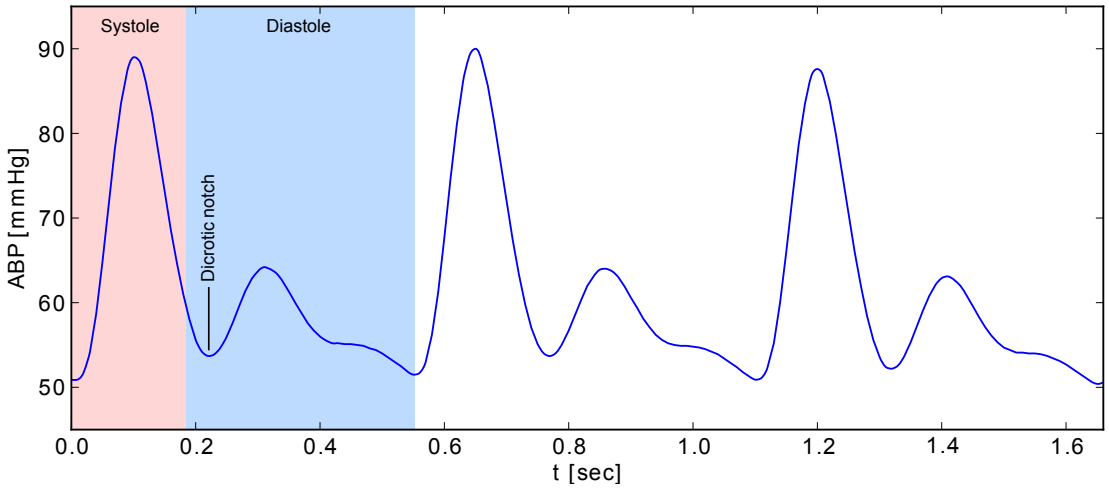

Figure 1.3: Peripheral arterial blood pressure curve taken invasively from the radial artery. The red shaded region marks a phase called *Systole* in which the heart ejects blood into the circulatory system. The blue shaded region marks a phase called *Diastole* in which the aortic valve is closed and the heart is not ejecting blood into the system.

The blood pressure curve in figure 1.3 shows the characteristic pulse wave that emerges in a peripheral artery during a heart stroke. The stroke is typically temporarily divided into a systolic and a diastolic phase.

The systolic phase is the time period in which the heart contracts. Parts of the blood within the right ventricle (chamber filled with blood) are pressed into the aortic arch and further into the arterial tree. In this period (marked red in the figure) the arterial pressure rises to the peak systolic pressure and falls as the heart contraction fades out.

At a certain point after the heart contraction the pressure in the right ventricle drops below the aortic pressure. As a result the semilunar aortic valve closes and the blood flow is interrupted. This event marks the beginning of the diastolic phase. During the diastolic phase (marked blue in the figure) the vascular system can be seen as free dynamical system without a power source. The curve clearly shows the characteristics of a decaying oscillation. The minimum between the systolic maximum and the dicrotic bump (first diastole maximum) is called dicrotic notch. It is typically distinctive and therefore often used to identify the beginning of the diastole. It is important to note that the dicrotic notch and the closing of the aortic valve is not equivalent. Depending on the "niceness" of the pulse wave it is particularly tough to find the closing point of the aortic valve.

1.2.1 Modified Windkessel Model

Over the last century a lot of models where proposed to describe the cardiovascular system in a mathematical way. The most popular model turned out to be the *Windkessel* model. The origin of the model is a method to regulate pressure in a hydraulic system utilizing a reservoir named after the german word "Windkessel". The basic structure and functionality is depicted in Figure 1.4. In this example a "Windkessel" is used to balance the water pressure that drives a turbine. The Windkessel has an inlet supplying the system with water and an outlet with a built in turbine. The Windkessel is partially filled with air. As the inlet pressure increases more water fills the chamber and compresses the air cushion. Thus the pressure cannot jump and pressure variations are smoothed out. If the air cushion was not there pressure peaks would directly effect the turbine. The equivalent electric circuit would be a capacitor for the Windkessel and in parallel a resistor for the turbine.

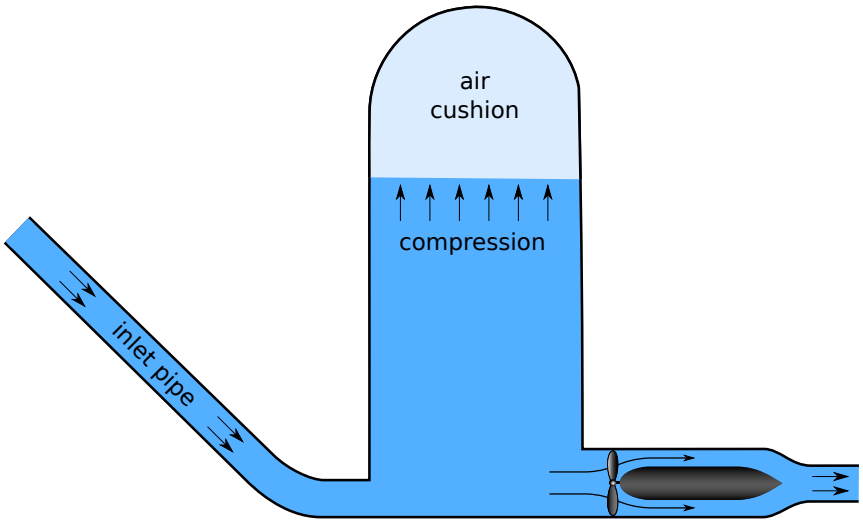

Figure 1.4: Schematics of a "Windkessel"

The heart as source and the aortic arch as elastic liquid reservoir behave in a way very similar to the Windkessel model. To cover the whole circulatory system the model was refined as depicted in Figure 1.5. The big and the small blood vessels are modeled separately as lumped capacitors. The inertia of the blood is introduced as an inductance. The *total peripheral resistance* is represented as a resistor. This refined model is known in the literature as the *modified Windkessel* model. Only arterial blood vessels are covered in this particular version of the model. Dynamics that arise from the venous vessels are rather small and can be neglected. The easiest way to include them in the model would be a small constant voltage drop after the TPR. In the modified Windkessel model the peripheral arterial blood pressure (ABP) appears as voltage drop at the TPR. Cohn *et al.* [2] demonstrated methods on how to use this type of model for early detection of vascular diseases.

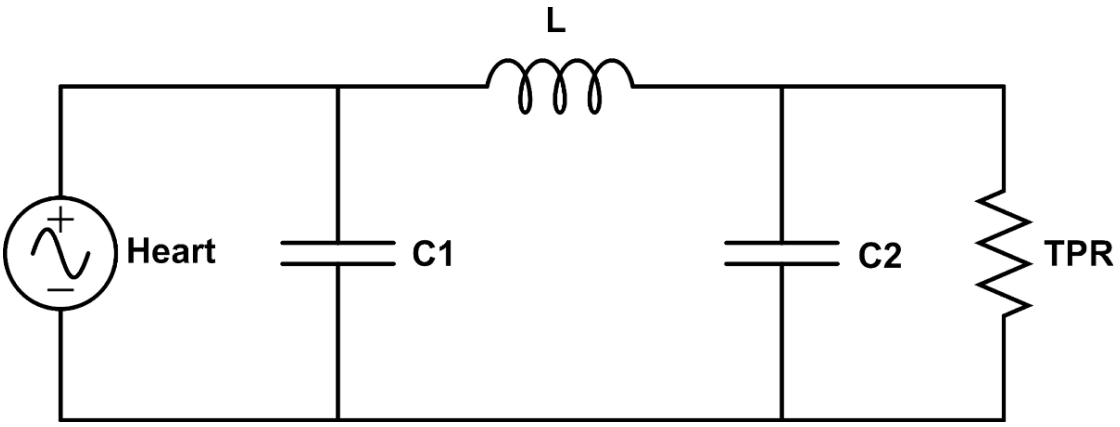

Figure 1.5: Schematics of a modified Windkessel system. $C1$ represents the elasticity of the big blood vessels close to the heart. L represents the blood's inertia. $C2$ represents the elasticity of the small blood vessels. TPR is the total peripheral resistance and represents the combined flow resistance

The modified Windkessel model is a third degree linear model and can be expressed in the standard state-space form

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{b}u, \\ y &= \mathbf{c}^T\mathbf{x}, \end{aligned} \tag{1.1}$$

with

$$\mathbf{x} = \begin{pmatrix} u_{C_1} \\ i_L \\ u_{TPR} \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} 0 & -\frac{1}{C_1} & 0 \\ \frac{1}{L} & 0 & -\frac{1}{L} \\ 0 & \frac{1}{C_2} & -\frac{1}{RC_2} \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} \frac{1}{C_1} \\ 0 \\ 0 \end{pmatrix}, \quad u = i,$$

$$\mathbf{c} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \qquad y = ABP.$$

u_{C_1} is the voltage drop at C_1, i_L the current through L and u_{TPR} the voltage drop at TPR. i is the current that flows into the system and represents the blood flow from the heart. The output y of the linear system is the ABP which happens to be equal to u_{TPR}.

The modified Windkessel model is used to the present day to determine cardiac output by a method called *diastolic decay* which will be introduced in chapter 1.3.3.

1.2.2 Vascular Unloading Technique

Peripheral arterial blood pressure is usually measured in the radial, brachial or femoral artery directly using a catheter with a pressure sensor. The resulting continuous measurement is often referred to as *A-Line* in the literature. Although minimal invasive it is

not without risks and requires trained personnel. In 1973 a new method for measuring arterial blood pressure curves non-invasively called the *vascular unloading technique*[30] was proposed.

The basic idea is to illuminate a limb with infrared light. Research has shown[34] that a finger is the optimal spot for this task. The light gets partially absorbed by the blood in the finger. The more blood is present the more light gets absorbed. A photoelectric detector measures the light that escapes the finger and therefore indirectly measures the blood volume. Obviously blood pressure is directly related to blood volume. Due to the physiology of blood vessels this relationship is non-linear, hard to determine and usually varies substantially between subjects. Thus it is very hard to derive pressure from volume directly.

To circumvent the unknown volume-pressure-relationship the approach suggests to apply a feedback-controlled counter pressure on the finger in order to keep the blood volume in the vessel constant. As a result the counter pressure is proportional to the pressure within the vessel, although a part of the counter pressure is absorbed by the tissue surrounding the vessel which can be a source of error. Ideally the pressure applied to the vessel from the outside should equal the pressure within the vessel to achieve true unloading of the vessel's wall.

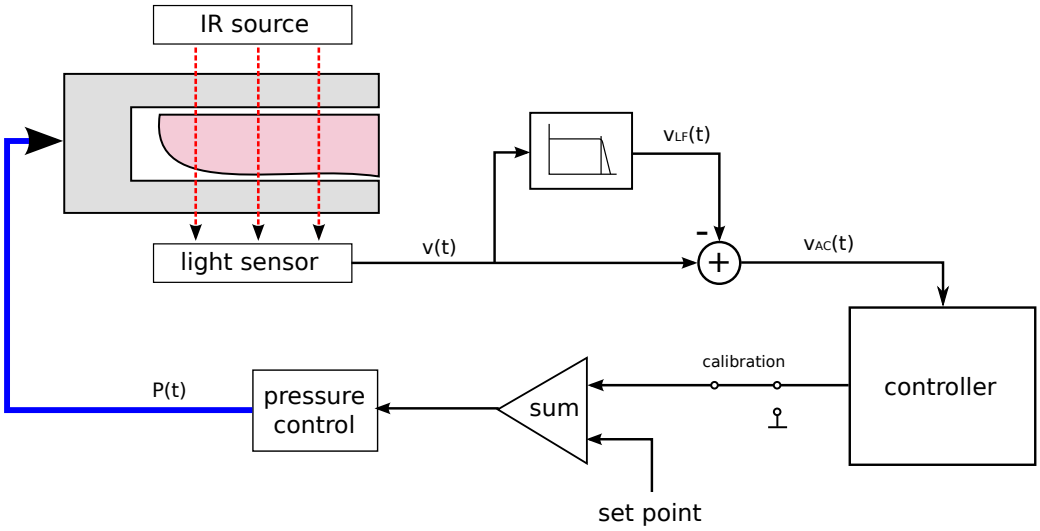

Figure 1.6: Schematics of the opened/closed control loop of a vascular unloading pressure sensor. The choice of the set point is essential to achieve true unloading and valid absolute pressure values.

Figure 1.6 shows the basic schematics of the vascular unloading technique (VUT). A pneumatic finger cuff is used to apply pressure to the finger. Built into the cuff is an infrared light source and a photoelectric light sensor. Together they measure the blood volume in the finger. The light signal has an inverse relationship with the blood volume. The subtraction of the low frequency parts of the volume signal ensures independence

from the set point used to calibrate the feedback control loop. The controller shows proportional, differential and integral (PID) behavior and is optimized to ensure optimal trade off between speed and stability.

The set point is used to calibrate the point of operation and is one of the most critical parameters. A wrong set point leads to constant tension of the vessel's wall and therefore to distorted pressure curves. Also a correct set point is necessary to get valid absolute pressure values. Various vascular effects like vasoconstriction and vasodilation lead to a set point drift on a time scale of minutes. This drift is the most ambitious challenge to tackle in the VUT approach. Continuous tracking of the set point is a non-trivial problem and without it the VUT is almost useless.

1.2.3 CNAP sensor

The commercially available CNAP sensor developed by CNSystems Medizintechnik AG is a substantial advancement of the vascular unloading technique [9]. The most important improvements include an effective set point tracking mechanism.

Based on the vascular unloading theory the pressure-volume-relationship within a blood vessel follows a S-shaped function similar to a sigmoid as depicted in Figure 1.7. To calibrate the set point the feedback loop between the light sensor (volume signal) and the controller is opened (see figure 1.6). By varying the set point the range of operation is shifted along the S-shaped function. At the inflection point of the function the vascular wall is unloaded. In this range the pressure-volume-relationship is almost linear and the volume signal shows a maximum in variance.

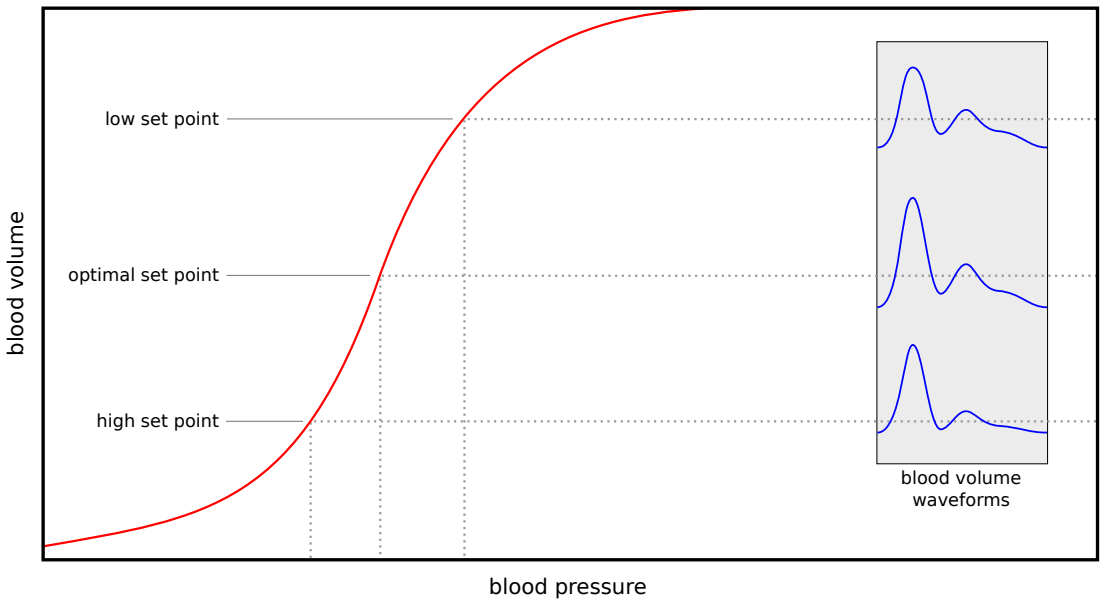

Figure 1.7: The red s-shaped function describes the non-linear relationship between blood volume and blood pressure in the finder. During the CNAP calibration phase this relationship is exploited to find the optimal set point.

Calibration

If the set point is too high the external pressure compresses the vessel until it starts to collapse. As a result the sub-mean parts of the volume signal gets compressed. If the set point is too low the vascular wall is exposed to increasing tension. This unveils the non-linear behavior at high tension and the super-mean part of the volume signal gets compressed. In both cases the variance of the volume signal is lower compared to the optimal set point.

This can be exploited to initialize the set point. The following steps are required:

- Open the feedback loop
- Perform a set point sweep
- Observe the variance of the light signal $v(t)$ (blood volume)
- Initialize the set point at the variance maximum

Continuous Tracking

During closed loop operation the set point needs continuous tracking since it is subject to drift. Depending mostly on the hemodynamics (vasoconstriction and vasodilation) the shape of the function in Figure 1.7 can change significantly within minutes or even seconds. The CNAP sensor tackles this problem by detecting basic contour changes of the pressure pulse wave.

The control error signal $v_{AC}(t)$ (see Figure 1.6) is bias free if hemodynamics are constant. If the hemodynamics change, either the sub-mean or the super-mean parts of $v_{AC}(t)$ get compressed. This results in a mean change of

$$P_n = \int_{T_{0,n}}^{T_n} v_{AC}(t)\,\mathrm{d}t \quad \neq 0, \tag{1.2}$$

where P_n is the mean of $v_{AC}(t)$ over the n-th beat. If P_n is bigger than zero the set point is too high and vice versa. An adaptive set point correction CD_n (correction for the n-th beat) is derived using P_n over a linear transfer function.

$$CD_n = T\{P_n\} \tag{1.3}$$

The pressure in the finger cuff is not exactly the same pressure applied to the arterial blood vessels in the finger. Tissue around the vessel of unknown shape and mechanical properties introduce their own transfer behavior. Assuming that this transfer behavior consists mostly of a constant and a proportional term and stays constant over a long period of time we can apply another simple correction. An ordinary sphygmometer is used to exactly measure the systolic and diastolic pressure. The known finger cuff pressure is then shifted and scaled according to the sphygmometer measurement.

$$ABP(t) = k \cdot P_{AC}(t) + P_{DC}(t) + d \tag{1.4}$$

$P_{DC}(t)$ refers to the constant component of the cuff pressure, $P_{AC}(t)$ to the alternating component, respectively.

1.3 Cardiac Output

The heart as central organ in the cardiovascular cycle is the only power source in terms of pump activity. The valves within the heart are passive elements that transform heart contractions into a pump mechanism. During the systolic phase blood is ejected into the aorta (systemic cycle) and the pulmonary artery (pulmonary cycle). The volume of ejected blood in a single stroke is usually about 70 ml. Assuming an average heart rate of 70 beats/min that amounts to almost 5 l/min. The average amount of ejected blood in liter per minute is called *cardiac output (CO)*. Under heavy physical stress the cardiac output can increase up to 25 l/min. This is an effect of elevated heart rates and more powerful heart contractions caused by the sympathetic nervous system. Additionally the heart has some self-regulatory mechanisms. By increasing venous pressure, more blood flows into the ventricle during the diastolic phase. The *end diastolic volume* and therefore the stroke volume increases as a result. Numbers and descriptions where taken from the book of Thews *et al.* [6].

Cardiac output is nowadays a crucial parameter in intensive care medicine. It is an indicator of a variety of cardiovascular diseases. Also during surgery cardiac output is among other measurements an important source of information about the patients condition. Especially anesthesia control and medication during surgery usually cause a change in cardiac output. Also reactions like septic shocks usually lead to a change in cardiac output.

The cardiac output is typically defined as heart rate times the stroke volume:

$$CO = HR \cdot SV \tag{1.5}$$

Most methods for measuring cardiac output directly focus on measuring the stroke volume. The heart rate can be derived straight forward from a variety of standard signals (EKG, ABP, etc.). A few standard approaches to measure cardiac output are sketched in the next view sections.

1.3.1 Echocardiography

Doppler Echocardiography is a method where ultrasound is used to measure the velocity of blood flow in a vessel. As a special application it can be used to measure cardiac output non-invasively. In a first step one has to estimate the cross-sectional area A of the aorta at a particular spot. Doppler Echocardiography is then used to derive the flow velocity $v(t)$ over a stroke. The stroke volume can then be derived by integrating over $v(t)$.

$$SV_n = A \int_{beat_n} v(t)\,\mathrm{d}t \tag{1.6}$$

Clinical tests were conducted [16, 33] comparing echocardiography with methods like the Fick method or thermodilution which where accepted standard procedures at that time. The conclusion was that doppler echocardiography is a valid non-invasive alternative to the established methods after addressing some difficulties that lie within this method:

- The cross-sectional area of the aorta is not constant due to its elastic characteristics. Therefore it is not necessarily clear where the optimal spot at the aorta is to conduct the measurement. There seems to be a consensus about choosing a spot as close as possible to the aortic valve. Erroneous estimation of the cross-sectional area can have a great impact on the results and usually leads to overestimation of CO.
- In order to measure blood stream velocity the direction of the ultrasound waves should be as parallel as possible to the flow direction of the blood. Since this is not possible (25° deviation is usual) the deviation angle must always be known and constant. In clinical situations this can be difficult to accomplish.

Nevertheless doppler echocardiography is today a well established non-invasive method for estimating cardiac output.

1.3.2 Indicator Dilution

Another indirect method of measuring cardiac output is through the injection of indicator dilutions. The injection takes place inside the left ventricle using a catheter. The dilution is mixing up with the blood. The concentration of the dilution in the mixture is then measured at a point where a homogeneous mixture of the blood and the dilution can be assumed. The dilution concentration decays over time. By injecting a dilution with the volume V_d at the beginning of a heart stroke and measuring its concentration $q(t)$ over the whole stroke, the stroke volume can be calculated as:

$$SV = \frac{V_d}{\int\limits_{beat} q(t)\,\mathrm{d}t} \tag{1.7}$$

Usually it is assumed that the dilution concentration decays within the duration of a heart stroke.

Over time different types of dilutions where used. Very early dye dilutions where proposed [7]. Later this method has been replaced with thermodilution [10], where a dilution is significantly colder than the blood. The temperature of the blood drops after injection as depicted in figure 1.8. Over time the temperature decays to normal. The dilution concentration is indicated by the temperature difference. Thermodilution measurements can be repeated quickly and indefinitely. This is not the case when using dye dilution.

Although highly invasive it remains one of the most accurate measurement of cardiac output to the present day. Due to the risks and efforts needed to apply thermodilution measurements it is only performed for patients whenever such an accurate measurement is indispensable.

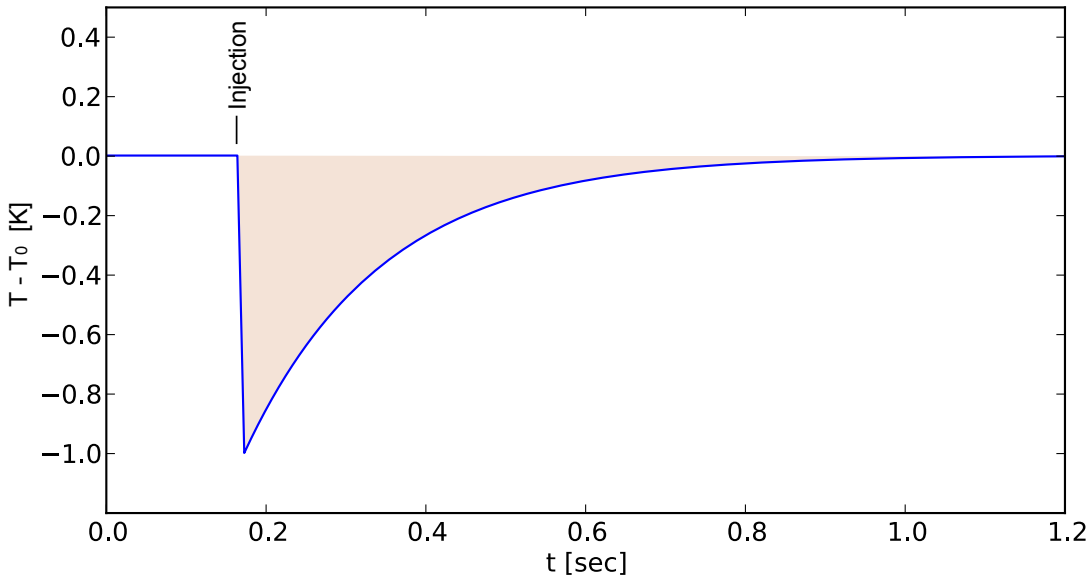

Figure 1.8: Temperature deviation from the nominal temperature. At a certain point a dilution with significantly lower temperature than the blood is injected into the blood stream. The graph represents the concentration of the dilution in the blood.

1.3.3 Diastolic Decay

Like the new approach presented in this work diastolic decay is a method where continuous arterial blood pressure measurements are used to estimate cardiac output [11]. The method is based on the modified Windkessel model presented in chapter 1.2.1. The model as shown in figure 1.5 implies that arterial blood pressure is the voltage drop u_{TPR} at the total peripheral resistance as a result of a pulsatile power source. As the source (the heart) is unknown - only u_{TPR} is known - the parameters of the model cannot be identified. But the source is only active during the systolic phase. After the aortic valve is closed the source is detached from the rest of the model. During the diastolic phase the model performs a free decaying oscillation. In the system equations (1.1) this is equivalent to **b** being the zero vector.

One can solve the homogeneous differential equation in order to find a solution for u_{TPR} in the source-free system. As u_{TPR} in the model is equivalent to ABP we get a solution for the diastolic arterial pressure:

$$ABP(t) = a_1 \cdot e^{-\frac{1}{\tau_1}t} + a_2 \cdot e^{-\frac{1}{\tau_2}t} \cdot \cos(\omega \cdot t + \phi) \tag{1.8}$$

This solution is only valid during the diastolic phase of each pulse wave and contains 6 unknown coefficients. The task at hand is to identify the diastole within a recorded pulse wave and fit the solution function into the blood pressure samples as shown in figure 1.9. This is a non-linear optimization problem with local minima and has no closed solution. A new iterative algorithm, designed exclusively for this task, was introduced 2000 by Heller *et al.* [13]. It is able to solve this fitting problem in significantly less iterations than standard approaches like the Levenberg-Marquardt-Algorithm [27].

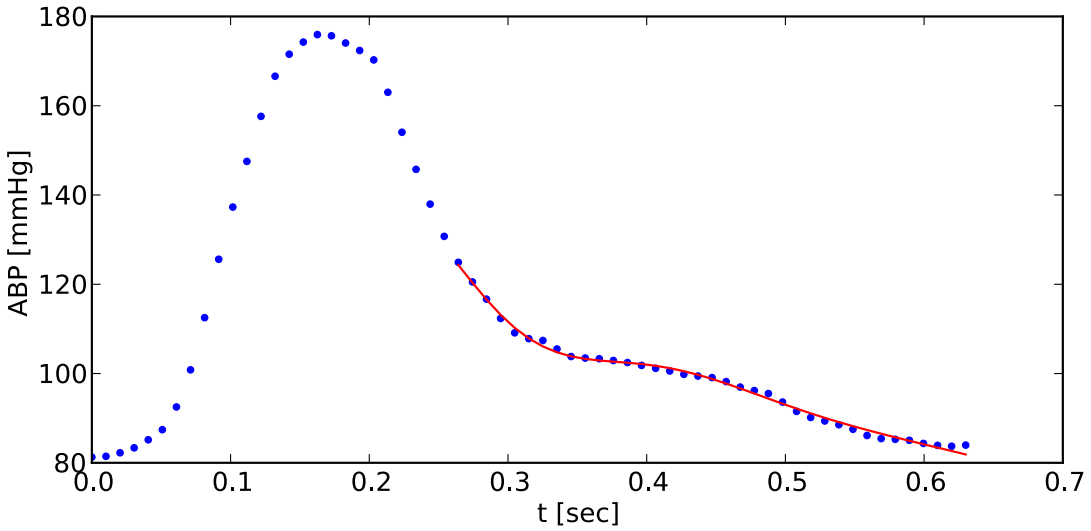

Figure 1.9: The solution of the homogeneous system of equations (red line) fitted into the diastolic part of the sampled blood pressure pulse wave (blue dots)

Cardiac output, defined as blood volume ejected by the heart per minute can be considered an averaged quantity and thus the very low frequency parts of the cardiac flow. At very low frequencies the capacitors are opened and the inductance is shortened. What remains is the mean arterial pressure (MAP) at the total peripheral resistance. Their relation is simply defined by ohm's law:

$$CO = \frac{MAP}{TPR} \tag{1.9}$$

Mean arterial pressure is a known quantity. What we have to gain from a parameter identification procedure is the total peripheral resistance. Three of the parameters in (1.8) are related to the initial condition of the system (namely a_1, a_2 and ϕ). The other three parameters encode the four circuit components (C_1, C_2, L). Thus we have four unknowns for only three equations. Only if one of the components is determined separately the TPR has a unique solution.

τ_1 appears as a factor in the TPR. Since all other components are more or less constant compared to the fast changing τ_1, the TPR can be interpreted as proportional to MAP over τ_1:

$$CO \propto \frac{MAP}{\tau_1} \tag{1.10}$$

In order to continuously estimate cardiac output from this model the following steps have to be performed:

- Estimate τ_1 according to the Windkessel model
- Use a different method to obtain a cardiac output reference measurement
- Use the reference value to determine the proportional factor between τ_1 and TPR

- Continuously redetermine τ_1 using the Windkessel model and derive cardiac output.

Another recent proposal by Mukkamala *et al.* [28] suggests to model the systoles as array of predefined pulses scaled with the pulse pressure. That makes it possible to determine the exact system parameters of the Windkessel model since its input is known. Further a more robust single estimation can be obtained using several successive beats.

1.3.4 Fick Method

The Fick principle to measure cardiac output is the oldest method, known for over 60 years [26]. It is based on the Fick method to measure the concentration of a substance within a mixture. In the present case one needs to know the oxygen concentration within the blood. At first the following three quantities must be obtained:

- Q_a: The concentration of oxygen in the blood after enrichment in the lungs in units of mg/liter. For convenience this is usually measured in a peripheral artery.
- Q_v: The venous concentration of oxygen in units of mg/liter. This measurement typically takes place in the pulmonary artery.
- O_c: The oxygen consumption in units of mg/min. This can be measured using an ordinary spirometer.

The cardiac output can then easily be calculated as:

$$CO = \frac{O_c}{Q_a - Q_v} \tag{1.11}$$

Likewise the carbon dioxide consumed by the lungs could be used as indicator substance.

This method has proven to be difficult due to the measurement of the indicator concentration. Since it is not less invasive it was mostly replaced by the thermodilution method.

2 Methods

Methods to estimate cardiac output (CO) from arterial blood pressure (ABP) were around for several decades. Numerous mechanical and computational models where proposed with various success, the most prominent example being the Windkessel model introduced in chapter 1.2.1.

The human circulatory system is a highly non-linear system. Non-linearities are caused by phenomena like wave reflections, pressure-volume-relationship in blood vessels or venous blood pockets. Therefore it has been attempted to derive non-linear models. These models' structures are typically complicated and have often proven not to yield better results than simpler linear models. One reason is the lack of reliable methods to identify model parameters in individual humans during a measurement.

In the present work I will show how an established machine learning method can be applied on the task at hand in order to train a generic, non-linear model. Such a model has no need for parameter identification and the theoretical potential to learn an arbitrary complex dynamic behavior.

2.1 Echo State Networks

Echo state networks (ESN) have simultaneously been developed by Herbert Jaeger [17] and Wolfgang Maass *et al.* [22] (Liquid State Machines) and are a novel approach to recurrent neural networks. The basic idea is to excite a recurrent neural network (RNN), the *reservoir*, with a set of input signals in order to transform these signals into a high-dimensional non-linear space. The states of the neurons are then interpreted in order to estimate a desired output signal.

A descriptive analogy is to see the reservoir as a *liquid*. Input signals excite the liquid in the reservoir and cause waves. The waves interfere as they propagate through the liquid and are reflected at the reservoir walls. A snapshot of the wave situation is a representation of the input signal's history. The goal is now to learn how to make quantitative or qualitative interpretation of the snapshots.

The structure of a simple ESN is depicted in figure 2.1. For simplicity all signals are considered to be sampled, time-discrete signals. A vector of input signals $\mathbf{u}[n]$ excites the neurons in the reservoir. The states of the neurons, combined in the state vector $\mathbf{x}[n]$, can be regarded as projected representations of the input signals in a high-dimensional, dynamic space. Since each neuron has a non-linear activation function, the neural network is non-linear as well.

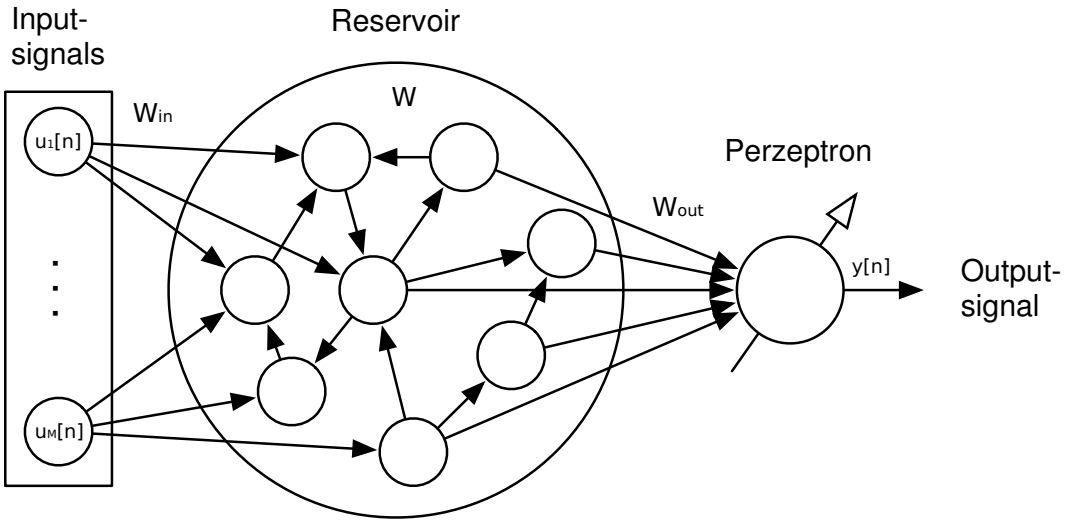

Figure 2.1: Schematics of a simple Echo State Network

The update equations of the ESN for each time step (defined by the sample rate of the input signals) are defined as:

$$\mathbf{x}[n] = f\left(\mathbf{W}_{in}\mathbf{u}[n] + \mathbf{W}\mathbf{x}[n-1] + \mathbf{w}_{bias}\right) \quad (2.1)$$

$$y[n] = g\left(\mathbf{w}_{out}^{T}\mathbf{x}[n] + w_{b,out}\right) \quad (2.2)$$

$\mathbf{W}_{in}$ is a matrix containing input weights. These weights are used to control the influence the input signals have on each neuron. $\mathbf{W}$ is a weight matrix defining how the neurons in the reservoir excite each other. $\mathbf{W}_{out}$ is an output weight matrix. It is the quasi-interpretation of the neurons' state. $\mathbf{w}_{bias}$ and $w_{b,out}$ are bias weights.

In each time step n each neuron is assigned a new state which depends on three terms. The first term is a weighted sum of all neurons' previous state $\mathbf{x}[n-1]$. The second term is a weighted sum of all input signals at time n. The third term is a bias which is different for each neuron. Finally the result is mapped by an activation function f which is typically a sigmoid function ($tanh$). The neurons' states are combined in a perceptron (simple linear combination) and yield a desired output signal $y[n]$. Optionally, the output function g can also be some non-linear activation function. The perceptron is often referred to as read-out.

More complex versions of ESNs often contain the following additional structures:

- The ESN's output is fed back into the reservoir using an additional weighting matrix/vector $\mathbf{W}_{fb}$. This can enhance the networks capabilities in order to model systems like oscillation generators. Since cardiac output is an averaged quantity it is not expected to oscillate.

- Additionally to driving the reservoir, the ESN's input signals are also fed directly into the output perceptron. ABP is a pulsatile signal on a much faster time-scale than CO and direct feed-forward of the input signals is not sensible.
- Holzmann *et al.* [14] introduced *sum&delay* readouts in order to enhance long-term memory capabilities of echo state networks. In this thesis memory behavior may sufficiently be controlled by using leaky-integrate neurons, which will be introduced in this chapter.

Besides the size of the reservoir the ESN is only defined by the weight matrices $\mathbf{W}_{in}$, $\mathbf{W}$ and $\mathbf{W}_{out}$. $\mathbf{w}_{bias}$ can be incorporated into the weight matrix $\mathbf{W}$ by introducing an additional state in $\mathbf{x}[n]$ which is always 1. The same way $w_{b,out}$ can be incorporated in $\mathbf{w}_{out}^T$, reducing the update equations to:

$$\mathbf{x}[n] = f\left(\mathbf{W}_{in}\mathbf{u}[n] + \mathbf{W}\mathbf{x}[n-1]\right) \tag{2.3}$$

$$y[n] = g\left(\mathbf{w}_{out}^T\mathbf{x}[n]\right) \tag{2.4}$$

Assuming N is the number of neurons in the reservoir and M the number of input signals. Further, $\mathbf{W}_{in}$ is an $(N+1)\times M$ matrix, $\mathbf{W}$ an $(N+1)\times(N+1)$ matrix and $\mathbf{w}_{out}^T$ an $(N+1)\times 1$ vector.

Previous approaches to recurrent neural networks (other than ESNs) typically aimed to learn the weights in the recurrent network. This has a couple of disadvantages as described by Lukoševičius *et al.* [21]:

- RNNs are usually trained using gradient descent methods. The gradient information can degrade during the learning process, thus convergence cannot always be guaranteed.
- The network weights can only be learned iteratively which makes it computationally expensive and only feasible for small networks.
- State of the art learning algorithms are hard to parameterize and require comprehensive experience.

As most important advantage of echo state networks, input weights $\mathbf{W}_{in}$ as well as the reservoir weights $\mathbf{W}$ are *generated randomly* at the beginning and remain constant. Only the output weights $\mathbf{w}_{out}^T$ are adjusted in order to minimize a certain error function.

Training

It is best practice to split a given data set into a training set and a test set in order to determine the true error. The model is then trained exclusively with the training set and tested with the test set. The input signals in the training set are here called $\mathbf{u}_{train}[n]$, the ones in the test set $\mathbf{u}_{test}[n]$. Respectively the output values (CO) in the training set are called $t[n]$ (*targets*) and in the test set $y[n]$.

Instead of having fixed training and test sets it is often feasible to split the whole data into several small subsets. The advantage is that one can conduct several trials in which the test set is always a different one of these small subsets. This concept is called *cross-validation* and is depicted in figure 2.2. The extreme version is to leave only a single sample out in each trial. This is called *leave-one-out cross-validation* and takes as many trials as there are samples in the data.

Figure 2.2: Cross-validation: In each trial a different subset of the data is excluded from training and afterwards used as test set.

During the training phase one has to simulate the reservoir with the given training input $\mathbf{u}_{train}[n]$ to obtain corresponding reservoir states $\mathbf{x}_{train}[n]$ using the reservoir update equation (2.3). Let $\mathbf{X}_{train}$ be a $K \times (N+1)$ matrix containing the combined reservoir states for all time steps. K is the number of time-steps and N the number of neurons in the reservoir. Respectively $\mathbf{t}$ is a $K \times 1$ vector containing the combined target values for all time-steps.

We are now using (2.4) and plugging in the targets and training sets in matrix format:

$$\begin{aligned} \mathbf{t} &= g\left(\mathbf{X}_{train} \cdot \mathbf{w}_{out}\right) && (2.5) \\ g^{-1}\left(\mathbf{t}\right) &= \mathbf{X}_{train} \cdot \mathbf{w}_{out} && (2.6) \end{aligned}$$

We are now seeking a solution for $\mathbf{w}_{out}$ which minimizes the error of this equation. We therefore define a quadratic error function:

$$\begin{aligned} E(\mathbf{w}_{out}) &= \sum_{k=1}^{K} \left(\mathbf{x}_{train}[k] \cdot \mathbf{w}_{out} - g^{-1}\left(t[k]\right)\right)^2 && (2.7) \\ &= \left(\mathbf{X}_{train} \cdot \mathbf{w}_{out} - g^{-1}\left(\mathbf{t}\right)\right)^T \left(\mathbf{X}_{train} \cdot \mathbf{w}_{out} - g^{-1}\left(\mathbf{t}\right)\right) && (2.8) \end{aligned}$$

Assuming the activation function g to be *bijective* (or at least *injective* in the range of the targets) the result (2.6) is a system of K linear equations. The *least-squares* optimal solution for (2.6) is simply obtained by constructing the *pseudo-inverse*.

$$\begin{aligned} \mathbf{w}_{out} &= \underbrace{\left(\mathbf{X}_{train}^T \mathbf{X}_{train}\right)^{-1} \mathbf{X}_{train}^T}_{\text{pseudoinverse } \mathbf{X}_{train}^+} \cdot g^{-1}\left(\mathbf{t}\right) \\ &= \mathbf{X}_{train}^+ \cdot g^{-1}\left(\mathbf{t}\right) \end{aligned} \qquad (2.9)$$

In machine learning tasks one must always carefully analyze the training data. A very important step is to estimate the needed model complexity in order to prevent *underfitting*, a situation in which the model is not complex enough to learn the desired behavior, regardless of the amount and quality of training samples.

On the other hand there exists the phenomenon of *overfitting*. This is the case when the amount of training samples is too small or the samples do not cover the desired input/output space thoroughly enough. As a result the model *specializes* on the training samples and performs poorly on new samples. In those cases the complexity of the model must be restrained in order to generalize.

The way at hand to control the complexity of an echo state network is to change the size of the reservoir. This however can have unwanted side-effects:

- Changing the reservoir size has a direct influence on the computation complexity. Adding training samples at a later point can make it necessary to increase the reservoir size and therefore the required computation capacity to achieve optimal results.
- Changing the reservoir size can have direct influence on the optimal settings for other reservoir properties like scaling of input weights or reservoir weights.

A better way to control the model's complexity is to use *ridge regression* instead of the standard *pseudo-inverse*. Ridge regression is a method that introduces a regularization term which favors solutions with smaller weights:

$$\mathbf{w}_{out} = \left(\mathbf{X}_{train}^T \mathbf{X}_{train} - \gamma \mathbf{I}\right)^{-1} \mathbf{X}_{train}^T \cdot g^{-1}\left(\mathbf{t}\right) \tag{2.10}$$

For $\gamma = 0$ the solution will be the *pseudo-inverse* solution. The bigger γ the more the optimal solution will tend towards smaller weights. This is an effective way to constrain the model's complexity without influencing the dynamic properties of the reservoir.

Leaky-Integrate Neurons

A major problem of RNNs is their poor capability to model long-term dependencies between input and output, since information within the reservoir decays exponentially. In order to control the memory behavior in echo state networks *leaky-integrate neurons* have been introduced [18]. In the recurrent network each neuron has integrative behavior with an adjustable leak-rate. This is achieved by a linear transfer function succeeding the activation function as depicted in figure 2.3.

The goal is to control the memory behavior of the neurons and to tune the reservoir dynamics to the spectrum of the input signals which presumably carries the essential information for the modeling task. In the analogy introduced earlier in this chapter, where reservoirs are regarded as a liquid, one could interpret the choice of the leaky-integrate parameters as adjusting the viscosity of the reservoirs liquid.

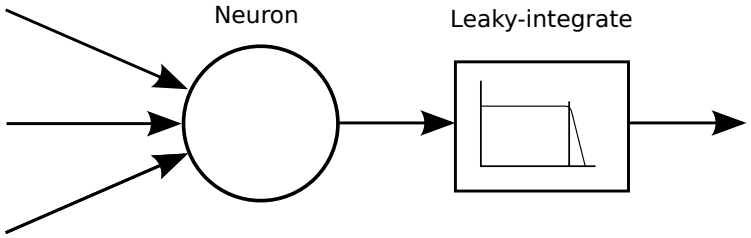

Figure 2.3: Leaky-integrate neuron. Each neuron is followed by a transfer function which typically has integrative behaviour with a specific leak-rate

By modifying (2.3) the leaky-integrate update equation is defined as:

$$\mathbf{x}[n] = \left(1 - a\frac{1}{c}\right)\mathbf{x}[n-1] + \frac{1}{c}f\left(\mathbf{W}_{in}\mathbf{u}[n] + \mathbf{W}\mathbf{x}[n-1]\right) \tag{2.11}$$

Hereby, a is the leak-rate and c a time-constant. In a special case where $a = 1$, the equation can be simplified to:

$$\mathbf{x}[n] = (1-\alpha)\,\mathbf{x}[n-1] + \alpha f\left(\mathbf{W}_{in}\mathbf{u}[n] + \mathbf{W}\mathbf{x}[n-1]\right) \tag{2.12}$$

In (2.12) the linear transfer function is a first order low-pass filter. For simplicity this form is used throughout the present work. For $0 < \alpha \leq 1$ BIBO-stability (bounded-input-bounded-output stability) can be guaranteed (for a single neuron, not the whole reservoir). Throughout the present work α will be referred to as *leak-rate*.

Before choosing a certain leak-rate one should estimate the expected time-constants of the target values. The reservoirs time-constants should approximately be in the order of the time in which the targets are expected to change significantly. After choosing a time-constant τ (measured in samples), the corresponding leak-rate follows as:

$$\alpha = 1 - e^{-\frac{1}{\tau}} \tag{2.13}$$

For example, if the desired time-constant is 5 seconds for a 100 Hz reservoir, the resulting leak-rate is approximately $\alpha = 2 \cdot 10^{-3}$.

Universal Approximator

As mentioned before, the main advantage of echo state networks compared to other recurrent neural network methods is the fact that only the output weights are learned. What seems at the first view to be a loss of generality is a clever way to circumvent most of the disadvantages that come along with training a RNN. Instead of a sophisticated iterative learning process the ESN learning problem is convex and has a closed solution.

The structure of an ESN can be generalized as is depicted in Figure 2.4. In a first step the input signals drive a non-linear, dynamic system. This system is parameterized randomly and remains unchanged during the training phase. In a second step the state variables of the first system are combined in a static, linear system in order to produce

an arbitrary number of output-signals. The second system's parameters are subject of optimization during the training phase. This structure is proved to be an universal approximator [17] for arbitrary complexity of the two systems, which means that any non-linear system can be approximated with an arbitrary small error.

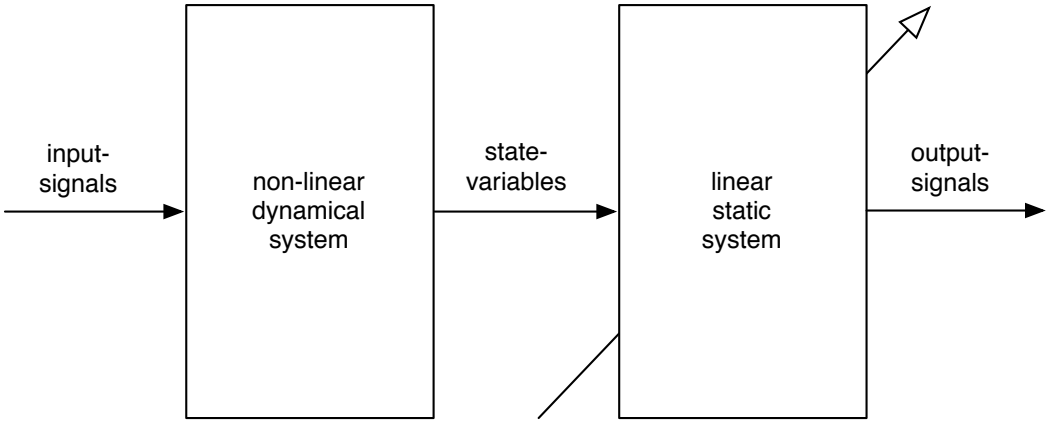

Figure 2.4: ESNs as universal approximator

Echo State Property

The reservoir as a dynamical system can become unstable, which has to be prevented. An unstable reservoir only jumps between the margins of the neurons' activation function and is not able to store additional information. Stability of non-linear dynamic systems can be shown for the class of fading-memory systems. There, the information about an arbitrary starting condition decays exponentially in a fading-memory system.

Herbert Jaeger [17] has pointed out under which circumstances an echo state network is a stable fading-memory system and called it *echo state property*. The central observation of the echo state property is the *spectral radius* (ρ) of the weight matrix $\mathbf{W}$. Essentially the spectral radius is nothing else than the biggest eigenvalue of $\mathbf{W}$. A sufficient condition for stability is:

$$|\rho| < 1 \tag{2.14}$$

This condition also holds for the simple leaky-integrate case (low-pass behavior) if the neurons' difference equation is stable. In the complex case the condition must be expanded with the leaky-integrate parameters in the following way:

$$\left|1 - \frac{1}{c}(a - \rho)\right| < 1 \tag{2.15}$$

Since only the simpler low-pass case is used in the present work the former condition (2.14) is sufficient.

Reservoir Weights

As mentioned previously the reservoir weights $\mathbf{W}$ are generated randomly beforehand following some random parameters. These parameters will be discussed in this chapter. The reservoir weights are equally distributed with zero mean, following two different parameters. The first parameter is the *connection probability* β. Experiments have shown that the modeling capabilities of ESNs can improve if the reservoir graph is not fully connected. Connection probabilities of $\beta < 0.4$ are very common, meaning that more than 60% of the connections (according to a fully connected graph) are missing. One can account for that by just randomly setting the desired amount of weights in $\mathbf{W}$ to zero.

The second parameter is the spectral radius ρ. The spectral radius specifies how the neurons' activation influence each other and is in general a parameter to control the memory behavior of the reservoir. A spectral radius close to zero means that an excitation decays rapidly in the reservoir. A spectral radius close to one means that an excitation propagates longer through the reservoir. For a spectral radius bigger than one, though useful in some cases, stability can not be guaranteed under all circumstances.

Let λ_i be the i-th eigenvalue of the weight matrix $\mathbf{W}$. The spectral radius of $\mathbf{W}$ is then defined as the biggest eigenvalue:

$$\rho = \max_i \left(|\lambda_i| \right) \tag{2.16}$$

By multiplying $\mathbf{W}$ with some factor $a \in \mathbb{R}$, the spectral radius changes by the same factor. We can use this to imprint the desired spectral radius onto the weight matrix. Let $\tilde{\mathbf{W}}$ be a weight matrix containing equally distributed weights with zero mean and random weights set to zero according to the connection probability β. Further let $\tilde{\lambda}_i$ be the i-th eigenvalue of $\tilde{\mathbf{W}}$. A weight matrix $\mathbf{W}$ with the desired spectral radius ρ can then be obtained by:

$$\mathbf{W} = \tilde{\mathbf{W}} \cdot \frac{\rho}{\max_i \left(\left| \tilde{\lambda}_i \right| \right)} \tag{2.17}$$

Each neuron also has a randomly generated bias. These bias weights are also equally distributed and of essential meaning for the reservoir. The input signals are typically normalized with zero mean. Without the bias the ESN's dynamics would have an entirely symmetric characteristic which makes it impossible to model biased output behavior. Thus one has to introduce a bias quantity of some sort during the process, either by using bias weights in the neurons or by adding a bias input signal which is mathematically the same.

Input Weights

The input weights $\mathbf{W}_{in}$ are equally distributed with zero mean and within a range referred to as *input scaling*. The input scaling directly influences how much the input signals excite the reservoir. This would be irrelevant if the reservoir was a linear system.

To see how the input scaling influences the reservoirs behavior one has to take a closer look at the non-linear activation function.

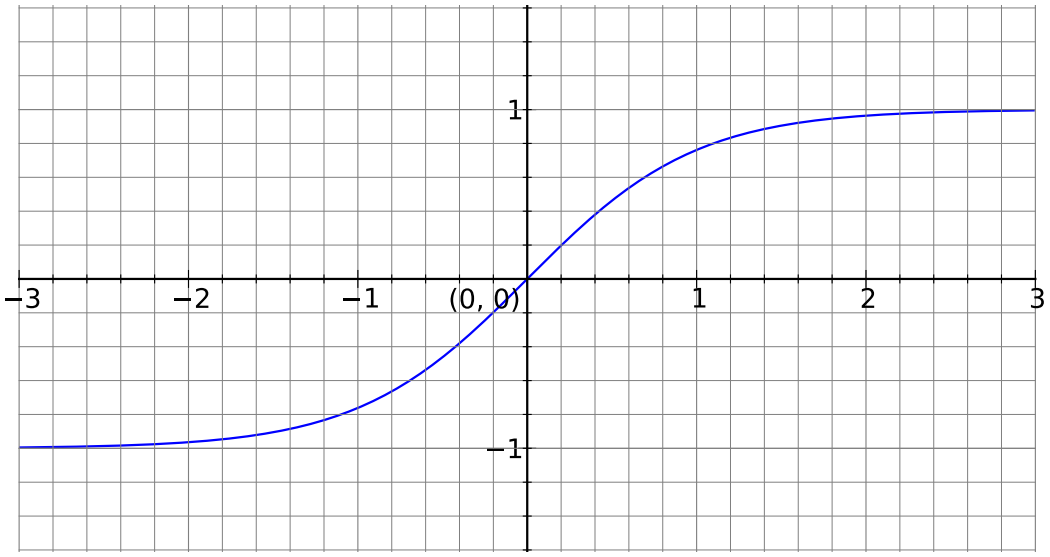

Figure 2.5: *tanh* - the activation function

The *tanh*-function as depicted in Figure 2.5 is bound between -1 and 1. If input weights are small, the excitation will be small and the neurons will operate in a small range around zero (assuming small biases). In that range the activation function is almost linear. This way the reservoir will behave mostly like a linear system. As the input weights increase the excitation will drive the neurons farther out to the marginal areas of the *tanh*-function which become increasingly non-linear. As a result the reservoir will become more of a non-linear system. In other words, the input scaling is used as a tool to adjust the non-linearity of the model.

Also the bias scaling and spectral radius are parameters that influence the neurons' operation range and therefore the non-linearity of the system. These parameters are not independent of each other and finding optimal settings for all parameters is usually non-trivial.

2.2 Data

Peripheral, arterial blood pressure (ABP) is a quantity that is typically measured at the limbs. Observation of the ABP is considered a standard procedure for intensive care unit (ICU) patients. The gold standard method to measure ABP is a pressure sensor at the end of an arterial catheter. Other, non-invasive methods like the CNAP device have emerged recently. Although less accurate, these methods are gaining acceptance for convenience reasons, especially for patients in non-serious conditions. Patient monitors

emerged in the last decade which are able to not only continuously measure ABP but also to estimate CO solely upon their ABP signals (CO-from-ABP) using various kinds of cardiovascular models.

These CO-from-ABP methods aim to partially replace the gold standards in CO measurement (e.g. thermodilution, echo cardiography) which are very difficult to apply and/or highly invasive. CO measurement of this kind is applied only on patients in critical condition or during surgery.

It is important to point out two substantial differences between the gold standard methods and the new CO-from-ABP methods:

- Indirect CO measurement methods, like echo cardiography or thermodilution, are in general more accurate than CO-from-ABP estimations but a lot more complex and/or invasive.
- Using indirect CO measurement methods, single measurements are usually triggered manually and only if needed. CO-from-ABP estimation on the other hand is a computational method and does only depend on sufficiently powerful hardware. As a result, if repeated fast enough, the estimation yields a continuous CO signal. This is an important improvement as it is feasible to automatically observe the CO signal (better trend visualization, alarm triggering, etc).

Verification of CO-from-ABP methods is performed by collecting real clinical data. When applying machine learning methods clinical data is also needed to train the model. This kind of data is usually very expensive as it requires to contract medical facilities in order to conduct medical trials, which can only be performed with the patients' consent.

I used two different medical databases in my experiments which I will introduce in this chapter. The first one is a proprietary database of the CNSystems Medizintechnik AG and is not publicly available. The second one is the "Multiparameter Intelligent Monitoring in Intensive Care" (MIMIC II) database version 3, a publicly available medical database governed by a MIT team.

2.2.1 CNAP/Vigileo Clinical Data

CNSystems Medizintechnik AG established a database during medical trials in order to verify a *pulse pressure variation* (PPV) algorithm in 2010 [32]. The database consists of records from over 60 individual patients.

The measurement setup as it was used in the medical trials is depicted in figure 2.6. It consists of 3 monitors, each recording in a different file format during the trials:

- *CNAP Monitor:* This monitor houses the CNAP sensor which measures ABP non-invasively as described in chapter 1.2.2 and 1.2.3. The monitor also records all quantities within the feedback-control mechanism of the CNAP sensor including the control deviation $v_{AC}(t)$ (see figure 1.6) and the set point correction P_n. These

two signals are believed to encode the state of the cardio-vascular system (especially vasoconstriction and vasodilation). All three signals were therefore used as *input signals* for the echo state network.

- *Vigileo/FloTrac:* The Vigileo monitor is a product of Edwards Lifesciences LLC. The FloTrac sensor is an extension to this monitor which is capable of measuring arterial pressure using an arterial catheter. The Vigileo monitor derives CO using a proprietary CO-from-ABP algorithm. The Vigileo monitor updates the CO every 20 seconds. This signal is used as targets for the ESN learning algorithm.
- *Dräger Monitor*: In this monitor the signals of both sub-monitors come together and are recorded in a single file. The only recorded signal that is coming from the CNAP monitor is ABP. Neither $v_{AC}(t)$ nor P_n are recorded which makes it impossible to work with a single recording file. As a consequence all recording files need to be aligned in a rather sophisticated process.

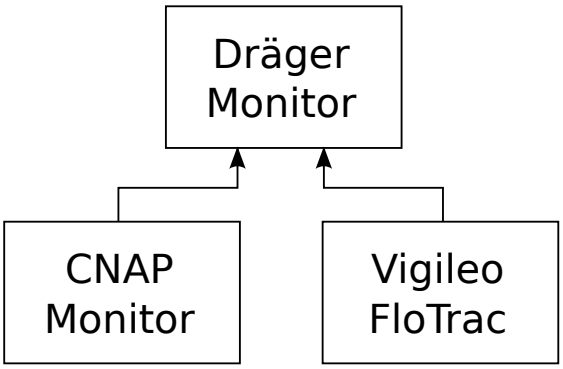

Figure 2.6: Monitor setup during the CNSystems Medizintechnik AG medical trials. The CNAP monitor measures ABP non-invasively; The Vigileo/FloTrac monitor estimates cardiac output using a CO-from-ABP method; The Dräger monitor collects ABP signals from both sub-monitors and continuous CO from the Vigileo/Flotrac monitor

Alignment

As mentioned before, parts of the ESN input signals ($v_{AC}(t)$ and P_n) are only present in the CNAP monitor recording file. The target values (CO) are only present in the Dräger monitor recording file (the Vigileo recording files are not available). Since the two monitors do not share a common time signal, the two records need to be aligned.

The signal used for alignment is the CNAP's ABP signal for it is present in both recordings. $ABP_C[n]$ denotes the discrete ABP signal from the CNAP recording whereas $ABP_D[n]$ denotes the same discrete ABP signal but from the Dräger recording file. The connections between the monitors in figure 2.6 are analog lines. That means that the two discrete signals are not identical. In fact even the sampling rate can differ slightly as the clocks of both monitors can deviate from each other. The goal now is to find the

optimal temporal offset between $ABP_C[n]$ and $ABP_D[n]$ that minimizes a predefined alignment error function:

$$e[k] = E\{(ABP_C[n] - ABP_D[n-k])^2\} \tag{2.18}$$

$$= \frac{1}{N-1}\sum_{n=0}^{N-1}(ABP_C[n] - ABP_D[n-k])^2 \tag{2.19}$$

Another intuitive decision for the alignment error would be to use the inverse of the normalized cross-correlation function. This choice has proved to lack robustness for two reasons:

- The ABP is a highly periodic function. Therefore the normalized cross-correlation function is periodic as well.
- The normalized cross-correlation function has shown to degrade rapidly in border regions. This is especially the case if the true offset is large and therefore the overlap regions of the two signals relatively small.

The expected quadratic error has shown to be the most robust choice.

CNAP Calibration

The set point tracking (chapter 1.2.3) is one of the CNAP's key features. Nevertheless the tracking cannot be kept up for unlimited time. Over tens of minutes the parameter drift can be too big to be compensated by the tracking mechanism. For this reason the CNAP monitor performs a calibration every 30 minutes which lasts for approximately 2 minutes. During this time the ABP signal is not available as shown in figure 2.7. The 30 minutes between two calibration phases are called a *stint*.

After a calibration the ABP signal tends to jump if a parameter drift occurred. These jumps are particularly difficult to handle since they are usually followed by jumps in the CO estimation. Erroneous parameter drifts are a phenomenon solely caused by the CNAP sensor and are not based on any physiological effects. Still they are yet unavoidable and worsen the performance of the CO estimation.

Not every record in the CNAP medical trials database is usable training/test data. Due to various reasons several measurements were excluded. The most common reasons are:

- The recording file itself can be corrupt or erroneous. In some cases the monitor ran into some error condition and failed to record signals without interruptions or didn't record any data at all.
- In some cases the ABP signal showed serious artifacts. This can happen due to bad handling of the monitor equipment (e.g. displaced sensor). Also the patient's movement can cause the equipment to malfunction.

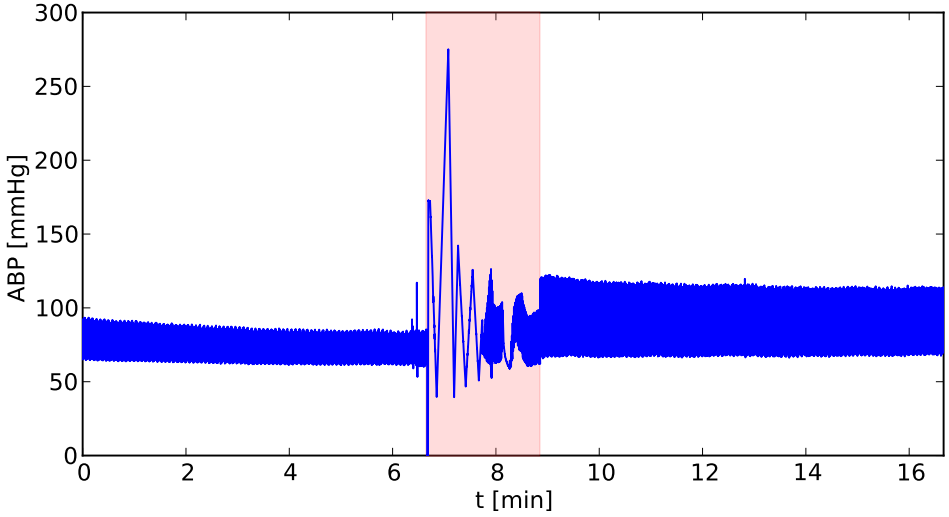

Figure 2.7: Marked in light red: The CNAP's calibration phase between two stints

- The Vigileo monitor itself relies on a clean ABP signal without artifacts. As the Vigileo monitor measures ABP using an arterial catheter, artifacts often occur if the patient lies on the catheter or it congests in some way.

All recordings were manually examined and evaluated after alignment. If a recording showed any of the above signs the corresponding stint or the whole patient was excluded. Finally 40 patients remained with an overall recording time of over 76 hours.

2.2.2 MIMIC II Medical Database

The second database used in the present work is the *Multiparameter Intelligent Monitoring in Intensive Care II* (MIMIC-II) database [19]. It is a publicly available database maintained by the *Massachusetts Institute of Technology* (MIT) and several other medical institutes all over the world. It contains patient records of over 25 thousand hospital stays. Most of the records include a variety of [continuous] life sign measurements, including arterial blood pressure and cardiac output.

Just recently the MIT released version 3 of the MIMIC II database increasing the amount of records from a couple of thousands to over 25 thousand. A lot of the records contain ABP recordings but only very few contain CO recordings. Only records that contain both signals are viable as training data. This is the case in more than 350 records (before validation) at the time of writing. With this amount of data, applying machine-learning methods on this estimation problem seems feasible.

The MIMIC II database consists of two parts:

- *MIMIC II DB:* Contains meta-data, doctors/nurse notes, discharge reports, etc.

- *MIMIC II waveform DB:* Contains high-resolution waveform records of various life signs like ECG, blood pressure, pulse oximetry, etc. This second part of the database is relevant for this work.

WFDB Format

The waveform database is stored in the WFDB format, which was first introduced in the early 1990ies. Its intention was to offer a simple format for the purpose of recording waveforms over a long time-range, with possible intermissions and recording setup changes. For simplicity the database is organized in files and folders and does not incorporate any relational database functionality in its format.

Multiple signals are stored in simple binary files. A record can consist of multiple of these binary files. Header files are used to keep meta-information about the binary files (sampling rates, signal types, bit-resolutions, etc.) and assign to each binary file a starting point in the record. All records are available on a public HTTP server hosted by PhysioNet[1].

PhysioNet also offers a C-library and a software package in order to access databases stored in the WFDB format. Wrappers for the C-library are available for MATLAB, Python and several other programming languages using the SWIG wrapper generator[2].

Measurement Methods

It is impossible to tell for certain which method, out of a number of methods, was applied in a specific record. The ABP signals were recorded invasively using an arterial catheter. The methods used to measure cardiac output are echo cardiography (see chapter 1.3.1) and thermodilution (see chapter 1.3.2). Both methods are highly sophisticated and complicated to apply. Additionally they cannot provide continuous measurements. This is maybe the most important difference to the CNAP/Vigileo database. Instead of having a CO value every 20 seconds, CO measurements are usually triggered manually and only if needed. In fact the interval between two CO measurements can be hours. Accordingly, the average record is over 40 hours long (125 Hz sample-rate) and contains less than 10 CO measurements.

ESN Burn-in Time

Considering the long intervals between CO measurements it would be unwise to simulate the ESN reservoir for the whole length of the records since we would waste a lot of unnecessary computation capacity. As pointed out in chapter 2.1, information about the initial condition of an ESN decays exponentially if it has the *echo state property*. The time after which the information about the initial condition has decayed to a degree

[1] http://www.physionet.org/physiobank/database/mimic2wdb/
[2] http://www.swig.org/

where it has become noneffective is referred to as the *burn-in time* (τ_b). Hence, If a CO measurement happened at time t_{CO} it is only necessary to simulate the reservoir beginning at the time $t_{CO} - \tau_b$.

2.3 ESN Configurations

A potential CO-from-ABP algorithm has to be evaluated with respect to two major categories:

- *Accuracy:* A potential model to estimate CO must be robust and accurate enough to be able to replace previous methods, though it does not necessarily need to be more accurate than the previous method. Application of the new method can be justified also if it is substantially easier to apply, involves less risk to the patient or improves other aspects of the measurement (e.g. interval between two successive measurements).
- *Resources:* Approval of a new medical device is a very time-consuming and costly process. Until a device can be applied in daily ICU routine the hardware design is in most cases already a few years old. Also the same hardware setup should be used as long as possible since any change must be followed by a new approval. For ICU medical devices it is therefore very important to keep the memory and computation requirements for new algorithms as low as possible. The number of addable features on a device is often very limited.

As echo state networks are numerically rather costly simulations, the *resource* argument is definitely an issue. In this chapter I will introduce different approaches in terms of signals and topology in order to trade between accuracy and resources. All of these approaches will later be evaluated using the two databases.

2.3.1 Continuous Signals

The simplest straight forward approach is to use the continuous signals directly. All used signals are therefore fed into the reservoir after a normalization step. The topology is depicted in figure 2.8.

CNAP/FloTrac

In case of the CNAP/FloTrac database the input contains the following signals:

- Arterial blood pressure ($ABP[n]$)
- CNAP control-deviation ($v_{AC}[n]$)
- Set point correction ($P_{\tilde{n}}[n]$)

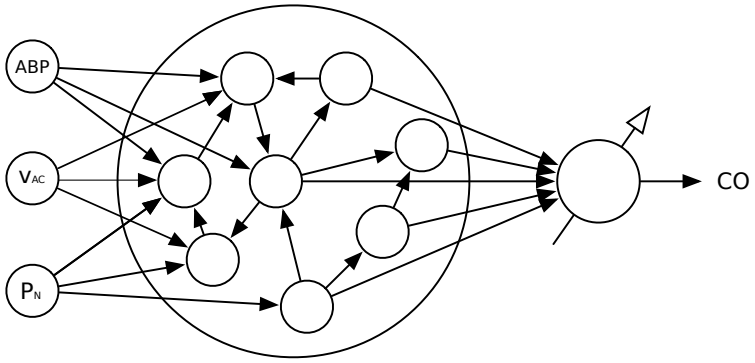

Figure 2.8: Schematics of an Echo State Network using the sampled signals as input directly

All of these signals are sampled with 100 Hz. The *control-deviation* and the *set point correction* are believed to carry information about the total peripheral resistance as described in chapter 1.2.1.

Normalization is done using a running average and variance for any signal $s[n]$ as follows:

$$\mu_s[n] = \frac{1}{K} \sum_{k=n-K}^{n} s[n-k] \tag{2.20}$$

$$\sigma_s^2[n] = \frac{1}{K-1} \sum_{k=n-K}^{n} (s[n-k] - \mu_s[n-k])^2 \tag{2.21}$$

The signals $\mu_s[n]$ and $\sigma_s[n]$ are then used to create an input signal with zero-mean and variance one:

$$\bar{s}[n] = \frac{s[n] - \mu_s[n]}{\sigma_s[n]} \tag{2.22}$$

The running normalization is an online normalization which is a requirement for an ICU monitor. The only exception is the arterial blood pressure. If a running average was applied here it would destroy the information about the mean arterial pressure (MAP). Therefore ABP is normalized statically using mean and variance obtained from all CNAP recordings in the database combined. The combined mean pressure is 88 mmHg, the combined standard deviation is 20 mmHg.

In another experiment the ABP is also normalized using a running normalization. The MAP is calculated as a continuous signal beforehand, statically normalized using the same empirical mean/variance and fed into the reservoir as a separate signal. Mathematically similar the only difference is that the MAP and the ABP signals have a different set of input weights. It turns out that this approach is slightly more accurate.

MIMIC II

The MIMIC II waveform database lacks the control signals of the CNAP's feedback-control loops. The only available signal is the invasive ABP with a sample rate of 125 Hz. Also the ABP is normalized using running normalization and a separate MAP signal carries information about the mean pressure.

2.3.2 Beat-to-Beat Features

Iterating a reservoir with a few hundred neurons at a sample rate around 100 Hz is costly in terms of computation. Such a simulation is hard to implement on the limited hardware of an ICU monitor. It may be necessary to develop a computationally *cheaper* model in order to fit it in a monitor's hardware.

Arterial blood pressure is a periodical signal. A series of pulse waves can be represented as coefficients of a Fourier series (or the DFT of a sampled pulse wave respectively). A pulse wave can also be statistically described using moments. These descriptive quantities (among others) can be interpreted as *features* of a pulse wave. By limiting the number of features one can obtain a reduced representation of a pulse wave.

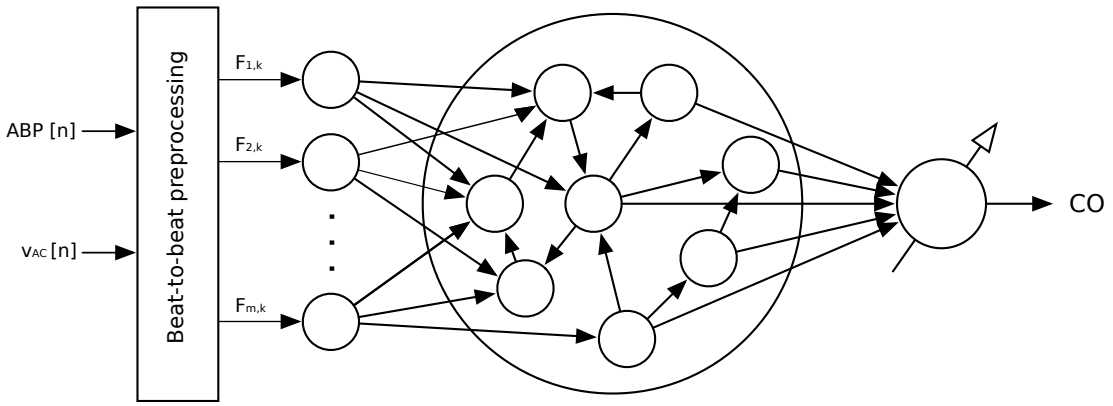

Figure 2.9: Schematics of an Echo State Network using beat-to-beat features as input. $F_{m,k}$ is the m-th feature of the k-th beat

Figure 2.9 shows the schematics of this approach. A pre-processing unit detects beats in the $ABP[n]$ signal and extracts beat-to-beat features $F_{m,k}$ (m-th feature of the k-th beat). Features are extracted only from signals that are beat-periodic. These signals include, among others, the arterial pressure and the CNAP control-deviation.

Assuming $s[n]$ being the samples of the k-th pulse with the length N_k of a pulse-periodic signal. The following features were considered in this approach:

- Central moments: $m_l = E\{(s[n] - E\{s[n]\})^l\}$ (l-th central moment)
- Spectral components: $S_l = \sum_{n=0}^{N} s[n] \cdot e^{2\pi j \frac{l \cdot n}{N}}$ (l-th frequency component)

- Heart rate: $HR_k = \frac{\text{samplerate}}{N_k}$
- Diastolic decay τ_1, τ_2 (see chapter 1.3.3)
- Pulse pressure: $PP_k = P_{sys,k} - P_{dia,k}$ (see figure 2.10)

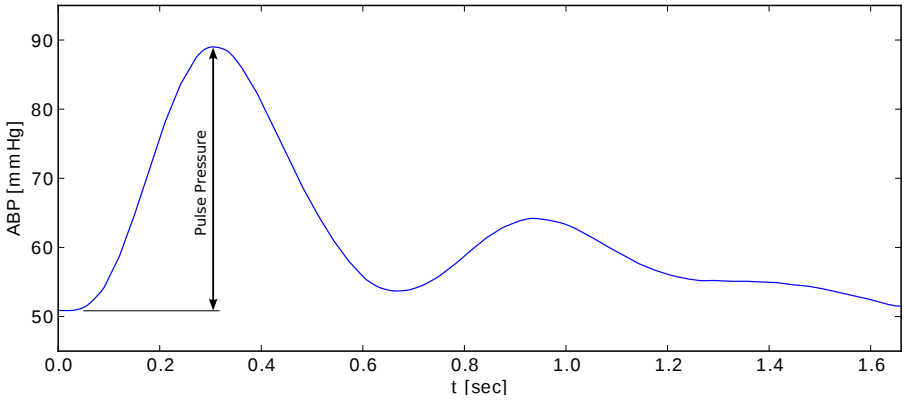

Figure 2.10: Pulse pressure is the difference between systolic pressure (high) and diastolic pressure (low)

All features are renewed on a beat-to-beat bases and fed into the reservoir as temporal input sequence. Compared to the direct approach the input vector has a lot more dimensions. The reservoir is iterated in sync with the heart rate. That means iteration frequencies around 1 Hz instead of 100 Hz like in the direct approach. As a result the frequency tuning of the reservoir has to be adjusted to match the dynamics of cardiac output.

This approach is of course a loss of generality and is expected to perform worse than the direct approach, yet it is a lot more efficient in terms of resources. Thus it is an interesting question how the performance of these two approaches compare.

The signal feature approach was tested on the CNAP/FloTrac database.

2.3.3 Multi-Reservoir

Early CO-from-ABP methods did not make use of extensive pulse contour analysis. Many approaches suggested that CO is proportional to mean arterial pressure or pulse pressure [8, 20]. The most common method to find the proportionality factor was calibration using more sophisticated measurement methods once. It is since well known that a deeper analysis of the pulse pressure waveforms is necessary to understand and model the underlying dynamics. Thus, modern approaches always rely on thorough pulse contour analysis.

Using leaky-integrate neurons in an ESN-based model gives one the possibility to control the memory behavior of the reservoir and tune the reservoir's dynamics to a certain frequency band in the signals. As CO is an averaged quantity it changes significantly on a time scale between seconds and minutes. That means relatively large time-constants

in the leaky-integrate neurons (up to 1 minute). This can lead to a poor impact of high-frequency components of the pulse contour in the output. Many approaches using pulse contour analysis point out the importance of these high-frequency components. It seems therefore sensible to extend the ESN's reservoir with additional neurons that represent the high-frequency components in the input signals.

To account for that a feasible and simple solution is to use two reservoirs, a slow one and a fast one as depicted in figure 2.11. While the slow reservoir is tuned to the dynamics of CO, the fast reservoir is tuned to the dynamics of pulse waves with time-constants around one second.

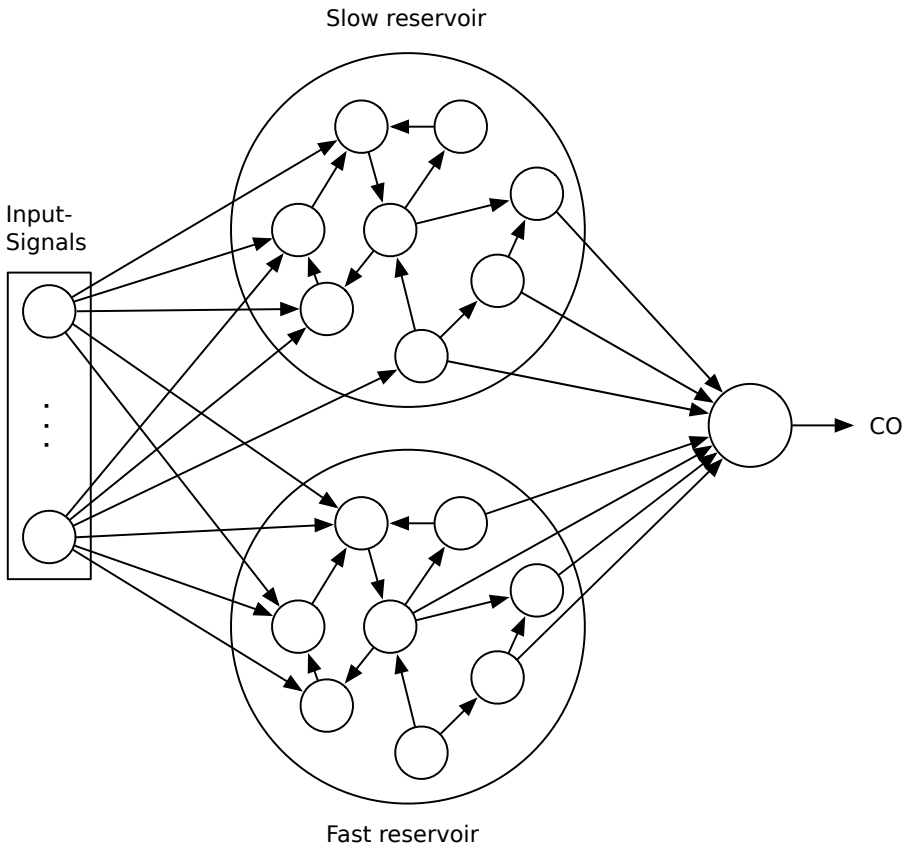

Figure 2.11: Schematics of an Echo State Network using two reservoirs tuned to different frequency bands in the signals

Since the dynamics of the fast reservoir contain components of higher frequencies it has pulsatile characteristics as well. Information stored in the reservoir decays within very view pulses. A different interpretation is that the reservoir accumulates information of a single beat and arrives in a final state at the end of the diastole. That state represents the condition of the circulatory system. That means that during the time of the stroke the state information is irrelevant. Only the state at the end of a beat contains useful information. This needs to be accounted for when the training data is collected.

A single training sample for the ESN consists of a CO target value (measured with a

reference method) and the reservoir's state at the exact point in time when the CO measurement took place. That point in time is arbitrary in case of the slow reservoir. In case of the fast reservoir the time of the CO measurement must be aligned with the end of the prior diastole. Therefore a beat-detection algorithm is applied to find the onsets in the sampled pulse wave $n_{sys,k}$. These onsets are the beginning of the k-th systole and the end of the preceding diastole.

Artifacts in the pulse waves can lead to a degeneration of the fast reservoir's state information. To prevent this kind of outliers the reservoir's state vector is averaged over the last 10 beats. Let $\mathbf{x}_m$ be the fast reservoir's state vector corresponding to the m-th training sample and let the k-th beat be the last beat preceding the m-th measurement event. $\mathbf{x}_m$ is then calculated as:

$$\mathbf{x}_m = \frac{1}{10} \sum_{i=0}^{9} x[n_{sys,k-i}] \tag{2.23}$$

2.3.4 Descriptive Data

Additionally to life sign waveforms the CNAP/Vigileo database contains descriptive information. Specifically this information includes weight, height, sex and age of the patients. It is well known that descriptive patient data improves the assessment of the patient's condition in various aspects. A few examples:

- The compliance of the arterial walls decreases with age. As the blood vessels get stiffer the mean arterial pressure increases causing hypertension.
- Men have substantially different physiological characteristics than woman.
- The larger a person is the more cardiac output is required to sufficiently supply the body.

A simple tool to quickly estimate the importance of a feature for a machine-learning task is to look at the correlation coefficient:

$$r_{x,y} = \frac{cov(x,y)}{\sigma_x \sigma_y} \tag{2.24}$$

cov is the co-variance and σ is the standard deviation of a signal, x and y are arbitrary signals. $r_{x,y}$ is a value between -1 and 1. If $r_{x,y}$ is 0 the signals are not correlated. The more the signals are correlated the closer $|r_{x,y}|$ is to 1. The sign of $r_{x,y}$ is therefore not important.

According to the CNAP/FloTrac database the patients' descriptive data is correlated with their mean cardiac output:

	age	height	sex	weight
$r_{\overline{CO},x}$	-0.59	0.46	-0.33	0.24

Sex was binary-encoded $\{-1, 1\}$.

Obviously descriptive patient data is a viable feature, at least to reduce bias in the CO estimation error. Two different approaches are made to integrate descriptive patient data into the model. In a first approach the descriptive quantities are used as additional constant input signals. Therefore they act as an additional bias to the reservoir neurons and are integrated into the non-linear component of the model. In a second approach the descriptive quantities act as *virtual neurons* with a constant state as depicted in figure 2.12. Therefore they are a linear component in the result and can only influence the bias of the estimation.

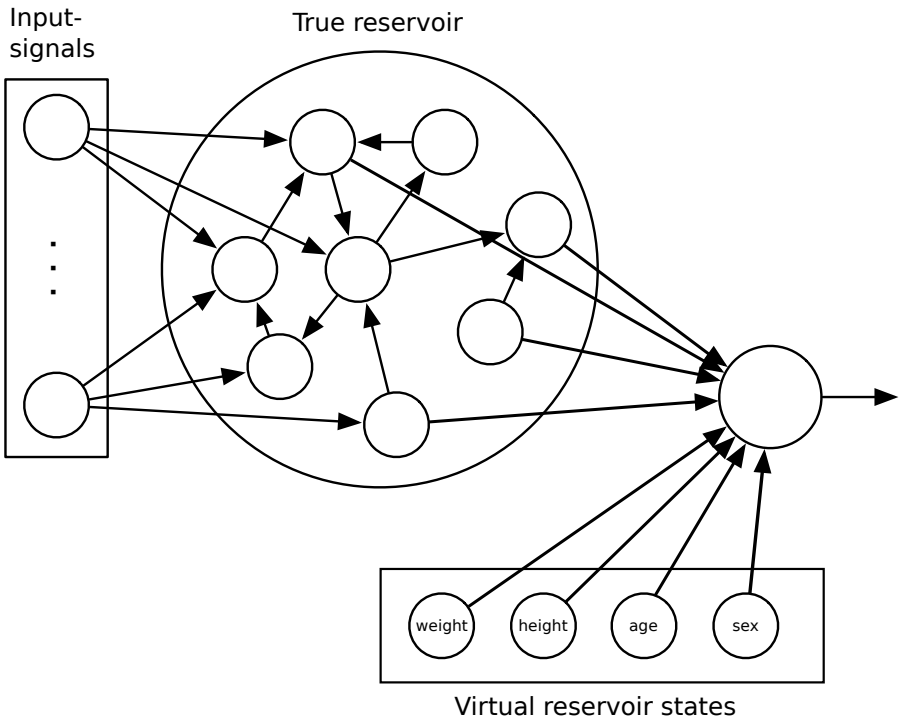

Figure 2.12: Echo state network with patient descriptive constants as virtual reservoir states.

As descriptive data is not present in the MIMIC II database version 3 at the time of writing this approach was only tested on the CNAP/FloTrac database. As soon as this situation changes this approach is going to be tested against the MIMIC II database.

2.4 Error Measurement

It is very difficult to choose sensible quality criteria in medical care. Unlike other fields of machine-learning applications the most important quality of a life sign measurement is in how far it can improve the assessment of the patients' condition and the decision-making process. Depending on the type of examination different quality criteria can apply on one and the same life sign recording.

Several error metrics were employed in order to assess the performance of new methods. Since there is no method to measure cardiac output in the human body directly evalu-

ation can only be a comparison to other indirect measurement or estimation methods. Some of these methods are continuous (Vigileo/FloTrac) others are not (thermodilution). In that two cases the same error measure can have a different meaning.

2.4.1 RMSE

The *root mean square error* (RMSE) should be almost self-explanatory. As a solid standard in the field of machine-learning it is also included in the evaluation:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{n=0}^{N-1}(t[n] - y[n])^2} \tag{2.25}$$

Whereby $t[n]$ are the target samples of the reference method, $y[n]$ are the output samples of the ESN.

$RMSE$ is not very common for quality assessment of life sign waveforms as it is a single value that does not differentiate between different types of deviations (bias, shape, scaling, etc).

2.4.2 Critchley Criterion

The Critchley criterion [3] is a quality estimator that emerged in 1989. It was supposed to be a standard criterion for comparing different life sign measuring methods among each other. It was later adapted for evaluation of CO measurement and ABP-from-CO methods.

The core of the criterion is a *percentage error* (PE) that is based on the assumption that two methods which are to be compared produce statistically independent measurements. It is then stated that the deviation between the two methods is defined by so-called *limits of agreement* which are supposed to enclose 95% of all deviations. Assuming the deviation between the two methods follows a normal distribution this can be simplified to twice the *standard deviation* of the error. The PE is calculated by normalizing that value with the mean of the reference method:

$$\mathrm{PE} = \frac{2 \cdot \mathrm{SD}(e[n])}{\mathrm{mean}(t[n])} \tag{2.26}$$

SD is the standard deviation, $t[n]$ are the samples of the reference method and $e[n]$ is the error between the two methods $(t[n] - y[n])$.

As mentioned before a reference method can also only be an indirect measurement. To make that clear: *There is no medically acceptable way to measure cardiac output directly.* Therefore the reference method itself naturally has a notable deviation from the real cardiac output. An estimate of the deviation in form of a PE is known for most methods. Thermodilution for instance has a known percentage error between 10% and 20%. By comparing an old method with a new one it is only possible to estimate

the PE between the two methods. As the errors of different methods are assumed to be statistically independent the estimated PE of the new method can be calculated by building the geometric sum:

$$\mathrm{PE}_{new} = \sqrt{\mathrm{PE}_{old}^2 + \mathrm{PE}_{diff}^2} \tag{2.27}$$

PE_{new} is the PE of the new method. PE_{old} is the PE of the reference method. PE_{diff} is the PE between the two methods. Figure 2.13 shows how the estimated PE of the new method changes with the PE of the reference method.

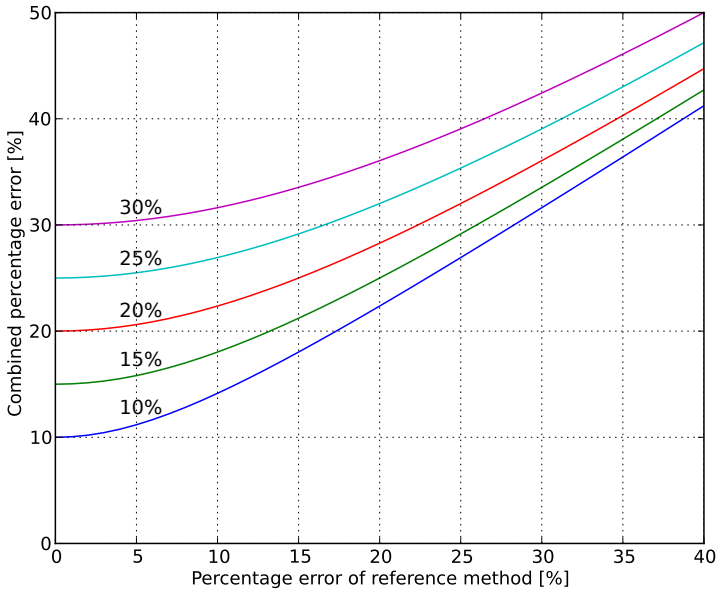

Figure 2.13: Estimation of the percentage error of a new method. Each graph represents a certain percentage error determined between the new and the old method. The characteristics of the graphs show the estimated percentage error of the new method plotted against the percentage error of the reference method.

Critchley states that, generally speaking, if a new method has a percentage error below 30% it is good enough to replace the old method if it has certain advantages.

The PE is a value that is entirely based on the error's standard deviation. Therefore the bias of the error does not influence the PE in any way. In other words PE is a quantity that assesses the quality of the signal's shape and scaling of the new method, and describes properties like tracking ability in the continuous case or ability to follow a trend in the non-continuous case. To account for the *bias* of the new method it is usually specified separately.

Critchley also describes the concept of *precision*. It is, in case of non-continuous measurements, the ability to hit the same target when repeatedly performing measurements. Even though a method can have a very small error when averaging over several successive measurements, a single measurement can be very imprecise. Since the ESN yields continuous measurements where the samples are not statistically independent from each other, the quantity of precision would not make sense.

2.5 Beat Detection

For the most models, especially the beat-to-beat features approach, all beat-periodic signals needs to be split into beats. In order to do that a beat-detection algorithm needs to be implemented. A very easy and effective algorithm to detect beats in ABP signals was introduced by Zong *et al.* [35]. It uses a *slope-sum-function* that enforces positive slopes in the blood pressure signal in order to detect systoles.

The diagram in figure 2.14 describes the design of the algorithm with its basic components. In a first step high frequency components in the ABP signal are eliminated using a second order low-pass filter with a cut-off frequency of 16 Hz.

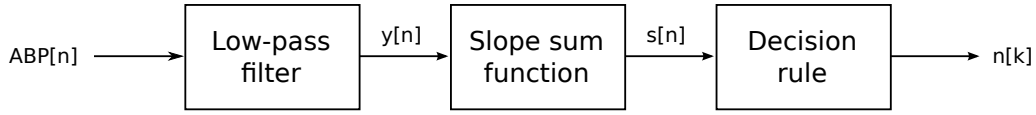

Figure 2.14: Information diagram of the beat-detection algorithm

The second stage consists of calculating a slope-sum-function (SSF). It simply sums up all positive slopes within a given window size w:

$$s[n] = \sum_{i=n-w}^{n} \Delta y[i], \qquad \Delta y[i] = \begin{cases} y[i] - y[i-1] & : \; y[i] - y[i-1] > 0 \\ 0 & : \; y[i] - y[i-1] \leq 0 \end{cases} \tag{2.28}$$

The SSF can be interpreted as "leaky integration" of positive slopes with the goal to emphasize these slopes.

The last stage is a decision making rule which executes algorithm 1 in order to detect beat-onsets. The algorithm is basically an adaptive thresholding with some additional constraints.

Depicted in figure 2.15 is a short clip of an ABP signal showing the detected onsets and the corresponding slope-sum-function.

A variety of different disorders can confuse the algorithm. Examples are artifacts, double systoles or alternating pulse pressures. Nevertheless it has been shown [15] that this simple algorithm correctly detects more than 99% of all beats from the MIT-BIH Polysomnographic Database with an insignificant amount of false positives. This algorithm was applied on the CNAP/FloTrac database. The performance was evaluated visually on random samples from the database and found sufficiently accurate for the task at hand.

2.6 Feature Selection

As described in chapter 2.3.2 one of the proposed approaches is to extract beat-to-beat features from beat-periodic signals like ABP or the *CNAP control deviation* in order

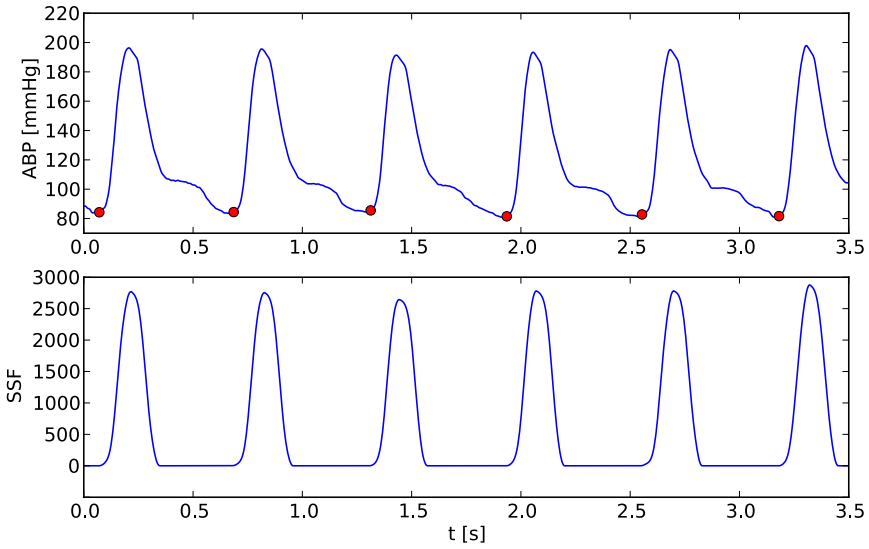

Figure 2.15: Example of a sum-slope-function. The red points are detected onsets in the ABP signal

Algorithm 1 Beat detection decision rule

$threshold \leftarrow 1.8 \cdot \text{mean}(s[n])$
$last \leftarrow -\infty$
for $n = 0 \rightarrow N$ **do**
    if $n - last > 75$ **then**
        if $s[n]$ just passed $threshold$ **then**
            $s_{max} \leftarrow$ succeeding maximum in $s[n]$
            $s_{min} \leftarrow$ preceeding minimum in $s[n]$
            if $s_{max} - s_{min} > sensitivity$ **then**
                $n_{onset} \leftarrow$ Search for beat onset
                Store n_{onset} in list of beats
                $threshold \leftarrow 0.6 \cdot s_{max}$
                $last \leftarrow n_{onset}$
            end if
        end if
    end if
end for

to reduce the computational complexity of the model. At this point there is a variety of potential features but yet no information about the relevance of each feature for the machine-learning task. Therefore a feature selection procedure must be applied to assess each feature's relevance.

In the present task we face a number of 20 features extracted from 2 beat-periodic signals ($ABP[n]$ and $v_{AC}[n]$ (CNAP control deviation)):

- First 4 central moments of both signals
- First 4 frequency components (DFT) of both signals
- Diastolic decay (2 time constants)
- Heart rate
- Pulse pressure (see figure 2.10)

The importance of each feature is unknown a priori. If important features are left out the model may be inaccurate. If unimportant features are included the model may generalize too much. Some combinations of features may perform better than others. In an optimal assessment one would evaluate all possible combinations of features. At this point one stumbles upon a common problem which is called the *curse of dimensionality*. The number of possible combination grows exponentially with the amount of features. For 20 features that would mean 2^{20} (1048576) possible combinations. Thus, one needs a more efficient way to search for optimums in the feature space.

This states a classical optimization problem in a discrete parameter space. What we seek to optimize is a predefined error function that is defined over the whole parameter space. In the present case the input to the optimization problem is the feature subset. The *RMSE* was chosen for the error function as it respects both shape and bias of the error. Now we seek the optimal feature subset which minimizes the error function. A feature subset is evaluated by performing a *leave-one-out cross-validation* with the model over all patients. This means that a new model is trained for each patient. The model is trained on all remaining patients and evaluated on the left out patient. This procedure is repeated for all patients and the error is averaged. This is a practical way to estimate the true error of a trained model.

Several standard methods are well known for these kinds of optimization problems including genetic algorithms, particle swarm optimization or gradient methods. In case of the present problem a both simple and common optimization method is a *sequential forward/backward feature selection* (SFBS). It is computationally inexpensive compared to other methods which is important for the present task as the evaluation of the model is relatively costly due to the cross-validation.

The principle of the SFBS is demonstrated in algorithm 2. Starting out with no features the process of adding features is divided into sequences. Each sequence consists of an adding phase and a removal phase.

In the adding phase n features are added to the set of used features sequentially. In each of the n steps each unused feature is evaluated together with the set of used features added so far. The best unused feature is then added to the set of used features.

The removal phase works similar. m features are to be removed from the set of used features. In each of the m steps the set of used features is evaluated several times, always with one of the features excluded. The feature which leads to the best performance upon its exclusion is then removed from the feature set.

So in each sequence n features are added to the set and afterwards m features are removed. $(n - m)$ features are added to the set of used features in each sequence. For the set to grow $n > m$ must hold.

Algorithm 2 Sequential forward feature selection

1: $n \leftarrow 3$
2: $m \leftarrow 2$
3: $features \leftarrow \{\ \}$
4: $unused \leftarrow \{\text{all features}\}$
5: $K \leftarrow |unused|$
6: **repeat**
7: **for** $1 \rightarrow n$ **do**
8: $best\_feature \leftarrow \underset{f \in unused}{\text{argmin}}\ \text{error}(features \cup \{f\})$
9: $features \leftarrow features\ \cup\ \{best\_feature\}$
10: $unused \leftarrow unused \setminus \{best\_feature\}$
11: **if** $|features| == K$ **then**
12: break
13: **end if**
14: **end for**
15: **for** $1 \rightarrow m$ **do**
16: $worst\_feature \leftarrow \underset{f \in features}{\text{argmin}}\ \text{error}(features \setminus \{f\})$
17: $features \leftarrow features \setminus \{worst\_feature\}$
18: $unused \leftarrow unused\ \cup\ \{worst\_feature\}$
19: **end for**
20: **until** some break condition

The smaller $(n - m)$ is the more sequences have to be tried. $(m + n)$ evaluations have to be made in each sequence. $n = 2$ and $m = 1$ were found to be a reasonable choice in the trade-off between computation time and thoroughness of the feature search.

3 Results

In this chapter all results are displayed and compared. The plots throughout the chapter are the results of leave-one-out cross-validations. In each turn one patient is left out during training and then tested against. This is repeated over multiple trials and the outcomes are averaged, which is an accepted way to estimate the true error of a learned model. The number of trials is between three and five depending on the computational complexity of the problem.

It is not helpful to present results for every single patient. Instead, in case of the CNAP/FloTrac database 10 meaningful patients are picked in order to compare the different approaches and explain the effects in the characteristics. In case of the MIMIC II database it amounts to 16 patients. Further more, several quality measures introduced in chapter 2.4 are applied and compared among all approaches.

Bland-Altman (BA) plots [24] are presented along with plots of the characteristics of CO and ABP. BA-plots provide information about the distribution of the error depending on its magnitude.

3.1 CNAP/Vigileo

3.1.1 CNAP Continuous Signals (Experiments 1 & 2)

The signals from the CNAP/Vigileo database (ABP, v_{AC}, $P_{\tilde{n}}[n]$) were taken directly at a sample rate of 100 Hz with minimum preprocessing. All input signals were normalized and the mean arterial pressure (MAP) was added as a separate signal as discussed in chapter 2.3.1. The results in this chapter includes two experiments as mentioned before. The only difference between the first and the second experiment is that the latter makes use of the descriptive data (DD) incorporating them as virtual neurons in the reservoir state.

What follows are plots of the 10 selected patients like the one in figure 3.1. Each page represents one patient and consists of four plots depicting the results of both experiments:

1. The first plot shows cardiac output graphs. The blue graph represent the targets as they were measured using the Vigileo/FloTrac device by Edwards Lifesciences LLC. Targets are emitted every 20 seconds by the Vigileo device. The green graph is the output of the ESN without descriptive data (first experiment). The red graph is the output of the ESN including descriptive data.

2. The second plot shows the underlying blood pressure signal as taken by the CNAP monitor. The duration of a single recording is usually at least one hour. Since a heart beat takes typically around one second the ABP signal appears heavily compressed in the plot, single beats cannot be distinguished. The purpose of this plot is rather to get an idea of systolic and diastolic pressure trends as well as mean arterial pressure. One can see the calibration phases of the CNAP monitor between the *stints*. They take place where the ABP signal jumps to zero for a moment.
3. The third plot is a BA-diagram for the first experiment showing the distribution of the error against it's absolute value. BA-diagrams are useful to determine how well different ranges of CO are covered by the model. Also one can easily discover linear dependencies between the error and the absolute CO value.
4. The last plot is a BA-diagram of the second experiment.

The plots can provide the following observations:

- How well the models perform in terms of tracking ability (e.g. the ability to follow trends over a long-time period).
- The bias of the error and how it is different in the two experiments.
- The reflection of pulse pressure and mean arterial pressure in the cardiac output.
- The dependency between error distribution and absolute error in the BA-diagrams.

In the upcoming two experiments in which a feature representation of the beat-periodic signals is used as reservoir input the same 10 patients are depicted for comparison.

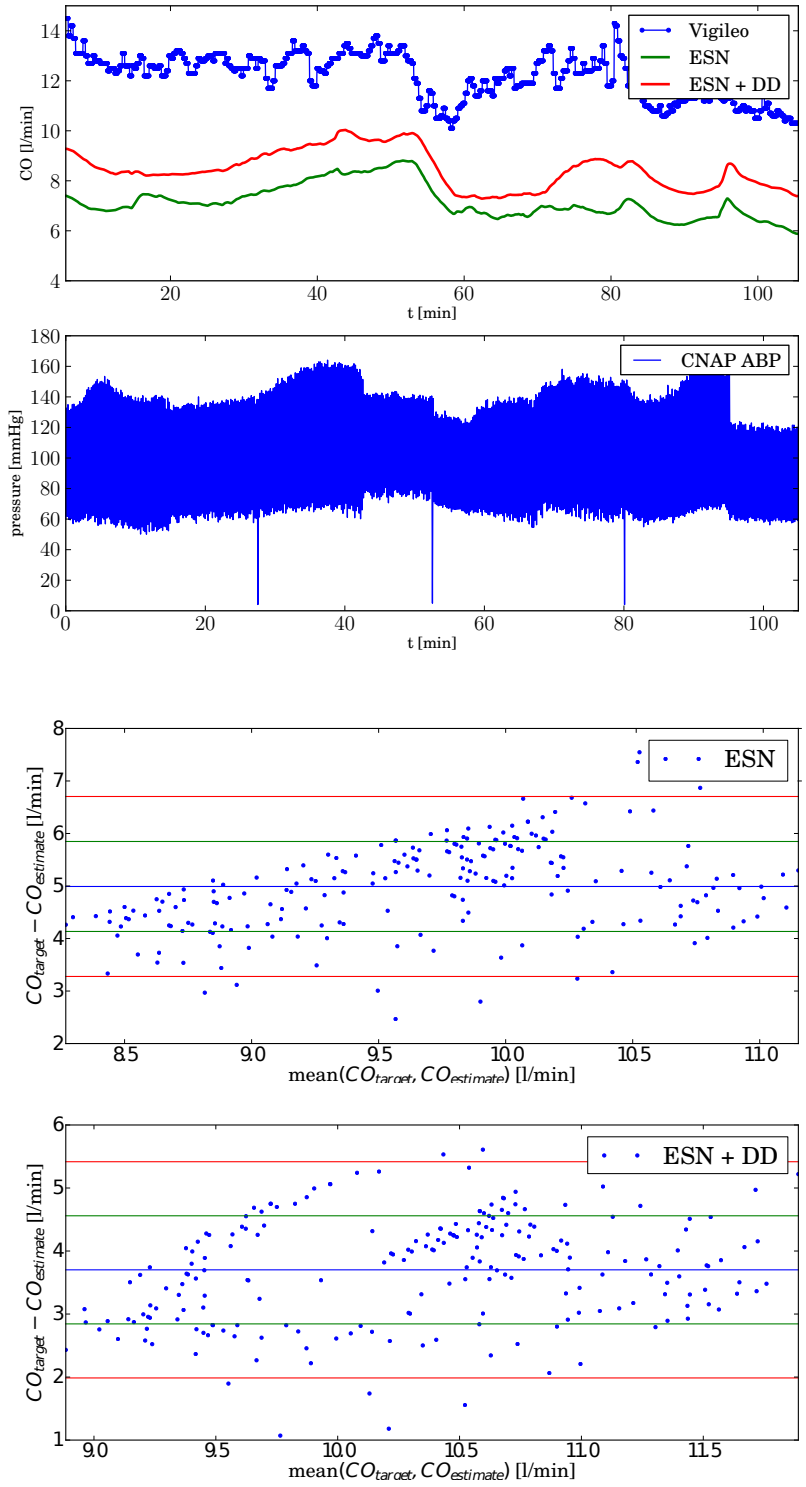

Figure 3.1: Patient 07: The first plot shows the targets (Vigileo/FloTrac) in blue. The green graph is the output of the ESN in the first experiment (without DD). The red graph is the output of the second experiment (including DD as virtual neurons). The second plot shows the corresponding ABP signal recorded by the CNAP sensor heavily compressed. Third and fourth plot show Blant-Altman-diagrams for both experiments.

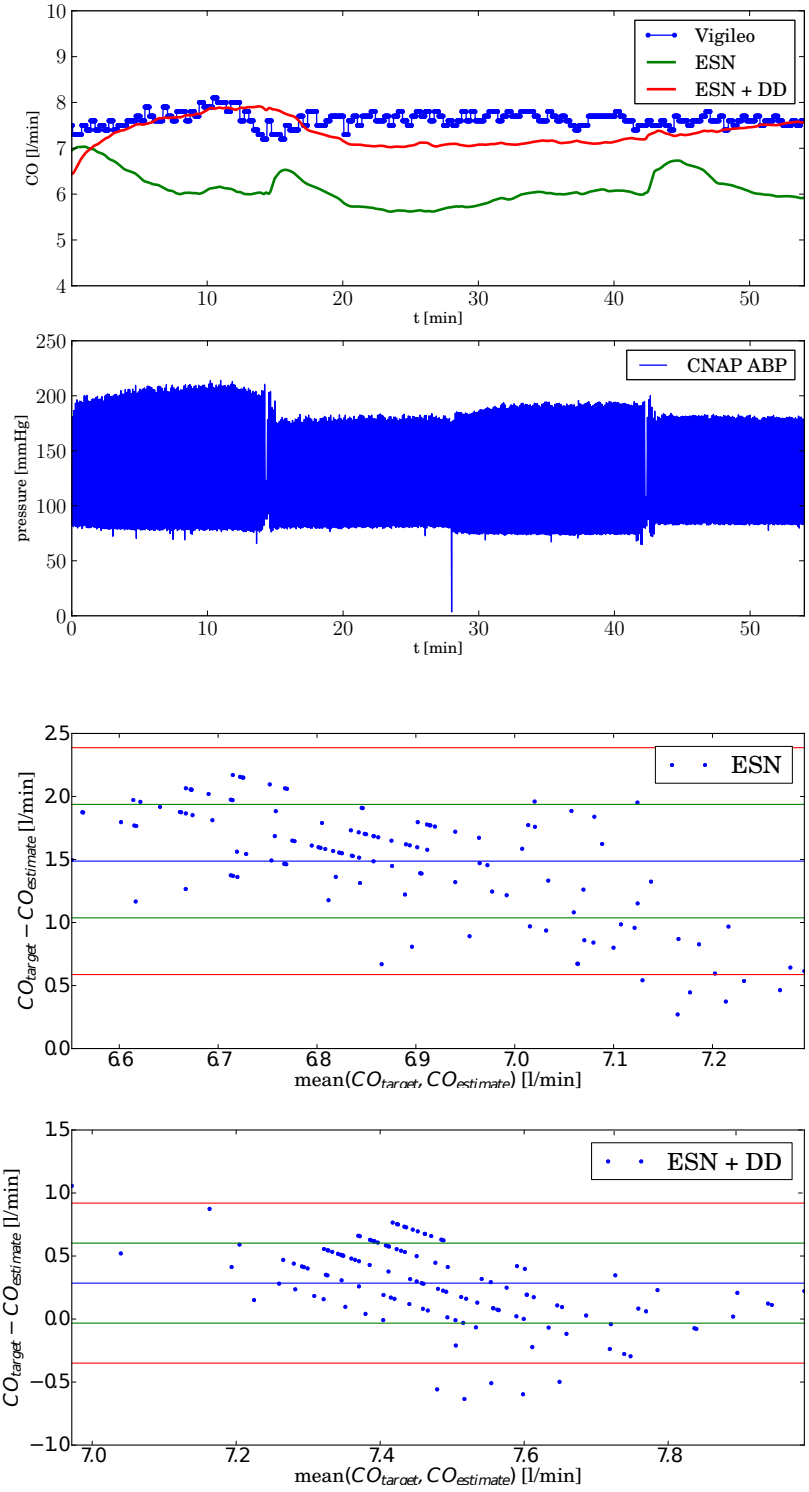

Figure 3.2: Patient 12: The first plot shows the targets (Vigileo/FloTrac) in blue. The green graph is the output of the ESN in the first experiment (without DD). The red graph is the output of the second experiment (including DD as virtual neurons). The second plot shows the corresponding ABP signal recorded by the CNAP sensor heavily compressed. Third and fourth plot show Blant-Altman-diagrams for both experiments.

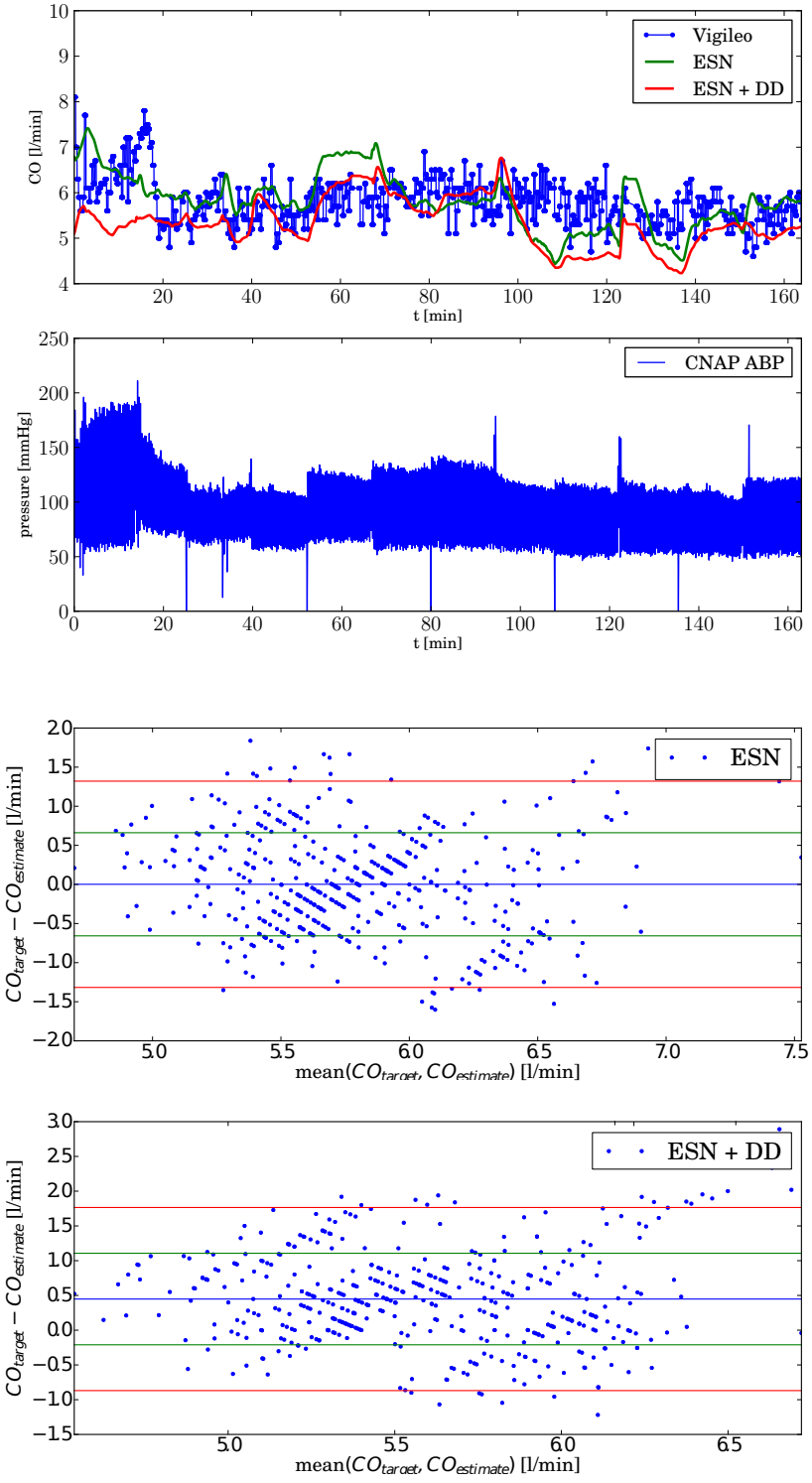

Figure 3.3: Patient 14: The first plot shows the targets (Vigileo/FloTrac) in blue. The green graph is the output of the ESN in the first experiment (without DD). The red graph is the output of the second experiment (including DD as virtual neurons). The second plot shows the corresponding ABP signal recorded by the CNAP sensor heavily compressed. Third and fourth plot show Blant-Altman-diagrams for both experiments.

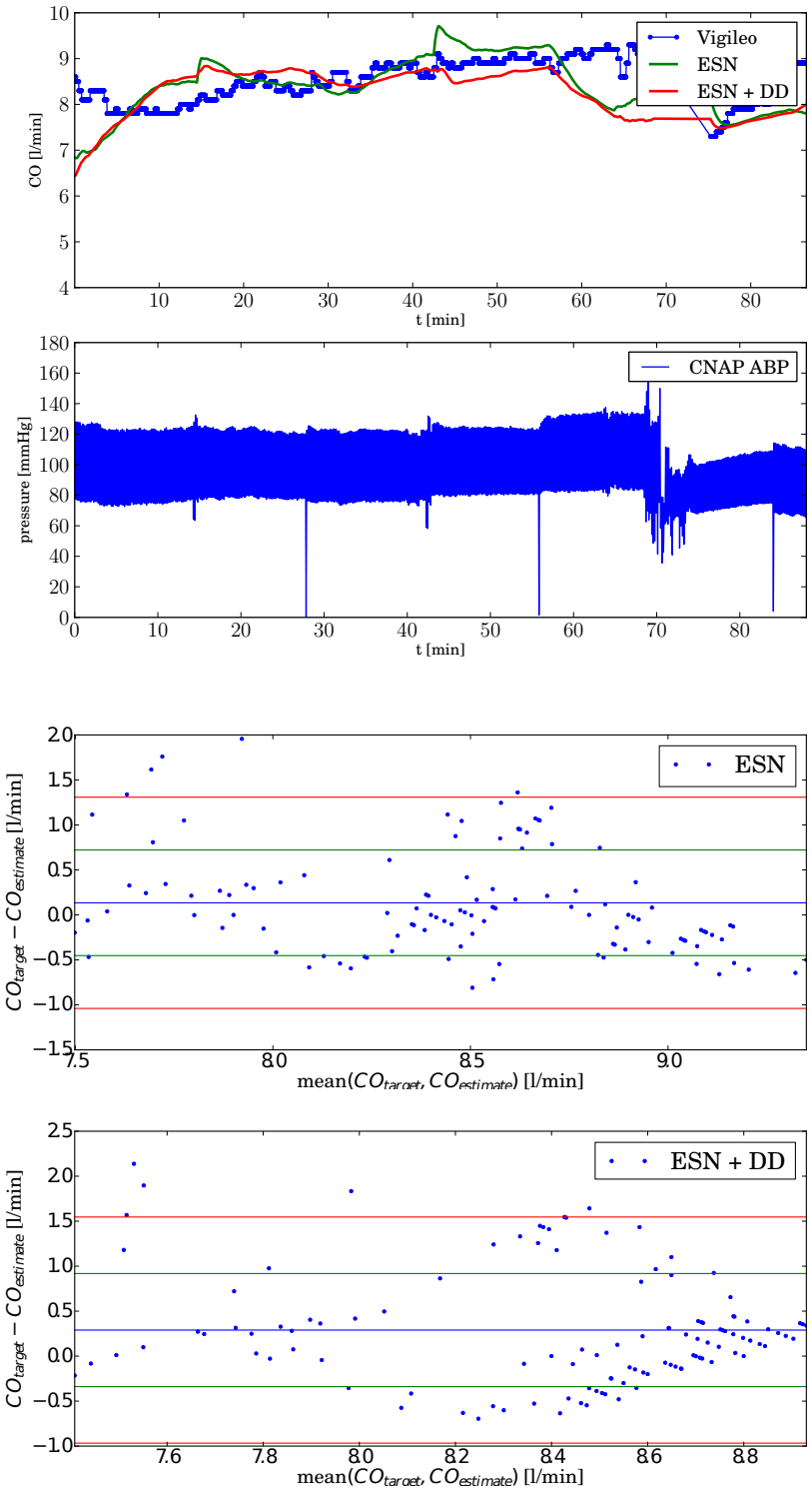

Figure 3.4: Patient 15: The first plot shows the targets (Vigileo/FloTrac) in blue. The green graph is the output of the ESN in the first experiment (without DD). The red graph is the output of the second experiment (including DD as virtual neurons). The second plot shows the corresponding ABP signal recorded by the CNAP sensor heavily compressed. Third and fourth plot show Blant-Altman-diagrams for both experiments.

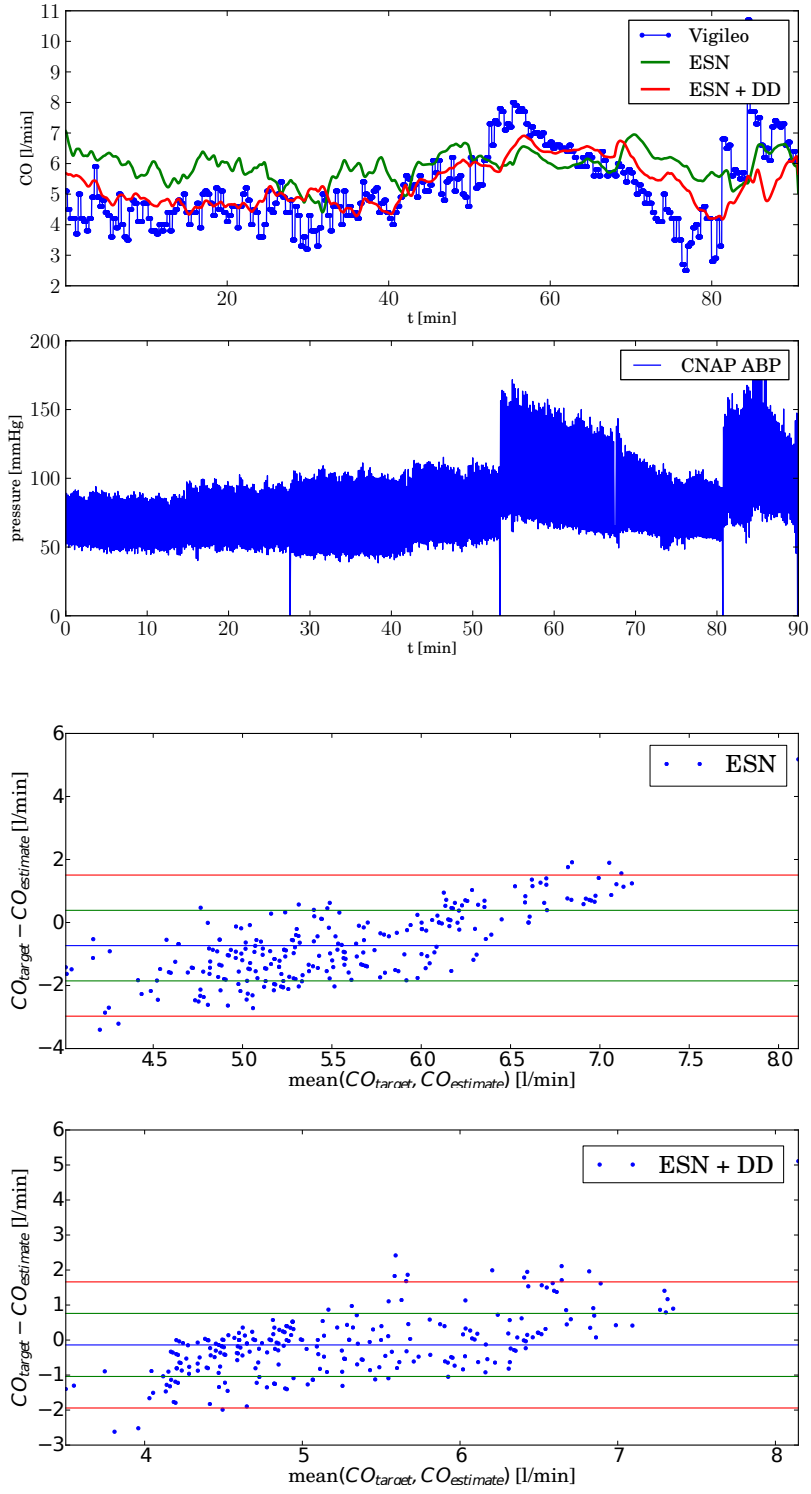

Figure 3.5: Patient 18: The first plot shows the targets (Vigileo/FloTrac) in blue. The green graph is the output of the ESN in the first experiment (without DD). The red graph is the output of the second experiment (including DD as virtual neurons). The second plot shows the corresponding ABP signal recorded by the CNAP sensor heavily compressed. Third and fourth plot show Blant-Altman-diagrams for both experiments.

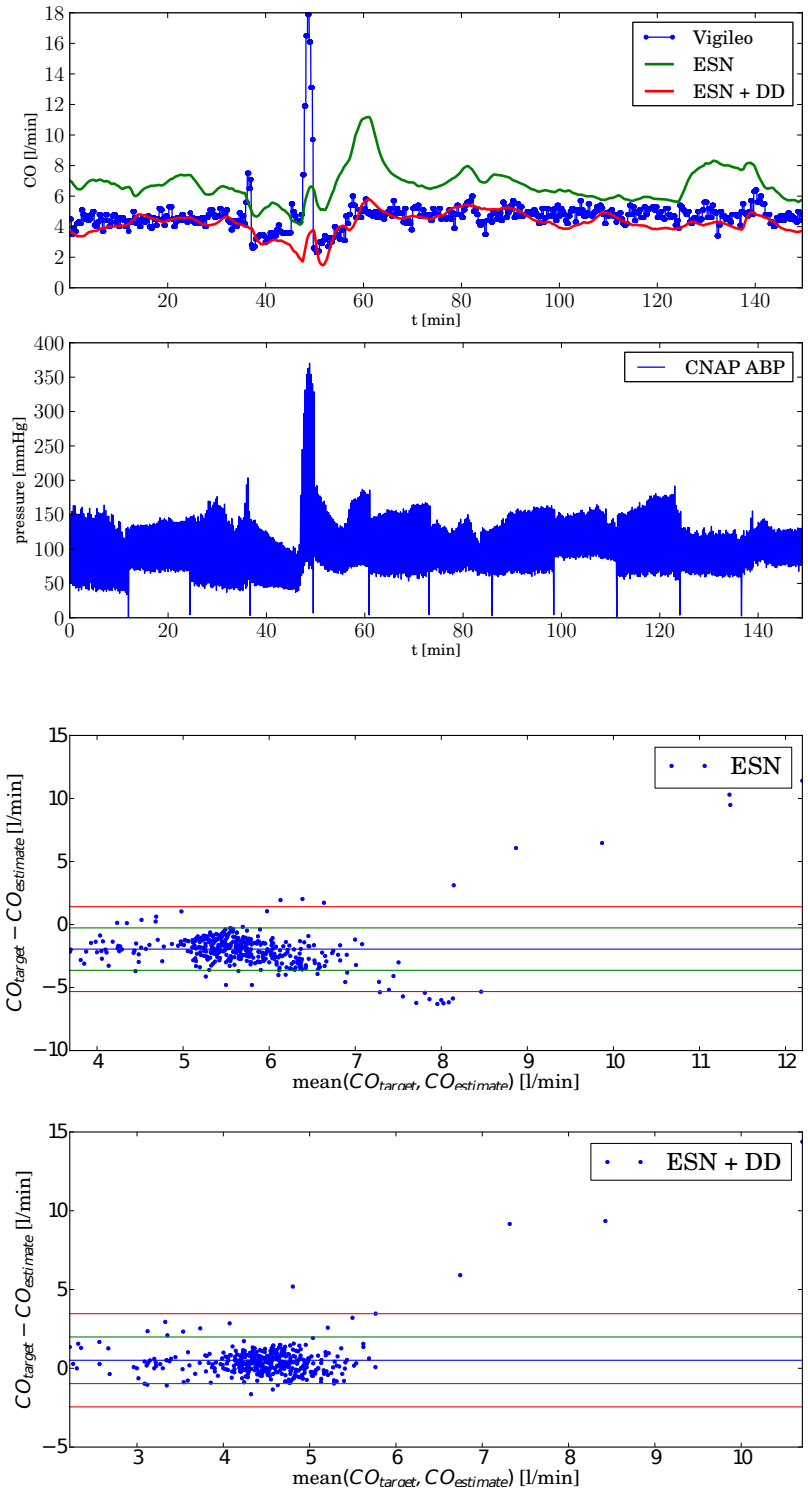

Figure 3.6: Patient 25: The first plot shows the targets (Vigileo/FloTrac) in blue. The green graph is the output of the ESN in the first experiment (without DD). The red graph is the output of the second experiment (including DD as virtual neurons). The second plot shows the corresponding ABP signal recorded by the CNAP sensor heavily compressed. Third and fourth plot show Blant-Altman-diagrams for both experiments.

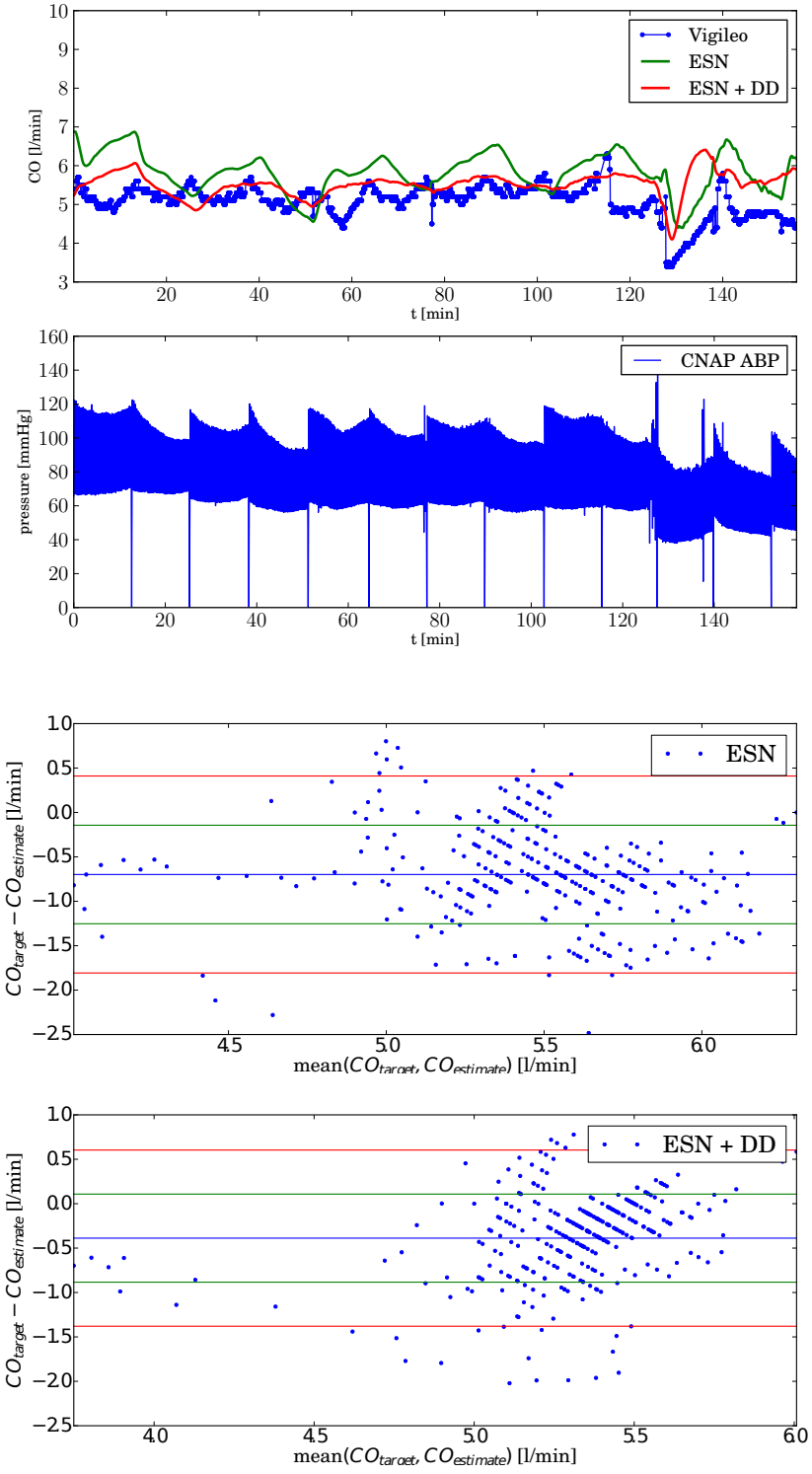

Figure 3.7: Patient 32: The first plot shows the targets (Vigileo/FloTrac) in blue. The green graph is the output of the ESN in the first experiment (without DD). The red graph is the output of the second experiment (including DD as virtual neurons). The second plot shows the corresponding ABP signal recorded by the CNAP sensor heavily compressed. Third and fourth plot show Blant-Altman-diagrams for both experiments.

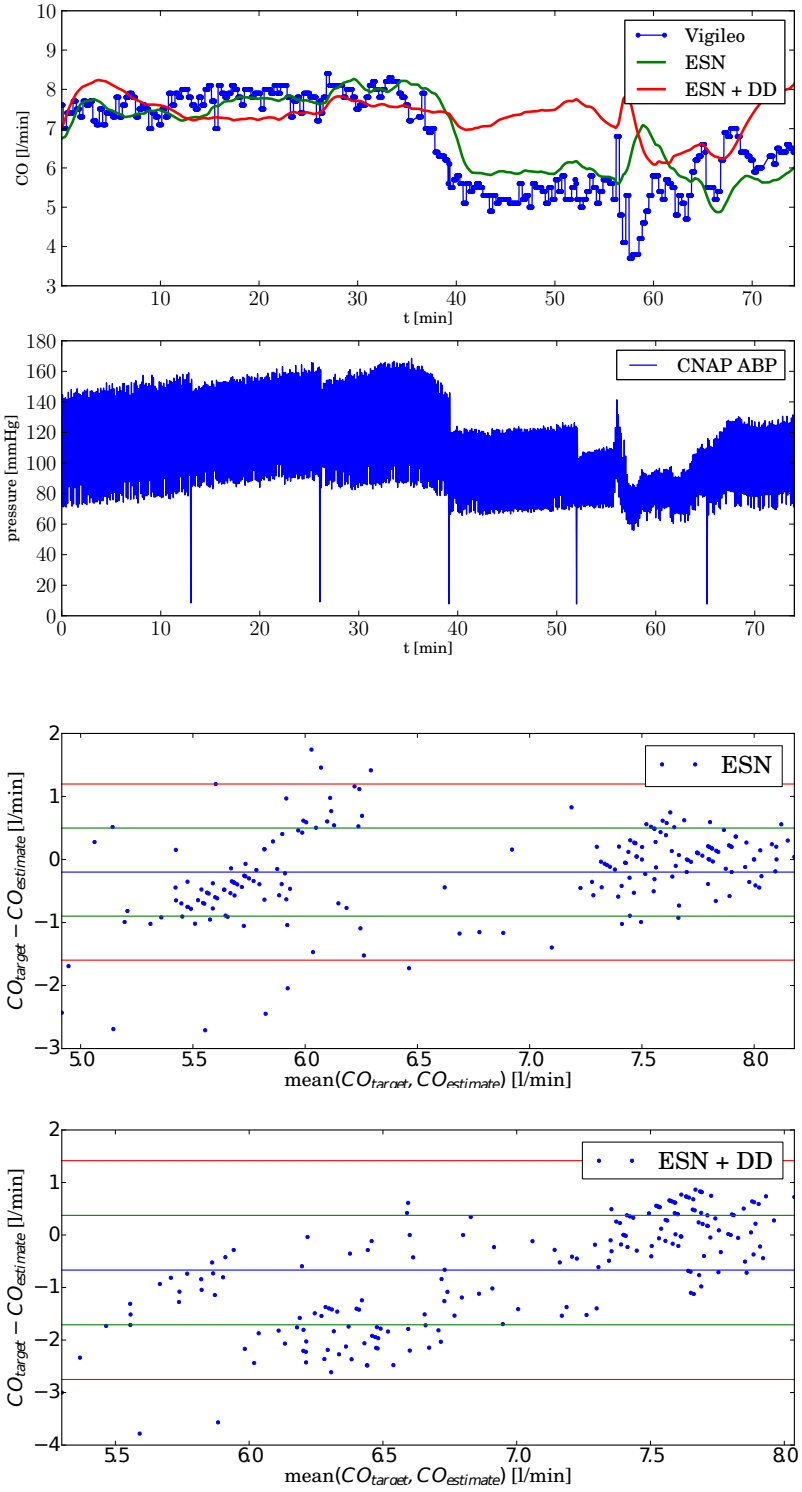

Figure 3.8: Patient 37: The first plot shows the targets (Vigileo/FloTrac) in blue. The green graph is the output of the ESN in the first experiment (without DD). The red graph is the output of the second experiment (including DD as virtual neurons). The second plot shows the corresponding ABP signal recorded by the CNAP sensor heavily compressed. Third and fourth plot show Blant-Altman-diagrams for both experiments.

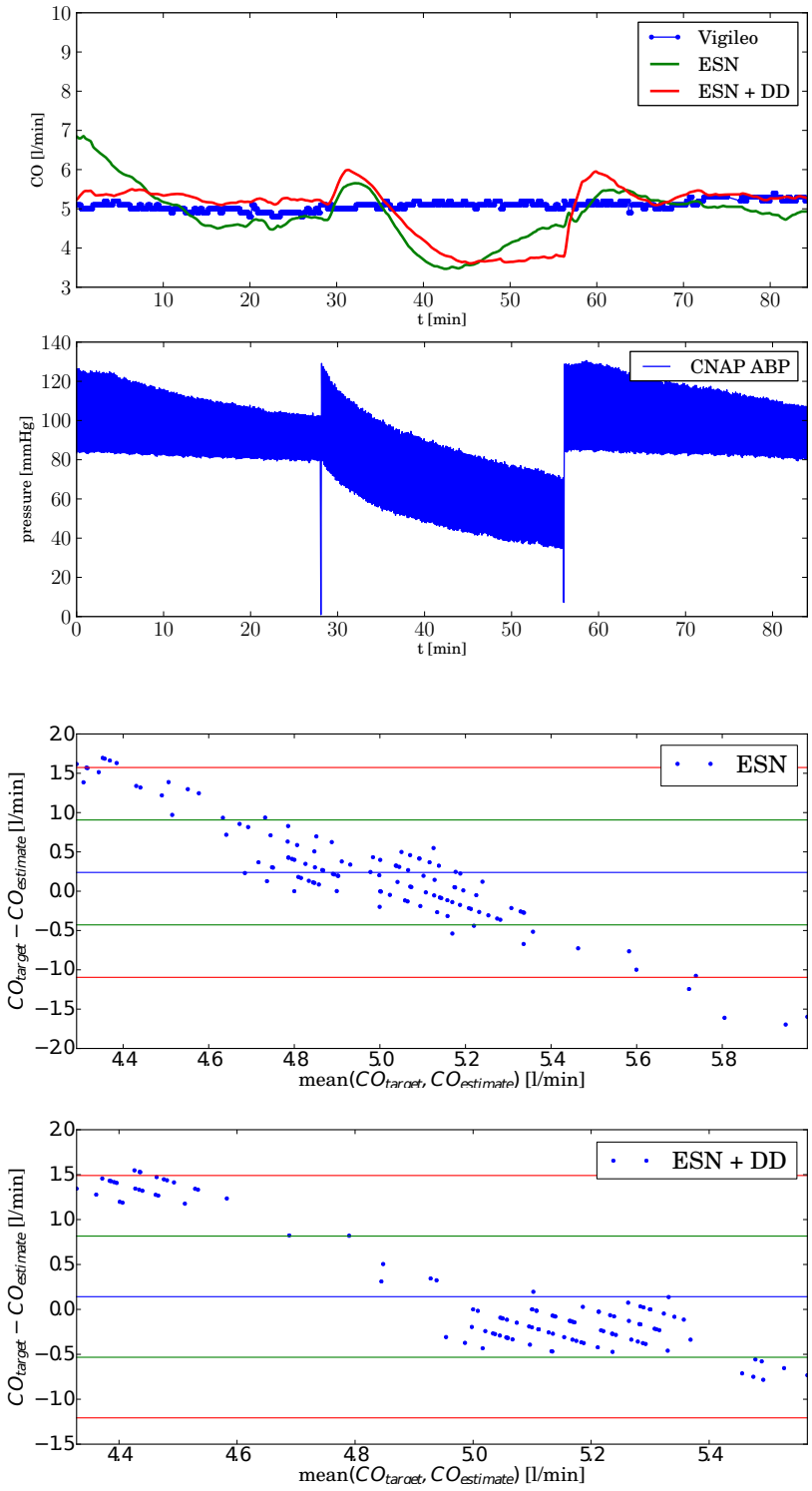

Figure 3.9: Patient 45: The first plot shows the targets (Vigileo/FloTrac) in blue. The green graph is the output of the ESN in the first experiment (without DD). The red graph is the output of the second experiment (including DD as virtual neurons). The second plot shows the corresponding ABP signal recorded by the CNAP sensor heavily compressed. Third and fourth plot show Blant-Altman-diagrams for both experiments.

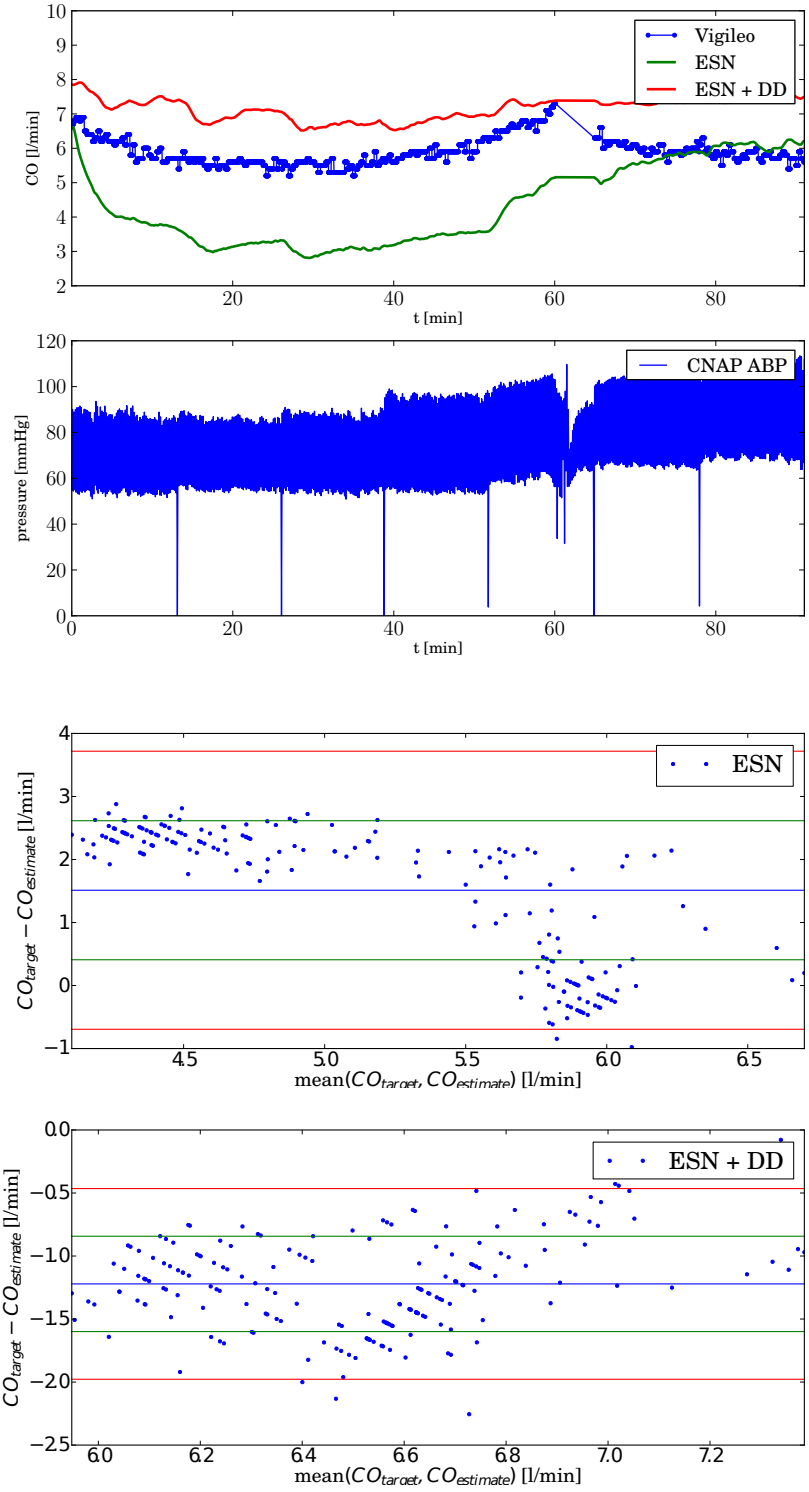

Figure 3.10: Patient 57: The first plot shows the targets (Vigileo/FloTrac) in blue. The green graph is the output of the ESN in the first experiment (without DD). The red graph is the output of the second experiment (including DD as virtual neurons). The second plot shows the corresponding ABP signal recorded by the CNAP sensor heavily compressed. Third and fourth plot show Blant-Altman-diagrams for both experiments.

Bias

What can be seen in the results is that the CO estimation is often inflicted with bias. Very good examples are patients 7 (figure 3.1) and 12 (figure 3.2). In general the bias seems to improve in most cases when descriptive data is added to the model. Obviously the Vigileo/FloTrac device uses descriptive data for calibration purposes. It is unknown if the four sets of descriptive data (height, weight, age, sex) suffice to model the Vigileo's calibration method. A case where the bias strongly improves is patient 25 (figure 3.6)

Tracking

The tracking ability also seems to improve when descriptive data is added to the model. An example where this is most obvious is patient 37 (figure 3.8). Sometimes the estimator struggles with jumps in the ABP signal after a calibration phase. A good example is patient 32 (figure 3.7). After each calibration the pulse pressure as well as the systolic pressure jumps. The pulse pressure almost doubles each time. Each jump is followed by a jump in the CO estimation. A prime example of how parameter drifts affect the estimation can be seen with patient 45 (figure 3.9). Although there is almost no bias in the error the tracking ability is non-existing. Another source of error is a target signal with heavy noise like the one of patient 14 (figure 3.3).

3.1.2 CNAP Beat-to-Beat Features (Experiments 3 & 4)

The plots in this chapter depict results of the beat-to-beat features approach as discussed in chapter 2.3.2. In a first step features are extracted on a beat-to-beat basis. These features are normalized and used as temporal signals in an ESN.

Both approaches in this chapter utilize descriptive patient data. What distinguishes the two approaches is the way the descriptive data is used. In experiment 3 the descriptive data is used as additional constant input signal acting as additional bias to the reservoir as described in chapter 2.3.4. The bias varies from patient to patient but is constant over all training samples of one patient.

In experiment 4 the descriptive data is used as virtual neurons just like in experiment 2 from the previous section. Also the upcoming plots are the same subset of patients as in the previous section to be able to compare all 4 experiments among each other.

Again each page represents a single patient including four plots:

1. The target samples as recorded by the Vigileo/FloTrac device.
2. A heavily compressed view of the ABP signal.
3. A BA-diagram of experiment 3
4. A BA-diagram of experiment 4

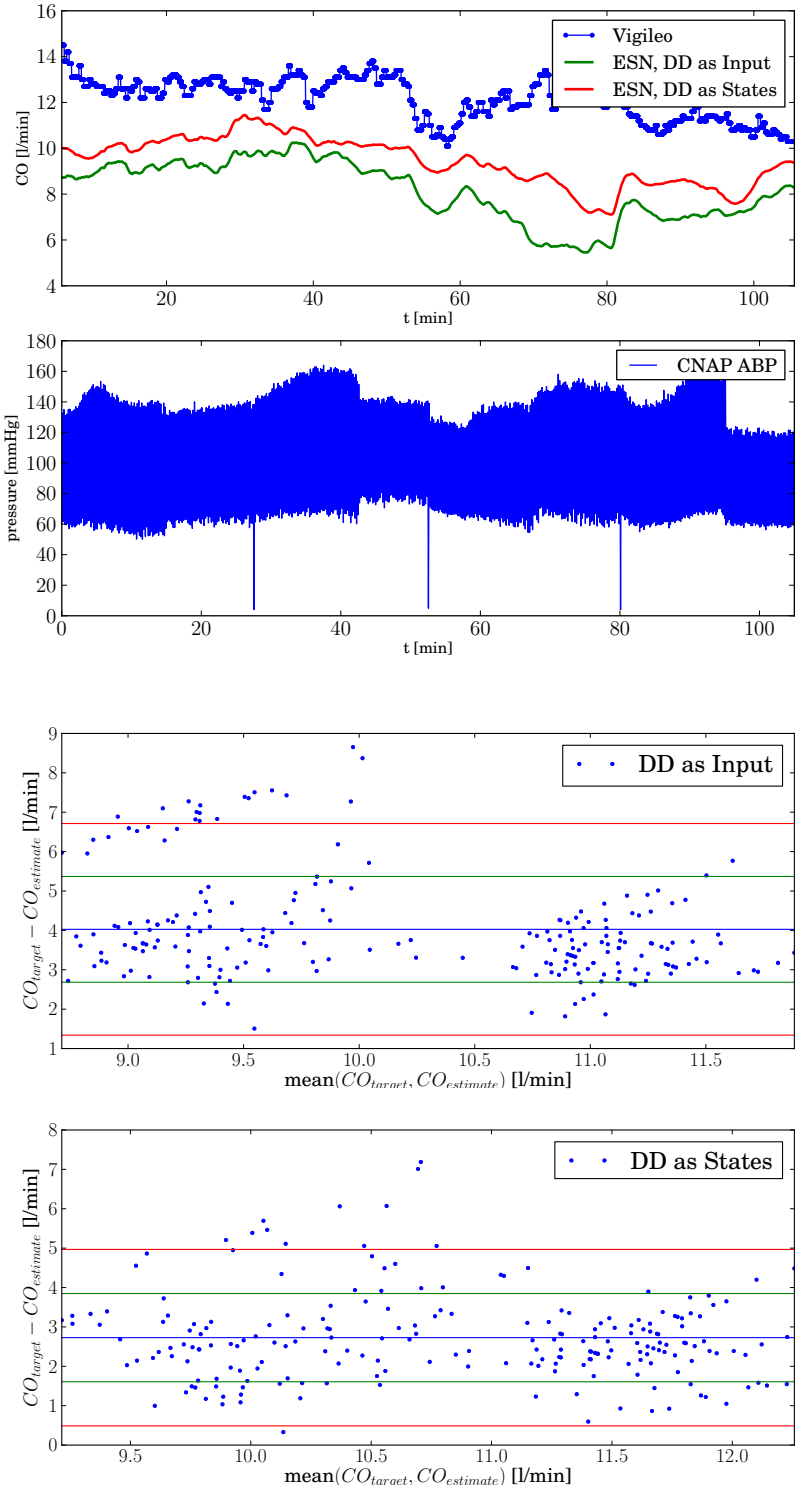

Figure 3.11: Patient 07: The first plot shows the targets (Vigileo/FloTrac) in blue. The green graph is the output of the ESN in the 3. experiment (Signal features with DD as input). The red graph is the output of the 4. experiment (Signal features with DD as virtual neurons). The second plot shows the corresponding ABP signal recorded by the CNAP sensor heavily compressed. Third and fourth plot show Blant-Altman-diagrams for both experiments.

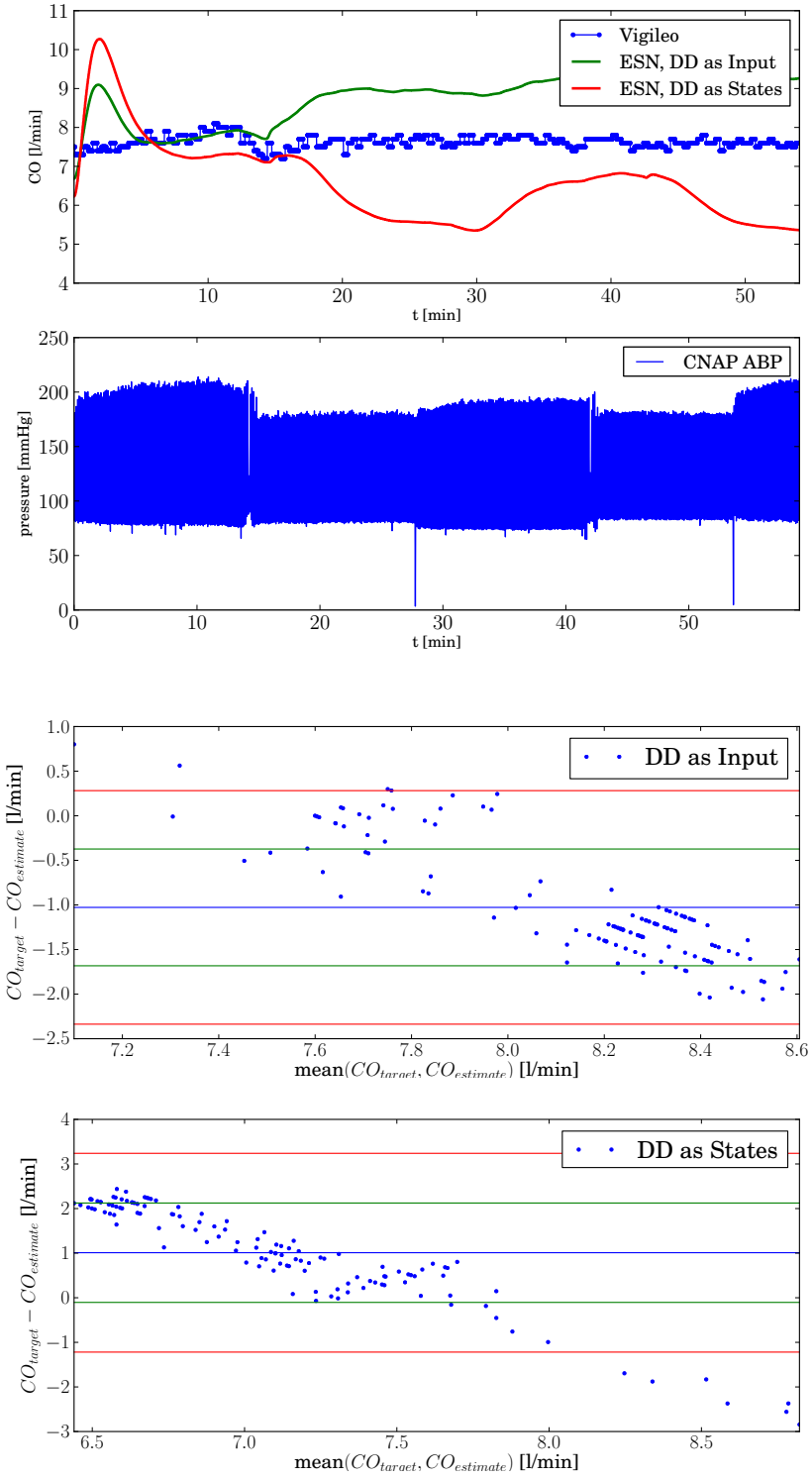

Figure 3.12: Patient 12: The first plot shows the targets (Vigileo/FloTrac) in blue. The green graph is the output of the ESN in the 3. experiment (Signal features with DD as input). The red graph is the output of the 4. experiment (Signal features with DD as virtual neurons). The second plot shows the corresponding ABP signal recorded by the CNAP sensor heavily compressed. Third and fourth plot show Blant-Altman-diagrams for both experiments.

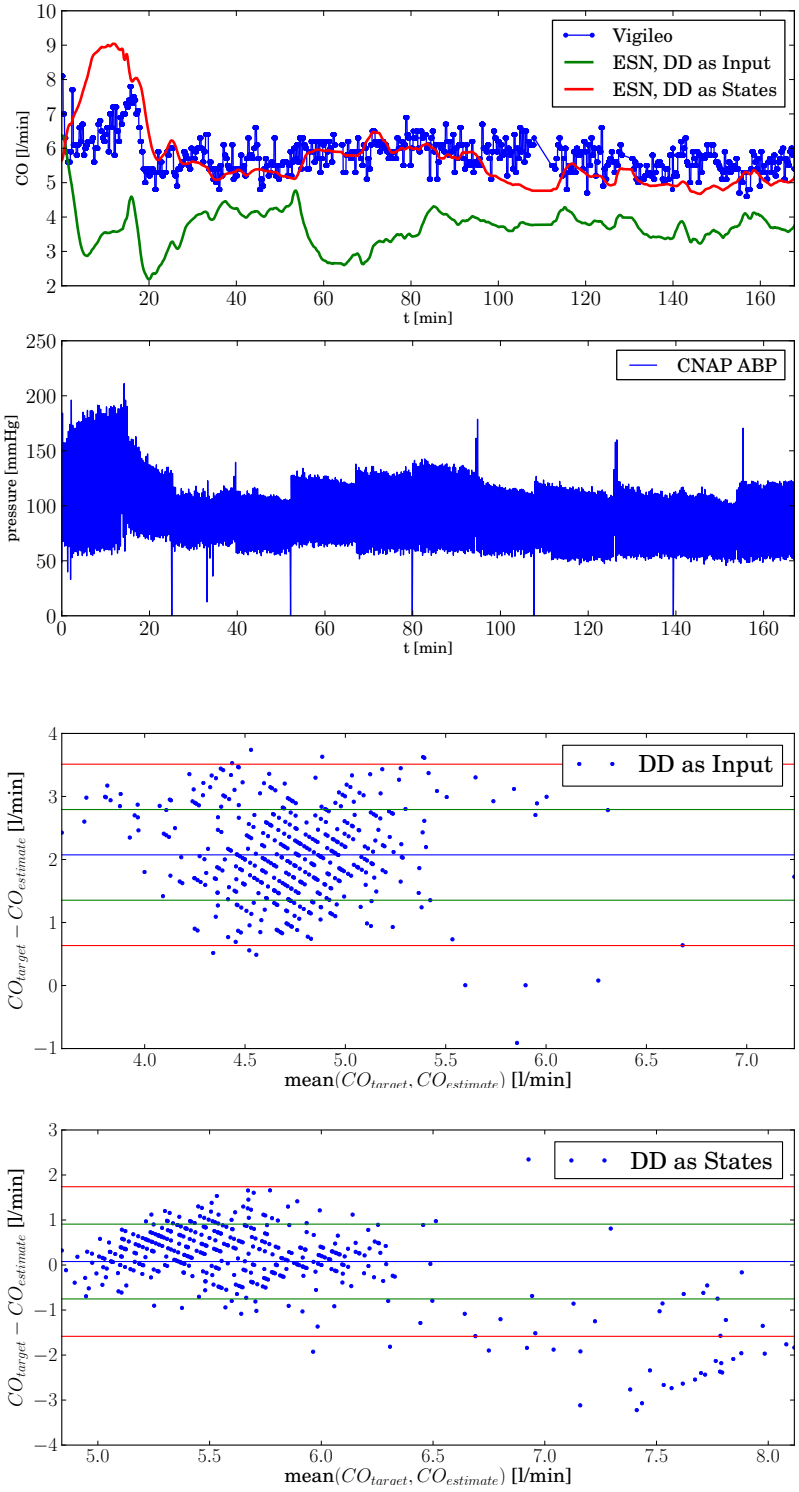

Figure 3.13: Patient 14: The first plot shows the targets (Vigileo/FloTrac) in blue. The green graph is the output of the ESN in the 3. experiment (Signal features with DD as input). The red graph is the output of the 4. experiment (Signal features with DD as virtual neurons). The second plot shows the corresponding ABP signal recorded by the CNAP sensor heavily compressed. Third and fourth plot show Blant-Altman-diagrams for both experiments.

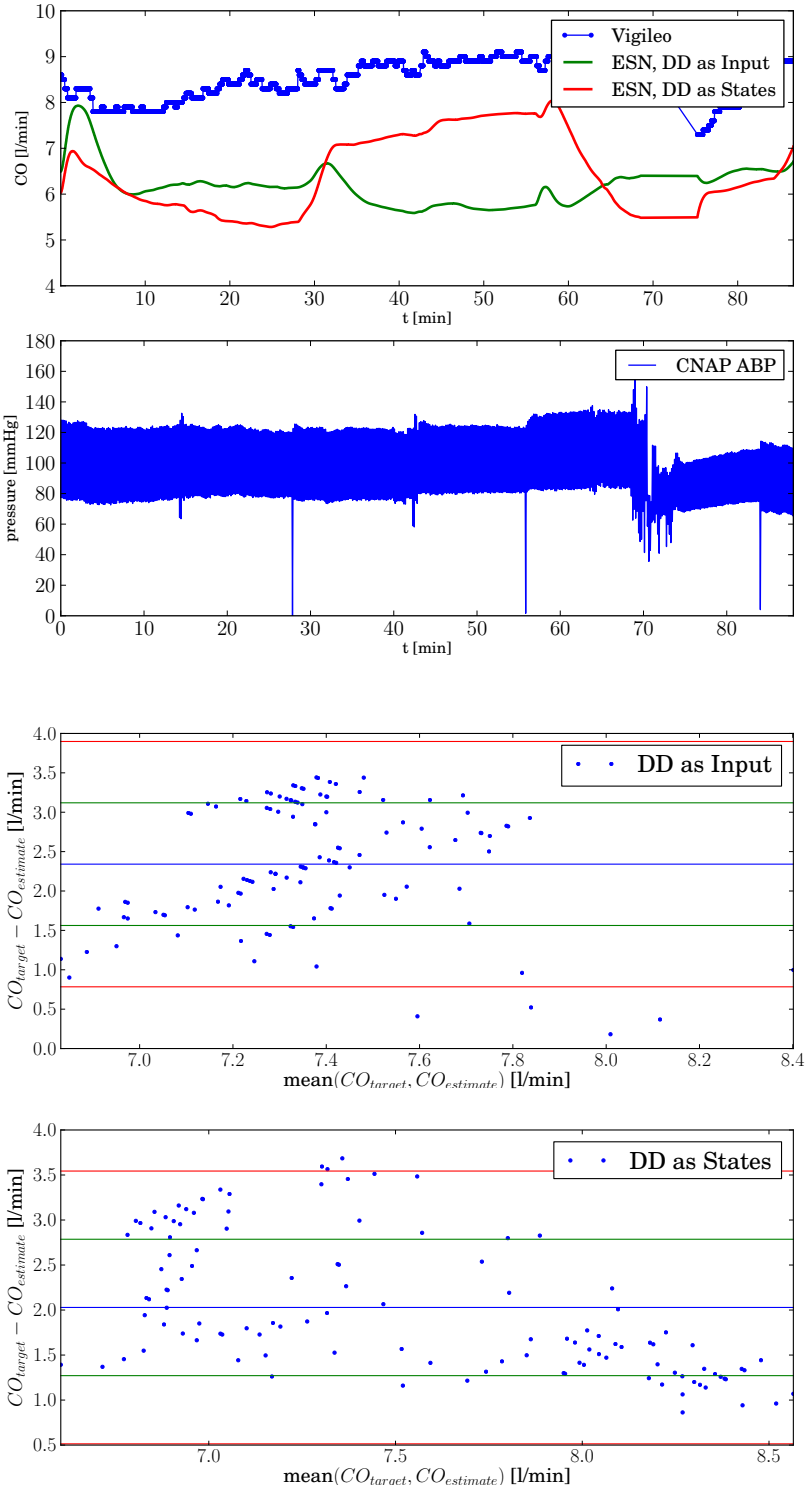

Figure 3.14: Patient 15: The first plot shows the targets (Vigileo/FloTrac) in blue. The green graph is the output of the ESN in the 3. experiment (Signal features with DD as input). The red graph is the output of the 4. experiment (Signal features with DD as virtual neurons). The second plot shows the corresponding ABP signal recorded by the CNAP sensor heavily compressed. Third and fourth plot show Blant-Altman-diagrams for both experiments.

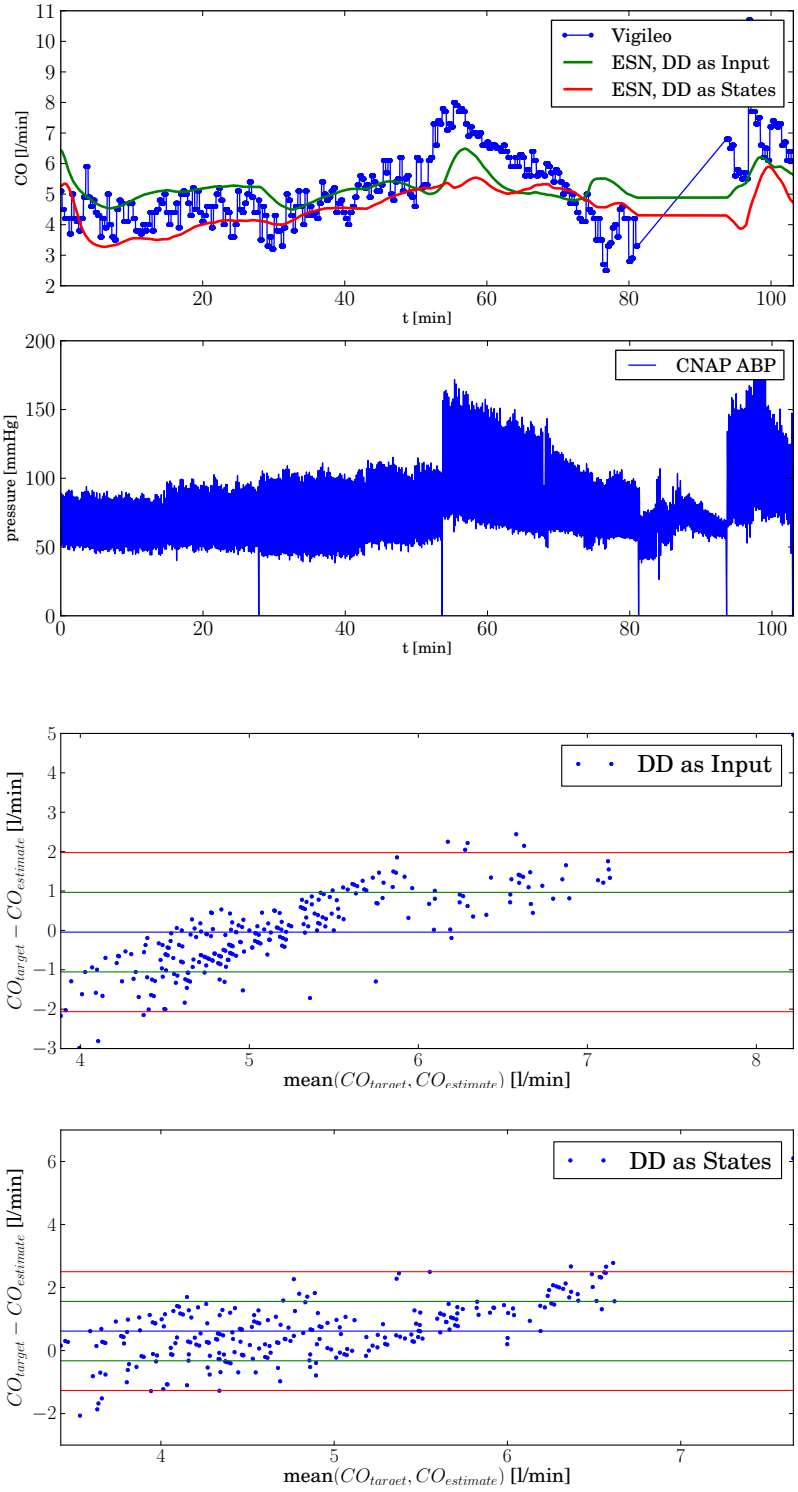

Figure 3.15: Patient 18: The first plot shows the targets (Vigileo/FloTrac) in blue. The green graph is the output of the ESN in the 3. experiment (Signal features with DD as input). The red graph is the output of the 4. experiment (Signal features with DD as virtual neurons). The second plot shows the corresponding ABP signal recorded by the CNAP sensor heavily compressed. Third and fourth plot show Blant-Altman-diagrams for both experiments.

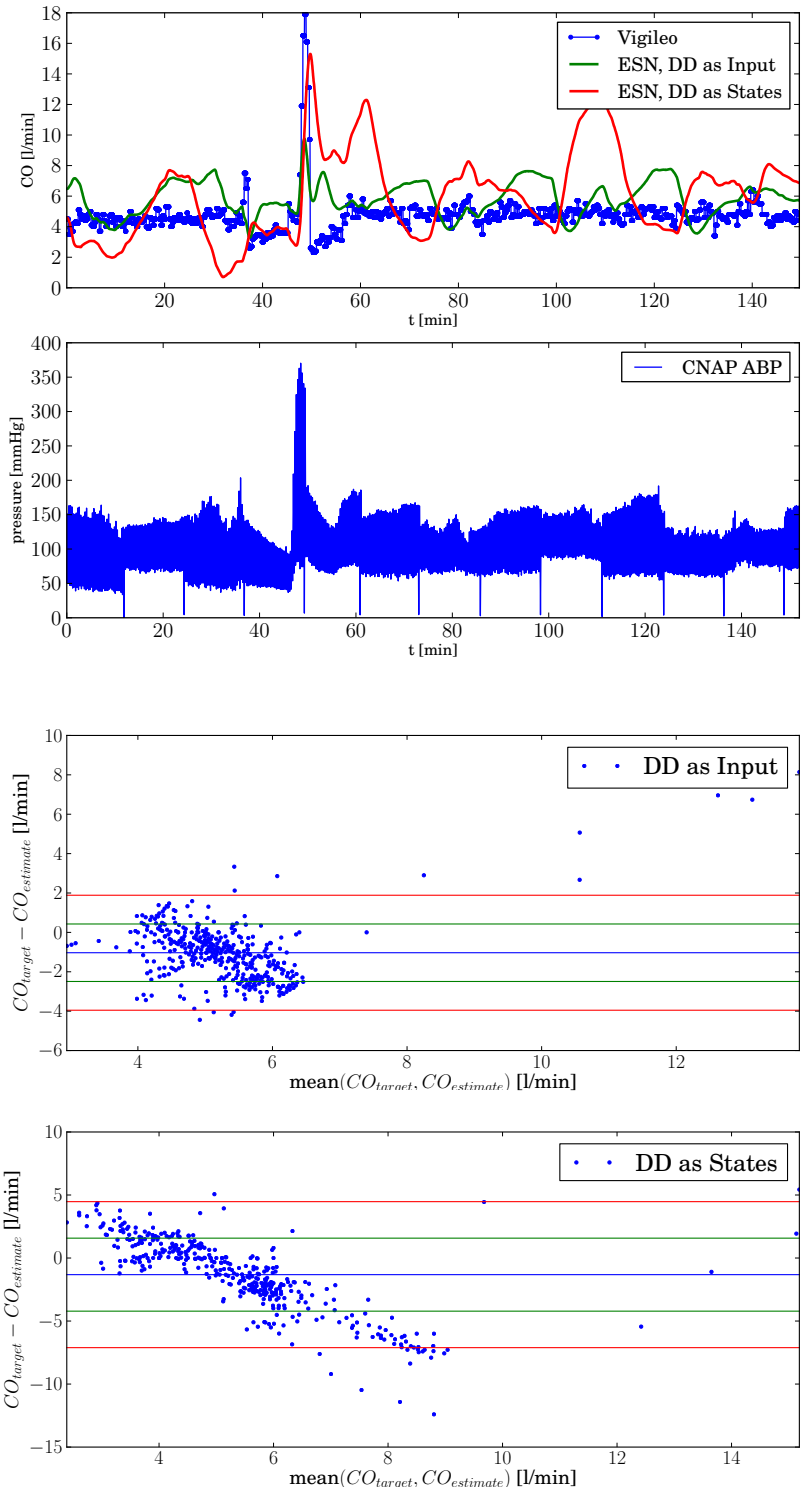

Figure 3.16: Patient 25: The first plot shows the targets (Vigileo/FloTrac) in blue. The green graph is the output of the ESN in the 3. experiment (Signal features with DD as input). The red graph is the output of the 4. experiment (Signal features with DD as virtual neurons). The second plot shows the corresponding ABP signal recorded by the CNAP sensor heavily compressed. Third and fourth plot show Blant-Altman-diagrams for both experiments.

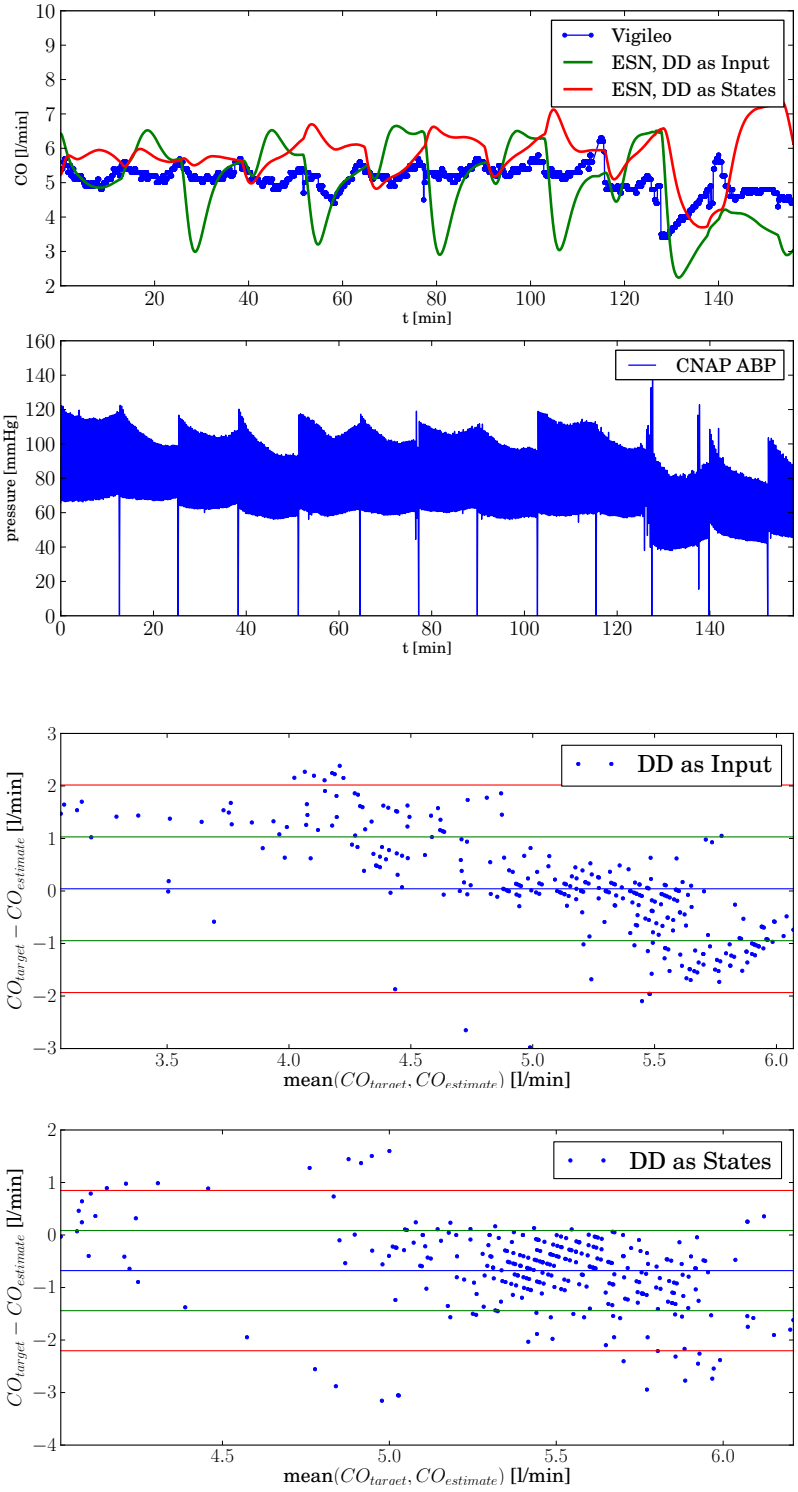

Figure 3.17: Patient 32: The first plot shows the targets (Vigileo/FloTrac) in blue. The green graph is the output of the ESN in the 3. experiment (Signal features with DD as input). The red graph is the output of the 4. experiment (Signal features with DD as virtual neurons). The second plot shows the corresponding ABP signal recorded by the CNAP sensor heavily compressed. Third and fourth plot show Blant-Altman-diagrams for both experiments.

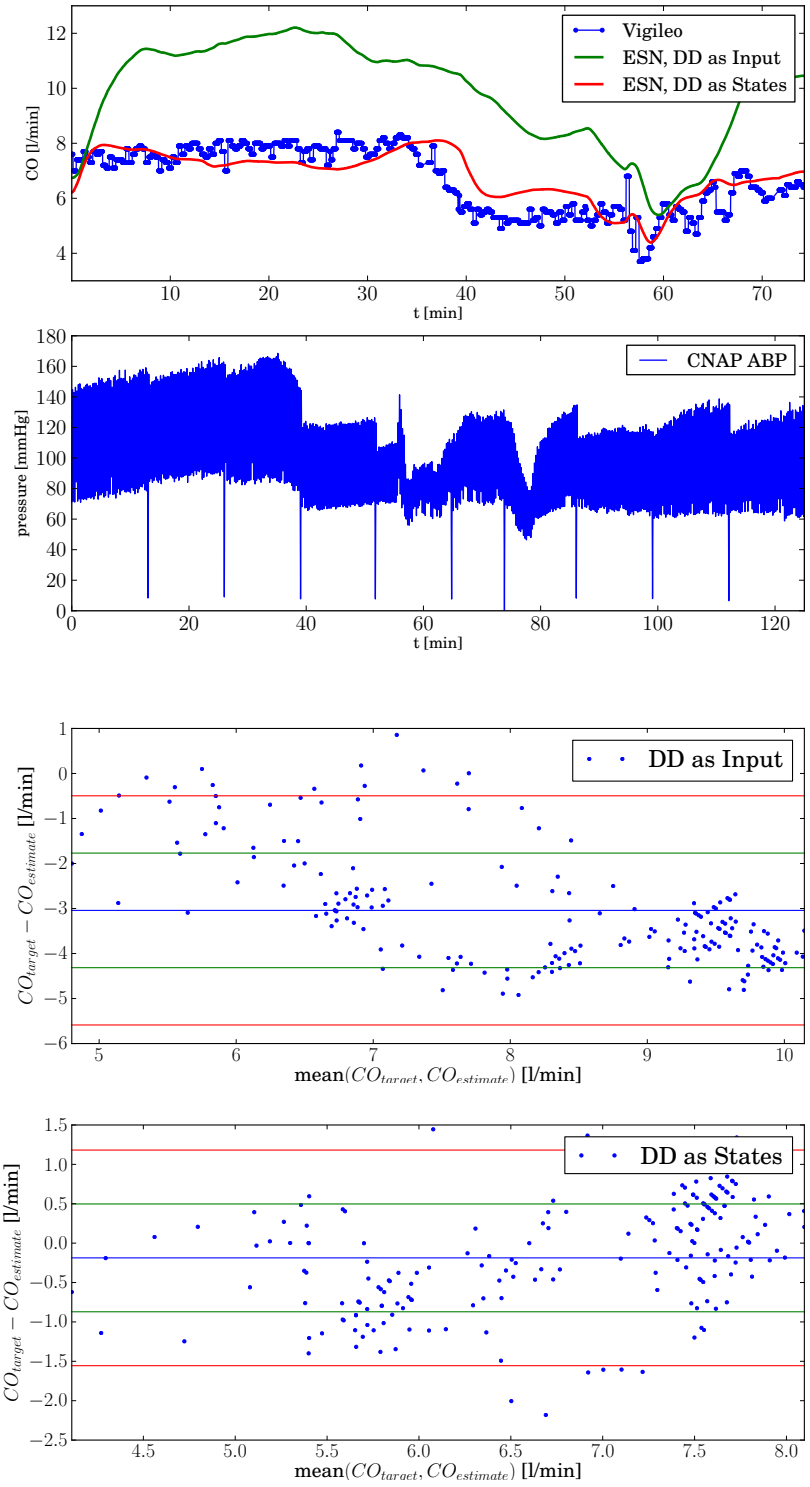

Figure 3.18: Patient 37: The first plot shows the targets (Vigileo/FloTrac) in blue. The green graph is the output of the ESN in the 3. experiment (Signal features with DD as input). The red graph is the output of the 4. experiment (Signal features with DD as virtual neurons). The second plot shows the corresponding ABP signal recorded by the CNAP sensor heavily compressed. Third and fourth plot show Blant-Altman-diagrams for both experiments.

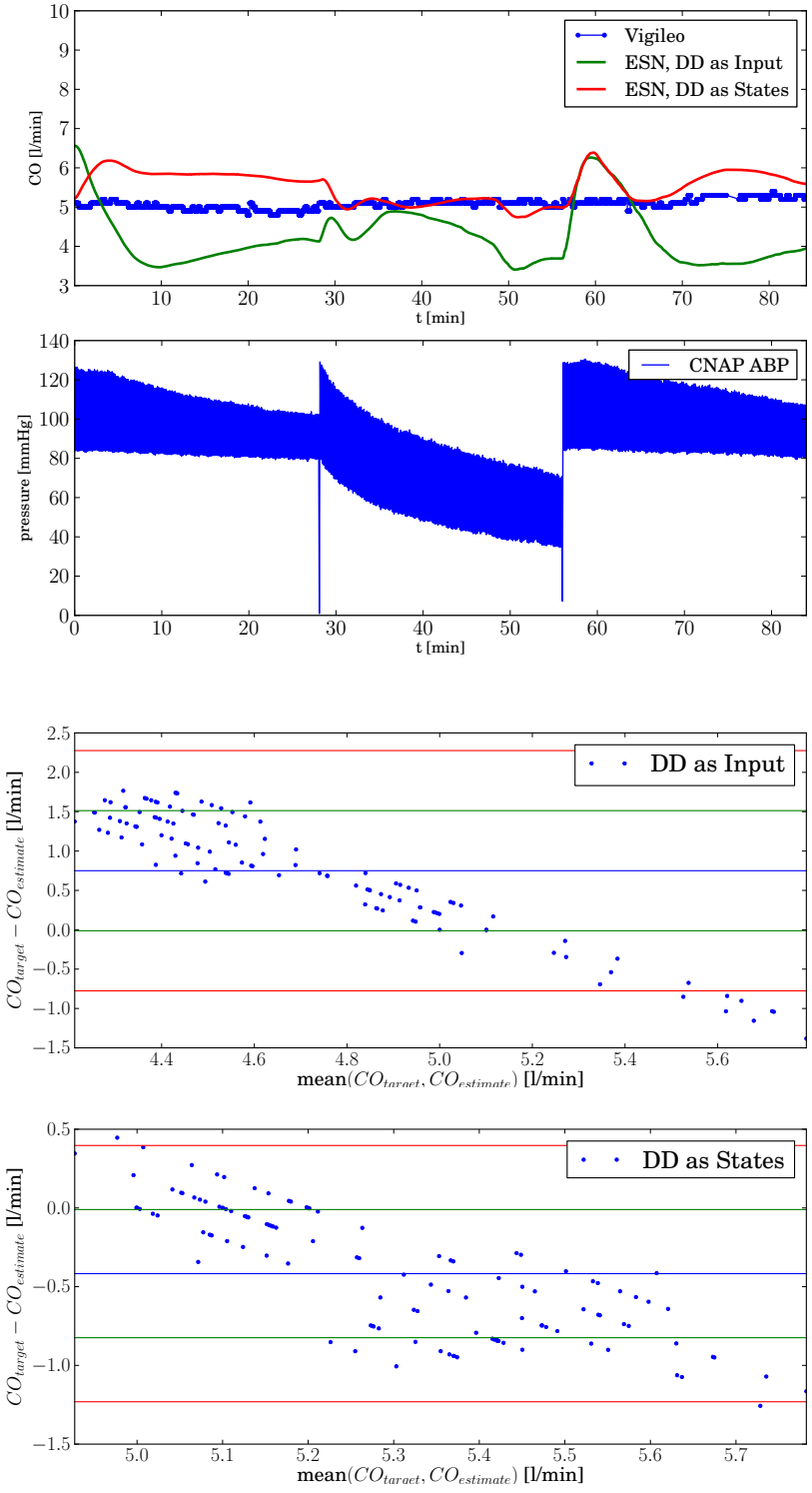

Figure 3.19: Patient 45: The first plot shows the targets (Vigileo/FloTrac) in blue. The green graph is the output of the ESN in the 3. experiment (Signal features with DD as input). The red graph is the output of the 4. experiment (Signal features with DD as virtual neurons). The second plot shows the corresponding ABP signal recorded by the CNAP sensor heavily compressed. Third and fourth plot show Blant-Altman-diagrams for both experiments.

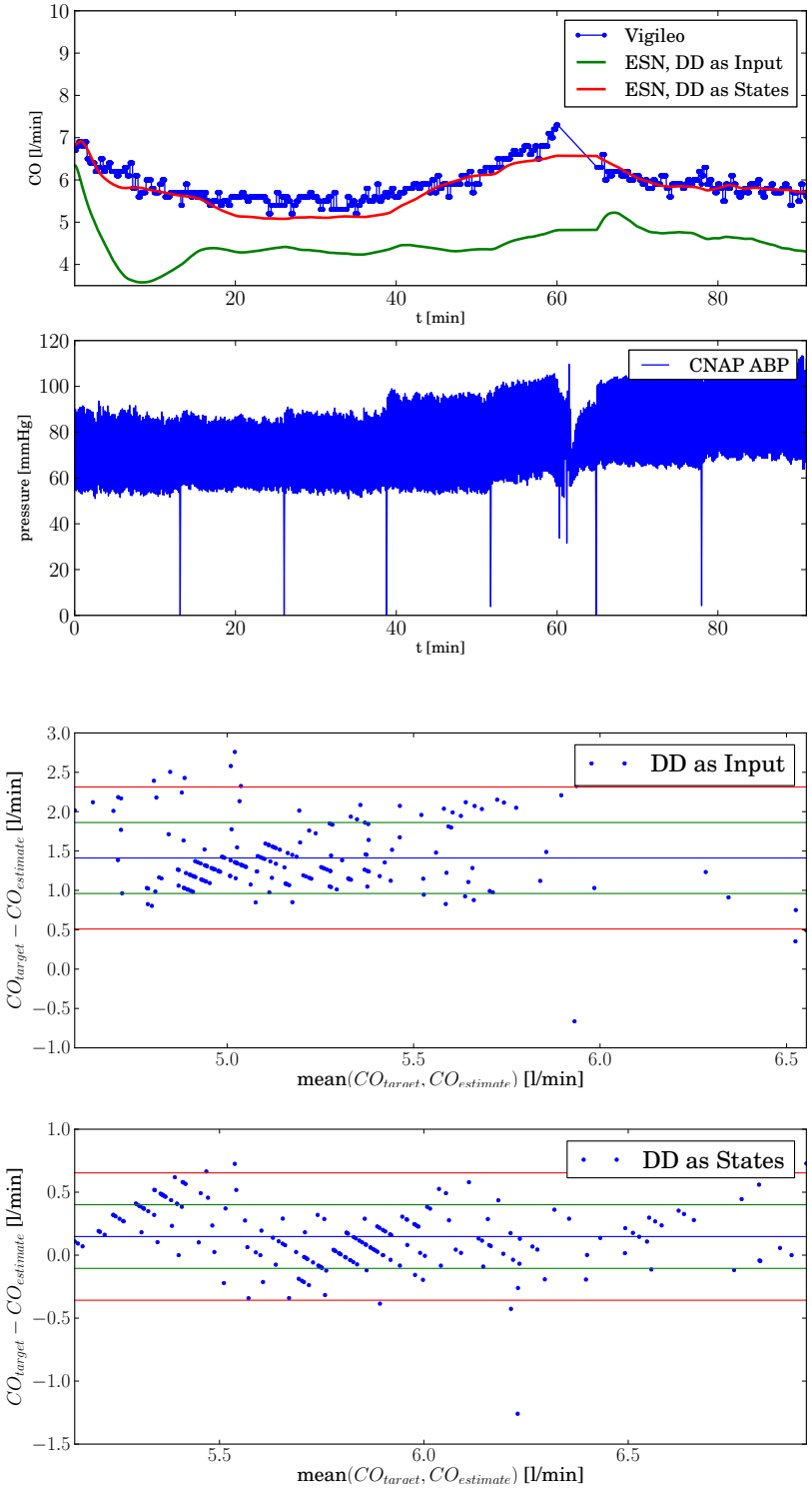

Figure 3.20: Patient 57: The first plot shows the targets (Vigileo/FloTrac) in blue. The green graph is the output of the ESN in the 3. experiment (Signal features with DD as input). The red graph is the output of the 4. experiment (Signal features with DD as virtual neurons). The second plot shows the corresponding ABP signal recorded by the CNAP sensor heavily compressed. Third and fourth plot show Blant-Altman-diagrams for both experiments.

Bias

The results in terms of bias are very similar to the results from experiment 1 & 2. In general the bias is smaller when using DD as virtual neurons than as input signal. In some cases the difference is quite big as for patient 14 (figure 3.13), 37 (figure 3.18) or 57 (figure 3.20).

When compared to the experiments of the previous section (1 & 2) the bias is worse except for some special cases like patient 37 (figure 3.18) where experiment 4 performed exceptionally good.

Tracking

In general the methods using signal features for inputs are much more sensitive to calibration jumps and parameter drifts than the direct methods (experiment 1 & 2). This can easily be seen in the case of patients 25 (figure 3.16) and 32 (figure 3.13). The cases in which these methods are performing well are those with steady smooth target trends as for patients 37 (figure 3.18) and 57 (figure 3.13), but only if DD is used as virtual neurons. Target signals including heavy noise are better matched using features with DD as virtual neurons (experiment 4) as shown in the case of patient 14 (figure 3.13). In this manner experiment 4 performed best out of all 4 experiments. Experiment 4 outperforms experiment 3 in general. Therefore it seams to be feasible to use descriptive data as virtual neurons (e.g. to add it as linear combination to the result).

3.1.3 Sequential Forward/Backward Feature Selection

As described in chapter 2.6 the optimal feature set was determined using a sequential feature selection algorithm. Figure 3.21 depicts sequences found by this algorithm. Generally, two steps are forward, every third step is backward. Each step consists of 4 trials of leave-one-out cross-validations.

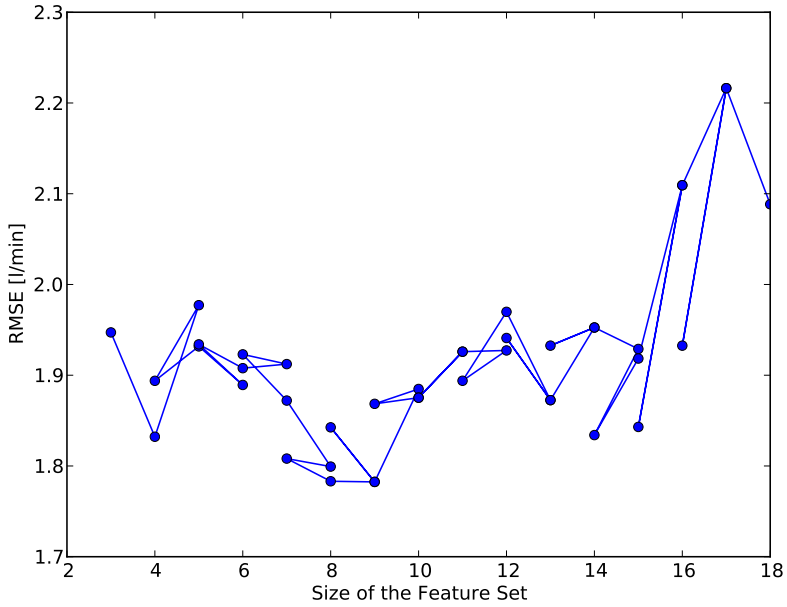

Figure 3.21: Sequential forward/backward feature selection. Starting at a set of 3 features the sequential feature selection was performed to optimize the RMSE.

Starting out with a set of 3 features the optimum was reached with a set of 9 features including:

- Mean and variance of ABP
- Mean and variance of blood volume (control deviation)
- Second harmonic of ABP
- First and second harmonic of blood volume
- Heart rate
- Height

3.1.4 Comparison of CNAP Results

All four approaches utilizing the CNAP/FloTrac database are compared in terms of average performance (experiments 1 through 4). Figure 3.22 represents the essence of this study since the *percentage error* (PE) is considered as the best established quantity to assess CO estimation methods. Figure 3.23 depicts statistics about the distribution of bias which complements the PE assessment.

The performance of all four experiments are compared using the measurement of the Vigileo/FloTrac device as ground truth. Thus all results can only show the difference of

the approaches presented here to the Vigileo system as reference. As the Vigileo/FloTrac device uses an entirely different measurement as bases (non-invasive ABP on the finger versus arterial catheter) this is an additional source of error which has to be taken into account.

It shows that a continuous approach (experiments 1 & 2) outperforms an approach based on beat-to-beat features (experiments 3 & 4) both in terms of PE and bias. Also the use of *descriptive data* (DD) as virtual states increases the performance in all approaches with the peak in the continuous case. Only in some particular cases the approach in experiment 4 outperforms all other. This is mostly the case when calibration jumps are small and the target signal is rather steady.

It is not surprising that the average bias is low in all approaches. It is an easy task to find the mean of a target quantity. The more interesting information is the standard deviation of the bias. With 1.56 l/min in the best case it is quite high. This reflects one of the main problems of CO estimation: It is extremely hard to get the bias right with uncalibrated methods. Even using descriptive data as virtual neurons did not improve the bias significantly. An additional source of error is the fact that the nature of the calibration of the Vigileo/FloTrac device is in general unknown. If important information is missing, calibration might be almost impossible.

The RMSE comparison in figure 3.24 shows a similar image as the PE comparison but with less variation between the different approaches. RMSE represents the errors in both terms, shape and bias.

Tabel 3.1 shows the same results in numbers.

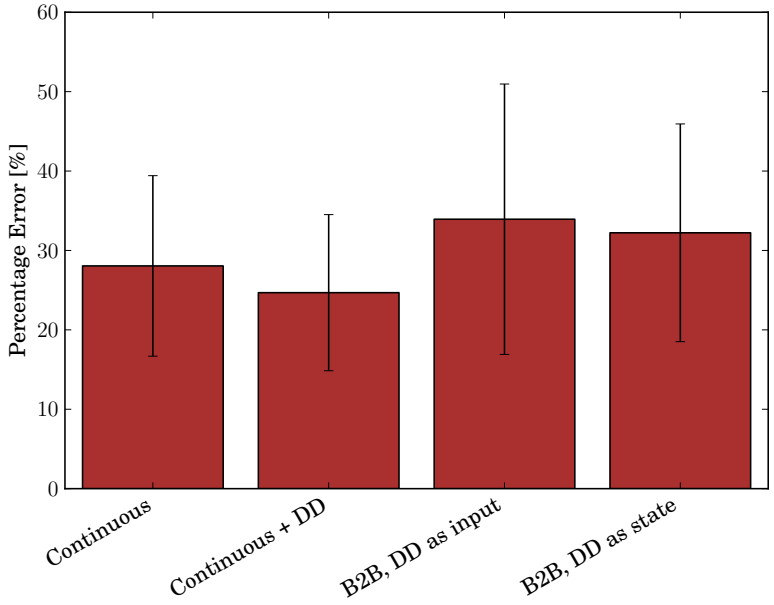

Figure 3.22: Comparison in terms of percentage error of all CNAP methods according to the critchley criterion. All approaches are compared with the Vigileo/FloTrac device as ground truth. Each bar represents an experiment in ascending order (1 through 4). *"Continuous"* refers to experiment 1, *"Continuous + DD"* refers to experiment 2, *"B2B, DD as input"* refers to experiment 3 and *"B2B, DD as state"* refers to experiment 4

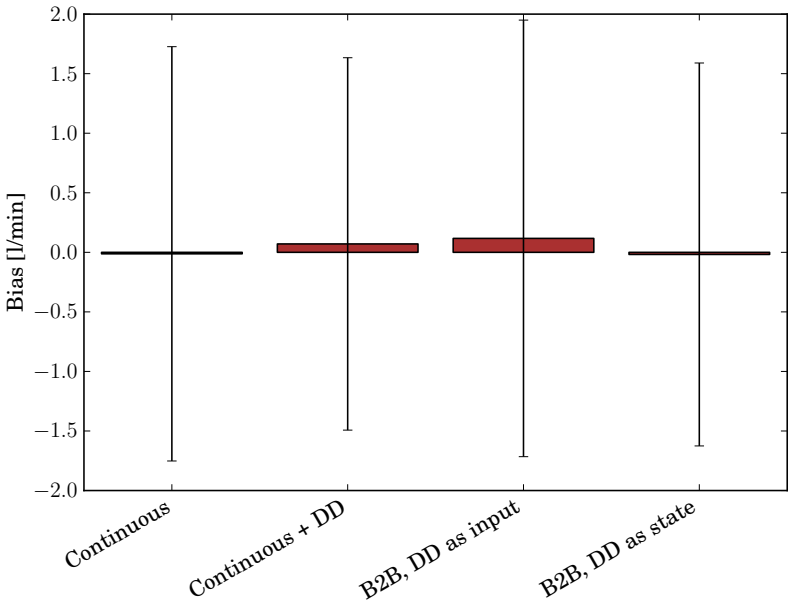

Figure 3.23: Comparison in terms of bias of all CNAP methods according to the critchley criterion. All approaches are compared with the Vigileo/FloTrac device as ground truth. Each bar represents an experiment in ascending order (1 through 4). *"Continuous"* refers to experiment 1, *"Continuous + DD"* refers to experiment 2, *"B2B, DD as input"* refers to experiment 3 and *"B2B, DD as state"* refers to experiment 4

	PE (%)		Bias (l/min)		RMSE (l/min)	
	mean	SD	mean	SD	mean	SD
Continuous	28.05	11.37	-0.01	1.74	1.78	0.94
Continuous + DD	24.68	9.83	0.07	1.56	1.54	0.93
B2B, DD as Input	33.93	17.03	0.12	1.83	1.83	1.17
B2B, DD as State	32.22	13.71	-0.02	1.61	1.80	0.97

Table 3.1: Performance comparison of all approaches utilizing the CNAP/FloTrac database. All approaches are compared with the Vigileo/FloTrac device as ground truth. Each line represents an experiment in ascending order (1 through 4). *"Continuous"* refers to experiment 1, *"Continuous + DD"* refers to experiment 2, *"B2B, DD as input"* refers to experiment 3 and *"B2B, DD as state"* refers to experiment 4

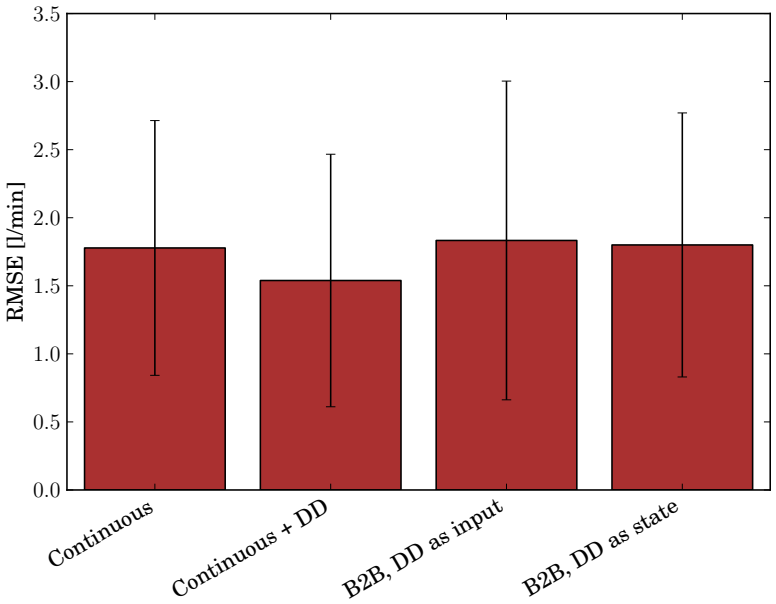

Figure 3.24: Comparison of RMSE of all CNAP methods. All approaches are compared with the Vigileo/FloTrac device as ground truth. Each bar represents an experiment in ascending order (1 through 4). *"Continuous"* refers to experiment 1, *"Continuous + DD"* refers to experiment 2, *"B2B, DD as input"* refers to experiment 3 and *"B2B, DD as state"* refers to experiment 4

3.2 MIMIC II (Experiments 5 & 6)

The CO measurements in the MIMIC II database are substantially different as compared to the CNAP/FloTrac database. While the FloTrac sensor produces a new measurement every 20 seconds, the gap between CO measurements in the MIMIC II database are often hours. Therefore the following plots represent rather CO trends than signals.

As described in chapter 2.2.2 the nature of the MIMIC database leads to a couple of consequences. For instance the reservoir is only simulated a view minutes prior to a measurement, enough to satisfy the echo state property.

Experiments for two approaches were conducted. The first approach (experiment 5) is the straight forward model as described in chapter 2.3.1. In contrast to the previous experiments only the ABP signal is present in the MIMIC II records, and of course also no CNAP controller signals. The whole model is based solely on the ABP signal.

The second approach (experiment 6) employs an additional reservoir to capture the fast dynamics of the pulsatile waveforms as described in chapter 2.3.3. The slow reservoir was tuned to a time-constant around one minute. The fast reservoir was tuned to a time-constant around one second.

Also there are no BA-diagrams given in this chapter since they are of little interest for only a few samples per records.

The following pages contain results of 16 patient records (note that two different records might not be two different patients). These 16 records were chosen to represent a meaningful subset of all patients. Each plot shows the CO targets, measured with a reference method (usually thermodilution) along with the results of both experiments.

The plots show the ability of the estimator to follow trends in the CO as well es bias in the results.

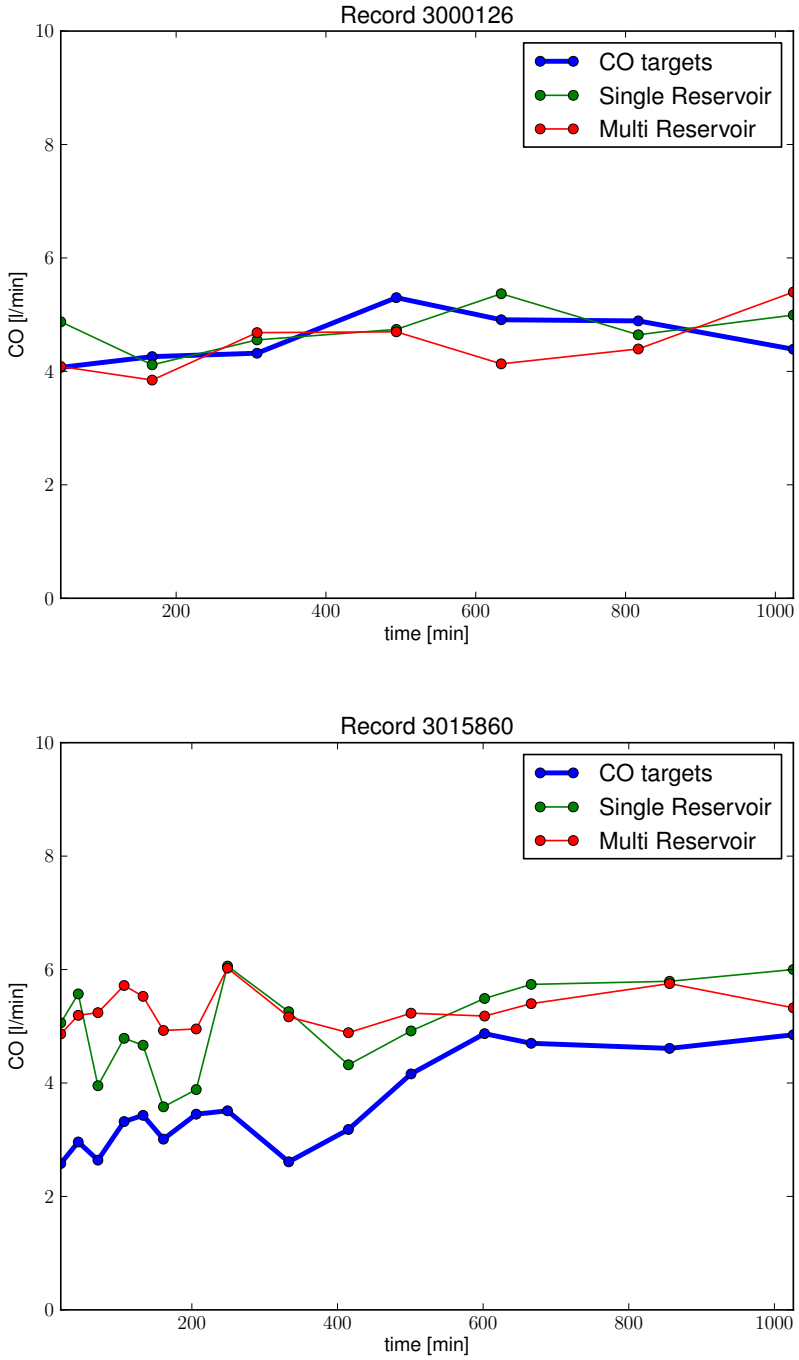

Figure 3.25: Records 3000126 and 3015860: The blue graphs are the target samples measured by a reference method (either thermodilution or echocardiography). The green graphs are the estimation by an ordinary single reservoir ESN (experiment 5). The red graphs represent the estimation of an ESN with an additional faster reservoir to capture characteristics in the fast dynamics (experiment 6).

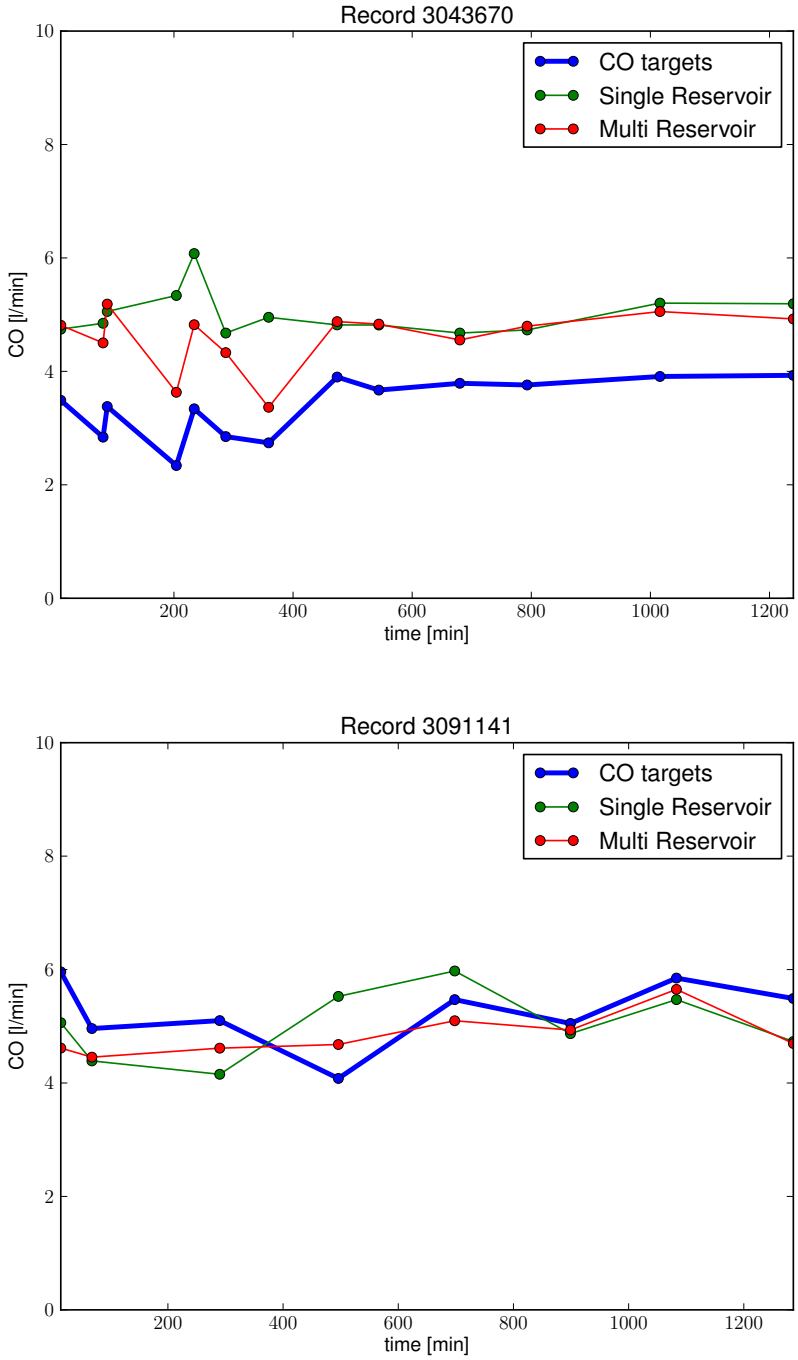

Figure 3.26: Records 3043670 and 3091141: The blue graphs are the target samples measured by a reference method (either thermodilution or echocardiography). The green graphs are the estimation by an ordinary single reservoir ESN (experiment 5). The red graphs represent the estimation of an ESN with an additional faster reservoir to capture characteristics in the fast dynamics (experiment 6).

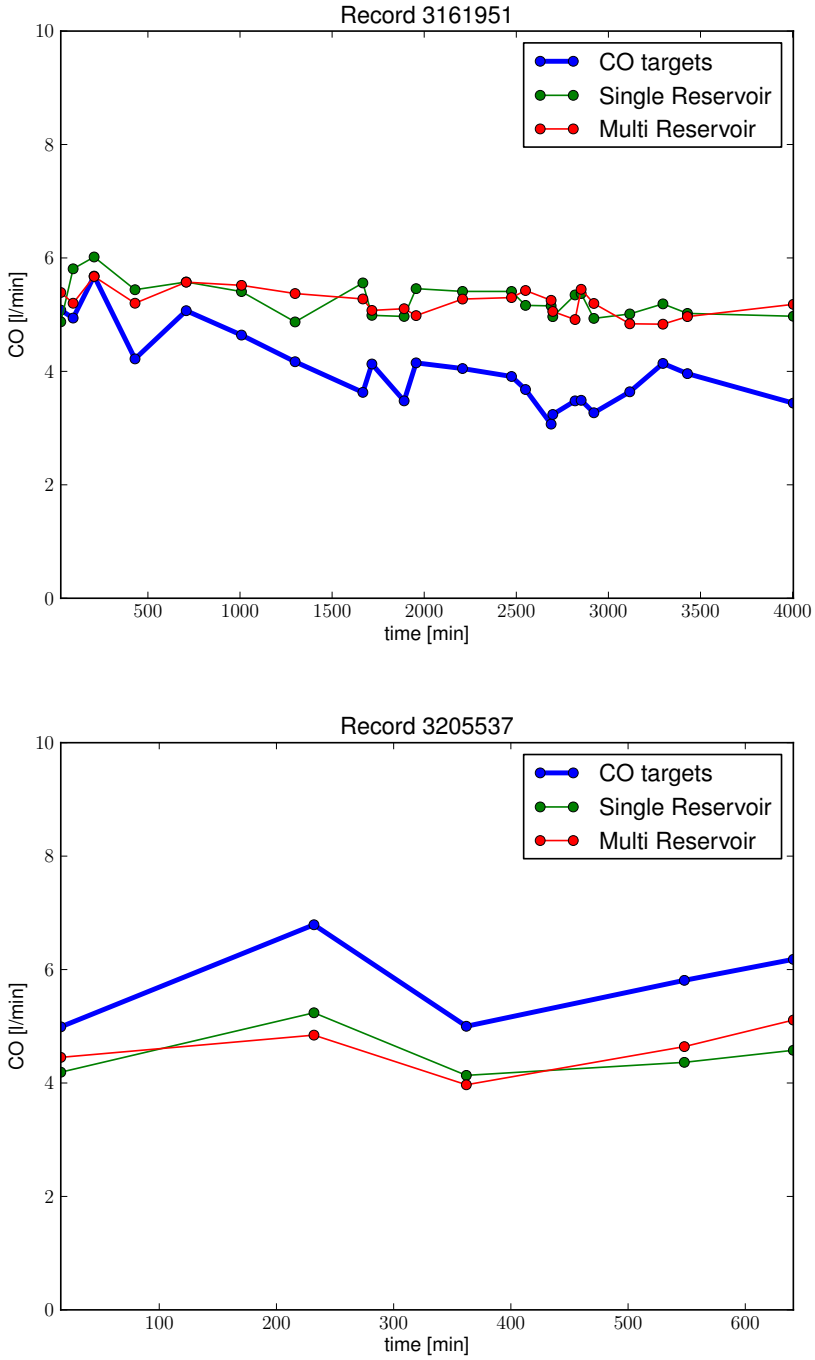

Figure 3.27: Records 3161951 and 3205537: The blue graphs are the target samples measured by a reference method (either thermodilution or echocardiography). The green graphs are the estimation by an ordinary single reservoir ESN (experiment 5). The red graphs represent the estimation of an ESN with an additional faster reservoir to capture characteristics in the fast dynamics (experiment 6).

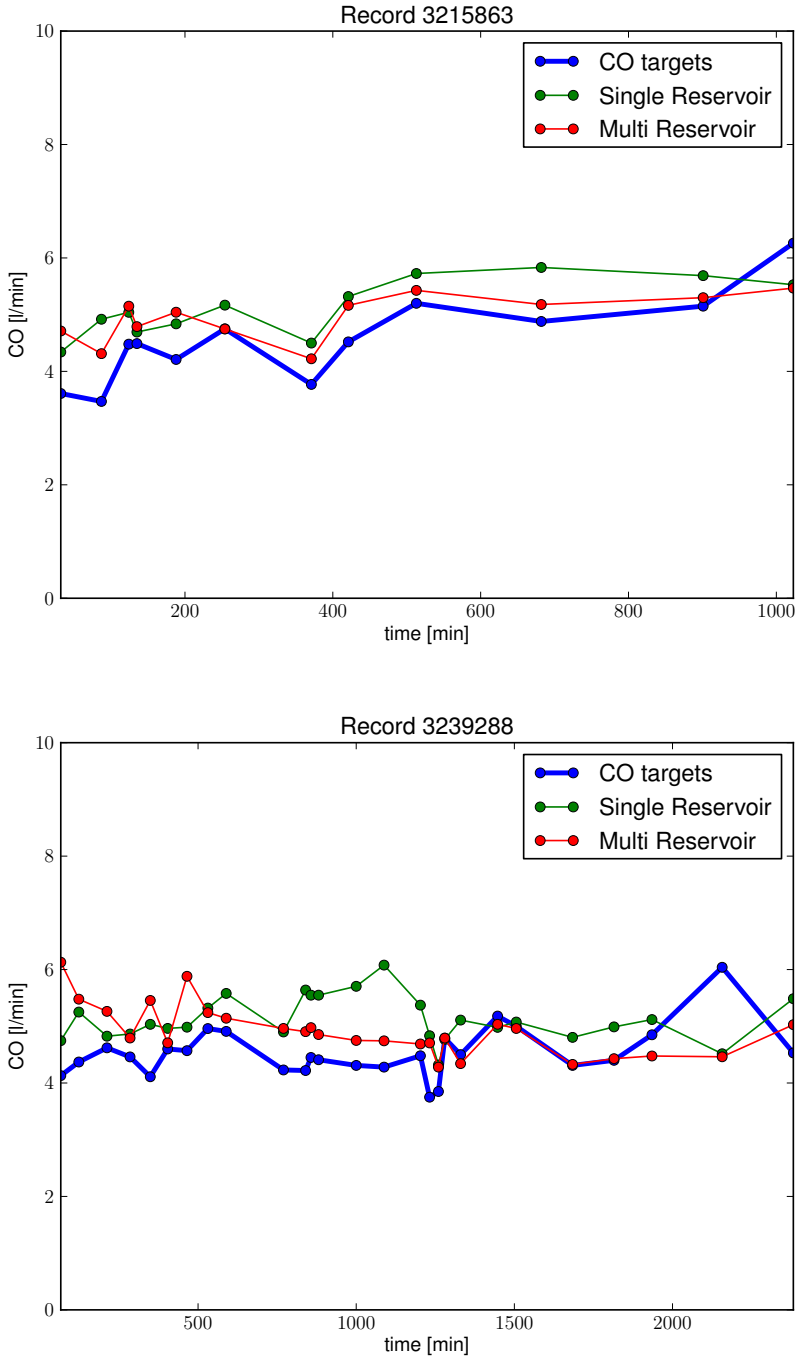

Figure 3.28: Records 3215863 and 3239288: The blue graphs are the target samples measured by a reference method (either thermodilution or echocardiography). The green graphs are the estimation by an ordinary single reservoir ESN (experiment 5). The red graphs represent the estimation of an ESN with an additional faster reservoir to capture characteristics in the fast dynamics (experiment 6).

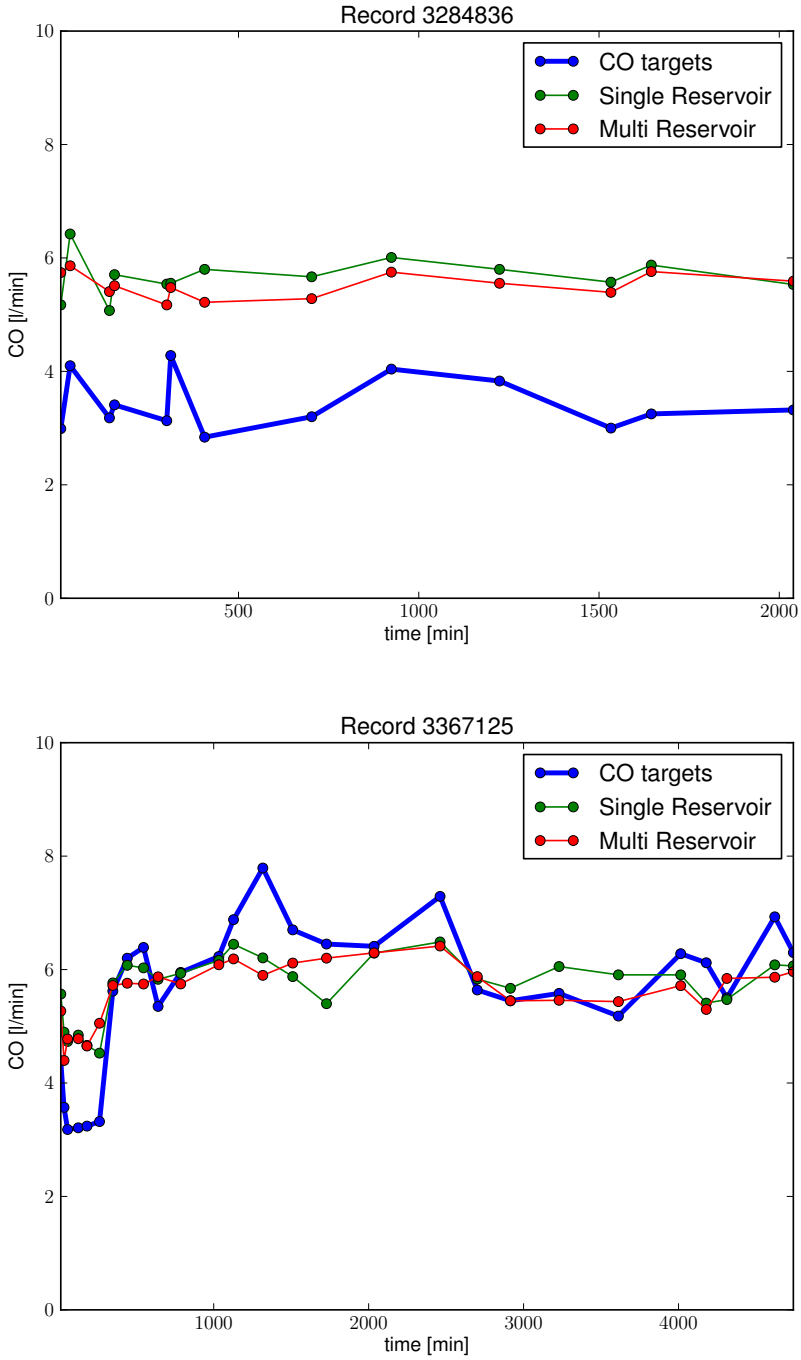

Figure 3.29: Records 3284836 and 3367125: The blue graphs are the target samples measured by a reference method (either thermodilution or echocardiography). The green graphs are the estimation by an ordinary single reservoir ESN (experiment 5). The red graphs represent the estimation of an ESN with an additional faster reservoir to capture characteristics in the fast dynamics (experiment 6).

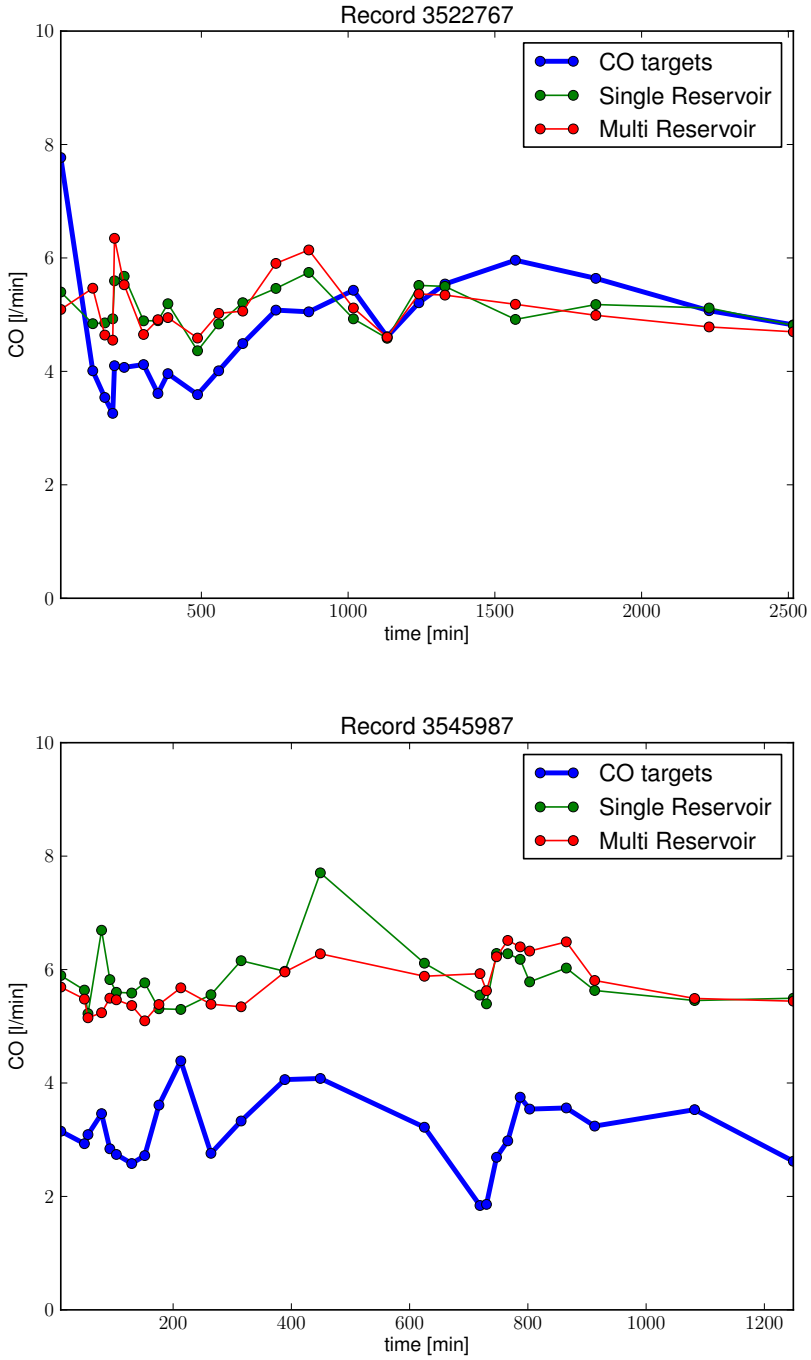

Figure 3.30: Records 3522767 and 3545987: The blue graphs are the target samples measured by a reference method (either thermodilution or echocardiography). The green graphs are the estimation by an ordinary single reservoir ESN (experiment 5). The red graphs represent the estimation of an ESN with an additional faster reservoir to capture characteristics in the fast dynamics (experiment 6).

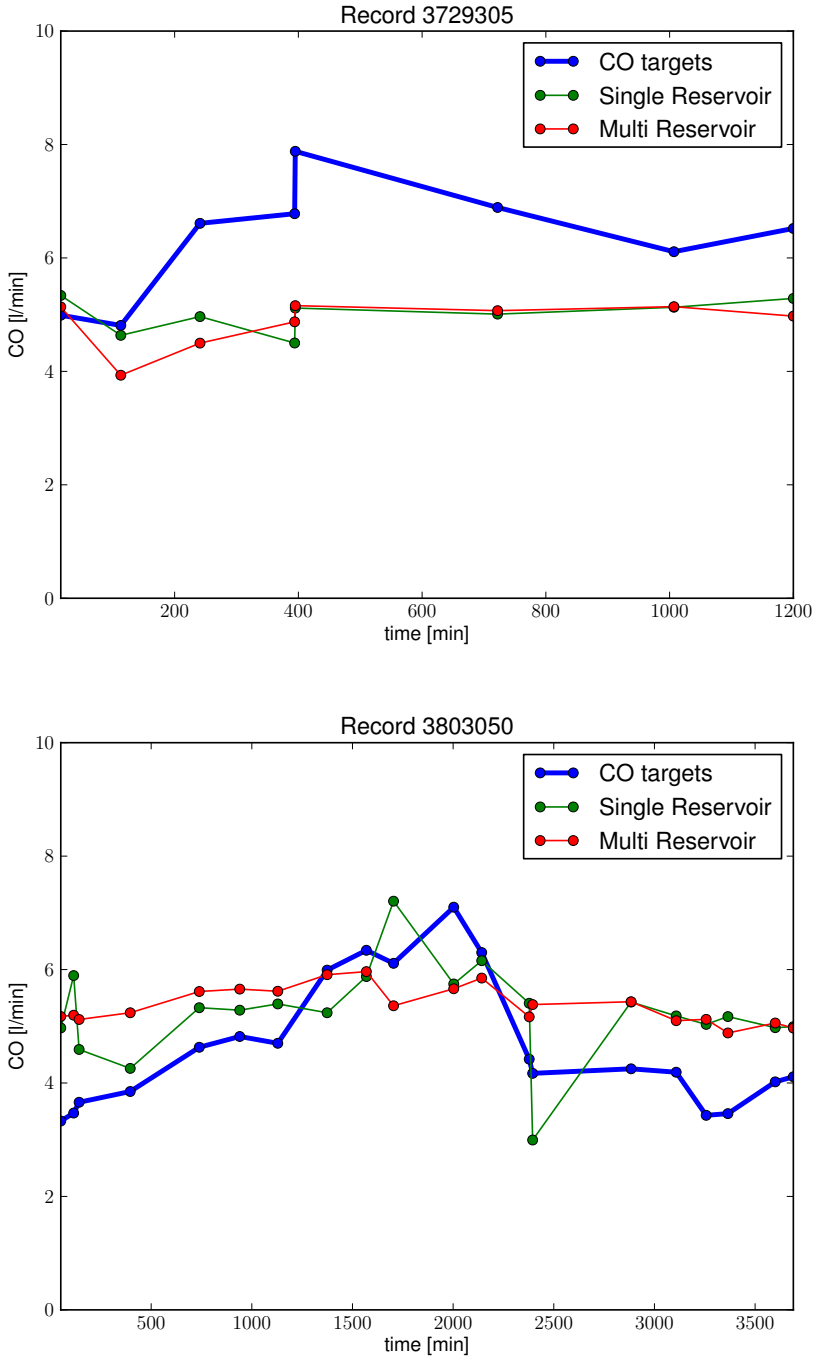

Figure 3.31: Records 3729305 and 3803050: The blue graphs are the target samples measured by a reference method (either thermodilution or echocardiography). The green graphs are the estimation by an ordinary single reservoir ESN (experiment 5). The red graphs represent the estimation of an ESN with an additional faster reservoir to capture characteristics in the fast dynamics (experiment 6).

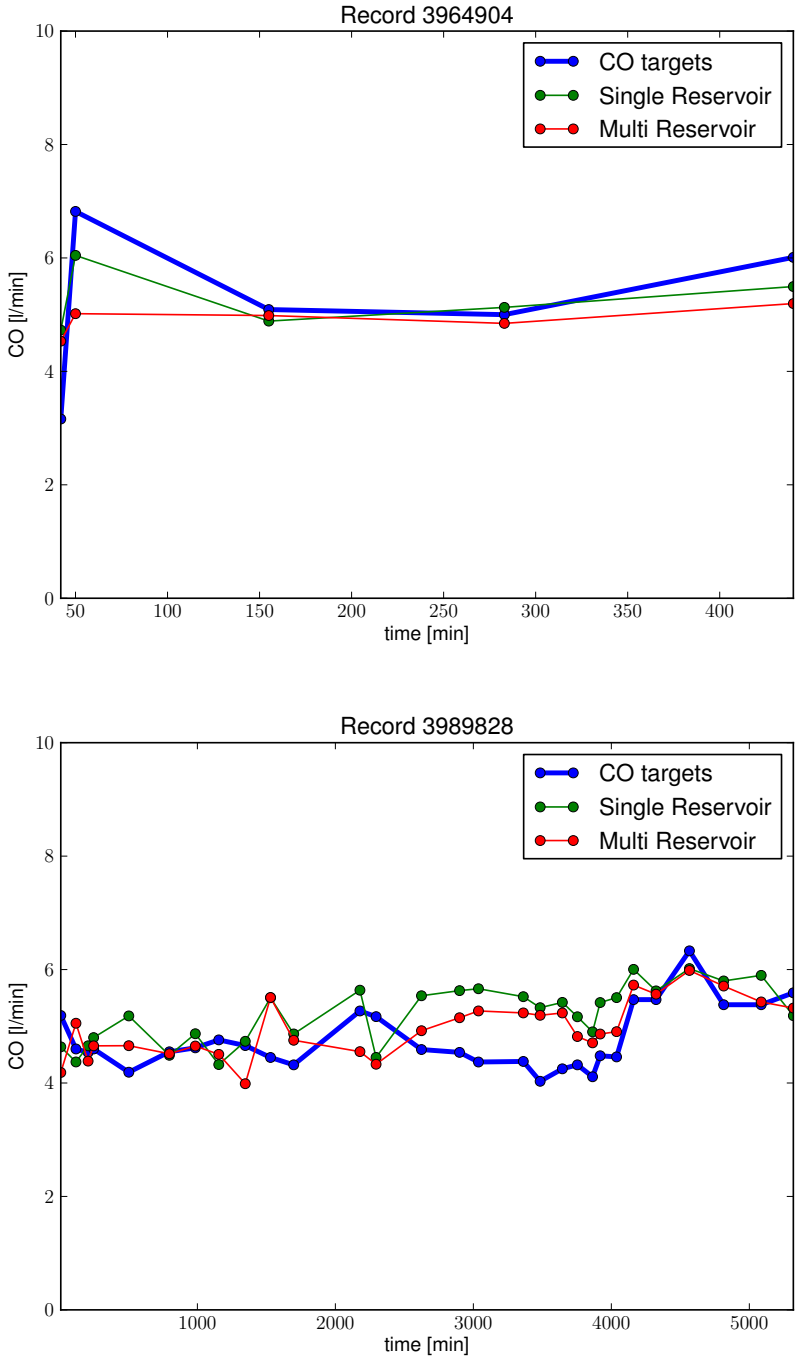

Figure 3.32: Records 3964904 and 3989828: The blue graphs are the target samples measured by a reference method (either thermodilution or echocardiography). The green graphs are the estimation by an ordinary single reservoir ESN (experiment 5). The red graphs represent the estimation of an ESN with an additional faster reservoir to capture characteristics in the fast dynamics (experiment 6).

Bias

As with the previous experiments the estimations are conflicted with bias in several cases. This bias could possibly be reduced if descriptive data is added as virtual neurons to the ESNs. This measure has shown to improve performance in previous experiments (2 & 4). In some cases the bias can be explained with outliers in the reference method as in record 3522767 (figure 3.30) or 3367125 (figure 3.29).

Tracking

The tracking ability is the ability to follow trends in CO. Important are the direction of the trend as well as the amount of change. In terms of direction the results improve in the dual reservoir approach. Prime examples for that are recording 3043670 (figure 3.26) or the two recordings in figure 3.28.

One big uncertainty in the assessment of the tracking behavior is the unknown precision of the reference method. Thermodilution for instance is known to have a rather poor precision.

3.2.1 Comparison of MIMIC II Results

The statistics of error measurements are assessed using the quality criteria from chapter 2.4. None of the methods could be tested including descriptive patient data since none of it is available in the records of the MIMIC II database at the time of writing. The purpose of these experiments is to evaluate the performance of the ESN method against clinical data with established CO measurement methods and the improvement of this method when including an additional fast reservoir.

Depicted in figure 3.33 is a comparison in terms of percentage error as defined by Critchley. Experiment 6 is able to reduce the average PE by 3% and the standard deviation by 5.8% over experiment 5. This is due to the improved tracking abilities of the dual reservoir approach.

The bias plot in figure 3.34 suggests almost no improvement of the bias behavior in experiment 6 over experiment 5. Further both methods are constantly over-estimating the cardiac output. A possible explanation could be positive outliers in the target values like those seen in figure 3.29 or figure 3.30.

Similar results to the PE but with less improvement can be seen in figure 3.35 comparing the RMSE. The RMSE represents the total error in respect to shape and bias. Since the improvement of bias is small it also leads to a smaller improvement in the RMSE results.

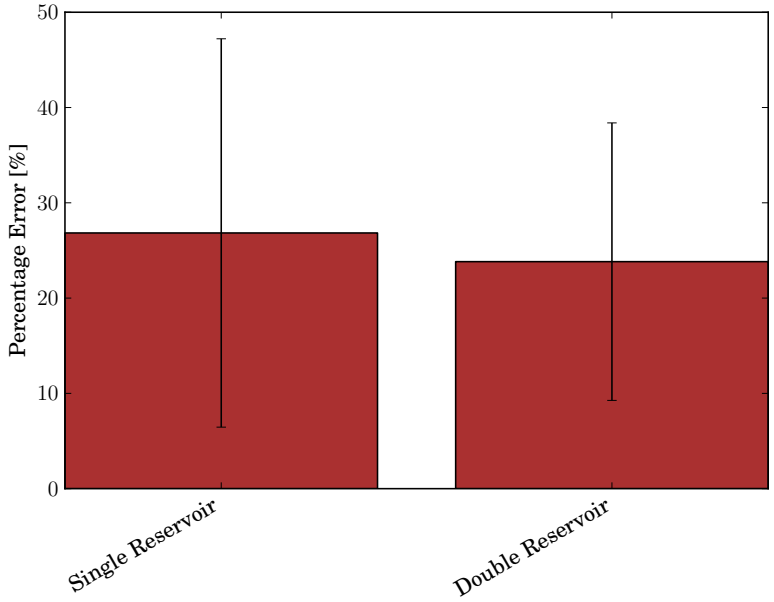

Figure 3.33: Comparison in terms of percentage error according to the critchley criterion of all experiments with the MIMIC DB as basis. All approaches are compared with the CO measurements in the MIMIC DB as ground truth. Each bar represents one experiment. *"Single Reservoir"* refers to experiment 5, *"Double Reservoir"* refers to experiment 6

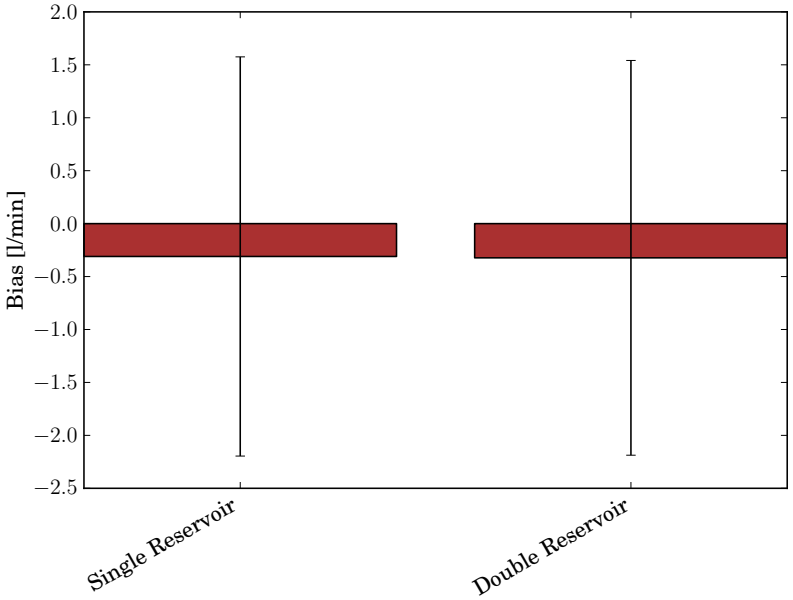

Figure 3.34: Comparison in terms of bias according to the critchley criterion of all experiments with the MIMIC DB as basis. All approaches are compared with the CO measurements in the MIMIC DB as ground truth. Each bar represents one experiment. *"Single Reservoir"* refers to experiment 5, *"Double Reservoir"* refers to experiment 6

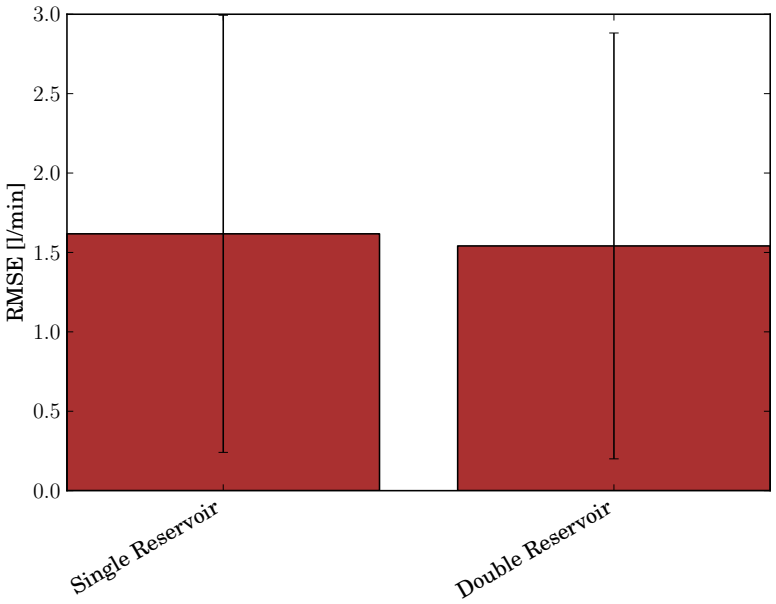

Figure 3.35: Comparison in terms of RMSE of all experiments with the MIMIC DB as basis. All approaches are compared with the CO measurements in the MIMIC DB as ground truth. Each bar represents one experiment. *"Single Reservoir"* refers to experiment 5, *"Double Reservoir"* refers to experiment 6

	PE (%)		Bias (l/min)		RMSE (l/min)	
	mean	SD	mean	SD	mean	SD
Single Reservoir	26.83	20.38	-0.31	1.89	1.62	1.38
Double Reservoir	23.82	14.56	-0.32	1.86	1.54	1.34

Table 3.2: Performance comparison of all approaches utilizing the MIMIC database. All approaches are compared with the CO measurements in the MIMIC DB as ground truth. Each line represents one experiment. *"Single Reservoir"* refers to experiment 5, *"Double Reservoir"* refers to experiment 6

4 Discussion

In the present thesis, six different variants were presented of applying echo state networks to the task of learning a non-linear model in order to estimate cardiac output from peripheral arterial blood pressure. Four of these variants were applied on the proprietary CNAP/FloTrac database, the other two were applied on the open MIMIC II waveforms database.

All approaches were compared among each other using statistical observation of three error measures (percentage error, bias, RMSE). Exemplary plots of blood pressure, targets and estimations were given for several patients.

4.1 Comparison with other CO-from-ABP Methods

CO-from-ABP methods made it possible for the first time to monitor CO on a continuous basis. Several devices incorporating CO-from-ABP emerged in the last years. In a meta-analysis [5] by de Wilde *et al.*, five different devices where compared against thermodilution.

The study includes the cZ method by Wesseling, the PiCCO system by Philips, the Hemac method, the Modelflow method and the system by LiDCO. The results in the study are presented somewhat differently but from the numbers one can derive bias and percentage error as it is used in the present work. Table 4.1 shows the results in terms of bias and percentage error of these 5 methods compared with the results of the ESN approach against the MIMIC II database.

	Bias (l/min)	PE (%)
ESN	-0.32	23.82
cZ	0.23	21.41
PiCCO	0.14	36.28
Hemac	0.06	18.09
Modelflow	0	15.38
LiDCO	-0.17	28.59

Table 4.1: A comparison of different CO-from-ABP methods against thermodilution. The new ESN approach is placed in the center of these methods. As the results in the reference method were all calibrated, they have a clear advantage in terms of bias.

The comparison shows that the new approach using ESNs to estimate CO is able to compete with current [commercially available] systems. The ESN approach is placed in the center of these methods. In the study by de Wilde *et al.* all methods were calibrated using four successive thermodiluation measurements. This gives them a great advantage in terms of bias.

The best performing methods Hemac and Modelflow are both based on non-linear Windkessel models demonstrating the power of simple models. With the increasing amount of open clinical data the ESN approach is expected to improve as it relies on training data.

4.2 Limits of the Critchley Criterion

Patient monitors in general are tools for patient monitoring in the ICU and for physicians to decide over a patients therapy. They are not meant to make decision themselves which would be far beyond the abilities of patient monitors in a lot of cases. Therefore a physician should be able to get the vital information out of a measured signal that helps him reasoning. The optimal quality assessment for patient monitors would be to evaluate their ability of emphasizing this information.

The percentage error as defined by Critchley estimates the expected limits of agreement for which 95% of the measurements lie within those limits. This quantity is based solely on the standard deviation of the error, which leads to a low percentage error if shape and scale of the compared signals are similar. In practice there are patient records where the model simply doesn't perform well. The reason could be too little training data (overfitting), artifacts in the signals or simply underrepresentation of the target effects in the given input signals. In those cases the Critchley criterion usually distorts the results as a less complex model automatically leads to a better percentage error.

Figure 4.1 shows the application of a simple and a complex model. While the simple model leads to underfitting the complex model does not. The complex model shows a rather good performance in terms of signal shape. The simple model is not able to represent the trends in the targets as good as the complex model.

The plot in figure 4.2 shows something different. The complex model performs bad. This is due to the bad quality of the targets. The cardiac output jumping rapidly within tens of seconds is physiologically implausible. But while the complex model doesn't show any correlation with the targets the results of the simple model are helped by underfitting. Basically they are just less wrong and as a result the percentage error is lower.

Nevertheless, the percentage error as defined by Critchley is the accepted way to compare this kind of studies. The RMSE for instance suffers from exactly the same shortcomings when it comes down to assessment of particular patients. Thorough evaluation of a new method must therefore always include practical testing to assess its usefulness in the field.

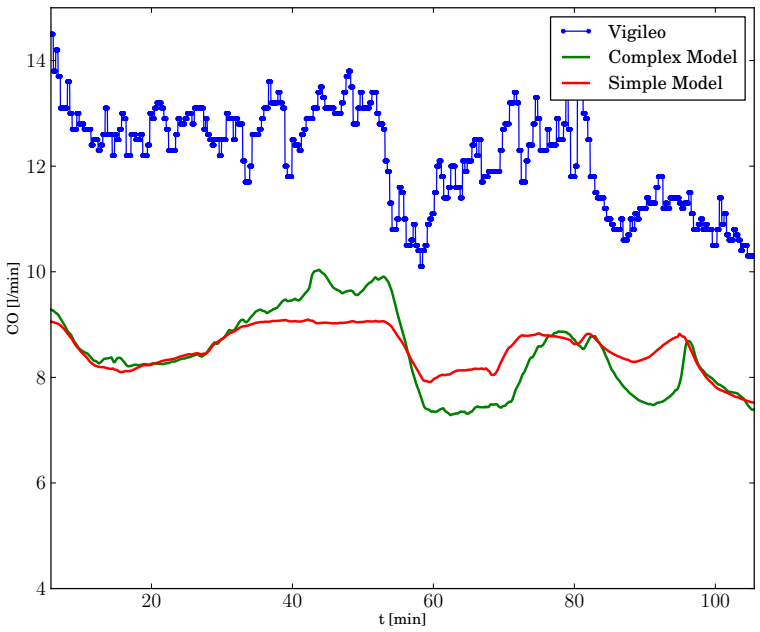

Figure 4.1: Application of a simple (red graph) and a complex (green graph) model on a high quality target signal (blue graph). The plot shows the cross-validation results of patient 07 from the CNAP/FloTrac database. While the complex model is able to express the trends in the target signal, the simple model suffers from underfitting.

4.3 Limits of the Clinical Data

The application of sophisticated machine-learning methods in medical care is a novelty. An important reason for the lack of these methods in medical care is the cost of training data. In case of medical monitoring training data is obtained through medical trials in hospitals which are usually complicated to conduct and extremely expensive. Projects like the MIMIC II medical database aim to improve the availability of such data.

With the constantly increasing amount of freely available patient records machine-learning methods in medical care will probably be seen more frequently.

4.3.1 Precision of the Reference Method

A direct measurement of CO would require a flow-probe inserted into the aorta. Although sometimes done in trials with animals[1], this comes with an unacceptable medical risk for humans. Thus, all methods that can be used as reference, respectively training

[1]The author of this thesis and everybody who supported it rejects this kind of trials with animals as unethical and unnecessary cruelty to animals.

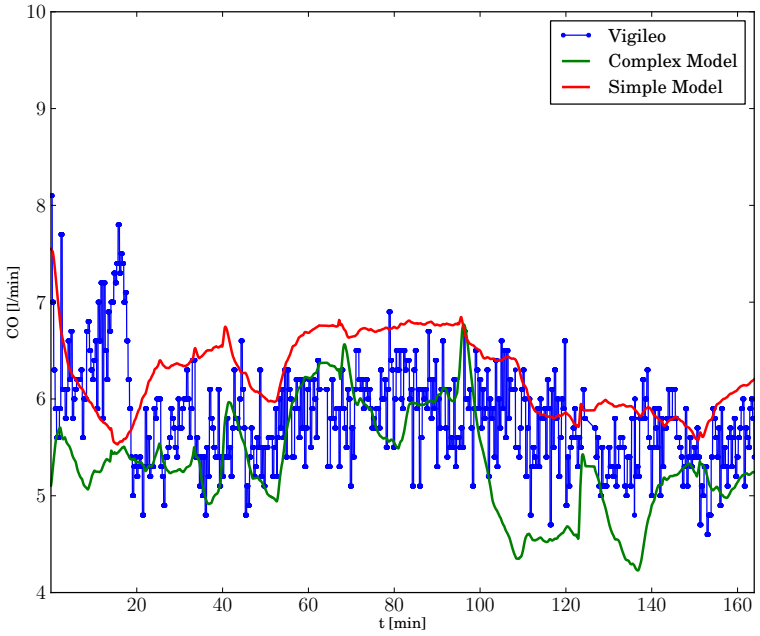

Figure 4.2: Application of a simple (red graph) and a complex (green graph) model on a poor quality target signal (blue graph). The plot shows the cross-validation results of patient 14 from the CNAP/FloTrac database. The simple model, which in general suffers from underfitting, performs better than the complex model on the poor quality recording.

data, are indirect measurements or estimations. These methods can be conflicted with poor precision which in turn leads to poor quality training data.

MIMIC II Database

It is known that the MIMIC II database contains CO measurements obtained with the methods of *echocardiography* as well as *thermodilution*. Both methods are known to be accurate ($10 - 20\,\%$ percentage error) but rather imprecise (meaning the variance of multiple successive measurements). We can thus safely assume the average percentage error to be $20\,\%$. With the requirement of the total percentage error to be below $30\,\%$ one can calculate the needed percentage error between the new method and reference method (PE_{diff}) using the relation (2.27) from chapter 2.4.2:

$$\mathrm{PE}_{diff} = \sqrt{30^2 - 20^2} = 22.36\,\% \tag{4.1}$$

The best percentage error that was achieved was $23.82\,\%$ which can be considered close enough due to the safe assumption of the percentage error of the reference method.

CNAP/FloTrac Database

In this case, the reference method for cardiac output is the FloTrac sensor by Edwards Lifesciences LLC. In the literature this method achieves different performance results ranging from 24.6 % to 52 % error between Vigileo/FloTrac and thermodilution CO [1, 23, 25, 31]. One can conclude that the total percentage error of any new method based on this reference has a bigger total error.

The FloTrac data is based on entirely independent measurements (i.e. not the CNAP measurements but invasive arterial lines). Those two signals are representing pressures at different locations. Also it can be problematic in terms of artifacts that are showing in CNAP recordings but not in invasive recordings and vice versa. Estimating the same quantity using two entirely different measurements is especially hard.

The CNAP sensor employs a set point tracking algorithm as described in chapter 1.2.3. Erroneous parameter drift can occur in special cases. After not later than 30 minutes into the measurement sequence (*stint*) a recalibration takes place. The parameter drift then leads to a jump in the ABP signals. An extreme example of this phenomenon can be seen in figures 3.9 and 3.19. Drifts and jumps of this magnitude render an accurate CO estimation almost impossible.

4.4 Limits of the Model

Echo state networks are known to be *universal approximators* for non-linear dynamic systems (see chapter 2.1). As a consequence ESNs are best suited to process temporal data (temporal in terms of a constantly changing parameter such as time). Several arguments would suggest that an ESN is a good choice for the task at hand:

- The circulatory system is a temporal dynamic system
- There are no long-term relations between input signals and targets
- The problem is [highly] non-linear
- ESNs can capture the complexity of the problem

On the other hand ESNs have a number of disadvantages that come with the simulation of the reservoir:

- The computational efforts to simulate an ESN are considerably higher than for other models (like fitted linear models).
- It is not possible, without further ado, to simulate a complex reservoir on the limited hardware of current patient monitors.
- Massive amounts of medical data must be maintained for the training of an ESN.

Echo state networks are known to require a lot of training samples. Facing the increasing amount of available records in the MIMIC II database performance results may improve in future attempts. Also hardware performance of patient monitors will increase in the future enabling them to simulate ESNs online.

5 Summery & Conclusion

A novel approach was presented to estimate cardiac output from peripheral arterial blood pressure using echo state networks. Two databases were used to evaluate this approach:

- CNAP/FloTrac database
- MIMIC II medical waveform database

The CNAP/Flotrac database consists of over 60 records containing a non-invasively measured arterial pressure signal and a continuous cardiac output target signal. The arterial pressure signal was measured using a CNAP sensor by CNSystems Medizintechnik AG . The cardiac output signal was measured using a FloTrac sensor by Edwards Lifesciences LLC. The latter is a CO-from-ABP method utilizing an invasively measured arterial pressure.

The MIMIC II database is a vast medical database containing close to 400 patients records with arterial blood pressure (ABP) signals and cardiac output (CO) measurements. In general the pressure signals were taken using an arterial catheter. The cardiac output measurements were taken using echocardiography or thermodiluation, both well established standard methods to measure cardiac output. In contrast to the CNAP/FloTrac database, the cardiac output measurements are not continuous but rather single measurements with irregular gaps of minutes up to hours in between.

Six different computational experiments were conducted. Four of which using the CNAP/FloTrac database:

- Normalized continuous CNAP signals as inputs and an additional mean arterial pressure signal.
- The first experiment but with descriptive patient data as virtual reservoir states.
- Beat-to-beat features as signal preprocessing. Features from all beat-periodic signals were used for this experiment.
- The beat-to-beat features experiment but again including descriptive data

Two of the experiments were conducted using the MIMIC II database:

- Normalized arterial blood pressure and mean arterial pressure as input signals.
- The first experiment but with two reservoirs tuned to different bands of the signal spectrum.

The feature selection process in the two experiments that employ beat-to-beat features was done using a sequential forward/backward feature selection algorithm.

It was shown that this machine-learning approach is able to match the performance of commercially available application of CO-from-ABP methods.

With a performance of 23.82 % error according to the Critchley criterion on the MIMIC II database it is a valid alternative to invasive cardiac output measurement methods.

With 24.68 % error between the FloTrac sensor and the new method it has been shown that this approach is able to perform the same task upon a CNAP signal which is obtained with a vascular unloading technique.

In two experiments descriptive patient data (i.e. age, height, weight, sex) was included in the model and lead to a slightly better performance. The reduction of bias in the error due to the descriptive data was minimal. In general bias in the error turned out to be very difficult to reduce. Fortunately this bias is not essential in the assessment of a cardiac output.

All experiments based in beat-to-beat features performed worse than their continuous counterparts.

6 Future Work

Several experiments and improvements are planned to increase the performance and the quality of the evaluation. These improvements are not subject of the present work.

- If an echo state network is driven with beat-to-beat features of blood pressure signals (as it was in two of the experiments in the present work) each reservoir iteration does not exactly represent a constant time-interval. Since each stroke delivers a new pulse-wave it is iterated with the heart-rate. The heart-rate is subject to change. This is mathematically equivalent to having a time-varying system. A possible solution to come by this is to align the reservoir iterations on a fixed time-interval and interpolate the features between heart-beats.
- All MIMIC II-based experiments applied life-sign signals directly without preprocessing. Further experiments should be conducted based on beat-to-beat signal features.
- The basic structure of an echo state network consists of two stages. The first stage is a constant dynamic non-linear system. The second stage is an optimized static linear system. In a future experiment the order of linear and non-linear stage could be swapped in the following way: A constant dynamic linear system in the first stage and an optimized static non-linear system in the second stage. A possible setting could be the following: Beat-to-beat features are filtered using a linear filter (low-pass in the simplest case). The second stage is a non-linear regression method (e.g. neural networks or support vector regression).
- While the Critchley criterion is currently established to compare different cardiac output estimation methods it has been under a lot of criticism due to it's lack of capturing the ability to follow significant trends in continuous CO estimations. It is commonly accepted now that the bias in a CO estimation is of little relevance to the physician. What is important is the ability to follow trends relative to the true cardiac output. Thus Critchley published a new method [4] to evaluate CO estimation based *only* in the trend, its direction and value. Evaluating the results on the present thesis using the new criterion is subject of future work.
- A new clinical database is expected in the near future including ABP measurements of a CNAP monitor along with CO estimations performed by a PiCCO-Plus monitor by Pulsion Medical Systems. This monitor is calibrated to a thermodilution reference measurement and known to perform better than the Vigileo/FloTrac system by Edwards Lifesciences LLC.

- Another interesting experiment is to train a model on the MIMIC II database and evaluate it on a CNAP-based database. This way one can assess the estimation error that arises from the artifacts of the non-invasive ABP measurement.

Bibliography

[1] G Biancofiore, LAH Critchley, and A Lee. Evaluation of a new software version of the FloTrac/Vigileo (version 3.02) and a comparison with previous data in cirrhotic patients undergoing liver transplant surgery. *Anesthesia & ...*, 2011.

[2] J N Cohn, S Finkelstein, G McVeigh, D Morgan, L LeMay, J Robinson, and J Mock. Noninvasive pulse wave analysis for the early detection of vascular disease. *Hypertension*, 26(3):503–508, 1995.

[3] L a Critchley and J a Critchley. A meta-analysis of studies using bias and precision statistics to compare cardiac output measurement techniques. *Journal of clinical monitoring and computing*, 15(2):85–91, February 1999.

[4] Lester a Critchley, Xiao X Yang, and Anna Lee. Assessment of trending ability of cardiac output monitors by polar plot methodology. *Journal of cardiothoracic and vascular anesthesia*, 25(3):536–46, June 2011.

[5] R B P de Wilde, J J Schreuder, P C M van den Berg, and J R C Jansen. An evaluation of cardiac output by five arterial pulse contour techniques during cardiac surgery. *Anaesthesia*, 62(8):760–8, August 2007.

[6] Theo Dingermann. Anatomie, Physiologie, Pathophysiologie des Menschen. Von Ernst Mutschler, Hans-Georg Schaible und Peter Vaupel. *Pharmazie In Unserer Zeit*, 36(5):406–407, 2007.

[7] P Dow. Estimations of cardiac output and central blood volume by dye dilution. *Physiological Reviews*, 36(1):77–102, 1956.

[8] J Erlanger and D R Hooker. An experimental study of blood pressure and of pulse pressure in man. *Johns Hopkins Hosp Rep*, 12(145-378):145–378, 1904.

[9] J Fortin, W Marte, R Grüllenberger, A Hacker, W Habenbacher, A Heller, Ch Wagner, P Wach, and F Skrabal. Continuous non-invasive blood pressure monitoring using concentrically interlocking control loops. *Computers in Biology and Medicine*, 36(9):941–957, 2006.

[10] W Ganz, R Donoso, H S Marcus, J S Forrester, and H J Swan. A new technique for measurement of cardiac output by thermodilution in man. *The American Journal of Cardiology*, 27(4):392–396, 1971.

[11] Roger M Goldwyn and Thomas B Watt. Arterial Pressure Pulse Contour Analysis Via a Mathematical Model for the Clinical Quantification of Human Vascular Properties, 1967.

[12] Peter Chris Harpas and Ahmad M. Qasem. A Method of Estimating Pulse Wave Velocety, 2007.

[13] A Heller. *A Fast Parameter Estimation Method for Pulse Contour Analysis*. Master thesis, Technical University of Graz, 2000.

[14] Georg Holzmann and Helmut Hauser. Echo state networks with filter neurons and a delay&sum readout. *Neural Networks*, 23(2):244–256, 2010.

[15] Y Ichimaru and G B Moody. Development of the polysomnographic database on CD-ROM. *Psychiatry and Clinical Neurosciences*, 53(2):175–177, 1999.

[16] H Ihlen, J P Amlie, J Dale, K Forfang, S Nitter-Hauge, J E Otterstad, S Simonsen, and E Myhre. Determination of cardiac output by Doppler echocardiography. *British heart journal*, 51(1):54–60, January 1984.

[17] Herbert Jaeger. The " echo state " approach to analysing and training recurrent neural networks. pages 1–47, 2001.

[18] Herbert Jaeger, Mantas Lukosevicius, Dan Popovici, and Udo Siewert. Optimization and applications of echo state networks with leaky-integrator neurons. *Neural networks : the official journal of the International Neural Network Society*, 20(3):335–52, April 2007.

[19] Li-wei Lehman, George Moody, Thomas Heldt, and Tin H Kyaw. Multiparameter intelligent monitoring in intensive care II (MIMIC-II): A public-access ICU database. *Critical Care*, 39(February 2010):952–960, 2011.

[20] G Liljestrand and E Zander. Vergleich der Bestimmung des Minutenvolumens des Herzens beim Menschen mittels der Stichoxydulmethode und durch Blutdruckmessung. *Zeitschrift Für Die Gesamte Experimentelle Medizin*, 59(1):105–122, 1928.

[21] Mantas Lukoševičius and Herbert Jaeger. Reservoir computing approaches to recurrent neural network training. *Computer Science Review*, 3(3):127–149, August 2009.

[22] Wolfgang Maass, Thomas Natschläger, and Henry Markram. Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural computation*, 2002.

[23] Gerard R Manecke. Edwards FloTrac sensor and Vigileo monitor: easy, accurate, reliable cardiac output assessment using the arterial pulse wave. *Expert review of medical devices*, 2(5):523–7, September 2005.

[24] J. Martin Bland and D.G. Altman. Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, 327(8476):307–310, 1986.

[25] Jochen Mayer, Joachim Boldt, Michael W Wolf, Johannes Lang, and Stefan Suttner. Cardiac output derived from arterial pressure waveform analysis in patients undergoing cardiac surgery: validity of a second generation device. *Anesthesia and analgesia*, 106(3):867–72, table of contents, March 2008.

[26] John Mcmichael. CARDIAC OUTPUT IN MAN BY A DIRECT FICK METHOD. *Change*, pages 33–40, 1943.

[27] J More. The Levenberg-Marquardt algorithm: Implementation and theory. In G A Watson, editor, *Numerical Analysis*, volume 630 of *Lecture Notes in Mathematics*, chapter 10, pages 105–116. Springer, 1978.

[28] Ramakrishna Mukkamala, Andrew T Reisner, Horacio M Hojman, Roger G Mark, and Richard J Cohen. Continuous cardiac output monitoring by peripheral blood pressure waveform analysis. *IEEE Transactions on Biomedical Engineering*, 53(3):459–467, 2006.

[29] Michael Francis O'Rourke. Non-Invasive Determination of Aortic Flow Velocety Waveforms, 2004.

[30] J Penaz. Photoelectric measurement of blood pressure, volume and flow in the finger. *Digest 10th Int Conf Med Biol EngDresden*, 104:104, 1973.

[31] Christopher Prasser, Sylvia Bele, Cornelius Keyl, Stefan Schweiger, Benedikt Trabold, Matthias Amann, Julia Welnhofer, and Christoph Wiesenack. Evaluation of a new arterial pressure-based cardiac output device requiring no external calibration. *BMC Anesthesiology*, 7(1):9, 2007.

[32] Martin Schmidt. *Evaluation of a non-invasive dynamic Parameter for optimal Fluid Management.* Master thesis, Technical University Graz, 2011.

[33] T. Skjaerpe, L. Hegrenaes, and L. Hatle. Noninvasive estimation of valve area in patients with aortic stenosis by Doppler ultrasound and two-dimensional echocardiography. *Circulation*, 72(4):810–818, October 1985.

[34] A. J. H. Vendrik and R. R. Vierhout. Die unblutige Registrierung des Blutdrucks. *European Journal of Physiology*, 268(5):496–509, 1959.

[35] W. Zong, T. Heldt, G.B. Moody, and R.G. Mark. An open-source algorithm to detect onset of arterial blood pressure pulses. *Computers in Cardiology, 2003*, pages 259–262, 2003.