List of Figures

1 Introduction

Capnography is a non-invasive monitoring technique of the carbon dioxide (CO_2) concentration from the exhaled breath. It is extensively used in clinical environments for mechanically ventilated patients that undergo anesthesia. Capnography has the potential to identify hypoventilation or hyperventilation in emergency care settings, to aid resuscitation and to monitor or diagnose respiratory dysfunctions. In capnography, the expired carbon dioxide can be measured as a function of time (time capnography) or as a function of expired volume (volumetric capnography). The recording produced by a capnography monitoring device is a waveform known as capnogram which reveals information about the underlying cardiopulmonary condition of the patient.

The aim of this study is to monitor the respiratory profile of spontaneously breathing patients through volumetric capnography. Because "knowledge of physiology is the key to the interpretation of capnograms (Kalenda)", the respiratory process alongside the morphology of the volumetric capnogram is briefly presented in Chapter 2. Additionally, the most common capnography monitoring devices and a concise literature research are included in this chapter.

Capnography provides a clear picture of the patient's cardiorespiratory dynamics, but the technique is still underutilized due to a lack of robust means for assessing it. Nowadays, most methods of analyzing the capnogram are based on clinicians' visual assessments which are qualitative and prone to errors. Since a capnographic measurement accurately reflects the pulmonary physiology and pathophysiology through steady changes in the shape of the volumetric capnogram, a quantitative analysis of the capnogram is proposed.

Quantitative analysis of the capnogram would allow capnography to be used as a monitoring and diagnostic tool. In this thesis, this type of analysis was performed by decomposing the volumetric capnogram into physiological features. Feature extraction methods from literature are described in Chapter 3. Subsequently, in Chapter 4 a detailed description of machine learning decision trees algorithms is presented. Moreover, methods for assessing the classifier performance of decision trees are discussed.

Typical values for volumetric capnography features in healthy patients and patients with lung disease that allow a statement about quantifiable differences among cardiorespiratory states are presented in Chapter 5. Perhaps the greatest advantage of a quantitative analysis of the volumetric capnogram lies in the different shapes of capnograms specific to distinct disease classes. Volumetric capnogram features that describe the length, height and inclination of the capnogram are known to vary in the presence of a lung dysfunction. Therefore, volumetric capnography may be well-suited for discriminating among cardiorespiratory states. In order to characterize changes in the shape of the capnogram a single-breath mathematical model was implemented and is presented in Chapter 5.

The features are used further in a decision making process for automatic discrimination of cardiorespiratory states in three different clinical scenarios. For this task, an adequate set of features was selected and used by a decision trees algorithm. The classification performance is reported in Chapter 6.

2 Background

Current trends in monitoring and patient safety confirmed capnography as a standard of care in many clinical environments. Capnography has been shown to be effective in the early detection of adverse respiratory events as it provides a rapid and reliable method that detects life-threatening conditions. Medical societies representing anesthesiology, cardiology, critical care, pediatrics, respiratory therapy and emergency medicine have strongly recommended the use of capnography for patient monitoring during general anesthesia, conscious sedation, resuscitation or intubation [1]. Volumetric capnography, the monitoring of expired CO_2 and volume, serves as an important diagnostic tool in a number of areas including ventilator management, intraoperative assessment, patient transport, cardiopulmonary resuscitation and pulmonary embolism management [1].

During anesthesia there is interaction between two components; the patient and the anesthesia administration device which is usually a circuit and a ventilator. The critical connection between the two modules is either an endotracheal tube or a mask where the CO_2 is typically monitored. In expired respiratory gases, capnography directly reflects the elimination of CO_2 by the lungs to the anesthesia device. Moreover, it reflects the production of CO_2 by tissues and the circulatory transport of CO_2 to the lungs. Hence, capnography is an important non-invasive technique that provides information about CO_2 production, pulmonary perfusion and alveolar ventilation, respiratory patterns as well as elimination of CO_2 from the anesthesia circuit and ventilator [2]. Capnography provides medical personnel instantaneous feedback whether the endotracheal tube used for mechanical ventilation is correctly positioned or not to deliver gas to the lungs [3]. If the endotracheal tube is correctly placed in the trachea, CO_2 will be immediately detected and a capnogram generated, otherwise no signal will be generated [4]. Also, during conscious sedation capnography will provide an early detection of ventilation depression and respiratory failure [1].

Nowadays capnography monitors are small, very robust and easy to use with non-invasiveness being their greatest advantage. This allows easy integration into patient monitors, defibrillators and ventilators, making capnography facile and inexpensive to use [3]. Despite the broad clinical utility and various advantages mentioned earlier, capnography still remains underutilized due to the lack of education concerning the wide range of applications of the technology. Thus, a more robust way of using capnography would be in the field of diagnostics. Since the transport of volume is central to the function of the lungs, volumetric capnography reveals comprehensive information about the underlying state of the cardiopulmonary system through a generated graphical waveform known as a capnogram [1].

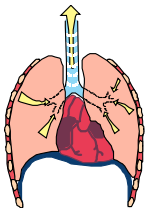
The content of this chapter is organized as follows: Chapter 2.1 explains the respiratory process with a focus on the dynamics between heart and lungs; Chapter 2.2 presents the most common capnography measuring devices, Chapter 2.3 summarizes the definition of the volumetric capnogram, while Chapter 2.4 provides an overview of capnography analyses carried out so far.

2.1 Respiratory Physiology

During normal breathing, air is regularly inhaled and exhaled, oxygen is delivered to the body and carbon dioxide is eliminated. The respiratory process consists of three main events [5]:

1. Cellular metabolism of food into energy – oxygen consumption and carbon dioxide production

2. Transport of oxygen and carbon dioxide between cells and pulmonary capillaries and diffusion from/into alveoli



3. Ventilation between alveoli and surrounding atmosphere

Because all three components of the respiratory process are involved in the appearance of CO_2 in the exhaled gas, capnography provides a clear picture of the patient's cardiorespiratory system. Understanding the relation between the heart and lungs sheds light on the respiratory process and how cardiopulmonary diseases affect the lungs.

The heart is composed of two upper chambers, the right and left atria, and two lower chambers, the right and left ventricles. Deoxygenated blood from all over the body enters the right atrium and flows into the right ventricle, where it's pumped through the pulmonary arteries into the lungs. There, the blood releases carbon dioxide and picks up oxygen. The oxygen-rich blood then returns to the left atrium through the pulmonary veins, flows into the left ventricle, and leaves the heart through the aorta. From the aorta, the blood travels to the rest of the body [6].

During inspiration, the oxygen diffuses from the membranes of the elastic air sacs, i.e. the alveoli, into the capillaries [4]. The alveoli are small air-containing compartments of the lungs in which the bronchioles terminate and from which respiratory gases are exchanged with the pulmonary capillaries. Capillaries are the smallest kind of blood vessels that form networks where the arterial and venous circulation meet for exchange of oxygen, nutrients, and wastes with body tissues (Concise Encyclopedia). The alveoli are surrounded by the alveolar-capillary membrane, which normally prevents liquid in the capillaries from entering the air sacs [6].

With each breath, the alveoli take in oxygen and release carbon dioxide. Normally, the exchange of gases takes place without problems [6]. But in certain circumstances, the alveoli are filled with an excess of fluid, which prevents oxygen from being absorbed into the bloodstream [6]. Most commonly this fluid accumulation in the lungs is caused by cardiac failure also known as heart failure (HF). Cardiac failure is a chronic condition in which a structural abnormality or disease limits the heart's capacity to pump an adequate amount of

blood to meet the oxygenation demands of the body's tissues [7]. Heart failure is characterized by repeating acute deterioration with pulmonary congestion and edema in lungs and legs that often result in hospitalization.

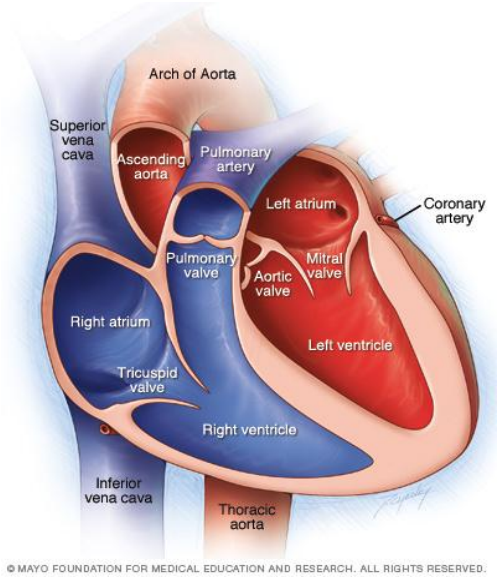


Figure 2.1. Chambers and valves of the heart. The blue part represents deoxygenated blood going to the lungs and the red part represents oxygenated blood coming from the lungs (source: [6])

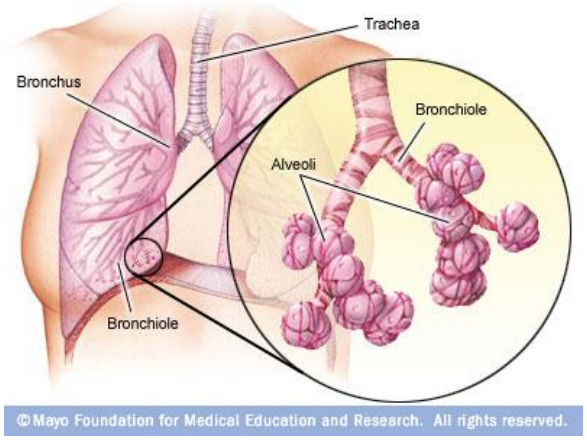


Figure 2.2. Airways (trachea, bronchus, bronchiole) and air sacs of the lungs (alveoli). (source: [6])

Cardiac pulmonary edema, resulting from congestive heart failure (CHF), occurs when the diseased or overworked left ventricle is not able to pump out enough blood from the lungs. Consequently, pressure increases inside the left atrium and then in the veins and capillaries from the lungs, causing fluid to be pushed through the capillary walls into the air sacs [6]. The increase in interstitial fluid causes shortness of breath, reduced cardiopulmonary efficiency, respiratory dyspnea and possibly cardiac arrest.

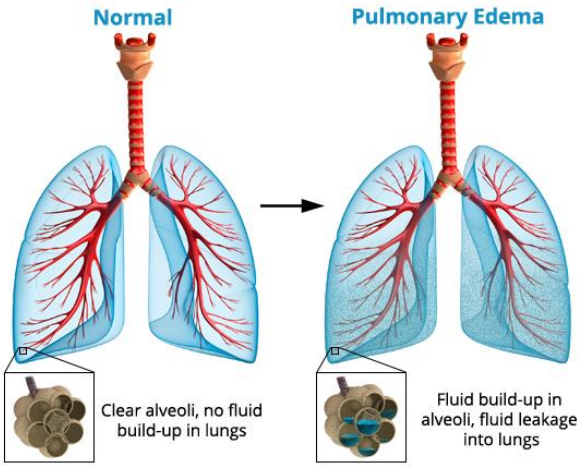


Figure 2.3. Normal lungs (left) and lungs with edema (right) with a close-up of alveoli. (source : [8])

2.2 Capnogram Technology

Infrared sensors (IR) are the principal means used for detecting CO_2 in modern capnography devices. There are two main types of capnography sensors: mainstream and sidestream sensors, with two examples being displayed in Figure 2.4. They both measure the number of CO_2 particles in the exhaled gas by using infrared light. A broad band of light wavelengths are emitted by an infrared source which is then focused by a lens and transmitted across the airway adapter (mainstream sensor) or the sample cell (sidestream sensor). CO_2 that passes through the light absorbs some of this light at a specific wavelength. The light is then focused through a filter set onto a detector at the CO_2 absorption wavelength. The amount of CO_2 in the airways is then determined by the measuring the light that was absorbed [3].

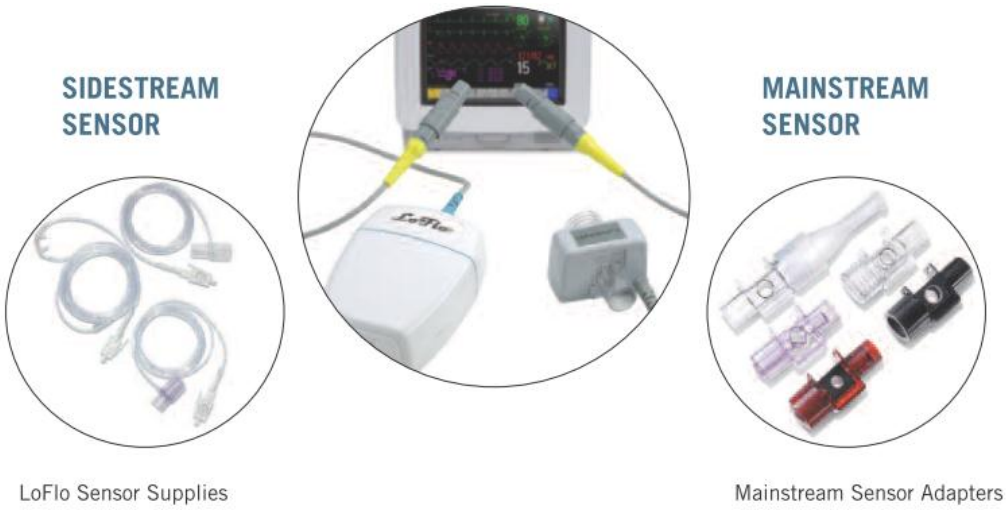


Figure 2.4. Sidestream versus mainstream capnography adopted from [3]

The sidestream sensor is commonly used in emergency medicine or procedural sedation, and the mainstream sensor is generally used on intubated patients [3]. Mainstream sensors are placed at the airway of the patient allowing the inspired and expired gas to pass directly across the IR light path [5]. Sidestream sensors are located away from the airway requiring a gas sample to be continuously aspirated from the breathing circuit and transported to the sensor by means of pump [5].

A commercially available, CE marked Philips NM3 respiratory profile monitor was used in this study for acquiring the respiratory measurements. This device is intended for use with ventilated patients but can also be used with subjects with spontaneous breathing. The breathing gas is analyzed and recorded with the NM3 monitor (Figure 2.5) and the Capnostat 5 mainstream CO_2 sensor accessory shown in Figure 2.6.

Figure 2.5. Philips Respironics NM3 Monitor

Figure 2.6. Capnostat 5 Mainstream CO_2 Sensor

2.3 The Volumetric Capnogram

Capnography is the monitoring of the concentration or partial pressure of carbon dioxide in the respiratory gases. The integration of flow and volume signals with the CO_2 signal and the measurement of the parameters characterizing this curve are widely known as volumetric capnography [1]. The volumetric capnogram is visualized as a graph that displays the elimination of CO_2 concentration plotted against the volume expired in one breath. It has also been referred as "the single breath test for carbon dioxide", a term used initially by Fowler [9] to describe the "single breath test for nitrogen" because the resulting curve resembles Fowler's curve in shape [1]. The volumetric capnogram is generally divided in three phases as shown in Figure 2.7. The meaning of each capnogram phase was explained below in this chapter.

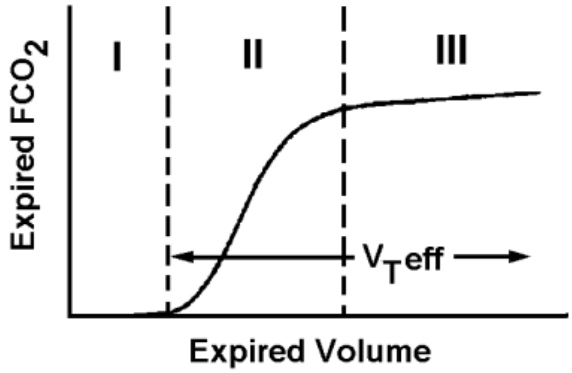
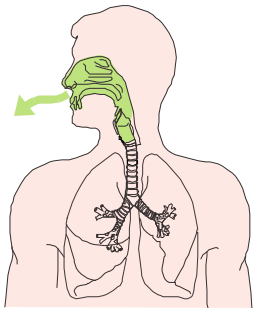


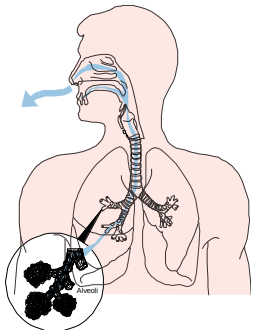
Figure 2.7. The three phases of a volumetric capnogram adopted from [1]

A volumetric capnography recording is similar to an electrocardiogram, only that it provides a breath by breath visualization of the patient's respiratory pattern. A section from a typical volumetric capnography waveform looks like in Figure 2.8.

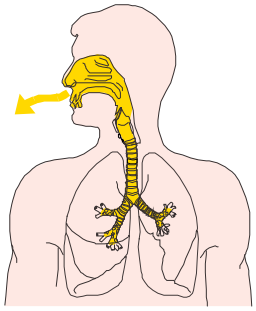
Figure 2.8. Volumetric capnogram waveform recording



Phase I is the beginning of exhalation which represents the CO_2-free ineffective tidal volume region as shown in Figure 2.7 [1]. There is no CO_2 at the beginning of exhalation. The air is from the trachea, mouth and nose. This upper airway area is often called "dead space" because there is no gas exchange present in the upper airway.



Phase II compromises the transitional region characterized by a S – shaped upswing on the tracing caused by a rapidly increasing CO_2 concentration resulting from progressive emptying of the alveoli [1] (see Figure 2.7). The deadspace gas starts mixing with the alveolar gas and CO_2 is detected in the exhaled air [1].



Phase III, the alveolar plateau representing the CO_2 rich gas from the alveoli, is flat with a slight upward tilt towards the end (see Figure 2.7) [1]. This plateau illustrates the uniform concentration of carbon dioxide in the pulmonary system. The ventilation-perfusion relationships of the lung are more accurately reflected in the slope of the phase III (illustrated in in Figure 3.4) [1].

2.4 Prior Analyses

One of the earliest descriptions of the volumetric capnogram and a method to determine airway deadspace is that of Aitken and Clark Kennedy (1928). Later in 1948, Fowler described the single breath test for the nitrogen curve, divided the curve into four phases and proposed an algorithm for the determination of deadspace [9]. While single breath curves for CO_2 appeared in 1961 in literature it was not until 1980 when Fletcher presented in a broad work the concepts of CO_2 elimination, divided the capnogram in three phases and proposed a method for calculating slope phase III [10]. Later research in the 2000's focused on giving reference values for diagnostic purposes. A staple publication frequently cited in literature is that of Astrom [11], who examined the concepts of deadspace, explaining the differences in deadspace volumes between genders, also noting the correlation between volume expired and slope phase III and describing it visually through an exponential model.

Unfortunately most studies investigate mechanical ventilation and respiratory dysfunctions such as asthma, sleep apnea, ARDS (Acute Respiratory Distress Syndrome), COPD (Chronic Obstructive Pulmonary Disease) and there is a lack of information concerning spontaneously breathing patients, especially those suffering from CHF and pulmonary edema. The first study discussing volumetric capnography in a group of spontaneously breathing patients, which also included seven patients suffering from CHF, is

that of Verschuren et al. from 2005 [12]. In this article, Fletcher's principles were implemented in order to verify the reliability and reproducibility of a new capnographic program in various types of clinical condition. Besides providing reference values for patients with several cardiopulmonary conditions it was concluded that volumetric capnography might be considered an easy non-invasive bedside gold-standard method for deadspace measurements.

More recently, another work that discusses patients with heart failure is that of Arena et al. from 2012 [13] who studied the prognostic value of capnography during rest and exercise in patients with heart failure (HF). In his study, the ability of resting and exercise $ETCO_2$ (end-tidal CO_2) to predict major cardiac events in heart failure was assessed through cardiopulmonary exercise testing. Thus, 963 patients with systolic HF were followed for major adverse events up to four years post the cardiopulmonary exercise analysis and the conclusions were that $ETCO_2$ (both at rest and during exercise) reflects disease severity and cardiac function and is a significant predictor of adverse events.

Last, but not least, two studies often used as a reference basis in this thesis, were those of Tusman et al. from 2009 and 2013. In the study from 2009 [14], the performance of the FA-LMA Algorithm is compared with Fowler's method to calculate volume of airway deadspace and slope phase III. The continuous and varied changes in the morphology of the curve are addressed by optimally fitting a custom mathematical model to the volumetric capnogram curve using a Functional Approximation Levenberg-Marquardt Function (FA-LMA) in order to extract the volumetric capnogram features from the fitted model curve. This custom mathematical model was implemented in Chapter 3 in order to check if the volumetric capnograms analyzed could provide reference values comparable with other studies. In the study from 2013 [15], volumetric capnography reference values are presented for healthy individuals based on the FA-LMA method described previously. The aim of this study was to provide normal values for most of the non-invasive parameters in healthy patients in order to interpret results in clinical patients. Furthermore is was aimed to determine the lung function in patients undergoing mechanical ventilation in order to adjust the ventilator settings and detect a pathological condition [15].

3 Processing and Feature Extraction Methods

Processing of the volumetric capnogram and feature extraction methods are applied in order to prepare the data for visual analysis and later classification. After the patient breathes in the measuring device, a respiratory file, also known as a waveform file, containing raw continuous signals, is generated. The waveform file contains 6 basic waveform signals, "Flow [LPM]", "Pressure" [cmH_2O], "CO_2" [%], "Volume" [ml],"Expired Volume" [ml] and "CO_2 Volume (VCO_2)" [ml], as illustrated in Figure 3.1. In this study the focus is on CO_2 [%] and Expired Volume [ml] signals. These continuous signals contain samples of volume [ml] with corresponding samples of CO_2 [%] at a unique timestamp. From the continuous waveform strip, the start and the end of each expiration period from a measurement were extracted in order to select each breath individually. Next, the start and the end of each expiration period were identified in the Expired Volume [ml] and CO_2 [%] signals. The CO_2 [%] signal was plotted versus the Expired Volume [ml] for each expiratory period. One graphic plot represents the volumetric capnogram.

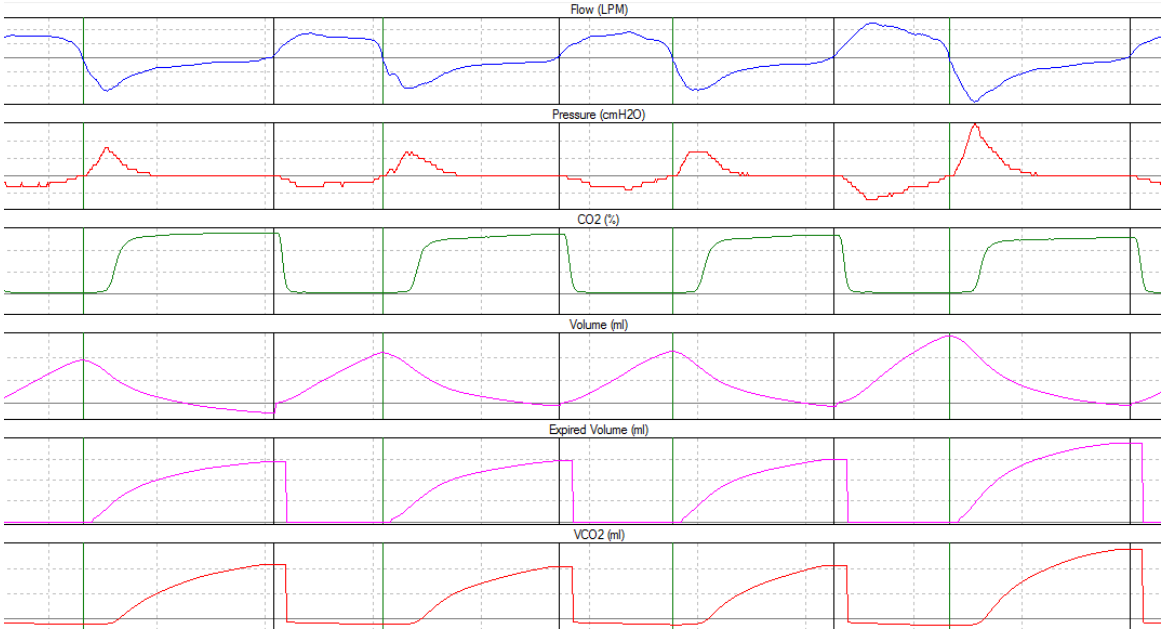


Figure 3.1. Respiratory waveform signals as displayed by FlowTool Physiologic Viewer©

Volumetric capnogram shapes differ accordingly to the respiratory disorders. The characteristic of each condition is reflected in the varying morphology of the capnogram. Physiological features give the typical shape of the capnogram, but underlying respiratory conditions change this form. To this end, it is important to notice and measure the variations between the volumetric capnograms and this is why feature extraction methods are applied.

The volumetric capnogram is processed in order to remove invalid breaths from the analysis so that the respiratory parameters represent normal-range values and not extreme values. Firstly, the analysis is started with pre-processing by removing breaths which have an expired volume outside the range of a chosen threshold. This threshold is set accordingly to

data distribution and to the medical background of the patients. Another approach to an initial processing is applying a low-pass filter that evens the volumetric capnogram so that a non-linear model curve could be fitted on it with the smallest possible error. Figure 3.2 shows a logical diagram with the topics discussed in this chapter.

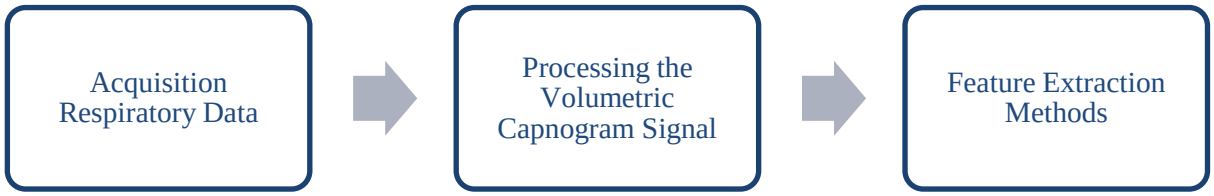


Figure 3.2. Analysis steps for studying volumetric capnograms

3.1 Volumetric Capnogram Processing

Real world measurements may contain erroneous data which appear due to human anxiety, improper handling or positioning of the measurement device, errors in the acquisition system or not respecting the measurement protocol. This erroneous data renders outlier values, i.e. extreme values, very small or very large values that do not reflect the average characteristic of the parameter discussed. Therefore, before any analysis is performed, outliers should be removed because they have a high error influence upon the feature values.

The volumetric capnogram is after all a signal, and as all real signals is subject to noise – an unexpected and unwanted change that appears during the recording of the respiratory gases with the measurement device. If curve fitting models are imposed on the raw points of the signal, signal processing techniques filters should be applied in order to obtain a good model fit that follow closely the recognized shape of the capnogram.

3.1.1 Discarding Outlier Breaths

In the present study, outlier breaths appear because in some cases it takes time to get accustomed with the device and meanwhile the patient feels nervous or is too focused on his respiration which becomes forced instead of spontaneous. Moreover, the device might not be placed correctly and thus it doesn't record properly the respiration. These particular situations, as well as the normal breathing pattern which tends to be irregular, lead to very high or very low expired volumes.

In order to avoid the inclusion of outlier breaths in further analysis, all breaths which were outside the ±50% median expired volume of the measurement in cause were removed as suggested by Tusman et al. [15]. Next, a fixed threshold of minimum 200 ml of expired volume per breath was applied. The expired volume is chosen as a discarding threshold parameter because it reflects the breathing pattern of a patient. This method works very well if the patient is healthy, but for CHF the average expired volume is much lower than in the case of healthy patients due to breathing impairment caused by an improper filling of the lungs. In the latter case, setting a ±50% median expired volume limit for discarding outliers might reject normally expired volumes corresponding to the upper median volume threshold. Therefore, firstly breaths outside the range -50%/+95% of the median expired volume were

taken out and then a threshold of minimum 200 ml of expired volume is set. The threshold of 200 ml was chosen because it contains the minimum amount of data samples from which respiratory parameters could be calculated correctly.

3.1.2 Signal Processing Techniques

Real world measurements are subject to noise which is expressed in different forms according to the data that is part of. Noise in a signal is described as unwanted and unexpected change that appears. In order to remove the gratuitous information from the capnogram signal while preserving its shape, a low-pass filter was applied. A second order Butterworth filter was chosen, its design criteria being explained below. By applying this technique unwanted frequencies are removed from the respiratory signal and the final filtered signal results in a "smoother" curve. The morphology of the capnogram is kept intact and thus all the useful information is preserved.

The frequency response of the Butterworth filter is maximally flat, i.e. has no ripples, in the pass-band and is monotonic overall [16]. These filters are defined by two parameters, the cutoff frequency, ω_c and the order of the filter *n*. The filter order controls the sharpness of the signal, or the roll-off rate that becomes steeper as the filter order increases. The roll-off rate is defined as 20**n* dB/decade, where *n* is the order of the filter [17]. In this case, a filter of order 2 was appropriate, with the transfer function in discrete time being the following:

$$H(z) = \frac{b_1 + b_2 z^{-1} + b_3 z^{-2}}{1 + a_2 z^{-1} + a_3 z^{-2}} \tag{3.1}$$

The cutoff frequency is the frequency where the magnitude response of the filter is $\sqrt{1/2}$ [16]. The normalized cutoff frequency, used in the design of the filter, is chosen with respect to the Nyquist frequency of the signal. The sampling interval from the volumetric capnogram was 0.1 [ms] therefore the sampling frequency of the respiratory signal was 100 [Hz]. This resulted in a Nyquist frequency of 50 [Hz].

Besides the typical noise that appears in the volumetric capnogram, sometimes, sudden drops in the CO_2 [%] rate, also known as "dips", appear in capnogram. Their cause could be the mouthpiece the patient breathes in spontaneously. Towards the end of the expiration the patient might feel the need to swallow and the "dips" appear. As these dips are not a part of the capnogram norm and they could render miscalculations in the parameters that will be extracted, their removal was needed. In order to accomplish this, the final part of the CO_2 signal was differentiated and the inflection points (the local minima and maxima) were identified. Starting from the inflection points, the values of the left cusp and right cusp of the dip were searched by turns. After the first and last points of the dip were found, the signal was extrapolated between them. In Figure 3.3 a volumetric capnogram signal is first shown without any filters (blue) and after the dip was removed and the Butterworth filter was applied (green). Note that the filtered signal is smoother and that the dip disappeared.

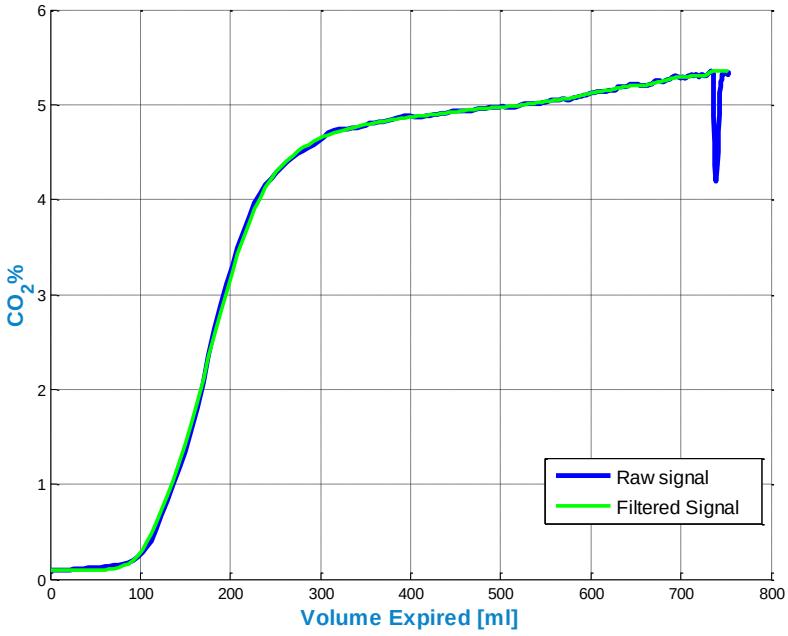


Figure 3.3. Volumetric capnogram signal. The blue capnogram represents the unfiltered signal and the green capnogram shows the filtered signal after the Butterworth filter was applied and the "dips" removed.

3.2 Algorithms for Feature Extraction

In order to describe the capnogram waveform quantitatively with the aim of analyzing the respiratory status of the patient, physiological features are extracted from the capnogram signal. To achieve this, firstly, the shape of the capnogram is followed closely and features that define its metrics such as length, height and slope are extracted. These metrics have a connection with the respiratory process. These features also express physiological relations and are studied with respect to the underlying health condition since respiration impairments are reflected as changes in the feature values.

Feature extraction is performed automatically by a waveform processing tool (FlowTool Physiologic Waveform Viewer © Version 3.0.3 from Philips Respironics) that displays the values of the features. These values were extracted from the available respiratory signals contained in the waveform files. This tool is used to validate the results of the feature extraction algorithms presented in this chapter. The same algorithms employed by the waveform processing tool were implemented in MATLAB® and used afterwards for further analysis.

3.2.1 Physiological Features

The most important features related with the respiratory state are those that describe the shape of the capnogram signal - the expired volume (V_{exp}) [ml], the end-tidal CO_2 [%] or [mmHg] ($ETCO_2$) and the inclination angles (slopes) of the capnogram. The values of the features represent the characteristic of a single breath.

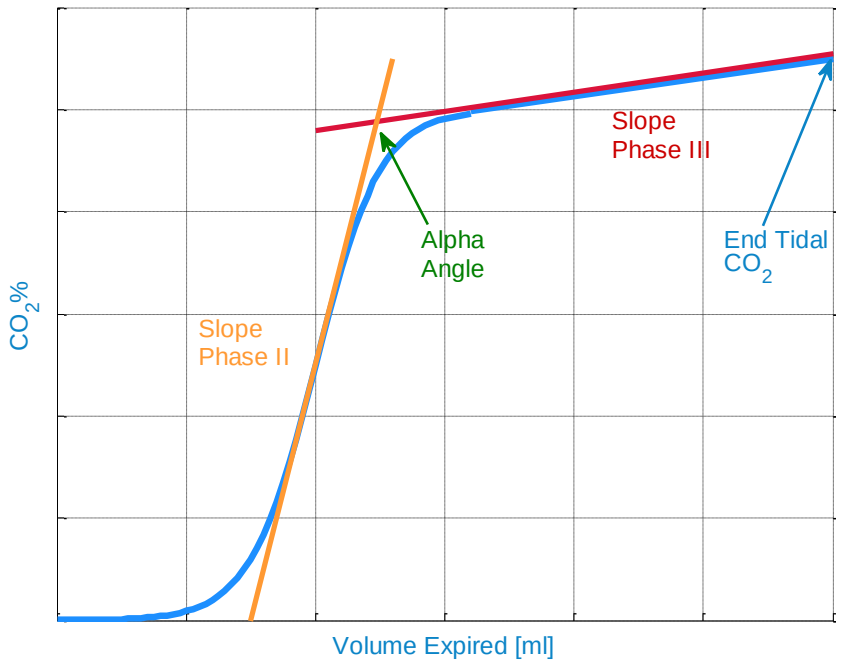


Figure 3.4. Volumetric capnogram signal, abstract representation. The slopes phase II [%/L] and III [%/L] are marked as extrapolated straight lines. Alpha Angle [°] is the angle at their intersection. The $ETCO_2$ [%] value is the final CO_2 [%] sample from the volumetric capnogram

Expired Volume (V_{exp}) [ml]

Expired volume, also known as tidal volume [ml], is recorded during the expiratory phase (expiration) and represents the part of the ventilatory cycle from the beginning of the expiratory flow to the beginning of inspiratory flow. More simply put, V_{exp} represents the total volume expired by the subject during a breath. As shown in Figure 3.4, V_{exp} represents the length of the capnogram. Patients with CHF, usually breathe at low expired volumes due to breathing impairment caused by accumulation of fluid in the lungs. V_{exp} is the final volume sample from the expired volume signal.

Expired Volume of CO_2 (VCO_2) [ml]

Another feature related with V_{exp} is the expired volume of CO_2 (VCO_2) [ml] which represents the carbon dioxide elimination during a breath. With no rebreathing this feature is exactly the area under the volumetric capnogram [18].

Slope Phase II (S_{II}) [%/L]

A feature that describes the inclination of the capnogram is slope phase II [%/L] (S_{II}) (see Figure 3.4), which was calculated by least-squares curve fitting on the points bounded by 40% to 60% of the phase II volume expired (VPh2). The phase II of the capnogram is represented by a rapid rise in CO_2 concentration, resulting from the progressive emptying of the alveoli [18]. The volume of the phase II constitutes the portion of the tidal volume where increasing amounts of CO_2 are leaving the lungs at different rates of ventilation and perfusion [15]. Due to the fact that CHF causes fluid leakage into the lungs, changes in S_{II} are expected to be observed.

Slope Phase III (S_{III}) [%/L]

Furthermore, another feature related with the inclination of the capnogram is slope phase III [%/L] (S_{III}), as described in Figure 3.4. This feature was computed by linear least-squares regression on the points bounded by 30% to 70% of expired CO_2 volume [19]. There are many definitions regarding the implications of slope phase III, but most clinicians agree that it reflects ventilation inhomogeneities within the lungs, also known as the ventilation/perfusion ratio (V/Q) mismatch [15]. In the case of patients with a respiratory disease, the value of the S_{III} is higher than at normal patients. The higher S_{III} values show that the V/Q ratio is not equilibrated and that the amount of air reaching the alveoli is too small (ventilation) because blood backs-up and pools in the lungs (perfusion) which leads to increased pressure that could move the interstitial fluid completely in the alveoli [20]. Thus, the alveoli do not function at their maximum capacity and less CO_2 is expired in the end. Usually, a low S_{III} value indicates a better matching between the ventilation and perfusion [15].

V/Q mismatch may be either due to incomplete gas mixing (alveolar mixing effect) or it could result from a delayed gas distribution to the alveoli which render a low V/Q mismatch [21]. Slope phase III is based on the gas-emptying sequence of the alveolar components, thus if the components empty at the same time, the CO_2 expired results in a smooth, flat or slightly up-inclined phase III [21]. In patients with greater V/Q mismatch, the CO_2 emptying is sequential which results in a steeper and more up-inclined phase III [21]. The implications of S_{III} make this feature very useful for diagnosing respiratory abnormalities within the lungs.

End-tidal CO_2 ($ETCO_2$) [%]

In assessing the height of the capnogram, end-tidal CO_2 ($ETCO_2$), expressed either as a fraction [%] or as a pressure [mmHg] is used. This value is computed as the last expiratory CO_2 value from the CO_2 [%] waveform signal, just immediately before the start of the next inspiration. Obstructive disease patients are generally seen to exhibit high $ETCO_2$ values reflecting high arterial CO_2, as they do not satisfactorily expel CO_2 during exhalation. Restrictive lung disease like CHF, result in decreased $ETCO_2$ levels since perfusion of CO_2 into the alveoli is impaired [4].

Volume of airway deadspace (VD_{aw}) [ml]

The volume of airway deadspace (VD_{aw}) [ml], is the volume of the conducting airways at the ‘midpoint’ of the transition from dead space to alveolar gas [19]. Deadspace refers to ventilated areas from the respiratory apparatus which do not participate in gas exchange [5]. Since these parts do not contribute to CO_2 eliminations wasted ventilation appears [5]. From another point of view, deadspace could be considered conceptually as areas of ventilation without any pulmonary perfusion [21]. Total or physiologic dead space is the sum of the following two components (only when mechanical breathing is not involved): anatomic dead space and alveolar dead space. The anatomical deadspace refers to the deadspace caused by the airways leading to the alveoli [5]. Alveolar deadspace can be defined as the portion of the inspired volume that passes through the anatomical deadspace to mix with gas at the alveolar level, but does not participate in the gas exchange [21].

In order to extract VD_{aw}, Fowler's method [9] was implemented in MATLAB ®. This is a geometrical method which was initially applied by Fowler by hand on the capnogram signal printed on paper (see Figure 3.6). For illustration purposes, in Figure 3.6, the extrapolated slope phase III line was depicted as a straight line. Fowler's method consists of the following steps which were computationally implemented (see Figure 3.5):

1) Firstly, slope phase III is extracted from the volumetric capnogram according to the algorithm mentioned previously
2) A vertical line is placed on the x-axis at the beginning of the phase II volume (VPh2) and moved iteratively towards the right boundary of the VPh2 until the area of the upper triangle (*q)* formed by this line and the extrapolated line of slope phase III becomes equal with the lower area triangle *(p)* formed by the vertical line and the CO_2 signal.
3) When the areas of the triangles (see Equation (3.3) and Equation (3.4) for the areas' formulas) become equal, stop the iteration and save the x-coordinate of the vertical line. This represents the value of the VD_{aw} [ml].

The vertical line was moved iteratively with a step of 1ml towards the right limit. The areas of the triangles are never going to be exactly equal due to trapezoidal integration and noise in the capnogram signal, so the algorithm could not have as a stopping condition "the equal area" rule. After all, the capnogram is not a perfectly smooth sigmoid curve. The iteration loop will stop immediately after the lower triangle area has become greater than the upper triangle area.

The areas of the triangles have been computed by numerical integration using the trapezoidal rule. This method approximates the integration over an interval by breaking the area down into trapezoids with more easily computable area. The trapezoidal rule approximates the region under the graph of the function *f(x)* as a trapezoid and calculates its area. For integration with N+1 points, where the spacing between points is not constant, the general formula [22] is the one from Equation (3.2):

$$\int_a^b f(x)dx = \frac{1}{2}\sum_{n=1}^{N}(x_{n+1} - x_n)[f(x_n) + f(x_{n+1})] \tag{3.2}$$

Making the equations particular to our data, with *f* being the concentration of CO_2 and S_{III}' the extrapolated regression line defining the slope of the phase III (S_{III}), the equations for the lower area triangle (*p*) and upper area triangle (*q*), will be the following:

$$p = \int_{VPh1}^{VD_{aw}} f(V\exp)dV\exp \tag{3.3}$$

$$q = \int_{VD_{aw}}^{V\exp} (S_{III}' - f(V\exp))dV\exp \tag{3.4}$$

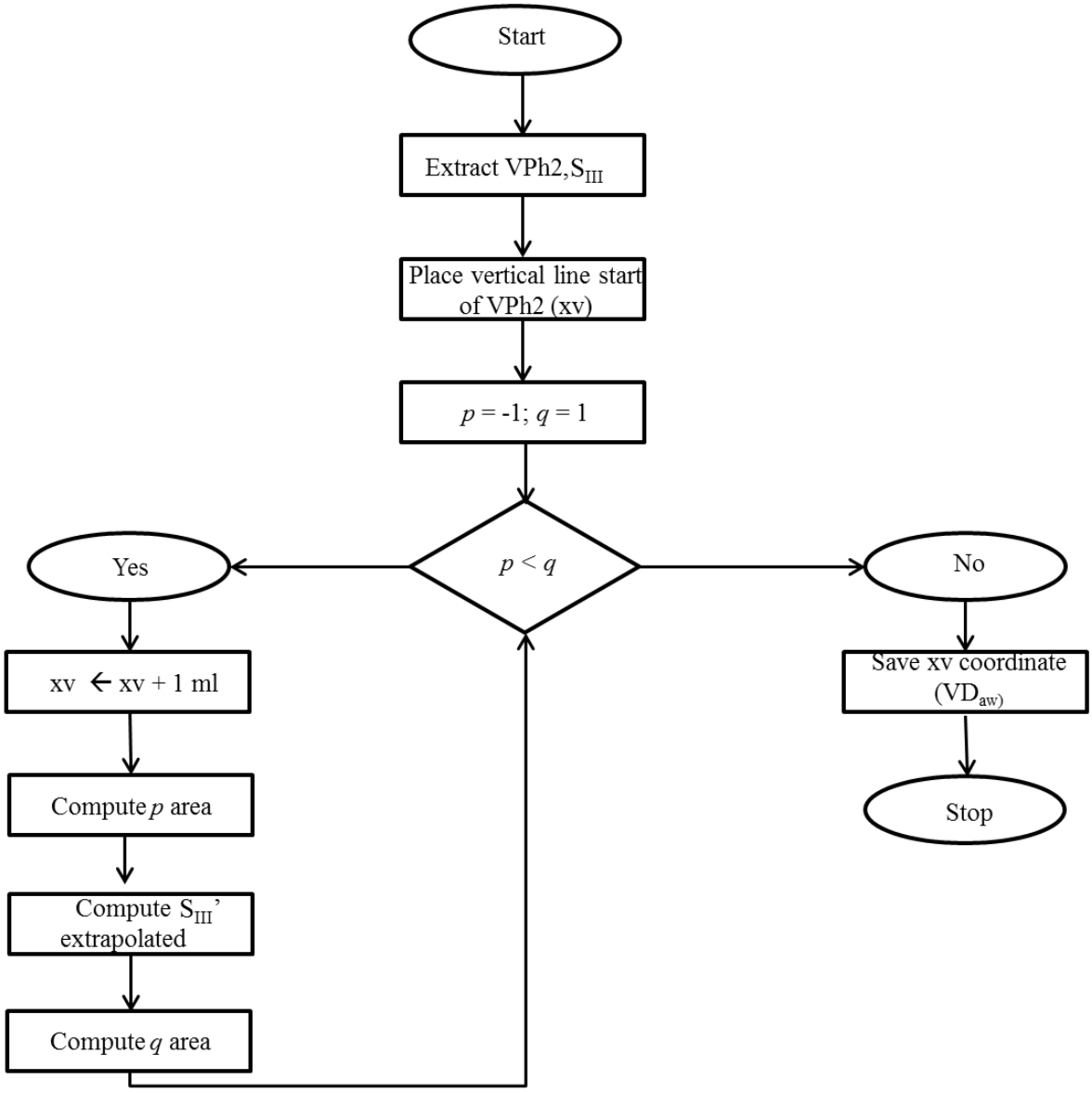


Figure 3.5. Logical diagram for calculating VD_{aw}

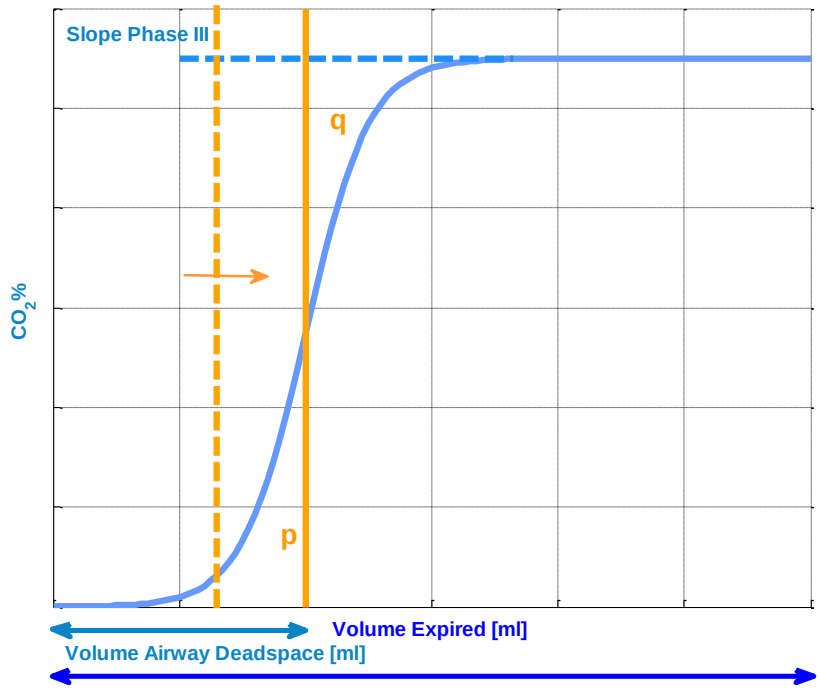


Figure 3.6. Abstract representation of Fowler's method "equal area triangles" (*p*,*q*) for calculating the volume of airway deadspace (VD_{aw}). Slope phase III was represented as an extrapolated straight line

Alpha Angle (α) [°]

Moving on to another feature which describes the inclination of the capnogram, i.e. the angle between the intersection of the extrapolated regression lines of the slopes of the phase II and III, known as the alpha angle, α [°] (shown in Figure 3.4**)**. Since the slope of a line is the tangent of its angle of elevation, the angle between two lines, is the difference between the two angles. After computing the difference, the inverse tangent function (arctangent) is applied. Because the alpha angle is represented by the obtuse angle of the intersection, the arctangent result will be subtracted from 180 degrees. The alpha angle increases as the slope of phase III increases and as slope of the phase II increases. This angle represents the connection between the slope of the phase III and $ETCO_2$ [21]. Changes of the alpha angle correlate with sequential emptying of the alveoli in obstructive lung diseases and in the case of bronchospasm [21]. It is generally accepted that the greater the alpha angle, the more heterogeneous are the functional units of the lungs [21].

Spontaneous Respiration Rate (RR) [breaths/min]

Spontaneous respiration rate (RR) [breaths/min] was calculated by measuring the time interval between the start of an inspiration and the start of the next consecutive inspiration. It was computed as 60 / (elapsed time interval in seconds). The respiration rate is generally inversely proportional with V_{exp}, which means the larger the expired volume, the lower the respiration rate.

3.2.2 Non-Linear Model Function

For the extraction of features from the volumetric capnogram, recently, a new method has been proposed that approximates the capnogram with a custom mathematical function as defined by Tusman et al [14]. In this study, a Functional Approximation Levenberg-Marquardt algorithm for the non-linear fitting of the volumetric capnogram was used. Next, a non-linear fitting algorithm was employed in order to extract the capnogram features from the fitted model curve.

In this work, the custom non-linear model function was applied on the volumetric capnograms available in order to test if it could be used as a discriminating method between healthy and diseased patients. The hypothesis was that in the case of CHF patients, the average root mean squared error statistic (RMSE) of the model will be higher than in the case of healthy patients due to a greater variability in the breathing pattern.

In Tusman's study [14], the volumetric capnogram was considered to be an asymmetrical sigmoidal curve which could be expressed by the sum of a few mathematical functions (see Equation (3.8)), defined by $f(t,\mathbf{x})$ where f is the concentration of CO_2, t the expired volume and $\mathbf{x}$ the coefficients vector. These functions describe parts of the volumetric capnogram as illustrated in Figure 3.7. The corresponding equations have been implemented in MATLAB® and applied on the respiratory measurements available.

$$f_0(t,\mathbf{x}) = x_1 \tag{3.5}$$

$$f_1(t,\mathbf{x}) = \frac{(x_2 - x_1)x_3}{(1 + e^{(-x-x_4)/x_5}} \tag{3.6}$$

$$f_2(t,\mathbf{x}) = \frac{(x_2 - x_1)(1 - x_3)}{(1 + e^{(-x-x_6)/x_7})} \tag{3.7}$$

$$f(t,\mathbf{x}) = f_0(t,\mathbf{x}) + f_1(t,\mathbf{x}) + f_2(t,\mathbf{x}) \tag{3.8}$$

In this problem, the vector coefficients $\mathbf{x}$ has 7 parameters $(\mathbf{x} = [x_1, x_2, x_3, x_4, x_5, x_6, x_7])$, t represents the independent variable and f represents the dependent function. In a least-squares problem a function that is a sum of squares is minimized so that the result of the fitting process shows an estimate of the "true", but not known coefficients from the model [14]. Given m data points (t_i, y_i), the aim was to find the vector $\mathbf{x}$ of n parameters that gave the best fit in the least-squares sense to the model function $f(t, \mathbf{x})$. If the components of the residual vector are defined as in Equation (3.9) then the goal is to minimize the function $S(\mathbf{x})$ from Equation (3.10):

$$F_i(\mathbf{x}) = y_i - f(t_i,\mathbf{x}), i = 1,...m, \text{ where } m \text{ is the number of data points} \tag{3.9}$$

$$S(\mathbf{x}) = \min_{x} \frac{1}{2} \sum_{i=1}^{m} F_i(\mathbf{x})^2 \tag{3.10}$$

The Levenberg-Marquardt curve-fitting method is actually a combination of two minimization methods: the gradient descent method and the Gauss-Newton method. The gradient descent method is a local minimization method which updates parameter values by taking direction steps proportional to the negative gradient. In the Gauss-Newton method, the sum of the squared errors is reduced by assuming the least squares function is locally quadratic, and finding the minimum of the quadratic by solving a set of linear equations. The solution of the quadratic problem gives a local optimum for the non-linear problem. Thus, if is $\mathbf{x_k}$ is an approximate solution and *J(x)* the Jacobian matrix of the vector function *F(**x**)* then Equation (3.11) is solved for the appropriate Newton step d_k at each iteration [14]. The next appropriate solution will be given by Equation (3.12) and the process will be iterated until convergence to the solution is achieved [14].

$$J^T(x_k)J(x_k)d_k = -J^T(x_k)F(x_k) \tag{3.11}$$

$$x_{k+1} = x_k + d_k \tag{3.12}$$

The Levenberg-Marquardt uses a search direction that is a solution of the linear equation [23]:

$$\left(J(x_k)^T J(x_k) + \lambda_k I\right)d_k = -J(x_k)^T F(x_k) \tag{3.13}$$

where the scalar parameter λ_k controls both the magnitude and the direction of d_k and I is the identity matrix. When λ_k is zero, the direction d_k is identical to that of the Gauss-Newton method. As λ_k tends to infinity, d_k tends towards the steepest descent direction, with magnitude tending to zero[23].

When trying to apply the custom non-linear model on the volumetric capnograms available, it was observed that the algorithm often failed to converge to a solution and thus it did not succeed fitting a curve that followed the volumetric capnogram. A possible reason behind this behavior could be the fact that the volumetric capnogram contains noisy frequencies that affect the functioning of the algorithm. Although Tusman et al. [14] does not mention prior signal processing techniques, after the Butterworth filter described in Chapter 3.1.2 was applied on the volumetric capnogram the fitting algorithm always converged to a solution. This proves that in order to extract features from a fitted curve that replaces the volumetric capnogram, filtering techniques that smooth out the capnogram without damaging its shape, must be applied beforehand.

After fitting the custom non-linear model function on the capnograms, the hypothesis formulated at the beginning of this section was tested. The root mean squared error statistic (RMSE) shown in Equation (3.14) is an estimate of the standard deviation of the random component in the data that was used to evaluate the goodness of fit [24]. A value closer to 0 indicates that the model has a smaller random error component [24]. The average RMSE of the fitted non-linear model was 0.1419 for healthy patients and 0.1175 for CHF patients. These results reject the hypothesis that the model function proposed by Tusman et al could

serve as a discriminating method and show instead that the model function could be generalized to healthy patients as well as to patients with breathing impairment.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{3.14}$$

In Figure 3.8 the fitted model functions were displayed for three breaths of a patient, recreate the volumetric capnogram behind them very well, keeping only the useful information and ignoring the unnecessary variability of the signal.

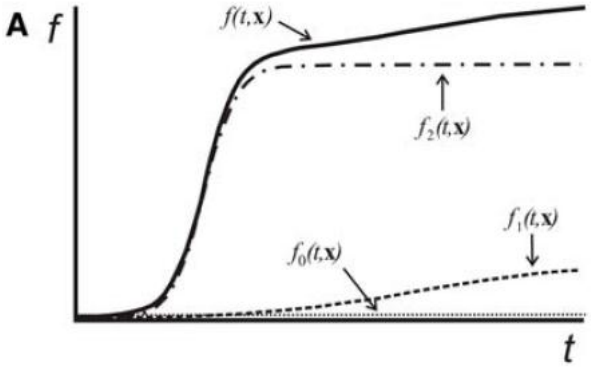


Figure 3.7. Parameterization of the volumetric capnogram as functions of CO_2 and Vexp. The parameters vector is denoted "**x**" and V_{exp} is defined by "*t*", whereas *f* is the concentration of CO_2 (adopted from [14])

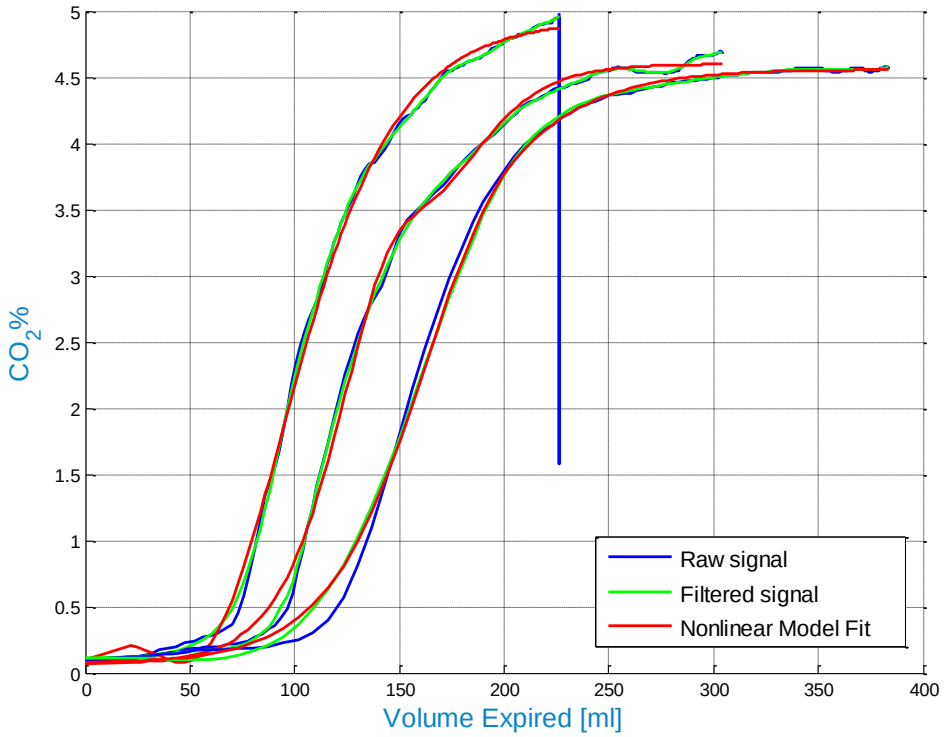


Figure 3.8. Levenberg - Marquardt algorithm for non-linear fitting of the volumetric capnogram. Three overlaid breaths from a subject with CHF are presented with their corresponding filtered signal and non-linear model fit

3.3 Algorithm for Time-Independent Features

In the current study not all the CHF patients had the same number of respiratory measurements because of the different hospitalization times. The CHF patients underwent a specific treatment corresponding to their medical background. Also, the patients were admitted to the hospital with different degrees of decompensation, i.e. a deterioration of the CHF condition. Finally, not all of them responded to treatment with the same rate of improvement. Hence, due to all this factors, the patients had varying hospitalization periods. In order to overcome this particularity when employing a classification algorithm, a method to extract time-independent features from the volumetric capnograms was implemented.

To begin with, the first step consisted of extracting a single parameter from a respiratory measurement. A single parameter value was defined by the median value of the selected feature across one measurement. For each day of measurement, a set of parameters were extracted for a set of chosen features. The only exception to the parameter extraction method was for the features V_{exp} and RR, where the average of all breaths was used as a parameter instead of the median. This metric was preferable due to the inversely proportional relationship between the V_{exp} and the RR [breaths/min].

Secondly, the parameter values extracted from each day of measurement of a patient were decomposed into four features. This way, the whole measurement period was reduced to only four features, no matter the number of parameter points available. Consequently, the mean value $\bar{x}$, the standard deviation σ_x, the slope b_x and the y-intercept a_x of time-depending regression line ($x_i = b_x \times t_i + a_x$) form the 4 features set describing a single parameter [7]. The general formulas describing these features are expressed below, where n denotes the number of measurement points (respiratory measurements), x the corresponding parameter extracted from the feature values and t_i is the corresponding measurement time:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{3.15}$$

$$\sigma_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{3.16}$$

$$b_x = \frac{n\sum_{i=1}^{n}(t_i \times x_i) - (\sum_{i=1}^{n} t_i) \times (\sum_{i=1}^{n} x_i)}{n\sum_{i=1}^{n} t_i^2 - (\sum_{i=1}^{n} x_i)^2} \tag{3.17}$$

$$a_x = \frac{\sum_{i=1}^{n} x_i - b\sum_{i=1}^{n} t_i}{n} \tag{3.18}$$

This procedure for obtaining time-independent features was repeated for all the features extracted in Chapter 3.2.1. The set of four features removes the time-dependence of the parameter values making the data of patients with variable measurements time comparable and suitable for the classification algorithm reported in Chapter 6. The advantage of this method is that it can be used independently of the measurement period on all datasets of different patients because it always returns a number of 4 features describing the evolution of the respiratory status [7]. It is assumed that all patients have more than 1 measurement otherwise this method will return a zero value for the slope and the standard deviation of the parameter, making the feature set useless for the purpose of the classification analysis.

4 Feature-based Classification. Decision Trees Methods

The field of machine learning is inspired from human reasoning for constructing an artificial intelligent algorithm that could learn from data the way humans do. Based on the discovered properties learned from the data, the machine learning algorithm tries to use this knowledge to make future predictions about the data behavior. The learning problems that arise are grouped in two main categories known as supervised and unsupervised learning. In supervised learning, the algorithm learns about the data from a training set with known response values and based on this information constructs a model that makes predictions about future data outcome values. In unsupervised learning there is no outcome measure as the goal is to describe the patterns among the input data.

This chapter explains the basic machine learning principles for supervised learning with a focus on the "Tree-based" classification methods that will be applied in Chapter 6. In Chapter 4.1 the basis for constructing a classification algorithm and cross-validation methods for assessing its performance are given. Next, in Chapter 4.2 the types of "Tree-based" methods with a focus on classification trees and their properties are thoroughly discussed. In Chapter 4.2.1 methods for assessing the performance of a decision tree classifier are discussed.

4.1 Background

For a feature approach analysis a machine learning classification algorithm could be applied. To this end, classification is often carried out by first decomposing the available dataset into features that are quantifiable and classifiable. The features are then partitioned into training and test set to evaluate the performance of a certain classifier.

In supervised learning, a labeled feature set, i.e. a feature set whose class is decided beforehand through one or several gold standard methods, is used as an input for the algorithm to learn the model. This is considered the training part of the model. The learning procedure is inductive which means that if a model estimates well the labels for a satisfactory amount of data, then it will also estimate well the labels for future unobserved data. Next, the model is tested using the test feature set to assess the truth of the predicted class labels. In the procedure described in Figure 4.1 a learning algorithm is applied to create a classifier using a training set. The classifier is then used to predict labels for a test set.

To get a better sense of the predictive accuracy of the classifier for new data, the model has to be validated. In the ideal situation, if there was enough data, a validation test set would be set aside and used to estimate the quality of the classification as shown in Figure 4.1. Since the data is usually sparse, this method is not applicable. For small datasets, with few observations, is not feasible to hold out a partition of the data to use as a test set. K-fold cross-validation is the solution to this limitation. This method directly estimates the expected

sample error, which is the average generalization error when the prediction model is applied to an independent test sample [25, p. 241]. Every "K-fold" method uses models trained on in-fold observations to predict labels for out-of-fold observations [26]. The dataset is split into *K* roughly equal sized-parts, also known as subsets or folds.

For example, by taking $K = 5$, every training fold contains roughly $K-1$ parts of the data, which makes 4 folds for the given example. Every test fold contains $1/K$, which is 1/5 parts of the data for this example which is shown in Figure (4.2). The first model is trained on the data with the first 1/5 excluded; the second model is trained on the data with the second 1/5 excluded, and so on. The labels for each observation contained in the test fold are computed by using the model trained without this fold [26]. Thus, *N* different models are produced, each one having the ability to be tested against an independent subset of the data [27]. The average performance of these *N* models is an accurate estimate of the performance of the original model (created by using the entire dataset), on a prospective independent test set [27].

Leave-one-out cross-validation (LOOCV) is a special case of cross-validation when *K* is equal with the number of observations from the dataset. In LOOCV each training set is created by taking all observations except one by turns. The left out observation is the test set. For *N* observations, *N* different training sets with *N* different test sets are created. In Figure 4.2 a cross-validation example for $K = 5$ is shown. The number of validation rounds is equal to the number of folds. For LOOCV the folds in Figure 4.2 are replaced with single observations. The average misclassification error is calculated with the misclassification error of each round. This average misclassification error is an excellent estimate of the performance of the initial model created on the entire dataset.

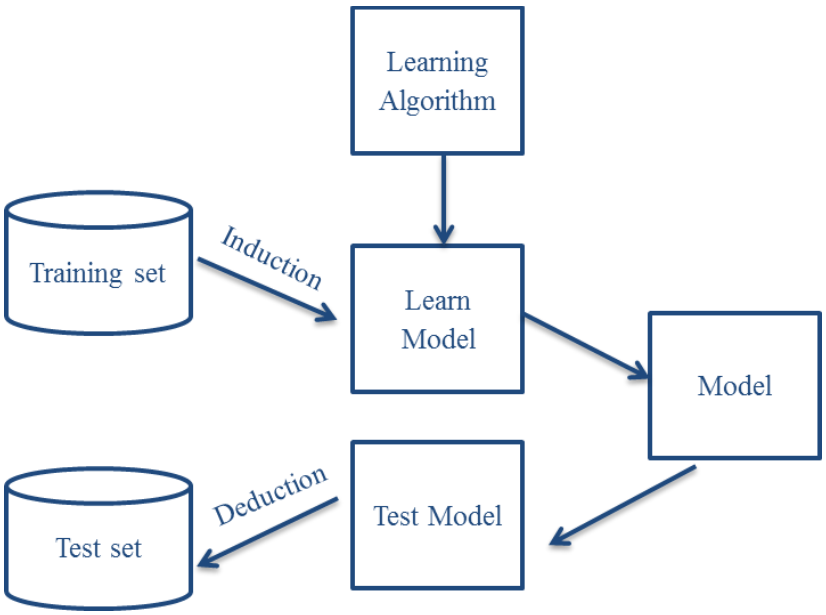


Figure 4.1. Supervised learning method procedure

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Round 1	Validation	Train	Train	Train	Train
Round 2	Train	Validation	Train	Train	Train
Round 3	Train	Train	Validation	Train	Train
Round 4	Train	Train	Train	Validation	Train
Round 5	Train	Train	Train	Train	Validation

Figure 4.2. Cross-validation example with $K = 5$ for the number of folds. Each fold represents a subset from the data. In the LOOCV, each fold represents a single observation

4.2 Classification Tree-Based Methods

In this thesis the aim is to automatically discriminate between, CHF patients with pulmonary edema, CHF patients without pulmonary edema and healthy subjects from volumetric capnography data. As the medical condition of each patient was established through several examination methods which are accepted as gold standards, the problem posed was to classify categorical dependent variables. A categorical variable is defined as a variable that has two or more categories as possible outcomes, but there is no inherent ordering to the categories [28]. The proposed solution for this problem consisted in applying decision trees. The main advantage of decision trees is that they have a non-parametric nature which means that they do not make any assumptions about the distribution of the data [27]. Additionally, they are quickly implemented and effective in creating clinical decision rules which are easily understood by everyone.

Decision tree inducers are algorithms that automatically construct a decision tree from a given dataset. Classification trees and regression trees predict labels or response for the data. Classification trees give responses that are nominal, such as “class 1” or “class 2”, “true” or “false”. Regression trees give numeric responses. The CART (classification and regression tree) algorithm as described in Breiman et al. (1984) was applied in this work using the MATLAB® Statistics Toolbox.

Classification trees are a binary classification method which predicts responses for the data, i.e. the input predictor variables. In order to predict the response for an instance or observation, the classification tree is followed top-down from the root node (the beginning node) through every branch which fulfills the criteria of the observation until a leaf node (terminal node) is reached. The leaf node on which the observation lands, contains the response label. Classification trees divide the data based on the input predictor values. Each node from tree, aside from the leaf nodes, is split into exactly two child nodes. The split is made based on the feature values. Each path from the root of the tree to one of the leaf nodes represents a set of rules of the form “if $x <$ value then $node_i$ else if $x \geq$ value then $node_{i+1}$”.

Tree-based method partition the feature space into a set of rectangles and then fit a simple model, like a constant in each one[25, p. 305]. The main idea behind a binary tree is to form the tree and then minimize the error in each leaf of the tree. A CART regression problem with continuous response Y and inputs X_1 and X_2, each taking values in the unit interval is considered in Figure 4.3 The feature space is partitioned by lines that are parallel to the coordinate axes in such a way that a certain type of error is minimal [25, p. 305] as shown in Figure 4.3, top-left picture. In each partition element, Y could be modeled with a different constant. However, there is a problem: although each partitioning line has a simple description like $X_1 = c$, some of the resulting regions are too complex to be described [25, p. 305]. To simplify matters, recursive binary partitions are chosen as exemplified in the top right panel of Figure 4.3. The feature space is first split into two regions, and the response is modeled by the mean of Y in each region. In case of a regression problem, the mean of Y is chosen; otherwise, for classification, the label of the majority class would be selected. Next, the variable and split-point are chosen in order to achieve the best fit with the minimum sum of squared distance error. Then, one or both of these regions are split into two more regions, and this process continues until some stopping rule is applied. For example, in the top right picture from Figure 4.3 the first split is made at $X_1 = t_1$. Then the region $X_1 < t_1$ is split at $X_2 = t_2$ and the region $X_1 > t_1$ is split at $X_3 = t_3$. Finally, the region $X_1 > t_3$ is split at $X_2 = t_4$. The result of this process is a partition into five regions (R_1, …, R_5) as shown in Figure 4.3 [25, p. 305]. The same model can be represented by the binary tree in the bottom left picture from Figure 4.3. Observations fulfilling the condition at each connection are assigned to the left branch, and the others to the right branch. The terminal nodes or leaves of the tree correspond to the regions R_1, …, R_5 [25, p. 305].

In Figure 4.4 a partition of the feature space is depicted for a classification tree with two nominal responses, “0” and “1”. The only difference from the algorithm of a regression tree is that instead of the mean of the points from a region, the dominant class label from the respective region gives the response. For example, a data point that will fall into the first region R_1, for which the split was made at $X_1 < t_1$, will have the class label “1”.

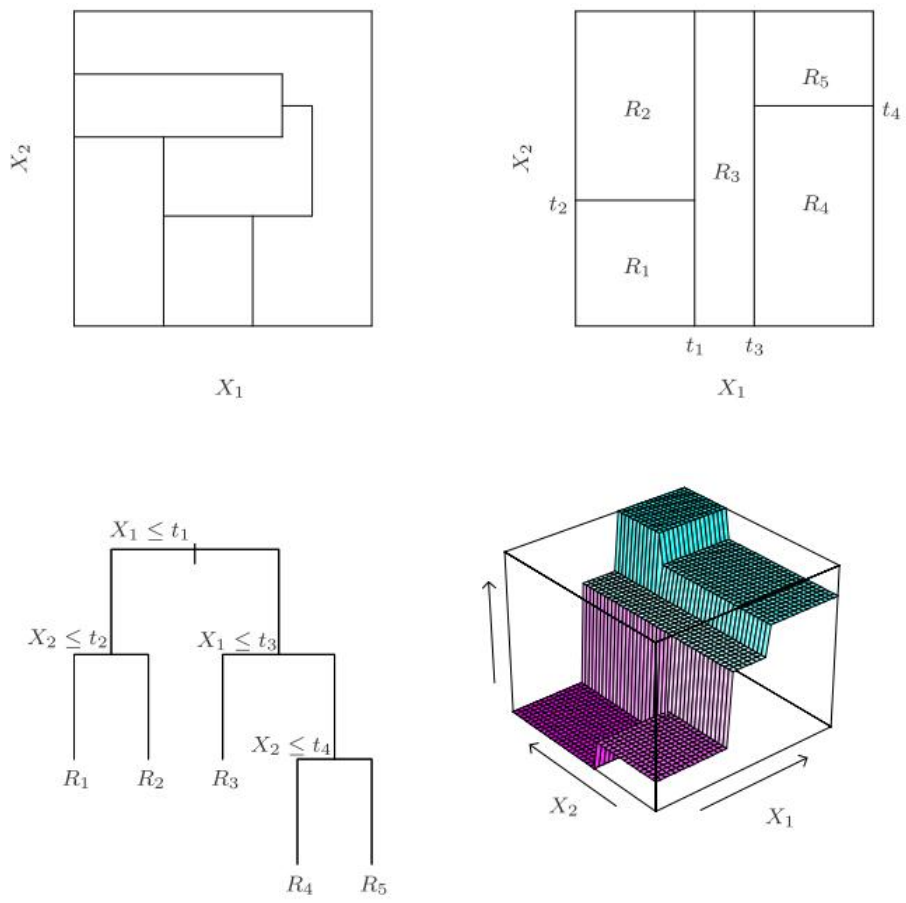


Figure 4.3. CART regression tree example. Top right figure shows a partition of a two-dimensional feature space by recursive binary splitting applied to some invented data. Top left panel shows a general partition that cannot be obtained from recursive binary splitting. Bottom left panel shows the corresponding tree, and a perspective plot of the prediction surface appears in the bottom right panel (adopted from [25, p. 306])

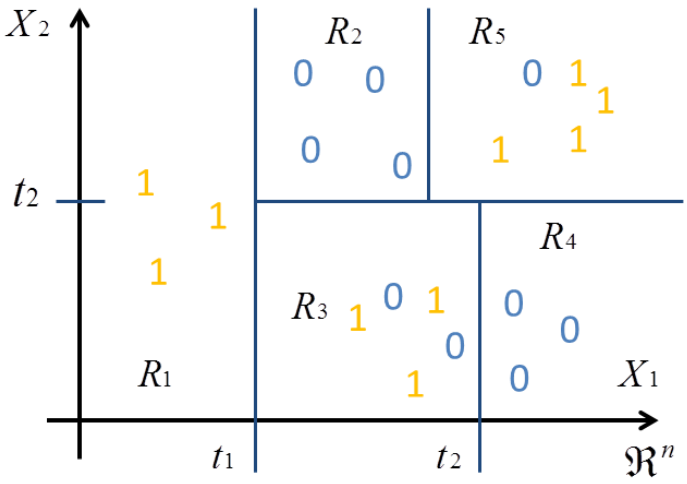


Figure 4.4. CART classification tree example. The two class labels are represented by "1" and "0". The feature space was split into five regions (R_1, …, R_5) in such a way that a region contains only one type of label

The following steps are performed in order to grow a decision tree [29]:

1. Start with all input data and examine all possible binary splits on every predictor
2. Select a split by applying the best optimization criterion
    - Regression trees: Mean squared error
    - Classification trees: Gini's diversity index
3. Impose the split
4. Repeat recursively for the two child nodes
5. Stop growing the tree if any of the following criteria are satisfied:
    - The node is pure. For classification, a node is pure if it contains only observations from one class
    - The size of this node is less than the allowed minimal size of a parent node
    - The maximal allowed number of leaf nodes has been reached
    - The maximal allowed tree depth has been reached

In order to grow the entire decision tree, a greedy heuristic "divide and conquer" algorithm is applied. For this particular case, the heuristic is to find the best split by minimizing the chosen measure for node impurity. At start, all the training examples are at the root of the tree and then it starts to grow top-down in a recursive manner by replacing all the leaf nodes with simple nodes. All the nodes are created through a series of optimally local selections of selected predictors on which the partitioning or splitting of the dataset is made. The aim is to partition the dataset in subsets more and more homogenous so that the impurity is minimal.

Further on, the definitions of a few of the properties of a classification tree are listed in the Table 4.1.

Table 4.1. Classification trees properties

Parameter Name	Definition
Gini's diversity index	"Gini's diversity index" is a measure of node impurity. The Gini index of a node has the formula: $$I_G(p) = 1 - \sum_i (p_i)^2 \quad (4.1)$$ where the sum is over the classes *i* at the node, and p_i is the observed fraction of classes with class *i* that reach the node. A node with just one class (that is a pure node) has the Gini index equal with 0; otherwise the index is positive. A node with maximum impurity has the Gini index equal with 0.5 which means that the classes are equally distributed within the node [30].
Node Probability	Let t_0 be the parent node. The left and right child nodes are then denoted by t_L and t_R. Let N_0, N_L and N_R be the number of training observations in the parent, left, and right node, respectively. The probability of a node is defined by a fraction of training observations landing on this node, $P(t) = N_t / N,\ t \in \{0, L, R\}$. N is the total number of observations in the training set [31, pp. 307–317].
Node posterior probability	For simplicity, the case of two classes, A and B is considered. If node *t* contains N_A observations of class A and N_B observations of class B, the class posterior probability is estimated by putting $P(A/t) = N_A / N_t$, and similarly for B. The sum of posterior probabilities equals one [31, pp. 307–317].
Score	The score of a classification of a leaf node is the posterior probability of the classification at that node. The posterior probability of the classification at a node is the number of training sequences that lead to that node with the classification, divided by the number of training sequences that lead to that node [32].
Node Error	The misclassification probability for a given node [30].
Classification error	The default classification error is the fraction of the training data that the tree misclassifies [33]. The classification error implements "the minimum cost function". This function predicts the label with the smallest expected misclassification cost, with expectation taken over the posterior probability, and cost as given by the square *Cost matrix* property of the classifier [33]. The value *Cost(i,j)* is the cost of classifying a point into class *j* if the true class is *i*. The loss is then the true misclassification cost averaged over the observations [33].

4.2.1 Methods for Assessing Classifier Performance

The performance of the classification tree was assessed through several metrics which include the resubstitution error, the cross-validation error, sensitivity and specificity. The reproducibility of the classification-tree was analyzed by resubstitution error and the leave-one-out cross validation error. These methods, which were tailored for decision trees are described next.

Resubstitution error

Resubstitution error is the difference between the response training data and the predictions the tree makes of the response based on the input training data [29]. If the resubstitution error is high, the predictions of the tree are expected to be weak. Still, having low resubstitution error does not guarantee good predictions for new data [29]. Resubstitution error is often a too optimistic estimate of the predictive error on new data because the same training data was used for learning [29].

Leave-one-out cross-validation (LOOCV)

In order to study how the model tree obtained will generalize to an independent dataset, the leave-one-out cross-validation error was used. Cross-validation is used when there is no test set available to validate the model tree. Thus, to get a better sense of the predictive accuracy of the classification tree for new data [29], the classification trees implemented in this thesis will be evaluated by leave-one-out cross-validation in Chapter 6. Leave-one-out cross-validation was chosen because the available dataset contained very few observations.

Cross validation splits the training data into K parts or folds at random. For LOOCV, K is equal with the number of observations from the training dataset. Hence, cross-validation trains K new trees, each one on $K-1$ parts of the data [29]. The predictive accuracy of each new tree is then examined on the data not included in training that tree. This method gives a good estimate of the predictive accuracy of the original tree, since it tests the new trees on new data [29].

Moving on to the topology of the tree, if the tree is leafy, i.e. has many terminal leaf nodes, a low training error (resubstitution error) is achieved, but its generalization (test) error is often large [29]. In other words, a deep tree with many leaves is usually highly accurate on predicting the training data, but is not guaranteed to show a comparable accuracy on a future independent test set [29]. In contrast, a shallow tree, with few leaves does not obtain high training accuracy [29]. However, a shallow tree can be more robust and its training accuracy could be close to that of a representative test set [29]. Additionally, a shallow tree is easy to follow and interpret in a clinical environment.

Receiver Operating Characteristics (ROC)

Receiver Operating Characteristics (ROC) graphs are a useful technique frequently employed in clinical decision making for visualizing, organizing and selecting classifiers based on their performance [34]. A ROC curve is a graphical plot which shows the performance of a binary classifier for different thresholds of the classifier output. It is created by plotting the true positive rate (TPR) versus the false positive rate (FPR) at different

threshold values. TPR is also known as sensitivity and FPR as (1- specificity). A ROC curve is used mainly to find the threshold that maximizes the classifier prediction rate or to assess how the classifier performs in regions of high sensitivity and high specificity [35].

A classification model is a mapping between each instance to one of the elements from the predicted classes [34]. A classifier result could give a continuous output or it can be a discrete label indicating one of the classes. Considering a binary classification problem, with outcomes positive (p) and negative (n) and an instance, there are four possible outcome values. If the instance is positive and is classified as positive, then it is a "true positive" (TP). More concretely, a TP value is achieved when a sick person is correctly classified as sick. Next, if the instance is negative and is classified as negative, then is called "true negative" (TN). A TN value stands for a healthy person correctly identified as being healthy. However, if the instance is classified as negative, but its actual value is positive, then is counted as a "false negative" (FN). In the real world, a FN value represents a sick person who was incorrectly diagnosed as being healthy. A "false positive" (FP) value appears when the instance is classified as positive when its real value is negative. FP refers to a healthy person incorrectly identified as sick [34].

Given a classifier and a set of instances, a two-by-two confusion matrix (also called a contingency table), shown in Figure 4.5, can be constructed to represent the outcomes of the set of instances [34]. This matrix forms the basis for many common metrics.

		True Class	
		Positive	**Negative**
Classifier Outcome	**Positive**	True Positive	False Positive
	Negative	False Negative	True Negative

Figure 4.5. Confusion Matrix as defined by [34]

The two common performance metrics used in plotting a ROC curve that could be defined from it are sensitivity and specificity. Sensitivity refers to the percentage of actual positives which are correctly classified as positive. Specificity measures the proportion of actual negatives which are correctly classified as negative.

$$Sensitivity = \frac{TP}{TP + FN} \tag{4.2}$$

$$Specificity = \frac{TN}{TN + FP} \tag{4.3}$$

A ROC curve could be applied on any classifier that returns a numeric score for an instance of input data. In the case of decision trees, the score represents the posterior probability of observing an instance of the positive class at a certain point, i.e. the fractions of

positive observations in a leaf of a decision tree [35]. Scores yield values into the range [0-1] and scores from the positive and negative class add up to unity [35]. By the convention adopted here, a high score returned by a classifier on any given instance is likely to belong to the positive class and a low score suggests that the instance is likely from the negative class [35].

Such a ranking or scoring classifier could be used with a threshold to produce a discrete binary classifier [34]. If the classifier output is above the threshold it gives a positive response, otherwise a negative response [34]. Each threshold value produces a different point in the ROC space [34]. For each threshold, TP is the count of true positive observations with scores greater or equal to this threshold and FP is the count of FP observations with scores greater or equal to this threshold [36]. The negative counts, TN and FN are expressed in a similar way [36]. The thresholds are then sorted in the descending order which corresponds to the ascending order of positive counts [27]. By convention, the first threshold value corresponds to the highest "reject all" threshold and the corresponding values for (1-specificity) and sensitivity are computed for TP=0 and FP=0 [36]. The last threshold value is the lowest "accept all" threshold for which TN=0 and FN=0 [36].

Several points in the ROC spaces have a significant meaning. ROC graphs are two-dimensional graphs in which the TP rate or sensitivity is plotted on the Y axis and FP rate (1-specifificity) is plotted on the X axis. The lower left point (0, 0) represents the strategy of never returning a positive classification; such a classifier assumes no false positive errors but also gains no true positives [34]. In other words, (0, 0) are the coordinates where the classifier finds no positives instances so that it always gets the negative cases right but it gets all positive cases wrong. The opposite strategy, of unconditionally issuing positive classifications, is represented by the upper right point (1, 1) [34]. For the second coordinate point (1, 1), everything is classified as positive so the classifier gets all positive cases right but it gets all negative cases wrong [37]. One point in the ROC space is better than another if it is located more to the northwest (TP rate is higher, FP rate is lower, or both) than the other one [34]. The point (0, 1) represents perfect classification and this is the result all classifiers are aiming for.

The diagonal line y = x represents the random guess performance. For example, if a classifier guesses the positive class half the time, it can be expected to get half the positives and half the negatives correctly which produces the point (0.5, 0.5) in the ROC space [34]. Hence, a random classifier will produce a ROC point that slides "back and forth" on the diagonal [34]. Any classifier that appears in the lower right triangle performs worse than random guessing. Figure 4.6 shows a basic ROC graph with four classifiers labeled A to C'. Each classifier in is a discrete classifier, i.e. it only returns a class label. The red line represents the random guess output.

The area under the ROC curve, abbreviated AUC is a two-dimensional representation of classifier performance [34]. For an easier comparison between classifiers' performance, the ROC space is reduced to a single scalar value representing the expected prediction performance [34]. Since the AUC is a portion of the area of the unit square, its value will always be between 0 and 1.0. However, because random guessing produces the diagonal line between (0, 0) and (1, 1), which has an area of 0.5, no classifier with

discriminant power should have an AUC less than 0.5 [34]. The AUC has an important statistical property which says that is equivalent to the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance [34]. The AUC is calculated by trapezoidal approximation [36].

Youden's point is a metric that captures the performance of a diagnostic test through a single point which is the optimal operating point of the ROC curve [36]. The general formula used for Youden's point involves the slope "*S*" defined in Equation (4.4), where *cost(i|j)* represents the cost of assigning an instance of class "*j*" to class "*i*" and the total instance counts in each class are given by $P = TP + FN$ and $N = TN + FP$ [36]. The optimal operating point is found by moving the straight line with the slope "*S*" from the upper left corner of the ROC plot, i.e. (FPR = 0,TPR = 1) down and to the right until it intersects the ROC curve [36]. Usually a classifier is used with a particular threshold that gives the optimal operating point on the ROC curve. The best operating point is chosen so that the classifier gives the best trade-off between the costs of failing to detect positives against the costs of raising false positives [37]. These costs need not be equal; however this is a common assumption [37].

$$S = \frac{\cos t(P \mid N) - \cos t(N \mid N)}{\cos t(N \mid P) - \cos t(P \mid P)} \cdot \frac{N}{P} \tag{4.4}$$

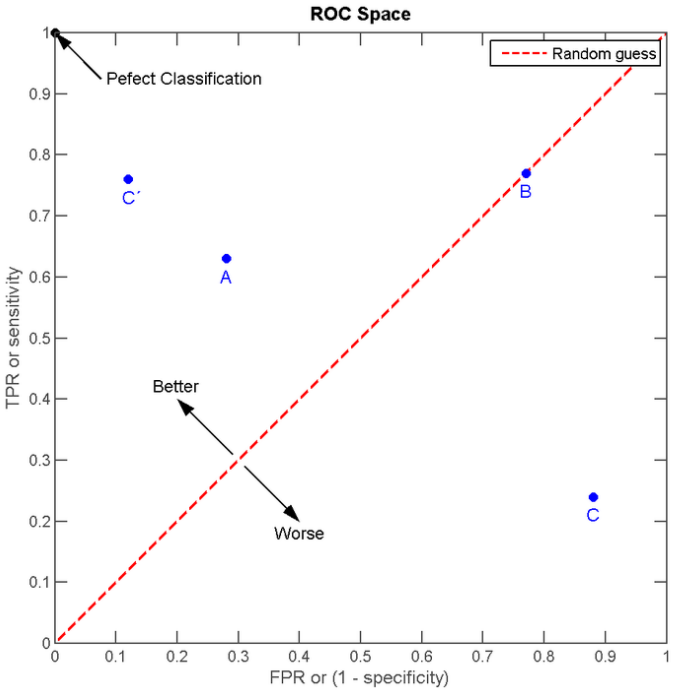


Figure 4.6. A basic ROC graph showing four discrete classifier noted A to C' (source: [38])

5 Volumetric Capnogram Assessment Results

Without any assessment, no method could be proven to be valid and accepted for use in other environments. There is no exception for the volumetric capnogram measurements as they need to be evaluated in order to check their viability: that they could serve as a reliable indicator of the respiratory status. To this aim, visual approaches and quantifiable methods to assess the volumetric capnogram are considered.

In Chapter 5.1.1 the reference values of the extracted capnogram features are compared with typical values obtained in other studies. Next, in Chapter 5.1.2 a comparison between different respiratory statuses is discussed also through visual means. The features visualization might be very versatile for clinicians who could spot out abnormal behaviors in the respiratory trends very quickly. Moreover, a single reference value adds discriminant power to a visual representation, thus setting certain expected values ranges, given the respiratory condition of the patient.

Chapter 5.2 presents the relations between features that describe the functioning of different parts of the respiratory apparatus. A power-two model detailed in Chapter 5.2.1 and a linear fit model presented in Chapter 5.2.2 are used to express the correlation between the volume expired and slope phase III. In Chapter 5.3 these model functions are further exploited to lead to the creation of a single breath mathematical model that describes a set of measurements. Figure 5.1 shows a logical diagram with the topics presented in this chapter.

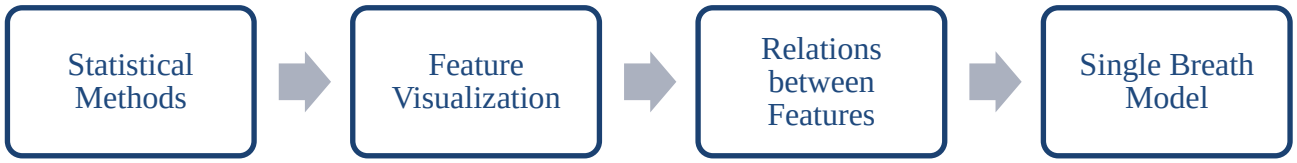


Figure 5.1. Analysis steps for assessing the volumetric capnograms

5.1 Statistical Methods

After the feature extraction step has been performed the values obtained have to be assessed so that their usability for diagnostic purposes is confirmed. To this end, basic statistical methods have been employed on the feature extracted previously from the filtered volumetric capnograms. The statistical methods that have been employed in this work are the ones which are commonly used in scientific journals. As seen in Table 5.1 and Table 5.2, the metrics used to report reference values were the mean ± standard deviation (SD), the median and the inter-quartile range (IQR).

The IQR, computed as the difference between the third quartile and first quartile, is a robust estimate of the spread of the data, since changes in the first quartile and third quartile

of the data do not influence it [39]. If there are outliers in the data, then the IQR is more representative than the standard deviation as an estimate of the spread of the body of the data [39]. The coefficient of variation (CV) is a normalized measure of dispersion of a probability distribution which is computed as the ratio of the standard deviation of the sample population to the population mean [40]. The benefit of the CV is that is a dimensionless number so it could be used for comparison between data sets with different units or different means [40].

The confidence interval (CI) gives an estimated range of values (interval) which is likely to include an unknown population parameter [41]. The estimated range is calculated from a given set of sample data [41]. The level of confidence of the confidence interval would indicate the probability that the confidence range captures this true population parameter given a distribution of samples, however it does not describe any single sample [42].

A simple way to check the outcome of the outlier discarding method explained in Chapter 3.1.1 is to look at the boxplots of the expired volumes of each patient (see Figure 5.2.). After the invalid breaths were removed, fewer outliers occur (red symbols, see Figure 5.2. right) and it could be seen more clearly that the medians of each patient for expired volumes differ. The advantage of using boxplots is that they are non-parametric, showing dissimilarities between populations without making any assumptions of the underlying statistical distribution [43].

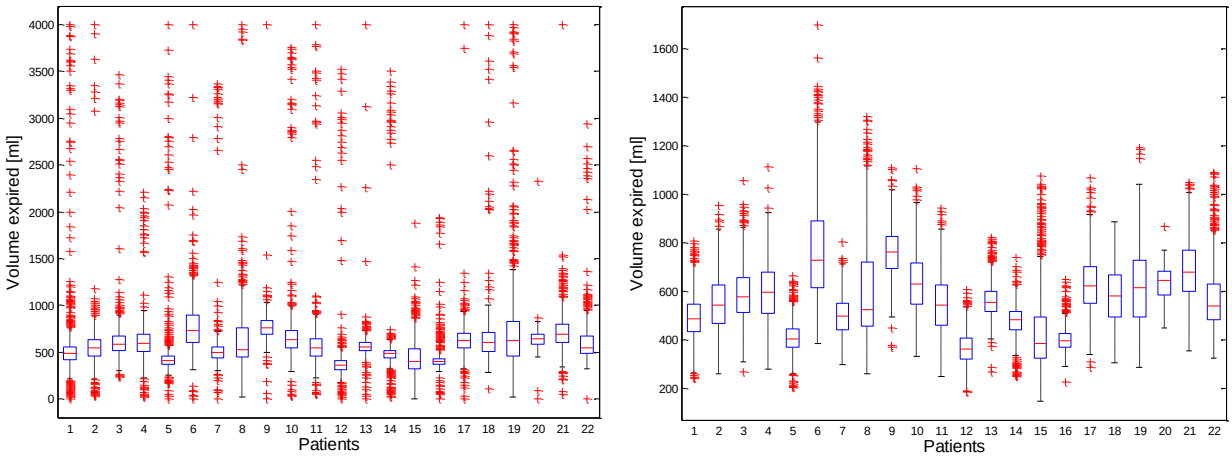


Figure 5.2. Boxplot comparison before (left picture) and after (right picture) discarding the invalid breaths from the measurements

5.1.1 Reference Values

The reference values that have been reported here could be used in a clinical setting to interpret cardiopulmonary conditions. The reference values of healthy patients could be used in contrast with the values of CHF patients.

The reference values extracted were compared with typical values from other reference studies of volumetric capnography. The results obtained were in concordance with those found in related literature. A discussion explaining the possible significance of each feature in relation with the respiratory morphology was carried out. Another aspect that renders the reference values obtained in this thesis useful is that there is a lack of information

concerning typical values for volumetric capnography. The reference values were extracted from 31 CHF patients and 21 healthy patients.

Starting with healthy patients, in Table 5.1, the median value for the expired volume (V_{exp}) is of 514.90 ml and the mean ± SD is 537.09 ±155.13 ml. These values are in concordance with those already found in literature, with Astrom et al. [11], giving a value of 648 ± 146 ml (mean ± SD) for the expired tidal volume of one breath and Tusman et al. [15], giving a value of 546(203) ml (median and inter-quartile range).

The $ETCO_2$ median value obtained was 5.38 [%], ($ETCO_2$ Table 5.1), a good value for healthy patients comparing with Tusman et al.[15], who obtained a value of 4.73 [%], for the selected barometric pressure 760 [mmHg].

Slope phase III (S_{III}, Table 5.1), the part of the capnogram that reflects the inhomogeneities of the lungs should be as flat as possible given a healthy lung. The median value obtained for S_{III} is 1.68 [%/L], the mean ± SD is 2.08 ± 1.64 [%/L]. In literature, Tusman et al. [15], obtained a normalized value of 0.26(0.17) [1/L] (median and interquartile range). In an earlier presentation of Tusman's [44], the following value is given for the mean of S_{III}, 0.009 [mmHg] which converted in [%/L] is 1.18 for a chosen barometric pressure of 760 [mmHg].

Another aspect worth mentioning is the high coefficient of variation (CV) we obtained in the case of S_{III}, 78.91 % which indicates a high dispersion in the values. Unfortunately, it seems that in literature there is a lack of data concerning reference values for S_{III} or a comprehending study dedicated in particular to its analysis, the most popular and accepted studies focusing on the volume of airway deadspace or $ETCO_2$ values.

Slope phase II (S_{II}, Table 5.1) is 50.63 (median) [%/L], close to the value found by Tusman et al. [15], who obtained 0.43[mmHg/ml] for the median which becomes 56.57 when converted to [%/L] with a chosen 760 [mmHg] barometric pressure.

The volume of airway deadspace (VD_{aw}, Table 5.1) is 132(50.8) ml or 135.28 ± 29.91 [ml] in healthy patients, these values being similar with those obtained by Astrom [11] 100 ± 14 [ml] for females and 136 ± 23 [ml] for males. Tusman et al [15] found a value of 92(39) ml for the median and interquartile range. A higher value, with pooled values from both genders, was obtained by Verschuren et al. [12], with 178 (mean) [ml].

Further on, the Alpha Angle [°] (α, Table 5.1), the angle at the slope II and slope III intersection has a median value of 150.51°, a value comparable with that of Tusman et al. [15] who obtained 157° .

Moving on Table 5.2, reference values for patients with CHF are presented. In literature there have not been reported so far any reference values for patients with this condition so this group was analyzed in contrast with the reference values of healthy patients to see if there appear any differences. As expected, the values obtained in Table 5.2 are a confirmation of the restrictive respiratory condition existent in the patients.

Thus, starting with the V_{exp} value, this feature presents smaller values both for the median and the mean of the population group compared with the values of the healthy group

(457.50 [ml] versus 514.90 [ml] for the median values and 508.23 [ml] versus 537.09 [ml] for the mean values). The respiration rate is much higher, 21.25 [breaths/min] which could be a cue for difficult breathing that needs a lot of effort. Next, the lower $ETCO_2$ value (4.55 [%] compared with 5.42 [%] for the mean value) points out to two particularities: a lower area under the curve (the capnogram) and less CO_2 produced in the expiration. A higher ventilation/perfusion (V/Q) ratio will cause a decrease in the amount of CO_2 which will lead eventually to an increased slope of the phase III (S_{III}), as the perfusion in the lungs decreases. There is quite a large difference between S_{III} of healthy patients (1.68 [%/L] for the median) and the S_{III} of the patients with CHF (3.49 [%/L] for the median). The greater value of the S_{III} is a good indicator of the status of the alveoli which are flooded by fluid from the capillaries when the CHF is present.

Slope of the phase II (S_{II}) is much smaller than that of the healthy patient (38.94 [%/L] versus 50.63 [%/L] for the median values). Finally, the values of the slopes in CHF lead to greater Alpha Angle by approximately 15° (see Table 5.2, last column). The volume of airway deadspace (VD_{aw}) also increases (151.10 [ml]) compared to the healthy group (132 [ml]).

All in all, these references value represent the first glimpse into discriminating healthy patients from those with CHF by showing comparable differences between the main volumetric capnography features. The values obtained confirm the fact that alterations in the cardiopulmonary system are reflected immediately and accurately in the volumetric capnogram features.

Table 5.1 Reference values for volumetric capnogram features in healthy patients

	V_{exp} [ml]	RR [breaths/min]	$ETCO_2$ [%]	S_{III} [%/L]	S_{II} [%/L]	VD_{aw} [ml]	α Angle [°]
Median	514.90	14.56	5.38	1.68	50.63	132.00	150.51
IQR	197.78	4.78	0.58	1.11	20.36	50.80	15.88
Mean	537.09	14.61	5.42	2.08	53.32	135.84	149.83
SD	155.13	4.43	0.44	1.64	25.54	29.91	12.61
Lower 95% CI	534.92	14.55	5.41	2.05	52.96	135.41	149.65
Upper 95% CI	539.25	14.67	5.42	2.10	53.67	136.26	150.01
CV	28.88	30.35	8.13	78.91	47.90	22.02	8.41

Table 5.2 Reference values for volumetric capnogram features in CHF patients

	V_{exp} [ml]	RR [breaths/min]	$ETCO_2$ [%]	S_{III} [%/L]	S_{II} [%/L]	VD_{aw} [ml]	α Angle [°]
Median	457.50	21.35	4.58	3.49	38.94	151.10	165.48
IQR	272.85	7.61	0.92	3.42	25.29	67.60	16.19
Mean	508.23	21.76	4.55	4.10	43.65	154.34	161.64
SD	231.5	7.12	0.71	2.95	34.44	45.53	12.36
Lower 95% CI	500.68	21.53	4.52	4.00	42.52	152.82	161.24
Upper 95% CI	515.78	21.99	4.57	4.20	44.77	155.87	162.04
CV	45.54	32.74	15.64	71.94	78.90	29.50	7.65

5.1.2 Feature Visualization

Feature visualization refers to displaying features in a manner which is easy to comprehend both by researchers and medical personnel. By visually presenting the features extracted, it was aimed to highlight the patterns and trends undergoing in the respiratory process. At a first glance, a picture speaks more than a thousand values and could point to the cause of the respiratory problem. Thus, these means of visualizing data could prove very useful for clinicians who could readily see distinctions among different states of the cardiopulmonary system.

To begin with, in Figure 5.4 and Figure 5.5 a notched boxplot approach was proposed for visualization and comparing statistical data samples [45]. Notched box plots are advantageous for determining if two random samples were drawn from the same population [45]. Similar notches of boxes indicate that the data on which they were based have the same distribution [45]. In Figure 5.4, the boxplot notch for slope phase III (S_{III}) from the measurements of all the patients from both groups was plotted. It can be observed in both figures that the notches of the boxplots for the healthy and for the CHF group do not overlap. According to [46] if two boxes' notches do not overlap there is strong evidence (95% confidence) that their medians differ significantly**.** This result suggests that these extracted features are able to separate different types of respiration patterns. Moreover, the boxplot notch visualization might show immediately which features could have a potential discriminative power when used in a classification algorithm.

Another visualization method implemented is the "plot by day-median" graphic which shows the feature values with the corresponding median value plotted for each measurement. In Figure 5.3 , the median line is depicted as a horizontal red line and the data points (blue squares) represent the values extracted from each breath. This type of representation is useful when it is desired to see what stands behind a single value representing a measurement day. In Figure 5.3, the patient exhibits an increasing trend marked by the red median line in the values of S_{III}.

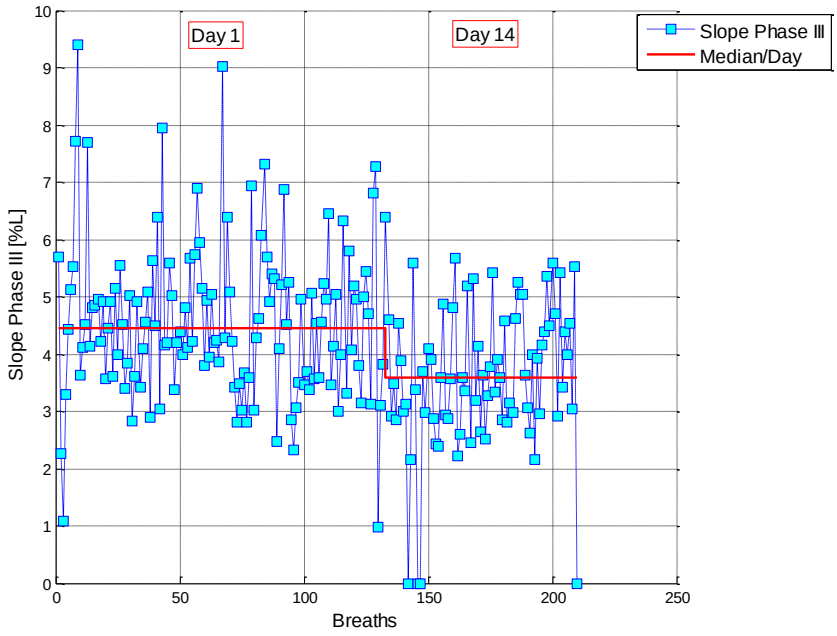


Figure 5.3. Slope Phase III – "Plot by day-median" visualization

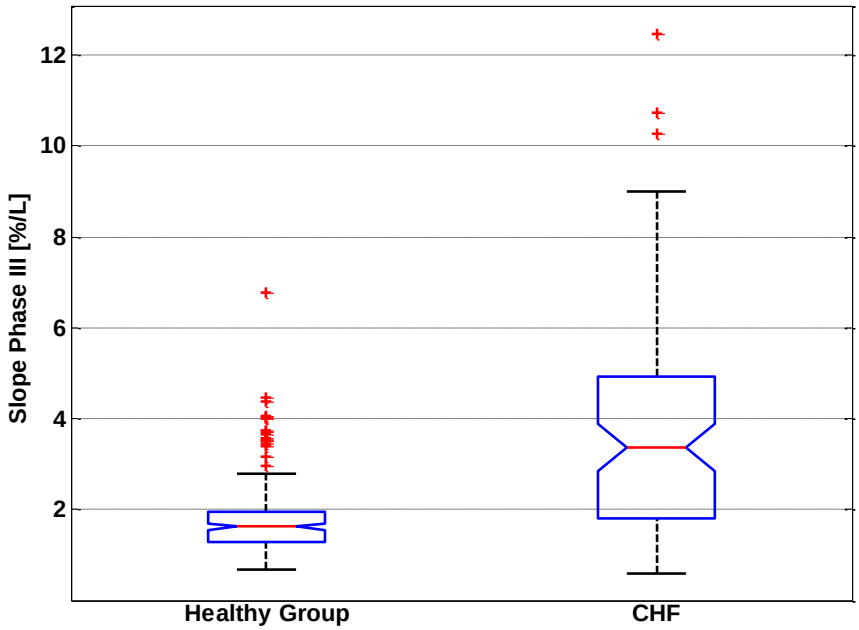


Figure 5.4. Slope Phase III. Boxplot Notch Comparison. The horizontal red midline is the median value, while the horizontal lower blue line represents the 1st quartile and the upper blue line the 3rd quartile. The extreme values represent the maximum and the minimum values

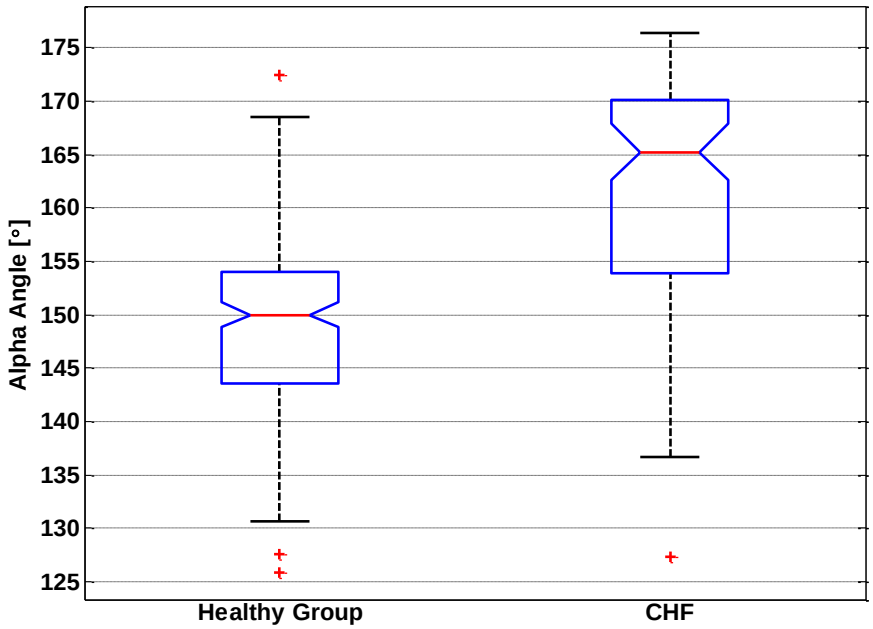


Figure 5.5. Alpha Angle. Boxplot Notch Comparison. The horizontal red midline is the median value, while the horizontal lower blue line represents the 1st quartile and the upper blue line the 3rd quartile. The extreme values represent the maximum and the minimum values

Furthermore, in Figure 5.6 the temporal evolution trends of several volumetric capnography parameters are presented. The median value from each measurement day was extracted and displayed along the corresponding value extracted from the single breath model detailed in Chapter 5.3. This type of visualization is convenient for viewing the general trend of the features. For example, a decreasing trend in the values of the S_{III} might point to an improvement in the respiratory condition of a CHF patient (first upper-left graph, Figure 5.6). Moreover, the last two graphs show a decrease in the values of VD_{aw} and Alpha Angle.

The plots shown in Figure 5.6 are a visual representation of the time-independent features described in Chapter 3.3. Thus, the slope and the intercept of the regression line traced across the median points from a plot were extracted. The median and the standard deviation of the points from a plot were also computed.

For a better visualization of the inversely proportional relation that often appears between V_{exp} and RR, the mean value from a measurement was extracted instead of the median value.

In Figure 5.7 the data of the CHF patient shown in Figure 5.6 is presented as a percentage change over a period of time. The relative change was calculated by subtracting from each feature value the initial value and then dividing the result by the initial value and multiplying by 100. In this type of plot an increasing or decreasing trend in the features is better assessed because the difference from one measurement day to another is shown on the y-label. Therefore, from a percentage-plot view it could be decided whether the change in a value is significant enough to tell if the patient improved or not.

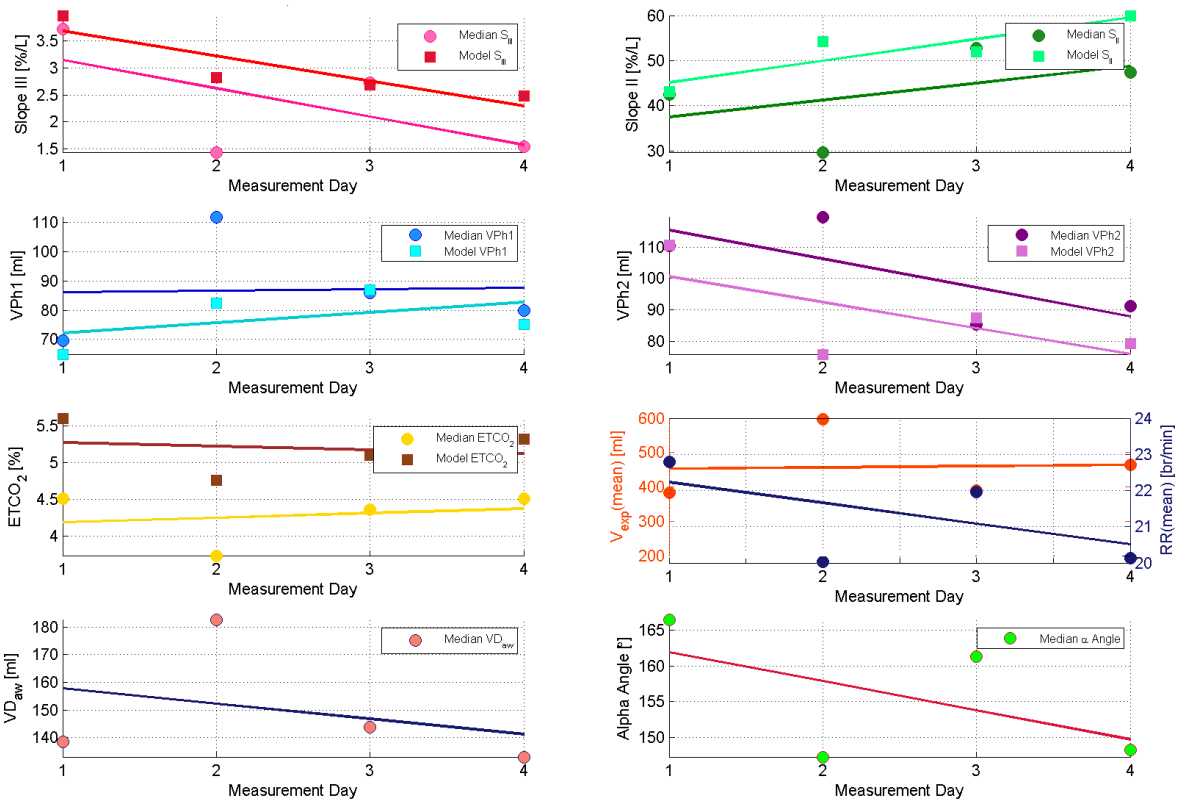


Figure 5.6. Volumetric capnography temporal evolution trends. Each dot represents the median value and each square represents the value extracted from the single breath model described in Chapter 5.3. For each set of points the linear regression fit was plotted

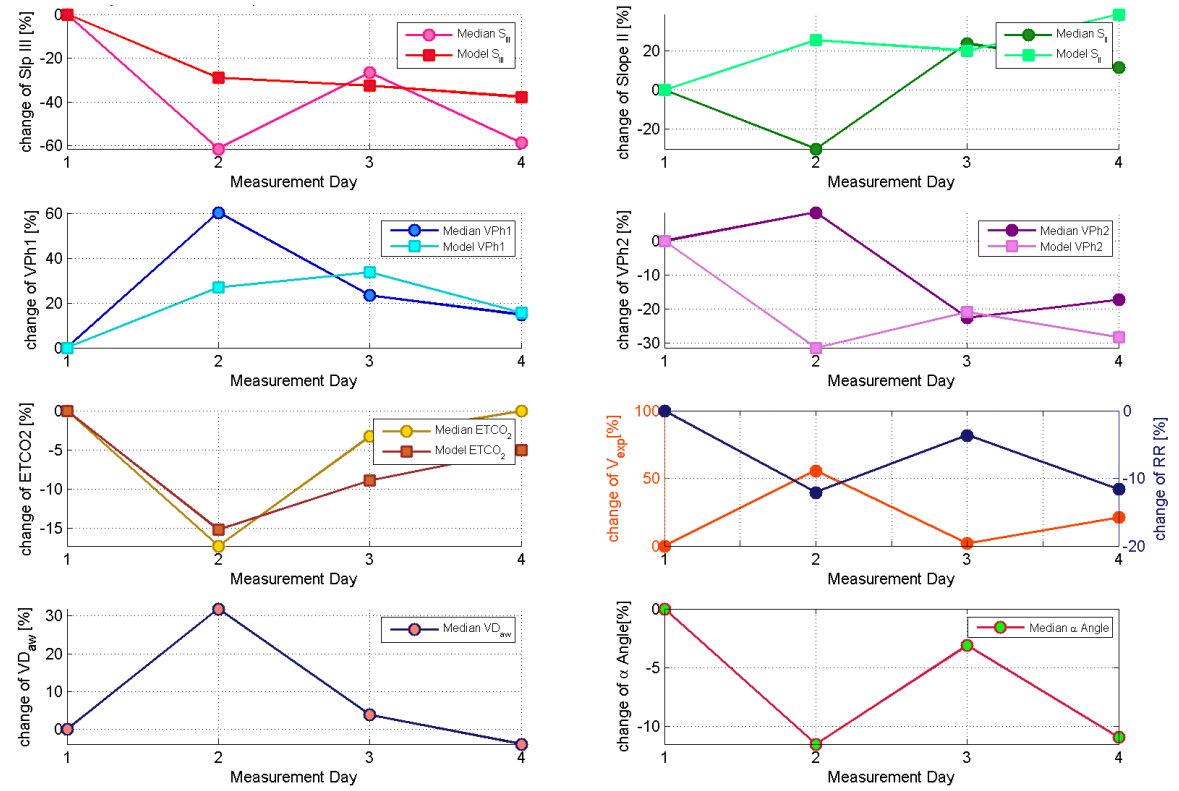


Figure 5.7. Percentage change of several features over a period of time. Each point from a plot represents the relative change reported to the first measurement value

5.2 Correlations Between Features

Functional correlations between volumetric capnographic features were computed by regression analysis to fit data with a model that explains its variability and gives the smallest possible error between the model and the predictor variable. The interactions between features show correlations that could be extrapolated to correlations in the respiratory system. Finding strong correlations in the data that are supported by a successful fitted model (with a high coefficient of determination and a small error) indicate the factors that determine changes in a parameter.

Fitting requires a parametric model that relates the response data to the predictor data with one or more coefficients. The result of the fitting process is an estimate of the model coefficients [47]. To obtain the coefficient estimates, the least-squares method minimizes the summed square of residuals [47]. There are different algorithms depending on the type of the model that is the most appropriate to be fitted on the data. Firstly, relations available in literature were implemented on the available features. Afterwards these models were adapted accordingly so that a large number of breaths could be fitted with a simpler model that was used to design of breath model that describes a set of breaths.

5.2.1 Power-Two Model Function

A ubiquitous correlation found in scientific literature is the relation between slope phase III (S_{III}) and volume expired (V_{exp}). This correlation was first described by Astrom et al. [11] and identified later by Tusman et al. [15] as a power-two model. A power-two model has the general formula in Equation (5.1). According to its definition, a power law is a functional relationship between two quantities, where one quantity varies as a power of another [48]. Identifying a power-law relation in the data can point to specific kinds of mechanisms that might underlie the respiratory morphology in question. A power-two model is a nonlinear model, an equation that is nonlinear in the coefficients or a combination of linear and nonlinear in the coefficients [48]. Non-linear models have the general formula from Equation (5.2).

$$y = a \cdot x^{b} + c \tag{5.1}$$

$$y = f(X, \beta) + \varepsilon \tag{5.2}$$

Here, in Equation (5.2) y is the vector of responses, *f* is the model function, defined by β the vector of coefficients, *X* the design matrix of the model and ε being the vector of errors. To solve the non-linear power-two function the Levenberg-Marquardt explained in Chapter 3.2.2 was applied.

First, in order to study the repeatability of measurements a new visual representation approach was implemented by using the two-term power model. Thus, the values of S_{III} [%/L] and the expired volume of each breath from different measurement days were plotted. The data points from each measurement day were fitted with the power-two model Equation (5.1). In Figure 5.8 and Figure 5.9 several measurements from two different healthy patients are

shown. The goodness of fit of the model expressed as the average r^2 was 0.7 suggesting that the fitted power-two model describes the variability of the data well.

R-square, denoted r^2 and known as the coefficient of determination represents the square of the sample correlation coefficient that is used for measuring how well the fitted model explains the variability in the data. An r^2 equal to zero indicates that there is no relationship between the response variable and the model function, while an r^2 equal to 1 indicates that the model function fits the response variable perfectly.

Further on, the model curves of healthy patients from each day (shown in Figure 5.8 and Figure 5.9) are almost overlaid which proves that the measurements are repeatable and their inter-day variability is not high. These results were expected from healthy patients who are presumed to have the same constant breathing pattern. Another aspect that could be noted in Figure 5.8 and Figure 5.9 is that at low expired volumes the slope tends to have higher and more variable values and at higher expired volumes the model converges towards a horizontal asymptote.

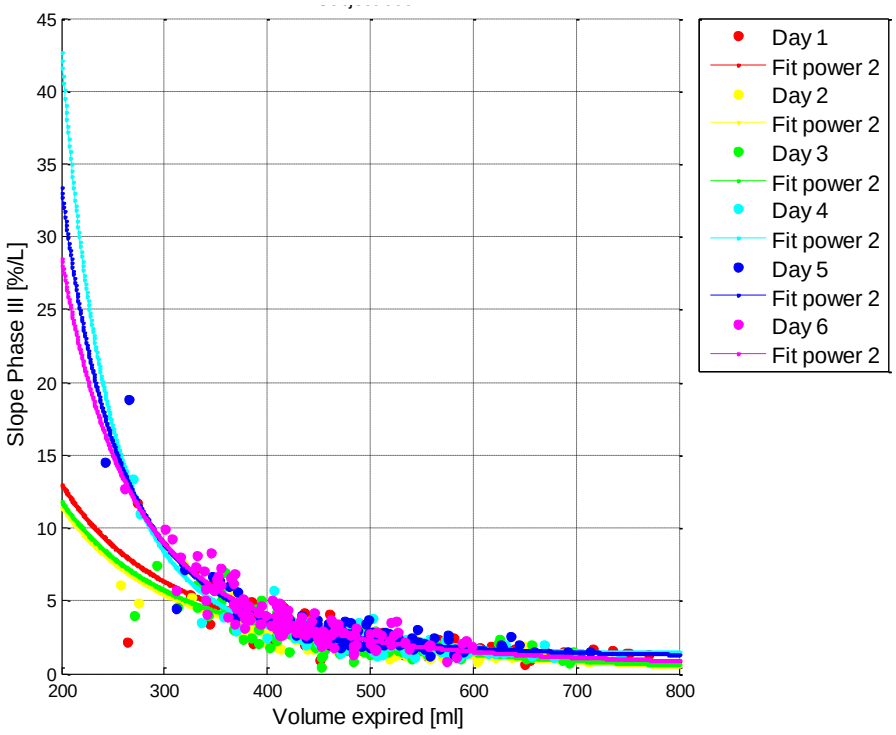


Figure 5.8. First example of a power-two model fitting for a healthy patient. Each dot represents the value from a single breath. Each line depicts the power-two model fit for each measurement day. Each measurement day is represented with a different color.

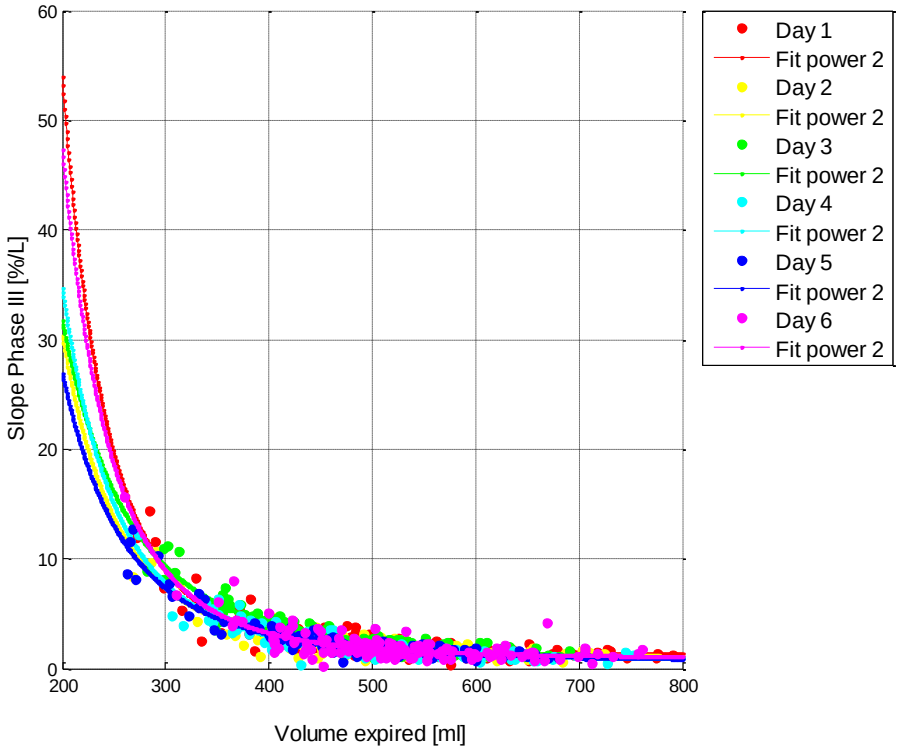


Figure 5.9. Second example of a power-two model fitting for a healthy patient. Each dot represents the value from a single breath. Each line depicts the power-two model fit for each measurement day. Each measurement day is represented with a different color.

The power-two model was applied for CHF patients and the results obtained were different than in the case of the healthy patients (shown in Figure 5.10 and Figure 5.11). Contrary to the overlaid model curves from the healthy patients, the model curves of the CHF patients show a displacement. This method of visualization could be used to assess the condition of the patient while he is undergoing treatment The downward shift of the model curves over time could be explained by the improvement of the patient due to the diuretics treatment. Thus, according to the power-two model curves fitted on the data from each measurement day, if the treatment received by the patient is effective, there should be a displacement in the model curves towards the lower left angle over time so that in the end they become almost overlaid. The CHF measurements present less data points because the patients have breathing difficulties and may not tolerate the device for a long time.

In Figure 5.10 and Figure 5.11 the power-two model fitting for slope phase III and volume expired of two patients with CHF and pulmonary edema are displayed. It is known from the medical examination documents of these patients that their condition improved after they received treatment. The excess fluid from their lungs was eliminated gradually and their improvement is shown in the downward-shifting of the curves, as marked by a black arrow (Figure 5.10 and Figure 5.11). Figure 5.12 shows the evolution of a CHF patient who according to the medical examinations did not improve and died. The power-two model curves have identical lengths and positions suggesting that his cardiopulmonary condition did not change between the measurements.

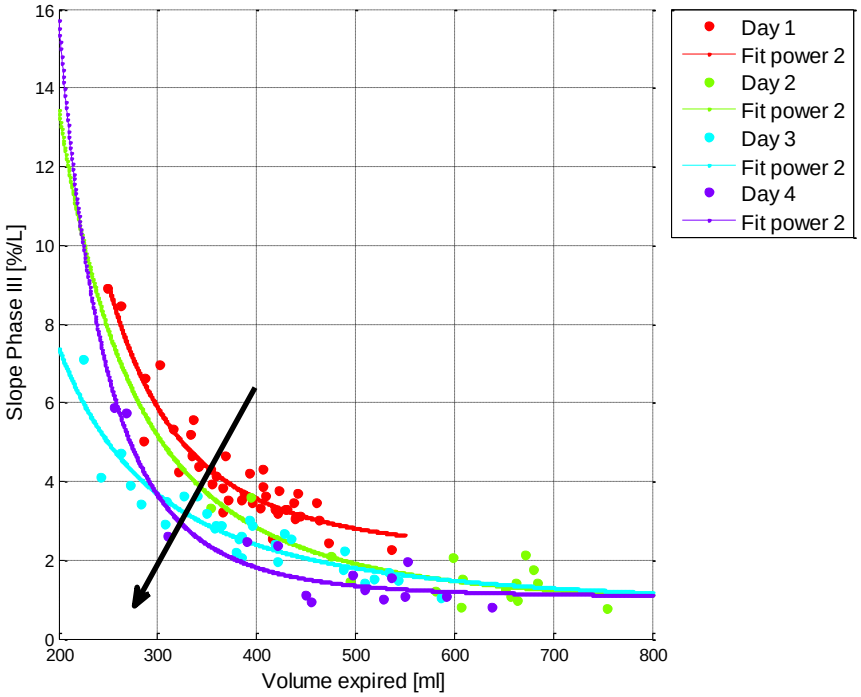


Figure 5.10. Power-two model curves for a CHF patient with four measurements. Each dot represents the value from a single breath. Each line depicts the power-two model fit for each measurement day. The black arrow shows the movement direction of the model curves as the patient improves

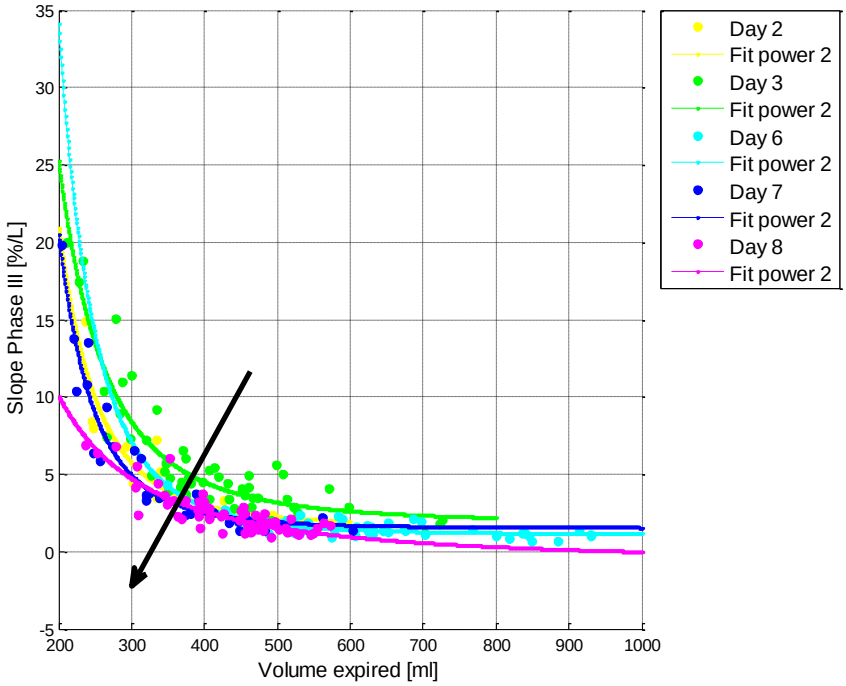


Figure 5.11. Power-two model curves for a CHF patient with five measurements. Each dot represents the value from a single breath. Each line depicts the power-two model fit for each measurement day. The black arrow shows the movement direction of the model curves as the patient improves

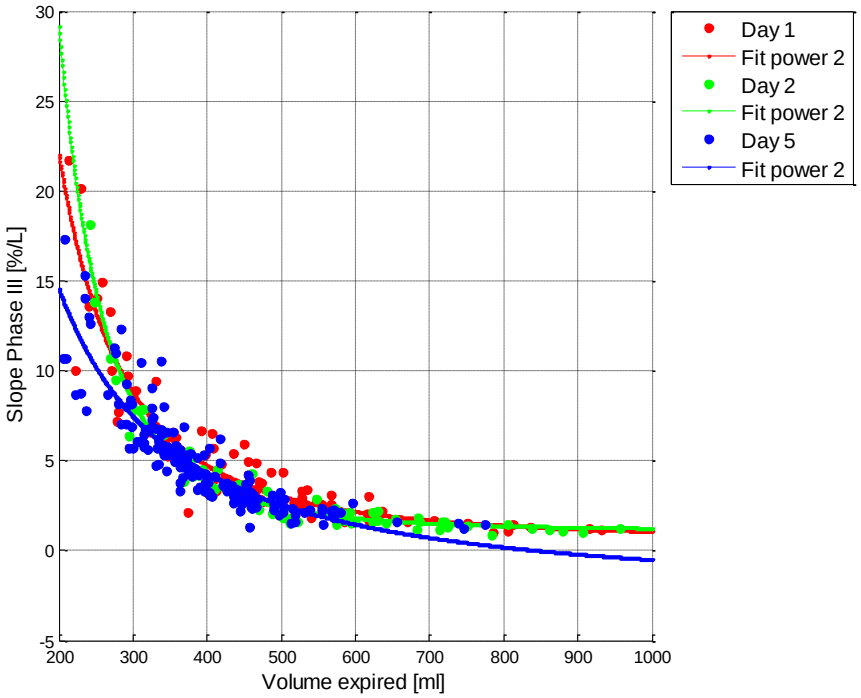


Figure 5.12. Power-two model curves for a CHF patient that did not respond to the treatment received. Each dot represents the value from a single breath. Each line depicts the power-two model fit for each measurement day. The model curves are overlaid since this patient did not present any improvement

5.2.2 Linear Regression Model

In order to simplify the correlation between slope phase III and the volume expired (shown in Figure 5.8 -Figure 5.12) it was aimed to reduce the scale of the model so that the influence of the outliers is minimized. The large values of slope phase III at small expired volumes determine a large range of values for the fitting interval of the two-term power function. For these reasons, the presentation of the data on a logarithmic scale was chosen. Using the logarithms of the values instead of the actual values reduces a wide range to a more convenient size [49]. When values that span large ranges need to be plotted, the logarithmic scale could provide a means of viewing the data that allows the values to be identified from the graph [49]. The logarithmic view becomes particular useful in the case of the CHF patients where there are a few data points, but many outliers the stretch the scale of the graphic.

According to N. Robbins, the author of [50], the main motive when using a logarithmic scale should be based on the distribution of the data and more specifically to its skewness. A logarithmic scale is a response to the much larger values in the data, i.e. the skewness. Skewness, defined in Equation (5.3) is a measure of the asymmetry of the data around the sample mean. If skewness is negative, the data are spread out more to the left of the mean than to the right [51]. Otherwise if the skewness is positive, the data are spread out more to the right [51]. The skewness of the normal distribution (or any perfectly symmetric distribution) is zero [51]. In Figure 5.13 the distribution of the values of slope phase III for CHF patients was plotted. The data distribution is right-skewed being influenced by the

greater slope values which appear as outliers. The skewness, calculated with the formula from the Equation (5.4) was 1.76.

$$s = \frac{E(x-\mu)^3}{\sigma^3} \tag{5.3}$$

$$s = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_{i-}\bar{x})^3}{\left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)^3} \tag{5.4}$$

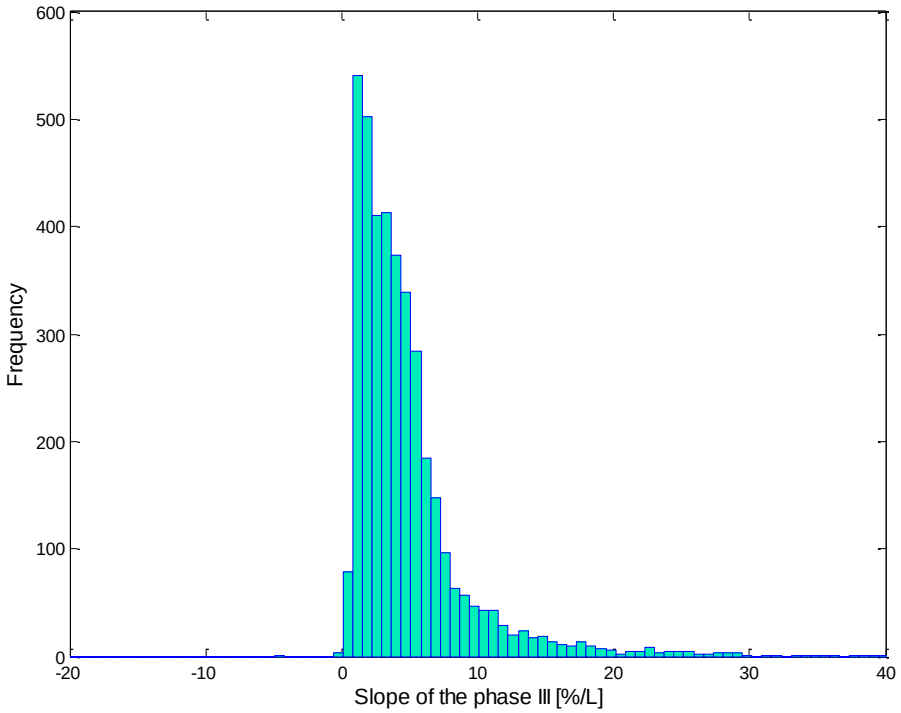


Figure 5.13. Distribution of slope phase III of CHF patients represented as a histogram

The equation $y = \log_{10}(S_{III})$ was used to express slope phase III in order to fit a linear regression model instead of a two-term power model. The base for the logarithmic scale was chosen 10 to take into consideration the data that might range over several orders of magnitude. The linear regression model was fitted with the least-squares method for linear models. In the matrix form, linear models are defined through the general formula:

$$y = X\beta + \varepsilon \tag{5.5}$$

where y is the vector of responses, β is the vector of coefficients, X is the design matrix for the model and ε is the vector of errors. For a first-degree polynomial the n equations in the two unknowns are expressed in terms of y, X and β as [47]:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ . \\ . \\ . \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ . \\ . \\ . \\ x_n \end{bmatrix} \times \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \tag{5.6}$$

The least-squares solution to the problem is the vector *b* which estimates the unknown vector of coefficients β. The normal equations are given by:

$$(X^T X)b = X^T y \tag{5.7}$$

$$b = \left(X^T X\right)^{-1} X^T y \tag{5.8}$$

After the Equation (5.8) for *b* is solved, *b* is replaced in the model formula to get the predicted response values:

$$\hat{y} = Xb = Hy \tag{5.9}$$

$$H = X\left(X^T X\right)^{-1} X^T \tag{5.10}$$

The projection matrix *H* is called the hat matrix because it puts the hat on *y*. The residuals are given by:

$$r = y - \hat{y} = (1 - H)y \tag{5.11}$$

To get the coefficients estimates, the least-squares method minimizes the summed square of residuals. The residual r_i for the i^{th} data point is the difference between the observed response value y_i and the estimated response value $\hat{y}_i$ [47]. A residual is defined as an observable estimate of an unobservable statistical error.

$$r_i = y_i - \hat{y}_i \tag{5.12}$$

The summed square of residuals is shown in Equation (5.13), where *n* is the number of data points used in the fitting:

$$S = \sum_{i=1}^{n} r_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i) \tag{5.13}$$

A linear model represents an equation that has two linear coefficients defined as in Equation (5.14), where p_1 is called the slope and p_2 the intercept:

$$y = p_1 x + p2 \tag{5.14}$$

In order to solve the above equation for the unknown coefficients, p_1 and p_2, S (Equation (5.15)) is written as a system of n simultaneous linear equations in two unknowns. If n is greater than the number of unknowns then the system of equations is over-determined.

$$S = \sum_{i=1}^{n} (y_{i-} (p_1 x_i + p_2))^2 \tag{5.15}$$

The least-squares fitting algorithm minimizes the summed square of the residuals, S, which means that the unknown coefficients are determined by differentiating S with respect to each coefficient and setting the result equal to zero.

$$\frac{\partial S}{\partial p_1} = -2\sum_{i=1}^{n} x_i (y_i - (p_1 x_i + p_2)) = 0 \tag{5.16}$$

$$\frac{\partial S}{\partial p_2} = -2\sum_{i=1}^{n} (y_i - (p_1 x_i + p_2)) = 0 \tag{5.17}$$

The estimates of the true parameters, represented by b, will be substituted in the Equations (5.16) and (5.17), where b_1 replaces p_1 and b_2 replaces p_2. The normal equations are defined as follows:

$$b_1 \sum_{i=1}^{n} x_i^{\,2} + b_2 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i \tag{5.18}$$

$$b_1 \sum_{i=1}^{n} x_i \; + n b_2 = \sum_{i=1}^{n} y_i \tag{5.19}$$

Solving b_1 and b_2 by using the b_1 value the next equations will result:

$$b_1 = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^{\,2} - (\sum_{i=1}^{n} x_i)^2} \tag{5.20}$$

$$b_2 = \frac{1}{n}(\sum_{i=1}^{n} y_i - b_1 \sum_{i=1}^{n} x_i) \tag{5.21}$$

By calculating b_1 and b_2 we estimate the coefficients p_1 and p_2. For a higher degree polynomial an additional normal equation is required for each linear term added to the model [47].

In Figure 5.14 and Figure 5.15 the data points of two CHF patients describing the correlation between slope phase III [%/L] and the expired volume [ml] was plotted on a logarithmic scale. The excluded data points, marked with the “star” symbol were also plotted to show that the power-two model is a better fit only if no methods for discarding outliers were applied. The excluded data points were identified through the filtering methods

described in Chapter 3.1.1. The linear regression model was fitted on the data points marked with the "dot" symbol. Unlike the power-two model graphic, this visualization does not show any obvious displacement of the lines in a direction. However, this visualization is useful for discerning the values plotted on the graph and it was used further in the design of a single breath mathematical model because of the simplicity of the linear regression model.

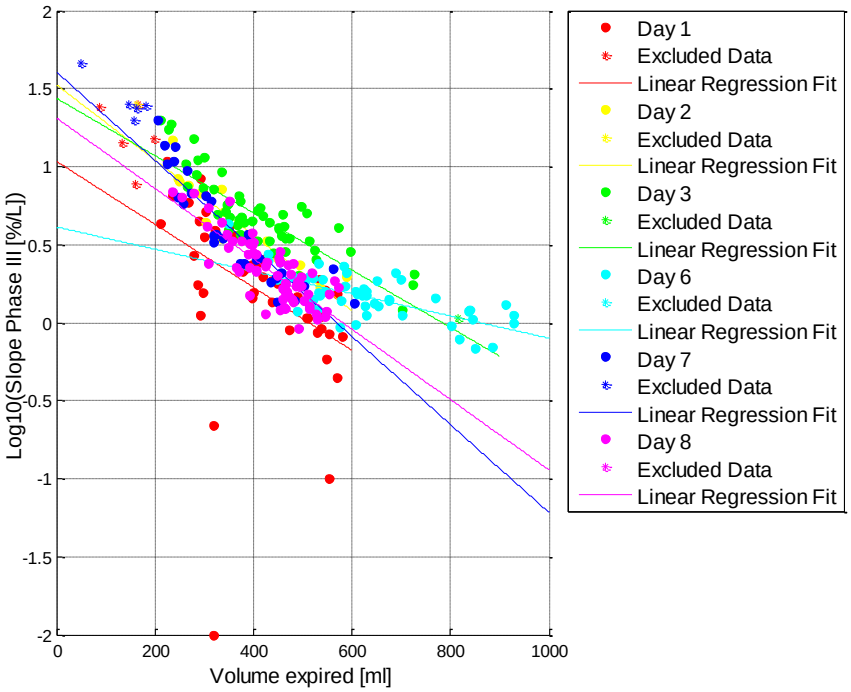


Figure 5.14. First example of a linear regression model fit for a CHF patient. Each dot represents the value from a single breath. Each star represents the excluded data

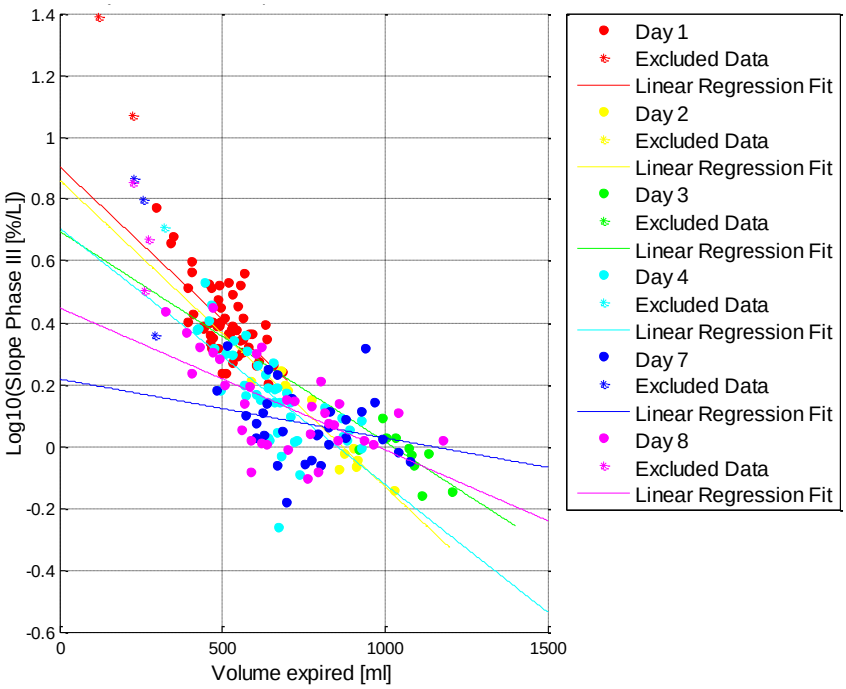


Figure 5.15. Second example of a linear regression model fit for a CHF patient. Each dot represents the value from a single breath. Each star represents the excluded data

5.3 Single Breath Model for Capnography Measurements

In this chapter a mathematical model composed from a series of volumetric capnograms is explained. The model breath was designed with the following aims in mind:

- describe a set of breaths (a measurement) by a single representative breath
- design a model capnogram adaptable for different expired volumes
- extract an average capnogram of a set of breaths

To this end, a mathematical model breath based on the relationship between different features of the volumetric capnogram was created. These features have been chosen because they define the shape of the capnogram. Their detailed definition and physiological significance was detailed in Chapter 3.2.1. The model breath used 5 input parameters representing all the feature values from one measurement:

1. Volume Phase I (VPh1)
2. Volume Phase II (VPh2)
3. Slope Phase II (S_{II})
4. Slope Phase III (S_{III})
5. Volume Expired (V_{exp})

The algorithm implemented had several steps which were explained in the logical diagram shown Figure 5.16. For the power-two model fitting a non-linear least-squares "Levenberg-Marquardt" algorithm was used. For the linear-regression fit a linear least-squares algorithm was used. The correlation between the volume of the phase I (and II) and the expired volume was chosen in order to construct the first part of the model breath, starting from the left side. The correlations between the volumes composing the phases of capnogram were described with a power-two model. Next, the correlation between slope phase II and expired volume was also expressed through a power-two model function. Finally, taking into account the logarithmic representation from Chapter 5.2.2 used for the correlation between slope phase III and expired volume, a linear regression model fit was computed. In all the fitting algorithms the independent variable was represented by the expired volume.

For each respiratory measurement the model coefficients of each function, rendered by the fitting algorithm were saved. Next, these coefficients were introduced in the functions which were used before for estimating them – the power-two model function and the linear regression function. This time, the input x of the functions was the average expired volume of the patient and the outputs were the volumes of the phase I and II, slope phase II and slope phase III. Subsequently, the outputs rendered by the model functions and the estimated coefficients were used to draw the single breath model function.

After fitting the data with the models, the coefficients of determination, r^2, used for measuring how well the fitted model explains the variability in the data, were between 0.54 (minimum value) and 0.9 (maximum value). These values suggest the chosen models fitted the data well.

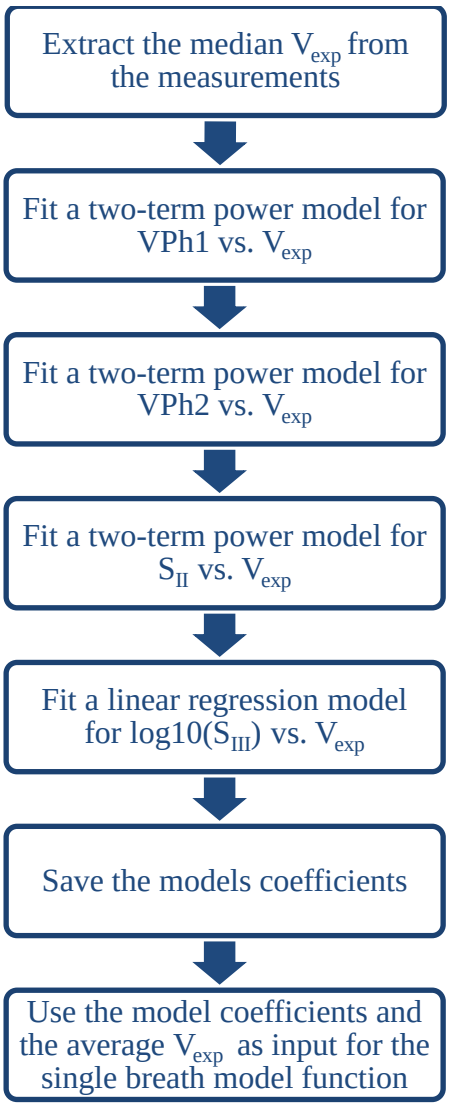


Figure 5.16. Logical workflow for computing the single breath model of a patient

In Figure 5.17 the single breath models for several measurements of a healthy patient are displayed. Note the overlaid capnogram models from each day which suggest the same breathing pattern throughout several different measurements. Moreover, observe that the slope lines, both for the phase II and phase III are parallel. As the patient is healthy, his respiratory parameters stay in the same range over time and as a result the single breath model does not change.

The single breath model shows very well the variation in the respiratory condition of the CHF patients as they receive the necessary treatment. Thus, in Figure 5.18 and Figure 5.19 the breath models of two different CHF patients with pulmonary edema are displayed. Note the upward movement direction of the single breath models as the patient improves. This upward movement corresponds to a higher $ETCO_2$ value which is supported by the reference value from Table 5.1. If the patient improves, his $ETCO_2$ value ought to approach the value of a healthy patient. After several days of treatment, the CHF patient should exhibit a condition close to a healthy respiratory state. In Figure 5.18 observe the purple line corresponding to the model breath from the last day of measurement. This line has a similar shape and height with the single breath models of the healthy patient from Figure 5.17. As not all the CHF patients

are hospitalized with the same disease degree and also because they do not respond in the same way to the treatments given, the movement pattern of the single breath model are unique for each patient. This difference could be observed in the different improvement trends of the patients from Figure 5.18 and Figure 5.19 where the distance between the single breath models is not the same.

To conclude, this type of representation, with an average breath model that describes an entire capnographic measurement, could be used to check if the treatment received by the patient is effective or it needs to be switched to another type of medication. Moreover, the gradual improvement, worsening or constant state of the respiratory apparatus is reflected in the shape of the single breath model and a clinician could take a look at it before taking a decision concerning the patient. Still, this single breath model needs further development in the parameterization of the model functions used in order to design a breath model not influenced by possible outliers.

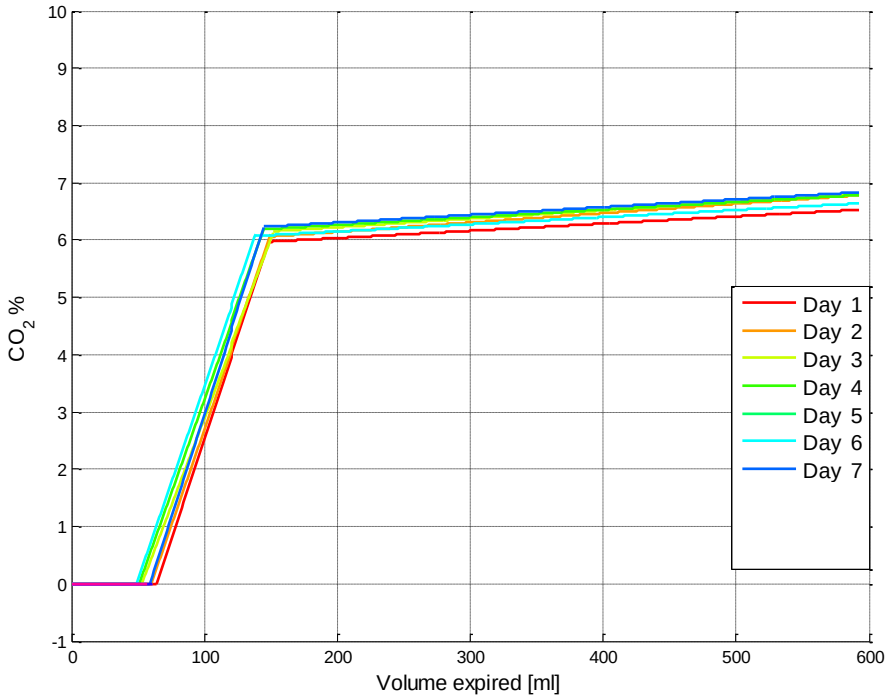


Figure 5.17. Single breath model for a healthy patient. Each colored line represents the average volumetric capnogram of a particular measurement. Note the overlaid single breath models from each day suggesting a constant cardiopulmonary state

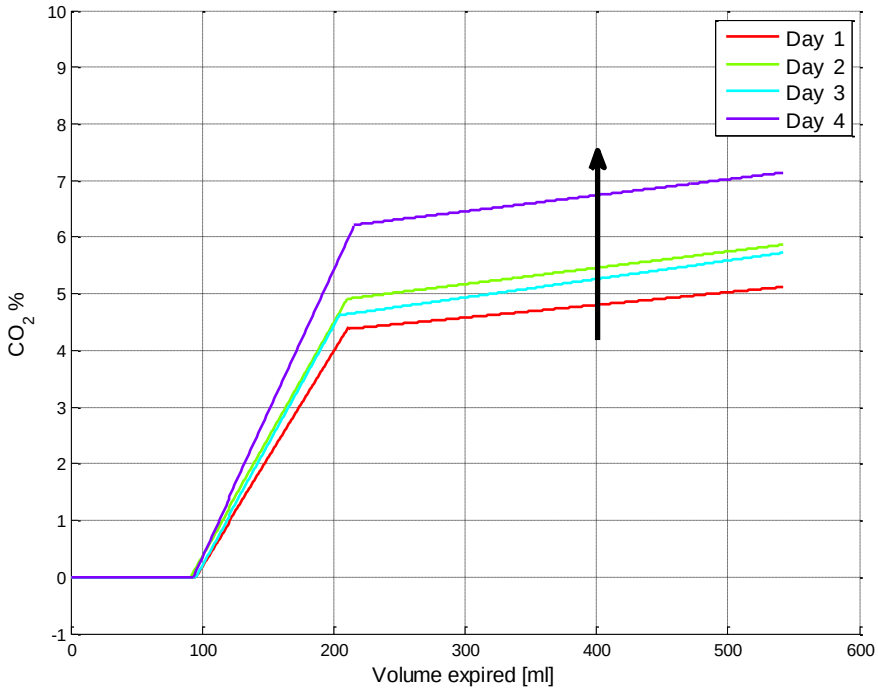


Figure 5.18. Single breath model for a CHF patient with four measurements. Each colored line represents the average volumetric capnogram of a particular measurement. Note the upwards shifting of the breath models (marked by the black arrow) suggesting the improvement of the patient

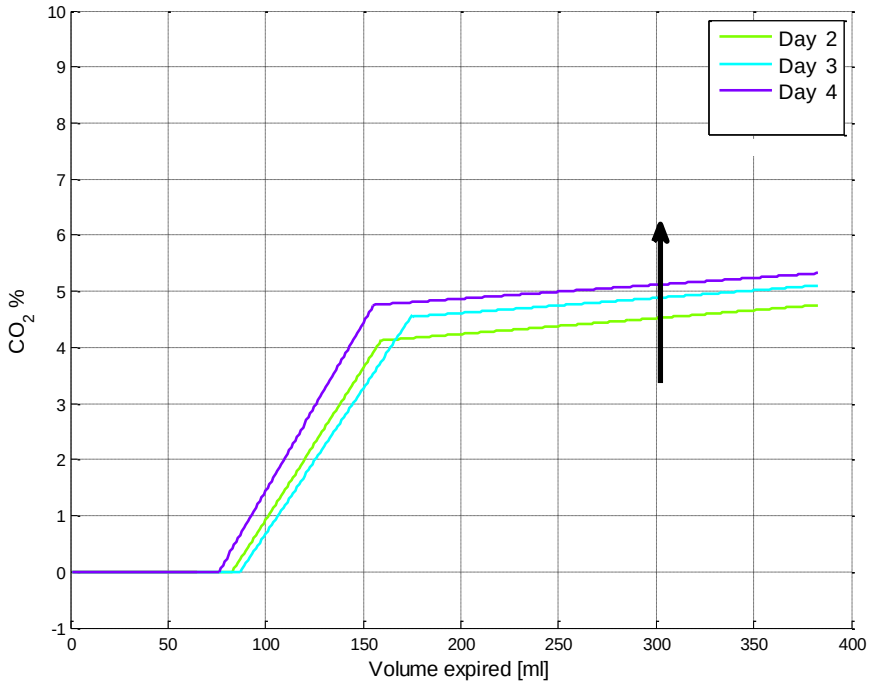


Figure 5.19. Single breath model for a CHF patient with three measurements. Each colored line represents the average volumetric capnogram of a particular measurement. Note the upwards shifting of the breath models suggesting the improvement of the patient

6 Classification Results

In the clinical trial, patients who were admitted with an initial diagnosis of Acute Decompensated Heart Failure (ADHF) were monitored throughout the recompensation process. In the recompensation process, which is the "reverse" process of decompensation, the excess body fluid of the patients is reduced through a treatment with diuretics medication. In order to monitor the effect of the recompensation therapy and to identify the degree of decompensation, the changes in the pulmonary fluid of the patient were assessed through volumetric capnography measurements. Patients participated between three and fourteen days to follow the whole recompensation process. The symptoms of the patients were recorded as part of the routine clinical diagnostic confirmation using x-ray, ultrasound and physical examinations. These clinical examinations will serve as "golden standards" for confirming pulmonary edema.

As the medical condition of each patient was established through several examination methods which are considered to be gold standards, the problem posed was to classify categorical dependent variables. To this end, classification trees as described in Chapter 4 were applied. Thus, Chapter 6.1 summarizes the steps followed in order to apply decision trees analysis. In Chapter 6.2 the results of the classification analysis are reported. Only the CHF patients who had a confident medical diagnosis were included in this classification analysis.

6.1 Classification Stages

The classification process involves several steps which are detailed in Table 6.1. These classification steps were applied for the entire dataset comprising volumetric capnographic measurements from CHF patients and healthy patients. The following clinical measuring routines were proposed for applying the decision trees classifiers:

A. The patient does not have a prior diagnosis of CHF. Then, a respiratory measurement is recorded and a classification tree for detecting a CHF condition is applied.
B. The patient has a confirmed diagnosis of CHF; consequently one respiratory measurement is recorded and a classification tree that predicts pulmonary edema and no pulmonary edema is employed.
C. If the diagnosis of pulmonary edema is confirmed and the patient undergoes treatment, it is desired to monitor his improvement or worsening. The time-independent feature extraction method and a classification tree for monitoring the temporal evolution of edema are applied.

The fourth step of classification, the extraction of time-independent features was only carried out for monitoring the status of pulmonary edema in CHF patients. This step was implemented in order to address the different hospitalization periods of patients.

Table 6.1. Classification steps for decision trees analysis

Classification steps	Description	Chapter
1) Outliers removal	Discard invalid breaths by applying the "median method" and setting a lower threshold limit for V_{exp}= 200 ml.	Chapter 3.1.1
2) Feature extraction	Extract physiological features representing the characteristics of a single breath from the volumetric capnogram.	Chapter 3.2.1
3) Parameter extraction	Extract the median (or mean, depending on the feature) feature value from each measurement day.	Chapter 3.3
4) Time-independent feature extraction	Reduce the whole measurement period of the patient to four features: mean, standard deviation and regression line coefficients.	Chapter 3.3
5) Dataset selection	Group the features into classification datasets according to the classification aim.	Chapter 6.2
6) Model classification tree	Grow a classification tree based on the input dataset.	Chapter 4.2
7) Leave-one-out cross-validation	Evaluate the accuracy of the model tree created by splitting the dataset into *N* folds, where *N* is the number of observations from the dataset.	Chapter 4.2.1
8) Classifier performance	Assess the performance of the classification tree in classifying future values.	Chapter 4.2.1

6.2 Results

A. Classification tree for detecting CHF

The formulated problem was to be able to identify automatically if a patient has CHF when he has not been previously diagnosed with it. A possible scenario where this classification problem has applicability is when a patient with symptoms of CHF comes to the hospital and a confirmation for his diagnosis is needed in a short time without the involvement of invasive measurement tools. The solution would be a volumetric capnographic respiratory measurement lasting a few minutes. Then, this measurement will serve as an input observation for the classification algorithm which will render right away a response about the patient's condition: whether he suffers from CHF or not.

The classification tree was trained on the data of the CHF and healthy patients. The input dataset distribution is shown in Table 6.2. In the end, there were 46 patients who were analyzed in this study, 21 healthy patients and 26 CHF patients. The training data consisted of volumetric capnography parameters as detailed in Table 6.1.

Table 6.2. Dataset distribution for classification case A

Patient Class Label	Number of observations
CHF	26
Healthy	21

In Figure 6.1 the decision tree for the identification of CHF is shown. The resubstitution error was 8.51% and the LVOOCV error was 21.27%. Moreover, the features selected as splitting nodes in the classification tree in Figure 6.1, S_{III} [%/L] and $ETCO_2$ are accepted in related literature as important discriminator features for discerning CHF patients from healthy patients. The decision tree in Figure 6.1 classified all patients with a smaller $ETCO_2$ as CHF patients, a choice supported by the reference values obtained in Chapter 5.1.1. Additionally, the decision trees classified all patients with a higher $ETCO_2$ value and a small S_{III} as healthy. In Figure 6.2 the ROC curve obtained for the decision tree in Figure 6.1 shows its performance across several thresholds obtained from the validation steps. The AUC for LVOOCV was 0.72 and the AUC for the training data was 0.92. The small error values and the large AUCs suggest that the classification tree is able to classify accurately future observations.

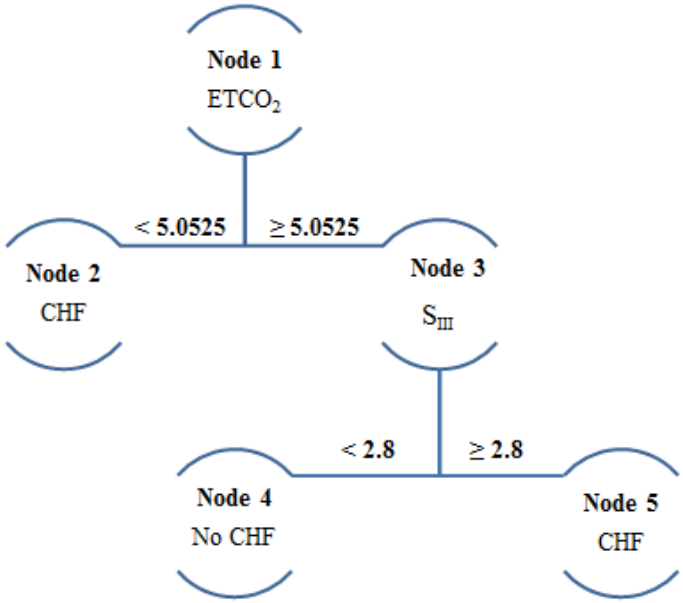


Figure 6.1. Case A. Classification tree for identifying CHF from a respiratory measurement. $ETCO_2$ is the median $ETCO_2$ value and S_{III} [%/L] represents the median value

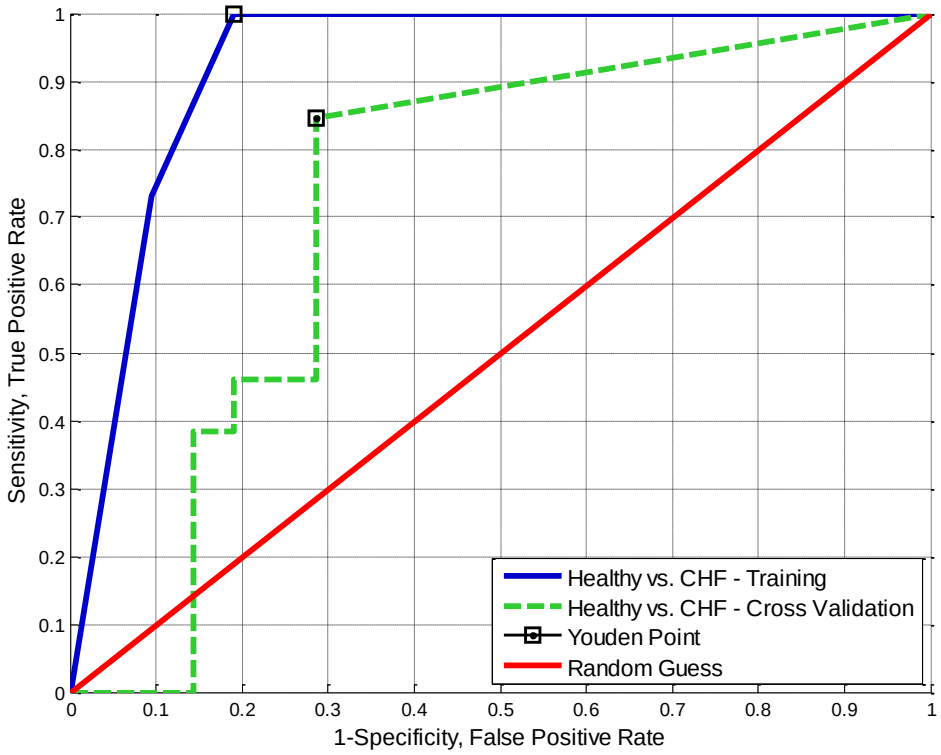


Figure 6.2. Case A. ROC curve obtained from the LVOOCV prediction scores (green line) and resubstitution prediction scores (blue line)

B. Classification tree for detecting pulmonary edema in CHF patients

If the patient has a confirmed diagnosis of CHF then is desired to know whether he has pulmonary edema or not so that the adequate treatment is applied. To this end, one respiratory measurement of the patient is recorded and a classification tree that automatically predicts pulmonary edema and no pulmonary edema is employed. In this scenario, only confirmed CHF patients were used for training the classification tree. Accordingly, the dataset was divided into 2 groups, patients with CHF but no pulmonary edema and patients with CHF and pulmonary edema.

Table 6.3. Dataset distribution for classification case B

Patient Class Label	Number of observations
CHF Edema	13
CHF No Edema	13

In Figure 6.3 the classification tree for identifying pulmonary edema is shown. The resubstitution error was 11.53% and the LVOOCV error was 30.76%, values slightly larger than the errors of the classification tree from Figure 6.1. The root of the decision tree was Alpha Angle [°] and the third node was Volume of the Phase II (VPh2) [ml]. Interestingly, the decision tree classified patients with a smaller value for the Alpha Angle [°] as having pulmonary edema, although the reference values from Chapter 5.1.1 suggested that CHF patients had a greater Alpha Angle [°]. Also, the decision tree classified patients with a larger Alpha Angle [°] as having pulmonary edema (see Node 4, Figure 6.3). An explication for this splitting could be the fact that when pooled together the CHF values render on average a larger value of the Alpha Angle [°].

In Figure 6.4 the ROC curve obtained for the decision tree in Figure 6.3 shows its performance across several thresholds obtained from the validation steps. The AUC for LVOOCV was 0.64 and the AUC for the training data was 0.88. All in all, the high LVOOCV error and the AUC value which is quite close to that of a random classifier points to the fact that this classification tree might not give accurate results. Further explanations for this result could lie in the medical background of the patients for which more medical opinions are needed. Still, the performance of this classification tree is better than random chance so there is no reason for which it should not be employed.

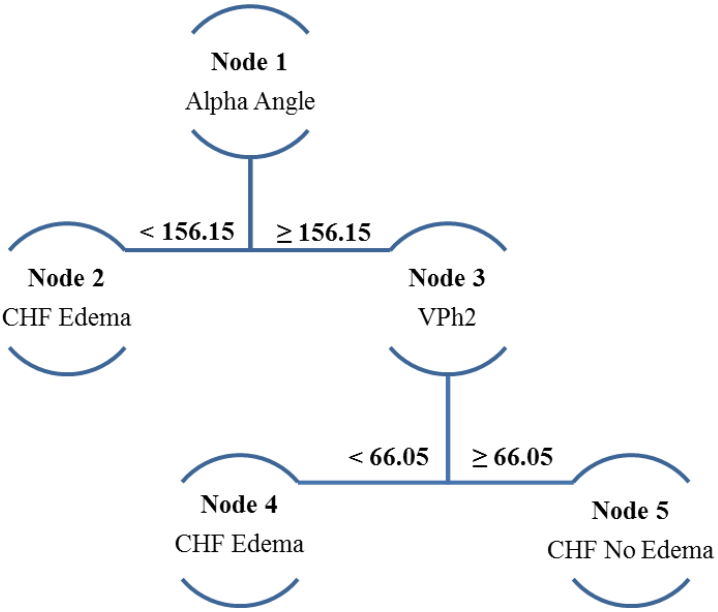


Figure 6.3. Case B. Classification tree for identifying pulmonary edema. Alpha Angle [°] is the median value from a measurement and VPh2 [ml] is the median value for the volume phase II

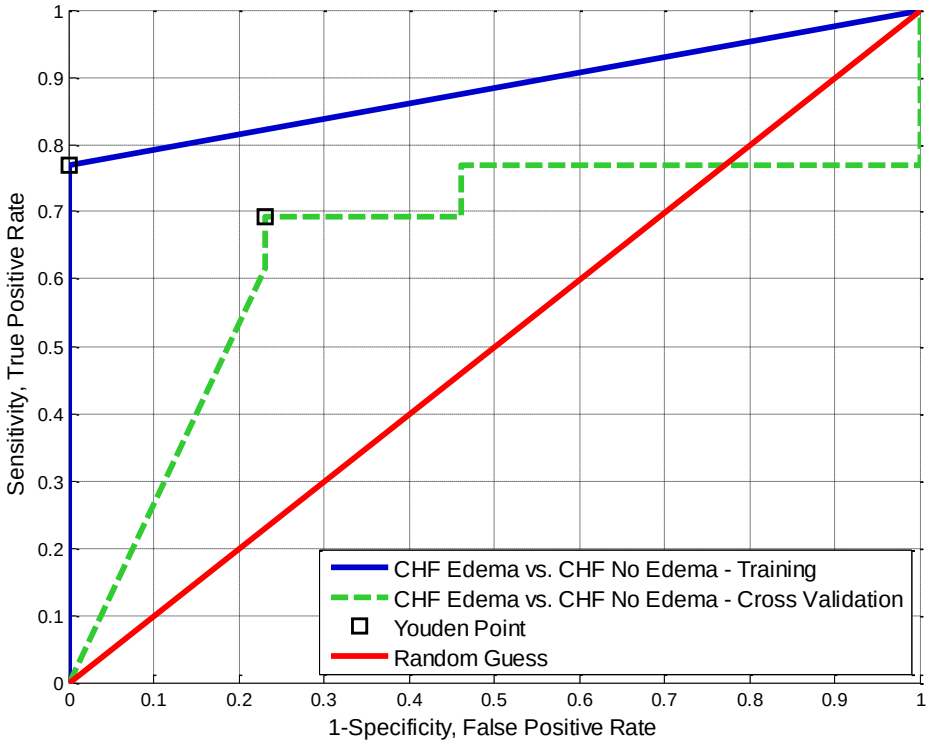


Figure 6.4. Case B. ROC curve obtained from the LVOOCV prediction scores (green line) and resubstitution prediction scores (blue line)

C. Classification tree for monitoring the temporal evolution of pulmonary edema patients

If the patient has been diagnosed with pulmonary edema then it is desired to monitor his responsiveness to the treatment received. A classification tree should be able to tell whether the treatment received is effective and this means that it will give a "No Edema" result for the patient or if the treatment needs to be switched and the result will be "Edema" class.

In the current study not all the CHF patients had the same number of respiratory measurements because of the different hospitalization times and for this reason the time-independent feature extraction method described in Chapter 3.3 was used. Therefore, the input dataset for this classification tree was represented by the mean value $\bar{x}$, the standard deviation σ_x, the slope b_x and the y-intercept a_x of time-depending regression line $(x_i = b_x \times t_i + a_x)$, where x represents a parameter value extracted from the feature values. The dataset has the distribution from Table 6.4. One patient diagnosed with "CHF Edema" was removed from the group due to his worsening condition that eventually led to death. Another patient was removed since he had only one respiratory measurement which renders a $\sigma_x = 0$ and $b_x = 0$.

Table 6.4. Dataset distribution for classification case C

Patient Class Label	Number of observations
CHF Edema	12
CHF No Edema	12

In Figure 6.5 the decision tree for identifying the presence of pulmonary edema in CHF patients undergoing targeted treatment is shown. The root of the classification tree is $\sigma_{AlphaAngle}$ which makes a split at the value 5.1042 classifying all observations larger or equal than it as having pulmonary edema. An explanation for the selection of $\sigma_{AlphaAngle}$ as a node in this classification task, could be the higher variability found in the breathing pattern of sick patients. This makes the values of Alpha Angle span across a large range. The next node chosen in the classification tree from Figure 6.5 is the slope of the VD_{aw} linear regression line. The slope of a parameter shows a change in the respiratory status pointing to an increasing or decreasing trend in the values of the feature given by responsiveness to the treatment.

The resubstitution error was 4.16% and the LOOCV error was 25 %. In Figure 6.6 the ROC curve obtained for the decision tree in Figure 6.5 shows its performance across several thresholds obtained from the validation steps. The AUC for LVOOCV was 0.73 and the AUC for the training data was 0.98611. The small errors and the AUC close to 1 suggest that this classification tree is likely to give highly accurate classifications for future unobserved data.

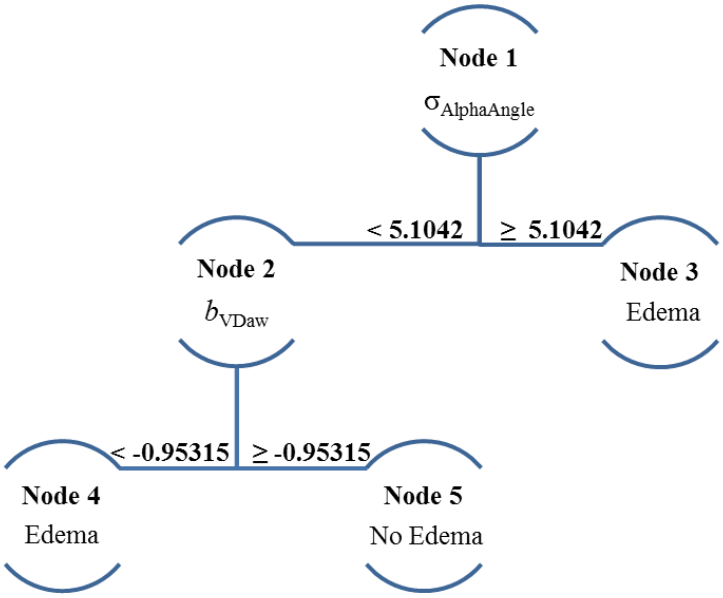


Figure 6.5. Classification tree for monitoring the temporal evolution of pulmonary edema in CHF patients. $\sigma_{AlphaAngle}$ is the standard deviation of Alpha Angle [°] and b_{VDaw} is the slope of the regression line traced across all the VD_{aw} [ml] measurement points

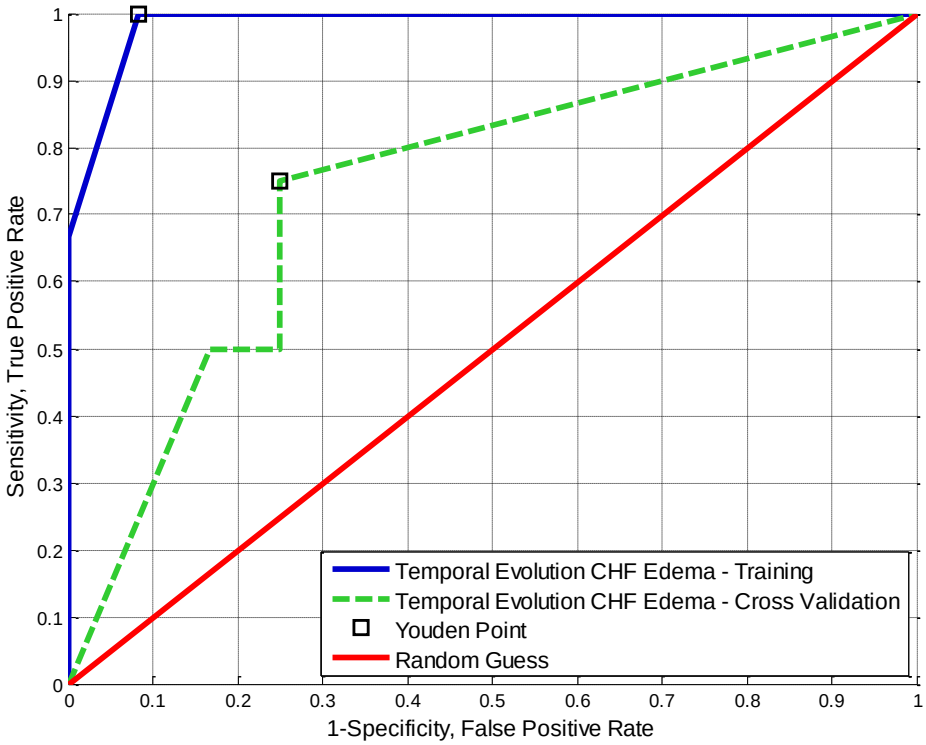


Figure 6.6. Case C. ROC curve obtained from the LVOOCV prediction scores (green line) and resubstitution prediction scores (blue line)

7 Conclusions and Perspectives

Towards the goal of monitoring the respiratory profile of spontaneously breathing patients quantitative capnogram analysis has the potential to distinguish different cardiopulmonary states. A quantitative analysis can be achieved through feature extraction and further by feature-based classification.

The first task of this work was to implement a method for pre-processing the volumetric capnogram in order to detect outlier breaths. This task was followed by a feature extraction step that rendered quantifiable typical values for volumetric capnography. The feature values extracted are comparable with values reported in related literature. Hence, these reference values can serve as a starting point for discriminating healthy patients from patients with lungs dysfunctions by highlighting physiological differences.

Furthermore, the correlations between features were exploited. In the correlation between the expired volume and slope phase III, the power-two model curves from each measurement day were plotted. The results suggested that in the case of healthy patients, the model curves from each day will be overlaid. In the case of CHF patients, the result was a downward displacement of the model curves that could be related to a loss of fluid from the lungs. Moreover, in order to address the slope phase III skewed distribution, a logarithmic scale was used. Next, the correlation was described by a linear regression model. All these steps led further to the design of a mathematical breath-model that can possibly summarize a set of breaths and discriminate among various cardiopulmonary states.

The last task was to distinguish automatically between patients with different cardiopulmonary states. To this end, a classification task involving three categories was proposed for three clinical scenarios. An adequate set of features and classification-trees as a machine learning algorithm were selected. The varying hospitalization periods was handled by implementing an algorithm for rendering the features values time-independent. The changes in the pulmonary fluid of the patients with pulmonary edema were assessed through a temporal evolution classification tree. The good performance of the classification trees is indicated by obtained AUC values close to 1. The small resubstitution and LOOCV errors of the classification trees obtained suggest that they could be applied in the future to larger datasets for which they will render accurate classifications results.

A possible prospective goal would be a grouping separation by the severity of HF, e.g. NYHA class. This way there would be a clearer separation between the cardiopulmonary states of the patients. Moreover, a medication correlation could be studied in order to see exactly how the medicine intake influences the respiratory parameters. By extending the study to other respiratory diseases like asthma, pneumonia or COPD, classification trees could prove useful to label patients with CHF among other respiratory dysfunctions. Another prospective approach could be the implementation of a different supervised machine learning algorithm like “Principal Component Analysis” or “Quadratic Discriminant Analysis” to discriminate among different respiratory states.

Bibliography

[1] M. B. Jaffe and Respironics, Inc., 'Volumetric Capnography - The next advance in CO2 monitoring'. Respironics, Inc., Jan-2012.

[2] 'Why Capnography - Capnography'. [Online]. Available: http://www.capnography.com/new/capnography-introduction/why-capnography. [Accessed: 29-Jun-2014].

[3] American Society of Anesthesiologists (ASA). Basic Standards for, M. B. Jaffe, and J. S. Gravenstein, 'Respironics CO2 Solutions. Quick Reference Guide'. 2009.

[4] R. J. Asher, 'Capnographic Analysis for Disease Classification', Master of Science, Massachusetts Institute of Technology, 2012.

[5] Respironics, Inc., 'Capnography Reference Handbook'. Respironics, Inc., 29-Jun-2004.

[6] 'Pulmonary edema Causes - Diseases and Conditions - Mayo Clinic'. [Online]. Available: http://www.mayoclinic.org/diseases-conditions/pulmonary-edema/basics/causes/con-20022485. [Accessed: 20-Jun-2014].

[7] S. Weyer, M. D. Zink, T. Wartzek, L. Leicht, K. Mischke, T. Vollmer, and S. Leonhardt, 'Bioelectrical impedance spectroscopy as a fluid management system in heart failure', *Physiol. Meas.*, vol. 35, no. 6, pp. 917–930, Jun. 2014.

[8] 'Lungs | Heart Failure Online'. [Online]. Available: http://www.heartfailure.org/heart-failure/lungs/. [Accessed: 29-Jun-2014].

[9] W. S. Fowler, 'Lung function studies. II. The respiratory dead space', *Am. J. Physiol. Content*, vol. 154, no. 3, pp. 405–416, 1948.

[10] R. Fletcher, B. Jonson, G. Cumming, and J. Brew, 'The Concept of Deadspace with Special Reference to the Single Breath Test for Carbon Dioxide', *Br. J. Anaesth.*, vol. 53, no. 1, pp. 77–88, Jan. 1981.

[11] N. E. Astrom, 'Partinioning of dead space - a method and reference values in the awake human', *The European Respiratory Journal*, 2000.

[12] F. Verschuren, E. Heinonen, D. Clause, F. Zech, M. S. Reynaert, and G. Liistro, 'Volumetric capnography: reliability and reproducibility in spontaneously breathing patients', *Clin. Physiol. Funct. Imaging*, vol. 25, no. 5, pp. 275–280, 2005.

[13] R. Arena, M. Guazzi, J. Myers, P. Chase, D. Bensimhon, L. P. Cahalin, M. A. Peberdy, E. Ashley, E. West, and D. E. Forman, 'Prognostic Value of Capnography During Rest and Exercise in Patients With Heart Failure', *Congest. Heart Fail.*, vol. 18, no. 6, pp. 302–307, 2012.

[14] G. Tusman, A. Scandurra, S. H. Böhm, F. Suarez-Sipmann, and F. Clara, 'Model fitting of volumetric capnograms improves calculations of airway dead space and slope of phase III', *J. Clin. Monit. Comput.*, vol. 23, no. 4, pp. 197–206, Jun. 2009.

[15] G. Tusman, E. Gogniat, S. H. Bohm, A. Scandurra, F. Suarez-Sipmann, A. Torroba, F. Casella, S. Giannasi, and E. S. Roman, 'Reference values for volumetric capnography-derived non-invasive parameters in healthy individuals', *J. Clin. Monit. Comput.*, vol. 27, no. 3, pp. 281–288, Feb. 2013.

[16] 'Butterworth filter design - MATLAB butter'. [Online]. Available: http://www.mathworks.com/help/signal/ref/butter.html. [Accessed: 18-Jun-2014].

[17] 'Butterworth Filter Design and Low Pass Butterworth Filters'. [Online]. Available: http://www.electronics-tutorials.ws/filter/filter_8.html. [Accessed: 18-Jun-2014].

[18] Respironics, Inc. and M. B. Jaffe, ‘VolumetricCapnography - The next advance in CO_2 monitoring Respironics’. Respironics, Inc.
[19] Respironic Novametrix, Inc., Wallingford CT 06492, ‘Parameter Calculations in CO2SMO Plus! Respiratory Profile Monitor and NICO Cardiopulmonary Monitor’. .
[20] ‘ventilation-perfusion relationships’. [Online]. Available: http://courses.kcumb.edu/physio/adaptations/vq.htm. [Accessed: 18-Jun-2014].
[21] *Capnography*, 2nd ed. Cambridge ; New York: Cambridge University Press, 2011.
[22] ‘Trapezoidal numerical integration - MATLAB trapz’. [Online]. Available: http://www.mathworks.com/help/matlab/ref/trapz.html. [Accessed: 18-Jun-2014].
[23] ‘Least-Squares (Model Fitting) Algorithms - MATLAB & Simulink’. [Online]. Available: http://www.mathworks.com/help/optim/ug/least-squares-model-fitting-algorithms.html. [Accessed: 21-Jun-2014].
[24] ‘Evaluating Goodness of Fit - MATLAB & Simulink’. [Online]. Available: http://www.mathworks.com/help/curvefit/evaluating-goodness-of-fit.html. [Accessed: 18-Jun-2014].
[25] J. Friedman, T. Hastie, and R. Tibshirani, *The elements of statistical learning*, vol. 1. Springer Series in Statistics New York, 2001.
[26] ‘Cross-validated classification model - MATLAB’. [Online]. Available: http://www.mathworks.com/help/stats/classificationpartitionedmodel-class.html. [Accessed: 25-Jun-2014].
[27] R. J. Lewis, ‘An introduction to classification and regression tree (CART) analysis’, in *Annual Meeting of the Society for Academic Emergency Medicine in San Francisco, California*, 2000, pp. 1–14.
[28] ‘What is the difference between categorical, ordinal and interval variables?’ [Online]. Available: http://www.ats.ucla.edu/stat/mult_pkg/whatstat/nominal_ordinal_interval.htm. [Accessed: 27-Jun-2014].
[29] ‘Classification Trees and Regression Trees - MATLAB & Simulink’. [Online]. Available: http://www.mathworks.com/help/stats/classification-trees-and-regression-trees.html. [Accessed: 19-Jun-2014].
[30] ‘Binary decision tree for classification - MATLAB’. [Online]. Available: http://www.mathworks.com/help/stats/classificationtree-class.html#bsv83uo-1. [Accessed: 26-Jun-2014].
[31] I. Naski and F. C. Porter, *Statistical Analysis Techniques in Particle Physics*. 2013.
[32] ‘Predict response for observations not used for training - MATLAB’. [Online]. Available: http://www.mathworks.com/help/stats/classificationpartitionedmodel.kfoldpredict.html. [Accessed: 26-Jun-2014].
[33] ‘Classification error by resubstitution - MATLAB’. [Online]. Available: http://www.mathworks.com/help/stats/classificationtree.resubloss.html. [Accessed: 26-Jun-2014].
[34] T. Fawcett, ‘An introduction to ROC analysis’, *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, 2006.
[35] ‘Performance Curves - MATLAB & Simulink’. [Online]. Available: http://www.mathworks.com/help/stats/performance-curves.html. [Accessed: 27-Jun-2014].

[36] ‘Compute Receiver Operating Characteristic (ROC) curve or other performance curve for classifier output - MATLAB perfcurve’. [Online]. Available: http://www.mathworks.com/help/stats/perfcurve.html. [Accessed: 27-Jun-2014].
[37] ‘Receiver Operating Characteristics’. [Online]. Available: http://www0.cs.ucl.ac.uk/staff/ucacbbl/roc/. [Accessed: 27-Jun-2014].
[38] ‘Receiver operating characteristic - Wikipedia, the free encyclopedia’. [Online]. Available: http://en.wikipedia.org/wiki/Receiver_operating_characteristic. [Accessed: 27-Jun-2014].
[39] ‘Interquartile range - MATLAB iqr’. [Online]. Available: http://www.mathworks.com/help/stats/iqr.html. [Accessed: 19-Jun-2014].
[40] ‘Coefficient of variation - Wikipedia, the free encyclopedia’. [Online]. Available: http://en.wikipedia.org/wiki/Coefficient_of_variation. [Accessed: 19-Jun-2014].
[41] ‘Confidence Intervals’. [Online]. Available: http://www.stat.yale.edu/Courses/1997-98/101/confint.htm. [Accessed: 19-Jun-2014].
[42] ‘Confidence interval - Wikipedia, the free encyclopedia’. [Online]. Available: http://en.wikipedia.org/wiki/Confidence_interval. [Accessed: 19-Jun-2014].
[43] ‘Box plot - Wikipedia, the free encyclopedia’. [Online]. Available: http://en.wikipedia.org/wiki/Box_plot. [Accessed: 19-Jun-2014].
[44] G. Tusman, E. Gogniat, F. Suarez-Sipmann, S. Bohm, E. San Roman, and A. Santos, ‘Reference Values For Non-Invasive Volumetric Capnographic Parameters In Awake And Anesthetized Humans’, *Am J Respir Crit Care Med*, vol. 185, p. A1663, 2012.
[45] ‘Notched boxes in box plots - MATLAB’. [Online]. Available: http://www.mathworks.com/help/symbolic/mupad_ref/notched.html. [Accessed: 20-Jun-2014].
[46] J. M. Chambers, *Graphical methods for data analysis*. Wadsworth International Group, 1983.
[47] ‘Least-Squares Fitting - MATLAB & Simulink’. [Online]. Available: http://www.mathworks.com/help/curvefit/least-squares-fitting.html. [Accessed: 21-Jun-2014].
[48] ‘Power law - Wikipedia, the free encyclopedia’. [Online]. Available: http://en.wikipedia.org/wiki/Power_law. [Accessed: 21-Jun-2014].
[49] ‘Logarithmic scale - Wikipedia, the free encyclopedia’. [Online]. Available: http://en.wikipedia.org/wiki/Logarithmic_scale. [Accessed: 21-Jun-2014].
[50] ‘When Should I Use Logarithmic Scales in My Charts and Graphs?’ [Online]. Available: http://www.forbes.com/sites/naomirobbins/2012/01/19/when-should-i-use-logarithmic-scales-in-my-charts-and-graphs/. [Accessed: 21-Jun-2014].
[51] ‘Skewness - MATLAB skewness’. [Online]. Available: http://www.mathworks.com/help/stats/skewness.html. [Accessed: 21-Jun-2014].