

---

# Capnographic Analysis for Disease Classification

by

Rebecca J. Asher

B.S., Electrical and Computer Engineering, Carnegie Mellon University, 2010

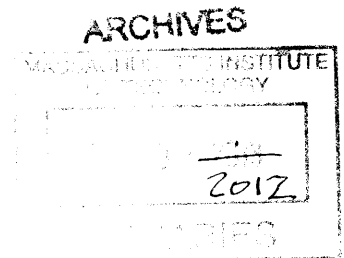
---

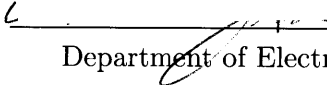
Submitted to the Department of Electrical Engineering and Computer Science  
in partial fulfillment of the requirements for the degree of

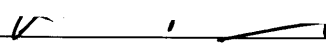
Master of Science  
in Electrical Engineering and Computer Science  
at the Massachusetts Institute of Technology

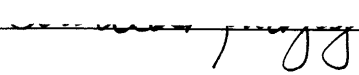
June 2012

© 2012 Massachusetts Institute of Technology  
All Rights Reserved.



Signature of Author:  \_\_\_\_\_  
Department of Electrical Engineering and Computer Science  
May 23, 2012

Certified by:  \_\_\_\_\_  
George C. Verghese  
Henry Ellis Warren Professor of Electrical Engineering  
Thesis Supervisor

Accepted by:  \_\_\_\_\_  
Leslie A. Kolodziej  
Professor of Electrical Engineering  
Chair, Committee for Graduate Students



---

---

## Capnographic Analysis for Disease Classification

by Rebecca J. Asher

Submitted to the Department of Electrical Engineering and Computer Science  
on May 23, 2012, in partial fulfillment of the requirements for the degree of  
Master of Science

### Abstract

Existing methods for extracting diagnostic information from carbon dioxide in the exhaled breath are qualitative, through visual inspection, and therefore imprecise. In this thesis, we quantify the CO<sub>2</sub> waveform, or capnogram, in order to discriminate among various lung disorders. Quantitative analyses of the capnogram are conducted by extracting several physiological waveform features and performing classification by discriminant analysis with voting. Our classification methods are tested in distinguishing between records from subjects with normal lung function and patients with cardiorespiratory disease. In a second step, we discriminate between capnograms from patients with obstructive lung disease (chronic obstructive pulmonary disease) and those with restrictive lung disease (congestive heart failure). Our results demonstrate the diagnostic potential of capnography.

---

Thesis Supervisor: George C. Verghese

Title: Henry Ellis Warren Professor of Electrical Engineering



---

---

# Acknowledgments

This work would not have been possible without the support and encouragement of several mentors. First and foremost, I thank Professor George Verghese for his patience and wisdom in guiding this thesis. His clear teaching style has elucidated many concepts I thought I would never understand. Thanks also go to Dr. Baruch Krauss of Boston Children's Hospital for initiating this research, for providing many capnographic records, and for tireless revision of this document. Baruch's enthusiasm for and knowledge of capnography have truly sustained this work and are helping to move it forward. Additional thanks go to Dr. Thomas Heldt for his advice in medical modeling, assistance in writing data collection protocols, and helpful instruction in respiratory physiology.

Many others have helped make life great over these past two years and have indirectly fueled this work in the process. To Mom, Dad, and Laura: you are amazing! Thanks for always being there and for your patience with me. I also thank Jolaade and Kate for the many conversations that have helped make my time at MIT so much more enjoyable. Thanks go to Alex for regularly feeding me, motivating me, and providing the encouragement I needed to complete this thesis. And I am extremely grateful to Sarah for her creative advice and support. Thanks also go to all my officemates and to the members of the Computational Physiology and Clinical Inference group for their constantly helpful feedback and insight!



---

---

# Contents

<b>List of Figures</b>	<b>9</b>
<b>1 Introduction</b>	<b>11</b>
<b>2 Background</b>	<b>15</b>
2.1 Respiratory Physiology . . . . .	16
2.2 Capnography Technology . . . . .	17
2.3 The Capnogram Signal . . . . .	19
2.4 Prior Analyses . . . . .	21
<b>3 Approaches to Feature-based Classification</b>	<b>23</b>
3.1 Preparing the Dataset . . . . .	24
3.2 Supervised Learning Methods . . . . .	26
<b>4 Discriminant Analysis Methods</b>	<b>29</b>
4.1 Linear Discriminant Analysis . . . . .	29
4.2 Diagonal Quadratic Discriminant Analysis . . . . .	31
<b>5 From Time Series to Feature Space</b>	<b>39</b>
5.1 Capnogram Pre-processing . . . . .	40
5.1.1 Exhalation Detection . . . . .	41
5.1.2 Template View . . . . .	42
5.1.3 Discarding Outlier Exhalations . . . . .	47
5.2 Feature Extraction . . . . .	50
5.2.1 Curve Fitting Parameters . . . . .	51
5.2.2 Physiological Features . . . . .	55
<b>6 Classification Results</b>	<b>59</b>
6.1 Dataset . . . . .	59
6.2 Voting Schema . . . . .	61
6.3 Results . . . . .	63

6.4 Misclassified Records . . . . .	67
<b>7 Conclusions</b>	<b>71</b>
<b>Bibliography</b>	<b>73</b>



---

---

## List of Figures

2.1	Alveolar gas exchange . . . . .	17
2.2	Dead-space ventilation . . . . .	18
2.3	Mainstream vs. sidestream capnography . . . . .	18
2.4	Normal capnogram appearance . . . . .	19
2.5	Obstructive capnogram appearance . . . . .	20
2.6	Capnographic shape indices . . . . .	21
3.1	Classification accuracy plateaus as the number of features increases . . . . .	24
3.2	Supervised learning process . . . . .	25
4.1	Choosing the LDA projection vector . . . . .	30
4.2	LDA thresholds . . . . .	30
4.3	Data separated better with a quadratic boundary . . . . .	31
4.4	Quadratic separator with 3 features . . . . .	36
4.5	Quadratic separator with 2 features . . . . .	37
5.1	Normal and abnormal capnogram morphologies . . . . .	40
5.2	Exhalation detection . . . . .	41
5.3	Anchoring exhalations . . . . .	42
5.4	Computing the template exhalation . . . . .	43
5.5	Computing standard deviation from the template exhalation . . . . .	44
5.6	Template view of an obstructive capnogram . . . . .	45
5.7	CHF templates . . . . .	46
5.8	COPD templates . . . . .	46
5.9	Normal templates . . . . .	47
5.10	Template views of normal and abnormal capnograms . . . . .	47
5.11	Normal record in which two breaths are excluded . . . . .	48
5.12	Normal record in which no breaths are excluded . . . . .	49
5.13	Waveform strips from normal and abnormal capnograms . . . . .	50
5.14	Analyzing the exhalation upslope . . . . .	51
5.15	Fitting the 2-parameter model to CHF data . . . . .	53

---

5.16	Pathologic data fitted with the 2-parameter model . . . . .	54
5.17	Attempting to fit the 2-parameter model to Normal data . . . . .	54
5.18	Measuring exhalation duration . . . . .	56
5.19	Measuring $ETCO_2$ . . . . .	57
5.20	Measuring end-exhalation slope . . . . .	57
6.1	Severity levels of CHF patients in the dataset . . . . .	60
6.2	Voting schema . . . . .	62
6.3	Testing on 35 exhalations from each record . . . . .	63
6.4	Sensitivity attained while varying the number of exhalations considered during training . . . . .	65
6.5	Sensitivity attained while varying the training set size . . . . .	66
6.6	Misclassified COPD waveform . . . . .	67
6.7	Misclassified CHF waveform . . . . .	68
7.1	Possible future capnography interface . . . . .	72

# Introduction

**C**APNOGRAPHY refers to the non-invasive measurement of the concentration of carbon dioxide exhaled in the breath. Carbon dioxide is a byproduct of tissue metabolism, and its concentration,  $[\text{CO}_2]$ , can be measured as a function of time or as a function of exhaled volume. These two types of capnography are described as time-based and volumetric, with time-based capnography being the common type in clinical use. Capnography was first introduced in 1943, when carbon dioxide was observed to absorb infrared radiation [12]. Although some modern devices now utilize Raman spectroscopy or photoacoustic methods to detect  $[\text{CO}_2]$ , the vast majority use infrared detection techniques.

Capnography monitors can be found in every operating room as monitoring  $[\text{CO}_2]$  in patients is an essential aspect of modern anesthesia and respiratory care. With the advent of more portable devices, or capnographs, capnography can now be used in ambulatory settings as well [1].

The waveform produced during time-based capnography is called a capnogram and contains much information about underlying respiratory dynamics. However, current methods for assessing the capnogram are based on subjective and qualitative pattern recognition, with the clinician observing the capnogram to see if it appears roughly

normal or abnormal.

Quantitative analysis of the capnogram would allow capnography to be used as a diagnostic tool. Developing a capnography-based monitoring system that could quantitatively classify different states of respiratory disease would constitute a significant improvement in diagnostics.

Several factors make capnography an attractive respiratory diagnostic tool. First, as a measure of ventilation, it accurately reflects underlying pulmonary physiology and pathophysiology. Second, capnography is an effort-independent measurement since it simply entails breathing normally through a nasal cannula. Unlike spirometry, the gold standard for measurement of airway obstruction, capnography does not require forced exhalation, which many children and subjects in respiratory distress are unable to perform. Third, with mathematical modeling, capnography provides an objective test: rather than relying on subjective qualitative observation for disease state classification, capnography allows for a quantitative respiratory assessment. However, present methods of inspecting the capnogram are not quantitative in nature and result in an underutilization of the capabilities of capnographic monitoring.

In order to investigate the efficacy of capnography in diagnostic settings, we analyze the capnogram by first decomposing the waveform into physiological features. These features are readily quantifiable and correspond to respiratory status. We then implement statistical classification methods to label, or “diagnose,” various patient records.

The success of these classification methods hinges upon there being discernible differences among capnograms representative of distinct disease classes. Obstructive and restrictive lung disease constitute the main types of respiratory pathology. In obstructive lung disease, exhalation becomes difficult with increased resistance in the

---

airway. Restrictive lung disease instead limits the amount of air that can be inhaled and does not allow for proper filling of the lungs. An obstructive capnogram appears very differently from its restrictive counterpart. Capnogram features such as roundedness, duration, and height can be seen to vary among diseases [18, 28].

We quantify these features to separate capnographic measurements collected from three types of patients: those that are in a normal state of health, those with congestive heart failure (CHF), and those with chronic obstructive pulmonary disease (COPD). These classes comprise a broad range of pulmonary states since CHF can be classified as a restrictive disease and COPD as an obstructive disease [23].

Statistical classification begins after feature extraction and pre-processing of the capnogram. In the context of quadratic discriminant analysis, we generate probabilistic models to construct a quadratic boundary separating different classes of exhalations in feature space. Individual exhalations classified in this manner then “vote” on the label of their corresponding patient record. In turn, several different quadratic classifiers constructed by training on different partitions of the training set then vote on the final classification of a patient record. This voting scheme is found to boost classification performance without demanding the use of more computationally intensive classification techniques. Resulting test record labels are found to compare well with clinicians’ diagnoses. Such performance motivates the use of capnography in diagnostics.

The remainder of this thesis is organized as follows. Background information necessary to the understanding of capnogram analysis is presented in Chapter 2. Chapter 3 provides an overview of feature-based classification, with an emphasis on supervised learning methods. A detailed description of discriminant analysis is presented in Chapter 4. The subsequent two chapters are specific to the patient dataset used during classification: Chapter 5 details pre-processing and feature extraction methods, while

Chapter 6 reports classification results. Finally, Chapter 7 summarizes the contributions of the thesis and touches on the possible directions of future work.

# Background

**T**ODAY , monitoring with capnography is standard practice for many aspects of clinical care. Smaller, more portable capnography devices have made it feasible to use capnography in ambulatory settings for many different clinical applications. These modern capnographs employ highly specific carbon dioxide sensors and allow for use in spontaneously breathing subjects via nasal cannulae. Historically, though, capnography has been limited to the operating room and intensive care unit to confirm correct endotracheal tube (ETT) position. If the ETT is correctly placed in the trachea, CO<sub>2</sub> will immediately be detected and a capnogram generated. If incorrectly placed in the esophagus, where there is no significant CO<sub>2</sub> content, no capnogram will be observed. Accidental esophageal intubation represents a real danger during anesthesia and must be detected immediately [8]. The rapid nature of capnographic monitoring proves useful for this task.

Although capnography has filled a monitoring need in the operating room, current bedside CO<sub>2</sub> monitors typically consider only a few sparsely sampled values of the exhaled CO<sub>2</sub> concentration and constitute an underutilization of the capabilities of capnography. A more robust use of capnography would be in the field of diagnostics since capnography reveals much information about the underlying state of the cardiorespiratory system.

Toward the goal of improving the diagnostic precision of capnography through capnogram quantification, a place to start is with two of the most common cardiopulmonary diseases, CHF and COPD. These two diseases may present with similar symptoms such as shortness of breath and difficulty breathing, but have very different physiological implications. COPD results from obstruction of the airway and limits exhalation, while CHF typically causes fluid buildup in the lungs and limits the air that can be taken in during inhalation. Accurate diagnoses are critical in initiating effective treatment for these conditions, but because of similar presenting symptoms, rapid diagnosis is not always straightforward [22].

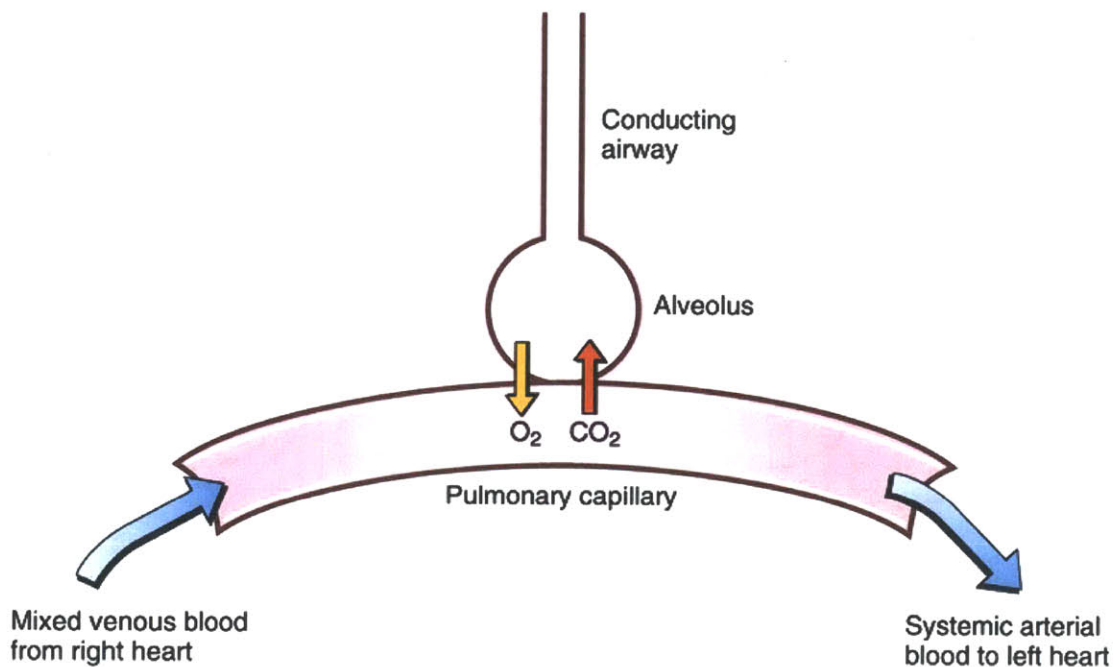
## ■ 2.1 Respiratory Physiology

During normal breathing, air is regularly inhaled and exhaled. In this way, oxygen is delivered to the body and carbon dioxide is removed. This gas exchange occurs as blood passes through the pulmonary capillary bed into the alveoli (Figure 2.1). During inspiration, oxygen diffuses from the alveoli into the capillaries. Because the concentration of carbon dioxide in room air is very close to zero,  $[\text{CO}_2]$  is insignificant during inspiration. During exhalation, carbon dioxide is transported from deoxygenated venous blood into the alveoli.

Initially, gas expelled during exhalation originates from the upper airway and contains no  $\text{CO}_2$ , corresponding to dead-space ventilation. A significant volume of  $\text{CO}_2$ -free air is exhaled in each breath since inspired air filling the upper conducting airways does not participate in gas exchange (Figure 2.2).

Dead-space volume is approximately 150 mL in an adult and constitutes roughly one third of the tidal volume, which is the amount inhaled air per breath. After exhalation, dead space again comes into play when residual alveolar gas is expelled from





**Figure 2.1.** Diffusion of oxygen out of the alveolus into the blood during inspiration and of carbon dioxide from the blood into the alveolus during exhalation [9].

the alveoli, but remains in the conducting airways. This residual gas is then inspired during the next inhalation.

## ■ 2.2 Capnography Technology

Infrared sensors are the main means of carbon dioxide detection in modern capnographs. Because carbon dioxide exhibits a very specific absorption at a wavelength of  $4.26 \mu\text{m}$ , the sensors function well in detecting  $\text{CO}_2$  [30].

Portable capnographs typically perform sidestream capnography, in which the infrared sensor responsible for detecting carbon dioxide is located in the monitor. Exhaled air is actively aspirated to reach the  $\text{CO}_2$  sensor. In mainstream capnography,

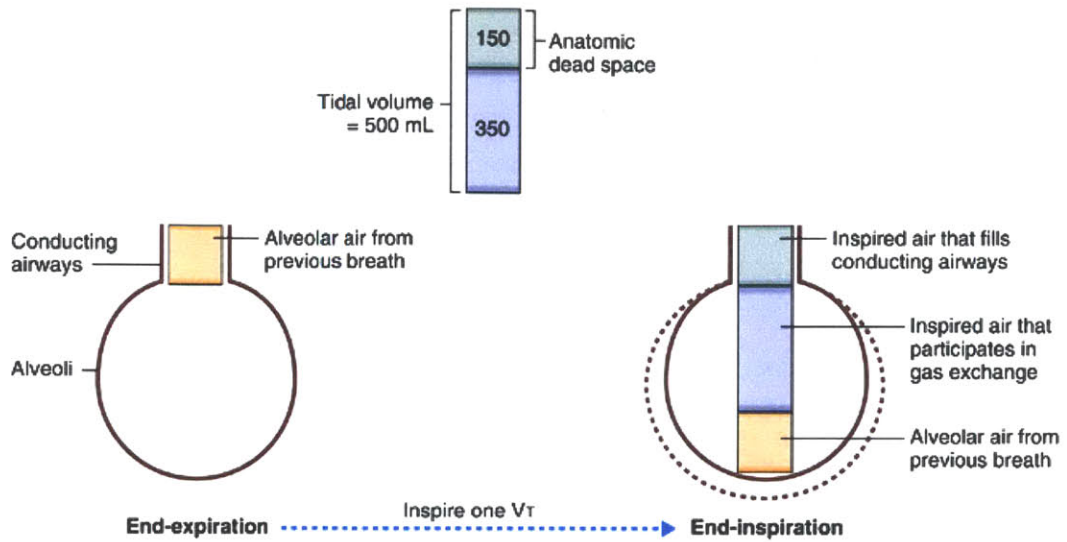


Figure 2.2. Dead-space ventilation [9].

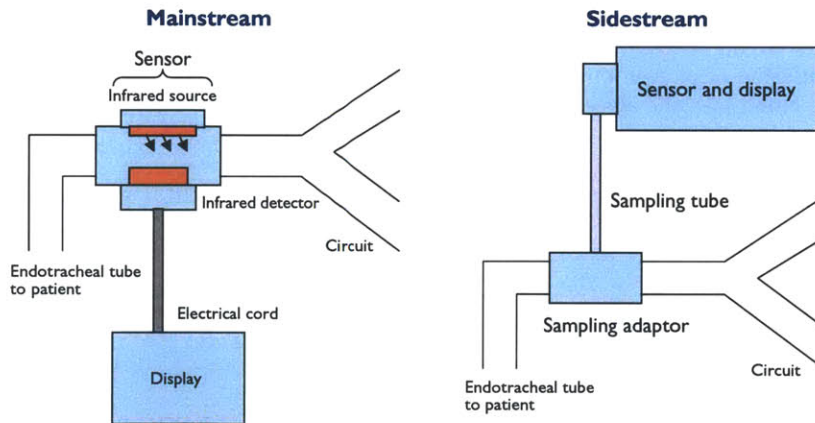


Figure 2.3. Mainstream vs. sidestream capnography [24]. The depicted setups are specific to ventilated patients. On the righthand side of each configuration, the “Circuit” sections indicate ventilator connections that will not be present during non-intubated capnography.

the sensor is located in line with the breathing circuit and is therefore reserved for intubated patients [16]. Figure 2.3 illustrates the distinction between mainstream and sidestream capnography.

## ■ 2.3 The Capnogram Signal

Each phase of the capnogram corresponds to a specific segment of breathing. The normal capnogram is roughly trapezoidal in appearance and is typically divided into four phases (Figure 2.4). Dead-space ventilation occurs during the first phase of exhalation, the start of alveolar gas exhalation during the second, an alveolar plateau during the third, and an inspiratory downstroke constitutes the fourth phase to complete the waveform. Each of these phases can be estimated as a straight line segment in the normal subject, and the terminal value of alveolar  $[\text{CO}_2]$  is defined as the End-Tidal  $\text{CO}_2$  (ET $\text{CO}_2$ ), the maximum  $[\text{CO}_2]$  in each breath.

In order to be classified as normal, a capnogram must exhibit the aforementioned

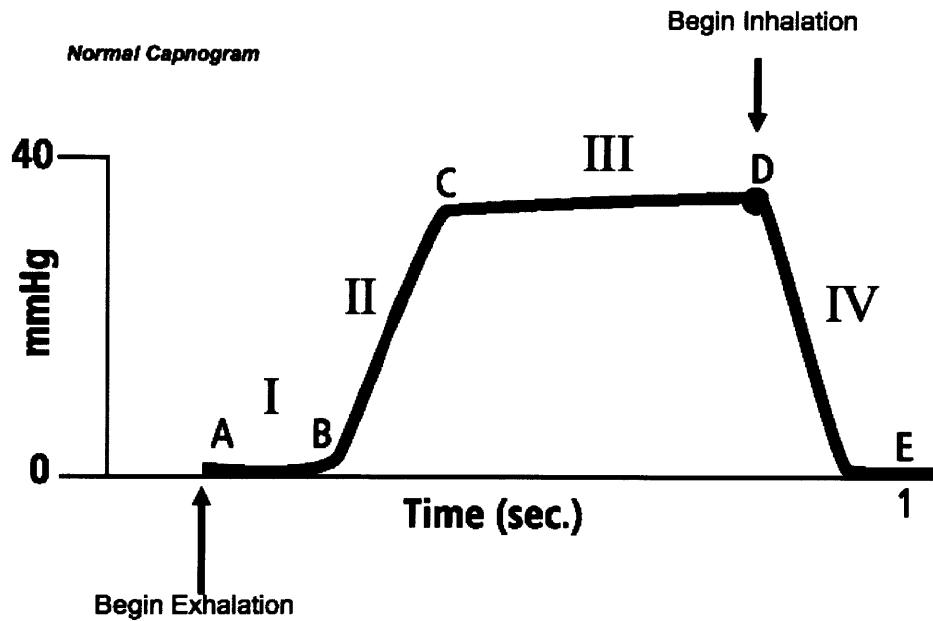
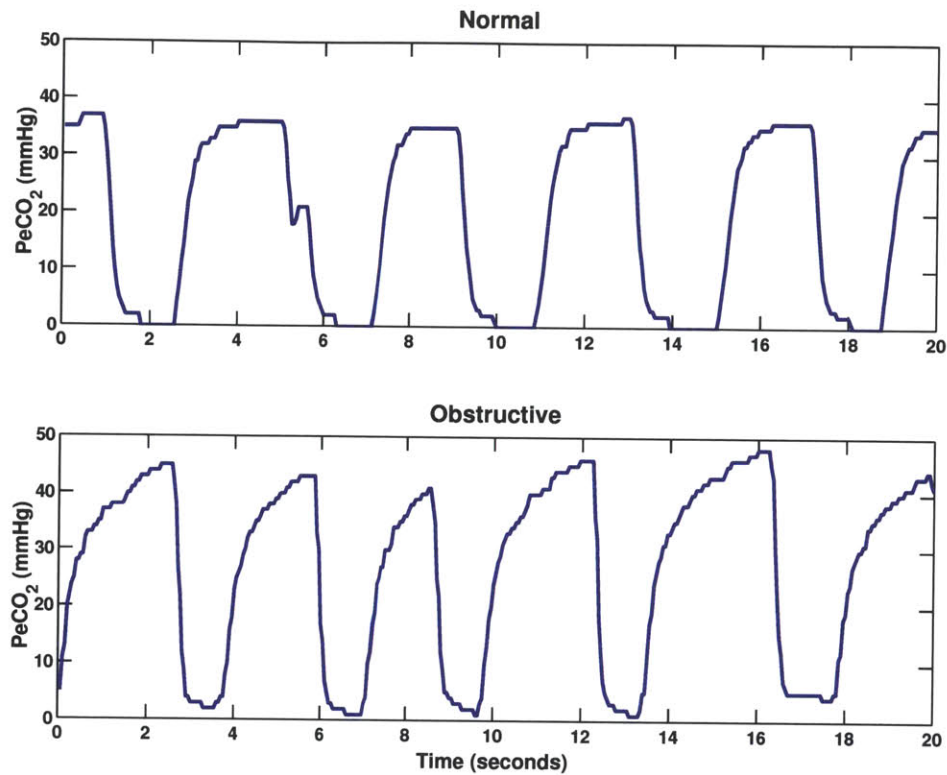


Figure 2.4. Normal capnogram appearance [19].



**Figure 2.5.** The normal capnogram (*above*) is distinctly different from the obstructive capnogram (*below*).

four phases, the  $[\text{CO}_2]$  must be zero at the start and end of the breath, and the  $\text{ETCO}_2$  must reach a normal level of 35-40 mmHg during each breath [12]. Several key clues from the capnogram can be used to assess underlying respiratory function.

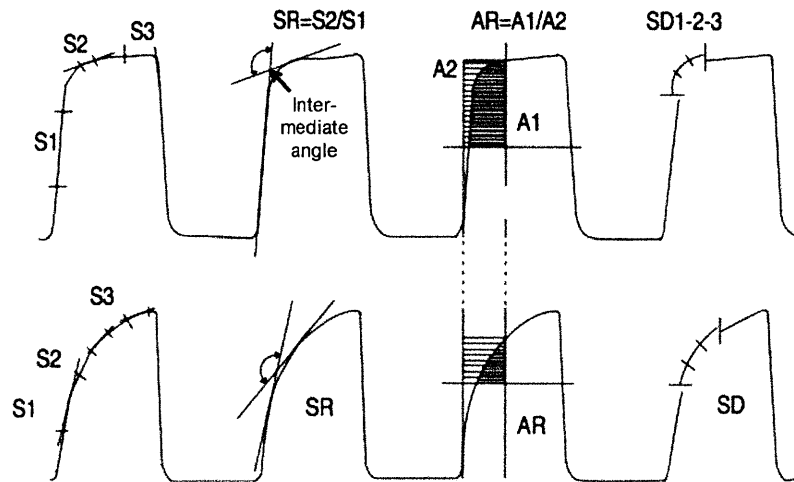
Figure 2.4 depicts the normal capnogram. However, in diseased states, the capnogram can take on a very different morphology. The two main classes of respiratory diseases considered in this thesis are obstructive and restrictive. Airway obstruction in diseases like asthma or COPD can cause a curved, “shark’s fin” appearance to the capnogram. The change in capnogram shape in obstructive lung disease correlates with a reduction in spirometric measures [18]. In restrictive lung disease, the waveform tends

to appear more compact, as the exhalation duration is shorter and  $\text{ETCO}_2$  is lower [4].

### ■ 2.4 Prior Analyses

In the 1950s and 1960s, early studies investigated the appearance of the normal capnogram and sought to model alveolar  $\text{CO}_2$  levels using time-dependent dilution equations [7, 32]. Linear segments were also fitted to the capnogram, and the canonical four phases associated with the normal capnogram were established. During this period, few researchers considered the appearance of the abnormal capnogram. However, ventilation-perfusion mismatch was noted to produce non-uniformity in the capnogram [15], and some work in the 1950s alluded to the possibility of using time-based capnography in the diagnosis of obstructive lung disease [13].

Later research during the 1990s examined the shape of abnormal capnograms more closely. A study conducted in mechanically ventilated dogs found that a sloping



**Figure 2.6.** Capnographic shape indices are defined in both normal (*above*) and asthmatic (*below*) conditions. Parameters considered include tangent slopes  $S_1$ ,  $S_2$ , and  $S_3$ , slope ratio  $SR$ , areas  $A_1$  and  $A_2$ , area ratio  $AR$ , and second derivatives  $SD$  [34].

alveolar plateau was produced as a result of uneven ventilation and prolonged gas exchange from the blood to the alveoli during exhalation [25]. In 1994, a French study investigated the links between capnogram shape indices and spirometric parameters in normal and asthmatic subjects. In the asthmatic subjects, the capnogram intermediate angle, measured between the initial expiratory upstroke and the alveolar plateau, was found to correlate with severity of airway obstruction [34]. Depicted in Figure 2.6, several capnogram shape and angle indices were defined in a small sample of subjects.

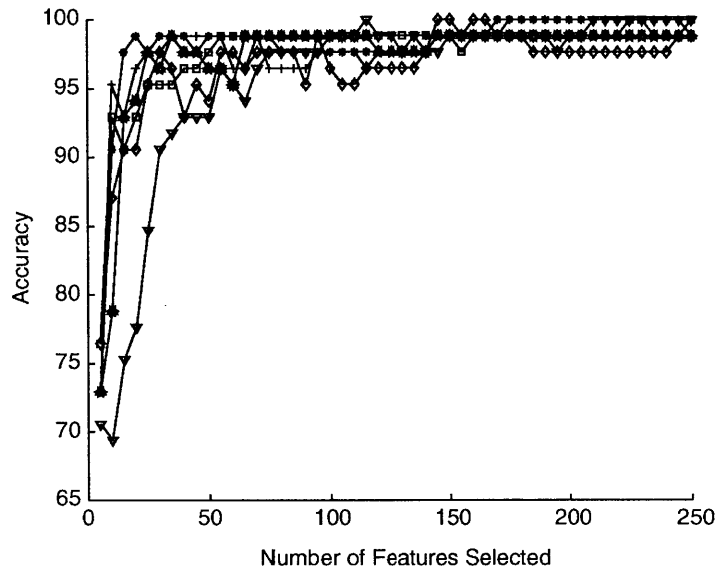
Previous research efforts have investigated the shape of both the normal and the abnormal capnogram. In these investigations, line segments were frequently fitted to the capnogram, and various geometric quantities were estimated. Although general features of the capnogram were observed, attempts were not made to develop classification rules based on these features. Our classification work, including feature-based classification methods (Chapter 3) and discriminant analysis techniques (Chapter 4), moves beyond basic feature identification and uses such features to classify normal and abnormal capnograms. While several small quantitative studies of the capnogram have been conducted, none has examined restrictive and obstructive lung disease in the hopes of distinguishing them. We expand quantitative capnogram analysis by incorporating modern classification tools that may prove useful in enhancing the diagnostic capabilities of capnography.

# Approaches to Feature-based Classification

**C**LASSIFICATION is frequently carried out by first decomposing training and test data into features that are readily quantified and classifiable. In supervised learning, a labeled training feature set is used to allow the model to learn. Then labels or predictions are assigned to the unlabeled test dataset. Choosing a robust set of features is thus very important.

Interestingly, using a large number of dataset features is not always advantageous. This is because features may be dependent upon one another, rendering some irrelevant. Including many correlated features results in poorly conditioned learning. In Figure 3.1, classification accuracy plateaus after a certain number of features is used. Much research has been done on how to select the most appropriate features for solving a supervised learning problem [11, 17].

Time-series data features could be constructed by taking a signal's average value over time or its wavelet transform decomposition, by recording peak values, or in many other ways. Various gene expressions are commonly used as features in genomic datasets generated by DNA microarrays. In ECG analysis, feature extraction typically involves



**Figure 3.1.** Classification accuracy is plotted as a function of the number of features present in the feature set. Different symbols correspond to different classification methods. After roughly 50 features, the accuracy no longer increases with the feature set size. Adapted from [20].

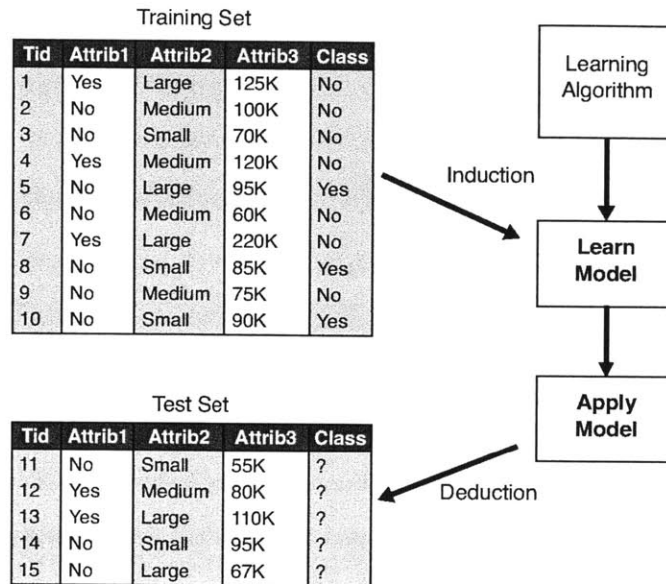
recording characteristics of the QRS complex or other marked segments of the waveform.

An essential task in capnogram classification is to select a small set of powerful features that perform well in separating different disease classes. As will be seen in Chapter 5, these features also turn out to be physiologically rooted.

### ■ 3.1 Preparing the Dataset

Since in supervised learning, we first start with a labeled training set and aim to subsequently classify an unlabeled test set, the relative sizes of the two sets will impact classifier performance. Various training and test set partitions have been used in the literature. In the process displayed in Figure 3.2, a learning algorithm is applied to develop a classifier using a training set. The classifier is then implemented to label a





**Figure 3.2.** Supervised learning conducted on a training set that is larger than the test set. Adapted from [29].

smaller test set.

Generally, the training set will comprise many more data samples than the test set. A 70%/30% partition is used in Chapter 6 for our classification experiments. In order to avoid overfitting during the supervised learning process, it is helpful to maintain a diverse test set containing a broadly representative sample of the data.

The test set size is often chosen to be inversely proportional to classifier accuracy [14]. Occasionally, a single test set is treated as a “holdout” set that is kept aside and used only for error estimation. In other settings, a validation set is also formed and used to choose which model will be implemented.

### ■ 3.2 Supervised Learning Methods

Many techniques have been developed to tackle the problem of supervised learning. Common methods include classifiers such as decision trees, probabilistic graphical models, k-nearest neighbors, discriminant analysis, and support vector machines.

- **Tree-based methods:** Decision trees sequentially divide the data based on feature values. At each node in the tree, the dataset is split based on a given feature. Thus, the features that best classify the data will be placed at the root of the tree. Trees can be quickly implemented and are easily understood, but they tend to overfit the data.
- **Bayesian networks:** Probabilistic graphical models like Bayesian networks use graphs to represent probabilistic influences among events. These types of directed graphical models are frequently used to identify signaling pathways and in fault diagnosis [5, 35].
- **Clustering methods:** Techniques such as the k-nearest neighbor algorithm assume that like data will be proximal in feature space and classify unlabeled test points based on which labeled training points are the closest neighbors in feature space. These methods perform surprisingly well, but they tend to demand large datasets.
- **Discriminant analysis:** Discriminant analysis methods take into account a prior probability distribution on the dataset and map the data to lower dimensions in feature space to select the classifier that maximizes interclass variance and minimizes interclass variance. The technique is straightforward and works even on small datasets.

- **Kernel-based methods:** Support vector machines (SVMs) search for the optimal hyperplane that will maximize the margin between classes of the dataset. Although this optimal boundary may sometimes be linear, kernel functions allow for nonlinear feature combinations [26]. Kernel-based methods can produce strong classifiers, but can also demand large training sets and prove to be computationally intensive.

Many classification methods exist, and different problems demand different learning techniques. Rule-based learners such as decision trees tend to work best with discrete feature values [21]. For continuous features, such as those extracted during capnogram classification, the robust classification options are narrowed down to k-nearest neighbors, discriminant analysis, or kernel-based techniques.

We opt to implement discriminant analysis methods and describe more about the technique in Chapter 4. Alternatives such as kernel-based techniques typically require a very large dataset in order to produce reliable classification. As will be seen in Chapter 6, discriminant analysis proves to be a suitable method in the context of our supervised learning problem given that the dataset is relatively small and is well-described by a small number of features.



# Discriminant Analysis Methods

**D**ISCRIMINANT analysis is unlike more modern methods of classification, such as support vector machines, in that it is very straightforward and performs well on small datasets, as mentioned in Chapter 3. Many challenging supervised learning problems in genomics and imaging still implement discriminant analysis methods.

In particular, we focus on the subtopic of quadratic discriminant analysis, which will be implemented in Chapter 6. Quadratic discriminants result from an extension of linear discriminant analysis. They allow for quadratic combinations of features and yield more robust classification schemes than linear discriminants.

## ■ 4.1 Linear Discriminant Analysis

First proposed by Ronald Fisher in 1936 for use in taxonomic classification, linear discriminant analysis (LDA) projects high-dimensional data to one dimension in feature space [33]. Given a particular feature space, an initial task is choosing the projection vector. Data are projected onto this vector for subsequent classification. Figure 4.1 displays two different projection vectors for the same dataset. There exist both good and bad projection vectors, as some vectors allow the data to be readily separable while others are not helpful in classification of the presented datasets.

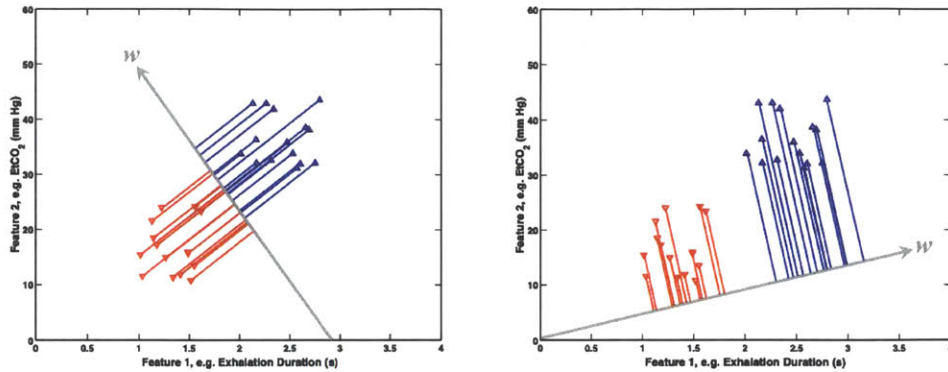


Figure 4.1. Data projected to a poorly chosen vector (*left*) and a more appropriate vector (*right*).

After selecting an appropriate projection vector, the next step is to decide where to bisect the vector and effectively create a separating plane. Figure 4.2 shows thresholds corresponding to the two projection vectors shown in Figure 4.1. The first projection vector produces poor classification while the second performs well.

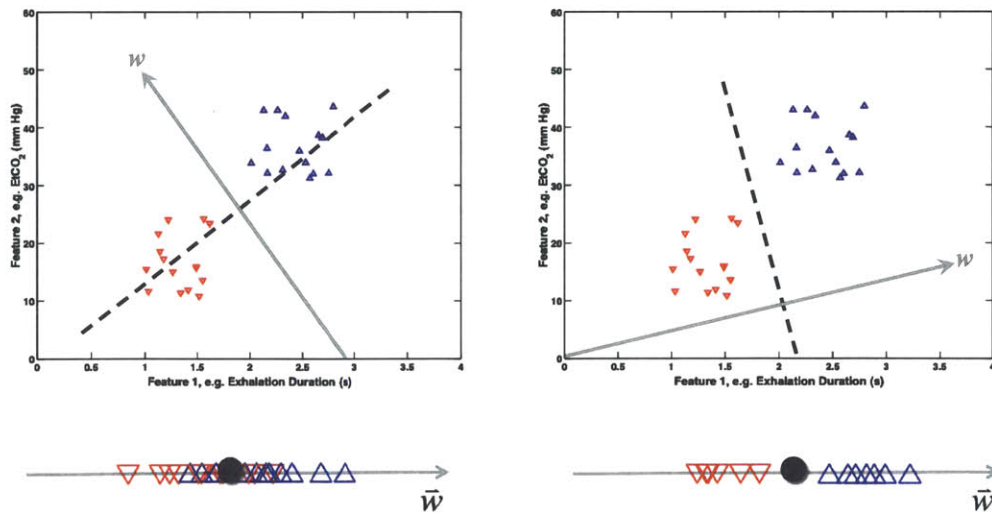
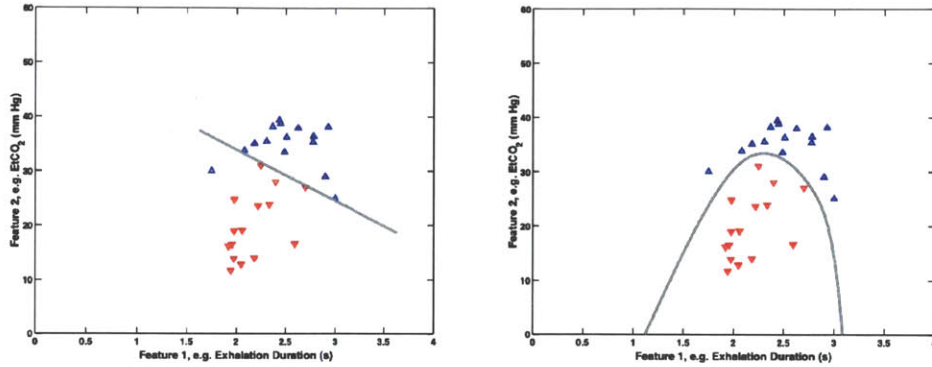


Figure 4.2. Thresholds bisecting the poor-performing vector (*left*) and the appropriate vector (*right*).

Occasionally, datasets are not separated well by a linear boundary, but do much better when a quadratic boundary is used. Figure 4.3 shows a dataset partitioned by

a linear separator and a quadratic separator. In this case, using a form of discriminant analysis more elaborate than LDA, namely quadratic discriminant analysis (QDA), yields better performance.



**Figure 4.3.** Data are separated poorly with a linear boundary (*left*) and much better with a quadratic separator (*right*).

## ■ 4.2 Diagonal Quadratic Discriminant Analysis

In conducting QDA, a multivariate normal distribution is assumed. The training phase involves estimating the means and covariance matrices of each class in the training set. If the features are relatively uncorrelated within each class, then a more robust classification is obtained by constraining the covariance matrix estimates to be diagonal.

Say we have three features and two classes. Let us first compute the means of each training set. Mean  $\vec{\mu}_1$  corresponds to training data from class  $C_1$  and mean  $\vec{\mu}_2$  is computed from class  $C_2$ 's training samples.

$$\vec{\mu}_1 = \begin{bmatrix} \mu_{1,Feature1} & \mu_{1,Feature2} & \mu_{1,Feature3} \end{bmatrix} \quad (4.1)$$

and

$$\vec{\mu}_2 = \begin{bmatrix} \mu_{2,Feature1} & \mu_{2,Feature2} & \mu_{2,Feature3} \end{bmatrix}. \quad (4.2)$$

We next compute standard deviation vectors. Let vector  $\vec{\sigma}_1$  represent the standard deviation of the training data from the first class,  $C_1$ , and  $\vec{\sigma}_2$  be the standard deviation of the training data from the second class,  $C_2$ . Then

$$\vec{\sigma}_1 = \begin{bmatrix} \sigma_{1,Feature1} & \sigma_{1,Feature2} & \sigma_{1,Feature3} \end{bmatrix} \quad (4.3)$$

and

$$\vec{\sigma}_2 = \begin{bmatrix} \sigma_{2,Feature1} & \sigma_{2,Feature2} & \sigma_{2,Feature3} \end{bmatrix}. \quad (4.4)$$

The diagonal covariance matrices  $\Sigma_1$  and  $\Sigma_2$  can be expressed as

$$\Sigma_1 = \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_f^2 \end{bmatrix}_{C_1} \quad \Sigma_2 = \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_f^2 \end{bmatrix}_{C_2}, \quad (4.5)$$

where  $f$  represents the feature space dimension. Denote the feature vector of a typical data point in the test set by the row vector  $\vec{x}$ . In the case of three features, we have



$$\vec{x} = \begin{bmatrix} x_{Feat\ 1} & x_{Feat\ 2} & x_{Feat\ 3} \end{bmatrix} \quad (4.6)$$

We describe the class-conditional density,  $p(\vec{x}|C_k)$ , of a given sample  $\vec{x}$  using a multivariate normal distribution. Then we compute a ratio similar to the log-likelihood ratio, but also containing the class priors,  $p(C_k)$  [2]. This ratio is used as a test to determine whether the sample is a member of one class or another. For each class,  $C_k$  (with  $k = 1, 2$ ),

$$p(\vec{x}|C_k) = \frac{1}{(2\pi)^{\frac{f}{2}}} \frac{1}{|\Sigma_k|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (\vec{x} - \vec{\mu}_k) \Sigma_k^{-1} (\vec{x} - \vec{\mu}_k)^T \right\}, \quad (4.7)$$

The modified log-likelihood ratio then becomes

$$\begin{aligned} \ln \frac{p(\vec{x}|C_1)p(C_1)}{p(\vec{x}|C_2)p(C_2)} &= \ln \left\{ \frac{\frac{p(C_1)}{(2\pi)^{\frac{f}{2}} |\Sigma_1|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (\vec{x} - \vec{\mu}_1) \Sigma_1^{-1} (\vec{x} - \vec{\mu}_1)^T \right\}}{\frac{p(C_2)}{(2\pi)^{\frac{f}{2}} |\Sigma_2|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (\vec{x} - \vec{\mu}_2) \Sigma_2^{-1} (\vec{x} - \vec{\mu}_2)^T \right\}} \right\} \\ &= \ln \left\{ \frac{p(C_1)}{p(C_2)} \right\} + \frac{1}{2} \ln \left\{ \frac{|\Sigma_2|}{|\Sigma_1|} \right\} - \frac{1}{2} \left\{ (\vec{x} - \vec{\mu}_1) \Sigma_1^{-1} (\vec{x} - \vec{\mu}_1)^T - (\vec{x} - \vec{\mu}_2) \Sigma_2^{-1} (\vec{x} - \vec{\mu}_2)^T \right\} \\ &= \vec{x} W \vec{x}^T + 2\vec{x} \vec{w}^T + w_0 \end{aligned} \quad (4.8)$$

where

$$W = \frac{1}{2} (\Sigma_2^{-1} - \Sigma_1^{-1}) \quad (4.9)$$

$$\vec{w} = \frac{1}{2} (\Sigma_1^{-1} \vec{\mu}_1 - \Sigma_2^{-1} \vec{\mu}_2) \quad (4.10)$$

$$w_0 = \frac{1}{2} \left( \vec{\mu}_2 \Sigma_2^{-1} \vec{\mu}_2^T - \vec{\mu}_1 \Sigma_1^{-1} \vec{\mu}_1^T \right) + \frac{1}{2} \ln \left\{ \frac{|\Sigma_2|}{|\Sigma_1|} \right\} + \ln \left\{ \frac{p(C_1)}{p(C_2)} \right\}. \quad (4.11)$$

The decision boundary described in Equation 4.8 describes a quadratic decision surface of the form:

$$K + \vec{x}L + \vec{x}Q\vec{x}^T \geq 0 \quad (4.12)$$

When the argument takes a value greater than 0, sample  $\vec{x}$  is labeled as a member of  $C_1$ . If the argument is less than 0 for a given  $\vec{x}$ , that  $\vec{x}$  is labeled a member of  $C_2$ . Because this discriminant compares posterior densities in order to classify a sample, it looks very much like a maximum a posteriori (MAP) estimator.

MAP estimators typically use the ratio presented in Equation 4.8 in order to determine an unknown  $\vec{x}$  [10]. Our application is a bit different in that we know the location of test samples  $\vec{x}$  in feature space and are estimating which class membership is most likely. The resulting decision boundary has coefficients  $K$ ,  $L$ , and  $Q$ , corresponding to constant, linear, and quadratic terms, respectively. The constant coefficient,  $K$ , is a scalar and is expressed as:

$$\begin{aligned}
K = \frac{1}{2} & \left\{ \sum_i \left( \frac{1}{\sigma_{2,Feat\ i} \mu_{2,Feat\ i}} \right)^2 - \sum_i \left( \frac{1}{\sigma_{1,Feat\ i} \mu_{1,Feat\ i}} \right)^2 \right\} \\
& + \sum_i \ln \sigma_{2,Feat\ i} - \sum_i \ln \sigma_{1,Feat\ i}
\end{aligned} \tag{4.13}$$

$L$ , the linear coefficient, is represented by a column vector:

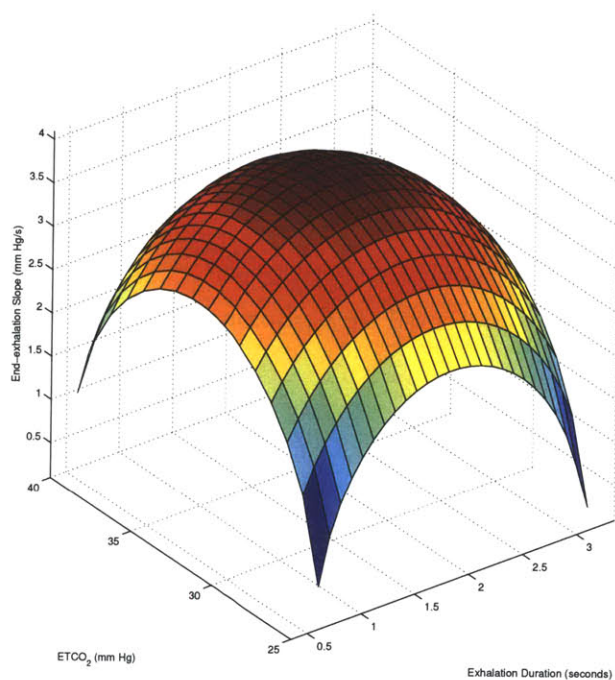
$$L = \begin{bmatrix} \left( \frac{1}{\sigma_{1,Feat\ 1}^2 \mu_{1,Feat\ 1}} \right) - \left( \frac{1}{\sigma_{2,Feat\ 1}^2 \mu_{2,Feat\ 1}} \right) \\ \left( \frac{1}{\sigma_{1,Feat\ 2}^2 \mu_{1,Feat\ 2}} \right) - \left( \frac{1}{\sigma_{2,Feat\ 2}^2 \mu_{2,Feat\ 2}} \right) \\ \left( \frac{1}{\sigma_{1,Feat\ 3}^2 \mu_{1,Feat\ 3}} \right) - \left( \frac{1}{\sigma_{2,Feat\ 3}^2 \mu_{2,Feat\ 3}} \right) \end{bmatrix} \tag{4.14}$$

Finally, the quadratic coefficient,  $Q$ , is a diagonal matrix:

$$Q = \begin{bmatrix} \left( \frac{1}{\sigma_{1,Feat\ 1}^2} - \frac{1}{\sigma_{2,Feat\ 1}^2} \right) & 0 & 0 \\ 0 & \left( \frac{1}{\sigma_{1,Feat\ 2}^2} - \frac{1}{\sigma_{2,Feat\ 2}^2} \right) & 0 \\ 0 & 0 & \left( \frac{1}{\sigma_{1,Feat\ 3}^2} - \frac{1}{\sigma_{2,Feat\ 3}^2} \right) \end{bmatrix}, \tag{4.15}$$

resulting in the separator from Equation 4.12.

Figure 4.4 shows such a quadratic classification boundary in three dimensions. Quadratic combinations of the three features are considered, but because our analysis considers a diagonal quadratic classifier, there are no cross-terms and, where  $x$  represents Feature 1,  $y$  is Feature 2, and  $z$  is Feature 3, the surface will take the form:



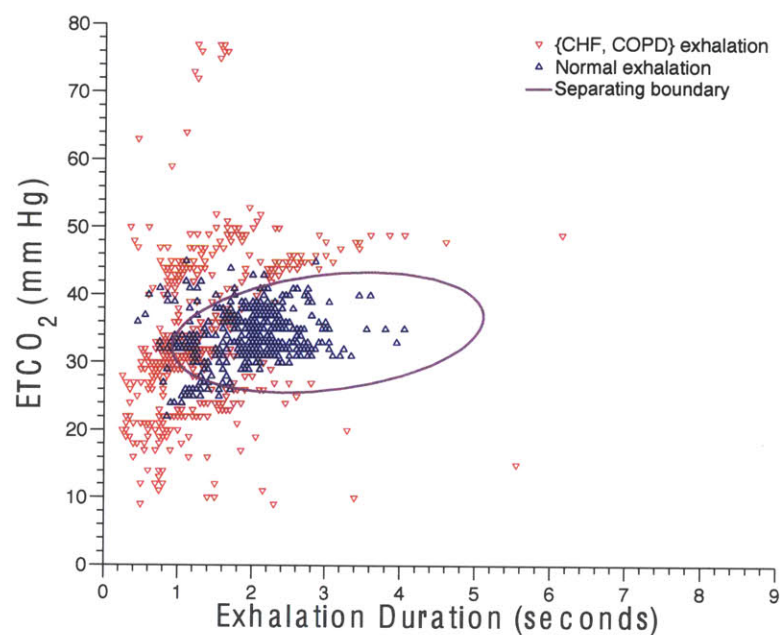
**Figure 4.4.** A quadratic separator with 3 features.

$$0 = K + L_1x + L_2y + L_3z + Q_1x^2 + Q_2y^2 + Q_3z^2 \quad (4.16)$$

Individual exhalations falling on one side of the separator will be classified as one class, while exhalations falling on the other side will be considered as members of the other patient class.

In Figure 4.5, we see a quadratic separator projected into 2 dimensions. Here there are only two features considered in classification, and the separator does a reasonable job distinguishing between the two patient classes when deciding on individual exhalations.

Discriminant analysis, although an old method for classification, holds up well with small datasets and is useful in classifying individual exhalations of the capnogram.



**Figure 4.5.** Quadratic separator with 2 features.

The next chapter describes how to extract features from these individual exhalations. In Chapter 6, subsequent voting methods will be outlined that combine the verdicts of several classifiers created using quadratic discriminant analysis. In this way, even if one classifier is not performing well individually, the ensemble of various classifiers can exhibit boosted accuracy.

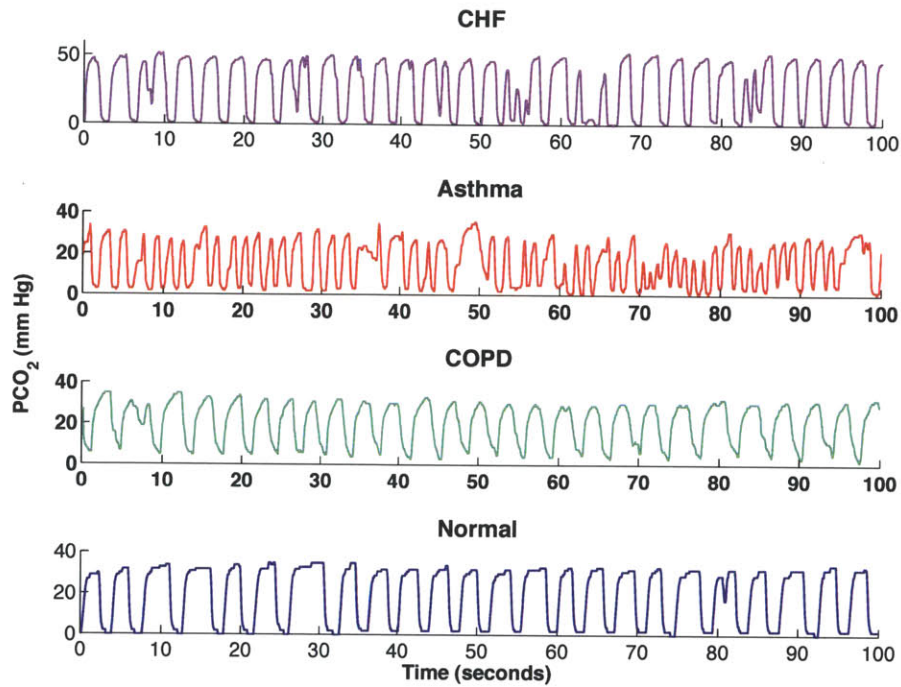


# From Time Series to Feature Space

**B**OTH pre-processing and feature extraction are performed to prepare the data for discriminant analysis and later consideration. Toward the goal of robustly detecting outlier exhalations and excluding them from consideration, a template exhalation of each record is constructed. This template also proves useful in visual inspection of the capnogram and could be used as an alternative way to examine records, instead of the conventional waveform strip.

Time-series capnographic data of varying morphologies are shown in Figure 5.1. It can be seen that capnogram shapes vary greatly from class to class. This fact is useful in later classification of capnograms by disorder. In order to implement feature-based classification, distinct features must be formulated and extracted.

Features formulated during curve-fitting and those rooted in physiology are considered when extracting features from the capnogram. Curve-fitting features include parameters that correspond to fitting the capnogram with exponential curves. More relevantly, physiological features directly relate to the respiratory system and are also useful in classification. These include exhalation duration,  $ETCO_2$  levels, and end-exhalation slope.



**Figure 5.1.** Capnograms from various patient classes, including restrictive disease (CHF), obstructive disease (Asthma, COPD), and Normal.

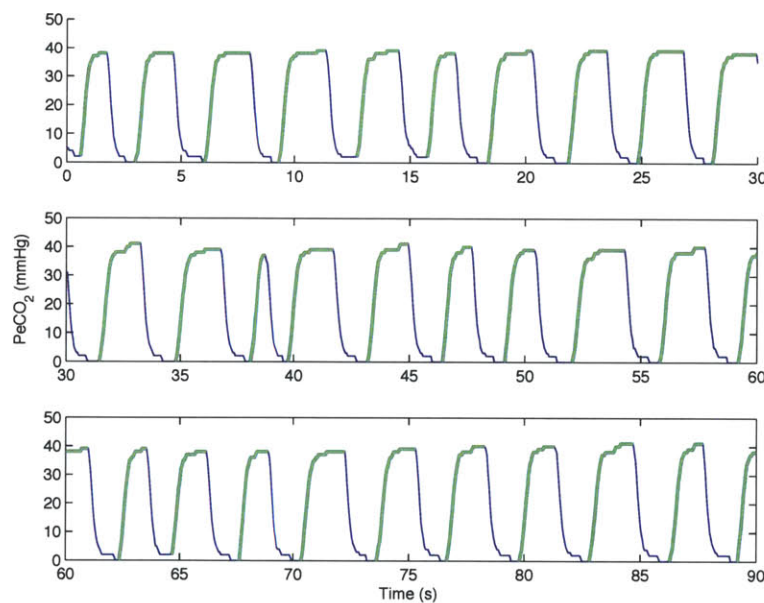
## ■ 5.1 Capnogram Pre-processing

In preparing capnograms for classification, exhalations must be quantified. Accurately detecting the beginning and end of exhalation is crucial to the robustness of subsequent analyses. After exhalation detection, a composite view is formulated in which capnograms are described by a single exemplary exhalation. This template is then used to judge exhalations and determine which should be excluded from analysis.



### ■ 5.1.1 Exhalation Detection

Specifying the start and stop of exhalation is essential to processing a capnogram signal. Although prior studies have used more complex schemes such as artificial neural networks for breath detection, looking for a slope change from negative to positive at the beginning of exhalation and from positive to negative at the end of exhalation seems to mark breaths reasonably well. Figure 5.2 displays the capnogram and detected exhalations.



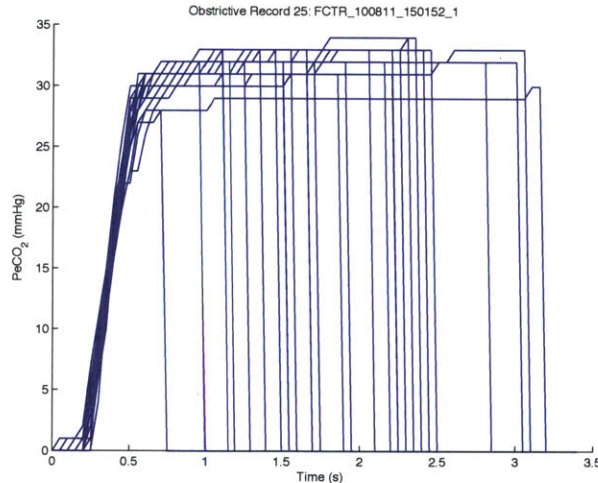
**Figure 5.2.** Using slope changes to indicate the start and stop of exhalations. The capnogram is shown in *blue*, and detected exhalation segments are marked in *green*.

When attempting to detect exhalations with the slope change method, challenges include sections of breathing that oscillate or change slope outside of the start and stop of exhalation. These occurrences can be quite frequent and are mitigated by imposing  $[\text{CO}_2]$  thresholds on what constitutes the beginning and the end of an exhalation.

### ■ 5.1.2 Template View

Capnograms are most commonly viewed as a waveform strip on a long time-axis. Examining records that are 20-30 minutes long in this fashion can be a daunting task. To facilitate the viewing of the capnogram, a template view presents an alternative to the ordinary sequential waveform view of the capnogram signal. Breaths from a single record are overlaid over the duration of the longest exhalation in the record.

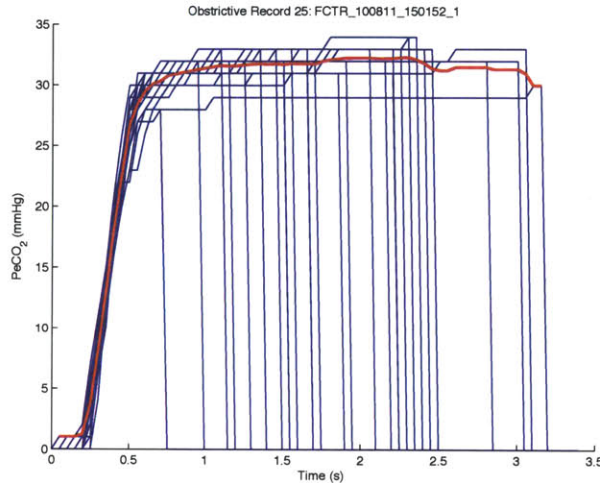
To do this, the exhalations are anchored at some fixed value of  $\text{PeCO}_2$ , the partial pressure of carbon dioxide in the exhalate; we chose 15 mmHg. Thus, all exhalations cross 15 mmHg at the same time in the template view. This anchoring is shown in Figure 5.3. In this way, more of a composite breath is seen and the eye is not thrown off by outlier breaths as much as when viewing the time-based capnogram in the ordinary way.



**Figure 5.3.** Anchoring all exhalations of a single record at a  $\text{PeCO}_2$  of 15 mmHg.

Then, the average exhalation is computed. This is done by taking the mean  $\text{PeCO}_2$  at every time sample in the overlaid exhalations. The average exhalation can be

thought of as a composite breath that is representative of the record as a whole. Rather than paying too much attention to outlier exhalations, the composite exhalation, shown in Figure 5.4, allows for quick viewing of each record.

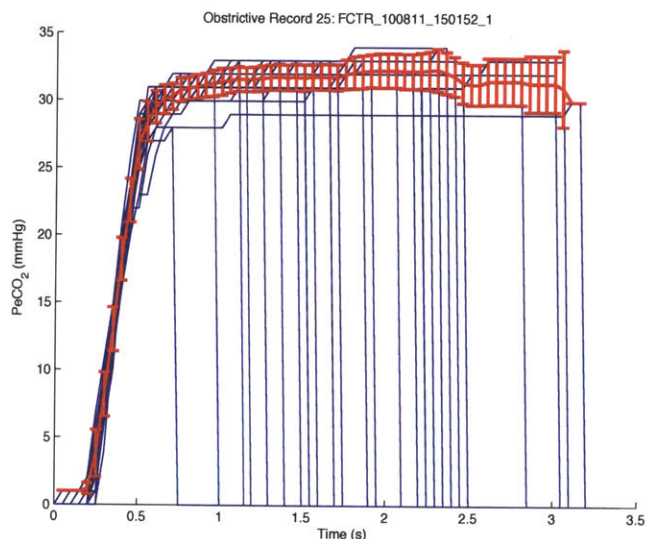


**Figure 5.4.** Computing the template as the average exhalation.

Toward the goal of cropping outlier breaths, the standard deviation from the average exhalation, or template, is computed. At each sample in the exhalation, the standard deviation bars can be seen in Figure 5.5.

Again, the template view proves very useful in that a general, or exemplary, exhalation can be seen when all the exhalations from a single record are overlaid. Figure 5.6 displays one obstructive disease record in the template view. All the exhalations are plotted, with the average or template exhalation highlighted. Since most of the exhalations cluster around a 2.5-second duration, it is easy for the eye to discard the several outlier exhalations when examining the general trend of the record.

In classification quizzes administered to a knowledgeable physician and two biomedical researchers first with the standard view and then later with the template view, the template seemed to be the preferred way to view the capnogram. The quiz setup in-



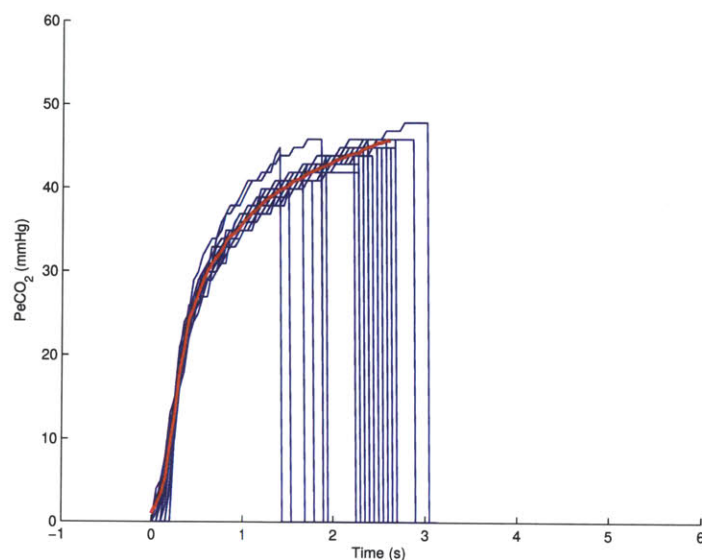
**Figure 5.5.** Computing the standard deviation of the template exhalation.

	Waveform View	Template View
Average Performance	70.4%	77.8%

**Table 5.1.** Quiz performance on 9 records from CHF, COPD, and Normal classes.

volved nine records being presented to evaluators familiar with capnogram analysis. Quiz records were either CHF, COPD, or Normal. The evaluators classified each of the records as one of these types and, as shown in Table 5.1, performed nearly 10% better when presented with the template view.

Sample CHF, COPD, and Normal records are shown in Figures 5.7, 5.8, and 5.9. To readily see the distinctions among different disease states, records are shown in template view. Note the generally shorter exhalation duration and smaller  $\text{ETCO}_2$  of the CHF capnograms, the longer duration and larger  $\text{ETCO}_2$  of the COPD capnograms, and the moderate appearance of the Normal capnograms. Viewing the templates in this



**Figure 5.6.** Template view of an obstructive capnogram. The curved shape is seen, and the template exhalation is shown in *red*.

way presents a clearer picture of the distinctions among different classes.

In an even more condensed presentation, Figure 5.10 shows all quiz records overlaid by class. One can readily pick out the waveform differences and see the rounded shape of the obstructive capnograms, the compact and shortened form of the CHF records, and the moderate, almost rectangular, morphology of the Normal set. The template view provides one way to better view the capnogram.

The canonical capnogram waveform view is informative, but can be overwhelming when examining long records. By distilling capnogram information to just a single template, humans seem to perform better during classification and perceive a more accurate representation of the general waveform trend.

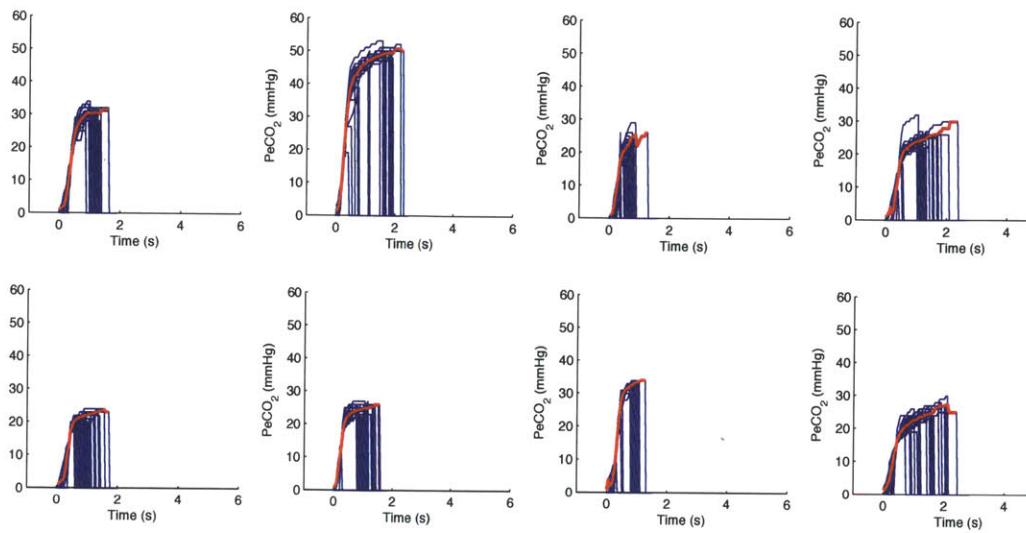


Figure 5.7. CHF templates.

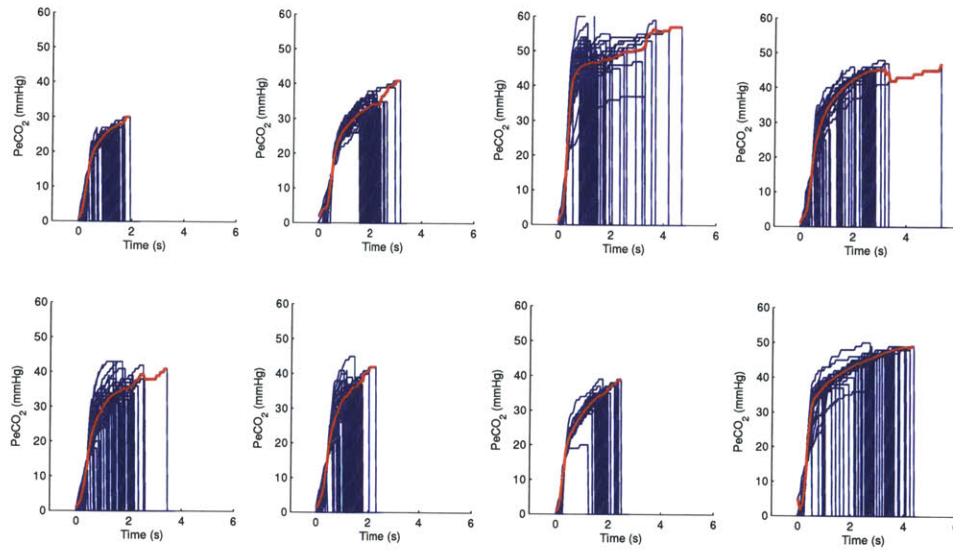


Figure 5.8. COPD templates.

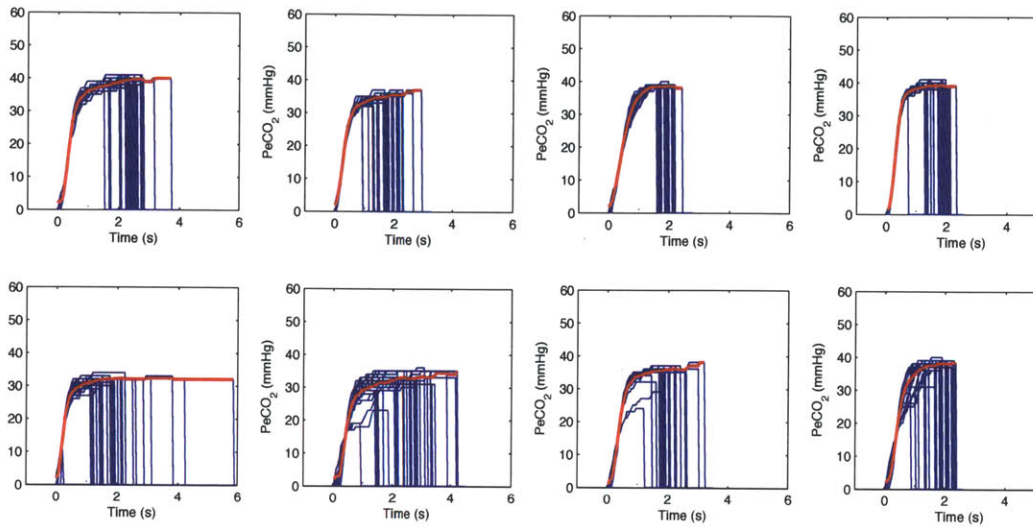


Figure 5.9. Normal templates.

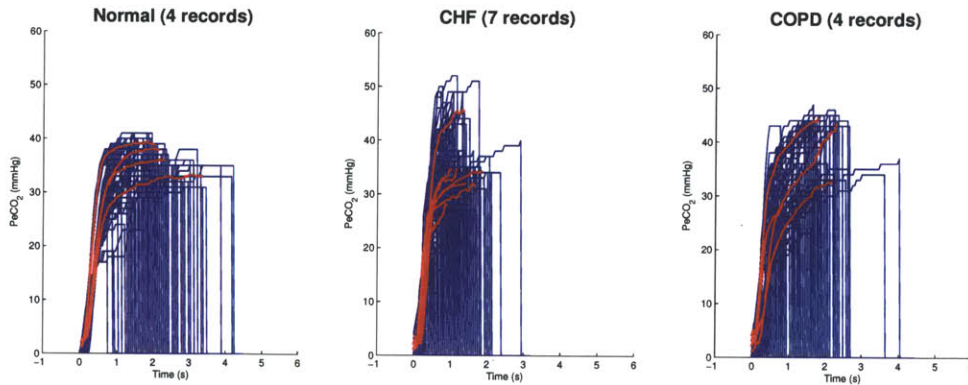
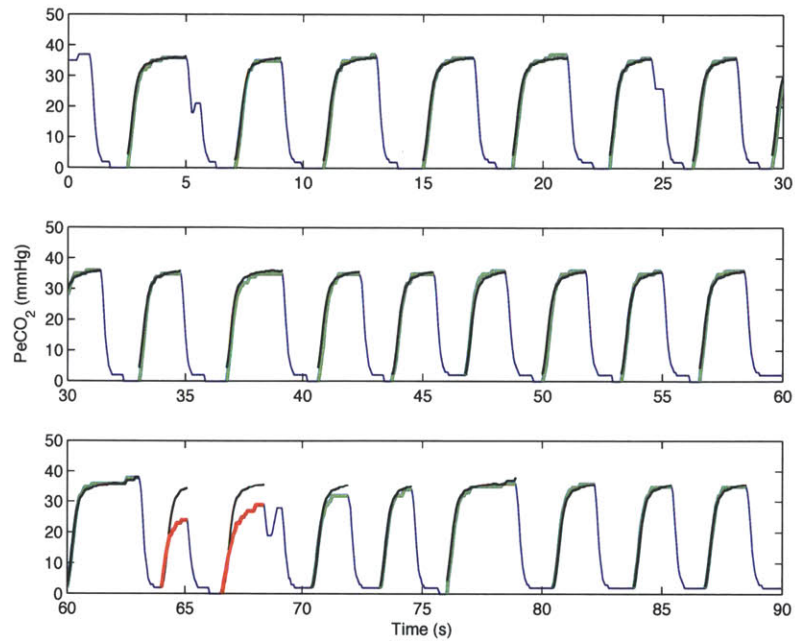


Figure 5.10. Template views of pathologic and normal capnograms.

### ■ 5.1.3 Discarding Outlier Exhalations

To further eliminate the effect of outlier exhalations, atypical breaths must occasionally be dropped from consideration. In order to complete the cropping process, the capnogram template is used as an exemplary breath. Exhalations that deviate significantly from the template are then excluded from analysis and classification.



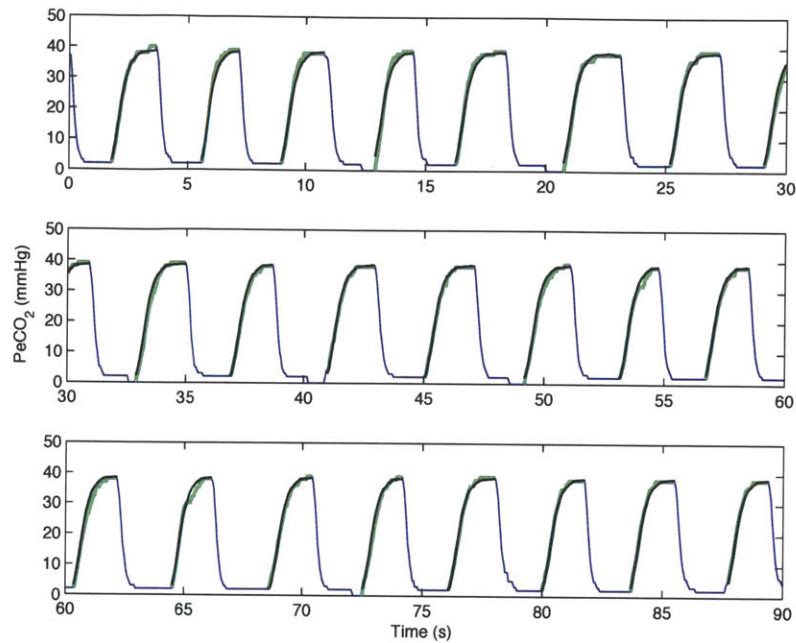
**Figure 5.11.** Normal record in which two breaths are excluded.

Figure 5.11 shows a Normal record in which only two outlier exhalations are detected. Displaying an even more consistent record, Figure 5.12 shows a Normal record in which every exhalation matched the template well and none were recommended for exclusion.

In practice, not very many exhalations are excluded from each record. Since pathologic exhalations tend to be more disordered and irregular, more outliers tend to occur in those records. Normal records, though, are generally more consistent and require less exhalations to be cropped.

Exhalation exclusion criteria include exhibiting a standard deviation from the template above a certain threshold. Additionally, if  $\text{ETCO}_2$  or exhalation duration deviates greatly from the record mean, the exhalation will be deemed an outlier.

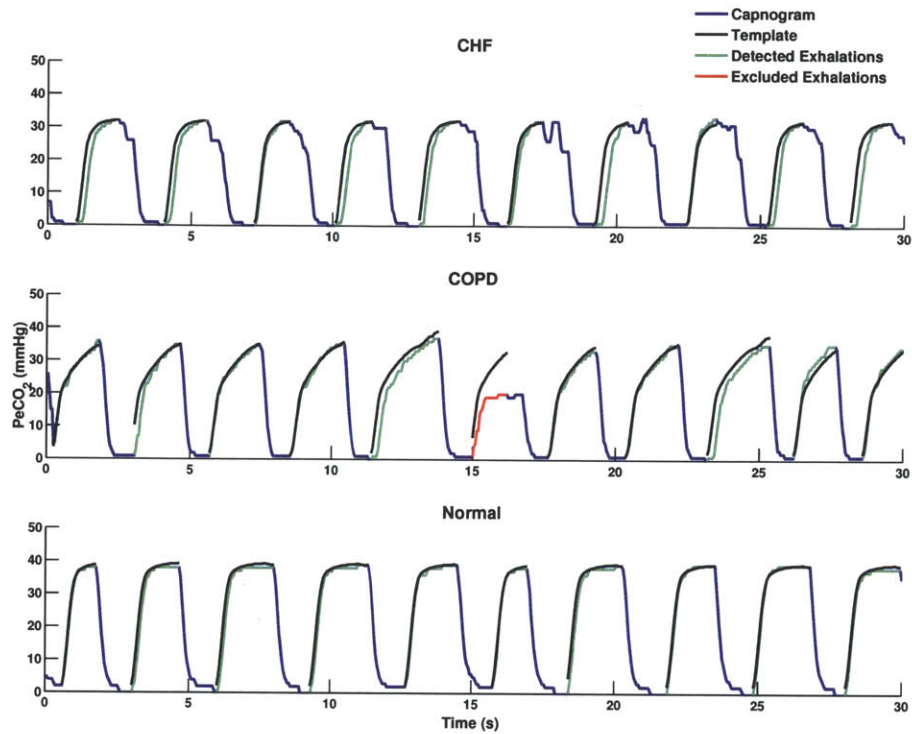




**Figure 5.12.** Normal record in which no breaths are excluded.

To better see the differences among detected exhalations, templates, and excluded exhalations in different classes, Figure 5.13 shows brief 30-second waveform strips from each of the patient classes considered. Exhalations that are kept in consideration seem to line up fairly well with the template exhalation, while those that are cropped deviate significantly.

Excluding atypical exhalations is an important part of preparing the waveform for feature extraction and further analysis. Outlier breaths can significantly impair the performance of classifiers if not cropped from consideration. The template formulation proves useful here in determining which breaths to crop and which exhalations to keep.



**Figure 5.13.** Waveform strips from normal and abnormal capnograms. Detected exhalations (*green*) are overlaid with the record's template exhalation (*black*) and outlier exhalations are displayed in *red*.

## ■ 5.2 Feature Extraction

Both analytic features and features that directly correspond to lung health are extracted from the capnogram. The first type of feature corresponds to parameters found during analytic curve fitting of the capnogram. Other features link directly to physiology and are extracted from the waveform by computing straightforward metrics such as duration, height, and slope of the capnogram.

### ■ 5.2.1 Curve Fitting Parameters

Our initial exploration with the respiratory waveform dataset involved examination of the upward exhalation slopes. Changes in lung dynamics will mostly affect the second and third capnogram phases, corresponding to the rise of  $[\text{CO}_2]$  during exhalation. By further examining expiratory rise behavior, much can be inferred about internal lung function. After exhalation rises around a semi-arbitrarily chosen point that they all cross, in this case  $[\text{CO}_2] = 15\text{mmHg}$ , the average rises of four different patients, one from each class, can be compared, as shown in Figure 5.14.

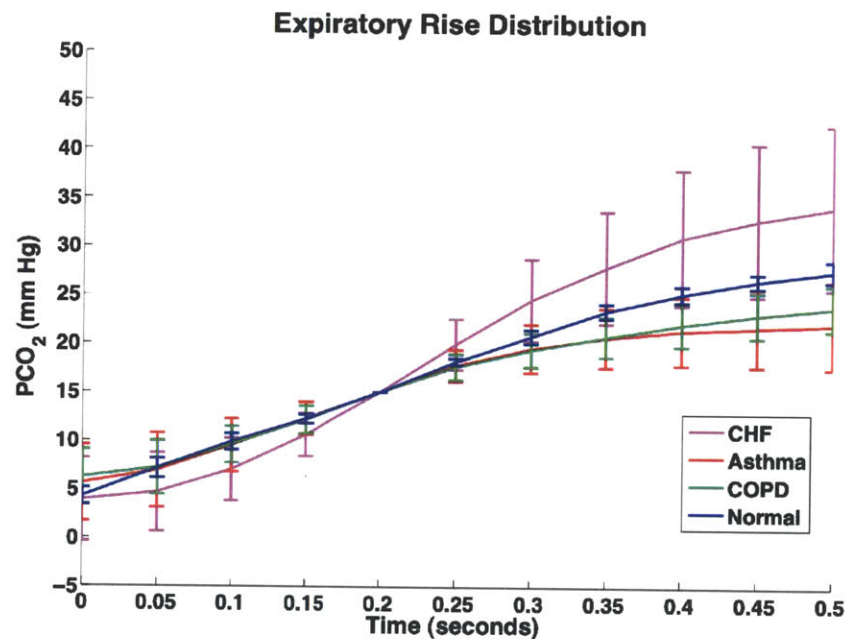


Figure 5.14. Analyzing the exhalation upslope.

Variance in the predicted exhalation time constants for different disease states is already evident in the expiratory rise diagram. Restrictive lung disease, expected to have a shorter time constant, rises very quickly during exhalation. The obstructive diseases, COPD and asthma, exhibit rise morphologies that lie on the other side of the

normal rise. These disease states were predicted to exhibit a longer time constant and indeed rise more slowly to  $ETCO_2$ . The expiratory rise plot is encouraging in that it shows good agreement with the predicted time constants.

First-order systems are often characterized by exponential transients. In an effort to identify a straightforward way of fitting the training data, an exponential rising to a final value of  $ETCO_2$  was formulated. The time constant  $\tau$  appears as the only unknown in this model:

$$[CO_2](t) = ETCO_2(1 - e^{-\frac{t}{\tau}}) \quad (5.1)$$

Rearranging Equation 5.1 and taking logarithms allows for linear least squares fitting to find a suitable  $\tau$  value. However, in performing this linear least squares fitting of expiratory rises, the lines obtained do not cross through the origin:

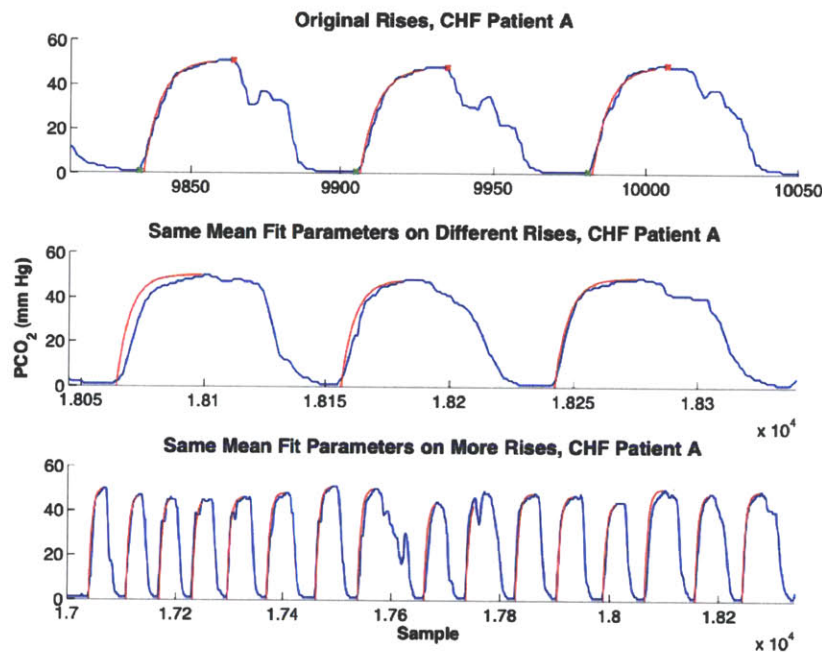
$$\ln\left(1 - \frac{[CO_2](t)}{ETCO_2}\right) = \frac{-t}{\tau} + b \quad (5.2)$$

Stated differently, there exists a non-zero y-intercept in the model. This y-intercept, termed  $b$ , shows up in the exponential model as an additional constant  $A$ . The value of  $A$  is unknown, meaning that the resulting exponential model contains two unknown values:

$$[CO_2](t) = ETCO_2(1 - Ae^{-\frac{t}{\tau}}) \quad (5.3)$$

Preliminary fits of this model performed very well even with the two unknowns. In order to perform the fits, expiratory rises were first detected and extracted from the time-series data. Rises were detected by monitoring changes in the sign of the slope

of the waveform. Once rises were extracted, each of three adjacent rises was fitted to the model. Thus, each rise was assigned its own  $A$  value and  $\tau$  value. These values were then averaged to achieve the mean rise parameters. Averaged parameter values constituting the mean rise were then used to plot the resulting exponential fits, which matched the data surprisingly well.



**Figure 5.15.** Fitting the 2-parameter model (*red*) to CHF data (*blue*).

In order to see if the results could be replicated on more data, the same mean fit parameters were used to identify exponentials for more expiratory rises from the same patient. This time, the exponentials still fit the data well. Testing on even more rises also revealed good agreement. During a more stringent test, the same exact mean  $A$  and  $\tau$  parameters as found in the original three expiratory rises evaluated were found to also fit the rises of different patients within the same disease class extremely well. These were surprising results that revealed the robustness of the fit and showed that

there may be physiological validity to the model.

The two-parameter model fit all the tested pathological disorders fairly well. New  $A$  and  $\tau$  parameters were identified for each disorder and were found to match up consistently with the data. However, the fits were not as good with data from normal subjects.

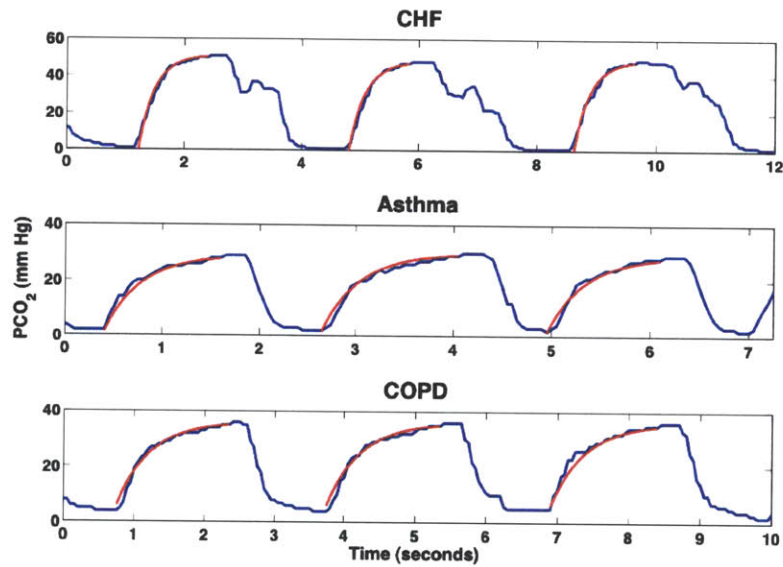


Figure 5.16. Pathologic data (*blue*) is fitted with the 2-parameter model (*red*).

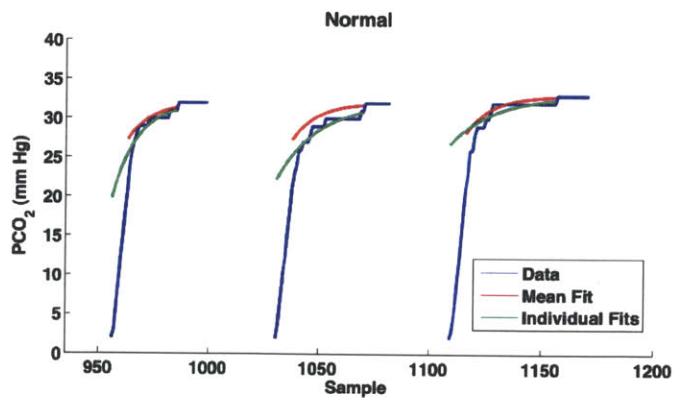


Figure 5.17. Attempting to fit the 2-parameter model to Normal data.

The above result prompted the search for a new model to used for normal subjects. There was some behavior in the two-parameter model indicating that it may not be the best choice. For instance, in the cases where exhalation corresponded to a long, continued effort, the terminal value of the expiration may not be accurately reflected by the  $ETCO_2$  detected. Perhaps such exponentials would continue to rise if inhalation were not to bring  $[CO_2]$  down quickly. The  $ETCO_2$  value then simply represents the  $[CO_2]$  value that occurs at some time  $t_E$ . The new model takes the form of a different-looking exponential, now with  $\tau$  as the only unknown:

$$[CO_2](t) = ETCO_2 \left( \frac{1 - e^{-\frac{t}{\tau}}}{1 - e^{-\frac{t_E}{\tau}}} \right) \quad (5.4)$$

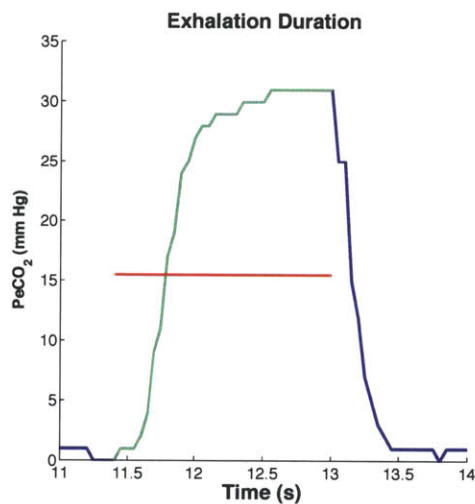
This new model with only one unknown can no longer be rearranged into a form that allows linear least squares fitting. Thus, a line search was conducted to find the value of  $\tau$  that would minimize the squared difference between the data and model values.

The 1-parameter model exhibited a uniformly higher mean squared error across all disease states. Normal records, however, always showed smaller error with the 1-parameter model. This result allows for quick separation of normal states from pathologic disease states.

### ■ 5.2.2 Physiological Features

Physiological features are characteristics of the waveform that directly correlate to physiological processes. Features such as respiratory rate, exhalation duration,  $ETCO_2$ , and end-exhalation slope are rooted in respiratory function. Besides being more understandable and easier to link to physiologic function, these features are also straightforward to extract from the waveform.

Exhalation duration is measured from the onset of exhalation, the time at which the capnogram slope becomes positive, until the last time at which  $[\text{CO}_2]$  is recorded on the exhalation (yielding  $\text{ETCO}_2$ ). The measurement is depicted in Figure 5.18. Patients with restrictive lung disease tend to exhibit shorter exhalation durations while those with obstructive disease have longer exhalations as they attempt to forcibly exhale air from the lungs against an obstruction.

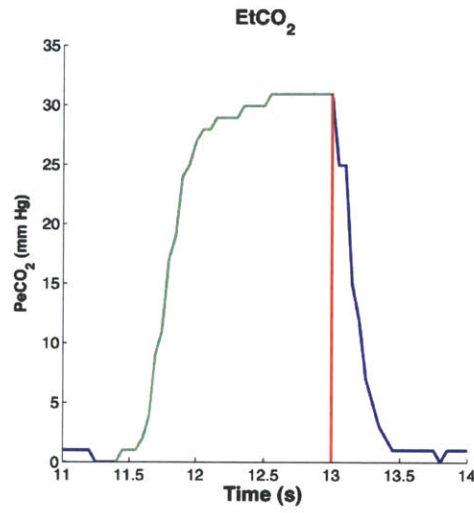


**Figure 5.18.** Measuring the first physiological feature, exhalation duration.

A second important physiological feature, shown in Figure 5.19, represents  $\text{ETCO}_2$ , the terminal value of the capnogram on exhalation. This value is captured just before the signal begins decreasing and is labeled as the  $\text{ETCO}_2$  value. This quantity is highly correlated with respiratory health. Obstructive disease patients are generally seen to exhibit high  $\text{ETCO}_2$  values, reflecting high arterial  $[\text{CO}_2]$ , as they do not satisfactorily expel carbon dioxide during exhalation. On the other hand, restrictive lung disease results in decreased  $\text{ETCO}_2$  levels since perfusion of carbon dioxide into the alveoli is impaired.

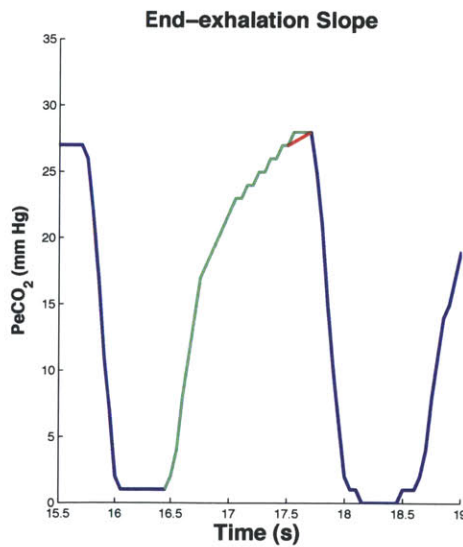
End-exhalation slope, seen in Figure 5.20, represents the third feature that we





**Figure 5.19.** The second physiological feature, ETCO<sub>2</sub>.

shall use, not so much because of its physiological significance as because of its prominence as a distinctive feature of COPD. In order to extract the end-exhalation slope, a linear regression is implemented over the last fifth of the capnogram exhalation. The



**Figure 5.20.** End-exhalation slope, the third feature used in our analysis.

slope of this tangent is then taken as the end-exhalation slope. Because normal breathing results in a relatively flat alveolar plateau and obstructive disease yields a more rounded shape, the end-exhalation slope feature is especially useful in distinguishing obstructive from normal exhalations.

In summary, several steps are necessary in the pre-processing of capnographic data. First, properly detecting exhalations is essential to accurately analyzing the waveforms. Formulating a record template is helpful in both viewing the record in a composite manner and in developing a normal standard for later exclusion of outlier exhalations. Cropping outlier breaths is important to clean the capnogram data before feature extraction. By considering a few important features, including those corresponding to curve analysis and those connected directly to lung physiology, we expect the classification process to prove robust.

# Classification Results

**O**UR classification process involves several steps. First, we examine the dataset and what types of patient records it contains. We then describe the partitioning of the dataset into training and test sets, each comprising a small number of successive breaths from different patients. Several quadratic discriminators are then trained implementing the three physiologic features highlighted in Chapter 5, using different training set partitions. Two levels of voting are now employed to classify each record in the test set. For each discriminator, the individual exhalation classifications vote on the preliminary classification of their corresponding test record. The various classifier verdicts then vote on the final record classification. Subsequent analyses show classification sensitivity and specificity while varying the number of exhalation votes required for classification on each test record. We also examine the effect of varying the size of the training set.

### ■ 6.1 Dataset

The dataset comprises records from 128 patients having CHF, COPD, or diagnosed as Normal. Records originate from both Albert Einstein Medical Center (Philadelphia, PA), and Brigham and Women's Hospital (Boston, MA). Severity metrics are only

---

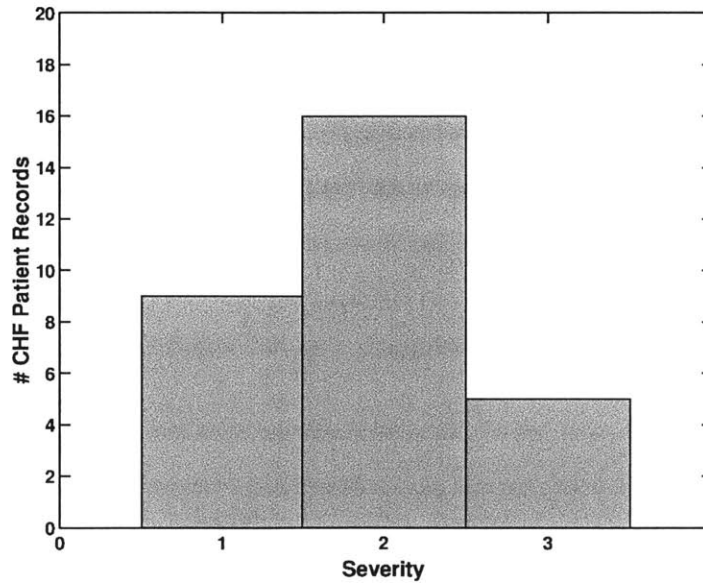
Class	# Patient Records
CHF	31
COPD	33
Normal	64

---

**Table 6.1.** Dataset distribution.

known for the CHF patients from Albert Einstein. Table 6.1 displays the number of patients in each category.

Although patients are lumped into the three general categories of CHF, COPD, or Normal, not all patients in the pathologic classes exhibit the same severity level. Indeed, in the case of CHF, for which most records have a physician-assigned severity score, the majority of patients exhibit moderate severity. Figure 6.1 summarizes the CHF severity spectrum in histogram form. There are a total of 30 patient records



**Figure 6.1.** Severity levels of CHF patients in the dataset. Most patients lie in the moderate category.

depicted in Figure 6.1 since one CHF record did not have an associated severity score.

Selecting the proportions of the training set and test set is another consideration. Roughly a 70%/30% training/test partition is used. Table 6.2 provides the record numbers in each partition. In this experiment, our classification algorithm takes into account 35 exhalations from each patient record. Short records containing less than 35 exhalations are removed from consideration during training and test. For this reason, not all patient records in the dataset are members of the training and test sets summarized in Table 6.2.

	Training Set Size (Records)	Test Set Size (Records)
Normal vs. {CHF, COPD}	67	30
CHF vs. COPD	44	20

**Table 6.2.** Dataset partition when considering 35 exhalations from each patient record.

## ■ 6.2 Voting Schema

To help improve the performance of the classifiers produced by quadratic discriminant analysis, two levels of voting are performed. This is similar in many ways to boosting, the method of creating combined classifiers from several base classifiers by means of voting [27].

Specifically, the two levels of voting employed in this process are the following:

- The classifications of the individual exhalations by a given classifier are taken as individual votes on the eventual classification of the record. The proportion of votes needed for a classification in one direction or the other may be varied.

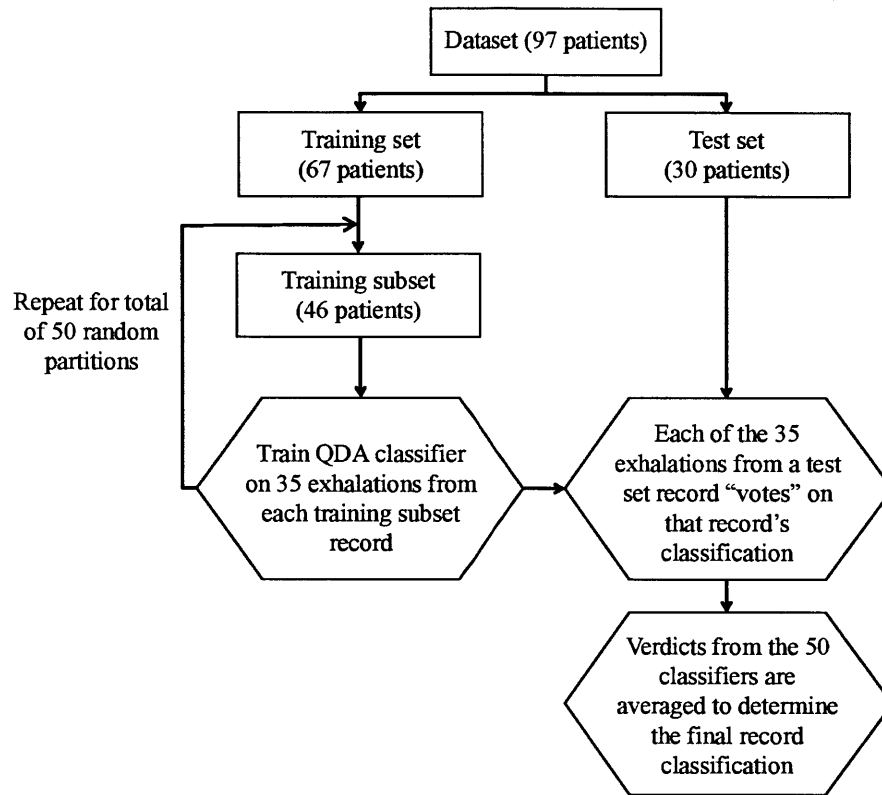


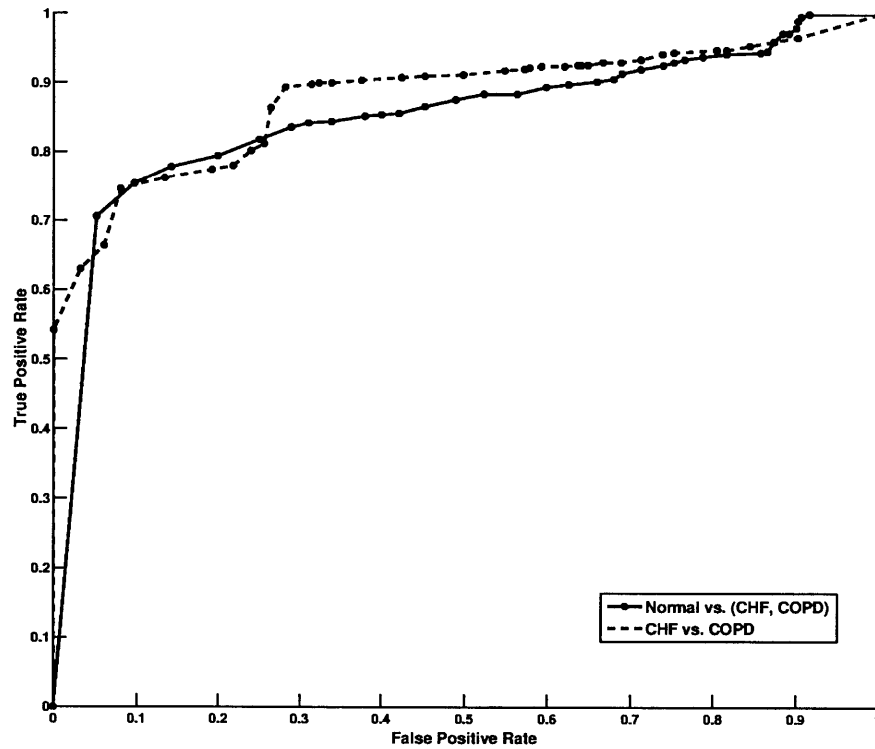
Figure 6.2. Voting schema, illustrated for the case of Normal vs. {CHF, COPD} discrimination.

- Each classifier votes on the final label of a record.

Figure 6.2 summarizes the voting process to determine final record classifications. Base classifiers are trained on different partitions of the training set. Each partition is a training subset comprising roughly 70% of the patient records in the full training set.

## ■ 6.3 Results

After proceeding through two stages of classification, {Normal vs. {CHF, COPD}} and {CHF vs. COPD}, results are reported for both. The tasks are certainly not of equal difficulty.



**Figure 6.3.** ROC Curves obtained testing on 35 exhalations from each record. The threshold is varied from 0 to 35 exhalations to reach a record verdict.

The ROC curves in Figure 6.3 display the true positive rate vs. the false positive rate over a variety of thresholds. In Normal vs. {CHF, COPD} classification, classifying a record as Normal represents a positive detection. During CHF vs. COPD classification, a CHF classification is considered a positive detection. The threshold

	<i>AUC</i>
Normal vs. {CHF, COPD}	0.85
CHF vs. COPD	0.88

**Table 6.3.** Area under the curves.

varied is the number of exhalations (out of the 35 total) required to reach a verdict on the classification of a given record, for a given classifier. When this threshold is low, sensitivity is increased as almost all positive instances of records are detected. As the threshold is raised, less positive instances are detected, but the specificity is increased.

Area under the curve (AUC) is a typical way to evaluate the ROC curve's quality [3]. A perfect classifier has an AUC of 1, while a coin-flip random classifier has an AUC of 0.5. Table 6.3 shows the area under both displayed ROC curves. The {CHF vs. COPD} classifier appears to perform slightly better than {Normal vs. {CHF, COPD}}.

Of course, a single threshold value must be exported as the final classifier. A threshold of 15 exhalations is used in the subsequent test investigating detection sensitivity as a function of the number of exhalations considered from each record.

In plotting the ROC curves of Figure 6.3, 35 exhalations from each patient record were used. However, decreasing the number of breaths necessary would improve capnography's effectiveness as a short-term monitor. For patients affected with CHF, this sort of short-term monitoring is frequently not available [6]. Toward this end, we evaluated the sensitivity as a function of the number of exhalations considered during both training and test.

Figure 6.4 shows that as the number of exhalations increases, so does the sensitivity. In order to achieve good performance, it appears that more than 25 exhalations should be employed. Because the threshold for a positive instance is 15 breaths, con-



sidering a number of exhalations less than 15 does not make sense. As the number of exhalations considered from each record varies, the training and test set sizes also shifts slightly. When 35 exhalations are considered from each record, the dataset is partitioned as in Table 6.2. Moving to 40 exhalations, only dataset records containing at least 40 exhalations were placed in the training and test sets. Similarly, when 20 exhalations from each record were considered, all dataset records containing 20 exhalations or more were used during training and test. This variation in the training and test set sizes may also alter classification performance.

A related question is how many records must be present in the training set to

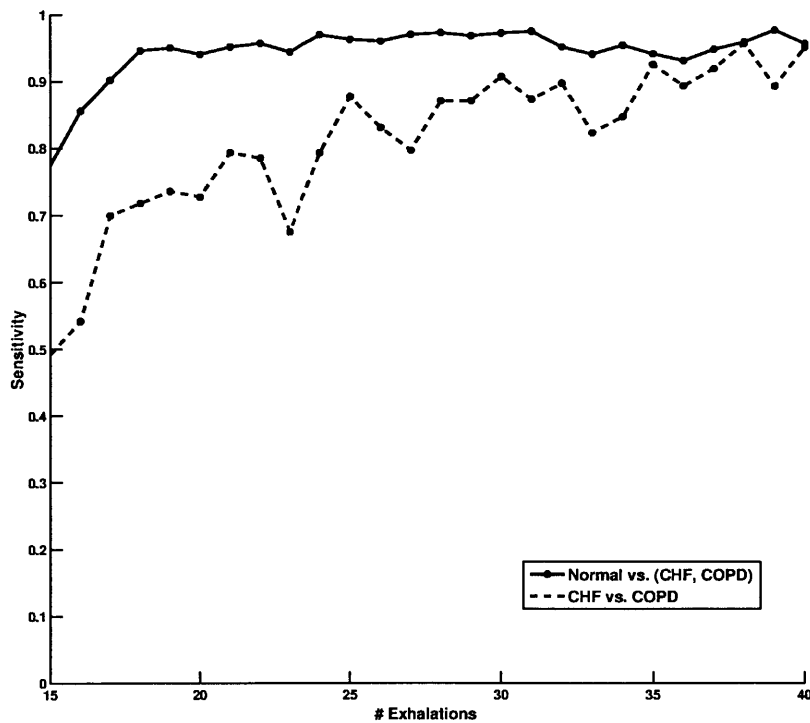
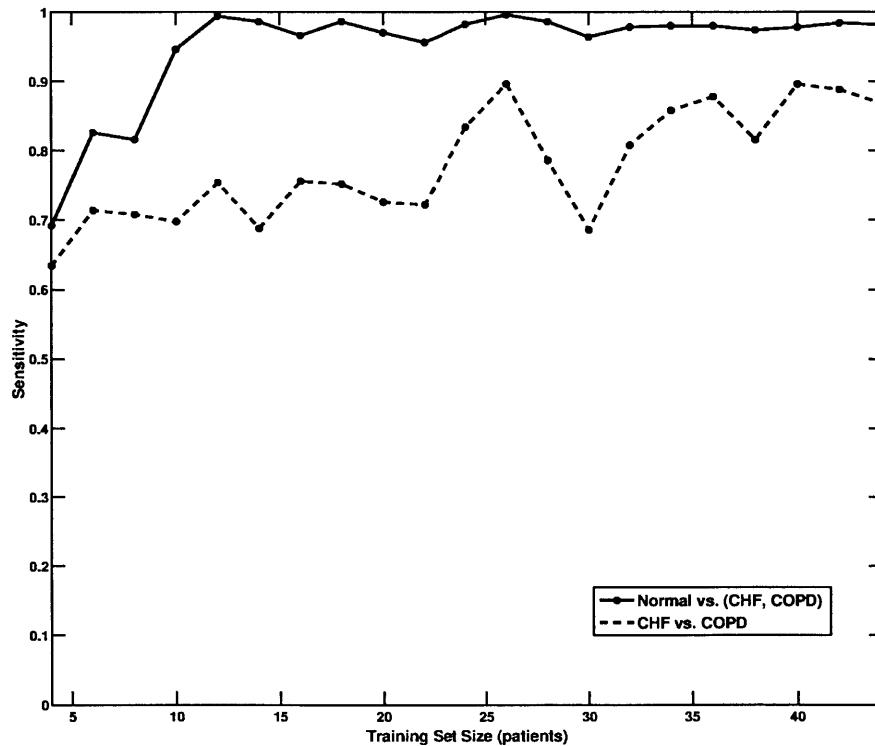


Figure 6.4. Sensitivity attained when training on a different number of exhalations per record.

yield good performance, i.e. how much data is needed to produce good classification. To investigate this, we determine the sensitivity as a function of training set size. The number of exhalations considered is set to 35 per record and the threshold for a positive verdict to 15 exhalations. For the Normal vs. {CHF, COPD} classification, the test set size is kept at 30 records, while a test set size of 20 records is used for CHF vs. COPD classification. Results of this examination are summarized in Figure 6.5. The sensitivity appears to increase with the size of the training set. {Normal vs. {CHF, COPD}} classification appears to be more insensitive to changes in the training set

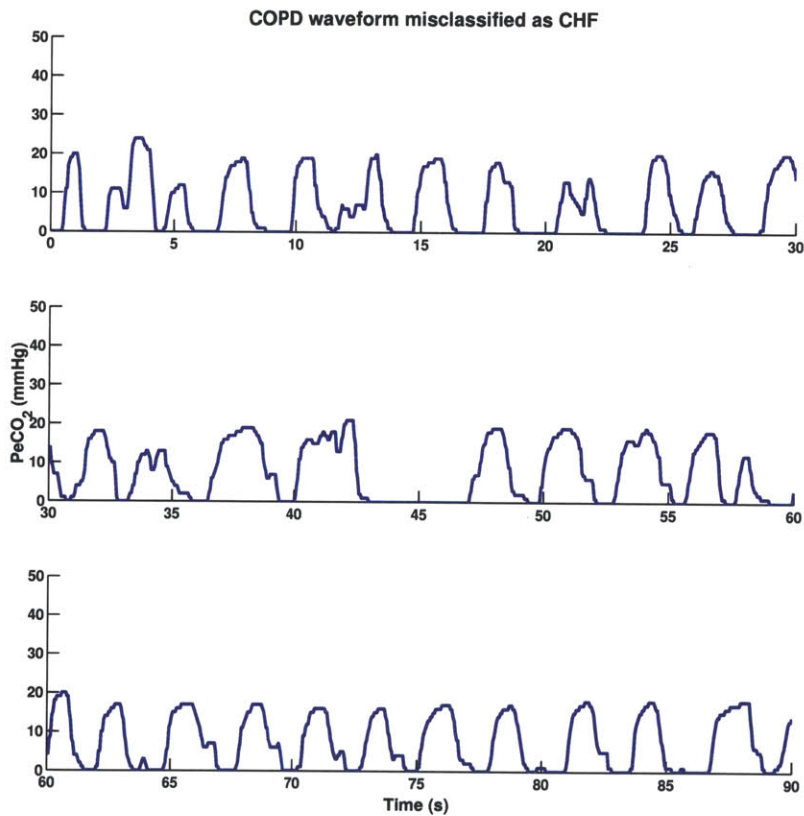


**Figure 6.5.** Sensitivity attained when training on different record sizes. Using a threshold of 15 breaths positive.

size.

## ■ 6.4 Misclassified Records

Although the outlined classification process performs well, it is not perfect, and some patient records are misclassified. We would like to determine whether we are at least correctly classifying the most severe of cases. If our classification is robust, the severest of CHF records with a score of 3 should be properly classified. We expect misclassified records to exhibit a lower severity and lie closer to the decision boundary. In order to

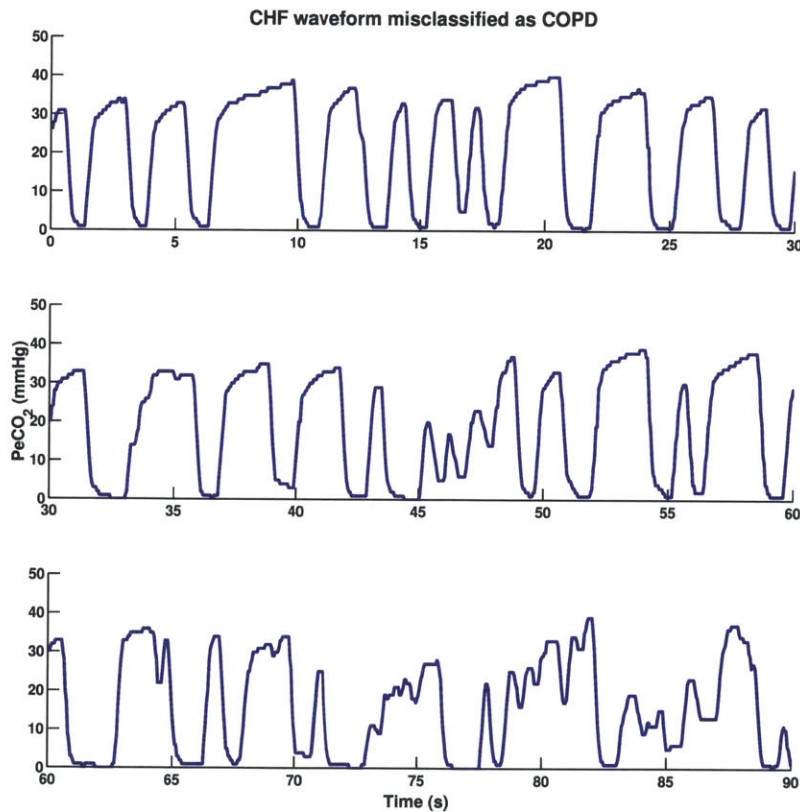


**Figure 6.6.** Misclassified COPD waveform.

isolate the worst offenders, or those records that continue to be misclassified even as the threshold for a positive instance is increased, we increase the positive verdict threshold to 30 exhalations and observe which records are still misclassified.

When a threshold of 30 exhalations is used for Normal vs. {CHF, COPD} classification, the one CHF waveform misclassified as Normal is of moderate severity. It is not severe. This is encouraging in that none of the severest CHF cases are misclassified.

Of the 2 COPD waveforms misclassified as CHF during CHF vs. COPD classification with a threshold of 30 exhalations to be classified as CHF, one is shown in Figure



**Figure 6.7.** Misclassified CHF waveform.

6.6. Although no physician-assigned severity score is available for the COPD records, it can be seen that the waveform does not prominently display the characteristic shark's fin COPD morphology or the higher  $\text{ETCO}_2$  values. In fact, the misclassified capnogram's features, such as low  $\text{ETCO}_2$ , short exhalation duration, and flat end-exhalation slope, much more closely resemble a CHF waveform. This COPD waveform is quite likely to be misclassified by a clinician as well.

The only CHF waveform misclassified as COPD in CHF vs. COPD classification using the same threshold has a severity level of 1 (mild). Its waveform is plotted in Figure 6.7. As can be seen, it exhibits the morphology and features more typical of a COPD waveform.

In summary, classification via discriminant analysis was performed on a dataset of patients exhibiting either restrictive, obstructive, or healthy capnograms. Patient records were classified by disease, and classifiers were trained on roughly 70% of the dataset. A number of different classifiers were trained on individual exhalations and then allowed to vote on the final outcome of the record. This method of classification and boosting performs reasonably well and further confirms capnography's potential as a diagnostic tool.



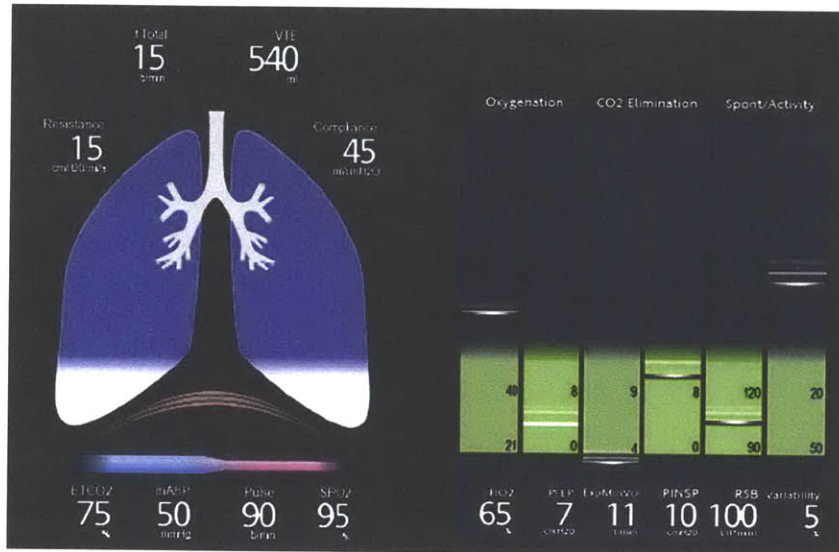
# Conclusions

**M**ATHEMATICAL capnogram analysis exhibits potential for distinguishing lung disease states based on objective measures. Toward the goal of separating obstructive from restrictive lung disease records, training datasets have been examined to assess the general behavior of the capnographic time series. Preliminary tests indicate reasonable performance of the models already developed.

In analyzing a dataset containing capnograms from CHF, COPD, and Normal patients, a new method of quantitatively characterizing the capnogram is proposed in the hopes of improving respiratory diagnostics and better quantifying the change in lung parameters during disease conditions. Voting methods are implemented on top of discriminant analysis in order to boost classification performance.

The small size of the dataset hampers classification performance and feature analysis. In the future, more pathologic data will be collected and will hopefully lead to better classification performance. Successful discriminant analysis classification schemes employ many more records and many more features. In extending to other disease states, semi-supervised techniques may prove useful in learning the patterns of previously untrained classes such as asthma, cystic fibrosis, or pneumonia.

Another future goal would be to describe parameters of the lung, such as compli-



**Figure 7.1.** Possible future descriptive capnography interface, including information about respiratory resistance, compliance, oxygenation, and CO<sub>2</sub> content [31].

ance and airway resistance, in a very specific way via capnography. These quantitative descriptors would lend an even more complete assessment of lung state and would allow clinicians to view more information than a single disease classification with a confidence interval. An example of a dashboard view of lung parameters that could be developed is shown in Figure 7.1. Reporting such information would necessarily involve more modeling of the lung to understand how respiratory parameter values are modified by disease.

Future modeling efforts could thus focus on learning new classification techniques based on statistical prediction methods and on how physical lung parameters change during disease. Acquiring more data will undoubtedly allow for better capnogram characterization and classification. Progress with the various records already tested has been encouraging. Capnographic monitoring harbors the potential to become a useful diagnostic tool in the assessment of lung disease.



---

---

## Bibliography

- [1] T. Ahrens and C. Sona. Capnography application in acute and critical care. *AACN Clinical Issues*, 14(2):123 – 132, 2003.
- [2] C. Bishop. *Pattern Recognition And Machine Learning*. Information Science and Statistics. Springer, 2006.
- [3] A. P. Bradley. The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7):1145 – 1159, 1997.
- [4] L. H. Brown, J. E. Gough, and R. H. Seim. Can quantitative capnometry differentiate between cardiac and obstructive causes of respiratory distress? *Chest*, 113(2):323 – 326, 1998.
- [5] J. Cheng and R. Greiner. Comparing bayesian network classifiers. pages 101 – 108, 1999.
- [6] V. Cheng, R. Kazanagra, A. Garcia, L. Lenert, P. Krishnaswamy, N. Gardetto, P. Clopton, and A. Maisel. A rapid bedside test for b-type peptide predicts treatment outcomes in patients admitted for decompensated heart failure: a pilot study. *Journal of the American College of Cardiology*, 37(2):386 – 391, 2001.
- [7] A. B. Chilton and R. W. Stacy. A mathematical analysis of carbon dioxide respiration in man. *Developmental neurobiology*, 14(1):1 – 18, 1952.

- 
- [8] P. Clyburn and M. Rosen. Accidental oesophageal intubation. *British Journal of Anaesthesia*, 73(1):55 – 63, 1994.
- [9] L. Costanzo. *Physiology*. Costanzo Physiology. Saunders/Elsevier, 2010.
- [10] J. Crassidis and J. Junkins. *Optimal Estimation of Dynamic Systems*. Chapman & Hall/CRC Applied Mathematics & Nonlinear Science. Taylor & Francis, 2011.
- [11] M. Dash and H. Liu. Feature selection for classification. *Intelligent Data Analysis*, 1(14):131 – 156, 1997.
- [12] J. D’Mello and M. Butani. Capnography. *Indian Journal of Anesthesia*, 46(4):269 – 278, 2002.
- [13] A. C. Dornhorst, S. J. G. Semple, and I. M. Young. Automatic fractional analysis of expired air as a clinical test. *Lancet*, 1(6756):370 – 372, 1953.
- [14] I. Guyon, J. Makhoul, R. Schwartz, and V. Vapnik. What size test set gives good error rate estimates? In *IEEE Trans PAMI*, pages 52 – 64, 1996.
- [15] B. I. Hoffbrand. The expiratory capnogram: a measure of ventilation-perfusion inequalities. *Thorax*, 21(6):518 – 523, 1966.
- [16] Z. Kalenda. *Mastering infra-red capnography*. Kerckebosch BV, 1989.
- [17] D. Koller and M. Sahami. Toward optimal feature selection. *13th International Conference on Machine Learning*, pages 284 – 292, 1995.
- [18] B. Krauss, A. Deykin, A. Lam, J. J. Ryoo, D. R. Hampton, P. W. Schmitt, and J. L. Falk. Capnogram shape in obstructive lung disease. *Anesthesia & Analgesia*, 100(3):884 – 888, 2005.
- [19] B. Krauss and D. R. Hess. Capnography for procedural sedation and analgesia in the emergency department. *Annals of Emergency Medicine*, 50(2):172 – 181, 2007.

- 
- [20] T. Li, C. Zhang, and M. Ogihara. A comparative study of feature selection and multiclass classification methods for tissue classification based on gene expression. *Bioinformatics*, 20(15):2429 – 2437, 2004.
- [21] I. Maglogiannis. *Emerging Artificial Intelligence Applications in Computer Engineering: Real World Ai Systems With Applications in Ehealth, Hci, Information Retrieval and Pervasive Technologies*. Frontiers in Artificial Intelligence and Applications. IOS Press, 2007.
- [22] A. S. Maisel, P. Krishnaswamy, R. M. Nowak, J. McCord, J. E. Hollander, P. Duc, T. Omland, A. B. Storrow, W. T. Abraham, A. H. Wu, P. Clopton, P. G. Steg, A. Westheim, C. W. Knudsen, A. Perez, R. Kazanegra, H. C. Herrmann, and P. A. McCullough. Rapid measurement of b-type natriuretic peptide in the emergency diagnosis of heart failure. *New England Journal of Medicine*, 347(3):161 – 167, 2002.
- [23] D. M. Mannino, E. S. Ford, and S. C. Redd. Obstructive and restrictive lung disease and functional limitation: data from the third national health and nutrition examination. *Journal of Internal Medicine*, 254(6):540 – 547, 2003.
- [24] M. Marshall. Capnography in dogs. *Compendium*, 26(10):761 – 778, 2004.
- [25] M. Meyer, M. Mohr, H. Schulz, and J. Piiper. Sloping alveolar plateaus of co<sub>2</sub>, o<sub>2</sub> and intravenously infused c<sub>2</sub>h<sub>2</sub> and chcl<sub>2</sub> in the dog. *Respiration Physiology*, 81(2):137 – 151, 1990.
- [26] E. Osuna, R. Freund, and F. Girosi. An improved training algorithm for support vector machines. In *Neural Networks for Signal Processing [1997] VII. Proceedings of the 1997 IEEE Workshop*, pages 276 – 285, 1997.
- [27] R. E. Schapire and Y. Freund. Boosting the margin: a new explanation for the

- 
- effectiveness of voting methods. *The Annals of Statistics*, 26:322 – 330, 1998.
- [28] U. Smidt. Emphysema as possible explanation for the alteration of expiratory po<sub>2</sub> and pco<sub>2</sub> curves. *Bull Eur Physiopathol Respir*, 12(5):605 – 624.
- [29] P. Tan, M. Steinbach, and V. Kumar. *Introduction to Data Mining*. Pearson Addison Wesley, 2006.
- [30] K. R. Ward and D. M. Yealy. End-tidal carbon dioxide monitoring in emergency medicine, part 1: Basic principles. *Academic Emergency Medicine*, 5(6):628 – 636, 1998.
- [31] M. Wysocki and J. X. Brunner. Closed-loop ventilation: An emerging standard of care? *Critical Care Clinics*, 23(2):223 – 240, 2007.
- [32] W. S. Yamamoto. Mathematical analysis of the time course of alveolar carbon dioxide. 15:215 – 219, 1960.
- [33] J. Ye, R. Janardan, and Q. Li. Two-dimensional linear discriminant analysis. *The Eighteenth Annual Conference on Neural Information Processing Systems*, pages 1569 – 1576, 2004.
- [34] B. You, R. Peslin, C. Duvivier, V. Vu, and J. Grilliat. Expiratory capnography in asthma: evaluation of various shape indices. *European Respiratory Journal*, 7(2):318 – 323, 1994.
- [35] M. Zou and S. D. Conzen. A new dynamic bayesian network (dbn) approach for identifying gene regulatory networks from time course microarray data. *Bioinformatics*, 21(1):71 – 79, 2005.