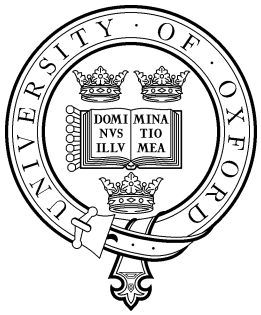
Methods for Determining Lung Function from Tracer Gas Concentrations



Chris Harrison

Wadham College

University of Oxford

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

Hilary Term 2016

To Dr Margaret Haire MD FRCPath (1921-2013), my grandmother.

I hope this partly makes up for me not being a medic!

Acknowledgements

I would like to thank my supervisor, Stephen Payne, for his excellent swift advice and encouragement during my research, which has ultimately led to the submission of this thesis.

Thanks are also due to Andrew Farmery, Phi-Anh Phan and Clive Hahn, my clinical collaborators from the Nuffield Division of Anaesthetics, Nuffield Department of Clinical Neurosciences, who have provided valuable advice and a clinical perspective. I am also grateful to Cathy Zhang and Daniel Geer, medical students, who collected the homogeneous bench lung and human data.

The staff of the CDT in Healthcare Innovation are thanked for providing an informative and stimulating first year course and for their ongoing support. The RCUK Digital Economy program is thanked for funding the CDT, which provided my studentship (EP/G036861/1).

The work would not have been possible without the help of the IBME IT staff, especially Vernon Bailey and David Simpson who got the required software running and kept it running on the ageing departmental cluster. Use of the University of Oxford Advanced Research Computing (ARC) facility is also acknowledged.

The pleasant working environment of the PUMMA research group has been appreciated, which is down to everyone who has been part of it over the years, including Claire Lucas, Amit Mehndiratta, Yee Kai Tee, Sheldon Hall, Flora Kennedy McConnell and Wahbi El-Bouri.

Finally I would like to thank the most special people. Louise in general and for putting up with me in the last few months. Linda, my sister, for some statistical advice and (heated) debate. Angie and Kevin, my parents, for general support over the years, but especially during my biggest academic challenge yet, my GCSEs, without which I would not be submitting this thesis.

Abstract

Methods for Determining Lung Function from Tracer Gas Concentrations

Chris Harrison
Wadham College

Doctor of Philosophy
Hilary Term 2016

The ability to routinely determine Functional Residual Capacity (FRC) or the ventilation-volume homogeneity of ventilated patients has been a long held goal. Such measures have the potential to greatly improve the treatment of patients by preventing ventilator related lung injuries. Many methods have been proposed, but none have seen routine clinical use. This thesis develops methods of determining such information from tracer gas concentrations, modelling the concentration of the entire expired breath in an attempt to extract more information from the data.

A method of using such models to produce a distribution of alveolar volumes, which broadens with an increase in the inhomogeneity of simulated data, was developed. When compared to the Homogeneity Index [Whiteley et al., Respir. Physiol., 124(1):65-83.], the method determines the homogeneity more robustly, particularly when noise is present in the flow signal. In real data from a water-displacement bench lung it also produced tentatively better results. However it did not perform as expected for 25% of the data sets.

Investigations into determining FRC, highlighted the need to include mixing within the dead space, which is a move away from traditional dead space models. In human data a dead space with two mixing compartments provided the best FRC results, reducing the mean limits of agreement with plethysmography by 32% when compared to the other practical methods investigated. Compared to data from the literature, this method was no worse than Helium Dilution (the bias and limits of agreement were within the 95% confidence interval) and the limits of agreement were significantly better than for the LUFU device. Thus, if the results are replicated in clinical practice, the method is likely to be robust enough to positively influence patient care. Attempts to correct for time offsets between the flow and concentrations signals did not improve results for humans, despite showing some improvements for a water-displacement bench lung.

Contents

1 Introduction **1**
- 1.1 Clinical Motivation . . . 1
- 1.2 Technical Challenges . . . 3
- 1.3 Aims of this Thesis . . . 4
- 1.4 Outline of this Thesis . . . 5

2 Literature Review **7**
- 2.1 Introduction . . . 7
- 2.2 Lung Physiology . . . 7
  - 2.2.1 Lung Volumes . . . 8
  - 2.2.2 Lung Inhomogeneity . . . 10
- 2.3 Ventilation Models . . . 11
  - 2.3.1 Bohr Equation . . . 11
  - 2.3.2 Compartment Models . . . 12
  - 2.3.3 Trumpet Model . . . 17
  - 2.3.4 Complex Anatomical Models . . . 18
- 2.4 Estimating Lung Parameters using Tracer Gases . . . 19
  - 2.4.1 Wash-out Methods . . . 19
  - 2.4.2 Helium Dilution . . . 27
  - 2.4.3 Forcing Sinusoids . . . 28
  - 2.4.4 Breath-by-Breath Analysis . . . 34
  - 2.4.5 Multiple Inert Gas Elimination Technique (MIGET) . . . 35
- 2.5 Estimating Lung Parameters by Other Means . . . 37
  - 2.5.1 Plethysmography . . . 38
  - 2.5.2 Determining Lung Volume from Images . . . 40
  - 2.5.3 Determining Ventilation Information from Images . . . 41
- 2.6 Signal Alignment and Sensor Response Time . . . 42
- 2.7 Estimating Model Parameters . . . 46
  - 2.7.1 Non-Linear Least Squares Optimisation . . . 46
  - 2.7.2 Approximate Bayesian Computation (ABC) . . . 47
- 2.8 Conclusion . . . 51

3 Methods **52**
3.1 Introduction . . . 52
3.2 Data . . . 52
3.2.1 Sinusoidal Delivery System . . . 54
3.2.2 Simulated Data . . . 55
3.2.3 Water-Displacement Bench Lung Data . . . 57
3.2.4 Human Lung Data . . . 58
3.2.5 Preprocessing . . . 58
3.3 Computational Models . . . 59
3.3.1 Homogeneous 'Balloon on a Straw' Model . . . 59
3.3.2 Modelling Mixing within the Dead Space . . . 61
3.3.3 Inhomogeneous Alveolar Compartment . . . 63
3.3.4 Soluble Gases . . . 67
3.3.5 Solving the Model . . . 68
3.4 Parameter Estimation Techniques . . . 70
3.4.1 Breath-by-Breath Method . . . 70
3.4.2 Fitting the Entire Signal . . . 73
3.5 Conclusion . . . 76

4 Simulated Data **77**
4.1 Introduction . . . 77
4.2 Simulated Data . . . 77
4.2.1 Experimental Error . . . 78
4.3 Log-Normal Distribution of Ventilation-Volume Ratios . . . 80
4.3.1 Homogeneity Index . . . 81
4.3.2 Inter-breath Variation . . . 90
4.3.3 Directly Recovering Volume-Ventilation Distribution . . . 96
4.4 Multi-Compartment Lungs . . . 102
4.4.1 Homogeneity Index . . . 103
4.4.2 Inter-breath Variation . . . 110
4.4.3 Directly Recovering the Compartment Volumes . . . 113
4.5 Independently Recovering the Dead Space Volume . . . 118
4.5.1 Results . . . 118
4.5.2 Discussion . . . 119
4.6 Alignment of the Signals . . . 121
4.6.1 Simulated Data . . . 121
4.6.2 Effect of an Offset . . . 121
4.6.3 Recovering the Offset . . . 126
4.7 Soluble Gas Data . . . 133
4.7.1 Simulated Data . . . 134
4.7.2 Fixed Venous Return . . . 135
4.7.3 Assuming Body Compartments of a Standard Man . . . 138
4.7.4 Discussion . . . 140
4.8 Summary . . . 142

4.8.1 Limitation of the Simulated Data . . . 142
4.8.2 Lung Homogeneity . . . 144
4.8.3 Other Findings . . . 146
4.9 Key Findings . . . 148
4.10 Conclusion . . . 149

5 Water-Displacement Bench Lung Data 151
5.1 Introduction . . . 151
5.2 Homogeneous Bench Lung . . . 152
5.2.1 Data . . . 152
5.2.2 Breath-by-Breath Method . . . 153
5.2.3 Models for Mixing within the Dead Space . . . 156
5.2.4 Signal Offset Correction . . . 174
5.2.5 Recovering the Dead Space Volume Only . . . 178
5.3 Multi-compartment Lung Data . . . 183
5.3.1 Data . . . 183
5.3.2 Recovering the Dead Space Volume . . . 184
5.3.3 Determining the Homogeneity of the Alveoli . . . 185
5.3.4 Discussion . . . 190
5.4 Summary . . . 192
5.5 Key Findings . . . 194
5.6 Conclusion . . . 196

6 Human Data 197
6.1 Introduction . . . 197
6.2 Data . . . 198
6.3 Statistical Analysis . . . 199
6.4 Method Comparison . . . 200
6.4.1 Breath-by-Breath Method . . . 200
6.4.2 Fixed Venous Return . . . 203
6.4.3 Standard Man . . . 214
6.4.4 Summary . . . 222
6.5 Comparison with Published Results . . . 227
6.5.1 Repeatability . . . 227
6.5.2 Comparisons with Plethysmograph . . . 229
6.5.3 Discussion . . . 230
6.6 Key Findings . . . 235
6.7 Conclusion . . . 235

7 Conclusions 237
7.1 Summary . . . 237
7.2 Conclusion . . . 240
7.3 Future Work . . . 241

A Symbols and Abbreviations **244**
A.1 List of Symbols 244
A.2 List of Subscripts 245
A.3 List of Abbreviations 246

References **247**

Chapter 1

Introduction

Oxygen is essential to aerobic respiration and therefore is critical to the existence of human life. This makes the passage of oxygen from the air to the organs of the body and the return of the by-product of respiration, carbon dioxide, very important. A critical part of that process is the transfer of gases between air and blood, which occurs in the lung. Therefore being able to assess the state of the lung, particularly in pathological conditions, can be very important clinically.

Although there are several methods of assessing lung function they usually require the cooperation of the subject or are only suitable for use within a dedicated lung function laboratory or both. This excludes their use in unresponsive patients or for routine use throughout a hospital in general, at less specialist centres and in primary care. For example routine monitoring of lung function within the intensive care unit (ICU) is not commonplace, even though such patients could benefit greatly from routine monitoring of their lung function.

Colleagues at the Nuffield Division of Anaesthetics, Nuffield Department of Clinical Neuroscience, have developed equipment and methods to determine the lung volume of subjects in such situations. This thesis develops new methods for processing the data collected by such equipment, in order to better determine meaningful information about the lung for use in clinical diagnosis and therapy.

1.1 Clinical Motivation

There are many applications where a readily available measure of functional residual capacity (FRC), a specifically defined lung volume, or ventilation homogeneity could be beneficial. For example these could be used for monitoring lung disease within cystic fibrosis patients, particularly in young subjects, or progressive obstructive diseases, such as asthma and chronic obstructive pulmonary disease (COPD), in the context of primary care. However the primary application that motivates the research in this

thesis is the monitoring of ventilated patients in order to assist decisions about their care, particularly the selection of the best ventilator settings.

Selection of ventilator settings is always a compromise. It is important to keep the lung open: if parts of the lung are allowed to close they can be injured when they snap back open [70, 77]. Over-inflating the lung, however, can also cause damage. The main ways in which lung inflation can be controlled are by adjusting the ventilator settings for tidal volume and positive end-expiratory pressure (PEEP), the amount of additional pressure (above atmospheric pressure) maintained at the end of the breath. In most normal situations a small amount of PEEP is applied but higher levels may be applied, as determined by a clinician.

24% of patients who start ventilation with no signs of acute lung injury (ALI) will go on to develop ventilator-associated lung injuries (VALI)[1] within the first five days of ventilation [81]. Therefore being able to better determine ventilator settings has the potential to greatly improve patient outcomes. For patients with ALI, and particularly the severe cases of ALI known as Acute Respiratory Distress Syndrome (ARDS), the choice of ventilator settings can be even more critical [70].

Low tidal volume ventilation in these patients is now widely accepted as the best strategy, but the optimum level of PEEP remains a controversial subject [70]. Higher levels of PEEP can help to reopen parts of the lung, recruiting alveolar units and reducing alveoli collapse. However the optimum level of PEEP required to achieve this without over-inflating the lung has not yet been established. While trials have shown clinical improvements when using high PEEP within ALI patients, they have generally not also shown improvements in mortality [70, 171].

Computed Tomography (CT) and ultrasound can be used to assess the potential recruitability of the lung and the response post application of higher PEEP [70, 171]. However these are discrete forms of monitoring, rather than continuous ones, and require the support of trained radiologists to be carried out. There are also obvious disadvantages associated with CT: it can not be performed at the bedside and exposes the patient to a high radiation dose. Some success has been reported by using combined chest wall and lung pressure-volume analysis, however there are still many uncertainties in relation to these techniques [77]. This has led to much debate about how the optimum PEEP level should be determined: via pressure-volume analysis or via simple but crude arterial oxygenation tables [133, 134, 175, 176].

[1]VALI is also known as ventilator-induced lung injury (VILI). The distinction being that VILI is proven to be caused by a mechanical ventilator, while VALI are injuries which develop during ventilation but the causes of which are unknown.

The use of FRC, in addition to pressure-volume analysis, has been advocated for some time [77, 94, 171]; however the ability to effectively determine FRC at the bedside was not possible until relatively recently. A number of commercially available products for measuring FRC are now on the market. Although they show promise, further validation studies are required before they can be recommended for routine clinical use [171].

The inhomogeneity of the lung is an important factor in VALI, as the forces acting on the lung have been shown to be several times higher within an inhomogeneous lung, due to the greater mismatch in the pressures experienced by neighbouring alveoli [35, 77]. Therefore measures of inhomogeneity would be valuable when determining optimum PEEP levels. Although methods for determining inhomogeneity have been applied to ventilated patients in research settings and PEEP has been shown to affect the inhomogeneity, the use of inhomogeneity measures to inform the level of PEEP has not yet seen clinical use [35, 137].

1.2 Technical Challenges

There are a number of technical issues that have prevented the routine monitoring of FRC and lung inhomogeneity in ventilated patients. Firstly it is important to note that the volume of the lung is not constant and can change throughout the breath by more than a litre. In ventilated subjects the breathing process is, at least, more controlled and so the lung will return to approximately the same volume at the end of each breath. However determining a volume which is constantly changing still presents significant issues.

A number of methods used within dedicated hospital lung laboratories are not suitable for bedside use due to the bulky nature of the equipment [77, 95], for example methods that enclose subjects within an air sealed box are not practical for use throughout a hospital. Also a number of these techniques require the cooperation of the subject and so are not feasible in unconscious patients, which is by far the most common state of ventilated patients [58]. Although there are methods that overcome these issues none of them have yet made it into mainstream use for various reasons.

Several methods rely on the measurement of a tracer gas flow in and out of the body [77, 135]. These flows are generally determined by measuring the total flow rate at the mouth and the fractional concentration of the tracer gas within that flow. Although there are issues in determining the total flow rate, for example the flow rate being affected by temperature and the water content, the main issues come from determining the concentration of the tracer gas and the alignment of the two signals [77, 102].

Mass spectrometers can provide accurate good quality fractional concentration measurements of tracer gases that have been used for determining FRC and homogeneity measures. They can also be used for measuring the concentration of a wide variety of gases. They are, however, not ideal for such applications. Their bulky and temperamental nature means they are not ideally suited for use throughout a hospital [89, 102]. It is also becoming harder to source suitable respiratory mass spectrometers as the number of manufacturers reduces [21, 102].

Mass spectrometers require the use of sampling tubes which suck a small amount of air from the main flow into a side stream flow which is then analysed. This can cause a number of its own issues. The gas takes a finite amount of time to travel down the sampling tube creating a delay and so causing a misalignment with the flow signal [27, 102, 191]. This delay can be affected by pressure differences between the ends of the sampling tube. The delay can also be affected by the viscosity of the sampled air, which will change with the concentration of the measured gas [47]. The sampled gas will also disperse slightly as it travels down the sampling tube corrupting the concentration measurements to some degree [22].

The amount of available data can be limited when techniques rely on a small number of observations per breath, as a patient's breathing rate can not be significantly increased. It can be further curtailed in certain techniques by the assumptions which rely on a particular form of perturbation. For example, in the case of multiple breath wash out methods, as described later, the information content of the data diminishes after about seventeen breaths [207, 216]. A reasonable amount of time is then required before further data collection to allow the gas content of the lung, and in some situations the body, to re-equilibrate. This can be particularly problematic when determining information about the homogeneity of the lung. Many methods have run into issues when trying to determine information about a large number of parameters from little data [117, 206, 216]. This has led to a general criticism of these methods and the measures being easily corrupted by noise [104, 107, 149, 207, 216].

1.3 Aims of this Thesis

The ultimate aim of the research of which this thesis forms part is to be able to better monitor lung function in all care settings. As already mentioned, the initial focus is the monitoring of lung function in ventilated patients in order to better determine PEEP. This thesis concentrates particularly on determining FRC and ventilation-volume measures of homogeneity.

The data used within this thesis come from equipment developed by colleagues at the Nuffield Division of Anaesthetics, Nuffield Department of Clinical Neuroscience, which overcomes a number of the the practical issues highlighted above. It does not use a mass spectrometer or sampling tube, using instead a small sensor that samples the main flow stream.

The thesis has three primary aims:

- To propose models of the lung that will enable the entire expired signal to be used when inferring information about the lung; thus potentially increasing the information content available from a set of data.
- To use these models to better determine the FRC of the lung.
- To use these models to determine information about the inhomogeneity of the lung.

In proposing new models of the lung, the aim will not be to model all the intricate details of the lung, but only to advance the models in ways that will enable the FRC and indicators of the inhomogeneity of the lung to be determined in a more accurate, precise and repeatable manner. These aims, if fulfilled, should overcome some of the issues relating to the information content of the data outlined above. They should also help to further research towards the ultimate aim of better lung function monitoring in the ICU and other care settings.

1.4 Outline of this Thesis

After providing background information on some basics of lung physiology and the research relating to the subject, the methods employed later in the thesis are detailed. These methods detail development of a model for the entire expired breath, which is then applied later in the thesis to different sets of data.

The chapters are structured as follows:

- Chapter 2 provides some details about the basic physiology of the lung, introduces the concepts that will be discussed later in the thesis and reviews the relevant literature related to the subject.
- Chapter 3 provides details of the methods that are applied in the thesis. In particular it introduces the data that will be used, presents the computational models that will be applied to the data and describes the parameter estimation techniques that will be used.

- Chapter 4 investigates the ability of these methods to estimate parameters from simulated data. Simulated data have the advantage of not being corrupted by noise or have a known amount of noise. This allows the estimation techniques to be investigated under known conditions, allowing valuable information about their ability to infer parameters to be obtained.

- Chapter 5 investigates data from a water displacement bench lung. The advantage of this lung is that the volumes within it are known, allowing the techniques to be evaluated effectively in data which include many of the issues also associated with human subject data.

- Chapter 6 investigates data from naturally breathing healthy human subjects. Although not providing some of the complications of data collected from ventilated patients, these data incorporate all the issues associated with determining information about real lungs. This allows the robustness of the techniques to be investigated further.

- Chapter 7 summarises and concludes the work described in this thesis. It also provides some direction on possible future research in order to fulfil the ultimate aims.

Chapter 2

Literature Review

2.1 Introduction

This chapter introduces the concepts that will be discussed and used later in the thesis, in addition to reviewing the literature related to these subjects. The chapter opens with some of the basic physiology of the lung before detailing how this is translated into mathematical models of lung ventilation. How these models, and other models, are used to determine physiological properties in conjunction with tracer gases is then discussed; in addition to a more general discussion on determining information about lung parameters and general methods of parameter estimation.

Some of the practical issues of collecting the two signals within the data, the concentration of the tracer gas and the flow rate into the lung, are also covered. In particular the techniques used to align the two signals in time and to compensate for the response time of the concentration sensor are detailed. Together these topics form the background to the research which is covered later in the thesis as well as providing further details surrounding the motivation covered in the first chapter.

2.2 Lung Physiology

The lung is the means by which oxygen is transferred from the air into the blood stream and carbon dioxide is transferred back in the opposite direction. Air enters at the mouth and nose and travels down the trachea. The trachea divides into two primary bronchi which serve the left and right lung, as shown in Figure 2.1. These primary bronchi divide into smaller secondary bronchi and tertiary bronchi and so on for up to an average of sixteen generations. This leads to an inverted tree like structure.

The bronchioles keep dividing into smaller generations until they become terminal bronchioles, the smallest size bronchioles not to be involved in gas exchange. These terminal bronchioles subdivide further into respiratory bronchioles, which are the first part of the lung where gas exchange can occur. The respiratory bronchioles proceed to

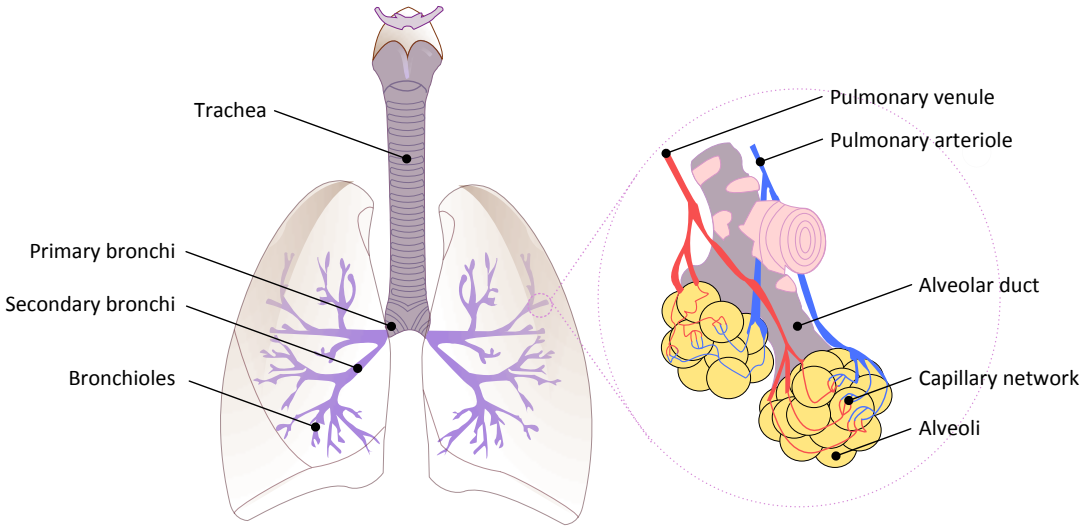


Figure 2.1: Structural diagram of the respiratory system. The inert shows the detail of the alveolar ducts and sacs, where gas exchange takes place. Adapted from [164].

divide into alveolar ducts which end in one or more alveolar sacs. The walls of these sacs consist of numerous alveoli grouped together like bunches of grapes. On average there are twenty three generations in total from the primary bronchi.

A similar structure of branching occurs to the pulmonary arteries and veins which bring blood to and from the lungs, but numerous supernumerary connections are also included, allowing blood to flow via alternative paths, as required. These ultimately terminate in a network of capillaries which surround the respiratory bronchioles, alveolar sacs and alveoli.

Travelling down the lung from the primary bronchi into deeper branches of the lung, the structure changes from a solid tube made of cartilaginous rings to less solid walls, with less cartilage within it, until the cartilage disappears entirely. By the time the respiratory bronchioles are reached only a very thin layer of cells remain, allowing for easy exchange of gases with the blood. The alveoli are the primary source of gas exchange in the lungs, with around 300 million alveoli present. The surface tension of the water present within these very small alveoli would cause them to collapse, were it not for the surfactant which coats their walls.

2.2.1 Lung Volumes

The volume of the lung is constantly changing, which is part of the reason that determining an exact volume of the lung is not a simple task. Therefore when reporting a technique it is important to be precise about which volume of the lung being deter-

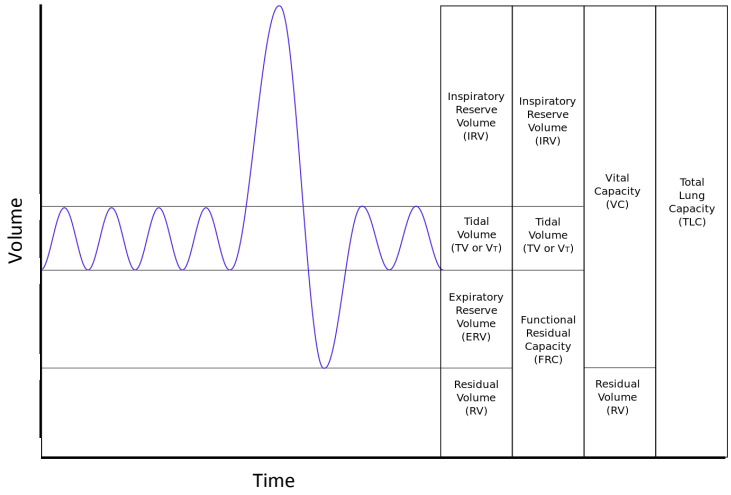


Figure 2.2: A volume-time spirogram of showing static lung volumes. Adapted from [172].

mined. Within this thesis the widely accepted definitions of particular lung volumes will be used [5, 6, 161, 162, 210], as illustrated in Figure 2.2.

During normal breathing at rest, from an end-expiratory level an amount of air will be inspired and then another amount of air expired. The amount of air will vary slightly from inspiration to expiration and between breaths; therefore so will the end-expiratory level. The average volume of air inspired/expired is the tidal volume (V_T). The functional residual capacity (FRC) is the volume of air in the lung at the average end-expiratory level. From the average end-expiratory level it is possible to exhale more air than would be normal until a point at which no more air can be exhaled. Once this point is reached the volume of air exhaled is the expiratory reserve volume (ERV) and the volume remaining in the lung is the residual volume (RV).

If the maximum volume of air is inhaled, the volume in the lung is the total lung capacity (TLC), with the volume of air inhaled above the average end-inspiratory level being the inspiratory reserve volume (IRV). Vital capacity (VC) is the volume change between maximum inspiration and maximum expiration.

The average end-expiratory level in ventilated patients is often referred to as the end-expiratory lung volume (EELV), rather than the FRC. This is considered a more precise term by some, as the volume can be effected by PEEP [153]. Within this thesis FRC will be used, as is commonly seen in the literature [136, 148], explicitly stating when ventilation is used. FRC is usually more easily determined in ventilated patients as they return to a very similar end-expiratory level, compared with the more erratic end-expiratory levels observed in spontaneously breathing subjects. It is relatively easy

to determine all volumes that do not include the RV using conventional spirometry. In the review that follows, and indeed the thesis as a whole, the focus will be on determining volumes that include the RV, particularly FRC.

2.2.2 Lung Inhomogeneity

The volume, the ventilation (the amount of air reaching part of the lung) and the perfusion (the blood flow through part of the lung) will vary throughout the lung. In a young healthy lung, although these will vary individually, the ventilation-volume and ventilation-perfusion ratios are relatively well matched throughout the lung. In fact, when in an upright position, the lower parts of the lung generally have higher perfusion rates, due to the effects of gravity. This is matched by higher ventilation, due to the position of the diaphragm, leading to a relatively consistent ventilation-perfusion ratio throughout the lung. However with age and in certain disease conditions, mismatches in these ratios will build up leading to an inhomogeneous lung.

At the extreme this leads to parts of the lung which are either not ventilated or not perfused. The non-perfused parts contribute to the alveolar dead space. The make up of blood passing through non-ventilated parts will not change and so it is as if bypasses the lung. This is referred to as pulmonary shunt. Although parts of the lung may exhibit these two extremes, there will also be parts of the lung that will exhibit a range of ventilation-volume and ventilation-perfusion ratios.

Methods that specifically account for these inhomogeneities exist and generally either use a number of discrete compartments or a continuous distribution to do this. Previous work has generally made an explicit or implied assumption of a log-normal distribution, or a series of log-normal distributions, throughout the lungs [42, 73, 105, 107, 117, 137, 163, 187, 197, 207, 214, 216–218]. However no underlying physiological reason for this type of distribution has been proposed.

This review will generally concentrate on determining the ventilation-volume distribution, as this is the focus of the thesis, but also discuss methods for determining the ventilation-perfusion distribution where directly relevant. In addition to these inhomogeneities, a distribution in the emptying rates across the lung can exist. This form of inhomogeneity will not be the primary focus of this thesis. This is consistent with much of the literature [62, 85, 117, 139, 147, 170, 199, 216, 218], as this type of inhomogeneity is complex to model and likely to further complicate parameter estimation. There is also some evidence, in normally breathing healthy subjects at least, to suggest that this is not the dominant form of inhomogeneity [152, 235].

2.3 Ventilation Models

Although an attempt could be made to understand and to model all the intricate details of this branching structure, the millions of alveoli and other structures that make up the lung, in reality this is not possible and, as discussed further below, it would be impossible to invert these models anyway. Therefore over the years simplified ways of explaining and modelling how the air travels through and mixes within the lung have been devised; these are now described.

2.3.1 Bohr Equation

In 1891 Bohr introduced the concept that the quantity of carbon dioxide eliminated from the alveolar compartment is equal to the quantity of carbon dioxide eliminated in the mixed expired gas. i.e. it can be assumed there is no carbon dioxide present in the air. Two equations have resulted from this and both have been referred to as the Bohr equation [58, 144, 148, 202, 233, 234].

The first equates the carbon dioxide in the alveolar compartment and the mixed-expired concentration of carbon dioxide over several breaths assuming a continuous breathing process. This allows the alveolar flow rate to be determined [144, 148]:

$$\dot{V}_A = \frac{\dot{V}_I . F_{\bar{E},CO_2}}{F_{A,CO_2}} \ , \tag{2.1}$$

where $\dot{V}_I$ and $\dot{V}_A$ are the air flow rate into the mouth and alveolar compartment and $F_{\bar{E},CO_2}$ and F_{A,CO_2} are the mixed-expired and alveolar fractional concentration of carbon dioxide, respectively. The second equates them over a single breath [144, 202]:

$$\frac{V_D}{V_T} = \frac{F_{A,CO_2} - F_{\bar{E},CO_2}}{F_{A,CO_2}} \ , \tag{2.2}$$

where V_D and V_T are the dead space and tidal volume, respectively. Subsequently this has been adapted for use with tracer gases that are included on the inspired breath [58]:

$$\frac{V_D}{V_T} = \frac{F_A - F_{\bar{E}}}{F_A - F_I} \ , \tag{2.3}$$

where F_A, $F_{\bar{E}}$ and F_I are the alveolar, mixed-expired and inspired fractional concentration of any tracer gas.

Bohr did not proposed a method of determining the alveolar fractional concentration in order to solve these equations [144]. In the early 1900's it was realised that the end-expired concentration could be used to estimate the alveolar concentration, allowing the alveolar flow rate and dead space volume to be determined. This dead space has become known as the anatomical dead space. Later a method was proposed for

determining the physiological dead space, those parts within which no gas exchange occurs with the blood. The concentration of carbon dioxide in the alveolar blood is used to determine the 'ideal' alveolar concentration, which is then substituted into Equation 2.2. The physiological dead space is made up of the anatomical dead space and the parts of the alveoli within which no gas exchange occurs, the alveolar dead space.

These equations in essence are a basic ventilation model and gave rise to the concept of a dead space compartment, which is used in a number of the models described below. In addition, they have been widely applied in their own right with some example uses described below.

2.3.2 Compartment Models

2.3.2.1 Continuous Model

The continuous model is analogous to the gills of fish [90, 212], which have a continuous flow of water over them, leading to the main assumptions of the model: continuous gas flow within the lungs. Figure 2.3a is a schematic representation of the classic continuous model. The mass balance equation for the alveolar compartment is [90]:

$$V_A \frac{dF_A}{dt} = \dot{V}_A(F_I - F_A) + \dot{Q}_P(C_{\bar{v}} - C_a) \ , \tag{2.4}$$

where V_A is the volume of the alveolar compartment; F_I is the fractional concentration of the inspired gas; $\dot{Q}_P$ is the pulmonary blood flow; $C_{\bar{v}}$ and C_a are the mixed venous and pulmonary end-capillary gas concentrations, respectively [1].

The model normally includes a compartment with ventilation but no perfusion, effectively an air flow path that bypasses the lung, as shown in Figure 2.3a. This represents the physiological dead space. The proportion of flow passing through the dead space may be set or derived from other measurements depending on the application, with the relationship between flow and the fractional concentration being [90]:

$$\frac{\dot{V}_D}{\dot{V}_I} = \frac{F_E - F_A}{F_I - F_A} \ , \tag{2.5}$$

where $\dot{V}_D$ is the flow rate through the dead space and F_E is the expired fractional concentration. In most situations it is assumed that the ratio of gas flow is equal to the ratio of volumes in the Bohr equation (see section 2.3.1) [90, 91].

[1] Within this thesis the convention observed elsewhere [5, 6, 83, 87, 91, 144, 233] is maintained, that is blood originating from the body's veins is referred to as venous blood, despite entering the lung in the pulmonary artery. Likewise blood leaving the lung in the pulmonary vein is referred to as arterial blood as it is ultimately destined for the body's arteries

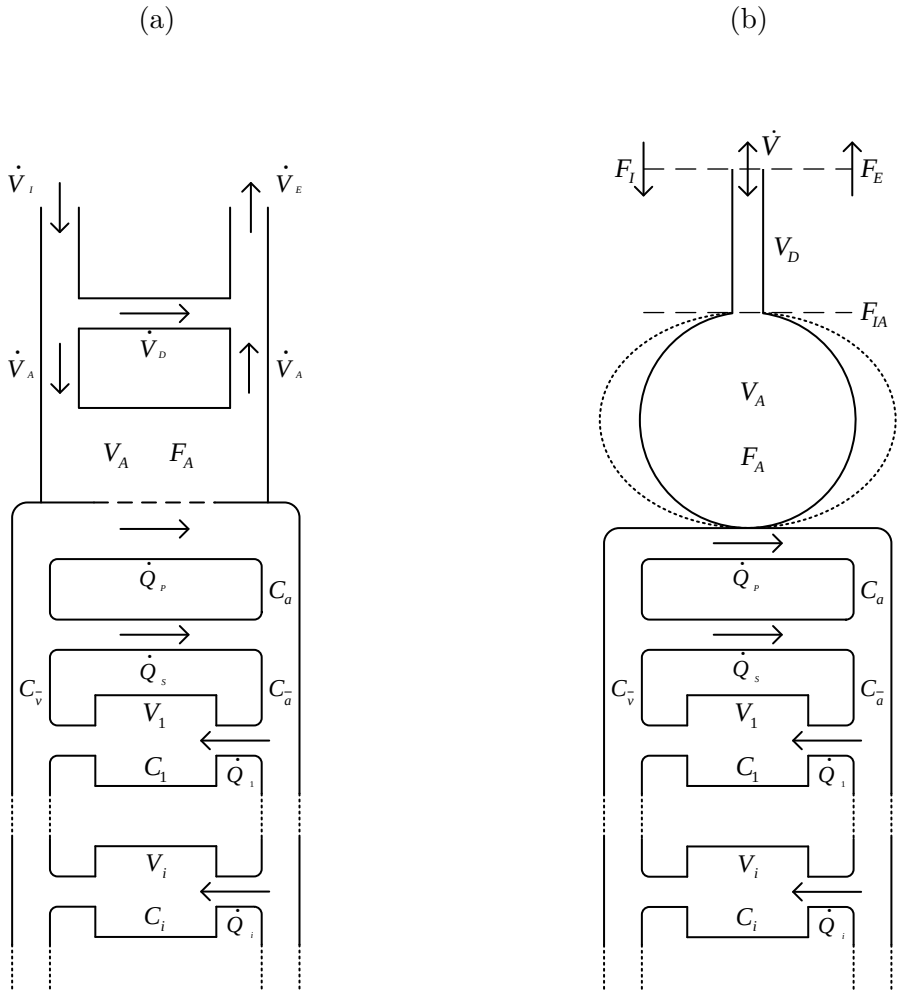


Figure 2.3: Schematic diagram for the two types of compartment model considered here. (a) Continuous ventilation model. (b) Tidal ventilation model. The symbols are defined in the text along with further explanation of the two models.

Although the flow through the dead space compartment is usually from the inspired air to the expired air, there are examples in which flow in the opposite direction has been applied [142]. This means that air exiting the alveolar compartment is returned to the alveoli, as happens in reality during tidal breathing. However in this case, perhaps counter intuitively, the flow rate inspired and expired is less than the flow through the alveolar compartment.

The assumptions within this model produce equations which can be solved analytically without the need to resort to numerical solutions [84, 90]. This coupled with the ability of this simple model to produce meaningful results, has led to its widespread use [90]. However, there are several applications when this model is not able to explain physiological observations.

2.3.2.2 Tidal Model

The tidal model separates breathing into two phases, inspiration and expiration, with the air entering through a common dead space. The simplest form of this model is the 'balloon on a straw' tidal model, as shown in Figure 2.3b [90]. This model consists of an anatomical dead space of fixed volume, the 'straw', through which the air passes before entering the alveolar compartment. On reaching the alveolar compartment the air mixes with the air already in the compartment. The volume of the alveolar compartment may change, like a balloon.

In this model breathing is separated into two phases governed by two different mass balance equations [90]:

Inspiration:

$$\frac{\mathrm{d}}{\mathrm{d}t}(V_A(t).F_A(t)) = \frac{\mathrm{d}V_A(t)}{\mathrm{d}t}F_{IA}(t) + \dot{Q}_P(C_{\bar{v}}(t) - C_a(t)) \ , \tag{2.6}$$

Expiration:

$$\frac{\mathrm{d}}{\mathrm{d}t}(V_A(t).F_A(t)) = \frac{\mathrm{d}V_A(t)}{\mathrm{d}t}F_A(t) + \dot{Q}_P(C_{\bar{v}}(t) - C_a(t)) \ , \tag{2.7}$$

where F_{IA} is the fractional concentration of the air entering the alveolar compartment. Changes in volume due to the uptake and return of the tracer gas from the blood are generally ignored [84, 216, 218], as is the case in the continuous model [25, 26, 83, 91, 233], so:

$$\frac{\mathrm{d}V_A}{\mathrm{d}t} = \dot{V}(t) \ , \tag{2.8}$$

where $\dot{V}(t)$ is the flow rate into the lung.

The delay through the anatomical dead space is dependent on the flow rate into or out of the lung. A fixed breathing pattern is usually assumed [84, 218, 219], with a constant flow rate during inspiration, meaning:

$$F_{IA}(t) = \quad F_{EDS,n-1} \qquad T_n < t < T_n + t_{ID} \tag{2.9}$$

$$F_{IA}(t) = \quad F_I(t - t_{ID}) \qquad T_n + t_{ID} < t < T_n + t_I \ , \tag{2.10}$$

where t_{ID} is the time taken for inspired air to pass through the anatomical dead space, T_n is the time at the beginning of breath n, t_I is the duration of inspiration and $F_{EDS,n-1}$ is the mixed gas concentration expired into the dead space during the previous breath.

A continuous signal for expired concentration at the mouth is not normally modelled with two discreet concentrations determined for each breath instead [84, 218]. The end-expired fractional concentration, $F'_{E,n}$, is:

$$F'_{E,n} = F_A(T_n + t_I + t_{ED}) \ , \tag{2.11}$$

where t_{ED} is the time taken for $V_T - V_D$ to be expired from the lung. Thus $(T_n + t_I + t_{ED})$ is the time point at which the last air reaching the mouth left the alveoli. The mixed-expired fractional concentration, $F_{\bar{E},n}$, is:

$$F_{\bar{E},n} = \frac{V_D}{V_T}.F_{IDS,n} + \left(1 - \frac{V_D}{V_T}\right).F_{EM,n} \ , \tag{2.12}$$

where $F_{IDS,n}$ is the flow weighted average fractional concentration of the tracer gas left in the dead space during inspiration and $F_{EM,n}$ is the flow weighted average fractional concentration of the tracer gas that leaves the alveoli during expiration but that is not left in the dead space. The alveolar dead space has generally not specifically been dealt with in tidal models [84, 218, 219], but theoretically it could be included as an additional non-perfused alveolar compartment.

The tidal model has been adapted for different applications. For example it has been used to investigate oxygenation during respiratory failure [219]. In this case, a continuous model would not be sufficient to explain within breath oxygen fluctuations. Although this type of model produces equations that need to be solved numerically, this can be done quickly and efficiently by modern desktop computers; therefore its use has now become more widespread, even for the inverse process [90]. It was argued by Hahn and Farmery [90], back in 2003, that the continuous models should now be abandoned in favour of tidal models.

2.3.2.3 Oswald Coefficient

The Oswald coefficient is used to link the concentration of the tracer gas in the blood to the fractional concentration in the lung. It is the steady state ratio of the gas volume contained within a volume of liquid [124, 144] and is expressed in terms of the gas volume at a given temperature. For an ideal gas the coefficient is independent of pressure, due to the combined effect of Henry's and Boyle's laws.

Normally equilibrium is assumed to be reached within the lung [25, 26, 83, 84, 91, 216, 218, 233], thus concentration of the gas dissolved in the pulmonary end-capillary blood (gas volume/liquid volume) is:

$$C_a = \lambda.F_A \ , \tag{2.13}$$

where λ is the blood-gas partition coefficient, which is the number of molecules of gas in the blood to that in gas phase. As one phase is gas and the other liquid, the partition coefficient is identical to the Oswald coefficient [144]. The fractional concentration of the alveoli (F_A) is the gas volume/total gas volume concentration of tracer gas.

All Oswald coefficients quoted in this thesis are the coefficient at body temperature, generally 37^oC.

On occasions throughout the thesis an insoluble gas is referred to. In reality all gases are soluble in blood to some extent. In these cases, a gas with very low Oswald coefficient is implied, such as argon ($\lambda \approx 0.003$). In the context to which they are referred these gases will act essentially as if they are insoluble.

2.3.2.4 Shunt

Pulmonary shunt can be included as a compartment that receives perfusion but no ventilation, effectively a blood flow path that bypasses the lung. This is shown in Figure 2.3. The mixed-arterial concentration, $C_{\bar{a}}$, is then determined by the combination of the venous blood and that from the lung [90, 216]:

$$C_{\bar{a}}(t) = \frac{\dot{Q}_S}{\dot{Q}} C_{\bar{v}}(t) + \frac{\dot{Q}_P}{\dot{Q}} C_a(t) \ , \tag{2.14}$$

where $\dot{Q}_S$ and $\dot{Q}$ are the shunt and total flow rates, respectively. Normally the shunt fraction, $\dot{Q}_S/\dot{Q}$, is given instead of the shunt flow rate, with the shunt flow rate then determined for a particular total flow rate.

2.3.2.5 Body Compartments

The body may be modelled by a number of compartments, as shown in Figure 2.3. The mass balance equation of these compartments is [84, 90, 233]:

$$V_i^* . \frac{\mathrm{d}C_i(t)}{\mathrm{d}t} = \dot{Q}_i (C_{\bar{a}}(t) - C_i(t)) \ , \tag{2.15}$$

where V_i^*, C_i and $\dot{Q}_i$ are the equivalent blood volume, concentration of the tracer gas and flow rate of compartment i. The equivalent blood volume of the compartment is given by:

$$V_i^* = V_{b,i} + \frac{\lambda_{t,i}}{\lambda} V_{t,i}, \tag{2.16}$$

where $V_{b,i}$ and $V_{t,i}$ are the blood and tissue volumes of each body compartment, $\lambda_{t,i}$ is the Oswald tissue-gas coefficient of each body compartment.

The number of compartments varies between application. For simplicity sometimes a single compartment is used [89, 229], but normally more compartments are used [91, 127]. In the modelling of anaesthetic agent uptake and the applications described below, organs with similar perfusion-volume ratios and tissue-gas partition coefficients are combined into a single compartment. The compartments used in this thesis are based on a 'standard' 70kg man and detailed in Table 2.1. They are based on information compiled by Mapleson [127] from many sources and previous proposals. This was

	Compartment	Blood Volume (l)	Tissue Volume (l)	Blood Flow (l/min)	Oswald Tissue-Gas Coefficient of Nitrous Oxide
	Lung	0.37	0.55	5.34	0.47
1	Viscera – Fast	0.87	0.35	1.39	0.40
2	Viscera – Liver	0.98	3.73	1.48	0.43
3	Viscera – Rest	0.62	1.89	1.09	0.51
4	Lean	0.53	37.4	1.13	0.40
5	Fat	0.16	12.8	0.25	1.40

Table 2.1: Body compartment parameter for the $70kg$ standard man. These values are taken from [91].

adapted to its current form by Hahn [91] and has been used by a number of authors since [77, 83, 84].

2.3.2.6 Multiple Alveolar Compartments

So far homogeneous compartment models of the lung have been discussed, where the ventilation-volume and ventilation-perfusion ratios are constant throughout the lung. In order to model the mismatches which may build up, multi-compartment models of the lung can be used. These are made up of a number of compartments, each with a unique volume, ventilation and perfusion. A common approach is to concentrate either on the ventilation-volume [117] or ventilation-perfusion [206] mismatch, with a number of compartments, normally fifty, being set-up on a log scale, as discussed further below.

An alternative approach is to setup a continuous distribution of the lung. This has been done simultaneously for both the ventilation-volume and ventilation-perfusion ratios [216, 218]. Both distributions were assumed to follow a log-normal distribution. The model was solved using a numerical process, which effectively uses a large number of small discrete compartments in order to solve the model.

2.3.3 Trumpet Model

In the trumpet model the lung is modelled as a series of generations. The transport of gas molecules gradually changes from a process dominated by convection to one dominated by diffusion deeper into the lung. Generally it is based on an anatomical model published by Weibel [211]. When the cross-sectional area is plotted against distance for the last 13 generations of branching the result resembles a trumpet shape [150], hence the name.

The model splits the lung into two parts: the first 10 generations, within which diffusion is ignored and mass transport is assumed to be governed by convection only, and the last 13 generations, within which diffusion is included [151]. Mass transport

within the last 13 generations is modelled by a mass transport equation:

$$\frac{\partial F}{\partial t} = D\frac{s}{S}\frac{\partial^2 F}{\partial l^2} + \frac{D}{S}\frac{ds}{dl}\frac{\partial F}{\partial l} - u\frac{\partial F}{\partial l} + G \ . \tag{2.17}$$

F is the fractional gas concentration; s is the cross-sectional area of the conducting airway only; S is the total cross-sectional area including the airway and the alveoli; l is the distance from the terminal alveoli of the last generation; D is the diffusion coefficient; u is the convection velocity; G is the gas source term; t is time. This equation is based on six assumptions described in detail by Paiva [151]. The G term accounts for gas that diffuses across the alveoli boundary, to/from the blood. For an insoluble gas this would be zero throughout the whole lung. The $u\frac{\partial F}{\partial l}$ term describes the convectional transport, while the remaining terms describe the diffusional transport.

This model allows the fractional concentration of a gas to be determined at all depths within the lung. Deeper into the lung, diffusion becomes increasingly dominant. In fact a point of marked change between parts of the lung dominated by convection and those dominated by diffusion has been identified, which is known as the diffusion front [151]. This transition can be observed during the expired breath and is the equivalent to the transition between air from the dead space and air from the alveolar compartment in the tidal model. The position of the diffusion front will change with breathing pattern and the diffusion coefficient of the gas.

This model has been used to examine how breathing patterns affect the performance of the lung. It has been shown that the breathing pattern has an effect on conventional models used to predict blood flow, airway dead space and end-tidal volume [90]. The solutions to the mass transport equations are complex and until recently would normally require more computing power than is found on a desktop computer [90]. Therefore it is currently only well suited for the forward process rather than the inverse process, as described below.

2.3.4 Complex Anatomical Models

The trumpet model treats the lung as symmetrical with every path to the terminal bronchial being the same. So the trumpet model can not be used to investigate gas transport problems affected by asymmetries within the lung [192] and the effects of gravity, which causes the ventilation-volume and ventilation-perfusion ratios to differ throughout the lung [48, 144]. Therefore models which replicate the branching structure of the bronchial tree lung have been developed. These can be based on scans from living humans [49]. The first ten generations of the bronchial tree can be taken directly from the scans. The remaining tree is then grown to fill the lobe dimensions of the lung

(also determined from the scans), using an algorithm that mimics the lung branching structure [193]. These models have been coupled with patient specific geometries to investigate effects of the carrier gas properties on aerosol deposits [114], amongst other applications. A similar process has been applied to pulmonary arteries and veins, allowing the effects of pulmonary embolism investigated [49, 50].

These models form part of a wider Human Physiome Project which aims to provide a "quantitative description of physiological dynamics and functional behaviour of the intact organs" [3]. Work on the respiratory system lags behind the most successful work, which has been applied to the cardiac system; however, ultimately the aim is to link models of different spatial and temporal scales, as has been achieved with the heart [49, 90]. This may lead to the possibility of modelling disease processes and interventions on a computer, allowing so called 'in silico' experimentation [90]. These models will lead to an alternative way of validating methods that inverse much simpler models; however, due to the very large number of parameters and very substantial computational power required to solve them, they are unlikely to prove useful in the inverse process themselves.

2.4 Estimating Lung Parameters using Tracer Gases

Within medicine models have commonly been used for the forward process [90]; where predictions of measurable quantities, such as the blood oxygen concentration, are calculated by inputting the physiological state of the patient. The power of models, however, primarily comes from using the inverse process. This allows unknown physiological parameters, such as the pulmonary blood flow, to be extracted from measurable quantities.

Uses of the inverse process are discussed in this and the following sections, starting with the applications where physiological parameters are determined using the concentration of tracer gases. A tracer gas may be a gas that naturally occurs within the air (oxygen, carbon dioxide or nitrogen) or one that is specifically inhaled or infused into the blood of a patient, such as nitrous oxide.

2.4.1 Wash-out Methods

An early use of a wash-out method was to determine the FRC. The initial technique involves the washing out of nitrogen by 100% oxygen consumed for seven minutes [141, 210]. The concentration of nitrogen is taken at the start and end of the process. Air expired is also collected in order to determine its volume. This allows the volume of nitrogen in the lung at the start to be balanced with the volume of nitrogen in the

lung and collection bag at the end of the process. The volume of nitrogen excreted from the body tissues also needs to be accounted for, initially assumed to be 220 ml for all subjects [141].

A number of developments have subsequently improved the technique [141, 210]. The volume of nitrogen excreted from the tissue can now be adjusted depending on the height and weight of the subject. The advent of the mass spectrometer and personal computers has allowed the development of an open circuit technique, which removes the need for the cumbersome collection bag. The fixed seven minute time period for collection is no longer recommended, as in some subjects this will result in low concentrations at the end of the process which are prone to error. For other subjects, for example emphysema patients, the lung will require longer to equilibrate. Therefore the current recommendation is to stop the test once the nitrogen concentration falls below 1.5% for three consecutive breaths.

Reliable and repeatable results can be determined via this method. However it has a number of specific limitations [141]. Only freely communicating air is measured with air trapped behind obstructions not included. However in certain situations where the amount of functioning lung is important, this may be preferable. There is a lack of standardisation in equipment and procedures used between centres making comparison difficult. Many centres still use mass spectrometers. Multiple measurements can not be made easily, due both to the duration of the test and the time required between tests (15 $mins$ generally but 1 hr for patients with severe obstructions). This technique can not be used for patients on high levels of oxygen. At oxygen concentrations above 70% (and even 50% in premature infants) the amount of nitrogen in the lung is too small to produce accurate results.

Several methods for using this technique in ventilated patients have been developed, making this one of the most appropriate techniques for determining FRC at the bedside [141, 210]. However in ventilated patients it still has a number of specific limitations. Two ventilators are required. Certain forms of ventilation cause additional errors, particularly in infants. Patients, including infants, also have to be intubated using tracheal tubes which include a cuff to ensure no leaks.

The use of oxygen as a proxy for the nitrogen concentration is possible. The initial method was proposed by Mitchell *et al.* in 1982 [136]. The oxygen consumption is calculated during steady state, prior to implementation of the step change in oxygen concentration. It is then assumed that the oxygen consumption remains constant and that the respiratory quotient is one. The method was compared favourably to helium

dilution (detailed further below) in a artificial bench lung made from anaesthetic bags and 21 healthy ventilated human subjects. In healthy human subjects the method overestimated FRC by $21 \pm 213ml$ (mean $\pm$ standard deviation (SD)). No details are given about how the oxygen concentration and flow rates were determined.

This method has received renewed interest with several papers published since 2006. It is implemented in the LUFU system (Dräger Medical, Lubeck, Germany) [95, 125, 153, 213]. The device was verified in a commercially available electromechanical test lung [213] and in human subjects [95, 125, 153]. Two of the studies compared this method to helium dilution and plethysmography in spontaneously breathing patients. In healthy volunteers the underestimate observed using this method varied between the studies; 0.2 ± 0.4 and $0.40 \pm 0.50l$ when compared to helium dilution and -0.1 ± 0.6 and $0.43 \pm 0.54l$ when compared to plethysmography ([95] and [125], respectively).

In the second study [125] unhealthy subjects were also compared with a underestimate -0.02 ± 0.95 and $0.18 \pm 0.61l$ for helium dilution and plethysmography respectively. Results in ventilated patients were significantly better when compared to helium dilution, with an overestimate of $0.008 \pm 0.161l$ observed [153]. This system uses a side stream analyser to determine the oxygen concentration.

An alternative method has also been proposed and is now implemented in a commercially available ventilator, the Engstrom Carestation (GE Healthcare) [53, 148]. This method monitors both the oxygen and carbon dioxide concentrations, removing the need for the assumption of a unity respiratory quotient. Also the method is less vulnerable to offsets in the flow and concentration signals, as the volume of nitrogen is balanced based only on the end-expired concentration and average flow rates in/out of the alveoli. The method has been compared with helium dilution and CT imaging techniques [53]. When determining FRC in ventilated patients the average overestimation was found to be 0.229 ± 0.164 and $0.093 \pm 0.143l$ respectively.

Fowler [79] described a single breath nitrogen washout method for determining the dead space volume. The method allows the dead space volume to be determined graphically over a single breath. The subject inhales 100% oxygen for a breath, ensuring that the inspired nitrogen concentration is zero. The nitrogen content and volume of the expired breath are recorded, as shown in Figure 2.4. The point of transition between air originating from the dead space and the alveoli is determined by matching the areas A and B in Phase II of the breath, as shown in Figure 2.4. The dead space volume is the volume of air expired up to this point. The volume obtained via this method is often referred to as Fowler's Dead Space and is an anatomical dead space.

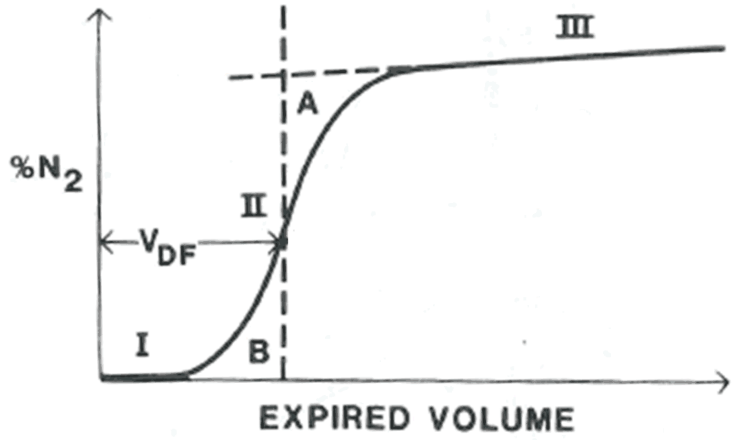


Figure 2.4: A typical plot of nitrogen concentration during expiration after a single inspired breath of 100% oxygen. The dashed lines show how the Fowler Dead Space Volume, V_{DF}, can be calculated graphically, with areas A and B being equal. Three phases of the breath are identified; Phase I: air coming from the dead space, Phase II: the transition from dead space to alveolar air and Phase III: alveolar air. Taken from [71].

For a homogeneous lung this would be identical to the volume acquired using the single breath Bohr Equation. However in reality a slope is observed in Phase III of the breath. Fowler chose to extrapolate this backwards to determine area A, as shown in Figure 2.4. This leads to slightly different volumes being obtained. The volume of the dead space is not of much clinical significance itself [77], however it may be useful in estimating other lung parameters.

Washout methods can also be used to provide information about the ventilation-volume homogeneity of the lung. The simplest indication of inhomogeneity is contained in the concentration of the nitrogen coming from the alveoli (phase III in Figure 2.4). The slope observed in this section represents the fact that the alveolar gas is not uniformly distributed and the magnitude of this slope can be used as a measure of lung inhomogeneity [60, 71, 167]. Inhomogeneities caused by variations in emptying rates within the lung can be a cause of this slope in addition to volume-ventilation disparities. In normally breathing healthy subjects, there is some evidence that variations in emptying rates is not the main contributor to the slope [152, 235].

Various measures of the slope have been proposed. For example one of the simplest was to take the magnitude of the rise between $750ml$ and $1250ml$ of a full expired VC breath [60]. A summary of others can be found in [71]. Via these single breath tests it is possible to distinguish between the age of the group observed and between smokers

and non-smokers [71].

An alternative approach is to consider multiple breaths. If the lungs were homogeneous the proportion of nitrogen washed out in subsequent breaths would be the same as the first. However in reality this is not the case as the inspired gas is unevenly distributed within the lung due to inhomogeneities [71, 167, 216]. The tracer gas expired in the first breath preferentially comes from the well ventilated parts of the lung. In subsequent breaths the nitrogen is washed out, so the concentrations of nitrogen progressively reflect these areas less and poorly ventilated areas more.

Several methods of modelling this process have been proposed. The method, initially proposed by Cournand *et al.* [62] assumes the fractional concentration of nitrogen in the alveolar compartment at breath n, $F_{A,n}$, to be [66, 71]:

$$F_{A,n} = F_{A,0}.W^n \ , \tag{2.18}$$

where W is the factor by which the nitrogen concentration is decreased each breath, which is [66, 71]:

$$W = \frac{V'_A}{V'_A + V_T - V_D} \ , \tag{2.19}$$

where V'_A is the end-expired alveolar volume.

Substituting in the single breath Bohr Equation (Equation 2.2) and then taking the logarithm of both sides the following relationship can be obtained:

$$\log F_{\bar{E},n} = \log \left(F_{A,0}.\frac{V_T - V_D}{V_T} \right) + n \log W \ . \tag{2.20}$$

Therefore, if the lung were homogeneous, when $\log F_{\bar{E},n}$ is plotted against n a straight line with intercept $\log \left(F_{A,0}.\frac{V_T - V_D}{V_T} \right)$ and slope $\log W$ will be obtained. In reality straight lines are rarely found even in normal subjects [71].

The first attempt to analyse the curves obtained in real data was by Robertson *et al.* [166]. The curves were assumed to be made up of two or more exponential functions corresponding to two or more lung compartments. The gradient of the curve towards the end of the process is extrapolated back to the ordinate and the characteristics of the compartment with the lowest ventilation-volume ratio can be determined. This is then subtracted from the original curve and the process repeated until a straight line is obtained [71]. This 'exponential stripping' method has been applied by a number of researchers [145, 166, 180] but has subsequently been found to be extremely sensitive to experimental error [216].

This method essentially represents the lung as a small number of discrete compartments of different ventilation-volume ratios. Rossing [170] formalised this, describing a

model consisting of a number of compartments with different ventilation-volume ratios which all contribute a proportion of the end-expired nitrogen concentration. Starting from the initial guess provided by the exponential stripping method, the volume, ventilation-volume ratio and number of compartments are determined by minimising the following function, E:

$$E = \frac{\sum_{l=1}^{n} (\frac{x_l - m_l}{m_l})^2}{n - k}, \tag{2.21}$$

where x_l and m_l are the end-expired concentrations in the data and those predicted by the model, respectively, n is the total number of breaths and k is the total number of parameters fitted.

A similar method, which included a dead space, was used by Tsunoda *et al.* [199] to investigate six points through the breath. Unsurprisingly the initial part of the breath was dominated by the dead space compartments in all subjects. Within older subjects the contribution of the dead space compartment continues to have an effect further into the breath. In addition a smaller contribution from well ventilated compartments was observed in the early part of the breath.

More recently a three compartment model has been used to determine FRC [41]. Like in the methods described above, it used the oxygen and carbon dioxide concentrations as a proxy for the nitrogen concentration. The volume of the three compartments was determined using a least squares minimisation, with the total volume of the three compartments becoming the FRC estimate. This was compared to plethysmograph FRC in 20 healthy non-smoking adults, with the method underestimating by $0.004 \pm 0.209l$. In ICU patients receiving controlled or assisted ventilation the reliability coefficient was $0.31l$. No analysis of the individual compartment volumes was performed.

The information content within the washout data rapidly diminishes as the test progresses. Methods have been proposed to overcome this, by either re-breathing air on random or pre-determined breaths. This increases the nitrogen concentration of compartments with a high ventilation-volume ratio, which are washed-out early in the test [118]. An alternative approach is to introduce a second tracer gas which is breathed at a constant concentration during certain breaths [23].

In reality the lung is not made up of a small number of discrete compartments, so theoretically at least it would be better modelled as a continuous distribution. The first attempt to recover a continuous ventilation-volume distribution was by Gomez *et al.* [85] who fitted a four parameter function for the continuous distribution. This did not provide convincing results, possibly because the function did not describe the

underlying distribution adequately [71]. For example it was not able to describe the bi- and tri-modal patterns observed by later researchers.

Another method, which used a direct inversion of the Laplace transform, was applied to experimental data by two authors [139, 147] but with limited success. A subsequent theoretical study [155] concluded that the method used was not suitable for inferring continuous distributions and was highly sensitive to experimental error.

Lewis *et al.* [117] introduced an alternative approach based on a fifty compartment model, similar to the Multiple Inert Gas Elimination Technique (MIGET) method of recovering ventilation-perfusion distribution (described below in section 2.4.5). When applied to human data this approach recovered a uni-modal distribution for healthy volunteers, which were broader for the normal subjects over 40 years old. Patients with pulmonary abnormalities displayed a variety of uni-, bi- or tri-modal distributions all with larger dead space compartments. This result was generally confirmed in a later more comprehensive study [105]. A range of distributions were seen in both the normal and cystic fibrosis groups, but the cystic fibrosis group showed a tendency to display wider distributions or distributions with more modes.

Mitchell *et al.* [137] investigated how the level of PEEP effected the ventilation-volume distributions of four post surgery cardiac patients. They used a similar method but with a oxygen tracer gas, accounting for oxygen consumption in a similar way to their FRC method [136], described above. Increasing the level of PEEP tended to increase the amount of normal ventilation observed, from 26% within the normal range at zero PEEP to 49% at 15 cmH_2O [137]. However there were some spurious modes of ventilation at high ventilation-volume ratios and the dead space fraction was lower than expected.

In a theoretical study Wagner [207] used a similar model and error-free simulated data from a fifty compartment model. Wagner showed that measurements from the first few breaths are critical to the recovery of compartments with a higher ventilation-volume ratio and that there was only minimal improvement from measuring more than 17 breaths. Once experimental error was added by Kapitan [107], it was concluded that only two modes of ventilation (plus the physiological dead space) could be resolved reliably.

Subsequent investigations using error free simulated data from a model with a continuous ventilation-volume distribution, showed that this method could locate modes of ventilation, but gave very little indication as to their width [216]. The ability to locate the modes was maintained once simulated experimental error was added. As the

width is the primary interest of this test, this does cast some doubt on the usefulness of the method.

Although the above methods can be useful in determining the condition of the lung, it is important to note that caution should be exercised when trying to directly relate the parameters recovered for a small number of discrete compartments or even the continuous distributions to the actual anatomical or physiological reality [71]. For example, two compartments being recovered using the 'exponential stripping' method does not necessarily indicate a bi-modal distribution. In fact analysis by Van Liew [204] has shown that the addition of up to twenty compartments with a uni-modal continuous distribution will result in two (or three) compartments when recovered in this way. It has also been shown for a model made up of any number of series and parallel compartments, when analysed using a washout test, an equivalent indistinguishable model can be found consisting of a collection of only parallel compartments [71, 72].

Many empirical indices have been proposed. Details of all these will not be provided here, but good summaries can be found in [71] and the online appendix to [167]. The most commonly applied index, however, is the Lung Clearance Index, which is defined as the number of FRC turnovers required to reduce the alveolar nitrogen concentration to a given fraction of the original concentration, usually between 1.5% and 2.5% of the original concentration [71, 167].

This index was shown to increase with age, although this result may be effected by concomitant effects of smoking in some studies [71], and to be raised in COPD patients [64]. A recent renewed interested in this measure has also shown it to be a useful method of determine the progression of airway disease, particular in cystic fibrosis patients [67, 100, 101] but also in other similar groups [86, 185]. This work used a compact side stream photoacoustic gas analyser, Innocor (Innovision ApS, Denmark). The LUFU system, described above, has also been used recently to identify changes in Lung Clearance Index in ventilated pigs [35]. Changes were observed when the level of PEEP was adjusted and when acute lung injury was induced.

The above discussion has focused on the use of nitrogen as the tracer gas, which is an obvious choice because oxygen is a readily available washout agent and also no pre-test wash in procedure is required as nitrogen is naturally present within the air [167]. However it does have disadvantages. Subjects' breathing patterns can be affected by consuming 100% oxygen, especially young children. Although nitrogen ($\lambda \approx 0.015$) defuses slowly out of the blood, the high concentrations within atmospheric air can cause large concentration gradients to be set up, potentially affecting results [167].

Therefore other gases can be used, which was the case for a number of the studies cited above [23, 86, 101, 139, 185]. Any non toxic and lowly soluble gas that does not participate in gas exchange is a possible candidate, such as argon ($\lambda \approx 0.003$), sulphur hexafluoride ($\lambda \approx 0.006$), helium ($\lambda \approx 0.010$) or methane ($\lambda \approx 0.014$) [167]. The difference in the value calculated for a particular index or volume may differ between gases for various reasons but primarily because the difference in molar mass effects the diffusion rates of the gas [167].

Almost all the work described in this section has made use of mass spectrometers. The exception to this is some of the recent work carried out using the Innocor device, the LUFU device and the Engstrom Carestation. The Innocor and LUFU are still side stream analysers and so still suffer from the delays and response time issues associated with these devices (discussed further below). The Engstrom Carestation uses a method for determining lung volume that is less affected by delays and response time issues, but it is unlikely to show similar performance when applied to other wash out techniques because of the approximations and averaging used. As of 2014 [102] none of the commercially available washout devices meet the recommendations of the consensus statement on washout testing [167].

2.4.2 Helium Dilution

In the section above, wash out techniques have been discussed. Similar wash in techniques are possible and have been included in some studies already cited [35, 152, 167], however these are generally far less studied than wash out techniques. The exception to this is helium dilution, a technique for determining FRC [44, 210].

Helium dilution is based on a technique developed about 200 years ago using hydrogen, which has been replaced for obvious safety reasons by helium. It is a closed system technique. A known volume of helium is added to the system to bring the concentration to about 10% by volume. The subject is then switched to breathe from the closed system until equilibrium is achieved. The time taken for equilibrium to occur will differ between subject. Thus the current recommendation is to wait until the concentration change is less than 0.02% for $30s$ [44, 210]. The volume of helium is then balanced between the start and end to determine the total volume of the lung.

The accuracy of the technique depends on a number of factors [44, 210]. For subjects that reach equilibrium quickly, it is possible to complete the procedure without the removal of carbon dioxide but in reality it needs to be removed by absorbers. In order to replace this carbon dioxide, oxygen is added to the system. Care has to be taken to add oxygen at the rate carbon dioxide is removed from the system or to account

for any difference in order to avoid errors. Helium is absorbed into the body during the process. The effect of this has been predicted to be about 100 ml for a normal procedure. However it varies between subjects, so due to lack of good evidence on how best to compensate, the current recommendation is not to make a correction.

The amount of nitrogen in the system is lower than in air, due to the oxygen added during the procedure, which causes nitrogen to be given up by the body. This error is, however, less than for helium and has the opposite effect and is often accounted for when dealing with any error in oxygen consumption. The point in the respiratory cycle that the subject is switched to breathing from the closed system affects the volume measured. Therefore care needs to be taken to switch at the correct point (usually the end of expiration) or to correct for errors in the switching point using a spirometry trace of the test. Errors caused by the respiratory quotient varying from unity are considered to be negligible.

Helium dilution is the most widely used technique for measuring FRC in spontaneously breathing subjects [141]. In addition it can be used on patients breathing higher concentrations of oxygen than the nitrogen wash out technique, up to 97% oxygen if required [141]. However it requires a skilled operator to perform [141] and, although much less cumbersome than plethysmography (disscussed below), it is still a reasonably bulky piece of equipment for easy bedside use. Therefore, although the method has been adapted for use with mechanical ventilation, it is still not widely used in such situations [141]. Like the nitrogen washout, it is not easily repeatable due to the time it takes for helium to wash out of the lung between tests and does not account for any trapped air within the lung.

2.4.3 Forcing Sinusoids

An alternative to washout and washin methods, which is essentially an analysis of the step response, is to analyse the response to a form of sinewave. An insoluble tracer gas can be used to determine information about the lung, but if information about the pulmonary blood flow is required a soluble gas needs to be used. Generally the use of a soluble gas relies on the fact that at certain forcing time periods the effects of recirculating tracer gas are negligible or at least if ignored do not significantly affect the parameters estimated. Therefore the key to this technique is to choose a forcing period of the sinusoid and solubility of the gas in order to minimise the recirculation effect, but still to be able to extract information about the parameters being estimated. In addition, the forcing period must not be too short and become aliased by the cyclical nature of breathing.

Zwart *et al.* [233, 234] pioneered this work using very low concentrations of halothane (0.02% by volume); a concentration at which it will not act as an anaesthetic agent. They showed analytically, using a continuous based lung model with no shunt, that the recirculation effects were negligible for forcing periods smaller than 5 minutes for halothane [233]. This resulted in a simplification to the solution for the continuous model, which results in the following relationship:

$$\frac{\mid F_A \mid}{\mid F_I \mid} = \left(1 + \lambda \frac{\dot{Q}_P}{\dot{V}_A}\right)^{-1} .(1 - \omega^2\tau^2)^{-\frac{1}{2}}, \tag{2.22}$$

where ω is the angular frequency of the sinsoidal forcing wave and τ is the time constant of the lung.

Two methods were proposed in order to determine the ventilation-perfusion ratio. The first [233] varied the inspired concentration of the halothane tracer gas by the sum of two sinusoids of different frequencies (0.6 mins and 3.0 mins). This allowed $\frac{|F_A|}{|F_I|}$ to be determined at both frequencies, using the assumption that the end-expired concentration is equal to the alveolar concentration. Two simultaneous equations can then be set up, eliminating τ and allowing the $\dot{Q}_P/\dot{V}_A$ to be determined. Carbon dioxide concentration measurements were then used to determine $\dot{V}_A$, using the continuous version of the Bohr Equation (Equation 2.1). This allowed the pulmonary blood flow ($\dot{Q}_P$) to be determined. This method was verified in exercising humans, with the pulmonary blood flow increasing with work load.

The second method [234] varies the inspired concentration by a sinusoid with a three minute period. The amplitude ratio and phase difference are measured and used to determine the $\dot{Q}_P/\dot{V}_A$, via an extension of the original model. The pulmonary blood flow can then be determined as before. This method was verified in pigs against results obtained using Fick's Principle [74] and found to underestimate, with this underestimate increasing with pulmonary blood flow.

In a 2000 study (also in pigs), results were compared with thermodilution [106]. It found good agreement between the trend of the results acquired using both techniques, even when lung injuries were induced, to simulate acute respiratory distress syndrome. However the offset between the two techniques varied between animals. In addition, it was shown that measurements made by this technique could be used to optimise PEEP ventilator settings.

The use of acetylene ($\lambda \approx 0.95$) and halothane ($\lambda \approx 2.3$) as tracer gases has been compared [122]. The lower solubility of acetylene will result in a larger recirculation effect. Simulated results predicted an overestimate of 0 to 10% in pulmonary blood

flow, due to this. Results from a small number (n=2) of anaesthetised dogs and an unknown number of humans at rest showed qualitative agreement with this prediction. In exercising humans, where the recirculation effect is more pronounced, a 7% underestimate was observed using acetylene. Therefore this study concluded that use of halothane was preferential to less soluble gases, like acetylene.

Brovko *et al.* [43] derived an electrical analogy of Zwart's model with shunt and physiological dead space added. This was adapted to allow the ventilation rate and pulmonary blood flow to vary over time. The method was verified in simulated data and in three cardiac patients, against the thermal dilution technique. Little information about these investigations was given, other than that the predictions were on average within 6% of thermal dilution.

Barton *et al.* [25, 26] developed Zwart's original model to include a dead space, inhomogeneous alveolar compartment and shunt. The model was solved in the forward direction using an analogue computer, which modelled compartments by a number of electrical components. Using this model it was possible to show several results. In particular, within Zwart's original model the change in pressure in the alveolar caused by the uptake/return of gases from the blood are ignored. Using this model it was possible to show that the errors caused by this simplification are minimal, if the mean of the forcing period is small or the frequency is high enough for the amplitude of the alveolar signal to be small.

The other significant change was the use of nitrous oxide as a tracer gas. The primary advantage of nitrous oxide is that it can be used in significantly higher concentrations (5 to 7% by volume) than the other gases discussed so far. The disadvantage is that a lower solubility ($\lambda \approx 0.47$) increases the recirculation effect.

Hahn *et al.* [91] simplified the model to a three compartment model (dead space, homogeneous alveolar compartment and shunt), allowing the inverse process to be applied. The method proposed was able to determine a subject's dead space volume, alveolar compartment volume and pulmonary blood flow using the amplitude of the mixed-expired sinewave in addition to the end-expired sinewave. This was primarily a graphical method based on the asymptotes of the end-expired/inspired and mixed-expired/inspired amplitude ratios plotted again forcing period on a logarithmic scale (a Bode Plot). Therefore data had to be collected at a series of forcing periods. A nitrous oxide tracer gas was used with Hahn *et al.* [91] theoretically showing that the effect of this on the parameter values recovered was negligible for forcing periods under 1 minute. The method was qualitatively tested in dogs.

The method was developed by Williams *et al.* [222, 224, 225] with the phase change also incorporated into the calculations. Two simultaneously inhaled tracer gases were also used. Argon was used to determine the dead space and alveolar volume using a single forcing period. The pulmonary blood flow was then calculated using nitrous oxide. The method was tested in a water-displacement bench lung and dogs [222, 224, 225]. Results in dogs compared favourably with those determined via alternative techniques [224].

Gavaghan and Hahn [83] carried out a theoretical study into four tracer gases: argon, nitrous oxide, acetylene and halothane, using a 'standard man' continuous homogeneous model of the lung. A number of previous assumptions were questioned, in particular Zwart's [122, 233, 234] assumption that longer time periods and more soluble gases were better, as this reduced the recirculation effect. It was pointed out that although these factors reduce systematic error due to recirculation, the information content within the alveolar signal is also reduced and therefore parameter recovery is more prone to experimental error.

The paper also proposed the use of several forcing periods (0.5, 1.0, 1.5, 2.0 and 3.0 *mins*) in order to allow parameters to be estimated using a single tracer gas and to make the process more robust to experimental noise. The parameters were estimated by minimising the least squares error of two amplitude ratios: the end-expired/inspired and mixed-expired/inspired amplitude ratios. The phase difference was not included as part of the estimation technique.

It showed that in the absence of noise there is theoretically little difference between the three soluble gases (nitrous oxide, acetylene and halothane) when recovering parameters. When experimental error is added the recovery of pulmonary blood flow is similar for all three gases but the recovery of the alveolar volume was significantly more robust when using nitrous oxide. However if only ventilation related parameters are required argon performed best.

An alternative to using inert tracer gases, which are not involved in respiration, is to use oxygen. This was first proposed in a pair of simultaneously published papers [89, 226]. The theoretical paper [89] showed that oxygen acted almost identically to a lowly soluble gas like argon. This may seem counter intuitive, as the purpose of the lungs is to transport oxygen into the blood, but oxygen has almost the same solubility in blood as argon. It is haemoglobin that primarily carries the oxygen within blood. This means that the attenuation of the alveolar sinewave amplitude is virtually unaffected

by uptake into the blood even in hypoxic conditions. However the offset in mean due to the consumption of oxygen does need to be taken into account.

The theoretical results were confirmed, in dogs [226, 227], with attenuation of argon and oxygen alveolar amplitudes being identical. The theoretical predictions for concentrations within the blood were also confirmed, with oscillations in mixed-arterial blood tension observable in all conditions, but only observable in the mixed-venous blood in the presents of hypoxia. In humans the dead space volumes recovered compared favourably with the those determined using the single-breath Bohr Equation (Equation 2.2) [228]. FRC (dead space plus alveolar volume) predictions were comparable to those determined via plethysmography and nitrogen washout; although with a constant overestimate in the second case.

The obvious advantage of using oxygen is that it is already in equilibrium throughout the body so an initial stage to bring the body into equilibrium is not required. The other advantage cited by the initial theoretical paper [89] is that smaller, reliable and more readily available oxygen sensors can be used instead of a mass spectrometer. However all three experimental papers [226–228] used a mass spectrometer.

The use of a tidal ventilation model was introduced by Gavaghan and Hahn [84]. The model provided a continuous time varying signal for the alveolar concentration but not for the expired signal at the mouth. Instead it provides two discrete output values for each breath: the mixed-expired concentration and the end-expired concentration. The end-expired concentration was assumed to be equal to the alveolar concentration. This is enough information to use the continuous model recovery method described in their previous paper [83].

The continuous method was compared to a similar recovery approach using the new tidal model. Recovery of the pulmonary blood flow using the continuous model was robust with tidal breathing only having minimal influence on the recovered value, even in the presence of simulated noise. However the recovery of the dead space and alveolar volumes was influenced by tidal breathing, particularly when the lung varied from a 'normal' lung to one which is closer to those observed in a diseased state. The most robust estimates were made using an insoluble tracer gas.

A later experimental study [229], in dogs, showed contradictory results. The pulmonary blood flow recovered using the tidal model method was in much closer agreement with the cardiac output determined by thermodilution than that determined using the continuous model. Here, however, the tidal model did include recirculation effects from a single compartment unlike Gavaghan and Hahn's [84] recovery process.

The compartment's size and perfusion rate were matched to the compartment with the smallest time constant, Mapleson's viscera/vessel rich group [126, 127].

Sainsbury *et al.* [173] investigated the differences between the alveolar volumes determined via the continuous and tidal model. They proposed that an adjustment factor of $(V_T + V_D)/2$ be subtracted from the volume recovered using the continuous model. This was determined with the use of a number of assumptions and verified in experiments using a water-displacement bench lung. A more rigorous derivation of this correction factor was provided later [216].

Zwart *et al.*'s amplitude and phase method [234], based on a continuous model, and Gavaghan and Hahn's tidal ventilation method [84] have also been verified against simulated data from the trumpet model [220]. The tracer gases argon, nitrous oxide and halothane were considered. This showed that the dead space volume recovered was dependent on the breathing pattern and structure of the lung; but was not dependent on the recovery method nor the pulmonary blood flow recovered. The alveolar volume was recovered more robustly when using the tidal ventilation model. The recovered pulmonary blood flow was highly dependent on the method used, tracer gas and forcing period, but all methods identified the trend of changes in pulmonary blood flow.

Argon was the best gas for determining the dead space and alveolar volumes. Pulmonary blood flow is best determined using nitrous oxide when the tidal model is used, as halothane grossly underestimates this value. However when the continuous model is used halothane is better, as parameters recovered with nitrous oxide vary with forcing period. Another finding was a significant difference in the end-alveolar and end-expired concentrations when using soluble gases. The differences were more pronounced when the inspired flow was not fixed and for halothane. This was due to the higher solubility and lower diffusion coefficient of halothane.

Limited effort has been made to infer information about the ventilation-volume distribution. Zwart extended his model from a single compartment model to a multi-compartment model [216]. This was done using a similar approach to Wagner for MIGET (see section 2.4.5), but with seven instead of fifty compartments. The use of several tracer gases was proposed, in order to determine a measure for the range of ventilation-perfusion ratios across the lung. However the method was unstable and too prone to experimental error for practical use.

Whiteley *et al.* [216, 218] used simulated data from a lung with a continuous log-normal ventilation-volume and ventilation-perfusion ratios to show that recovery of physiological parameters using a homogeneous model was prone to systematic error.

This varied predictably with the forcing period leading to the proposal of an empirical index, I:

$$I = \frac{V'_A(5) - V'_A(1)}{V'_A(1)} \ , \tag{2.23}$$

where $V'_A(1)$ and $V'_A(5)$ are the alveolar volumes recovered using the homogeneous model at forcing periods of 1 and 5 minutes, respectively. This would then give an indication of the inhomogeneity without directly inferring the actual distributions. This index was found to be fairly robust to simulated noise on the concentration signal.

2.4.4 Breath-by-Breath Analysis

Determining information about the alveolar volume and pulmonary blood flow using breath-by-breath analysis was first proposed by Farmery [77]. The volume of the dead space is determined independently from the the other parameters and then subsequently used to determine other lung parameters. Two methods of determining the dead space are those based on the Bohr Equation (see section 2.3.1) and Fowler's wash-out method (see section 2.4.1). Both have been modified to allow the inspired concentration of the tracer gas to be non-zero.

In Section 2.3.8 of Farmery [77] the following equation is derived by balancing the mass of the tracer gas over a breath:

$$V'_A.F_{A,n} = V'_A.F_{A,n-1} + V_D.F_{A,n-1} + (V_{T,n} - V_D).F_{I,n} - V_{T,n}.F_{A,n} - \dot{Q}_P \lambda (F_{A,n} - F_{\bar{v},n}) \Delta t_n \ , \tag{2.24}$$

where $F_{\bar{v},n}$ and Δt_n are the mixed venous fractional gas concentration and duration of breath n; i.e. the tracer gas in the alveolar compartment at the end of a breath ($V'_A.F_{A,n}$) is the tracer at the end of the previous breath ($V'_A.F_{A,n-1}$), plus that re-inhaled from the dead space ($V_D.F_{A,n-1}$), plus that which reaches the alveolar compartment from the mouth ($(V_{T,n} - V_D).F_{I,n}$), minus that exhaled ($V_{T,n}.F_{A,n}$), minus that taken up by the blood ($\dot{Q}_P \lambda (F_{A,n} - F_{\bar{v},i}) \Delta t_n$).

$F_{A,n}$ is assumed to be equal to the end-expired fractional concentration, as assumed by many other techniques already described. V_D can be determined as described above. $V_{T,n}$ and $F_{I,n}$ can be determined directly from the data. For a insoluble gas the final term ($\dot{Q}_P \lambda (F_{A,n} - F_{\bar{v},i}) \Delta t_n$) can be ignored. Thus leaving V'_A as the only unknown, which in theory can be determined from two breaths' worth of data. In reality, however, the use of least squared regression over a series of breaths provides a much more reliable prediction. This method can be applied to any form of perturbation function. Farmery [77] experimented with step functions (wash-in and wash-out) and sinusoid functions in a water-displacement bench lung.

When soluble gas is used the final term can no longer be ignored. In these cases the equation has been solved using a sinusoidal perturbation function, allowing $F_{\bar{v},i}$ to be assumed to be equal to the mean of the forcing sinusoid [57, 58, 77]. This is the standard assumption made by many sinusoidal based techniques (see section 2.4.3 above), which has been shown to induce little error at certain forcing periods. This is assumed to hold true for this technique, but has not been specifically examined.

By including these assumption two unknowns remain within Equation 2.24: V'_A and $\dot{Q}_P$. These could be determined using two simultaneous equations formed from data for three breaths. In reality data from more breaths are used. There is an implicit assumption here that the concentration in the alveolar compartment is constant during expiration. The validity of this is not addressed in the literature [57, 58, 77].

Experiments in small sheep by Farmery [77] did not provide convincing quantitative results, but there did appear to be some qualitative relationship with values determined via thermodilution. Recently more favourable results have been reported [57, 58]. In a study that included limited data from an artificial bench lung, it was shown that it was possible to determine the dead space volume, alveolar compartment volume and to show that no blood flow was present [57]. However as blood flow is not present in the artificial bench lung nor is the systematic error due to recirculation. Another study compared predictions to those made by a continuous model in human subjects (n=22) [58]. Comparable results were obtained although with pulmonary blood flow predictions on average 0.35 l/min above that determined using the continuous model. All these results were obtained using a small main-stream infra-red sensor, not a mass spectrometer.

2.4.5 Multiple Inert Gas Elimination Technique (MIGET)

The MIGET involves the infusion of six inert gases of varying solubility (λ from 0.005 to 300) into the blood stream for about half an hour [98]. Several measurements are then taken [215]: (a) concentration of the inert gases within the expired breath, the arterial blood and the mixed venous blood; (b) cardiac output; (c) ventilation rate.

There are several ways in which the measurements can be used to determine information about the ventilation-perfusion distribution. The most widely used and relevant here was a method originally described by Wagner *et al.* [206]. They described that for a multi-compartment model:

$$\frac{P_{\bar{a},j}}{P_{\bar{v},j}} = \frac{1}{\dot{Q}_P}.\sum_{i=1}^{n}\left[\frac{\dot{Q}_i.\lambda_j}{\lambda_j + \frac{\dot{V}_i}{\dot{Q}_i}}\right] \tag{2.25}$$

$$\frac{P_{A,j}}{P_{\bar{v},j}} = \frac{1}{\dot{V}_A} . \sum_{i=1}^{n} \left[\frac{\dot{V}_i . \lambda_j}{\lambda_j + \frac{\dot{V}_i}{\dot{Q}_i}} \right] , \tag{2.26}$$

where $P_{A,j}$, $P_{\bar{a},j}$ and $P_{\bar{v},j}$ are the alveolar, mixed arterial and mixed venous partial pressures of gas j, respectively and λ_j is the Oswald blood-gas partition coefficient of gas j. $\dot{Q}_i$ and $\dot{V}_i$ are pulmonary blood flow rate and ventilation of lung compartment i, respectively, which are the only unknowns within these equations.

A multi-compartment model with fifty compartments, separated on a logarithmic scale across the range of perfusion-ventilation ratios, was then used. Compartments with zero and infinite perfusion-ventilation ratios were included to represent shunt and dead space, respectively. A gradient based optimisation technique was then used to minimize the sum of squares error between the solutions of Equation 2.25 and the measured retention values, by varying $\dot{Q}_i$ for each compartment. This was then repeated with Eq. 2.26, by varying $\dot{V}_i$. The optimisation was constrained to ensure $\dot{Q}_i$ and $\dot{V}_i$ are positive. The result is a fifty compartment distribution of ventilation-perfusion ratios, like those shown in Figure 2.5.

Some controversy arose around the validity of this method because it seemed unlikely that unique information about fifty compartments could be obtained from just six gases [104, 149]. To partly answer this criticism Evans and Wagner [73] produced a more mathematically robust analysis of the technique, proposing a ridge regression method. Ridge regression is a standard mathematical approach for ill posed problems, which effectively required compartments of similar $\dot{V}_i/\dot{Q}_i$ to have similar $\dot{Q}_i$ in this case. In addition, a method of combining the retention and excretion values was proposed in order to allow the estimate for perfusion and ventilation of each compartment to be determined simultaneously.

This method has been applied to both healthy subjects and those with pulmonary disease [215]. Healthy subjects showed a consistent broadening of the distributions with age. Patients with lung disease show a variety of different distributions, often with multiple peaks. Patients with the same condition do not necessarily show similar distributions. For example the distribution seen in Figure 2.5a is from a COPD patient, where the disease appears to be due to poorly perfused areas of lung, while for the COPD patient in Figure 2.5b it appears to be due to poorly ventilated areas.

There are several limitations to the MIGET technique, which have meant that the technique has not been used in clinical practice [208]. It is an invasive procedure which is relatively complex to perform, requiring specialised personnel. It requires a reasonable amount of time to perform and so cannot realistically be used for continuous

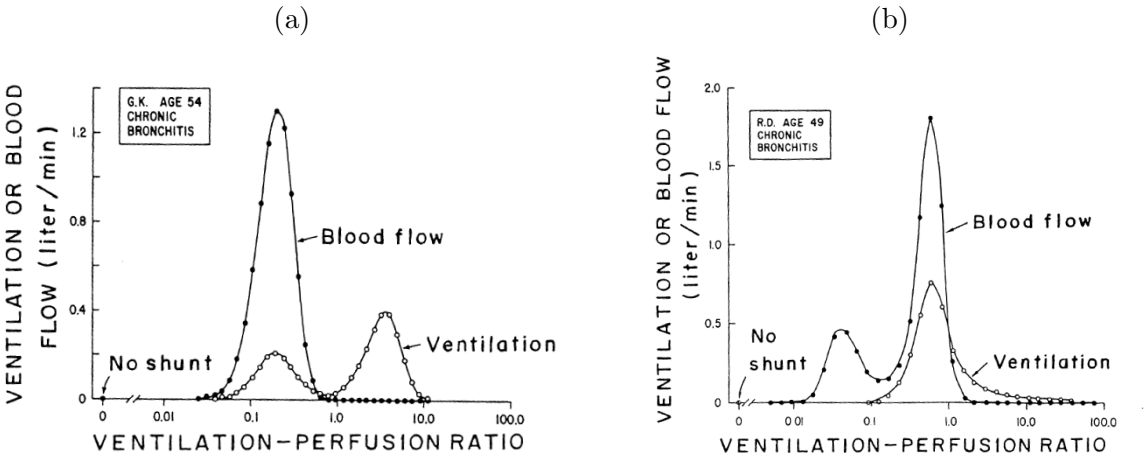


Figure 2.5: Distributions recovered from two different COPD patients using the MIGET. Figures taken from [215].

monitoring of a patient's lung function. Also the physiological dead space region is evaluated from an extrapolation, which is largely dependent on measurements for gases with high λ [98]. These gases are at very low concentrations within the blood and so sensitive to errors. A similar effect occurs when extrapolating to determine shunt.

2.5 Estimating Lung Parameters by Other Means

There are numerous alternative methods of determining information about lung function and pulmonary blood flow. However only additional methods for determining lung volume and ventilation-volume distributions which are in wide spread clinical use will be discussed in detail here. Respiratory physiology text books [124, 144] are a good source for further information on other techniques, such as those for determining pulmonary blood flow.

In addition to nitrogen wash out and helium dilution, there are two widely accepted methods of determining lung volumes that include the RV: plethysmography and imaging techniques. These four techniques are those recommended by the American Thoracic Society (ATS)/European Respiratory Society (ERS) Task Force on pulmonary function standards [55], which were based on the National Heart, Lung and Blood Institute (NHLBI) workshop on the subject [44, 54, 59, 141]. None of the the associated publications recommend one technique over the others. In addition, nuclear medicine based imaging techniques for determining information about the lungs ventilation-volume distribution will be briefly detailed.

There are two other techniques which are worth mentioning, but are still at the research stage and so will not be covered in detail. Hyperpolarised gas magnetic resonance imaging (MRI), which shows promise in determining information about

ventilation-volume and ventilation-perfusion distributions [75, 76]. However it suffers from the general disadvantages of all imaging techniques (discussed below) and is also currently confined to the research environment; no hyperpolarised gas has been approved for general clinical use by the United States Food and Drug Administration [9, 11–13, 15, 17, 19, 75, 111]. Electric impedance tomography allows low spatial resolution images of the lung to be presented at the bedside and can allow FRC and information about the ventilation-volume distribution to be determined [108, 116]. It has been promoted as a potential method for monitoring the health of ventilated patients' lungs, but it is yet to see any routine clinical use and studies show mixed results for its potential benefits [52].

2.5.1 Plethysmography

Plethysmography involves measuring the pressure changes at the same time as volume change [59, 63, 210]. In order to do this subjects are placed in a box about the size of a telephone box and breathe through a tube. A shutter is then closed to occlude the tube and the subject is asked to breathe against it. This causes the pressure in the lungs to increase at the same time as the volume decreases. The pressure change in the lungs is measured by a sensor at the mouth, while the volume change in the box is measured to determine change in volume of the lung. Boyle's law for ideal gases is then used to determine the volume of the lung from changes in the pressure and volume. Correction for the water vapour present is required, as it does not act as an ideal gas with its volume only dependent on temperature not pressure.

In reality a single measurement is not made, instead subjects are asked to slowly pant against the shutter, at a frequency of about $1Hz$, for three to five breaths [59, 63, 210]. This allows a series of data points to be built up to which a straight line is fitted, usually manually by the operator. In addition, it is recommended that the complete procedure is repeated until at least three tests agree within 5%. Thermal drift can occur due to changes in temperature within the box, usually due to the subject warming the air within the box. Assuming the drift is not too large and constant, correction can be made for this.

The plethysmograph is often been referred to as the 'gold standard' for determining the total lung volumes [88, 141]. However there has been some doubt in the literature that it is any more reliable than other techniques, such as helium dilution [44, 141, 146]. Without a true 'gold standard' it is hard to compare techniques. However there are several factors which are known to lead to errors when using plethysmography. The pressure at the shutter is assumed to be equal to that throughout the lung, which is

true if there is no flow in the lungs. However during the panting procedure differences in pressure can occur, due to the ability of lung tissues to stretch, particularly upper airways and cheeks, causing a capacitive effect [44, 59]. This is particularly significant in subjects with high airway resistance, such as asthma sufferers. Ensuring a slow panting motion of $1Hz$ can reduce any errors, with differences as much as $2l$ being reported between $1Hz$ and the natural panting frequency of about 2 to $4Hz$.

Any trapped air in the thorax that is able to change volume will be included in the predicted volume [44, 59]. This means air trapped behind obstructions within the lung is included, unlike in the nitrogen wash out or helium dilution techniques. However so is other air trapped within the abdomen, such as that within the digestive tract. Conversely air trapped within areas of lung not compliant enough to allow them to expand and contract as required may not be accounted for. This has been suggested as a possible source for lower than expected lung volume values determined by plethysmography in some infants [59].

Overestimates made by plethysmography (or underestimates by other techniques) have been widely reported since its early use. Studies comparing plethysmography to helium dilution [44, 56, 131, 146, 168], CT imaging techniques [82, 146] and other techniques [56, 131] have shown an overestimate. A possible explanation in early studies was the lack of adherence to low panting frequencies. Once this was recognised as a possible source of error, procedures were put in place to ensure panting frequencies of less than $1Hz$ in future studies. Trapped gas within the thoracic cavity, either within the lung or not, was then offered as an explanation.

In 2010 O'Donnell *et al.* [146] suggested that this may not be the source of the error when they compared plethysmography, helium dilution and CT imaging. This was a extensive study carried out on 132 patients, 79 of which had indication of a obstruction, across three hospitals. Little offset was found between helium dilution and CT but similar offsets when compared to plethysmography, suggesting that the offset was due to other sources than trapped air. This paper has, however, attracted some criticism [186, 190] mainly for assuming that CT imaging is a proven method: criticism the authors responded to by pointing out its wide spread use for many decades and that all three techniques provided similar results in non-obstructed patients.

Plethysmography requires the subject to be enclosed within a sealed box, therefore it is not suited to bed side use. Also an amount of cooperation is required when using the conventional panting method. Therefore it cannot be used for ventilated patients, young children or those who are particularly unwell [59, 141]. Alternative techniques

have been devised for young children, where the measurements are done during natural sleep (up to about 1 month old) or during sedation [59]. However they are only used in a few specialised centres using non-standardised 'homemade' equipment. This limits the use of plethysmography and excludes routine use in ventilated patients.

2.5.2 Determining Lung Volume from Images

It is possible to determine the volume of the lung from conventional radiographs, CT and MRI [54, 210]. The state of inflation has to be maintained at a known point for the duration of image acquisition, with TLC generally being determined. Several techniques exist for conventional radiography [54, 210] which rely on making measurements of a number of dimensions or areas within two planar chest images (anteroposterior and lateral). These are translated into volume by assuming particular geometries or by empirically relating them to volumes determined by other means.

Comparison of the volume determined by these methods and plethysmography have shown high correlation coefficient, however, as pointed out by Bland and Altman [37], this does not necessarily mean that the techniques are equivalent [54]. Further investigations revealed that the standard error for the estimate for regression was between 0.28 and 0.51l, depending on the method used, which meant between 6 and 25% of measurements differed by more than 10%. Inter-observer reproducibility was also reported to be high for volumes determined from the same images.

3D CT and MRI imaging are relatively new ways of determining lung volume when compared to some of the more established techniques. They provide the ability to determine the individual geometries of subject's lungs in more detail than 2D techniques. These methods usually rely on automatic segmentation techniques to define a boundary between the thoracic cage and surrounding tissue [54]. The volume inside is then computed. Exact details of how this is done are normally limited and likely to differ between proprietary software [54]. It is also likely that measurements may differ between scanners of differing quality and type.

The ATS/ERS Task Force [210] concluded that there is insignificant evidence to support one imaging technique over the others and commented that "the larger standard deviation of radiographic measurements may limit their clinical usefulness" [210]. Their recommendation was based on the NHLBI workshop publication [54]. This cited a single CT comparative study in 21 children comparing two conventional radiography techniques and CT with results from plethysmography as ground truth [174]. CT did not out perform the conventional radiography technique in this study.

Since this recommendation a number of other CT studies have been reported. Chiumello *et al.* [53] studied 30 ICU ventilated patients without chronic lung disease. It assumed CT to be the ground truth and compared it to the FRC determined using the Engstrom Carestation and helium dilution. The overestimate of CT was -0.094±0.143 and 0.136±0.133l, respectively. O'Donnell *et al.'s* [146] extensive study has already been discussed. They mention that for non-obstructive subjects there was not significant difference between TLC determined by CT, helium dilution and plethysmograph, but no detailed results were given. However when all data were included the Bland-Altman plots showed a reasonable amount of scatter, with a standard deviation of about 0.8l (estimated from the plot).

Another study [82] compared high resolution CT to plethysmography in COPD patients (n= 59), reporting significant variation when individual results were compared on a Bland-Altman plot with a mean difference of 1.12l and a standard deviation of about 0.6l (estimated from the plot). A recent study [88] concentrated on repeatability comparing the lung volume of prostate cancer patients, without any lung disease, at 3 mouth intervals. The median individual standard deviation was found to be 0.85l, which results in a 16% coefficient of variation. The Chiumello *et al.* [53] and O'Donnell *et al.* [146] studies may suggest that for subjects without obstruction CT is in a closer agreement with other widely accepted method than previously thought. However, taken as a whole these studies probably do not provide sufficient evidence to break from the position of the ATS/ERS Task Force.

In addition to these issues, imaging techniques also have other limitations. They rely on well trained radiographers to be performed. CT and MRI can not be performed at the bedside and can not readily be repeated due to the high radiation dose and cost, respectively. However routine chest radiography is still the most common method of assessing lung volume, either through conscious or unconscious qualitative assessments or via the quantitative methods described above [55].

2.5.3 Determining Ventilation Information from Images

By far the most routine way of determining information about the distribution of ventilation within the lung is nuclear medicine. Nuclear medicine can be thought of as radiology done with a source of radiation from inside the body, rather than an external source. Ventilation imaging involves the inhalation of a radio active gas, such as Xenon-133, or radioactive labelled aerosols which are then imaged externally [132, 203]. Gases and aerosols of different sizes have different properties which can be used as an advantage in particular diagnostic tests. However noble gases have the

advantage that they distribute uniformly throughout the lung without being affected by the state of the bronchial mucosa and so have been most widely studied.

Ventilation images on their own have been used to assess regional ventilation, in order to detect early air flow obstructions, and in the assessment of small airway function [132]. However they are most widely applied in conjunction with separately acquired perfusion images, which are obtained by injecting radio active particles that get stuck in the lung's capillaries [203]. This allows for a qualitative assessment of the ventilation-perfusion ratio throughout the lung. The main application of this is in the diagnosis of pulmonary embolism, but it has been used in the assessment of other diseases [132, 203]. Most commonly 2D images are used but it is also possible to obtain 3D images. These methods have similar limitations as radiography, as discussed above.

2.6 Signal Alignment and Sensor Response Time

Many of the techniques of determining lung parameters from tracer gases rely on the measurement of the total flow rate of all the gases into the lung and the fractional concentration of the particular tracer gas. It has been recognised for some time that the alignment of the flow and concentration signals is critical to the success of these methods [22, 27, 47, 102, 135, 167, 184].

Generally the concentration signal lags behind the flow signal, especially when the concentration measurements are taken via a sampling tube, which it takes a finite time for the gas to flow down [27, 47, 102]. This effect can be seen in Figure 2.6. In addition, when a parameter changes state it takes a finite amount of time for a sensor to reflect that change; this time is known as the response time. The response time of flow sensors is generally significantly quicker than that of concentration sensors [46, 78, 102, 230]. The typical effect of the concentration sensor's response time can also be seen in Figure 2.6.

The alignment of the signals and the response time of the sensors is important because many of the techniques for determining lung parameters rely on numerically integrating the product of the flow and concentration signals. If the flow rate is constant and a square wave change in concentration is observed, then any delay in the concentration sensor changing state (either due to delay or the response of the sensor) will not result in error as the lag in the rise and fall will cancel each other out [22]. However flow is not constant during normal tidal breathing and so errors occur. In fact Farmery [77] showed that if large step changes in concentration occur at points of increased flow, as happens during the beginning of expiration in many techniques, these errors are exacerbated.

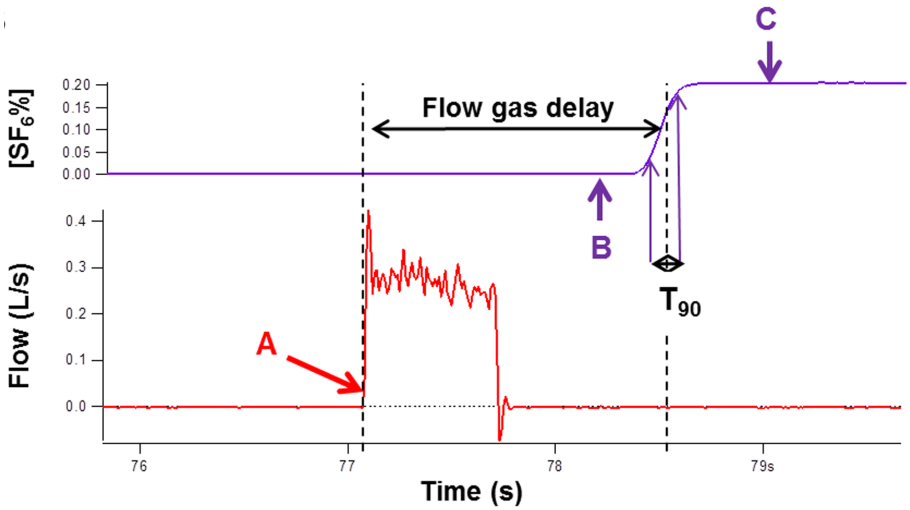


Figure 2.6: Example response of a concentration and flow sensor to a square wave input of 0.2% of Sulphur Hexafluoride (SF_6) gas. The traces are recorded using Innocor photo-acoustic spectrophotometer gas analyser via a sampling line. The offset between the two signals is labelled as the 'Flow gas delay' and in this case is measured from the initial up spike in the flow signal (labelled A) to the 50% rise point of the concentration signal. The 10% to 90% response time of the concentration sensor is labelled T_{90}. Taken from [102].

Errors that occur due to the simple offset in time, as opposed to the response time of the concentration sensor, could easily be adjusted for by simply advancing the concentration signal in time if the time delay is known. However determining the delay precisely is not possible without knowledge of the real concentration signal, unaffected by the response time of the sensor, as the real end point for measuring the delay is unknown. Many studies that advance the concentration signal do not detail how the end point is determined [22]. One simple approach is to take the point at which the concentration passes through 50% of the final concentration in response to a step function [102], as shown in Figure 2.6. Another, when the response is not symmetrical about the 50% point, is to match areas above and below the concentration curve [46, 188]. In reality, these methods also partly correct for the response time too (as detailed further below). A potential solution to this is to include the delay as an additional parameter to be fitted during calibration for response time [27].

To complicate matters further, when collecting measurements via a sampling tube, the rate at which the gas travels through the tube varies depending on the composition of the gas and the pressure gradient through the tube. Brunner *et al.* [47] investigated the effect of the nitrogen concentration on the delay when performing a nitrogen wash out. They showed that dynamically correcting the delay in accordance with the

gas composition improved the dead space and alveolar volumes recovered from water-displacement bench lung data. Ideally the flow rate through a sampling tube should be independent of airway pressure which is generally the case in mass-spectrometers due to the specification and set up of the sampling pumps but for equipment with lower specification pumps this can become an issue [46]. For example dynamic correction for the time delay through the sampling tube dependent on the pressure gradient is required in the case of the LUFU system [95, 213].

There are several methods of correcting for the response time of the sensor. In main stream analysers only the inherent response time of the sensor needs to be taken into account. In side stream analysers, which have a sampling tube that the gas travels through before being sensed at a remote point, both the sensor response time and that caused by dispersion though the sampling tube need to be taken into account [27, 46, 78]. The sensor response to a step function is normally characterised as an exponential [22, 46, 135] while the combination of the two affects in side stream methods has been characterised as a sigmoid [22, 46, 78].

One simple method of compensating for the response time is to advance the concentration sensor further than would be required by the non-alignment of the two signals [27, 77]. The amount of advancement is determined empirically during a form of calibration. This may be done by minimising the difference between the true and measured signal [27], or by minimising the error in the parameter being predicted [77].

Even though this is effective it does not take into account the shape of the sensor's impulse response; methods that do this were first developed in the late 1970's. The initial work concentrated on side stream mass spectrometers with exponential models of the response developed, despite the response being sigmoidal. A first-order exponential correction method was described by Mitchell [135] and Noguchi *et al.* [143]:

$$F^*(t) = F(t) + \tau_1.\dot{F}(t) \ , \tag{2.27}$$

where F is the concentration measured by the sensor, $\dot{F}$ is the first derivative of F with respect to time, F^* is the corrected signal and τ_1 a time constant, which is selected during a calibration step. This improves the the response rate of the signal, however a second-order exponential model:

$$F^*(t) = F(t) + (\tau_1 + \tau_2).\dot{F}(t) + \tau_1.\tau_2.\ddot{F}(t) \ , \tag{2.28}$$

was shown to provide better correction [22]. $\ddot{F}$ is the second derivative and τ_2 is the second time constant selected during calibration.

An empirical calibration method was initially proposed for the second order method, that fitted the two time constants and the delay in turn [22]. Bates *et al.* [27] subsequently showed that a least-squares fit of the three parameters (τ_1, τ_2 and the delay) provided better correction; however the corrected signal exhibits more of an overshoot [27, 78]. If a short line segment is used in the numerical differentiation within these methods, then the noise in the original signal will cause a noisy oscillating output. To avoid this a longer segment length is used, without extending the segment too much and so losing too much signal detail [22, 27]. The optimum length will depend on the nature of the data but Bates *et al.* [27] found that $20ms$ provided acceptable results for their apparatus.

An alternative, more sophisticated, method using a Wiener filter, which is based on Fourier transforms and routinely used in signal processing, was described by Bates *et al.* [27]. It performed better than the second-order exponential model even when the parameters were set using their least squares method. However this method is more computationally expensive to perform and has more of an overshoot than the second-order exponential model with the parameters determined via the Arieli and Van Liew method [22].

These second-order methods have more recently been applied to newer more portable clinical analysers [230], in particular to a side stream infra-red absorption analyser, the Datex Ultima. The corrected signal improved the response time and brought it into line with the response usually observed by a mass spectrometer. Farmery and Hahn [78] proposed an alternative model based on a Gompertz Sigmoid, which is first-order. This means the method is potentially less sensitive to noise and has one fewer parameter to be determined during calibration. When applied to a signal from the same analyser, it performed better than the second-order exponential model. However even with this improvement the response time is still only similar to the response time of a mass spectrometer.

The importance of the alignment of the signals and the response time of the concentration sensor has received increased interest in recent years with renewed interest in the washout methods, as detailed above. A 2013 consensus statement [167] recognised the importance of both factors when determining lung parameters accurately. It recommended that alignment of the signals should be within $10ms$ with a concentration sensor 10% to 90% rise time of less than $100ms$. As mentioned above, as of 2014 none of the commercially available Multiple Breath Inert Washout devices meet the requirements of this statement, many because they do not meet the response time

requirements [102]. In addition, the 'gold standard' systems used in clinical studies still rely on mass spectrometers.

2.7 Estimating Model Parameters

Thus far parameter estimation techniques specific to the lungs have been covered; however parameter estimation is a common problem that has concerned scientists, engineers and medics for many years. There are many methods for making parameter estimations for both linear and non-linear systems, therefore only those methods with direct relevance to this thesis will be discussed here. Justification for using these particular methods is given in Section 3.4.2. If more information is required on parameter estimation techniques in general there are many good texts on the subject [32, 36, 115, 123, 183].

2.7.1 Non-Linear Least Squares Optimisation

Since the early 1800s the use of the least squares method has been a standard approach for determining estimates of model parameters. The method relies on minimising the sum of the squared residuals. In simple linear cases this minimisation can be done analytically, however in more complex non-linear cases an iterative optimisation approach is required. There are many methods of optimisation which generally follow the same approach. The value of an object function (in this case the sum of the squared residuals) is determined for a selected starting point and a number of points in parameter space round that point. The various algorithms then use a process of iterating towards a point with minimum (or in some cases maximum) value.

Optimisation methods can be categorised into derivative based methods, which require a smooth first or higher derivatives of the object function and non-derivative based methods. Derivative based methods are generally more computationally efficient and provide clear methods of determining convergence, but they do have drawbacks [232]. They are generally intolerant to discontinuities or inaccuracies in the derivatives caused by noisy object functions or other means. Also they are only suited to finding local minima within a continuous parameter space.

Non-derivative methods can be categorised into those which will find local minima and those which attempt to find global minima. Those which find local minima will find minima close to the starting point but this may not be the overall minima within the parameter space. Global techniques will, given sufficient time and computational power, find the global minima, although in reality they may not due to computational

constraints. The general disadvantage of global methods is that they are computationally less efficient, as they have to evaluate the object function at more points, and they can be slow to converge, especially near the global minima [232].

Within this thesis two non-derivative based optimisation methods are used. The first is the Nelder-Mead Simplex method [140], which is a local minima method. When performed in two dimensional parameter space a triangle is formed. The object function is evaluated at each of the three points of the triangle. Then the processes of reflection, expansion and contraction are preformed in turn, as required. This results in a new triangle. Over a number of iterations the triangle progresses towards the minima until convergence criteria are met and the algorithm is terminated. The method can be generalised to be performed on any number of dimensions.

This method has generally been found to be quite efficient [32] and has been widely implemented [113]. Although there are extensions available, the basic algorithm does not allow the searchable parameter space to be constrained and will only find local minima [32]. It is also not suited to discrete parameters.

The second method is a genetic algorithm, which is based on the process of genetic reproduction relying on the concept of 'the survival of the fittest' [32]. The parameters can be discrete or a bounded continuous variable. These are coded into a single binary number to create members of the population. Initially a random population is created and the object function corresponding to each member is evaluated. A mating pool is then created via a random selection process that favours the 'fitter' members, those with a lower object function value. Two of the mating pool, the parents, are then 'mated' via a crossover operation to produce two children. This operation randomly selects a set of bits that represent a parameter to be swapped between the parents, creating children with 'genes' from both parents. It is possible that mating pool selection and crossover can result in similar or identical members which are not necessarily the 'fittest'. Therefore a mutation step is carried out within which the bits, which represent the 'genes' of each child, are randomly swapped. The probability of a swap is chosen to be relatively low, normally in the range 0.005 to 0.01 [32].

Genetic algorithms are well suited to discrete and bounded continuous parameters. Given a sufficient population size and computational power they can find global minima. Also the algorithm naturally lends itself to a parallel processing approach.

2.7.2 Approximate Bayesian Computation (ABC)

Bayes' theorem allows prior information about a system to be updated based on the evidence presented. The challenge in Bayesian Inference comes from the fact that

analytical forms of the posterior are not generally available, so an approximate solution has to suffice [24].

The methods of finding an approximate solution generally fall into two categories: deterministic or stochastic [36]. Deterministic methods rely on analytical approximations, such as assuming a specific form and so inherently will always produce an approximation. In general they are computationally less intensive. On the other hand, stochastic methods, such as Markov Chain Monte Carlo (MCMC), will produce the exact solution given sufficient computational power; but in reality must always be an approximation.

Many of these methods rely on having at least some information about the likelihood. However in many cases, it may not be possible to determine this information or the information may be too computationally expensive to obtain [189, 198]. Therefore a set of methods, collectively known as Approximate Bayesian Computation (ABC), that directly approximate the posterior distribution, have been developed.

ABC algorithms were first used in DNA sequencing applications [160]. They eliminated the need to determine the likelihood by replacing it with a comparison between observed and simulated data. All ABC algorithms have the same three stage process [29, 34, 189, 198]:

1. A set of parameters (known as a particle) is sampled from proposed prior distributions;
2. The proposed model for the system is then used to simulate a data set for that particle;
3. The simulated data set is compared to the experimental data set using a distance function (usually the Euclidean distance). If below a tolerance, ϵ, it is accepted, otherwise it is rejected.

This is repeated for a number of particles and results in a selection of particles with associated parameters. The distribution of these particles will be a good approximation to the posterior, if the number of particles is sufficiently large and ϵ sufficiently small.

These key steps on their own do not result in an efficient algorithm, as the particle acceptance rate is low, especially when the prior and posterior differ to any great degree, i.e. the data is informative. The algorithm has been improved in several different ways which can be grouped into three major approaches: regression based methods, MCMC and Sequential Monte Carlo (SMC).

Beaumont *et al.* [30] introduced regression to reduce the discrepancies between the simulated and actual data. In addition, weighting dependent on the distance function

was introduced and this has become common in many other methods. Together these modifications allow ϵ to be increased, increasing acceptance rates, while not adversely affecting the outcome [29]. This approach is not, however, without difficulties. In certain situations the algorithm can be forced to extrapolate rather than interpolate, which is known to be error prone [29]. This error is exacerbated if the relationship is non-linear, which is usually the case. Problems also occur if the projected particles lie outside the bounds of the priors. Several methods have been proposed to overcome this, such as using a transformation, but these are all somewhat ad hoc [29].

The above improvements still have the inherent disadvantage that parameter values are sampled from the prior and so are still prone to high rejection rates. To improve on this the parameters need to be sampled from a distribution closer to the posterior [29]. MCMC simulation is one method of achieving this aim. MCMC ABC originates from work by Marjoram *et al.* [129]. The work presents a method based on the general MCMC Metropolis-Hastings method, replacing the likelihood term in the Metropolis-Hastings ratio with an indicator function [34]. This function is based on the fit of the simulated data and affects the next step in the chain, meaning that particle selection is no longer a random sample from the prior; but based on the results of previous simulations.

Although MCMC ABC improves the efficiency of the algorithm there are a number of disadvantages to this approach. A major one is that if the proposed distribution and ϵ are not chosen carefully the algorithm can get stuck in areas of low probability for long periods of time [29, 34, 198]. However a number of techniques have been proposed to reduce this problem, such as setting a high value of ϵ which is reduced after an initial burn in phase [29]. Also the dependence of one particle selection on the previous selection makes the use of parallel computing techniques (which is one of the main attractions of ABC) impossible [34].

The other approach to ensure that samples come from a distribution closer to the posterior is ABC SMC [181, 182, 198], which passes through several populations with ϵ starting off large and being slowly reduced. At each stage a small number of particles are rejected with the output distribution of one population used as the sampling procedure of the next. The result is distributions that slowly evolve towards the ultimate approximation of the posterior.

One of the issues with ABC SMC is as the scope to improve the fit between the data and simulation decreases, while ϵ continues to be decreased, the number of particle rejections becomes large [29]. This slows the algorithm without major improvements

in the approximation. ABC SMC can be improved at this stage by adding a regression step. The major advantage to ABC SMC is, however, that it can be applied to a diverse range of models without substantial modification [198].

The development and use of ABC algorithms has not been without criticism [165, 194–196], with the main criticism coming from its use in phylogeography. Although directed at ABC, this criticism is not specific to ABC methods but applies to Bayesian methods in general [189], for example criticising the choice of prior distributions. This criticism has been addressed in detail by Beaumont *et al.* [31].

2.7.2.1 Model Selection

The task of model selection is to select the best model from a given set of candidate models given a particular set of data. The challenge of model selection is that more complex models are more likely to fit a given set of data than simpler models, i.e. for a model with more parameters there is more opportunity to change those parameters in order to get a better fit. However Occam's razor, which is often quoted in English as 'it is vain to do with more what can be done with fewer' [183], would guide against having an undue preference for more complex models.

Therefore methods, such as the Akaike information criterion [18], have been developed which provide a trade off between goodness of fit and complexity by the inclusion of a penalty term. This penalises models with a larger number of parameters. Generally Bayesian model selection methods are thought to automatically penalise models with more parameters, thus do not require a separate penalty term. This is because a more complex model is capable of making a greater variety of predictions and so must spread its predictive probabilities more thinly over the data space[2]. However one of the main issues with Bayesian model selection is that the model selection can be particularly sensitive to the choice of priors for parameters within the models [109, 198]. Therefore further investigation is usually required to determine how sensitive the results are to the chosen priors.

Model selection is possible using the original ABC framework and was used in some of the early work [160]. Assuming that the same value of ϵ is used for each model, an approximation for the marginal likelihoods can be made from the ratio of each model's acceptance rates. Combining this information with a sampling procedure that takes account of the priors allows the posterior distributions to be calculated [29]. Models with more parameters are penalised as particles assigned to those models

[2]A good source of further information on Bayesian methods for model selection in the context of Occam's razors is Chapter 28 of [123].

are less likely to be accepted, due to the larger parameter space [29]. In order to improve acceptance rates of the ABC algorithms it has been common to use summary statistics to describe more complex data, particularly in phylogeography and other genetic based applications [29, 189]. Care has to be taken in choosing these summary statistics when carrying out model selection, however, as they can significantly effect the result [165, 189].

It is not possible to use the standard ABC model selection approach with the ABC SMC algorithms, as acceptance rates will depend on the ϵ schedule. Instead model selection is implemented by considering the choice of models as another discrete parameter [198]. This method does not rely on the use of summary statistics and so does not suffer from the issues relating to this, described above [165, 189]. The method can, however, still be sensitive to the choice of the priors for model parameters, like Bayesian model selection methods in general [128, 198]. In a general paper on ABC Marin *et al.* [128] also raised some concerns about this method, highlighting the fact that the results vary with the value of ϵ and the variance of the perturbation kernel, a dependence they believed should not occur, but they did not elaborate greatly about this. Although theoretically possible, no model selection techniques have yet been implemented for MCMC ABC [29].

2.8 Conclusion

Despite all the research described above, no method for routinely monitoring FRC or determining a measure of homogeneity in ventilated patients has seen widespread use. There have been some recent advances in determining FRC with the development of commercially available devices, such as the LUFU and Engstrom Carestation. These may prove suitable and effective for use in ICU, however full scale trials are still required. The challenge of developing a reliable and robust method for determining ventilation-volume inhomogeneity still remains.

Within this thesis the tidal models of the lung described above will now be developed to allow the entire expired concentration signal to be modelled and used in the inverse process for the first time. This will allow the the entire expired concentration signal to be used with non-linear optimisation techniques and ABC to determine information about the lung. Although sampling tubes will not be used in the collection of the data, the alignment of the two signals is still important. Methods of combining alignment correction with parameter estimation are investigated for the first time. These advances may allow FRC to be determined more robustly and will enable alternative methods for inferring information about the ventilation-volume distribution to be proposed.

Chapter 3

Methods

3.1 Introduction

Within this chapter the data, models and parameter estimation techniques that will be used in subsequent chapters are presented. The chapter starts by outlining the three sources of data, which will be analysed in the following three chapters. The models used in this analysis are then presented. First a standard 'balloon on a straw' model of tidal lung ventilation is described. This model is developed in three ways: the dead space is then developed to include mixing, the alveolar compartment is developed from the homogeneous case into two possible models of inhomogeneity and a method of modelling the body as a series of compartments is described.

The chapter closes by presenting the parameter estimation techniques that will be used within the thesis. A method devised by clinical colleagues is first described as this provides a baseline technique against which to compare other methods developed in this thesis. These new methods are based on two standard techniques of parameter estimation: least squares optimisation and Approximate Bayesian ABC SMC.

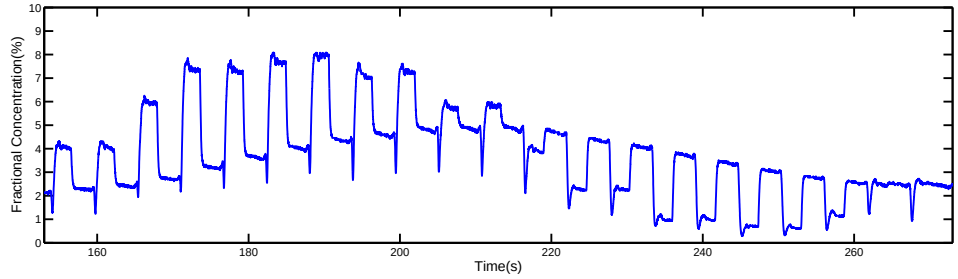
3.2 Data

The data used in this thesis come from three sources:

- Simulated data – data created using the models described later in this chapter.
- Bench lung data – data collected from a water-displacement bench lung, which is described further below.
- Human data – data collected from healthy subjects.

In all cases each data set is made up of two signals: the concentration of the tracer gas at the mouth and the flow rate into the lungs. The concentration sensor measures the the fractional concentration of the tracer gas and the flow sensor measures the amount of air inhaled or exhaled from the lungs. Positive flows are defined as being into the

(a) Concentration Signal



(b) Flow Signal

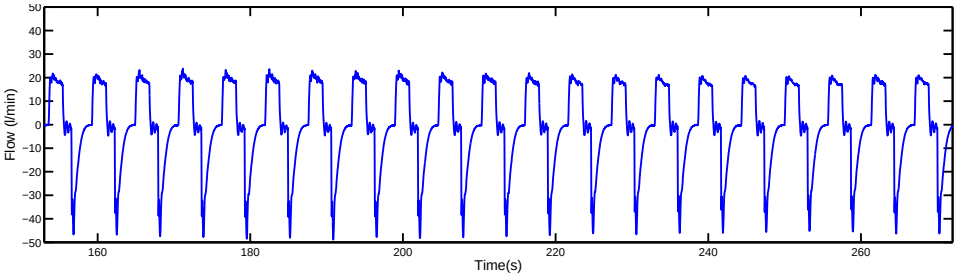


Figure 3.1: Example data taken from a homogeneous bench lung with a dead space and alveolar volume of $100ml$ and $2400ml$, respectively. The forcing sinusoid has a nominal period of $120s$.

lungs and negative out. Figure 3.1 is an example of the typical concentration and flow signals found within the data.

The inspired concentration and the total flow rate at the mouth are taken as inputs to the system with the resulting output seen in the expired concentration signal. In theory it would be possible to control the flow rates into and out of the lung for a ventilated subject, however this thesis does not attempt to investigate the effects of different flow patterns. Instead it concentrates on the manipulation of the system by varying the inspired concentration of the tracer gas. Most commonly the concentration is varied sinusoidally, with the inspired fractional concentration of the tracer gas, F_I, being:

$$F_I(t) = F_{IM} + \Delta F_I.\sin(2\pi f t) \ , \tag{3.1}$$

where F_{IM}, ΔF_I, f are the mean, amplitude and frequency of the inspired tracer gas forcing signal, receptively. In other situations, in order to increase the broadband nature of the input, the inspired signal can be varied using a frequency modulated sinusoidal baseband signal, which is given by formula [93]:

$$F_I(t) = F_{IM} + \Delta F_I.\cos\left(2\pi f_c t + \frac{f_\Delta}{f_m}\cos(2\pi f_m t)\right) \ , \tag{3.2}$$

where f_c, f_m and f_Δ are the frequency of the carrier sinusoid wave, modulated sinusoid wave and frequency deviation, respectively.

The use of a modulated waveform allows the time between successive peaks of the wave to vary. The amount of variation is governed by f_c and f_Δ with the same wave form repeated at a time period of $1/f_m$. Increasing f_Δ will increase the deviation from carrier frequency cosinewave and so increase the broadband nature of the signal; however if f_Δ is increased above $f_c - f_m$ the individual components of the wave will have amplitudes of less than ΔF_I. Therefore within this thesis f_Δ will be set as follows:

$$f_\Delta = f_c - f_m. \tag{3.3}$$

Within this thesis a f_m and f_c of 1/15.00 and $1/2.63 min^{-1}$ are used respectively. This produces a waveform that starts at the origin, has approximately from 1.5 to $6min$ between successive peaks and repeats every $15min$.

3.2.1 Sinusoidal Delivery System

In order to deliver a varying input in situations where the flow rates are dynamic or unpredictable, such as spontaneous breathing or when using modern ventilators, the tracer gas needs to be delivered 'on the fly'. A system to do this has been developed by collaborators in the Nuffield Division of Anaesthetics, Nuffield Department of Clinical Neurosciences. The basic set-up, shown in Figure 3.2, consists of the following components:

- A flow sensor (THOR D-22 by Thor [4]), which monitors the total air flow rate into and out of the mouth.
- An in-line concentration sensor (a non-dispersive infra-red sensor by SquareOne, now discontinued), which monitors the fractional concentration of the tracer gas at the mouth.
- A mass flow controller (MC-10SLPM by Alicat Scientific [1]), which injects a known amount of tracer gas into the airflow path being consumed by the subject.

- A Personal Computer (PC) running a LabVIEW virtual instrument [7].

The control of the system is provided by LabVIEW which allows the user to set-up a virtual instrument [7]; this mimics a real instrument but has the obvious advantage that it can be easily reconfigured. The virtual instrument provides a graphical interface to the user, logs the data and controls the amount of tracer gas injected. At the beginning of each breath the control system uses a forcing function (described in the previous

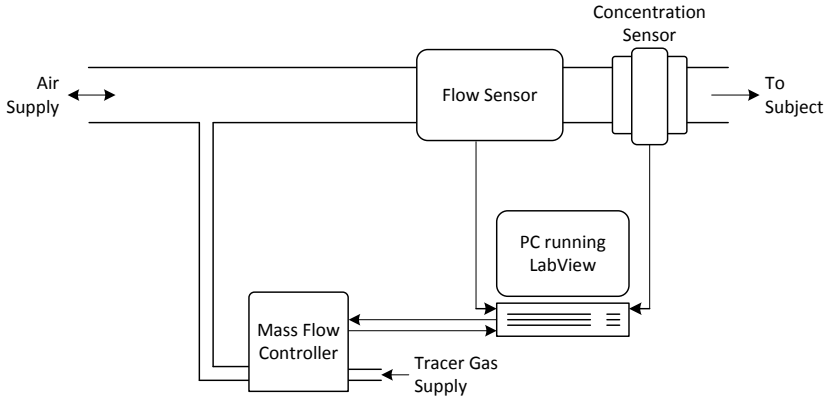


Figure 3.2: Block diagram of the sinusoid delivery system.

section) to determine the required fractional concentration of the tracer gas required during inspiration.

The system then attempts to deliver this concentration throughout inspiration. This is done by varying the output of the mass flow controller depending on the current flow rate and feedback about the actual amount of gas injected. Feedback from the concentration sensor is not used in the control process. The result is a fractional concentration which is roughly constant throughout the inspiration phase of the breath but which varies sinusoidally between breaths, as can be seen in Figure 3.1a. Further background information on this set-up can be found in Chapter 9 of [77], with the current set-up described in a patent for the system [158].

3.2.2 Simulated Data

In chapter 4, simulated data, created using models described in the next section, are used. This allows different techniques for determining information about the lung, such as a measure of homogeneity, to be explored on error free data with known parameters, before using these techniques on real data. In order to produce simulated data, the two input signals, the fractional concentration of the tracer gas during inspiration and the flow rate into the lung, are required.

The inspired fractional concentration comes from a forcing function (Equation 3.1 or 3.2) but normally with the assumption that the mean values of the inspired and expired sine waves cancel out. This assumption is supported by other work [84, 91, 218, 224, 229, 233] and simplifies the input to either:

$$F_I(t) = \Delta F_I.\sin(2\pi f t) \ , \tag{3.4}$$

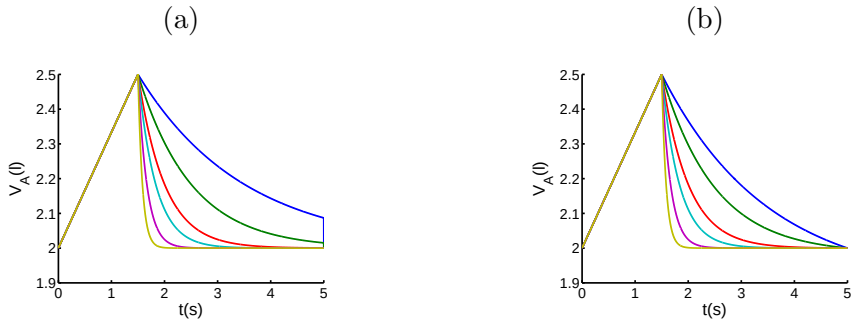


Figure 3.3: Total volume of the alveolar compartments (V_A) during a five second breath cycle. End expired alveolar (V'_A) and tidal (V_T) volume are $2000ml$ and $500ml$ respectively. The decay constant (γ) varies from 0.5 (dark blue) to 12.0 (yellow). (a) Breathing pattern proposed by Gavaghan and Hahn [84]. (b) Breathing pattern used within this thesis.

or

$$F_I(t) = \Delta F_I.\cos\left(2\pi f_c t + \frac{f_\Delta}{f_m}\cos(2\pi f_m t)\right) \ . \tag{3.5}$$

These forcing functions are used as they mimic those used when collecting the bench lung and human data described later. However any forcing function could be used with the models that follow, such as a step function to mimic a nitrogen washout test.

In order to create the flow pattern, it is assumed the subject is ventilated and the breathing pattern proposed by Gavaghan and Hahn [84] is thus used as a starting point. The alveolar volume, V_A, is assumed to increase linearly during inspiration and then follow an exponential decay during expiration:

$$V_A(t) = \quad V'_A + \tfrac{V_T}{t_I}(t - T_n) \qquad \text{on inspiration} \tag{3.6}$$

$$V_A(t) = \quad V'_A + V_T.e^{-\gamma(t-(T_n+t_I))} \qquad \text{on expiration,} \tag{3.7}$$

where $\gamma = 1/R.K$ (R and K are the resistance and compliance of the lung). This breathing pattern is shown in Figure 3.3a.

The disadvantage of this breathing pattern is that the amount of air expired is slightly less than the tidal volume. The error will depend on γ and is not significant for larger values; however the error becomes larger as γ decreases. This can be seen in Figure 3.3a. As lower values of γ more closely match the real emptying pattern of ventilated lungs, a breathing pattern with an expired tidal volume of exactly V_T is required.

The derivation of this breathing pattern starts with a modified version of Equation 3.7:

$$V_A(t) = V'_A + V^*_T.e^{-\gamma(t-(T_n+t_I))} \ , \tag{3.8}$$

where V_T is replaced by V_T^*. For the expired volume to be equal to V_T, the following must be true:

$$V_A(T_n + t_I) - V_A(T_n + t_I + t_E) = V_T \ , \tag{3.9}$$

where t_E is the duration of expiration. Substituting Equation 3.8 into Equation 3.9 gives:

$$V_T^* = \frac{V_T}{1 - e^{-\gamma t_E}} \ . \tag{3.10}$$

This ensures that the correct amount of air is expired. However expiration begins at $V_A' + V_T^*$ which is above the volume at which inspiration ends: $V_A' + V_T$, therefore $(V_T^* - V_T)$ needs to be removed, making the alveolar volume during expiration:

$$V_A(t) = V_A' + V_T^* \left(e^{-\gamma(t-(T_n+t_I))} - 1\right) + V_T \ . \tag{3.11}$$

This new breathing pattern can be seen in Figure 3.3b. The rate of change in the alveolar compartment volume (which is usually equal to the flow rate at the mouth) can then be determined by differentiating Equation 3.6 and 3.11:

$$\frac{\mathrm{d}V_A(t)}{\mathrm{d}t} = \frac{V_T}{t_I} \quad \text{on inspiration} \tag{3.12}$$

$$\frac{\mathrm{d}V_A(t)}{\mathrm{d}t} = -\gamma . V_T^* . e^{-\gamma(t-(T_n+t_I))} \quad \text{on expiration.} \tag{3.13}$$

This breathing pattern assumes a respiratory quotient of one, as is generally assumed in the computational models used within this thesis.

3.2.3 Water-Displacement Bench Lung Data

A homogeneous water-displacement bench model of the lung, as shown in Figure 3.4a, was developed by Williams *et al.* [223]. It consists of a large outer cylinder, with a lid, and a smaller inner cylinder, which are connected in such a way that water can pass between them at the base. Water is added to the cylinders so as to cover the connection between the two compartments. Air is then pumped into the outer cylinder, using a ventilator, causing the water to be displaced into the inner cylinder. In order to change the compliance of the lung a smaller solid cylinder can be added to the inner cylinder reducing its volume.

This bench model serves as a valuable way of verifying techniques for estimating lung volumes because it has the advantage of having known volumes for the dead space and alveolar compartment. The dead space volume can be varied by adding tubing of known volume between the mouth sensors and the connection to the alveolar compartment, and the alveolar volume can be changed by modifying the water level.

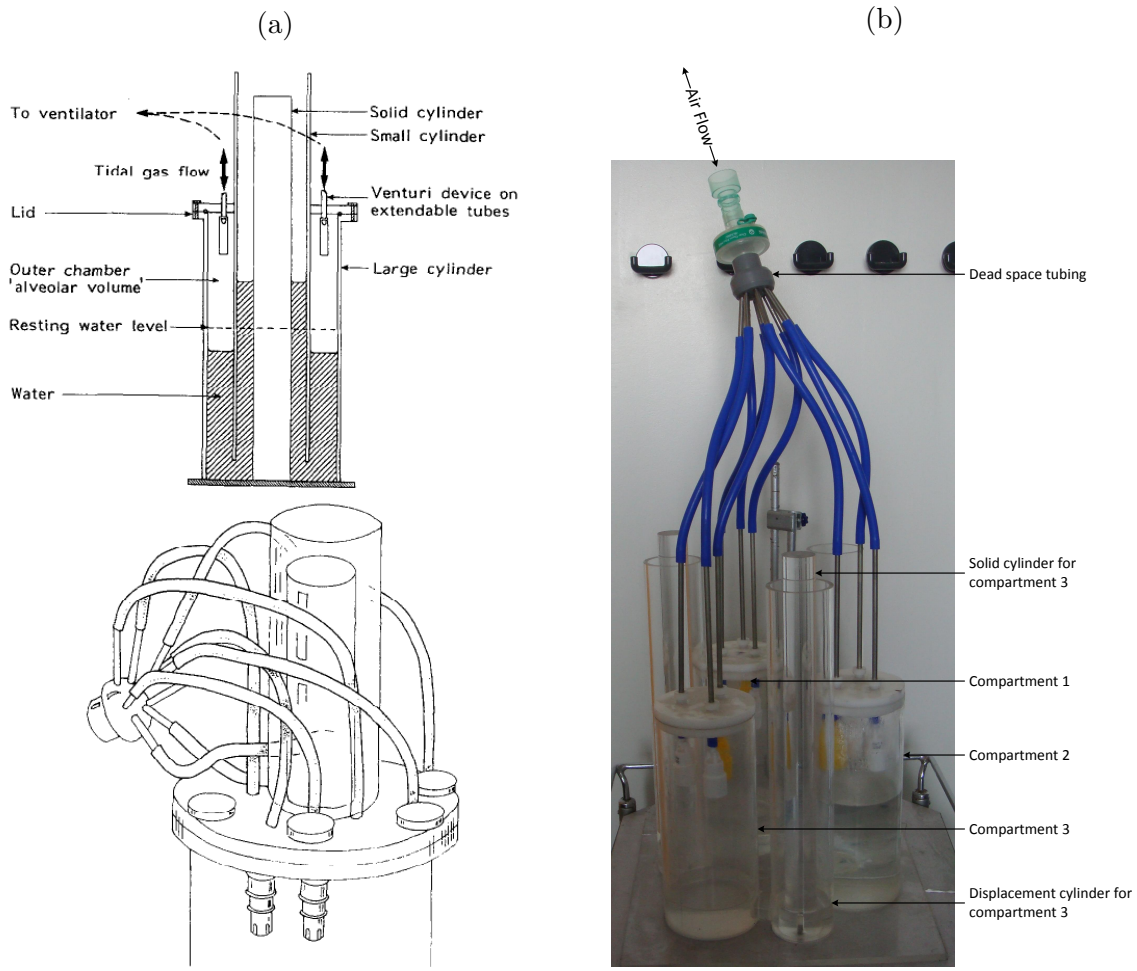


Figure 3.4: The water-displacement bench lung. (a) A diagram of a homogeneous lung. Taken from [223]. (b) A photograph showing a three compartment lung.

In addition a multi-compartment version of a water-displacement lung has been developed, which has three separate alveolar compartments each made up of two cylinders. Figure 3.4b shows this set-up within which the inner cylinder has been placed to one side of the outer one to ease construction. This bench model can be used to verify techniques for determining the homogeneity of the lung.

3.2.4 Human Lung Data

Data from 18 healthy volunteers are used to verify the methods developed using the simulated and bench lung data. A number of repeated measurements are made allowing the readability of the various methods to be assessed. Also FRC is determined using the plethysmograph in order to allow the performance to be compared against a benchmark volume.

3.2.5 Preprocessing

A number of preprocessing steps are carried out on the bench lung and human data before further analysis is performed. Firstly both the concentration and flow data are

linearly interpolated onto the same time series. A time step of $10ms$ is used, which is comparable to the time steps found within the data, however neither of the signals are evenly sampled throughout. Linear interpolation is used as regions of rapid change are observed in both signals, causing other interpolation techniques with higher degrees of freedom to potentially produce large errors. Median filters are then applied to both signals to reduce noise. A window size of nine samples is used, as this was found to be a good compromise between smoothing local fluctuations in the data while not curtailing short lived peaks within the data.

Finally it is assumed that the volume of the lung does not consistently increase or decrease in value, so any long term linear trend is removed. In order to do this the flow signal is first numerically integrated using the trapezoidal rule and the resulting volume signal linearly detrended. The resulting volume signal is then numerically differentiated to produce a modified flow signal.

3.3 Computational Models

As detailed in the previous chapter (see Section 2.3.2.2), tidal compartment models of lung ventilation have previously only attempted to predict the average expired tracer gas concentration of a breath or the concentration at particular points within the breath. Here models which predict the concentration throughout expiration are presented.

3.3.1 Homogeneous 'Balloon on a Straw' Model

The simplest form of the tidal breathing model, the 'Balloon on a Straw' model, was illustrated previously in Figure 2.3b. The lung is modelled by two compartments; a fixed volume compartment which represents the dead space, the 'Straw', and a compartment which expands and contracts, the 'Balloon'.

3.3.1.1 The Dead Space Compartment

It is assumed here that no mixing or diffusion occurs within the dead space and here gas that enters the dead space, either at the mouth (during inspiration) or from the alveolar compartment (during expiration), later leaves the dead space at the same concentration. Therefore during inspiration, assuming that the volume of the dead space has already been inspired (i.e. $d_{v_D} < d_0$), the fractional concentration of a tracer gas that enters the alveolar compartment, F_{IA}, is:

$$F_{IA}(t) = F_I(t - d_{v_D}) \ , \tag{3.14}$$

where d_{v_D} is the time delay caused by the time taken for gas to travel through the dead space compartment which is dependent on the volume of the dead space, V_D. If the volume of the dead space has not yet been inspired (i.e. $d_{v_D} > d_0$), then the gas entering the alveolar compartment is gas that exited the same compartment previously, so:

$$F_{IA}(t) = F_A(t - d_0) \ , \tag{3.15}$$

where d_0 is the time delay for gas to leave and then re-enter the alveolar compartment.

The total flow rate of gas entering at the mouth ($\dot{V}$) is measured. This allows d_{v_D} to be determined from knowledge of V_D, where d_{v_D} is the lowest value solution of the integral:

$$\int_{t-d_{v_D}}^{t} \dot{V}(\Lambda) \ \mathrm{d}\Lambda = V_D \ . \tag{3.16}$$

Similarly, d_0 is the lowest value solution to the integral:

$$\int_{t-d_0}^{t} \dot{V}(\Lambda) \ \mathrm{d}\Lambda = 0 \ . \tag{3.17}$$

During expiration the fractional concentration at the mouth (F_E) can be determined in a similar way to F_{IA}:

$$F_E(t) = F_A(t - d_{v_D}) \qquad \text{if } d_{v_D} < d_0 \tag{3.18}$$

$$F_E(t) = F_I(t - d_0) \qquad \text{if } d_{v_D} > d_0 \ , \tag{3.19}$$

where the time delays are determined by the lowest value solutions of the following integrals:

$$\int_{t-d_{v_D}}^{t} \dot{V}(\Lambda) \ \mathrm{d}\Lambda \ = \ -V_D \tag{3.20}$$

$$\int_{t-d_0}^{t} \dot{V}(\Lambda) \ \mathrm{d}\Lambda \ = \ 0 \ . \tag{3.21}$$

In equation 3.20 the integral is now equated to $-V_D$ as $\dot{V}$ is negative during expiration.

3.3.1.2 The Alveolar Compartment

The volume of the alveolar compartment changes throughout the breath. A insoluble gas is assumed here and the dead space volume is fixed, so the volume of the alveolar compartment (V_A) is only dependent on the flow:

$$\frac{\mathrm{d}V_A}{\mathrm{d}t} = \dot{V}(t) \ . \tag{3.22}$$

The mass of the tracer gas within the alveoli compartment is governed by two mass balance equations (similar to those described in Section 2.3.2.2):

$$\frac{\mathrm{d}}{\mathrm{d}t}(V_A.F_A) = \dot{V}(t).F_{IA}(t) \quad \text{for inspiration} \tag{3.23}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(V_A.F_A) = \dot{V}(t).F_A(t) \quad \text{for expiration .} \tag{3.24}$$

3.3.2 Modelling Mixing within the Dead Space

Generally models of the dead space have assumed that it is a transitory compartment within which no mixing occurs. In reality, as gases diffuses and there is likely to areas of turbulent flow due to the branching structure of the lungs, mixing of the gas within the dead space is likely to occur. Within this section several possible ways of modelling this mixing are described.

One way to include mixing is to add an additional compartment at the mouth, as shown in Figure 3.5a. Air entering this compartment fully mixes with the air already within it, with the mass balance equations for the compartment being:

$$V_M.\frac{\mathrm{d}F_M}{\mathrm{d}t} = \dot{V}(t)(F_I(t) - F_M(t)) \quad \text{for inspiration} \tag{3.25}$$

$$V_M.\frac{\mathrm{d}F_M}{\mathrm{d}t} = \dot{V}(t).(F_M(t) - F_{ER}(t)) \quad \text{for expiration ,} \tag{3.26}$$

where F_M is the fractional concentration of tracer gas, V_M is the fixed volume of the mixing compartment and F_{ER} is the fractional concentration expired from the transitory compartment into the mixing compartment. Note that flow is negative during expiration, hence F_M is positive and F_{ER} is negative in the Equation 3.26. This additional compartment is a simple way to model the mixing which in reality will occur throughout the whole dead space compartment, rather than only at the mouth.

The equations for the transitory compartment are similar to that of the dead space in the simple 'balloon and straw' model. During inspiration:

$$F_{IA}(t) = F_M(t - d_{v_R}) \quad \text{if } d_{v_R} < d_0 \tag{3.27}$$

$$F_{IA}(t) = F_A(t - d_0) \quad \text{if } d_{v_R} > d_0 \ , \tag{3.28}$$

with the delays being the lowest value solutions of:

$$\int_{t-d_{v_R}}^{t} \dot{V}(\Lambda)\ \mathrm{d}\Lambda = V_R \tag{3.29}$$

$$\int_{t-d_0}^{t} \dot{V}(\Lambda)\ \mathrm{d}\Lambda = 0 \ , \tag{3.30}$$

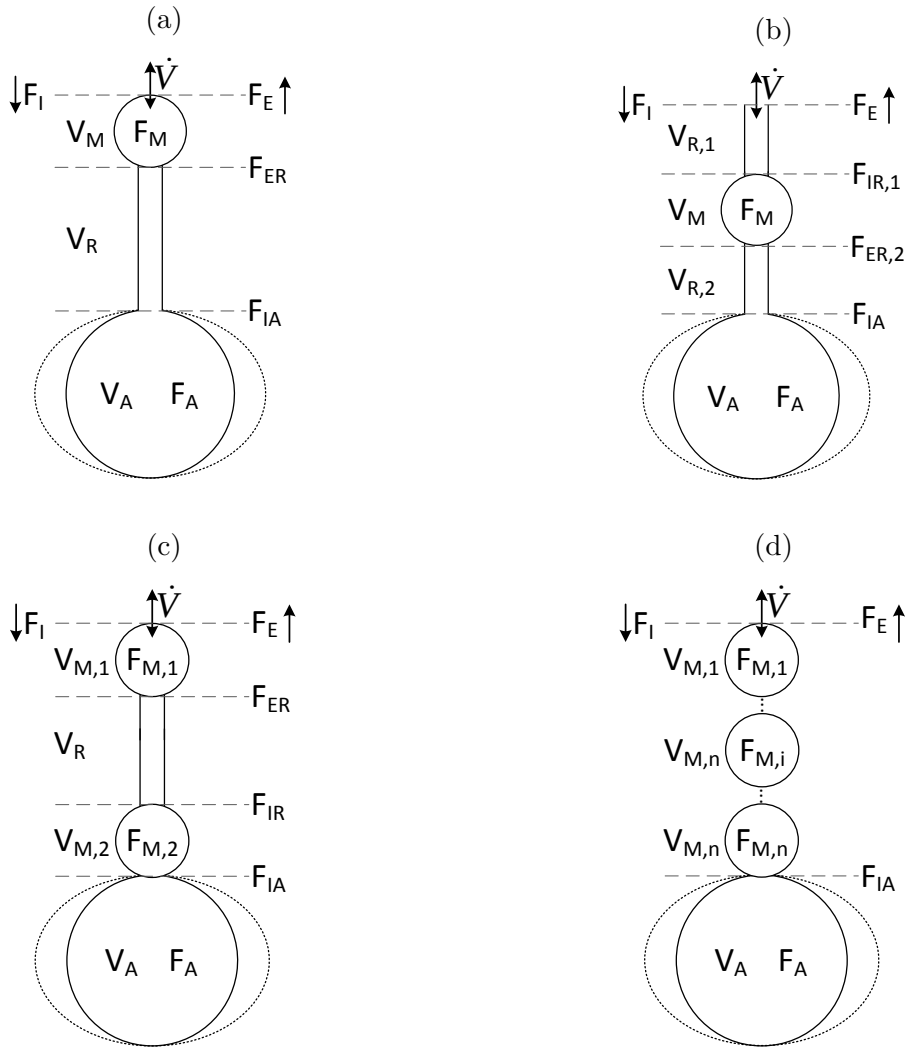


Figure 3.5: Schematic diagrams of possible models of mixing within the dead space. (a) Mixing Compartment at the mouth. (b) Mixing Compartment Between Two Transitory Compartments. (c) Transitory Compartment Between Two Mixing Compartments. (d) Series of Mixing Compartments.

where V_R is the volume of the transitory compartment and d_{v_R} is the delay induced by that volume. During expiration:

$$F_{ER}(t) = F_A(t - d_{v_R}) \qquad \text{if } d_{v_R} < d_0 \tag{3.31}$$

$$F_{ER}(t) = F_I(t - d_0) \qquad \text{if } d_{v_R} > d_0 \ , \tag{3.32}$$

with the delays being the lowest value solutions of:

$$\int_{t-d_{v_R}}^{t} \dot{V}(\Lambda)\, \mathrm{d}\Lambda \;=\; -V_R \tag{3.33}$$

$$\int_{t-d_0}^{t} \dot{V}(\Lambda)\, \mathrm{d}\Lambda \;=\; 0 \ . \tag{3.34}$$

The expired fractional concentration at the mouth (F_E) is equal to F_M during expiration.

This is one option for the make up of the dead space; however in reality there are many possible ways of modelling the dead space with a number of mixing and transitory compartments. Three others which are investigated within this thesis are: a mixing compartment between two transitory compartments, a transitory compartment between two mixing compartments and a series of mixing compartments. These models are illustrated in Figure 3.5 and their equations are listed in Table 3.1.

3.3.3 Inhomogeneous Alveolar Compartment

So far in this chapter it has been assumed that the alveolar compartment is homogeneous and so can be modelled by a single expanding compartment (a single balloon). However in reality the lung is not homogeneous and so the model of the alveolar compartment needs to be adjusted to allow inhomogeneities.

As previously discussed, the investigations carried out in this thesis are confined to the effects of volume-ventilation inhomogeneities, with two methods of modelling variation in the end-expired alveolar volume/tidal volume ratio throughout the lung now presented. The effects of variation in the emptying rates throughout the lung are not investigated, i.e. γ is assumed to be constant throughout the lung. The first method, the multi-compartment model, replaces the single alveolar compartment with a discrete number of compartments, while the second, the continuous log-normal distribution model, replaces it with a continuous distribution of compartments across a range of ventilation-volume ratios.

3.3.3.1 Multi-Compartment Model

In the multi-compartment model the alveolar compartment is modelled by a number of discrete compartments. Each of these compartments has a tidal volume, $V_{T,i}$, and an end expired alveolar volume, $V'_{A,i}$, associated with it, allowing the ventilation-volume ratio of each compartment to differ. The assumption of constant emptying rates throughout the lung means the inspired and expired flow is distributed evenly to all compartments in proportion to the tidal volume associated with that compartment ($V_{T,i}$). This means the flow into a compartment, $\dot{V}_i$, is:

$$\dot{V}_i(t) = \frac{V_{T,i}}{V_T}.\dot{V}(t) \ , \tag{3.35}$$

where V_T is the total tidal volume. Therefore the mass balance equations for each compartment are:

$$\frac{\mathrm{d}}{\mathrm{d}t}(V_{A,i}.F_{A,i}) = \dot{V}_i(t).F_{IA}(t) \qquad \text{for inspiration} \tag{3.36}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(V_{A,i}.F_{A,i}) = \dot{V}_i(t).F_{A,i}(t) \qquad \text{for expiration} \ , \tag{3.37}$$

Mixing Compartment Between Two Transitory Compartments

	Inspiration	Expiration	
Compartment 1 (Transitory)	$F_{IR,1}(t) = F_I(t - d_{v_R,1})$ $F_{IR,1}(t) = F_M(t - d_0)$ $\int_{t-d_{v_R,1}}^{t} \dot{V}(\Lambda)\ \mathrm{d}\Lambda = V_{R,1}$ $\int_{t-d_0}^{t} \dot{V}(\Lambda)\ \mathrm{d}\Lambda = 0$	$F_E(t) = F_M(t - d_{v_R,1})$ $F_E(t) = F_I(t - d_0)$ $\int_{t-d_{v_R,1}}^{t} \dot{V}(\Lambda)\ \mathrm{d}\Lambda = -V_{R,1}$ $\int_{t-d_0}^{t} \dot{V}(\Lambda)\ \mathrm{d}\Lambda = 0$	if $d_{v_R,1} < d_0$ if $d_{v_R,1} > d_0$
Compartment 2 (Mixing)	$V_M.\frac{\mathrm{d}F_M}{\mathrm{d}t} = \dot{V}(t)(F_{IR,1}(t) - F_M(t))$	$V_M.\frac{\mathrm{d}F_M}{\mathrm{d}t} = \dot{V}(t)(F_M - F_{ER,2})$	
Compartment 3 (Transitory)	$F_{IA}(t) = F_M(t - d_{v_R,2})$ $F_{IA}(t) = F_A(t - d_0)$ $\int_{t-d_{v_R,2}}^{t} \dot{V}(\Lambda)\ \mathrm{d}\Lambda = V_{R,2}$ $\int_{t-d_0}^{t} \dot{V}(\Lambda)\ \mathrm{d}\Lambda = 0$	$F_{ER,2}(t) = F_A(t - d_{v_R,2})$ $F_{ER,2}(t) = F_M(t - d_0)$ $\int_{t-d_{v_R,2}}^{t} \dot{V}(\Lambda)\ \mathrm{d}\Lambda = -V_{R,2}$ $\int_{t-d_0}^{t} \dot{V}(\Lambda)\ \mathrm{d}\Lambda = 0$	if $d_{v_R,2} < d_0$ if $d_{v_R,2} > d_0$

Transitory Compartment Between Two Mixing Compartments

	Inspiration	Expiration	
Compartment 1 (Mixing)	$V_{M,1}.\frac{\mathrm{d}F_{M,1}}{\mathrm{d}t} = \dot{V}(t)(F_I(t) - F_{M,1}(t))$	$V_{M,1}.\frac{\mathrm{d}F_{M,1}}{\mathrm{d}t} = \dot{V}(t)(F_{M,1}(t) - F_{ET}(t))$	
Compartment 2 (Transitory)	$F_{IR}(t) = F_{M,1}(t - d_{v_R})$ $F_{IR}(t) = F_{M,2}(t - d_0)$ $\int_{t-d_{v_R}}^{t} \dot{V}(\Lambda)\ \mathrm{d}\Lambda = V_R$ $\int_{t-d_0}^{t} \dot{V}(\Lambda)\ \mathrm{d}\Lambda = 0$	$F_{ER}(t) = F_{M,3}(t - d_{v_R})$ $F_{ER}(t) = F_{M,1}(t - d_0)$ $\int_{t-d_{v_R}}^{t} \dot{V}(\Lambda)\ \mathrm{d}\Lambda = -V_R$ $\int_{t-d_0}^{t} \dot{V}(\Lambda)\ \mathrm{d}\Lambda = 0$	if $d_{v_R} < d_0$ if $d_{v_R} > d_0$
Compartment 3 (Mixing)	$V_{M,2}.\frac{\mathrm{d}F_{M,2}}{\mathrm{d}t} = \dot{V}(t)(F_{IR}(t) - F_{M,2}(t))$	$V_{M,2}.\frac{\mathrm{d}F_{M,2}}{\mathrm{d}t} = \dot{V}(t)(F_{M,2}(t) - F_A(t))$	
	$F_{IA}(t) = F_{M,2}(t)$	$F_E(t) = F_{M,1}(t)$	

Series of Mixing Compartments

	Inspiration	Expiration
Compartment 1 (First)	$V_{M,1}.\frac{\mathrm{d}F_{M,1}}{\mathrm{d}t} = \dot{V}(t)(F_I(t) - F_{M,1}(t))$	$V_{M,1}.\frac{\mathrm{d}F_{M,1}}{\mathrm{d}t} = \dot{V}(t)(F_{M,1}(t) - F_{M,2}(t))$
Compartment i (Intermediary)	$V_{M,i}.\frac{\mathrm{d}F_{M,i}}{\mathrm{d}t} = \dot{V}(t)(F_{M,i-1}(t) - F_{M,i}(t))$	$V_{M,i}.\frac{\mathrm{d}F_{M,i}}{\mathrm{d}t} = \dot{V}(t)(F_{M,i}(t) - F_{M,i+1}(t))$
Compartment n (Last)	$V_{M,n}.\frac{\mathrm{d}F_{M,n}}{\mathrm{d}t} = \dot{V}(t)(F_{M,n-1}(t) - F_{M,n}(t))$	$V_{M,n}.\frac{\mathrm{d}F_{M,n}}{\mathrm{d}t} = \dot{V}(t)(F_{M,n}(t) - F_A(t))$
	$F_{IA}(t) = F_{M,n}(t)$	$F_E(t) = F_{M,1}(t)$

Table 3.1: Formula for dead space models with mixing compartments. F_{IR} is the fractional concentration inspired into the transitory compartment.

which are equivalent equations to 3.23 and 3.24 for the homogeneous case. $F_{A,i}$ is the fractional concentration of the tracer gas within compartment i.

The fractional concentration expired into the dead space is now no longer simply the fractional concentration of a single alveolar compartment but a mixture of the fractional concentration from all the individual discrete compartments:

$$F_{\bar{A}}(t) = \frac{1}{V_T} \sum_{i=1}^{n} V_{T,i}.F_{A,i}(t) \ , \tag{3.38}$$

where $F_{\bar{A}}$ is the mixed alveolar fractional concentration and n is the total number of compartments.

3.3.3.2 Continuous Log-Normal Distribution Model

There has been a general assumption made within the literature [42, 73, 117, 137, 163, 197, 214, 217, 218] that the ventilation-volume and ventilation-perfusion ratios of a human lung exhibit a distribution that can be made up of a number of log-normal distributions. Within these works no justification is given for the use of such distributions. However when simulating data some form of distribution is required. This is the most logical starting point given its previous use and allow confidence in the methods to be built up before potential further testing in humans, which may exhibit more complex distributions. Also this type of distribution appears to mimic some of the real characteristics of the inhomogeneous human lung, as discussed later.

Whiteley *et al.* [218] used this assumption to create a model which includes a continuous distribution of compartments. This model allowed both ventilation-volume and ventilation-perfusion distributions; here however, as insoluble gases are being considered, the model has been modified to consider only a distribution of the ventilation-volume ratio.

Following Whiteley's method, a distribution is set up by defining:

$$x = \ln\left(\frac{V'_{Ad}}{V_{Td}}\right) \qquad -\infty < x < \infty \ , \tag{3.39}$$

and the tidal volume associated with point x, V_{Td}, as:

$$V_{Td}(x) = \sum_{p=1}^{n} \frac{V_{T,p}}{\sqrt{2\pi}.\sigma_p}.e^{-\frac{(x-\bar{x_p})^2}{2\sigma_p^2}} \ . \tag{3.40}$$

V'_{Ad} is the end-expired alveolar volume associated with point x; $V_{T,p}$, $\bar{x_p}$ and σ_p are the tidal volume, mean and standard deviation of the individual peak p, respectively.

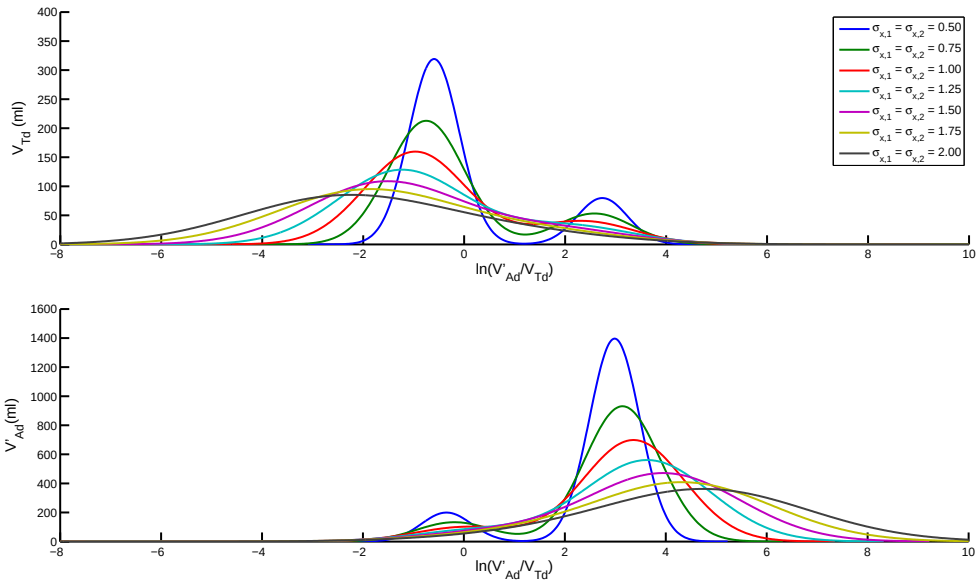


Figure 3.6: Example of a continuous log-normal distribution with two peaks. The standard deviation of both peaks are the same and varied together as indicated in the key above. $V'_{A,1} = 250ml$, $V'_{A,2} = 1750ml$, $V'_{T,1} = 400ml$ and $V'_{T,2} = 100ml$.

Noting that $V'_{Ad} = V_{Td}.e^x$, from Equation 3.39, it follows that the end-expired alveolar volume associated with peak p, $V'_{A,p}$, is:

$$V'_{A,p} = \int_{-\infty}^{\infty} V'_{Ad}(x)\ \mathrm{d}x = \frac{V_{T,p}}{\sqrt{2\pi}.\sigma_p} \int_{-\infty}^{\infty} e^{-\frac{(x-\bar{x_p})^2}{2\sigma_p^2}} e^x \mathrm{d}x\ , \tag{3.41}$$

thus:

$$V'_{A,p} = V_{T,p}.e^{\left(\frac{\sigma_p^2}{2}+\bar{x_p}\right)} \tag{3.42}$$

therefore:

$$\bar{x_p} = \ln\left(\frac{V'_{A,p}}{V_{T,p}}\right) - \frac{\sigma_p^2}{2}\ . \tag{3.43}$$

Equations 3.40 and 3.43 allow a lung containing alveoli with a distribution of tidal volumes (V_{Td}) to be modelled. These distributions are made up of a number of log-normal distribution peaks, each of which is defined by three values: $V_{T,p}$, $V'_{A,p}$ and σ_p. For each point on the distribution there is also an associated value of V'_{Ad}, which can be obtained using Equation 3.39. More inhomogeneous lungs are represented by distributions with more peaks, which have higher standard deviations (σ_p) or have means ($\bar{x_p}$) that are well separated. Figure 3.6 shows an example distribution with two peaks.

As assumed by Whiteley *et al.* [218] and previously for the multi-compartment model, the flow is distributed in proportions to the associated tidal volume. This means that the flow into a compartment at a point on the distribution, $\dot{V}_{Ad}$, is:

$$\dot{V}_{Ad}(x,t) = \frac{V'_{Td}(x)}{V_T}\ \dot{V}(t)\ , \tag{3.44}$$

and so the volume associated with a point on the distribution, V_{Ad}, is:

$$V_{Ad}(x,t) = V'_{Ad}(x) + \int_{t_0}^{t} \dot{V}_{Ad}(x,\Lambda)\ \mathrm{d}\Lambda\ , \tag{3.45}$$

where t_0 is the initial time point (i.e. $V_{Ad}(x,t_0) = V'_{Ad}(x)$). The mass balance equations mimic Equations 3.36 and 3.37 for the multi-compartment model:

$$\frac{\mathrm{d}}{\mathrm{d}t}(V_{Ad}(x,t).F_{Ad}(x,t)) = \dot{V}_{Ad}(x,t).F_{IA}(t) \qquad \text{for inspiration} \tag{3.46}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(V_{Ad}(x,t).F_{Ad}(x,t)) = \dot{V}_{Ad}(x,t).F_{Ad}(x,t) \qquad \text{for expiration,} \tag{3.47}$$

where F_{Ad} is the fractional concentration of the tracer gas at a point on the distribution. Similarly, Equation 3.38 becomes a flow weighted integral, rather than a weighted sum:

$$F_{\bar{A}}(t) = \frac{1}{V_T}\int_{-\infty}^{+\infty} V_{Td}(x).F_{Ad}(x,t)\ \mathrm{d}x\ . \tag{3.48}$$

3.3.4 Soluble Gases

When a soluble gas is considered the uptake and return of tracer gas to and from the body has to be taken into account. When using a soluble gas the lung will be treated as homogeneous throughout this thesis and so only a homogeneous model is described here. The model used is based on the model described in Section 2.3.2.2 (Equations 2.6 to 2.8 and 2.17 to 2.16).

This traditional model assumes that the volume change of the alveolar compartment due to the exchange of tracer gas with blood has little effect and so can be ignored. In some cases described later this same assumption will be made. However when simulating data and testing this assumption the model will be modified to ensure that the mass of the tracer gas is balanced throughout the system. The mass balance equations become:

Inspiration:

$$\frac{\mathrm{d}}{\mathrm{d}t}(V_A(t).F_A(t)) = \dot{V}(t).F_{IA}(t) + \dot{Q}_P(C_{\bar{v}}(t) - C_a(t))\ , \tag{3.49}$$

Expiration:

$$\frac{\mathrm{d}}{\mathrm{d}t}(V_A(t).F_A(t)) = \dot{V}(t).F_A(t) + \dot{Q}_P(C_{\bar{v}}(t) - C_a(t))\ , \tag{3.50}$$

where:

$$\dot{V}(t) = \frac{\mathrm{d}V_A(t)}{\mathrm{d}t} + \dot{Q}_P(C_a(t) - C_{\bar{v}}(t))\ . \tag{3.51}$$

The model is also modified to include the transit times between the lung and the body compartments (and return). These delays have been included in several previously published models [25, 83, 127]. The mass balance equation for the body compartments becomes:

$$V_i^* . \frac{\mathrm{d}C_i(t)}{\mathrm{d}t} = \dot{Q}_i(C_{\bar{a}}(t - d_a) - C_i(t)) \ , \tag{3.52}$$

and the mixed venous return concentration is:

$$C_{\bar{v}} = \frac{1}{\dot{Q}} \sum_{i=1}^{n} \dot{Q}_i . C_i(t - d_v), \tag{3.53}$$

where d_a and d_v are the time delays associated with travelling through the arteries and veins, respectively. All other equations remain as previously described.

3.3.5 Solving the Model

All the models can be solved using a delay differential equation (DDE) solver that has the ability to cope with general delays, i.e. delays which change throughout the evaluation of the model. When using Matlab a 4th order Runge-Kutta method with residual control was used, *ddesd* [14, 177], with the default settings (relative tolerance of 10^{-3}). When a transitory compartment is not included within the dead space the delays do not change during evaluation so a more computationally efficient DDE solver for fixed delays was used. A Runge-Kutta DDE solver method based on the Bogack-Shampine (2,3) pair [40], was used: *dde23* [14, 178] (with the default settings).

When no body compartments and no transitory compartments are present no delays are present so an ordinary differential equation (ODE) solver can be used, which is even more computationally efficient. An ODE solver can also be used for the 'balloon on a straw' model. This is done by recognising that F_A is constant during expiration. As a result only past values of F_I are required, not F_A. Since F_I is an input and not a solution of the solver, all past values are available, allowing the model to be solved using an ODE solver. Several ODE solvers are available within Matlab [14, 179]. The use of a Runge-Kutta ODE solver based on the Bogack-Shampine (2,3) pair [40], *ode23* [14, 179], with a relative tolerance of 10^{-4} was found to be a good compromise between speed and accuracy, producing a similar output to the DDE solvers.

For the reasons described below, it was also necessary to use a DDE solver written in C. Solv95 [231] (a DDE solver for general delays) was used to solve all models. An initial time step of $10^{-4}s$ (which ensures a maximum time step of $10^{-2}s$) and a relative tolerance of $5\text{x}10^{-5}$, provided similar accuracy to the Matlab solvers.

When creating simulated data using these models (as described in Section 3.2.2), the breathing pattern was fixed. This means that the source of air exiting the compartment and the delay through the compartment are known. The breath breaks down into four stages:

- Phase one – from the beginning of inspiration until the volume of the transitory compartment has been inspired;
- Phase two – from the end of phase one until expiration begins;
- Phase three – from the beginning of expiration until the volume of the transitory compartment has been expired;
- Phase four – from the end of phase three until inspiration begins again.

Although it is possible to create a simulated flow signal and then to solve the models using the delays already described, it is significantly more computationally efficient to break the breath down into these four phases and to use the known delays for each phase to solve each stage in turn.

These can be derived in a similar way to the delays used by Gavaghan and Hahn [84], but for the new breathing pattern, resulting in the following delays:

Phase One ($T_n < t < T_n + t_{IR}$):

$$d_1 = t - T_n + t_E + \frac{1}{\gamma} \ln \left[\left(1 - e^{-\gamma t_E}\right) \left(\frac{t - T_n}{t_I} - 1 \right) + 1 \right] \ . \tag{3.54}$$

Phase Two ($T_n + t_{IR} < t < T_n + t_I$):

$$d_2 = \frac{V_R}{V_T} t_I \ . \tag{3.55}$$

Phase Three ($T_n + t_I < t < T_n + t_I + t_{ER}$):

$$d_3 = t - T_n - t_I - \frac{t_I}{1 - e^{-\gamma t_E}} \left(e^{-\gamma (t - (T_n + t_I))} - 1 \right) \ . \tag{3.56}$$

Phase Four ($T_n + t_I + t_{ER} < t < T_n + t_I + t_E$):

$$d_4 = t - T_n - t_I + \frac{1}{\gamma} \ln \left(\frac{V_R}{V_T^*} + e^{-\gamma (t - (T_n + t_I))} \right) \ . \tag{3.57}$$

The transition between the phases is determined by the times given in brackets above, with the transition between re-inspiring air and inspiring new air determined by:

$$t_{IR} = \frac{V_R}{V_T} t_I \ , \tag{3.58}$$

$$t_{ER} = -\frac{1}{\gamma} \ln \left(1 - \frac{V_R}{V_T^*} \right) \ . \tag{3.59}$$

3.3.5.1 Initial Conditions

In general the model is solved from the beginning of the sinewave input. Prior to that point the lung is manually brought up into equilibrium, so the inspired and expired concentrations are constant. Thus the end-expired concentration of the last breath is taken as the initial concentration for all lung compartments. When estimating parameters the initial end-expired alveolar volume is unknown and so is one of the parameters determined by the parameter estimation methods, described below. The model is then solved for the entire length of the data starting from the first inspiration of the sinewave. In a number of cases the initial conditions are also reset for each breath or the model is solved from the start of the data; as described in more detail later.

3.4 Parameter Estimation Techniques

Initially an improved breath-by-breath method is detailed. This method is generally used as a benchmark to which new methods proposed in the thesis are compared. These new methods are based on using the models described above to infer information about the lung of individual subjects. Two broad types of parameter estimation are employed: non-linear least squares optimization and ABC SMC.

3.4.1 Breath-by-Breath Method

This method was devised by collaborators in Oxford and determines the volumes associated with a homogeneous lung. It is based on the breath-by-breath methods described earlier in Section 2.4.4. The method consists of two stages; the first determines the volume of the dead space independently of the end-expired alveolar volume and the second recovers the end-expired alveolar volume using the dead space volume recovered in the first stage. Further details can be found in the patent relating to the work [158].

3.4.1.1 Dead Space Volume

Two methods of determining the dead space volume have been proposed, one based on the Bohr equation and the other on the Fowler method. The Bohr method is based on a rearranged version of Equation 2.3:

$$V_{D,n}.F_{I,n} - V_{D,n}.F'_{E,n} = V_{T,n}.(F_{\bar{E},n} - F'_{E,n}) \ , \tag{3.60}$$

for each breath n within the data. In reality however, $F_{I,n}$ is not constant throughout inspiration and therefore the component relating to the mass of tracer gas left in the dead space after inspiration, $V_{D,n}.F_{I,n}$, is replaced by an integral:

$$\int_{T_{E,n}-d_{V_D,n}}^{T_{E,n}} \dot{V}(t).F_I(t) \ \mathrm{d}t - V_{D,n}.F_{E',n} = V_{T,n}.(F_{\bar{E},n} - F'_{E,n}) \ , \tag{3.61}$$

where $T_{E,n}$ is the time at which expiration started and $d_{V_D,n}$ is the lowest value solution of the integral:

$$\int_{T_{E,n}-d_{V_D,n}}^{T_{E,n}} \dot{V}(t)\ \mathrm{d}t = V_{D,n}\ . \tag{3.62}$$

In order to determine $V_{D,n}$, first the right hand side of Equation 3.61 is evaluated for each breath. $V_{T,n}$ is determined by numerical integration of the expired flow. $F'_{E,n}$ is determined directly from the data by averaging the recorded value of the fractional concentration at the mouth over the last 20 data measurements of the breath (approximately $0.2s$). $F_{\bar{E},n}$ is determined using:

$$F_{\bar{E},n} = \frac{1}{V_{T,n}} \int_{T_{E,n}}^{T_{n+1}} \dot{V}(t).F_E(t)\ \mathrm{d}t\ , \tag{3.63}$$

which is evaluated numerically from the data.

$V_{D,n}$ is then determined for each breath by minimising the difference between the left hand side of Equation 3.61 and the value determined for right hand side. This is done by evaluating the left hand side of the equation for $V_{D,n}$ between $1ml$ and $V_{T,n}$ in $1ml$ steps with the first term evaluated by numerical integration. The estimated value of $V_{D,n}$ is taken to be the lowest value at which the right hand side minus the left hand side is of a different sign to that of the next value.

The Fowler method matched areas A and B, as shown in Figure 2.4, for each breath in turn through the following process:

- The total volume of air expired is determined by numerical integration of the flow signal using the rectangular method.
- The average level in Phase I is determined by taking the mean concentration between 30 and $60ml$ of expired volume[1].
- The slope in Phase III is determined by ordinary least squares linear regression of the concentration signal from the point at which 70% of the air has been expired to the end of the breath.
- Areas A and B are determined by numerical integration using the rectangular method for all time points in the breath.
- The dead space volume for a breath is the volume expired at the first time point in the data at which area B is larger than area A.

[1]This is required as the concentration of inspired tracer gas is not zero like in the standard Fowler breath test. Further details can be found in [77].

In both cases a global value for the dead space volume is then determined from the individual breath predictions. First any predictions which are either $150ml$ below or above the median value are removed. $y = x$ is then fitted using robust ordinary least squares linear regression, where the y values are the individual breath predictions and x is taken as the global prediction of V_D.

3.4.1.2 Alveolar Volume

The method for determining the end-expired alveolar volume using an insoluble gas starts with Equation 2.24, but with the solubility term removed:

$$V'_A.F_{A,n} = V'_A.F_{A,n-1} + V_D.F_{A,n-1} + (V_{T,n} - V_D).F_{I,n} - V_{T,n}.F_{A,n} \ . \tag{3.64}$$

In reality $F_{I,n}$ is not constant throughout the breath. Here it is replaced by the average inspired concentration, $F_{\bar{I},n}$, which is:

$$F_{\bar{I},n} = \frac{1}{V_{T,n} - V_D} \int_{T_n}^{T_{LM}} \dot{V}(t).F_I(t) \ \mathrm{d}t \ , \tag{3.65}$$

where T_{LM} is the first time after T_n which satisfies the expression:

$$\int_{T_n}^{T_{LM}} \dot{V}(t) \ \mathrm{d}t = V_{T,n} - V_D \ . \tag{3.66}$$

Also, $F_{A,n}$ is assumed to be equal to $F'_{E,n}$, as previously discussed in Section 2.4.4. Substituting these two terms into Equation 3.64 and rearranging gives:

$$V_D(F_{\bar{I},n} - F'_{E,n-1}) + V_{T,n}(F'_{E,n} - F_{\bar{I},n}) = V'_A.(F'_{E,n-1} - F'_{E,n}) \ . \tag{3.67}$$

This is in the form $y = a.x + b$, where: $y = V_D(F_{\bar{I},n} - F'_{E,n-1}) + V_{T,n}(F'_{E,n} - F_{\bar{I},n})$, $a = V'_A$, $x = (F'_{E,n-1} - F'_{E,n})$ and $b = 0$.

$F'_{E,n}$ and $F_{\bar{I},n}$ are determined from the data. These observations are assumed to follow a sinewave, so a sinewave is then fitted to the observations. The acquired function is used to calculate an adjusted value for $F'_{E,n}$ and $F_{\bar{I},n}$ for each breath. $V_{T,n}$ is determined by numerical integration of the flow signal during inspiration. These values, along with the final prediction of V_D made in the previous section, are used to determine x and y for each breath. Robust ordinary least squares linear regression is used to fit these observations to $y = a.x + b$ with a being the prediction of V'_A.

Two methods of removing outliers prior to fitting the linear function were proposed by collaborators. Investigations, in simulated data with noise and bench lung data, showed that these methods had very limited effect on the results and also that this was sometimes a negative effect. In noise free simulated data one of the methods was also shown to have a negative impact on the homogeneity index recovered from the most inhomogeneous lungs. Thus in the results presented later no outliers were removed.

3.4.1.3 Alveolar Volume and Pulmonary Blood Flow

When recovering parameters for soluble data, the solubility term can not be ignored and Equation 3.67 becomes:

$$V_D(F_{\bar{I},n} - F'_{E,n-1}) + V_{T,n}(F'_{E,n} - F_{\bar{I},n}) = V'_A.(F'_{E,n-1} - F_{E,n}) + \dot{Q}_P\lambda(F_{\bar{v},n} - F_{A,n})\Delta t_n \ . \tag{3.68}$$

This is in the form $z = ax + by + c$, where: $z = V_D(F_{\bar{I},n} - F'_{E,n-1}) + V_{T,n}(F'_{E,n} - F_{\bar{I},n})$, $x = (F'_{E,n-1} - F'_{E,n})$, $y = \lambda(\bar{F'}_E - F'_{E,n})\Delta t_n$, $a = V'_A$, $b = \dot{Q}_P$ and $c = 0$.

A sinewave is again fitted to $F'_{E,n}$ and $F_{\bar{I},n}$, but in this case the mean of the $F_{E',n}$ sinewave is adjusted to be equal to that of the $F_{\bar{I},n}$ sinewave. This removes any offset due to the body not being brought into equilibrium prior to the test, as detailed further in Section 4.7. The adjusted value for $F_{E',n}$ and $F_{\bar{I},n}$ are determined from these sinewaves and $F_{\bar{v},n}$ is assumed to equal the mean of these sinewaves. Δt_n is calculated directly from the data. These values, along with the $V_{T,n}$ and V_D (which were determined as above) are used to determine x, y and z for each breath. Robust ordinary least squares linear regression is used to fit the breath predictions to $z = ax + by + c$. Coefficient a and b are taken as the global predictions of V'_A and $\dot{Q}_P$.

3.4.2 Fitting the Entire Signal

3.4.2.1 Least Squares Parameter Estimation

As detailed in the previous chapter (see Section 2.7.1), using the Least Squares approach to estimate parameters of overdetermined non-linear systems is very common with a multitude of possible implementations. Within this thesis two approaches will be used: the Nelder-Mead simplex and a genetic algorithm. These methods are used because they are both non-derivative based methods and so less prone to errors caused by artefacts found in real data. The Nelder-Mead simplex has been found to be an efficient method of finding local minima, with many implementations available [32, 113].

However in order to find the global minimum other techniques are required. Within this thesis a genetic algorithm is used as it allows discrete parameters to be minimised. This is preferable when attempting to recover offsets between the two signals because continuous parameter estimation requires the signals to be interpolated. This could potentially add additional local minima, which all types of optimisation techniques can potentially get stuck in.

A object function the Root Mean Squared (RMS) error, E_{RMS}, between the model prediction and the actual expired breath output within the data is used:

$$E_{RMS} = \sqrt{\frac{\sum_{l=1}^{n}(x_l - m_l)^2}{n}}, \tag{3.69}$$

Only points during expiration are included with inspiration ignored.

This object function is minimised using the Matlab function *fminsearch* [14] which performs a minimisation using the Nelder-Mead simplex. This is an unconstrained multidimensional minimisation from a single starting point, which is detailed in the following chapters.

Although generally an unbounded minimisation will be used, the method can be bounded by transforming the parameter space. Function *fminsearchbnd*, available from the Matlab Central file exchange [69], performs such a transform allowing a bounded minimisation. As detailed later in the thesis this method is occasionally used to ensure that very low or negative volumes are not proposed.

The second method uses a genetic algorithm to minimise the cost function. The Matlab function *ga* [14] is used with lower and upper bounds for each parameter. In all cases, the default Matlab settings are used for the convergence and other criteria, unless otherwise stated.

3.4.2.2 Approximate Bayesian Computation

As discussed in section 2.7.2 of the previous chapter, there are several Bayesian parameter estimation techniques that could be applied. Here an ABC SMC approach has been implemented as the method is generally applicable to any kind of complex model without the need to determine the form of the likelihood function or make any assumptions about the underlining form of the distribution. Also the algorithm is generally applicable without making any substantial changes for the particular application.

In addition ABC SMC can be used to perform a model selection at the same time as estimating parameters for the selected models. This may enable a model for the underlying distribution to be determined (for example a uni-modal or bi-modal log-normal distribution) at the same time as the model parameters. Variations in predictions of the alveolar volume could be used as a measure of lung homogeneity. Thus use of a Bayesian approach in general may be beneficial as it may be possible to create a distribution for these predictions and so to gain information about their variability.

The ABC SMC algorithm used was chosen due to the availability of a tool, ABC-SysBio [119], produced by the Theoretical Systems Biology Group at Imperial College London. It also includes model selections within the standard set-up. The underlying algorithm used was first developed by Del Moral *et al.* [68], before being adapted to its current form by Toni *et al.* [198]. ABC-SysBio is written in Python and developed for use with models defined in the Systems Biology Markup Language (SBML), so does not

readily accept models of other formats; especially where the input and output signal are combined, as is the case here. In addition, the current implementation does not include a DDE solver. Therefore several modifications to the tool have been made for use with these particular data and models, without changing the underlying algorithm.

There is no readily available, reliably and currently maintained DDE solvers with general delays available for Python. Therefore the models have been implemented in C, using Solv95 [231]. Both ABC-SysBio and Solv95 have been modified to allow the models to be called from ABC-SysBio with various parameters. The Solv95 algorithm is then used to compute the results of the model which are returned for analysis by ABC-SysBio.

A uniform prior is used throughout this thesis with bounds reflecting a reasonable range for the parameters, as detailed further later. Once the priors have been set-up, ABC-SysBio is run using the automatic algorithm to reduce the tolerance, with the parameter alpha set to 0.9. The algorithm is terminated once the distributions of successive populations have converged. How best to detect convergence is investigated in the subsequent chapters. In addition to simply estimating parameters using a single model, ABC-SysBio is used with multiple models to preform a model selection. A uniform distribution for the model priors is used, this being the only option available within ABC-SysBio.

3.4.2.3 Estimating the Dead Space Volume Only

The ability to determine the dead space volumes independently relies on the assumption that the end-expired concentration is equal to the alveolar concentration, a widely applied assumption, which is used to solve the Bohr Equation (see Section 2.3.1) for example. This assumption is unlikely to introduce large errors for insoluble gases where the alveolar concentration is constant during expiration, but this assumption has also been applied to soluble gases where larger errors would be expected [58, 77, 122, 233, 234].

Using these assumptions it is possible to solve the models of the dead space (as described above in Sections 3.3.1.1 and 3.3.2) without having to specifically model the alveolar compartment. F_A in these models is simply set to the end expired fractional concentration during all periods of expiration. This has the advantage of allowing the volumes that make up the dead space to be determined independently of the alveolar compartment volume. The implications of making these assumptions is investigated in the subsequent chapters.

3.4.2.4 Post-Processing

The end-expired alveolar volume recovered by these methods is that for the point at which the initial conditions were set, usually the end of the last breath before the start of the sinewave. For simulated and bench lung data, the end-expired volume is the same for each breath, but for spontaneously breathing human subjects it will be different for each breath. Thus the end-expired alveolar volume is determined for all breaths by solving the model using the estimated parameters. The mean value is then added to the other lung compartment volumes in order to determine FRC, which is defined as the average end-expired lung volume (as detailed in section 2.2.1). In many cases, where the alveolar volume is governed by Equation 3.22, the end-expired alveolar volumes can be determined by solving that equation, rather than the full model.

3.5 Conclusion

Within this chapter the three types of data, which the following three chapters will address in turn, have been outlined. The models for lung ventilation that will be used in the analysis were also outlined. These models are adapted from the simple 'balloon on a straw' model in three ways: mixing is included within the dead space, the alveolar compartment is developed from the homogeneous case to two possible inhomogeneous representations and the implementation of body compartments, which are necessary for soluble gases, was described. How these models will be used to estimate parameters was also presented. These methods will be used in the remaining chapters to investigate the appropriateness of the models and other methods in fulfilling the aims of the thesis.

Chapter 4

Simulated Data

4.1 Introduction

The ultimate test of any method is how it performs on real data. However using a computational model to produce simulated data to which possible methods can then be applied is a common technique of assessing their viability, accuracy and robustness. Such data may also allow investigations which are not be possible in real data. Within this chapter simulated data will be used to assess methods for determining lung volumes and measures of ventilation-volume homogeneity.

In a number of cases, the simulated data will be created using a more complex model with a simpler model then used to recover the parameters. The ability of the simpler model, which will rely on particular simplifying assumptions, to recover particular parameters can then be assessed. Simulated data will also be used to investigate the effect that a known amount of noise will have on a particular method.

4.2 Simulated Data

Simulated data were created using the models of the lung described in the previous chapter. A single version of the dead space was investigated: a dead space, made up of a single mixing compartment at the mouth and a transitory compartment. Exhaustive investigation of all possible dead space configurations would be excessive here, especially as a large part of the chapter concentrates on inferring information about the alveoli compartment. The reason for choosing a model that includes a mixing compartment within the dead space is to represent the mixing that will come from several different sources, including regions of turbulence (due to major changes in diameter at the mouth and in the branching structure of the lung) and difference in the flow velocities at the centre of the bronchi and close to the walls. The possible limitation of modelling the dead space in this way will be discussed later in the thesis. Further in-

vestigation into various models for the dead space is presented later when investigating data from the artificial bench lung.

In most cases a mixing compartment volume (V_M) of $50ml$ and a transitory compartment (V_R) of $100ml$ will be used, resulting in a total dead space volume of $150ml$. This is consistent with the volumes used in previous research [84, 216, 218] and those found in the general population [79]. The proportion allocated to each of these two compartments is based on the early findings within the artificial bench lung, presented later, where mixing compartment volumes of approximately $50ml$ are observed. In some of the work described below a transitory compartment of $150ml$ is used. These simulations were carried out when limited data from alternative less robust sensors were available and the mixing compartment was thought to be present but not to contribute to the dead space volume. The implications of using these volumes will be discussed further below.

Unless stated otherwise, the total end-expired alveoli volume of $2.0l$ is used. The breathing pattern, unless otherwise stated, has a tidal volume (V_T) of $500ml$, decay constant (γ) of 1, inspired breath time (t_I) of $1.5s$ and expired breath time (t_E) of $3.5s$; this results in a breathing rate of twelve breaths a minute. These values are consistent with those used in simulations previously reported in the literature [84, 216, 218] and are within the ranges observed in the human population [154].

4.2.1 Experimental Error

Within real data experimental error is present. In order to investigate the natural variations found in the data noise is added to the simulated data. It is assumed that the noise is independent of signal strength and normally distributed with a zero mean. While this will not simulate all possible forms of error present, for example it will not account for any bias present on the signal, it will provide valuable information about how the methods behave in the presence of experimental noise. There are two signals present within the data which may experience some form of experimental error. An estimate of the noise present within these signals will be helpful in gauging the viability of the methods presented.

The flow sensor's data sheet [4] provides two pieces of information as to the noise associated with its measurements: flow precision is 3% and flow resolution is $3ml/s$. No details of what values the 3% relates to are given. Since it is unlikely to be the current reading as noise is generally interdependent of signal level, it is assumed here that it is 3% of the mid point of the flow meter range, $1.5l/s$; this would then be a precision of $45ml/s$. Assuming that approximately 95% of measurements are within

this range, then an estimate for the standard deviation of the noise would be $22.5ml/s$. Therefore the noise associated with the flow sensor will be assumed to have a typical standard deviation of $25ml/s$ ($1.5l/min$) throughout this chapter. There will be some additional noise associated with the $3ml/s$ resolution. This could be considered a form of continuous uniform noise to the limits of $\pm$ $1.5ml/s$, the standard deviation of which will be $0.87ml/s$. This is negligible compared with noise induced by the precision of the meter and so can be ignored.

The data sheet for the concentration sensor was not available, as the sensor has been discontinued. However, two similar sensors from alternative manufacturers quote the accuracy as "$\pm$ (2vol% + 2% of reading)" [2, 10] and the relevant standard requires an accuracy of "$\pm$ (volume fraction of 2.0% + 8% of gas level)" [8]. It is not clear if this is accuracy in the pure engineering sense, i.e. only concerned with the systematic errors of the sensor, or a form of overall measurement of the accuracy as the term can sometimes be used. In the later case this measurement of accuracy will contain an element of precision. However if this is the case and the sensor has similar properties to those currently available, it is clear from the data (see Figure 3.1a for example) that the precision only constitutes a small proportion of the accuracy value, as a signal with a 4 to 5% amplitude is clearly discernible.

Therefore an alternative method of estimating the random errors associated with the sensor is required. An estimate has been made using data from the homogeneous water displacement bench lung data (see Chapter 5 for details). When collecting this data the operator initially brings the concentration of the tracer gas throughout the lung into equilibrium. Once an equilibrium is reached there are periods of time, particularly during expiration from the alveolar compartment, where the concentration of the tracer gas passing through the sensor should be constant. In order to get an estimate of the variation 76 sections of constant concentration were identified manually in 26 of the 52 data sets; this amounted to 3723 data points (49 on average per section).

Figure 4.1 is a histogram showing the difference between the data points and the mean for that section. This distribution has a standard deviation of 0.012% and a mean of zero. A normal distribution with those parameters is also shown in Figure 4.1. This could be used as a rough estimate of the natural variation of the concentration sensor. The resolution of the sensor is around 0.045% and so has the potential to cause an underestimate by masking some of the natural variation, however the real concentration is likely not to be truly constant within the sections and so can potentially cause an overestimate. This also seems particularly low considering the accuracy quoted

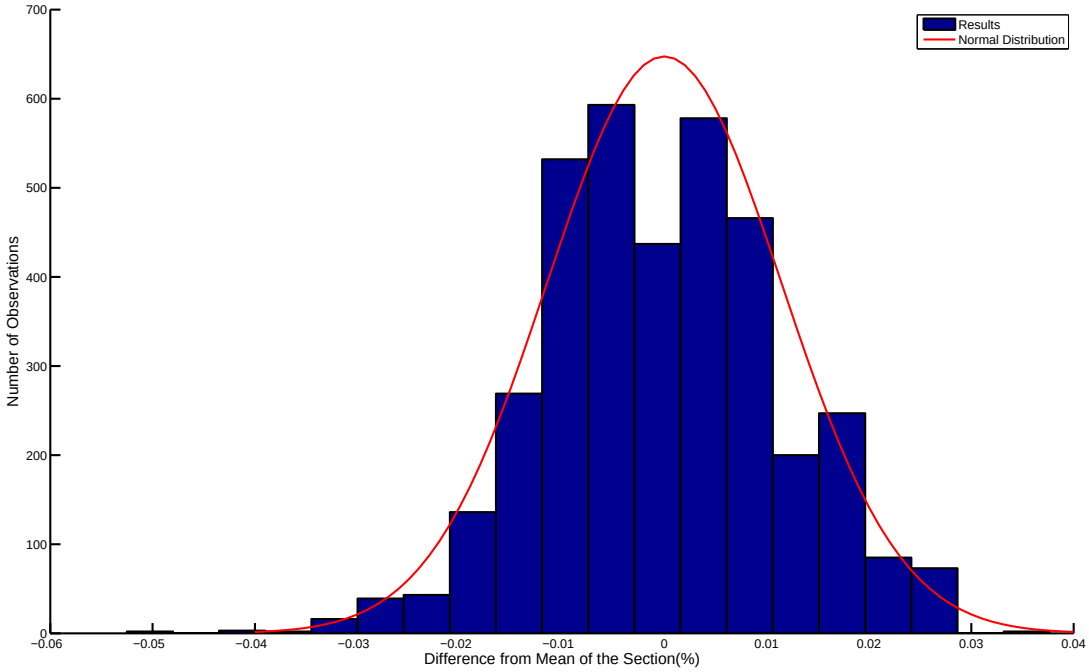


Figure 4.1: Histogram showing the difference between a data point and the mean concentrations during sections of assumed constant concentration within data from the homogeneous water-displacement bench lung. A normal distribution with a mean of zero and standard deviation of 0.012% is also shown.

for alternative sensors. Therefore the noise associated with the concentration sensor will be considered to be typically about 0.02% within this chapter. In the real data presented later the amplitude of the forcing sinewave is about 4 to 5%, thus where noise is added the amplitude of the concentration signal will be adjusted to 5% to ensure it is consistent with the noise levels used.

4.3 Log-Normal Distribution of Ventilation-Volume Ratios

Within this section the ability to determine information about the homogeneity of the lung is investigated. The data were simulated by using a continuous distribution of ventilation-volume ratios with a number of log-normal peaks for the make up of the alveolar compartment. Initially the homogeneity index proposed by Whiteley *et al.* [218] was investigated with the end-expired alveolar volumes determined using the breath-by-breath methods devised by colleagues.

In addition other methods of determining the homogeneity of the lung have been proposed. The first method proposed also uses the homogeneous model of the lung, proposing the use of the variation in alveolar volumes recovered at different breaths as

a measure of lung homogeneity. Methods of directly determining the underlying log-normal distribution that gave rise to the data are also investigated along with the ability of ABC SMC model selection to determine the number of peaks within a distribution. An insoluble tracer gas is assumed throughout this section, thus eliminating the need to account for the uptake and return of the tracer gas from the blood.

Throughout this section the homogeneity is varied by changing the standard deviation of the distribution. There is little evidence to help to define a typical range of distributions that reflect the lungs found in both healthy and diseased subjects. When simulating data previous publications have used σ_x from 0.0625 to 1.5[1] for uni-modal distributions [117, 137, 207, 216]. Lewis *et al.* [117] reported standard deviations on the log-scale from 0.30 to 2.04 when recovering distributions from both healthy subjects and patients with lung disease; normal young subjects exhibited standard deviations on the log scale of 0.53 ± 0.13 and cystic fibrosis patients exhibited 1.23 ± 0.24. However these distributions may not have been uni-modal and, as previously discussed, the ability of their method to determine the width of the distribution has been questioned in later research [216]. Within this section values of σ_x from 0.001 to 2 are used, with 0.001 included as a reference of what is essentially a homogeneous lung.

4.3.1 Homogeneity Index

The homogeneity index proposed by Whiteley *et al.* [216, 218] is investigated here. They observed that the end-expired alveolar volume recovered using a homogeneous model varied with the forcing period and proposed a homogeneity index based on this observation (see Equation 2.23). The various breath-by-breath methods devised by collaborators are used here to estimate the end-expired alveolar volumes, which are then used to determine the homogeneity index.

The homogeneity is varied by changing the standard deviation of the distribution while retaining the tidal and end-expired alveolar volumes of $500ml$ and $2.0l$. Mixed and transitory compartments of $50ml$ and $150ml$ are used, resulting in a total dead space volume of $200ml$.

4.3.1.1 Uni-Modal Distribution

Firstly the effect of changing the homogeneity of uni-modal log-normal distributions is investigated. Figure 4.2 shows the dead space and end-expired alveolar volumes recovered at various forcing periods and homogeneities using both the Bohr and Fowler methods of determining the dead space. No outliers were removed when determining

[1]Estimated from distribution 12 in Figure 1 of [207].

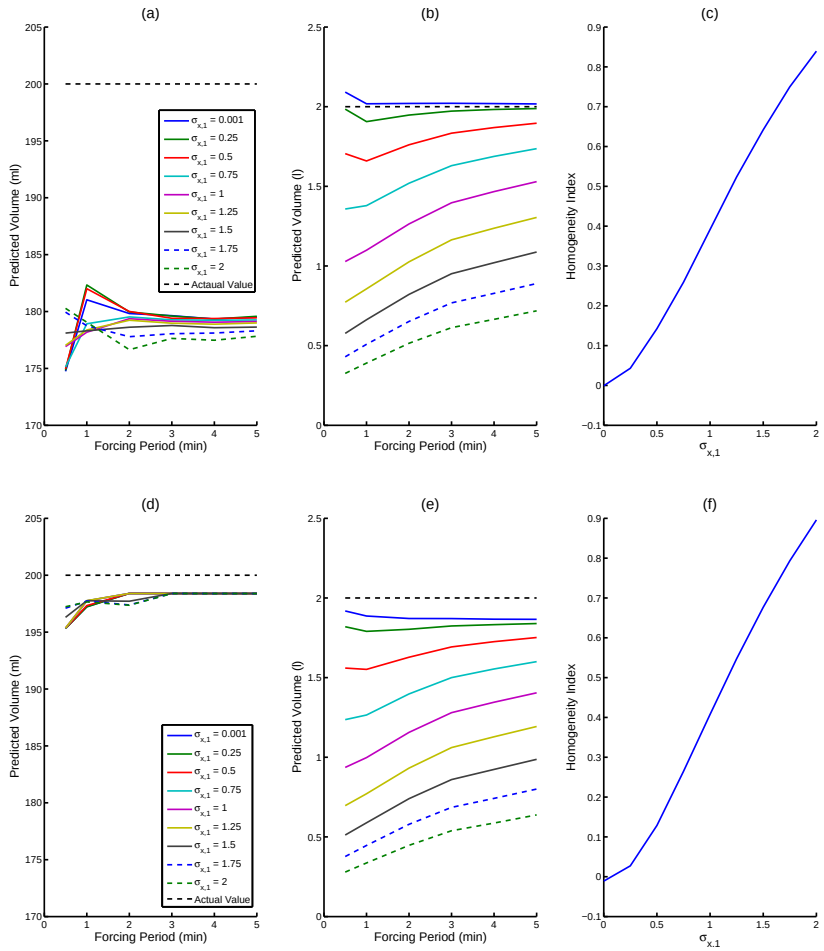


Figure 4.2: Dead space volumes as recovered using the (a) Bohr and (d) Fowler method. These dead space volumes are then used to recover the end-expired alveolar volumes by the breath-by-breath method, as shown in (b) and (e), respectively. (c) and (f) are the respective plots of the homogeneity indices. The simulated data used came from a uni-modal log-normal distribution of ventilation-volume ratios.

the end-expired alveolar volumes. The dead space volume recovered via both methods is relatively consistent for all forcing periods and homogeneities, with the slightly larger range of results observed for the Bohr method. It is underestimated by both methods. In the case of the Fowler method this is minimal with a maximum underestimate of approximately $5ml$, while for the Bohr method it is more pronounced at approximately $20ml$.

The end-expired alveolar volumes recovered vary with forcing period and homogeneity of the lung in both cases. The gradient of these plots increases noticeably with homogeneity, which is observed in an increase in the associated homogeneity index. In addition, the volume recovered at all forcing periods decreases with inhomogeneity. At low values of homogeneity ($\sigma_x = 0.001$ to 0.75) the volume recovered at a forcing period of $0.5min$ (and in one other case for $1min$ forcing period: $\sigma_x = 0.001$ by the

Fowler method), goes against the trend of the rest of the line, being noticeably higher than would be expected. This is not reflected in the homogeneity index (except slightly in the Fowler method case near zero) as the value relies on the volumes recovered at a forcing period of $1min$.

The other noticeable change with increased inhomogeneity is that the end-expired alveolar volume values recovered decrease at all forcing periods. At $\sigma_x = 0.001$, which is close to a homogeneous lung, the Bohr method recovers the actual end-expired alveolar volume ($2.0l$) at all forcing periods, except $0.5min$, while the Fowler method underestimates this by approximately $0.1l$. This underestimate in the results determined with the Fowler dead space method appears to be reflected in the end-expired alveolar volumes recovered at all homogeneities, with the results consistently $0.1l$ below the volume recovered when using the Bohr dead space method.

4.3.1.2 Experimental Error

Here the effect of adding noise to the simulated data is investigated. Initially the typical level of noise determined in Section 4.2.1 (normally distributed noise with standard deviation of 0.02% for the concentration sensor and $1.5l/min$ for the flow sensors) was added to each of the two signals. Twenty five repeated simulations were then carried out to get a indication of the variability of the results in the presence of such experimental error. Figure 4.3 show these results when the Bohr and Fowler methods were used to determine the dead space volume.

The dead space volumes recovered by the Bohr method are more variable than those recovered using the Fowler method. This was the case in the absence of noise however the variability when using the Bohr method increased to approximately 15 ml at $0.5min$ and to $10ml$ at other forcing periods. The presence of noise increased the average volume recovered by the Bohr method, so that it is closer to the actual volume. The variability of the dead space volumes recovered by the Fowler method has increased in the presence of noise, when compared to Figure 4.2, but not as significantly as the Bohr method. The average volume recovered is still approximately the same as that recovered when no noise was present, at around $197ml$.

The range of end-expired alveolar volumes recovered, using both the Fowler and Bohr Methods, is notably higher when the lung is more homogeneous. This results in a wider range for the homogeneity index for $\sigma_{x,1} = 0.001$ than at other levels of homogeneity. The range of results for $\sigma_{x,1} = 0.001$ and 0.5 when using the Bohr method is larger than when using the Fowler method but otherwise the two methods produce remarkably similar results for the end-expired alveolar volumes, which result

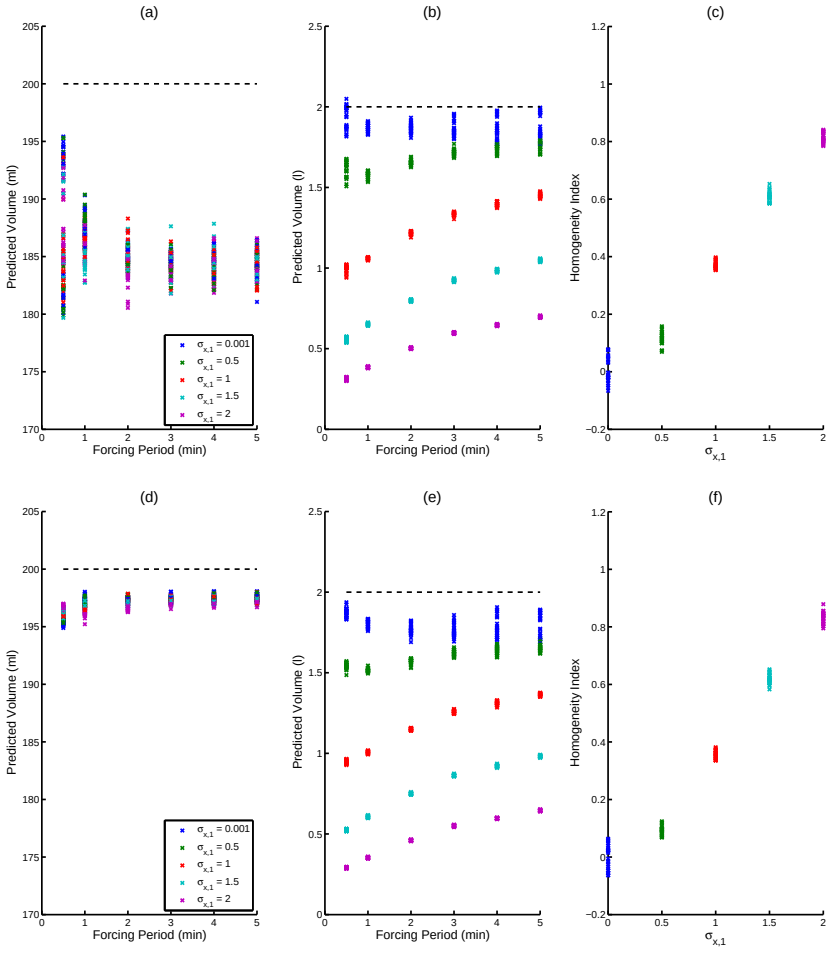


Figure 4.3: Complement volume predictions for 25 sets of data with the typical amount of noise added to both signals. Dead space volumes were recovered using the (a) Bohr and (d) Fowler method. These values were then used to recover the end-expired alveolar volumes by the breath-by-breath method with out any outliers removed as shown in (b) and (e), respectively. (c) and (f) show the respective homogeneity indices.

in similar ranges for the homogeneity index. There is no overlap in the homogeneity indices recovered at each distinct values of homogeneity, except at $\sigma_{x,1} = 0.001$ and 0.5 when both dead space methods are used.

Table 4.1 details the homogeneity index recovered in the presence of different levels of noise. In the presence of noise in just the concentration signal the standard deviation is small and the mean is close to that recovered when no noise was applied. When noise is added to the flow signal alone the results are remarkably similar to those observed when noise is added to both signals. These two observations suggest that noise on the flow signal is the main contributor to any errors in the homogeneity index. At a typical level noise, the standard deviation at $\sigma_{x,1} = 0.001$ is approximately double that observed at other homogeneities.

SD of Noise Conc. %	SD of Noise Flow l/min	Bohr Method $\sigma_{x,1} = 0.001$	$\sigma_{x,1} = 0.5$	$\sigma_{x,1} = 1.0$	$\sigma_{x,1} = 1.5$	$\sigma_{x,1} = 2.0$
-	-	0.00	0.14	0.39	0.64	0.84
0.02	-	-0.00 ± 0.001	0.14 ± 0.006	0.39 ± 0.002	0.64 ± 0.003	0.84 ± 0.004
-	1.5	0.04 ± 0.036	0.11 ± 0.016	0.37 ± 0.015	0.61 ± 0.011	0.82 ± 0.015
0.02	1.5	0.00 ± 0.044	0.12 ± 0.023	0.37 ± 0.013	0.61 ± 0.018	0.81 ± 0.017
0.04	3.0	-0.01 ± 0.030	0.11 ± 0.027	0.35 ± 0.017	0.60 ± 0.020	0.80 ± 0.021

SD of Noise Conc. %	SD of Noise Flow l/min	Fowler Method $\sigma_{x,1} = 0.001$	$\sigma_{x,1} = 0.5$	$\sigma_{x,1} = 1.0$	$\sigma_{x,1} = 1.5$	$\sigma_{x,1} = 2.0$
-	-	-0.01	0.12	0.41	0.67	0.90
0.02	-	-0.01 ± 0.001	0.13 ± 0.002	0.41 ± 0.001	0.67 ± 0.002	0.89 ± 0.004
-	1.5	0.01 ± 0.044	0.09 ± 0.016	0.36 ± 0.014	0.62 ± 0.016	0.83 ± 0.017
0.02	1.5	0.00 ± 0.043	0.09 ± 0.016	0.36 ± 0.014	0.62 ± 0.018	0.83 ± 0.020
0.04	3.0	-0.03 ± 0.023	0.09 ± 0.022	0.34 ± 0.018	0.60 ± 0.021	0.80 ± 0.030

Table 4.1: Mean and standard deviation of the homogeneity index determined from uni-modal log-normal simulated data in the presence of different levels of noise.

Doubling the amount of noise on both signal in general slightly increases the error in the mean value and increases the standard deviation of the results but to a minimal amount compared to the actual increase in noise. The mean of the homogeneity indices consistently reduces with the level of noise in all cases except when $\sigma_{x,1} = 0.001$. Similar mean and standard deviation values are observed when using both the Bohr and Fowler methods.

4.3.1.3 Bi-Modal Distribution

Here the effect of changing the homogeneity of the bi-modal log-normal ventilation-volume distributions is investigated. The first set of simulated data comes from a distribution with two peaks both with the same standard deviation ($\sigma_{x,1} = \sigma_{x,2} = 1.0$) and tidal volume ($V_{T,1} = V_{T,2} = 250ml$). The distance between the means ($\bar{x}_i$) is increased by changing the mismatch in end-expired alveolar volumes of each peak while keeping the total end-expired alveolar volume constant at $2000ml$. $V'_{A,1}$ is varied between $1000ml$ and $1800ml$, with $V'_{A,2} = V'_A - V'_{A,1}$. The resulting distributions are shown in Figure 4.4. The dead space volumes are then recovered via the Bohr method and used to determine the end-expired alveolar volume via the breath-by-breath method.

The results are shown in Figure 4.5, when the Bohr dead space method is used. The dead space volume predictions are approximately $20ml$ below the actual value. They do not vary greatly with forcing period but have a slight dip at lower forcing periods.

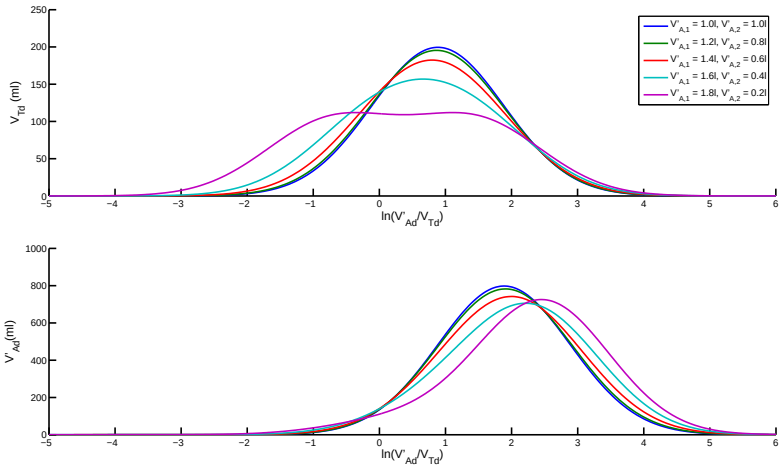


Figure 4.4: A continuous log-normal distribution with two peaks. The distance between the means of the two peaks is increased by changing the alveolar end-expired volumes associated with the peaks as detailed in the key. $\sigma_{x,1} = \sigma_{x,2} = 1.0$ and $V_{T,1} = V_{T,2} = 250ml$.

The end-expired alveolar volume predictions where the volumes associated with each peak are identical ($V'_{A,1} = V'_{A,2} = 1000ml$) are the same as those seen for the uni-modal distribution with $\sigma_{x,1} = 1.0$ (see Figure 4.2b). This would be expected as two smaller compartments with the same ventilation-volume ratio are equivalent to a larger one. As $V'_{A,1}$ increases and the distance between the means of the peaks increases, the overall V'_A predictions decrease and the difference between predictions for low and high forcing periods increase. This is in keeping with the effect of increasing $\sigma_{x,1}$ in the uni-modal case and results in an increase in homogeneity index.

A second set of simulated data was produced where the standard deviations of both peaks were increased while keeping the end-expired alveolar and tidal volumes constant ($V'_{A,1} = 250ml$, $V'_{A,2} = 1750ml$, $V_{T,1} = 400ml$ and $V_{T,2} = 100ml$; this makes $V'_A = 2000ml$ and $V_T = 500ml$). These distributions are shown in Figure 3.6. The volumes recovered are shown in Figure 4.6.

The predictions of dead space volume are again around $20ml$ below the actual volume. The range is larger than for the previous bi-modal distribution but around the level observed in the uni-modal case above and below $10ml$. The end-expired alveolar predictions are well below the actual value and the underestimation increases with the standard deviation of the peaks. There is a notable difference between the volume values recovered at different forcing periods, with larger volumes recovered at higher forcing periods. This difference appears to be roughly the same for all standard deviations of the peaks. This is reflected in a homogeneity index which does not change greatly with the peaks' standard deviation, but with a slight upwards trend between

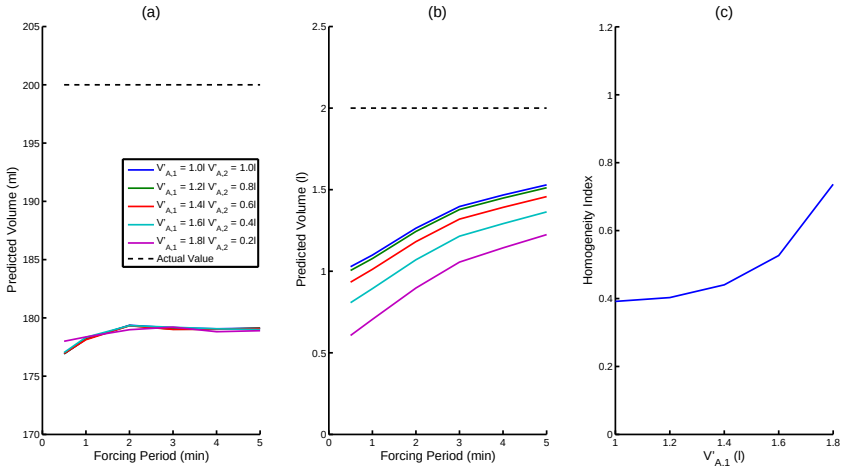


Figure 4.5: (a) Dead space volumes recovered using the Bohr method. These volumes are then used to recover the end-expired alveolar volumes by the breath-by-breath method as shown in (b). (c) is a plot of the respective homogeneity indices. The simulated data used came from a bi-modal log-normal distribution with different distances between the peaks.

$\sigma_{x,1} = 1$ and 1.75.

4.3.1.4 Discussion

These results show that the values of dead space volume recovered via both the Bohr and Fowler methods are not affected greatly by the forcing period or homogeneity. The dead space volumes recovered are closer to the actual volume (mixed plus transitory compartment) when using the Fowler method, however a larger underestimate in the end-expired alveolar volume recovered is then observed, which is not the case when using the Bohr method. The method of recovering the alveolar volume is the same, so it would appear that the underestimate in dead space by the Bohr method results in a corresponding improvement in the end-expired alveolar volume recovered. Neither method specifically models mixing within the dead space which is present in the simulated data, so anomalies in the dead space volumes recovered are not unexpected.

Both methods produce results that allow the homogeneity of the lung to be determined by observing the variation of the end-expired alveolar volume with forcing period. In this respect there is very little to warrant the choice of one method over the other. The Fowler method does result in a slightly higher range in the homogeneity index but it is slightly under zero for the near homogeneous lung. Both these factors are unlikely to be significant in real data.

In the uni-modal case, the cause of the alveolar volume recovered being higher than expected at a forcing period of $0.5min$ at low inhomogeneities is difficult to explain.

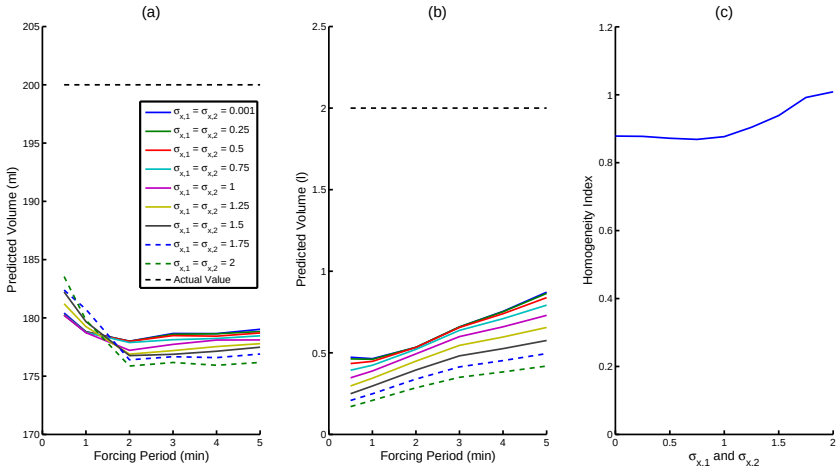


Figure 4.6: (a) Dead space volumes recovered using the Bohr method. These volumes are then used to recover the end-expired alveolar volumes by the breath-by-breath method as shown in (b). (c) is the plot of the respective homogeneity indices. The simulated data used came from a bi-modal log-normal distribution with different standard deviations.

It may in part be due to the slightly lower dead space volume values being recovered, but this is unlikely to be the main cause, especially as similar dead space volumes are recovered across the range of homogeneities in the Fowler case. Another factor may be the aliasing that will occur at such a small forcing period (there are only six breaths per period). This may affect the the fitting of sinewaves to the mixed and end-expired concentrations which occur as part of this method. The lower weighting that robust linear regression gives to outlying results could be another cause. Further investigations would be required to determine the contributions of these factors, if any.

For bi-modal distributions, increasing the distance between the means of both peaks has a similar effect on the plots of the end-expired alveolar volume against forcing period as widening the width of a uni-modal distribution; the volume recovered decreases at all forcing periods as the homogeneity of the distribution increases and the rate of this decrease is greater at lower forcing periods. As a result the homogeneity index increases. The increase in homogeneity index for a uni-modal distribution appears to be roughly linear; for this bi-modal distribution it appears to be of higher order, however this may be a fair reflection of the increased inhomogeneity within the distributions. The distributions are shown in Figure 4.4 and show a noticeable rate of increase in width with the change in $V'_{A,1}$, especially in relation to the ventilation (V_{Td}).

The change in the standard deviation of both peaks is not reflected in the homogeneity index to the same degree. Although there is some increase, this is modest and decreases slightly initially. However reference to a plot of the distributions (see Fig-

ure 3.6) does give a possible explanation into this; it is unclear from this as to which distribution is most inhomogeneous. For example it is hard easily to determine if the two distinct separated peaks in the case of $\sigma_{x,1} = 0.5$ or the broader almost single peak observed in the case of $\sigma_{x,1} = 1.5$ is more inhomogeneous. Interestingly the trend of lower end-expired alveolar volume recovered as the inhomogeneity increases is still observed although the rate of decrease in the volumes recovered is lower.

When noise is added to uni-modal data, the results indicate that although the dead space volumes recovered by the Fowler method are less affected by noise on the data, this does not feed through to the homogeneity index with both methods consistently showing similar results for the homogeneity index. In all cases the standard deviations of the homogeneity indices recovered for the near homogeneous case ($\sigma_{x,1} = 0.001$) were significantly larger than those at other levels of homogeneity, with standard deviations of generally about double that observed at other homogeneity levels. This is likely to mean that the ability to distinguish between values of homogeneity up to about $\sigma_{x,1} = 0.5$ is more difficult. However this may not be an issue in practice as Lewis *et al.* [117] results suggest that even young healthy subjects do not have perfectly homogeneous lungs. Distinguishing between higher levels of homogeneity at a resolution of $\sigma_{x,1} = 0.5$ should be possible at the noise levels presented. This is also true at double the typical noise level.

These results are not as favourable as those reported by Whiteley *et al.* [216, 218], but they only considered noise in the concentration signal. The results presented here show that the homogeneity index is more prone to error due to noise in the flow signal. Thus this is the likely cause of the less favourable results, rather than the different method used to calculate the end-expired alveolar volumes, which are in turn used to determine the homogeneity index.

The general underestimate in the end-expired alveolar volume with an increase in homogeneity is not a surprising result as it was also observed by Whiteley *et al.* [216, 218]. It is likely to be due to the nature of the log-normal distribution: as the standard deviation is increased while not changing the tidal or end-expired alveolar volume, the mean volume-ventilation ratio of the peaks shifts. The mean of the tidal volume shifts to the lower volume-ventilation ratios (i.e. to the left), while the mean of the end-expired alveolar volume shifts to the higher volume-ventilation ratios (i.e. to the right)[2]. This results in areas of the lung which have a high tidal volume but a low end-expired volume (i.e. they have a high time constant) and areas of the lung with

[2]This effect can be seen in the bi-modal distribution shown in Figure 3.6.

high end-expired volumes but which receive little tidal volume (i.e. they have a low time constant). Thus the lung appears smaller from a homogeneous point of view because there are parts which have little or no tidal volume and so will not be accounted for, while in those parts of the lung that do receive tidal volume the concentration changes quickly, so appearing to come from a smaller lung.

How well a log-normal distribution models the true distributions found within the human lung is discussed below. The results here are, however, consistent with observations made within real lungs. i.e. the FRC of the lungs decreases with the severity of disease when reported by many other techniques [77, 94, 95, 125, 136, 148, 153, 171, 213]. Hence the proposed use of FRC to monitor ALI in ventilated patients, which may also be possible using this technique.

These results support the use of the homogeneity index to determine information about the underlying homogeneity of the lung in the presence of typical experimental noise. However there are some limitations. Experimental error is likely to mean that it is difficult to distinguish between lungs with low homogeneities. Also in bi-modal distributions, some situations where the homogeneity index fails were identified. However it could be reasonably argued that even if homogeneity index does not reflect every change, it may still provide additional valuable information on which to base clinical decisions.

In addition the results support the use of the breath-by-breath methods as a method of determining the dead space and end-expired alveolar volumes, with the Fowler method providing more robust estimates of the dead space volume but the Bohr method resulting in better end-expired alveolar volumes in the near homogeneous case.

4.3.2 Inter-breath Variation

The homogeneous model of the lung is used here, with the possibility of using the variability in the end-expired alveolar volumes predictions as a measure of homogeneity being investigated. Two methods are proposed. The first uses the ABC SMC technique to make predictions of the end-expired alveoli volume, which results in a distribution of possible values. This distribution may vary with the level of homogeneity in the lungs.

The second method investigates the inter-breath variation in the predictions. Resetting the initial conditions after each breath allows a prediction of the end-expired alveoli volume to be made for each breath. This prediction is made by minimising the least squares error using the Nelder-Mead simplex. The variability in the predictions made may provide some indication of the homogeneity of the lung. Throughout it is

assumed that the volume of the dead space compartments it known. It is assumed that they can be determined independently of the homogeneity of the lung, using the methods detailed below in Section 4.5.

4.3.2.1 Uni-Modal Distribution

Data were created using the uni-modal continuous log-normal distribution of ventilation-volume ratios. Data with three different standard deviations (σ_x of 0.05, 0.50 and 1.00) are analysed. Figures 4.7 (a to c) show the continuous distributions that give rise to this data.

Figures 4.7 (d to f) show the distributions of the end-expired alveolar volume determined by ABC SMC. The convergence of these distributions was determined by manually monitoring the mean and standard deviation as shown in Figure 4.8. There is a very slight increase in the standard deviation with an increase in inhomogeneity, as detailed in the text, but this is not observable in the histograms. The mean of the distribution does change significantly however, reducing as the inhomogeneity increases. The tolerances (ϵ) required for a particle to be accepted at population 55 were: 1.14, 3.72 and 10.58 for σ_x values of 0.05, 0.50 and 1.00, respectively.

Using the least squares minimisation approach, an end-expired alveolar volume was recovered for each breath. A number of outlying results occur which skew the distribution extremely and so have a large effect on the mean and standard deviation. This is especially the case for $\sigma_x = 0.50$ and 1.00, where the maximum volumes recovered were $6.7\text{x}10^7 l$ and $4.3\text{x}10^9 l$, respectively. These outliers were removed by excluding any results were which below and above the lower and upper quartile by more than 1.5 times the inter-quartile range. This is the common definition of an outlier in the case of a box plot [80, 103, 200].

Figures 4.7 (g to i) show these results when a sinewave with a forcing period of $3min$ was used. The mean of the distribution decreases with an increase in the inhomogeneity, as previously seen. In addition the width of these distributions increases more notably with an increase in the inhomogeneity. However the distribution appears somewhat discontinuous with a number of spurious high peaks especially towards the edges of the histogram. A possible explanation for the spurious high peaks is that a sinewave input results in a number of breaths with similar characteristics over the repeated cycles, i.e. the breaths where the concentration within the alveolar compartment changes by a similar amount.

Therefore in order to provide a more broadband input signal a frequency modulated sinusoidal baseband signal (Equation 3.5), with a modulation and carrier period of

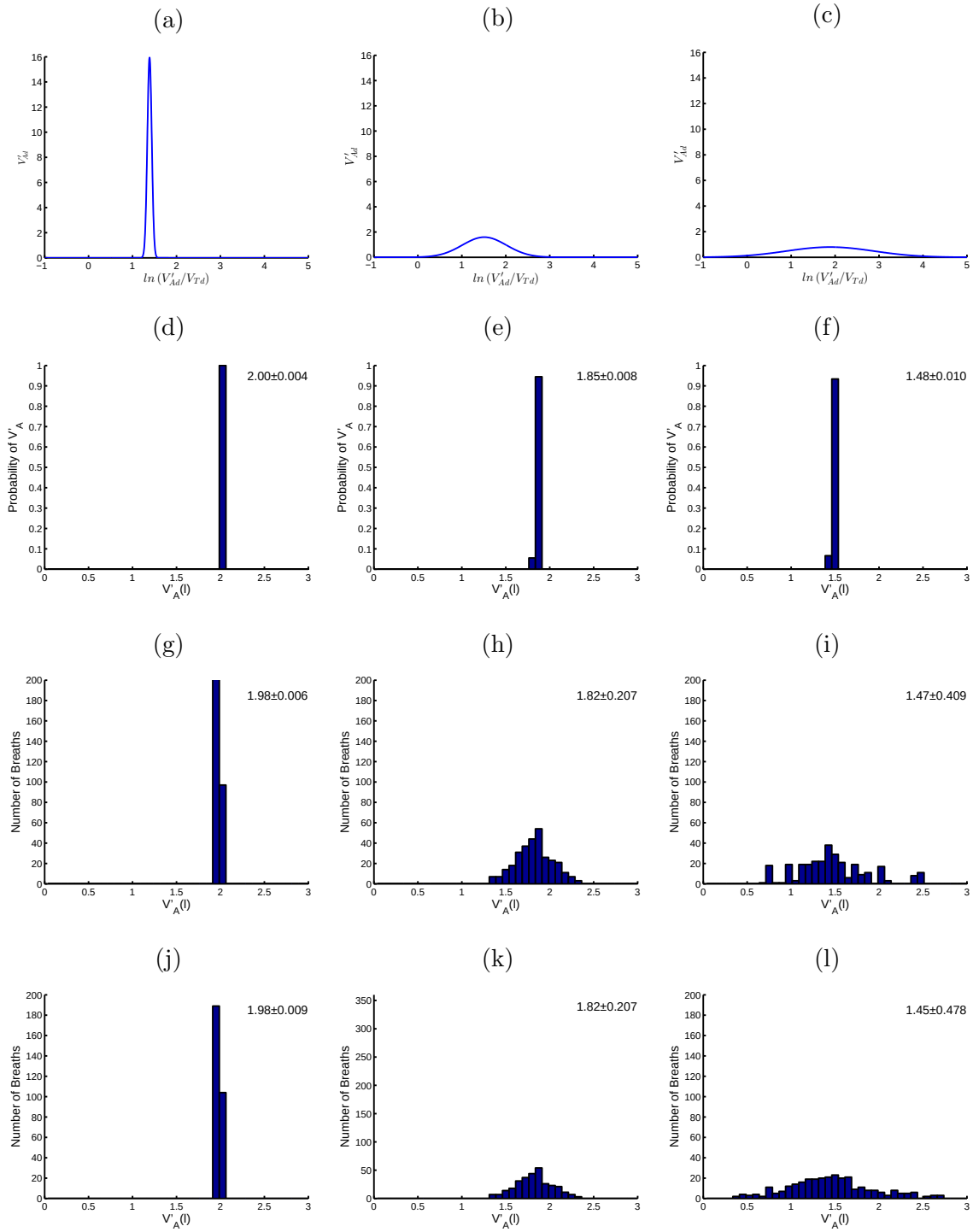


Figure 4.7: Recovery of the end-expire alveolar volumes using a homogeneous model of the lung from simulated data with a uni-modal log-normal distribution. The log-normal distributions with σ_x = (a) 0.05, (b) 0.50 and (c) 1.00 are shown. The histograms of the volumes recovered using ABC SMC are shown in (d) to (f). The histograms for the individual breath prediction using least squares minimisation by the Nelder-Mead simplex method are shown in (g) to (i) for a standard sinewave forcing function and in (j) to (l) for a modulated sinewave.

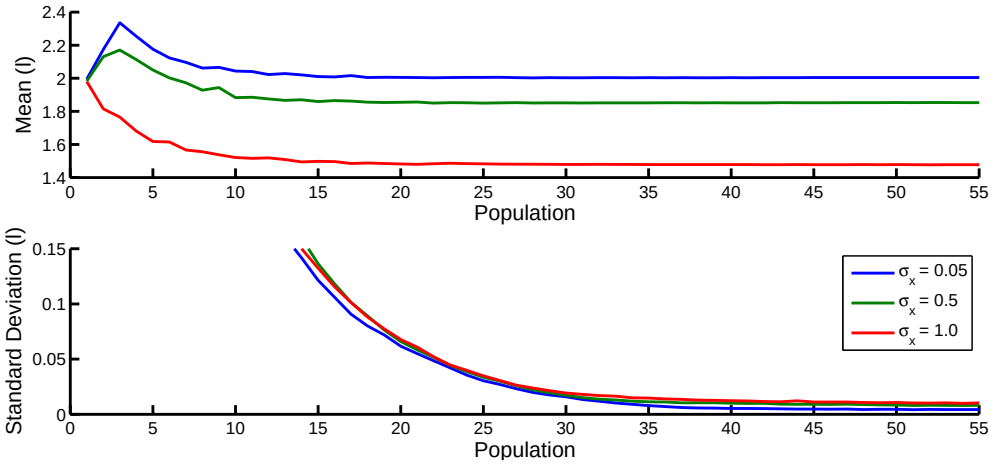


Figure 4.8: Mean and standard deviation of the homogeneous end-expired alveolar volume distribution recovered at successive populations of the ABC SMC algorithm. The simulated data came from a uni-modal log-normal distribution.

15.00 and 2.63$mins$, was used. This results in a wave form that starts at the origin and has peak to peak separation of between approximately 1.5 and 6min. This results in histograms with similar means and standard deviations but which are noticeably smoother across the range of predictions, as shown in Figures 4.7 (j to l).

4.3.2.2 Experimental Error

Experimental noise on both signals within the data has potential to affect the least squares minimisation approach in a similar way to an increase in the inhomogeneity of the lung; both are likely to increase the width of the distributions. In order to investigate this simulated data were created for the same three values of σ_x, using the same modulated sinewave as detailed above. Noise was then added to the two signals. The least squares minimisation approach was used to infer the end-expired alveolar volumes of each individual breath, assuming knowledge of the dead space compartment volumes. Outliers were removed using the same approach as above. This process was repeated twenty five times with a unique set of noise each time.

Figure 4.9 shows the resulting histograms in the first of these repeats, for noise of 0.02% and 1.5l/min on the concentration and flow signals, respectively. The noticeable change is an increase in the width of the distribution for $\sigma_x = 0.05$, the mean however has not changed to any significant extent. The other two distributions have the same means and have not changed greatly in width or shape, although the distributions are slightly smoother.

Table 4.2 provides a more detailed summary of the effect of noise. The addition of noise has little effect on the mean of the recovered distributions. The means are close

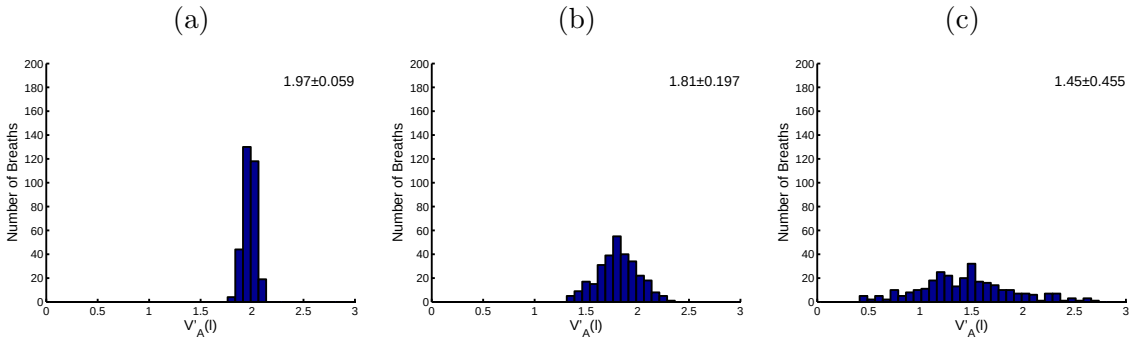


Figure 4.9: Recovery of the end-expired alveolar volumes using a homogeneous model of the lung in the presents of noise. The simulated data came from uni-modal log-normal distribution with σ_x values of : (a) 0.05, (b) 0.50 and (c) 1.00. The noise on the concentration and flow signals was 0.02% and $1.5l/min$, respectively.

SD of Noise		Recovered			
Conc. %	Flow l/min	Distribution Property	$\sigma_{x,1} = 0.05$	$\sigma_{x,1} = 0.5$	$\sigma_{x,1} = 1.0$
-	-	Mean (l)	1.98	1.82	1.45
		SD (l)	0.01	0.20	0.48
0.02	-	Mean (l)	1.98 ± 0.003	1.82 ± 0.005	1.45 ± 0.008
		SD (l)	0.04 ± 0.003	0.20 ± 0.007	0.50 ± 0.015
-	1.5	Mean (l)	1.98 ± 0.002	1.81 ± 0.004	1.46 ± 0.006
		SD (l)	0.03 ± 0.002	0.21 ± 0.005	0.43 ± 0.013
0.02	1.5	Mean (l)	1.98 ± 0.004	1.81 ± 0.006	1.45 ± 0.007
		SD (l)	0.05 ± 0.003	0.20 ± 0.006	0.47 ± 0.010
0.04	3.0	Mean (l)	1.97 ± 0.006	1.81 ± 0.008	1.45 ± 0.012
		SD (l)	0.11 ± 0.005	0.22 ± 0.011	0.46 ± 0.018

Table 4.2: Mean and standard deviation for the mean and standard deviation of the distributions of individual breath predictions recovered using a homogeneous model in the presents of noise. The simulated data came from a lung with a uni-modal log normal distribution.

to those recovered when no noise is present for all homogeneities and noise levels. The standard deviation of these means is also low, $0.012l$ at maximum, which is less than 1% of the mean.

The standard deviations of the recovered distributions are affected to a greater degree by the noise. For $\sigma_{x,1} = 0.05$, the means of the standard deviations for the 25 recovered distributions increase with the level of noise to $0.11l$ at double the typical noise level (11 times that when no noise is present), but the standard deviation is low at $0.005l$. In contrast for $\sigma_{x,1} = 0.5$ and 1.0, the means are little affected by the noise but the standard deviations are larger, up to $0.018l$ at maximum, but this is still relatively low at 3.9% of the mean. The effect of noise in each of the two signals alone seems

to be comparable, with little difference in the means and standard deviations of the recovered distributions.

4.3.2.3 Discussion

ABC SMC does not reflect the increased inhomogeneity with an noticeable increase in the uncertainty of the predictions of the end-expired volume. Therefore the use of the distribution's width as an indirect measure for the homogeneity of the lung is not viable. In general this may also present some slight concerns about the ABC SMC method as more uncertainty about the results might reasonably be expected. However using a recovery model which does not match the model which produced the data does make it hard to predict how this method will react. It may also indicate an ability to reliably determine the reduction in end-expired alveolar volume and so an ability also to reliably determine the reduction in FRC associated with inhomogeneity.

The tolerances achieved at population 55 in each case give some indication of a possible explanation. The tolerances achieved increase as the inhomogeneity increases, meaning that the model predictions of the accepted particle are further from the simulated data. This indicates that the homogeneous model is not fitting the data as well when the inhomogeneity increases, as would be expected, but that this does not result in a change in the width of the distributions. The likely explanation for this is that although an increase in homogeneity effects the absolute value of the minimum tolerance, the profile of tolerances around that minimum point is relatively unchanged, thus resulting in distributions with similar widths.

The least squares minimisation approaches in contrast are better at capturing the changes in homogeneity through an uncertainty in the predictions. They result in a distribution which increases in width with inhomogeneity, thus the distribution appears to mimic the underlying ventilation-volume distribution in terms of width and to some extent shape, although not being strictly related to each other. A smoother distribution which mimics the underlying distribution shape better is achieved by using a modulated sinewave, although some distributions still appear to have a number of somewhat spurious peaks. A smoother distribution may be achieved by increasing the broadband nature of the forcing wave.

One disadvantage to this method is that experimental error will also result in a widening of the distribution. However, as with any method, the real issue is whether the effect of experimental error is great enough to obscure the change that is being observed. These results indicate that it is likely to be possible to determine change in homogeneity shown in the presence of noise. The difference in the mean of the 25

standard deviations of the recovered distributions in the case of $\sigma_{x,1}$ of 0.5 and 1.0 is roughly comparable to that seen for the homogeneity index, but the standard deviation of the 25 results is lower, particularly in the case or $\sigma_{x,1}$ of 0.5. This would indicate that this method is less susceptible to noise. However knowledge of the dead space make up has been assumed here, but not in the breath-by-breath method. Although methods of independently determining this information are discussed further below, errors in determining this information are likely to follow through into determining the end-expired alveolar volume of each breath.

In common with the homogeneity index, the identification of slight changes in lung homogeneity in lungs which are near to being homogeneous may be more difficult, as noise in the sensor signals appears to increase the absolute value at low homogeneities. This method is also unlikely to be so easily applicable to soluble gases where the concentration of the alveoli compartment will change during the breath, as the use of the end expired concentration to reset its initial conditions is less likely to be valid.

All three methods presented here result in similar mean results that decrease as the homogeneity increases. This is a similar observation to that observed when using the breath-by-breath methods and might be used as an alternative measure of homogeneity.

4.3.3 Directly Recovering Volume-Ventilation Distribution

So far a homogeneous model of the lung has been used to indirectly infer information about the homogeneity of the lung. Here methods which specifically model the inhomogeneities within the lung are proposed and investigated. Attempts are made to infer information from data of lungs made up of uni-modal log-normal and bi-modal distributions. In addition, the possibility of using model selection to identify the type of distribution that was most likely to give rise to a given set of data is investigated.

4.3.3.1 Uni-Modal Distribution

Here the results of inferring a uni-modal log-normal distribution are presented. The data were simulated using a single peak with parameters: $V'_{A,1} = V'_A = 2000ml$, $V_{T,1} = V_T = 500ml$ and $\sigma_{x,1} = 1.0$. ABC SMC was then used to estimate between two and five parameters. Initially the two parameters recovered were the end-expired alveolar volume and standard deviation of the distribution, as it is unlikely to be possible to determine them by other means. The ability to also recover the dead space volumes (mixed and then transitory compartment) was then investigated as it may be possible to determine these independently of the alveoli make up, as described further below. Finally the tidal volume was introduced, as it is relatively easy to determine by integrating the flow signal.

One Peak Model

Parameters	Distribution	Lower Limit	Upper Limit
V'_A (ml)	Uniform	1500	2500
$\sigma_{x,1}$	Uniform	0.5	2.5
V_M (ml)	Uniform	35	65
V_R (ml)	Uniform	100	200
V_T (ml)	Uniform	400	600

Two Peak Model

Parameters	Distribution	Lower Limit	Upper Limit
V'_A (ml)	Uniform	1500	2500
$V'_{A,2}$ (ml)	Uniform	0	1125
$V_{T,2}$ (ml)	Uniform	0	500
$\sigma_{x,1}$	Uniform	0.5	2.5
$\sigma_{x,2}$	Uniform	0.5	2.5

Table 4.3: Priors used for ABC SMC parameter estimation and model selection when recovering log-normal distributions.

In all cases it was assumed that the flow signal follows the pattern defined for simulated data (see Section 3.5) and that the three parameters that define that pattern (t_I, t_E and λ) are known. This is likely to be a reasonable assumption for ventilated patients and it should be possible to determine these parameters with relative ease from the flow signal and knowledge of the ventilator settings, however it is unlikely to be an assumption that can be made for spontaneously breathing subjects. A uniform prior is used for of all parameters with the limits detailed in Table 4.3.

As can be seen in Figure 4.10, the means of the recovered distribution are good predictions of the actual value no matter how many parameters are being predicted. The error in the mean is however significantly greater when five parameters are predicted and the ability to make a good prediction may in part be due to the fact the distributions are flat and the limits of the prior are symmetrical about the mean. The spreads of the distributions are relatively small when predicting up to four parameters but greatly increase when all five parameters are considered with the edges of the distributions coinciding with the limits of the priors, except in the case of $\sigma_{x,1}$. Interestingly the standard deviations of $V'_{A,1}$ and $\sigma_{x,1}$ when two parameters are predicted are about an order of magnitude greater than when three or four parameters are predicted.

The progress through successive populations towards convergence is shown in Figure 4.11. It is clear that in the case where five parameters are inferred none of the distributions, except that for $\sigma_{x,1}$, have converged. This is despite processing time in excess of 60 days being applied to the task. When two parameters were inferred it

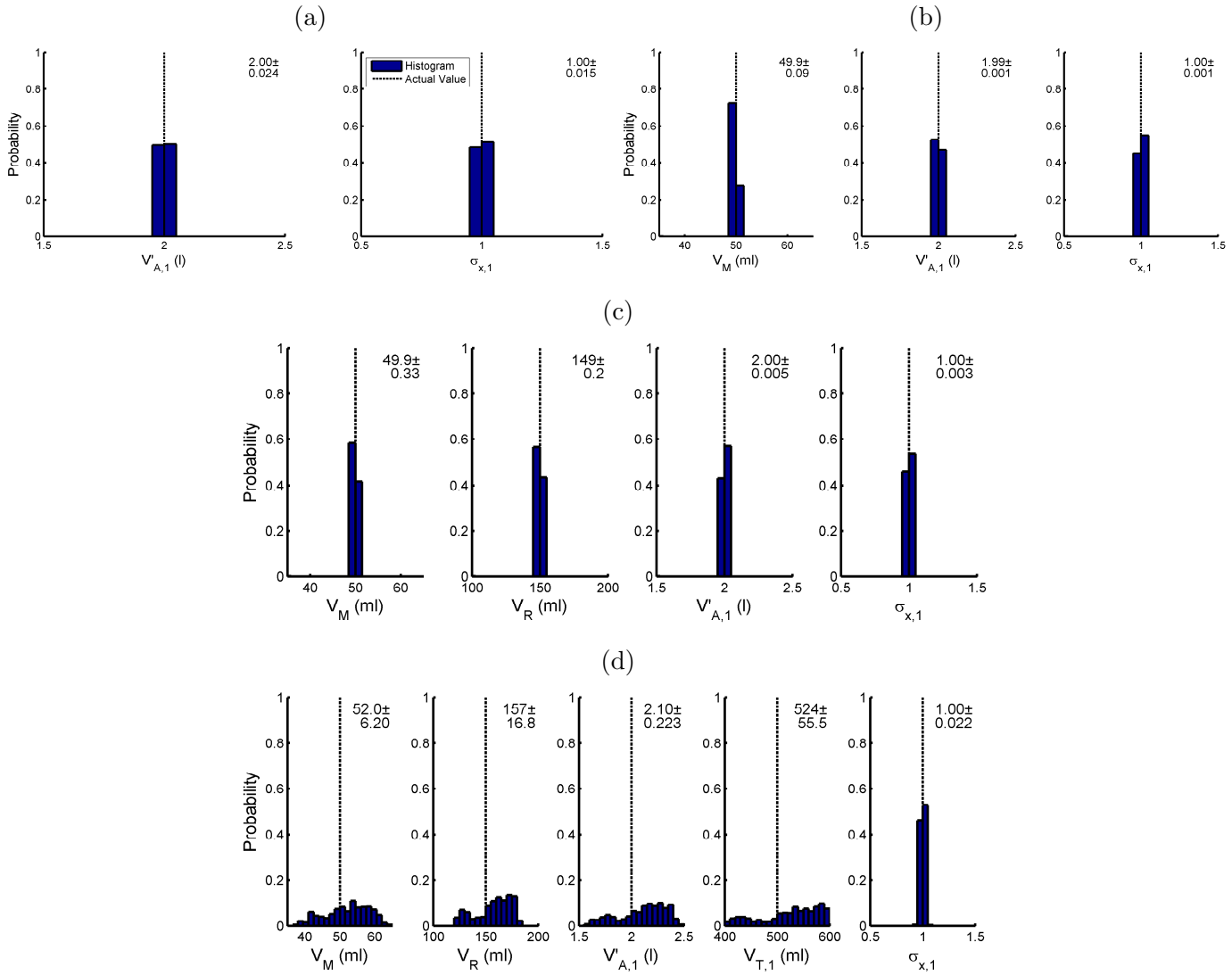


Figure 4.10: Distributions inferred for each parameter from simulated data with a uni-modal log-normal distribution using ABC SMC. Number of parameters inferred: (a) two (b) three (c) four and (e) five. In the five parameter case convergence was not achieved after 60 days of processing time, so the non-converged results of the final population are shown here.

took 49 populations (less than 1 day) before all parameter distributions were assessed manually to have converged. For three parameters it took 141 populations (4 days), but convergence could have potentially been assessed to have been reached at a earlier stage. Four parameters took 160 populations (13 days).

4.3.3.2 Bi-Modal Distribution

A log-normal distribution with two peaks was set up with the following parameters: $V'_{A,1} = 1.5l$, $V_{T,1} = 200ml$, $\sigma_{x,1} = 1.0$, $V'_{A,2} = 0.5l$, $V_{T,2} = 300ml$ and $\sigma_{x,2} = 1.0$. This results in total volumes of: $V'_A = 2000ml$, $V_T = 500ml$. In order to reduce the number of parameters to be inferred, knowledge of the mixed and transitory dead space volumes

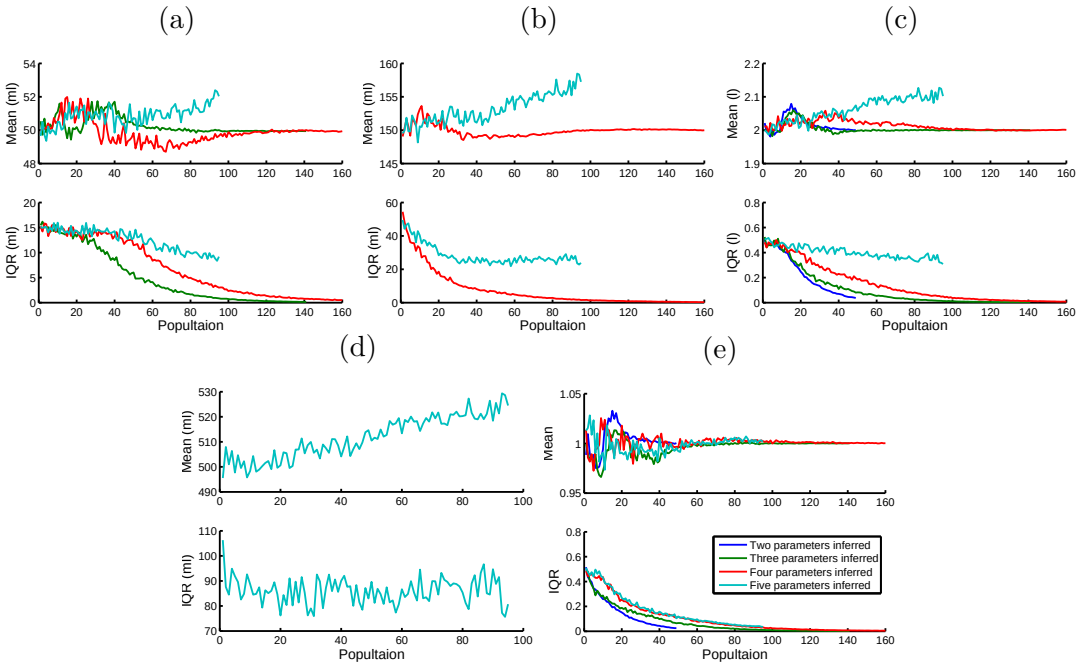


Figure 4.11: Convergence of each parameter when using ABC SMC, for simulated data with a uni-modal log-normal distribution. Parameters inferred: (a) V_M (b) V_R (c) $V'_{A,1}$, (d) $V_{T,1}$ and (d) $\sigma_{x,1}$.

is assumed. Also it was assumed that the total tidal volume can be determined directly from the flow signal, thus only the $V_{T,i}$ associated with one peak had to be inferred, for example $V_{T,2}$, with $V_{T,1} = V_T - V_{T,2}$. A similar approach was taken for V'_A with the total volume and the volume of the second peak being inferred. The priors were set-up as detailed in Table 4.3.

As can be seen in Figure 4.12, a wide distribution is recovered for all parameters, which extends to the limits of the prior in most cases. This is reflected in the large standard deviations. The mean value of the $V'_{A,1}$ and $\sigma_{x,1}$ distributions are close to the actual value, but this is not the case for the other distribution. The difficulties in inferring these parameters may be because of an interaction between the end-expired alveolar and tidal volumes of both peaks. An ABC SMC parameter estimation was also carried out for a narrower prior around the actual volume of $V_{T,2}$, which did not include the actual volume of $V_{T,1}$. Similar results were seen with wide distributions and no better mean predictions.

Figure 4.13 shows the mean and interquartile range of the parameters over the populations evaluated. This shows that the parameter distributions have not converged. This is despite over 70 days of computational effort being applied. Although the mean values are relatively stable the variation in interquartile range is still high, especially for

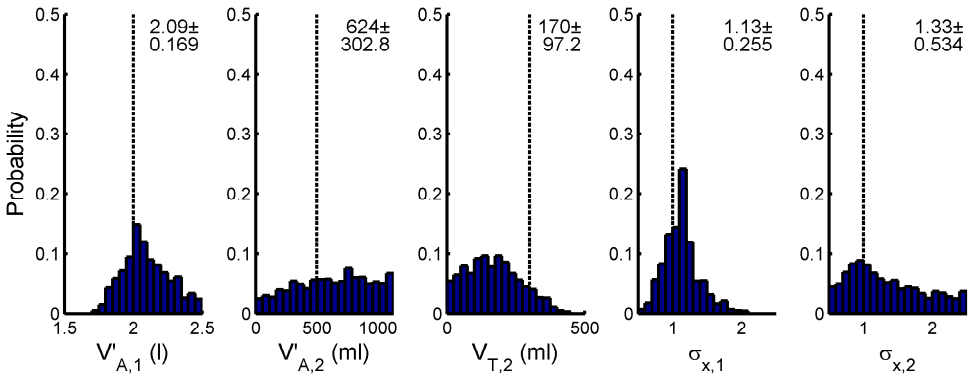


Figure 4.12: Distributions inferred for each parameter using ABC SMC, when the ventilation-volume distribution is bi-modal. The actual value of the parameters are: $V'_A = 2000ml$, $V'_{A,2} = 500ml$, $V_{T,2} = 300ml$, $\sigma_{x,1} = 1.0$ and $\sigma_{x,2} = 1.0$. The vertical dashed line indicates the actual value.

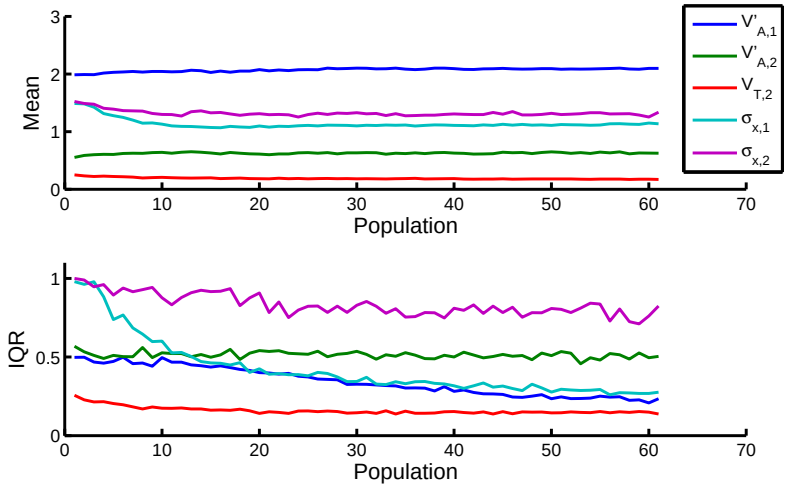


Figure 4.13: Convergence of each parameter when using ABC SMC, for simulated data with a bi-modal log-normal distribution. All volumes are in litres.

$\sigma_{x,2}$ and $V'_{A,2}$. The interquartile ranges of $\sigma_{x,1}$ and $V'_{A,1}$ also appear to have a downward trend.

4.3.3.3 Identifying the Distribution

Whether it is possible to determine the number of peaks within a distribution using ABC SMC model selection was next investigated. A model selection was carried out between two possible models: a model with one peak, which only inferred the two parameters ($V'_{A,1}$ and $\sigma_{x,1}$) and a model with 2 peaks, which inferred five parameters (V'_A, $V'_{A,2}$, $V_{T,2}$, $\sigma_{x,1}$ and $\sigma_{x,2}$). The make-up of the dead space and the breathing pattern, including the total tidal volume, was assumed to be known. The priors used for each parameter were the same as the two previous sections and detailed in Table

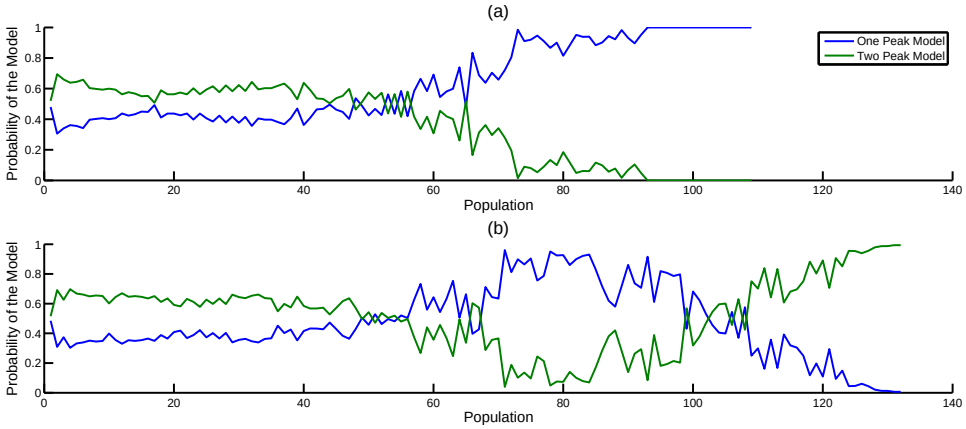


Figure 4.14: Model selection between two models including one or two log-normal peaks performed on simulated data from a lung with (a) a uni-modal log-normal distribution and (b) a bi-modal log-normal distribution.

4.3. The prior assumption that both models are equally likely is made (the only option in the ABC-SysBio implementation).

The two sets of simulated data from the previous two sections were used (i.e. one originating from a lung with a uni-modal distribution and one originating from a bi-modal distribution). The results of the model selection are shown in Figure 4.14. Ultimately the correct model is chosen with 100% certainty in both cases. In the case of the uni-modal distribution the two peak model is initially favoured before the one peak model rises in probability and is ultimately chosen with full certainty at population 94 ($\epsilon = 0.71$). In the bi-modal distribution case, initially the two peak model is favoured before the probability falls substantially to 0.04. It then rises again to be selected with full certainty at population 133 ($\epsilon = 12.58$). There is a relatively large variation in probability between one population and the next throughout both model selections.

4.3.3.4 Discussion

The results presented here show that it is likely to be possible to infer a uni-modal log-normal distribution assuming that the tidal volume is known, which is not unreasonable as the tidal volume is relatively easily determined from the flow signal. The means of the distributions are good predictors of the actual value and the standard deviations are small. It may be that tidal volume could also be determined using this method given sufficient computational power for the results to converge. However this was not achieved despite a considerable amount of computational power being expended and the prior distribution being relatively narrow for a number of the parameters.

It is interesting to note that the standard deviations of both distributions in the two parameter recovery case are larger than when three or four parameters are determined. This may be due to the algorithm being terminated too soon in the two parameter case. On the scale plotted in Figure 4.11 and when plotted against the other results, their does appear to be some potential for additional convergence. The periodic and subjective nature of determining the convergence manually may have played a factor here. Thus this may be resolved by using a automatic convergence algorithm, as will be discussed later in the thesis.

There is still potential for parameter estimations with more parameters to continue for longer resulting in narrower distributions. This is evident by looking at the results for four parameters. On the basis of the results for $V'_{A,1}$ and $\sigma_{x,1}$ convergence could have been reasonably determined as being achieved by population 120 by either a manual or automatic method, but the need to wait for V_M also to converge means waiting till a later population. This means the width of the $V'_{A,1}$ and $\sigma_{x,1}$ distributions can reduce. A potential cause for this is the fact this is simulated data with no experimental error and so there is a solution to the recovery model that will fit the results exactly. Therefore the distribution should theoretically, given enough time, converge on one result with no width in the distribution. This is a situation that would not be true in real data.

The results for a bi-modal distribution are not so favourable. This is unlikely to be because ABC SMC is being disrupted by the possibility of selecting parameters for the wrong peak because a narrower prior for $V_{T,2}$ did not improve the results. It is likely that the presence of two peaks has made it difficult or possibly impossible to infer the parameters that make up the distribution. Continuing the estimation for longer may result in better predictions, however considering that each population normally takes longer to compute, gaining much better results in this way is unlikely to be feasible. This means that inferring a bi-modal log-normal distribution in this way is unlikely to be possible.

Although the parameters that made up the bi-modal distribution can not be inferred, these results at least imply that it should be possible to determine a bi-modal distribution from a uni-modal distribution using model selection. The correct models were ultimately selected with full certainty in both model selections, which may be plausible given the fact these are simulated data with no noise added.

4.4 Multi-Compartment Lungs

Within this section many of the investigations carried out for data from a log-normal distribution are repeated for data from multi-compartment lungs. Data were simulated

	Compartment 1			Compartment 2		
Data Set	$V'_{A,1}(ml)$	$V_{T,1}(ml)$	$V'_{A,1}/V_{T,1}$	$V'_{A,2}(ml)$	$V_{T,2}(ml)$	$V'_{A,2}/V_{T,2}$
A	1080	195	5.54	1080	195	5.54
B	1270	195	6.51	900	195	4.62
C	1480	195	7.59	690	195	3.54
D	1480	173	8.55	680	217	3.13

Table 4.4: Details of the multi-compartment lung configurations.

for the four multi-compartment lungs detailed in Table 4.4, which are equivalent to the multi-compartment bench lung setups investigated later, in Section 5.3. Mixed and transitory compartment volumes of 50 and 100 ml were used.

4.4.1 Homogeneity Index

First results for the homogeneity index are presented. The dead space volumes were recovered using both the Bohr and Fowler methods, which were then used to determine the end-expired alveolar volumes via the breath-by-breath method. The results are shown in Figure 4.15.

The dead space volumes recovered using the Bohr method show some variation with forcing period. In all cases the volume recovered at a $30s$ forcing period is significantly below the other volumes, however the range of values is still below $10ml$. There is, however, little variability in the dead space volumes recovered for the four different multi-compartment lungs with near identical plots in all cases. The volume recovered is approximately $20ml$ under the actual volume. In comparison the dead space volume recovered using the Fowler method is very consistent, with a volume approximately $5ml$ under the actual volume recovered for all multi-compartment lungs at all forcing periods.

As the disparity between end-expired alveolar volumes of the two compartments is increased without changing the tidal volume, the volume recovered decreases at all forcing periods when the Bohr method is used. This can be seen in the plots for configurations A to C in Figure 4.15b. In addition the disparity between the volumes recovered at lower and higher forcing periods increases, which results in an increase in homogeneity index. However it should be noted that the alveolar volume recovered from set A is not consistent across the range of forcing periods even though it comes from what is essentially a homogeneous lung, as both compartments are identical. It is also interesting to note that the disparity between the volumes recovered at different forcing periods, the homogeneity index and the underestimate in end-expired alveolar volume does not increase by the same amount between A and B as between B and C,

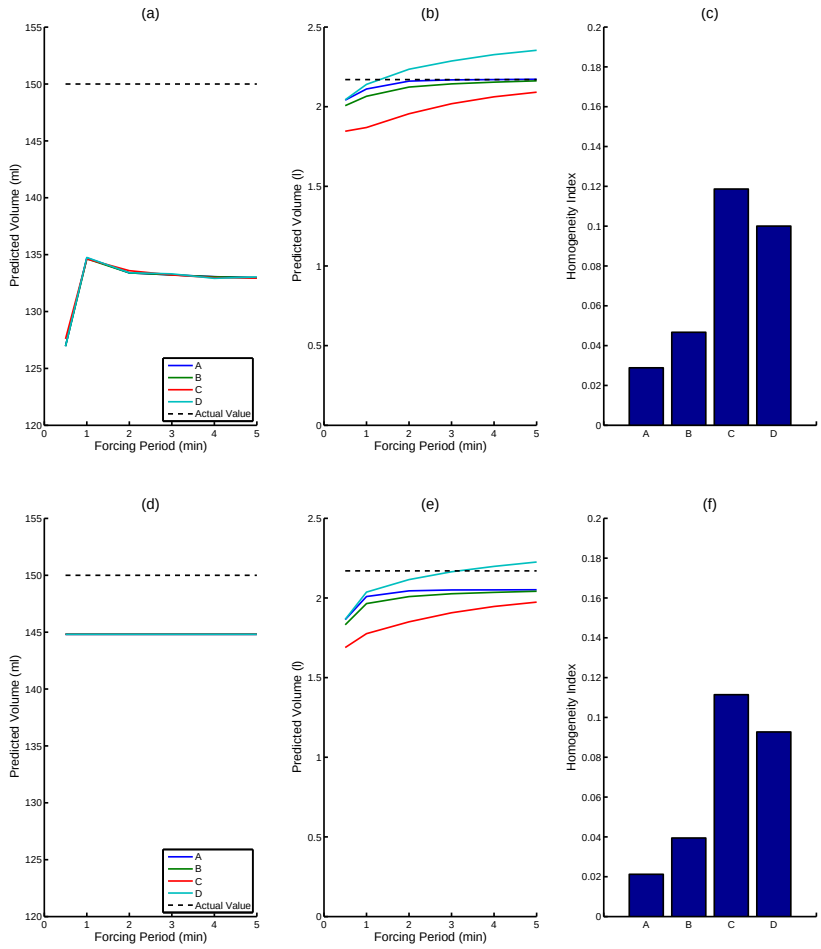


Figure 4.15: Dead space volumes were recovered using the (a) Bohr and (d) Fowler methods. These dead space volumes are then used to recover the end-expired alveolar volumes by the breath-by-breath method, as shown in (b) and (e), respectively. (c) and (f) show the respective homogeneity indices. The simulated data used came from four multi-compartment lungs as detailed in Table 4.4.

even though the disparity in the alveolar volumes is increased by approximately the same proportion.

Configuration D has the same end-expired alveolar volumes as C with a change in tidal volumes adjusting the ratio between tidal and alveolar volumes. This causes the volume recovered to increase across all forcing periods rather than to decrease. Also the difference in volume recovered at different forcing periods does not change greatly which results in a similar lower homogeneity index, even though the homogeneity has essentially increased.

A similar pattern of results are seen in Figure 4.15d when the Fowler method was used to recover the dead space volume. The end-expired alveolar volumes recovered are however slightly below those seen in Figure 4.15b. This results in a similar trend in homogeneity indices but with values slightly below those recovered using the Bohr

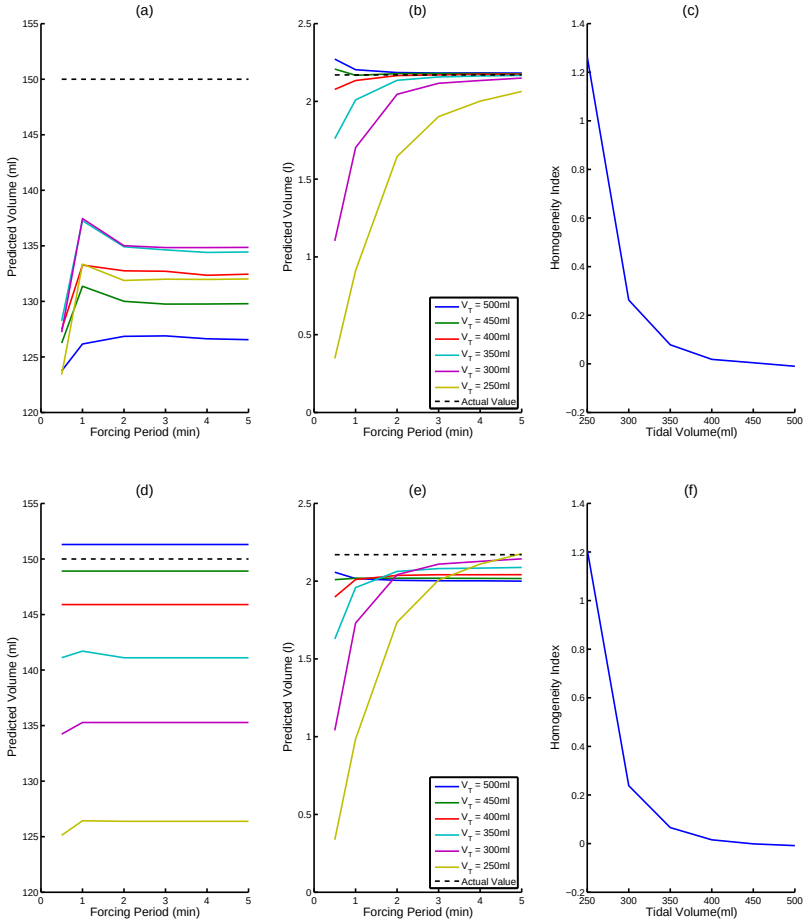


Figure 4.16: Dead space volumes were recovered using the (a) Bohr and (d) Fowler method. These dead space volumes are then used to recover the end-expired alveolar volumes by the breath-by-breath method, as shown in (b) and (e), respectively. (c) and (f) show the respective homogeneity indices. The simulated data used came from a multi-compartment lung with two identical compartments. The end-expired alveolar concentration for both compartments was $1080ml$ and the tidal volume was varied as detailed in the key, with half the tidal volume entering each compartment.

method.

In order to investigate the disparity between the volumes recovered at lower forcing periods and the actual alveolar volume of configuration A (the homogeneous case), simulated data for a variety of tidal and mixed compartment volumes were created. The volumes recovered are shown in Figures 4.16 and 4.17.

The tidal volume has an effect on the dead space and alveolar volumes recovered using both the Bohr and Fowler methods. As previously observed, the dead space volumes recovered using the Fowler method are more consistent across the range of forcing periods. However the range of the dead space volumes recovered for different tidal volumes is much wider than for the Bohr method, with the dead space volumes recovered getting closer to the actual volume as the tidal volume increases. The oppo-

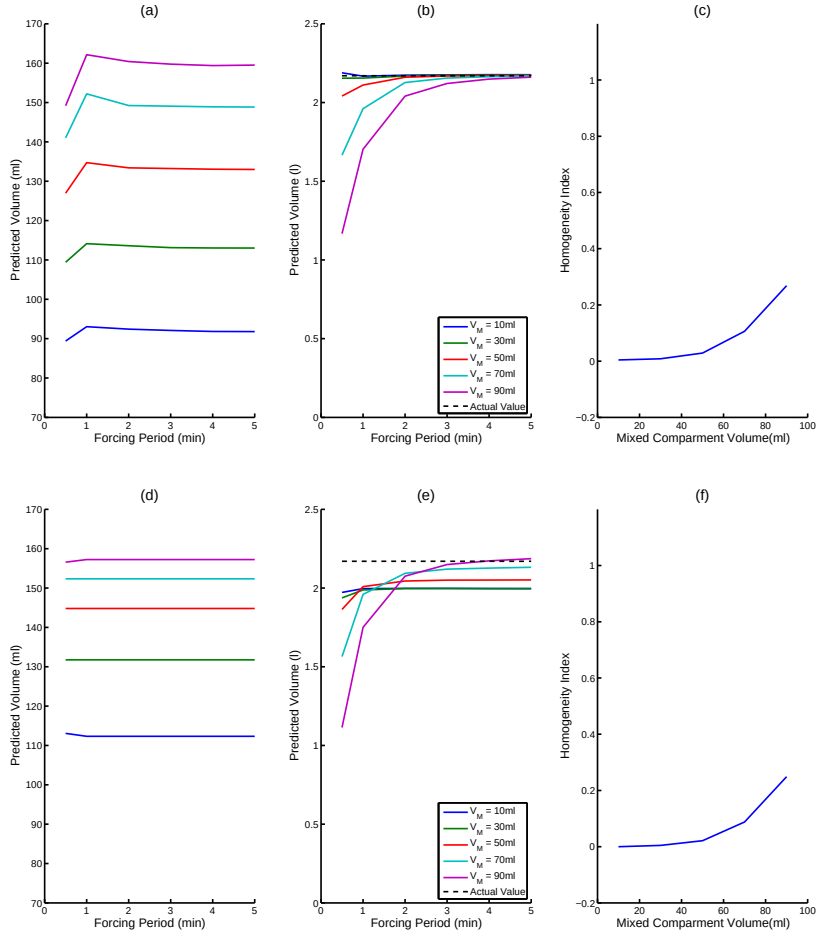


Figure 4.17: Dead space volumes were recovered using the (a) Bohr and (d) Fowler method. These dead space volumes were then used to recover the end-expired alveolar volumes by the breath-by-breath method, as shown in (b) and (e), respectively. (c) and (f) show the respective homogeneity indices. The simulated data used came from a multi-compartment lung with configuration A in Table 4.4. The dead space mixed compartment volume was varied as detailed in the key. A $100ml$ transitory dead space compartment was used.

site is generally true for the Bohr method with the underestimate generally increasing, although not in all cases. The Bohr method is further from the actual value at all tidal volumes, except $250ml$.

The disparity in the end-expired alveolar volume recovered at different forcing periods follows a similar pattern when both the Bohr and Fowler methods are used. In both cases, as the tidal volume decreases the disparity in the volumes recovered at $1min$ and $5min$ increases, which yields a high homogeneity index at low tidal volumes, even though the homogeneity of the alveolar compartment has not changed. When the Bohr method is used, the end-expired alveolar volume recovered is lower as the tidal volume decreases at all frequencies. This is a similar pattern to that seen when the homogeneity of the lung is increased. In contrast, the volume recovered does not

SD of Noise Conc. %	SD of Noise Flow l/min	Bohr Method A	Bohr Method B	Bohr Method C	Bohr Method D
-	-	0.03	0.05	0.12	0.10
0.02	-	0.03 ± 0.001	0.05 ± 0.001	0.12 ± 0.001	0.10 ± 0.001
-	1.5	0.02 ± 0.063	0.03 ± 0.044	0.07 ± 0.019	0.05 ± 0.029
0.02	1.5	0.02 ± 0.059	0.03 ± 0.043	0.07 ± 0.019	0.05 ± 0.028
0.04	3.0	0.00 ± 0.038	0.01 ± 0.030	0.07 ± 0.022	0.04 ± 0.033

SD of Noise Conc. %	SD of Noise Flow l/min	Fowler Method A	Fowler Method B	Fowler Method C	Fowler Method D
-	-	0.02	0.04	0.11	0.09
0.02	-	0.02 ± 0.001	0.04 ± 0.001	0.11 ± 0.001	0.09 ± 0.001
-	1.5	0.01 ± 0.061	0.01 ± 0.041	0.06 ± 0.018	0.03 ± 0.026
0.02	1.5	0.01 ± 0.056	0.01 ± 0.040	0.06 ± 0.019	0.03 ± 0.026
0.04	3.0	-0.01 ± 0.037	0.00 ± 0.029	0.06 ± 0.022	0.03 ± 0.030

Table 4.5: Mean and standard deviation of the homogeneity index determined from the simulated multi-compartment lungs detailed in Table 4.4. Noise is added to the data as detailed in the table.

consistently decrease for all forcing periods when the Fowler method is used, with the alveolar volume increasing at higher forcing periods.

When the mixed compartment volume is varied in size, the dead space volumes recovered are again more consistent across forcing periods when the Fowler method is used, although the volumes recovered with the Bohr method do become more consistent at lower mixed compartment volumes. The Bohr method always underestimates the total volume (mixed compartment plus transitory compartment) with the underestimate increasing with mixed compartment volume. The Fowler method is closer to the actual volume in all cases, slightly overestimating at lower mixed compartment volumes but underestimating at higher volumes.

The end-expired alveolar volumes recovered echo the results when varying the tidal volume, but in this case the disparity between the end-expired alveolar volumes at different forcing periods increases as the mixed compartment volume increases, rather that decreasing, as in the tidal volume cases. This is reflected in an increase in homogeneity index with mixed compartment volume. The decrease in the volume recovered at all forcing periods is also maintained for the Bohr method, as are the increases in the volumes recovered at higher forcing periods and the decreases at lower volumes for the Fowler method.

4.4.1.1 Experimental Error

In reality the ability to determine the homogeneity of a given multi-compartment lung will depend on the amount of noise present within the two signals. Table 4.5 provides details of the homogeneity indices determined in the presence of noise. As previously observed for lungs with log-normal distributions, noise in the concentration signal at the typical level has minimal effect on the results. The mean of the homogeneity indices recovered is identical to that observed when no noise was present and the standard deviations are low at 0.001 in all cases.

The effect of the typical noise in the flow sensor however is more significant. The mean homogeneity index determined in the presence of noise is reduced compared to the value recovered without noise, with the larger reductions occurring for the more inhomogeneous configurations (C and D). The standard deviation also increases to between 0.018 and 0.061, with the largest rises in the homogeneous configurations (A).

The combined effect of typical noise in both signals is similar to noise in the flow signal alone, as may be expected. The one noticeable exception to this is a reduction in the standard deviation for the homogeneous configuration (A) in both the Bohr and Fowler cases, but these are small at 0.004 and 0.005, respectively. Doubling the noise present in both signals continues to reduce the mean value for the homogeneity index for configurations A and B, but also appears to cause a slight reduction in the standard deviations. In contrast the means are unaffected for configurations C and D but the standard deviation increases slightly. In all cases, there is little difference between the Bohr and Fowler methods, except that the mean homogeneity indices are consistently slightly lower in the latter case.

4.4.1.2 Discussion

These results show that levels of inhomogeneities in a multi-compartment lung may be determined using homogeneity index in the absence of noise but that the method may be limited under certain circumstances. Where the increase in homogeneity is caused by an increase in the disparity between the end-expired alveolar volumes the homogeneity index appears to increase as expected. Increasing the disparity of the compartments ventilation-volume ratios by changing the tidal volume does not necessarily have the same effect. This is observed by the decrease in the homogeneity index in the case of configuration D.

Assuming that a mixed compartment models the mixing that happens in human dead spaces well (an assumption which is discussed further below), the homogeneity index may over indicate the homogeneity present within the alveoli at low tidal volumes

or where a high amount of mixing occurs within the dead space. This may not be unexpected as the breath-by-breath method for determining the alveolar volume relies on the assumption that the end-expired tracer gas concentration is equal to the alveolar concentration.

This is a reasonable assumption in the presence of a small mixed compartment volume, as the concentration of that compartment will reflect the concentration of air entering the compartment relatively quickly. This is also the case for higher tidal volumes, as the larger air flows will result in any compartment being well flushed through and so closely reflecting the concentration of the alveoli. However if this is not the case the assumption will start to break down.

This may be an issue if the homogeneity indices of different subjects are to be compared at different tidal volumes. However if the homogeneity index is used for ongoing monitoring of a single subject this may be less important, assuming that the same tidal volume can be replicated in later tests. Assuming that the mixing in the dead space does not change substantially it may then be possible to determine the relative homogeneity of the alveoli. Alternatively the homogeneity index could possibly be used as an overall index which combines the homogeneity of the alveoli with a measure of the relative amount of mixing in the dead space, however the results would still be dependent on the tidal volume.

These results support the previous indication from log-normal distribution data, i.e. that the dead space volumes recovered by both methods are only affected to a small degree by the make up of the alveoli and that the Fowler method is more consistent across the range of forcing periods. However the errors in the predictions made by the Fowler method are affected more by the relative tidal and mixed compartment volumes. The Bohr method provides a better estimation of the actual end-expired alveolar volume for a homogeneous lung when the tidal volume is sufficient to ensure that any mixing in the dead space compartment does not corrupt the end-expired concentration.

When no noise is present these results would indicate that it is possible to identify the difference between the homogeneity of configurations A, B and C using the homogeneity index, particularly when configuration C is compared to the other two. However in the presence of typical levels of noise the combination of the reduction in the mean results at higher inhomogeneities, which reduces the separation in the values of the homogeneity index, and the increase in the value of the standard deviation is

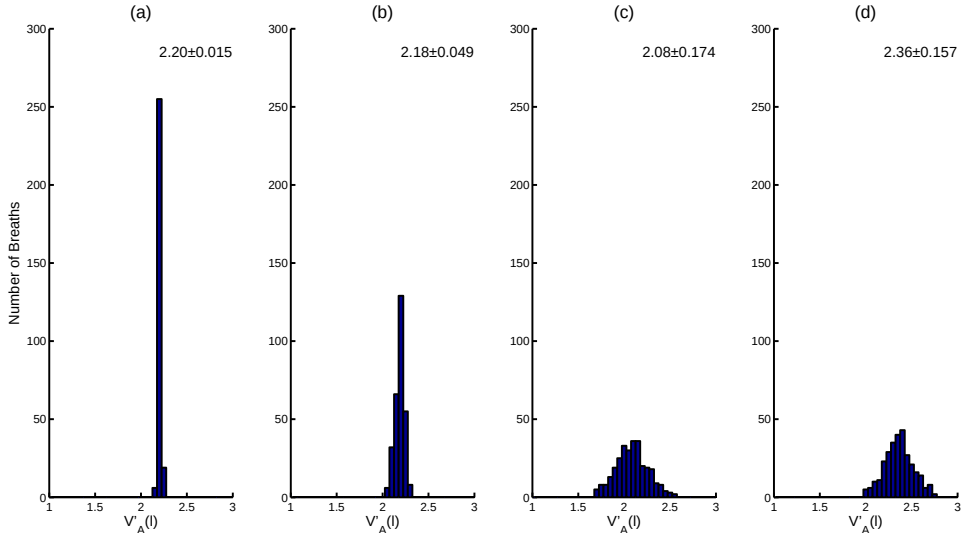


Figure 4.18: Recovery of the end-expired alveolar volumes for each breath using a homogeneous model from multi-compartment lung data. The simulated data came from the following multi-compartment lungs: (a) A, (b) B, (c) C and (d) D, as detailed in Table 4.4.

likely to make it difficult to distinguish between any of these configurations. For example, the difference in homogeneity index between configuration A and C is 0.05 which is less than the standard deviation observed for configuration A. The variability in the predicted homogeneity index will thus be more than any observed differences due to changes in the homogeneity of the alveoli.

Unlike for the log-normal data, increasing the inhomogeneity of the lung did not always result in a lower end-expired alveolar volume being recovered. This possibly adds some doubt about using this method to inform about individual homogeneity using a measure of FRC determined in this way.

Theoretically this method will provide information about changes in homogeneity of multi-compartment lungs under certain situations but there are some limitations in practice. The relative effect of the mixing within the dead space to size of the tidal volume may cause issues if the tidal volume is not large enough to flush residual gas out of it. Also the presence of experimental noise is likely to mask any changes in the homogeneity index.

4.4.2 Inter-breath Variation

Here the ability to distinguish between the four configurations of multi-compartment lung by examining inter-breath variability will be investigated. Figure 4.18 shows a histogram of the end-expired alveolar volumes recovered for each breath. The volumes of the mixed and transitory compartment are assumed to be known. A modulated

SD of Noise Conc. %	SD of Noise Flow l/min	Recovered Distribution Property	A	B	C	D
-	-	Mean (l)	2.20	2.18	2.08	2.36
		SD (l)	0.02	0.05	0.17	0.16
0.02	-	Mean (l)	2.19 ± 0.003	2.17 ± 0.004	2.07 ± 0.006	2.35 ± 0.005
		SD (l)	0.05 ± 0.003	0.07 ± 0.005	0.17 ± 0.006	0.16 ± 0.006
-	1.5	Mean (l)	2.20 ± 0.003	2.18 ± 0.004	2.08 ± 0.007	2.37 ± 0.006
		SD (l)	0.06 ± 0.004	0.07 ± 0.005	0.17 ± 0.008	0.16 ± 0.009
0.02	1.5	Mean (l)	2.19 ± 0.004	2.17 ± 0.007	2.07 ± 0.008	2.35 ± 0.007
		SD (l)	0.07 ± 0.005	0.08 ± 0.007	0.18 ± 0.010	0.17 ± 0.008
0.04	3.0	Mean (l)	2.19 ± 0.011	2.17 ± 0.010	2.06 ± 0.011	2.34 ± 0.014
		SD (l)	0.14 ± 0.011	0.15 ± 0.011	0.22 ± 0.014	0.21 ± 0.018

Table 4.6: For each of the noise levels detailed in the table, 25 sets of simulated data were created for each of the multi-compartment lung configurations. The end-expired alveolar volume of each of the breaths within these data sets was determined and the mean and standard deviation of their distribution taken. The results in this table are the mean and standard deviation of the 25 distributions' mean and standard deviation.

forcing sinewave with the parameters described previously was used when simulating this data.

The distributions recovered from the multi-compartment lung data with configurations A to C, show the two characteristics observed in simulated data from lungs with log-normal distributions. The mean volume decreases and the standard deviation increases as the mismatch in the end-expired alveolar volumes increases (i.e. the inhomogeneity increases). This is not the case for configuration D where the mismatch in ventilation-volume ratios is increased by changing the tidal volumes of the two compartments. The mean of the recovered distribution is larger than the mean in the homogeneous case (configuration A). The standard deviation is still larger than for configurations A and B but not C, which has the same end-expired alveolar volumes for both compartments.

4.4.2.1 Experimental Error

As previously highlighted, changes in the inhomogeneity of the lung and the noise in either of the signals are likely to have similar effects on the width of the distribution, so it is important to assess the effect of noise. Table 4.6 summarises results from 25 simulated sets of data in the presences of noise. The mean value for the mean of the recovered distributions is relatively consistent at all levels of noise, with the values very close to that recovered when no noise is present. The standard deviations for these means increases as the amount of noise is increased for all configurations. There is little

difference, however, between the standard deviation in the presence of typical noise in both the concentration and flow signals.

The mean results for the standard deviations of the recovered distribution, in contrast, do vary with noise, with the mean values increasing with noise for all configurations. The variation in the results recovered also increases with noise, as indicated by an increase in the standard deviations. There is again little difference between the results when the typical level of noise is only present on one of the signals, however both the mean and standard deviations are occasionally higher when the noise is only found on the flow signals.

Concentrating on the results for the combined effect of typical noise in both signals, it would seem unlikely that the difference between configuration A and B could be identified by either using the mean or width of the distribution, as the difference in the mean values is on a par with the standard deviations. However it would be plausible to identify differences between configurations A or B and configuration C as the separations in the means of both mean and standard deviation is $0.1l$ while the standard deviations in all cases are at least an order of magnitude smaller at less than $0.01l$.

At double the typical noise levels, the situation with regard to the mean of the distributions is similar, but the standard deviation has increased slightly. The mean and standard deviation of the standard deviations of the recovered distributions have increased in all cases. This reduces the separation between the standard deviations but this is still at least 5 times the standard deviation of the standard deviations of the recovered values.

4.4.2.2 Discussion

When no noise is present these results are similar to those observed for the homogeneity index. The proposed method for measuring the homogeneity, the width of the distribution, works well for configurations A to C but does not pick up the changes in configuration D. A decrease in the mean end-expired alveolar volume recovered, which is observed in the other cases, also does not occur for configuration D. Thus, in the absence of noise, it should be possible to infer information about the changes that occur in configurations A to C but not necessarily to be able to distinguish them if mismatches in tidal volumes occur, as is the case for configuration D.

When experimental noise is present the situation is somewhat different to the homogeneity index. The mean and standard deviation of the standard deviation for the distribution recovered increase with noise, especially in the case of configuration A (the

One Compartment Model

Parameters	Distribution	Lower Limit	Upper Limit
V'_A (ml)	Uniform	1000	3000

Two Compartment Model

Parameters	Distribution	Lower Limit	Upper Limit
$V'_{A,1}$ (ml)	Uniform	900	1900
$V'_{A,2}$ (ml)	Uniform	200	1200

Table 4.7: Priors used for ABC SMC parameter estimation and model selection for simulated multi-compartment data.

homogeneous case). This is likely to make it indistinguishable from configuration B as the variation in the width of the distribution is likely to mask any differences in the standard deviation of the recovered distribution. However it should still be possible to infer differences in the homogeneity equivalent to configuration C. Also, unlike the homogeneity index, the effects of typical levels of noise in the concentration and flow signals alone seem to have a roughly similar effect.

It should be acknowledged that knowledge of the make up of the dead space has been assumed here, but not in the breath-by-breath method. Methods of independently determining that information are discussed further below. Also this method specifically models mixing within the dead space so is likely not to suffer to the same extent should a significant amount of mixing occur within the dead space.

These results imply that this method should provide a more sensitive measure of the homogeneity of the multi-compartment lung than the homogeneity index. Although the method is affected by noise on the signals, especially for configuration A, it should at least be possible to identify between configuration A and C in the presence of noise, a task that the homogeneity index would struggle with.

4.4.3 Directly Recovering the Compartment Volumes

Within this section the ability to directly infer the volumes of the compartments is assessed. An ABC SMC model selection was carried out between a lung with one compartment and a lung with two compartments. Within the two compartment model it is assumed that both compartments receive the same ventilation allowing the flow rate into each compartment to be determined directly from the flow signal (i.e. half of the total flow rate) without the need for additional parameters. This is true for the data created for three of the configurations but not the fourth, configuration D. It is assumed that the make up of the dead space compartments is known. Details of model parameter priors are given in Table 4.7.

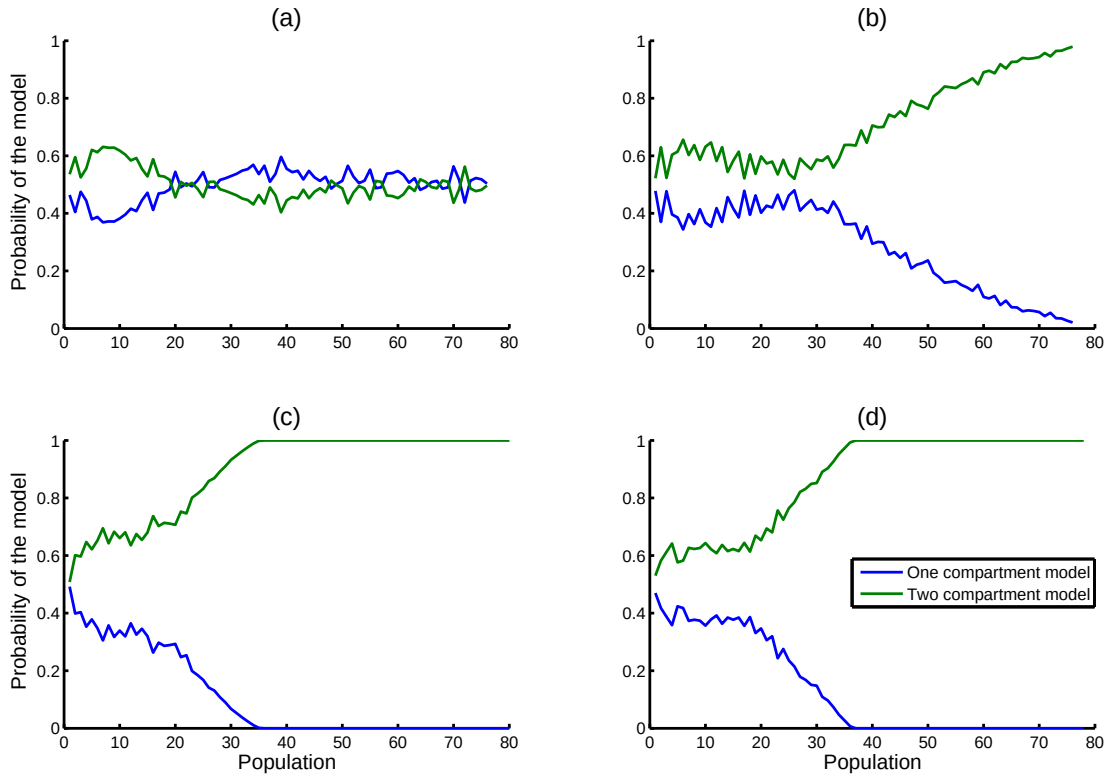


Figure 4.19: Convergence of ABC SMC model selection. Simulated data came from multi-compartment lungs with configuration: (a) A, (b) B, (c) C and (d) D, as detailed in Table 4.4.

Figure 4.19 shows the results of the model selection. In the case of configuration C and D, a selection is made with full certainty after approximately 35 populations that the data has come from a lung made up of two compartments. For configuration A and B the results of the model selection have not converged, despite over 80 days of computational time.

In the case of configuration A, both models are approximately as likely to have produced the data across all populations, however there is much variation in the results at each population. Configuration B, starts with similar probabilities which slightly favour the two compartment model but after about 30 populations these probabilities diverge with a two compartment model gradually becoming favoured. There is also significant variation in the probability from one population to the next. At the final population presented here the probability of each model appears to still be diverging.

Figure 4.20 shows the mean and interquartile range of each of the parameters within the two models. In all the cases the distributions of end-expired alveolar volume recovered by the single compartment model appears to have converged, even if the model then dies out. The mean value of the final population is close to the actual value in the

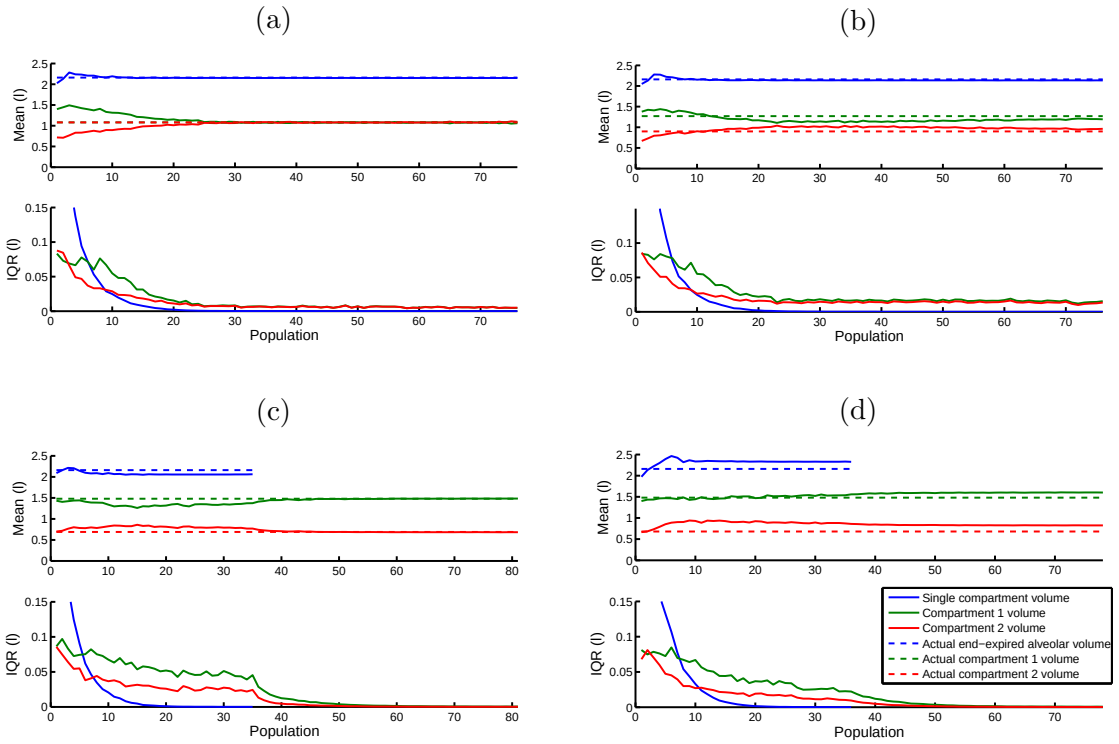


Figure 4.20: Convergence of all parameters during an ABC SMC model selection. Simulated data came from multi-compartment lungs with configuration: (a) A, (b) B, (c) C and (d) D, as detailed in Table 4.4. In a number of cases, the single compartment model results are only shown up to the point the model dies out.

case of configuration A (the homogeneous case), with configuration B underestimating to a small degree, with an estimate of $2.14l$. Configuration C underestimates more significantly, with a result of $2.05l$, while configuration D overestimates, with a result of $2.33l$. These results are similar to those seen for the breath-by-breath method at longer forcing periods and for the mean of the distributions in the inter-breath variation results.

For the two compartment model the parameter distributions have converged for all but configuration B, where the means of the two compartment volumes may still be slowly diverging from each other. The interquartile ranges of configuration C and D are very close to zero by the final population. Both show a significant decline in the interquartile ranges from the point at which the one compartment model dies out. In comparison the interquartile ranges of configuration A and B are significantly higher in the final population presented. For configurations A and C the mean values for the two compartment volumes at the final population shown are close to the actual volume. This is not the case for configuration B and D. This might be expected for configuration D as the model does not account for the different tidal volumes with

Compartment	A	B	C	D
1 (l)	1.09 ± 0.030	1.23 ± 0.097	1.47 ± 0.016	1.60 ± 0.015
2 (l)	1.07 ± 0.007	0.96 ± 0.065	0.68 ± 0.014	0.82 ± 0.016

Table 4.8: Effect of typical noise levels on the compartment volumes recovered using an ABC SMC parameter estimation. Simulated data came from multi-compartment lungs with the four configurations detailed in Table 4.4. In two cases the configuration B results had not converged and so were excluded when calculating these values.

compartment 1 expected to overestimate (as the volume-ventilation is higher than for compartment C) and compartment 2 expected to underestimate. This is not the case with both volumes overestimated.

4.4.3.1 Experimental Error

In order to investigate the effect of noise, ten sets of simulated data were created for each of the four configurations with the typical amount of noise on the concentration and flow signals (0.02% and $1.5l/min$ respectively). ABC SMC parameter estimation was then carried out with one model, the two compartment model. Table 4.8 shows the mean and standard deviations of all ten sets of data, where the mean of each parameter distribution has been taken as the estimate. Convergence was monitored manually.

The mean values recovered by configurations A and C are close to the actual volumes of the compartments. The standard deviations are generally low but do rise to $30ml$ in one case. The mean values for configuration B are further from the actual and the standard deviation is higher. In addition in two cases the parameter distributions did not converge within 80 days. The average computational time for those that did convergence was 55 days compared to 23 days and 14 days for configurations A and C, respectively. It should be noted, of course, that convergence is determined manually and so these times will be affected by the nature of the process.

The size of the standard deviations is likely to make it difficult to identify configuration B from A or C using a single set of data, especially by using the values recovered for compartment 1, as mean results are within one standard deviation of each other. Distinguishing configuration A from C should be possible as the standard deviations are much lower and there is a larger difference in the mean volumes recovered.

As would be expected, because the recovery model does not allow for changes in the tidal volume, the means of the results for configuration D are not close to the actual volumes. They are however close to the values recovered when no noise was present and have low standard deviations, comparable to that of configuration C.

4.4.3.2 Discussion

These results show that it may be possible to identify between two compartment lungs with configurations A and C using an ABC SMC parameter estimation with a two compartment model. There were however a few issues when it comes to configuration B. Both in the absence and presence of noise the convergence of the ABC SMC algorithm was slower than in the other cases. This may be due in part to the fact that model selection is also being performed in the no noise case, with the categorisation of which model the data are likely to come from being less apparent. However in general it is unclear why this is the case. In the present noise the results also show a larger error in the mean and substantially higher standard deviation, making its identification from configuration A and C unlikely.

Although configuration D may be distinguishable from the other configurations, the volumes recovered are not representative of the actual volumes of the compartments or the total end-expired alveolar, with $2.42l$ total volume recovered. The separation in volumes recovered for the two compartments is actually very similar to that observed for configuration C, even though the disparity in the ventilation-volume ratios has increased. Further investigations could be carried out using a recovery model that allowed the amount of tidal volume distributed to each compartment to be altered but this would require an additional parameter.

The lack of convergence of the ABC SMC model selection in two cases makes it difficult to draw firm conclusions. The results for configurations C and D indicate that these data sets are certain to come from a two compartment lung, as may be expected considering the disparity between the volume of the two compartments and the data has no form of noise add to it. It may be expected that the results for configuration A would favour the one compartment lung, especially initially where the two compartment model would be expected to be penalised by its larger parameter space. However this does not seem to be the case with a two compartment lung initially marginally favoured and both models receiving similar probabilities throughout the populations shown. Ultimately the result of the model selection is unclear, as it does not converge within the the time available.

In many ways configuration B is behaving as could be expected. Initially both models are favoured to a comparable degree, reflecting the fact that this is a marginal case. Then in subsequent populations the two compartment model becomes more favoured reflecting the fact that the data are only likely to be fully explained by a two compartment model. Although it does not fully converge it would seem likely that the

two compartment model would ultimately be selected with full certainty. This may be a fair result considering that these are simulated data without noise.

These results are dependent on the dead space compartment volumes being known and errors in these predictions may induce further errors. Also, ideally more sets of simulated data with noise than the ten used here would have been used but the computationally intensive nature of this method introduces a limit.

Like the inter-breath method it should be possible to identify between configurations A and C even in the presence of noise. Unlike the homogeneity index, a noticeably higher increase in the relative standard deviations for the configuration A results is not seen here, the standard deviation for compartment 1 is about double that for configuration C and D but for compartment 2 it is about half. However the standard deviation of the configuration B results is significantly larger, meaning that it will still be difficult to identify configuration B from A or C.

4.5 Independently Recovering the Dead Space Volume

Within this section the ability to determine the volume of the dead space compartments independently of the make up of the alveolar compartments is investigated. Although it may be useful to independently determine these volumes in their own right, the primary motive is to use the volumes in further analysis, as described above.

Simulated data were created using alveoli with a continuous log-normal distribution of ventilation volume ratios, with an alveolar volume of $2.0l$ and a tidal volume of $500ml$. The mixed and transitory volumes of the dead space were then estimated independently by using the Nelder-Mead simplex, using the model described in Section 3.4.2.3. The starting points used in all cases were $V_M = 40ml$ and $V_R = 200ml$.

4.5.1 Results

Figure 4.21 shows the dead space volumes recovered as the homogeneity of the alveolar compartment is varied by changing the standard deviation of a single peak. The volumes recovered vary no more than $1ml$ from the actual volume for both the mixed and transitory compartments.

The effect of varying the volume of the mixed and transitory dead space compartments is shown in Figure 4.22, when σ_x is 1.0. The mixed compartment volume is increasingly underestimated as it increases in volume, with a $14ml$ underestimate by $100ml$. The overestimate of the transitory compartment does increase slightly at the

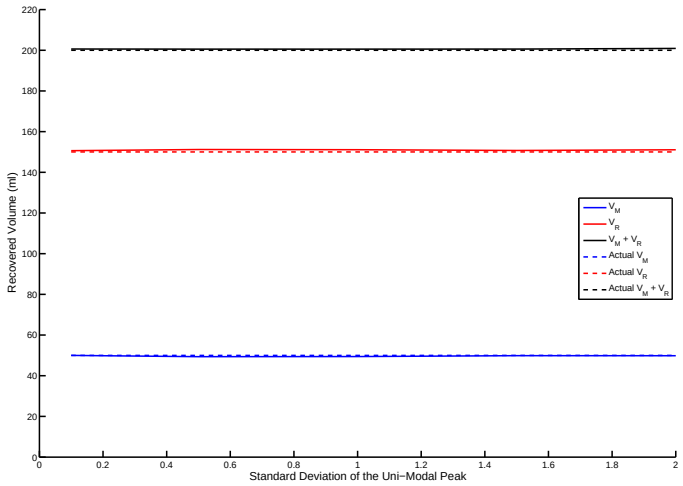


Figure 4.21: The volumes recovered for the mixed (V_M) and transitory (V_R) compartment volumes of the dead space independently of the make-up of the alveolar compartment when varying the homogeneity of the alveolar compartment.

same time to $4ml$ for a $100ml$ mixed compartment volume. This results in a maximum underestimate in the total dead space volume of $10ml$.

A similar trend is seen when the transitory compartment volume is increased, with the mixed compartment underestimated by $18ml$ and the transitory compartment overestimated by $3ml$ for a transitory compartment of $400ml$. This results in a maximum underestimate of the total dead space volume of $15ml$.

4.5.2 Discussion

The mixed and transitory compartment volumes recovered are unaffected by changes in the homogeneity of the lung. However, increasing the volume of the mixed compartment causes an underestimate in its volume. This is likely to be caused by the breakdown of the assumption that the end-expired and alveolar concentrations are equal, in a similar way to the breath-by-breath method, described above. As the mixed compartment volume is increased, the time constant of the compartment increases, meaning that the concentration of the compartment takes longer to approach the concentration of the air entering. Thus the concentration of the mixed compartment takes longer to approach the alveolar compartment concentration during expiration, as it is still being influenced by the concentration of the tracer gas left in the mixed and transitory compartments during the previous inspiration phase.

The mixed compartment naturally induces a delay in the tracer gas passing through it, as it takes time for the concentration of the compartment to increase/decrease to that of the air passing through it. Therefore an underestimate in the mixed compartment

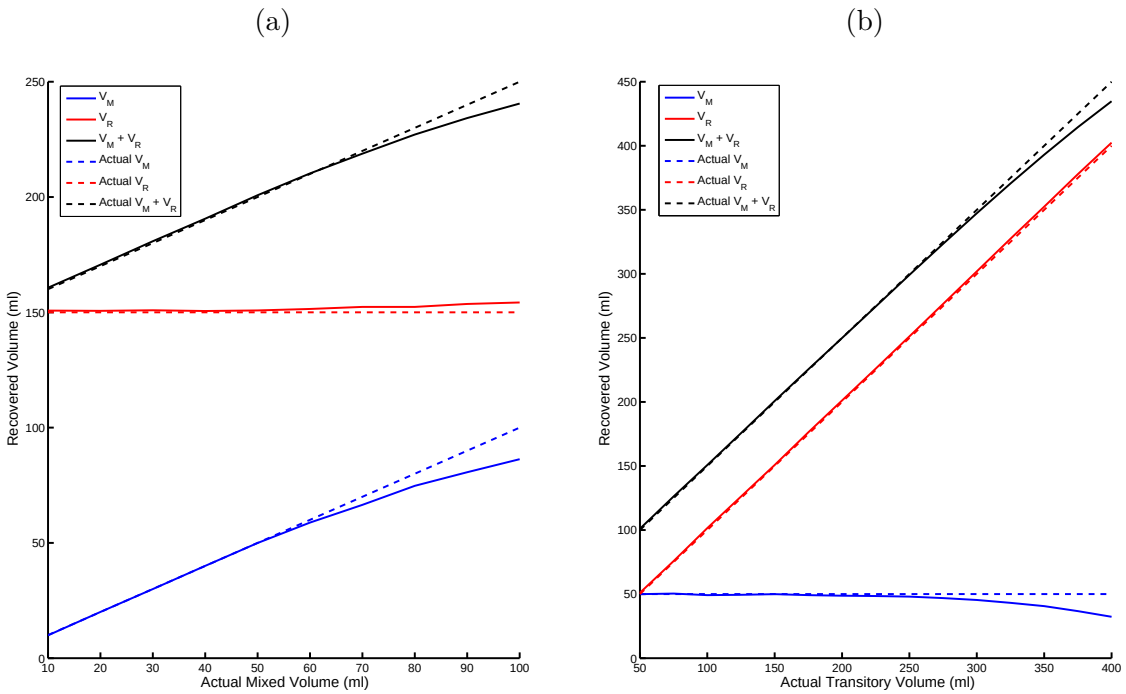


Figure 4.22: The volumes recovered for the mixed (V_M) and transitory (V_R) compartment volumes of the dead space independently of the make up of the alveolar compartment. The volumes of the (a) mixed compartment and (b) transitory compartment are varied.

volume is likely to result in an increase in the estimate for the delay and so also an increase in the transitory compartment volume. This is observed in the results but the increase is less than the underestimate of the mixed compartment volume.

A similar explanation would explain the results of increasing the volume of the transitory compartment. As the transitory compartment volume is increased, without increasing the tidal volume, the volume of alveolar air reaching the mixing compartment is decreased (only $100ml$ will reach it in the case of a $400ml$ transitory compartment). Therefore the air in the mixing compartment is not fully flushed out and so the end-expired fractional concentration is still influenced by the inspired concentration and thus the mixed compartment appears smaller. Again a slight increase in the transitory compartment volume is observed which counteracts the smaller mixed compartment, however the increase is much smaller here.

It is acknowledged that only uni-modal distributions, a limited range of dead space compartment volumes, one tidal volume and one alveolar volume have been investigated here, but these results imply that the two dead space volumes can be determined independently of the alveolar make up with little error. However errors are likely to increase with an increase in both the mixed and transitory compartment volumes. These results suggest that increases in the mixed compartment volume are likely to

affect the result to a greater degree than increases in the transitory volume. Assuming that the explanation given above is correct, these errors may partly be reduced by an increase in tidal volume that is likely to occur in subjects with increased dead space volume in order to maintain the level of oxygen delivered to the alveoli.

4.6 Alignment of the Signals

As detailed in the literature review, the alignment of the two signals (the flow and concentration signals) can have a significant effect on these type of methods. Within this section, initially the effect of misalignment on the volumes recovered by the breath-by-breath methods and least square optimisation methods is investigated. Following this, methods of recovering the offset, by treating it as another parameter which requires estimating, are investigated.

4.6.1 Simulated Data

Data were simulated using the homogeneous model of the lung. The concentration signal was shifted forwards or backwards in time in order to simulate the misalignment which is likely to be found in real data. The data were simulated at a $10ms$ sampling rate, which is the most common separation between data points in the real data. In order to shift the concentration signal it was first shifted forwards or backwards by the given offset, positive values indicating that the concentration signal had been moved forwards in time, while negative values indicated that it had been moved backwards in time.

Below attempts are made to determine the offset present within a set of data. During the recovery process the negative of the original offset is expected to be recovered, as the concentration signal will need to be shifted back in time in order to be aligned with the flow signal again. In the results presented below the sign of the recovered values will be inverted so that they match the original shift in the data. Where a continuous range of offsets is required, the concentration signal is linearly interpolated onto the same time series as the flow signal, thus allowing offset values other than at $10ms$ intervals. Unless otherwise stated, the baseline compartment volumes used elsewhere in this chapter are used ($V_M = 50ml$, $V_R = 100ml$ and $V'_A = 2.0l$).

4.6.2 Effect of an Offset

Here the effects of an offset between the two signals on the breath-by-breath methods and a least squares optimisation by the Nelder-Mead simplex are investigated. Simulated data are created for offsets between -100ms and 100ms and the volume of the

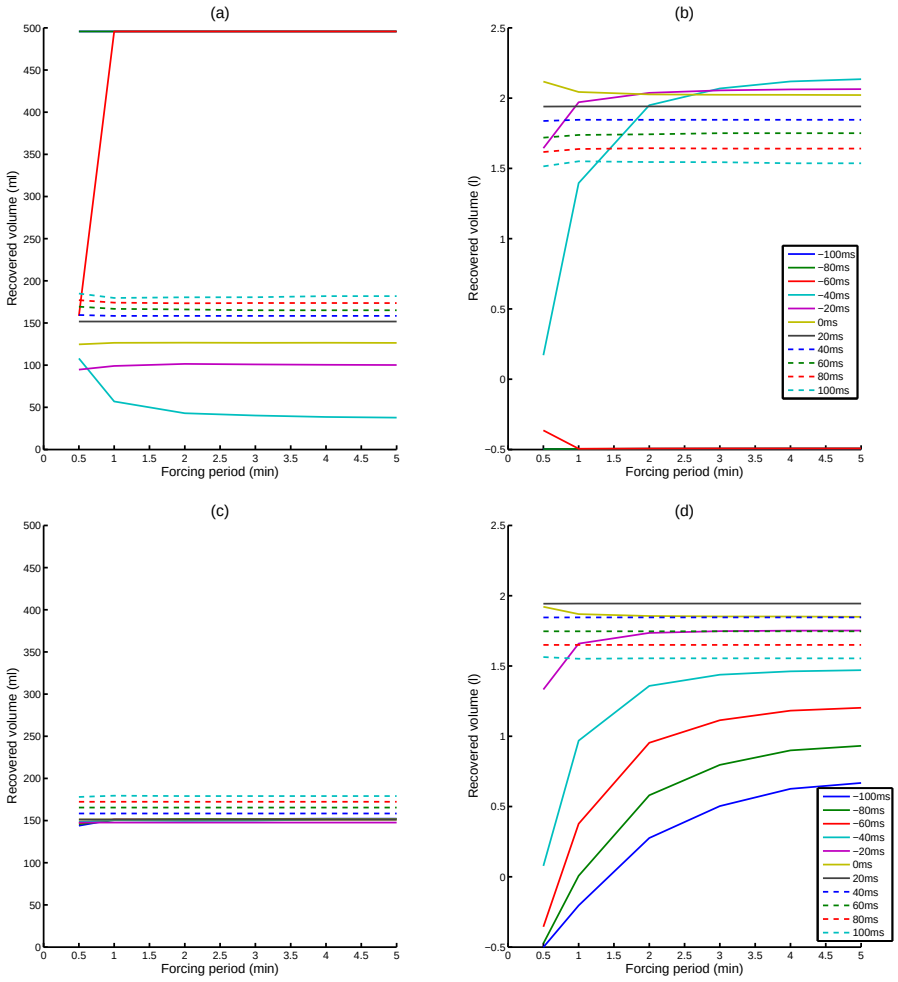


Figure 4.23: The plot show the (a & c) dead space and (b & d) end-expired alveolar volumes determined using the breath-by-breath method when an offset between the two signals is present. Results are shown when the (a & b) Bohr and (c & d) Fowler methods are used to determine the dead space volume at a range of forcing periods.

compartments inferred from these data, making no correction for the offset between the time signals.

4.6.2.1 Breath by Breath Method

The dead space and end-expired alveolar volumes determined by the breath-by-breath methods are shown in Figure 4.23. When no offset is present the Bohr method underestimates the dead space volume by approximately $25ml$ across the range of forcing periods. The recovered end-expired alveolar volume is close to the actual volume at all forcing periods, except at $30s$. In contrast the Fowler method estimates the dead space volume more accurately but the end-expired alveolar volume is underestimated by approximately $150ml$. This is consistent with previous observations.

The Bohr and Fowler methods show a similar pattern for positive offsets (where

the flow proceeds the concentration signal) with the predictions for the dead space volume increasing as the offset increases. A similar pattern is seen for the end-expired alveolar volume which decreases as the offset increases, except in the the Fowler case at $20ms$ where an increase is observed but the recovered volume is still about $50ml$ under the actual volume. The volumes recovered are consistent across the range of forcing periods, except for minor differences observed at a $30s$ forcing period in a number of cases.

When the Bohr method is used on data with negative offsets (where the concentration proceeds the flow signal) the dead space volume is initially underestimated until an offset of -$60ms$. From that offset onwards a dead space volume of approximately $500ml$ is predicted in all but one case. This happens to be the tidal volume of the data. A similar divide in the end-expired alveolar volume predictions occurs, with offsets of -$60ms$ and under producing a result of -$0.5l$, except for one exception. Offsets above -$60ms$ result in predictions which vary significantly with forcing period, with lower forcing periods underestimating and higher volumes overestimating. This would result in a significant homogeneity index if calculated, despite the simulated data coming from a homogeneous lung.

The Fowler method does not show similar variations in the dead space volumes calculated from data with negative offsets, with the volume recovered being consistently close to the actual volume at all negative offsets and forcing periods. The end-expired alveolar volumes recovered are underestimated at all forcing periods, with the underestimate observed at lower forcing periods being greater than at higher forcing periods. This would again result in a significant homogeneity index. The underestimate at a particular forcing period generally increases as the offset decreases.

4.6.2.2 Least Squared Optimisation

Figure 4.24 shows the results of least squares optimization using the Nelder-Mead simplex. The recovery model used was a homogeneous alveolar compartment with mixed and transitory dead space compartments. Two starting points are shown. In the case of the lower starting point ($V_M = 25ml$, $V_R = 50ml$ and $V'_A = 1.0l$) the algorithm attempted to evaluate the model at very low or negative volumes causing errors, predominantly in low or negative mixed compartment volumes. To prevent this occurring, the bounded version of the Nelder-Mead simplex was used limiting the minimum volume of all compartments to $1ml$.

Both starting points show similar results indicating that the algorithm is likely to be finding a global minimum. There is a more noticeable difference in the mixed

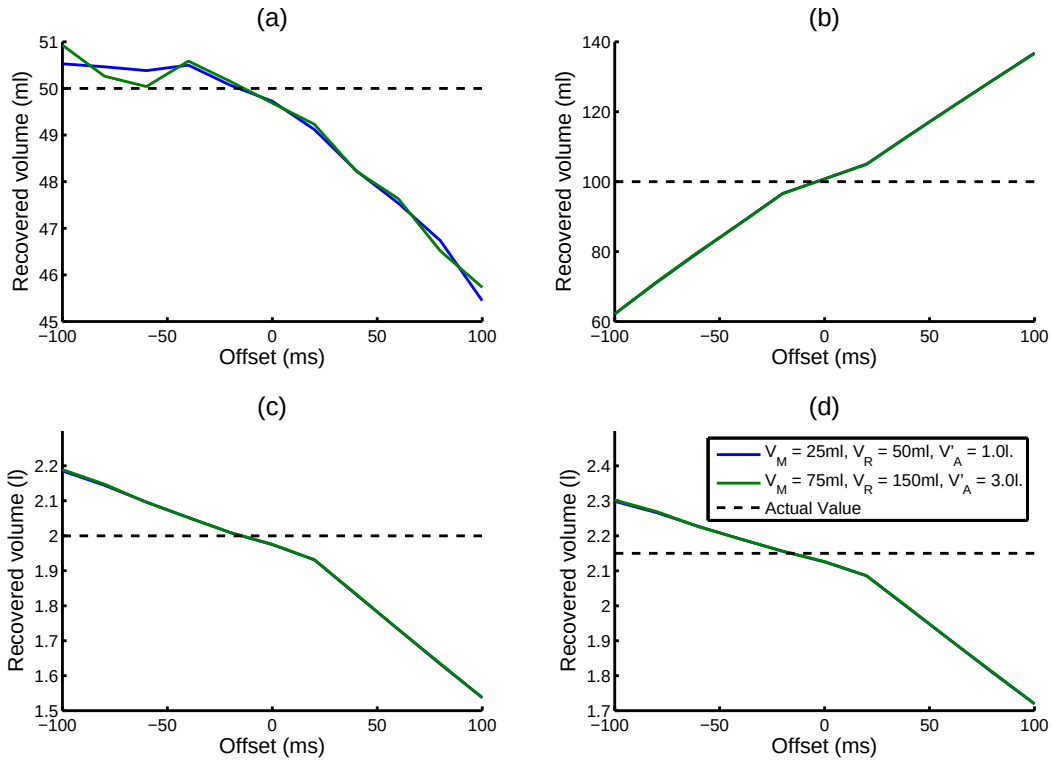


Figure 4.24: Volumes recovered using least squared optimization via the Nelder-Mead simplex when there is an offset between the two signals. (a) Mixed compartment volume, (b) Transitory compartment volume, (b) End-expired alveolar volume and (d) FRC.

compartment volumes recovered but the errors in the mixed compartment volumes are less in both absolute and percentage terms. The results show that the recovered mixed compartment volume is influenced little by negative offsets, as in fact a number of the volumes recovered are closer to the actual volume than at zero offset. Positive offsets have more of an effect but the effect is still less in absolute and percentage terms than for the other volumes recovered, at 4.3ml at 100ms (8.5%).

The transitory dead space volume recovered is significantly affected by both positive and negative offsets with the error being roughly linear; the error is 36ml (36%) at 100ms and -38ml (38%) at -100ms. The errors in the end-expired alveolar volume are in the opposite direction to that for the transitory volume; -0.46l (22%) at 100ms and 0.19l (10%) at -100ms. Thus the error in FRC is generally less; -0.48l (23%) at 100ms and 0.10l (5%) at -100ms. In fact the least error in FRC is found at around -40ms.

4.6.2.3 Discussion

The end-expired alveolar volume values recovered are significantly affected by the offsets between the two signals regardless of the direction of the offset or the method used, with an error of at least 10% at an offset of $\pm$100ms. These feed through to the

FRC predictions. Having no offsets generally produces better results, except in some specific cases for specific compartments.

For the breath-by-breath methods positive offsets are generally preferable, as the errors caused are lower and the results are not influenced by forcing period. This is the same direction to what generally occurs, due to the concentration sensor's slower response time. The error in the end-expired alveolar volume predictions is generally less, except for in some specific cases of low negative offsets, and the predictions are not effected by the forcing period. Negative offsets do however give better results in the Fowler case when determining dead space and in the Bohr case when recovering the end-expired alveolar volume at mid to high forcing periods and offsets lower than -60ms.

The 500ml dead space volume observed at offsets of -60ms and less when the Bohr method is used can be explained by examining the implementation of the algorithm. As the offset value becomes increasingly more negative, the recovered value for the dead space reduces; i.e. the last dead space volume at which the right hand side minus the left hand side of Equation 3.61 is of a different sign is reducing. When this goes below zero then no point will be found and the algorithm returns by default the final value, i.e. the tidal volume. Thus the tidal volume is given as the estimate for the dead space as is the case for offsets under -60ms. This could be overcome by updating the algorithm to ignore breaths where no change in sign is observed or by minimising the difference between the left and right hand sides.

This in turn explains the reason why the end-expired alveolar volume is estimated as -0.5l in these cases. Taking Equation 3.67 and assuming that V_D is equal to the tidal volume ($V_{T,i}$) allows a number of terms to be eliminated from the left hand side:

$$V_{T,i}(F_{E',i} - F_{E',i-1}) = V'_A.(F_{E',i-1} - F_{E',i}) \ . \tag{4.1}$$

Rearranging to give V'_A and simplifying further gives:

$$V'_A = -V_{T,i} \ ; \tag{4.2}$$

i.e. the prediction for end-expired alveolar volume will be the negative of the tidal volume: -500ml. Calculations of the tidal volume will not be affected by any offsets between the two signals as they are derived from the flow signal alone.

In the least squares optimisation case if an offset must be present then a negative one is preferable to a positive one, except where the transitory volume is concerned.

In some cases, particularly for a FRC, a small offset of -20 to -40ms produced better results, but this may be specific to the particular lung simulated.

These results back up the previous research, detailed in section 2.6, which shows that the offset between the two signals can significantly affect the values of volume recovered. Determining the offset and correcting for it in a preprocessing calibration step would be the preferable method of dealing with this issue. Alternatively if the offset could be determined by other means from the data this may also improve the results.

4.6.3 Recovering the Offset

One method that may enable the offset to be determined is to add an additional parameter for the offset. The logic here is that the best fit will be associated with the correct offset and so should result in an improvement in the other parameters being estimated. Three methods of implementing this parameter estimation are investigated here: least square optimisation using the Nelder-Mead Simplex, least squares optimisation using a genetic algorithm and ABC SMC parameter estimation.

4.6.3.1 Nelder-Mead Simplex

Figure 4.25 shows the predictions for the four parameters over a range of offsets when using the Nelder-Mead simplex. The bounded algorithm was used with a lower bound of 1ml on all compartment volumes but no bounds on the offset. Three different starting points were used for the offset value, as detailed in the figure but the compartment starting volumes were the same: mixed compartment of 30ml, transitory compartment of 150ml and end-expired alveolar volume of 1.5l.

The offset is predicted best across the range when a starting offset of 100ms is used, but the maximum absolute error is still relatively large at 38ms. This however still results in a significant improvement in the compartment volumes recovered compared to those seen in Figure 4.24. The volumes recovered for the mixed compartment show least error, with a maximum absolute error of 2.6ml (5.2%). A maximum error of 16.5% in the other compartment volumes is observed but this is still an improvement. The errors in the volumes predicted appear to be correlated with the errors in the offset prediction, with both increasing at the same actual offset.

This is also true for the results from the two alternative starting points (0 and -100ms), but the errors in the predicted offset are greater at higher offsets, so they result in higher errors in the predicted volumes. This results in higher errors at 80ms than observed when no offset correction was made. The exceptions to this are the

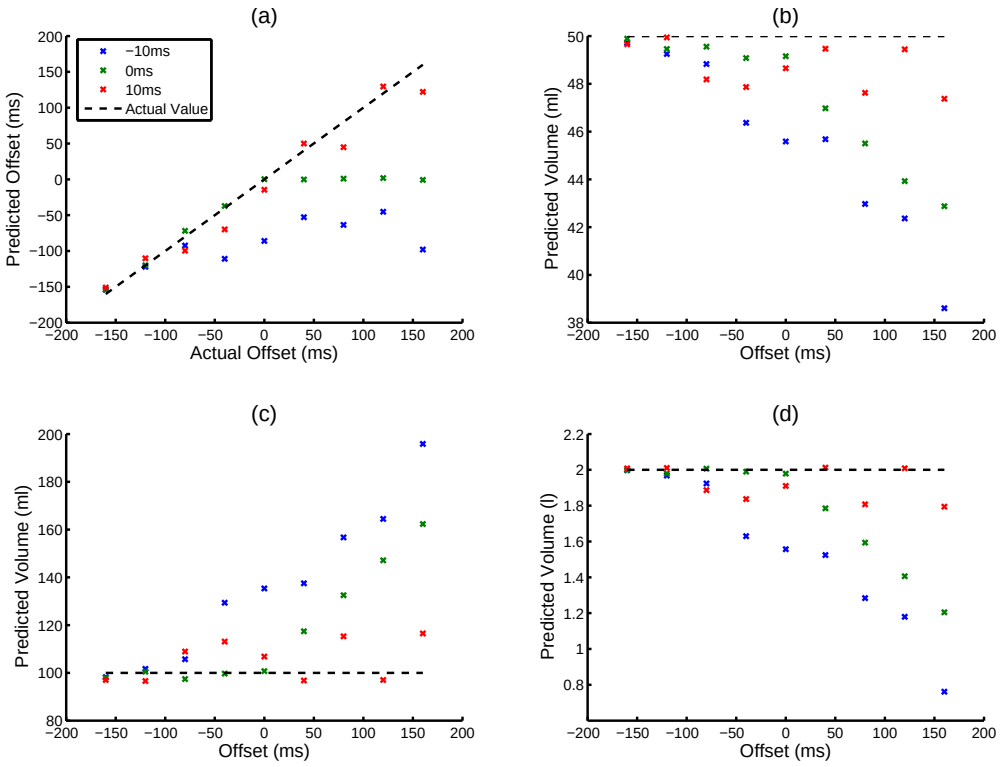


Figure 4.25: Parameter recovery when including the offset between the flow and concentration signals as an additional parameter to be recovered. The Nelder-Mead Simplex was used to recover the parameters: (a) offset in time between the two signals, (b) mixed compartment volume, (c) transitory compartment volume and (d) end-expired alveolar compartment volume. The simulated data came from a homogeneous lung with the actual volumes shown in the figures. The offset starting point was different between the three plots shown, as detailed in the key.

results below the starting point in both cases which show lower errors in the offset, a maximum absolute error of $8ms$. This in turn results in better volume predictions, with maximum error of 2.6% in the volume predictions.

4.6.3.2 Genetic Algorithm

Figure 4.26 shows the results when using the genetic algorithm with the default convergence criteria: stop when the change in RMS error is less than 10^{-6} for 50 generations or when the 100th generation is reached. The 100th generation was not reached in any of the cases presented here. The ranges for each of the parameters are detailed in

Parameters	Lower Limit	Upper Limit
Signal offset (ms)	-200	200
V_M (ml)	5	200
V_R (ml)	5	200
V'_A (l)	1.0	3.0

Table 4.9: Bounds for the genetic algorithm and the ABC SMC uniform priors.

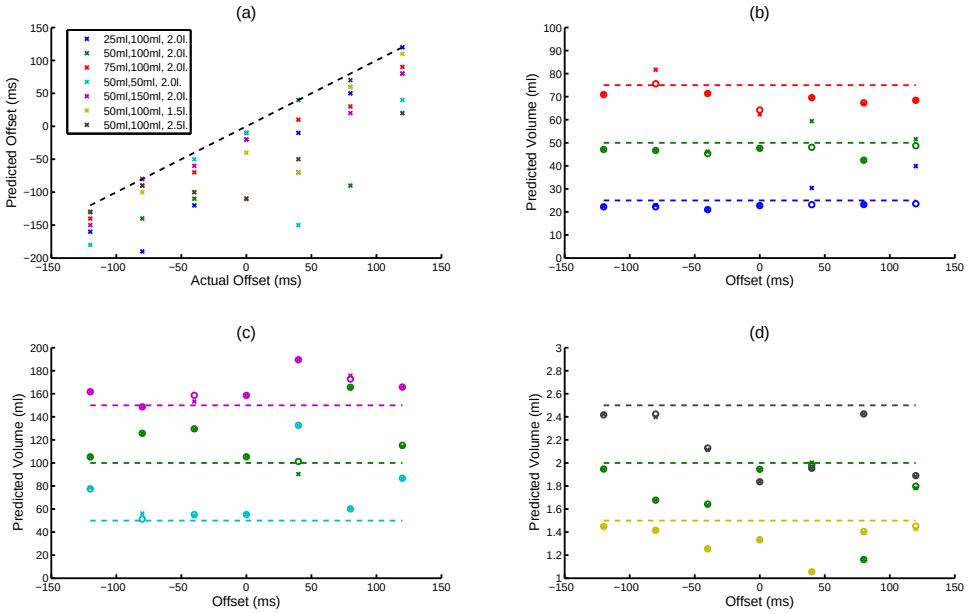


Figure 4.26: Parameter recovery when including the offset between the flow and concentration signals as an additional parameter. A genetic algorithm was used to recover the parameters: (a) offset in time between the two signals, (b) mixed compartment volume, (c) transitory compartment volume and (d) alveolar compartment volume. The compartment volumes are given in the key in the order: mixed, transitory and end-expired alveolar volume. The crosses show the results of the genetic algorithm, the circles show the results of the Nelder-Mead simplex when the results of the genetic algorithm are used as the starting point and the dashed line shows the actual values.

Table 4.9. A discrete set of values is used for the offset parameter at $10ms$ intervals, the sampling interval of the simulated data. The offset in the data is not recovered well with a significant variation between the offset recovered for each different simulated lung. There does, however, seem to be some linear trend in the results with higher actual offsets resulting in higher predicted offsets. The predicted offsets are never an overestimate.

The compartment volume predictions are generally quite variable. In general the mixed compartment volumes are predicted better than the other volumes with less absolute and percentage error. A Nelder-Mead simplex algorithm was used in an attempt to improve the genetic algorithm results. The offset was fixed at the result of the genetic algorithm and only the compartment volumes estimated, with the genetic algorithm results used as the starting point. The default convergence criteria were used: stop when the differences in all parameter values and the RMS error on all points of the simplex are less than 10^{-4}. In most cases this does not change the result to any significant degree but in a few cases a large improvement in the results is observed. In all cases except one (where it results in a additional error of $5ml$ in the transitory

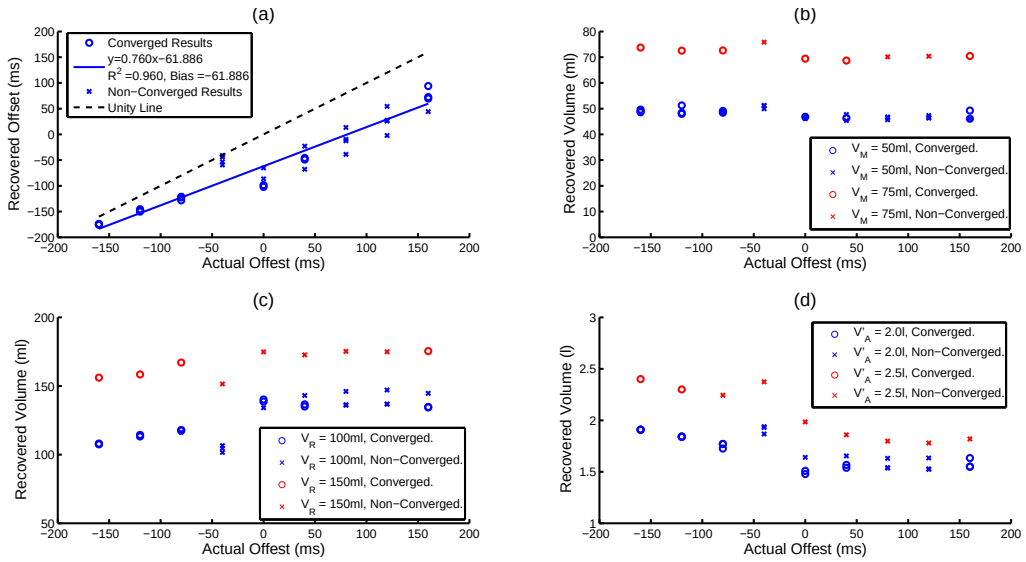


Figure 4.27: Parameters recovered when including the offset between the flow and concentration signals as an additional parameter to be recovered. ABC SMC was used to recover the parameters: (a) offset in time between the two signals, (b) mixed compartment volume, (c) transitory compartment volume and (d) alveolar compartment volume.

volume prediction) the simplex results are an improvement on the genetic algorithm result.

Where the genetic algorithm predicts an offset with little or no error, the results of the Nelder-Mead simplex are close to the actual compartment volumes, but this is not always the case for the genetic algorithm results. In general, larger errors in the offset result in larger errors in the compartment volumes determined by the simplex. For example the offset for the baseline volumes (shown in green) at $40ms$ is predicted correctly and so are all the compartment volumes but at $50ms$ there is a large error in the offset prediction which results in errors in the compartment volumes predicted, especially in this case for the transitory and end-expired alveolar volumes.

To avoid overcrowding the figures making them unreadable the results for all sets of simulated data are not shown in the compartment volume sub-figures in Figure 4.26. Similar results are seen in the cases not shown. In addition similar results are seen, for a lung with baseline compartment volumes, when the convergence criteria of the genetic algorithm is decreased by an order of magnitude to 10^{-7} for 50 generations and when noise at the typical level is added to both signals.

4.6.3.3 ABC SMC

ABC SMC parameter estimation was also used to determine the offset and other parameters; the mean results are shown in Figure 4.27. Uniform priors were used for all

parameters with the limits given in Table 4.9. Four sets of data were simulated at each offset: one where baseline compartment volumes were used and three sets where the mixed, transitory and end-expired alveolar volumes were increased in turn to $75ml$, $150ml$ and $2.0l$. This is a subset of the simulation used in the genetic algorithm case.

Convergence was determined manually by monitoring the mean and variance of all parameter distributions over successive populations. In a number of cases, as indicated in the figure, convergence was not achieved after at least 60 days of processing time. In these cases the results of the final population are still shown but with an indication that convergence was not judged to have occurred.

The recovered offset is below the actual offset in all cases. The amount by which the results are underestimated increases as the actual offset increases, as can be observed by the gradient of the regression line for converged results being under one. Also the number of converged results decreases as the actual offset increases. The exceptions to this are the results for -$40ms$ which are much closer to the actual value, but none of these results were judged to have converged.

The error in the estimate of the offset appears to affect all the volumes recovered, with the volumes recovered at an actual offset of -$160ms$ being closest to the actual volume in all cases and gradually getting further from the true volume as the actual offset is increased. The non-converged results for -$40ms$ are again the exception with the results being close to the actual volumes. The results for the mixed compartment are least affected by these errors with the percentage error at all offsets being low, while the end-expired alveolar volumes are most affected.

The full distributions for the recovered offset are shown in Figure 4.28 for two of the cases, along with scatter plots showing the particles accepted at given populations. The width of the distribution for data with an offset of -$160ms$ is quite narrow. The histogram and scatter plots show few particles accepted above -$160ms$, apart from in the early populations, while particles are accepted down to -$200ms$, the lower limit of the prior, at all populations. A similar pattern is seen in the $150ms$ offset data with few particles accepted above the offset of $160ms$ but a large number for lower offsets. This is not curtailed by the prior, resulting in a wide distribution for the possible offset values. Similar results are seen for the full range of offsets (not shown here) with distributions of results gradually widening as the offset increases. The exception to this is an offset of -$40ms$ which includes a similar range of accepted particles to $0ms$ but with a higher number of particles accepted close to -$40ms$ with generally higher weights.

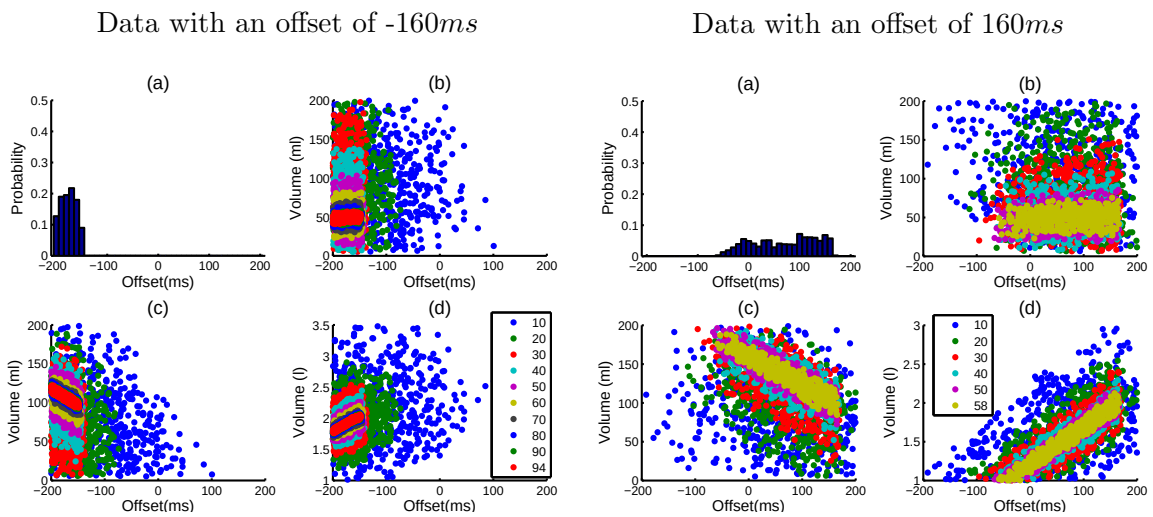


Figure 4.28: Plots showing the offset recovered from data with an offset of -160ms and 160ms. (a) Weighted histogram of the recovered offset. (b) to (d) are scatter plots showing the accepted particles at the populations detailed in the key. Offset is plotted against recovered (a) mixed, (b) transitory and (c) end-expired alveolar compartment volume. The ABC SMC algorithm was manually judged to have converged in both cases. The highest population shown is the final output population before the algorithm terminated.

The scatter plot results for 160ms show that accepted particles come from a relatively large area of parameter space compared with -160ms. This indicates that there are more possible locations for an accepted particle at a given tolerance, indicating plateaus of tolerance within the parameter space. The scatter plots indicate that some correlation exists in the tolerance values between the offsets and the transitory dead space and end-expired alveolar volumes, but this is not the case for the mixed compartment volume.

4.6.3.4 Discussion

All the methods for recovering the offset show one universal result: if the offset is recovered with minimal error then the compartment volumes recovered also show minimal error. Therefore the challenge is to recover the offset correctly. However, the only method that managed to recover the offset sufficiently well to improve the results across the range, compared to the volumes recovered when no correction for the offset was made, was the Nelder-Mead simplex at a 100ms starting point. The results for other starting points did not provide similar results, suggesting that there are multiple minima within the parameter space which would be a general concern when using this method. The exception to this is cases where the actual offset is below the starting point, in which case the results from different starting points are closer to each other

and the actual volume.

The results from the genetic algorithm and ABC SMC show similar results. Both show better results at lower offsets, especially ABC SMC, but this may in part be due to the lower limit for the offset curtailing the parameter space. Both also find minima away from the actual offset (or at least areas of similar tolerance in ABC SMC's case) which are exclusively at lower offsets than the actual offset, the same location as the most erroneous Nelder-Mead simplex results.

The ABC SMC scatter plots of the accepted particles at a selection of populations provide further information. Although slightly different metrics are used by the least squares and ABC SMC methods (root mean squared error and the Euclidean distance) the relative magnitudes will be the same in both cases and so information gained from these plots is relevant to all methods. The plots show an area of similar tolerance over a wide range of offsets below the actual offsets but a much steeper increase in the tolerance for offsets above the actual tolerance.

This provides an explanation of why only underestimates are made. In the case of ABC SMC this results in the distributions being biased towards lower offsets as they are able to achieve the required tolerance but higher offsets are not. For the Nelder-Mead simplex and the genetic algorithm, although in theory they should find minima within these plateau, in reality if the RMS error is not changing at a sufficient rate, then the convergence criteria may be met at a point away from an actual minima. Making the convergence criteria more stringent may result in better results, however reducing the convergence criteria of the genetic algorithm by an order of magnitude was not sufficient to result in a noticeable improvement in the results. This will also increase the computational requirements. It is also possible that the Nelder-Mead simplex is getting stuck in local minima. There is no evidence of local minima within the plots but, if present, changes to the convergence requirement would make no difference in this case.

The plots also show why errors in offset do not affect the mixed compartment volume to the same extent as other volume predictions. The plot for the mixed compartment does not show any noticeable correlation between the offset and mixed compartment tolerances, while there is an obvious correlation in the other two cases; thus lower offsets result in over and under estimates in the transitory and end-expired alveolar volumes respectively. This would also explain why starting points above the actual offset result in better predictions when using the Nelder-Mead simplex. In these cases the area of similar tolerance is likely to be approached from the higher offset side and

so converge on a minimum within the higher offset area of the plateau closer to the actual offset.

The better results, for all methods, in the case of lower actual offsets are not likely to be due to a genuine ability to recover these offsets better. Instead it is likely that the plateau is being approached from the right direction in the case of the Nelder-Mead simplex and that the lower limit reduces the size of the plateau in the other two cases thus reducing the scope of possible errors.

If a reasonable estimate of the offset was available in advance it may be possible to get better results from these techniques. In the Nelder-Mead simplex this could be used to select the starting point, but this may be no better than choosing a high offset. In the case of the genetic algorithm this could be used to inform the bounds of the parameter space better. In the case of ABC SMC it would be used to inform the prior, for example the limits of a uniform prior or the mean and standard deviation of a normal prior. This may however provide no better results than selecting a reasonably good estimate for the offset.

These results indicate that correcting for any offset in the two signals will result in a noticeable improvement in the predicted compartment volumes, however that correcting for the offset at the same time as inferring the parameters of the lung may not be a simple process. Uses of the Nelder-Mead simplex with a starting offset above those expected in the data may provide improved predictions for the lung volume. The ability to improve results by correcting the offset between the signals will be investigated further in real data later in the thesis.

4.7 Soluble Gas Data

So far this chapter has concentrated on insoluble gases; within this section the focus now shifts to the use of a soluble gas. This is motivated by the fact that all the available data are collected using a nitrous oxide tracer gas. Our collaborators continue to work with another sensor that can collect nitrous oxide, oxygen and carbon dioxide data. Ultimately this may allow the above techniques to be used by determining the nitrogen concentration as the remaining constituent of air or by developing the techniques for use with oxygen, which can act like an insoluble gas under certain conditions. However in the absence of this data the possibility of making simplifying assumptions in order to determine the volume of the lung is investigated.

As detailed previously, the assumption that the sinusoidal component of the venous return has little effect on the results derived from sinewave techniques has received much

research. However this relies on the fact that the tracer gas concentrations within the lung, and more importantly the body, can be brought into equilibrium. In reality for nitrous oxide, with a λ value of 0.47, this may take a considerable amount of time to achieve. For example in experiments by Williams *et al.* [224] in dogs this took longer than 20 minutes to achieve.

This may not be such a limitation when monitoring ICU patients who are ventilated for days, however it is not practicable for other shorter term monitoring. For example, equilibrium was not achieved in the human data used later in this thesis, as it was not felt that the non-ventilated volunteers would tolerate an experimental procedure that would take that long. In this section, two alternative methods of dealing with the fact that steady state has not been achieved are presented and their robustness investigated.

4.7.1 Simulated Data

Here data are simulated using the homogeneous model of the lung with a dead space made up of a transitory and mixed compartment. The model described in the previous chapter (see Section 3.3.4) which accounts for the volume of tracer gas taken or returned from the blood is used. The body is modelled by the standard five compartment model of Hahn *et al.* [91], described earlier, with the baseline values given in Table 2.1. The baseline lung volumes ($V_M = 50ml$, $V_R = 100ml$ and $V'_A = 2.0l$) and breathing pattern are the same as used elsewhere in this chapter. A baseline shunt fraction of 3%, typical for healthy subjects [221], and circulation time of $14s$ ($7s$ arterial and $7s$ venous), as justified by Gavaghan and Hahn in [83], are used.

A forcing sinewave of unit amplitude about a zero mean is not used here, as the equilibrium process is being investigated. A sine wave with a mean and amplitude of 5% and 4% is used, which is typical of the human data presented later in Chapter 6. A manual process is used within the human data to bring the lung into equilibrium. This process changes between data set and operator; a typical example is shown in Figure 4.29a. This process is standardised for all simulated sets of data. A $100s$ equilibrium time was assumed; $60s$ or $100s$ equilibrium times are observed in the data. A 10% inspired concentration of tracer gas is implemented for the first $20s$, followed by 7% for the remaining $80s$. This brings the end-expired concentration up to approximately 5%. This process can be seen in Figure 4.29b.

The process brings the air in the alveoli to approximately the mean concentration of the forcing sinewave (5%) but the concentration of the body is not in equilibrium with the lung. This is clear from the fact that the inspired concentration has to be

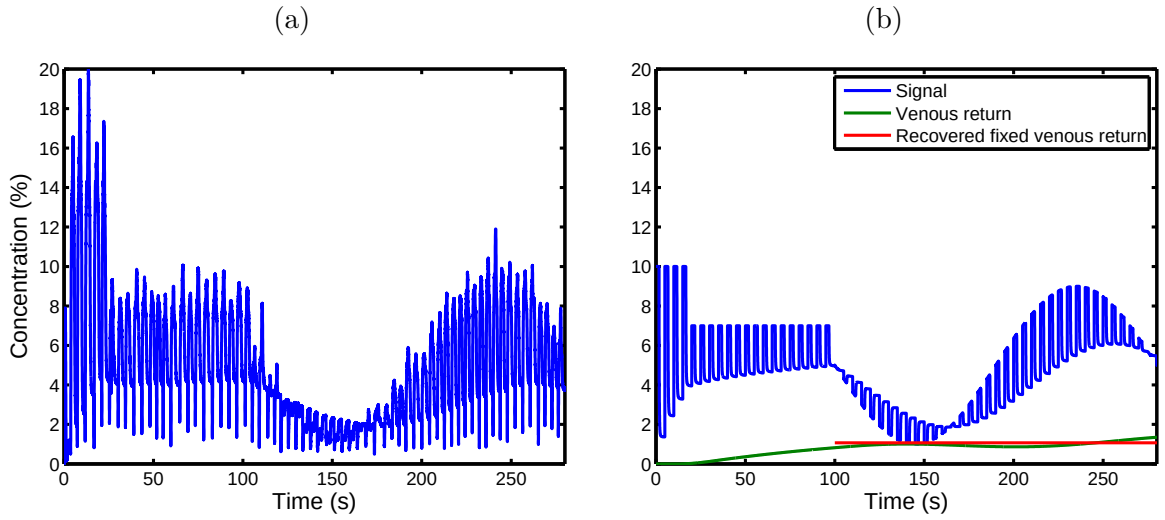


Figure 4.29: Typical data showing the alveolar compartment being brought into equilibrium over $100s$. (a) Human data, described further in Chapter 6. (b) Simulated data, created using the model and baseline values in the text. In addition simulated mixed venous concentration (which is not available in the human data) and the fixed venous return, determined via the method described below, are shown. In both case only the first $280s$ of data are shown.

maintained above 5% in order to bring the end-expired concentration to that level. This is also the case in the real data, as can be seen in Figure 4.29a. The volumes and other parameters will be changed from the baseline values during these investigations. This will affect the equilibrium process, however the end-expired concentration at the end of the process was not observed to change greatly, and significantly less than the difference observed in the human data, therefore the same equilibrium process was maintained for all simulated data.

4.7.2 Fixed Venous Return

As the body has not been brought into equilibrium with the lung, it will not be possible to assume that venous return will not be at a concentration equivalent to the mean of the forcing signal, as assumed by previous research (as described in Sections 2.4.3 and 2.4.4). Instead the concentration will start below that values and generally increase during data collection, as shown for the simulated data in Figure 4.29b. This method makes the assumption that the mixed venous blood is at a constant concentration, which is included as an additional parameter when estimating parameters of the model. Figure 4.29b shows the fixed venous return concentration recovered using ABC SMC for data with the baseline parameter values.

In order to test the robustness of this method to changes in the physiology of a particular subject, data were simulated using the baseline parameters detailed above.

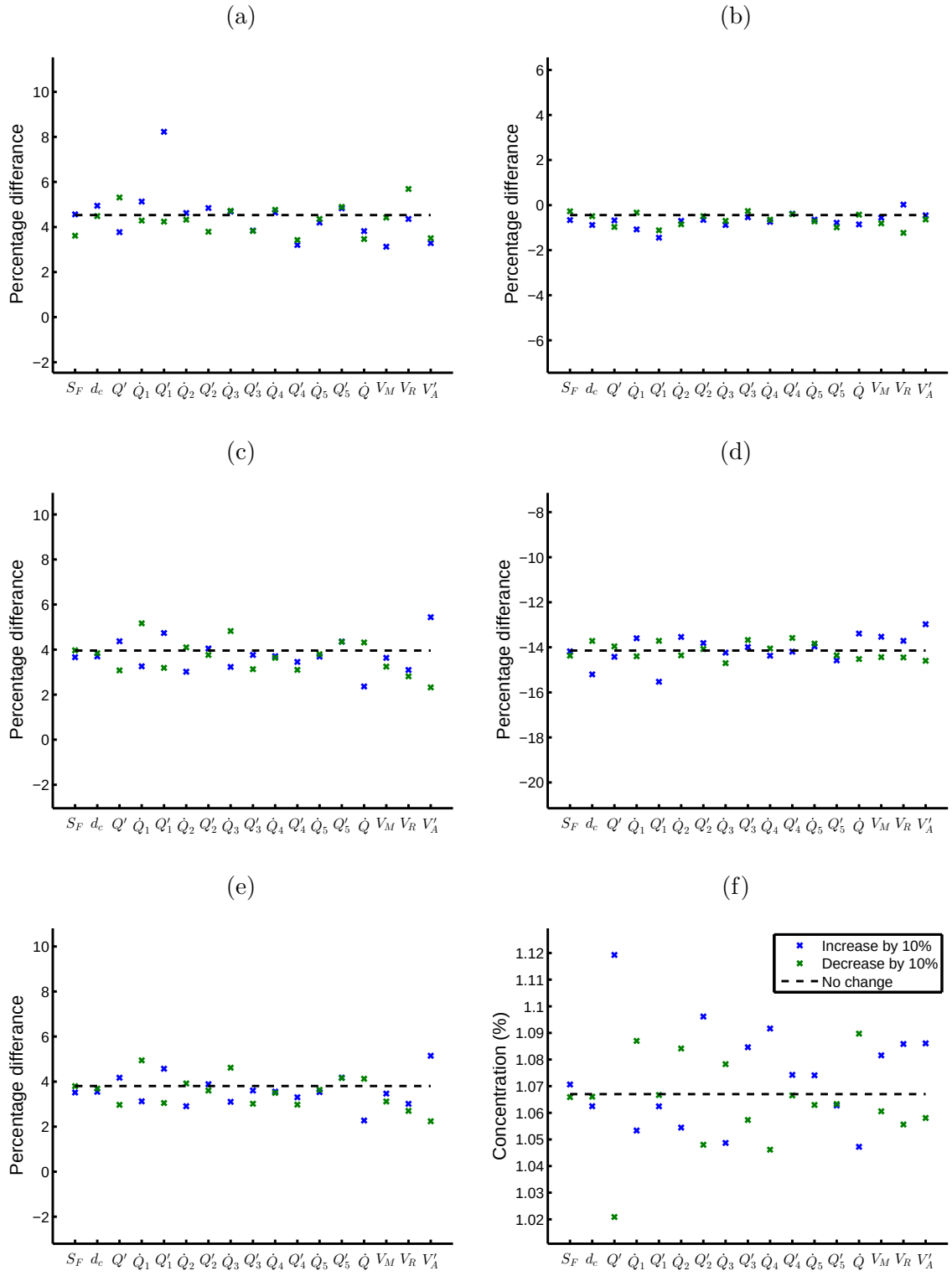


Figure 4.30: Plots showing the recovered values when the venous return is assumed to be at a constant concentration. The simulated data came from the model described within the text. Results are shown when the parameter values are increased and decreased by 10% from the baseline values. The percentage difference of the actual value of parameters is shown in: (a) Mixed compartment volume, (b) Transitory compartment volume, (c) End-expired alveoli volume and (d) Pulmonary blood flow. (e) shows the percentage difference in the FRC (sum of the recovered volumes for the three lung air compartments). (f) shows the actual recovered value for the fixed venous return concentration, as there is no equivalent within the model used to create the data.

The make up of the lung and body was then changed by increasing and decreasing the following values by 10% in order to observe the associated effect on the results:

- The shunt fraction.
- The circulation time delay.
- The total equivalent blood volume (the equivalent blood volume of all the body compartments was adjusted by 10%).
- The equivalent blood volume (blood volume plus the equivalent blood volume of the tissue) of each body compartment.
- The blood flow through each body compartment.
- The total pulmonary blood flow (the flow to all body compartments is adjusted by 10%).
- The mixed dead space compartment volume.
- The transitory dead space compartment volume.
- The alveolar compartment volume.

Figure 4.30 shows the results. For data simulated using the baseline parameters, the mixed compartment and end-expired alveolar volumes were overestimated by approximately 4%. The volume recovered for the transitory compartment were close to the actual value, with an underestimate of 0.4%. This results in the FRC being overestimated by approximately 4%, as the transitory volume is significantly smaller than the end-expired alveolar volume so errors in the recovered end-expired alveolar volume dominate. The recovered pulmonary flow was underestimated to a more significant degree by approximately 14%. It is impossible to compare the fixed venous return to any actual volume as an equivalent parameter is not present within the model used to create the data.

When the parameters were adjusted from the baseline by 10%, the over and under estimates generally remained approximately the same. The lung volumes and pulmonary flow are within $\pm 1.5\%$ of the values recovered at baseline, except in the following cases:

- The mixed compartment volume when the blood flow to body compartment 1 (fast viscera compartment) is increased by 10% (3.7% above baseline);
- The end-expired compartment volume when the total blood flow is increased by 10% (1.6% below baseline);
- The end-expired compartment volume is decreased by 10% (1.6% below baseline).

Although the latter two are a lower deviation from the values recorded at baseline they have a greater effect on the recovered FRC, as the end-expired alveolar volume is significantly larger than the mixed compartment. Deviations from baseline of up to $\pm 0.05\%$ are seen in the recovered fixed venous return values.

4.7.3 Assuming Body Compartments of a Standard Man

An alternative approach is to include recirculation within the model. Attempting to infer all of the model parameters of the five body compartments, the shunt and circulation delays within the model of a standard man is not realistic. Therefore one approach is to investigate whether assuming the parameters of a 'standard man' will allow the volumes of the lung to be estimated without a significant error.

The three lung volumes and the pulmonary flow were inferred while all the other parameters were assumed to be constant at the baseline values. Simulated data were then created with a 10% increase or decrease in the same parameters as for the fixed venous return case. This method has the advantage of requiring one fewer parameter to be inferred than in the last method.

In this case the initial conditions for the body compartments are required in addition to the initial conditions of the lung compartments. Although it is fair to assume that the lung compartments are at the end-expired concentration at the end of the equilibrium process, this assumption cannot be made for the body compartments; therefore the recovery process is run from time zero assuming that all compartments are initially at zero concentration.

The results when the body of a 'standard man' are assumed are shown in Figure 4.31. The recovered parameter values from the baseline data are close to the actual values, as would be expected because the assumed body parameters match those which produced the data. The deviation of the recovered values when the parameters are adjusted by 10% is roughly comparable to the fixed venous return results for the transitory and end-expired alveolar volumes. The one exception to this is when the transitory compartment volume is adjusted, with the recovered transitory compartment volumes 1.7% and 4.2% being greater than that recovered at baseline.

The deviation from the volumes recovered at baseline appear in general to have slightly increased for the recovered mixed compartment volumes. In the cases when the transitory compartment volume of the simulated data is adjusted, this deviation is more significant at 5.9% and 6.9% below that recovered for the baseline data. The slightly larger deviations of the recovered mixed compartment volumes from baseline values do not however affect the FRC results, because the mixed volume has less effect

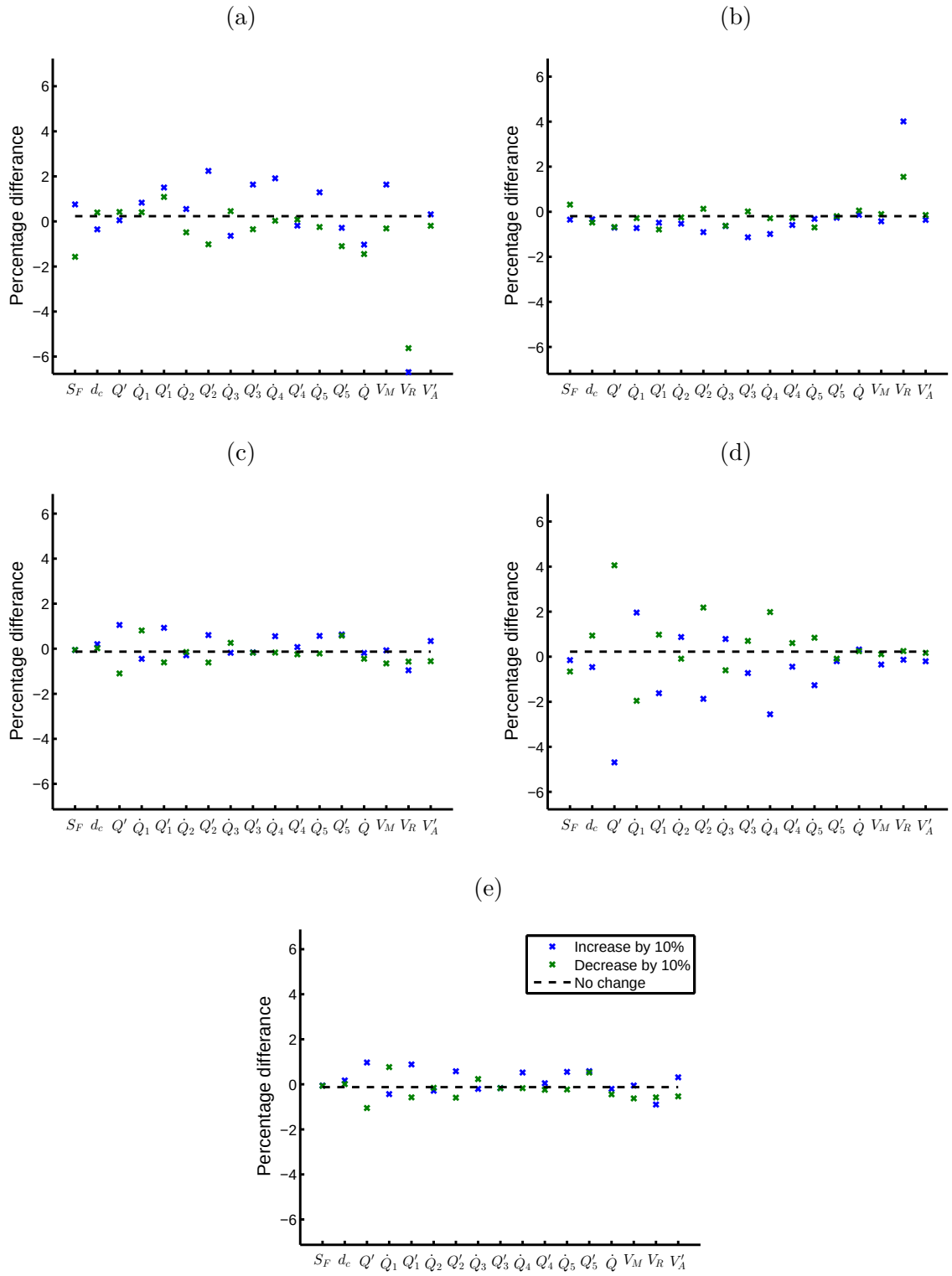


Figure 4.31: Plots showing differences in the recovered parameters when the body of a 'standard man' is assumed. The simulated data came from the model described in the text with the parameters increased and decreased by 10%. The percentage difference of the recovered parameters are shown in: (a) mixed compartment volume, (b) transitory compartment volume, (c) end-expired alveoli volume and (d) pulmonary blood flow. (e) shows the percentage difference in the FRC.

on the FRC than the end-expired alveolar volume. Also the large errors observed when recovering mixed and transitory volumes when the transitory volume is adjusted in the simulated data do not affect the FRC, because they are in opposite directions. This results in a FRC which has little error and comparable deviation from the values recovered at baseline as the fixed venous return method.

In all cases the deviation from the actual pulmonary blood flow is much lower, with a maximum deviation of -4.7% compared with the general deviation of approximately -14% in the case of a fixed venous return. However the deviation from the flow recovered at baseline is generally greater than that observed in the fixed venous return results.

4.7.4 Discussion

If the determination of FRC is of primary importance then these results would imply that the 'standard man' method is better than the fixed venous return method, as the percentage difference from the actual FRC is less in all cases. The maximum difference is 1.1%, which is equivalent to $25ml$ error in the FRC. This is unlikely to be clinically significant and less than that caused by other forms of error. Also the deviation from the result recovered at baseline is comparable to the fixed venous return results.

Use of the fixed venous return method may still provide sufficiently accurate results. The percentage difference of 4% (approximately present in all results) will result in an error of less than $100ml$. This is comparable to systematic errors found in other techniques, for example when implementing helium dilution the uptake of helium into the body generally ignored but thought to result in a error of approximately $100ml$ [44]. If the application is only concerned with the trend of any changes in FRC then these results indicate that the fixed venous return method is likely to perform similarly to the standard man method, as deviation from the values recovered at baseline are comparable.

These observations are also true if the end-expired alveoli volume is being determined as the results show a similar pattern, which is not surprising as it provided the majority of the volume within the FRC results. This is not the case if a good representation of the mixed or transitory volumes is required. Although in general the mixed compartment volume recovered using the standard man method shows less error than the approximate 4% systematic error seen in all the fixed venous return cases, when the transitory volume is adjusted in the simulated data this method underestimates the value by approximately 6%. Therefore if the trend in the mixed compartment volume is required then the fixed venous return method is likely to be the preferable method.

For the recovery transitory volume little systemic error is observed in all results when using the fixed venous return, but a overestimate of 2 to 4 % is seen in the standard man case, suggesting that the fixed venous return method would be prefered when determining the transitory volume.

The ability to determine the pulmonary blood flow via either method is not as favourable as for the air volumes of the lung. The fixed venous return method generally underestimated the flow by approximately 14%, which is around $0.75l/min$. However this underestimate is not very sensitive to changes in the parameters of the simulated data, varying by a maximum of 1.4% about $0.08l/min$, therefore this method may have the ability to determine the trend of any changes in the blood flow while being subject to a general systematic error. The errors when the standard man method is used are less in absolute terms, with a maximum absolute error of 4.7%, but this method is more sensitive to changes in the parameters of the simulated data, meaning that this method may not be as robust at determining the general trend of any changes in blood flow.

There are a number of limitations in the results presented here. The investigations carried out here are a basic form of sensitivity analysis, however it is acknowledged that more robust forms of sensitivity analysis are available. These would require additional simulations to be carried out including those which investigate the implications of changing multiple parameters at the same time. In addition the relative variability of each parameter would be taken into account, for example the range of tissue volumes for the body fat compartment is likely to be higher than the lean compartment within a standard population. The investigations carried out here are only an attempt to assess the basic viability of these methods before testing them in human data later in the thesis.

The baseline value used for the shut fraction is 3%, which is a typical value for healthy subjects but higher shunt fractions, typically around 20%, can be observed in certain patients [221]. Further work would be required to investigate the sensitivity of the methods to these higher levels of shunt. These investigations are based on the standard parameters of a man, as they are readily available within the literature, and so may not be as applicable to women. This will be addressed further later in the thesis when they are applied to human data from both men and women.

A single equilibrium process has been assessed here, however as a manual process this will inevitably vary between data sets. The manual nature of this process makes it difficult to simulate all possible equilibrium scenarios, but investigations could be

carried out to inform the implementation of better manual or automatic equilibrium processes. These could bring the whole body into a state nearer equilibrium in a timely manor or at least ensure that any systematic errors are comparable between successive tests. Experimental error has not been accounted for within these simulations. Two methods of simplification have been investigated here, but there are others that could be investigated, such as the use of a single compartment to represent the body, as used previously by Williams *et al.* [229].

These investigations support the possibility of using either method to determine the trend of any changes in FRC, with the standard man method more likely to perform better in determining the actual value. Further investigation in human subjects would be prudent, as will be presented later. The results for determining the pulmonary blood flow are not as favourable, however it may be possible to determine changes in blood flow rate using the fixed venous return method, if not the actual value reliably.

4.8 Summary

4.8.1 Limitation of the Simulated Data

Although there are many advantages to investigating methods using simulated data this does have some limitations. These come from two primary sources: limitations of the model used to create the data and the ability to simulate the experimental errors present in real data. There are many assumptions made in the models used to create the data investigated throughout this chapter.

In all cases the dead space is assumed to be made up of two compartments; a mixed compartment at the mouth followed by a transitory compartment. In many ways this breaks from the traditional tidal models of the lung [84, 90, 216–218], which in general would be analogous with the single transitory model of the dead space, however these models are conceptually different, only modelling a few fixed concentrations at particular time points not the entire breathing cycle. These models also assume that some mixing occurs within the dead space; the air left in the dead space at the end of inspiration and expiration is assumed to leave the compartment, not in the concentration in which it entered, but at the flow weighted average concentration of all air present within the dead space at the end of inspiration or expiration.

In reality, mixing of the air within the dead space will occur to some degree, even within the water displacement bench lung, which is set up to mimic the concept of a lung made up of two well defined compartments. There are still likely to be some areas

of turbulence which will cause mixing of the air, particularly at abrupt changes in the pipe diameter.

Ultimately the question is whether the addition of a mixing compartment improves the ability to infer information about the lungs. Attempts to answer this question will be presented later in the thesis. The purpose of the work carried out in this chapter is not to answer that question, but to assess the viability of the above methods when mixing occurs within the dead space.

There is of course an issue about what size of mixing compartment is required to be representative of the amount of mixing that occurs within a human dead space. The mixing compartment is to some extent a conceptual compartment and therefore it would be difficult to determine a representative size from anatomical measurements, but that is also the case for the division between the dead space and alveolar compartments. Here an estimate based on the results from the homogeneous bench lung was used ($50ml$), which was the best available estimate at the time at which the investigations were carried out.

A larger transitory compartment volume was used in conjunction with this value when investigating the log-normal distribution results. This is only likely to affect the homogeneity index results as the dead space was assumed to have a mixing compartment in the other cases, as were the volumes of both dead space compartments. In the homogeneity index case other investigations would suggest that the slightly larger transitory compartment volume might cause the index to be slightly overestimated, but this is likely to be minimal as the tidal volume is over twice the combined dead space volume.

The alveoli are modelled in various ways throughout this chapter. Where the homogeneity of the lung was investigated, two models were used: a multi-compartment model of the lung and a series of peaks with a log-normal distribution of ventilation-volume ratios. The multi-compartment model of the lung is unlikely to model the realities of a human lung apart from in a very general sense. Modelling the inhomogeneities in this way has found moderate success in previous research [199], however it is primarily included here to provide better understanding of the data from the multi-compartment bench lung, which is analysed later in the thesis.

The log-normal distributions are unlikely to model completely the complex ventilation-volume distributions found in human subjects, especially those with a form of lung disease. However as described and discussed already, many previous researchers have found them to provide a good representation of the behaviour found in inhomogeneous

lungs [117, 137, 207, 216]. They therefore provide a good starting point from which to assess the viability of a particular method.

Within real data there are many forms of experimental error which have potential to affect the results, for example calibration errors or non-linearities within the sensors. Here only two of the most problematic sources of error have been investigated: the noise within the two signals and the misalignment between the two signals. Noise is random in nature, making it problematic to correct for, therefore it is valuable to assess the robustness of the proposed methods to given levels of noise. Determining a reasonable expectation for the levels of noise within a signal is difficult; here an attempt has been made to determine the typical noise level, however to guard against underestimates the results for double the expected noise level have been presented, where computationally feasible to do so.

4.8.2 Lung Homogeneity

A large part of this chapter is concerned with methods to determine the homogeneity of the lung. These investigations will be summarised here. Three broad methods of determining the homogeneity of the lung were investigated: the homogeneity index, the inter-breath variation method and using ABC to specifically infer the homogeneity present.

The homogeneity index and inter-breath variation method were applied in the same way to both log-normal and multi-compartment data. The homogeneity index is 0.12 (Bohr Method) and the standard deviation of the inter-breath variation is 0.16 for configuration C of the multi-compartment lung, this compares with 0.14[3] and 0.21, respectively for a uni-modal log-normal distribution with $\sigma_{x,1} = 0.5$. This indicates that the homogeneity of configuration C is less than the homogeneity of the $\sigma_{x,1} = 0.5$ log-normal distribution, by these measures. Thus on these measures, the magnitude of configuration C's homogeneity is less than would be expected to be observed in young healthy subjects, assuming that the widths of the distributions determined by Lewis *et al.* [117] are representative and that they translate into the standard deviation of a uni-modal log-normal distribution.

This has two implications. Firstly, near homogeneous lungs may not be found in reality, so some of the absolute errors and increases in standard deviation found in the homogeneity measures when applied to near homogeneous lungs may be less of an issue. Secondly the largest homogeneity found in the multi-compartment lung may be less than half that found in actual clinical cases.

[3]This value may be slightly inflated by the transitory volume used.

Assuming that the standard deviation of the distribution is used as the measure of homogeneity, the inter-breath method generally performed better than the homogeneity index. The magnitude of the change in standard deviation of the distribution was slightly more than for any given change in homogeneity. The standard deviation of the measure was also generally less in the presence of noise, especially for homogeneous or near homogeneous lungs. The mean of the measure is also less affected by noise, except in the case of a homogeneous or near homogeneous lung. The one exception to this is when only noise in the concentration sensor is present, when the homogeneity index is little affected.

A number of situations in which the homogeneity index could indicate that a homogeneous lung was inhomogeneous were also identified: when the mixing compartment was sufficiently large, the tidal volume particularly low or the concentration signal lagged behind the flow signal. It should be noted however that the inter-breath method was not specifically tested at these extremes but is likely to perform better due to the dead space mixing being specifically modelled. Also the method assumes knowledge of the dead space volumes. The method of determining the dead space volumes, proposed above, is also affected in a similar way by large mixing compartments. The transitory dead space results also indicate that it is likely to be affected by low tidal volumes. In addition the effect of noise on interdependently recovering the dead space volumes has not been investigated.

The models proposed for directly inferring the homogeneity using ABC SMC differ between the log-normal and multi-compartment lung due to the nature of the data. The results show that it is possible to infer the parameters of a uni-modal log-normal distribution in the absence of noise and configurations A and C of the multi-compartment lung in both the absence and presence of noise. However inferring the parameters of a bi-modal log-normal distribution was not possible even in the absence of noise. The situation for configuration B in the absence of noise is not clear due to the lack of convergence in the results, but where noise is added the results were not favourable. It should however be noted that homogeneities found in configuration B may be below that expected in human lungs.

The use of ABC SMC model selection was also investigated. Although the results for the log-normal distribution data indicate the possibility of identifying the number of peaks within the distribution, if not the parameters of the distribution, the results from the multi-compartment model are less favourable. Configurations C and D do converge on the model as expected within a reasonable time limit, but the lack of

convergence even after a considerable amount of time of the other results indicates that the method may be unusable in practice. In separate research, not detailed here, polynomial model results also cast doubt on the general abilities of ABC SMC model selection to determine a good approximation of the posterior distribution when applied to real data with some form of experimental noise.

This leaves the possibility of characterising the lung as a single log-normal peak or two compartments, although the issues with configuration B may suggest that problems could occur in the latter case. Characterising the lung as a number of discrete compartments has had some limited success in wash-out data [170, 199], but a method that restricted the lung to a uni-modal distribution, which was not necessarily log-normal, was not successful [85]. The information content of successive breaths is unlikely to diminish in the same way as for wash-out techniques when using a sinewave. Along with the use of the entire expired signal this is likely to reduce the susceptibility of these techniques to experimental noise. However if a uni-modal distribution is fundamentally not enough to characterise the homogeneities of human lungs then any method that relies on one may be unsuccessful.

The aim here is not to determine the full distribution of any underlying ventilation-volume inhomogeneity, but merely to gain enough information to better guide clinical decisions. To that extent, on the basis of the evidence provided, any of the three possible methods presented above (the homogeneity index, the inter-breath variation method and the identification of a uni-model peak) could potentially fulfil this aim. However for the reasons discussed above further investigations would be required into the sensitivity of these measures to real changes in homogeneity of human lungs.

4.8.3 Other Findings

In addition to investigating methods of determining the ventilation-volume homogeneity the breath-by-breath methods proposed by collaborators were analysed more generally. The Fowler method generally determined the dead space volume more accurately than the Bohr method. It was also less susceptible to changes in the forcing period used and misalignment of the two signals. However the end-expired alveolar volumes recovered for homogeneous or near homogeneous lungs using the dead space volume determined using the Bohr method were closer to the actual volume, except in some specific misalignment cases.

The ABC SMC method was investigated in several ways throughout this chapter. In a number of cases good results were seen when estimating parameters using the mean of the distributions, however there were some occasions when this was not the

case. The offset of the two signals was not recovered well, but nor was it by the two least squares optimisation methods, so it is likely to be a general identifiability problem related to the model rather than one exclusive to ABC SMC. Also the parameters of the bi-modal log-normal distribution were not recovered well, but this is also likely to be a general identifiability problem.

The greater concerns are about the ability of ABC SMC to correctly determine the uncertainty in an estimation of a parameter, which comes from two sources: the lack of increase in the width of the distribution for lung inhomogeneity increases when the end-expired alveolar volume is recovered using the homogeneous model and the fact that the widths of the parameter distributions for the $V'_{A,1}$ and $\sigma_{x,1}$ appear to some extent to depend on the computational effort expended. Alternative possibilities of why these observations are made, other than a fundamental issue with ABC SMC, are expressed above but these observations still seem somewhat counter intuitive.

The lack of convergence in the model selection without selecting a model with full certainty is also unexpected, however this may be not surprising as the data are free from noise. The lack of convergence within a reasonable time scale for some of the multi-compartment lung configurations is likely to make this method unworkable. Further investigations in simulated data with noise from a simple polynomial model (not shown here) provided further concerns about ABC SMC model selection, showing that the algorithm does not truly converge unless one model is selected with full certainty and the model ultimately selected is more complex than may be reasonably expected given the model that produced the data. These concerns will be discussed further in Section 5.3.2.1, when also discussing ABC SMC model selection in bench lung data.

Investigations into the issues of misalignment between the signals confirmed previous research that the alignment of the signals is critical for obtaining accurate results. The methods of correcting for this misalignment, while also determining the other parameters, confirm that if the misalignment could be correctly identified, then the predictions of the other parameters would be likely to improve, but that a robust method of doing so was not forthcoming. Thus on the basis of these results the inclusion of an offset parameter is unlikely to fulfil the aim of the thesis to better determine the FRC from the available data.

Finally standalone investigations into dealing with the practical issue of reaching equilibrium in soluble gas data were investigated. Two methods by which assumptions can be made in order to accommodate for the body not being in equilibrium, showed that the trend in compartment volumes and blood flow changes should be determinable

without a large amount of additional error. Based on the investigations, the standard man method is likely to provide better absolute values for most of the parameters inferred, but the fixed venous return method is likely to be better for determining changes in the pulmonary blood flow rate. Confirmation in human data would be valuable, due to the limitations of the preliminary sensitivity analysis carried out here.

4.9 Key Findings

The key findings within this chapter fall into two categories: those relating to the methods for providing information about the homogeneity of the lung and those relating to some of the practical issues found within real data. In the first case the following key findings are made:

- Even with the potential additional information available from modelling the entire expired concentration signal, only the parameters of a uni-model log-normal distribution can be determined robustly from simulated data without noise.
- Using both the homogeneity index and the inter-breath variation method (which do not attempt to model the inhomogeneities directly) it is possible to gain information about the lung homogeneity from log-normal data, even when typical noise is present on both the flow and concentration signals.
- The inter-breath variation method provides a homogeneity measure with better resolution than the homogeneity index. This is particularly the case when noise is added to the flow signal and for multi-compartment data.
- The homogeneities produced by the simulated multi-compartment lung (which mimics the homogeneities found in the actual multi-compartment bench lung used later) are not as wide ranging as those which were previously reported in humans.
- The homogeneities produced by the multi-compartment bench lung are closer to being fully homogeneous than those previously reported in young healthy subjects. This also happens to be the range of homogeneity at which the inter-breath variation and homogeneity index methods are poor at distinguishing change in homogeneity.
- In simulated data without noise it is possible to recover the characteristics of the dead space independently of the alveolar make up across a range of homogeneities.

In the second case the key findings are:

- An offset in time between the two signals in the data has a substantial effect on the accuracy of the volumes recovered.
- The inclusion of an additional offset parameter in the recovery model improves the results when the correct offset is determined, however for the four different methods tried the offset was not determined robustly enough to provide a general improvement in the volume predictions.
- Based on the sensitivity analysis carried out, both the proposed methods (the fixed venous return and standard man methods) of accommodating for the lack of equilibrium in the body can be used without greatly affecting the FRC predictions.

In addition to these primary findings, some concerns about the ABC SMC method are identified:

- ABC SMC model selection appears not to converge until only one model remains.
- ABC SMC parameter estimation does not appear to capture the uncertainty in a particular parameter with the width of the posterior distribution appearing to continually reduce at higher populations rather than truly converge.
- The technique is several orders of magnitude more computationally intensive than the other parameter estimation techniques used.

4.10 Conclusion

The aim was not to fully characterise the ventilation-volume distribution present within a subjects lung but merely to provide a measure which is sensitive enough to aid clinical decision making. To that extent recovering a uni-model distribution as a proxy for more complex underlying distributions may be sufficient, however given the limitations found when doing so in previous work (albeit using different types of data) it would seem sensible to concentrate on methods that use simpler homogeneous recovery models. This is particularly true as the work carried out here was restricted to an idealised breathing pattern, which was taken as prior knowledge, and did not investigate the effects of noise.

This leaves two methods of which the inter-breath variation method proves more robust, especially when typical noise is added to both signals. This method does rely on the assumption that the dead space volumes can be determined independently of the alveolar make up, however the results here show that to be possible for a range of homogeneities. In the next chapter, these methods will be examined in real data from a multi-compartment bench lung.

The results also show that if the correct time offsets between the two signals can be determined this improves the FRC prediction. However a method of correctly determining the offset was not forthcoming. Both methods of dealing with the lack of equilibrium in the body compartments provided promising results, showing that they can be used without adversely affecting the FRC recovered.

The issues with ABC SMC have meant that in a number of cases it was not possible to use it as initially envisaged. However for simple parameter estimation it did not perform any worse than the other methods used, apart from in terms of computational efficiency. ABC SMC will be investigated further in the next chapter, especially in relation to concerns about the convergence of the model selection algorithm. All these results provide an insight into the benefits and potential pitfalls of the methods proposed; however simulated data can never mimic all the aspects found in real data. Therefore further investigations are required in real data and these follow in the subsequent chapters, both in artificial bench lung data and human data.

Chapter 5

Water-Displacement Bench Lung Data

5.1 Introduction

Within this chapter the methods developed in the previous chapter using simulated data are tested and also developed further using data from a water-displacement bench lung. Data presented in this chapter come from two sources: a homogeneous bench lung with a single alveolar compartment and a multi-compartment bench lung which can have up to three alveolar compartments.

The advantage of data from the bench lung is that the volume of the dead space and alveolar compartments can easily be obtained. The slight exception to that, however, is the minimum volume of the dead space. Although the volume of all the individual pipes that make up the dead space can be determined, this cannot be done to the same accuracy as for the volume of the alveolar compartment or the additional pipes added to the dead space to increase its volume.

The initial part of this chapter describes the investigations completed using the single compartment lung. First the breath-by-breath methods are investigated, before investigations into the different compartment models for the dead space are presented. The aim of these investigations is to determine which of the possible dead space models allows information about the lung to be inferred most robustly. Methods for correcting for offsets in time between the two data signals are then investigated. Finally the method of determining the dead space volume independently of alveolar configuration is investigated.

The second part of this chapter describes investigations into data which came from the multi-compartment lung. First, the ability to determine the dead space volume independently of the alveolar compartments is investigated. The chapter then details the attempts to recover information about the homogeneity of the alveoli, using three of

Dead Space Volume (ml)	End-Expired Alveolar Volume (l)	Forcing Time Period (s)	n
88	2.4	30	2
88	2.4	60	2
88	2.4	120	2
88	2.4	180	2
88	2.4	240	2
88	2.4	300	2
138	2.4	30	2
138	2.4	60	2
138	2.4	120	2
138	2.4	180	2
138	2.4	240	2
138	2.4	300	2
188	1.8	120	2
188	2.1	120	1
188	2.4	30	2
188	2.4	60	3
188	2.4	120	4
188	2.4	180	2
188	2.4	240	3
188	2.4	300	2
188	2.7	120	2
188	3.0	120	2
238	2.4	120	5

Table 5.1: Details of the homogeneous water-displacement bench lung configurations.

the methods described and investigated during the previous chapter: the homogeneity index, the inter-breath variation method and the ability to recover volumes of the two compartments directly. Together these investigations will help build an understanding of how the previously proposed methods perform with real data.

5.2 Homogeneous Bench Lung

5.2.1 Data

The data used to perform this analysis were collected from the homogeneous single compartment bench lung (as detailed in section 3.2.3) using a sinusoidal delivery system (as described in section 3.2.1). In total 52 sets of data were collected. The data were collected for a range of dead space and end-expired alveolar volumes using a selection of different forcing time periods, as detailed in Table 5.1. The data were collected using a nitrous oxide tracer gas. This is a soluble gas, however, as there is no gas exchange within the bench lung, it will act like a insoluble gas. Thus the analysis presented here was conducted as for an insoluble gas.

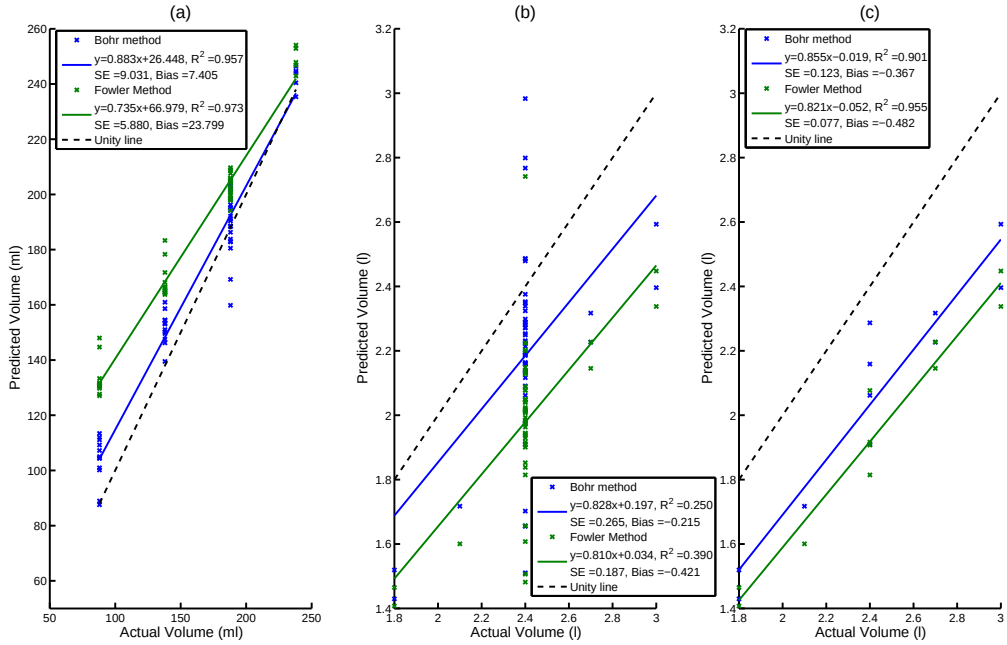


Figure 5.1: Regression plots comparing the actual volume and predictions made using the breath-by-breath method. (a) Dead space volume, (b) end-expired alveolar volume and (c) end-expired alveolar volume for the subset of data described within the text.

5.2.2 Breath-by-Breath Method

How various methods perform on homogeneous bench lung data is investigated throughout this chapter. Two forms of analysis for assessing the performance of each method will be used. The first is a regression analysis; Figure 5.1 shows these results. Here the ordinary least squares method will be used, as the error in the actual bench lung volumes is assumed to be negligible in comparison to the predictions made by the various methods. How well the method has performed can be assessed using four criteria:

- the linearity of the prediction – how close the slope of the regression line is to one and so the slope of the ideal unity line;
- the bias in the prediction – the difference between the prediction and the actual value at the mid-point of the range;
- the precision of the prediction – indicated by the coefficient of determination, R^2, and the standard error (SE) of the estimate [33, 51].

When determining the dead space volume the Bohr method is more linear and results in less bias than the Fowler method. The Fowler method however has a better R^2 and standard error of the estimate. Similar results are seen for the end-expired alveolar volume with the bias and the linearity being slightly better when dead space values are recovered using the Bohr method but a better R^2 and standard error of the estimate are observed when the Fowler method is used.

Method	$V_D(ml)$	V'_A (l)	Subset of V'_A (l)
Bohr	8.0±10.5	-0.22±0.27	-0.37±0.13
Fowler	25.1±13.5	-0.42±0.19	-0.49±0.10

Table 5.2: Mean and standard deviation of the difference between the actual and predicted volumes when using the breath-by-breath methods.

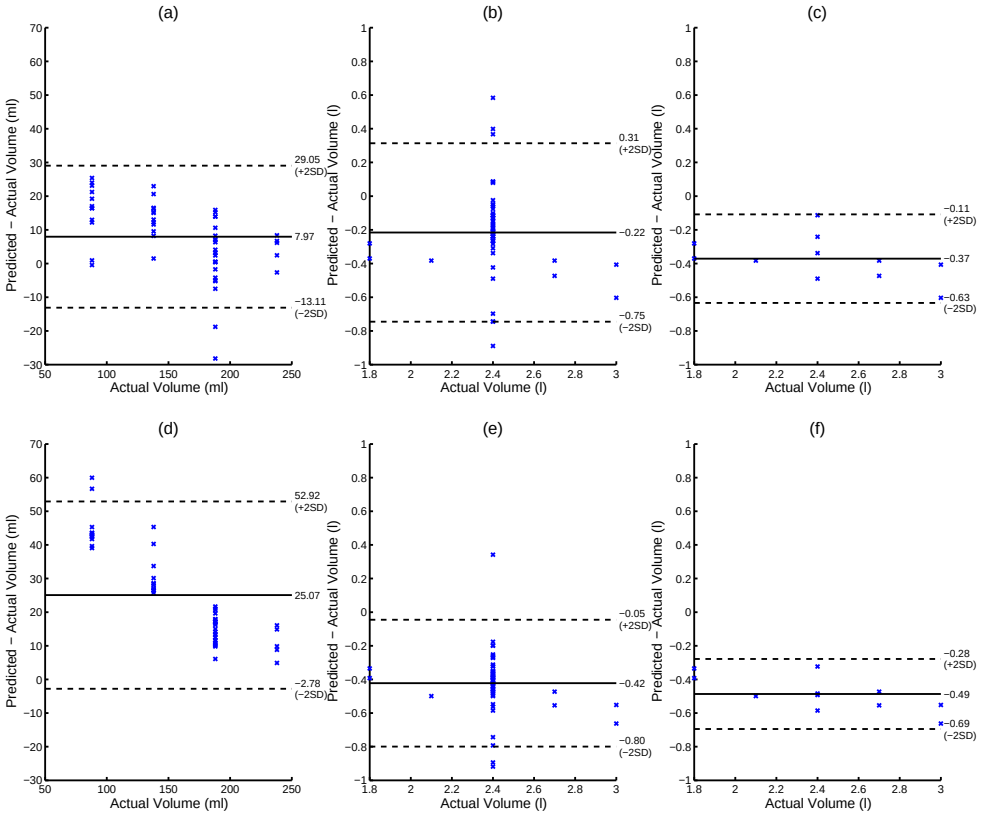


Figure 5.2: Modified Bland-Altman plots (as described in the text) comparing the actual volume to predictions made using the breath-by-breath methods. (a) & (d) Dead space volume, (b) & (e) end-expired alveolar volume and (c) & (f) end-expired alveolar volume for the subset of data described within the text. (a) to (c) show the predictions made using the Bohr method and (d) to (f) the Fowler method.

Throughout this chapter in addition to plotting all end-expired alveolar results, a subset of the results are plotted where the dead space volume and forcing period are $188ml$ and $120s$ respectively. This provides a range of end-expired alveolar volumes which is not dominated by a volume of $2.4l$. This will be referred to as the subset of end-expired alveolar results. The plots of this subset, shown in Figure 5.1 (c), have a similar pattern with bias and linearity favouring the Bohr method but the R^2 and standard error of the estimate values favouring the Fowler method.

The second set of analysis will use statistics from work by Bland and Altman [37], as detailed in Table 5.2. The mean of the error (predicted value minus actual value) is an alternative measure of bias and standard deviation of this error is another measure

for the precision of the method.

In addition Bland and Altman proposed a plot to determine any trend in the error visually. A trend would indicate a form of systematic error, rather than purely random error which for example could be caused by noise in the data. Bland-Altman plots are effective when comparing measurements which have approximately the same amount of variation and will be used later. However there is some debate about what to plot on the x-axis when the comparison is made to a more precise measurement or reference method [28, 38, 65, 99, 112]. It is assumed here that the error in the volumes determined from the bench lung is negligible compared with the predictions made. Therefore the actual volume is plotted on the x-axis here, instead of the mean of the two measurements as proposed by Bland and Altman. This results in the plots shown in Figure 5.2.

The results for both the Bohr and Fowler methods show similar biases to those observed in the regression analysis, with the Bohr method again showing less bias. The limits of agreement are wider for the Fowler dead space volume than for the Bohr volume, which is the opposite observation to that of the regression analysis precision measures. However the limits of agreement of the Fowler method are narrower for the end-expired alveolar volume predictions. All results show a trend in the error with the error reducing as the volume increases, which is what would be expected as the linear term in the regression analysis is under one. This trend seems similar in both cases for the end-expired alveolar volume, but is more pronounced in the Fowler dead space case.

5.2.2.1 Discussion

These results do not provide an unambiguous indication that either the Bohr or the Fowler method is better. The choice will depend on the weight placed on the various indicators presented. For example when considering the end-expired alveolar volume the Bohr method would be considered better for both the bias and linearity measures, but the precision measures would all favour the Fowler method.

Errors in the bias may be considered to be less important, particularly in the case of the dead space where the accuracy of the original volume is not as great as that of the additional volume. In addition it would be possible to scale the results in order to remove the errors in the linearity and bias, which would imply that precision is the most important factor; in which case the results would suggest that the Fowler method was more preferable.

However some caution is required as these scalings are unlikely to remove all linearity and bias errors in future results because the scaling factor will have some error, particularly given the poor precision indicators for the full set of end-expired alveolar results. Further verification in new data, or machine learning type methods of splitting the data into training and verification sets, would be required to assess the benefits of such methods.

5.2.3 Models for Mixing within the Dead Space

Within this section possible models for the mixing of the gases within the dead space are investigated. These models are based on models of the dead space detailed previously in section 3.3. Results are presented for eleven different dead space configurations:

- Model 1 – a single transitory dead space compartment within which no mixing occurs;
- Model 2 – a single transitory dead space compartment with a single mixing compartment at the mouth;
- Model 3 – two transitory dead space compartments of equal volume with a single mixing compartment between them (making the transitory compartments equal in volume means that the dead space can be characterised by two parameters);
- Model 4 – a transitory dead space compartment between two equal volume mixing compartments (this dead space can again be characterised by two parameters);
- Model 5 – two transitory dead space compartments of different volume with a single mixing compartment between them;
- Model 6 – a transitory dead space compartment between two mixing compartments of different volumes;
- Model 7 to 11 – a one to five sequential number of mixing compartments, respectively.

The most appropriate models are determined in two ways, firstly by estimating the dead space and the end-expired alveolar volume and comparing these estimates to the actual volume. The second method was to use ABC SMC model selection to determine which model was most likely to have given rise to the data.

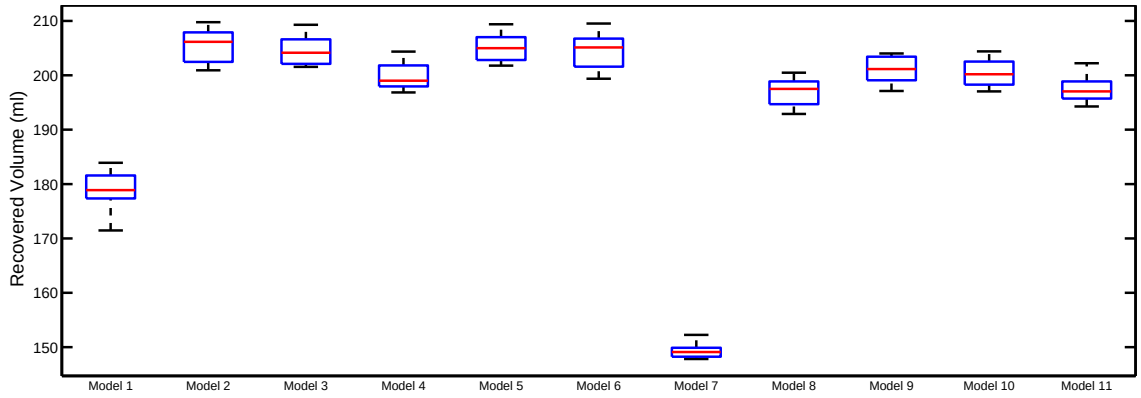


Figure 5.3: The dead space volumes recovered for the subset of the data using all eleven models.

5.2.3.1 ABC SMC Parameter Estimation

ABC SMC was used to estimate the volume of all the compartments within the models described above. A full parameter estimation for all models on all 52 sets of data would be computationally expensive, therefore only a subset of the data was analysed for all models. A uniform prior was used for all parameters with bounds of $5ml$ and $200ml$ for all dead space compartments and $1.0l$ to $3.5l$ for the alveolar compartment.

The means of the distributions for each of the dead space compartments within the individual models are added together to give the overall estimate for the dead space volume. The results are presented as box plots in Figure 5.3. For models with at least two compartments (Models 2 to 6 and 8 to 11) the median is roughly the same at around $200ml$, as is the interquartile range at around 3 to $5ml$. This is above the $188ml$ nominal volume of the dead space. Model 1 (the single transitory compartment model) has a median volume of $179ml$ which is slightly closer to the actual value. The interquartile range is similar, at 4 ml. Model 7 has a lower interquartile range of just over $1ml$, but has a median which is further from the actual value at $149ml$.

The results for the end-expired alveolar volume are shown in Figure 5.4a. When the dead space model includes a transitory compartment, the predictions follow a similar linear trend with a similar slope of around 0.90. The offset bias at the mid-point is roughly the same; Model 1 shows a slightly lower bias of approximately -360ml, while the other models have a similar bias of around -410ml. The R^2 value varies from around 0.92 to 0.94 and the standard error of the estimate is approximately $0.09l$, with Model 2 performing best in both measures.

Where only mixing compartments are included, as shown in Figure 5.4b, the slope of the line varies over a slightly wider range, from around 0.93 and 1.13, but is generally

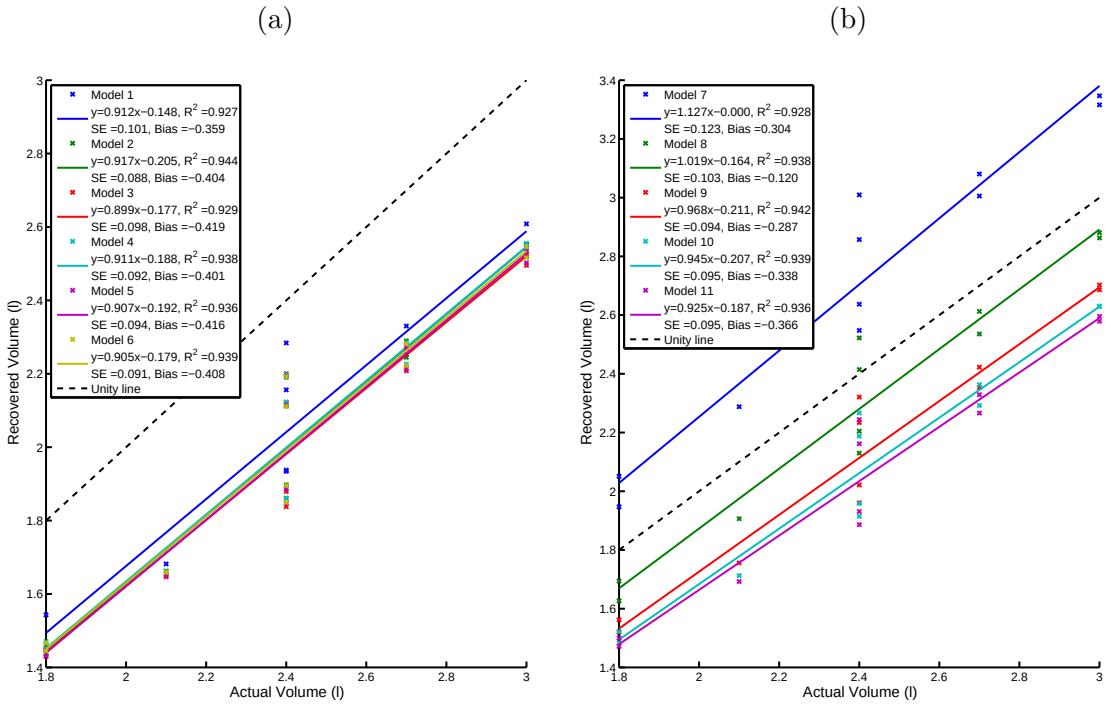


Figure 5.4: The end-expired alveolar volume recovered for the subset of the data using all eleven models described in the text. (a) Models including transitory compartments and (b) models including only mixed compartments.

closer to unity, with Model 8 (two mixed compartments) being the closest of all models. A similar range of R^2 values, from around 0.92 to 0.94, is observed. The standard error of the estimate is again approximately $0.09l$, except in the case of Model 7 where it increases slightly to $0.12l$. The bias changes as the number of compartments increases from $304ml$ (Model 7 – one mixed compartment) to -$366ml$ (Model 10 – five mixed compartments) at the mid-point. Model 8 (two mixed compartments) has the least bias of all the models, including those with a transitory compartment, at -$120ml$.

Due to the large computational requirement, only three models were analysed on the full set of data. Model 1 was selected as the predicted value for the dead space was different to the other models but still close to the actual value. In addition, it is similar to the approach used in previous studies [79, 84, 90, 216–218] and is a model with only two parameters. Model 2 was chosen to represent those models with both a mixing and a transitory compartment which all produced similar results. Model 8 was chosen from models which only included mixed compartments because it shows good linearity with little bias when predicting the end-expired alveolar volume and reasonable dead space prediction. It also has a low number of parameters. As some of the additional data had a dead space volume over $200ml$, the upper bound of the prior for Model 1 was extended to $300ml$. The results of this analysis are shown in Figure 5.5.

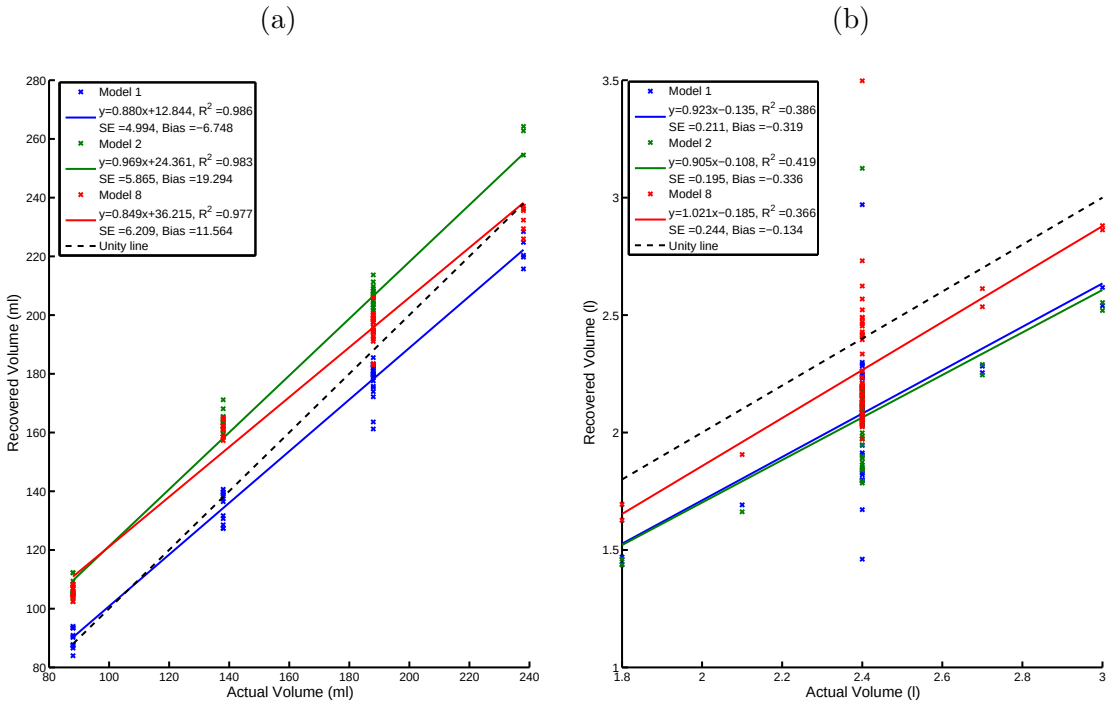


Figure 5.5: The volumes recovered for all datasets using three of the models described in the text. (a) The dead space volume and (b) the end-expired alveolar volume.

Unlike in the previous results a range of dead space volumes is included, so regression analysis is performed. For all three models, the R^2 values are above 0.97 indicating that the precision of the results is good. The standard error of the estimate confirms these results with a value of less than $7ml$ in all cases. In both cases Model 1 performs slightly better than the other models, but the difference is small especially when compared to Model 2. In terms of linearity Model 2 performs best with a slope of 0.97 while Model 1 has the least bias at the mid-point of about -$7ml$. Model 7 has a worse bias than Model 1 and has linear regression term further from unity than both of the other models.

The noticeable change in the end-expired alveolar results is the increase in variation of the $2.4l$ results when the additional data is included. Figure 5.6 shows box plots of the results for datasets with a end-expired alveolar volume of $2.4l$ for each dead space volume and forcing period. These show no obvious trend, which would explain this increase in variation. There are slight differences observed in the median and interquartile range with dead space volume, which may cause some of the additional variation. However the $188ml$ dead space used in the subset exhibits a higher interquartile range for all models. Also these differences are much less pronounced for Model 2, which also has an increase in variation.

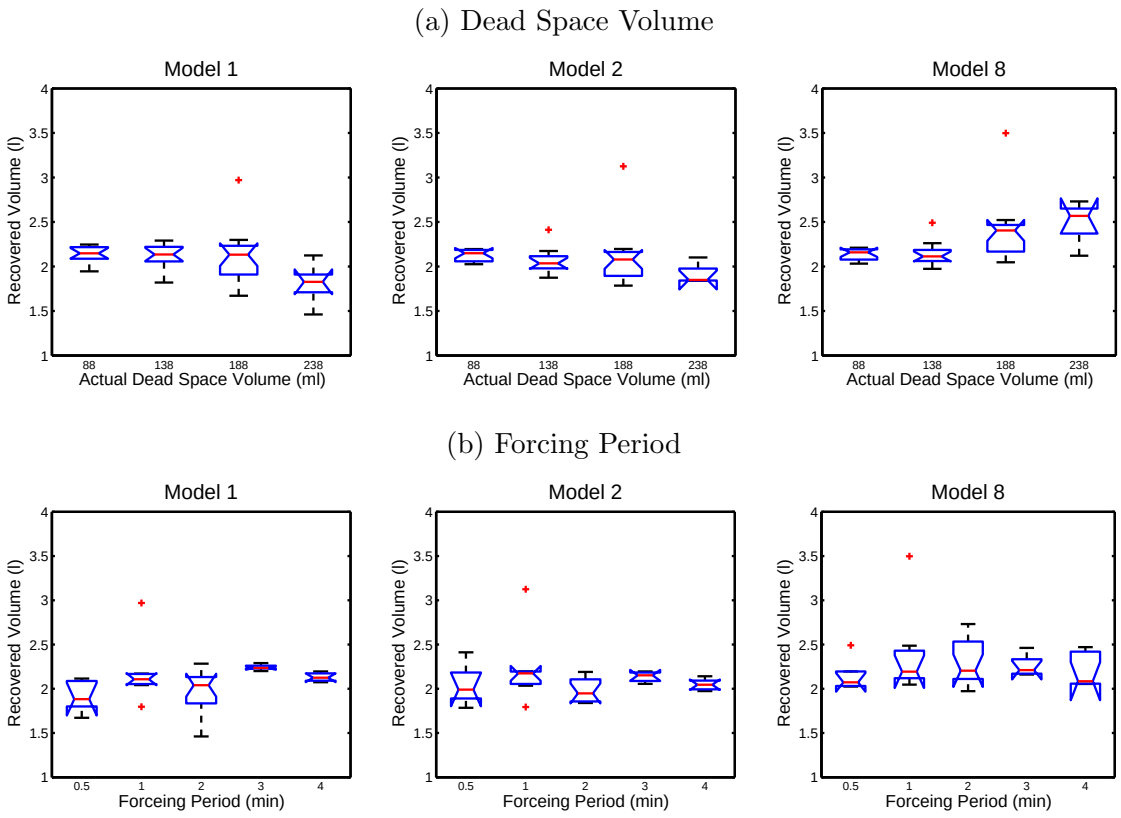


Figure 5.6: Box plots of end-expired alveolar volume recovered for the different dead space volumes and forcing periods. Only datasets with an actual volume of $2.4l$ are shown.

The forcing period results show no obvious trend in the median values, with many of the notches overlapping. For Models 1 and 2 there is noticeably less variation in the 3 and $4min$ results. However the subset has a forcing period of $2min$, which shows one of the highest interquartile ranges throughout. This suggests that the increase in variation is not related to these factors. When plotted against the time at which the data were collected, no obvious trend is observed either.

Retuning to the results presented in Figure 5.5b, Model 7 has the least bias at the mid-point and the most linear results. Models 1 and 2 linearity are still close to 1 and both have better R^2 values and standard errors of regression than Model 7.

It is not possible to include all nine modified Bland-Altman plots here. The trend in error is less than that observed in Figure 5.2 for the breath-by-breath methods, with the dead space volume predicted using Model 2 and the end-expired alveolar volume using Model 8 showing very little change, if any. This is as would be expected, given the linearity terms in the regression analysis.

The error observed in the predictions is detailed in Table 5.3. The bias observed is

	$V_D(ml)$	V'_A (l)	Subset of V'_A (l)
Model 1	-6.17±7.64	-0.32±0.21	-0.40±0.04
Model 2	19.49±6.10	-0.34±0.20	-0.42±0.05
Model 8	12.29±9.56	-0.13±0.25	-0.14±0.04

Table 5.3: Mean and standard deviation of the difference between the actual and predicted volumes. ABC SMC parameter estimation is used to make the predictions using three different models.

similar to the regression analysis. The standard deviations of the error provide a similar indication to the R^2 and standard error of the estimate terms. Models 1 and 2 generally perform slightly better than Model 8. There is a slight difference when comparing the dead space volume of Models 1 and 2, with Model 2's standard deviation being slightly lower, but the difference in the standard deviations is minimal.

5.2.3.2 Convergence

Thus far convergence of the ABC SMC parameter estimation algorithm was determined manually. In order to effectively use ABC SMC the convergence of the distributions has to be monitored automatically in order to determine a stopping point for the algorithm. ABC-SysBio [119] provides no automated method of detecting convergence. When proposing this method Toni *et al.* [198] did not provide details of how they monitored for convergence, but limited advice is given in the discussion that monitoring the interquartile range or goodness of fit for the parameter distribution of successive populations could be a potential method. It is also commented that the number of proposals required at each population is another "crucial signature". Here possible stopping criteria are investigated.

Figures 5.7 and 5.8 show a number of statistics for the populations for parameter estimation using Model 1: the mean, interquartile range, variation and number of proposals required. In addition a Kolmogorov-Smirnov (KS) test [130], a non-parametric goodness of fit test, was performed between the current and previous distributions, with the KS statistic and asymptotic p-value of this test shown. The examples shown here are not necessarily typical but do highlight some of the difficulties of choosing convergence criteria.

The results of the KS test do not provide an obvious solution to determining convergence as both the KS statistic and p-value vary considerably across the range of populations with no obvious difference between earlier and later populations. This may in part be because the distributions do not vary considerably between successive populations, which is not unexpected as part of the motivation of the SMC technique

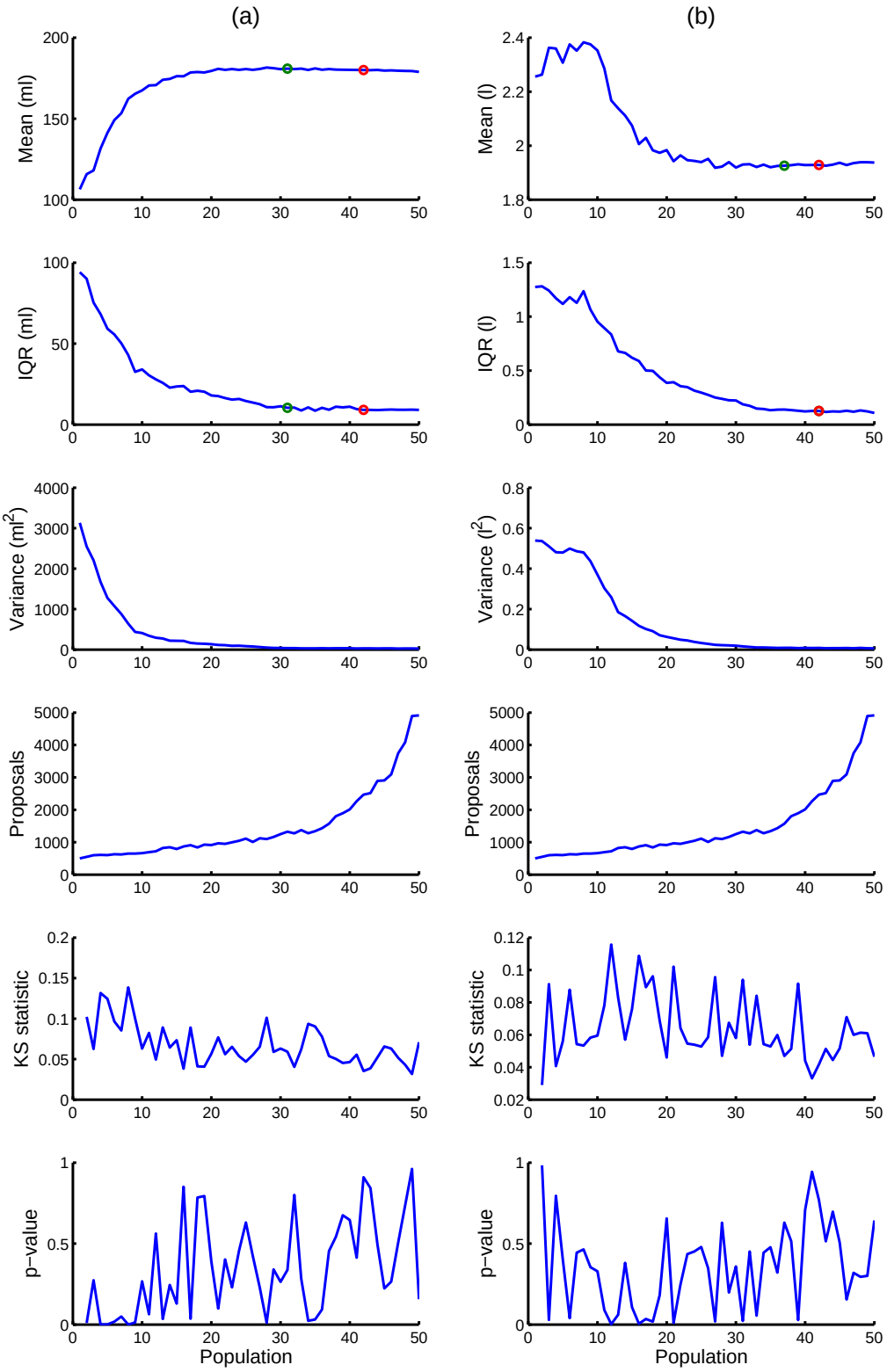


Figure 5.7: Example 1; the statistics for the parameter distributions of each population during a ABC SMC parameter estimation using Model 1. The data used came from a homogeneous bench lung. Column (a) and (b) show the dead space and end-expired alveolar volume, respectively. Green circles show when the tolerance given in the text is achieved for ten populations for the individual statistic. Red circles show when the tolerances given in the text are achieved for ten populations for the mean and interquartile range of both parameters.

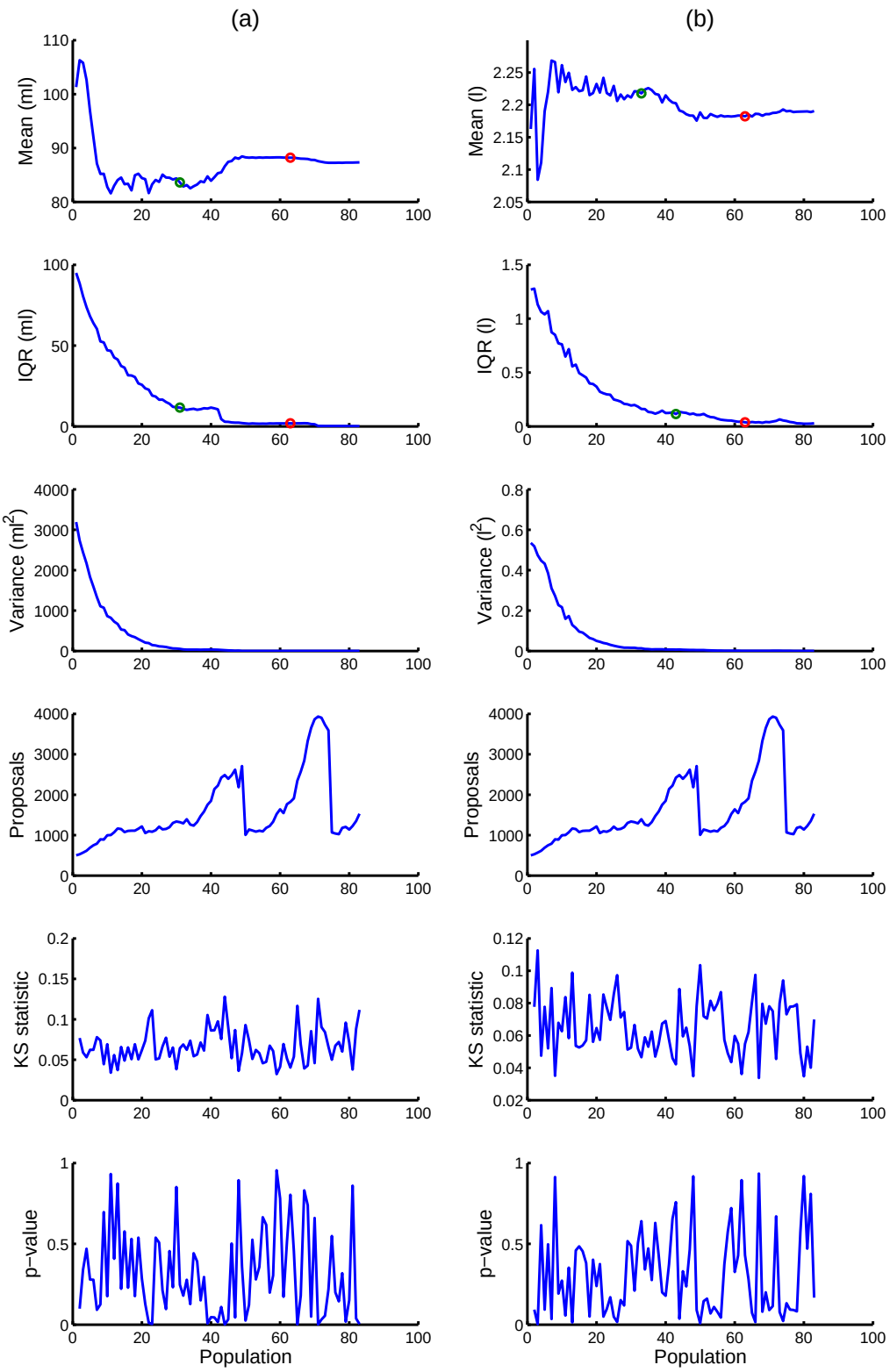


Figure 5.8: Example 2; the statistics for the parameter distributions of each population during a ABC SMC parameter estimation using Model 1. The data used came from a homogeneous bench lung. Column (a) and (b) show the dead space and end-expired alveolar volume, respectively. Green circles show when the tolerance given in the text is achieved for ten populations for the individual statistic. Red circles show when the tolerances given in the text are achieved for ten populations for the mean and interquartile range of both parameters.

is to limit the change from one population to the next. Increasing the number of populations between which the test is carried out was investigated with a gap of up to 20 populations analysed. This generally increases the KS statistic and so reduces the p-value. Populations at which the KS statistic is lower are then observed in later populations at which convergence appears to have been achieved, but this is not sustained over multiple populations and similar instances of low KS statistic values still occasionally occur at low populations where convergence has clearly not occurred.

It would seem plausible to use the interquartile range to determine convergence, stopping at a point where it has not changed by a given amount for a number of populations. As an example the first population above 30 at which the interquartile ranges of the 10 previous populations are within $3ml$ and $50ml$ is shown in the figures. Generally this would be successful as the interquartile range becomes constant as convergence occurs (see Figure 5.7), however Figure 5.8 does highlight that this can occur in other situations, such as just before a change in the mean and interquartile range of the distribution as seen in the dead space case.

Use of the number of proposals required at each populations is unlikely to provide a solution. In general the number of proposals increases and the rate of increase goes up with population number, as seen in Figure 5.7. However this is not universal, with various patterns being seen and a drop in the number of proposals required often being associated with points at which convergence may have occurred, as in Figure 5.8. Also the number of proposals required and the rate of increase associated with convergence both vary considerably between datasets. Using the variance instead of the interquartile range would seem a plausible alternative as it follows a similar pattern, but is generally smoother. This could has the advantage that convergence may be determined in a more timely manner, however it be unlikely to be any better at determining convergence at the correct point than the interquartile range.

An alternative would be to use the mean of the distribution to determine convergence. A possible criterion for doing this is shown in the Figures: the first population after 30 at which the mean of the 10 previous populations are within $1ml$ and $15ml$ for the dead space and end-expired alveolar volumes respectively. The stopping point determined by these criteria is shown in the example figures. Figure 5.8 highlights that using the convergence of the means alone is also not always effective, with the population at which both the interquartile range and mean criteria are met for the last 10 populations being a better measure of convergence. These criteria gave similar results as determining the convergence manually with the absolute error in the distribution

Model	All Dead Space Compartments		Alveolar Compartment	
	Lower Bound	Upper Bound	Lower Bound	Upper Bound
	ml	ml	l	l
1	10	300	1.0	3.5
2	10	300	1.0	3.5
3	10	300	1.0	3.5
4	10	300	1.0	3.5
5	5	200	1.0	3.5
6	5	200	1.0	3.5
7	10	300	1.0	3.5
8	10	300	1.0	3.5
9	5	200	1.0	3.5
10	5	200	1.0	3.5
11	1	100	1.0	3.5

Table 5.4: Bounds for the uniform prior used in the ABC SMC analysis of possible models of dead space mixing.

means being $0.9\pm1.0ml$ and $9.2\pm9.7ml$ (dead space and end-expired alveolar volume) and the absolute error for the standard deviation being $1.1\pm1.2ml$ and $34.8\pm23.9ml$.

Figure 5.8 does also highlight, however, that this method of determining convergence is not always effective, as the mean value of subsequent populations does ultimately change. No method of detecting convergence can fully guard against this happening, including determining convergence manually. Reducing the tolerance levels or increasing the number of populations at which they have to be met will reduce the likelihood of this happening, but there will be a computational price to pay for this. This is higher for ABC SMC than for a genetic algorithm or the Nelder-Mead simplex as the number of functional evaluations is higher and so not necessarily such a viable option. These convergence plots are typical of those observed for other models and when using data from other sources, but the best convergence criteria are likely to change between applications.

5.2.3.3 ABC SMC Model Selection

An alternative approach of selecting the most appropriate model was also investigated. A model selection using ABC SMC was carried out between the models described above. A uniform prior was used for each of the model parameters, the ranges of which are detailed in Table 5.4. The range of the prior limits for the dead space volumes reduces as the number of compartments making up the dead space is increased. The rationale behind this is to reduce the size of the parameter space in early populations to ensure that models with higher numbers of parameters do not die out at the early

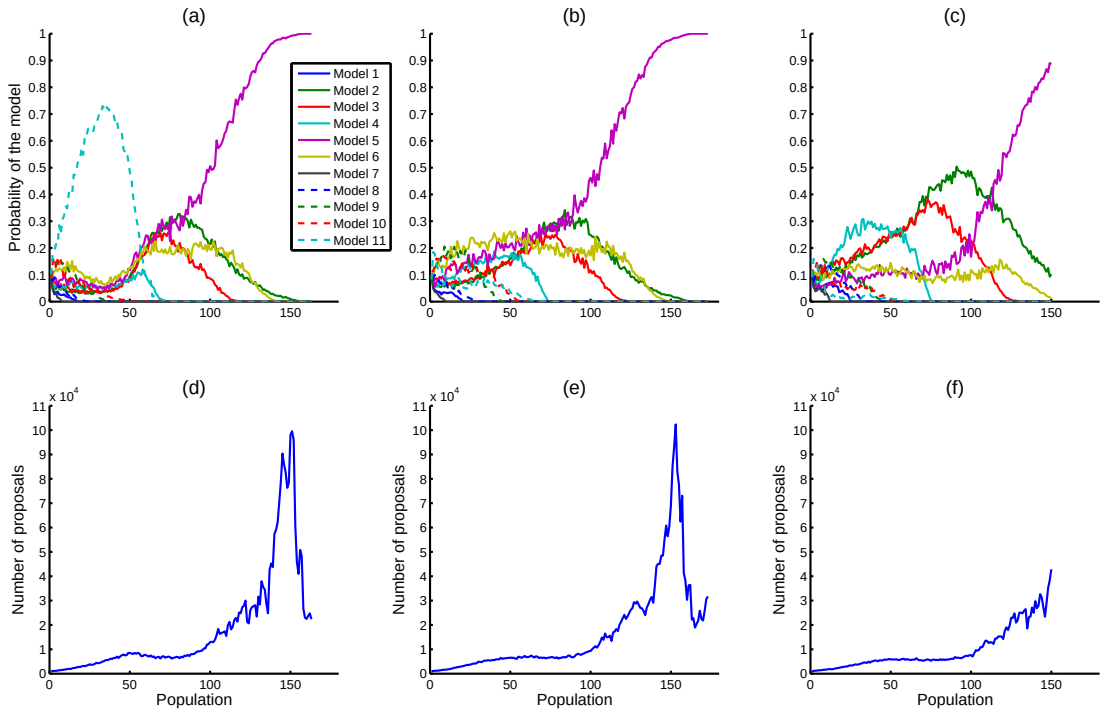


Figure 5.9: Plots showing convergence when comparing all 11 model using ABC SMC model selection for dataset I. The data came from a bench lung with a dead space of $188ml$ and an alveolar compartment of $2.4l$. The forcing period was $240s$. (a) Using priors given in Table 5.4 (112), (b) the dead space priors of Model 11 were adjusted to $5ml$ to $200ml$ (110) and (c) the dead space priors of all models were adjusted to $10ml$ to $300ml$ (59). (d) to (f) show the number of proposals required in each respective model selection. (The number in brackets is the processing time in days.)

stages. However this will affect the relative parameter space penalty associated with models with more parameters, potentially unfairly favouring these models.

Five datasets were randomly selected for analysis. Sub-figure (a) in Figures 5.9 to 5.13 shows the results. The probabilities of each model were monitored manually for convergence. In two cases (dataset I and II), after a considerable amount of processing time, convergence was ultimately achieved but only once one model remains. In the other cases convergence was not observed even though a very substantial amount of processing time had lapsed.

In all cases it can be seen that Model 11 is initially favoured but then loses favour to other models. In some cases the model then dies out. In order to investigate whether this is causing other models to be under represented, and to die out at a early population, the priors on the dead space compartments of Model 11 were extended to between 5 and 200ml and the model selection rerun. These results are shown in sub-figure (b).

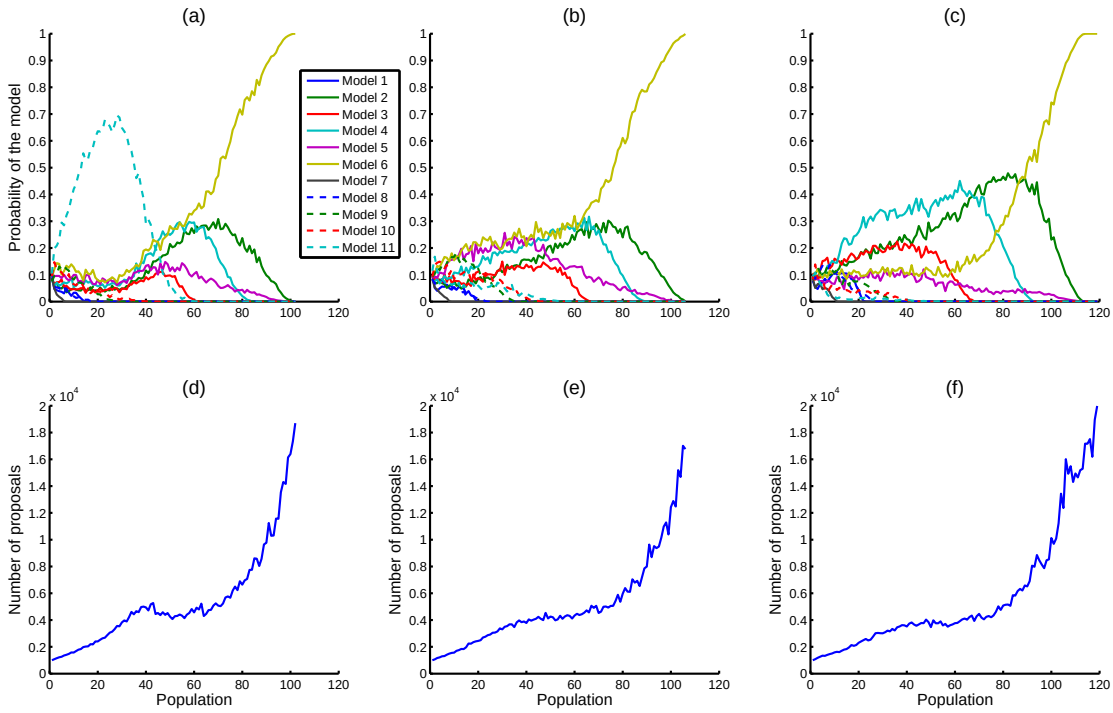


Figure 5.10: ABC SMC 11 model selection for dataset II. The data came from a bench lung with a dead space of 188ml and an alveolar compartment of 2.4l. The forcing period was 60s. (a) Using priors given in Table 5.4 (19), (b) the dead space priors of Model 11 were adjusted to 5ml to 200ml (18) and (c) the dead space priors of all models were adjusted to 10ml to 300ml (24). (d) to (f) show the number of proposals required. (Processing time in days.)

The final results of dataset I and II are not affected, however the relative probabilities on the path to that result are affected. The final results for sets III to V are unknown due to the restriction in computational power and time, but the probabilities at the completed populations change significantly. Model 11 is much less favoured throughout but increases in favour at the higher populations. Whether this model would ultimately become most probable if the process continued is however unknown.

An alternative approach was investigated in datasets I and II, chosen because they were quickest to converge rather than necessarily being the best candidates for further investigation. The same prior was used for all dead space compartments: a uniform prior from 10 to 300ml. The motivation for this was to have a parameter space which increases in size with the number of parameters, thus penalising more complex models to a fairer degree. These results are shown in sub-figure (c).

This affects the relative probabilities up to about population 100, with models 2, 3 and 4 receiving higher probabilities in both cases. Models 5 and 6 then become more probable in the case of dataset I and II, respectively, as previously seen. In the dataset

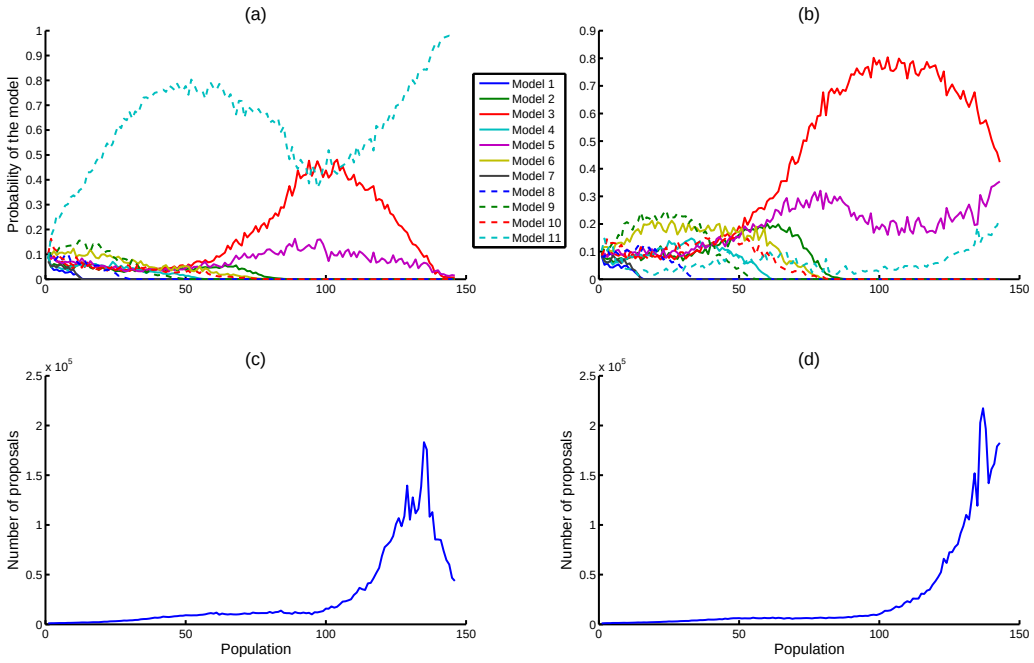


Figure 5.11: ABC SMC 11 model selection for dataset III. The data came from bench lung with a dead space of 138ml and an alveolar compartment of 2.4l. The forcing period was 30s. (a) Using priors given in Table 5.4 (75) and (b) the dead space priors of Model 11 were adjusted to 5ml to 200ml (81). (c) and (d) show the number of proposals required. (Processing time in days.)

II case, model 6 is ultimately chosen with 100% certainty. It is likely that the same will be true for Model 5 in the dataset I case.

5.2.3.4 Discussion

The parameter estimation results for all of the possible models of the dead space do not provide a clear indication of a best model for determining both the dead space and end-expired alveolar volume. Concentrating first on the dead space volume, the results for all models show little evidence for choosing any of the models with two or more compartments over another as all models show similar results. The results from single compartment dead space models are below those of models with more compartments with Model 1 closer to the estimated actual volume than Model 7, but the interquartile range for Model 7 is much lower.

When a trend is investigated over a range of dead space volumes for three of the models, selection of the best candidate model depends on the weight placed on the individual indicators of performance. From one viewpoint the linearity of the results could be considered more important than bias within the results, because of the issues of measuring the initial dead space volume. In this case the results from Model 2 would

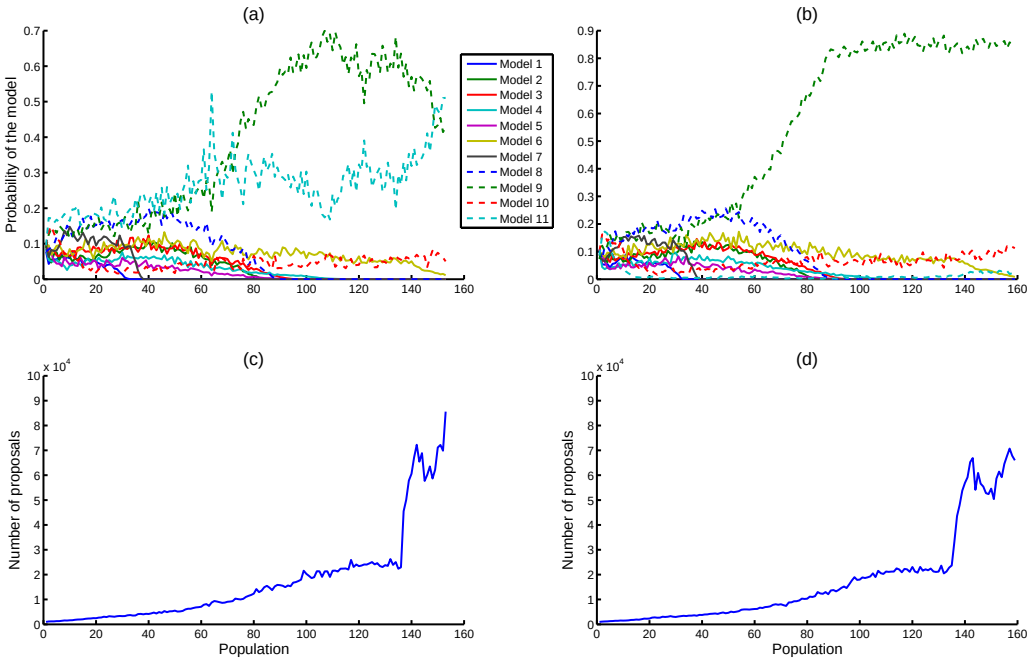


Figure 5.12: ABC SMC 11 model selection for dataset IV. The data came from a bench lung with a dead space of $88ml$ and an alveolar compartment of $2.4l$. The forcing period was $240s$. (a) Using priors given in Table 5.4 (125) and (b) the dead space priors of Model 11 were adjusted to $5ml$ to $200ml$ (123). (c) and (d) show the number of proposals required. (Processing time in days.)

be considered the best, as the slope is closest to unity and the R^2 and standard error of the estimate are better than for the other models. However if all factors are taken into account equally then Model 1 may be considered the best compromise as it has less bias. The modified Bland-Altman analysis concurs with this regression analysis.

When considering the end-expired alveolar volume recovered the best model again depends on how the individual indicators of performance are taken into account. Models of the dead space with at least one mixing and one transitory compartment (Models 2 to 6) provide very similar results with Model 1 also showing similar results although with very slightly less bias. The results from models with only mixed compartments (Models 7 to 11) show a different trend. Initially the bias is positive and the slope above 1 but these both reduce as more compartments are added with the two compartment model (Model 8) resulting in the best slope with least bias. The R^2 values and the standard error of the estimate are roughly the same throughout.

Similar results are seen when the three models are investigated on all datasets, although with a noticeable increase in the variability. Here the linearity of the slope and bias indicate Model 8 to be the best model, but this is not reflected in the R^2 and

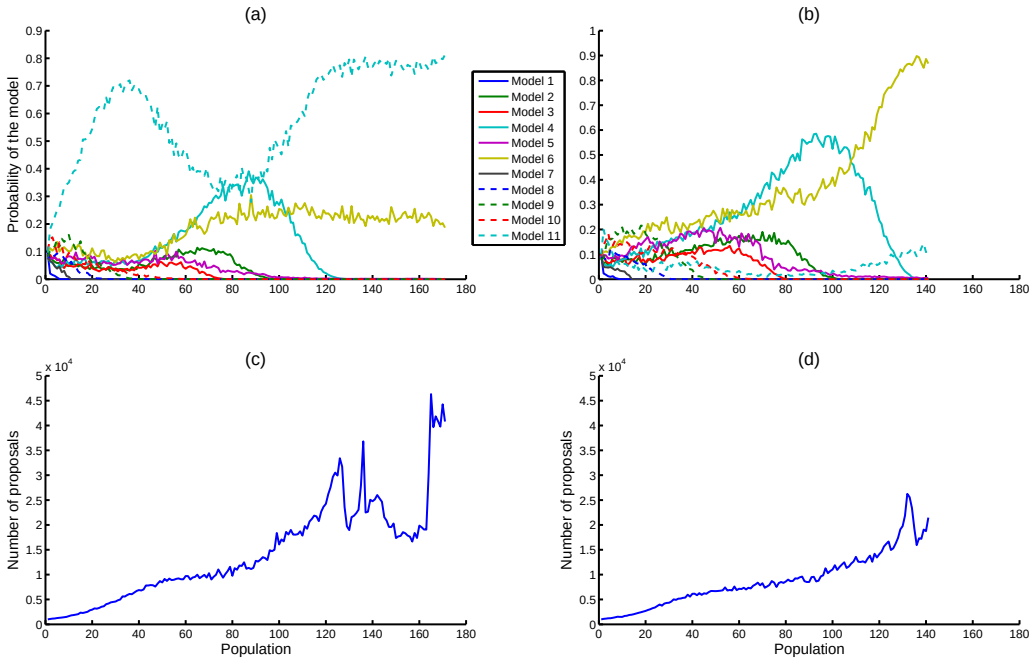


Figure 5.13: ABC SMC 11 model selection for dataset V. The data came from a bench lung with a dead space of $188ml$ and an alveolar compartment of $2.7l$. The forcing period was $120s$. (a) Using priors given in Table 5.4 (81) and (b) the dead space priors of Model 11 were adjusted to $5ml$ to $200ml$ (43). (c) and (d) show the number of proposals required. (Processing time in days.)

standard error of the estimate values which would favour Model 2. Similar results are seen for the Bland-Altman analysis with little difference observed between Models 1 and 2 and Model 8 showing less bias but a slightly higher standard deviation.

It is interesting to note that as compartments are added to the models which only have mixing compartments the trend of the results appears to head to that of a model with a transitory compartment. As the number of mixing compartments is increased the effective amount of mixing within the dead space decreases, assuming that one of the compartments is not significantly bigger than the others, because the amount of air held in each compartment is less and so less change in composition occurs as the air passes through each individual compartment. It may be the case that a dead space made up of a infinite number of infinitely small mixing compartments is equivalent to a transitory compartment, however this is not covered here. These results may indicate this to be true, with Model 11 (five mixing compartments) and Model 1 (a single transitory compartment) showing similar results. If true this may lead to an alternative method of solving models with transitory compartments allowing the use of an ODE solver instead of a DDE solver.

Why the variation in the results increases when all the data are analysed is obviously an interesting question because if it were possible to identify the origins of this variation it may help to eliminate it and so to make a much more reliable technique. The marked change in the R^2 and standard error of the estimate values between the results from the subset of the data and those for all of the data would indicate that this may not purely be due to natural variation in the noise on the signals. The fact that all of the datasets within the subset, except two (both of which had a volume of $2.4l$) were collected within a 2.5 hour period, while the data as a whole were collected over 3 days may also add weight to this argument.

However some type of trend would then be expected when the results are plotted against collection time. The results also indicate that it is not due to the forcing period used or the dead space volume. Whether this is due to how the sensors are calibrated for a particular period of testing, the ambient conditions on the day, variations in the initial manual equilibrium or some other factor is however impossible to determine from this dataset. Further investigations into the cause of the variation could prove valuable in improving the precision of the results.

ABC SMC model selection does not provide a clear answer to the respective probabilities of the data coming from the proposed 11 models. Unequivocal convergence is only achieved in two of the five cases analysed and in those cases a single model is chosen with full certainty, which would seem unlikely to be a good approximation of the posterior distribution, especially as a different model is also chosen in the two cases which do converge.

The parameter estimation results do not provide a convincing argument for choosing a particular model, which may reflect the difficulty in making a selection. This does not provide a satisfactory reason why ABC SMC model selection is also struggling. In fact it provides more reason to doubt the results as all models should have a similar probability if this is the case, rather than one being selected with 100% certainty.

Simulated investigations using a polynomial model were also carried out, but are not presented as they are not directly relevant to the thesis, showing that ABC SMC model selection method does not converge until only one model is remaining. That model is also not the model that would be expected to be the most probable, but the one with the largest number of parameters which did not die out in the earlier stages, i.e. it appears the penalty for having a larger number of parameters is not working effectively.

This work indicates that there may be a point earlier in the model selection at which a better approximation of the posterior can be found, however a method of determining this point was not clear. When proposing this method, Toni *et al.* [198] did not advocate using a large number of populations to push the tolerances towards successively lower values so eventually only one model is left, but neither did they propose any detailed stopping criteria. Instead they appeared to stop at an arbitrary population and tolerance.

As already discussed, limited advice is given in the discussion about monitoring the interquartile range or goodness of fit of successive distributions for convergence. No information is given about how best to do this and it is not clear whether this is in relation to parameter estimation, model selection or both. Also it is not clear from their results that convergence has been reached in the examples given, with the final two distributions being noticeably different. They also comment that the number of proposals required at each population is another crucial signature.

It is clear from the results presented here that the stopping point will affect the results. For example if population 40, 80 or 120 was chosen as a stopping point for the results presented in Figure 5.13, the probabilities obtained would be very different. It is unlikely that monitoring for a point of near convergence would adequately determine an earlier stopping point robustly. In a number of cases convergence criteria may be met, for example convergence may be deemed to have been achieved at some point after population 100 in Figure 5.12b, but this unlikely to be reliably replicated across datasets.

Monitoring the number of proposals may be another method of determining a stopping point and larger increases in the proposal required are observed in the results presented. For example, there is a large increase in the number of proposals in Figure 5.12 (c) and (d) at about population 135. However such clear changes are not seen in other cases, for example in Figure 5.10(e) where increases seem to happen more smoothly. Setting an absolute threshold is unlikely to be a satisfactory solution either, due to the differing magnitudes of the proposals required; case II only requires $2\text{x}10^4$ proposals to make a single selection, while the number of proposals required in other cases can be an order of magnitude larger. Ultimately choice of stopping criteria is going to result in very different results for the five examples shown.

A solution to this may be to use another model selection technique, such as the Akaike information criterion [18]. This could present the solution to this particular question, but this is not ideal for the model selection tasks required when determining

the homogeneity of the lung, as described in Sections 4.3.3.3 and 4.4.3. This is because these methods do not generally implement model selection and parameter estimation at the same time.

A possible method of determining if the ABC SMC parameter estimation method had converged was also proposed. These investigations suggest that methods using the KS test and the number of proposals are unlikely to be successful. Another method of testing for goodness of fit may prove to work better, but the nature of ABC SMC is likely generally to cause issues with this type of method. The SMC process purposely attempts to shift the parameter distributions from the prior to the posterior using a gradual process. Therefore it is unlikely that a noticeable change in goodness of fit will occur when convergence is achieved.

Convergence criteria that are based on the mean or the width of the distribution (in this case the interquartile range) were more successful. Alone they were prone to detecting convergence early which may again be due to the distributions not changing greatly between populations. However together they were more successful, with little difference from that determined using the manual method. It should be acknowledged that this comparison may be slightly biased as the convergence criteria were determined to mimic the manual decision and criteria which selected a stopping point after the last population of available results were rejected. Ultimately however the time taken for ABC SMC to reach a form of convergence, whether determined manually or automatically, is likely to mean it cannot be used in practice.

These results do highlight the same possible issue, as seen in the simulated data. The widths of the parameter distributions appear to continue towards zero as ABC SMC progresses. It would seem plausible that if sufficient computational time was available all particles of a distribution would ultimately end up with the same parameters, those of a minima. Thus the width of the distribution would be zero, which is clearly not a good representation of the posterior distribution where all aspects of the system are not perfectly modelled, including any noise. This would be at odds with the general (but not universal) description of ABC where the results of lower tolerances are thought to be better representations of the posterior distribution.

A better approximation to the posterior distribution may be found at an earlier population with a higher tolerance, much like in the model selection case, but the same issue arises: how can that population or tolerance be determined. Although it is unlikely to be possible to demonstrate this due to computational limitations, investigations using the polynomial model like those completed for model selection

could provide further insight into this issue. For the rest of this thesis, this is not necessarily important as the methods used will not use the width of the distribution rather just the mean as an estimated value for the parameters.

The results from both approaches investigated here do not provide a clear conclusion as to the best model of the dead space. ABC SMC model selection does not appear to be a robust method and ABC SMC parameter estimation for all of the models compared did not provide a clear answer to the best model. It is probably the case that the model of the dead space has little effect on the results. This will be investigated further below when also determining the offset between the signals.

5.2.4 Signal Offset Correction

As detailed previously (see Sections 2.6 and 4.6.2) the alignment of the two signals in the data has a significant effect on the recovered lung volumes. Methods of correcting the alignment of the signals by adding another parameter for the offset between the signals were also proposed (see Section 4.6.3), although there was some difficulty recovering this parameter in simulated data. Here further investigations are made in data from the bench lung.

Note that the convention used elsewhere in the thesis is also used here when quoting recovered offsets. The sign of the recovered offset is inverted, so that the offset quoted here is that found in the original signal. A positive offset indicates that changes in the flow signal are seen earlier in the original data than changes in the concentration signal.

5.2.4.1 Results

Figure 5.14 shows the linear regression results and box plots of the offset when ABC SMC parameter estimation is used. The main difference observed between these results and those when no offset correction is implemented (see Figures 5.4 and 5.5) is the bias of the results. For the dead space volume the bias is reduced in all cases. This results in less absolute bias in the case of Model 2, but an increase for the other two models. In the case of the end-expired alveolar volume the bias is increased in all cases with all methods now over- instead of under- estimating. This results in a reduction in the absolute bias for Models 1 and 2 but an increase for Model 8.

For the dead space volume the linearity and precision (by both measures) are not as favourable. However when Models 1 and 8 are used the linearity and precision both improve for the end-expired alveolar volumes. The results for Model 2 are more mixed,

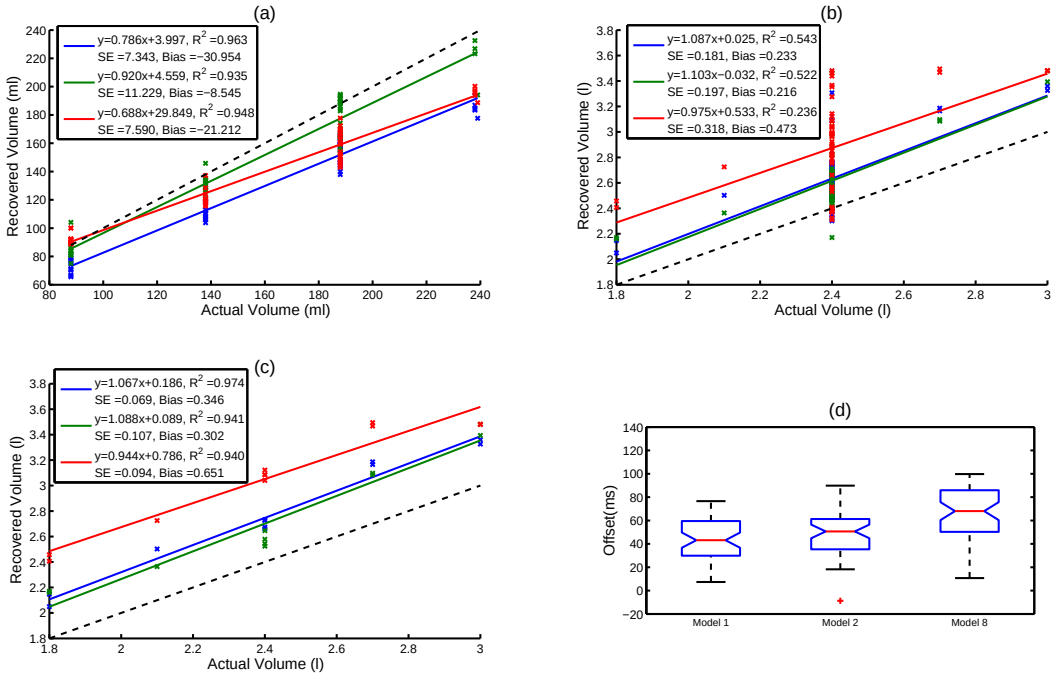


Figure 5.14: Parameters recovered from homogeneous bench lung data using ABC SMC when the offset between the signals is included as an additional parameter. (a) the dead space volume, (b) the end-expired alveolar volume, (c) the end-expired alveolar volume for the subset of the data and (d) the offset between the two signals. Model 1 (blue), Model 2 (green) and Model 8 (red).

	$V_D(ml)$	V'_A (l)	Subset of V'_A (l)
Model 1	-29.82±12.62	0.23±0.18	0.35±0.08
Model 2	-8.06±11.94	0.22±0.20	0.30±0.12
Model 8	-19.56±16.78	0.47±0.32	0.65±0.10

Table 5.5: Mean and standard deviation of the difference between the actual and predicted volumes. ABC SMC parameter estimation is used to make the predictions when the offset between the signals is included as a parameter.

with the linearity term and precision results improving in some cases but getting worse in others.

Table 5.5 shows the statistics for the error in the results. The bias here is similar to the regression analysis, improving in some cases, especially for Model 2, but getting worse in others. The standard deviation of the error is worse in all cases except for two of the end-expired alveolar results when all data are analysed. It reduces by $0.03l$ for Model 1 and is the same for Model 2. The modified Bland-Altman plots, although not shown here, show no recognisable trend in the end-expired alveolar results. Some trend is seen in the dead space results, with the error becoming more negative with increases in volume. This trend is substantially less apparent in the Model 2 results, with the results almost appearing completely random.

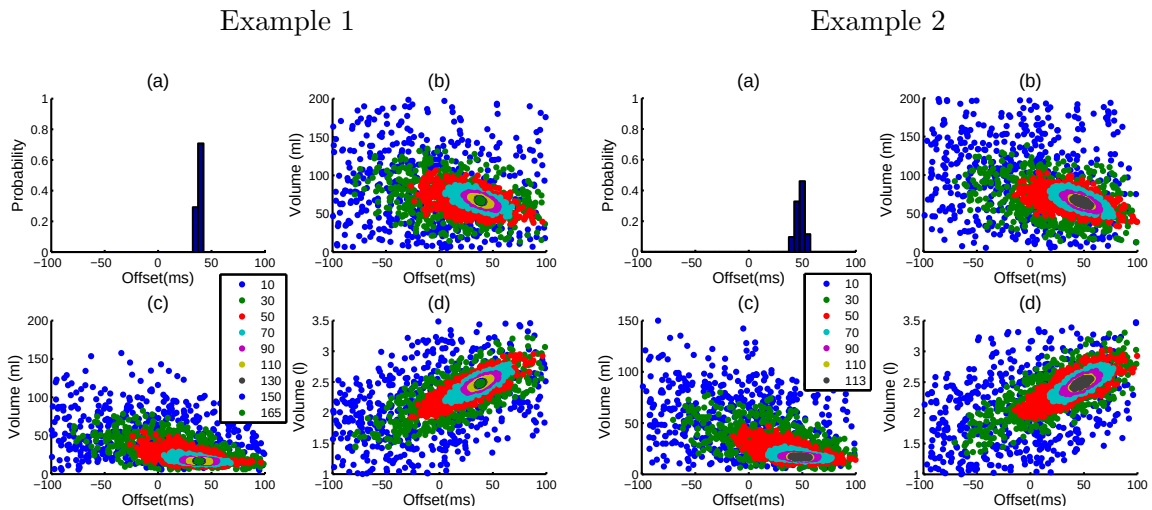


Figure 5.15: Plots showing the offset recovered from homogeneous bench lung data using Model 2. (a) Weighted histogram of the recovered offset. (b) to (d) are scatter plots showing the accepted particles at the populations detailed in the key. Offset is plotted against recovered (a) mixed, (b) transitory and (c) end-expired alveolar compartment volume. The data for both examples came from homogeneous lungs with $V_D = 88ml$ and $V'_A = 2.4l$. The forcing period was $3min$ and $2min$ for example 1 and 2, respectively. The highest population shown is the final output population before the algorithm was manually judged to have converged.

The box plot of the offset for Model 1 seems reasonably symmetrical with a median value of $43ms$. For Model 2 this is slightly higher at $49ms$ and there is possibly a slight skew towards higher values. The interquartile range is $5ms$ less but the range of the whiskers is the same. Model 8 results in a higher median of $68ms$ and has a larger spread of results with a higher interquartile range and wider whisker limits, which appear to be slightly curtailed by the upper bound of the prior.

Figure 5.15 shows typical particle scatter plots for two of the datasets. Although the particles are spread over a relatively large range of offset values at lower population numbers, the range of spread reduces as the population number increases. This can be compared to Figure 4.28 which shows similar plots for simulated data, where the spread of particles across offset values did not reduce with population number to the same degree, even when substantially more computational effort was employed.

Although the ability to identify the volume parameters and offset seems to be more achievable in this data than in simulated data, there is a computational cost to also determining the offset, especially in the case of Model 1. The results took on average 35.4±25.9, 14.4±7.4 and 7.5±3.5 days to converge (determined manually) in the case of Models 1, 2 and 8, compared to 2.9±0.4, 4.3±1.9 and 3.9±1.1 days when the offset was not included.

5.2.4.2 Discussion

Inferring the offset in addition to the other parameters provided more favourable results here than in simulated data. A reason for this is not immediately obvious but must be because the simulated data do not completely mimic what is happening in the real data. Two things that may play a part are the fact the simulated data had more instantaneous changes in concentration than found in the real data and no attempt was made to model the noise, including errors in the time point at which the sensor data was recorded. It would seem reasonable that these factors may affect the parameter space in such a way that makes the minimum more distinct.

The actual offset is not known and so the recovered value cannot be compared with it. However the direction of the recovered offset is as expected with the concentration signal having a longer delay or response time. This is usually the case, as detailed in Section 2.6. The recovered offset is likely in part to be adjusting for the response time of the concentration sensor in addition to any delay, as slightly advancing the arrival of the concentration sensor has previously been shown to partially compensate for the response [27, 77]. Similar analysis was carried out using a first-order exponential response time correction (Equation 2.27) instead of offset correction, but the noisy oscillations in the output increase with time constant, causing low time constants to be recovered in all cases. This results in very little change to the recovered volumes.

The motivation for including the offset within the parameter estimation was not to determine the offset itself but to improve the recovery of the other parameters. In general, however, the inclusion of an offset does not improve the results determined using Model 8. The results for the dead space volumes are worse by all measures. The linearity improves for the end-expired alveolar volumes but the bias is worse and the precision measures improve slightly for the regression analysis but the standard deviation of the error is worse. The use of offset correction with Models 1 and 2 seems slightly more favourable. For recovery of the end-expired alveolar volume using Model 1 all measures improve, except the standard deviation of the subset of results. However the dead space volume results are worse and there is a substantial computational price to pay.

The major improvement for Model 2 is the reduction in bias in both the dead space and end-expired alveolar volumes. However the dead space volume results do show slightly worse linearity and the R^2 value. The end-expired alveolar volume measures of precision do not change to a large degree with some showing better results while other are slightly worse. The linearity is now greater than 1 rather than being below

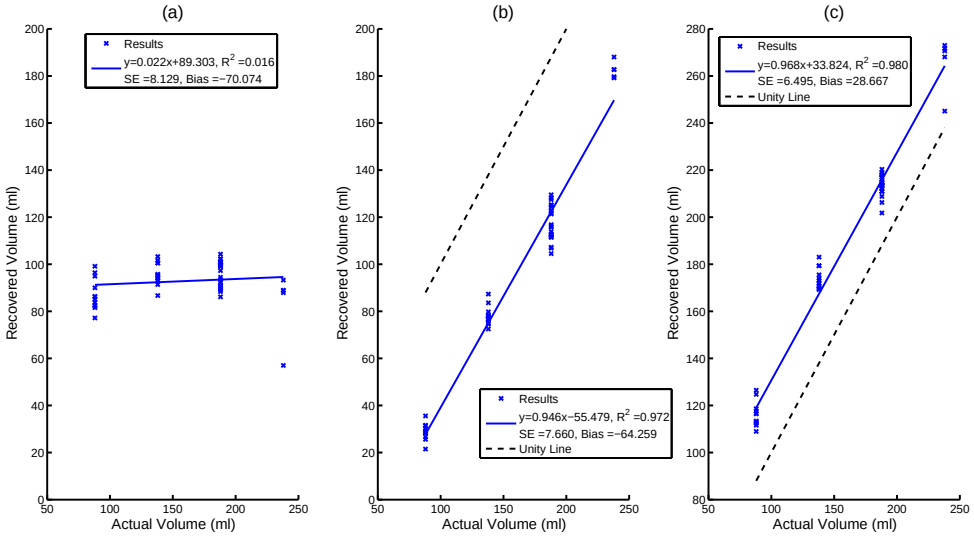


Figure 5.16: Dead space volume recovered independently of the alveolar volume using ABC SMC: (a) mixed, (b) transitory and (c) total dead space volume.

1, but by about the same amount. There is a computational cost to including offset correction but this is less than in the case of Model 1.

If offset correction could only be included in the parameter estimation when using ABC SMC, then it is unlikely to be suitable for real applications. Model 8 does not benefit to a degree where the use of offset correction would be desirable. Models 1 and 2 suffer from the computational requirements needed to determine the offset. A final device is likely to benefit from improvements to the efficiency of the code, the implementation of parallel processing and improvements in technology. These will bring the processing requirements down, however to bring it down from a mean of 35 and 14 days to a time in minutes, which is likely to be required, is probably unrealistic. This is especially the case as the model evaluations are already implemented in C, a much lower level language than Matlab and Python, which are used elsewhere in the thesis.

5.2.5 Recovering the Dead Space Volume Only

As shown in simulated results (see section 4.5), it may be possible to determine the volume of the dead space compartments independently of the alveolar compartment. The results presented are restricted to Model 2, as this was the model used in the simulated data and this will be investigated further below. Figure 5.16 shows the results when using this model.

The volumes recovered for the mixed compartment change little as the actual dead space volume of the bench lung is changed, with a volume of around $90ml$ recovered. The volume of the transitory dead space compartment appears to capture much of the

change in the volume of the dead space with a linear term close to that of the combined volume. When combined the total volume recovered shows good linearity but with a overestimate of $29ml$.

The variability of the results appears to be about the same for each of the four dead spaces within the data and for the two volumes recovered as well as the total volume. This is reflected in the mean interquartile range for the volume recovered at each of the four dead space volumes, which are: $10ml$ (range of 8 to $10ml$), $6ml$ (3 to $12ml$) and $7ml$ (5 to $10ml$) for the mixed, transitory and total volumes, respectively. Also the standard errors of regression are similar.

The total dead space volumes recovered are similar to those when an alveolar compartment is included within the recovery method (see Figure 5.5a). The linearity and R^2 of the regression line are almost identical in both cases. There is a slight difference in the bias at the midpoint, which is about $9ml$ when the alveolar compartment is not included.

Figure 5.17 shows the regression and Bland-Altman plots comparing the volumes recovered using this assumption and those recovered when an alveolar compartment is included within the model. Deming regression [61, 209] is used here as it allows for errors in both variables, unlike ordinary least squares regression which assumes no error in the x variable. The error in both measurements is assumed to be the same here, making it equivalent to orthogonal regression. Also standard Bland-Altman plots [37] are used rather than the modified version used earlier in the chapter.

Comparison of the total dead space volume shows both methods to be almost equivalent but with a bias of $9ml$. The transitory compartment shows similar results; there is a slight increase in the standard error of the estimate by about $0.5ml$ and the Bland-Altman limits of agreement by $1ml$, but the bias is reduced to approximately $6ml$. The overestimate however does not appear to be completely independent of its value, increasing slightly with volume. For the comparison of mixing compartment the standard error of the estimate increases again by about $0.5ml$ and the Bland-Altman limits of agreement improve back to that of the total dead space volume. The linearity coefficient decreases becoming further from unity but the bias is reduced to about $4ml$.

When compared with results recovered using a model that includes an alveolar compartment and offset correction, the volumes do not compare as well, as shown in Figure 5.18. The standard error of the estimate and width of the Bland-Altman limits of agreement increase substantially, especially in the case of the transitory and total dead space volumes. The linearity term is further from unity, especially in the case of

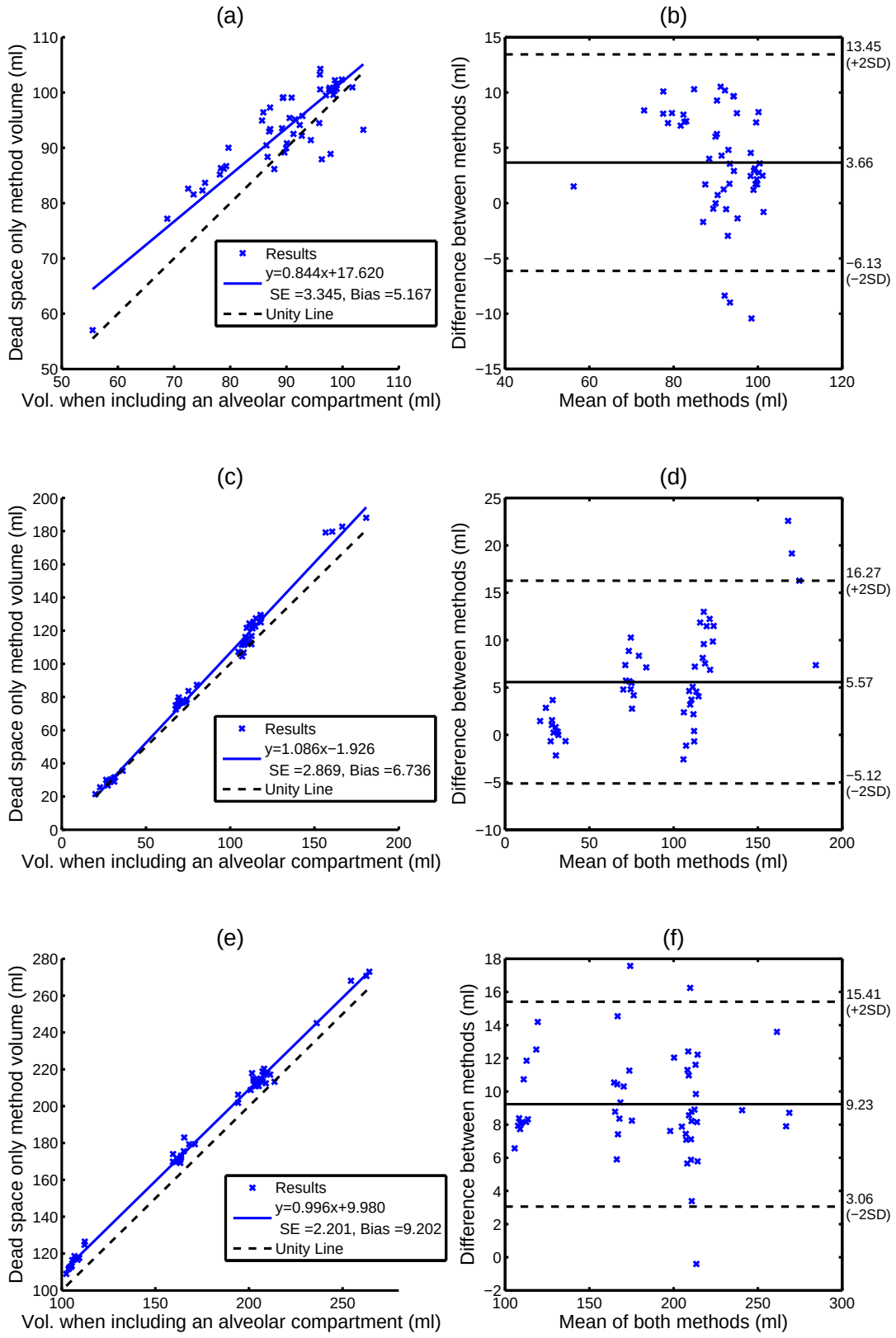


Figure 5.17: Comparison of the individual dead space compartment volumes recovered independently of the alveolar volume with the model that includes alveolar compartment. Both recovered using ABC SMC. Regression and Bland-Altman plots for: (a) & (b) the mixing volume, (c) & (d) the transitory volume, (e) & (f) the total dead space volume.

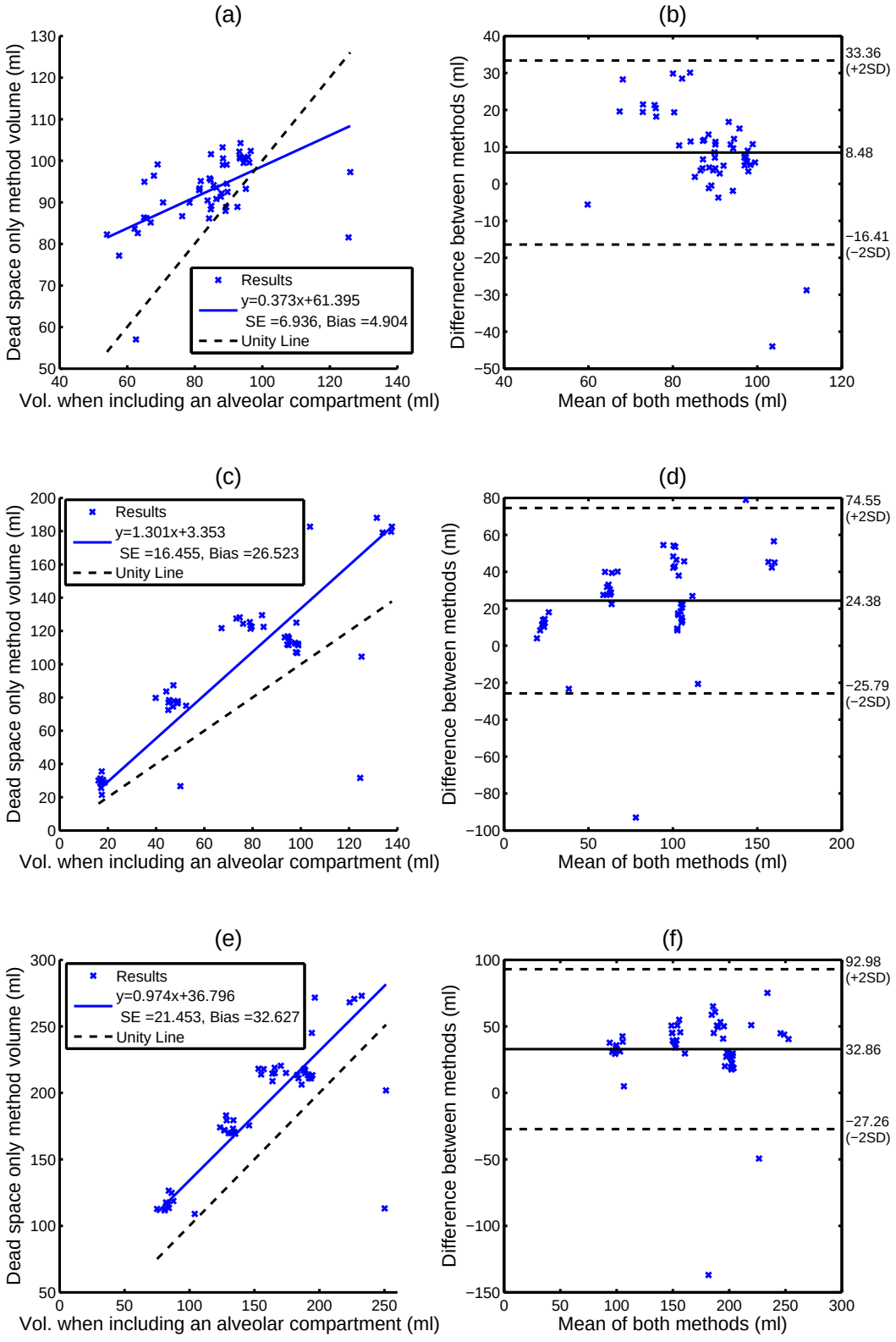


Figure 5.18: Comparison of the individual dead space compartment volumes recovered independently of the alveolar volume with the model that includes alveolar compartment and correction of the signal offset. Both are recovered using ABC SMC. Regression and Bland-Altman plots for: (a) & (b) the mixing volume, (c) & (d) the transitory volume, (e) & (f) the total dead space volume.

the mixed compartment, and the bias increases for the transitory and total dead space volumes.

An attempt was then made to recover the offset along with the two dead space compartment volumes independently of the alveolar configuration. This was not successful. The results recovered showed minimal resemblance to those recovered when an alveolar compartment was included.

5.2.5.1 Discussion

These results show that the dead space volume can be inferred independently of the alveolar volume. There is a slight increase in the overestimate but otherwise the results are remarkably similar to those including the alveolar compartment. However the interest in determining the characteristics of the dead space independently of the alveoli make up is not driven by the need to determine the dead space volume itself but to use this information in order to determine other information about the lung more effectively.

When no offset is included the results are comparable to those when an alveolar compartment is included, but with a slight bias. Therefore it is likely that it would be possible to use these values either to fix the volume of these compartments when determining other parameters or to provide a more informative prior distribution when using ABC SMC. If the simulated results are representative it is likely also to be possible for inhomogeneous alveoli.

When an offset is included in the recovery the results are not so comparable, thus it is unlikely to be possible to use this method when also inferring an offset. Including the offset as an additional parameter obtained worse results. Therefore this method of recovering the dead space volumes independently of the alveolar compartment volume is only likely to be of value in cases where an attempt to recover the offset is not also made. However it may still be possible if the offset could be determined in advance.

Additionally these results show that the mixed compartment volume recovered does not appear to change with an increase in the dead space volume, but that the transitory compartment does appear to increase in volume. This seems to be in keeping with the change that is happening within the bench lung. When the volume is increased an additional section of smooth pipe is added which is unlikely to increase mixing greatly, especially when compared to the mixing that is likely to occur when the pipe diameter changes abruptly, as can be found elsewhere in the bench lung or when passing through the concentration sensor and other equipment.

5.3 Multi-compartment Lung Data

Within this section the methods developed in the previous chapter for determining the dead space volume and the inhomogeneity of the lung are applied to data from the multi-compartment water displacement bench lung. Although the results were promising for simulated data, it will never mimic the complete process or all of the errors that may occur. Thus data from the multi-compartment bench lung allow the methods to be investigated in conditions more like that found in real lungs, but where the nature of the underlying inhomogeneity is known.

5.3.1 Data

Data used within this section come from a multi-compartment bench lung, as detailed in section 3.2.3. Data were collected from four different configurations of two compartments, which are the same configurations simulated earlier and detailed in Table 4.4. In the first three configurations (A, B and C) the compliance of both compartments was maintained by including the same $350mm$ diameter solid rods in both the $600mm$ displacement cylinders. In the final configuration (D) the rod was removed from one of the displacement cylinders, thus changing the compliance and so the tidal volume. This will also affect the emptying characteristics of this compartment. The approximate volume of the dead space was $140ml$ for all configurations.

For each configuration, data were collected for two different forcing wave forms: a sinewave and a modulated sinewave. The sinewave data were collected for forcing periods of 0.5, 1, 2, 3, 4 and $5min$. The modulated sinewave had a carrier and modulated time period of 2.626 and $15min$ as previously simulated. The amplitude of forcing signal was 3% with a mean of 5%. During an initial 100 seconds a manual process was used to bring the fractional concentration within the lung up to the mean value. Sinewave data were then collected for $12mins$ for forcing periods of 3 and $4min$ and $10min$ for all other forcing periods. The modulated data were collected for $15min$.

Each data collection was repeated once, but with the following exceptions:

- the sinewave data collection for configuration D were not repeated at any forcing period;
- the sinewave data collection for configuration C was not repeated for $3min$ but a third dataset was collected for 0.5 min.

Compartment	Lower Bound	Upper Bound
V_M	$5ml$	$200ml$
V_R	$5ml$	$250ml$
$V'_{A,1}$	$200ml$	$1200ml$
$V'_{A,2}$	$900ml$	$1900ml$

Table 5.6: Bounds for the uniform prior used in the ABC SMC analysis of the multi-compartment lung.

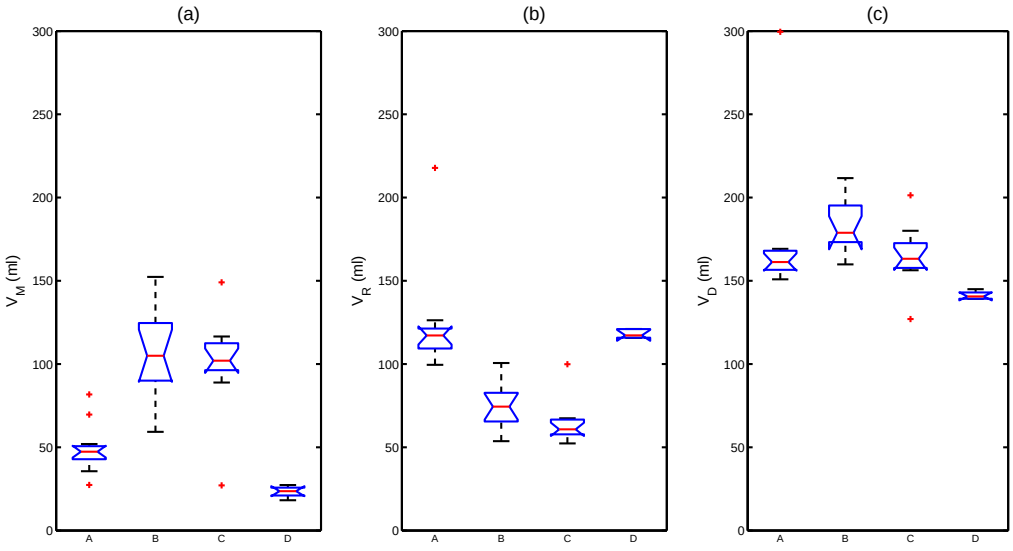


Figure 5.19: (a) Mixed and (b) transitory dead space volumes recovered from multi-compartment bench lung data using ABC SMC; (c) is the sum of both volumes, the total dead space. The alveolar concentration was assumed to equal the end-expired concentration.

5.3.2 Recovering the Dead Space Volume

5.3.2.1 ABC SMC Model Selection

First an attempt to determine the dead space volume independently of the alveolar make up was made using ABC SMC parameter estimation. Uniform priors were used as detailed in Table 5.6. The results for the sinewave data are shown in Figure 5.19.

The results for configuration D show the least variability for both compartments. This feeds through into the combined value for the total dead space volume, which has a median value of $141ml$. This is very close to the estimate for the actual volume. The mixed compartment volume, with a median of $24ml$, has a low volume compared to that observed in the homogeneous bench lung.

Configuration A results have a larger recovered mixed compartment, with a median value of $47ml$, but more variation is observed. The transitory compartment volumes recovered have a similar median value, but again show more variation. This results in a

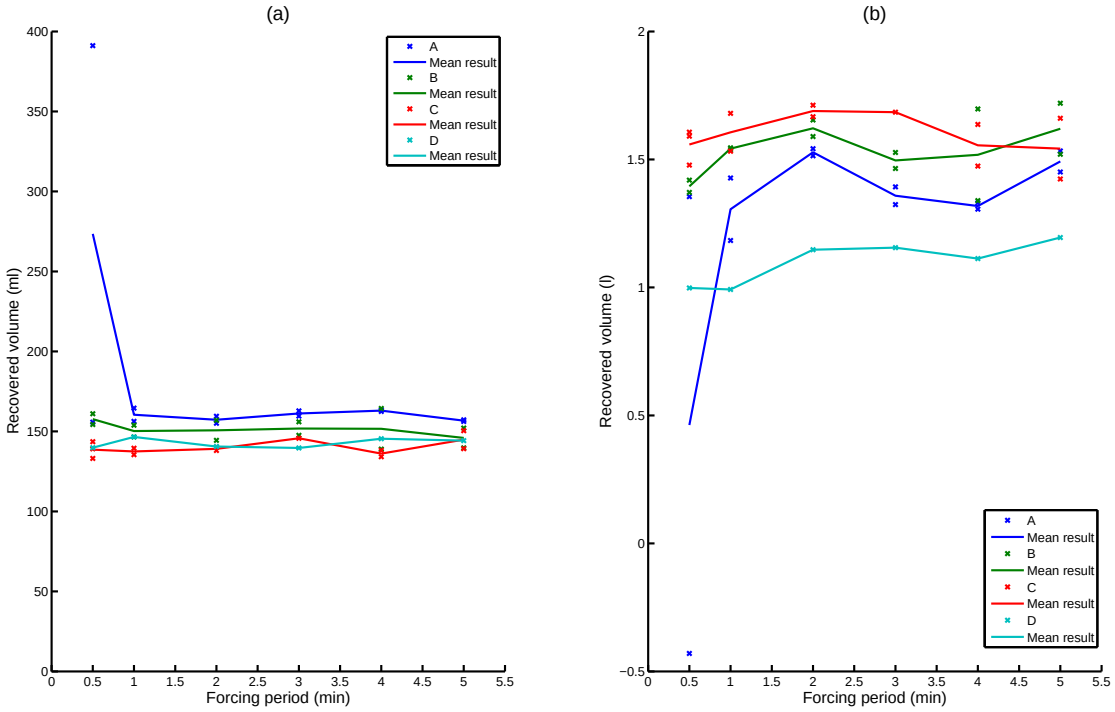


Figure 5.20: (a) Dead space and (b) end-expired alveolar volumes recovered from multi-compartment bench lung data using the Bohr breath-by breath method.

larger median value for the total dead space volume, which also shows more variation. However the interquartile range is still around $10ml$, which is similar to that of the homogeneous bench lung data.

The results recovered for configuration B and C have larger mixed compartment volumes and lower transitory volumes. These results are significantly different from those observed for configuration A and D. This results in a combined total volume which has a higher median. However they are not significantly higher than the median of configuration A. The variability of the configuration C results is approximately the same as configuration A, but the variation of the configuration B results is larger for both the recovered volumes and the combined volume. This results in an interquartile range for the total dead space of around double that seen in the homogeneous bench lung at $22ml$.

5.3.3 Determining the Homogeneity of the Alveoli

5.3.3.1 Breath-by-Breath Method

Here the data collected for sinewaves with different forcing periods are analysed using the breath-by-breath method. If it is possible to determine information about the homogeneity of the lung, a variation in the end-expired alveolar volume recovered at different forcing periods would be expected, as this would result in a change to the

homogeneity index. Figure 5.20 shows these results.

The dead space volume recovered is relatively consistent across the forcing periods between approximately 135 and 165ml, which is around the actual value of approximately 140ml. There is possibly a slight trend of the recovered dead space volume decreasing with increase in inhomogeneity, but this is only of order 20ml. There is one anomalous result of 390ml.

The end-expired alveolar volumes recovered show a wide range of results of approximately 0.5l for configuration A, B and C. In these cases, there is also no evidence of a reduction in the volume recovered at lower forcing periods, with any changes appearing to be within the natural variation observed. There is again an anomalous result, -0.43l, which is a prediction from the same set of data that produced the anomalous result in the dead space volume. Dataset D estimated a volume below that of all other configurations with all results approximately 0.5l below that observed for the other configurations. There is a slight reduction in the volume recovered at forcing periods of 0.5 and 1min, but this is within the natural variation observed elsewhere. All results are below the actual combined volume of the two compartments.

Similar results were observed when the Fowler method was used but with a slightly smaller range for the results in the dead space volumes and no noticeable change of the dead space volume recovered with an increase in inhomogeneity. There was a slightly larger range and variation between the repeated results for the end-expired alveolar volume but the separation between the configuration D and other results remained. An anomalous result also occurred for the same set of data that produced an anomalous result when using the Bohr method.

5.3.3.2 Inter-Breath Variation

The method of using the inter-breath variation as a proxy for inhomogeneity, as proposed in Section 4.3.2 of the previous chapter, was applied to these data. The method relies on knowledge of both the dead space compartment volumes; however the results presented in section 5.3.2 show results that vary with the configuration despite the dead space remaining unchanged between configurations. Therefore the median results for all data were used: 85 and 98ml for the mixed and transitory compartments, respectively.

These results are shown in Figure 5.21 for the two modulated sinewave datasets. The standard deviation of the results as seen in the sub-figures (a) and (d) to (h) are as expected given the results seen in simulated data with noise (see Table 4.6). A lower standard deviation is seen for configurations A and B compared to C and D.

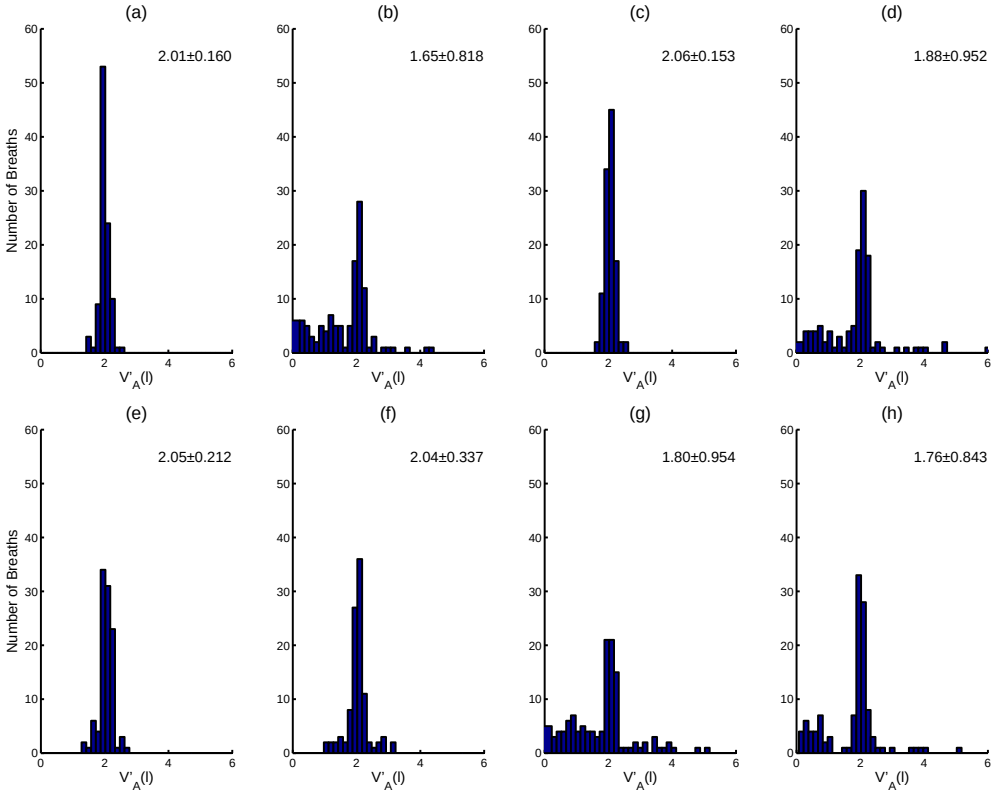


Figure 5.21: Histograms showing the individual breath predictions for data from the multi-compartment bench lung, with configurations: (a & e) A, (b & f) B, (c & g) C and (d & h) D. The data used had a modulated sinewave forcing waveform.

This is reflected in much narrower and more uniform distributions for configurations A and B than for C and D. Both the difference between the standard deviations and the standard deviations themselves are however much higher than in the simulated case. The difference was approximately 0.1 in the simulated case, compared to 0.7 here. The standard deviations were all less than 0.2 in the simulated case, whereas most of the values here are greater than 0.2.

The means follow a similar pattern to simulated data for configuration A to C. They are approximately the same for A and B but reduce for C. However the mean of configuration D is approximately the same as configuration C, unlike in simulated data where configuration D had a higher mean than any other configuration. The means are lower here for all configurations than in simulated data. The results in sub-figures (b) and (c) do not fit the pattern either for the standard deviation or the mean, with the results for configuration B similar to those seen for C and D and those for configuration C resembling those seen for A and B.

The results were also analysed using the dead space volumes acquired for the individual datasets. This resulted in a very different set of distributions with a wider

	(a)	(b)	(c)	(d)	(e)	(f)	(g)
(b)	**0.0000**						
(c)	**0.0018**	**0.0000**					
(d)	**0.0018**	**0.0344**	**0.0001**				
(e)	**0.0146**	**0.0000**	0.8662	**0.0030**			
(f)	**0.0181**	**0.0000**	0.3677	**0.0179**	0.4761		
(g)	**0.0000**	0.7510	**0.0000**	0.2311	**0.0000**	**0.0001**	
(h)	**0.0009**	0.1176	**0.0001**	0.1922	**0.0012**	**0.0022**	0.3011

Table 5.7: p-values when the distributions shown in Figure 5.21 are compared using the KS test.

range of means and standard deviations, which did not follow the same trends. The configuration B distributions had large standard deviations of 35.8 and $4.34l$, which were larger than the standard deviations observed for any other configuration.

The results for the non-modulated sinewave data were also analysed in two ways. The results from the same forcing period were compared and the results were combined, with a weighting to ensure that results from each of the forcing periods were equally represented. However there was no noticeable difference between the standard deviation of the configuration A and B results compared to C and D results.

Table 5.7 shows the results of KS tests carried out between the distributions shown in Figure 5.21. The null hypothesis is that the distributions are the same. The ideal results would be that the null hypothesis is not rejected when the distributions came from data with the same configuration, but rejected when the configurations were different. However given the simulated results it would be reasonable to only expect rejections when configuration A or B are compared to C or D.

Assuming a 5% significance threshold, when comparing the two datasets from the same configuration the only case of the null hypothesis not being rejected is for configuration D. This is probably not surprising for configurations B and C given that (b) and (c) do not appear to follow the pattern. When comparing A or B to C or D there are 5 times when the null hypothesis is not rejected; when (b) is compared to (d), (g) and (h) and when (c) is compared to (e) and (f). This is again probably not surprising as all these comparisons include (b) and (c).

It is not advisable to change the threshold for significance retrospectively. However changing the threshold to 1% would result in the null hypothesis not being rejected when the two configuration A distributions are compared, but it would also result in it no longer being rejected when (d) and (f) are compared.

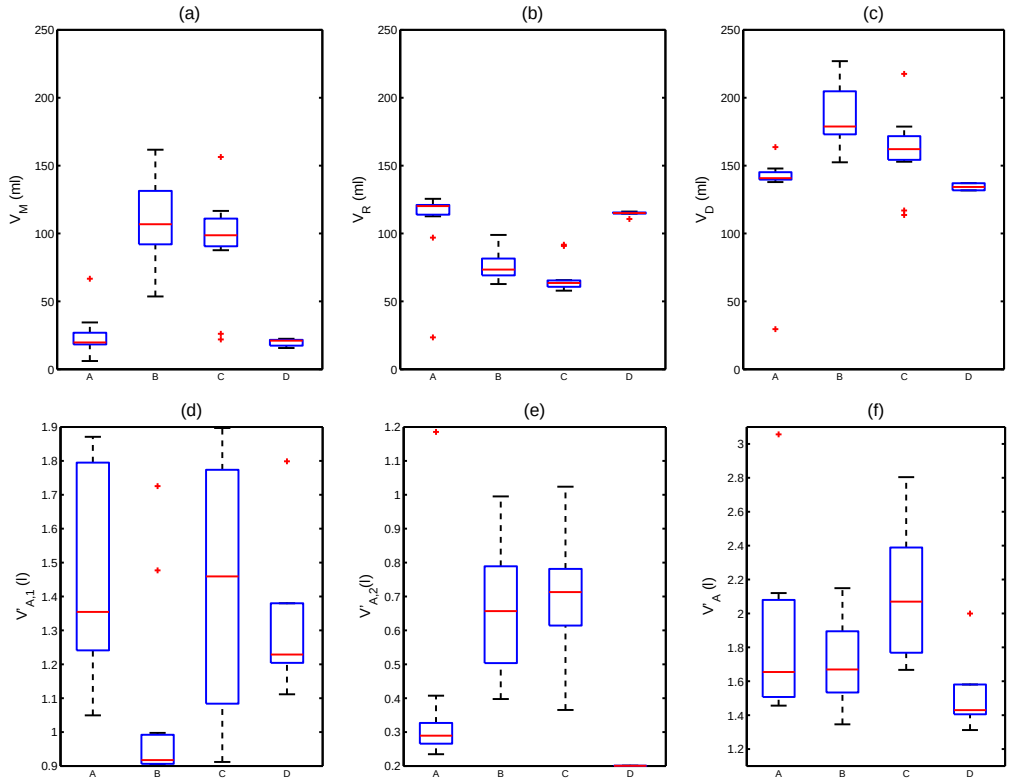


Figure 5.22: (a) Mixed, (b) transitory, (d) compartment 1 end-expired alveolar and (e) compartment 2 end-expired alveolar volumes recovered from multi-compartment bench lung data using ABC SMC. (c) and (f) are the total dead space and end-expired alveolar volumes.

5.3.3.3 Recovering Volume-ventilation Ratio of Multiple Compartments

Here ABC SMC parameter estimation is used in an attempt to recover the volumes of both compartments. The tidal volume is assumed to be equally distributed between the two compartments. The priors used are detailed in Table 5.6.

Figure 5.22 shows the results when the two dead space and two end-expired alveolar compartment volumes are recovered for the sinewave data. The two dead space volumes recovered for configuration D are most consistent, but with a low mixed compartment volume of approximately $20ml$. This results in a fairly consistent total dead space volume of approximately $140ml$, which is close to the estimate of the actual volume. Similar results are seen for configuration A but all three sets of results show more variation. A number of outliers are observed, with the $30ml$ total dead space outlier coming from the same dataset that caused the anomalous results seen for the Bohr method.

The results for configurations B and C have significantly higher mixed but lower transitory volumes. This results in combined volumes which are larger, with median results of 179 and $162ml$, respectively. In general results from these two configura-

tions are more variable, especially for configuration B and in both cases for the mixed compartment volumes. Overall these results have a similar pattern to the results seen when the dead space is determined alone (see Figure 5.19).

The end-expired alveolar volumes recovered for both compartments show a large amount of variability. In three cases a lower amount of variations is observed, but the median value is well below the actual volume of these compartments and in one of these cases ($V'_{A,2}$ for configuration D) the limit of the prior has been reached in all cases. In general the recovered combined volume is also very variable. The variability of the results for configuration D is less but this may be due to the lack of variability in the volume of $V'_{A,1}$. The volume recovered is less than the actual combined volume, except in the case of configuration C where the median is $0.10l$ below the actual volume, but the variability of these results is high.

Similar results were seen for the modulated sinewave data (not shown here). The results from the two datasets produced similar volumes in the case of configuration A and D for both the mixed and transitory compartments, which results in total dead space volumes within $20ml$ of the estimated volume. For configurations B and C, the value and difference between two volumes recovered for the mixed compartment is larger, which results in an overestimate of the total dead space volume of 33 to 83 ml.

The difference between the end-expired volumes for each of the two datasets is not as great as may have been expected from the sinewave data results with a maximum difference of $68ml$. However the volumes recovered are not close to the actual volumes of the compartments. The combined volume is also below the actual volume, with a maximum volume of $1.56l$ recovered.

Assuming that both the dead space compartments were equal to the median values obtained in section 5.3.2 and then only estimating the alveolar volumes produced results with lower variation. However these results were no better at determining the actual volumes of the compartments. All configuration resulted in $V'_{A,1}$ volumes of just over $0.9l$. A variety of results were obtained for $V'_{A,2}$ with the value for configuration C generally being greater than the results for configuration A and B. The combined total volume was consistently below $1.9l$ and generally around $1.5l$.

5.3.4 Discussion

The results from the multi-compartment bench lung are generally not as positive as those from the homogeneous bench lung. The attempt to recover the dead space volumes independently of the alveolar for configuration B were not as successful as for the homogeneous case with an increase in variation observed. It is unlikely to be the

characteristics of the flow or concentration sensors that cause this as the same sensors were used. The range of the results observed for the other configurations is about the same as that observed in the homogeneous bench lung and even substantially less in the case of configuration D. Thus the change in the bench lung does not seem to be the cause of more variation.

The surprising observation is the distinction between the results for the mixed and transitory volumes in the case of configuration A when compared to B and C, given that the simulated results for independently determining dead space showed virtually no change in the volumes recovered. The simulations were for a log-normal distribution rather than the multi-compartment lung used here, however this is unlikely to be the source of the differing results. It is plausible to expect configuration D to act differently, as the compliance of one of the compartments is adjusted so that the emptying characteristics of the compartment will change, a factor not included within the simulated data. However this is not the case for configuration A.

The breath-by-breath results confirm the findings in simulated data that the noise will make any changes in homogeneity index on the scale present within this data impossible to discern. However, with the exception of what appears to be a single anomalous result, the dead space volumes recovered do not show as much variation or as much difference between the four configurations, especially when the Fowler method is used. The one noticeable difference from the simulated data is that the end-expired alveolar volume results for configuration D were consistently above the values recovered for the other configurations in the simulated results, however here they are consistently below those values. This may be due to the change in compliance of one of the compartments which was not modelled in the simulated data.

Directly determining the volumes of the two compartments is not effective either. This is at odds with the results from the simulated data which suggested that it should at least be possible to determine between configurations A and C. The wide range of the recovered volumes along with the fact that the median result is far from the actual value would suggest that this method is unlikely to be successful without a very significant improvement in data quality.

Although results from the inter-breath variation method still have a number of issues, they do give the most promising results. The width of the distribution does appear to act as expected in most of the cases for the modulated sinewave data. However in two cases (25% of the cases) it does not and the way the dead space volumes were determined would not generally be possible as it relies on several hours of data for

all the different configurations. Further investigations would be required in data collected with modulated sinewaves to determine if the good results were an indication of promise in the technique or just occurred by chance. Even if additional investigations were positive a method of reliably independently determining the dead space would be required. One solution may be to assume a single transitory compartment dead space. Its volume could then either be determined using a similar method to that used above for the mixed and transitory compartment or the volume determined using one of the breath-by-breath methods could be used.

The ultimate question is what these results mean for determining the homogeneity of human lungs using these apparatus. It could be argued that, given the difficulties experienced here, it is unlikely that a measure for the homogeneity could be determined from human lungs, given that inferring information from human data is likely to be more complex than inferring information from in this ideal set-up.

However as discussed in the previous chapter the homogeneities found within the multi-compartment bench lung appear to be lower and to have a smaller range to those found in human subjects. The simulated data showed that the inter-breath method and homogeneity index are also more effective at higher homogeneities in the presence of noise. Thus it may in fact be more difficult to determine between the configurations of the multi-compartment lungs than inhomogeneities in human lungs. Also the effects of the offset between the signals have not been taken into account. Given the results from the simulated data and the tentatively more positive results presented here, it would seem that the inter-breath method is the best candidate for further investigation in both multi-compartment lungs and humans.

5.4 Summary

Many different methods have been explored throughout this chapter. None of the methods applied to the homogeneous data have provided a clear indication that they are universally better than all other methods. However all of the methods generally show better results when determining the dead space volume than the end-expired alveolar volume, especially when comparing the various precision measurements.

Another general observation is that the precision of the end-expired alveolar volume results when considering all of the data is consistently significantly worse than when considering the subset of the data. Investigations into reasons for this did not provide any further indications as to the source of the additional variation in the results. Within any dataset there are likely to be subsets which produce more consistent results than

the whole dataset. Also the results from each method are not independent from each other as they came from the same dataset. If a source of this variation could be identified and measures put in place to compensate for errors, this would significantly improve the precision of all of the methods.

The benefits and shortcomings of many of the methods have already been discussed with Model 2 appearing to generally produce more precise results and Model 8 producing more linear end-expired alveolar results with less bias. Little discussion has so far been dedicated to the benefits of the breath-by-breath methods in comparison to the other methods. The other methods do not universally produce better results than the breath-by-breath methods.

The Fowler method performs better in terms of precision for the end-expired alveolar volume results than many of the newly proposed methods, although the results are quite similar when using Model 2, with the Model 2 results having slightly better precision. For the dead space results the Fowler method precision is generally outperformed when Model 2 is used and is comparable to the results for the other models. The Fowler method does not however perform as well in terms of linearity and bias. The ABC SMC results are all more linear for both the dead space and end-expired alveolar volume, especially when Models 2 and 8 are used. The absolute bias is also always worse when using the Fowler method, with the Model 8 results having significantly less bias.

The precision of the Bohr method results is not as high as those when using ABC SMC to determine both volumes. It does have slightly better linearity for the dead space volume when compared to Models 1 and 8, but not for Model 2 or in any case for the end-expired alveolar volume. The bias results are more mixed, with the Bohr method having less bias in the dead space results in all cases, except for ABC SMC when using Model 1. The absolute bias is also less for the end-expired alveolar volume when compared to that for Models 1 and 2, but is worse when using Model 8.

The inclusion of an offset as an additional parameter can provide improvements in the results. In general Model 1 results appear to improve most, with an improvement in the bias also seen for Model 2. When using Model 8, inclusion of the offset parameter did not generally improve the results.

Putting the results from the homogeneous bench lung data into the context of the thesis aims, the modelling of the entire expired signal does appear to have improved the results acquired from this data set. Although the breath-by-breath method when using the Fowler dead space method (which only models particular points in the breath) is not outperformed in all measures by all of the proposed methods, the use of Model

2 with or without offset correction does outperform the results in many ways and so is likely to result in better FRC predictions. The other models also generally show improved linearity and offset when compared to the breath-by-breath method. The results presented within this chapter do not provide a clear answer as to which model of the dead space should be used in order to obtain the best FRC results. However they do indicate that use of the newly proposed methods is more likely to result in a more robust measure of FRC. This could be used to better monitor the FRC of patients within the ICU, which has long been touted as a possible method of determining the appropriate level of PEEP [77, 94, 171].

Recovering the dead space independently of the alveoli make up was successful in the homogeneous data with only a slight change in bias occurring. The results from the multi-compartment lung were not so promising. In three of the cases (configurations A, C and D) the variation was similar to that seen for the homogeneous bench lung data but in the fourth case (configuration B) more variability was seen. In addition there was a marked difference in the mixed and transitory volumes recovered from the different configurations despite the simulated results suggesting that the volumes should be unaffected by the alveoli configuration.

The results for determining information about the homogeneity of the lung were not very promising. This may in part be due to the issues in independently determining the dead space volumes, especially when using the inter-breath variation method which relies on knowing information about the dead space. However the other two methods, which did not necessarily rely on this information, performed no better. This was not necessarily surprising for the breath-by-breath method given the simulated results. For the reasons discussed above, this does not necessarily exclude the use of these methods to determine information about the homogeneity of human lungs and so improving clinical decision making within the ICU, however it does highlight that this may be a reasonably difficult task and that resolution of the methods may be poor. The tentative good results for the inter-breath variation method may suggest that it is the best candidate for further investigation, if a method for determining information about the dead space can be devised.

5.5 Key Findings

Data from a homogeneous and multi-compartment bench lung are analysed within this chapter. The key findings from the homogeneous bench lung data are:

- The new methods, which includes a model for the entire expired concentrations, generally outperform the previously used breath-by-breath method, which only uses particular points within the breath.
- No one model for the dead space outperforms the others in all the statistics calculated.
- In general, Models 2 and 8 (which have mixing within the dead space) perform better than Model 1 (which models the dead space as a single transitory compartment).
- Model 2 performs better than Model 8 in terms of the precision statistics, while Model 8 performs better in terms of linearity and bias.
- An additional offset parameter within the model generally improves the results when using Model 2, particularly in terms of bias. Although there is a significant computational price to pay.
- An additional offset parameter does not improve the Model 8 results.
- Analysis using Model 2 (with no offset parameter) shows that it is possible to determine the volume of the two dead space compartments independently of the alveolar volume without a major impact on the accuracy and precision of the recovered volumes.

The key findings from the multi-compartment bench lung data are:

- Determining the dead space volumes (using Model 2) independently of the alveolar make up was not successful in the data available, with largely varying volumes recovered from a bench lung with the same dead space.
- None of the methods of determining the homogeneity provided convincing results, but the inter-breath variation results were tentatively better than the others.

In addition to these specific findings relating to the determination of lung function, a number of secondary findings about convergence of ABC SMC are made:

- When making a model selection true convergence is only achieved when only one model remains, which is unlikely to be a good representation of the true posterior distribution for the proposed models.
- When making parameter estimations there are difficulties determining when convergence has occurred automatically (and even manually), thus automating this process is unlikely to be a trivial task.

5.6 Conclusion

These findings suggest more accurate and precise lung volumes are more likely to be determined using the new methods (which model the concentration of the entire expired breath) than by using the breath-by-breath methods. In addition the use of a model that includes mixing within the dead space (Model 2 or 8) is preferable to one without (Model 1). Including an offset parameter can improve the Model 2 results, but not the Model 8 results. These findings will be used in the next chapter to guide the analysis carried out in human data.

The findings from the multi-compartment lung are less favourable and suggest that all of the proposed methods may struggle to be successful in humans. However there are a number of reasons why the multi-compartment bench lung may not mimic the homogeneities of a human lung well, so further investigations in human data could possibly be beneficial. The tentatively more positive results for the inter-breath variation method concur with results of the previous chapter suggesting it is likely to provide the more promising results in any further investigations.

These results would suggest concentrating on predicting FRC may be more fruitful than attempting to determine a measure of ventilation-volume inhomogeneity. This conclusion may appear to suggest that the ability to determine abnormalities in ventilated patients' lungs, and so influence their care, will be diminished. However the use of FRC in influencing the choice of ventilator settings, and so avoiding VALI, has long been advocated [77, 94, 171]. Thus the ability to robustly determine FRC could have as much or even more potential to positively influencing patient care than a poor measure of homogeneity. The relative robustness of these methods in determining FRC will be investigated further in the chapter that follows.

Chapter 6

Human Data

6.1 Introduction

The ultimate test of any biomedical method is how it performs on human data. Within this chapter the methods previously described will be applied to human data in order to investigate how they perform. Performance is analysed by examining the variability of repeated measurements in the same subject and by comparing the FRC (sum of all lung compartments) with that determined by a body plethysmograph.

The available data use a nitrous oxide tracer gas, which is relatively soluble with a Oswald blood-gas coefficient of 0.47. Therefore the uptake and return from the body have to be considered. Due to the practical time constraints of reaching equilibrium throughout the body, only the equilibrium of the lung compartment is achieved prior to starting the individual tests. Previously, in Section 4.7, methods to deal with this were devised and investigated in simulated data. Those two methods, the fixed venous return and standard man methods, will thus be used in this analysis.

The better performing dead space models from the previous chapter will be used here. Dead space Models 2 and 8 will be used when not correcting for the offset between the two signals and Models 1 and 2 when the offset is included as an additional parameter. Generally a genetic algorithm will be used as investigations in bench lung data (not detailed within this thesis) found it to be less prone to getting stuck in local minima than the Nelder-Mead simplex and to be more computationally efficient than ABC SMC. However some comparisons with ABC SMC will also be made and the Nelder-Mead simplex will be used to improve the results obtained by the genetic algorithm.

Note that the methods described earlier for determining the homogeneity of the lung assume the use of an insoluble gas and would require further modification in order to

Subject ID	No. of Datasets	Plethysmograph FRC (l)	Sex	Age	Smoker (no. per day)	Pulmonary or Cardiovascular Health Problems
F2	6	3.60	M	20	Never	Asthma as child (mild).
F5	9*†	3.90	M	21	Never	N/A.
F12	2	3.27	M	30	Current (10)	N/A
R2	6	4.06	M	30	Never	Bronchitis as child.
R3	5	2.78	F	19	Never	N/A
R4	7	3.07	M	20	Never	N/A.
R5	6†	2.43	F	23	Never	N/A.
R6	6	3.46	M	20	Never	N/A
R7	6	2.15	F	20	Never	N/A
R8	5	3.64	M	21	Never	N/A
R9	6	2.89	F	26	Never	Heart mummer.
R10	6†	3.21	F	22	Never	N/A.
R11	6	3.25	M	31	Never	N/A
R12	4	2.35	F	23	Current (<1)	N/A
S1	6	3.49	M	23	Never	N/A
S4	6	4.93	M	21	Current (1)	N/A
S10	6	3.00	F	20	Never	N/A

Table 6.1: Details of the human data. *These datasets were obtained at two different sessions on different days. With three datasets taken on one occasion and six on the other. †For one of these datasets the concentration signal is corrupt at the start whilst the lung is being brought into equilibrium.

be applied to the data available. Thus this chapter will concentrate on determining the FRC assuming that the lung is homogeneous.

6.2 Data

The data available are from spontaneously breathing human subjects. Although the primary aim is to use these methods on ventilated patients, only limited data were available with no comparative plethysmograph volume. There are some practical issues associated with the collection of data from ventilated patients, however the effective use of the equipment has been demonstrated in these patients by the clinical collaborators. They hope to publish this work in due course. Otherwise the main difference between data collected in ventilated and in spontaneously breathing subjects is the breathing pattern. The breathing pattern is more predictable in the ventilated case and so FRC should be less complex to determine in ventilated patients, particularly as the lung returns to the same volume at the end of each breath.

The data used in this chapter are detailed in Table 6.1. All 18 participants self declared that their general health was good; of these 11 were male and seven were female. The mean age was 23±4 years. Participants had repeated measurements taken during a single session lasting approximately two hours. This was preceded or immediately followed by the use of a body plethysmograph (MasterScreen Body, Cardinal Health; Lung Function Laboratory, Churchill Hospital, Oxford) to determine the FRC.

The order of the tests was randomly selected. In a number of cases plethysmograph readings were not taken and so these data were excluded from the analysis. The study was approved by the research ethics committee of the Medical Sciences Division, University of Oxford (Ref: MSD/IDREC/C1/2012/35) and written informed consent was obtained from all subjects.

6.3 Statistical Analysis

The repeatability of a particular method was assessed using the coefficient of variation. There are two subjects who have a significantly different number of datasets than the others: F5, who has nine datasets from two different sessions (3 and 6 for each session) and F12 who only has two datasets. It is not ideal to have three results from a different session when this not the case for any other subject. The significant difference in the number of datasets is also problematic for the bootstrap method described below. Therefore a subset is also analysed separately, which excludes the three repeats from the separate session of F5 and both F12 results. F12 is the heaviest smoker within the group with all other subjects smoking one cigarettes or fewer a day and having no current lung condition, which makes for a young healthy nearly smoker free cohort. This is both advantageous and disadvantageous, as discussed further below.

The coefficient of variation provides information about the repeatability of the methods but provides no information about how well they perform in comparison with other methods. Therefore the results of regression analysis, comparing the FRC values to that determined by plethysmograph are also presented. As previously discussed in Section 2.5.1, FRC values determined by the plethysmograph have a reasonable amount of variation and therefore ordinary least squares is unlikely to provide a reliable regression analysis for the reasons described elsewhere [20, 37, 61]. Therefore Deming regression [61, 209] is used here with a measurement error ratio of 4 (variance of method/variance of plethysmograph). This is based on the assumption that the approximate coefficient of variation for the plethysmograph is 5% [92] and that the coefficient of variation for the various methods presented below is 10%. Note that the mean values are also assumed to be the same.

A bootstrapping approach [45, 156, 157] is used here to determine the mean and standard deviation of the regression coefficients. This is done by randomly selecting a result for each subject and then fitting a regression line through those points. Where the number of repeated measurements for a subject is below the maximum number of repeats, a subject is excluded from the analysis when data are missing. The subset,

	V_D	V'_A	$\dot{Q}_P$	FRC
All data	7.28±4.19	12.64±15.24	18.26±18.26	11.40±13.87
Subset	7.09±4.23	10.71±13.34	15.06±15.06	9.70±12.22

Table 6.2: Mean and standard deviation of the coefficient of variation (expressed as a percentage) for the parameters determined using the Fowler breath-by-breath method.

described above, is used to ensure a similar number of datasets for each subject. Using the subset does mean that the plethysmograph reading for F5 was taken seven days earlier than data were collected, but this is felt to be better than excluding this subject in the analysis as would be necessary otherwise. The FRC is unlikely to have changed significantly over that time and if it has then this is only likely to add more variation into the results of the study.

100,000 bootstrap iterations were normally enough to get values for the mean and standard deviation consistent to two decimal points; however where there is a larger amount of variability present this is not always the case. Thus when two repeats of this bootstrap do not agree both will be presented and discussed. In addition to the regression information, the mean and standard deviation of the difference between the method and plethysmograph FRC will also be determined as part of the bootstrap.

6.4 Method Comparison

6.4.1 Breath-by-Breath Method

Here the results for the Fowler method are presented, rather than the Bohr method, as the end-expired alveolar volumes were shown to have good precision indicators in the previous chapter. Table 6.2 shows the coefficient of variation results. The dead space volume shows the best repeatability with the lowest mean and standard deviation. The end-expired alveolar volume has a mean value of nearly twice that of the dead space and a standard deviation of about three times the value. When these volumes are combined to give a value for FRC, the mean and standard deviation are in between the values observed when the volumes were individually determined. However the values are closer to those for the end-expired alveolar volume, as might be expected as it constitutes more of the volume of FRC. The mean and standard deviation for the pulmonary blood flow are larger than those observed for the other parameters. The results for the subset are slightly more favourable in all cases, except for the standard deviation of the dead space volume which increases slightly.

Table 6.3 shows the regression analysis and error results when comparing the FRC results to those determined by plethysmography. The results from two separate runs

Regression Analysis				Error	
Gradient	Intersect(l)	SE(l)	Bias(l)	Mean(l)	SD(l)
0.94±0.23	-0.25±0.71	0.32±0.19	-0.44±0.12	-0.44±0.11	0.44±0.26
0.94±0.25	-0.26±0.78	0.32±0.20	-0.44±0.12	-0.44±0.11	0.44±0.27

Table 6.3: Results of the bootstrapped regression and error analysis when using the Fowler breath-by-breath method to determine FRC.

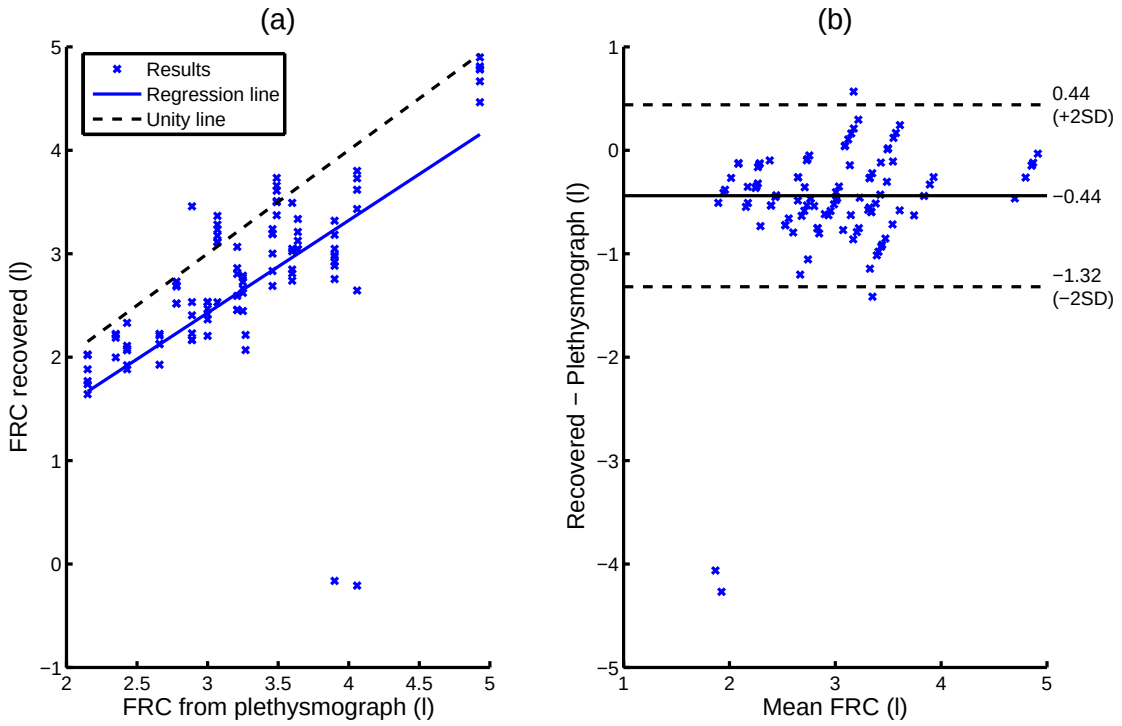


Figure 6.1: Plots of the results determined by the Fowler breath-by-breath method. (a) scatter plot and (b) Bland-Altman plot. The regression line and Bland-Altman mean difference and limits of agreement are determined from the mean values in Table 6.3.

of the bootstrapping analysis are shown. The results are similar for the two runs but show some minor differences for the gradient and intersect. Figure 6.1 shows both a scatter and a Bland-Altman plot. The regression line on the scatter plot and the mean and limits of agreement on the Bland-Altman plot were determined using the values in Table 6.3.

These results show a general bias of -0.44l with the regression and Bland-Altman error analysis showing virtually the same results. The linearity of the results is quite close to unity but does have a reasonably high standard deviation. The standard error of the estimate and the standard deviation of the error both have relatively high means, with the mean values about two and four times that seen for the bench lung, respectively. Both plots back up the results in Table 6.3 with the results generally

below unity in both cases and showing a reasonable amount of scatter. Apart from the bias, there are no obvious trends which would suggest a systematic form of error. The plots do however reveal two anomalous results from two different subjects. Further investigations into these results reveals dead space volumes of 285 and $231ml$, which are within the range of those observed in other cases. This would suggest issues in recovering the end-expired alveolar volume rather than the dead space volumes.

6.4.1.1 Discussion

The coefficient of variation results indicate that more variation is present within the repeated recovered volumes for the end-expired alveolar volume than for the dead space. This is consistent with the results from the homogeneous bench lung where the precision of the dead space volume was higher. Also within this data, because the subjects are breathing naturally, the end-expired alveolar volume changes to a greater degree than when using a ventilator. The approximate mean absolute difference from the mean end-expired volume is $81ml$ but with a maximum absolute difference of $1.4l$, based on the flow sensor data alone. These differences are not accounted for by this method, but doing so could potentially improve these results.

The coefficient of variation is even higher for the pulmonary blood volume with a higher mean and standard deviation observed. There are no comparative results from the bench lung, but more variation is consistent with the expected results. The pulmonary blood flow is likely to be affected by errors in the estimate of the dead space volume in a similar way to the end-expired alveolar volume. Although this method makes an effort to correct for the fact that the body is not brought into equilibrium with the lung concentration at the beginning of the test, there will still be some errors due to this. These are likely to affect the pulmonary blood flow estimates more than other parameters as the method of determining the blood flow is directly influenced by the gas concentration in the body and lung. Finally more variation in the blood flow rate would be expected as the cardiac output will vary to a much greater degree than the air volume of the lung during the session.

With little to compare the results against it is difficult to comment on the precision of this method. However the standard error of the estimate, the standard deviation of the error and the standard deviation for all values determined by the bootstrapping approach are all relatively high, seeming to indicate a reasonable amount of variation within the results. This may in part be due to the two anomalous results. Determining what caused these anomalies could be a route to improving this method.

Compartment	Lower Bound	Upper Bound
V_M	$5ml$	$300ml$
V_R	$5ml$	$350ml$
V'_A	$0.5l$*	$7.0l$*
$\dot{Q}$	$3.0l/min$	$18l/min$
$C_{\bar{V}}$	0%	5%

Table 6.4: Bounds for the ABC SMC uniform prior and genetic algorithm when processing human data. *Different bounds were used for each dataset when using ABC SMC as described in the text.

The mean of the gradient is 0.94 which seems relatively good and which is consistent with the results seen for the bench lung, however the standard deviation is relatively large. The mean bias is also large at -0.44l, however this is similar to that observed for the end-expired alveolar volume in the bench lung. Similar analysis by the collaborators suggested that the source of this error could be a general overestimation by the plethysmograph [159]. The possible sources for this error will be discussed in more detail below.

6.4.2 Fixed Venous Return

The fixed venous return method as proposed in Section 4.7.2 is used here to analyse these results. Both the ABC SMC and a genetic algorithm are used, in addition to methods which include the offset as an additional parameter.

6.4.2.1 ABC SMC

The use of ABC SMC is restricted here to one model of the dead space, Model 2, as it is computationally intensive. Two models of gas exchange with the blood were used in this analysis: one which accounts for the volume change due to the uptake and return of tracer gas from the blood (i.e. the model balances the mass of tracer gas within the body) and one that assumes no volume change as a result of the tracer gas exchange with the blood. These models were described in Section 3.3.4.

The priors used are detailed in Table 6.4. The one exception is the prior for the end-expired alveolar volume, since if the volume of the alveolar compartment approached zero the DDE solver implemented in C crashed (getting stuck in a never-ending loop). In order to guard against this the prior was set to be 0.75l (lower limit) and 5.75l (upper limit) above the minimum alveolar volume. This was determined by direct numerical integration of the flow signal. When the alveolar compartment volume is adjusted, due to the exchange of the tracer gas with the body, low volumes can still occur under certain combinations of parameters and so the alveolar volume was held at 0.1ml, should the compartment get that small.

Volume Change Due to Exchange with Blood Included

	V_M	V_R	V'_A	$\dot{Q}_P$	V_D	FRC
All data	17.55±8.98	15.28±10.73	10.88±4.89	18.54±13.05	7.18±4.50	9.37±3.98
Subset	16.72±9.09	14.83±11.03	11.25±4.66	18.51±13.23	7.37±4.41	9.70±3.76

Volume Change Due to Exchange with Blood Ignored

	V_M	V_R	V'_A	$\dot{Q}_P$	V_D	FRC
All data	17.58±8.95	15.27±10.67	11.13±5.29	18.43±12.85	7.17±4.48	9.60±4.32
Subset	16.80±9.04	14.84±10.97	11.53±5.06	18.33±13.08	7.35±4.40	9.94±4.09

Table 6.5: Mean and standard deviation of the coefficient of variation (expressed as a percentage) for the parameters determined using ABC SMC with Model 2.

	Regression Analysis				Error	
Volume Change	Gradient	Intersect(l)	SE(l)	Bias(l)	Mean(l)	SD(l)
Included	1.11±0.11	-0.95±0.37	0.30±0.04	-0.55±0.08	-0.58±0.07	0.42±0.06
Ignored	1.12±0.11	-0.96±0.38	0.30±0.04	-0.54±0.08	-0.57±0.07	0.42±0.06

Table 6.6: Results of the bootstrapped regression and error analysis when using ABC SMC with Model 2 to determine FRC.

Tables 6.5 shows the coefficient of variation results. Of all the directly recovered parameters, the end-expired alveolar volume has the lowest variation with the lowest mean and standard deviation in all cases. The pulmonary blood flow has the largest variation, with the dead space volumes in between. Interestingly when the two dead space compartment volumes are combined to give the total dead space volume, the variation reduces with a lower mean and standard deviation observed. The same is true when this volume is combined with the end-expired alveolar volume to form the FRC. The mean of this is slightly higher than the total dead space volume, but the standard deviation is marginally lower. There are some minor differences when the volume change due to exchange with the blood is ignored, but these are small with a maximum difference of 0.5%.

Table 6.6 shows the results when the FRC is compared to the plethysmograph FRC. The second run of the bootstrap was identical and so is not shown. Figure 6.2 shows the results plotted on a scatter and a Bland-Altman plot; the regression line, mean difference and limits of agreement are determined from the mean values in Table 6.6. The mean and limits of agreement are those when the volume change is included, as showing multiple lines on the Bland-Altman plot is confusing.

Whether the volume changes associated with tracer gas exchange are included or not has little effect on the results. Some minor differences can be seen, especially on the Bland-Altman plot, but these do not appear to be systematic. The variation

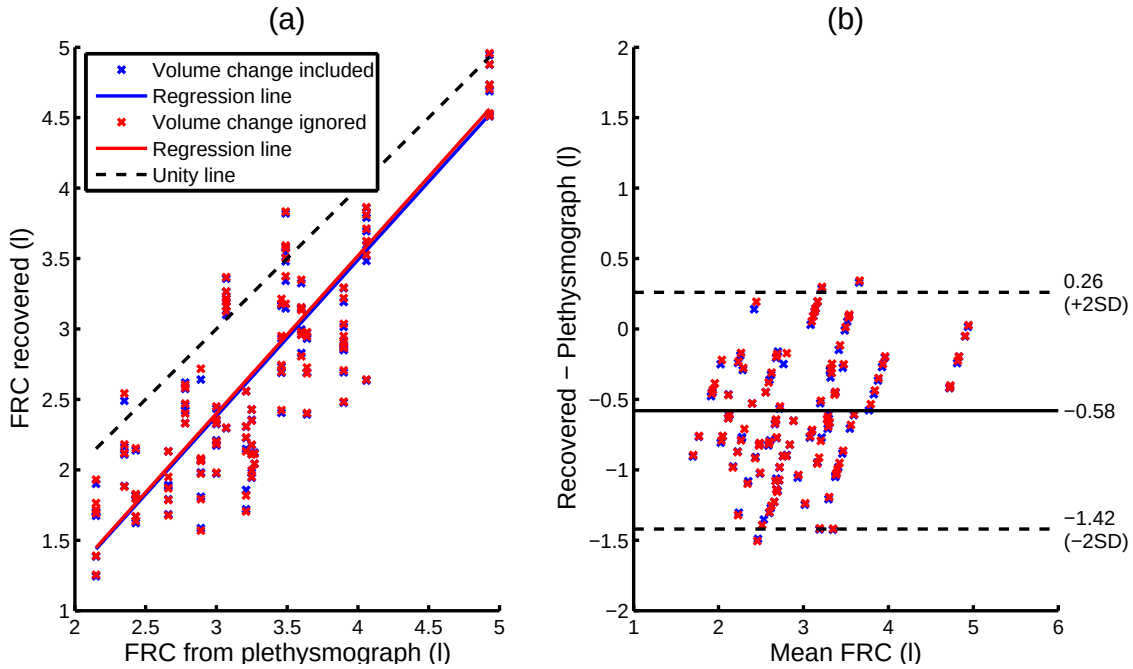


Figure 6.2: Plots of the results determined by ABC SMC with Model 2: (a) scatter plot and (b) Bland-Altman plot. The regression line and Bland-Altman mean difference and limits of agreement are determined from the mean values in Table 6.6.

		V_M	V_R	V'_A	$\dot{Q}_P$	V_D	FRC
Model 2	All data	21.94±13.63	15.48±10.37	15.80±6.26	35.58±18.46	7.67±4.03	13.93±5.30
	Subset	21.78±13.98	15.33±10.47	16.25±5.94	34.74±19.25	7.91±3.78	14.33±4.98
		$V_{M,1}$	$V_{M,2}$	V'_A	$\dot{Q}_P$	V_D	FRC
Model 8	All data	12.95±8.19	12.37±8.26	15.46±6.24	26.78±10.55	8.66±3.39	14.21±5.92
	Subset	13.05±8.41	12.20±8.44	15.95±6.04	24.98±8.66	8.24±3.28	14.70±5.69

Table 6.7: Mean and standard deviation of the coefficient of variation (expressed as a percentage) for the parameters determined using a genetic algorithm.

between the two methods is much less than for the repeated measurements within a subject. The standard deviations of the results in Table 6.6 are not as large as the breath-by-breath method. The one larger standard deviation is that of the intercept, however this is not reflected in the bias, which is effectively the intercept at the mid point of the results, and so a better measure of the variation that is likely to affect the results.

6.4.2.2 Genetic Algorithm

Coefficient of variation results when using a genetic algorithm with dead space Models 2 and 8 are given in Table 6.7. The bounds given in Table 6.4 were used. In general the coefficients of variation are quite similar when using Model 8 and Model 2. The major difference is that the two dead space mixing compartments of Model 8 have lower

	Regression Analysis				Error	
Model	Gradient	Intersect(l)	SE(l)	Bias(l)	Mean(l)	SD(l)
2	0.99±0.23	-0.65±0.71	0.35±0.06	-0.69±0.14	-0.69±0.11	0.49±0.08
8	0.86±0.22	0.22±0.70	0.35±0.08	-0.27±0.15	-0.23±0.12	0.51±0.11

Table 6.8: Results of the bootstrapped regression and error analysis when using a genetic algorithm to determine FRC. The results when the analysis was rerun were identical.

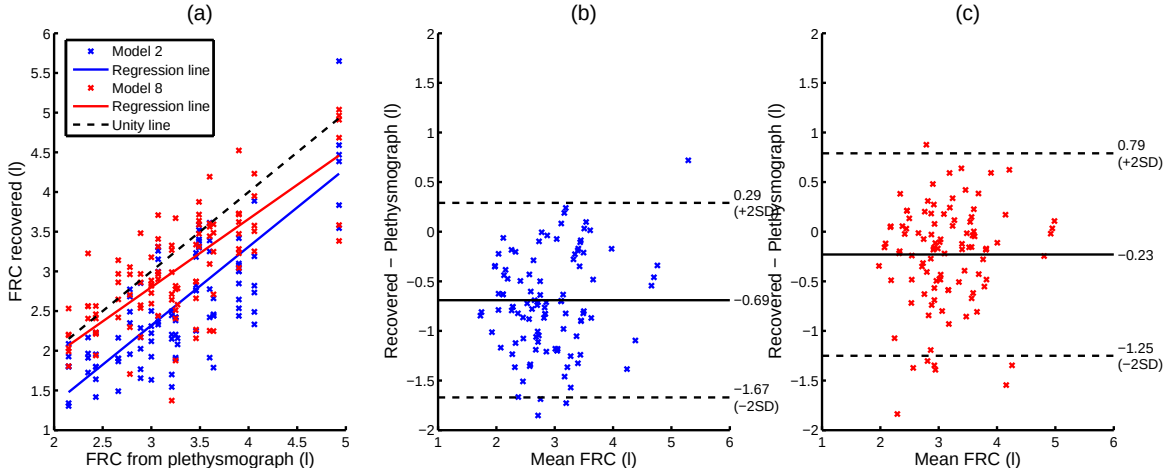


Figure 6.3: Plots of the results determined by a genetic algorithm: (a) scatter plot and Bland-Altman plots for (b) Model 2 and (c) Model 8. The regression line and Bland-Altman mean difference and limits of agreement are determined from the mean values in Table 6.8.

means and standard deviations than the mixed and transitory compartments used for Model 2. However this does not translate into a better coefficient of variation for the combined value, with the total dead space showing similar values for both models.

Table 6.8 and Figure 6.3 show a comparison with the plethysmograph results. The precision indicators are similar for both models, with Model 2 showing a slightly better mean for the standard deviation of the error. The biases determined by regression and error analysis have similar means and standard deviations. The mean value for Model 8 is closer to zero. For the gradient Model 2 has a better mean value, which is very close to unity. However the standard deviation is reasonably large at about 0.23. Thus the repeatability of these result is unlikely to be high.

The Bland-Altman plots show no obvious trend, signifying that the errors are unlikely to be systematic in nature. There are more points below the limits of agreement in Figure 6.3 (c) than may normally be expected in a Bland-Altman plot. However the limits have been determined using the mean values from the bootstrap procedure and so these repeated observations from a small number of subjects have less of an effect on the limits of agreement.

		V_M	V_R	V'_A	$\dot{Q}_P$	V_D	FRC
Model 2	All data	17.42±9.13	15.19±10.42	11.78±6.33	27.37±22.69	7.21±4.54	10.16±5.18
	Subset	16.69±9.29	14.77±10.74	12.12±6.27	26.94±22.95	7.39±4.46	10.45±5.11
		$V_{M,1}$	$V_{M,2}$	V'_A	$\dot{Q}_P$	V_D	FRC
Model 8	All data	10.74±5.14	8.87±4.34	8.09±3.21	20.82±12.91	8.05±3.87	7.42±2.83
	Subset	10.89±5.18	8.80±3.58	8.35±2.82	20.04±12.06	7.81±3.75	7.62±2.49

Table 6.9: Mean and standard deviation of the coefficient of variation (expressed as a percentage). The parameters are determined using a genetic algorithm followed by refinement using the Nelder-Mead Simplex.

	Regression Analysis				Error	
Model	Gradient	Intersect(l)	SE(l)	Bias(l)	Mean(l)	SD(l)
2	1.12±0.12	-0.98±0.39	0.31±0.05	-0.55±0.09	-0.58±0.08	0.44±0.06
8	1.03±0.09	-0.16±0.29	0.21±0.03	-0.05±0.07	-0.06±0.06	0.30±0.05

Table 6.10: Results of the bootstrapped regression and error analysis when using a genetic algorithm followed by refinement using the Nelder-Mead Simplex to determine FRC. The results when the analysis was rerun were identical.

The results from the genetic algorithm were refined using the Nelder-Mead Simplex. The results of the genetic algorithm were used as the starting point. The coefficients of variation results are shown in Table 6.9. The means and standard deviations are generally lower than when no refinement is implemented. The coefficient of variation are lower when using Model 8 than Model 2. The one exception is the combined dead space volume where the mean is slightly lower for Model 2, but the standard deviation is still lower for Model 8. In both cases the total dead space and FRC have lower means and standard deviations than the volumes that combined to determine them. In both cases, there is little difference between the results for all the data and the subset.

Table 6.10 and Figure 6.4 show a comparison with the plethysmograph results. The results generally all improve when compared to the pure genetic algorithm, with a notable reduction in the standard deviations. The one exception is the mean of the gradient for Model 2 which is further from unity. For all of the results Model 8 outperforms Model 2 with better means and lower standard deviations. The intersect results are larger in both cases, especially the standard deviation, but this is not reflected in the bias at the mid point.

Comparing both sets of results for Model 2 to the ABC SMC results shows that the results are similar when the Nelder-Mead simplex refinement is included. For the coefficient of variation of the dead space volumes little difference is seen. For the end-expired alveolar volume the results are slightly worse when using the genetic algorithm

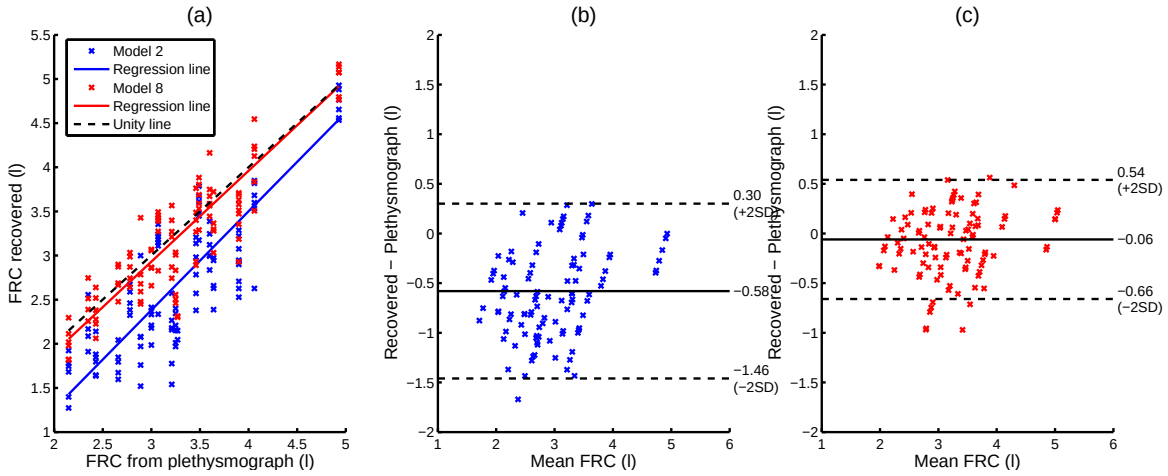


Figure 6.4: Plots of the FRC results determined by genetic algorithm followed by refinement using the Nelder-Mead Simplex: (a) scatter plot and Bland-Altman plots for (b) Model 2 and (c) Model 8. The regression line and Bland-Altman mean difference and limits of agreement are determined from the mean values in Table 6.10.

with Nelder-Mead Simplex refinement, with the standard deviation showing the most difference. This appears slightly to feed through into the FRC coefficient of variation. There is a more significant increase for the pulmonary flow with an increase of nearly 10% in both the mean and standard deviation. The results for the comparison of FRC with the plethysmograph FRC show some small differences, but are generally quite consistent.

The results when using the genetic algorithm alone show more significant differences to the ABC SMC results. The coefficient of variation results are generally worse. The exceptions are the pulmonary blood flow, which has a slightly lower mean and a substantially lower standard deviation, and the transitory and total dead space volumes, where the results are similar. For the comparison with the plethysmograph FRC the results are better in the ABC SMC case, except for the mean of the gradient but this has a reasonably high standard deviation.

6.4.2.3 Signal Offset Correction

The genetic algorithm was used also to infer the offset in addition to the other parameters. The resolution of the recovered offset was limited to $10ms$, as was previously the case when using a genetic algorithm. Models 1 and 2 were used as they showed the largest improvements when an offset parameter is included in the bench lung data. The coefficients of variation result are shown in Table 6.11.

The results for Model 1 are generally large, the means and standard deviations being higher than seen for Model 2 and those results presented so far. The exception

		V_R	V'_A	$\dot{Q}_P$	FRC
Model 1	All data	6.38±5.66	19.64±14.26	32.70±22.19	16.40±11.76
	Subset	5.48±2.81	19.47±10.43	29.48±18.10	16.19±8.32

		V_M	V_R	V'_A	$\dot{Q}_P$	V_D	FRC
Model 2	All data	14.52±7.67	29.74±19.03	15.18±7.64	25.13±11.47	10.07±5.64	13.92±7.15
	Subset	15.03±7.62	29.70±18.97	15.08±7.86	25.31±11.84	9.92±5.19	13.81±7.35

Table 6.11: Mean and standard deviation of the coefficient of variation (expressed as a percentage) when an offset parameter is included. The parameters were determined using a genetic algorithm.

	Regression Analysis				Error	
Model	Gradient	Intersect(l)	SE(l)	Bias(l)	Mean(l)	SD(l)
1	1.17±0.20	-1.76±0.64	0.51±0.07	-1.13±0.12	-1.18±0.10	0.71±0.09
1	1.18±0.37	-1.77±1.21	0.51±0.07	-1.13±0.12	-1.18±0.10	0.71±0.09
2	1.18±0.27	-0.18±0.81	0.38±0.10	0.46±0.17	0.42±0.14	0.54±0.15
2	1.18±0.26	-0.16±0.78	0.38±0.10	0.45±0.17	0.42±0.14	0.54±0.15

Table 6.12: Results of the bootstrapped regression and error analysis when an offset parameter is included. A genetic algorithm was used to determine FRC.

is the transitory dead space volume, which is the only dead space compartment. It has a considerably lower mean and standard deviation, which are lower than for any of the dead space compartment results presented so far. For the subset, the means and standard deviations are also lower than for Model 2 and the other results presented so far.

The Model 2 results are generally also worse than those seen so far. The exception to this is the mixed dead space compartment, which has both lower means and standard deviations. This is not reflected in the total dead space volume. Also the pulmonary flow results are sometimes better, with both the mean and standard deviation being less than those when the genetic algorithm is used, with or without Nelder-Mead simplex refinement.

Table 6.12 and Figure 6.5 show the results when these methods are compared to plethysmograph FRC. For the first time since the breath-by-breath method, differences between the bootstrap runs are seen. These affect the standard deviations recovered for the gradient and intercept and to a minor degree the mean of the intercept. This is likely to be related to the increase in the standard deviation in these cases; i.e. in general the variation in the results has increased, so there is an increase in variability between each iteration of the bootstrap, so the results do not converge within the 100,000 iterations used.

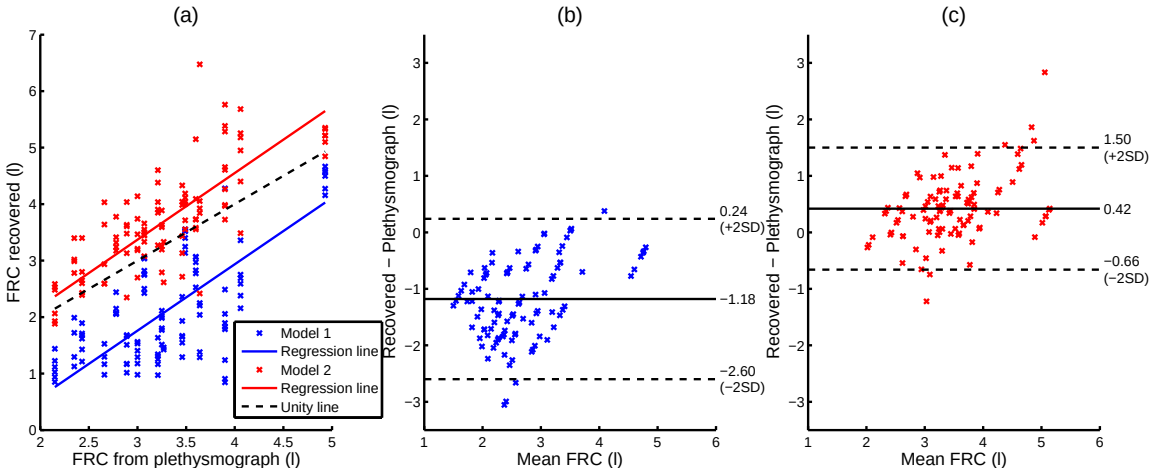


Figure 6.5: Plots of the FRC results determined by genetic algorithm when the offset is included as a parameter: (a) scatter plot and Bland-Altman plots for (b) Model 1 and (c) Model 2. The regression line and Bland-Altman mean difference and limits of agreement are determined from the mean values in Table 6.12.

The mean gradients are near identical for both models, with a reasonably high standard deviation in both cases. The bias however is quite different with Model 1 underestimating and Model 2 overestimating. The absolute mean bias for Model 1 is over double that of Model 2 for both regression and error, however the standard deviations are slightly lower for Model 1. The same is true for the standard error of the estimate and standard deviation of the error; Model 2 has a lower mean but slightly higher standard deviation.

When comparing the Model 2 mean results to the other Model 2 results presented so far the main change is the bias. It is now positive whereas it was previously always negative with the absolute value $100ml$ closer to zero bias. However the standard deviation has increased by $90ml$ when compared to the best results. This is part of a general increase in the standard deviation in all cases which indicates that the results are less repeatable. This is coupled with an increase in the mean values of the other indicators, particularly the two precision indicators with the mean standard deviation of the error increasing by $100ml$ when compared to the best results.

Figure 6.6 shows box plots for the offset determined here. The median results are significantly different from each other at -20ms and 90ms for Model 1 and 2, respectively. The spread of the results is approximately the same with an interquartile range of 50 and 60ms and a whisker range of 170 and 180ms respectively. There are a number of outliers in each case, with a single outlier at the upper limit of the parameter range in each case.

The Nelder-Mead simplex was used to refine the results of the genetic algorithm.

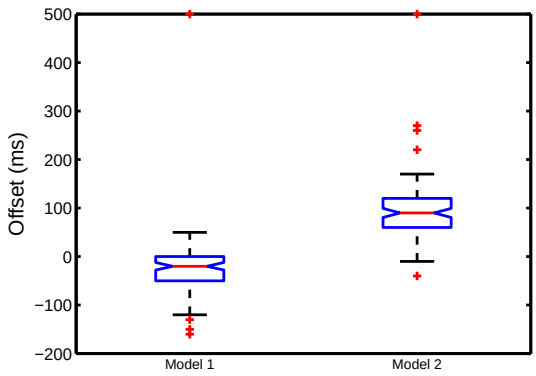


Figure 6.6: Box plots of the offset recovered by a genetic algorithm.

The offset was fixed at the result of the genetic algorithm and not refined, as in previous implementations. For the other parameters the genetic algorithm results were used as the starting point. This did not result in any improvement, with the results being close to those observed when no refinement was made. Attempts were also made to infer the offset by ABC SMC, but in the majority of cases the results had not converged after 60 days.

6.4.2.4 Discussion

None of these methods outperform all the other methods in all cases. However the use of Model 8 with the genetic algorithm followed by refinement with the Nelder-Mead simplex generally performed better than all the other fixed venous return methods. When FRC is considered it does outperform the other methods. The coefficient of variation has a lower mean and standard deviation.

It also performs better when compared to the plethysmograph FRC. The mean and standard deviation of absolute bias determined by regression and error analysis are less than in any other case, with a mean of less than $60ml$. Both the precision statistics are also better. The $300ml$ mean for the standard deviation of the error may initially seem relatively large. However as the plethysmograph has a coefficient of variation of approximately 5% [92], which would result in a standard deviation of $170ml$ at the mid point of the results presented here, this is a respectable level of precision. The one statistic in which this method is outperformed by another method is the mean of the gradient. When the genetic algorithm alone is used with Model 2 it achieves a mean of 0.99 (compared to 1.03), but the standard deviation is significantly greater (0.23 compared with 0.03).

The coefficient of variation for the total dead space volume does not have the lowest mean, with both Models 1 and 2 having lower values in a number of cases. However

the standard deviations are lower, except for the subset in one case. As discussed previously (see Section 2.4.1), the dead space is easily determined by other methods and is not of much physiological significance, so is only important if it also assists with determining FRC better. Also in some ways Model 8 is moving away from the concept of modelling the lung as made up of a dead space and an alveolar compartment to one made up of a series of mixing compartments. It should also be noted however that the two individual mixing compartments do in all cases have lower coefficients of variation than the mixing and transitory compartments of Model 2.

The coefficient of variation for pulmonary flow is also not better than all other cases, with ABC SMC (with Model 2) producing better results. A better coefficient of variation does not mean the flow rates determined are any more representative of the actual flow rates, especially as the within subject variation of the flow rate is expected to be much larger than the air volumes of the lung. Some comparative values for the flow rate would also be required to determine how representative the recovered flow rates are. Also if the primary focus is determining FRC then the additional variation and even less accurate estimates of pulmonary flow rate may be acceptable if better FRC estimates are achieved.

Only Model 2 was analysed in all cases. The ABC SMC results were similar to those for the genetic algorithm with the Nelder-Mead simplex refinement for the air volumes with ABC SMC only showing very limited improvements in the FRC estimate, which are unlikely to be significant. The more notable improvement was in the pulmonary blood flow which had a coefficient of variation with means and standard deviations which were just under 10% better. If further investigations into determining the pulmonary blood flow were being made, it may be worth investigating if similar improvements could be made when using Model 8.

However there is a significant computational cost to using ABC SMC. Detailed results for these are not given here as comparisons are difficult due to the differing equipment used to compute results. However the ABC SMC results all took more than 10 days to compute, while the other results all took less than 10 hours. The ABC SMC implementation has also already benefited from the significant efficiencies that the use of C can bring.

The results when the volume changes associated with the exchange of tracer gas with the blood is included within in the model or not included, do not differ greatly. Thus this volume change can be ignored without greatly effecting the results. This is likely to be because the volume changes associated with this exchange are minor and

likely to be negligible compared with other sources of error. For example, the fact that the respiratory quotient is unlikely to be unity and that the volume change associated with changes in temperature are likely to be more significant. Also it is not clear if detrending the inspired volume when pre-processing the flow signal eliminates some of the changes in volume associated with this exchange, especially as the body is not in equilibrium with the lung and so there is a net uptake of tracer gas during each test.

This is consistent with the previous simulated results of Barton *et al.* [25, 26]. However, as detailed by Barton *et al.* [25, 26], this is likely to be constrained to these particular forcing waves (including their forcing period, mean and amplitude) and the tracer gas (nitrous oxide) and not generally applicable across all forcing waves and concentrations.

Including the offset as an additional parameter does not generally improve the results. The coefficient of variation results for dead space volume of Model 1 are the only real improvement and then the standard deviation for all of the data does not improve. No analysis of Model 1 alone has been carried out here, so no comparison can be made. However the results in the previous chapter showed no real improvement in the dead space results, only the end-expired alveolar results, when an offset was included.

The median offsets are significantly different when using the two models, which was not the case for bench lung data, presented in the previous chapter. For Model 1 the median offset is negative. This implies that the change in concentration registered before the change in flow, which is contrary to what has been previously reported in the literature (see Section 2.6) and the previous chapters. For Model 2 all but a few of the offsets are positive, as would be expected. However the median and interquartile range are about double that seen in the bench lung data. If the true offset was being determined similar median results would be expected as the same equipment was used. These changes could be down to other factors such as the offset correction associated with the response time of the concentration sensor, which is inseparable from the true offset. Changes in nature of the breathing pattern, as it is spontaneous here rather than controlled by a ventilator, could also have been a cause of this.

However, ultimately the aim is not to determine the offset itself but to improve the robustness of other parameters, which has not been achieved. Allowing a higher resolution for offsets, as suggested in the previous chapter, could improve the results. However based on the results of ABC SMC in the previous chapter, where a continuous range was used, any improvement is unlikely to be large and less than that already seen

	V_M	V_R	V'_A	$\dot{Q}_T$	V_D	FRC
All data	9.84±5.76	12.27±7.46	10.28±4.37	39.84±22.33	6.75±3.66	9.38±3.98
Subset	10.21±5.51	11.99±7.41	10.49±4.49	40.31±22.95	6.68±3.71	9.58±4.07

Table 6.13: Mean and standard deviation of the coefficient of variation (expressed as a percentage). The parameters were determined using ABC SMC and assuming the body of a standard man.

by using Model 8 instead of Model 2. Use of Model 8 with offset correction would also seem unlikely to improve results given that no improvement was seen in the previous chapter.

One parameter which has been estimated in all cases but not presented here is the fixed concentration of the venous return. This is because there is little expectation for the value of that parameter, other than that it should be less than the value associated with the mean of the forcing sine wave. This was the case, except for up to two outliers, in all the results above where an offset is not included. A significant amount of variation in the value would be expected dependent on the particular manual equilibrium process used for that dataset, whether any tracer gas remains within the bodies of subjects from previous experiments and any particular characteristics of that subject's body.

6.4.3 Standard Man

As described in Section 4.7.3, an alternative to assuming a fixed concentration of the tracer gas in the venous obtained return blood is to assume the body of the standard man. Here results when using Models 2 and 8 for the dead space and making this assumption will be presented. No attempt will be made to correct the offset between the signals as this was found to provide only limited benefit in the bench lung data and no benefits when using the fixed venous return method.

The modelling of the lung starts from time zero in this case, when the concentration of tracer gas within the body is assumed to be zero. For three of the datasets (detailed in Table 6.1) the signal was not recorded correctly from time zero, so these three datasets were excluded from this analysis.

6.4.3.1 ABC SMC

Table 6.13 details the coefficient of variation results when using ABC SMC with Model 2 for the dead space. The volume change of the alveolar compartment due to tracer gas exchange with the blood was taken into account. Comparing these results to those determined using the ABC SMC and assuming a fixed venous return shows an improvement in the mean and standard deviation for the mixed and transitory dead

Regression Analysis				Error	
Gradient	Intersect(l)	SE(l)	Bias(l)	Mean(l)	SD(l)
0.88±0.10	-0.31±0.34	0.26±0.05	-0.74±0.07	-0.71± 0.07	0.37±0.07

Table 6.14: Results of the bootstrapped regression and error analysis. FRC was determined using ABC SMC and assuming the body of a standard man.

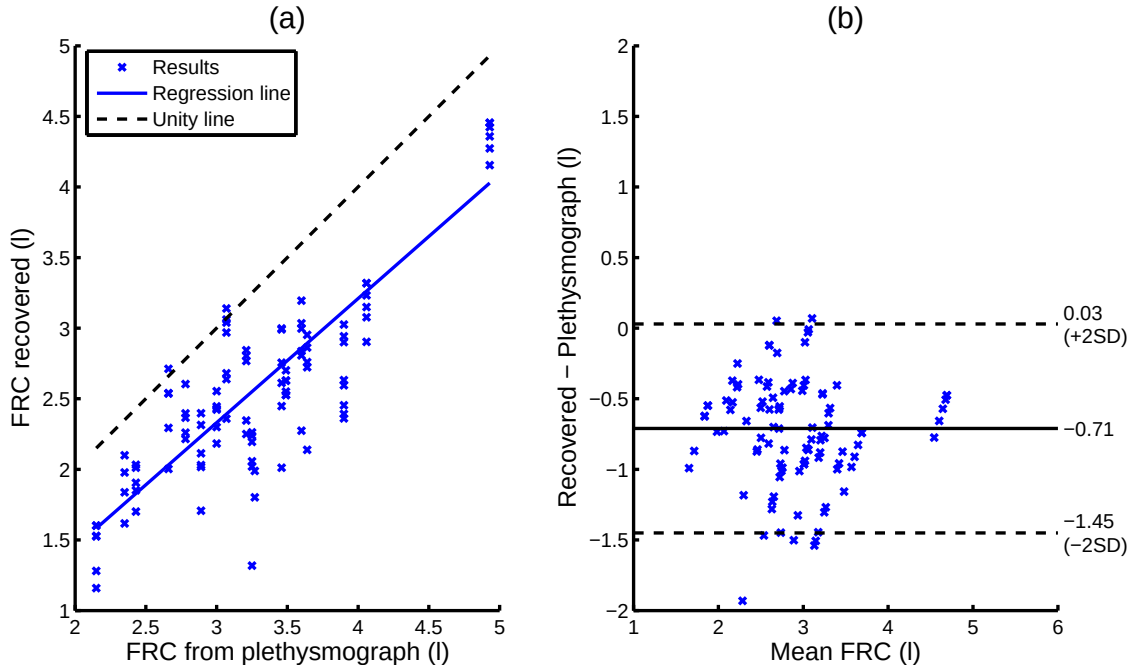


Figure 6.7: Plots of the results determined by ABC SMC with Model 2 and assuming the body of a standard man: (a) scatter plot and (b) Bland-Altman plot. The regression line and Bland-Altman mean difference and limits of agreement are determined from the mean values in Table 6.14.

space compartments. To some extent this appears to feed through into the results for the total dead space volume, with a slightly lower mean and standard deviation observed. However the improvement is not as large as for the compartments on their own.

The end-expired alveolar results are marginally better but the improvement is unlikely to be significant. The improvement to the individual volumes that make up the FRC does not result in an improvement in the FRC results. Also the variation in blood flow rate has increased substantially[1]. The coefficient of variation mean for the total arterial blood flow rate is more than double that previously observed for the pulmonary blood flow and the standard deviation has increased by nearly 10%.

[1]The total arterial blood flow ($\dot{Q}_T$) is the parameter recovered here, rather than the pulmonary blood flow ($\dot{Q}_P$), as the model for a standard man's body includes shunt. The coefficient of variation is unlikely to be affected as the value is normalised against the mean. In addition a relatively low amount of shunt (3%) is assumed in the model. Thus the values for both flow rates are assumed to be comparable.

	V_M	V_R	V'_A	$\dot{Q}_T$	V_D	FRC
Male	10.92±5.69	11.68±8.04	10.93±5.44	40.18±26.98	6.18±3.43	9.78±4.86
Female	9.19±5.50	12.43±7.00	9.86±2.95	40.49±17.68	7.39±4.26	9.31±2.93

Table 6.15: Mean and standard deviation of the coefficient of variation (expressed as a percentage) for both sexes. The parameters were determined using ABC SMC and assuming the body of a standard man. The results are for the subset detailed in the text.

	Regression Analysis				Error	
	Gradient	Intersect(l)	SE(l)	Bias(l)	Mean(l)	SD(l)
Male	0.91±0.70	-0.46±2.41	0.31±0.08	-0.78±0.09	-0.78±0.10	0.43±0.10
	0.91±1.17	-0.48±4.09	0.31±0.08	-0.78±0.09	-0.78±0.10	0.43±0.10
Female	1.03±0.23	-0.69±0.60	0.15±0.05	-0.61±0.07	-0.60±0.07	0.21±0.07
	1.03±0.23	-0.70±0.60	0.15±0.05	-0.61±0.07	-0.60±0.07	0.21±0.07

Table 6.16: Results of the bootstrapped regression and error analysis for both sexes. FRC was determined using ABC SMC and assuming the body of a standard man.

Table 6.14 and Figure 6.7 show a comparison of these results with the plethysmograph FRC. Comparing the results to those determined by the fixed venous return method, the standard deviations are approximately the same across the range of statistics. The mean values vary more substantially. The mean gradient is below unity, instead of above and the absolute difference from a value of one is higher. There is also a larger bias, which is 130 to 200ml larger in absolute terms depending on the statistic compared. The mean standard error of the estimate and standard deviation of the error are however both lower by 40 and 50ml, respectively.

The assumption of the body of a standard man may result in this method not being applicable to women. In a attempt to investigate if this is the case, the analysis was carried out on the 10 males and 7 females within the datasets separately. Table 6.15 details the coefficient of variation for the subset. Comparing these values to where all subjects are included, as detailed in Table 6.13, the results for the male subjects generally show a slight increase in the mean and standard deviation, while for the female subject a slight decrease is observed. The main exception to this is the combined total dead space where the opposite is observed. The results for all data were similar.

Table 6.16 and Figure 6.8 show the results when a comparison is made with the plethysmograph FRC. These results are more favourable for the female subjects. The standard deviations for all of the statistics are lower and there are no differences between the reruns of the bootstrap. The gradient is closer to unity and there is slightly less bias. The mean of the standard error of the estimate and the standard deviation

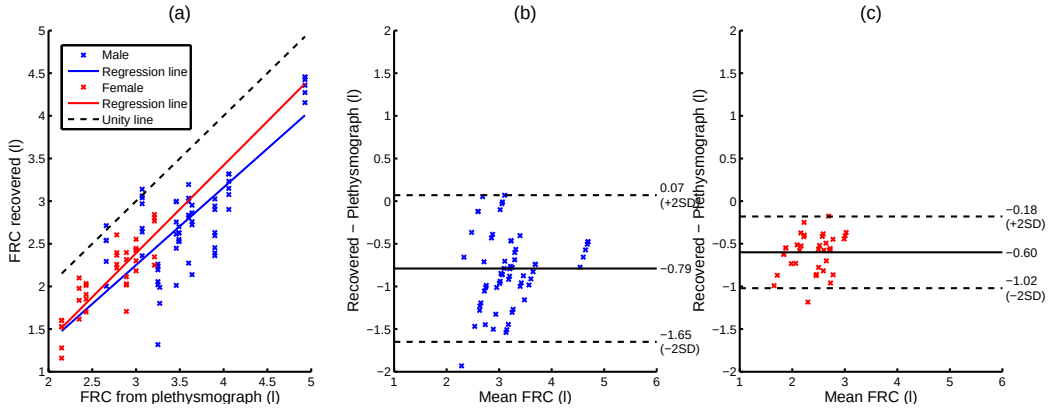


Figure 6.8: Plots of the results determined for both sexes: (a) scatter plot and Bland-Altman plot for (B) male and (c) female subjects. ABC SMC was used with Model 2 and the assumption of a standard man's body to determine the FRC. The regression line and Bland-Altman mean difference and limits of agreement are determined from the mean values in Table 6.16.

of the error are also about half that seen for the male results.

An attempt to investigate if the first test for a subject produced better results was made. A dataset was classified as the first dataset if it both had the earliest time stamp and also started with a concentration below 0.25%. However only eight datasets matched this criteria, which was considered to be too few to produce any meaningful results.

6.4.3.2 Genetic Algorithm

Table 6.17 details the coefficient of variation results when the genetic algorithm is used alone and when the results are refined using the Nelder-Mead simplex. Results are shown for Models 2 and 8 of the dead space. The volume change of the alveolar compartment due to the exchange of gas with the blood was not accounted for.

When using Model 2, there is little difference between the results of the genetic algorithm alone and when refinement using the Nelder-Mead simplex is included. When using Model 8 there is little difference in the results for the two dead space mixing compartments and their combined volume. However the end-expired alveolar and arterial blood flow results show more differences, especially in the case of the standard deviation of the end-expired alveolar volume which increases dramatically to above 42%.

Comparing the unrefined results for both models, the dead space results show less variation for Model 2. The mean of the end-expired alveolar results and FRC are approximately the same but the standard deviations for Model 2 are about 5% lower.

Model 2

		V_M	V_R	V'_A	$\dot{Q}_T$	V_D	FRC
Without refinement	All data	15.01±12.26	14.22±7.43	26.90±10.91	43.52±15.63	7.63±4.11	24.32±9.85
	Subset	15.40±12.51	14.01±7.33	27.76±10.23	43.94±15.96	7.66±4.06	25.15±9.19
With refinement	All data	15.41±12.84	14.27±7.58	27.06±10.32	44.95±15.83	7.71±4.30	24.03±9.04
	Subset	15.76±13.13	14.07±7.47	28.06±9.66	45.42±16.14	7.77±4.22	24.96±8.38

Model 8

		$V_{M,1}$	$V_{M,2}$	V'_A	$\dot{Q}_T$	V_D	FRC
Without refinement	All data	17.74±20.09	11.67±10.58	25.99±16.17	33.02±14.60	11.50±11.11	24.02±14.95
	Subset	18.68±20.53	12.15±10.69	27.16±15.68	33.89±14.54	11.83±11.32	25.13±14.47
With refinement	All data	16.70±19.80	11.77±10.09	28.31±42.06	35.75±18.52	11.82±10.86	24.55±31.84
	Subset	17.13±20.32	12.17±10.21	29.81±42.86	36.76±18.59	12.07±11.13	25.85±32.34

Table 6.17: Mean and standard deviation of the coefficient of variation (expressed as a percentage). The parameters determined using a genetic algorithm alone or followed by refinement using the Nelder-Mead Simplex, as detailed above. The body of a standard man was assumed.

The arterial blood flow rates are better for Model 8, with a mean approximately 10% lower but a similar standard deviation.

Table 6.18 shows the results for two repeats of the regression and error analysis. There are some differences between the two repeats of the regression analysis. This may be due to the assumptions made when choosing the measurement error ratio for Deming regression. Previously a coefficient of variation of 10% was assumed for the method being analysed, but the mean coefficient of variation is much higher here at approximately 25%. Therefore the measurement error ratio is adjusted to 25 (variance of method/variance of plethysmograph). Table 6.19 shows the results.

These results are more consistent between the two runs of the bootstrap and are thus likely to be a better representation of the true linear relationship. There are however still some differences in the results, which are likely to be due to the larger amount of variation seen. For Model 2 these new regression results have lower standard deviations when refinement by the Nelder-Mead simplex is included. The opposite is true for Model 8, with lower standard deviations when refinement is not used. Model 8 generally has higher standard deviations and more disparity between the two runs.

The newer regression analysis slightly favours Model 2. The mean gradient is closer to unity and has lower standard deviation. The mean standard error of the estimate is lower when refinement is used as the standard deviation. The standard deviation of the bias is lower, but the absolute mean value is about double that observed in other cases when refinement is used. The error analysis results, presented in Table 6.18, support those results. The bias is similar, with Model 2 with refinement having highest mean

Model 2

Nelder-Mead Simplex Refinement	Regression Analysis				Error	
	Gradient	Intersect(l)	SE(l)	Bias(l)	Mean(l)	SD(l)
No	1.09±0.46	-0.84±1.43	0.56±0.12	-0.51±0.24	-0.53±0.19	0.80±0.16
	1.10±0.55	-0.85±1.74	0.56±0.12	-0.51±0.24	-0.53±0.19	0.80±0.17
Yes	0.84±0.28	-0.42±0.90	0.45±0.10	-0.96±0.18	-0.92±0.16	0.64±0.14
	0.84±0.32	-0.42±1.00	0.45±0.10	-0.96±0.18	-0.93±0.16	0.64±0.14

Model 8

Nelder-Mead Simplex Refinement	Regression Analysis				Error	
	Gradient	Intersect(l)	SE(l)	Bias(l)	Mean(l)	SD(l)
No	0.67±0.88	0.65±2.78	0.57±0.16	-0.51±0.37	-0.43±0.22	0.85±0.31
	0.67±1.30	0.64±4.18	0.57±0.16	-0.52±0.38	-0.43±0.22	0.85±0.31
Yes	0.98±4.69	-0.50±14.75	0.55±0.27	-0.63±0.61	-0.58±0.23	0.84±0.49
	1.20±27.56	-1.17±84.31	0.55±0.27	-0.64±0.82	-0.58±0.23	0.83±0.49

Table 6.18: Results of the bootstrapped regression and error analysis. A genetic algorithm alone or followed by refinement using the Nelder-Mead Simplex was used to determine FRC.

absolute bias but the lowest standard deviation. The standard deviation of the error has the lowest mean and standard deviation in that case. However the generally high standard deviation throughout all the results does mean that these slight differences are unlikely to be significant.

6.4.3.3 Discussion

The results presented here clearly show that when using the standard man method to determine FRC the use of ABC SMC is required. The coefficient of variation results for FRC are significantly worse when the genetic algorithm is used. The source of this increase appears to be due to issues in determining the end-expired alveolar volume,

	Nelder-Mead Simplex Refinement	Gradient	Intersect(l)	SE(l)	Bias(l)
Model 2	No	0.83±0.31	0.01±0.96	0.56±0.13	-0.56±0.23
		0.82±0.31	0.04±0.96	0.56±0.13	-0.57±0.23
	Yes	0.72±0.25	-0.03±0.78	0.44±0.09	-0.98±0.19
		0.73±0.25	-0.03±0.78	0.44±0.10	-0.98±0.19
Model 8	No	0.56±0.52	0.99±1.60	0.55±0.15	-0.53±0.33
		0.55±0.54	1.03±1.64	0.55±0.15	-0.53±0.33
	Yes	0.68±0.60	0.46±1.85	0.56±0.30	-0.65±0.37
		0.68±0.60	0.46±1.82	0.56±0.30	-0.65±0.37

Table 6.19: Results of the bootstrapped regression analysis when the measurement error ratio is adjusted to 25. A genetic algorithm alone or followed by refinement using the Nelder-Mead Simplex was used to determine FRC.

where a large increase in the mean and standard deviation of the coefficient of variation is seen.

The comparison with the plethysmograph FRC confirms that the FRC is not determined well using a genetic algorithm. There is a substantial increase in the standard deviations determined via the bootstrap method for all of the statistics, even when the Deming measurement error ratio is adjusted to take account for the higher variation observed. In addition the mean for the two precision indicators increases. In a number of cases the mean gradient and bias are better than for ABC SMC, but given the large standard deviations this does not indicate that the genetic algorithm is performing any better than ABC SMC.

Instead of a poor performance by the genetic algorithm, another explanation for these results is that different models were used. The model used with ABC SMC took account of changes in the volume due to exchange of tracer gas with the blood, but the model used with the genetic algorithm did not. However this would seem to be an unlikely cause of these very different results. Only minor differences are seen in the fixed venous return case and the previous analysis by Barton *et al.* [25, 26]. Also subjective empirical observations of the raw output showed little difference. Further analysis could be carried out to ensure that this was not the cause.

Comparing the ABC SMC results with those for the fixed venous return highlights that some improvements may be possible using this method. The variation of the individual air compartment volumes is less, so, should determining them individually be important, this method then may produce better results. The precision indicators for FRC are slightly better, so should precision/repeatability be more important than linearity and bias, then this method may also be preferable. However this will be the case only if similar improvements are seen for Model 8, as the precision of the fixed venous return with Model 8 is higher. The use of ABC SMC with Model 8 may produce similar improvements and possibly others, however, given the computational requirements of ABC SMC, this is unlikely to result in a practical method, so no further analysis has been carried out here.

The separation into male and female subjects does not show that the standard man method favoured the male subjects. In fact the female results are better for FRC. The major improvement is the repeatability of the results with the mean and standard deviation of the coefficient of variation, standard deviation of the bootstrapped statistics, the mean standard error of the estimate and the mean standard deviation of the error

all being lower. The number of subjects within the two sets of analysis is small however and so these results should be treated with some caution.

This may be why the male results are less favourable. Alternatively it is possible that the $70kg$ standard man body compartments are actually more representative of female subjects than the male subjects. The compartments were devised by Mapleson in a paper published in 1973 [127], which were based on older sources. The population is likely to have changed since then. For example, the mean weight of a male in the United Kingdom is currently $84.1kg$ [16] compared to $73.9kg$ [169] in 1980, while for a female it is currently $69.7kg$ compared to $62.0kg$. It is likely that further increases in weight occurred between the original observation on which the compartment are based and 1980. Further investigations could be carried out to determine the most appropriate size for the body compartments in both male and female subjects, which may improve the results. Alternatively a link between the accuracy of the results and body mass index (BMI) could be investigated to determine if the body type is a more important factor than the sex. This could potentially be developed into a method of varying the size of the body compartments depending on the subjects BMI in order to get a more accurate measure of FRC.

The increase in the variation of the blood flow rate results may be because that parameter is compensating for the fact that the subject's body does not match that of the standard man. For example the subject may have a body which has a larger volume, in which case, reducing the blood flow rate is likely to result in changes in venous return which are closer to that within the subject's body, as reducing the blood flow is likely to slow the changes in the concentration of the body compartment in a similar way to an increased volume. A solution to this may be to include a parameter for the total volume of all body compartments. However the parameter estimation techniques may struggle to determine this value and the blood flow rate, as varying both parameters are likely to have similar effects on the observed data.

In addition to correcting for long term changes in the rise and fall of the venous return, deviation from the actual blood flow rate of the subject may also correct any differences in the phase of the venous return. The fact that variations in the blood flow rate may result in these complex forms of feedback may also result in a more complex parameter space. This may explain why ABC SMC produces better results than a genetic algorithm, as it has coped better with identifying the correct parameter values within this space.

There is a range of alternative models for the body that may improve the results obtained, such as the use of a single compartment to represent the body [229]. Inferring additional parameters of the standard man body compartments may also result in improvements, but care would be required to choose easily identifiable parameters and to keep the total number of inferred parameters low. Further investigation into convergence and other criteria used by the genetic algorithm could improve the results obtained by that method, while still being relatively computationally efficient.

6.4.4 Summary

Within this section several different methods have been analysed in an attempt to discover which worked best in human data. Here a summary of those findings will be made and a single method and model selected for further investigation.

The bootstrapping approach does allow information about the likely results from many repeated tests providing the most likely value of the statistical indicators (the mean result) and their variation (the standard deviation). However the way missing data are dealt with has the potential to influence the results. Above, if a set of missing data was selected then that subject was left out of that iteration. This has the potential to favour a subject with more datasets over those with fewer and so potentially to bias the results. An alternative would be to ensure that all subjects were included in each iteration by always randomly selecting one of their datasets. However this would potentially bias the results to the datasets which would be chosen more often, because a subject had few datasets.

Another alternative would be to ensure that each subject had the same number of datasets by randomly selecting them before the bootstrapping procedure. However this would reduce the data used and so the power of the analysis. It could also increase the variation between runs. Ultimately each method has its advantages and disadvantages and further analysis could be carried out using the alternative methods proposed to see if the approach used changed the results.

Errors in the Deming measurement error ratio could cause errors in the regression analysis. Its value was determined on the basis of a 5% coefficient of variation for the plethysmograph, based on information from the literature which gave a range of 3.5 to 6.8% [92]. The actual variation is unknown as only one plethysmograph reading was available for each subject. As the coefficient of variation of the plethysmograph is an approximate value, the measurement error ratio has not been adjusted for each analysis, with a value of 4 generally used. Also normally the assumed 10% coefficient of variation for the method being analysed was within one standard deviation of the

mean. For the one occasion when this was not the case, results for a value close to the mean were also presented.

In many ways the ABC SMC results for the fixed venous return and standard man method match the expectation from the simulated sensitivity analysis in Section 4.7. The end-expired alveolar volumes and FRC show about the same amount of variation. The trend in FRC is better determined by the fixed venous return method. As also expected, the variation in the blood flow rate is higher when using the standard man method. However as no alternative measure for the flow rate was available it was impossible to verify if, despite the higher variation, the average value was closer to the real value, as predicted. The exception is the absolute bias in FRC which was expected to be lower when using the standard man method, but the human results show that it is in fact lower for the fixed venous return method.

From the results above, use of the fixed venous return method would seem preferable, especially as the use of a genetic algorithm with the standard man method does not produce good predictions. This by no means rules out further investigation using this and similar methods, which could prove fruitful. The method does have the major advantage that the body does not need to be brought into equilibrium prior to the start of the test. However this may mean that repeated tests are not possible without a break to allow the tracer gas levels within the body to return to zero. This may in part be the cause of some of the less favourable results here.

Better defining the body compartments used, and perhaps adjusting them depending on the weight of the subject, could improve the results. However the complex feedback situation with regards to the blood flow rate may limit this approach. Firstly, because this may result in higher variation in the recovered flow rate, which may limit the ability to determine that parameter. Secondly, because allowing the body compartment size to change is unlikely to improve the results as both the volume of the compartment and the flow rate are unlikely to be easily identifiable together. Thirdly, because it may be what causes the poor results from the genetic algorithm.

Possible improvements are not confined to the standard man method. Better controlling the initial equilibrium process, potentially by automating it, could improve the results from the fixed venous return method. Further investigations, possibly using simulated data, could help bring the body to an equilibrium state in a timely manner. This could remove the requirement for the fixed venous return parameter with the mean sinewave concentration used instead, thus reducing the number of parameters to

be estimated and so potentially improving the results for the other parameters. However if true equilibrium (or something close enough to remove this parameter) could not be achieved, bringing the body closer to an equilibrium state should still improve the results, as the return concentration will be better approximated by a fixed value.

Although much discussion has been dedicated to the merits of the different methods, little comparison has been made between the breath-by-breath method and the other methods. Here it will be compared to the fixed venous return results without any offset correction. The dead space coefficient of variation results are similar to those seen when using the fixed venous return method. The other results all have a much larger standard deviation than seen when using any of the fixed venous return method and model combinations. The means for the end-expired alveolar volume and FRC are comparable to that seen for Model 2, but the Model 8 results outperform the breath-by-breath method. The mean pulmonary flow rate is, however, better for the breath-by-breath method than those values determined using the genetic algorithm with refinement using the Nelder-Mead simplex and similar to the ABC SMC results.

When comparing the breath-by-breath FRC results with plethysmograph results, the standard deviations of the bootstrap statistics are also found to be high: double that seen when using the fixed venous return method, except for the bias, but which is still higher. The two anomalous results may be the cause of this and also the high standard deviations seen in the coefficient of variation. Removing them may improve the standard deviation, however these anomalous results are a failure of the method rather than a poor dataset because similar anomalous results are not seen in any of the other methods. Dealing with the source of these anomalous results may improve the repeatability of the breath-by-breath method.

The mean results are at least on par and sometimes better than that seen for Model 2, but Model 8 has better mean values for all statistics. Thus in general the fixed venous return method seems to outperform the breath-by-breath method when using both models of the dead space, but definitely when using Model 8. This means that FRC is better determined by methods which include a model for the entire expired concentration signal than the breath-by-breath method which does not. Also these results point towards the use of Model 8 for robustly determining FRC here and so potentially being more sensitive to changes within ventilated patients.

As discussed above, the use of the fixed venous return method with Model 8 outperforms all the other methods when determining FRC. This was not the case in the bench lung data, where Model 2 and Model 8 performed similarly with each having

different pros and cons. A reason for this might be that the bench lung was devised to mimic a lung with a transitory dead space and alveolar compartment, which favours models with a transitory compartment.

In reality the human lung is not made of these two defined compartments, but has a less well defined transition from gas transport dominated by convection to gas transport dominated by diffusion. This is what causes the sigmoidal change in the concentration observed in the expired breath of real data, rather than the abrupt change which would occur with a transitory dead space and alveolar compartment. It would seem that Model 8 models the realities of the human lung to a better degree and so produces better results.

The absolute bias of the Model 8 results when compared to the plethysmograph FRC is less than $60ml$. It has been assumed to be better than the higher absolute bias seen for other methods, which can be above $500ml$. However it is not clear what bias should be expected. Many authors have reported that the plethysmograph overestimates FRC when compared to other methods [44, 56, 95, 125, 138]. This was presented as the probable cause for these differences in similar analysis of the breath-by-breath method by collaborators [159]. This may be the source of the error and errors of up to 10% have previously been estimated to occur within healthy subjects due to trapped air and pressure differences in the lung [44].

However as discussed in Section 2.5.1, the major source of overestimation is pressure losses that occur due to obstructive disease. Some more recent studies have only reported an overestimate in patients with obstructive disease, with healthy subjects reporting minimal disparity [95, 125, 138, 146]. The data used here are for young subjects (age 30 or below) none of whom currently suffer from respiratory diseases. Also, once F12 is excluded from the subset, only smokers of less than one cigarette a day remain.

The bench lung results support a systematic underestimate by the breath-by-breath method and Model 2, rather than an error solely due to the plethysmograph. The results for those methods have a consistently higher underestimate for the end-expired alveolar volume, the main constituent of FRC, than when Model 8 was used. The underestimate is in some cases as large as that observed here, but not always. Thus some doubt is raised as to whether the bias is all down to an overestimate by the plethysmograph.

It is also unclear the effect that nitrous oxide dissolved in the tissue of the lungs will have. Oxygen can be directly taken up by the tissue of the lung, without passing

through the bloodstream [120] and so it would seem plausible that nitrous oxide can be exchanged in that way. Based on the compartments of the standard man (see Table 2.1), the tissues within the lung would contain approximately the equivalent nitrous oxide as contained in $0.25l$ of the lung at steady state. If this 'acts' although it is within the lung then an overestimate of approximately $0.25l$ would be expected.

In reality how the nitrous oxide within the lung tissue is perceived, as being in the lung or part of the buffer within the tissue of the body, will depend on the rate at which it is exchanged with gas in the lung. This is likely to be somewhere in between the instantaneous exchange required to act as additional lung gas and the slower exchange which ultimately happens with body compartments. It is therefore likely to cause an overestimate, but by less than $0.25l$.

Assuming that the bias should be around zero, a possible second source for the error seen when using the breath-by-breath method and Model 2 is the offset between the signals. A positive offset was shown to reduce the FRC volume recovered in Section 4.6.2. A offset of $100ms$ produced an underestimate in the FRC of just under $0.5l$, approximately that seen here. The median offset recovered for Model 2 when it was included as a parameter was $90ms$, which may suggest this as a possible source. The interquartile range of that result is high and the result for Model 1 was different. However if this is a true representation of the offset then it could suggest that the offset is the cause of the bias observed when using the breath-by-breath method and Model 2. Also it may suggest that Model 8 is more robust when offsets are present; however no simulations have yet been done to support that observation.

The expected bias is unclear and is dependent on a number of factors. However there are a number of factors that suggest that the bias seen for Model 8 is likely to be closer to the actual bias than those seen in the other methods. Even if the bias results are discounted the Model 8 fixed venous return results are still better based on the other statistics, with better linearity and precision/repeatability.

These methods are all based on a homogeneous model of the lung and as such any inhomogeneities caused by restrictive or obstructive lung disease are likely to affect the prediction of FRC. However, as previously discussed, there is potential to use the decrease in FRC in those cases as indications of these underlying issues allowing clinicians to alter their care, particularly when applied to ventilated patients. Here data from spontaneously breathing subjects has been used, however in many ways this is more complex than the more controlled breathing patterns found in ventilated patients;

thus these methods should produce more favourable results if applied to mechanical ventilated patients.

Further investigation into the standard man and other methods may improve the results. Also, Model 8 and the parameter estimation techniques used with it could be improved. For example the use of three mixing compartments could improve the results. However the fixed venous return method with Model 8 is the best performing dead space model and method combination investigated here. Thus it will be used in comparisons with other published methods, below.

6.5 Comparison with Published Results

Thus far the focus has been on the comparison of the different methods using the data collected via the sinusoidal delivery system. However it is also important to assess the performance of the proposed methods when compared to other available methods. Within this section the FRC results from the fixed venous return method with Model 8, detailed above, will be compared to data available within the literature.

Comparisons will be made with the performance of the plethysmograph, the helium dilution technique and the LUFU device. Comparisons with the Engstrom Carestation would also be of interest, however fewer published data are available and those available are for ventilated patients only [41, 53]. There are also many other comparisons that could be made to other techniques that have be proposed over the years [41, 228], however these are felt to be the most clinically relevant due to their current clinical use or recently studies showing they have potential to improve the care of ventilated patients [96, 97, 108].

6.5.1 Repeatability

The coefficient of variation for FRC using the fixed venous return method with Model 8 is 7.42±2.83 for all of the data, which results in a 95% confidence interval of 6.01 to 8.83% assuming that the results are normally distributed. For the subset the results were 7.62±2.49%, which results in a confidence interval of 6.34 to 8.90%.

In response to the NHLBI workshop (discussed earlier in Section 2.5) a paper was published summarising the repeatability of the various lung volume measurements [92]. The paper summarises previous findings on the coefficient of variation for FRC when determined by plethysmograph and helium dilution. As previously stated, the mean coefficient of variation for the plethysmograph ranged from 3.5 to 6.8% for the nine sources included within the review.

Method	Repeatability coefficient (l)
Fixed venous return (all data)	0.71
Fixed venous return (subset)	0.72
Plethysmograph	0.72
Helium dilution	0.74
LUFU device	0.60

Table 6.20: Repeatability coefficients for the fixed venous return method and the other methods investigated by Maisch *et al.* [125].

These studies include both healthy and unhealthy subjects, however the report does comment that the health of the subjects within the studies did not appear to influence the results, with inter-laboratory factors appearing to be more important. In fact an extensive 1982 study (which was not included in the review) carried out 50 repeated measurements in 512 healthy non-smokers, with a reported value towards the upper limit of 6.7% [205].

The reported mean coefficient of variation for helium dilution ranged from 4.9 to 10.4% [92] based on three sources within the review. One study reported both the lower and upper limit of the range, comparing the results for 40 laboratory workers tested in a laboratory setting (4.9%) and 38 coal miners in the 'field' (10.4%) [121]. This partly highlights the issues of using the helium dilution technique outside the lung function laboratory and in some ways the higher value could be considered the better comparator here. The participants were selected from the working population and not excluded based on their smoking habits or having current respiratory symptoms.

Another study reported a value of 7.6% for 40 healthy volunteers, which included 27 smokers [110]. The final study was actually for a multi-breath Nitrogen washout test [64], which it is assumed was included as the authors felt it was representative of the repeatability of helium dilution. This study reported 8.1±8.1% for 14 healthy subjects (including seven smokers). This would result in a 95% confidence interval for the mean of 3.4 to 12.8% for the healthy subjects, assuming a normal distribution for the results. The distribution is however unlikely to be normal, given that a coefficient of variations below 0% is not possible and that is within one standard deviation of the mean.

An alternative measure is the repeatability coefficient [37, 39]. This was calculated using the one way analysis of variation (ANOVA) method, as described by Bland [39] and is given in Table 6.20. Maisch *et al.* [125] carried out a similar analysis on a similar cohort of healthy subjects: 20 healthy volunteers with a mean age of 27±2 years, including 10 current smokers. This cohort is of similar age (on average 4 years

Method	Approach /Source	Bias	95%CI Bias	Limits of Agreement	95% CI Lower Limit	95% CI Upper Limit
Fixed venous return	All sets	-0.06	-0.18 to 0.07	-0.66 to 0.54	-0.86 to -0.42	0.32 to 0.75
	Five sets	-0.05	-0.16 to 0.06	-0.69 to 0.58	-0.85 to -0.48	0.38 to 0.77
Helium dilution	Heinze *et al.*	-0.3	-0.5 to 0.1	-1.3 to 0.7	-1.7 to -0.9	0.3 to 1.0
	Maisch *et al.*	-0.03	-0.23 to 0.17	-0.89 to 0.83	-1.24 to -0.54	0.48 to 1.18
LUFU	Heinze *et al.*	-0.1	-0.4 to 0.2	-1.3 to 1.1	-1.8 to -0.9	0.7 to 1.6
	Maisch *et al.*	-0.43	-0.68 to -0.18	-1.51 to 0.65	-1.95 to -1.07	0.21 to 1.09

Table 6.21: Comparison of the bias and limits of agreement of the fixed venous return, helium dilution and LUFU methods when compared to the plethysmograph. Where known the values given are for the method FRC minus plethysmograph FRC, however this direction is not given by Maisch *et al.* therefore results are presented as given in the paper.

older), but has a higher number of smokers than, the cohort investigated earlier in this chapter. The study carried out duplicated results for each of the methods it analysed: plethysmograph, helium dilution and the LUFU device. The repeatability coefficients[2] for these results are also given in Table 6.20.

6.5.2 Comparisons with Plethysmograph

Two papers were identified which made a comparison of FRC determined by plethysmograph with helium dilution and the LUFU device using a similar methodology and which also had a cohort of similar participants [95, 125]. Heinze *et al.* [95] presented results for 23 healthy subjects with a mean age of 35±10 years, which is an average age of 12 years older than for the data presented in this chapter. No details are given about the smoking habits of these subjects. A single measurement was taken by each method from each subject. The second was the same paper by Maisch *et al.* [125] that was used above for comparison of the repeatability coefficients. The paper does not make it clear how the duplicated results were treated. If an average was taken this may not be a direct comparison, however this should only favour the results in the paper.

The results from these papers are detailed in Table 6.21, along with the results of the fixed venous return method. Not all the values shown were given by the authors, so they were calculated from the mean and standard deviation, as described by Bland and Altman [37]. For the fixed venous return method the bias and limits of agreement are mean values determined by the bootstrapping approach described in Section 6.3

[2]The repeatability coefficients given by Maisch *et al.* [125] do not match what would be expected given the standard deviation of the difference quoted in the paper, i.e it is not twice that standard deviation as it should be according to Bland and Altman [37]. Therefore twice the standard deviation of the difference is given for comparison in Table 6.20.

Method	Source	Direction	Bias	Lower Limit	Upper Limit
Helium dilution	Heinze *et al.*	As paper	0.050	**0.003**	0.437
	Maisch *et al.*	As paper	0.795	0.247	0.140
		Inverted	0.436	0.389	0.076
LUFU	Heinze *et al.*	As paper	0.776	**0.009**	**0.020**
	Maisch *et al.*	As paper	**0.007**	**0.000**	0.628
		Inverted	**0.000**	0.982	**0.000**

Table 6.22: p-values for the two sample z-test comparing the bias and the limits of agreement for the fixed venous return method to those given in the literature. The direction of the bias is not given by Maisch *et al.* therefore results are also given for the inverted bias.

(all sets approach). The 2.5 and 97.5 percentile of the 100,000 iterations are used as an estimate of the confidence intervals [45, 156, 157].

In order to ensure that the missing data were not biasing the results, the analysis was also carried out on five randomly selected datasets for each subject, ensuring no missing data but potentially reducing the power of the study (five sets approach). These results are also shown in Table 6.21 and are the results for the first time the analysis was run. Each run will produce different results depending on the initial datasets chosen. Two additional runs produced similar results, which were not sufficiently different to affect the conclusions drawn below. The subset was still used here as reducing the analysis to only two repeats could significantly reduce the power of the study and is likely to produce widely differing results between runs.

In order to determine if the difference in the bias and limits of agreement are significant, a two sample z-test [39] was carried out between the results for all the sets approach and the results are presented in Table 6.22. This test is used due to the available information for the results derived from the literature and the fact that the sample size is unclear for the bootstrap method. The standard deviation of the 100,000 bootstrap iterations is used as an approximation of the standard error [45, 156, 157] for the fixed venous return method and the standard error is determined as described by Bland and Altman [37] for the literature results. The test assumes a large sample size, meaning that the p-values may contain some error as the sample sizes are below 23 here. As the direction of the difference is unknown for the Maisch *et al.* results, the p-values for the inverted mean bias are also given.

6.5.3 Discussion

Showing if one method is superior to another (to a particular significance level) is a common statistical task and in a number of cases detailed above this has been possible.

Showing that two methods perform to the same standard, which may also be useful here, is a more difficult task. However it may at least be possible to show that a particular method is not any worse given the evidence available. In the case of the repeatability results, where little information about the variability is available in the literature, the comparison of the mean results will be much more subjective and it may not be possible to draw firm conclusions.

When comparing the results for the repeatability measures, it is possible to say that the fixed venous return method is at best able to achieve similar results to the plethysmograph with some of the reported mean values within the confidence intervals. Also the repeatability coefficient is approximately the same, although this is based on only two repeats for the plethysmograph. However it would seem quite likely that the plethysmograph has better repeatability, given that the mean results for the fixed venous return method are above the upper limit of the range and that some of the lower mean results in the literature are from large same sizes.

This is what would be expected, considering that the plethysmograph is generally thought to produce the most repeatable measurements for FRC [92]; hence it has been used as the benchmark in many studies (including here) even though it is thought to overestimate FRC in certain situations. Ultimately the newly proposed methods do not seek to replicate the plethysmograph, but instead to be practical enough to be implemented throughout the hospital and then to be precise enough to positively influence clinical care.

The mean coefficient of variation for the fixed venous return method is within the reported range for the helium dilution technique. The confidence intervals also fit within the range. In addition, the only reported value which is lower than the mean fixed venous return result is the 4.9% result. A similar repeatability coefficient is reported. This would suggest that the repeatability of the fixed venous return method is likely to be no worse than that of the helium dilution technique. The repeatability coefficient for the LUFU device is better than that for the fixed venous return method by just under $100ml$. However the fact that this is a single measure and the lack of any information on the confidence intervals make it hard to draw any conclusions.

The results for the comparisons with the plethysmograph provide some more certain conclusions. There is no evidence that the fixed venous return method is any worse than either helium dilution or the LUFU device based on the information presented. The fixed venous return method has a tighter confidence interval around zero difference. Also in the one situation where the mean bias is better, for helium dilution results of

Maisch *et al.* [125], the z-test shows that this is not significant with a p-value above 0.4 for both possible directions. Thus there is no evidence that the accuracy of the fixed venous return method is any worse.

For the limits of agreement the mean values for the fixed venous return method are closer to zero than the limits of any other method. The one exception to that is if the Maisch *et al.* results were inverted for the LUFU device, which would mean that the lower limit could be $10ml$ better for the LUFU device. However the z-test results show that this is not significant, with a p-value of 0.982. Also the opposite limit for the LUFU device is $970ml$ worse. The fact that the fixed venous return results are no worse is confirmed by the outer values of the confidence intervals for the upper and lower limits of agreement, which are closer to zero for the fixed venous return method than for any of the other methods. Thus the precision of the fixed venous return method is no worse than the other two methods.

Shifting focus to whether there is any evidence that the fixed venous return method is any better than the other methods, in terms of bias, when the results are compared to the Heinze *et al.* [95] results for helium dilution the bias is on the bounds of being better with a p-value of 0.05. However given the possible errors due to the approximations and assumptions already discussed and the fact that the values in the paper are only given to one decimal point, it would be sensible not to draw any conclusions from that value. Also the Maisch *et al.* results do not support an improvement. For the LUFU device, the Maisch *et al.* results suggest that the fixed venous return method has a significantly worse bias, but Heinze *et al.'s* results do not support this fact.

For the limits of agreement, the Heinze *et al.* results for helium dilution suggest that the lower limit is significantly worse and so the precision is very likely to be significantly worse. However this is not supported by results from Maisch *et al.*. For the LUFU device both the upper and lower limits are significantly better for the fixed venous return method when compared to the Heinze *et al.* results. For the Maisch *et al.* results only the lower limit (or upper limit when inverted) is significantly better, however this is only because the bias is significantly different, which has brought the upper limit close to that of the fixed venous return upper limit. Given that and the fact that the p-value for the lower limit is less than 0.001, it would seem fair to conclude that that the results from both papers show the precision of the LUFU device is significantly worse than that of the fixed venous return method.

This may seem slightly at odds with the better repeatability results for the LUFU device, however that result is unlikely to be significant. Even if it was, the within

subject repeatability of the results could be good but this does not necessarily mean that the value of the measurement is anything like the actual volume, so the between subject precision and accuracy compared to a benchmark result can still be worse. The above results are based on the all sets approach, which may be affected by the way outliers are dealt with, but the five sets results are similar and would be likely to result in the same conclusions.

These results show that the fixed venous return method is no worse than the helium dilution technique when compared to the plethysmograph. Both the Heinze *et al.* and Maisch *et al.* papers used a similar observation (but determined in a different way) to promote the use of the LUFU device[3]. This is a reassuring result as the helium dilution technique is widely used in clinical practice [44]. Thus the accuracy and precision of the fixed venous return method is also likely to be good enough to aid clinical decisions.

In addition, it was shown that the accuracy is no worse than the LUFU device and that the precision is better. Use of this device to aid clinical decisions within ventilated patients has been reported [96, 97, 108]. Thus it is likely that the fixed venous return method could be used to similarly assist in such decision making. If the results presented here were translated into ventilated patients the use of these methods would be likely to perform more robustly when determining FRC than the LUFU device. An ability to more robustly determine FRC is likely to pick up more subtle changes in ventilated patients' lungs and so better assist clinicians decision making. Ultimately more research would be required to show whether this is true.

These comparisons do also have some limitations. One of the difficulties is that the effect the methodological differences between the studies will have. For example, the repeatability data used within this chapter were within session repeatability, however some of the studies are for multiple sessions. For example, in the laboratory worker/coal miner study [121] the repeats were carried out anything from one day to several weeks apart. In theory this is only likely to result in worse results, but this is likely to be only one of many methodological differences that could affect the comparison.

In the comparisons with the plethysmograph there is also potential for the differences in the accuracy and precision of the different plethysmographs used to affect the the results, especially given that the plethysmograph measurements are not significantly more accurate or precise than the other methods. These errors would be easier to control in a study using a single device.

[3]These papers also compared the LUFU device directly to the helium dilution technique, which obviously cannot be done for the fixed venous return method as no helium dilution results are available.

Differences between the cohort of subjects used also have the potential to affect the results. In general all the results compared are for healthy subjects, apart from the coefficient of variation for the plethysmograph where there is evidence that the health of the subjects has a limited effect. None of the cohorts specifically excluded smokers, but where information is given the number of smokers is greater than for the results presented in this chapter. Also once F12 is excluded the subset only has two smokers, which both smoke one or less a day. The cohort is younger than for all the other studies that give an age of their cohort. However the average age in the Maisch *et al.* study is not much different. Both these factors have the potential to produce more favourable results for the fixed venous return method.

The fact that the same data were used to choose the best method also has the potential to produce favourable results for the fixed venous return method. As the number of methods attempted is increased the probability of finding one that fits the data and produces good results also increases. However this model may not actually model the underlying process better, rather just fit the data available. There are a number of factors that lower that risk here. The number of different model/method combinations tried is relatively low; 14 in total when counting all method and model combinations. Also a number of these use very similar models or methods. Although there are only 18 subjects, including all the repeats there are 98 datasets in total within the subset. The models are based on what is thought to happen physiologically within the body and they have shown similar results in separate data from the bench lung. No additional parameters, other than those discussed and presented, have been fitted using the data.

Results from another dataset did not produce such favourable results which could also cause some concern. However that data had a range of forcing periods and it was not clear what effect the forcing period had on the results. The results for a $3min$ forcing period seemed to show less bias and variability than many of the other forcing periods. This may be the cause of the worse results, however there were too few observations to conclude anything about the effect of the forcing period or to investigate the $3min$ results alone.

The lower variability of the results at $3min$ is similar to the effect seen in the bench lung data, however those results showed the most significant reduction in variability at $3min$ when using Models 1 and 2. Further investigations into the variability at different forcing periods or even with different forcing waves may provide further insight into reducing the variability more generally.

In general, all the limitations described can be overcome by further investigations allowing direct comparison in the same cohort. In any further investigations direct comparison with the helium dilution technique would be preferable if an attempt to show similar accuracy and precision is being made. Further testing in patients with some form of lung disease would also be required to ensure that the method is more generally applicable to determining FRC in all states of health.

6.6 Key Findings

The key findings of this chapter are:

- Using the fixed venous return method, Model 8 and a genetic algorithm followed by the Nelder-Mead simplex produces FRC results in humans which are better than any other methods proposed within this thesis.
- When compared to plethysmography, these results were no worse in terms of accuracy and precision than when using the Helium Dilution technique.
- These results also have significantly better precision than the LUFU device.
- The FRC results are also better than those for the breath-by-breath method, which only models particular points within the breath.
- The addition of an offset parameter does not improve the FRC results.
- The standard man method results only begin to approach the precision of the fixed venous return method when ABC SMC is used.
- ABC SMC is too computationally expensive for practical use in determining FRC within real clinical settings.

6.7 Conclusion

These findings confirm what was previously seen in the bench lung data, i.e. that the inclusion of mixing within the dead space seems to be an important factor in determining volumes of the lung more accurately. They in fact go further, favouring a model of the dead space that includes no transitory compartment. This would suggest that moving away from the two compartment classic 'balloon on a straw' model, which models an abrupt transition from convection to diffusion, to one which models the process as a more gradual transition from convection to diffusion (like many of the more complex forward models), is likely to improve the results of applications that use models to infer information via the inverse process. This observation has potential to feed back into the models for determining lung ventilation-volume homogeneity

proposed earlier in the thesis, improving their ability to determine information from human data.

The comparison to the Helium Dilution technique shows that the sinusoidal delivery system combined with the methods developed in this thesis can determine FRC to an accuracy and precision level comparable to that found in current clinical practice. In addition the comparison to the LUFU device shows the proposed method to be significantly more precise. The LUFU device has already shown some potential to improve ventilated patient outcomes, thus if these unventilated healthy patient results translate into ventilated patients it is likely to be possible to provide a more robust measure of FRC. This could be more effective at guiding clinical care and so have the potential to further improve the outcomes of ventilated patients.

Chapter 7

Conclusions

7.1 Summary

For many years researchers have attempted, and also succeeded, to use tracer gases to measure FRC or the ventilation-volume distribution of the lungs. However none of the methods developed have seen routine clinical use in ventilated patients, where such measures have the potential to guide clinical decisions and so improve the patient outcomes. There are various reasons for this. One reason is the amount of data available, which can not be provided at a higher rate than we breathe.

An aim of the thesis was to develop a model which predicts the concentration of the entire expired breath in an attempt to extract more information from the available data. Various models were developed and have been applied to the data. These have generally performed better than the main benchmark method, the breath-by-breath method, which relies on models of the lung that only analyse the concentrations of particular points in the breath (the end-expired concentration) or average values over the breath (for example the mixed expired concentration). Although the assumptions on which this method rely are different, the improvements are likely in part to be because modelling the entire breath has allowed more information to be extracted from the data.

One of the assumptions made by the breath-by-breath method, and many of the previous models, is that the end-expired concentration equals the alveolar concentration. This assumption is likely to only introduce small errors for insoluble gases where the concentration in the alveoli does not change during expiration. However for soluble gases larger errors are likely to be introduced, as the concentration does change. This assumption has generally not been made throughout this thesis and is also likely to have led to the improvements seen, particularly in human data.

The other main development is in the compartments used to model the lung. Modelling the lung as two compartments, a dead space and an alveolar compartment, originates from Bohr's work. Using that concept of how the lung is made up has allowed many useful methods to be devised. However in reality the lung is more complex. This is reflected in many of the more complex models of the lung, which model a more gradual transition from gas transport dominated by convection to gas transport dominated by diffusion. Here the challenge has been how to model these more complex modes of gas transport without making the model too complex and parameter heavy that it is impossible to infer information from the data.

This was done by adding mixing compartments to the upper airways, with various configurations of transitory and mixing compartments investigated. Ultimately in human data a series of three mixing compartments (Model 8) proved to be the best at inferring FRC. Two initial smaller fixed volume compartments, which could be thought of as representing the dead space, followed by a larger expendable compartment which allowed gas exchange, the alveolar compartment. This model was not the worst performing model in water-displacement bench lung data, but models which include a transitory compartment within the dead space lung performed better than they did in human data. This suggests that the bench lung may be limited in the ability to produce data which matches the realities of a human lung, possibly mimicking Bohr's concept of the lung too much.

The possibility of improving the results by correcting offset between the concentration and flow signal was also investigated. Although some promising results were seen in bench lung data when using the ABC SMC technique with models that include transitory compartments within the dead space (Models 1 and 2), similar results were not replicated for more computationally efficient least squares optimization methods, when using Model 8 or in human data. The results, including those in simulated data, suggest that this is because the additional offset is not easily identifiable even in data not corrupted by noise. Thus correcting for the offset in this way is generally unlikely to be successful in improving results.

In addition to outperforming the breath-by-breath method, Model 8 (the best of the proposed methods) also performs better than the LUFU device in terms of precision and is no worse in terms of accuracy. It has been shown that this device has the potential to positively influence decisions about the care of ventilated patients. In addition, this method has an accuracy and precision comparable to that of the Helium dilution technique, which is routinely used within lung laboratories to aid clinical decisions.

Together these results show that this method should be able to determine FRC at a precision and accuracy which is capable of providing valuable information to clinicians.

In addition to FRC, this thesis also investigated the ability to determine information about the ventilation-volume distribution of the lung. Two broad approaches were investigated: directly inferring information about the ventilation-volume distribution and using the difference in the volumes determined using the homogeneous model of the lung to indirectly infer information about the distribution. Directly inferring information about the uni-modal log-normal distribution was possible and inferring some information from multi-compartment lungs was also possible in simulated data. However inferring information about bi-modal log-normal distributions was not successful, even in the absence of noise. As the distributions found within the human lung, particularly unhealthy lungs, are likely to be more complex than can be represented as uni-modal log-normal distributions, using this method is unlikely to be successful.

Two methods that assumed a homogeneous lung model were investigated; a previously proposed method (the homogeneity index) and a newly proposed method (the inter-breath variation method), which forms a distribution of the homogeneous alveolar volumes recovered at each breath. In simulated data both were shown to be susceptible to noise, but the inter-breath variation method was less susceptible than the homogeneity index. It also performed better when inferring information about the homogeneity of the multi-compartment bench lung. These better results were however very tentative. This was partly due to issues determining the make up of the dead space independently of the alveoli. Also the results from the bench lung may not be a true reflection of the homogeneities found within the lung. This does however also highlight both the difficulties of inferring this kind of information from tracer gas concentrations and the need to extract all the available information from the data.

In addition to investigating least squares optimization methods, which are much more established, the more untested ABC SMC method was investigated. Two of the main motivations for doing so were to investigate the possibility of using the uncertainty in the recovered alveolar volumes as a measure of homogeneity and also to use its combined model selection and parameter estimation capabilities. ABC SMC failed to deliver as expected on both counts. ABC SMC seems to be limited in both situations by a lack of clear criteria for choosing a population (and by proxy a tolerance) at which to stop, with lower tolerances not necessarily providing the best approximation to the posterior distribution. The investigations also showed that in the presence of noise

the model selection tended to choose models with more parameters, rather than those which gave rise to data, seeming not to properly penalise models with more parameters.

Estimating parameters using ABC SMC does seem to be successful. Although not always outperforming least squares optimisation methods, ABC SMC often performed better, especially when the parameter space appeared to be more complex. There is, however, a significant computational cost to using ABC SMC. This in part may be overcome by better determining the tolerance schedule. However the number of function evaluations required is several orders of magnitude higher. This means that the wall clock times are significantly higher, in the order of days rather than hours, even with the substantial efficiencies that using C instead of Matlab brings. Making sure that convergence has been achieved is also difficult as later iterations (populations) normally take longer to evaluate, so evaluating additional iterations to ensure convergence becomes infeasible. These computational constraints are likely to mean the use of ABC SMC is not practical when making such parameter estimations. A number of possible issues have been identified with regards the width of the parameter distributions determined by ABC SMC and with the model selection process. These generally seem to be related to determining a suitable point at which to stop the algorithm and methods for determining that point would be required before ABC SMC can be used in the ways originally envisaged at the beginning of this research.

These issues and limitations mean the use of least squared optimisation methods is likely to be preferable in any future bedside device. The results show methods which only find local minima are unlikely to be successful, as there has been evidence of more than one minima existing within the parameter space. However the improvement seen when using the Nelder-Mead simplex after the genetic algorithm suggests that method has not converged to sufficient accuracy. The motivation behind using a genetic algorithm was its ability to cope with discrete parameters. This functionality is no longer required as accounting for time offsets between the two signals did not produce better results. Thus the use other global optimisation methods could be investigated and may provide better performance.

7.2 Conclusion

Within this thesis a model for the entire expired concentration signal has been proposed and evaluated. In order to produce better results in bench lung and human data, it was necessary to move beyond a model of the lung that includes an anatomical dead space which contains no mixing to a model that introduces mixing within the upper airways.

In doing so it has been possible to better determine the FRC using data from the sinusoidal delivery system developed by collaborators, producing results comparable to devices used in routine clinical practice. Further use and development of these models has the potential to improve the ability to determine FRC and information about lung homogeneity. These should hopefully improve clinical diagnosis and so improve patient outcomes.

7.3 Future Work

It was possible to show that the methods and models developed in the thesis performed well in 18 healthy subjects. Further validation in more subjects and those with lung disease would be required in any further developments of this technique. Any additional validation in new data would provide reassurance that the method is generally applicable and not a statistical anomaly from evaluating many methods in the same dataset. Further investigation into the effects of forcing period may also help understand the source of the variation within results. It may also be possible to better determine if a forcing period of $3mins$ does produce more repeatable results and if so why.

It may also be possible to improve the results further by improving the model used. The decision to investigate Model 8 further was based on the ABC SMC bench lung results, which favoured a two mixing compartment dead space. However, the human data shows different results than those previously seen in the bench lung for Models 2 and 8, so evaluating models with different numbers of mixing compartments in human data may provide different results. It may be the case that more compartments are required in patients with lung disease, as more complex forms of mixing occurs within their lungs. However, as previously discussed, too many small compartments may cause the response of the model to tend towards that of a transitory dead space compartment.

Model 8 could be seen as a development towards a model more like the trumpet model. An alternative to developing the compartment model further, would be to start with the trumpet model and simplify it to a few key parameters, thus potentially allowing the inverse process to be used. This is likely to also involve simplifying the model in such a way as to ensure the parameters can be determined within a reasonable time frame. However this may be no more of a challenge than already overcome with this thesis for the models with a transitory compartment.

The simulations and inhomogeneous bench lung analysis were based on Model 2. Re-evaluating these results in light of the findings in human data may provide further

insights into improving the results for FRC and lung homogeneity measures. There is also potential that the relative volumes determined for the three mixing compartments within Model 8 may provide information about the health of the lung. However the coefficient of variation of individual compartments was not as favourable as the combined volumes, which may reduce the reliability of a measure derived in this way.

Although the use of the water-displacement bench lung was valuable in developing the methods and models used within the thesis, the difference between the bench lung results and the human results when comparing dead space models highlights the limitations of the data from this device. Similarly, so does the probability that the multi-compartment bench lung does not reflect the level of homogeneity found within the human lung. This highlights the need to test further in the human subjects (or possibly animals), however relying on human testing alone does have cost implications and will always be limited by the lack of true gold standard for both FRC and homogeneity.

The wall clock time for producing the results is currently in the region of one to two hours. This is unlikely to be satisfactory in any final product and future work would be required to improve this time. However improvements to times less than five minutes should be easy to achieve. Implementation of the models in C has seen execution times of the models reduce, with times one hundredth of the Matlab execution time seen. A prove of concept was implemented using the Nelder-Mead simplex in Python with the C version of Model 2 and completed in less that $5min$ for every case. Also the genetic algorithm easily benefits from any increase in the number of available processors, due to its parallel nature.

Investigations into the best way to bring the whole body into equilibrium in a timely manor are likely to improve the technique. For example it may be possible to have a longer period at which the concentration of nitrous oxide is kept high, causing a higher concentration in the lungs and so bringing the body to the mean sinewave concentration quicker. The concentration would then be lowered at a later point so as to equilibrate the lung compartments before starting the sinewave. There is however a possibility of the body compartments overshooting the equilibrium point. A sensitivity analysis could be used to determine the optimum amount of time and concentration to use during the equilibrium process, so as to bring the body closer to equilibrium without drastically overshooting. Even without this analysis an automated equilibrium process is very likely to improve the repeatability of the results, removing the random nature of the current manual process.

An alternative would be to use an insoluble gas (or gas with low solubility) where bringing the body into equilibrium would not be an issue as no exchange (or very little exchange) would happen. However the availability of suitable sensors may prevent this. In which case, the model could be adapted for oxygen where the availability of sensors is less problematic and equilibrium is not an issue. Two reasons for this are: oxygen has been shown to act much like a insoluble gas [89, 226–228] and anyway the body will be at equilibrium state already (assuming the mean concentration is matched to that of the previous supply) .

Better aligning the flow and concentration sensors are very likely to improve the results acquired, as demonstrated in simulated data. However the results from this thesis would suggest any effort should be focused towards correcting the alignment at source.

Some issues with the ability of ABC SMC to perform a model selection or determine the uncertainty of parameters have been highlighted. Further investigations into how best to determine the ideal population or tolerance at which to stop would be essential for any future general applications, however the computational requirements of ABC SMC are still likely to exclude it from use as described within this thesis.

As already discussed, further investigations into other global optimisation methods may produce more robust results and may converge more readily than a genetic algorithm. For example a multi start approach [201] could be used with the Nelder-Mead simplex as the local optimiser from each starting point. Use of a derivative based local optimiser could improve the efficiency of the algorithm, however discontinuities in the derivatives may cause issues, especially when using real data.

Appendix A

Symbols and Abbreviations

A.1 List of Symbols

a	Constant as defined at that point the text
b	Constant as defined at that point the text
c	Constant as defined at that point the text
d	Delay
e	Euler's number/Napier's constant
f	Frequency
k	Number of parameters
l	Distance from the terminal alveoli
m	Model prediction
n	Total number of an item (implicit or described in the text)
t	Time
s	Cross-sectional area of the conducting airway
u	Convection velocity
x	Point on the x-axis
$\bar{x}$	Mean of x
y	Point on the y-axis
z	Point on the z-axis
γ	Decay constant
ϵ	ABC tolerance
λ	Partition coefficient/Oswald coefficient
σ	Standard deviation
τ	Time constant
ω	Angular frequency
C	Concentration in the blood phase
D	Diffusion coefficient
E	Error function
F	Fractional concentration in the gas phase
F'	End-expired fractional concentration
F^*	Corrected fractional concentration
$\dot{F}$	1st derivative of the fractional concentration
$\ddot{F}$	2nd derivative of the fractional concentration
G	Gas source term
I	Homogeneity index

K	Capacitance (of the lung)
P	Partial pressure
$\dot{Q}$	Blood flow
R	Resistance (of the lung)
S	Total cross-sectional area of the airway
T	Time point within a breath
V	Volume
V^*	Equivalent volume
V'	End-expired volume
$\dot{V}$	Gas flow
W	Concentration reduction factor
Δt	Breath duration
ΔF	Amplitude of the fractional concentration
Λ	Dummy variable of integration

A.2 List of Subscripts

a	Arterial
$\bar{a}$	Mixed-arterial
b	Blood
c	Carrier
i	Compartment number
j	Gas number
l	Data point number
m	Modulated
n	Breath number
p	Peak number
t	Tissue
v	Venous
$\bar{v}$	Mixed-venous
v_D	Caused by the dead space volume
v_R	Caused by the dead transitory compartment volume
A	Alveolar
$\bar{A}$	Mixed-alveolar
Ad	Alveolar distribution
CO_2	Carbon dioxide (tracer gas)
D	Dead space
DF	Fowler dead space
E	Expired
$\bar{E}$	Mixed-expired
EM	Expired at the mouth
ED	Expired through the dead space
EDS	Expired into the dead space
ER	Expired from the transitory compartment
I	Inspired
$\bar{I}$	Mixed-inspired
IA	Inspired into the alveoli
ID	Inspired through the dead space

IDS	Inspired into the dead space
IM	Mean inspired
IR	Inspired from the transitory compartment
LM	Last air past the mouth
M	Mixed compartment
P	Pulmonary
R	Transitory compartment
RMS	Root mean square
S	Shunt
T	Tidal
Td	Tidal distribution
Δ	Deviation

A.3 List of Abbreviations

ABC	Approximate Bayesian computation
ALI	Acute lung injury
ARDS	Acute respiratory distress syndrome
ATS	American thoracic society
BMI	Body mass index
COPD	Chronic obstructive pulmonary disease
CT	Computed tomography
DDE	Delay differential equation
EELV	End-expiratory lung volume
ERS	European respiratory society
ERV	Expiratory reserve volume
FRC	Functional residual capacity
ICU	Intensive care unit
IRV	Inspiratory reserve volume
KS	Kolmogorov-Smirnov
LUFU	(A device produced by Dräger Medical)
MCMC	Markov chain Monte Carlo
MIGET	Multiple inert gas elimination technique
MRI	Magnetic resonance imaging
NHLBI	National heart, lung and blood institute
ODE	Ordinary differential equation
PEEP	Positive end-expiratory pressure
RMS	Root mean squared
RV	Residual volume
SD	Standard deviation
SE	Standard error
SMC	Sequential Monte Carlo
TLC	Total lung capacity
VALI	Ventilator-associated lung injuries
VC	Vital capacity
VILI	Ventilator-induced lung injuries

References

[1] *Data sheet: MC series of digital mass flow control.* Alicat Scientific. URL www.alicat.com/documents/sales_sheets/MC-Series-Sales-Sheet.pdf. [Accessed 29-09-2014].

[2] *Data Sheet: PHASEIN Multigas Technologies – Which one?* PHASEIN AB, Danderyd, Sweden. URL www.masimo.com/pdf/phasein/IRMA-ISA-brochure-web.pdf. [Accessed 18-08-2015].

[3] Physiome Project. URL physiomeproject.org. [Accessed 20-02-2015].

[4] *Data sheet: THOR D-22.* THOR Research & Development. URL www.thormed.com/index.php?page=products&id=flow22. [Accessed 29-09-2014].

[5] Standardization of definitions and symbols in respiratory physiology. *Federation proceedings*, 9(3):602–5, sep 1950.

[6] Pulmonary terms and symbols. A report of the ACCP-ATS Joint Committee on Pulmonary Nomenclature. *CHEST*, 67(5):583–593, may 1975.

[7] *White paper: Getting started with LabVIEW virtual instruments.* National Instruments, 2008. URL www.ni.com/white-paper/7001/en/. [Accessed 04-09-2014].

[8] *BS EN ISO 80601-2-55: Medical electrical equipment – Particular requirements for the basic safety and essential performance of respiratiry gas monitors.* The British Standards Institution, London, 2011.

[9] Novel Drug Approvals for 2011, 2011. URL www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugInnovation/ucm285554.htm. [Accessed 28-05-2016].

[10] *Data Sheet: CARESCAPE Respiratory Modules.* GE Healthcare, 2012. URL www.gehealthcare.co.uk/en-gb/products/categories/patient_monitoring. [Accessed 18-08-2015].

[11] Novel Drug Approvals for 2012, 2012. URL www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugInnovation/ucm336115.htm. [Accessed 28-05-2016].

[12] Novel Drug Approvals for 2013, 2013. URL www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugInnovation/ucm381263.htm. [Accessed 28-05-2016].

[13] Novel Drug Approvals for 2014, 2014. URL www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugInnovation/ucm429247.htm. [Accessed 28-05-2016].

[14] *Matlab R2014a documentation.* MathWorks (UK), 2014. URL http://www.mathworks.co.uk/help/matlab/. [Accessed 11-09-2014].

[15] Novel Drug Approvals for 2015, 2015. URL

www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugInnovation/ucm430302.htm. [Accessed 28-05-2016].

[16] Health indicators for the United Kingdom and its constituent countries based on the 2013 to 2014 European Health Interview Survey. Technical report, Office for National Statistics, London, UK, 2015.

[17] Novel Drug Approvals for 2016, 2016. URL www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugInnovation/ucm483775.htm. [Accessed 28-05-2016].

[18] H. Akaike. Information theory and an extension of the maximum likelihood principle. In B. Petrov and F. Csaki, editors, *Second international symposium on information theory*, pages 267–281, Budapest, 1973. Academiai Kiado.

[19] S. K. Alford and E. A. Hoffman. Measurement of Pulmonary Vascular Structure and Pulmonary Blood Distribution by Multidetector-Row Computed Tomography and Magnetic Resonance Imaging Techniques. In J. X. J. Yuan, J. G. Garcia, C. A. Hales, S. Rich, S. L. Archer, and J. B. West, editors, *Textbook of Pulmonary Vascular Disease*, pages 511–524. Springer, Boston, MA, 2011.

[20] D. G. Altman and J. M. Bland. Measurement in medicine: The analysis of method comparison studies on JSTOR. *Journal of the Royal Statistical Society. Series D (The statistician)*, 32(3):307–317, 1983.

[21] R. Arieli. Mass spectrometer for respiratory research. *Respiratory physiology & neurobiology*, 170(2):183–4, feb 2010.

[22] R. Arieli and H. D. Van Liew. Corrections for the response time and delay of mass spectrometers. *Journal of applied physiology*, 51(6):1417–22, dec 1981.

[23] K. I. Arita, S. M. Lewis, and C. Mittman. Bolus technique for assessing distribution of inspired gas during tidal breathing. *Journal of applied physiology*, 64(2):681–8, feb 1988.

[24] C. P. Barnes, D. Silk, X. Sheng, and M. P. H. Stumpf. Bayesian design of synthetic biological systems. *Proceedings of the National Academy of Sciences of the United States of America*, 108(37):15190–5, sep 2011.

[25] S. A. Barton, A. M. S. Black, and C. E. W. Hahn. Dynamic models and solutions for evaluating ventilation, perfusion, and mass transfer in the lung - Part I: The models. *IEEE transactions on biomedical engineering*, 35(6):450–457, jun 1988.

[26] S. A. Barton, A. M. S. Black, and C. E. W. Hahn. Dynamic models and solutions for evaluating ventilation, perfusion, and mass transfer in the lung - Part II: Analog solutions. *IEEE transactions on biomedical engineering*, 35(6):458–465, jun 1988.

[27] J. H. Bates, G. K. Prisk, T. E. Tanner, and A. E. McKinnon. Correcting for the dynamic response of a respiratory mass spectrometer. *Journal of applied physiology*, 55 (3):1015–22, sep 1983.

[28] A. M. Batterham. Commentary on Bias in Bland-Altman but not Regression Validity Analyses. *Sportscience*, 8:47–49, 2004.

[29] M. A. Beaumont. Approximate Bayesian Computation in evolution and ecology. *Annual review of ecology, evolution, and systematics*, 41(1):379–406, dec 2010.

[30] M. A. Beaumont, W. Zhang, and D. J. Balding. Approximate Bayesian Computation in population genetics. *Genetics*, 162(4):2025–2035, dec 2002.

[31] M. A. Beaumont, R. Nielsen, C. P. Robert, J. Hey, O. O. Gaggiotti, L. Knowles,

A. Estoup, M. Panchal, J. Corander, M. Hickerson, S. Sisson, N. Fagundes, L. L. Chikhi, P. Beerli, R. Vitalis, C. Jean-Marie, J. Huelsenbeck, M. Foll, Z. Yang, F. Rousset, D. J. Balding, L. Excoffier, P. Be, R. Vitalis, J.-M. Cornuet, J. Huelsenbeck, M. Foll, Z. Yang, F. Rousett, D. J. Balding, and L. Excoffier. In defence of model-based inference in phylogeography. *Molecular ecology*, 19(3):436–446, jan 2010.

[32] A. D. Belegundu and T. R. Chandrupatla. *Optimization Concepts and Applications in Engineering*. Cambridge University Press, 2011.

[33] R. A. Berk. *Regression Analysis: A Constructive Critique*. SAGE Publications, 2004.

[34] G. Bertorelle, A. Benazzo, and S. Mona. ABC as a flexible framework to estimate demography over space and time: some cons, many pros. *Molecular ecology*, 19(13): 2609–25, jul 2010.

[35] I. G. Bikker, W. Holland, P. Specht, C. Ince, and D. Gommers. Assessment of ventilation inhomogeneity during mechanical ventilation using a rapid-response oxygen sensor-based oxygen washout method. *Intensive care medicine experimental*, 2(1):14, apr 2014.

[36] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

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

[38] J. M. Bland and D. G. Altman. Comparing methods of measurement: why plotting difference against standard method is misleading. *Lancet*, 346(8982):1085–7, oct 1995.

[39] M. Bland. *An Introduction to Medical Statistics*. Oxford University Press, Oxford, 3rd edition, 2000.

[40] P. Bogacki and L. Shampine. A 3(2) pair of Runge - Kutta formulas. *Applied Mathematics Letters*, 2(4):321–325, 1989.

[41] L. M. Brewer, J. A. Orr, M. R. Sherman, E. H. Fulcher, and B. A. Markewitz. Measurement of functional residual capacity by modified multiple breath nitrogen washout for spontaneously breathing and mechanically ventilated patients. *British journal of anaesthesia*, 107(5):796–805, nov 2011.

[42] B. S. Brook, C. M. Murphy, D. Breen, A. W. Miles, D. G. Tilley, and A. J. Wilson. Theoretical models for the quantification of lung injury using ventilation and perfusion distributions. *Computational and mathematical methods in medicine*, 10(2):139–154, apr 2009.

[43] O. Brovko, D. Wiberg, L. Arena, and J. Bellville. The extended Kalman filter as a pulmonary blood flow estimator. *Automatica*, 17(1):213–220, jan 1981.

[44] R. Brown, D. E. Leith, and P. L. Enright. Multiple breath helium dilution measurement of lung volumes in adults. *The European respiratory journal*, 11(1):246–55, jan 1998.

[45] P. C. Bruce. *Introductory Statistics and Analytics: A Resampling Perspective*. Wiley, 2015.

[46] J. X. Brunner and D. R. Westenskow. How the rise time of carbon dioxide analysers influences the accuracy of carbon dioxide measurements. *British journal of anaesthesia*, 61(5):628–38, nov 1988.

[47] J. X. Brunner, G. Wolff, G. Cumming, and H. Langenstein. Accurate measurement of N2 volumes during N2 washout requires dynamic adjustment of delay time. *Journal of applied physiology*, 59(3):1008–1012, sep 1985.

[48] K. S. Burrowes and M. H. Tawhai. The effect of lung orientation on functional imaging

of blood flow. In A. Manduca and X. P. Hu, editors, *Progress in Biomedical Optics and Imaging, Medical Imaging 2007: Physiology, Function, and Structure from Medical Images*, page 65110Z, San Diego, 2007. International Society for Optics and Photonics.

[49] K. S. Burrowes, A. J. Swan, N. J. Warren, and M. H. Tawhai. Towards a virtual lung: multi-scale, multi-physics modelling of the pulmonary system. *Philosophical transactions. Series A (Mathematical, physical, and engineering sciences)*, 366(1879): 3247–63, sep 2008.

[50] K. S. Burrowes, A. R. Clark, A. Marcinkowski, M. L. Wilsher, D. G. Milne, and M. H. Tawhai. Pulmonary embolism: predicting disease severity. *Philosophical transactions. Series A (Mathematical, physical, and engineering sciences)*, 369(1954):4255–77, nov 2011.

[51] S. Chatterjee and J. S. Simonoff. *Handbook of Regression Analysis*. John Wiley & Sons, 2013.

[52] L. Chen and L. Brochard. Lung volume assessment in acute respiratory distress syndrome. *Current opinion in critical care*, 21(3):1, mar 2015.

[53] D. Chiumello, M. Cressoni, M. Chierichetti, F. Tallarini, M. Botticelli, V. Berto, C. Mietto, and L. Gattinoni. Nitrogen washout/washin, helium dilution and computed tomography in the assessment of end expiratory lung volume. *Critical care*, 12(6):R150, jan 2008.

[54] J. Clausen. Measurement of absolute lung volumes by imaging techniques. *The European respiratory journal*, 10(10):2427–31, oct 1997.

[55] J. Clausen and J. Wagner. Consensus statement on measurements of lung volumes in humans, 2003. URL `www.thoracic.org/adobe/lungvolume.pdf`. [Accessed 04-06-2015].

[56] I. J. Cliff, A. H. Evans, C. F. A. Pantin, and D. R. Baldwin. Comparison of two new methods for the measurement of lung volumes with two standard methods. *Thorax*, 54 (4):329–333, apr 1999.

[57] L. A. Clifton, A. D. Farmery, and C. E. W. Hahn. A non-invasive method for estimating lung function. In *The sixth international conference on condition monitoring and machinery failure prevention technologies*, pages 509–518, Dublin, Ireland, 2009.

[58] L. A. Clifton, D. A. Clifton, C. E. W. Hahn, and A. D. Farmery. Assessment of lung function using a non-invasive oscillating gas-forcing technique. *Respiratory physiology & neurobiology*, 189(1):174–82, oct 2013.

[59] A. Coates, R. Peslin, D. Rodenstein, and J. Stocks. Measurement of lung volumes by plethysmography. *European respiratory journal*, 10(6):1415–1427, 1997.

[60] J. H. Comroe and W. S. Fowler. Lung function studies. *The American journal of medicine*, 10(4):408–413, apr 1951.

[61] P. J. Cornbleet and N. Gochman. Incorrect least-squares regression coefficients in method-comparison analysis. *Clinical chemistry*, 25(3):432–8, mar 1979.

[62] A. Cournand, E. D. Baldwin, R. C. Darling, and D. W. Richards. Studies on intrapulmonary mixture of gases. IV. The significance of the pulmonary emptying rate and a simplified open circuit measurement of residual air. *The journal of clinical investigation*, 20(6):681–9, nov 1941.

[63] C. P. Criée, S. Sorichter, H. J. Smith, P. Kardos, R. Merget, D. Heise, D. Berdel, D. Köhler, H. Magnussen, W. Marek, H. Mitfessel, K. Rasche, M. Rolke, H. Worth, and R. A. Jörres. Body plethysmography–its principles and clinical use. *Respiratory*

medicine, 105(7):959–71, jul 2011.

[64] A. Cutillo, R. Perondi, M. J. Turiel, M. Egger, S. Watanabe, and A. D. Renzetti. Reproducibility of multibreath nitrogen washout measurements'. *The American review of respiratory disease*, 124(4):505–507, 1981.

[65] B. S. Dahler-Eriksen, J. F. Lassen, P. H. Petersen, E. D. Lund, T. Lauritzen, and I. Brandslund. Evaluation of a near-patient test for C-reactive protein used in daily routine in primary healthcare by use of difference plots. *Clinical chemistry*, 43(11): 2064–75, nov 1997.

[66] R. C. Darling, A. Cournand, D. W. Richards, and B. Domanski. Studies on intrapulmonary mixture of gases. V. Forms of inadequate ventilation in normal and emphysematous lungs, analyzed by means of breathing pure oxygen. *Journal of clinical investigation*, 23(1):55–67, jan 1944.

[67] J. C. Davies, S. Cunningham, E. W. F. W. Alton, and J. A. Innes. Lung clearance index in CF: a sensitive marker of lung disease severity. *Thorax*, 63(2):96–7, feb 2008.

[68] P. Del Moral, A. Doucet, and A. Jasra. Sequential Monte Carlo samplers. *Journal of the Royal Statistical Society: Series B (Statistical methodology)*, 68(3):411–436, jun 2006.

[69] J. D'Errico. MATLAB CentralFile Exchange - fminsearchbnd, 2005. URL `uk.mathworks.com/matlabcentral/fileexchange/`. [Accessed 27-01-2016].

[70] A. Dushianthan, M. P. W. Grocott, A. D. Postle, and R. Cusack. Acute respiratory distress syndrome and acute lung injury. *Postgraduate medical journal*, 87(1031):612–22, sep 2011.

[71] L. A. Engel. Intraregional Gas Mixing and Distribution. In L. Engel and M. Paiva, editors, *Gas Mixing and Distribution in the Lung*, volume 25 of *Lung Biology in Health and Disease*, chapter 7, pages 287–358. Dekker, New York, 1 edition, mar 1985.

[72] J. W. Evans. The gas washout determination under a symmetry assumption. *The bulletin of mathematical biophysics*, 32(1):59–63, mar 1970.

[73] J. W. Evans and P. D. Wagner. Limits on VA/Q distributions from analysis of experimental inert gas elimination. *Journal of applied physiology*, 42(6):889–898, jun 1977.

[74] R. Fagard and J. Conway. Measurement of cardiac output: Fick principle using catheterization. *European heart journal*, 11 Suppl I:1–5, dec 1990.

[75] S. Fain, M. L. Schiebler, D. G. McCormack, and G. Parraga. Imaging of lung function using hyperpolarized helium-3 magnetic resonance imaging: Review of current and emerging translational methods and applications. *Journal of magnetic resonance imaging*, 32(6):1398–408, dec 2010.

[76] S. B. Fain, F. R. Korosec, J. H. Holmes, R. O'Halloran, R. L. Sorkness, and T. M. Grist. Functional lung imaging using hyperpolarized gas MRI. *Journal of magnetic resonance imaging*, 25(5):910–23, may 2007.

[77] A. Farmery. *Interrogation of the Cardiopulmonary System with Inspired Gas Tension Sinusoids*. Thesis, University of London, 2008.

[78] A. D. Farmery and C. E. W. Hahn. Response-time enhancement of a clinical gas analyzer facilitates measurement of breath-by-breath gas exchange. *Journal of applied physiology*, 89(2):581–589, aug 2000.

[79] W. Fowler. Lung function studies; the respiratory dead space. *The American journal of physiology*, 154(3):405–416, sep 1948.

[80] M. Frigge, D. C. Hoaglin, and B. Iglewicz. Some Implementations of the Boxplot. *The American statistician*, 43(1):50–54, feb 1989.

[81] O. Gajic, S. I. Dara, J. L. Mendez, A. O. Adesanya, E. Festic, S. M. Caples, R. Rana, J. L. St Sauver, J. F. Lymp, B. Afessa, and R. D. Hubmayr. Ventilator-associated lung injury in patients without acute lung injury at the onset of mechanical ventilation. *Critical care medicine*, 32(9):1817–24, sep 2004.

[82] J. L. Garfield, N. Marchetti, J. P. Gaughan, R. M. Steiner, and G. J. Criner. Total lung capacity by plethysmography and high-resolution computed tomography in COPD. *International journal of chronic obstructive pulmonary disease*, 7:119–26, jan 2012.

[83] D. J. Gavaghan and C. E. W. Hahn. A mathematical evaluation of the alveolar amplitude response technique. *Respiration physiology*, 102(1):105–120, oct 1995.

[84] D. J. Gavaghan and C. E. W. Hahn. A tidal breathing model of the forced inspired inert gas sinewave technique. *Respiration physiology*, 106(2):209–221, nov 1996.

[85] D. M. Gomez, W. A. Briscoe, and G. Cumming. Continuous distribution of specific tidal volume throughout the lung. *Journal of applied physiology*, 19(4):683–692, jul 1964.

[86] S. Gonem, A. Scadding, M. Soares, A. Singapuri, P. Gustafsson, C. Ohri, S. Range, C. E. Brightling, I. Pavord, A. R. Horsley, and S. Siddiqui. Lung clearance index in adults with non-cystic fibrosis bronchiectasis. *Respiratory research*, 15(1):59, jan 2014.

[87] F. S. Grodins, J. Buell, and A. J. Bart. Mathematical analysis and digital simulation of the respiratory control system. *Journal of applied physiology*, 22(2):260–276, feb 1967.

[88] M. Haas, B. Hamm, and S. M. Niehues. Automated lung volumetry from routine thoracic CT scans: How reliable is the result? *Academic radiology*, 21(5):633–8, may 2014.

[89] C. E. W. Hahn. Oxygen respiratory gas analysis by sine-wave measurement: a theoretical model. *Journal of applied physiology*, 81(2):985–97, aug 1996.

[90] C. E. W. Hahn and A. D. Farmery. Gas exchange modelling: No more gills, please. *British journal of anaesthesia*, 91(1):2 –15, jul 2003.

[91] C. E. W. Hahn, A. M. S. Black, S. A. Barton, and I. Scott. Gas exchange in a three-compartment lung model analyzed by forcing sinusoids of N2O. *Journal of applied physiology*, 75(4):1863 –1876, oct 1993.

[92] J. Hankinson, J. Stocks, and R. Peslin. Reproducibility of lung volume measurements. *European respiratory journal*, 11(3):787–790, mar 1998.

[93] S. S. Haykin. *Communication Systems*. Wiley, Chichester, 4th edition, 2001.

[94] G. Hedenstierna. The recording of FRC is it of importance and can it be made simple? *Intensive Care Medicine*, 19(7):365–366, jul 1993.

[95] H. Heinze, B. Schaaf, J. Grefer, K. Klotz, and W. Eichler. The accuracy of the oxygen washout technique for functional residual capacity assessment during spontaneous breathing. *Anesthesia and analgesia*, 104(3):598–604, mar 2007.

[96] H. Heinze, B. Sedemund-Adib, M. Heringlake, T. Meier, and W. Eichler. Changes in functional residual capacity during weaning from mechanical ventilation: a pilot study. *Anesthesia and analgesia*, 108(3):911–5, mar 2009.

[97] H. Heinze, W. Eichler, J. Karsten, B. Sedemund-Adib, M. Heringlake, and T. Meier. Functional residual capacity-guided alveolar recruitment strategy after endotracheal suctioning in cardiac surgery patients. *Critical care medicine*, 39(5):1042–9, may 2011.

[98] M. P. Hlastala. Multiple inert gas elimination technique. *Journal of applied physiology*, 56(1):1 –7, jan 1984.

[99] W. G. Hopkins. Bias in Bland-Altman but not regression validity analyses. *Sportscience*, 8:42–46, 2004.

[100] A. R. Horsley. Lung clearance index in the assessment of airways disease. *Respiratory medicine*, 103(6):793–9, jun 2009.

[101] A. R. Horsley, P. M. Gustafsson, K. A. Macleod, C. Saunders, A. P. Greening, D. J. Porteous, J. C. Davies, S. Cunningham, E. W. F. W. Alton, and J. A. Innes. Lung clearance index is a sensitive, repeatable and practical measure of airways disease in adults with cystic fibrosis. *Thorax*, 63(2):135–40, feb 2008.

[102] A. R. Horsley, K. Macleod, R. Gupta, N. Goddard, and N. Bell. Enhanced photoacoustic gas analyser response time and impact on accuracy at fast ventilation rates during multiple breath washout. *PloS one*, 9(6):e98487, jan 2014.

[103] M. Hubert and E. Vandervieren. An adjusted boxplot for skewed distributions. *Computational statistics & data analysis*, 52(12):5186–5201, aug 2008.

[104] S. A. Jaliwala, R. E. Mates, and F. J. Klocke. An efficient optimization technique for recovering ventilation-perfusion distributions from inert gas data. Effects of random experimental error. *The journal of clinical investigation*, 55(1):188–92, jan 1975.

[105] A. A. Jalowayskfr, S. M. Lewis, I. Young, N. Olmsted, and I. R. Harwood. Distribution of Specific Ventilation in Cystic Fibrosis. *Respiration*, 43(4):249–257, 1982.

[106] K. J. K. Kanhai, H. Strijdhorst, J. C. Pompe, W. P. J. Holland, E. G. J. Ong, and G. J. Puppels. Noninvasive monitoring of nonshunted pulmonary capillary blood flow in the acute respiratory distress syndrome. *Critical care medicine*, 28(4):1059–1067, apr 2000.

[107] K. S. Kapitan. Information content of the multibreath nitrogen washout: effects of experimental error. *Journal of applied physiology*, 68(4):1621–1627, apr 1990.

[108] J. Karsten, T. Meier, and H. Heinze. Bedside-measurements of electrical impedance tomography and functional residual capacity during positioning therapy in a case of acute respiratory failure. *Applied cardiopulmonary pathophysiology*, 15:81–86, 2011.

[109] R. E. Kass and A. E. Raftery. Bayes Factors. *Journal of the American Statistical Association*, 90(430):773–795, feb 1995.

[110] K. Kauppinen-Walin, A. R. Sovijärvi, A. Muittari, and A. Uusitalo. Determination of functional residual capacity with 133-xenon radiospirometry. Comparison with body plethysmography and helium spirometry. Effect of body position. *Scandinavian journal of clinical and laboratory investigation*, 40(4):347–54, jun 1980.

[111] P. Komlosi, J. L. Benjamin, and T. A. Altes. Hyperpolarized Gas MRI in Pediatric Lung Disease. In P. Garcia-Peña and R. P. Guillerman, editors, *Pediatric Chest Imaging*, pages 513–521. Springer, 3rd edition, 2014.

[112] J. S. Krouwer. Why Bland-Altman plots should use X, not (Y+X)/2 when X is a reference method. *Statistics in medicine*, 27(5):778–80, feb 2008.

[113] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright. Convergence Properties of the Nelder–Mead Simplex Method in Low Dimensions. *SIAM journal on optimization*, 9(1):112–147, jan 1998.

[114] A. R. Lambert, P. O'Shaughnessy, M. H. Tawhai, E. A. Hoffman, and C.-L. Lin. Regional deposition of particles in an image-based airway model: large-eddy simulation

and left-right lung ventilation asymmetry. *Aerosol science and technology*, 45(1):11–25, jan 2011.

[115] M. Lambert, A. Metcalfe, D. Walker, and M. Leonard. *Engineering Modelling and Analysis*. Taylor & Francis Group, London, 2011.

[116] S. Leonhardt and B. Lachmann. Electrical impedance tomography: the holy grail of ventilation and perfusion monitoring? *Intensive care medicine*, 38(12):1917–29, dec 2012.

[117] S. M. Lewis, J. W. Evans, and A. A. Jalowayski. Continuous distributions of specific ventilation recovered from inert gas washout. *Journal of applied physiology*, 44(3): 416–423, mar 1978.

[118] S. M. Lewis, D. Z. D'Argenio, G. A. Bekey, and C. Mittman. Optimal inputs for parameter determination of inert gas washout from the lung. *Respiration physiology*, 50(1):111–27, oct 1982.

[119] J. Liepe, C. Barnes, E. Cule, K. Erguler, P. Kirk, T. Toni, and M. P. H. Stumpf. ABC-SysBio–approximate Bayesian computation in Python with GPU support. *Bioinformatics*, 26(14):1797–9, jul 2010.

[120] S. A. Loer, T. W. Scheeren, and J. Tarnow. How much oxygen does the human lung consume? *Anesthesiology*, 86(3):532–7, mar 1997.

[121] R. G. Love, M. D. Attfield, and K. D. Isles. Reproducibility of pulmonary function tests under laboratory and field conditions. *British journal of industrial medicine*, 37 (1):63–9, feb 1980.

[122] S. C. Luijendijk, A. Zwart, A. M. van der Kooij, and W. R. de Vries. Evaluation of alveolar amplitude response technique for determination of lung perfusion in exercise. *Journal of applied physiology*, 50(5):1071–8, may 1981.

[123] D. J. C. MacKay. *Information Theory, Inference and Learning Algorithms*. Cambridge University Press, 2003.

[124] P. Magee and M. Tooley. *The Physics, Clinical Measurement and Equipment of Anaesthetic Practice for the FRCA*. Oxford University Press, 2nd edition, 2011.

[125] S. Maisch, S. H. Boehm, D. Weismann, H. Reissmann, M. Beckmann, B. Fuellekrug, A. Meyer, and J. Schulte Am Esch. Determination of functional residual capacity by oxygen washin-washout: a validation study. *Intensive care medicine*, 33(5):912–6, may 2007.

[126] W. W. Mapleson. An electric analogue for uptake and exchange of inert gases and other agents. *Journal of applied physiology*, 18(1):197–204, jan 1963.

[127] W. W. Mapleson. Circulation-time models of the uptake of inhaled anaesthetics and data for quantifying them. *British journal of anaesthesia*, 45(4):319–34, apr 1973.

[128] J.-M. Marin, P. Pudlo, C. P. Robert, and R. J. Ryder. Approximate Bayesian computational methods. *Statistics and computing*, 22(6):1167–1180, oct 2011.

[129] P. Marjoram, J. Molitor, V. Plagnol, and S. Tavare. Markov chain Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences of the United States of America*, 100(26):15324–8, dec 2003.

[130] F. J. Massey. The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of the American Statistical Association*, 46(253):68–78, apr 1951.

[131] F. Milite, D. J. Lederer, J. A. Weingarten, P. Fani, A. M. Mooney, and R. C. Basner. Quantification of single-breath underestimation of lung volume in emphysema.

Respiratory physiology & neurobiology, 165(2-3):215–20, feb 2009.

[132] R. F. Miller and M. J. O'Doherty. Pulmonary nuclear medicine. *European journal of nuclear medicine*, 19(5), may 1992.

[133] R. R. Miller, N. R. Macintyre, R. D. Hite, J. D. Truwit, R. G. Brower, and A. H. Morris. Point: should positive end-expiratory pressure in patients with ARDS be set on oxygenation? Yes. *Chest*, 141(6):1379–82, jun 2012.

[134] R. R. Miller, N. R. Macintyre, R. D. Hite, J. D. Truwit, R. G. Brower, and A. H. Morris. Rebuttal to: should positive end-expiratory pressure in patients with ARDS be set based on oxygenation? No. *Chest*, 141(6):1384, jun 2012.

[135] R. R. Mitchell. Incorporating the gas analyzer response time in gas exchange computations. *Journal of applied physiology*, 47(5):1118–1122, nov 1979.

[136] R. R. Mitchell, R. M. Wilson, L. Holzapfel, A. M. Benis, D. Sierra, and J. J. Osborn. Oxygen wash-in method for monitoring functional residual capacity. *Critical care medicine*, 10(8):529–33, aug 1982.

[137] R. R. Mitchell, R. M. Wilson, and D. Sierra. ICU monitoring of ventilation distribution. *International journal of clinical monitoring and computing*, 2(4):199–206, jan 1986.

[138] C. Mottram. *Ruppel's Manual of Pulmonary Function Testing*. Elsevier Health Sciences, 10th edition, 2013.

[139] T. Nakamura, T. Takishima, T. Okubo, T. Sasaki, and H. Takahashi. Distribution function of the clearance time constant in lungs. *Journal of applied physiology*, 21(1): 227–232, jan 1966.

[140] J. A. Nelder and R. Mead. A simplex method for function minimization. *The computer journal*, 7(4):308–313, jan 1965.

[141] C. J. Newth, P. Enright, and R. L. Johnson. Multiple-breath nitrogen washout techniques: including measurements with patients on ventilators. *The European respiratory journal*, 10(9):2174–85, sep 1997.

[142] N. Nielsen, J. G. Nielsen, and A. R. Horsley. Evaluation of the impact of alveolar nitrogen excretion on indices derived from multiple breath nitrogen washout. *PloS one*, 8(9):e73335, jan 2013.

[143] H. Noguchi, Y. Ogushi, I. Yoshiya, N. Itakura, and H. Yamabayashi. Breath-by-breath VCO2 and VO2 required compensation for transport delay and dynamic response. *Journal of applied physiology*, 52(1):79–84, jan 1982.

[144] J. F. Nunn. *Nunn's Applied Respiratory Physiology*. Elsevier, 4th edition, 1993.

[145] R. E. Nye. Theoretical limits to measurement of uneven ventilation. *Journal of applied physiology*, 16(6):1115–24, nov 1961.

[146] C. R. O'Donnell, A. A. Bankier, L. Stiebellehner, J. J. Reilly, R. Brown, and S. H. Loring. Comparison of plethysmographic and helium dilution lung volumes: Which is best for COPD? *Chest*, 137(5):1108–15, may 2010.

[147] T. Okubo and C. Lenfant. Distribution function of lung volume and ventilation determined by lung N2 washout. *Journal of applied physiology*, 24(5):658–67, may 1968. From Witeley's thesis.

[148] C. Olegård, S. Söndergaard, E. Houltz, S. Lundin, and O. Stenqvist. Estimation of functional residual capacity at the bedside using standard monitoring equipment: a modified nitrogen washout/washin technique requiring a small change of the inspired oxygen fraction. *Anesthesia and analgesia*, 101(1):206–12, table of contents, jul 2005.

[149] A. J. Olszowka. Can VA/Q distributions in the lung be recovered from inert gas retention data? *Respiration physiology*, 25(2):191–198, nov 1975.

[150] M. Paiva. Computation of the boundary conditions for diffusion in the human lung. *Computers and biomedical research*, 5(6):585–595, dec 1972.

[151] M. Paiva. Theoretical Studies of Gas Mixing in the Lung. In L. Engel and M. Paiva, editors, *Gas Mixing and Distribution in the Lung*, volume 25 of *Lung Biology in Health and Disease*, chapter 6, pages 221–285. Dekker, New York, 1985.

[152] M. Paiva, A. van Muylem, and L. A. Engel. Slope of phase III in multibreath nitrogen washout and washin. *Clinical respiratory physiology (Bulletin européen de physiopathologie respiratoire)*, 18(2):273–280, 1982.

[153] N. Patroniti, M. Saini, A. Zanella, D. Weismann, S. Isgrò, G. Bellani, G. Foti, and A. Pesenti. Measurement of end-expiratory lung volume by oxygen washin-washout in controlled and assisted mechanically ventilated patients. *Intensive care medicine*, 34 (12):2235–40, dec 2008.

[154] K. T. Patton and G. A. Thibodeau. *Anatomy and Physiology*. Elsevier, St. Louis, 7th edition, 2009.

[155] R. Peslin, S. Dawson, and J. Mead. Analysis of multicomponent exponential curves by the Post-Widder's equation. *Journal of Applied Physiology*, 30(4):462–472, apr 1971.

[156] J. Pezzullo. *Biostatistics For Dummies*. John Wiley & Sons, Somerset, NJ, USA, 2013.

[157] J. Pezzullo. The Bootstrap Method for Standard Errors and Confidence Intervals. In *Biostatistics For Dummies Extras*. 2013. URL `www.dummies.com/extras/biostatistics`. [Accessed 12-01-2016].

[158] P. Phan, A. D. Farmery, and C. E. W. Hahn. Method and Apparatus for Measurement of Cardiopulmonary Function, 2015.

[159] P. Phan, C. Zang, D. Geer, C. E. W. Hahn, and A. D. Farmery. Inspired sinewave technique: A novel technology to measure cardiopulmonary function. In V. V. Toi and T. H. Lien Phuong, editors, *5th international conference on biomedical engineering*, volume 46 of *IFMBE proceedings*, pages 83–86, Cham, 2015. Springer.

[160] J. K. Pritchard, M. T. Seielstad, A. Perez-Lezaun, and M. W. Feldman. Population growth of human Y chromosomes: a study of Y chromosome microsatellites. *Molecular biology and evolution*, 16(12):1791–1798, dec 1999.

[161] P. H. Quanjer, G. J. Tammeling, J. E. Cotes, L. M. Fabbri, H. Matthys, O. F. Pedersen, R. Peslin, J. Roca, P. J. Sterk, W. T. Ulmer, and J. C. Yernault. Symbols, abbreviations and units. *The European respiratory journal*, 6(Suppl 16):85–100, mar 1993.

[162] P. H. Quanjer, G. J. Tammeling, J. E. Cotes, O. F. Pedersen, R. Peslin, and J. C. Yernault. Lung volumes and forced ventilatory flows. *The European respiratory journal*, 6(Suppl 16):5–40, mar 1993.

[163] H. Rahn. A concept of mean alveolar air and the ventilation-bloodflow relationships during pulmonary gas exchange. *American jornnal physiology*, 158(1):21–30, jul 1949.

[164] Rastrojo. Wikimedia Commons: Diagrama de los pulmones.svg, 2007. URL `commons.wikimedia.org/wiki/File:Diagrama_de_los_pulmones.svg`. [Accessed 08-06-2015].

[165] C. P. Robert, J.-M. Cornuet, J.-M. Marin, and N. S. Pillai. Lack of confidence in approximate Bayesian computation model choice. *Proceedings of the National Academy of Sciences of the United States of America*, 108(37):15112–7, sep 2011.

[166] J. S. Robertson, W. E. Siri, and H. B. Jones. Lung ventilation patterns determined by analysis of nitrogen elimination rates; use of mass spectrometer as a continuous gas analyzer. *The journal of clinical investigation*, 29(5):577–90, may 1950.

[167] P. D. Robinson, P. Latzin, S. Verbanck, G. L. Hall, A. R. Horsley, M. Gappa, C. Thamrin, H. G. M. Arets, P. Aurora, S. I. Fuchs, G. G. King, S. Lum, K. Macleod, M. Paiva, J. J. Pillow, S. Ranganathan, S. Ranganathan, F. Ratjen, F. Singer, S. Sonnappa, J. Stocks, P. Subbarao, B. R. Thompson, and P. M. Gustafsson. Consensus statement for inert gas washout measurement using multiple- and single- breath tests. *The European respiratory journal*, 41(3):507–22, mar 2013.

[168] D. O. Rodenstein and D. C. Stnescu. Reassessment of lung volume measurement by helium dilution and by body plethysmography in chronic air-flow obstruction. *American review of respiratory disease*, 126(6):1040–1044, may 1982.

[169] S. Rosenbaum, R. Skinner, I. Knight, and J. Garrow. A survey of heights and weights of adults in Great Britain, 1980. *Annals of human biology*, jul 2009.

[170] R. G. Rossing. Evaluation of a computer solution of exponential decay or washout curves. *Journal of applied physiology*, 21(6):1907–10, nov 1966.

[171] J.-J. Rouby, C. Arbelot, H. Brisson, Q. Lu, and B. Bouhemad. Measurement of alveolar recruitment at the bedside: the beginning of a new era in respiratory monitoring? *Respiratory care*, 58(3):539–42, mar 2013.

[172] Rscottweekly. Wikimedia Commons::LungVolume.svg, 2011. URL `https://commons.wikimedia.org/wiki/File:Lungvolumes.svg`. [Accessed 14-01-2016].

[173] M. C. Sainsbury, A. Lorenzi, E. M. Williams, and C. E. W. Hahn. A reconciliation of continuous and tidal ventilation gas exchange models. *Respiration Physiology*, 108(1): 89–99, apr 1997.

[174] A. E. Schlesinger, D. K. White, G. B. Mallory, C. F. Hildeboldt, and C. B. Huddleston. Estimation of total lung capacity from chest radiography and chest CT in children: comparison with body plethysmography. *American journal of roentgenology*, 165(1): 151–4, jul 1995.

[175] G. A. Schmidt. Counterpoint: should positive end-expiratory pressure in patients with ARDS be set based on oxygenation? No. *Chest*, 141(6):1382–4, jun 2012.

[176] G. A. Schmidt. Rebuttal to: should positive end-expiratory pressure in patients with ARDS be set on oxygenation? Yes. *Chest*, 141(6):1386, jun 2012.

[177] L. Shampine. Solving ODEs and DDEs with residual control. *Applied Numerical Mathematics*, 52(1):113–127, 2005. ISSN 01689274. doi: 10.1016/j.apnum.2004.07.003.

[178] L. Shampine and S. Thompson. Solving DDEs in Matlab. *Applied Numerical Mathematics*, 37(4):441–458, 2001. ISSN 01689274. doi: 10.1016/S0168-9274(00)00055-6.

[179] L. F. Shampine and M. W. Reichelt. The MATLAB ODE Suite. *SIAM Journal on Scientific Computing*, 18:1–22, 1997.

[180] T. Shinozaki, J. C. Abajian, B. S. Tabakin, and J. S. Hanson. Theory and clinical application of a digital nitrogen washout computer. *Journal of applied physiology*, 21 (1):202–8, jan 1966.

[181] S. Sisson, Y. Fan, and M. M. Tanaka. Sequential Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences of the United States of America*, 104 (6):1760–5, feb 2007.

[182] S. Sisson, Y. Fan, and M. M. Tanaka. Correction to: Sequential Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences*, 106(39):16889–16889, sep 2009.

[183] D. S. Sivia and J. Skilling. *Data Analysis: A Bayesian Tutorial*. Oxford University Press, 2006.

[184] I. E. Sodal, G. D. Swanson, A. J. Micco, F. Sprague, and D. G. Ellis. A computerized mass spectrometer and flowmeter system for respiratory gas measurements. *Annals of biomedical engineering*, 11(2):83–99, jan 1983.

[185] P. Stafler, S. Weinreb, H. Mussaffi, M. Mei-Zahav, D. Prais, G. Steuer, O. Bar-On, and H. Blau. Lung clearance index – A new parameter for airway obstruction in acute bronchiolitis? *European respiratory journal*, 44(Suppl 58):188, sep 2014.

[186] D. C. Stanescu. Letter: How to measure lung volume? *Chest*, 138(5):1280–2, nov 2010.

[187] W. E. Stewart and S. M. Mastenbrook. Parametric estimation of ventilation-perfusion ratio distributions. *Journal of applied physiology*, 55(1 Pt 1):37–51, jul 1983.

[188] D. Y. Sue, J. E. Hansen, M. Blais, and K. Wasserman. Measurement and analysis of gas exchange during exercise using a programmable calculator. *Journal of applied physiology*, 49(3):456–61, sep 1980.

[189] M. Sunnåker, A. G. Busetto, E. Numminen, J. Corander, M. Foll, and C. Dessimoz. Approximate Bayesian Computation. *PLoS computational biology*, 9(1):e1002803, jan 2013.

[190] C. Tantucci. Letter: Comparison among different methods on measurement of total lung capacity in COPD. *Chest*, 138(2):458–60, aug 2010.

[191] S. Tavener, S. J. Withy, and E. A. Harris. Response characteristics of a mass spectrometer. *Medical & biological engineering & computing*, 22(6):493–498, nov 1984.

[192] M. H. Tawhai, A. J. Pullan, and P. J. Hunter. Generation of an anatomically based three-dimensional model of the conducting airways. *Annals of biomedical engineering*, 28(7):793–802, jul 2000.

[193] M. H. Tawhai, P. Hunter, J. Tschirren, J. Reinhardt, G. McLennan, and E. A. Hoffman. CT-based geometry analysis and finite element models of the human and ovine bronchial tree. *Journal of applied physiology*, 97(6):2310–21, dec 2004.

[194] A. R. Templeton. Nested clade analysis: an extensively validated method for strong phylogeographic inference. *Molecular ecology*, 17(8):1877–80, apr 2008.

[195] A. R. Templeton. Statistical hypothesis testing in intraspecific phylogeography: nested clade phylogeographical analysis vs. approximate Bayesian computation. *Molecular ecology*, 18(2):319–31, jan 2009.

[196] A. R. Templeton. Why does a method that fails continue to be used? The answer. *Evolution*, 63(4):807–12, apr 2009.

[197] M. K. Tham. Letter: Measurement of continuous distributions of ventilation perfusion ratios : theory. *Journal of applied physiology*, 38(5):950–951, may 1975.

[198] T. Toni, D. Welch, N. Strelkowa, A. Ipsen, and M. P. Stumpf. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of The Royal Society Interface*, 6(31):187–202, feb 2009.

[199] S. Tsunoda, A. C. Young, and C. J. Martin. Emptying pattern of lung compartments in normal man. *Journal of applied physiology*, 32(5):644–9, may 1972.

[200] J. Tukey. *Exploratory Data Analysis*. Addison-Wesley Publishing Company, London,

1st edition, 1977.

[201] Z. Ugray, L. Lasdon, J. Plummer, F. Glover, J. Kelly, and R. Martí. Scatter Search and Local NLP Solvers: A Multistart Framework for Global Optimization. *INFORMS Journal on Computing*, 19(3):328–340, 2007.

[202] J. S. Ultman. Gas Transport in the Conducting Airways. In L. Engel and M. Paiva, editors, *Gas Mixing and Distribution in the Lung*, volume 25 of *Lung Biology in Health and Disease*, chapter 3, pages 63–136. Dekker, New York, mar 1985.

[203] E. J. R. van Beek. Lung Scintigraphy. In H.-U. Kauczor, editor, *Functional Imaging of the Chest*, Medical Radiology, chapter 7.3, pages 149–153. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.

[204] H. D. Van Liew. Semilogarithmic plots of data which reflect a continuum of exponential processes. *Science*, 138(3541):682–3, nov 1962.

[205] A. A. Viljanen, B. C. Viljanen, P. K. Halttunen, and K.-E. Kreus. Body plethysmographic studies in non-smoking, healthy adults. *Scandinavian journal of Clinical and Laboratory Investigation*, 42(S195):35–50, 1982.

[206] P. D. Wagner. Measurement of continuous distributions of ventilation-perfusion ratios: theory. *Journal of applied physiology*, 36(5):588–599, may 1974.

[207] P. D. Wagner. Information content of the multibreath nitrogen washout. *Journal of applied physiology*, 46(3):579–587, mar 1979.

[208] P. D. Wagner. The multiple inert gas elimination technique (MIGET). *Intensive care medicine*, 34(6):994–1001, jun 2008.

[209] P. Wakkers, H. Hellendoorn, G. De Weegh, and W. Heerspink. Applications of statistics in clinical chemistry. *Clinica chimica acta*, 64(2):173–184, oct 1975.

[210] J. Wanger, J. L. Clausen, A. Coates, O. F. Pedersen, V. Brusasco, F. Burgos, R. Casaburi, R. Crapo, P. Enright, C. P. M. van der Grinten, P. Gustafsson, J. Hankinson, R. Jensen, D. Johnson, N. Macintyre, R. McKay, M. R. Miller, D. Navajas, R. Pellegrino, and G. Viegi. Standardisation of the measurement of lung volumes. *The European respiratory journal*, 26(3):511–22, sep 2005.

[211] E. R. Weibel. *Morphometry of the Human Lung*. Springer, Berlin, 1963.

[212] E. R. Weibel. *The Pathway for Oxygen*. Harvard University Press, jul 1984.

[213] D. Weismann, H. Reissmann, S. Maisch, B. Füllekrug, and J. Schulte. Monitoring of functional residual capacity by an oxygen washin/washout; technical description and evaluation. *Journal of clinical monitoring and computing*, 20(4):251–60, aug 2006.

[214] J. B. West. Causes of and compensations for hypoxemia and hypercapnia. *Comprehensive physiology*, 1(3):1541–53, jul 2011.

[215] J. B. West and P. D. Wagner. Pulmonary Gas Exchange. In J. B. West, editor, *Bioengineering Aspects of the Lung*, volume 3 of *Lung Biology in Health and Disease*, pages 361–457. Dekker, New York, 1977.

[216] J. P. Whiteley. *Biomathematical Models for Some Respiratory Forcing Techniques*. Thesis, University of Oxford, 1998.

[217] J. P. Whiteley, D. J. Gavaghan, and C. E. W. Hahn. A tidal breathing model for the multiple inert gas elimination technique. *Journal of applied physiology*, 87(1):161–9, jul 1999.

[218] J. P. Whiteley, D. J. Gavaghan, and C. E. W. Hahn. A tidal breathing model of the inert gas sinewave technique for inhomogeneous lungs. *Respiration physiology*, 124(1):

65–83, dec 2000.

[219] J. P. Whiteley, A. D. Farmery, D. J. Gavaghan, and C. E. W. Hahn. A tidal ventilation model for oxygenation in respiratory failure. *Respiratory physiology & neurobiology*, 136(1):77–88, jun 2003.

[220] J. P. Whiteley, D. J. Gavaghan, and C. E. W. Hahn. The effect of diffusion in the respiratory tree on the alveolar amplitude response technique (AART). *Respiratory physiology & neurobiology*, 137(1):81–96, aug 2003.

[221] M. K. Whyte, A. M. Peters, J. M. Hughes, B. L. Henderson, G. J. Bellingan, J. E. Jackson, and E. R. Chilvers. Quantification of right to left shunt at rest and during exercise in patients with pulmonary arteriovenous malformations. *Thorax*, 47(10):790–796, oct 1992.

[222] E. M. Williams and C. E. W. Hahn. Measurement of Cardio-Respiratory Function using Single Frequency Inspiratory Gas Concentration Forcing Signals. In M. Hogan, O. Mathieu-Costello, D. Poole, and P. D. Wagner, editors, *Oxygen Transport to Tissue XVI*, Advances in Experimental Medicine and Biology, chapter D, pages 187–195. Plenum Press, New York, 1994.

[223] E. M. Williams, L. B. Gale, P. A. Oakley, M. C. Sainsbury, and C. E. W. Hahn. Development of a concentric water-displacement model lung. *Journal of biomedical engineering*, 15(5):420–424, sep 1993.

[224] E. M. Williams, J. B. Aspel, S. M. Burrough, W. A. Ryder, M. C. Sainsbury, L. Sutton, L. Xiong, A. M. S. Black, and C. E. W. Hahn. Assessment of cardiorespiratory function using oscillating inert gas forcing signals. *Journal of applied physiology*, 76(5):2130 – 2139, may 1994.

[225] E. M. Williams, D. J. Gavaghan, P. A. Oakley, M. C. Sainsbury, L. Xiong, A. M. S. Black, and C. E. W. Hahn. Measurement of dead-space in a model lung using an oscillating inspired argon signal. *Acta anaesthesiologica scandinavica*, 38(2):126–129, feb 1994.

[226] E. M. Williams, R. Hamilton, L. Sutton, and C. E. W. Hahn. Measurement of respiratory parameters by using inspired oxygen sinusoidal forcing signals. *Journal of applied physiology*, 81(2):998–1006, aug 1996.

[227] E. M. Williams, R. Hamilton, L. Sutton, and C. E. W. Hahn. Oxygen transport with oscillations of inspired oxygen concentration. *Respiration physiology*, 108(1):79–87, apr 1997.

[228] E. M. Williams, R. M. Hamilton, L. Sutton, J. P. Viale, and C. E. W. Hahn. Alveolar and dead space volume measured by oscillations of inspired oxygen in awake adults. *American journal of respiratory and critical care medicine*, 156(6):1834–9, dec 1997.

[229] E. M. Williams, M. C. Sainsbury, L. Sutton, L. Xiong, A. M. S. Black, J. P. Whiteley, D. J. Gavaghan, and C. E. W. Hahn. Pulmonary blood flow measured by inspiratory inert gas concentration forcing oscillations. *Respiration physiology*, 113(1):47–56, jul 1998.

[230] L. Wong, R. Hamilton, E. Palayiwa, and C. E. W. Hahn. A real-time algorithm to improve the response time of a clinical multigas analyser. *Journal of clinical monitoring and computing*, 14(6):441–446, aug 1998.

[231] S. N. Wood. Solv95 C/C++ Windows based delay differential equation solver, 1995. URL `www.maths.bath.ac.uk/ sw283/simon/dde.html`. [Accessed 05-08-2013].

[232] D. W. Zingg, M. Nemec, and T. H. Pulliam. A comparative evaluation of genetic and

gradient-based algorithms applied to aerodynamic optimization. *European journal of computational mechanics (Revue Européenne de mécanique numérique)*, 17:103–126, 2008.

[233] A. Zwart, R. Seagrave, and A. Van Dieren. Ventilation-perfusion ratio obtained by a noninvasive frequency response technique. *Journal of applied physiology*, 41(3):419 –424, 1976.

[234] A. Zwart, J. Bogaard, J. Jansen, and A. Versprille. A non-invasive determination of lung perfusion compared with the direct Fick method. *Pflügers archiv*, 375(2):213–217, jul 1978.

[235] A. Zwart, J. R. Jansen, and S. C. Luijendijk. Bohr dead space during helium washout. *Clinical respiratory physiology (Bulletin européen de physiopathologie respiratoire)*, 18 (2):261–72, mar 1982.