Modelling and Prediction of Physiological Behaviour in Critically Ill Patients

Ruby Langdon

Submitted for the degree of Doctor of Philosophy
in Mechanical Engineering
at the University of Canterbury,
Christchurch, New Zealand

September 2017

Acknowledgements

I would like to thank my supervisor Dr. Paul Docherty for his expertise, as well as his constant support and motivation.

Thanks to my co-supervisors Dr. Yeong Shiong Chiew, Prof. Geoff Chase, and Prof. Knut Möller for their help and many valuable insights.

I would also like to acknowledge the Mechanical Engineering Department for providing funding.

Table of Contents

List of Figures

List of Tables

Abstract

For patients with acute respiratory distress syndrome (ARDS), mechanical ventilation (MV) is an essential therapy in the intensive care unit. ARDS is diverse condition, the impact of which varies across patients. Every patient has different optimal ventilator settings that may change over the course of treatment, and there is no consensus on how these optimal settings should be found. In particular, the optimal level of positive end expiratory pressure (PEEP) is widely debated. PEEP that is too high or low can cause damage to healthy alveoli, leading to ventilator induced lung injury (VILI). VILI is associated with increased mortality, extended ICU stay, and high cost.

The use of mathematical models to determine patient-specific ventilator settings can reduce the incidence of VILI. There have been many models developed to capture pulmonary mechanics, but they have limitations in lack of ability to capture all relevant physiology, or in complexity and difficulty of implementation. The focus of this research is the development of a model of pulmonary mechanics that does not suffer from many of the disadvantages of previous models.

A nonlinear autoregressive (NARX) model was developed using a complex data set, and contains terms that enable it to fit to all features of the pressure waveform. It captures recruitment and distension across many increasing PEEP steps via an elastance vs. pressure curve that is defined by basis functions. Flow dependent terms allow it to capture viscoelastic effects and fit to an end-inspiratory pause. This model, and slight variations on it were tested on three cohorts of data in total. In many cases the model was compared with the well validated and extensively used first order model (FOM).

Various investigations supported the choice of the NARX model terms. This included using the model for interpolation across a recruitment manoeuvre. The interpolated NARX model fit was consistent across different types of patients, while the FOM performed worse in patients experiencing over-distension at high pressure. Another comparison with the FOM found that the NARX model could more reliably capture expected changes in resistance with PEEP. The NARX model could also identify independent inspiratory and expiratory elastance, due to the flow dependent terms that the FOM does not have.

The NARX model is flexible in its implementation. While it is normally identified in real time using the simple linear least squares method, it was also able to be combined with a modified Gauss – Newton parameter identification method for a spontaneous breathing application. In this case, anomalies in the pressure waveform caused by intermittent patient efforts were able to be removed to enable a more accurate identification of patient parameters.

Aside from patient-specific parameter identification, the main potential clinical use of the NARX model is in predicting the effects of changes in PEEP. An extrapolation of the elastance curve allowed pressure at higher PEEP levels to be predicted. By using partial recruitment manoeuvres as the training data, the NARX model predicted pressure waveforms at higher PEEP levels with significantly lower residuals than the FOM. Since large PEEP changes are not recommended clinically, the most relevant results were the predictions for small PEEP increases of 2 cmH_2O. In this scenario the NARX model accuracy was very high.

A statistical classification analysis used the prediction methodology to test the ability of the NARX model to detect when alveolar over-distension is likely to occur with PEEP increases. The analysis considered a pressure threshold above which the risk of over-distension is high. False negatives are potentially much more harmful to patients than false positives, as a false negative means a failure to detect when over-distension will occur with a PEEP increase. Thus, sensitivity was a more important metric than specificity in the analysis. In most scenarios, the NARX model threshold detection had a very high sensitivity and outperformed the FOM, even when compared to a separate method designed to produce the best prediction outcomes from the FOM. However, on one cohort, the parameterisation of the NARX model had to be reduced by reducing the number of basis functions in order to outperform the FOM over large prediction horizons.

An adaptation of the NARX model aimed to capture differences between COPD patients with resultant high auto-PEEP, and non-COPD patients. The adaptation replaced the flow dependent terms with basis functions that enabled linear resistance changes to be captured throughout a recruitment manoeuvre. The model parameters were able to distinguish between the two groups. At low pressure, the high auto-PEEP group had significantly higher modelled resistance, and had elastance curves that indicated a greater proportion of un-recruited lung units. Both of these outcomes were expected due to the airway narrowing and airway closures known to occur in COPD patients.

As in MV, model based glycaemic control can allow personalised care, reduce mortality and improve clinical outcomes. Hence, a side project was undertaken to investigate whether the basis function approaches developed for MV could have potential applications in glycaemic control. The concept was applied to a glucose model to identify a time varying insulin sensitivity (*SI*) in ICU patients over multiple days. Parameterisation of the model was varied by varying the ratio of basis functions to data points, and this ratio influenced the identified *SI* profiles that

were used to build *SI* prediction distributions. An analysis determined the appropriate level of parameterisation that resulted in accurate and precise predictions.

The glucose model, the NARX model, and its adaptations all captured clinically relevant patient-specific parameters. The NARX model in particular overcame many of the limitations of previous models, due to the novel use of basis functions to describe elastance, and the use of terms that fit an end-inspiratory relaxation. It achieved this over a range of cohorts that represented a wide variety of patient physiologies and ventilation protocols. The data fitting and prediction outcomes indicate that it has high potential to be useful in diagnosis and disease tracking in a clinical setting.

Nomenclature

Abbreviations

Pulmonary and Glucose

AC	Assist control
ARDS	Acute respiratory distress syndrome
COPD	Chronic obstructive pulmonary disease
CURE	Clinical utilisation of respiratory elastance
GC	Glycaemic control
ICU	Intensive care unit
MV	Mechanical ventilation
PEEP	Positive end expiratory pressure
PIP	Peak inspiratory pressure
PP	Plateau pressure
RM	Recruitment manoeuvre
SB	Spontaneous breathing
SIMV	Synchronised intermittent mandatory ventilation
SPRINT	Specialised relative insulin nutrition table
VILI	Ventilator induced lung injury
WOB	Work of breathing
ZEEP	Zero end-expiratory pressure

Mathematics and Statistics

AIC	Akaike information criterion
ARX	Autoregressive model with exogenous inputs
AUC	Area under the curve
CDF	Cumulative distribution function
CI	Confidence interval
CT	Computed tomography
EIT	Electrical impedance tomography
EVD	Evaluation data
FEA	Finite element analysis

FN	False negative
FOM	First order model
FP	False positive
GN	Gauss-Newton
IDD	Identification data
IID	Independent and identically distributed
IQR	Interquartile range
NARX	Nonlinear autoregressive model with exogenous inputs
PRM	Pressure dependent recruitment model
RMS	Root mean square
ROC	Receiver operating characteristic
SD	Standard deviation
SE	Standard error
STAR	Stochastic targeted
TCP	Threshold closing pressure
TN	True negative
TOP	Threshold opening pressure
TP	True positive

Model Parameters

Pulmonary Modelling

Ø	Basis function
C	Compliance
d	Basis function degree
E	Elastance
E_e	Expiratory elastance
E_i	Inspiratory elastance
L	Number of NARX model flow dependent terms
M	Number of NARX model basis functions
P	Airway pressure
P_e	Expiratory pressure

P_i	Inspiratory pressure
P_0	Offset pressure = PEEP
ΔP	Change in pressure
R	Resistance
t	Time
V	Volume
$\dot{V}$	Flow
V_e	Expiratory volume
V_i	Inspiratory volume
V_t	Tidal volume

Glucose Modelling

$\emptyset$	Basis function
E_{GP}	Endogenous glucose production
G	Total plasma glucose
G_b	Basal glucose
I	Plasma insulin
p_G	Non-insulin mediated glucose uptake
P_x	Exogenous glucose
Q	Interstitial insulin
SI	Insulin sensitivity
U	Insulin input to plasma
V_G	Glucose distribution volume
V_p	Plasma insulin distribution volume

Part 1: Introduction

Chapter 1 Background

1.1 Respiration

In healthy humans, breathing is primarily accomplished by the contraction and relaxation of the diaphragm and intercostal muscles. In inspiration, contraction of these muscles causes the diaphragm to move down and the ribcage to move upwards and outwards. The resulting volume increase in the thoracic cavity causes a pressure drop, and air flows from the area of higher pressure outside the body to lower pressure in the lungs. Normal expiration is a passive process, and depends on the elasticity of the lungs and chest wall (Figure 1.1).

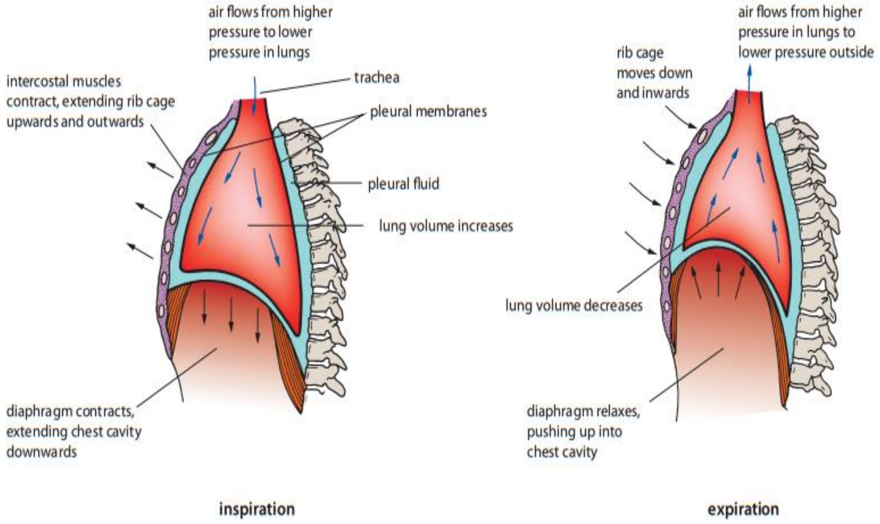


Figure 1.1. Inspiratory and expiratory processes (Won et al., 2014).

During breathing, air travels through the nose / mouth, through the pharynx, through the larynx, and then into the trachea. From the trachea, air passes into the bronchial tree, which consists of the right and left primary bronchi, branching into smaller secondary and tertiary bronchi. These branches lead to bronchioles and terminal bronchioles that connect to the alveoli (Figure 1.2). The trachea and primary bronchi have rings of cartilage that prevent the airways from collapsing. In the secondary and tertiary bronchi, the rings are replaced by strips of cartilage. Around the bronchioles, smooth muscle allows for constriction and dilation. A large amount of elastic connective tissue surrounds and supports the alveoli and other structures.

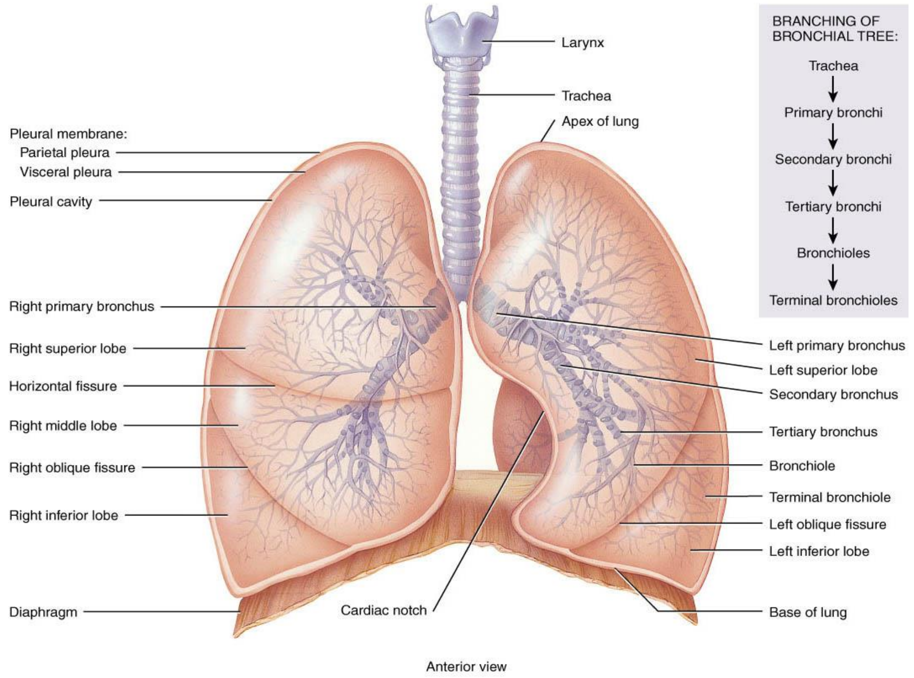


Figure 1.2. Anatomy of the lungs (G.J. Tortora, 2011).

Alveoli are the site of gas exchange of oxygen and carbon dioxide between air in the lungs and the blood capillaries in the walls of each alveolus. Each alveolus has a diameter of 0.2 to 0.3 mm, and are covered with a fluid film which contains surfactant, secreted by some alveolar cells. Surfactant decreases the surface tension of the water in the fluid film, which would otherwise provide a force that would collapse the alveoli at the end of expiration (Daniels and Orgeig, 2003).

By preventing the collapse of alveoli, the fluid film has the effect of increasing pulmonary compliance. Compliance (*C*) is the ability of the lungs to expand, and is defined as the change in volume (*V*) that can be achieved in the lungs per unit of pressure (*P*) change:

$$C = \frac{\Delta V}{\Delta P} \tag{1.1}$$

Low compliance represents a stiff lung that requires extra work to bring in a normal volume of air. Elastance (*E*) is the inverse of compliance, and is a measure of the tendency of the lung to recoil towards its original dimensions after a distending force is removed:

$$E = \frac{1}{C} = \frac{\Delta P}{\Delta V} \tag{1.2}$$

Elastance varies in healthy lungs for various reasons, such as age and body position. Elastance is an important metric as it increases as a result of many lung diseases, including acute respiratory distress syndrome (ARDS).

1.2 Acute Respiratory Distress Syndrome

ARDS was first described in 1967 by Ashbaugh *et al.* (1967). It is a diverse condition, but is generally characterised as an inflammatory response in the lungs, resulting in pulmonary oedema, alveolar injury, alveolar collapse, and abnormally low oxygen concentration in the blood (hypoxemia). It can involve loss of surfactant, fluid accumulation in airspaces, and the release of cytokines in response to the inflammation. The Berlin definition of ARDS is the latest iteration of the ARDS definition, published in 2012, and it classifies three categories of ARDS depending on the degree of hypoxemia: mild, moderate, and severe (The ARDS Definition Task Force, 2012).

ARDS can be caused by numerous factors, including smoke inhalation, pneumonia, sepsis, and trauma. ARDS is life threatening as it results in reduced gas exchange and poor oxygenation of the blood. Studies have found mortality rates that range from 30% to 70% (Bersten *et al.*, 2002, Phua *et al.*, 2009, Zambon and Vincent, 2008). Patients with ARDS often incur much greater medical costs, due to the need for mechanical ventilation to maintain adequate gas exchange (Dasta *et al.*, 2005, Zilberberg *et al.*, 2008).

In ARDS lungs, a convex pressure dependent elastance shape is common (Harris, 2005). For sedated ARDS patients, elastance initially decreases when pressure at the airway opening is increased from atmospheric level. The recruitment of lung units leads to lung volume increasing faster than what may be expected for equivalent pressure changes at higher ventilation pressures. When most of the recruitable lung volume has already been opened up, pressure begins to stretch the already recruited units and elastance begins to increase with pressure. Convex elastance is concurrent with a sigmoidal pressure – volume relationship (Figure 1.3).

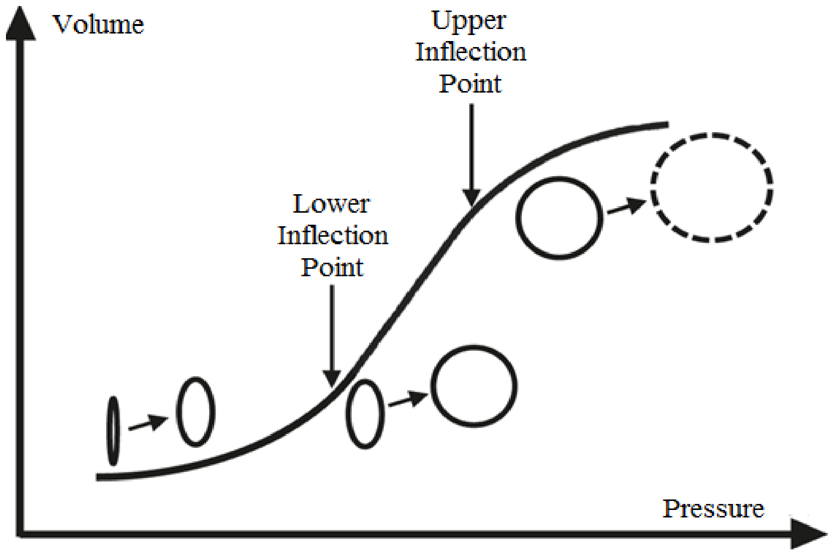


Figure 1.3. The inspiratory limb of the pressure – volume curve. Below the lower inflection point, increased pressure may recruit alveoli (represented by ovals and circles). Between the lower and upper inflection points, open alveoli distend with pressure. Above the upper inflection point, alveoli may over-distend and become injured (Miller et al., 2012).

Both the inter-patient and intra-patient physiology can vary greatly in ARDS, due to the disease progression over time, and due to different underlying causes of ARDS. At a particular time, each patient will have a unique elastance, or pressure – volume shape, but certain features can be used to characterise patient condition. For example, patients with emphysema may have decreased elastance due to the loss of elastic alveolar tissue, resulting in a lung that is easy to distend and a P – V curve that is concave to the pressure axis at low volume (Papandrinopoulou *et al.*, 2012). Patients with fibrosis or pneumonia may have a higher elastance due to a stiffer lung (Light, 1999). In this case, a higher pressure than normal is required to give the same increase in volume.

1.3 Mechanical Ventilation

Mechanical ventilation (MV) is critical in the treatment of ARDS. MV does not cure the underlying disease, but provides essential breathing assistance when the patient's spontaneous breathing efforts are not sufficient. Negative pressure ventilators first appeared in the 1800s, and consisted of a chamber that housed the whole patient apart from the head (Kacmarek, 2011). Sub-atmospheric pressure was delivered around the body to replace the work of respiratory muscles. Modern ventilators instead use positive pressure to push air into the airways. An endotracheal tube is inserted through the mouth and provides an unrestricted passageway for air into the trachea. Alternatively, a tube can be inserted through an artificial opening in the neck. Air is pushed into the airway until the ventilator terminates the breath, then the elastic recoil of the chest and lungs pushes air out by passive exhalation.

MV settings are often chosen based on ad hoc preferences of clinicians. Two frequently utilised MV modes are volume controlled ventilation and pressure controlled ventilation. In volume controlled mode, the volume of air displaced between inspiration and expiration (tidal volume, V_t) is fixed. Air is delivered with a constant flow, and when a flow sensor measures that the pre-set V_t has been applied, the ventilator cycles to exhalation. In pressure controlled mode, the pressure delivered is constant during a set inspiratory time, and the flow decelerates exponentially. For some patients, particularly those with high airway resistance, there is potential for dangerously high peak pressures to be reached in volume controlled mode (Campbell and Davis, 2002). In pressure controlled mode, peak airway pressure is limited, reducing the risk of high pressure injuries. Additionally, the decreasing flow pattern may allow air to be more evenly distributed in the lungs, resulting in improved gas exchange (Cadi *et al.*, 2008). However, V_t is variable in pressure controlled mode, and may change depending on changes to patient state. Patients must be monitored to ensure appropriate V_t is being delivered.

If a patient is able to initiate breaths, an assist control (AC) mode can be used in which the ventilator delivers support with a set tidal volume when it detects the patient's efforts. AC mode will periodically provide a breath if the patient does not initiate. Alternatively, synchronized intermittent mandatory ventilation (SIMV) assists some breaths and lets the patient draw some breaths on their own (Singer and Corbridge, 2009b). AC mode is undesirable for patients who breath rapidly as hyperinflation may occur (Groeger *et al.*, 1989). SIMV reduces this risk, however the patient work of breathing is increased compared to AC mode, and is thus unsuitable for patients with respiratory muscle weakness (Sassoon *et al.*, 1994).

In addition to the ventilation mode, a variety of other ventilator parameters can be set, including respiratory rate, fraction of inspired oxygen (FiO_2), inspiratory time, and peak allowable inspiratory pressure. An important parameter is positive end expiratory pressure (PEEP), which maintains positive airway pressure through the whole breathing cycle and helps to prevent alveolar collapse at the end of expiration.

A recruitment manoeuvre (RM) is a common MV procedure in ARDS patients, in which pressure is briefly increased to open collapsed alveoli. There are a variety of different RM techniques. In the sustained inflation technique, a continuous pressure of 40 cmH_2O is applied for up to 1 minute. Another common RM is a stepped increase in pressure every five to ten breaths until a peak pressure of 40 – 50 cmH_2O is reached. Then pressure is gradually stepped down again, and some PEEP is applied to keep the recruited lung units open. The most effective RM technique is unknown, and may vary in different circumstances (Pelosi *et al.*, 2010).

Since the degree of damage is different in every patient, and since ARDS is commonly combined with another underlying disease, standardising and optimising ventilator treatment is a challenge. Additionally, optimal ventilator settings will change over time as the disease progresses. Clinicians often rely on experience to select settings, and inexperienced clinicians may use a one-size-fits-all approach that will not be the optimal solution for most patients.

Figure 1.4 shows pressure, volume, and flow waveforms for a MV patient over a single breath in volume controlled mode. In this case, an end-inspiratory pause is performed, in which the flow is stopped for a short time at the end of inspiration, and expiration is prevented. This manoeuvre allows the plateau pressure (PP) to be determined. PP is the pressure applied to small airways and alveoli, and can be measured as the airway pressure at the end of the inspiratory pause.

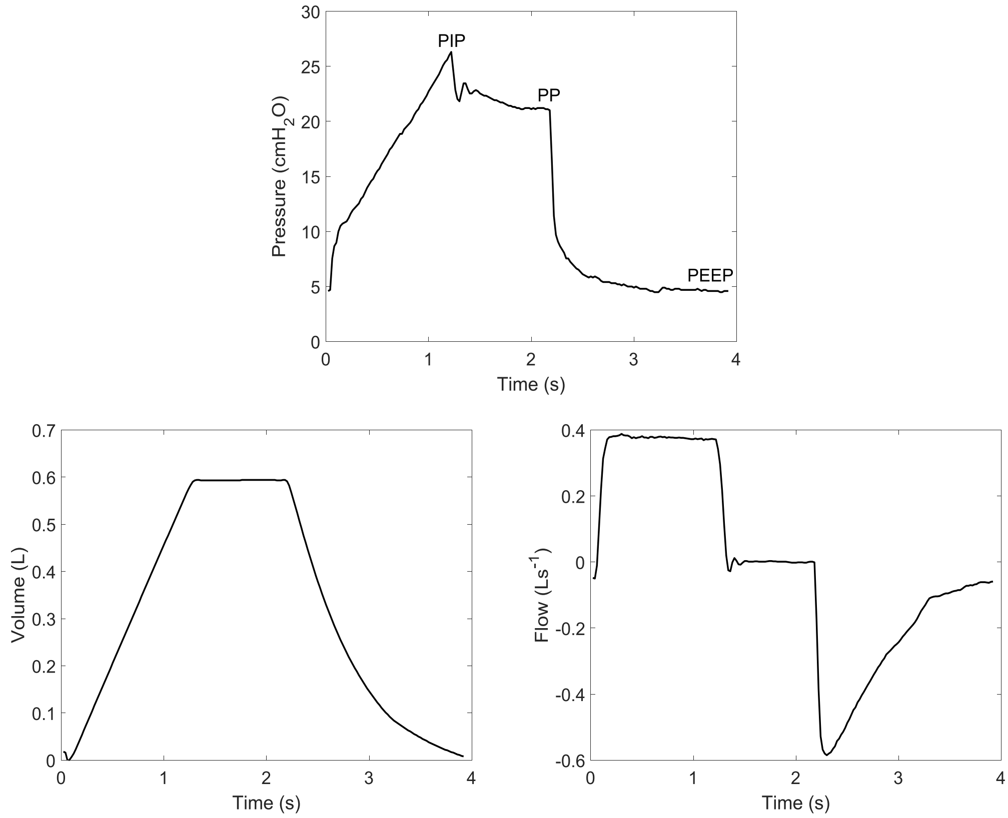


Figure 1.4. Mechanical ventilation waveforms.

1.4 Ventilator Induced Lung Injury

Ventilator induced lung injury (VILI) is damage that occurs due to the ventilator. There are two well-known ways in which VILI can occur. Firstly, alveoli can be over-distended and ruptured due to high tidal volumes (volutrauma) or high pressure (barotrauma). Secondly, an injury known as atelectrauma can be caused by shear forces from alveoli that cyclically open and close with each breath.

Damage to alveolar cells, and especially the rupture of cell membranes and cell death, results in local and systemic inflammatory responses (Ranieri *et al.*, 1999) (Rock and Kono, 2008). Biological mediators associated with inflammation can lead to organ failure, and ultimately death of the patient (Lionetti *et al.*, 2005), and this is most likely the method by which VILI increases mortality of ARDS patients (Whitehead and Slutsky, 2002). Most deaths in ARDS ultimately do result from multiple system organ failure, rather than respiratory failure (Ferring and Vincent, 1997). Alveolar rupture can additionally lead to conditions such as pneumothorax, pulmonary interstitial emphysema, and pneumomediastinum. In these disorders, air has escaped the lungs and has collected somewhere outside of the normal air spaces.

Patients with ARDS are at greater risk of VILI than others, due to the heterogeneous nature of ARDS lungs. Some areas may be collapsed or filled with fluid, while other lung units are healthy. For a given tidal volume, the collapsed areas may be underinflated, while the healthy alveoli are simultaneous overinflated (Gattinoni *et al.*, 2003).

Suboptimal PEEP is often the cause of VILI. In particular, PEEP that is too high can cause the over-distension of alveoli, and low PEEP can cause atelectrauma, as well as insufficient oxygenation. Other negative effects of PEEP can include disuse atrophy of the diaphragm muscle, and hypotension caused by impeded venous return due to increased pressure in the pleural cavity.

1.5 Lung Protective Strategies

A protective lung strategy to reduce the chances of VILI in ARDS patients uses a low tidal volume (V_t ~6 ml/kg predicted body weight), limited plateau pressure (PP < 30 cmH_2O), limited peak airway pressure (PIP < 50 cmH_2O) and a low level of PEEP. In particular, the ARDSNet trial found that a tidal volume of 6 ml/kg of predicted body weight reduced mortality by 22% compared to a V_t of 12 ml/kg . Several studies suggest that PP should not exceed 30 – 35 cmH_2O. (Hager *et al.*, 2005, Shiu and Rosen, 2006). PIP over 50 cmH_2O is associated with increased risk of alveolar rupture, and pneumothorax (Petersen and Baier, 1983). However, over-distension can often occur at 40 cmH_2O (Schranz *et al.*, 2012b).

The optimal PEEP for avoiding VILI is more ambiguous, however setting PEEP at the point of minimum elastance theoretically provides the best trade-off between maximising recruitment and avoiding over-distension. In particular, Carvalho *et al.* (2007) found that PEEP titrated to the point of minimum elastance corresponded to the greatest amount of normally aerated areas, by reducing the amount of lung units that were either collapsed or hyper-inflated.

Chapter 2 Models

2.1 First Order Model

The first order model (FOM), or single compartment model, describes airway pressure as a linear combination of a constant resistance and constant elastance (Bates, 2009). In an electrical analogy, this model consists of a resistor and capacitor in series, where the voltage represents airway pressure, and current represents air flow (Figure 2.1). The FOM is described by:

$$P = R\dot{V} + EV + P_0 \tag{2.1}$$

where E is respiratory system elastance (cmH_2O/L), R represents resistance of the trachea or endotracheal tube, bronchi, and bronchioles ($cmH_2O.s/L$), V is inspired volume (L), $\dot{V}$ is the flow (L/s), and P_0 represents PEEP (cmH_2O).

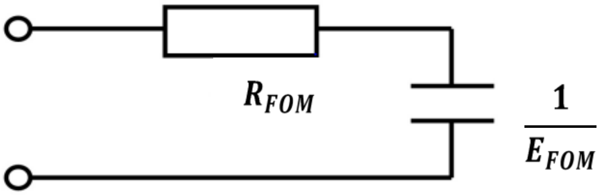


Figure 2.1. First order lung model.

A constant resistance is assumed because the endotracheal tube and bronchioles do not stretch much during MV. However, in some situations this may be an unreasonable assumption. If the severity of ARDS changes over time, small airways may collapse or open up, which could substantially alter the overall resistance. Variable resistance within one breathing cycle has been shown to occur in a high proportion of ARDS patients (Mols *et al.*, 2001). Variable resistance at different PEEP levels has been observed and attributed to the opening and closing of some airways (van Drunen *et al.*, 2013). Furthermore, a constant elastance does not match known patient behaviour. In particular, ARDS lungs have a very non-linear elastance due to the recruitment of collapsed alveoli at low pressure and distension at high pressure.

The FOM is the most frequently used model for most applications. This model, or simple derivatives of it, are the most likely to be used in current ventilators to calculate patient parameters. For example, in the CURE Soft trials the FOM is used over inspiratory data to calculate elastance for each breathing cycle (Szlavecz *et al.*, 2014). Despite the assumptions it makes, and despite a wealth of proposed alternative models, the FOM is very popular, implying that the alternative models are not providing much further benefit. Therefore, the performance of the pulmonary model developed in this thesis is frequently compared to the FOM.

2.2 Viscoelastic Model

As extensions to the single compartment model, several two compartment models were proposed (Otis *et al.*, 1956, Mead, 1969, Mount, 1955). Experimental data showed that the two compartment model proposed by Mount (1955) was found to be the most physiologically plausible of these options (Bates *et al.*, 1988). It describes a uniformly ventilated alveolar compartment surrounded by viscoelastic tissue, and the electrical analogy consists of an additional parallel circuit of a resistance and elastance in series with the FOM (Figure 2.2). Viscoelasticity means that the tension generated in lung tissue following a sudden stretch will adapt slowly over time. This property is due to elastic energy stored in the lung in inspiration that slowly dissipates over time as structural elements of the lung rearrange to lower energy levels. Viscoelasticity causes airway pressure to exponentially decrease during an end-inspiratory pause. Unlike the FOM, the viscoelastic model can capture this behaviour.

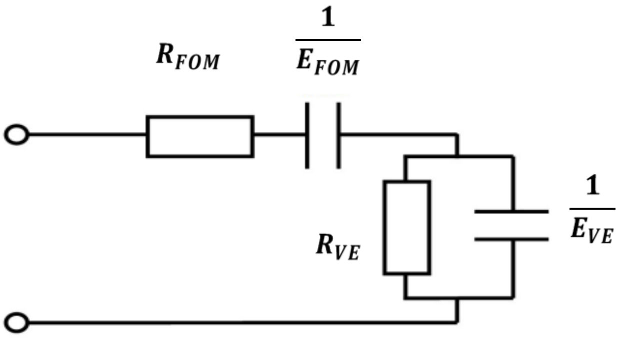


Figure 2.2. Viscoelastic lung model.

The viscoelastic model is limited in its abilities to describe all fundamental physiology in respiratory mechanics. Most importantly, it does not capture alveolar recruitment or over-distension (Schranz *et al.*, 2012a). Another disadvantage is that physiologically plausible values for resistance and elastance are not always identifiable, depending on the parameter identification method used, and the presence of noise in the data. Additionally, the model can be sensitive to the initial values used in gradient based identification, which can lead to incorrect parameters being computed (Schranz *et al.*, 2012a).

2.3 E_{drs} Model

The E_{drs} model is a different extension to the FOM, and includes a dynamic elastance term (Chiew *et al.*, 2011, van Drunen *et al.*, 2014):

$$P = R\dot{V} + E_{drs}V + P_0 \tag{2.2}$$

where E_{drs} is a time varying lung elastance that varies during each breath of MV. E_{drs} allows the changes occurring due to recruitment and distension to be seen dynamically within each breath. This gives a more detailed view of the patient's lung condition and severity of ARDS compared with the FOM. Minor extensions were later implemented by Knörzer *et al.* (2014) and Laufer *et al.* (2017) which added corrective terms to the dynamic elastance for different PEEP levels. Resistance is still assumed to be constant in the E_{drs} model. Additionally, the expiratory part of the breathing cycle is ignored, as E_{drs} is calculated from inspiration data only.

2.4 Expiratory Time Constant Model

The expiratory time constant model (van Drunen *et al.*, 2013) uses expiration data to determine a parameter (ζ) that is directly proportional to passive respiratory system elastance:

$$\dot{V}(t) = \dot{V}_0 e^{-\zeta t} \qquad 2.3$$

where $\dot{V}_0$ is the value of maximum expiratory flow. The system time constant (τ) is given by:

$$\tau = \frac{1}{\zeta} \qquad 2.4$$

If the parameter ζ is increasing, it implies that elastance is increasing. Setting PEEP at minimum τ (i.e. minimum elastance) would theoretically provide the optimal PEEP for that patient. However, the expiratory time constant model is only valid when resistance is constant. If this assumption is violated, there will be a poor correlation between ζ and respiratory elastance, and the model will not be useful. Additionally, the use of expiratory data only means that the model is ignoring information about patient state that exists in the inspiratory part of the breath.

2.5 Alveolar Recruitment Model

Hickling's alveolar recruitment model (Hickling, 1998) describes the lung as a collection of lung units, divided into horizontal compartments, where lower compartments experience higher superimposed pressure due to the weight of the lung (Figure 2.3). The opening and closing of alveoli in this model are assumed to be governed by normally distributed threshold opening pressures (TOP) and threshold closing pressures (TCP). TOP is the critical pressure at which a previously collapsed alveoli unit is recruited during inspiration. TCP is the critical pressure at which a previously recruited unit collapses during expiration. The mean and standard deviation of the TOP and TCP distributions change as the patient's condition evolves.

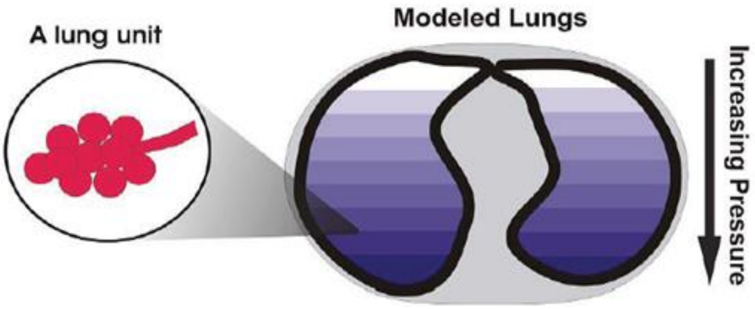


Figure 2.3. Graphical representation of the alveolar recruitment model (Sundaresan et al., 2009).

2.6 Minimal Model

The minimal model (Sundaresan *et al.*, 2009) is based on Hickling's description of the alveolar recruitment model. It also includes a sigmoid based curve that gives the volume of a lung unit once it is recruited:

$$V = \frac{a}{1 + e^{b(-P+c)}} + d \qquad 2.5$$

where *a* is the height of the curve, *b* defines the curvature, *c* is the midpoint, and *d* is a minimum volume. This curve is used to represent viscoelastic effects of recruited alveoli, improving the accuracy of the model.

For a given PEEP, pressure and volume measurements are used to calculate TOP and TCP. Changes in these values with PEEP yield information about patient state. For example, a decrease in mean TOP with increased PEEP implies recruitment is occurring. An increase in TCP with PEEP implies that PEEP is preventing de-recruitment. This information can be used to determine an optimal patient-specific PEEP.

An earlier version of the minimal model used four different types of lung units, instead of one (Yuta *et al.*, 2004). The model was highly physiologically representative, but required up to 42 parameters to be identified, most of which were impractical or impossible to obtain clinically. The minimal model was therefore created as a simplified version. The identification of the minimal model is greatly simplified. However, physiological detail is lost, such as the ability to identify the number of healthy and ARDS affected lung units for each patient.

Though the minimal model is comparatively easy to identify, it still has the disadvantage of requiring an unusual intervention from ICU staff that interrupts the patients' breathing pattern. In particular, the fitting method requires briefly deflating the lungs to atmospheric pressure. This interruption may be undesirable for some patients, and recording deflation data requires a separate data acquisition system, or specialised ventilator.

2.7 Pressure Dependent Recruitment Model

The pressure dependent recruitment model (PRM) (Schranz *et al.*, 2012b) uses the alveolar recruitment model and combines it with an alveolar distension model (Salazar and Knowles, 1964). The alveolar distension model describes the pressure – volume relationship using an exponential function:

$$V = V_{max}(1 - e^{-p_a K}) \tag{2.6}$$

where V_{max} is the maximum pulmonary volume, p_a represents alveolar pressure, and K describes the speed at which compliance decreases with increasing pressure.

The PRM is ultimately a pressure dependent compliance function embedded in the structure of the FOM. As in Hickling's model, it describes n = 30 layers of alveolar units that are recruited once the appropriate pressure is exceeded. Once a layer is recruited, the layer compliance contributes to the overall lung compliance. However, the compliance of each layer decreases exponentially as pressure increases. The PRM is shown in Equation 2.7.

$$P = R\dot{V} + p_a \tag{2.7a}$$

$$\dot{p}_a = \left[C_{FRC}e^{-Kp_a} + C_L \sum_{n=1}^{30} H_n e^{-K(p_a - TOP - SP_n)}\right]^{-1} \dot{V} \tag{2.7b}$$

C_{FRC} is the overall initial compliance of the alveoli that are open at the beginning of inspiration. C_L is the initial compliance of a recruited layer. H_n takes the value 1 to represent a layer of units that are recruited, or 0 if the layer is not recruited. In addition to the TOP, the superimposed pressure SP_n must be exceeded to recruit a specific layer. SP_n increases linearly as the layer number increases.

Resistance in this model is assumed constant, and the model is generally only identified on inspiration data. However, the main disadvantage of the PRM has been its issues with identifiability (Schranz *et al.*, 2012c, Raue *et al.*, 2009). Where recruitment or over-distension effects are not observed in the measured data, identifiability of one or more of the five PRM parameters is impaired. It is a particular disadvantage that the model requires evidence of over-distension of the lungs, as the high pressures that lead to over-distension may result in VILI.

2.8 Finite Element Modelling

In finite element analysis (FEA), a particular geometry is subdivided into smaller components that combine into a structure that represents the entire system. In pulmonary modelling, FEA has been used to model the physical structure of the lungs, including the bronchial tree and vascular structure (Perzl *et al.*, 1996, Bradley *et al.*, 1997, Tawhai *et al.*, 2000, Tawhai *et al.*, 2004, Burrowes *et al.*, 2005, Montesantos *et al.*, 2016). Information about patient-specific lung geometry must first be obtained via patient scans, such as magnetic resonance imaging (MRI) or more recently, computed tomography (CT) scans. This information is then used to discretise the domain into the small elements by a meshing process. The use of small elements can enable an accurate physical representation.

The mesh models can be used to simulate behaviour, such as ventilation distribution, and pulmonary dysfunctions (Tgavalekos *et al.*, 2005, Kim *et al.*, 2015). FE models of the lungs have also been used to model respiratory motion (Werner *et al.*, 2009, Villard *et al.*, 2005, Eom *et al.*, 2010, Al-Mayah *et al.*, 2008). 4D-CT scans are used to create a mesh, and FE analysis can then be used to solve a model that simulates the motion of lung inflation and deflation. This type of model has applications in tumour treatment. In radiation therapy, an assumption that the tumour location and shape remains constant will be violated by respiration. FE models can predict localised deformation of the lungs and surrounding tissues and reduce tumour targeting errors (Eom *et al.*, 2010).

FE models have also utilised electrical impedance tomography (EIT) scans (Gong *et al.*, 2015). EIT uses a number of electrodes placed on the patient's skin that measure the relationships between voltage and current across the geometry. For modelling pulmonary tissue, the conductivity distribution in a cross section of the thorax is estimated. EIT can visualise shifts in body fluids and lung aeration as pressure changes. Luepschen *et al.* (2008) uses EIT and FE analysis to model levels of fluid in the lungs.

The effectiveness of FE models depends heavily on the quality of the geometric information obtained by the scan, and the accuracy of the FE mesh. A good quality mesh can be difficult to create. Meshes generated from CT images especially can have a dense arrangement of elements with irregular geometries. Smoothing methods can be used to create a more regular mesh, though detail and accuracy will be lost (Brock *et al.*, 2005, Werner *et al.*, 2008, Villard *et al.*, 2004).

Compared with the lumped parameter pressure – flow based models, FE model validation can be a more ambiguous process. Validation involves comparison with the scan data, where anatomical reference points must be chosen, and their distance measured from corresponding points on the model. CT scan based FE models have been consistent with anatomic measurements. However, the generation of these models is more expensive and more dangerous due to the radiation used in CT scans. EIT based models are more appropriate for real time bedside monitoring of ARDS patients, though EIT has limitations in that it requires precise electrode placement (Boyle and Adler, 2011, Sabine *et al.*, 2015) and lacks the spatial resolution of CT scans.

2.9 Summary

There are a wide range of physiologically and clinically relevant pulmonary models. However, there remains significant debate about the best model that fits all modes of ventilation. The models presented here have not been able to consistently describe recruitment and distension effects across PEEP changes. The simpler models, such as the FOM and viscoelastic model, are limited in their ability to describe all relevant respiratory mechanics. As a general rule, more complex models can capture more physiological behaviour, but are much less practically useful. Complex models often have issues with identifiability, so they cannot be relied upon to generate accurate patient-specific models. Ambiguity in parameter values ultimately leads to lower confidence in the ability of a model to predict patient responses to changes in therapy, and thus lower confidence in its use for decision support.

Chapter 3 Mathematics and Statistics

3.1 Parameter Identification

Parameter identification (ID) methods are used to determine the parameters of a mathematical model that allow it to most closely describe measured data. This section describes the parameter ID methods used in the thesis. It is not intended to provide an overview of all parameter ID methods. Linear regression was used to identify the NARX model and FOM in almost all scenarios. The Gauss – Newton (GN) method was used to identify the NARX model and FOM in Chapter 9, and the glucose model in Chapter 13 and Chapter 14.

3.1.1 Linear Least Squares

Linear least squares aims to find the parameter values that result in the best fit between measured data and the model output. As a general example, consider a system of linear equations:

$$y_i = \sum_{j=1}^{m} a_j x_{i,j} \quad (i = 1,2,\ldots,N) \tag{3.1}$$

where there are *N* number of (*x*, *y*) paired data points and *m* unknown coefficients. For *N* > *m*, the system is overdetermined. The system can be defined in matrix form:

$$\mathbf{Ax} = \mathbf{b} \tag{3.2}$$

where

$$\mathbf{A} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,m} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N,1} & x_{N,2} & \cdots & x_{N,m} \end{bmatrix};$$

$$\mathbf{x} = [a_1 \quad a_2 \quad \cdots \quad a_m]^T ;$$

$$\text{and} = [y_1 \quad y_2 \quad \cdots \quad y_N]^T .$$

Then the objective is to find the coefficients which fit the equation best, i.e. to minimise the objective function:

$$\Psi = \sqrt{\sum\nolimits_{i=1}^{N} (y_i - \hat{y}_i)^2} \tag{3.3}$$

where y_i are the values of the measured data, and $\hat{y}_i$ are the values of the model at the corresponding times.

If $N < m$, the model is non-identifiable. If $N \geq m$, there are many methods of finding the optimal parameters. By iterating through values for each parameter, a contour plot of Ψ can be created, which can graphically show the minimum. If $N = m$, Equation 3.4 can be used. If **A** is non-square then Equation 3.5 can be used. Generally, when $N > m$, the most convenient method of solving a linear least squares problem is to use the Matlab backslash operator, which invokes different algorithms depending on the structure of **A**.

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} \tag{3.4}$$

$$\mathbf{x} = (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{b} \tag{3.5}$$

3.1.2 Gauss – Newton

Gauss – Newton (GN) is a gradient descent method, used for nonlinear least squares problems. It is an iterative method, in which parameter values are updated each iteration, depending on a rule that utilises the error between the measured data and model output.

The first step in GN is to choose initial parameter values ($\mathbf{x}_0$), and a small value for the perturbation (Δx). The model is first forward simulated at the initial values. Then, the iterative process that updates the parameter set (**x**) at each iteration, i, is given by:

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \left(\mathbf{J}^{\mathrm{T}}\mathbf{J}\right)^{-1}\mathbf{J}^{\mathrm{T}}\mathbf{\Psi} \tag{3.6}$$

where

$$\mathbf{J} = \begin{bmatrix} \frac{\delta\psi_1}{\delta x_1} & \frac{\delta\psi_1}{\delta x_2} & \cdots & \frac{\delta\psi_1}{\delta x_m} \\ \frac{\delta\psi_2}{\delta x_1} & \frac{\delta\psi_2}{\delta x_2} & \cdots & \frac{\delta\psi_2}{\delta x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\delta\psi_N}{\delta x_1} & \frac{\delta\psi_N}{\delta x_2} & \cdots & \frac{\delta\psi_N}{\delta x_m} \end{bmatrix};$$

$$\text{and } \mathbf{\Psi} = \begin{bmatrix} \hat{y}_1 - y_1 \\ \hat{y}_2 - y_2 \\ \vdots \\ \hat{y}_N - y_N \end{bmatrix}.$$

$\mathbf{\Psi}$ is the residual vector, calculated as the difference between the measured data ($y_1 \ldots y_N$) and the model output at the sample times ($\hat{y}_1 \ldots \hat{y}_N$). **J** is the Jacobian, which is a matrix of first order partial derivatives.

The GN algorithm steps the parameters in the direction that reduces $\mathbf{\Psi}$. This process repeats for a set number of iterations, chosen such that the parameter values have converged. GN typically converges quickly, however the method may not work well if the function is flat in the region around the minimum. The initial values, and the size of the perturbation are important settings, as they can determine whether or not convergence is achieved, and whether convergence occurs at a local or global minimum.

3.2 Autoregressive Modelling

A typical form of an autoregressive model with exogenous inputs (ARX) is given by (Billings, 2013):

$$Y(t) = \sum_i d_i Y(t_{-i}) + \sum_j \sum_i e_{j,i} X_j(t_{1-i}) + f \qquad 3.7$$

where *Y* is the measured output, *X* is a known system input, *j* is the index of system inputs, and *i* is the index of the past inputs and outputs. The coefficients to be identified are *d*, *e*, and *f*. The system output is essentially a linear combination of previous values of the output, and the current and previous values of the inputs.

To identify the unknown parameters, a matrix system of equations must be generated:

$$\mathbf{Ax} = \mathbf{b} \qquad 3.8$$

where

$$\mathbf{A} = \begin{bmatrix} Y(t_{-1}) & Y(t_{-2}) & \dots & Y(t_{-i-1}) & X_1(t_0) & X_1(t_{-1}) & \dots & X_1(t_{-i}) & X_2(t_0) & \dots & 1 \\ Y(t_0) & Y(t_{-1}) & \dots & Y(t_{-i}) & X_1(t_1) & X_1(t_0) & \dots & X_1(t_{1-i}) & X_2(t_1) & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ Y(t_{N-1}) & Y(t_{N-2}) & \dots & Y(t_{N-i-1}) & X_1(t_N) & X_1(t_{N-1}) & \dots & X_1(t_{N-i}) & X_2(t_N) & \dots & 1 \end{bmatrix};$$

$$\mathbf{x} = [d_1 \quad d_2 \quad \dots \quad d_i \quad e_{1,1} \quad e_{1,2} \quad \dots \quad e_{1,i} \quad e_{2,1} \quad \dots \quad f]^T;$$

$$\text{and } \mathbf{b} = [Y(t_0) \quad Y(t_1) \quad \dots \quad Y(t_N)]^T.$$

The number of data points is given by *N*. Terms that have negative values of *t* will be set to zero, as these measurements do not exist. The equation is easily solved for **x** using linear regression.

The number of previous input and output data points, *i*, must be decided. This value determines the number of model parameters, and is dependent on the nature of the particular system. A general rule is to use as few parameters as possible while ensuring the model adequately describes the system. If *i* is small, all important features may not be captured and incorporated into the model. However, too many parameters can cause the model to capture measurement noise, and will result in reduced predictive capabilities. Note also that the value of *i* may need to be different for the inputs and output.

Equation 3.1 has one output only, though some ARX systems have multiple outputs. The equation may also include linear or nonlinear functions of the inputs and outputs. The addition of these nonlinear terms, (generating a NARX model), will depend on characteristics of the system, which may or may not be initially known. By providing a number of candidate nonlinear terms, parameter ID can determine which terms are appropriate, and the underlying behaviours of the system can be determined.

3.3 ARX Model Structure Selection

There are many methods to choose between different ARX model structures. For linear systems, criteria such as the Akaike Information Criterion (AIC) or Minimum Descriptive Length (MDL) may be used, both of which consider a trade-off between complexity of the model and complexity of the data (Lind and Ljung, 2008).

In the AIC method, the AIC value is found for each of a set of candidate models. The formula takes into account the residuals of the model fit and the measurement noise assumed to be present. The AIC values give a relative measure of how probable each model is to be the model that minimises information loss, and is closest to the true process that generated the data (Burnham and Anderson, 2002). This method is discussed further in section 3.5.4. Closely related to AIC is the Bayesian Information Criterion (BIC), which gives a greater penalty for highly parameterised models compared to AIC (Schwarz, 1978).

MDL uses the idea of data compression (Rissanen, 1978). Regularity in data can be used to compress the data, and the model that captures the most regularity in the data can achieve the best compression. The best model of the candidate set is considered to be the one for which the description of the model plus the model’s description of the data is minimal.

Of these three methods, the most appropriate option is debated. If the true model is in the candidate set, BIC is more likely than AIC to select it (Burnham and Anderson, 2002). However, normally, a good assumption is that none of the candidates are the exact true model, in which case AIC is more appropriate (Yang, 2005). Some studies suggest AIC has performance advantages over BIC (Burnham and Anderson, 2004). However, AIC can have a tendency to favour more complex models that over-fit the data. MDL methods avoid over-fitting, (Grünwald, 2007) however they can be very computationally intensive.

An important criterion is the model – data residuals. Low residuals are important, but low systemic bias in residuals is also key. While one model might best fit the data, an alternative model with a worse fit might be more appropriate if residual bias is less. This is increasingly important if the model is to be used for prediction. Also, if residuals are too low, the model may be over parameterised. In this situation, the model parameters are capturing noise rather than useful physiological dynamics, and the model's predictive capability is likely to be poor. By testing different terms and analysing the model fit, parameter values, and predictive power, the most useful model structure can be found. Ultimately, the specific system and the purpose of the model influence how the structure will be determined.

3.4 Basis Functions

Basis functions are used to approximate multivariable functions by linear combinations of terms that are based on a single univariate function. They are often used to approximate nonlinear behaviour that might be too difficult to evaluate otherwise. Any continuous function can be fit with arbitrary accuracy by a sum of a sufficiently large number of basis functions. Global basis functions e.g. polynomials affect the whole input space. Often local basis functions are more appropriate, such as Gaussian functions, or basis – spine (b-spline) functions.

Zero order b-spline functions are square functions, defined:

$$\emptyset_{i,0}(P) = \begin{cases} 1 & if\ P_i \le P < P_{i+1} \\ 0 & otherwise \end{cases} \qquad 3.9$$

where P_i are division points (also known as knots) that subdivide the interval $0 \le P \le P_{\max}$. Basis functions of higher degrees (d) are defined recursively (de Boor, 1972):

$$\emptyset_{i,d}(P) = \frac{P - P_i}{P_{i+d} - P_i}\emptyset_{i,d-1}(P) + \frac{P_{i+d+1} - P}{P_{i+d+1} - P_{i+1}}\emptyset_{i+1,d-1}(P) \qquad 3.10$$

Figure 3.1 shows examples of zeroth, first, and second order basis functions. If the degree is zero, the basis functions are all step functions, each one spanning the distance between two adjacent knots. Basis functions of higher orders overlap, and a basis function of degree d is non-zero on $d + 1$ knot spans. The functions sum to 1 at every point.

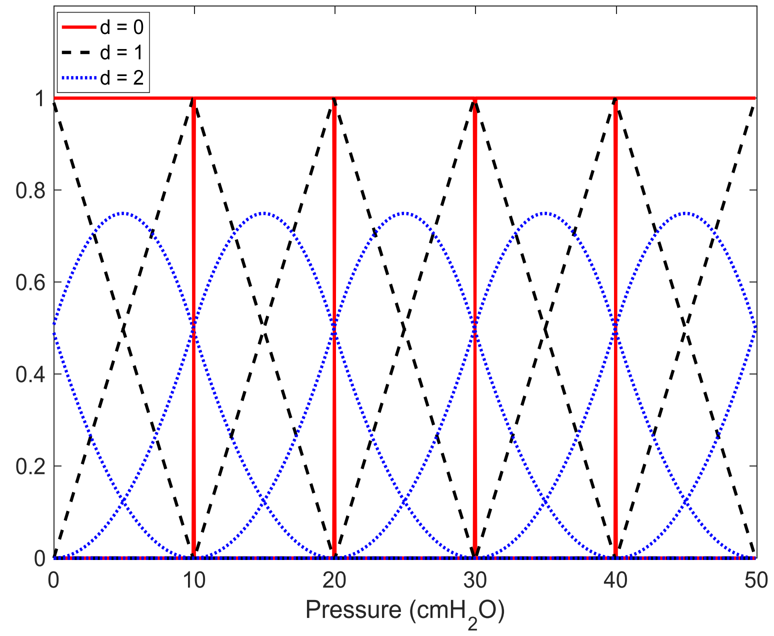


Figure 3.1. Basis functions for 0 ≤ P ≤ 50 with P_i = [0, 10, 20, 30, 40, 50] cmH$_2$O.

A model containing a linear combination of basis functions can also be interpreted as a simple type of artificial neural network (Coelho and Krohling, 2006, Botzheim *et al.*, 2007). The basis function parameters, determined by parameter ID, are like the neuron weights, and the basis functions are like the activation functions that convert weights into an output activation.

3.5 Statistics

3.5.1 Residuals

One of the most important methods of assessing model suitability is observation of the residuals. For a given sample point, the model – data residual is the difference between the model output at that point, and the measured data point. This is distinct from the error, which is the difference from the true value of a quantity, which is unobservable due to measurement noise.

The root mean square (RMS) residual is a commonly used measure in the analysis of model fits. It is defined:

$$\mathrm{RMS} = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N}} \qquad 3.11$$

The RMS value aggregates the magnitudes of the residuals for all the sample times into a single measure. When a model is used for prediction, the RMS difference between model prediction and measured data provides a single measure of predictive power. Comparison of RMS values between models must always be over the same variable, as the value is scale dependent.

3.5.2 Statistical Tests

Several statistical test are commonly used to compare two different sets of data. Popular test include the t-test (one sample, two sample, or paired), the Kolmogorov – Smirnov test (K-S test), and the Wilcoxon signed-rank test.

The t-test is used to determine if two sets of data are significantly unlikely to be drawn from the same normal distribution (Dodge, 2008). A one sample t-test evaluates whether the data is significantly unlikely to come from a distribution with a particular mean. This can be used to test for bias in residuals. In particular, biased residuals will have a mean that is significantly different from zero.

A paired t-test evaluates the hypothesis that two matched samples come from distributions with equal means. It can be used to compare something that has been tested twice e.g. the same subjects being tested on two different treatments, or to compare the performance of two different models on the same data set. In contrast, a two sample t-test tests the hypothesis that two independent samples come from distributions with equal means. This form of the test is used when two separate sets of samples are obtained, one from each of the two populations being compared. The two sample t-test could be used to compare model parameters for two groups of patients with different diagnoses.

The signed-rank test is a non-parametric test that can be used as an alternative to the t-test when the population cannot be assumed to be normally distributed (Wilcoxon, 1945). It can be used on model residuals in an interpolation or extrapolation context. If a model is being tested over data it has not been trained on, we cannot assume residuals will be normally distributed because the model fit to the unknown data might be bad. Thus, in such cases the signed-rank sum test is an appropriate test for bias.

The K-S test is used to compare distributions (Massey, 1951). A one sample K-S test compares a sample with a reference distribution, whereas a two sample K-S test compares two empirical distributions with each other. The null hypothesis is that the samples are drawn from the same distribution. The two sample K-S test may be used to strengthen an argument about the performance of two different models, based on an analysis of residuals that may or may not be normally distributed. A model with biased residuals will have a significantly different residual distribution to a model with unbiased residuals.

3.5.3 Correlations, R^2, and Bland-Altman

The Pearson correlation coefficient, *R*, is a measure of the linearity of the relationship between two variables (Sharma, 2005). It has a value between +1 and -1, where +1 indicates perfect positive linear correlation, 0 indicates no linear correlation, and -1 is perfect negative linear correlation. The interpretation of the correlation coefficient depends on the context, and the amount of noise or confounding factors expected in the data. The value of *R* can be misleading if outliers are present, or if performed on a small subgroup of the population that does not represent the full population. The formula for *R* is given by Equation 3.12, where *N* is the number of (*x*, *y*) data points.

$$R = \frac{N(\sum xy) - (\sum x)(\sum y)}{\sqrt{[N\sum x^2 - (\sum x)^2][N\sum y^2 - (\sum y)^2]}} \tag{3.12}$$

In regression analysis, the coefficient of determination (R^2) is often used. R^2 ranges from 0 to 1, and it indicates the proportion of variance in the dependent variable that is predictable by the independent variable. It indicates the strength of the equation which is used to predict the value of the dependent variable, and can give a measure of the goodness of fit of a model.

If there is a bias between two variables, a Bland – Altman plot allows this to be more easily observed (Altman and Bland, 1983). The *X* axis is the average of the paired measurements, and the *Y* axis is the difference between these measurements. This type of plot is commonly used to compare two methods that are designed to measure the same parameter. Even if there is a high correlation between the two methods, this does not necessarily mean the methods have a high agreement. A Bland – Altman plot quantifies agreement by constructing limits of agreement around the mean difference, or bias. Limits of agreement usually designate two standard deviations from the mean difference, within which 95% of the differences of the second method, compared to the first one, fall (Giavarina, 2015). The Bland – Altman method does not say whether these limits are clinically acceptable or not. Ultimately, clinical context determines whether the two methods have a high enough agreement.

3.5.4 Akaike Information Criterion

The Akaike information criterion (AIC) provides a theoretical method of model selection, based on the trade-off between the goodness-of-fit of the model, and model complexity (Burnham and Anderson, 2002):

$$\mathrm{AIC} = -2\ln(A) + 2B \tag{3.13}$$

where B is the number of model parameters and $ln(A)$ is the log-likelihood function. The MATLAB *normlike* function can be used to calculate the log likelihood, using model – data residuals and the expected measurement noise as inputs.

The AICc correction increases the penalty on higher model complexity, and decreases the probability of selecting a model with too many parameters:

$$\mathrm{AICc} = \mathrm{AIC} + \frac{2B(B+1)}{N-B-1} \tag{3.14}$$

where N is the number of data points. AICc should be used when the number of data points is not many times larger than B^2 (Hurvich and Tsai, 1989, Anderson *et al.*, 1994).

When comparing models, the model with the minimum AIC or AICc value is assumed to be the best model. It says nothing about the absolute quality of the models, only their relative quality.

3.5.5 Statistical Classification

A binary classifier is an algorithm that determines which of two categories that a certain observation belongs to (Ting, 2011). A confusion matrix is a 2 x 2 table that allows visualisation of the classifier algorithm's performance. The table reports the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). From this table, the accuracy of the classifier can be calculated as:

$$\mathrm{Accuracy} = \frac{\sum \mathrm{TP} + \sum \mathrm{TN}}{\sum \text{total population}} \tag{3.15}$$

The accuracy gives the proportion of tests that the classifier got correct. However, accuracy is not a reliable metric when the number of samples in different classes varies. Often, a more useful way of assessing the classifier is to calculate the sensitivity and specificity:

$$\mathrm{Sensitivity} = \frac{\sum \mathrm{TP}}{\sum \mathrm{TP} + \sum \mathrm{FN}} \tag{3.16}$$

$$\mathrm{Specificity} = \frac{\sum \mathrm{TN}}{\sum \mathrm{TN} + \sum \mathrm{FP}} \tag{3.17}$$

Sensitivity measures the proportion of positives that are correctly identified, and specificity measures the proportion of negatives that are correctly identified. A highly sensitive classifier rarely gives a false negative, and a highly specific classifier rarely gives a false positive.

In the context of model performance assessment, this classification method can be used to compare binary predictive capabilities. Each model is a different classifier, which may predict at different times if a certain clinically important threshold is exceeded.

The results of the confusion matrix can be used to create receiver operating characteristic (ROC) curves. A ROC curve plots the true positive rate (sensitivity) against the false positive rate (1 – specificity), for different discrimination thresholds. Points above the diagonal represent results that are better than random, and the best possible predictor would result in a ROC curve that consists of a point at the upper left corner, where sensitivity and specificity both equal 1.

When choosing between different classifiers, generally, the one with the largest area under the curve (AUC) has the best overall performance (Hanley and McNeil, 1982). If the AUC is similar, specific characteristics of the ROC curves may be considered. For a given AUC, one classifier might perform better when a high sensitivity is important, whereas a different classifier might perform better when high specificity is required. In a medical application, a highly sensitive test might be favoured for initial screening purposes, so that presence of a disease is rarely missed. Further tests might be more specific so as to avoid unnecessary treatment (Lalkhen and McCluskey, 2008).

For a given classifier, the optimal threshold can be calculated as the point on the ROC curve closest to (0, 1). A similar method, the Youden index, is the point on the ROC curve that is the furthest from the diagonal 1:1 line. In the context of optimal threshold choice, the costs of false positives and false negatives must again be considered, and setting the threshold to create either a more sensitive or more specific test might be appropriate (Schisterman *et al.*, 2005).

Chapter 4 Pulmonary Data

4.1 Cohort1

Cohort1 came from a clinical study that was conducted from 2000 to 2002 in the eight German hospital ICUs (Stahl *et al.*, 2006). Measurements were taken from 28 ARDS patients, aged 17 to 77 years. The cause of ARDS varied among patients, and included pneumonia, lung contusion, and brain injury. Patients were ventilated for varying lengths of time, with the minimum being two days.

The volume controlled mode of ventilation was used, with the tidal volume targeted at 8 ± 2 mL/kg body weight. Inspiratory time and flow rate were set to obtain and end-inspiratory pause of ≥ 0.2 seconds. The breathing rate varied. Sedatives were titrated to achieve a Ramsay sedation score of 4 – 5, and neuromuscular blocking drugs were administered as needed. Therefore, patient breathing efforts did not occur.

Patients were ventilated with Evita4Lab systems. Flow was measured with a calibrated Fleisch No. 2 pneumotachograph connected to a differential pressure transducer. Pressure at the airway opening was measured by a piezoresistive pressure transducer. Volume was calculated by continuous integration of the flow with adjustment for volume drift, so that volume at the end of expiration was zero. Data was sampled at 125 Hz, and then later reduced to 62.5 Hz to reduce noise in the data and improve the speed of simulations. This was achieved by decimating the signal by a factor of two. Low pass filtering was not required as the data was greatly oversampled according to the Nyquist criterion (Lyons, 2001).

Patients were first ventilated at zero PEEP for approximately five minutes. Then, 27 of the 28 patients underwent a RM in which PEEP was increased in steps of 2 cmH_2O after every 10 breaths. This process continued until a peak inspiratory pressure of approximately 50 cmH_2O was reached (Figure 4.1). The protocol was approved by the local ethics committee of each participating institution, and informed consent was obtained from the patient or their legally authorised representative.

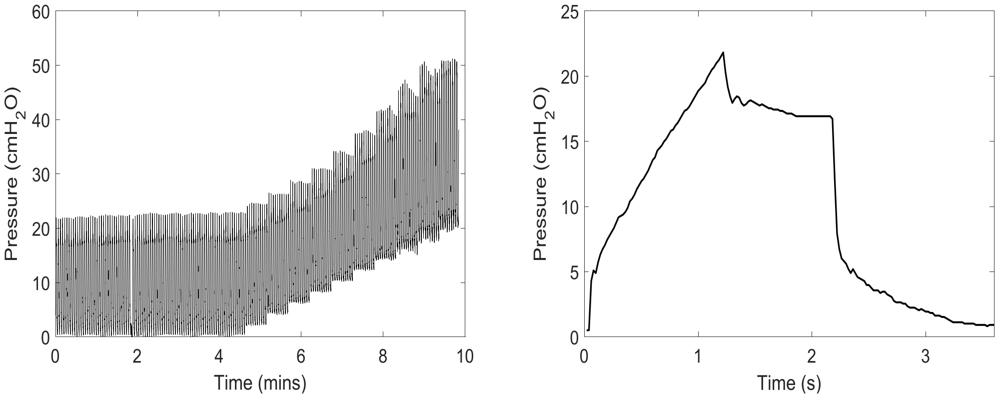


Figure 4.1. Airway pressure data for one patient of Cohort1. Full RM (left), and a single breath at PEEP = 0 cmH$_2$O (right).

4.2 Cohort2

Cohort2 came from a pilot Clinical Utilisation of Respiratory Elastance (CURE) software trial, conducted in the Christchurch Hospital ICU, New Zealand, between September 2013 and November 2014 (Davidson *et al.*, 2014). Measurements were collected from ten ARDS patients, aged 18 to 88 years with a mean of 50.3 years. Diagnoses included pneumonia, aspiration, and pancreatitis.

Pressure and flow were recorded with a Puritan Bennett 840 ventilator at a sampling rate of 50 Hz. Volume was calculated from continuous integration of the flow, with compensation for volume drift. Seven of the patients were ventilated in pressure controlled mode, and three in volume controlled mode. No end-inspiratory pause was used. The breathing rate was approximately 18 breaths per minute. Patients were fully sedated.

The patients underwent RMs in which PEEP was increased in steps of 2 cmH$_2$O. PEEP at the beginning of the RM varied between 8 cmH$_2$O and 16 cmH$_2$O for different patients. The RMs contained between four and nine PEEP step increases. The maximum pressure reached for each patient ranged between 36 cmH$_2$O and 52 cmH$_2$O. Some patients underwent multiple RMs during the trial, thus 19 data sets were obtained. Figure 4.2 shows one of the data sets. Ethics approval for the study and use of collected data was granted by the New Zealand South Regional Ethics Committee.

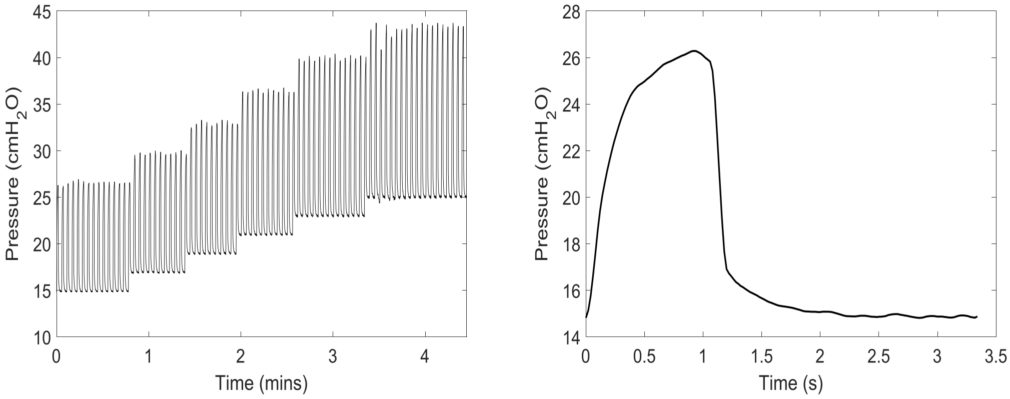


Figure 4.2. Airway pressure data for one patient of Cohort2. Full RM (left), and a single breath at PEEP = 15 cmH$_2$O (right).

One further data set was obtained from this study. This patient was ventilated using a SIMV ventilator in volume controlled mode, and exhibited rarely seen M-waves in the pressure data due to spontaneous breathing efforts (Figure 4.3). The data has a constant PEEP of 17 cmH$_2$O and spans approximately 80 minutes. Around 65% of the breathing cycles contained some degree of M-wave shape.

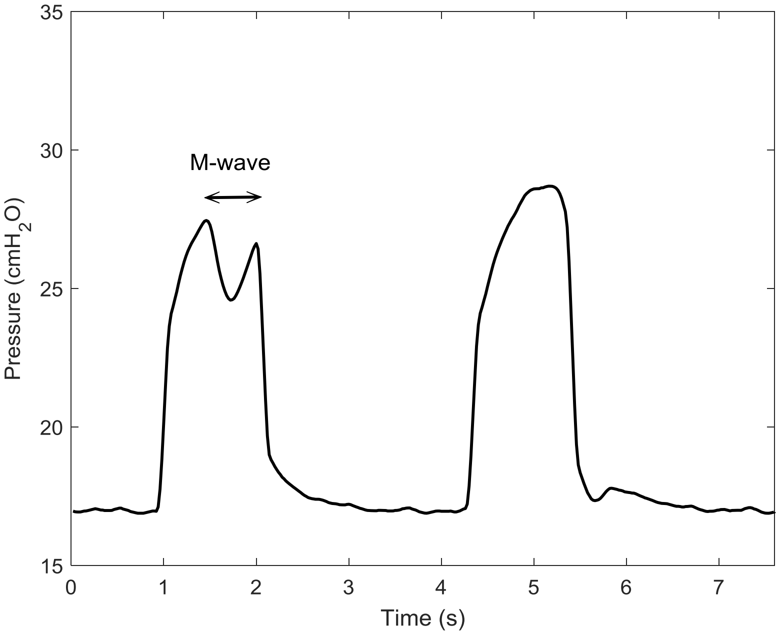


Figure 4.3. Cohort2 M-wave patient pressure data showing an M-wave and a normal breath.

4.3 Cohort3

Cohort3 came from a study conducted in the Christchurch Hospital ICU, New Zealand, between February 2010 and September 2010 (Sundaresan *et al.*, 2011). Measurements were collected from ten ARDS patients, aged 22 to 88. Four patients had a high auto-PEEP, which is a condition caused by incomplete expiration resulting in accumulation of air in the lungs. High auto-PEEP was defined as an auto-PEEP ≥ 5 cmH_2O. The remaining six patients had low auto-PEEP < 5 cmH_2O. All patients with high auto-PEEP were independently diagnosed with COPD, as was one patient with low auto-PEEP. Table 4.1 summarises patient information, including diagnoses and measured auto-PEEP.

Patients were ventilated in volume controlled mode, with a constant tidal volume of 6 ml/kg or 500 ml, whichever was lower. No end-inspiratory pause was used. The breathing rate varied between patients. Muscle relaxants were used to prevent spontaneous breathing efforts.

Patients were ventilated with a Puritan Bennett PB840 ventilator, and a Hamilton Medical flow sensor was attached to a calibrated pneumotachometer to obtain pressure and flow measurements at a sampling rate of 50 Hz. Volume was calculated by continuous integration of the flow, with compensation for volume drift.

At PEEP = 0 cmH_2O, an expiratory hold was performed to measure auto-PEEP. In this manoeuvre, the exhalation valve is closed so that the flow is stopped, and the airway pressure rises as alveolar and airway pressure equilibrates. The measured pressure at the airway opening is then the auto-PEEP. After the expiratory hold manoeuvre, a RM was performed, in which PEEP began at 0 cmH_20, and was increased in steps of 5 cmH_2O. Patients underwent PEEP increases between three and six times. Peak airway pressure was limited to 45 cmH_2O. Data was available for one breath per PEEP level, after steady state was achieved. Breaths at each PEEP were concatenated to obtain the data set for each patient (Figure 4.4). The study was approved by the Upper South Island Regional Ethics Committee.

Table 4.1. Cohort3 Patient Information.

Patient	Sex	Age (years)	Diagnosis	Auto-PEEP (cmH_2O)
1	F	61	Peritonitis, COPD	7
2	M	22	Trauma	2
3	M	55	Aspiration	0
4	M	88	Pneumonia, COPD	9
5	M	59	Pneumonia, COPD	8
6	M	69	Trauma	2.3
7	M	56	Legionnaires	2
8	F	45	Aspiration	0
9	M	37	H1N1, COPD	12
10	M	56	Legionnaires, COPD	3

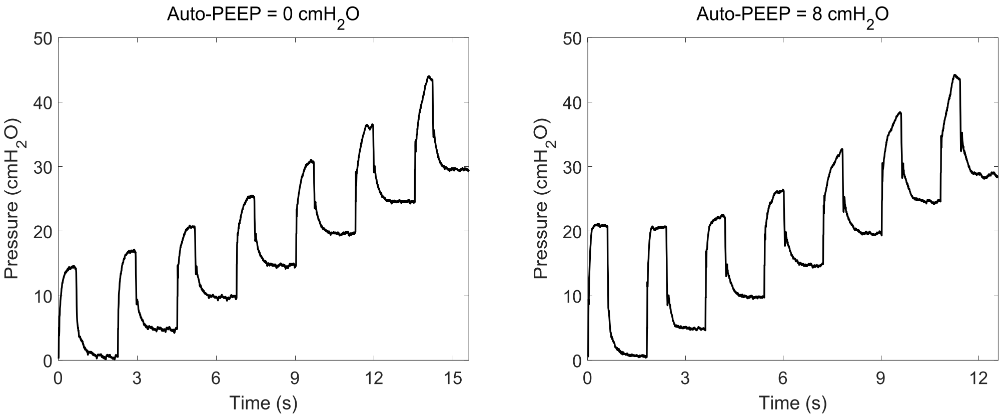


Figure 4.4. Cohort3 pressure data for one low auto-PEEP patient (left), and one high auto-PEEP patient (right).

4.4 Summary

All three data sets provide pressure and flow data from ARDS patients. Across the three cohorts, many different patient conditions are represented. There is a wide range of patient ages, diagnoses, and time spent on MV. Cohort3 in particular had a large number of COPD patients, which were much less common in Cohort1 and Cohort2. The cohorts also differ in the specific ventilation protocol and RM method, and both pressure controlled and volume controlled modes are represented.

Cohort1 is the most complex due to the administration of the end-inspiratory pause, and the longer RMs that generated airway pressure ranges from zero to 50 cmH$_2$O. Cohort1 provides a wealth of features for a potential new model to capture. However, it is an older data set, collected between 2000 and 2002, whereas the Cohort2 and Cohort3 data was collected much more recently (2013 and 2010 respectively). Cohort2 and Cohort3 are less complex but are useful as they may better represent modern ICU equipment and practices.

These data sets will be used throughout the thesis to develop a pulmonary model, validate model components, and assess predictive performance. A model capable of handling all patient conditions and ventilation scenarios present in the data would potentially be of significant clinical value.

Part 2: NARX Model

Chapter 5 NARX Model Design

5.1 Introduction

This chapter describes the development of the initial form of the NARX model. A justification of each of the NARX model terms is provided, and the result of fitting the NARX model to the Cohort1 data is shown. The FOM was used as a comparison for these results. The content of this chapter was published in Langdon *et al.* (2016b).

5.2 Model Description

The NARX model used here was based on the structure of the FOM. The FOM was adapted with pressure dependent b-spline basis functions, and time dependent resistance coefficients:

$$P(t) = \sum_{i=1}^{M} a_i \emptyset_{i,d}\big(P(t)\big)V(t) + \sum_{j=0}^{L} b_j \dot{V}(t_{-j}) + P_0(t) \qquad 5.1$$

where: a_i and b_j are the parameters to be identified. M is the number of basis-functions to be used, i is the index of a particular basis function of degree d, a_i is the coefficient for a given basis function, and $\emptyset_{i,d}\big(P(t)\big)$ is the basis function value for a given pressure measurement. The sum of the basis functions multiplied by their a_i coefficients defines elastance.

$$E(P) = \sum_{i=1}^{M} a_i \emptyset_i(P) \qquad 5.2$$

There are $j = 0 \ldots L$ b_j coefficients that capture the effect of airway resistance and pressure responses that occur due to changes in flow (e.g. pressure equalisation in the relaxation phase). The subscript *-j* in the second term refers to the previous time sample. Thus, each *P(t)* is calculated from information from the previous *L* data points. The FOM is replicated with *M* = *L* = 1, and *d* = 0.

Equation 5.1 can be rewritten over time to generate a matrix system of equations, and the coefficients (a_i, b_j) are determined by linear least squares.

$$\mathbf{Ax} = \mathbf{b} \qquad 5.3$$

where:

$$\mathbf{A} = \begin{bmatrix} \emptyset_{1,d}(P_{aw}(t_0))V(t_0) & \emptyset_{2,d}(P_{aw}(t_0))V(t_0) & \ldots & \emptyset_{M,d}(P_{aw}(t_0))V(t_0) & \dot{V}(t_0) & \dot{V}(t_{-1}) & \ldots & \dot{V}(t_{-L}) \\ \emptyset_{1,d}(P_{aw}(t_1))V(t_1) & \emptyset_{2,d}(P_{aw}(t_1))V(t_1) & \ldots & \emptyset_{M,d}(P_{aw}(t_1))V(t_1) & \dot{V}(t_1) & \dot{V}(t_0) & \ldots & \dot{V}(t_{1-L}) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \emptyset_{1,d}(P_{aw}(t_N))V(t_N) & \emptyset_{2,d}(P_{aw}(t_N))V(t_N) & \ldots & \emptyset_{M,d}(P_{aw}(t_N))V(t_N) & \dot{V}(t_N) & \dot{V}(t_{N-1}) & \ldots & \dot{V}(t_{N-L}) \end{bmatrix};$$

$$\mathbf{b} = \begin{bmatrix} P_{aw}(t_0) \\ P_{aw}(t_1) \\ \vdots \\ P_{aw}(t_N) \end{bmatrix}; \text{ and } \mathbf{x} = \begin{bmatrix} a_1 \\ \vdots \\ a_M \\ b_1 \\ \vdots \\ b_L \end{bmatrix}.$$

This approach contrasts substantially with existing pulmonary models. In particular, the treatment of elastance as a function of pressure using basis functions has not previously been undertaken.

5.3 Model Justification

Appropriate values for *L*, *d*, and *M* were determined, and the results are presented in this section. Undertaking a full grid search process was not necessary. Residuals were not the only factor considered in optimal parameter selection, and the consequences of varying *L*, *d*, and *M* were independent. In this section, each model component is justified separately, with the other parameters used in the final NARX model formulation. Ultimately, $L = 350$, $d = 1$, and $M = 5$ were the most appropriate values.

Two out of the 27 data sets in Cohort1 were excluded from the optimisation process, due to outlier behaviour. The two patients exhibited no end-inspiratory pressure changes to increases in PEEP up to 8 cmH$_2$O. After a PEEP of 8 cmH$_2$O, the end-inspiratory pressure increased with PEEP as expected. This behaviour was caused by an unusually high auto-PEEP of around 8 cmH$_2$O. It is likely that this behaviour was caused by chronic obstructive pulmonary disease (COPD). For an individual with COPD it is likely that blockages on the bronchial pathway would be altered during a RM. Thus the assumption that patient state would not change during the clinical protocol was violated. These patients were excluded to remove their influence on optimal model selection.

5.3.1 Justification of *L*

The value of *L* was varied from 1 to 500 in steps of 10 ($L = [1, 10, 20, \ldots, 500]$), with constant values of $M = 5$, and $d = 1$. Convergence was assumed when the RMS residuals reduced by less than 0.5% for each step in *L*. Table 5.1 shows the *L* value required to meet this condition for each of the 25 patient data sets. Table 5.1 also records the *L* value as a percentage of the number of data points in one breath for each patient. Overall, Table 5.1 indicates that there were generally very small reductions in the RMS residuals for $L > 350$. The residuals at $L = 350$ have zero bias overall (one sample t-test $p < 0.05$). The cumulative distribution (CDF) plot also indicates a difference between *L* = 1, *L* = 100, and *L* = 350, but no further improvement from increasing *L* to 500 (Figure 5.1).

The order of the model may be considered large, however the number of data points per data set was on the order of 10^4. Thus, the number of parameters used for fitting the model is small compared to the number of data points. Furthermore, the measurement noise was estimated using a steady state region at the end of expiration of a breath. The CDF plot (Figure 5.2) shows that the NARX model residuals are significantly larger than the estimated measurement noise. Therefore we are confident that the NARX model is not over-parameterised and is not fitting to noise.

Table 5.1. L value required for RMS residual convergence. M = 5, d = 1.

Patient	*L* required for RMS convergence	% of one breath
1	200	101
2	310	99
3	160	106
4	320	102
5	260	104
6	350	102
7	190	101
8	230	98
9	270	100
10	190	102
11	210	100
12	380	101
13	380	121
14	240	102
15	180	106
16	270	101
17	250	120
18	380	101
19	200	106
20	270	101
21	190	126
22	260	104
23	210	123
24	210	100
25	290	93

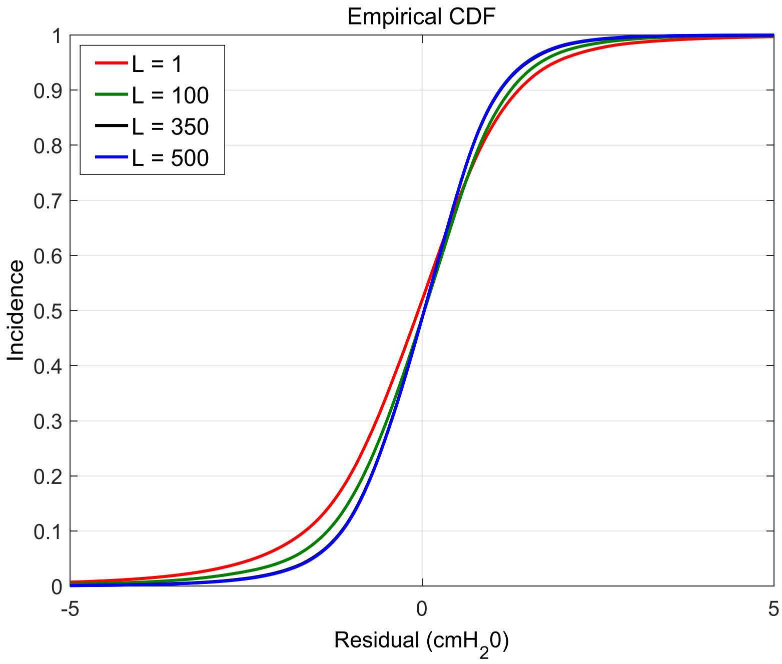


Figure 5.1. CDFs for the residuals of all 25 patients for L = 1, 100, 350, 500. M = 5, d = 1.The lines for L = 350 and L = 500 coincide.

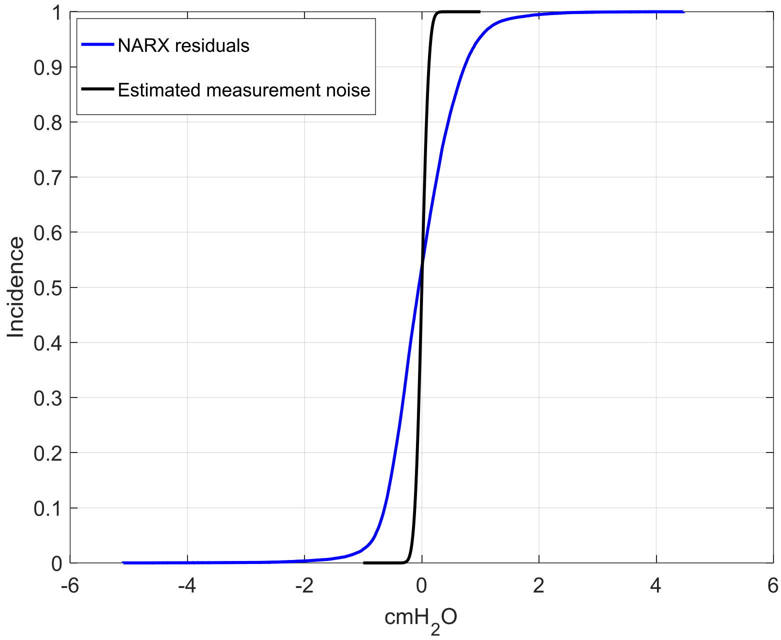


Figure 5.2. CDF of the estimated measurement noise, and the NARX residuals for all 25 patients. L = 350, M = 5, d = 1.

5.3.2 Justification of *d*

Zeroth, first, and second order basis functions were tested in the NARX model with constant values of *L* = 350, and *M = 5*. Table 5.2 shows that using first order basis functions reduce the RMS residuals by an average of 8.9% compared to zeroth order basis functions. There was on average only a 0.6% difference between the RMS residuals for first and second order basis functions. The paired signed-rank test confirmed that there were significant changes in the RMS values across zeroth and first order basis functions. However, the changes in RMS from first to second order basis functions were insignificant at $p = 0.05$. Figure 5.3 shows that there is little variation in the residuals for first and second order basis functions.

Figure 5.4 shows an example of the zeroth, first, and second order basis functions multiplied by their coefficients at each point. This figure again shows that first and second order basis functions produced a similar result, whereas the same number of zeroth order step functions was unable to capture the same elastance behaviour.

Table 5.2. RMS residuals (cmH_2O) for zeroth, first, and second order basis functions. $M = 5$, $L = 350$.

Patient	$d = 0$	$d = 1$	$d = 2$
1	0.9553	0.8578	0.8632
2	1.0023	0.9770	0.9653
3	0.6498	0.5258	0.5200
4	0.8370	0.7624	0.7587
5	1.0867	0.9712	0.9665
6	0.8003	0.7617	0.7605
7	1.0068	0.9568	0.9549
8	1.2049	1.1573	1.1612
9	0.9570	0.8942	0.9098
10	0.8442	0.7487	0.7491
11	1.2181	1.1894	1.1854
12	1.2362	1.1938	1.1932
13	1.7830	1.6986	1.7025
14	1.2728	1.0441	1.0348
15	2.3161	2.2983	2.2990
16	2.0036	1.7150	1.7045
17	1.3103	1.2568	1.2558
18	1.7823	1.7648	1.7610
19	0.7187	0.6692	0.6728
20	0.8891	0.8322	0.8365
21	0.8079	0.7239	0.7172
22	0.8524	0.7705	0.7694
23	0.8906	0.7866	0.7851
24	0.9510	0.9190	0.9192
25	0.9775	0.8632	0.8761
Paired signed-rank	0.00001	0.34	

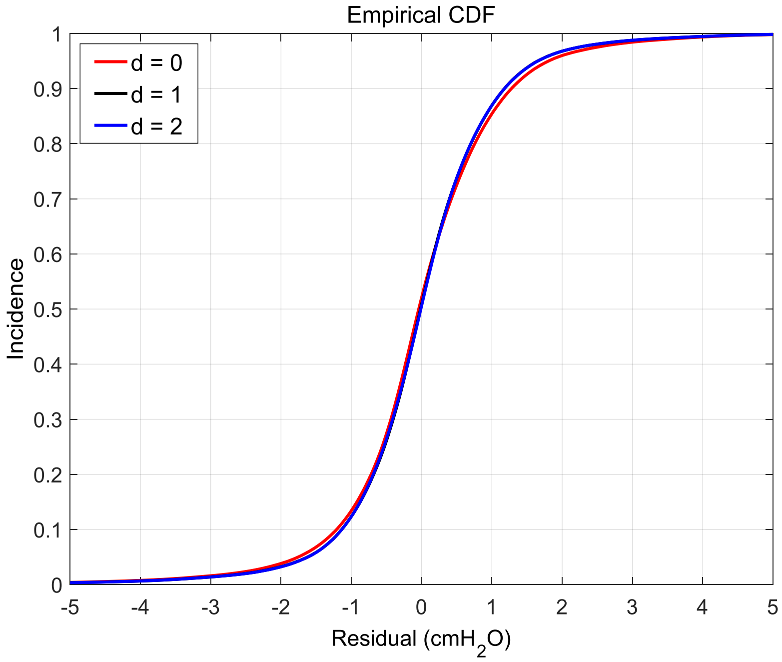


Figure 5.3. CDFs for the residuals of all 25 patients for d = 0, 1, 2. M = 5, L = 350.

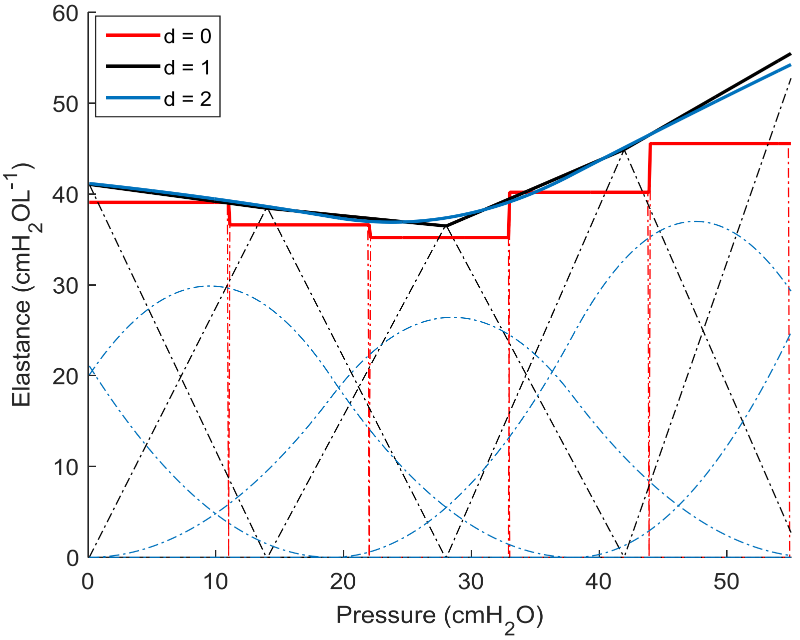


Figure 5.4. Elastance through pressure for one patient, d = 0, 1, 2. M = 5, L = 350.

5.3.3 Justification of *M*

The number of basis functions was tested with *M* = [2, 5, 10, 15], with constant values of *L* = 350, and *d* = 1. Table 5.3 shows that there was on average a 5.5% reduction in RMS residuals between *M* = 2 and *M* = 5, a 3.2% reduction between *M* = 5 and *M* = 10, and 1.2% between *M* = 10 and *M* = 15. The paired signed-rank test found significant changes in RMS across increasing resolution from 2 to 5, from 5 to 10, and from 10 to 15 ($p < 0.05$). However, Figure 5.5 shows that the CDFs for $M = 5$, 10, and 15 are almost identical.

While M had a significant effect on the size of the RMS residuals above $M = 2$, the magnitude of the changes was minimal. Hence, the stability and descriptive interpretation of the a_i coefficients were also investigated to determine an appropriate value for M. In particular, it is expected that in the range of typical MV, patient elastance will be a relatively smooth function of pressure. Figure 5.6 shows that the elastance shape becomes more erratic and unstable for $M = 10$ and 15, compared with $M = 5$. This behaviour does not match the expected mechanical and physiological behaviour of ARDS lungs during PEEP step increases, and suggests the model becomes practically non-identifiable with $M = 10$ and $M = 15$. Ultimately, a value of $M = 5$ was thus chosen as a reasonable value for the 25 patient data sets.

Table 5.3. RMS residuals (cmH$_2$O) for $M = 2, 5, 10, 15$. $L = 350$, $d = 1$.

Patient	$M = 2$	$M = 5$	$M = 10$	$M = 15$
1	1.0951	0.8578	0.8353	0.8378
2	1.0335	0.9770	0.9287	0.9043
3	0.8577	0.5258	0.5209	0.5229
4	0.9798	0.7624	0.7266	0.7281
5	1.2759	0.9712	0.9386	0.9338
6	0.9078	0.7617	0.7501	0.7521
7	1.0858	0.9568	0.9496	0.9337
8	1.3733	1.1573	1.1433	1.1284
9	1.1076	0.8942	0.8369	0.8390
10	0.9285	0.7487	0.7433	0.7451
11	1.2483	1.1894	1.1662	1.1644
12	1.2367	1.1938	1.1692	1.1592
13	2.0653	1.6986	1.6821	1.6804
14	1.3822	1.0441	1.0258	1.0217
15	2.3632	2.2983	2.2703	2.2587
16	1.8756	1.7150	1.5120	1.4746
17	1.2676	1.2568	1.2500	1.2486
18	1.7782	1.7648	1.7354	1.7308
19	0.8159	0.6692	0.6603	0.6643
20	0.9903	0.8322	0.8091	0.7995
21	0.9891	0.7239	0.7029	0.6940
22	0.8073	0.7705	0.7629	0.7656
23	0.8266	0.7866	0.7854	0.7984
24	1.0105	0.9190	0.9009	0.9005
25	1.1616	0.8632	0.8540	0.8563
Paired signed-rank	0.00001	0.00001	0.09	

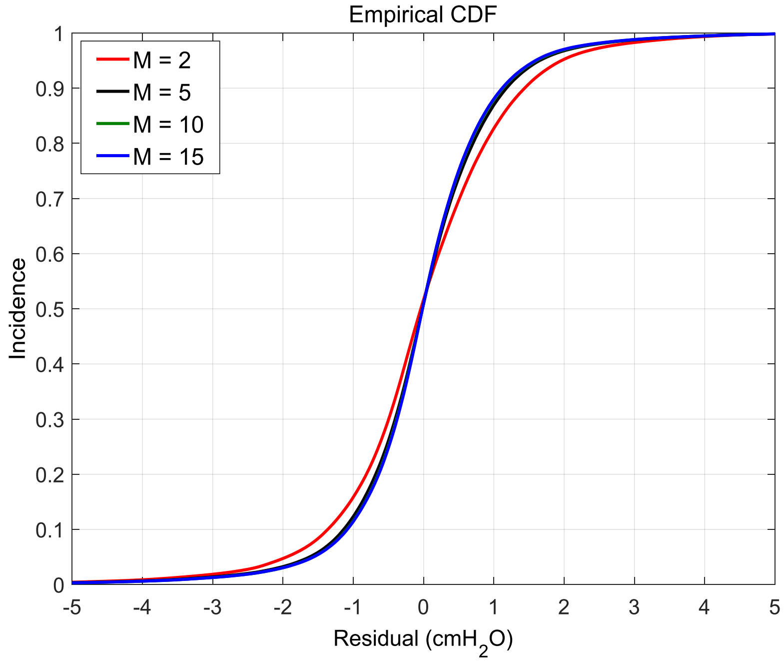


Figure 5.5. CDFs for the residuals of all 25 patients for M = 2, 5, 10, 15. L = 350, d = 1.

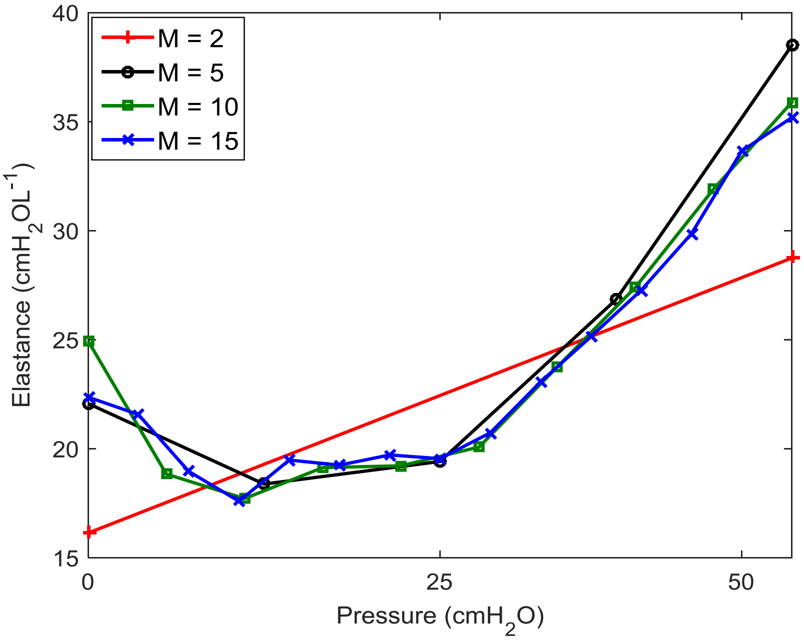


Figure 5.6. Elastance for one patient for M = 2, 5, 10, 15. L = 350, d = 1.

5.4 Modelling Results

The NARX model and FOM were identified on the entire 8 – 10 minutes of each data set in Cohort1. The NARX mean RMS residual was 1.05 (90% CI: 1.02 – 1.07) cmH_2O. The FOM mean RMS residual was 1.87 (90% CI: 1.83 – 1.91) cmH_2O. The two sample K-S test implied that the FOM and NARX model RMS residuals were from different distributions with $p < 0.01$. The CDF plots also demonstrate that residuals are reduced when the NARX model is used, compared with the FOM (Figure 5.7).

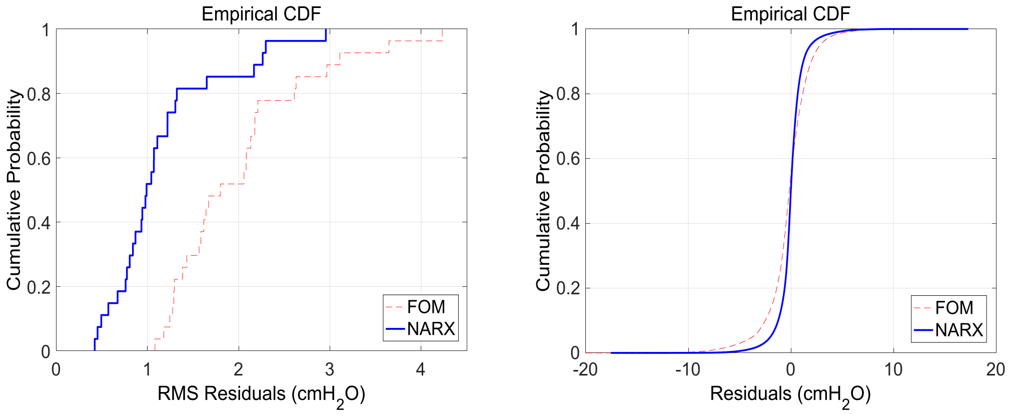


Figure 5.7. CDF for the RMS residual values (left), and exact residuals (right) for all 27 data sets.

For one patient, a breath that is representative of the whole data set was plotted by averaging all of the breaths. The mean residuals relative to the breath were superimposed at uniformly distributed points (Figure 5.8). The same process was carried out for particular PEEP steps (Figure 5.9). The FOM was unable to fit well to the expiratory part of the breath, or to the relaxation during the end-inspiratory pause. The NARX model was able to capture all parts of the breath equally well. The NARX model was also able to fit the data at each PEEP step equally well, whereas the FOM was unable to match the tidal pressure change as PEEP increased.

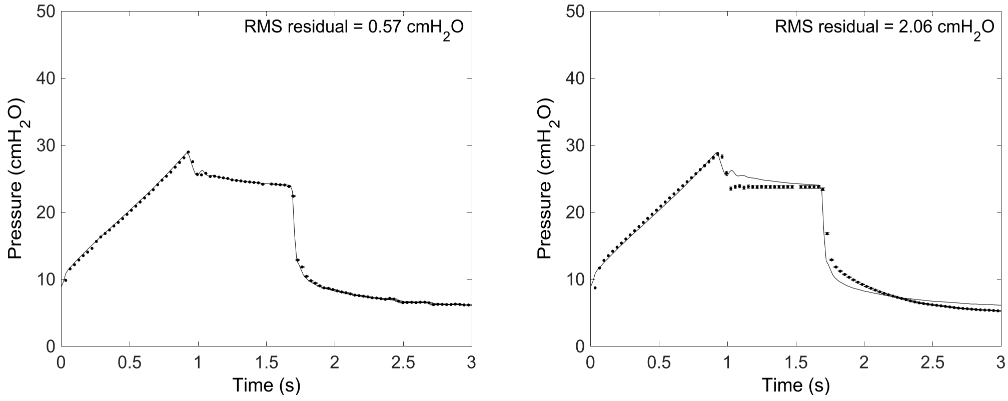


Figure 5.8. Mean residuals relative to one patient's average breath at all PEEP for the NARX model (left) and FOM (right). Error bars are the standard error.

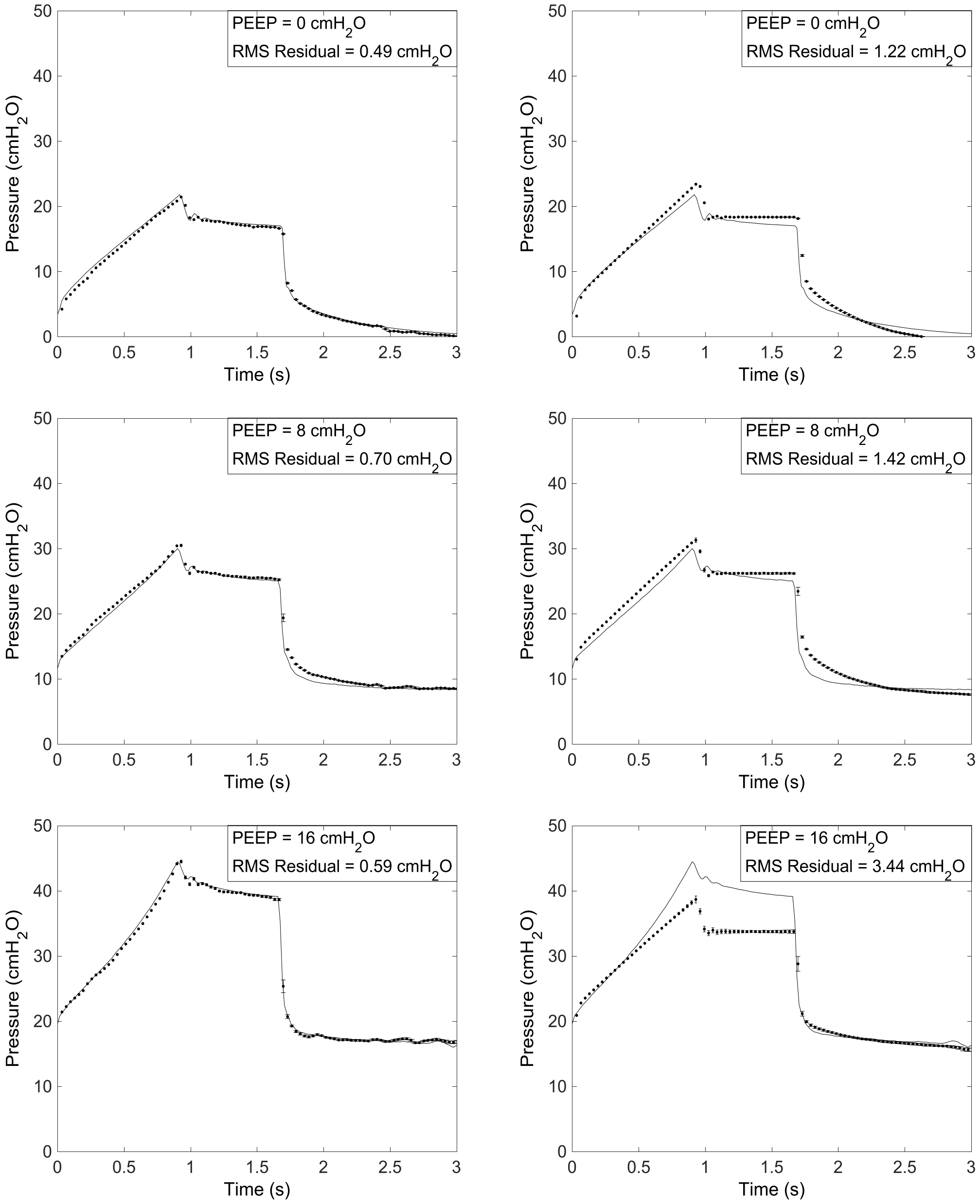
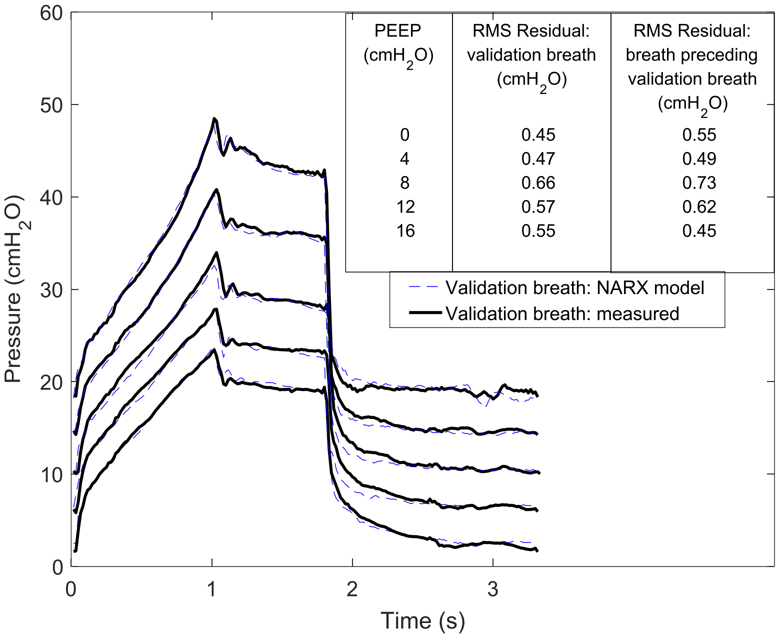


Figure 5.9. Mean residuals plotted relative to one patient's average breath at PEEP = 0, 8, 16 cmH₂O for the NARX model (left), and FOM (right).

As a further validation of the NARX model, the model was identified while excluding the final breath of each PEEP level. The measured data was them compared to the forward simulation of the model for those breaths. Figure 5.10 shows this result for one patient at a PEEP of 0, 4, 8, 12, and 16 cmH$_2$O, and the resulting RMS residuals for these individual breaths. The RMS residuals at each PEEP are comparable (within 0.05 cmH$_2$O) to the RMS residuals obtained when fitting the NARX model to the entire data set.



PEEP (cmH$_2$O)	RMS Residual: validation breath (cmH$_2$O)	RMS Residual: breath preceding validation breath (cmH$_2$O)
0	0.45	0.55
4	0.47	0.49
8	0.66	0.73
12	0.57	0.62
16	0.55	0.45

Figure 5.10. NARX model validation for one patient over a single breath at PEEP = 0, 4, 8, 12, 16 cmH$_2$O.

5.5 Discussion

The NARX model with parameters L = 350, d = 1, and M = 5 successfully fit 25 of the patient data sets with very low unbiased residuals compared to the FOM. The model also successfully predicted airway pressure at each PEEP for single breaths that were excluded from the training data sets (Figure 5.10). The model was able to describe all parts of the breath, in particular the end-inspiratory pause and expiratory relaxation, where the simple FOM typically fails to capture this nonlinear patient-specific behaviour. The NARX model was also able to fit inspiration, expiration, and relaxation during the end-inspiratory pause simultaneously, quickly, and easily. These capabilities exceed all other current identifiable models.

The NARX model was identified with linear least squares, which is a very simple and robust parameter ID method. Though the model is complex, the ID method is very simple compared to other complex models, such as the PRM (Schranz *et al.*, 2012b). It is thus much easier to apply within a computational framework for real-time use at the bedside. Other models with simple ID methods have generally lacked the parameterisation to capture all of the behaviour that the NARX model can, indicating the increased potential of this model and method.

The NARX model uses basis functions as a simple way to capture the nonlinear pressure dependent respiratory system elastance. Nonlinear elastance is related to the idea that different compartments of the lungs are recruited as airway pressure is increased, as proposed by (Hickling, 1998), and supported by multiple *in vivo* studies (Carney *et al.*, 1999, Schiller *et al.*, 2003). ARDS lungs commonly exhibit a sigmoidal pressure – volume response due to the recruitment phases of collapsed alveoli, and over-distension occurring at a lower volume compared to healthy lungs (Figure 1.3). This behaviour is concurrent with the convex elastance / pressure relationship (Figure 5.4) which was typical of all results.

Capturing nonlinear elastance is useful because it can be used to help identify the pressures at which there is a higher elastance that corresponds to a stiffer, less healthy lung. Setting PEEP at the point that results in the lowest elastance could potentially give the best compromise between recruitment and distension (Carvalho *et al.*, 2007). Thus, finding this elastance curve over a RM could result in setting PEEP at an optimal patient-specific level.

The *E(P)* curve enables the NARX model to fit the data well across all PEEP steps. As PEEP increases to the range where over-distension occurs, the tidal pressure increases, and the shape of the breath also changes. The inspiratory curve changes from concave downwards at low PEEP (Figure 5.9, PEEP = 0) to concave upwards at high PEEP (Figure 5.9, PEEP = 16). The NARX model fit equally well to the zero PEEP data as to the high PEEP data.

In contrast, the tidal pressure defined by the FOM was dependent on tidal volume alone, because of the single constant elastance term. Since a full 50% of each dataset was at zero end-expiratory pressure (ZEEP), the elastance term was identified such that the FOM tidal pressure was closer to the data tidal pressure at zero PEEP. This meant that at high PEEP the FOM was unable to come close to the tidal pressure in the data. Since the NARX model varies elastance with pressure, the tidal pressure at all PEEP levels can be matched. The FOM was not designed to capture breaths at different PEEP levels, and is thus more typically evaluated on a breath by breath basis. Thus, the comparison is made in an environment for which the FOM was not intended.

The NARX model uses 350 terms to capture the response to flow. When L is large, the b_j coefficients no longer strictly represent resistance. Instead they are model terms that have a net effect similar to resistance. These terms are able to capture the viscoelasticity, or gas redistribution phenomenon that happens during the end-inspiratory pause. During an inspiratory pause, when flow is stopped but expiration is not allowed to begin, the air will redistribute in the lungs over time, leading to an exponentially decreasing pressure. The NARX model is able to fit well to this nonlinear part of the breath, fitting both the exponential decrease and the oscillations. The NARX b terms also help the model fit during expiration (Figure 5.9), which is also a relaxation process.

L, *M*, and *d* were varied to determine values that give a best fit to the data. For different data and different ventilation modes, these model parameters may be customised if necessary, or potentially evolved for each patient. The large *L* was primarily useful for capturing end-inspiratory relaxation. Therefore, if the data had not contained this end-inspiratory pause, a large number of *b* coefficients in the model could potentially have caused the model to suffer from non-identifiability, requiring *L* to be reduced. $M = 5$ was chosen as the optimal number to capture an elastance curve while also avoiding non-identifiability. The data sets used in this study contained a pressure range from 0 cmH_2O to 50 cmH_2O. However, if a larger or smaller pressure range existed, *M* may have to be suitably adjusted.

The NARX approach did not produce particularly good outcomes for two of the 27 patients. When finding optimal model coefficients *L, M, d,* these two data sets were excluded due to the potential for this outlier behaviour to influence optimal model selection. The two patients had a high auto-PEEP, most likely caused by COPD. The model had to account for the lack of increase in peak pressure during early recruitment by failing to capture the pressure at other times. These two patients highlight the need to monitor model fits during clinical practice to ensure that changes in patient state are captured.

In Cohort1, the patients were sedated and given neuromuscular blocking drugs. If patients were less heavily sedated, their intermittent breathing efforts would result in anomalies in the breathing waveforms, and thus mask the underlying respiratory mechanics. The NARX model would be unable to identify useful elastance and resistance parameters under these conditions.

5.6 Summary

Basis functions were incorporated into a nonlinear autoregressive model, and the model was successful at describing all features of the airway pressure curve, for 25 patient data sets. By selecting a suitable number of appropriate basis functions, an elastance vs. pressure curve was able to be described for each patient. Coefficients that captured airway resistance effects enabled the end-inspiratory and expiratory behaviour to be described with very low residuals compared to the FOM.

While it is expected that a more complex model with more parameters will result in reduced residuals, the reduction in residuals compared to the FOM is a worthwhile result as the NARX parameters were robust. Figure 5.2 shows that the model is not fitting to noise, and Figure 5.10 shows that accurate predictions occurred when the model was validated on breaths that were not used to identify the model.

Chapter 6 Interpolation

6.1 Introduction

This section presents an extended validation of the NARX model, where the NARX model is identified on the first and last 20% of data, and then interpolated to cover the middle 60%. This is valuable because complete data for the full range of PEEP steps may not always be available for each patient. Having a method to estimate patient-specific respiratory mechanics when data is not available would provide clinicians with a more complete understanding of the individual patient's condition, and could potentially aid the selection of optimal PEEP or other MV settings. The 25 patients from Cohort1 were used, and the FOM provided a comparison for the results. The content of this chapter was published in Langdon *et al.* (2015b).

6.2 Methods

The FOM and NARX model coefficients were identified using 40% of the available data for each patient. This identification data (IDD40) was composed from the concatenation of the first 20% and the last 20% of data. The pressures present in the IDD40 covered the full range of 0 – 50 cmH_2O, which provided adequate information for the identification of all five b-spline basis function coefficients. The IDD40 consisted of approximately 100 seconds of zero PEEP, and data from the highest PEEP levels encountered. The remaining 60% was used as the evaluation data (EVD) to assess the interpolation capabilities of the models. The EVD consisted of 3 – 4 minutes of zero PEEP data, and approximately six increasing PEEP steps for each patient (Figure 6.1).

For each data set, the models derived from the IDD data were used with the measured volume and flow data from the EVD section to predict EVD pressure. The predicted EVD pressure was then compared to the measured EVD pressure, and the RMS residual was calculated. The RMS residual for the model fit to the IDD section was also calculated. To provide a comparison, the models were also identified on 100% of the data (IDD100).

The mean RMS residual of all data sets was calculated, along with the 90% confidence interval. Overlapping confidence intervals for IDD and EVD sections would imply that there is not a substantial difference in the residuals. Signed-rank tests were also performed, to test for statistically significant differences in RMS residuals at the 5% level.

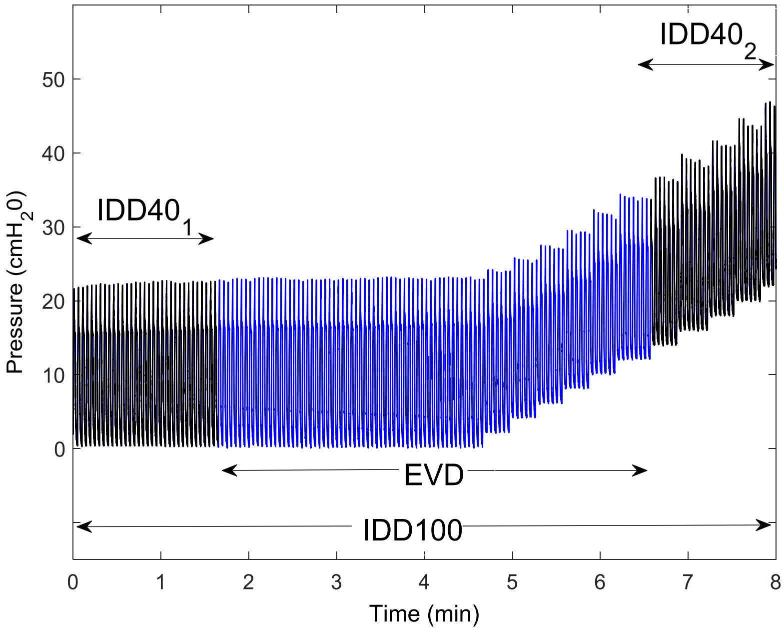


Figure 6.1. Pressure identification data and evaluation data for one patient. IDD40 was composed from the concatenation of IDD40$_1$ and IDD40$_2$.

6.3 Results

Table 6.1 shows the average of the RMS residuals across the 25 patient data sets. This information is also shown graphically in Figure 6.2. The error bars overlap in all three cases for both models. Table 6.2 shows the p values resulting from paired signed-rank tests. There is a significant difference in RMS residuals for most IDD100, IDD40, and EVD combinations at the 5% level.

Table 6.1. Mean RMS residuals and the 90% confidence intervals.

	RMS Mean ± Standard Error and 90% Confidence Interval (CI) (cmH$_2$O)
NARX IDD100	0.93 ± 0.05 (CI: 0.83 – 1.03)
NARX IDD40	1.00 ± 0.06 (CI: 0.88 – 1.12)
NARX EVD	1.08 ± 0.07 (CI: 0.93 – 1.23)
FOM IDD100	1.72 ± 0.10 (CI: 1.53 – 1.91)
FOM IDD40	1.89 ± 0.11 (CI: 1.67 – 2.11)
FOM EVD	1.69 ± 0.10 (CI: 1.49 – 1.89)

Table 6.2. Paired signed-rank test results for RMS comparison, for all patient data sets.

NARX		FOM	
IDD100, EVD	$p < 0.05$	IDD100, EVD	$p = 0.2$
IDD100, IDD40	$p < 0.05$	IDD100, IDD40	$p < 0.05$
IDD40, EVD	$p < 0.05$	IDD40, EVD	$p < 0.05$

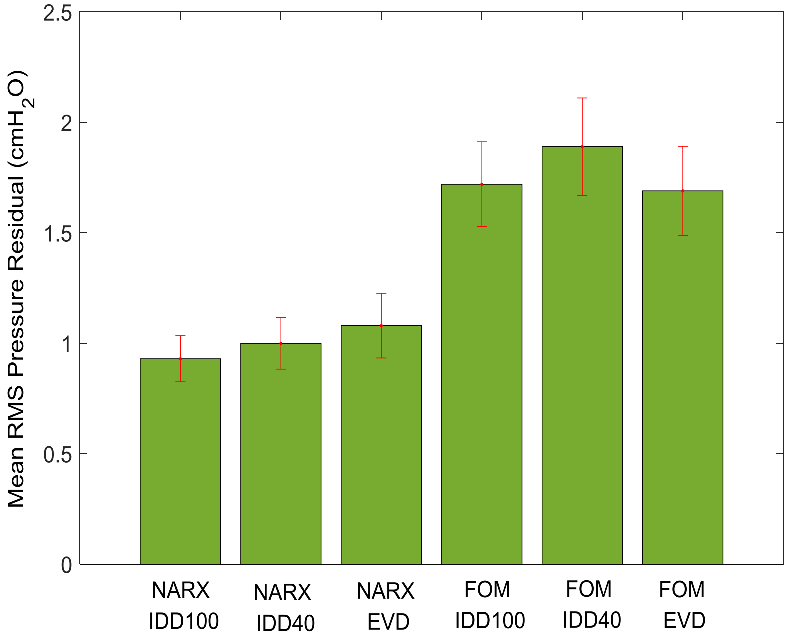


Figure 6.2. Mean RMS residuals and 90% confidence intervals for the NARX model and FOM, calculated from the identification data IDD100 and IDD40, and the evaluation data EVD.

Figure 6.3 to Figure 6.6 present results for two individual patient data sets, A and B. In these figures, the average breath in the IDD or EVD section is plotted. The mean residuals are then plotted relative to the average breath, at regularly spaced points. The error bars are the standard error. The first and second halves of IDD40 are denoted IDD40$_1$ and IDD40$_2$, respectively. Figure 6.7 shows how the elastance for these patients changes over pressure for the NARX model with five basis functions, compared to the FOM with a single elastance coefficient.

The tidal pressure (ΔP) is defined as the difference between the maximum and minimum pressure within a breath. This value is relevant because changes in ΔP are indicative of changes in elastance in volume controlled mode. Both patients A and B began with a ΔP of approximately 22 cmH$_2$O at zero PEEP. For patient A, the ΔP remained constant until PEEP = 16 cmH$_2$O, and then slowly increased to a maximum of 26 cmH$_2$O at PEEP = 24 cmH$_2$O. For patient B, the ΔP was constant until PEEP = 10 cmH$_2$O, then increased rapidly to a maximum of 30 cmH$_2$O at PEEP = 20 cmH$_2$O, at the end of the RM. Each of the other 23 patients showed similar behaviour to either patient A or B.

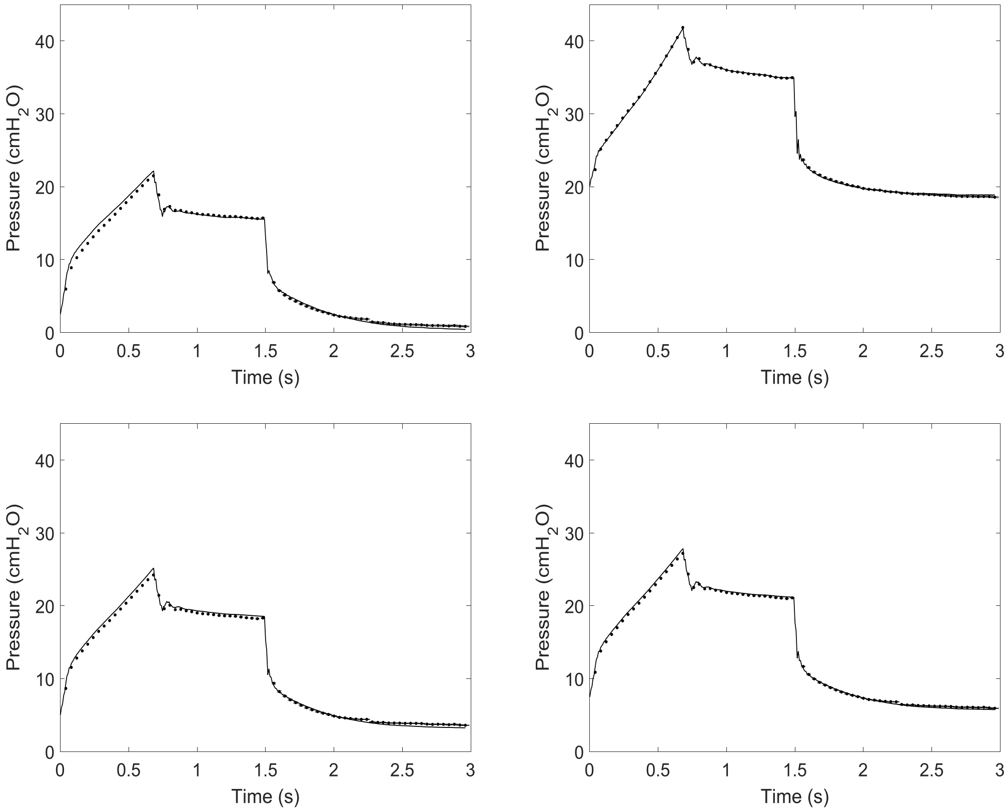


Figure 6.3. Patient A, NARX model: IDD40$_1$ (top left), IDD40$_2$ (top right), EVD (bottom left), IDD100 (bottom right).

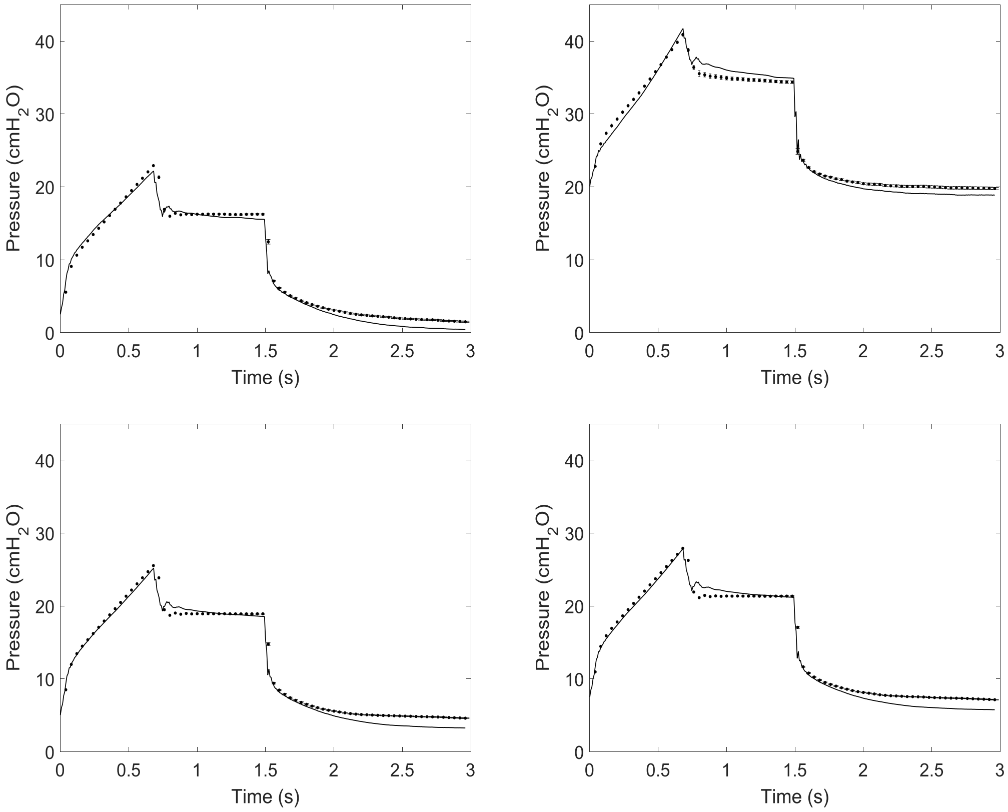


Figure 6.4. Patient A, FOM: IDD40$_1$ (top left), IDD40$_2$ (top right), EVD (bottom left), IDD100 (bottom right).

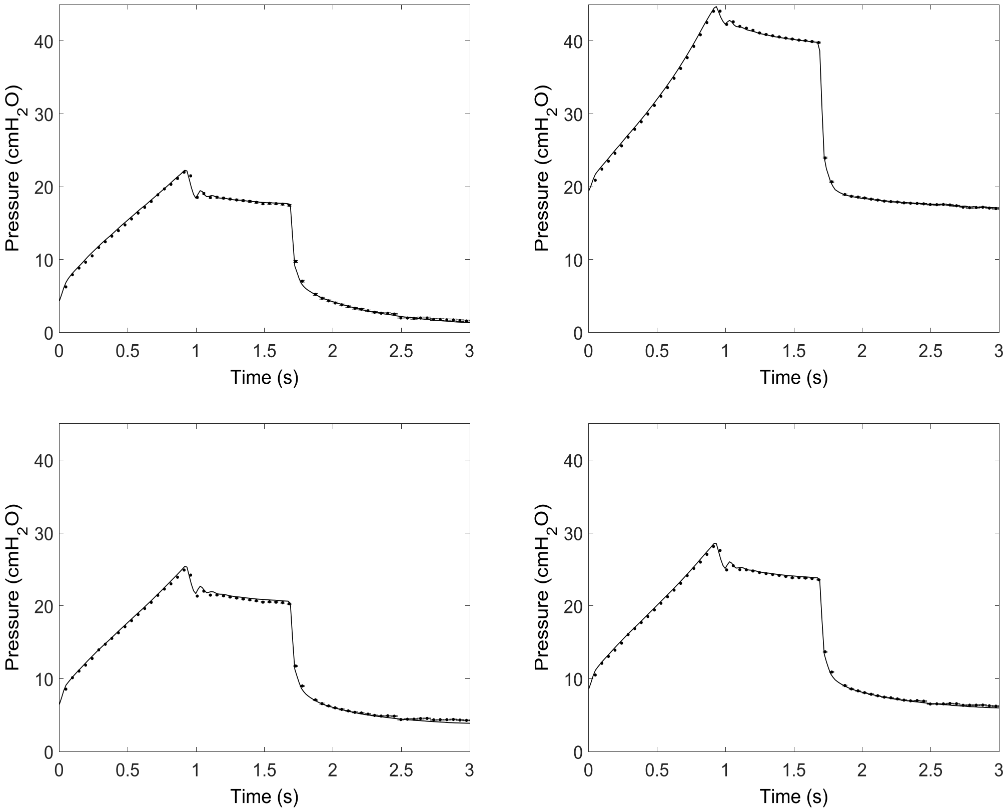


Figure 6.5. Patient B, NARX model: IDD40$_1$ (top left), IDD40$_2$ (top right), EVD (bottom left), IDD100 (bottom right).

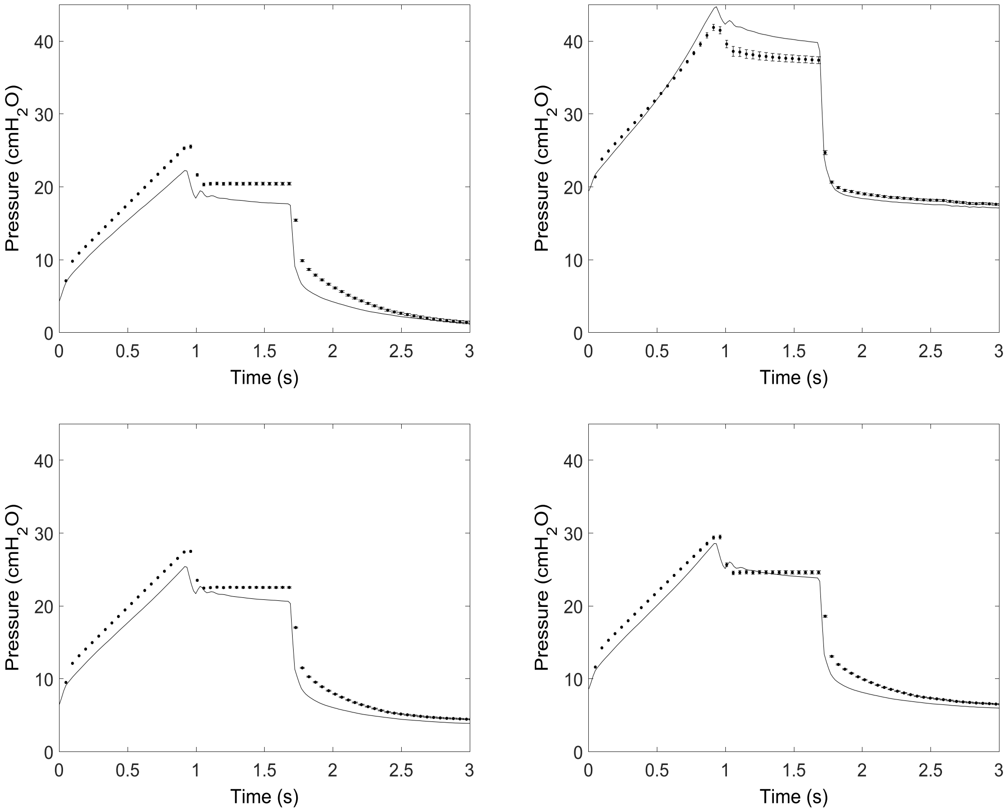


Figure 6.6. Patient B, FOM: IDD40$_1$ (top left), IDD40$_2$ (top right), EVD (bottom left), IDD100 (bottom right).

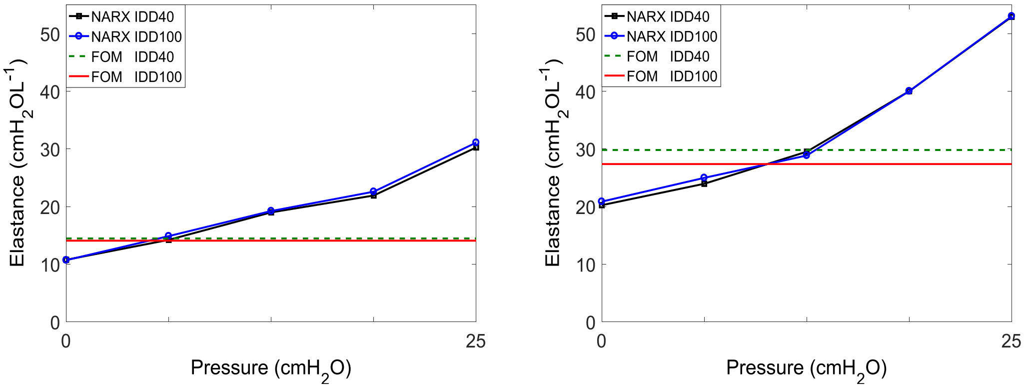


Figure 6.7. Elastance coefficients identified on 40% or 100% of data. Patient A (left), and patient B (right).

6.4 Discussion

Figure 6.2 shows the average RMS residuals for the various identification and evaluation sections of data. The NARX model residuals were smallest when the model was identified on 100% of the data, and largest when interpolated over the evaluation data, as expected. There were significant differences in RMS error across the IDD100, IDD40, and EVD for the whole cohort (signed-rank $p < 0.05$). However, the mean values are similar, and the 90% confidence intervals overlap for the three sections (Table 6.1). This result indicates that the outcome of interpolating the NARX model was not substantially worse than identifying the model on 100% of the data, for this group of 25 patients. Hence, the model can be effectively used to interpolate between measured, clinically relevant states.

For the FOM, the largest RMS residuals occurred for IDD40, and the smallest residuals occurred for EVD. This was because the ΔP tended to gradually increase over the course of the RM, and the FOM was unable to capture both the lower ΔP and higher ΔP that were both present in IDD40 with a single elastance. The parameter identification found a trade-off between the two states, resulting in high residuals for the whole IDD40 section. Since the ΔP in the EVD was generally in between the ΔP at the lowest and highest PEEPs, the trade-off allowed the FOM to be a better fit to the EVD section than the IDD40 section overall. The FOM IDD100 RMS residuals were slightly better than IDD40 on average because the model was fit to all of the data rather than just the extremes. There was a statistically insignificant difference between the FOM RMS residuals for IDD100 and EVD for this cohort (Table 6.2).

Similar to the NARX model result, the magnitude of the differences in RMS residuals for the FOM were not large. The 90% confidence intervals of the three cases overlapped, as shown in Table 6.1 and Figure 6.2. However, there was clear evidence that the NARX model resulted in significantly lower residuals than the FOM. Lower residuals were achieved largely because the

NARX model was more successful than the FOM at capturing the lung relaxation during expiration and the end-inspiratory pause. This result is illustrated in Figure 6.3 to Figure 6.6 and was expected because the FOM is too simple to fully capture the complex behaviour described by this data that includes viscoelastic effects.

Figure 6.3 to Figure 6.6 show results for two individual patients, A and B that represent the extremities of the range of ARDS patients tested. For patient B, the ΔP began increasing at a lower PEEP, and increased at a faster rate compared to patient A. The larger ΔP at the end of the RM was most likely caused by over-distension of alveoli in the lungs of patient B. This important characteristic implies that patient B may be more at risk of VILI at lower pressures than patient A (Dreyfuss and Saumon, 1998).

For both patients A and B, most of the ΔP increases occurred in the final 20% of the data, so the EVD measured pressure waveforms were more similar to those in the $IDD40_1$ section. Thus, the FOM EVD residuals were similar to the $IDD40_1$ residuals. For patient B, the more extreme ΔP increase in $IDD40_2$ caused the FOM to considerably overshoot the data in $IDD40_1$, capturing patient state comparatively poorly. Thus, the FOM EVD residuals for this patient are much larger than the IDD100 residuals. This result shows that in addition to failing to capture the end-inspiratory pause and expiration curve, the FOM is unsuitable for the type of interpolation performed here, when the patient exhibits over-distension at high PEEP levels. In contrast, the NARX model was capable of more accurate interpolation. There were no major differences between the IDD40, EVD, and IDD100 sections for the NARX model, for both patients A and B.

Figure 6.7 shows the NARX model and FOM elastance for patients A and B. The similarity of the NARX IDD40 and IDD100 elastance shapes shows that very similar models were identified for both patients using either 40% or 100% of the data. The result is also very similar for the FOM for patient A, but not for patient B, due to the over-distension present in the IDD40, as discussed above.

Model interpolation is valuable because data covering the full range of PEEP steps may not be available for some respiratory patients in the ICU. In practice, it is unlikely that clinicians would jump from zero PEEP to high PEEP without going through some intermediate stages. However, these results proved that 60% of the RM data was not necessary to identify an accurate NARX model. This result could lead to the implementation of more efficient RMs, wherein PEEP is increased as quickly as is safe for the patient.

6.5 Summary

The NARX model and FOM were identified on 40% of the data from 25 patients who underwent a RM, and interpolated over the remaining 60% of data. The NARX model was more successful than the FOM at fitting to the EVD. The pressure dependent elastance enabled ΔP to be captured at different PEEP levels simultaneously, and it was able to better capture the end-inspiratory pause and expiratory relaxation in each breathing cycle. The NARX model was particularly superior in cases where the data suggested over-distension was occurring at high PEEP levels. The consistency of the NARX model interpolation for both type A and B patients suggests it could be successfully used with a wide range of ARDS patients with different disease characteristics.

Chapter 7 Inspiratory vs. Expiratory Elastance

7.1 Introduction

In general, current pulmonary models are not able to uniquely identify inspiratory elastance E_i, and expiratory elastance E_e as independent variables. This is because the flow and volume both follow exponential decays during relaxed expiration of the sedated lung. As flow and volume are not linearly independent, single elastance and resistance terms are non-identifiable during expiration. The NARX model's use of multiple flow dependent terms means that it is able to identify independent inspiratory and expiratory elastances. This chapter demonstrates this ability, and shows the relationship between E_i and E_e. This helps to further validate the NARX model and its ability to accurately describe patient physiology. The 19 data sets from Cohort2 were used in this analysis. The content of this chapter was published in Langdon *et al.* (2016d).

7.2 Methods

The NARX model here uses separate sets of basis functions for the inspiratory and expiratory parts of the breath. The choice of the number of basis functions depends on the range of pressures in the data. Since the Cohort2 RMs start at non-zero PEEP, the range of pressures is smaller than that of Cohort1. Therefore, four basis functions were used, rather than five. The appropriate choice of *L* depends on the number of data points per breath (Table 5.1). As the Cohort2 sampling rate was lower than Cohort1, and there was no end-inspiratory pause, an *L* value of 170 was used in this analysis. Thus, the NARX model formulation was:

$$P(t) = \sum_{k=1}^{4} a_{ik} \emptyset_k\big(P_i(t)\big) V_i(t) + \sum_{k=1}^{4} a_{ek} \emptyset_k\big(P_e(t)\big) V_e(t) + \sum_{j=1}^{170} b_j \dot{V}(t_{-j}) + P_0(t) \qquad 7.1$$

where P_i and P_e are the measured inspiratory and expiratory airway pressures (cmH$_2$O):

$$P_i(t) = \begin{cases} P(t), & \dot{V} > 0 \\ 0, & \dot{V} < 0 \end{cases}$$
$$P_e(t) = \begin{cases} 0, & \dot{V} > 0 \\ P(t), & \dot{V} < 0 \end{cases} \qquad 7.2$$

and V_i and V_e are the inspiratory and expiratory volumes (L):

$$V_i(t) = \begin{cases} V(t), & \dot{V} > 0 \\ 0, & \dot{V} < 0 \end{cases}$$
$$V_e(t) = \begin{cases} 0, & \dot{V} > 0 \\ V(t), & \dot{V} < 0 \end{cases} \qquad 7.3$$

To identify the NARX model coefficients, a linear system of equations was generated and inverted to find:

$$\mathbf{x} = [a_{i1} \quad \dots \quad a_{i4} \quad a_{e1} \quad \dots \quad a_{e4} \quad b_1 \quad \cdots \quad b_{170}]^{\mathrm{T}} \qquad 7.4$$

where: $a_{i1} - a_{i4}$ are the inspiratory elastance coefficients, and $a_{e1} - a_{e4}$ are the expiratory elastance coefficients. The inspiratory elastance coefficients were identified based on inspiratory data only, and the expiratory elastance coefficients were identified on expiratory data only. The *b* coefficients were identified on both inspiratory and expiratory data.

The a_i, a_e, and *b* coefficients were identified for each data set. The basis functions multiplied by the *a* terms gives a continuous elastance across pressure (Equation 7.5).

$$E_i(P_i) = \sum_{k=1}^{4} a_{ik} \emptyset_k(P_i)$$
$$E_e(P_e) = \sum_{k=1}^{4} a_{ek} \emptyset_k(P_e) \qquad 7.5$$

The coefficient of determination (R^2) was calculated for the linear relationship between E_e and E_i at different pressures. Bland-Altman analysis was performed to determine any bias in the difference between E_e and E_i as pressure increased. To further quantify intra-patient versus inter-patient variability, the variance of E_i, E_e, and the difference ($E_i - E_e$) were calculated.

7.3 Results

The relationship between E_i and E_e at various pressures is shown in Figure 7.2. The plot at $P = 15$ cmH$_2$O contains 18 data points because one data set did not contain pressures as low as 15 cmH$_2$O. Similarly, two data sets did not contain pressures at 40 cmH$_2$O or greater. The agreement between E_i and E_e was generally poor at low pressure, and good at pressures greater than the second basis function knot. The behaviour of a typical patient is shown in Figure 7.1.

The R^2 value for the E_i to E_e linear relationship describes how well the E_i value predicts the E_e value. Figure 7.2 shows the strength of prediction is weak at $P = 15$ cmH$_2$O, but strong for $P \geq 25$

cmH_2O, with $R^2 \geq 0.78$. The 1:1 lines plotted show that there is a tendency for E_i to be greater than E_e at low pressures, and for E_e to be slightly higher than E_i at high pressures.

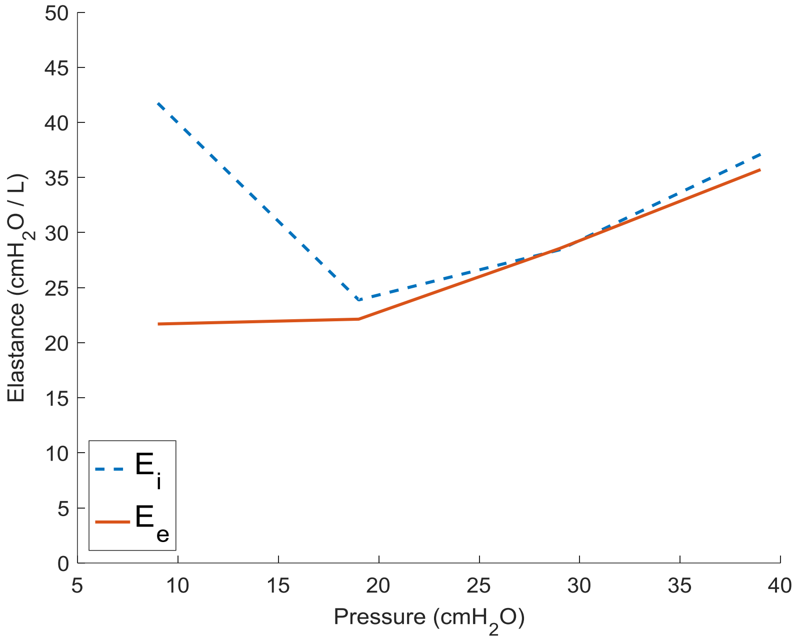


Figure 7.1. Inspiratory and expiratory elastance across pressure for one patient.

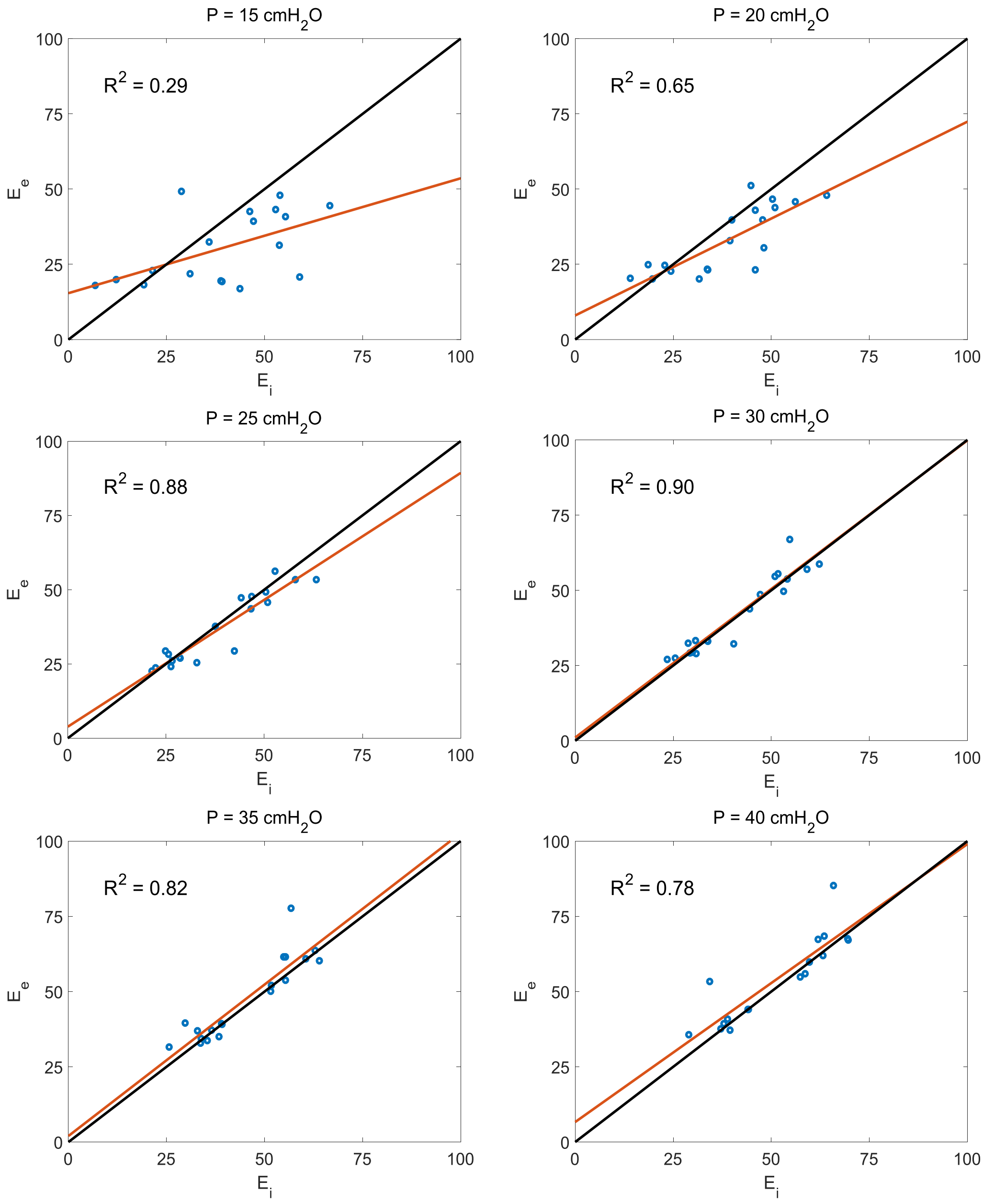


Figure 7.2. Inspiratory elastance vs. expiratory elastance for P = [15 20 25 30 35 40] cmH₂O. R^2 values are given for the linear relationship between E_i and E_e, plotted in red. A 1:1 line is plotted in black.

Bland-Altman plots allow any fixed bias between E_i and E_e to be more easily observed. The mean and corresponding p value are specified on each plot of Figure 7.3. A p value greater than 0.05 indicates that the mean difference is not significantly different from zero, based on a one sample t-test, thus there is no significant difference between E_i and E_e for this pressure. The analysis found that there was a significant difference between E_i and E_e for P = 15 cmH_2O and P = 20 cmH_2O only. At these low pressures, E_i tends to be larger than E_e.

Variances of E_i, E_e, and ($E_i - E_e$) were calculated to determine intra-patient versus inter-patient variability (Table 7.1). For pressures of 20 cmH_2O and above, the variance of ($E_i - E_e$) was smaller than the variance of both E_i and E_e at each measured pressure. Based on the t-test, the variance of E_i and E_e were not significantly different. The variance of (E_i – E_e) was significantly smaller than the variance of E_i (t-test, $p < 0.05$) and significantly smaller than the variance of E_e (t-test, $p < 0.05$).

Table 7.1. Mean variances of E_i, E_e, and the difference (E_i – E_e) across pressure for all data sets.

P (cmH_2O)	Variance of E_i	Variance of E_e	Variance of (E_i – E_e)
15	0.103	0.058	0.083
20	0.072	0.056	0.029
25	0.070	0.067	0.008
30	0.066	0.075	0.006
35	0.066	0.083	0.014
40	0.081	0.096	0.019

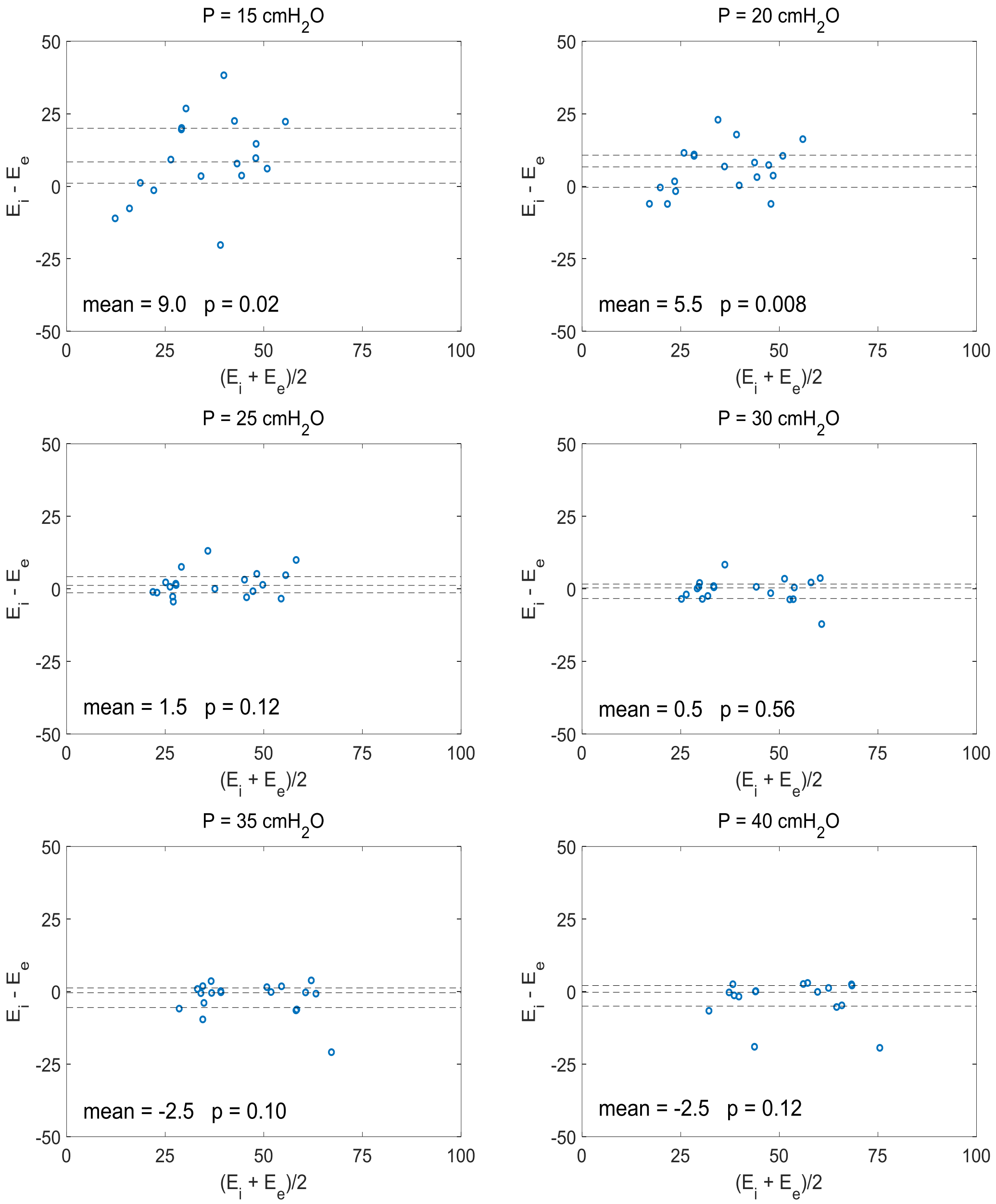


Figure 7.3. Bland-Altman plots for $P = [15\ 20\ 25\ 30\ 35\ 40]$ $\mathrm{cmH_2O}$. Dotted lines show the 25th, 50th, and 75th percentiles of the difference.

7.4 Discussion

This analysis shows that the NARX model is capable of identifying unique inspiratory and expiratory elastance profiles. E_i was a well-matched predictor of E_e for P = 25 to P = 40 $\mathrm{cmH_2O}$ (Figure 7.2**)**. There was no significant bias in the difference between E_i and E_e for P = 25 to P = 40 $\mathrm{cmH_2O}$ (Figure 7.3). The intra-patient variability was significantly lower than the inter-patient variability for P = 20 to P = 40 $\mathrm{cmH_2O}$ (Table 7.1). Overall, this indicates that for this cohort, E_i and E_e were comparable for $25 \le P \le 40$ $\mathrm{cmH_2O}$, and thus may be equally valuable as an indicator of patient condition.

Generally, there was low agreement between E_i and E_e at low pressures. There was a significant positive bias in ($E_i - E_e$) at P = 15 cmH$_2$O and P = 20 cmH$_2$O. In addition, the variance of ($E_i - E_e$) was larger than the variance of E_e at P = 15 cmH$_2$O. The cause of this behaviour relates to the use of distinct local basis functions used to define elastance. With M = 4, the first basis function ($\emptyset_1$) is non-zero for only the lowest third of the pressures present in the data set. Therefore $\emptyset_1$ is identified using only the volume data that exists when these low pressures are present in inspiration.

At the beginning of inspiration, the pressure rises very rapidly. There are relatively few data points here, as the gradient of the pressure increase is so steep. Thus there are relatively few data points used to identify $\emptyset_1$, compared to $\emptyset_2$, $\emptyset_3$, and $\emptyset_4$. There is not enough useful information for determining the elastance in this small portion of inspiratory data. Since the gradient of the pressure drop during expiration is shallower at lower pressures, there is more data available to identify expiratory elastance, and the issue is mitigated. Therefore E_e is likely to be more reliable than E_i at low pressure using the method presented in this chapter.

A similar problem with inspiratory elastance would be likely to occur if the NARX model was identified using this method on any single PEEP level, where a RM was not carried out. In this scenario, either the E_i and E_e should not be identified separately, or the E_i should not be relied on for diagnostic use.

This outcome shows that important information for defining elastance does exist in the expiratory segment of the breath. Many models ignore expiration (Szlavecz *et al.*, 2014, Chiew *et al.*, 2011). However, in this scenario, inspiratory data was not sufficient for defining elastance at low pressure but expiratory data was sufficient.

Separate inspiratory and expiratory elastances are not currently used as a diagnostic aid by clinicians. However, this analysis has shown that unique E_i and E_e values can be obtained using the NARX model, and that expiratory data contains useful information for defining elastance. The outcomes of this type of analysis may provide further insight into sedated ARDS patient conditions that other models cannot accomplish.

7.5 Summary

The NARX model was used to identify inspiratory and expiratory pressure dependent elastance as independent variables. At pressures P = 15 and 20 cmH$_2$O the agreement between E_i and E_e was low. However, E_i was a well-matched predictor of E_e for P = 25 to 40 cmH$_2$O, with $R^2 \geq 0.78$, and there was no significant bias in the difference between E_i and E_e. Since many other models cannot uniquely identify consistent E_i and E_e values, the outcome may provide further insight into the characteristics of ARDS lungs in sedated patients.

Chapter 8 Resistance vs. Pressure

8.1 Introduction

In this section, the changes in resistance captured by the NARX model in response to PEEP changes is evaluated. When PEEP is increased, resistance is expected to decrease as the higher pressure causes recruitment of alveoli, and a widening of the bronchial path vessels (Damanhuri *et al.*, 2014). The ability of the NARX model to capture this reduction in resistance at higher PEEP levels helps to further validate the physiological relevance of the model.

The NARX model b_1 coefficient corresponds most closely to the resistance term *R* in the FOM. Therefore the behaviour of the b_1 term was compared with the resistance term of the FOM, *R*. The 25 patient data sets of Cohort1 were used in this analysis. The content of this chapter was published in Langdon *et al.* (2016c).

8.2 Methods

The NARX model (5.1) was identified on a moving window of 20 breaths that shifted across the RM of each data set (Figure 8.1). Unique a_i and b_j coefficients were identified independently for each of these windows, allowing the trend of the resistance term b_1 to be observed over time as PEEP increased. The NARX model was identified on breaths at one to three adjacent PEEP levels, thus the pressure range was limited. Therefore, an *M* value of three was deemed appropriate to provide a robust result.

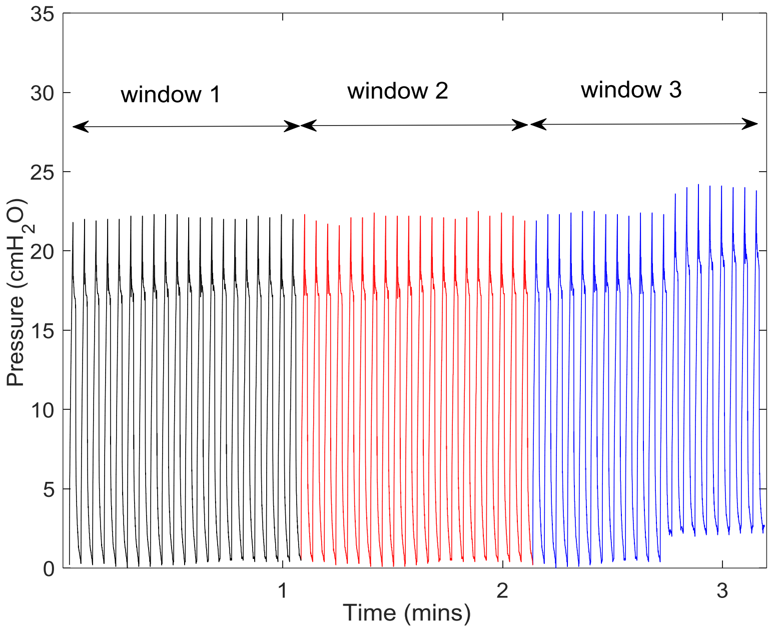


Figure 8.1. Three pressure windows of length 20 breaths.

The NARX model used in this analysis has 353 parameters (L = 350, M = 3). The number of data points per breath is in the range 190 – 380 for different patients, due to varied breathing rates of patients, and a constant sampling frequency of 62.5Hz. A window length of 20 breaths was chosen as it enabled practically identifiable b_1 values that exhibited the underlying behaviour of the patients.

To observe how b_1 varied with PEEP, boxplots for the 25 data sets were created. The average PEEP for each window was calculated, and the boxplots were plotted using data from windows where the average PEEP was 0.1 – 5, 5.1 – 10, 10.1 – 15, and 15.1 – 20 cmH$_2$O.

8.3 Results

Figure 8.2 shows how the NARX b_1 term and the FOM R term changed over time, for three data sets. Separate linear trend lines have been plotted over the ZEEP portion and over the RM portion. These three cases show the range of behaviours observed in this cohort. In the RM, the NARX b_1 tended to decrease. In contrast, the FOM resistance exhibited three different types of behaviour. The FOM resistance increased, decreased, or had a quadratic shape as PEEP increased.

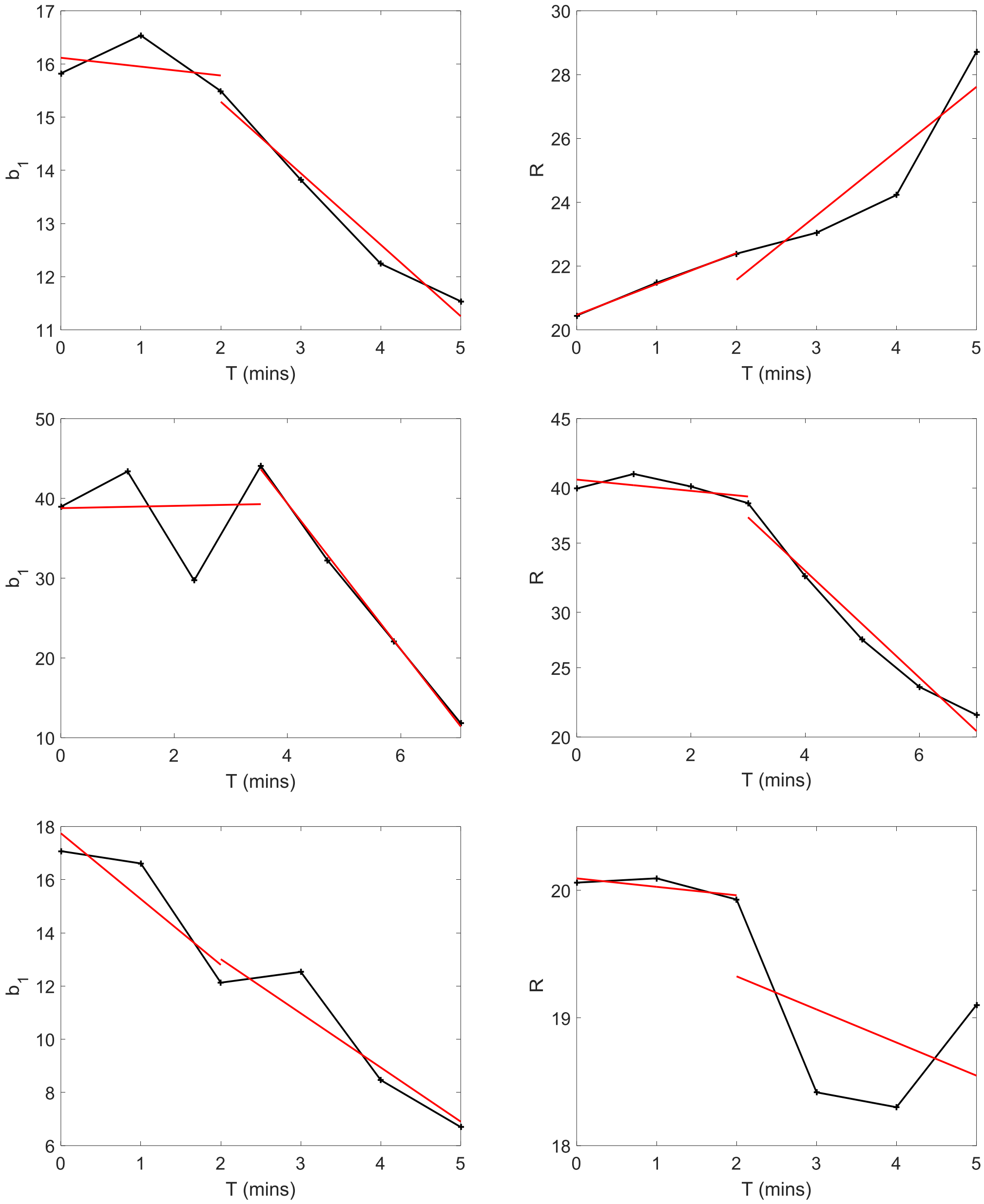


Figure 8.2. The NARX model b_1 (left) and the FOM R (right) for three patients. T = time at the start of the window.

Table 8.1 shows the mean gradient of the linear trend line applied to the ZEEP and RM portions for b_1 and R. Within the 90% confidence interval, the RM gradient of b_1 was negative, as b_1 consistently decreased with PEEP. In fact, there was only one patient for which the RM b_1 gradient was positive. In contrast, the varied behaviour of R as PEEP increased meant that the 90% confidence interval for the R gradient in the RM ranged across zero. The t-test revealed a significant difference between the mean gradient of b_1 and R during the RMs. There was no significant difference in the behaviour of R and b_1 at ZEEP.

Table 8.1. The mean gradient of the NARX b_1 and FOM R terms for the ZEEP and RM portions.

	Mean ZEEP Gradient and 90% CI	Mean RM Gradient and 90% CI
NARX b_1	-0.011 [-0.021, -0.0012]	-0.031 [-0.041, -0.021]
FOM R	-0.0012 [-0.0027, 0.0003]	-0.0035 [-0.011, 0.0036]
t-test	$p = 0.11$	$p = 0.0006$

The boxplots for the b_1 and R terms (Figure 8.3) show that the median of b_1 decreased as PEEP increased. The median of R also decreased with PEEP, but with a much shallower slope. The boxplot at PEEP = 10.1 – 15 cmH$_2$O contains 23 data sets out of the total of 25, because the highest PEEP for two patients was less than 10 cmH$_2$O. Similarly, the boxplot at PEEP = 15.1 – 20.1 cmH$_2$O contains data from 17 patients only, as eight data sets contained no breaths within this PEEP range.

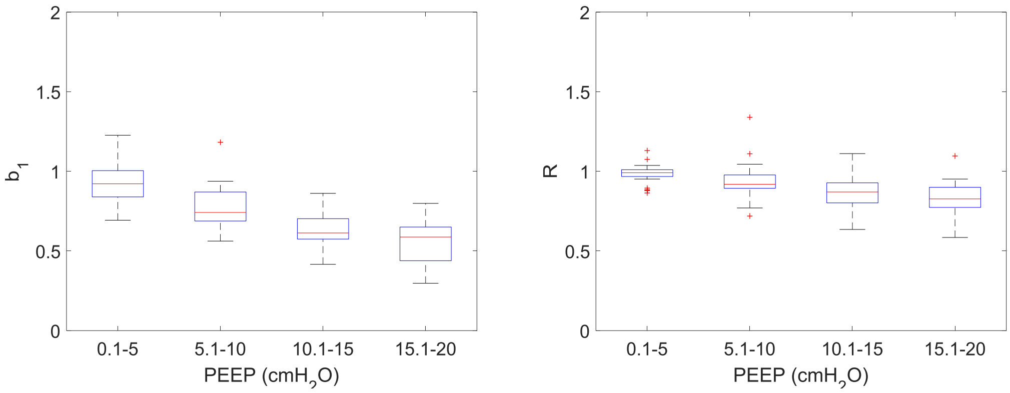


Figure 8.3. Boxplot for the NARX model b_1 (left) and the FOM R parameters (right), normalised to the average value at ZEEP. The box limits are the 25th and 75th percentiles, and whiskers show the range limited to data points that are within 1.5 IQR.

8.4 Discussion

The resistance to flow is primarily described by the NARX model via variance in the b_1 parameter. Figure 8.2 and Figure 8.3 show that the resistance captured by the NARX b_1 term consistently reduced as PEEP increased, over 25 RMs. This behaviour conforms to expected behaviour, as high pressures cause widening of airway passages, thus reducing resistance.

In contrast, the behaviour of the FOM R term was inconsistent. In response to the RM, the R term either decreased, increased, or had a quadratic shape (Figure 8.2**)**. While Figure 8.3 shows that the general trend of R as a function of PEEP was downwards, the variation in the FOM R behaviour meant that the strength of the downwards trend was much weaker than that of the NARX b_1 value.

The average trend in both b_1 and R during ZEEP was a very shallow negative slope. However, the 90% confidence interval indicated no consistent trend in R at ZEEP, and the t-test showed no significant difference in the gradient of R and b_1 at ZEEP (Table 8.1). This behaviour was expected, as patient's resistance at ZEEP would be likely to be roughly constant, unless patient condition changes.

The R term of the FOM had inconsistent behaviour across patients. Many patients experienced an increase in this modelled resistance as PEEP increased, from either the beginning of the RM or from midway through the RM. It is suspected that those patients that had an apparent increase in modelled resistance at higher pressures were actually exhibiting non-linear elastance behaviour that was being compensated for by increased resistance values. In contrast, the non-linear elastance behaviour was captured by the more complex NARX model, due to the multiple elastance parameters and the other flow-dependent terms. Thus, the b_1 term was robust to this non-linearity.

8.5 Summary

The NARX model b_1 term consistently decreased as PEEP increased, while the FOM resistance behaviour varied. Overall the NARX model behaviour is more in-line with expected trends in airway resistance over a RM. This work has further validated the descriptive capability of the NARX model for capturing changes in airway resistance over PEEP steps. This aspect of the NARX model may be particularly useful for observing changes in ARDS patients suffering from COPD, who often have notably higher resistance and greater potential for resistance changes during a RM as airway blockages shift.

Part 3: NARX Model Applications

Chapter 9 Capture of M-waves

9.1 Introduction

In MV, partially sedated spontaneously breathing (SB) patients apply their own inspiratory efforts on top of a ventilator supported breathing cycle. These SB efforts can sometimes result in abnormal airway pressure curves, or 'M' shaped pressure curves (M-waves) (Akoumianaki *et al.*, 2013), as shown in Figure 4.3. The M-wave pressure curve masks the underlying respiratory mechanics from identification since the exact SB effort is unknown and effectively random. Therefore, a method is required to overcome the impact of the M-waves to provide a consistent model-based estimation of respiratory mechanics for clinical use. Here, the NARX model is applied to M-wave data in conjunction with a modified version of the Gauss-Newton parameter ID algorithm (Gray *et al.*, 2016). The modified GN method has previously been used to ignore contributions from outlying data by finding the parameter set that fits the majority of the data points, rather than the least squares optima for all patients (Docherty *et al.*, 2014). The aim was to use these two approaches to model respiratory mechanics while effectively ignoring M-waves in the pressure signal of the M-wave patient of Cohort2. The content of this chapter was published in Langdon *et al.* (2015a).

9.2 Methods

The section of data used in this work had a constant PEEP of 17 cmH_2O. As PEEP was constant, the pressure range was small, at around 16 cmH_2O. Similarly to Chapter 8, an M value of three was used to provide a robust result. Since the data contained no RM and no end-expiratory pause, the NARX model was able to be further simplified with $L = 1$.

To identify the model, the original GN method was implemented (Equation 3.6), as well as a modified GN algorithm (Equation 9.1). The original GN method leads to a least squares optimisation. The adapted method replaces $\boldsymbol{\psi}$ with $\widehat{\boldsymbol{\psi}}$.

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \left(\mathbf{J}^{\mathrm{T}}\mathbf{J}\right)^{-1}\mathbf{J}^{\mathrm{T}}\widehat{\boldsymbol{\psi}} \tag{9.1}$$

$$\widehat{\boldsymbol{\psi}} = \left[\widehat{\psi}_j\right] = \left[\psi_j e^{\frac{-|\psi_j|}{\beta|\widetilde{\psi}|}}\right] \tag{9.2}$$

The median of the absolute values of the residuals is denoted by $|\widetilde{\psi}|$, and β is a scaling factor.

$\widehat{\boldsymbol{\psi}}$ changes how each residual error value contributes to the magnitude of the step to adjust **x**, compared to the original GN method. In the original GN, the contribution to the change in **x** at each iteration increases with the square of the error. Therefore, if undesired outliers exist in the data, they have a large effect on the direction of convergence, and the resulting model may not represent the majority of the data points.

When $\widehat{\boldsymbol{\Psi}}$ is used, the contribution of residuals greater than a certain value decreases exponentially. Therefore, large outliers will not greatly affect the result. The value of β determines where the exponential decrease becomes influential, with respect to model residuals at the *i*th iteration. A β value of infinity means that the original GN method is applied as ψ_j is multiplied by one. However, if β is too small, the approach will ignore important characteristics that define the system (Docherty *et al.*, 2014, Gray *et al.*, 2016). Figure 9.1 depicts how the contribution of residuals changes as residuals increase.

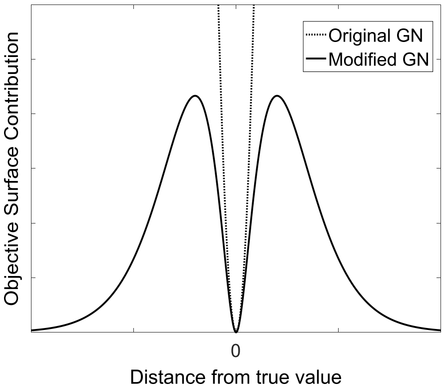


Figure 9.1. Relative contribution of residuals to the parameter step for the original and modified GN methods.

The FOM and the NARX model were identified using the entire ~80 minutes of ventilation data. The β parameter was varied to determine an optimal number for ignoring M-waves, and results were compared to the original GN method (β = Inf) for both the FOM and NARX model. The GN initial values were chosen by evaluating the FOM and the NARX model via direct inversion.

RMS residuals were calculated for each model fit. However, the aim was not to choose the model with the lowest residuals. The aim was to find the value of β for which the model outputs a breath that is close to the shape of normal breaths where M-waves exist in the measured pressure. Therefore the best model will have large residuals where M-waves exist in the data, and small residuals for normal breaths.

9.3 Results

Figure 9.2 shows a section of the data containing both M-wave breaths and normal breaths. The FOM produced similar results with the original GN method and modified GN with β = 4. In comparison with the FOM, the NARX model was able to better match the peak pressure in normal breaths and better fit the expiration curve in all breaths. The NARX model with β = 4 is also able to successfully ignore M-waves. When β = Inf, the NARX model provided a better fit to the data than the FOM, as expected, because the NARX model is more complex as it contains a larger number of identified parameters. However, in this special case, this improvement is undesirable as the aim was to ignore M-waves rather than capture them.

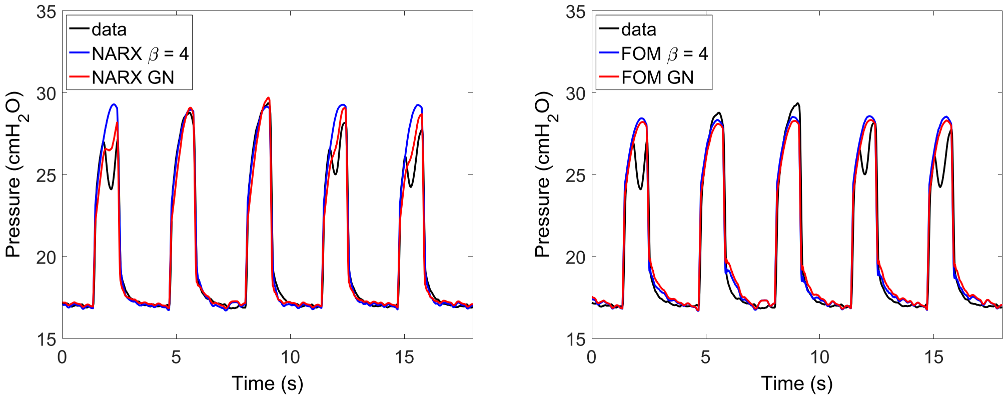


Figure 9.2. M-wave pressure data and the NARX model (left) and FOM (right) identified with the original GN method, and modified GN (β = 4).

Figure 9.3 shows the effect of varying β on the NARX model output pressure. A β value that is too low means that the model is unable to capture the shape of the inspiratory curve, and unable to reach the peak pressure in the breath. If β is too high, the model tends to start following the M-wave curve rather than ignoring it, and also tends to slightly undershoot the inspiratory curve. In this analysis, β = 4 was the optimal number that allowed the model to capture the shape of normal breaths as well as ignore M-waves.

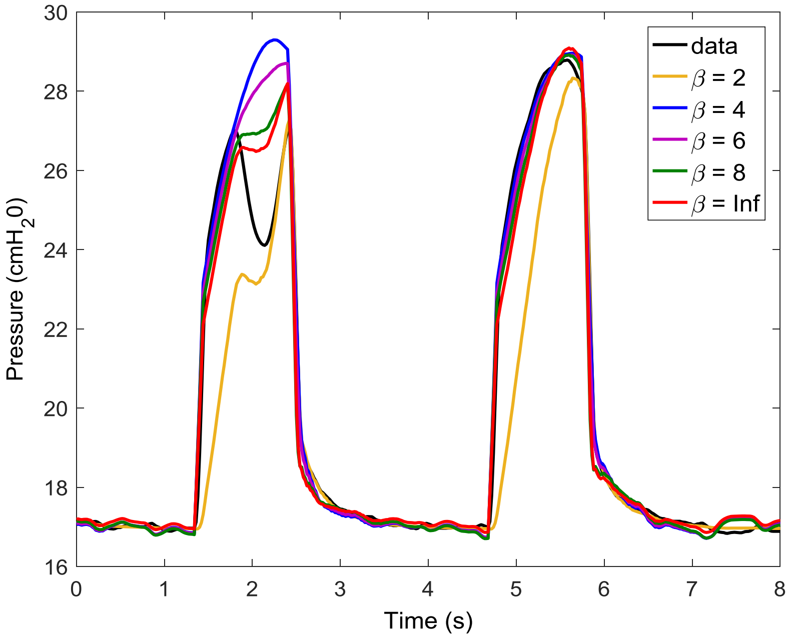


Figure 9.3. Airway pressure data and the NARX model identified with GN for β = 2, 4, 6, 8, and Inf.

Notably, for the NARX model, 2000 GN iterations were required for coefficient convergence for β = 2, whereas 150 iterations were sufficient for β = 4, 6, and 8. For the original GN method, only one iteration was required. For β = 1, the coefficients did not converge when tested up to 3000 iterations. Figure 9.4 verifies that the NARX model coefficients have converged after 150 iterations of the modified GN algorithm. 150 iterations took approximately 12.5 seconds to complete.

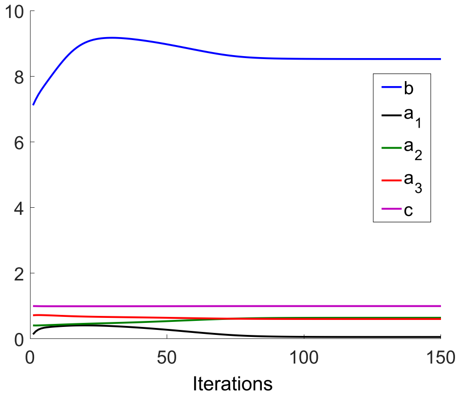


Figure 9.4. Convergence of the NARX model coefficients with β = 4.

Table 9.1 presents the root mean square (RMS) residuals for both the NARX model and FOM for the original and modified GN methods. Table 9.2 shows the identified coefficient values. Both methods applied to the FOM resulted in similar coefficients and residuals, giving rise to the similar plots in Figure 9.2. The FOM resistance coefficients are 25% or more higher than the corresponding NARX model resistance coefficients. Figure 9.5 shows how the elastance coefficients in the NARX models change through pressure. In contrast, one elastance parameter exists for all pressure in the FOM.

Table 9.1. FOM and NARX model RMS residuals for the original and modified GN method.

	RMS residual (cmH_2O)
NARX β = Inf	0.853
NARX β = 4	1.123
FOM β = Inf	1.128
FOM β = 4	1.182

Table 9.2. FOM and NARX model coefficients for the original and modified GN method.

	Elastance (cmH_2O/L)			Resistance (cmH_2Os/L)
NARX β = Inf	0.254	0.428	0.733	7.286
NARX β = 4	0.055	0.644	0.606	8.521
FOM β = Inf	0.521			9.547
FOM β = 4	0.525			10.765

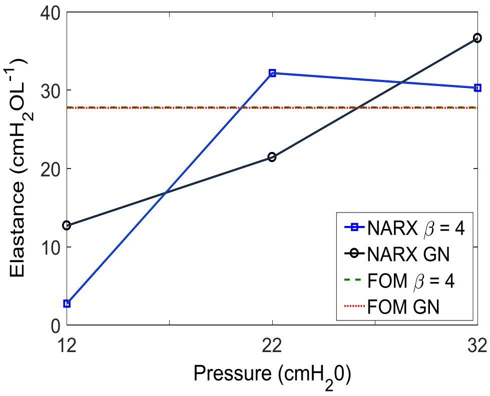


Figure 9.5. Elastance through pressure for the NARX model and FOM when identified with the original and modified GN methods.

Figure 9.6 shows undesirable behaviour in expiration that the models sometimes exhibit. The behaviour corresponds with oscillation in the flow measurements that are caused by the patient's SB efforts when the ventilator is in expiration mode. The upwards blips in the airway pressure during expiration exists in both the FOM and NARX models, and with both the original and modified GN methods.

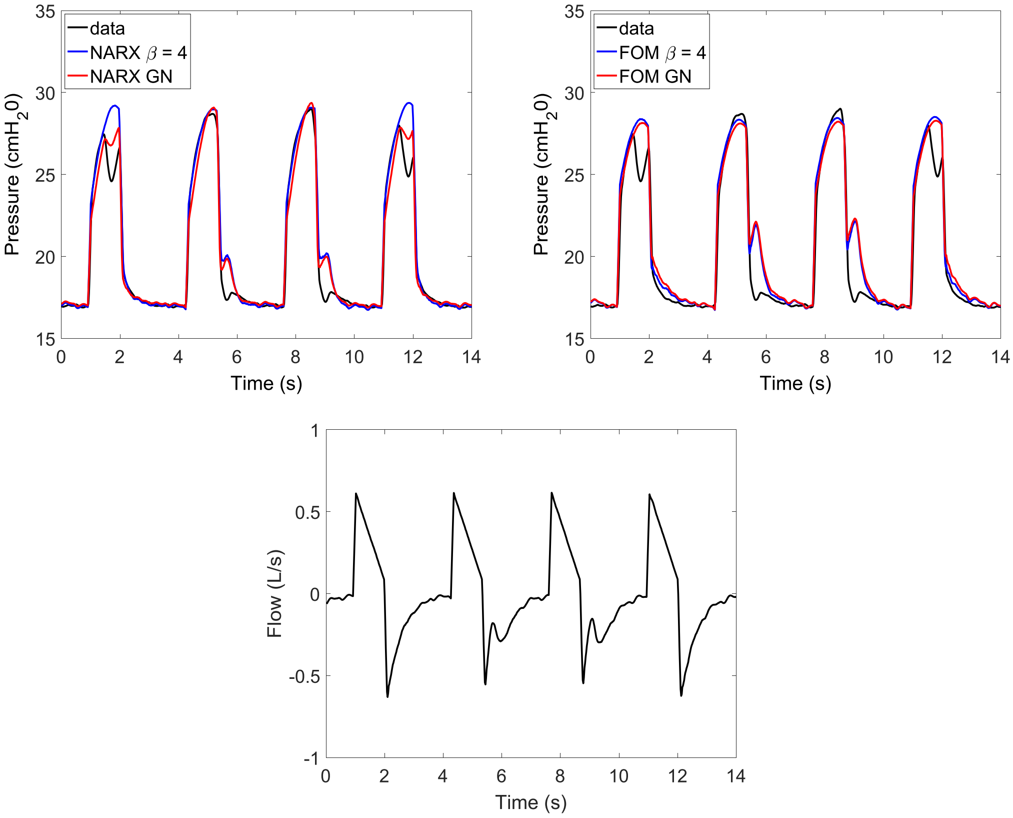


Figure 9.6. Airway pressure with undesired model behaviour in expiration for the NARX model (top left) and FOM (top right), and the corresponding flow measurement data (bottom).

9.4 Discussion

Figure 9.3 indicates that β = 4 was optimal in this analysis for effectively ignoring M-waves. Assuming the residual error is normally distributed, the value of β equals the number of standard deviations (SD) of the error distribution that is between the peaks of the objective contribution shape (Docherty *et al.*, 2014). For example, when β = 2, the largest contribution to the step in **x** happens for residuals that are two SDs from the mean. For residuals smaller than two SD, the contribution increases with the square of the residual, and for errors larger than two SD, the contributions decrease exponentially. Thus, when β was small, the contribution of much of the valuable information was small and the model was not able to capture the shape of the breaths. However, when β is large, the contributions from only very large residuals are preferentially reduced by GN. When β is large, the method approaches the original GN method, and the model becomes a better fit to the measured data so M-waves begin to be followed rather than ignored.

Table 9.1 indicates that the NARX model with the original GN method was better than the FOM at providing a best fit to the data, according to least squares criterion. This result is consistent with the findings of Chapter 5. This outcome is primarily facilitated by the use of three elastance coefficients that depend on pressure, compared to the single elastance FOM coefficient (Table 9.2, Figure 9.5). The difference in resistance coefficients of over 25% between the FOM and NARX model would have also played a role. These factors allowed the NARX model to partially capture M-waves, which the simpler FOM was unable to do (Figure 9.2).

The NARX model with the modified GN method resulted in a larger RMS residual value compared to the original GN method (Table 9.1). This outcome was an expected and desired result because the modified GN method allowed the M-waves in the data to be ignored. The model followed the shape of normal inspiration in breaths where M-waves were present. Thus, the model residuals were large in these regions. Since most breaths contain some degree of M-wave, the NARX RMS residual for the modified GN method was significantly larger than for the NARX identified with the original GN.

The NARX model with β = 4 resulted in an RMS residual that was similar to those of the FOM models. This result occurred because the FOM similarly did not fit the M-waves, as the model is too simple to capture this type of behaviour. However the FOM was not able to reach the peak inspiratory pressure in normal breaths, and also tended to provide a worse fit to the data during expiration, compared to the NARX model. Thus, the NARX model with β = 4 had a slightly lower RMS residual than both FOM methods.

Assuming the residuals are normally distributed, use of the modified GN method should have no negative effects when outlier behaviour does not exist in the data, and the result in this case should closely match the outcome of the original GN method (Docherty *et al.*, 2014). Thus, the method could safely be used to provide clinicians with patient-specific information in situations where the patient is not spontaneously breathing, as well as when SB is present. The parameter identification method for modified GN had a higher computational burden than the typical linear least squares approach for the NARX model. However, this added burden would be negligible in a clinical setting.

Figure 9.6 showed that oscillations in the flow measurements cause the models to fail to capture the appropriate expiratory curve. This flow characteristic is caused by the patient’s spontaneous breathing efforts while the ventilator is in the expiration part of the breathing cycle. As the inspiratory pressure curve is the important part of the breath used by clinicians to determine ventilator settings, this effect is not a significant problem. A method to smooth out the flow oscillation before identifying the model could be employed in future work, e.g. by using the expiratory time constant calculated from other breaths (van Drunen *et al.*, 2013).

The patient had many instances of SB during the ventilation period, as over 65% of breaths contained an M-wave. The success of the method under these conditions suggests it could be useful in monitoring many SB patients, though it is not clear whether the approach would still be useful in situations where an even larger percentage of breaths contain M-waves. The β value or NARX model parameters could potentially be adjusted to allow the model to successfully fit other situations such as this. Comparing the output pressure with the original M-wave pressure curves could also give clinicians an indication of the breathing effort of the patient, which can be useful in determining when to extubate the patient (Boles *et al.*, 2007).

The modified GN method cannot be used to identify the model across data where the patient state changes, e.g. due to lung recruitment or over-distension caused by a PEEP increase. The reason for this limitation is because breaths that have different characteristics to the majority of data will be treated as outliers and will not be tracked by the model. This issue could be reduced by accounting for known PEEP changes in the model and identification.

9.5 Summary

The NARX model was successfully fit to M-wave patient data using the modified GN method with β = 4. The modified GN method exponentially reduces the contribution of large residuals on the step in the coefficients at each GN iteration. This approach allowed the model to effectively ignore the anomaly in the pressure waveform due to SB efforts, while successfully describing the shape of normal breathing cycles. In comparison to the FOM, the NARX model was able to better capture the expiratory curve, and fit to the peak pressure of each normal breath. The successful elimination of M-waves allowed respiratory mechanics to be more accurately estimated.

Chapter 10 Extrapolation

10.1 Introduction

In this section the ability of the NARX model to extrapolate and predict the effects of changes in PEEP is determined. This type of analysis would allow clinicians to determine whether or not to alter, and in particular raise, the PEEP setting for a particular patient without the need for actually trialling these alternative settings. Thus, the risk of testing higher PEEP that could result in over-distension would be reduced. The analysis used the 19 data sets in Cohort2, and a comparison with the FOM was undertaken. The content of this chapter was published in Langdon *et al.* (2016a).

10.2 Methods

The original NARX model used first order b-spline basis functions to capture the elastance pressure shape. However, such functions are not suitable for extrapolation beyond the range of data used to train the model. To extrapolate the NARX model, elastance basis functions must extend to cover pressures not present in the training data.

When the pressure dependent elastance profiles were identified for the entire RM, several different types of pressure vs. elastance shapes were observed (Figure 10.1). The range of profiles observed could be constructed with linear combinations of constant, linear, exponential, and sigmoidal functions (Figure 10.2):

$$\emptyset_1 = 1 \tag{10.1}$$

$$\emptyset_2 = \frac{P}{50} \tag{10.2}$$

$$\emptyset_3 = e^{-0.04P} \tag{10.3}$$

$$\emptyset_4 = \frac{1}{1+ e^{-0.25(P-28)}} \tag{10.4}$$

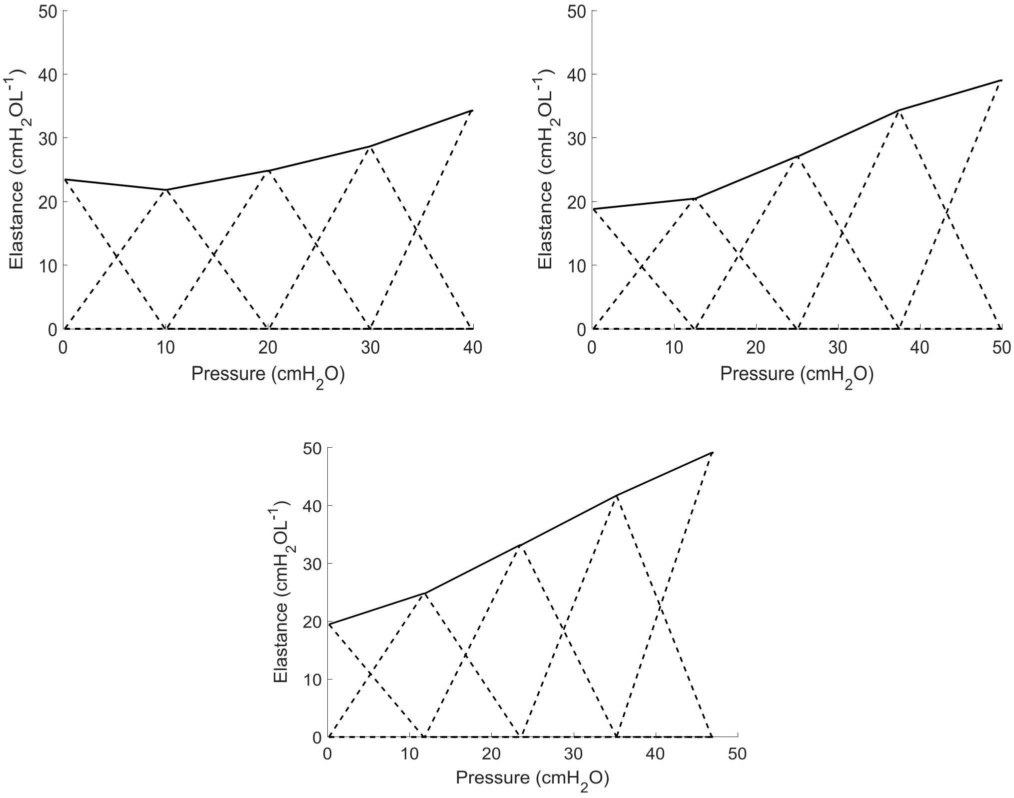


Figure 10.1. Elastance identified with five first order b-spline basis functions, for three indicative patient responses over 100% of available data.

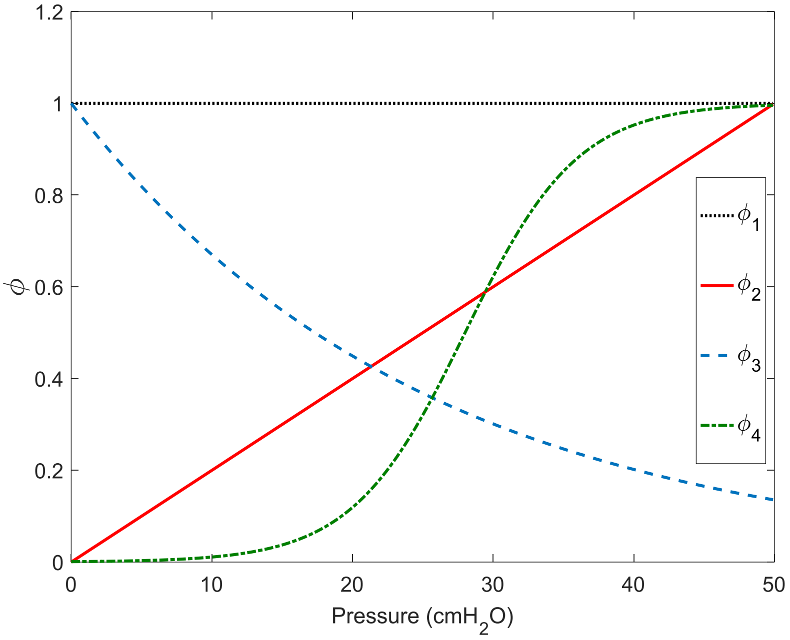


Figure 10.2. Constant, linear, exponential, and sigmoidal basis functions.

To assess the accuracy of extrapolation, the NARX model was trained on specific PEEP levels, and was then extrapolated to predict the pressure response across different PEEP levels. The identification data (IDD) was used to train the model, and the evaluation data (EVD) was used to test the prediction accuracy. Each data set contained at least four PEEP increases. Therefore only the first five PEEP levels of each data set were used in the analysis, so that the amount of data at each EVD level was consistent. Several scenarios were tested:

1. The model was trained on one PEEP level in the dataset ([IDD1, IDD2, … IDD5]). The accuracy of extrapolation was tested over each of the remaining four PEEP levels ([EVD1, EVD2, … EVD5]), both higher and/or lower levels.
2. The model was trained on two adjacent PEEP levels ([IDD1-2, IDD2-3, IDD3-4, IDD4-5]). Extrapolation was tested over the remaining three PEEP levels ([EVD1, EVD2, … EVD5]).
3. The model was trained on three adjacent PEEP levels ([IDD1-3, IDD2-4, IDD3-5]). Extrapolation was tested over the remaining two PEEP levels ([EVD1, EVD2, EVD4, EVD5]).
4. The model was trained on four adjacent PEEP levels ([IDD1-4, IDD2-5]). Extrapolation was tested over the remaining one PEEP level ([EVD1, EVD5]).

RMS residuals between the modelled pressure and the measured pressure were calculated to quantify the accuracy of the extrapolation. To provide a comparison, the RMS residuals for the IDD sections were also calculated. The paired signed-rank test was used to test for significant differences between the mean RMS residuals for the different EVD proportions.

10.3 Results

Figure 10.3 shows the increase in residuals as the EVD moves further from the training PEEP levels. Note that the residual gradients decrease as the IDD range increases. This result implies that the larger identification window with a greater number of PEEP levels allowed a more accurate model of patient behaviour to be captured. This improved prediction accuracy, particularly to PEEP levels several steps beyond the IDD. For example, the mean RMS residual for EVD5 was 1.25 (90% CI: 0.95 – 1.55) cmH_2O when the NARX model was trained on IDD1, but reduced to 0.71 (90% CI: 0.61 – 0.81) cmH_2O when the model was trained on IDD1-2.

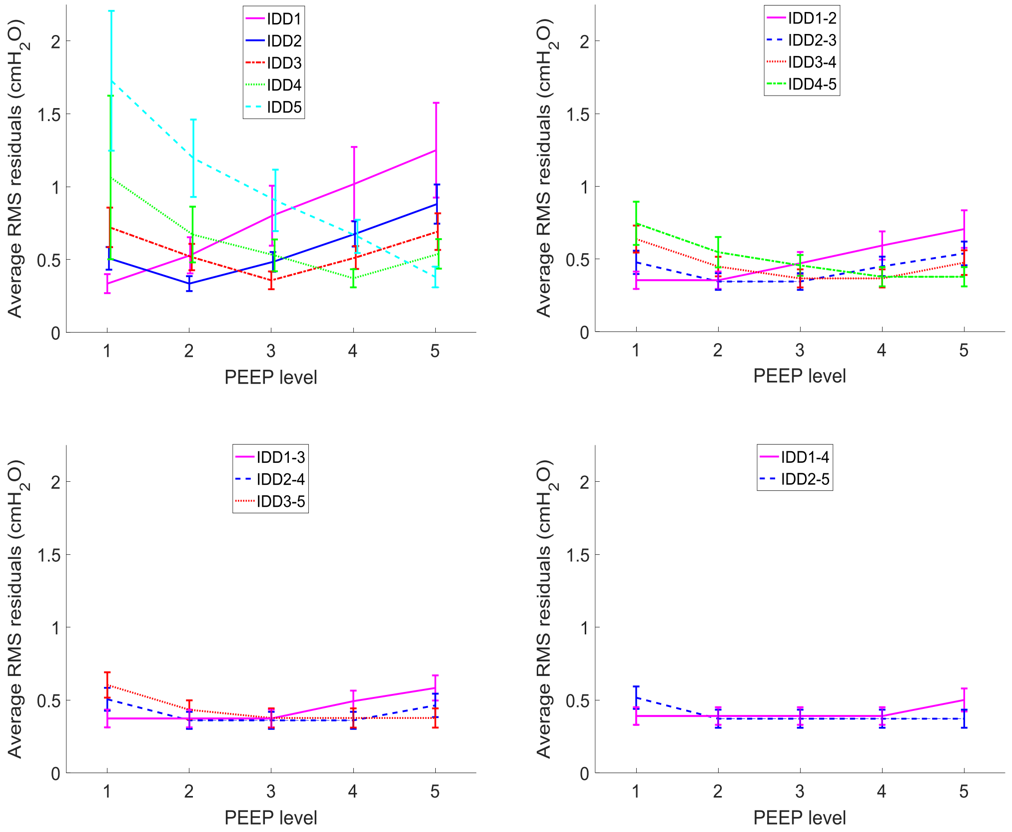


Figure 10.3. Mean pressure RMS residuals and 90% confidence intervals for the IDD and EVD regions. The NARX model was identified on one PEEP (top left), two PEEP levels (top right), three PEEP levels (bottom left), and four PEEP levels (bottom right). Extrapolation was implemented for PEEP above and below the IDD region, where available.

Table 10.1 shows the RMS residuals when the model extrapolation is assessed at the PEEP step immediately greater than the identification region. This case is the most clinically important case as it is very uncommon to increase PEEP in an ARDS patient by more than 5 cmH_2O in any single instance. In addition, once a PEEP level is predicted to be safe, and then tested, it could be used to aid further training and extrapolation.

Table 10.2 shows the changes in RMS residuals when the training data set is enlarged. Generally, Table 10.1 and Table 10.2 show that increasing the IDD range did not always reduce the RMS residuals for the PEEP step one level up from the training data. For IDD2-3 vs. IDD1-3, and IDD2-4 vs. IDD1-4, the EVD RMS residuals increased for the larger training data set. Importantly, in all scenarios, the NARX RMS residuals were significantly lower than the FOM residuals (paired signed-rank test, $p < 0.05$).

Table 10.1. IDD and EVD RMS residuals for extrapolation to one PEEP step up (2 cmH_2O) from the highest PEEP in the IDD.

	Mean RMS residual across all data sets (N = 19) and 90% confidence interval (cmH_2O)			
Identification Data -> Evaluation Data	NARX IDD	NARX EVD	FOM IDD	FOM EVD
IDD1 -> EVD2	0.33 ± 0.07	0.53 ± 0.12	1.30 ± 0.25	1.39 ± 0.25
IDD2 -> EVD3	0.33 ± 0.05	0.48 ± 0.07	1.34 ± 0.25	1.49 ± 0.24
IDD3 -> EVD4	0.36 ± 0.06	0.51 ± 0.08	1.40 ± 0.25	1.53 ± 0.25
IDD4 -> EVD5	0.37 ± 0.06	0.54 ± 0.10	1.45 ± 0.25	1.61 ± 0.24
Mean IDDi -> EVDi+1	0.35 ± 0.03	0.52 ± 0.05	1.37 ± 0.12	1.50 ± 0.12
IDD1-2 -> EVD3	0.35 ± 0.06	0.47 ± 0.08	1.32 ± 0.25	1.56 ± 0.24
IDD2-3 -> EVD4	0.34 ± 0.06	0.45 ± 0.07	1.40 ± 0.25	1.60 ± 0.24
IDD3-4 -> EVD5	0.37 ± 0.06	0.47 ± 0.08	1.45 ± 0.25	1.68 ± 0.24
Mean IDDi-i+1 -> EVDi+2	0.35 ± 0.03	0.46 ± 0.04	1.39 ± 0.14	1.61 ± 0.14
IDD1-3 -> EVD4	0.37 ± 0.06	0.49 ± 0.07	1.38 ± 0.24	1.73 ± 0.24
IDD2-4 -> EVD5	0.36 ± 0.06	0.46 ± 0.08	1.46 ± 0.24	1.77 ± 0.23
Mean IDDi-i+2 -> EVDi+3	0.37 ± 0.04	0.48 ± 0.05	1.42 ± 0.17	1.75 ± 0.17
IDD1-4 -> EVD5	0.39 ± 0.06	0.50 ± 0.08	1.44 ± 0.24	1.95 ± 0.24

Table 10.2. NARX model paired signed-rank test results for the EVD for increasing IDD.

Training PEEP range	Evaluation PEEP range	Mean EVD RMS residual and 90% confidence interval (cmH_2O)	Signed-rank test
IDD2 IDD1-2	EVD3	0.48 ± 0.07 0.47 ± 0.08	*p* = 0.63
IDD3 IDD2-3	EVD4	0.51 ± 0.08 0.45 ± 0.07	*p* = 0.003
IDD2-3 IDD1-3	EVD4	0.45 ± 0.07 0.49 ± 0.07	*p* = 0.002
IDD4 IDD3-4	EVD5	0.54 ± 0.10 0.47 ± 0.08	*p* = 0.024
IDD3-4 IDD2-4	EVD5	0.47 ± 0.08 0.46 ± 0.08	*p* = 0.28
IDD2-4 IDD1-4	EVD5	0.46 ± 0.08 0.50 ± 0.08	*p* = 0.02

Figure 10.4 shows the mean NARX pressure prediction at EVD levels that are immediately above the IDD PEEP levels for one patient. For EVD3, EVD4, and EVD5, the mean absolute residual

between the predicted pressure and the measured pressure was less than 0.5cmH_2O at all points in the breath. Figure 10.5 shows elastance curves identified across differing PEEP levels and their extrapolations compared to the elastance curve identified over all PEEP levels (IDD1-5). Figure 10.6 and Figure 10.7 show a case wherein the extrapolation yielded poor residuals. For this particular pressure controlled ventilation patient, anomalies in the flow profile caused by the patient's breathing efforts yield insight to the variations shown.

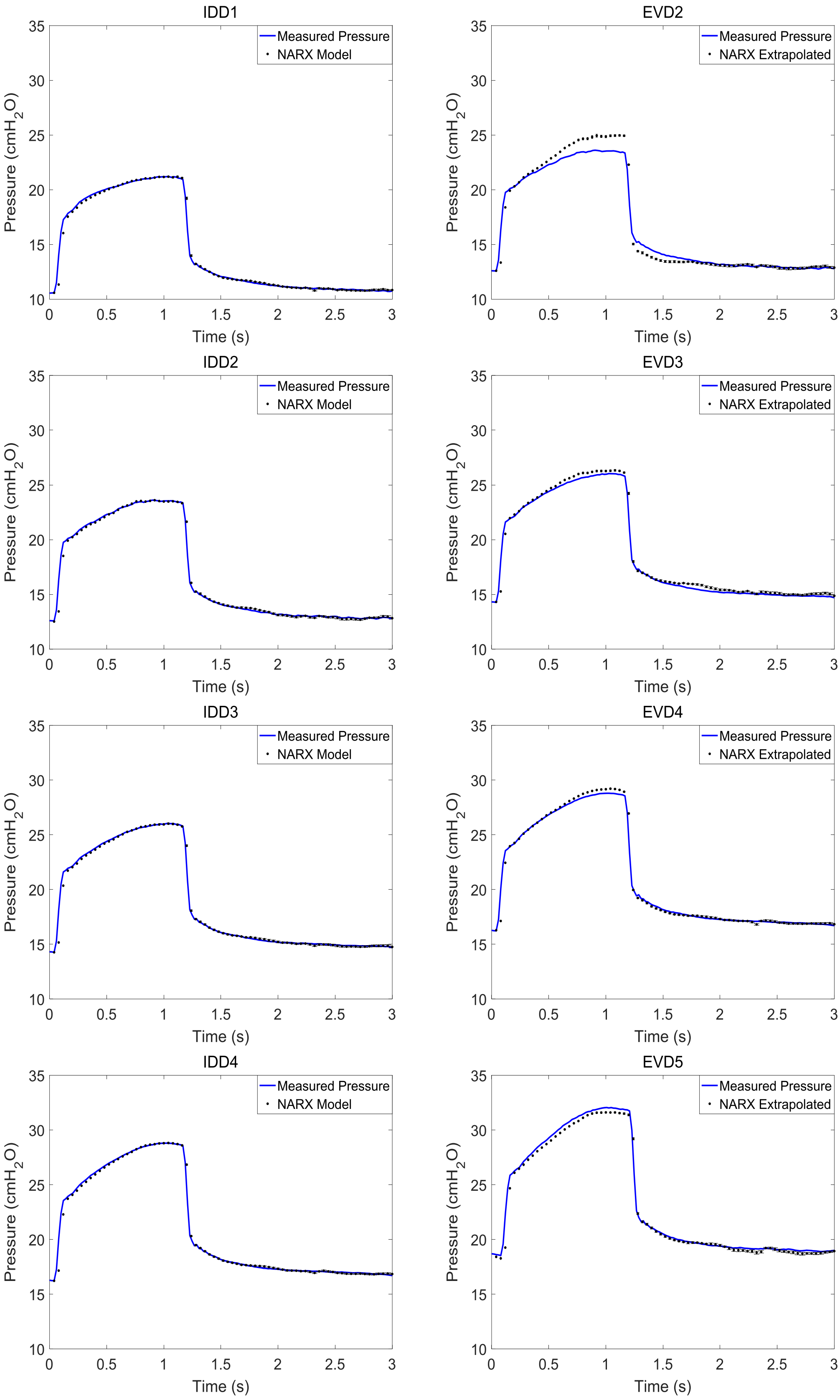


Figure 10.4. One patient's mean NARX model plotted relative to the average breath for IDD1 -> EVD2, IDD2 -> EVD3, IDD3 -> EVD4, and IDD4 -> EVD5. Error bars are the standard error (SE) = σ $/\sqrt{B_N}$, where σ = standard deviation, and B_N = number of breaths in the PEEP level.

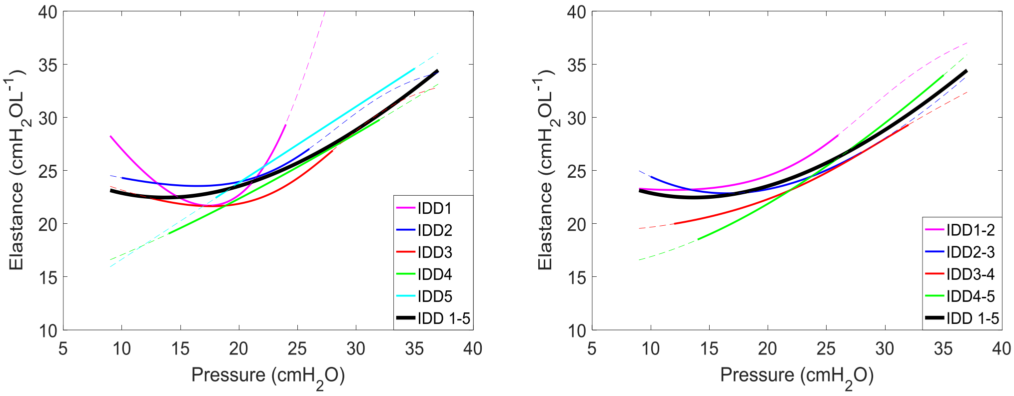


Figure 10.5. Elastance through pressure for one patient. Solid lines show the identified elastance and dashed lines show the extrapolation of the elastance shape. Left = NARX model identified on a single PEEP. Right = NARX model identified on two PEEP steps.

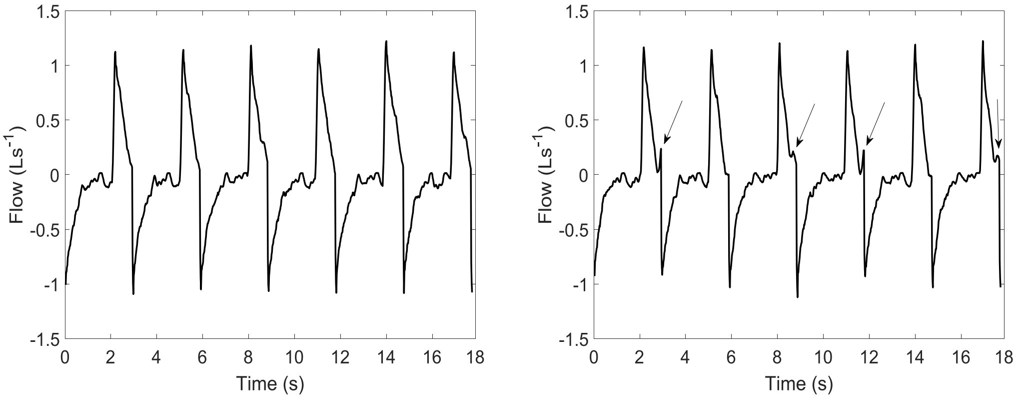


Figure 10.6. Flow signal for one patient at PEEP level 3 (left) and PEEP level 4 (right). Arrows show the location of the intermittent oscillations at the higher PEEP.

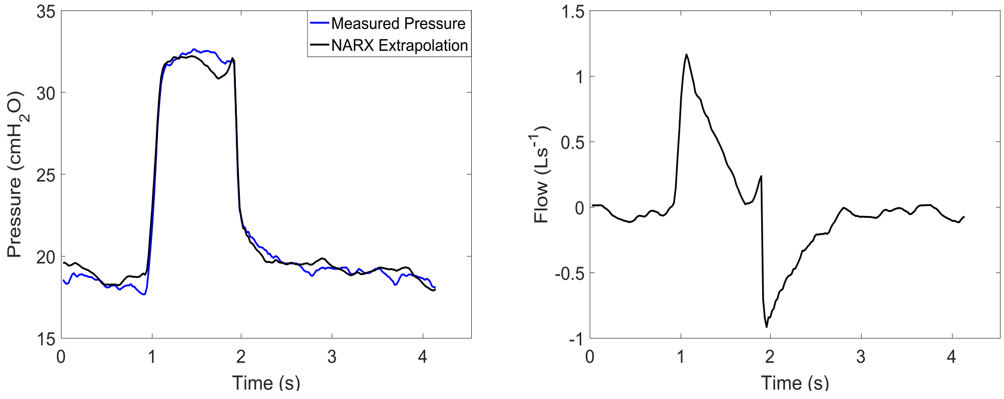


Figure 10.7. One patient's measured pressure and NARX extrapolation from IDD3 for a single breath during EVD4 (left), and the corresponding measured flow signal for this breath (right).

10.4 Discussion

The NARX extrapolation was more successful than the FOM in all scenarios that were tested. The mean RMS residual for prediction at the PEEP level following a single training level was 0.52 (90% CI: 0.47 – 0.57) cmH_2O for the NARX model and 1.50 (90% CI: 1.38 – 1.62) cmH_2O for the FOM (Table 10.1). The NARX model was similarly superior to the FOM when two, three, or four PEEP levels were used for training, and tested on the next PEEP step. Table 10.1 also shows that the NARX EVD RMS residuals were close to the IDD residuals, indicating that the NARX model predictions were not substantially worse than the model fit to the training data.

Figure 10.3 shows that prediction accuracy decreased as the EVD moved further away from the IDD, as expected. The RMS residuals of predictions up to four PEEP steps higher than the training set are shown. This result should not be interpreted as the RMS residual that may be expected if a single PEEP step of 8cm H_2O were undertaken. If a large PEEP jump was implemented, slow acting viscoelastic effects are unlikely to occur in the same manner as they do when smaller increments of pressure steps are undertaken. Furthermore, changes of such magnitude are likely to induce stress on pulmonary tissue and are unlikely to be clinically beneficial in any circumstance.

Figure 10.5 shows the elastance curves for one patient, and how they were extrapolated beyond the identification range, as IDD was increased from one to two PEEP steps. This figure shows that the NARX model elastance identified using only a single PEEP step does not closely match the elastance found when the model is identified on 100% of the data, as might be expected. In general, the elastance shapes identified in the middle range (IDD2, IDD3 and IDD4) were more accurate than the elastance shapes defined by the extreme PEEP levels (IDD1 and IDD5). When the training set included more than one PEEP level, the elastance shape was much closer to the shape identified on the whole dataset. This outcome ultimately means that the model was capable of improved prediction over all EVD PEEP levels when the training set was enlarged, as also shown in Figure 10.3.

Table 10.1 and Table 10.2 show the most clinically relevant results, which are the RMS residuals for the EVD one PEEP step higher than the IDD range. In two out of six cases, significantly smaller residuals were obtained in the EVD PEEP when the IDD contained more data ($p < 0.05$). In the other four cases, there was no significant difference, or the RMS residuals were worse on average for the larger IDD range. While this result was unexpected, it represents a potential benefit clinically. In particular, this outcome implies that less data would be required for a good prediction, and thus less time is needed for a useful prediction to be made. Overall, increasing the amount of training data did not result in any clear trend in prediction accuracy when the EVD was only one step up from the IDD.

Figure 10.4 shows a plot of the average breath in several IDD and EVD sections overlaid with the mean NARX model reconstruction, for one data set. While the error in EVD2 from IDD1 was substantial for this data set, there was a high accuracy in the prediction one PEEP level up from IDD2, IDD3, and IDD4. The error in the mean predicted pressure shown on the plot was less than 0.5 cmH_2O at each point of the breath, in these three cases. The overestimation of pressure for EVD2 indicates that the extrapolated elastance was too high. This result represents a potential for pulmonary over-distension to be predicted at lower than reasonable pressures. The poor elastance prediction from IDD1 is in concordance with Figure 10.5, which shows the higher accuracy of elastance shapes identified in the middle range (IDD2, IDD3, IDD4) compared to the elastance shapes defined by IDD1.

PEEP is an important model input that allows consistent model fits to the pressure troughs in each breath. The elastance metrics capture the tidal pressure (ΔP), and thus captures the peak inspiratory pressure. Capturing PIP is valuable as high PIP is a risk factor for VILI (Brower *et al.*, 2004). At pressure where the predicted elastance was higher than the elastance identified over the whole data set, the model tended to predict larger ΔP values, and thus predicted higher peak pressure in a breath. Similarly, when the predicted elastance was lower than the elastance for IDD1-5, the predicted peak pressure in a breath would be lower than the measured data. A predicted pressure that was much lower than what occurs in practice could lead to an unnecessarily high PEEP applied to a patient, which may lead to the onset of VILI. Hence, this model and methodology is recommended for predicting pressure responses due to PEEP increases of 2 cmH_2O or less, based on the data and results here. For this PEEP step magnitude, the methodology was observed to accurately predict peak pressure. Finally, as noted previously, once a further PEEP is tested it can be used in this method for further predictions before implementing subsequent PEEP levels.

Figure 10.6 and Figure 10.7 represent a scenario that was reasonably common at higher PEEPs for patients on the pressure-controlled ventilation mode. In pressure-controlled mode, the ventilator is directly controlling the pressure, whereas in volume-controlled mode the ventilator is directly controlling volume and flow. Therefore the flow signal in pressure-controlled mode tends to be much more variable in shape compared to volume-controlled mode. One factor that can cause discrepancy between the flow and pressure waveforms is SB efforts. These oscillations were observed in many patients, and were often much more frequent and had a much larger magnitude in the higher PEEP levels compared to lower PEEP (Figure 10.6). This behaviour occurs because high pressure can trigger SB efforts at certain levels of sedation (Major *et al.*, 2015). The incidence of SB is lower at lower ventilator induced pressure.

When the SB efforts in EVD flow data were not present in the IDD, the model was not trained to output normal pressure waveforms in the presence of these oscillations in the flow input. The result was oscillations in the predicted pressure waveforms that corresponded to the times that oscillations occurred in the flow. As a consequence, the correct inspiratory shape was not well captured (Figure 10.7). This issue resulted in much larger residuals for the EVD than would have been seen if this unmodelled effect was not present. This problem did not occur for the patients in volume controlled mode (e.g. Figure 10.4), because the flow is actively controlled by the ventilator. Thus, inspiratory SB efforts do not affect the flow waveform in volume controlled mode, and the pressure predictions are overall much more consistent.

The results of this analysis suggest that for this cohort of sedated ARDS patients, the full RM was not required to obtain an accurate estimate of lung mechanics across PEEP levels. However, there may be other clinical reasons for going through with the entire RM. For example, a RM may be beneficial to recruit alveoli. In contrast, an extrapolation from a smaller number of PEEP steps could be beneficial for patients who are identified as particularly susceptible to VILI at higher PEEP levels.

10.5 Summary

The NARX model predictions were more successful than the FOM. In particular, accurate predictions were made for PEEP steps of 2 cmH_2O higher than the PEEP in the identification data. While the model did not perform well when predicting over larger PEEP changes, such changes in PEEP are not recommended in a clinical setting due to the stress they put on pulmonary tissue. Furthermore, prediction of one PEEP step forward is all that is clinically necessary, as once the higher PEEP is applied, this data can be used in subsequent predictions of higher PEEPs.

The outcome is very important for the potential clinical use of the model. The primary benefit of the methodology would be the accurate prediction of over-distension prior to using a PEEP that causes this effect. The NARX model prediction provides important insight into the expected PIP that may be experienced by the patient if PEEP is increased. Since PIP is an important risk factor for VILI, the methodology could enable pulmonary treatment support algorithms to avoid harmful over-distension.

Chapter 11 Over-Distension Prediction

11.1 Introduction

In this section, the ability of the NARX model to predict high pressure is assessed via a statistical classification analysis. The pressures considered were the peak inspiratory airway pressure (PIP), and plateau pressure (PP). PP is a reflection of alveolar pressure, so high PP is an indicator of alveolar over-distension and likelihood of VILI (Morgan *et al.*, 2006). However, if an end-inspiratory pause is not performed, then PP is not directly measured. In this situation PIP can be considered. The PIP measured at the airway is generally higher than pressure experienced at the alveoli. However, since airway resistance reduces at high pressure, higher than expected PIP is likely to be due to increases in elastance. Thus, high PIP is indicative of alveolar over-distension, and can be used as a proxy metric for high alveolar pressure.

The detection of over-distension is of primary interest to clinicians, as the avoidance of over-distension will reduce incidence of VILI, and thus reduce mortality of ARDS patients. Just as optimal PEEP levels vary between patients, the pressures that cause VILI will also vary. However, PP < 30 cmH_2O is generally considered safe (Amato *et al.*, 2015, Jardin and Vieillard-Baron, 2007), as is PIP < 40 cmH_2O (Schranz *et al.*, 2012b).

Cohort1 was used to predict PP, and Cohort2 was used to predict PIP, as the Cohort2 protocol did not include an end-inspiratory pause. Prediction was performed over PEEP step increases of 2, 4, and 6 cmH_2O. The analyses considered whether Cohort1 PP exceeded 30 cmH_2O, and whether Cohort2 PIP exceeded 40 cmH_2O. Both analyses were compared with the FOM. The Cohort1 analysis is presented in Langdon *et al.* (2017c), and the Cohort2 analysis is in Langdon *et al.* (2017d).

11.2 Methods

For each RM, the NARX model was first trained on the baseline PEEP level, and was used to predict the PP or PIP at the following three PEEP steps, that were 2, 4, and 6 cmH_2O higher than the training PEEP. The predicted PP or PIP was compared with the measured value at the higher PEEP levels. Then the baseline and second PEEP level were used as the training set, and prediction was performed at the third, fourth, and fifth PEEP levels. The training set was increased until there was only a single PEEP step left for prediction. NARX model prediction was achieved by extrapolation of the basis functions to higher pressures. In contrast, the FOM used the constant identified elastance value from the training data to predict pressure.

A second method was used for the FOM evaluation. The FOM is generally considered to be most accurate over small pressure ranges. While the first FOM method, FOM(I), uses the same training set as the NARX model, FOM(II) uses the elastance determined from a single PEEP level to predict PIP or PP at one to three PEEP steps above that level. FOM(I) allows a direct comparison with the NARX model method, and FOM(II) allows the FOM to operate in a scenario that better represents the clinical use of the FOM.

For both cohort analyses, the true positive, true negative, false positive, and false negative incidences were recorded. Sensitivity and specificity were then calculated. ROC curves were generated by varying the discrimination threshold.

11.2.1 Cohort1 PP Analysis

Two different options for the basis functions were investigated for the application of the NARX model on Cohort1. NARX(I) uses the four basis functions described by Equations 10.1 to 10.4 and Figure 10.2. NARX(II) uses only two basis functions, the linear and the exponential functions, Equations 10.2 and 10.3. In the course of the analysis, it was determined that the NARX model *L* should be set to 1. This limited the model's ability to capture viscoelasticity. However, it allowed more successful PP predictions on Cohort1. Figure 11.1 shows one patient's identified and extrapolated elastance for the Cohort1 NARX and FOM methods.

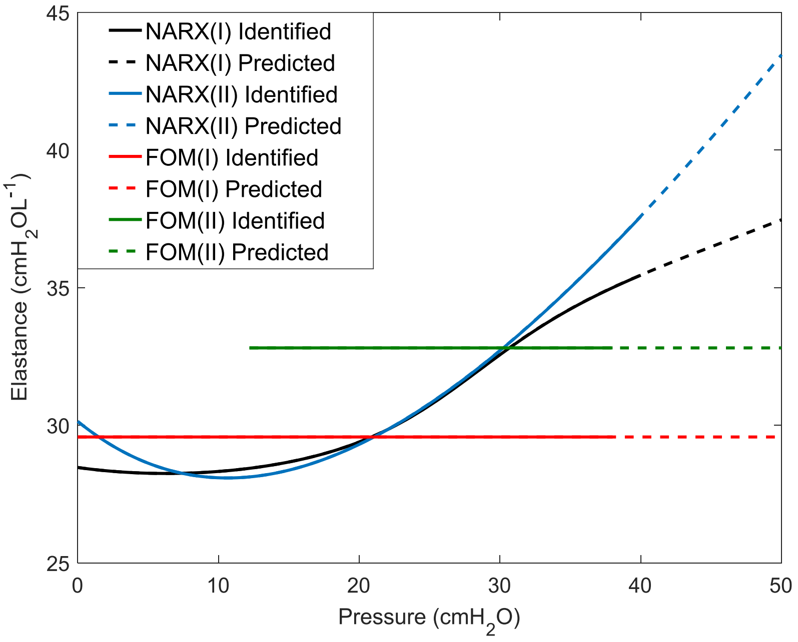


Figure 11.1. Identified and predicted elastance for a particular patient of Cohort1.

11.2.2 Cohort2 PIP Analysis

From Cohort2, one of the ten patients was excluded as they were not ventilated up to a PIP of 40 cmH_2O. In total, there were 16 data sets available for this analysis. The four basis functions described by Equation 10.1 to 10.4 were used, with L = 170. Figure 11.2 shows one patient's identified and extrapolated elastance for the NARX and FOM methods.

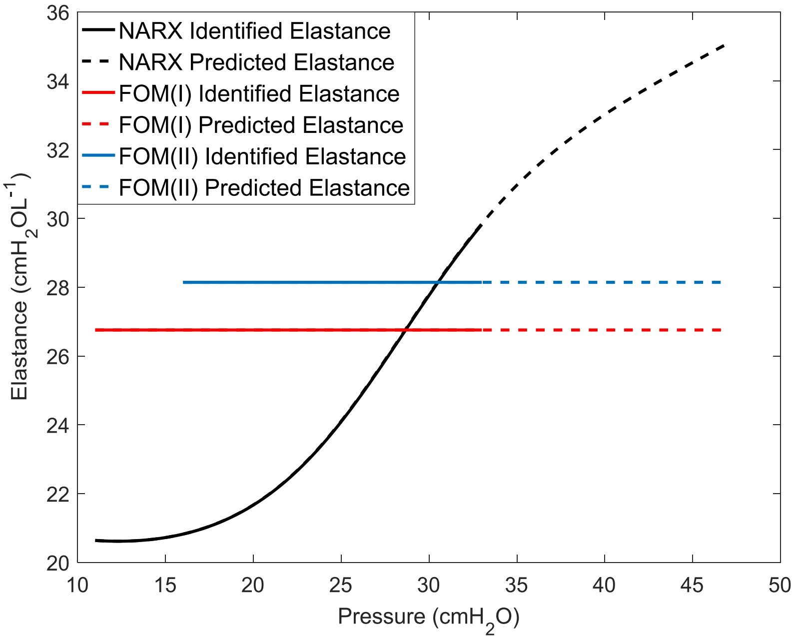


Figure 11.2. Identified and predicted elastance for a particular patient of Cohort2.

11.3 Results

11.3.1 Cohort1 PP Results

Table 11.1 lists the PP prediction results for each model. Compared with FOM(I), the FOM(II) specificity was consistently greater, averaging 6.3 percentage points higher. The NARX(I) and NARX(II) had very similar specificity in all cases. Overall, the specificity was fairly good for all models (> 0.90). Thus, the models can be more easily distinguished by looking at the sensitivity.

FOM(I) sensitivity was generally lower than the other models at ~0.85 at all prediction horizons. FOM(II) improved this measure for the prediction horizon of one PEEP step only. NARX(I) performed similarly to FOM(II), however the NARX(II) sensitivity was substantially higher, especially for prediction horizons of 4 and 6 cmH_2O. While NARX(I) sensitivity decreased with greater prediction horizons, the NARX(II) sensitivity remained >0.95.

Table 11.1. Cohort1 PP prediction results. Brackets denote the classifications for the 2, 4, and 6 cmH_2O prediction horizons.

		Predicted PP > 30 cmH_2O	Predicted PP < 30 cmH_2O	
NARX(I)	Measured PP > 30 cmH_2O	TP [86, 83, 76]	FN [9, 11, 15]	Sensitivity [0.91, 0.88, 0.84]
	Measured PP < 30 cmH_2O	FP [2, 4, 4]	TN [113, 89, 69]	Specificity [0.98, 0.96, 0.95]
NARX(II)	Measured PP > 30 cmH_2O	TP [90, 93, 88]	FN [5, 1, 3]	Sensitivity [0.95, 0.99, 0.97]
	Measured PP < 30 cmH_2O	FP [6, 4, 4]	TN [107, 89, 69]	Specificity [0.95, 0.96, 0.95]
FOM (I)	Measured PP > 30 cmH_2O	TP [81, 80, 78]	FN [14, 14, 13]	Sensitivity [0.85, 0.85, 0.86]
	Measured PP < 30 cmH_2O	FP [8, 7, 7]	TN [107, 86, 66]	Specificity [0.93, 0.92, 0.90]
FOM (II)	Measured PP > 30 cmH_2O	TP [86, 79, 76]	FN [9, 15, 15]	Sensitivity [0.91, 0.84, 0.84]
	Measured PP < 30 cmH_2O	FP [2, 1, 2]	TN [113, 92, 71]	Specificity [0.98, 0.99, 0.97]

Figure 11.3 and Figure 11.4 show the correlation between the measured and predicted plateau pressures for each model where prediction was a single PEEP step up from the training data, as well as the Bland-Altman analysis. ΔPP denotes the difference between measured and predicted PP. NARX(I) and NARX(II) had a smaller spread of ΔPP compared to the FOM, particularly when compared with FOM(I), indicating smaller errors in prediction overall. The predicted PPs that resulted in false negatives in both NARX models were very close to the 30 cmH_2O threshold. In contrast, the false negative results in the FOM often resulted from predictions that were much lower than 30 cmH_2O.

Dotted lines in the Bland-Altman plots give the 25th, 50th, and 75th percentiles of the difference in measured and predicted PP. The median bias in all cases was relatively small. NARX(II) had a negative median bias of -0.2 cmH_2O, while NARX(I) had a slightly larger negative bias at -0.8 cmH_2O. FOM(II) was similar with a median bias of -0.7 cmH_2O. FOM(I) had the largest prediction errors overall, but the bias was the smallest at 0.1 cmH_2O.

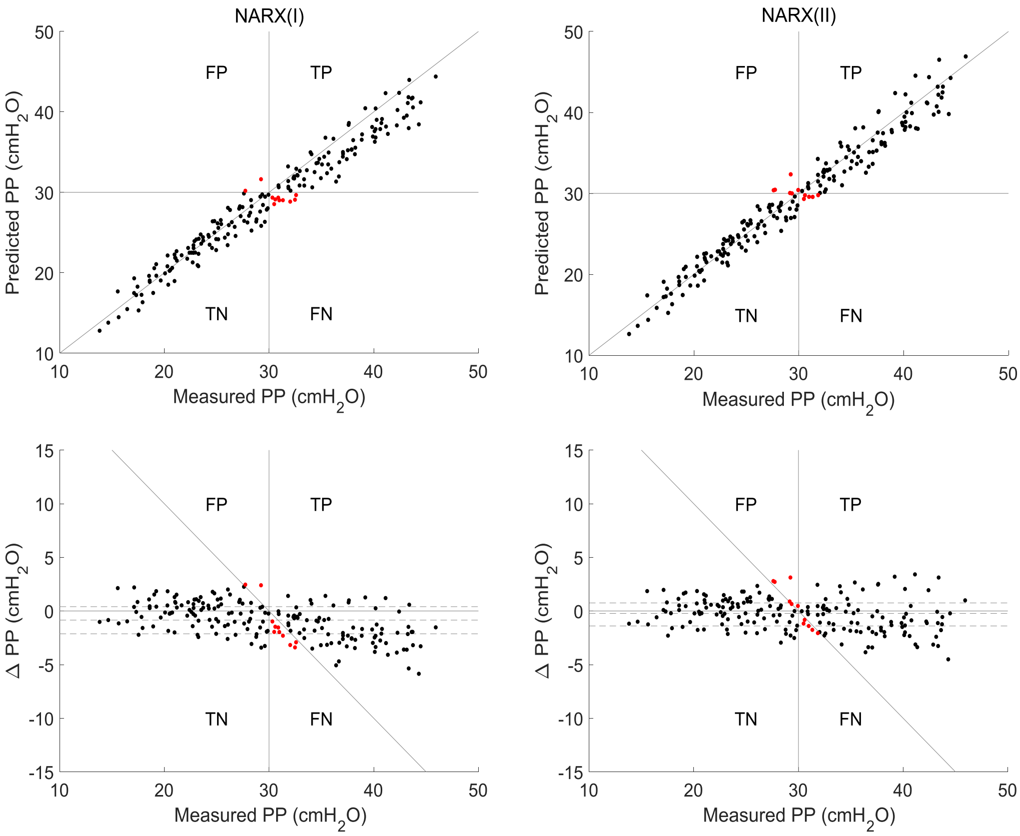


Figure 11.3. Measured and predicted PP relationship for NARX(I) (left) and NARX(II) (right). Prediction horizon = 2 cmH$_2$O.

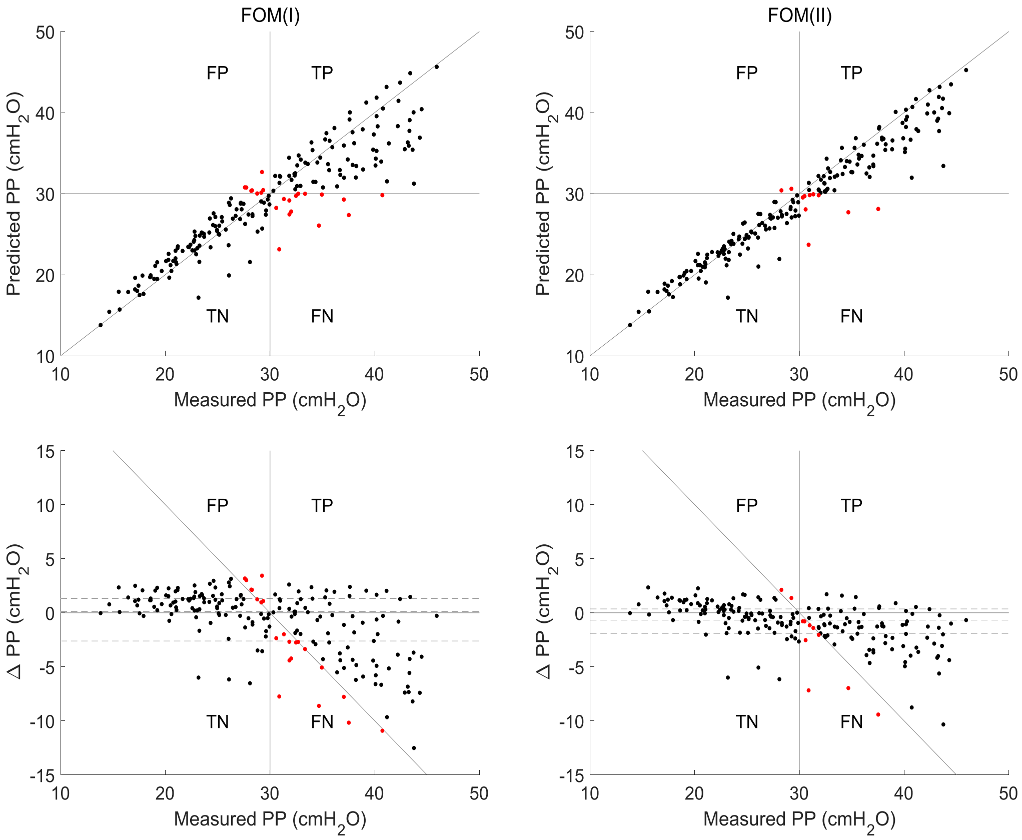


Figure 11.4. Measured and predicted PP relationship for FOM(I) (left) and FOM(II) (right). Prediction horizon = 2 cmH$_2$O.

The ROC curves in Figure 11.5 plot the TP rate for a single PEEP step prediction against the FP rate as the discrimination threshold was varied. This figure again shows that NARX(I), NARX(II), and FOM(II) were very similar predictors over a small prediction horizon. All models performed significantly better than a random classifier. The optimal threshold was 32.0 for NARX(I), 33.5 for NARX(II), 32.4 for FOM(I), and 31.5 for FOM(II).

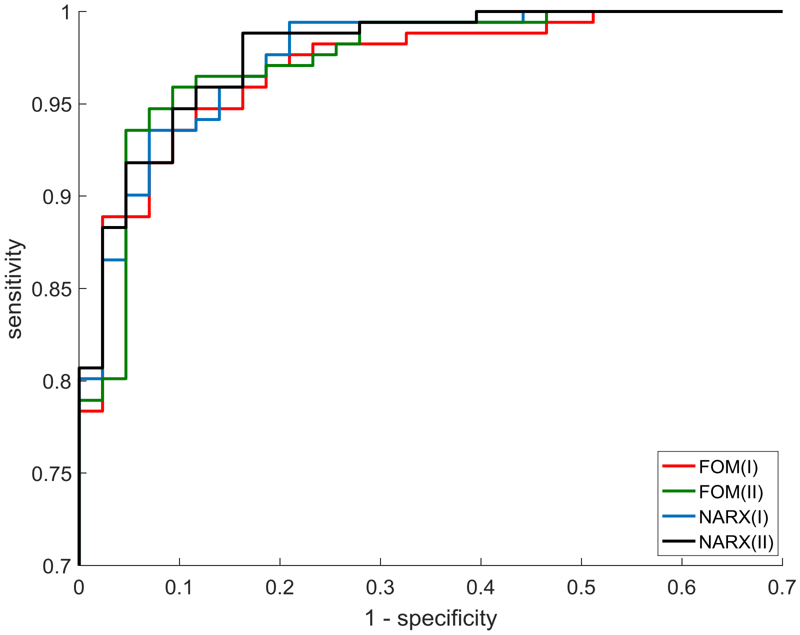


Figure 11.5. ROC curves for PP prediction.

11.3.2 Cohort2 PIP Results

Table 11.2 lists the PIP prediction results for each model. The NARX model sensitivity was substantially better than the FOM(I) or FOM(II) in all cases. As expected, FOM(II) was more sensitive than FOM(I). The specificity of the NARX model was lower than FOM(I) and FOM(II). There is a general, but imperfect, trend for the prediction metric scores of each model to be poorer as the prediction PEEP moves further from the training PEEP.

Table 11.2. Cohort2 PIP prediction results. Brackets denote the classifications for the 2, 4, and 6 cmH_2O prediction horizons.

		Predicted PIP > 40 cmH_2O	Predicted PIP < 40 cmH_2O	
NARX	Measured PIP > 40 cmH_2O	TP [45, 43, 42]	FN [0, 2, 1]	Sensitivity [1.00, 0.96, 0.98]
	Measured PIP < 40 cmH_2O	FP [3, 5, 3]	TN [54, 36, 24]	Specificity [0.95, 0.88, 0.89]
FOM(I)	Measured PIP > 40 cmH_2O	TP [31, 30, 29]	FN [14, 15, 14]	Sensitivity [0.69, 0.67, 0.67]
	Measured PIP < 40 cmH_2O	FP [0, 0, 1]	TN [57, 41, 26]	Specificity [1.00, 1.00, 0.96]
FOM (II)	Measured PIP > 40 cmH_2O	TP [39, 38, 32]	FN [6, 7, 11]	Sensitivity [0.87, 0.84, 0.74]
	Measured PIP < 40 cmH_2O	FP [0, 0, 1]	TN [57, 41, 26]	Specificity [1.00, 1.00, 0.96]

The relationship between measured PIP and predicted PIP across Cohort2 is given in Figure 11.6 and Figure 11.7, for the prediction at a 2 cmH_2O PEEP step. The three cases of false positive detection by the NARX model occurred when the measured PIP was very close to 40 cmH_2O. In all three of these cases, measured PIP was ≥ 38.6 cmH_2O, indicating the NARX model prediction was very close to the clinical outcome.

Bland-Altman plots show a bias in the NARX model that caused predictions to be higher than measured pressure (median bias = 0.8 cmH_2O). In contrast, FOM(I) had a bias towards low predictions (median bias = -1.2cmH_2O) , and FOM(II) was also slightly biased low at -0.07 cmH_2O.

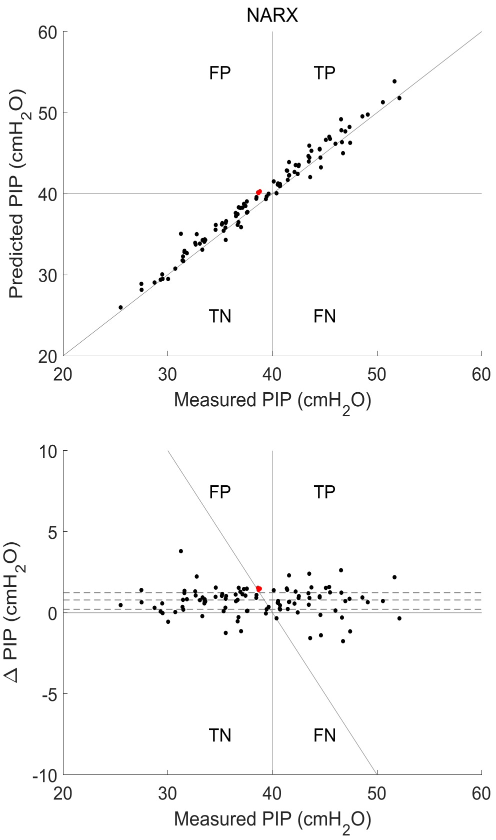


Figure 11.6. Measured and predicted PIP relationship for the NARX model. Prediction horizon = 2 cmH$_2$O.

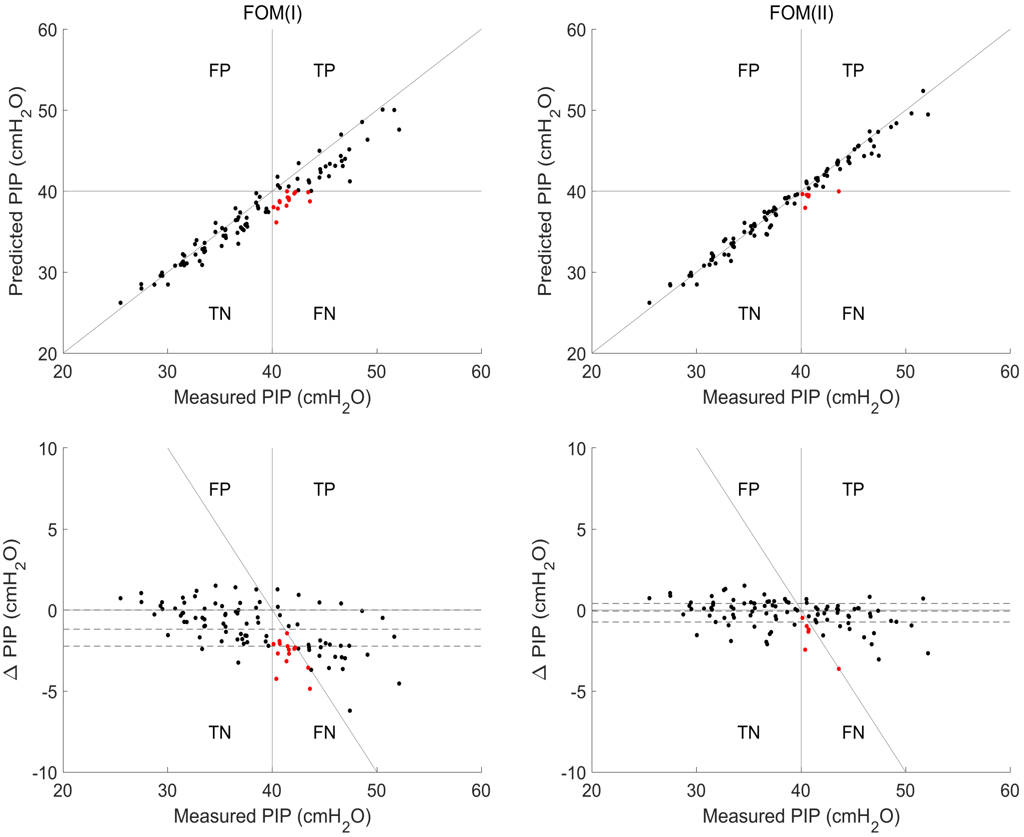


Figure 11.7. Measured and predicted PIP relationship for FOM(I) (left) and FOM(II) (right). Prediction horizon = 2 cmH$_2$O.

The ROC curves in Figure 11.8, plot the TP rate of prediction over a single PEEP step against the FP rate as the discrimination threshold was varied. While all models performed significantly better than a random classifier, the area under the curve was clearly greatest for the NARX model, indicating the best overall performance. The optimal threshold was 40.3 cmH_2O for the NARX model, 37.8 cmH_2O for FOM(I), and 39.3 cmH_2O for FOM(II).

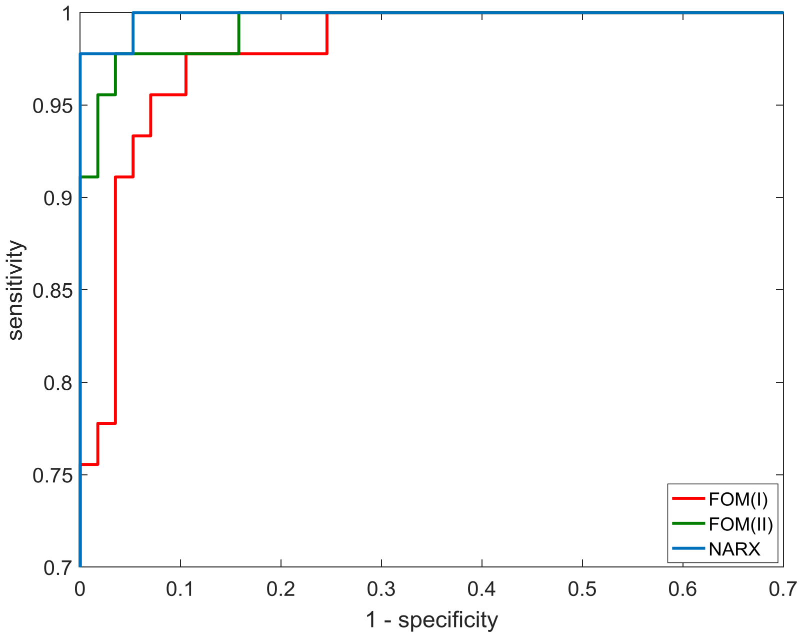


Figure 11.8. ROC curves for PIP prediction.

Figure 11.9 shows results that were typical for this cohort. In this particular instance, the FOM and NARX models were identified on six PEEP steps and extrapolated to the seventh PEEP step. The NARX model provided an accurate prediction of the pressure waveform and correctly determined that PIP at the seventh PEEP step would be greater than 40 cmH_2O. The FOM(I) prediction was substantially lower than the true PIP at the seventh PEEP step and gave a false negative result. FOM(II) was able to correctly predict that PIP would exceed 40 cmH_2O at the seventh PEEP step. However, the predicted waveform still undershot the PIP, and was clearly a worse fit than the NARX model.

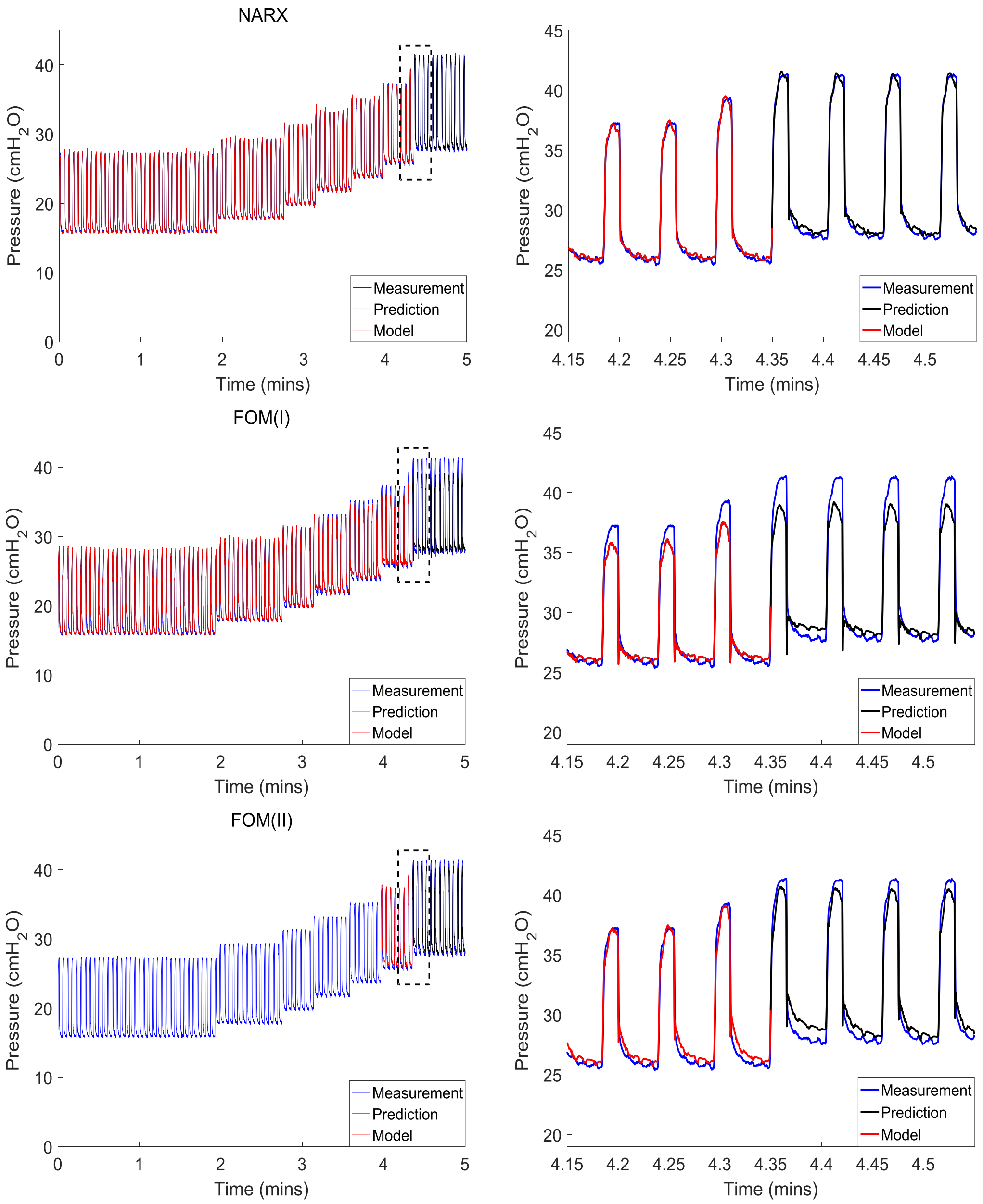


Figure 11.9. Training data and PIP prediction for the NARX model (top), FOM(I) (middle), and FOM(II) (bottom). Plots on the right are zoomed in to the box shown in the corresponding left figure.

11.4 Discussion

Clinically, false negatives are much more harmful to patients than false positives. In the context of pulmonary over-distension, a false negative may result in the application of high ventilation pressures that lead to over-distension and VILI. In contrast, a false positive result may mean clinicians do not increase PEEP, and miss out on the potential for improved recruitment at the higher PEEP level. Hence, the comparative impact of a false negative is stronger than the impact of a false positive, and thus, a high sensitivity should be favoured when predicting the outcomes of potential treatment methods.

The pressure dependent basis functions of the NARX model enable it to predict the respiratory mechanics of patients when they are ventilated at higher pressure levels than experienced in the training data. The Cohort2 PIP prediction analysis showed that the NARX model exhibited very high sensitivity in all cases (> 0.96), and a high specificity (> 0.88). The FOM had high specificity, but the sensitivity was much lower, especially for FOM(I) (Table 11.2). In Cohort1, the four basis function NARX model did not perform substantially better than FOM(I) or FOM(II) (Table 11.1). However, with reduced parameterisation, NARX(II) sensitivity remained high at prediction PEEP increases of up to 6 cmH_2O.

The high false negative rate in both PP and PIP prediction by the FOM was due to the constant elastance term. In reality, elastance changes according to recruitment and distension effects. At pressure levels near the 30 cmH_2O or 40 cmH_2O threshold, most recruitment will have already occurred. Thus, elastance is likely to be increasing with pressure, and the single elastance term identified in the FOM will lead to underestimated elastance and thus underestimated PIP or PP. This led to the high rate of false negatives given by FOM(I) and FOM(II).

FOM(II) yielded better results than FOM(I) even though it is trained on a reduced data set. This result occurred because it was trained in a region that is on average closer to the regions of prediction, and the single elastance of FOM(II) did not need to partially represent the lower elastance of the lower pressure ranges. However, the predicted pressures of FOM(II) still tended to be lower than the NARX predictions and had higher residuals than the NARX model (Figure 11.9).

The FOM(I) and FOM(II) specificities were often higher than the NARX model specificities for each prediction horizon. This was an expected result, as a FOM false positive would only occur if elastance was decreasing when PIP was near 40 cmH_2O. When PIP is close to 40 cmH_2O, elastance is very unlikely to decrease as all recruitable lung regions are likely to have already been recruited at lower pressures. FOM false positives would be more likely to occur in a decreasing PEEP scenario. Though there was no data available to confirm this, a single elastance identified from high pressure data is likely to be too high for an accurate prediction when PEEP is decreased. In this scenario, the NARX model would be expected to perform better than the FOM, due to the ability to easily extrapolate a continuous elastance to lower pressures.

In the Cohort2 PIP analysis, there were no false negative and only three false positive predictions from the NARX model over a prediction horizon of one PEEP step. In the three false positive cases, the mean difference between measured and predicted PIP was only 1.5 cmH_2O. For the NARX false positives that occurred at prediction horizons of two and three PEEP steps, the mean difference between measured and predicted PIP was 2.4 cmH_2O in both cases. Thus, the false positives did not actually represent poor prediction.

In Cohort2, there were seven patients on pressure controlled ventilation and three patients on volume controlled ventilation. In pressure controlled data, the PIP is a setting defined by the clinician, and thus it may seem strange to analyse the ability of the model to predict PIP in pressure controlled mode. However, none of the modelling approaches used in the analyses incorporated a-priori information on the applied ventilator settings. In contrast, the modelling approaches simply provide a transfer function between pressure and flow. Thus the ability to predict pressure from flow data remains scientifically valid, even in pressure controlled mode, when the model does not use the ventilator settings as an input.

In Cohort1, the NARX(I) and NARX(II) both predicted precise PP and thus the sensitivity and specificity at a single PEEP-step prediction threshold were ≥ 0.91. However, NARX(I) had reduced prediction precision over greater prediction horizons (Table 11.1). In contrast, the NARX(II) consistently maintained a high sensitivity (≥ 0.95) and specificity (≥ 0.95) up to a prediction horizon of 6 cmH_2O. The improved prediction performance of the NARX(II) is due to the reduced parameterisation of the model. NARX(I) utilises four basis-functions to capture the patient-specific pressure/elastance relationships whereas NARX(II) uses only two. Hence, while NARX(I) has more ability to fit accurately to the clinical data, small levels of parameter trade-off may propagate to generate lower prediction precision in the extrapolated elastance profile. NARX(II) only uses two basis functions and thus, the level of parameter trade-off is lowered.

The improved precision of NARX(II) over NARX(I) can be observed in the Bland-Altman plots in Figure 11.3. The NARX(I) model shows a general trend to under-predict PP, with a median bias of -0.8 cmH_2O. This led to the lower sensitivity observed in Table 11.1. The NARX(II) model resulted in a bias of much smaller magnitude (-0.2 cmH_2O) and thus retained the high sensitivity at high PEEP levels.

This work highlights the need for further investigation into the optimal level of parameterisation of the NARX model when it is used for prediction of pulmonary mechanics, in particular for patients with more complex breath waveforms such as those of Cohort1. While NARX(I) has more basis functions to fit the data, and thus can provide better model fits, its predictive ability on Cohort1 was lower than NARX(II) which utilises only two basis functions. Hence, the increased robustness of the NARX(II) model enables more accurate prediction when used on the complex end-inspiratory pause data.

Clinically, once a dangerous PIP or PP has be recorded, PEEP would not be increased further as the risk of over-distension and VILI would be high. Thus, outside of an initial RM, PEEP might not be increased to allow PIP or PP beyond the clinician's chosen threshold. In both the Cohort1 and Cohort2 analyses, we included all available data from the RMs and thus contradicted the clinical process. However, it was scientifically necessary to establish that the models did not erroneously predict low pressure at high PEEP. Furthermore, ceasing the evaluation at the first instance of PIP > 40 cmH_2O or PP > 30 cmH_2O would generate a low true positive rate, and thus the sensitivity calculation would be unrepresentative of the true abilities of the model.

11.5 Summary

In the PIP Cohort2 analysis, The NARX model predicted high peak pressures more accurately than the FOM, and importantly, had very low instances of false negatives. Zero false negatives occurred at the prediction horizon of one PEEP step. While FOM(II) improved the FOM sensitivity over FOM(I), it was still less successful than the NARX model. The error in the NARX false positive predictions was small, and the consequences of false positives are far less severe than the potential VILI that may result from false negative predictions.

In the PP Cohort1 analysis, all four models yielded high specificity in all scenarios (≥ 0.90). The FOM(I), FOM(II) and NARX(I) sensitivity were similar across all prediction horizons, and generally decreased as the prediction PEEP increased. However the reduced parameterization of NARX(II) compared with NARX(I) allowed the sensitivity to remain high at ≥ 0.95 at all prediction horizons. The results highlight the importance of finding a balance between model fitting and predictive ability when considering optimal model parameterization.

Chapter 12 Auto-PEEP Patients

12.1 Introduction

Some MV patients suffer from auto-PEEP, otherwise known as intrinsic PEEP. Auto-PEEP involves the accumulation of air in the lungs, caused by incomplete expiration, and results in increased alveolar pressure. Auto-PEEP is common in patients with COPD, as the narrow or blocked airways associated with COPD can cause expiratory flow restriction (Reddy, 2005). In partially sedated patients, this phenomenon increases WOB as auto-PEEP must be overcome by the patient to trigger a breath (Rossi *et al.*, 1990, Singer and Corbridge, 2009a). PEEP is especially important in these patients as increased PEEP levels can open blocked airways and reduce WOB. In this section, the NARX model is modified to identify clinically significant differences between patients with high and low auto-PEEP. In particular, the flow dependent terms were replaced with basis functions to capture linear resistance changes with pressure. Cohort3 data was used in this analysis, as four out of the ten patients in this cohort had high measured auto-PEEP (≥ 5 cmH_2O). The contents of this chapter is presented in Langdon *et al.* (2017a).

12.2 Methods

The adapted NARX model is given by:

$$P(t) = \sum_{i=1}^{3} a_i \emptyset_{i,d}\big(P(t)\big)V(t) + \sum_{j=0}^{2} b_j \varphi_j(P(t))\dot{V}(\mathrm{t}) + P_0(t) \qquad 12.1$$

where $\emptyset_i$ are the elastance basis functions, and φ_i are the resistance basis functions. The basis functions are defined:

$$\emptyset_1 = \tfrac{P(t)}{45};\ \emptyset_2 = e^{-0.04P(t)};\ \emptyset_3 = 1 \qquad 12.2$$

$$\varphi_1 = P(t);\ \varphi_2 = -P(t) + 45 \qquad 12.3$$

A linear combination of the elastance basis functions allows a convex pressure dependent elastance shape to be captured, and a linear combination of the resistance basis functions allows a linear pressure dependent resistance to be captured (Figure 12.1).

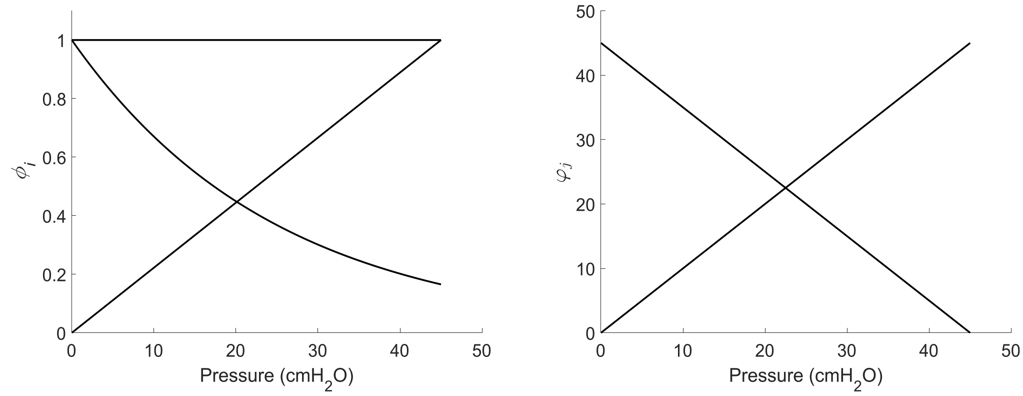


Figure 12.1. Elastance basis functions (left), and resistance basis functions (right).

The model was fit to the ten patient data sets of Cohort3. RMS residuals between measured and modelled pressure were calculated, and the elastance and resistance across pressure were plotted for each patient. The main aim was to compare the parameters identified for high and low auto-PEEP patients. The t-test was used to measure differences in elastance and resistance at high and low pressure, for each group.

12.3 Results

Figure 12.2 shows the typical result of fitting the basis function model to patients with low auto-PEEP and patients with high auto-PEEP. For patients with low auto-PEEP, the residuals were low and unbiased across the entire RM. For patients with high auto-PEEP, the model tended to undershoot measured pressure at the beginning of the RM, and overshoot measured pressure in the middle of the RM, while giving a good fit to the higher pressure breaths.

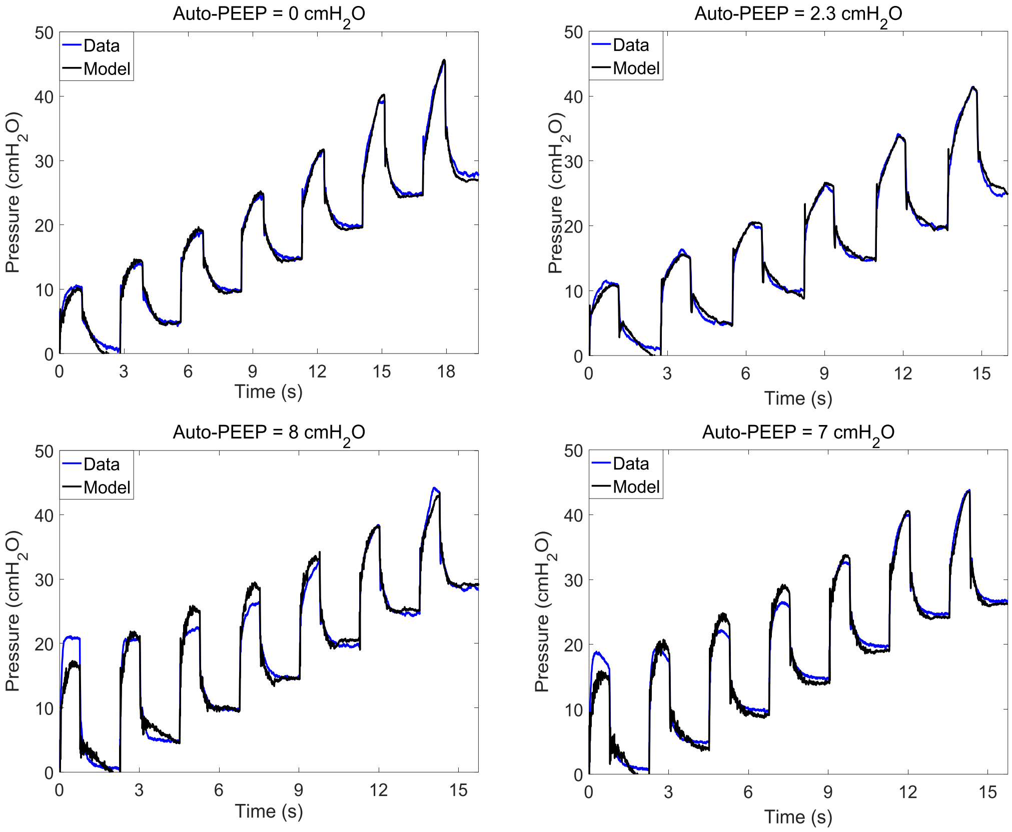


Figure 12.2. Model fit to measured pressure for two low auto-PEEP patients (top) and two high auto-PEEP patients (bottom).

The mean RMS residuals were 0.93 (range 0.73 – 1.18) cmH₂O in the low auto-PEEP group and 2.1 (range 1.42 – 2.91) cmH₂O in the high auto-PEEP group. The RMS residuals for the patients with low auto-PEEP were significantly lower than for the patients with high auto-PEEP (t-test, $p < 0.01$). The correlation between auto-PEEP and RMS error was $R = 0.86$, giving $R^2 = 0.74$ (Figure 12.3).

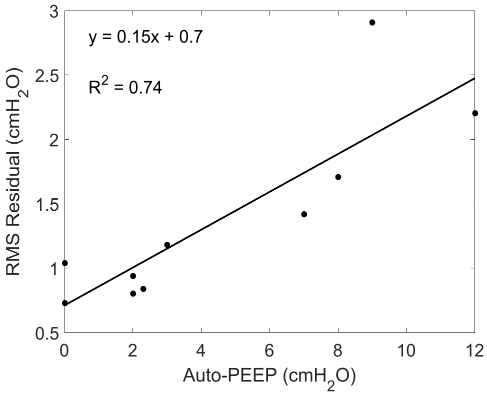


Figure 12.3. Relationship between auto-PEEP and RMS residuals for the NARX model.

Resistance decreased with pressure for nine out of 10 patients (Figure 12.4). Patients with high auto-PEEP had significantly higher modelled resistance at low pressure than patients with low auto-PEEP (t-test, $p < 0.01$). Additionally, the resistance gradients were significantly steeper in the patients with high auto-PEEP (t-test, $p < 0.01$), leading to similar resistance values at high pressure across both groups ($p = 0.5$).

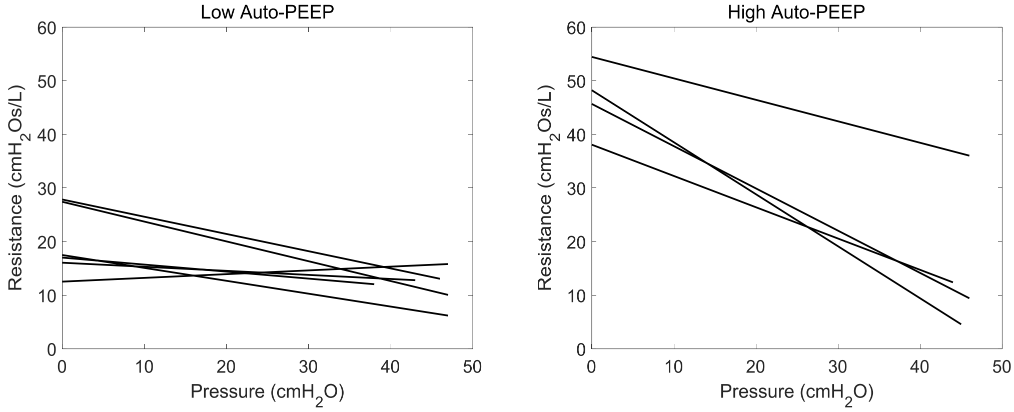


Figure 12.4. Modelled resistance of patients with low auto-PEEP (left) and high auto-PEEP (right).

In Figure 12.5, the elastance of patients with high and low auto-PEEP is plotted. While the difference was not significant, the high auto-PEEP group tended to have higher elastance at low pressure compared to the low auto-PEEP group. The low auto-PEEP group also tended to have the minimum elastance occur at lower pressures.

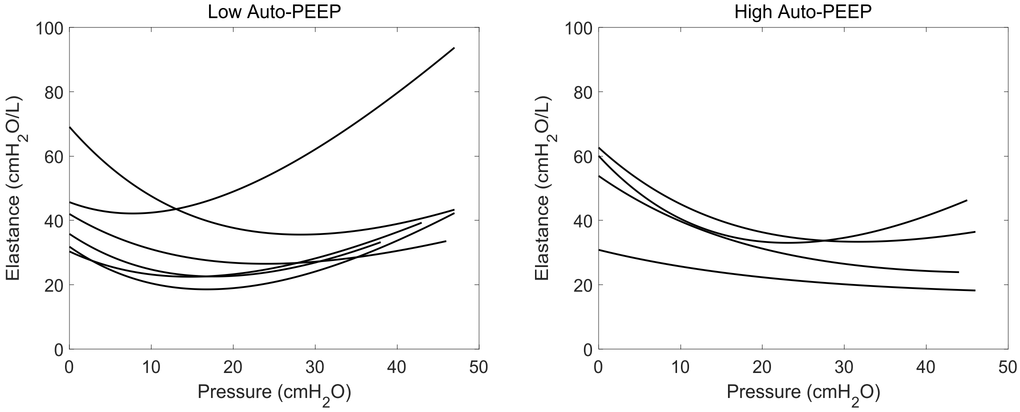


Figure 12.5. Modelled elastance of patients with low auto-PEEP (left) and high auto-PEEP (right).

12.4 Discussion

The model contains two basis functions that allow it to capture a linear resistance across pressure (Figure 12.4). Resistance is expected to decrease with pressure. In particular, higher pressure causes widening of bronchial passages, and the larger diameter lowers resistance according to the Poiseuille law (Sutera and Skalak, 1993). This behaviour was observed in nine out of the ten patients. However, one patient exhibited a mild positive resistance gradient. This unexpected result may be due to patient state changing within the RM, such as new obstructions to airflow arising, or potentially due to parameter trade-off, or other factors.

Patients with high auto-PEEP had significantly higher modelled resistance at low pressure, and significantly steeper resistance gradients overall (Figure 12.4). COPD is associated with higher than normal resistance at ZEEP due to blockages and narrowing of the bronchi and bronchioles. An increasing PEEP has the potential to lower resistance more dramatically in COPD lungs compared to healthier lungs, through opening these blockages. The RM was able to bring resistance down to a similar level as the patients with low auto-PEEP, in three out of the four cases. Overall, the resistance outcomes observed in this study were physiologically sensible.

Modelled elastance was not significantly distinct across patients with high auto-PEEP and patients with low auto-PEEP. However, there was an observed trend for patients with high auto-PEEP to have flatter elastance curves with higher elastance at low pressure (Figure 12.5). This implies that the patients with high auto-PEEP had lower rates of distension than those with low auto-PEEP, but that they also may have had a greater proportion of un-recruited alveoli at low pressure. This outcome concurs with the known physiology of COPD lungs. Firstly, airway blockages in COPD can cause a greater number of un-recruited lung units to exist at ZEEP. Secondly, in non-COPD lungs, maximum recruitment would be achievable at lower PEEP, and thus over-distension is more likely to occur at lower pressures compared to COPD lungs.

Model residuals were low and unbiased across the entire RM for patients with low auto-PEEP. However, the model fit was comparatively poor in the high auto-PEEP patient data sets (Figure 12.2). The model was biased towards lower than measured pressure at ZEEP, and higher than measured pressure in the middle of the RM. This result was the consequence of high peak pressure at ZEEP that did not increase with externally applied PEEP until PEEP exceeded the auto-PEEP. The unusually high peak pressure indicates that the true elastance at ZEEP is much higher than the model could account for. Therefore, modelled elastance at ZEEP was too low, and modelled pressure at ZEEP was substantially lower than measured pressure. This meant that the least squares approach required the model to overshoot measured pressure at higher PEEP to provide an overall best fit to the data.

Future work could involve a modification to the model to improve the fit for patients with high auto-PEEP. The modification may include a discontinuous term to fit low PEEP data that terminates once PEEP exceeds auto-PEEP. This may improve the accuracy of the identified resistance and elastance parameters. Regardless, the model in its current state was still able to identify clinically relevant differences in patients with high and low auto-PEEP.

12.5 Summary

The NARX model was adapted to capture changes in resistance and elastance noted to occur in COPD patients when pressure is increased. Model residuals were low and unbiased across the entire RM for patients with low auto-PEEP. While the high auto-PEEP model fit had higher residuals, the identified parameters captured the dysfunction expected in COPD and its effect on RMs. A statistically significant reduction in resistance at high pressure was found in high auto-PEEP patients compared to patients with different pulmonary dysfunctions. The elastance curve of COPD patients indicated a greater proportion of un-recruited pulmonary tissue at low pressure than the non-COPD group. The COPD group elastance also suggested a lower incidence of pulmonary over-distension than the non-COPD group. All of these outcomes imply COPD patients benefit from higher positive pressures than non-COPD patients. Ultimately, the basis function model was able to demonstrate clinically important differences between patients with high and low auto-PEEP.

Part 4: Glucose Modelling and *SI* Prediction

Chapter 13 Glucose Modelling with Basis Functions

13.1 Introduction

Hypoglycaemia, hyperglycaemia and glycaemic variability in critical care are each individually associated with increased mortality and morbidity (Christiansen *et al.*, 2004, Capes *et al.*, 2000, Kerby *et al.*, 2012, Bosarge *et al.*, 2015, Egi *et al.*, 2006, Krinsley, 2008). Many critically ill patients are admitted to the ICU with hyperglycaemia, very high insulin resistance, and very high glycaemic metabolic variance, which can be exacerbated by sepsis (Mizock, 2001, McCowen *et al.*, 2001, Marik and Raghavan, 2004). This combination of dysfunctions requires blood glucose to be intensively monitored and controlled. Glycaemic control (GC) algorithms have been proven to reduce blood glucose levels, variability, and the risk and incidence of hypoglycaemia (Evans *et al.*, 2011, Chase *et al.*, 2008, Morris *et al.*, 2008, Blaha *et al.*, 2005, Van den Berghe *et al.*, 2006, Lonergan *et al.*, 2006b, Wong *et al.*, 2006). Model based GC approaches can reduce the rate and severity of organ failure, and also reduce mortality. (Chase *et al.*, 2010, Stewart *et al.*, 2016, Van den Berghe *et al.*, 2006, Krinsley, 2004, Chase *et al.*, 2008).

In this section, the basis function concept is applied to a model of glucose and insulin kinetics. B-spline basis functions are used to identify *SI* over time. The way in which the number and order of basis functions affect *SI* variability and model fit is investigated. The majority of Chapter 13 and all of Chapter 14 is presented in Langdon *et al.* (2017b).

13.2 Methods

13.2.1 Data

This study uses 30 data sets covering 6785 hours of GC from patients treated using the Specialised Relative Insulin Nutrition Table (SPRINT) protocol for GC in critically ill patients (Chase *et al.*, 2008). Data was collected between August 2005 and May 2007. Measurements were usually hourly (2420 samples) or two hourly (1365 samples). There were occasional gaps of three hours (222 samples) or more (202 samples) between measurements. The patient age range was 37 – 86 years, with a mean of 65. Diagnoses included sepsis, respiratory failure and pneumonia. Baseline blood glucose varied between 3.3 and 19.7 mmol.L^{-1}, with a mean of 8.3 mmol.L^{-1}. Each patient's data spanned multiple days, and had a differing number of data points (*N*). Further details on SPRINT and its development can be found in (Lonergan *et al.*, 2006a, Lonergan *et al.*, 2006b, Chase *et al.*, 2008, Chase *et al.*, 2007). Ethics approval to collect, audit, and present these data was obtained from the South Island Regional Ethics Committee, New Zealand.

13.2.2 The Glucose Model

Linear differential equations were used to model insulin and glucose kinetics. The insulin model was a two compartment approach with clearance from both plasma and interstitial compartments (Equations 13.1 and 13.2) (Polonsky *et al.*, 1986). The glucose model (Equation 13.3) was a linear adaptation of a clinically validated single compartment approach (Lotz *et al.*, 2010). The linear adaptation allowed analytical forward simulations, and thus enabled quicker computational simulation.

$$\dot{I}(t) = -k_1 I(t) + k_2 Q(t) + \frac{U}{V_p} \qquad 13.1$$

$$\dot{Q}(t) = k_3 I(t) - k_4 Q(t) \qquad 13.2$$

$$\dot{G}(t) = p_G\big(G_b - G(t)\big) + \frac{P_x(t) + E_{GP}}{V_G} - SI(t)G(t)Q(t) \qquad 13.3$$

G(t) is the total plasma glucose (mmol.L^{-1}), $I(t)$ is the plasma insulin (mU.L^{-1}), $Q(t)$ is the interstitial insulin (mU.L^{-1}), $P_x(t)$ is exogenous glucose (mmol.min^{-1}), U is the insulin input to plasma (mU.min^{-1}), and $SI(t)$ is the identified insulin sensitivity profile (L.mU^{-1}.min^{-1}). The model constants are given in Table 13.1.

Table 13.1. Constant glucose model parameters

Variable	Description	Value
p_G	Non-insulin mediated glucose uptake	0.004 min^{-1}
V_G	Glucose distribution volume	9 L
V_P	Plasma insulin distribution volume	4 L
E_{GP}	Endogenous glucose production	1.5 mU.min^{-1}
G_b	Basal glucose	5 mmol.L^{-1}
k_1	Linear clearance of plasma insulin	0.05 min^{-1}
k_2	Transfer rate of insulin from interstitium to plasma	0.025 min^{-1}
k_3	Transfer rate of insulin from plasma to interstitium	0.025 min^{-1}
k_4	Clearance rate of insulin from interstitium	0.033 min^{-1}

Insulin sensitivity was modelled with a series of b-spline basis functions:

$$SI(t) = \sum_{i=1}^{M} SI_i \, \emptyset_{i,d}(t) \qquad 13.4$$

where $\emptyset_{i,d}$ are basis functions of degree *d* though time, and SI_i are the identified basis function coefficients. *M* is the number of basis functions, and determines the level of parameterisation of the model. A larger number of basis functions allows insulin sensitivity to vary more quickly and thus improves the fit of the model.

Basis functions of zeroth (*d* = 0), first (*d* = 1), and second (*d* = 2) order were used in this analysis. While the parameterisation is the same for a given *M* regardless of *d*, the added complexity of higher order basis functions allows *SI*(*t*) to vary more smoothly.

Since blood glucose sampling rates were not always consistent, the basis function knots were evenly distributed across the sample times. For example, when *M* = *N*, each basis function knot corresponds exactly to the time of each glucose measurement. When *M* = 0.5*N* the basis function knots corresponded with every second data point. This protocol removes the possibility of one or more basis functions existing in a large gap between samples and thus inducing non-identifiability in the coefficients of those basis functions. Figure 13.1 shows zeroth, first, and second order basis functions with a dataset containing a higher frequency of measurements in the later stages of the measurement window.

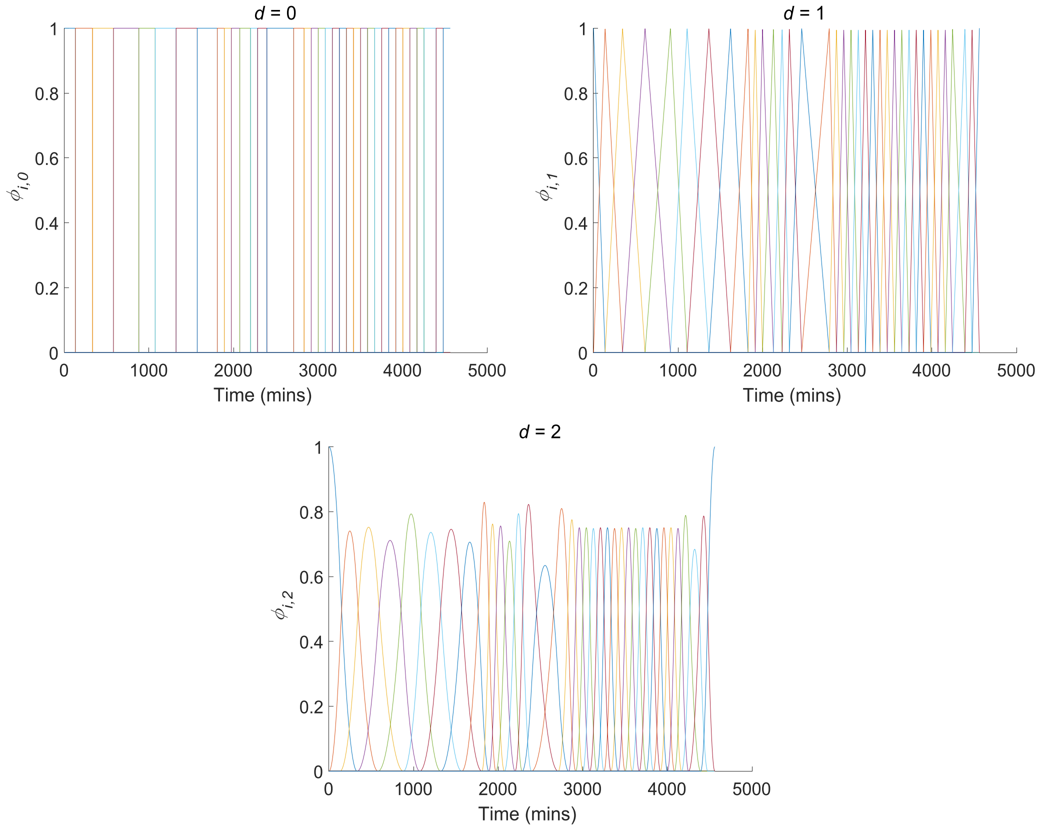


Figure 13.1. Basis function distribution for a particular patient episode with lower frequency sampling between t = 0 and 3000 minutes, and high frequency sampling between t = 3000 and 4500 minutes. M = 0.3N.

13.2.3 Mapping *SI* Variability

This analysis is based on the Stochastic Targeted (STAR) GC approach (Evans *et al.*, 2011, Evans *et al.*, 2012). *SI*(*t*) profiles were analysed for all data sets and used to build a stochastic map to describe the variation in *SI* for the cohort. Using the *SI* value at some time t_s, the stochastic map defines the distribution of *SI* at a particular time ($SI(t_s + \Delta t)$).

Stochastic maps for models with different combination of d = [0, 1, 2] and M = [0.3N, 0.65N, 0.85N, N], and Δt, were completed by identifying the full *SI*(*t*) profile for each of the 30 data sets. The Gauss-Newton parameter identification method was used with initial *SI* values of 10^{-5} L.mU^{-1}.min^{-1}, and a perturbation of 10^{-10} L.mU^{-1}.min^{-1}. Convergence of all parameters was reached within 50 iterations. The stochastic map was then built by plotting a one minute resolution $SI(t_s)$ vs. $SI(t_s + \Delta t)$ scatter plot, created using all *SI*(*t*) profiles from the 30 data sets.

13.2.4 Analysis

For each combination of d and M:N, the glucose model was fit to each of the 30 data sets, and RMS residuals were calculated. The t-test was used to test for differences in RMS residuals as the basis function order was increased, and as the parameterisation was increased. The identified *SI* from each model was used to build maps of *SI* variability. The aim was to analyse the effect of *d* and *M*:*N* on the model fit and on the corresponding *SI* maps.

13.3 Results

Figure 13.2 shows the comparative model fits for a section of one patient's data for each of the basis function options. Table 13.2 gives the mean RMS residual values for each of the *M*:*N* and *d* combinations. As *M* increased, the model fit improved. For a given *d*, the t-test found significant differences in RMS values for *M*:*N* = 0.3 compared with *M*:*N* = 0.65 ($p < 0.05$), and *M*:*N* = 0.65 compared with *M*:*N* = 1 ($p < 0.05$). When *M* = *N*, there is one basis function for every data point, so the model fit should be perfect. The non – zero residuals for *M*:*N* = 1 in Table 13.2 were the result of cases where the basis function knots did not precisely line up with the time of measurements, due to rounding errors. For a given level of parameterisation, the basis function order had less of an effect on the model fit. However, for *M*:*N* = 0.3, and *M*:*N* = 0.65, there were significantly higher residuals for *d* = 0 compared with *d* = 1 and *d* = 2 (t-test, $p < 0.05$).

Table 13.2. Mean RMS residual (cmH$_2$O) across the 30 data sets for each combination of d and M:N.

	d = 0	*d* = 1	*d* = 2
M:*N* = 0.3	0.73	0.64	0.64
M:*N* = 0.65	0.38	0.33	0.33
M:*N* = 1	0.022	0.00035	0.018

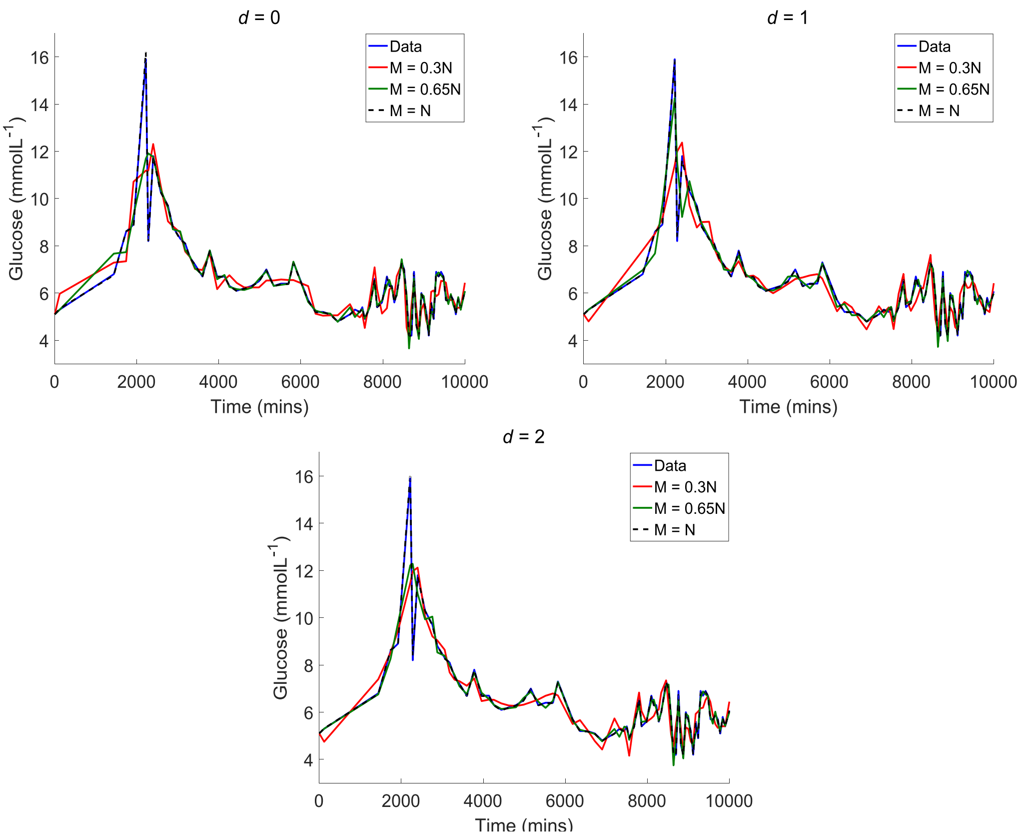


Figure 13.2. Glucose model fit for d = [0, 1, 2] and M = [0.3N, 0.6N, N].

Figure 13.3 to Figure 13.5 show the scatter plots for the combinations of *d* = [0, 1, 2], and *M* = [0.3N, 0.65N, N], when *Δt* was 0.5 hours, 1 hour, or 2 hours. Note the changes in the axis limits in the *M* = *N* plots, due to the much greater range of *SI*. As the number of basis functions increased, the scatter plots became wider. For a given *M*:*N* ratio, the scatter plots also widened as *Δt* increased. In terms of basis function order, there were noticeable decreases in scatter plot width between *d* = 0 and *d* = 1, and no substantial difference between *d* = 1 and *d* = 2.

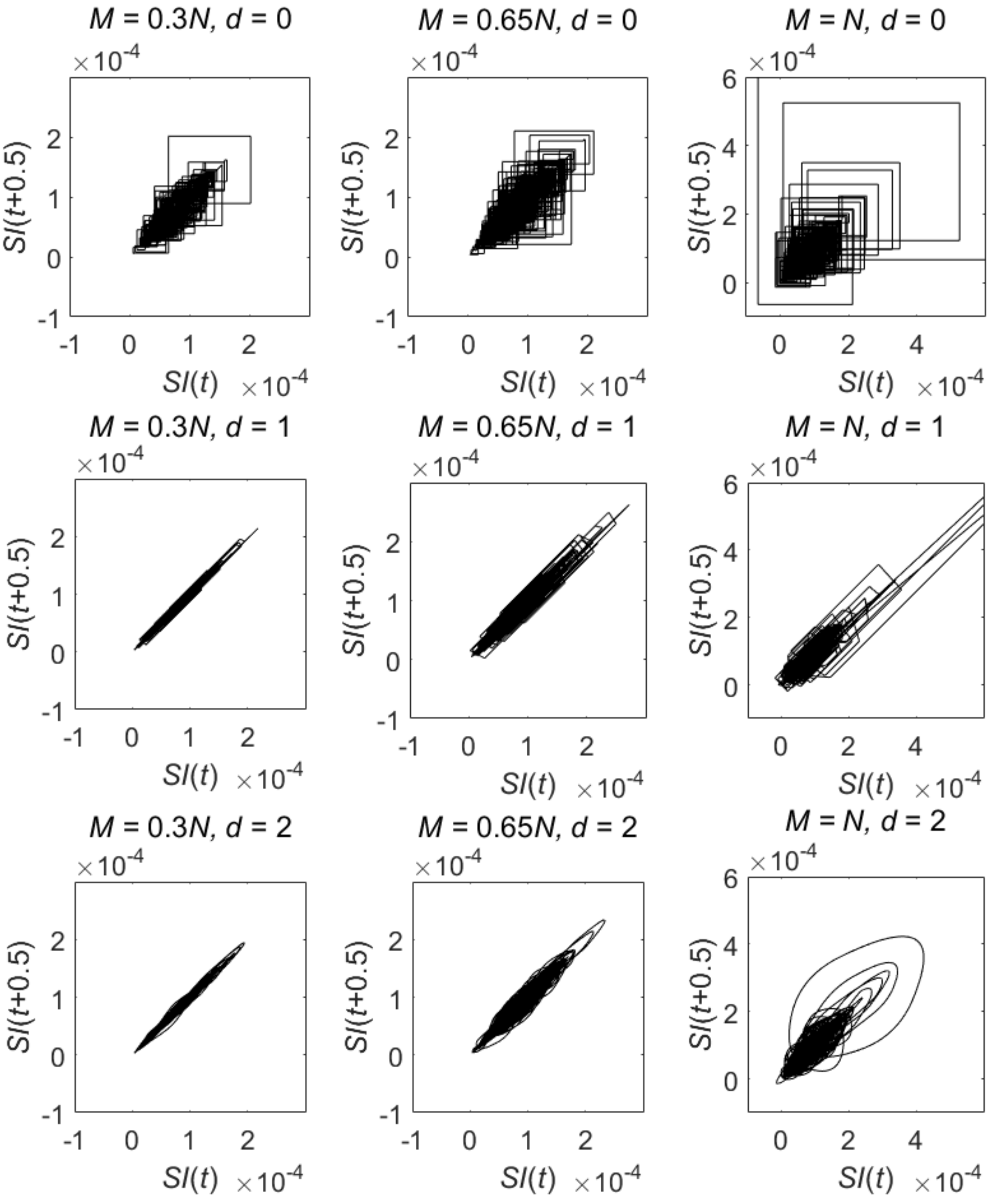


Figure 13.3. SI scatter plots for Δt = 0.5 hours.

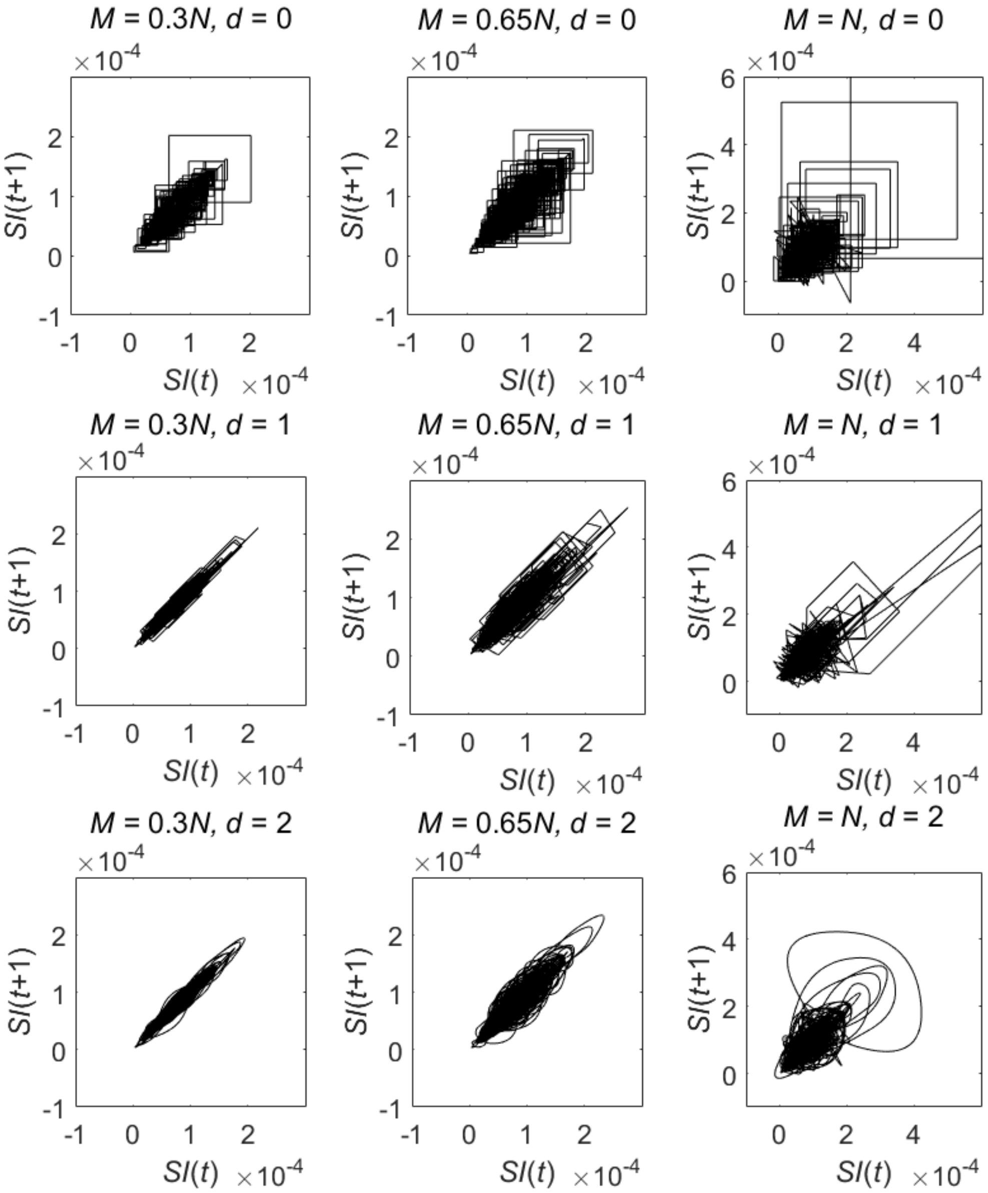


Figure 13.4. SI scatter plots for Δt = 1 hour.

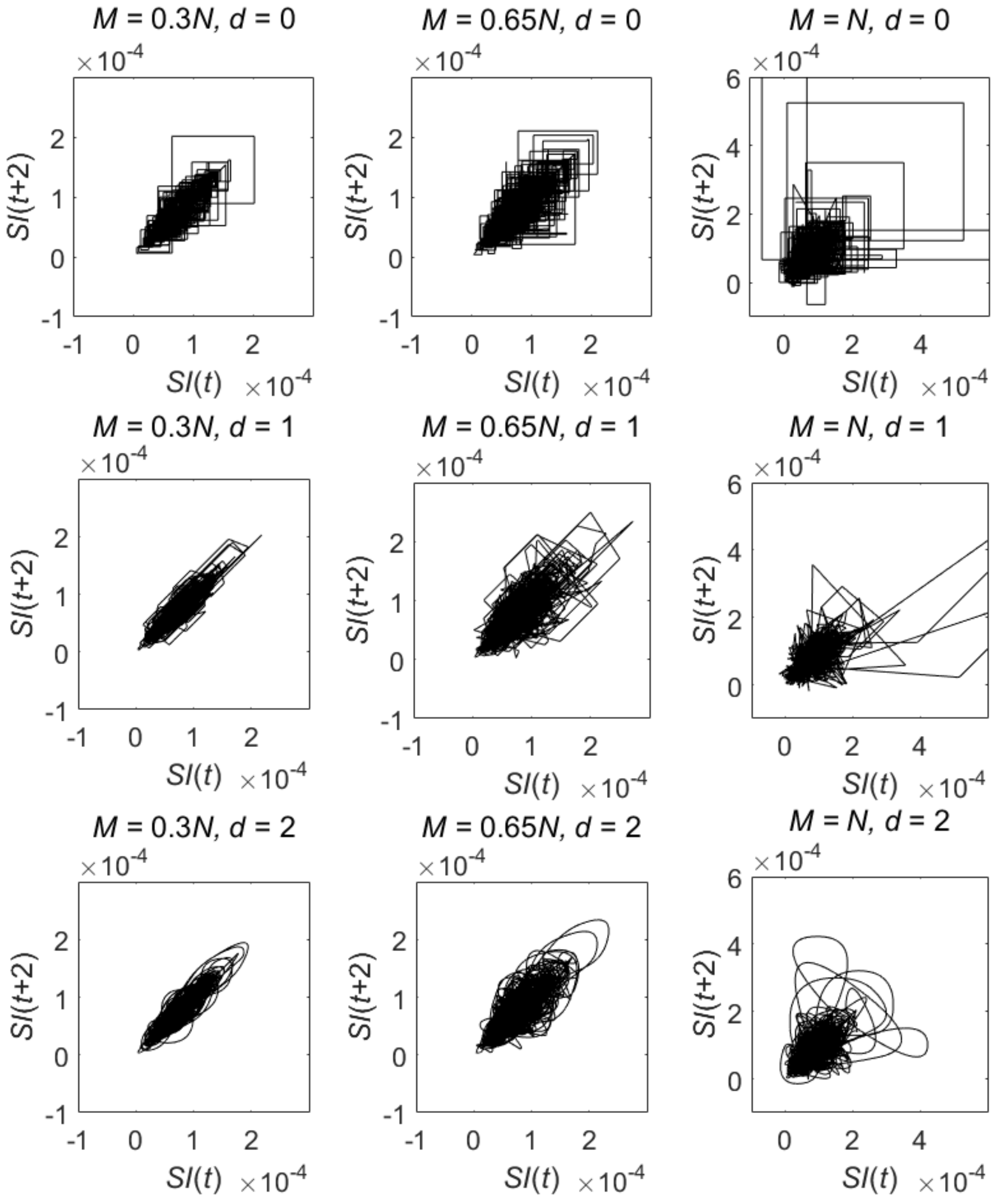


Figure 13.5. SI scatter plots for Δt = 2 hours.

13.4 Discussion

The consequence of changing the *M*:*N* ratio, in terms of model fit, is shown in Figure 13.2 and Table 13.2. When the *M*:*N* ratio is low, there are fewer parameters, and the glucose model – data residuals are comparatively high. When the *M*:*N* ratio is high, more parameters allow the dynamics in the measured glucose data to be captured more completely, and residuals are lower. The low residuals are not necessarily a positive outcome in this case, as it could mean the model is fitting to measurement noise, rather than patient dynamics. When *M*:*N* = 1, each data point is fit almost perfectly, because there is one basis function assigned to each data point. Since measurement noise always exists, the model is certainly fitting to noise in this scenario.

The scatter plots in Figure 13.3 to Figure 13.5 also show the consequence of increasing the M:N ratio, in terms of the variability in the identified SI. For a given time step, more parameters lead to a wider scatter plot, as a result of greater differences between $SI(t)$ and $SI(t + \Delta t)$. This occurred because gradients in the $SI(t)$ profile tended to be steeper at any given point, as $SI(t)$ varied more sharply to enable the model to be a closer fit to measured glucose.

As Δt increased from 0.5 to 2 hours, the scatter plots tended to widen for a given d and M:N ratio. This occurred because patient conditions are likely to be more different as the time between readings is increased. Significant changes in SI two hours in the future are much more likely to have occurred compared to half an hour in the future, as the patient is more likely to have eaten or undergone interventions.

The order of the basis functions did not have as great an effect on model fit as the number of basis functions. For a given M:N ratio, there were no major differences in the model fit between the $d = 1$ and $d = 2$ options, however the $d = 0$ residuals were significantly higher in the M:N = 0.3 and M:N = 0.65 cases (Table 13.2). Additionally, the SI variation was very similar for $d = 1$ and $d = 2$, but noticeably greater for $d = 0$ at every M:N level (Figure 13.3 to Figure 13.5). This is due to the step function nature of the zeroth order basis functions. The sharp jumps in SI when $d = 0$ meant that variability from one time point to the next was often greater, as compared to the smoother variation afforded by the $d = 1$, and especially the continuous $d = 2$ functions.

This analysis used data from critically ill patients from the Christchurch hospital, varying in age and diagnosis. Glycaemic regulation was generally achieved via enteral feed and bolus insulin injections in response to hourly blood glucose measurements. Hence, while this current analysis used data from a cohort that were representative of patients in most ICU settings, the administration method for insulin was distinct from some practices used in other ICU settings. However, the model based nature of this analysis does not rely on a specific insulin administration protocol. Thus, the findings of this analysis are likely to be applicable in glycaemic monitoring or control algorithms that utilise infusion or subcutaneous insulin administration.

As opposed to *in silico* simulated patients, the data used in this analysis came from real patients. The variability in these patients is representative of the type of critically ill patients who benefit from glucose control. Using real data meant that the time between measurements was not consistent. However the basis function implementation allowed the parameters to be evenly distributed across sample times.

While the model is complex, the computational burden is low, and a model fit is achievable in several seconds. This model could therefore be applicable in real time critical care environments. It thus has the potential to be integrated into a closed loop system for patient therapy optimisation. In such a system, blood glucose measurements would be taken and control decisions could be made based on current *SI* determined by the model and previously saved *SI* variability measurements.

All of the *M*:*N* ratios and basis function orders offered a reasonable model fit. Thus, the basis functions offer some malleability in the modelling approach. An analysis of *SI* prediction may provide an indication of the specific model formulation that could best suit clinical practice. In the following chapter, the *SI* variability maps are used to predict *SI* for different basis function *M*:*N* ratios.

13.5 Summary

A model of glucose and insulin kinetics used basis functions to identify insulin sensitivity through time. The analysis considered the outcomes of increasing basis function order and the resolution of basis functions. The results confirmed the expected outcome that increasing the parameterisation of the model resulted in lower model – data residuals, more variable *SI* profiles, and wider *SI* stochastic maps.

Chapter 14 Insulin Sensitivity Prediction

14.1 Introduction

Clinical data and models can be used retrospectively to determine the variability of *SI* in cohorts of critically ill patients (Lin *et al.*, 2008, Ferenci *et al.*, 2013, Dickson *et al.*, 2013, Thomas *et al.*, 2014). Such analyses can lead to stochastic models that can predict the patient *SI* or glucose likelihood distributions. It is critical that the prediction distribution is accurate. If the predicted likelihood distribution is too tight, the GC may be overly aggressive and risk unexpected hypoglycaemia. If the distribution is too wide, the GC may be too conservative. While this may be safe in terms of avoiding acute hypoglycaemia, patient outcomes may be worse than in a more tightly controlled system.

In this section, the appropriate level of basis function parameterisation of the glucose model (Equation 13.3) is determined. The *SI* variability for different levels of parameterisation determined in Chapter 13 is used to predict future *SI* distributions. The level of parameterisation is determined that allows the *SI* predictions to closely conform to the 'true' *SI* identified at future glucose measurements. The prediction analysis is compared with an AIC analysis, which estimates the level of parameterisation that provided the best trade-off between goodness-of-fit and model complexity.

14.2 Methods

For the AIC analysis, M was varied on the range $[0.1N, 0.9N]$, in steps of 1. The AICc correction was used (Equation 3.14), as M was relatively close to N. The expected measurement noise for the data is zero mean with a standard deviation of 0.35 mmol.L^{-1} (Arkray). For each data set, AICc was plotted against M:N ratio. The AIC theory implies that the model with the minimum AICc value should be the model with the best trade-off between fit and complexity, based on the estimated measurement noise.

For the prediction analysis, the accuracy and precision of the $SI(t_s + \Delta t)$ predictions was tested for each combination of $d = [0, 1, 2]$ and $M = [0.3N_s, 0.65N_s, 0.85N_s, N_s]$ on each of the 30 data sets. N_s is the number of samples in a particular subset of data. The method used the following steps:

1. A new $SI(t)$ profile was identified for a subset of the glucose data, containing $\text{N}_\text{s} = 0.5N$ data points ($G(t_{0\rightarrow s})$, $t_{0\rightarrow s} = t_0, t_1, \ldots t_{0.5N}$). The *SI* corresponding to the final point in the subset, $SI(t_s)$, was defined fully by a single basis function with no cross-over.

2. The appropriate value of Δt was determined. The value of Δt depends on how far into the future the prediction is made, so it depends on the time between t_s and the following measurement. The majority of data points were spaced by one hour. Thus, most predictions were made with Δt = 1 hour.
3. The relevant *SI* variability scatter plot for the particular Δt (created using the method in Chapter 13) was loaded. This scatter plot contains information from all 30 data sets.
4. A density surface of $SI(t_s)$ versus $SI(t_s + \Delta t)$ was created from the scatter plot using kernel density estimation. Gaussian kernels were defined over a 32 x 32 grid on the range of identified *SI*. Any slice of this density surface at a given $SI(t_s)$ provides a cumulative density function for $SI(t_s + \Delta t)$.
5. The modelled $SI(t_s)$ value at the final data point in the subset was used to retrieve the predicted CDF for $SI(t_s + \Delta t)$ from the density surface.
6. The true $SI(t_s + \Delta t)$ was identified using the $G(t_{0 \rightarrow s})$ data and $G(t_s + \Delta t)$ data point. The $SI(t)$ profile between t_s and $t_s + \Delta t$ used two basis functions with knots at t_s and $t_s + \Delta t$. To ensure that both $G(t_s)$ and $G(t_s + \Delta t)$ data points were fully represented by their respective $SI(t)$ values, these final basis functions both had a value of one at the knots.
7. The percentile on which the true $SI(t_s + \Delta t)$ fell on the prediction CDF was recorded.
8. Steps 1 – 7 were repeated as many times as possible for each data set, by shifting the subset in step 1 by one data point each time ($G(t_{0+j \rightarrow s+j})$), $j = [1, 2, \ldots 0.5N\text{-}1]$.
9. A CDF of all the percentiles from step 5 was plotted. This plot shows prediction incidence vs. actual incidence of *SI*.

Figure 14.1 shows an example of the CDFs resulting from the amalgamated $SI(t)$ profiles of the 30 data sets, and a particular CDF for $SI(t_s + 1)$ at $SI = 10^{-4}$ L.mU^{-1}.min^{-1}.

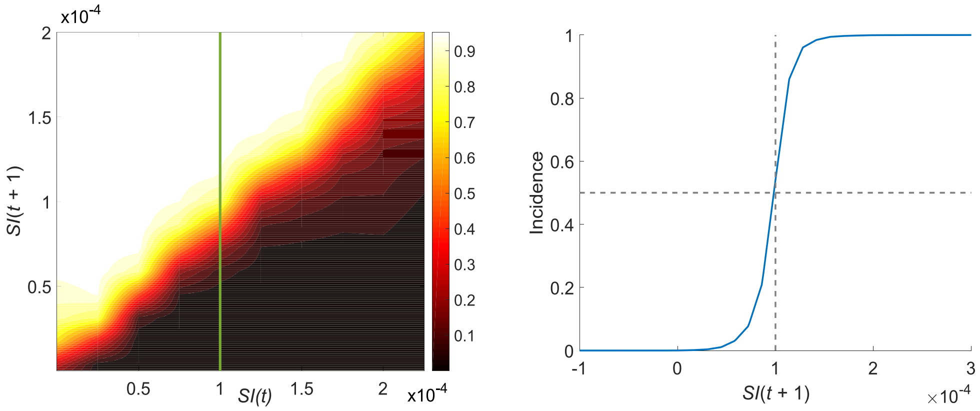


Figure 14.1. All prediction CDFs for d = 2, M = N, Δt = 1 hour (left), and the resulting prediction CDF of $SI(t_s + 1)$ when $SI(t_s) = 10^{-4}$ L.mU^{-1}.min^{-1} (right).

Figure 14.2 shows the range of possible outcomes of step 9. If the prediction is unbiased and appropriately precise, the true value of $SI(t_s + \Delta t)$ will be higher than the median of the SI prediction distributions exactly 50% of the time. Equally, the true value of $SI(t_s + \Delta t)$ will be below the i^{th} percentile of the prediction distributions exactly i% of the time. Ideal precision and zero bias will correspond to the 1:1 line on Figure 14.2.

Prediction distributions that are consistently too wide will result in prediction incidence vs. actual incidence existing in the green region in Figure 14.2. Overly tight prediction CDFs will result in the prediction incidence vs. actual incidence existing in the red region. Predictions biased high will result in a line completely below the 1:1 line, and predictions biased low will result in a line completely above the 1:1 line.

The ideal case is a line on the 1:1 line. However, clinically, the green area where prediction distributions are too wide is preferred. A clinical application favours the safety of a conservative prediction range and control, rather than the risk of a too tight prediction range.

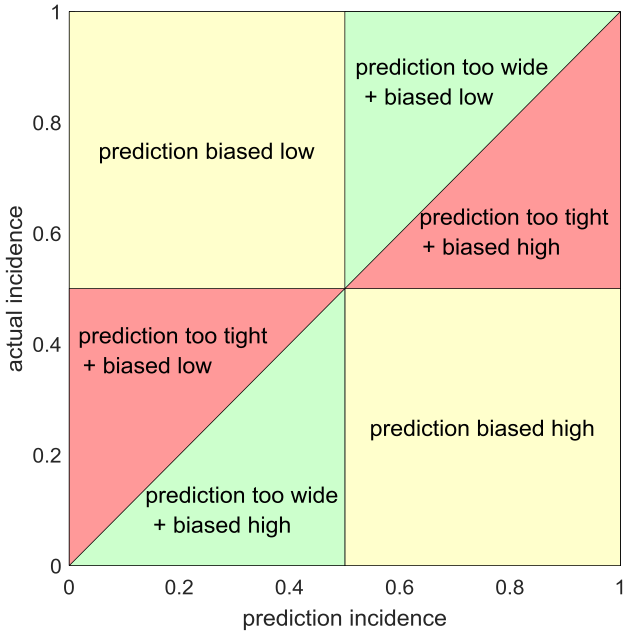


Figure 14.2. Regions of the prediction vs. outcome incidences showing the implications of prediction distributions that are consistently too wide, too tight, or biased.

14.3 Results

Figure 14.3 shows CDFs for the prediction distribution vs. the outcome distribution, created by plotting the percentile locations on the prediction CDFs at which the real $SI(t_s + \Delta t)$'s existed. Table 14.1 provides qualitative measures of the prediction outcomes. The absolute areas between the 1:1 line and the Figure 14.3 lines gives a measure of the closeness to the 1:1 line, and thus the relative accuracy of prediction distributions overall for different *M:N* ratios. The bias was calculated as the sum of the difference between the *x* and *y* axis, normalised to the number of data points. A positive value indicates that overall the predictions were biased high, while a negative value indicates a bias low. The maximum possible absolute area is 0.5, and the maximum bias is ± 0.5.

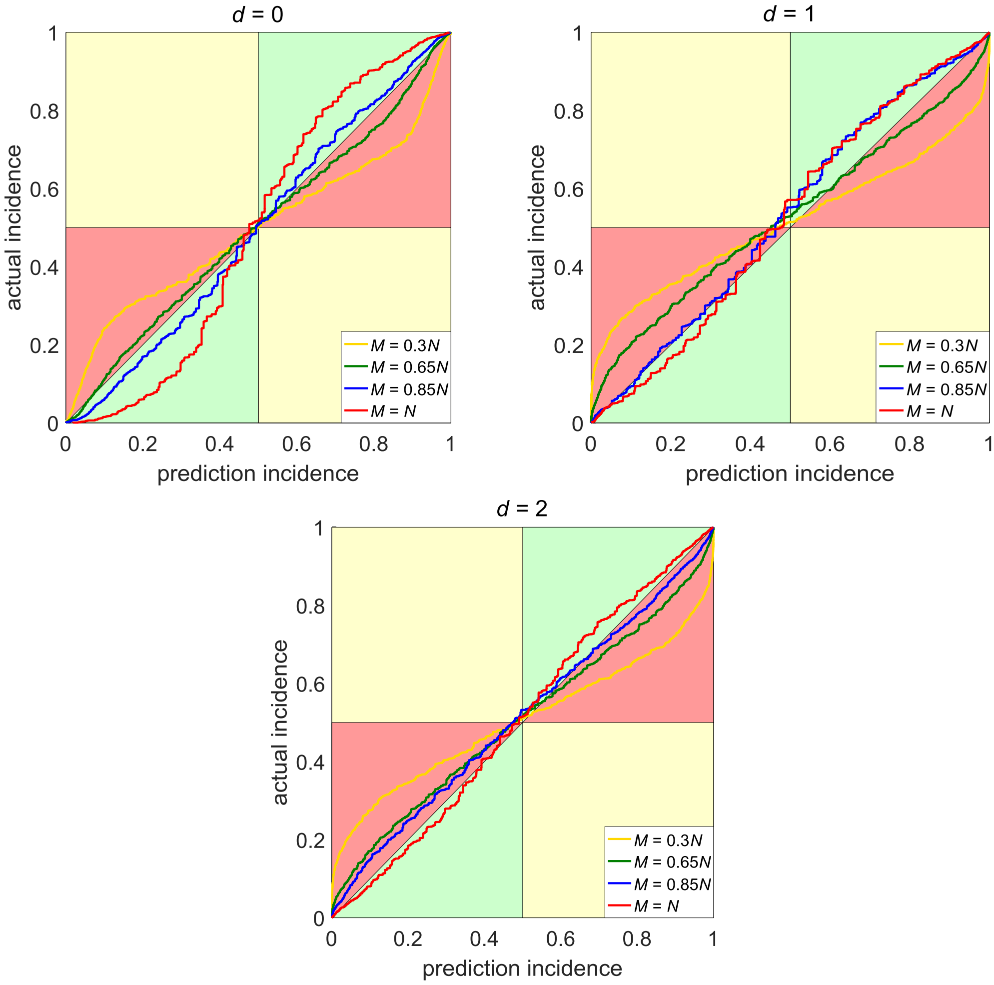


Figure 14.3. Prediction distribution vs. outcome distribution for d = 0 (top left), d = 1 (top right), and d = 2 (bottom).

Table 14.1. Precision and bias in the prediction distributions. T and W denote that the prediction distribution was too tight, and too wide, respectively.

	d = 0		*d* = 1		*d* = 2	
	Absolute Area	Bias	Absolute Area	Bias	Absolute Area	Bias
M = 0.3*N*	0.076 (T)	0.001	0.107 (T)	-0.011	0.108 (T)	-0.003
M = 0.65*N*	0.021 (T)	0.004	0.054 (T)	-0.022	0.045 (T)	-0.003
M = 0.85*N*	0.026 (W)	0.006	0.034 (W)	-0.023	0.027 (T)	-0.010
M = *N*	0.088 (W)	0.001	0.042 (W)	-0.016	0.023 (W)	-0.007

Figure 14.4 and Figure 14.5 show the results of the AIC analysis. AICc was normalised to the value at $M = 0.1N$. The median M:N ratio at which AICc was a minimum was [0.44, 0.43, 0.45] for d = [0, 1, 2].

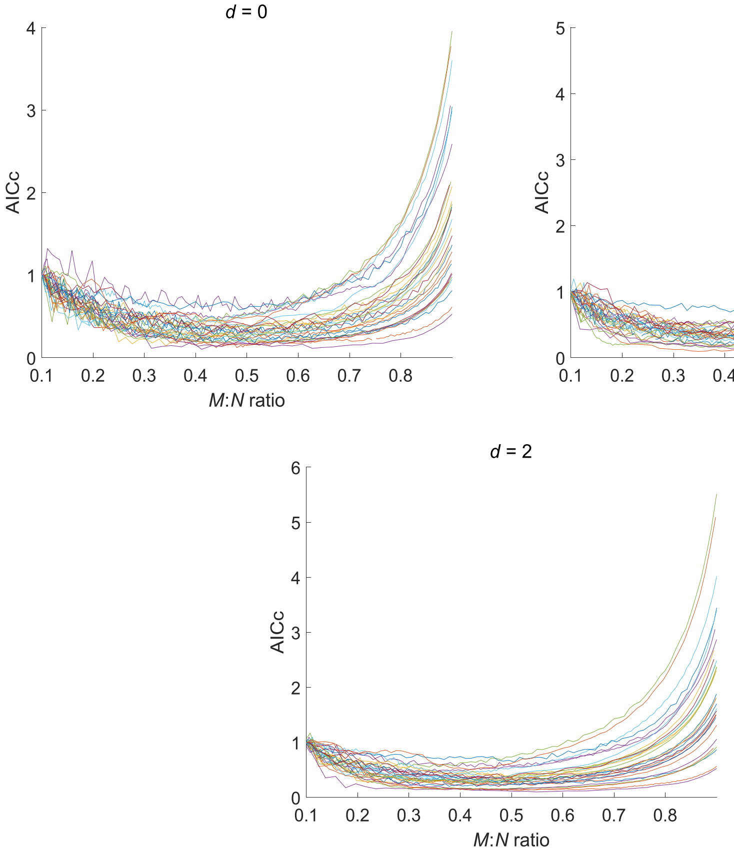


Figure 14.4. The change in AICc as the M:N ratio increased for the full data set of each patient episode, d = 0 (top left), d = 1 (top right), d = 2 (bottom).

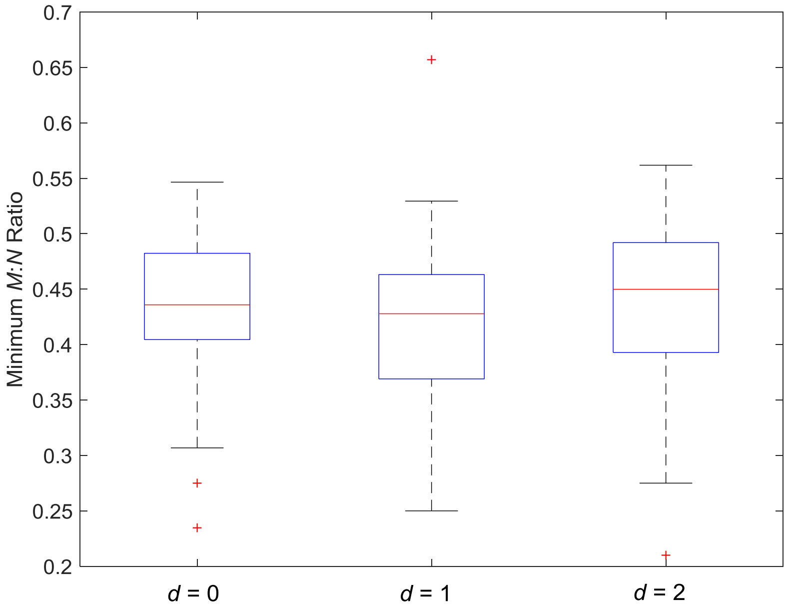


Figure 14.5. Boxplot of the minimum M:N ratio for all 30 data sets. The box limits are the 25th and 75th percentiles, and the whiskers show the range limited to data points that are within 1.5 IQR.

14.4 Discussion

Figure 14.3 shows that the parameterisation (M:N ratio) of the model had a much larger effect on prediction accuracy than the order of the basis functions. While AIC analysis found an average optimal M:N ratio of approximately of 0.45, the prediction analysis found that increasing the order of basis functions increased the optimal parameterisation. The ratios that yielded the smallest differences from perfect prediction distributions were M:N = [0.65, 0.85, 1] for d = [0, 1, 2], respectively. However, since predictions that are too tight are clinically more dangerous, the best ratios tested were M:N = [0.85, 0.85, 1] for d = [0, 1, 2].

Figure 14.3 indicates the accuracy and precision of the prediction CDFs compared to the distribution of the real $SI(t_s + \Delta t)$. If the prediction CDF is too tight, then the true $SI(t_s + \Delta t)$ value is likely to be distributed over a wider span of SI values outside the predicted distribution. The result of this narrow prediction profile is shown in Figure 14.3 for $M = 0.3N$, d = [0, 1, 2]. In contrast, if the prediction distribution is too wide, the true $SI(t_s + \Delta t)$ will usually lie near the middle. Such a result is most clearly shown in Figure 14.3 for $M = N$, $d = 0$.

While Table 14.1 shows that M:N = [0.85, 1] for d = [1, 2] resulted in the smallest difference from the 1:1 line, in these cases the predicted distributions were slightly wider than the true distributions of $SI(t_s + \Delta t)$. This result suggests that the model was fitting to measurement noise, rather than the true underlying patient glycaemic metabolic changes. Reducing the parameterisation of the model reduces the ability of the SI profile to adhere to measurement noise, and thus it may be hypothesized that a ratio between 0.85 and 1 might be optimal for *d* = 2, in terms of yielding a smaller absolute area in Table 14.1. Similarly, a ratio between 0.65 and 0.85 might be better for *d* = 1. Additionally, the best *M*:*N* ratio for *d* = 0 gave distributions that were slightly too tight, indicating that this model was not fully capturing glucose dynamics, and a ratio between 0.65 and 0.85 may be better for *d* = 0. Overall, the analysis has shown the expected trend towards wider prediction distributions as parameterisation increases. Additionally, the results showed that using higher order basis functions allowed a higher level of parameterisation before predictions became too wide.

Clinically, it is preferable for predictions to be too wide rather than too tight. Wide predictions will lead to more conservative and careful titration of treatment around the border of healthy glycaemia. This will ultimately lead to a greater proportion of time in the healthy range. In contrast, prediction that is too tight would lead to less conservative treatment when the patient glycaemia is close to dangerous levels, and thus, glycaemic excursions to unsafe regions are more likely. In the *d* = 0 case, the absolute areas for *M*:*N* = 0.85 and *M*:*N* = 0.65 were almost identical. However, the predictions for *M*:*N* = 0.85 were overall slightly too wide (in the green region), whereas the predictions for *M*:*N* = 0.65 were too tight (in the red region). Thus, while the *M*:*N* = 0.65 line was very slightly closer overall to 1:1, an *M*:*N* ratio of 0.85 could be preferable for *d* = 0 in clinical practice.

Table 14.1 indicates that first order basis functions resulted in comparatively high negative biases in the prediction distributions. The *d* = 0 predictions were biased high, and the *d* = 2 predictions were biased low, but the magnitude of these biases were relatively small. Thus, first order basis functions may be a worse option than zero or second order functions in obtaining accurate and precise predictions.

The AIC analysis showed that all datasets reached an AICc minimum between *M* = 0.1*N* and *M* = 0.9*N*. Figure 14.4 and Figure 14.5 showed that a number of model parameters corresponding to approximately 45% of the number of measurements gave the optimum trade-off between model complexity and model fit in all cases. However, the results of the *SI* prediction analysis clearly contradict the AICc results. For *d* = [0, 1, 2], an *M*:*N* ratio of 0.45 would result in prediction distributions that are too tight, meaning $SI(t_s + \Delta t)$ is more variable than predicted.

The reliability of the AICc criterion depends on the accuracy of the expected measurement noise, which cannot be exactly known in a clinical setting (Chase *et al.*, 2006). Additionally, the AIC analysis assumes independent and identically distributed (IID) random variables. While it is often assumed, it is not exactly the case for glucose data. Glucose data is not necessarily independent from point to point clinically, nor is it identically distributed, as error varies with blood glucose level (Chase *et al.*, 2006). However, the AIC has often been used in biomedical applications (Mueller *et al.*, 2011, Li *et al.,* Wilinska *et al.*, 2005) despite the difficulty in meeting the IID condition perfectly. Hence, the AIC analysis results should be considered to lead to an approximate optimal case rather than definitive optimal case. The numerical prediction analysis has a stronger basis for determining the definitive optimal case as it directly measured the ability of the models to forward predict measured data.

The analysis considered the prediction horizon of *SI* rather than glucose level, which is the key consideration of glycaemic control. However, the *SI* analysis was proximal to an analysis in glucose. The model strategy uses an identified $SI(t)$ profile to define glucose kinetics. No other parameters are identified. Hence, any shift in glucose must also be matched by a commensurate shift in *SI*. These changes are not necessarily equivalent in magnitude. However, equivalence of the changes in glucose and *SI* is not important since this analysis considered prediction percentiles alone. Importantly, using *SI* rather than glucose as the metric of interest allows a simpler computation of the protocol in section 14.2. In data not shown, generating Figure 14.3 using $G(t_s + \Delta t)$ yielded the same outcomes.

In glycaemic control, acute hypoglycaemia is generally considered more harmful than acute hyperglycaemia. Hence, most glycaemic regulation protocols attempt to maintain near normal glycaemic levels while strictly avoiding hypoglycaemia. In glycaemic control algorithms the accuracy and precision of the prediction distributions is critical. If the prediction distribution is too tight, the clinician may be too aggressive in treatment and maintain the patient at a low glycaemic level. Thus, they will be at risk of acute hypoglycaemia. In contrast, if the prediction distribution is too wide, the clinician is likely to be too conservative and maintain the patient glycaemia too high with the undue fear of inducing hypoglycaemia. By ensuring that the prediction is accurate and appropriately precise, the clinician can set the preferred glycaemic level with confidence that the incidence of hypoglycaemia will match their expectations. This analysis presented a new strategy for determining the precision and accuracy of these *SI* predictions. The strategy is robust to changes in the times between glucose samples, and could lead to alterations in the modelling of glycaemia in critical illness that improves predictive capability.

14.5 Summary

SI variability across 30 data sets was used to enable prediction distributions of *SI* over particular time steps. Increasing the parameterisation of the model resulted in wider prediction distributions, and increasing the basis function order resulted in tighter prediction distributions, such that accurate predictions were made with a higher M:N ratio. The ratios that gave the most accurate predictions were M:N = 0.65 when d = 0, M:N = 0.85 when d = 1, and M:N = 1 when d = 2. In contrast, the AIC analysis found that an M:N ratio of approximately 0.45 was optimal in all cases, though some assumptions required for a trustworthy AIC analysis result were breached.

The prediction analysis gives insight into the level of parameterisation needed for accurate future predictions of *SI* variability. Importantly, it was found that a maximum parameterisation of M:N = 1 caused the predictions to be wider than the true distribution of $SI(t_s + \Delta t)$. Accurate predictions would ultimately allow model-based GC algorithms to be implemented more effectively, increasing clinician confidence in treatment outcomes and improving safety for patients.

Part 5: Conclusions and Future Work

Chapter 15 Conclusions

This thesis primarily describes a novel pulmonary model that has potential to aid clinicians in mechanical ventilation of ARDS patients. Since patient physiology, ARDS dysfunctions, and ICU ventilation protocols are diverse, the NARX model was developed to be useful in a wide range of MV situations. Clinically relevant parameters describing patient-specific conditions were determined, and prediction of the patients' respiratory responses was implemented successfully, providing a method for potential avoidance of harmful over-distension. Basis functions are a key aspect of the NARX model, and have numerous possible applications in data fitting. In Part III of the thesis, basis functions were used to determine a suitable level of parameterisation of a blood glucose model, avoid over fitting, and determine accurate predictions of insulin sensitivity variance.

Chapters 5 to 12 concern novel pulmonary modelling using the NARX model. The main innovation of the NARX model was the pressure based basis functions used to describe elastance. Three cohorts of data were used to test the NARX model, each with a range of patient diagnoses, and representing different ventilation and recruitment manoeuvre protocols.

Chapter 5 describes the initial development of the NARX model and justifies the terms using an analysis of residuals. This work used complex patient data sets that have confounded previous modelling efforts. In particular, the cohort contained end-inspiratory pauses, and the RMs contained a large number of PEEP step increases, allowing opportunity for over-distension effects to be observed in some cases. The use of this complex data allowed a descriptive model to be developed, that could fit all features of the breath waveform across RMs. Despite a large number of model parameters, the measurement noise was significantly lower than the model – data residuals, implying that the NARX model was not over-parameterised.

Chapters 6 to 8 further validate the descriptive capabilities of the NARX model. The interpolation result in Chapter 6 demonstrates the NARX model ability to consistently capture recruitment and distension when supplied with less than a full RM of data, and its superiority to the FOM in this circumstance. Chapter 7 showed that the NARX model could identify independent inspiratory and expiratory elastances for a wide pressure range. This result is significant because expiratory elastance is non-identifiable for most models when expiration is a passive process, and flow and volume are not linearly independent. This limitation is overcome by the NARX model because it includes the flow dependent terms that incorporate information from the previous breath. Chapter 8 showed that the first flow dependent term in the NARX model captured the expected decrease in resistance across an RM. The FOM resistance term was unable to do the same, because the identified FOM resistance could be confounded by increases in elastance.

In Chapter 9, the NARX model was applied to a patient whose spontaneous breathing efforts caused M-waves in the pressure signal. A modified Gauss-Newton parameter ID method was used to effectively ignore the M-waves, so that the model exhibited the shape of a normal breath even when an SB effort occurred. The modified GN method achieved this by reducing the contribution of large residuals on the step in parameters at each GN iteration. Ignoring the random unknowable effect of SB on the pressure signal meant that respiratory mechanics for this patient could then be estimated under steady state ventilation.

Chapters 10 and 11 describe the most clinically significant result of the pulmonary modelling section of the thesis, the NARX model extrapolation and pressure prediction. In Chapter 10, a suitable extrapolation method for the NARX model was determined. This involved replacing the b-spline basis functions with four functions that are continuous across the entire pressure range considered, and can linearly combine to produce any of the common elastance shapes observed by the b-spline basis functions. Extrapolation to higher and lower PEEP from different amounts of data was significantly more successful than the FOM in all scenarios. The most relevant results were those that considered prediction one PEEP step up from the training data. The NARX model error in these predictions was small in a clinical sense.

Chapter 11 uses the same extrapolation method for step increases in PEEP to perform a statistical classification analysis and test the NARX model's ability to predict over-distension. High PP and PIP thresholds were used as measures of the occurrence of over-distension, depending on the availability of these measurements in Cohort1 and Cohort2. On the simpler Cohort2 data sets, the NARX model was significantly more successful than the FOM at predicting the high PIP, even under conditions designed to produce the best outcome from the FOM. Of particular significance was the extremely high sensitivity of the NARX model, meaning that instances of high PIP were almost never missed. On the more complex Cohort1 data, the NARX model parameterisation had to be reduced, by reducing the number of basis functions and flow dependent terms, to yield prediction outcomes that surpassed the performance of the FOM.

In the early stages of NARX model development, it was found that the model fit the RMs of high auto-PEEP patients poorly. In Chapter 12, a cohort of both high and low auto-PEEP patients was used, with a simplified version of the NARX model to examine whether the model could be useful for these patients. While the fit was still poor for high auto-PEEP patients, clinically significant differences between the parameters for the two groups were found. Specifically, the

COPD patients had a much higher identified resistance, and their elastance curves showed that higher pressures were required for recruitment. These outcomes made sense in terms of the known physiology of COPD patients.

Chapters 13 and 14 use basis functions for a different biomedical application. By using basis functions to identify insulin sensitivity, an investigation was undertaken into the appropriate level of parameterisation of a model of glucose and insulin kinetics. *SI* was predicted as a distribution, based on expected *SI* variability, and compared to the true *SI* outcomes. The ratio of basis functions to measured data points affected the width of the prediction distributions, and thus the accuracy of predictions. When determining an optimal level of parameterisation, consideration of the impact of prediction errors was crucial. Since acute hypoglycaemia is more harmful than acute hyperglycaemia, the importance of minimising prediction errors that would potentially lead a control system to induce hypoglycaemia in patients was noted.

Chapters 13 and 14 show the flexibility of the basis function concept, however the pulmonary NARX model is the main accomplishment of this thesis. The NARX model overcomes many of the disadvantages of previous models. It does not have the single constant resistance or elastance of the simpler models. Instead it has a time varying resistance and a pressure dependent elastance. It is able to fit to both inspiratory and expiratory data simultaneously, and it is quickly and easily identifiable using simple linear least squares parameter ID. It does not require patient intervention or measurements beyond those that are normally taken during mechanical ventilation. It has been able to fit to data containing both recruitment and distension across increasing PEEP steps. It has successfully predicted patient response to PEEP changes within a RM. It thus has high potential to be integrated into a PEEP optimisation system to avoid over-distension and improve patient outcomes.

Chapter 16 Future Work

The pulmonary NARX model has been successful in fitting and predicting behaviour in the cohorts tested. However, further testing to assess the clinical value of the model parameters, as well as its prediction capabilities, would be worthwhile in verifying the model's ability to improve patient outcomes. Similarly, the primary avenue for future work using the glucose model would be testing in a clinical setting. Further NARX model work may also include customisation for different types of patients, such as those with different disease states, or for spontaneous breathing, and further analysing the benefit of the flow dependent parameters.

16.1 Clinical Value of the NARX Model

Diagnosis and disease progression tracking are the basis of most model – based optimised control algorithms. The NARX model may be able to be used to discriminate between patients with specific disease characteristics. A possible method could involve using the identified elastance shape to indicate recruitability. For example, COPD patients were observed to have shallower elastance curves compared to non-COPD patients. Other features of elastance or flow terms may correlate with different patient attributes and diagnoses. The model parameters could also be used to monitor disease progression over time.

A randomised controlled trial could evaluate the NARX model's prediction capabilities. This would involve one patient cohort undergoing the standard PEEP selection process. In a second cohort, PEEP may be increased or decreased based on the model's prediction from recent pressure – flow data. Measurable outcomes such as mortality, oxygen saturation, and length of MV would be used to assess performance.

16.2 NARX Model Flow Dependent Parameters

The original NARX model contained 350 flow dependent parameters, or *b* coefficients. The *b* coefficients collectively capture information about the behaviour of the airways in response to air flow, including resistance, and the gas redistribution that happens during the end-inspiratory pause. Though we do not fully understand how these processes influence the NARX model outcomes, the grey box NARX modelling approach was able to successfully fit to data without needing to know what each parameter corresponds to physiologically.

There is value in the *b* coefficients because they enable the model to fit the data well, without overfitting, based on the estimation of measurement noise. However, the full clinical value they provide is currently unclear. Gaining a better understanding of what aspects these parameters are capturing would be useful in gaining clinician confidence in the model.

One possible method of extracting value from the *b* coefficients could be to look at the cardiogenic oscillations in the pressure during an end-inspiratory pause. The amplitude of these oscillations may indicate the amount of pulmonary blood flow (Tusman *et al.*, 2009). Using the model fit to obtain the amplitude may provide a useful measure to clinicians. Also, the oscillations may reflect something about respiratory system elastance, especially during spontaneous breathing (Schumann *et al.*, 2011).

16.3 Adaptation for High Auto-PEEP Patients

The two patients in Cohort1 and the four patients in Cohort3 who had unusually high auto-PEEP were fit poorly by both the original NARX model, and the simplified version discussed in Chapter 12. While the simplified NARX model resulted in useful parameter outputs for the high auto-PEEP patients, a better model fit could give a more accurate identification of patient-specific parameters.

A model adaptation could be implemented to enable it to better fit high auto-PEEP patients. The adaptation could use a discontinuous term that fits low PEEP data separately to high PEEP data. The discontinuity would occur at the point where PEEP exceeds auto-PEEP. The poor model fit occurred because modelled pressure was lower than measured pressure when PEEP < auto-PEEP. This implies the modelled elastance was too low, therefore the adaptation could aim to raise elastance for this period.

16.4 Treatment of Spontaneous Breathing Signals

For patients that are ventilated in partial rather than full sedation, SB efforts are common, and M-waves or similar anomalies frequently show up in the pressure and flow signals. These irregularities will reduce the usefulness of patient-specific values gained from parameter identification.

The modified GN method to ignore M-waves only works during steady state ventilation. It cannot deal with PEEP steps, or any other state change. A more robust method that allows the NARX model to handle SB during a recruitment manoeuvre would therefore be beneficial.

The treatment of the effects of SB on the pressure and flow signals could also lead to work in assessing the strength of patient SB efforts, as the strength of the effort effects the magnitude of the waveform anomaly. If embedded in ventilator software, this could help clinicians determine when to extubate patients, or provide a patient effort driven warning to adjust anaesthesia.

16.5 Trialling the Glucose Model

The glucose model work showed the effect of different levels of parameterisation on the accuracy of a predicted distribution of insulin sensitivity. However, only a few parameterisation options were tested. Further work may first aim to reach a conclusion as to the optimal model that is reasonably accurate and precise, and minimises the impact of prediction failures. Specifically, the model should place more emphasis on avoiding hypoglycaemia than hyperglycaemia. Future work may then integrate the model into existing glucose control algorithms, with the ultimate aim of testing in a randomised control trial.

References

Akoumianaki, E., Lyazidi, A., Rey, N., Matamis, D., Perez-Martinez, N., Giraud, R., Mancebo, J., Brochard, L. & Richard, J. C. M. 2013. Mechanical ventilation-induced reverse-triggered breaths: a frequently unrecognized form of neuromechanical coupling. *CHEST,* 143**,** 927-938.

Al-Mayah, A., Moseley, J. & Brock, K. K. 2008. Contact surface and material nonlinearity modeling of human lungs. *Physics in Medicine and Biology,* 53**,** 305-17.

Altman, D. G. & Bland, J. M. 1983. Measurement in medicine: the analysis of method comparison studies. *Journal of the Royal Statistical Society. Series D (The Statistician),* 32**,** 307-317.

Amato, M. B. P., Meade, M. O., Slutsky, A. S., Brochard, L., Costa, E. L. V., Schoenfeld, D. A., Stewart, T. E., Briel, M., Talmor, D., Mercat, A., Richard, J. C. M., Carvalho, C. R. R. & Brower, R. G. 2015. Driving pressure and survival in the acute respiratory distress syndrome. *New England Journal of Medicine,* 372**,** 747-755.

Anderson, D. R., Burnham, K. P. & White, G. C. 1994. AIC model selection in overdispersed capture-recapture data. *Ecology,* 75**,** 1780-1793.

The Acute Respiratory Distress Syndrome Network, Brower, R. G., Matthay M. A., Morris, A., Schoenfeld, D., Thompson B.T. & Wheeler, A. 2000. Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. *New England Journal of Medicine*, 342, 1301-1308.

Arkray 2001. "Glucocard TM Test Strip 2 Data Sheet," Supplied by Arkray, Inc. Japan.

Ashbaugh, D., Boyd Bigalow, D. B., Petty, T. & Levine, B. 1967. Acute respiratory distress in adults. *The Lancet,* 290**,** 319-323.

Bates, J. H., Ludwig, M. S., Sly, P. D., Brown, K., Martin, J. G. & Fredberg, J. J. 1988. Interrupter resistance elucidated by alveolar pressure measurement in open-chest normal dogs. *Journal of Applied Physiology,* 65**,** 408-414.

Bates, J. H. T. 2009. *Lung Mechanics: An Inverse Modeling Approach,* United States of America, New York, Cambridge University Press.

Bersten, A. D., Edibam, C., Hunt, T., Moran, J., Group, T. A. & New Zealand Intensive Care Society Clinical Trials Group. 2002. Incidence and mortality of acute lung injury and the acute respiratory distress syndrome in three australian states. *American Journal of Respiratory and Critical Care Medicine,* 165**,** 443-448.

Billings, S. A. 2013. *Nonlinear System Identification: NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains*, Wiley.

Blaha, J., Hovorka, R., Matias, M., Kotulak, T., Kremen, J., Sloukova, A., Svacina, S., Haluzik, M. 2005. Intensive insulin therapy in critically ill patients: comparison of standard and MPC protocols. *Intensive Care Medicine,* 31**,** S203.

Boles, J.M., Bion, J., Connors, A., Herridge, M., Marsh, B., Melot, C., Pearl, R., Silverman, H., Stanchina, M., Vieillard-Baron, A. & Welte, T. 2007. Weaning from mechanical ventilation. *European Respiratory Journal,* 29**,** 1033-1056.

Bosarge, P. L., Shoultz, T. H., Griffin, R. L. & Kerby, J. D. 2015. Stress-induced hyperglycemia is associated with higher mortality in severe traumatic brain injury. *Journal of Trauma and Acute Care Surgery,* 79**,** 289-294.

Botzheim, J., Cabrita, C., Kóczy, L. T. & Ruano, A. E. 2007. Genetic and bacterial programming for b-spline neural networks design. *Journal of Advanced Computational Intelligence and Intelligent Informatics,* 11**,** 220-231.

Boyle, A. & Adler, A. 2011. The impact of electrode area, contact impedance and boundary shape on EIT images. *Physiological Measurement,* 32**,** 745-54.

Bradley, C. P., Pullan, A. J. & Hunter, P. J. 1997. Geometric modeling of the human torso using cubic hermite elements. *Annals of Biomedical Engineering,* 25**,** 96-111.

Brock, K. K., Sharpe, M. B., Dawson, L. A., Kim, S. M. & Jaffray, D. A. 2005. Accuracy of finite element model-based multi-organ deformable image registration. *Medical Physics,* 32**,** 1647-59.

Brower, R. G., Lanken, P. N., Macintyre, N., Matthay, M. A., Morris, A., Ancukiewicz, M., Schoenfeld, D. & Thompson, B. T. 2004. Higher versus lower positive end-expiratory pressures in patients with the acute respiratory distress syndrome. *New England Journal of Medicine,* 351**,** 327-336.

Burnham, K. P. & Anderson, D. R. 2002. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach,* New York, Springer.

Burnham, K. P. & Anderson, D. R. 2004. Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Research,* 33**,** 261-304.

Burrowes, K. S., Hunter, P. J. & Tawhai, M. H. 2005. Anatomically based finite element models of the human pulmonary arterial and venous trees including supernumerary vessels. *Journal of Applied Physiology,* 99**,** 731-738.

Cadi, P., Guenoun, T., Journois, D., Chevallier, J. M., Diehl, J. L. & Safran, D. 2008. Pressure-controlled ventilation improves oxygenation during laparoscopic obesity surgery compared with volume-controlled ventilation. *British Journal of Anaesthesia,* 100**,** 709-716.

Campbell, R. S. & Davis, B. R. 2002. Pressure-controlled versus volume-controlled ventilation: does it matter? *Respiratory Care,* 47**,** 416-24.

Capes, S. E., Hunt, D., Malmberg, K. & Gerstein, H. C. 2000. Stress hyperglycaemia and increased risk of death after myocardial infarction in patients with and without diabetes: a systematic overview. *The Lancet,* 355**,** 773-778.

Carney, D. E., Bredenberg, C. E., Schiller, H. J., Picone, A. L., McCann, U. G., Gatto, L. A., Bailey, G., Fillinger, M. & Nieman, G. F. 1999. The mechanism of lung volume change during mechanical ventilation. *American Journal of Respiratory and Critical Care Medicine,* 160**,** 1697-1702.

Carvalho, A., Jandre, F., Pino, A., Bozza, F., Salluh, J., Rodrigues, R., Ascoli, F. & Giannella-Neto, A. 2007. Positive end-expiratory pressure at minimal respiratory elastance represents the best compromise between mechanical stress and lung aeration in oleic acid induced lung injury. *Critical Care,* 11**,** R86.

Chase, J. G., Hann, C. E., Jackson, M., Lin, J., Lotz, T., Wong, X.-W. & Shaw, G. M. 2006. Integral-based filtering of continuous glucose sensor measurements for glycaemic control in critical care. *Computer Methods and Programs in Biomedicine,* 82**,** 238-247.

Chase, J. G., Shaw, G., Le Compte, A., Lonergan, T., Willacy, M., Wong, X. W., Lin, J., Lotz, T., Lee, D. & Hann, C. 2008. Implementation and evaluation of the SPRINT protocol for tight glycaemic control in critically ill patients: a clinical practice change. *Critical Care,* 12**,** R49.

Chase, J. G., Shaw, G. M., Lotz, T., Le Compte, A., Wong, J., Lin, J., Lonergan, T., Willacy, M. & Hann, C. E. 2007. Model-based insulin and nutrition administration for tight glycaemic control in critical care. *Current Drug Delivery,* 4**,** 283-296.

Chase, J. G., Suhaimi, F., Penning, S., Preiser, J. C., Le Compte, A. J., Lin, J., Pretty, C. G., Shaw, G. M., Moorhead, K. T. & Desaive, T. 2010. Validation of a model-based virtual trials method for tight glycemic control in intensive care. *BioMedical Engineering Online,* 9**,** 84-84.

Chiew, Y. S., Chase, J. G., Shaw, G., Sundaresan, A. & Desaive, T. 2011. Model-based PEEP optimisation in mechanical ventilation. *BioMedical Engineering OnLine,* 10**,** 111.

Christiansen, C., Toft, P., Jørgensen, H. S., Andersen, S. K. & Tønnesen, E. 2004. Hyperglycaemia and mortality in critically ill patients: A prospective study. *Intensive Care Medicine,* 30**,** 1685-1688.

Coelho, L. D. S. & Krohling, R. A. 2006. Nonlinear system identification based on b-spline neural network and modified particle swarm optimization. *The 2006 IEEE International Joint Conference on Neural Network Proceedings*, 3748-3753.

Damanhuri, N. S., Docherty, P. D., Chiew, Y. S., Van Drunen, E. J., Desaive, T. & Chase, J. G. 2014. A patient-specific airway branching model for mechanically ventilated patients. *Computational and Mathematical Methods in Medicine,* 645732.

Daniels, C. B. & Orgeig, S. 2003. Pulmonary surfactant: the key to the evolution of air breathing. *News in Physiological Sciences,* 18**,** 151-7.

Dasta, J. F., Mclaughlin, T. P., Mody, S. H. & Piech, C. T. 2005. Daily cost of an intensive care unit day: the contribution of mechanical ventilation. *Critical Care Medicine,* 33**,** 1266-1271.

Davidson, S. M., Redmond, D. P., Laing, H., White, R., Radzi, F., Chiew, Y. S., Poole, S. F., Damanhuri, N. S., Desaive, T., Shaw, G. M. & Chase, J. G. 2014. Clinical utilisation of respiratory elastance (CURE): pilot trials for the optimisation of mechanical ventilation settings for the critically ill. *IFAC Proceedings Volumes,* 47**,** 8403-8408.

De Boor, C. 1972. On calculating with b-splines. *Journal of Approximation Theory,* 6**,** 50-62.

Dickson, J. L., Floyd, R. P., Le Compte, A. J., Fisk, L. M., Chase, J. G., Lynn, A. & Shaw, G. M. 2013. External validation and sub-cohort analysis of stochastic forecasting models in NICU cohorts. *Biomedical Signal Processing and Control,* 8**,** 409-419.

Docherty, P. D., Gray, R. A. L. & Mansell, E. J. 2014. Reducing the effect of outlying data on the identification of insulinaemic pharmacokinetic parameters with an adapted Gauss-Newton approach. *IFAC Proceedings Volumes,* 47**,** 5635-5640.

Dodge, Y. 2008. *The Concise Encyclopedia of Statistics*, Springer Science & Business Media.

Dreyfuss, D. & Saumon, G. 1998. Ventilator-induced lung injury: lessons from experimental studies. *American Journal of Respiratory and Critical Care Medicine,* 157**,** 294-323.

Egi, M., Bellomo, R., Stachowski, E., French, C. J. & Hart, G. 2006. Variability of blood glucose concentration and short-term mortality in critically ill patients. *Anesthesiology,* 105**,** 244-252.

Eom, J., Xu, X. G., De, S. & Shi, C. 2010. Predictive modeling of lung motion over the entire respiratory cycle using measured pressure-volume data, 4DCT images, and finite-element analysis. *Medical Physics,* 37**,** 4389-4400.

Evans, A., Le Compte, A., Tan, C. S., Ward, L., Steel, J., Pretty, C. G., Penning, S., Suhaimi, F., Shaw, G. M., Desaive, T. & Chase, J. G. 2012. Stochastic targeted (STAR) glycemic control: design, safety, and performance. *Journal of Diabetes Science and Technology,* 6**,** 102-115.

Evans, A., Shaw, G., Le Compte, A., Tan, C. S., Ward, L., Steel, J., Pretty, C., Pfeifer, L., Penning, S., Suhaimi, F., Signal, M., Desaive, T. & Chase, J. G. 2011. Pilot proof of concept clinical trials of stochastic targeted (STAR) glycemic control. *Annals of Intensive Care,* 1**,** 38.

Ferenci, T., Benyó, B., Kovács, L., Fisk, L., Shaw, G. M. & Chase, J. G. 2013. Daily evolution of insulin sensitivity variability with respect to diagnosis in the critically ill. *PLoS ONE,* 8**,** e57119.

Ferring, M. & Vincent, J. L. 1997. Is outcome from ARDS related to the severity of respiratory failure? *European Respiratory Journal,* 10**,** 1297-1300.

G.J. Tortora, B. D. 2011. *Principles of Anatomy and Physiology*, John Wiley and Sons Inc.

Gattinoni, L., Carlesso, E., Cadringher, P., Valenza, F., Vagginelli, F. & Chiumello, D. 2003. Physical and biological triggers of ventilator-induced lung injury and its prevention. *European Respiratory Journal,* 22**,** 15s-25.

Giavarina, D. 2015. Understanding Bland Altman analysis. *Biochemia Medica,* 25**,** 141-151.

Gong, B., Krueger-Ziolek, S., Moeller, K., Schullcke, B. & Zhao, Z. 2015. Electrical impedance tomography: functional lung imaging on its way to clinical practice? *Expert Review of Respiratory Medicine,* 9**,** 721-737.

Gray, R. A. L., Docherty, P. D., Fisk, L. M. & Murray, R. 2016. A modified approach to objective surface generation within the Gauss-Newton parameter identification to ignore outlier data points. *Biomedical Signal Processing and Control,* 30**,** 162-169.

Groeger, J. S., Levinson, M. R. & Carlon, G. C. 1989. Assist control versus synchronized intermittent mandatory ventilation during acute respiratory failure. *Critical Care Medicine,* 17**,** 607-12.

Grünwald, P. D. 2007. *The Minimum Description Length Principle*, MIT Press.

Hager, D. N., Krishnan, J. A., Hayden, D. L. & Brower, R. G. 2005. Tidal volume reduction in patients with acute lung injury when plateau pressures are not high. *American Journal of Respiratory and Critical Care Medicine,* 172**,** 1241-1245.

Hanley, J. A. & Mcneil, B. J. 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology,* 143**,** 29-36.

Harris, R. S. 2005. Pressure-volume curves of the respiratory system. *Respiratory Care,* 50**,** 78-98.

Hickling, K. G. 1998. The pressure-volume curve is greatly modified by recruitment: a mathematical model of ARDS lungs. *American Journal of Respiratory and Critical Care Medicine,* 158**,** 194-202.

Hurvich, C. M. & Tsai, C. L. 1989. Regression and time series model selection in small samples. *Biometrika,* 76**,** 297-307.

Jardin, F. & Vieillard-Baron, A. 2007. Is there a safe plateau pressure in ARDS? The right heart only knows. *Intensive Care Medicine,* 33**,** 444-447.

Kacmarek, R. M. 2011. The mechanical ventilator: past, present, and future. *Respiratory Care,* 56**,** 1170-80.

Kerby, J. D., Griffin, R. L., Maclennan, P. & Rue, L. W. 2012. Stress-induced hyperglycemia, not diabetic hyperglycemia, is associated with higher mortality in trauma. *Annals of Surgery,* 256**,** 446-452.

Kim, M., Bordas, R., Vos, W., Hartley, R. A., Brightling, C. E., Kay, D., Grau, V. & Burrowes, K. S. 2015. Dynamic flow characteristics in normal and asthmatic lungs. *International Journal for Numerical Methods in Biomedical Engineering,* 31.

Knörzer, A., Docherty, P. D., Chiew, Y. S., Chase, J. G. & Möller, K. 2014. An extension to the first order model of pulmonary mechanics to capture a pressure dependent elastance in the human lung. *IFAC Proceedings Volumes,* 47**,** 1176-1181.

Krinsley, J. S. 2004. Effect of an intensive glucose management protocol on the mortality of critically ill adult patients. *Mayo Clinic Proceedings,* 79**,** 992-1000.

Krinsley, J. S. 2008. Glycemic variability: A strong independent predictor of mortality in critically ill patients. *Critical Care Medicine,* 36**,** 3008-3013.

Lalkhen, A. G. & Mccluskey, A. 2008. Clinical tests: sensitivity and specificity. *Continuing Education in Anaesthesia Critical Care & Pain,* 8**,** 221-223.

Langdon, R., Docherty, P. D. & Chase, J. 2017a. Basis function modelling of respiratory patients with high or low auto-PEEP. *IFAC-PapersOnLine, accepted awaiting publication*.

Langdon, R., Docherty, P. D., Chiew, Y. S. & Chase, J. G. 2016a. Extrapolation of a non-linear autoregressive model of pulmonary mechanics. *Mathematical Biosciences*, 284, 32-39.

Langdon, R., Docherty, P. D., Chiew, Y. S., Damanhuri, N. S. & Chase, J. G. 2015a. Implementation of a non-linear autoregressive model with modified Gauss-Newton parameter identification to determine pulmonary mechanics of respiratory patients that are intermittently resisting ventilator flow patterns. *IFAC-PapersOnLine,* 48**,** 354-359.

Langdon, R., Docherty, P. D., Chiew, Y. S., Möller, K. & Chase, J. G. 2015b. Interpolation within a recruitment manoeuvre using a non-linear autoregressive model of pulmonary mechanics. *IFAC-PapersOnLine,* 48**,** 297-302.

Langdon, R., Docherty, P. D., Chiew, Y. S., Möller, K. & Chase, J. G. 2016b. Use of basis functions within a non-linear autoregressive model of pulmonary mechanics. *Biomedical Signal Processing and Control,* 27**,** 44-50.

Langdon, R., Docherty, P. D., Laufer, B. & Möller, K. 2016c. Resistance in a non-linear autoregressive model of pulmonary mechanics. *Current Directions in Biomedical Engineering,* 2**,** 623-627.

Langdon, R., Docherty, P. D., Mansell, E. J. & Chase, J. G. 2017b. Accurate and precise prediction of insulin sensitivity variance in critically ill patients. *Biomedical Signal Processing and Control,* 39, 327-335.

Langdon, R., Docherty, P. D. & Möller, K. 2016d. Inspiratory and expiratory elastance in a non-linear autoregressive model of pulmonary mechanics. *Current Directions in Biomedical Engineering,* 2**,** 629-632.

Langdon, R., Docherty, P. D., Schranz, C. & Chase, J. 2017c. Plateau pressure prediction in ARDS patients. *IFAC-PapersOnLine, accepted awaiting publication*.

Langdon, R., Docherty, P. D., Schranz, C. & Chase, J. G. 2017d. Prediction of high airway pressure using a non-linear autoregressive model of pulmonary mechanics. *BioMedical Engineering Online, in review*.

Laufer, B., Docherty, P. D., Knörzer, A., Chiew, Y. S., Langdon, R., Möller, K. & Chase, J. G. 2017. Performance of variations of the dynamic elastance model in lung mechanics. *Control Engineering Practice,* 58**,** 262-267.

Li, X., Chen, Y., Zhao, Z., Lu, W. & Zhou, T. Pharmacokinetic/pharmacodynamic analysis of metformin using different models in diabetic rats. 2016. *Drug Research (Stuttg)*, 66, 547-554.

Light, R. B. 1999. Pulmonary pathophysiology of pneumococcal pneumonia. *Seminars in Respiratory Infections,* 14**,** 218-26.

Lin, J., Lee, D., Chase, J. G., Shaw, G. M., Le Compte, A., Lotz, T., Wong, J., Lonergan, T. & Hann, C. E. 2008. Stochastic modelling of insulin sensitivity and adaptive glycemic control for critical care. *Computer Methods and Programs in Biomedicine,* 89**,** 141-152.

Lind, I. & Ljung, L. 2008. Regressor and structure selection in NARX models using a structured ANOVA approach. *Automatica,* 44**,** 383-395.

Lionetti, V., Recchia, F. A. & Ranieri, V. M. 2005. Overview of ventilator-induced lung injury mechanisms. *Current Opinion in Critical Care,* 11**,** 82-86.

Lonergan, T., Le Compte, A., Willacy, M., Chase, J. G., Shaw, G. M., Hann, C. E., Lotz, T., Lin, J. & Wong, X. W. 2006a. A pilot study of the SPRINT protocol for tight glycemic control in critically ill patients. *Diabetes Technology and Therapeutics,* 8**,** 449-462.

Lonergan, T., Le Compte, A., Willacy, M., Chase, J. G., Shaw, G. M., Wong, X. W., Lotz, T., Lin, J. & Hann, C. E. 2006b. A simple insulin-nutrition protocol for tight glycemic control in critical illness: development and protocol comparison. *Diabetes Technology and Therapeutics,* 8**,** 191-206.

Lotz, T. F., Chase, J. G., Mcauley, K. A., Shaw, G. M., Docherty, P. D., Berkeley, J. E., Williams, S. M., Hann, C. E. & Mann, J. I. 2010. Design and clinical pilot testing of the model-based dynamic insulin sensitivity and secretion test (DISST). *Journal of Diabetes Science and Technology,* 4**,** 1408-1423.

Luepschen, H., Riesen, D. V., Beckmann, L., Hameyer, K. & Leonhardt, S. 2008. Modeling of fluid shifts in the human thorax for electrical impedance tomography. *IEEE Transactions on Magnetics,* 44**,** 1450-1453.

Lyons, R. 2001. *Understanding Digital Signal Processing*, Prentice Hall.

Major, V., Simon, C., Redmond, D., Beatson, A., Glassenbury, D., Chiew, Y. S., Pretty, C., Desaive, T., Szlavecz, A., Benyo, B., Shaw, G. M. & Chase, J. G. 2015. Assessing respiratory mechanics of reverse-triggered breathing cycles - case study of two mechanically ventilated patients. *IFAC-PapersOnLine,* 48**,** 505-510.

Marik, P. E. & Raghavan, M. 2004. Stress-hyperglycemia, insulin and immunomodulation in sepsis. *Intensive Care Medicine,* 30**,** 748-756.

Massey, F. J. 1951. The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association,* 46**,** 68-78.

McCowen, K. C., Malhotra, A. & Bistrian, B. R. 2001. Stress-induced hyperglycemia. *Critical Care Clinics,* 17**,** 107-124.

Mead, J. 1969. Contribution of compliance of airways to frequency-dependent behavior of lungs. *Journal of Applied Physiology,* 26**,** 670-3.

Miller, R. R., Macintyre, N. R., Hite, R. D., Truwit, J. D., Brower, R. G. & Morris, A. H. 2012. Point: should positive end-expiratory pressure in patients with ARDS be set on oxygenation? Yes. *Chest,* 141**,** 1379-1382.

Mizock, B. A. 2001. Alterations in fuel metabolism in critical illness: hyperglycaemia. *Best Practice & Research Clinical Endocrinology & Metabolism,* 15**,** 533-551.

Mols, G., Kessler, V., Benzing, A., Lichtwarck-Aschoff, M., Geiger, K. & Guttmann, J. 2001. Is pulmonary resistance constant, within the range of tidal volume ventilation, in patients with ARDS? *British Journal of Anaesthesia,* 86**,** 176-182.

Montesantos, S., Katz, I., Pichelin, M. & Caillibotte, G. 2016. The creation and statistical evaluation of a deterministic model of the human bronchial tree from HRCT images. *PLoS One,* 11.

Morgan, E. G., Mikhail, M. & Murry, M. 2006. *Clinical Anesthesiology,* Philadelphia, McGraw Hill.

Morris, A. H., Orme, J., Truwit, J. D., Steingrub, J., Grissom, C., Lee, K. H., Li, G. L., Thompson, B. T., Brower, R., Tidswell, M., Bernard, G. R., Sorenson, D., Sward, K., Zheng, H., Schoenfeld, D. & Warner, H. 2008. A replicable method for blood glucose control in critically Ill patients. *Critical Care Medicine,* 36**,** 1787-1795.

Mount, L. E. 1955. The ventilation flow-resistance and compliance of rat lungs. *The Journal of Physiology,* 127**,** 157-167.

Mueller, M., Talary, M. S., Falco, L., De Feo, O., Stahel, W. A. & Caduff, A. 2011. Data processing for noninvasive continuous glucose monitoring with a multisensor device. *Journal of Diabetes Science and Technology,* 5**,** 694-702.

Otis, A. B., Mckerrow, C. B., Bartlett, R. A., Mead, J., Mcilroy, M. B., Selverstone, N. J. & Radford, E. P. 1956. Mechanical factors in distribution of pulmonary ventilation. *Journal of Applied Physiology,* 8**,** 427-443.

Papandrinopoulou, D., Tzouda, V. & Tsoukalas, G. 2012. Lung compliance and chronic obstructive pulmonary disease. *Pulmonary Medicine,* 2012**,** 6.

Pelosi, P., Gama De Abreu, M. & Rocco, P. R. 2010. New and conventional strategies for lung recruitment in acute respiratory distress syndrome. *Critical Care,* 14**,** 210.

Perzl, M. A., Schulz, H., Paretzke, H. G., Englmeier, K. H. & Heyder, J. 1996. Reconstruction of the lung geometry for the simulation of aerosol transport. *Journal of Aerosol Medicine: Deposition, Clearance, and Effects in the Lung,* 9**,** 409-418.

Petersen, G. W. & Baier, H. 1983. Incidence of pulmonary barotrauma in a medical ICU. *Critical Care Medicine,* 11**,** 67-9.

Phua, J., Badia, J. R., Adhikari, N. K. J., Friedrich, J. O., Fowler, R. A., Singh, J. M., Scales, D. C., Stather, D. R., Li, A., Jones, A., Gattas, D. J., Hallett, D., Tomlinson, G., Stewart, T. E. & Ferguson, N. D. 2009. Has mortality from acute respiratory distress syndrome decreased over time?: a systematic review. *American Journal of Respiratory and Critical Care Medicine,* 179**,** 220-227.

Polonsky, K. S., Given, B. D., Pugh, W., Licinio-Paixao, J., Thompson, J. E., Karrison, T. & Rubenstein, A. H. 1986. Calculation of the systemic delivery rate of insulin in normal man. *Journal of Clinical Endocrinology and Metabolism,* 63**,** 113-118.

Ranieri, V. M., Suter, P. M., Tortorella, C., De Tullio, R., Dayer, J. M., Brienza, A., Bruno, F. & Slutsky, A. S. 1999. Effect of mechanical ventilation on inflammatory mediators in patients with acute respiratory distress syndrome: a randomized controlled trial. *JAMA: The Journal of the American Medical Association,* 282**,** 54 - 61.

Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmuller, U. & Timmer, J. 2009. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics,* 25**,** 1923-9.

Reddy, V. G. 2005. Auto-PEEP: how to detect and how to prevent - a review. *Middle East Journal of Anaesthesiology,* 18**,** 293-312.

Rissanen, J. 1978. Modeling by shortest data description. *Automatica,* 14**,** 465-471.

Rock, K. L. & Kono, H. 2008. The inflammatory response to cell death. *Annual Review of Pathology,* 3**,** 99-126.

Rossi, A., Brandolese, R., Milic-Emili, J. & Gottfried, S. 1990. The role of PEEP in patients with chronic obstructive pulmonary disease during assisted ventilation. *European Respiratory Journal,* 3**,** 818-822.

Sabine, K. Z., Benjamin, S., Jörn, K., Ullrich, M. L., Knut, M. & Zhanqi, Z. 2015. Positioning of electrode plane systematically influences EIT imaging. *Physiological Measurement,* 36**,** 1109.

Salazar, E. & Knowles, J. H. 1964. An analysis of pressure-volume characteristics of the lungs. *Journal of Applied Physiology,* 19**,** 97-104.

Sassoon, C. S., Del Rosario, N., Fei, R., Rheeman, C. H., Gruer, S. E. & Mahutte, C. K. 1994. Influence of pressure- and flow-triggered synchronous intermittent mandatory ventilation on inspiratory muscle work. *Critical Care Medicine,* 22**,** 1933-41.

Schiller, H. J., Steinberg, J., Halter, J., McCann, U., DaSilva, M., Gatto, L. A., Carney, D. & Nieman, G. 2003. Alveolar inflation during generation of a quasi-static pressure/volume curve in the acutely injured lung. *Critical Care Medicine,* 31**,** 1126-1133.

Schisterman, E. F., Perkins, N. J., Liu, A. & Bondell, H. 2005. Optimal cut-point and its corresponding Youden Index to discriminate individuals using pooled blood samples. *Epidemiology,* 16**,** 73-81.

Schranz, C., Docherty, P., Chiew, Y. S., Moller, K. & Chase, J. G. 2012a. Iterative integral parameter identification of a respiratory mechanics model. *BioMedical Engineering OnLine,* 11**,** 38.

Schranz, C., Docherty, P. D., Chiew, Y. S., Chase, J. G. & Moller, K. 2012b. Structural identifiability and practical applicability of an alveolar recruitment model for ARDS patients. *IEEE Transactions on Biomedical Engineering,* 59**,** 3396-3404.

Schranz, C., Docherty, P. D., Chiew, Y. S., Möller, K. & Chase, J. G. 2012c. Identifiability analysis of a pressure-depending alveolar recruitment model. *IFAC Proceedings Volumes*, 137-142.

Schumann, S., Messmer, F., Lichtwarck-Aschoff, M., Haberthuer, C. & Guttmann, J. 2011. Cardiogenic oscillations in spontaneous breathing airway signal reflect respiratory system mechanics. *Acta Anaesthesiologica Scandinavica,* 55**,** 980-6.

Schwarz, G. 1978. Estimating the dimension of a model. *The Annals of Statistics*, 6, 461-464.

Sharma, A. K. 2005. *Text Book Of Correlations And Regression*, Discovery Publishing House.

Shiu, K. K. & Rosen, M. J. 2006. Is there a safe plateau pressure threshold for patients with acute lung injury and acute respiratory distress syndrome? *American Journal of Respiratory and Critical Care Medicine,* 173**,** 686.

Singer, B. D. & Corbridge, T. C. 2009a. Basic invasive mechanical ventilation. *Southern Medical Journal,* 102**,** 1238-1245.

Singer, B. D. & Corbridge, T. C. 2009b. Basic invasive mechanical ventilation. *Southern Medical Journal,* 102**,** 1238-45.

Stahl, C. A., Moller, K., Schumann, S., Kuhlen, R., Sydow, M., Putensen, C. & Guttmann, J. 2006. Dynamic versus static respiratory mechanics in acute lung injury and acute respiratory distress syndrome. *Critical Care Medicine,* 34**,** 2090-2098.

Stewart, K. W., Pretty, C. G., Tomlinson, H., Thomas, F. L., Homlok, J., Nemedi Noemi, S., Illyes, A., Shaw, G. M., Benyo, B. & Chase, J. G. 2016. Safety, efficacy and clinical generalization of the STAR protocol: a retrospective analysis. *Annals of Intensive Care*, 6, 24.

Sundaresan, A., Chase, J. G., Shaw, G., Chiew, Y. S. & Desaive, T. 2011. Model-based optimal PEEP in mechanically ventilated ARDS patients in the intensive care unit. *BioMedical Engineering OnLine,* 10**,** 64-81.

Sundaresan, A., Yuta, T., Hann, C. E., Chase, J. G. & Shaw, G. M. 2009. A minimal model of lung mechanics and model-based markers for optimizing ventilator treatment in ARDS patients. *Computer Methods and Programs in Biomedicine,* 95**,** 166-180.

Sutera, S. P. & Skalak, R. 1993. The history of Poiseuille's law. *Annual Review of Fluid Mechanics,* 25**,** 1-19.

Szlavecz, A., Chiew, Y. S., Redmond, D., Beatson, A., Glassenbury, D., Corbett, S., Major, V., Pretty, C., Shaw, G., Benyo, B., Desaive, T. & Chase, J. 2014. The clinical utilisation of respiratory elastance software (CURE Soft): a bedside software for real-time respiratory mechanics monitoring and mechanical ventilation management. *BioMedical Engineering OnLine,* 13**,** 140-153.

Tawhai, M. H., Hunter, P., Tschirren, J., Reinhardt, J., McLennan, G. & Hoffman, E. A. 2004. CT-based geometry analysis and finite element models of the human and ovine bronchial tree. *Journal of Applied Physiology,* 97**,** 2310-2321.

Tawhai, M. H., Pullan, A. J. & Hunter, P. J. 2000. Generation of an anatomically based three-dimensional model of the conducting airways. *Annals of Biomedical Engineering,* 28**,** 793-802.

Tgavalekos, N. T., Tawhai, M., Harris, R. S., Mush, G., Vidal-Melo, M., Venegas, J. G. & Lutchen, K. R. 2005. Identifying airways responsible for heterogeneous ventilation and mechanical dysfunction in asthma: an image functional modeling approach. *Journal of Applied Physiology,* 99**,** 2388-2397.

The Ards Definition Task Force. 2012. Acute respiratory distress syndrome: The Berlin definition. *JAMA: The Journal of the American Medical Association,* 307**,** 2526-2533.

Thomas, F., Pretty, C. G., Fisk, L., Shaw, G. M., Chase, J. G. & Desaive, T. 2014. Reducing the impact of insulin sensitivity variability on glycaemic outcomes using separate stochastic models within the STAR glycaemic protocol. *Biomedical Engineering Online,* 13**,** 43-43.

Ting, K. M. 2011. *Encyclopedia of Machine Learning*, Springer.

Tusman, G., Suarez-Sipmann, F., Peces-Barba, G., Climente, C., Areta, M., Arenas, P. G. & Bohm, S. H. 2009. Pulmonary blood flow generates cardiogenic oscillations. *Respiratory Physiology & Neurobiology,* 167**,** 247-54.

Van Den Berghe, G., Wilmer, A., Hermans, G., Meersseman, W., Wouters, P. J., Milants, I., Van Wijngaerden, E., Bobbaers, H. & Bouillon, R. 2006. Intensive insulin therapy in the medical ICU. *New England Journal of Medicine,* 354**,** 449-461.

Van Drunen, E., Chiew, Y. S., Chase, J., Shaw, G., Lambermont, B., Janssen, N., Damanhuri, N. & Desaive, T. 2013. Expiratory model-based method to monitor ARDS disease state. *BioMedical Engineering OnLine,* 12**,** 57-71.

Van Drunen, E., Chiew, Y. S., Pretty, C., Shaw, G., Lambermont, B., Janssen, N., Chase, J. & Desaive, T. 2014. Visualisation of time-varying respiratory system elastance in experimental ARDS animal models. *BMC Pulmonary Medicine,* 14**,** 33.

Villard, P. F., Beuve, M., Shariat, B., Baudet, V. & Jaillet, F. 2004. Lung mesh generation to simulate breathing motion with a finite element method. *Eighth International Conference on Information Visualisation*, 194-199.

Villard, P. F., Beuve, M., Shariat, B., Baudet, V. & Jaillet, F. 2005. Simulation of lung behaviour with finite elements: influence of bio-mechanical parameters. *Third International Conference on Medical Information Visualisation - BioMedical Visualisation*, 9-14.

Werner, R., Ehrhardt, J., Schmidt, R. & Handels, H. 2008. Modeling respiratory lung motion: a biophysical approach using finite element methods. *Proceedings of SPIE, Medical Imaging 2008: Physiology, Function, and Structure from Medical Images*, 6916.

Werner, R., Ehrhardt, J., Schmidt, R. & Handels, H. 2009. Patient-specific finite element modeling of respiratory lung motion using 4D CT image data. *Medical Physics,* 36**,** 1500-1511.

Whitehead, T. & Slutsky, A. 2002. The pulmonary physician in critical care • 7: Ventilator induced lung injury. *Thorax,* 57**,** 635-642.

Wilcoxon, F. 1945. Individual comparisons by ranking methods. *Biometrics Bulletin,* 1**,** 80-83.

Wilinska, M. E., Chassin, L. J., Schaller, H. C., Schaupp, L., Pieber, T. R. & Hovorka, R. 2005. Insulin kinetics in type-1 diabetes: continuous and bolus delivery of rapid acting insulin. *IEEE Transactions on Biomedical Engineering,* 52**,** 3-12.

Won, M., Yoon, H. & Treagust, D. F. 2014. Students' learning strategies with multiple representations: explanations of the human breathing mechanism. *Science Education,* 98**,** 840-866.

Wong, X. W., Singh-Levett, I., Hollingsworth, L. J., Shaw, G. M., Hann, C. E., Lotz, T., Lin, J., Wong, O. S. W. & Chase, J. G. 2006. A novel, model-based insulin and nutrition delivery controller for glycemic regulation in critically ill patients. *Diabetes Technology and Therapeutics,* 8**,** 174-190.

Yang, Y. 2005. Can the strengths of AIC and BIC be shared? A conflict between model indentification and regression estimation. *Biometrika,* 92**,** 937-950.

Yuta, T., Chase, J. G., Shaw, G. M. & Hann, C. 2004. Dynamic models of ARDS lung mechanics for optimal patient ventilation. *Proceedings of the 26th International Conference of IEEE Engineering in Medicine and Biology Society (EMBC 2004)*. 861-864.

Zambon, M. & Vincent, J. L. 2008. Mortality rates for patients with acute lung injury/ARDS have decreased over time. *Chest,* 133**,** 1120-1127.

Zilberberg, M. D., Luippold, R. S., Sulsky, S. & Shorr, A. F. 2008. Prolonged acute mechanical ventilation, hospital resource utilization, and mortality in the United States. *Critical Care Medicine,* 36**,** 724-730.