Traversing the Fuzzy Valley: Problems Caused by Reliance on Default Simulation and Parameter Identification Programs for Discontinuous Models

P.D. Docherty*, C. Schranz*,, J.G. Chase*, Y.S. Chiew* and K. Möller****

** Department of Mechanical Engineering, University of Canterbury, Christchurch, New Zealand (author email: paul.docherty@canterbury.ac.nz)*
*** Institute of Technical Medicine, Furtwangen University, Villingen-Schwenningen, Germany*

The Levenberg-Marquardt parameter identification method is often used in tandem with numerical Runge-Kutta model simulation to find optimal model parameter values to match measured data. However, these methods can potentially find erroneous parameter values. The problem is exacerbated when discontinuous models are analyzed.

A highly parameterized respiratory mechanics model defines a pressure-volume response to a low flow experiment in an acute respiratory distress syndrome patient. Levenberg-Marquardt parameter identification is used with various starting values and either a typical numerical integration model simulation or a novel error-stepping method.

Model parameter values from the error-stepping method were consistently located close to the error minima (median deviation: 0.4%). In contrast, model values from numerical integration were erratic and distinct from the error minima (median deviation: 1.4%).

The comparative failure of Runge-Kutta model simulation was due to the method's poor handling of model discontinuities and the resultant lack of smoothness in the error surface. As the Leven-berg-Marquardt identification system is an error gradient decent method, it depends on accurate measurement of the model-to-measured data error surface. Hence, the method failed to converge accurately due to poorly defined error surfaces.

When the error surface is imprecisely identified, the parameter identification process can produce sub-optimal results. Particular care must be used when gradient decent methods are used in conjunction with numerical integration model simulation methods and discontinuous models.

Parameter identification, Physiological modelling, Numerical integration, Hickling model, Alveolar recruitment.

1. INTRODUCTION

Model-based approaches can be used to characterize and quantify an individual's response to certain stimulus. This information can be used to guide therapy or aid diagnosis (Sundaresan et al. 2009; Lozano et al. 2008; Chase et al. 2011). However, fitting a model simulation to measured data often requires complex mathematical algorithms, and success is not guaranteed.

The default approach used by most researchers to identify model parameters involves iterations of error gradient descent with model simulation via time-stepping numerical integration (Schranz et al. 2012). IT is easily implemented in a range of existing software, but the final outcome can potentially be incorrect due to the effect of unsuitable model simulation techniques on the parameter identification algorithms perception of the error surface.

This study will identify parameters of an enhanced discontinuous model of alveolar recruitment as partially described by Hickling et al. (Hickling 1998; Hickling 2002). The model includes terms that describe the recruitment and expansion of alveoli that are open at end expiration and those that are recruited during inspiration.

II. METHODS

2.1. Subject and clinical protocol

One representative subject was selected from a study of 12 Acute Respiratory Distress Syndrome (ARDS) patients. Patients were mechanically ventilated with low-flow (LF) manoeuvres performed using a Evita4Lab-System (Stahl et al. 2006). During the LF-Manoeuvre, the lung is inflated by an extremely low constant gas flow of 35 mL/s over 55 s, yielding a quasi-static pressure/volume curve. Measurements consist of flow rate and airway pressure sampled at 125 Hz.

The study was approved by local ethics committees of all participating university hospitals. Informed consent was signed by patients or their legally authorized representative. Details of the experimental process are in Stahl et al. (Stahl et al. 2006).

2.2. Model description

The Hickling model used in this analysis assumes that recruitable alveoli open at discrete pressures, with common compliance and distensibility (Schranz et al. 2012). The lung is modelled as thirty discrete layers of alveoli that open at 0.5 mbar increments from *TOP* + 0.5 to *TOP* + 15 mbar. The model assumes distensibility with respect to impacting pressure of open alveoli is equal to that of the recruitable alveoli. The model is fully defined in (Schranz et al. 2012) and is presented:

$$p_{aw}(t) = R\dot{V}(t) + p_a(t) \qquad 1$$

$$\dot{p}_a(t) = \dot{V}(t)\left(C_1 e^{-kp_a(t)} + \sum_{\xi} \mathrm{H}_{\xi}(t) C_N e^{-k(pa(t)-TOP-\xi/2)} \right)^{-1} \qquad 2$$

$$\mathbf{X} = [R,\ C_1,\ C_N,\ k] \qquad 3$$

where: $\mathbf{X}$ is the identified set; $\xi = 1,2,3,...,30$; and

$$\mathrm{H}_{\xi}(t) = \begin{cases} 0, & p_a(t) \le \xi/2 + TOP \\ 1, & p_a(t) > \xi/2 + TOP \end{cases}$$

Equation nomenclature is defined in Table 1, and a typical model response is shown in Figure 1. Note that the total volume response is made up of the open alveoli volume (V_1) and the recruitable alveoli volume (V_N). V_N is made up of a series of 'layers', which are activated at discrete pressure intervals, and is the source of model discontinuity.

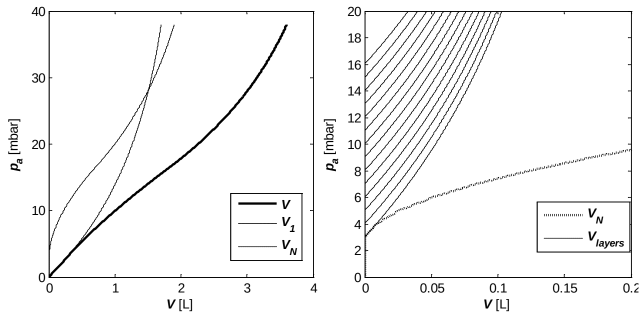

Figure 1. Pressure-volume curve defined by the model (left) and the contribution of the layers to the volume of the recruitable alveoli (right).

Table 1. Equation nomenclature

Symbol	Meaning	units	model role
p_{aw}	Airway pressure	mbar	measured
V	Lung volume	mL	$\dot{V}$ measured
p_a	Alveolar pressure	mbar	simulated
H_{ξ}	Heaviside function	{0, 1}	conditional
R	Flow resistance	mbar·mL^{-1}·s^{-1}	identified
C_1	Open alveolar compliance	mL·mbar^{-1}	identified
C_N	Recruitable alveolar compliance	mL·mbar^{-1}	identified
k	Distention	mbar^{-1}	identified
TOP	Threshold opening pressure	mbar	assumed

2.3. Parameter identification and model simulation

The values of the four 'identified' parameters in Table 1 were identified using the MATLAB^{TM} (Version R2011b 64-bit) proprietary Levenberg-Marquardt parameter identification method lsqnonlin.m (Levenberg 1944; Marquardt 1963). Model simulation was undertaken using two methods: 1) Time-stepping numerical integration using the MATLAB^{TM} Runge-Kutta function ode45.m with default settings (Dormand and Prince 1980); and 2) Iterative error-stepping method. The error-stepping method is novel and model dependent, and uses the algorithm defined in steps 1-5:

1. Initialize $p_a(t) = V(t)/C_1$
2. Derive the 30 Heaviside functions $\mathrm{H}_{1\text{-}30}(t)$ using $pa(t)$
3. Calculate $\dot{p}_a(t)$ directly with $\mathrm{H}_{1\text{-}30}(t)$, the measured $\dot{V}(t)$, and Equation 2.
4. Calculate $p_a(t)$ using the trapezium method (cumtrapz.m in MATLAB^{TM}) to integrate $\dot{p}_a(t)$.
5. Iterate across steps 2-4 updating $\mathrm{H}_{1\text{-}30}(t)$ at each iteration. Ten iterations are used.

A maximum pressure of p_a=1500 mbar is used in the error-stepping method to avoid simulation failure due to infinite pressure values, which can occur in early iterative steps, or by poorly estimated model parameter values yielding physiologically insufficient maximum lung volume.

2.4. Analysis

Initially, the participant's least-square model parameter values were identified using the Levenberg-Marquardt method, and the error-stepping model simulation method. The function tolerance ('tolfun') convergence setting was set to 10^{-12} and the maximum allowable iterations was increased to 10000. Initial values of R = 0.2; $C1$ = 10; CN = 3; and k = 0.01 were used. *TOP* was set to 10 mbar. These steps provide a set of highly accurate parameter values for the test participant ($\mathbf{X_T}$).

Second, model error surfaces ±5% around these $\mathbf{X_T}$ parameter values were obtained using both the numerical integration method and the error-stepping method. A series of two-dimensional error contour plots were defined for each possible 2-parameter combination. The natural logarithms of the error were recorded to allow comparable magnitudes of error across variations in model parameters with different model sensitivities. The contour plots had a 0.2% resolution and model error was calculated:

$$\mathrm{E}r = \ln\left(\sum_{t=0}^{50} \left(p_{aw}(t)_{measured} - p_{aw}(t)_{modeled} \right)^2 \right) \qquad 4$$

Finally, C_1 and C_N were re-identified in a Monte Carlo analysis. The values of k and R were set constant at the previously identified values ($k_{\mathbf{T}}$ and $R_{\mathbf{T}}$). Parameter identification was undertaken with the MATLAB^{TM} function lsqnonlin.m with convergence declared at 500 function evaluations. C_1 and C_N were identified using both the numerical integration and error-stepping methods, with the resultant parameter values recorded. Initial values of C_1 and C_N were randomly distributed on the range ±10% of the accurate $C_{1\mathrm{T}}$ and $C_{N\mathrm{T}}$ values. A total of 500 Monte Carlo iterations were undertaken. Normalized proportional differences ($\delta_{C1,CN}$) between the identified values and the minima value were measured with:

$$\delta_{C_1,C_N} = \sqrt{\left(\frac{C_1 - C_{1\mathrm{T}}}{C_{1\mathrm{T}}}\right)^2 + \left(\frac{C_N - C_{N\mathrm{T}}}{C_{N\mathrm{T}}}\right)^2} \qquad 5$$

III. RESULTS

The parameter identification method located an error minima with values of R = 0.121; C_1 = 32.706; C_N = 2.744; and k = 0.040, which are expected values for this type of model (Schranz et al. 2012). Figure 2 shows the error surface defined by the numerical integration method. Note that the surfaces show apparent wide scale trends, but have very erratic local error surfaces. Figure 3 shows the error surface of the error-stepping method. The overall shape of the error surface is similar to that of the numerical integration method. However, in contrast to the numerical integration method, the error-stepping method produced locally smooth contour surfaces.

Figure 4 shows the distribution of the C_1 and C_N values obtained after 500 Levenberg-Marquardt parameter identification iterations in relation to the error minima. Note that the locations of the error minima found with the error-stepping method were along the major axis of the error surface shown in Figure 3 (middle right). In contrast, the error minima located via the time-stepping method tended toward the major axis, but exhibited significant variation across the major axis.

Figure 5 shows the distribution of relative normalized discrepancy between the parameter values identified with the two model simulation methods and the parameter values at the true error minima. The median normalized parameter discrepancy caused by the error-stepping method was 0.4% (IQR: 0.3% to 1%). The median normalized parameter discrepancy caused by the numerical integration method was

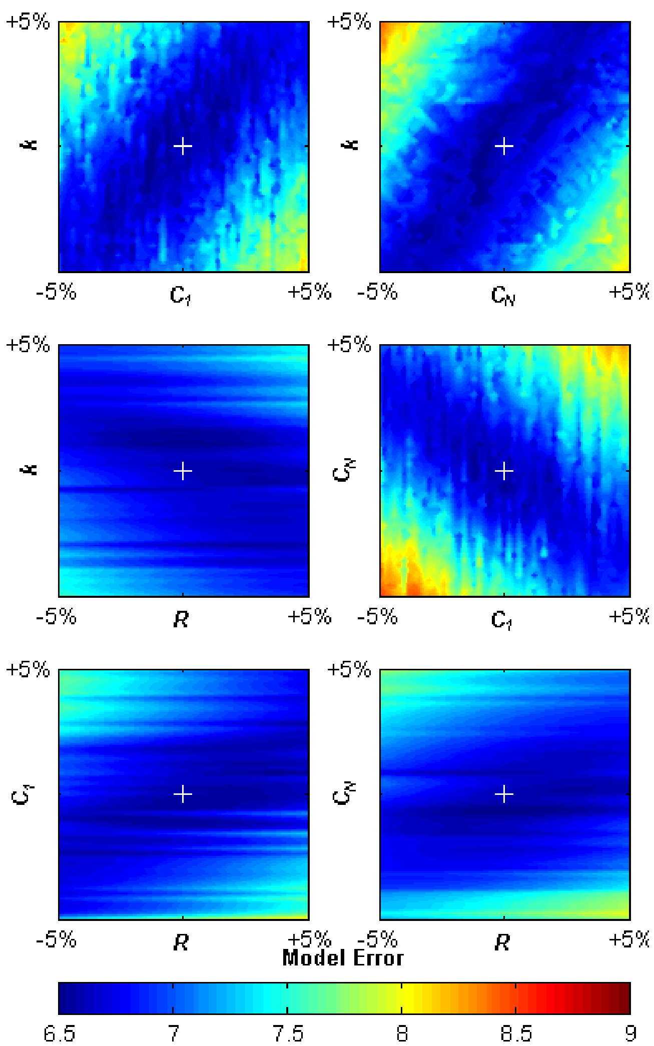

Figure 2. 2-D error maps across a 10% range about the error minima derived via numerical integration.

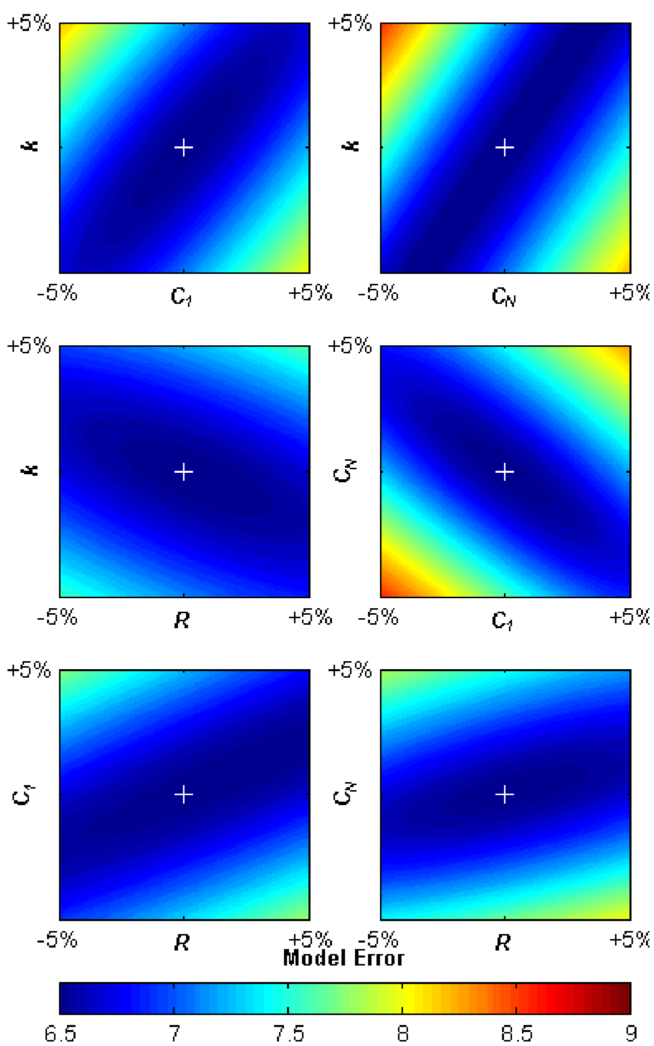

Figure 3. 2-D error maps across a 10% range about the error minima derived via error-stepping model simulation.

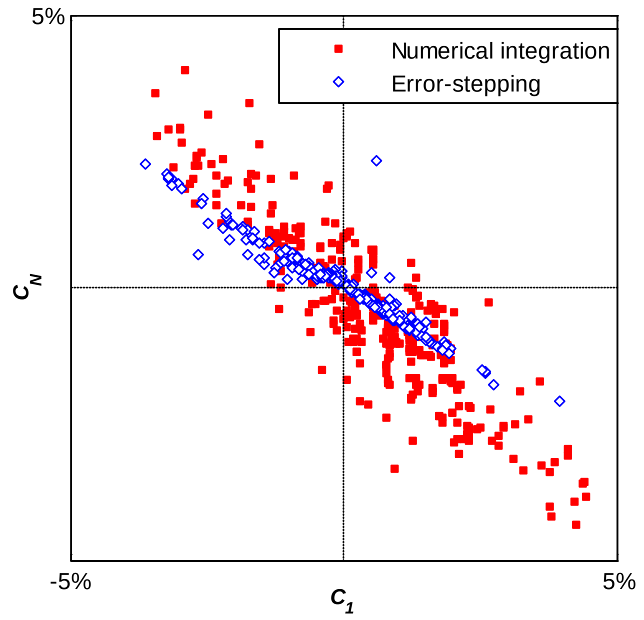

Figure 4. Distribution of parameter values identified via the Levenberg-Marquardt method with model simulation from the time-stepping and error-stepping methods

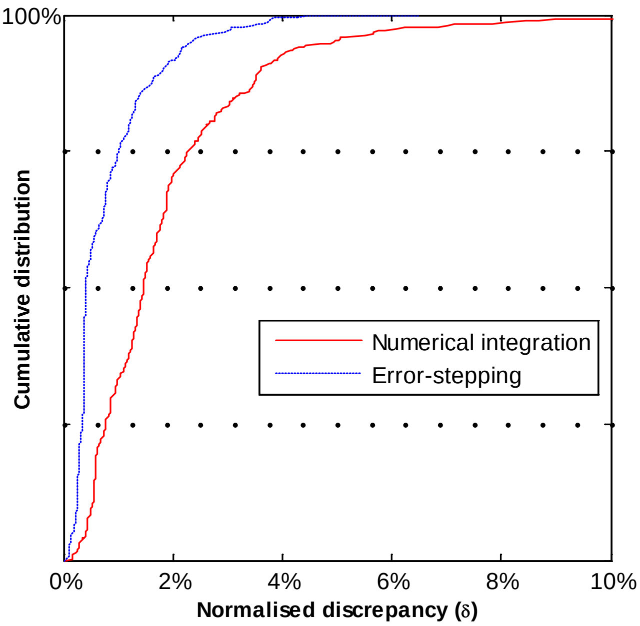

Figure 5. Distribution of the normalised parameter discrepancy from the two model simulation methods.

1.4% (IQR: 0.8% to 2.3%). Both Wilcoxon rank-sum and Kolmogorov-Smirnov tests found significant differences in the discrepancies across the methods ($p<0.00001$).

IV. DISCUSSION

This analysis demonstrated the potential for sub-optimal parameter identification when the Levenberg-Marquardt gradient decent method uses imprecise methods for model simulation of a discontinuous model. There was a significant improvement in the discrepancy between the ideal model parameter values and the values found via the precise error-stepping method over the values found with the more commonly used numerical integration method.

The model used incorporated a series of discontinuities that the MATLABTM numerical integration method ODE45.m failed to accurately define. Thus, small variance in parameter values sometimes caused disproportionately large changes in model simulation. This behaviour resulted in erratic local surfaces in the error plane (Figure 2). Gradient decent methods, such as the Levenberg-Marquardt method rely on accurate measurement of the gradient of the error plane to converge to a true error minima. As such, local inaccuracy in the error surface can cause erratic parameter stepping and failed convergence of parameter identification.

Figure 4 shows the distribution of minima located by the Levenberg-Marquardt identification method with model simulation from the error-stepping method and the time-stepping methods. As the error surface defined with the error-stepping method was locally smooth, parameter convergence was able to successfully descend the error plane. However, a number of outlying parameter values were distant from the true error minima identified in the initial stage of this investigation when the error-stepping method was used (Figure 3). Although this error is often referred to as 'identification of a local minima', there was actually no local minima at the identified values. Instead, the Levenberg-Marquardt identification method could not locate the proper direction of gradient decent.

Figure 6 shows the reason for the occasional halted convergence when parameter values fall on the major axis of an error contour. Point A in Figure 6 is situated on the major axis and thus has a limited range of directions with a lower error state. Thus, there is reduced likelihood of continued convergence. In contrast, point B has a much greater range of directions with a lower error-state, and thus convergence is much more likely to continue. Hence, in cases where the major axis of the error contours is much greater than the minor axis, convergence can be prematurely declared due to difficultly in locating a lower error state. In certain investigations, this occurs in cases where the true error minima is known (though another means), a false assumption of local minima is generally made.

Parameter values identified using the error-stepping method was generally closer to the error minima than values identified using numerical integration. The Levenberg-Marquardt method could not accurately recognize or define the local error-surface. This deleterious effect was most profound where the error surface was flattest; along the major axis close to the error minima. In regions where the error surface gradient was greatest, the gradient-to-surface noise ratio was such that the identification method could more readily obtain an appropriate direction for convergence. This is despite the local error surface noise caused by the numerical integration Runge-Kutta model simulation.

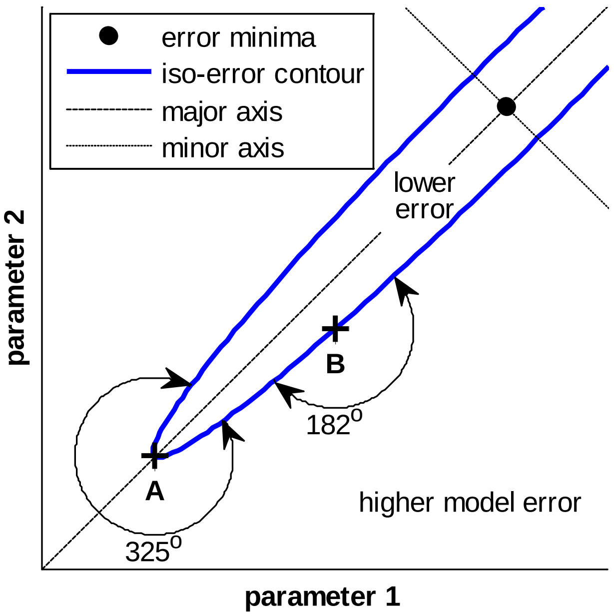

Figure 6. A typical error surface showing the range of directions of in-creased model error from two model parameter sets (A, B) with the same model error magnitude. Note that it is significantly less likely for a randomly selected convergence direction from point A to locate a lower error state than a randomly selected direction from point B.

The discrepancy measured by both methods was within tolerances that could be considered clinically acceptable. However, this investigation successfully demonstrates the potential for failed convergence, when erroneous model simulation techniques are used. The MATLABTM numerical integration function ode45.m was used with the default settings during this investigation. Altering the step-size of this algorithm is relatively simple and could significantly mitigate the errors encountered using the method. However, the reasons for this step are not immediately apparent, and may be overlooked. The primary focus of this investigation was to show the potentially deleterious effect of default use of a common model fitting strategy.

V. CONCLUSIONS

While numerical integration coupled with gradient based identification methods are frequently used successfully (Sundaresan et al. 2009; Schranz et al. 2011), it is important to know the limitations of such methods. In particular, this investigation has shown that numerical integration methods which handle discontinuities poorly, such as the proprietary MATLABTM methods, are likely to cause parameter identification failure when poorly applied. Thus, investigators using similar models must ensure that the model simulation methodology can define smooth error surfaces that do not hinder a gradient decent parameter identification method.

REFERENCES

Sundaresan A., Yuta T., Hann C.E., et al. (2009) A Minimal Model of Lung Mechanics and Model based Markers for Optimizing Ventilator Treatment in ARDS Patients, Comput Methods Programs Biomed., 95(2), 166-180

Lozano S., Möller K., Brendle A., et al. (2008) AUTOPILOT-BT: a system for knowledge and model based mechanical ventilation, Technol Health Care., 16(1), 1-11

Chase J.G., Le Compte A.J., Preiser J.C., et al. (2011) Physiological modeling, tight glycemic control, and the ICU clinician: what are models and how can they affect practice?, Ann Intensive Care, 1(1), 11

Schranz C., Docherty P.D., Chiew Y.S., et al.: A Time-Continuous Model of Respiratory Mechanics of ARDS Patients. In World Congress 2012 Medical Physics and Biomedical Engineering Beijing, China, submitted, 2012

Hickling K.G. (1998) The pressure-volume curve is greatly modified by recruitment. A mathematical model of ARDS lungs, Am J Respir Crit Care Med., 158(1), 194-202

Hickling K.G. (2002) Reinterpreting the pressure-volume curve in patients with acute respiratory distress syndrome, Curr Opin Crit Care, 8(1), 32-38

Stahl C.A., Möller K., Schumann S., et al. (2006) Dynamic versus static respiratory mechanics in acute lung injury and acute respiratory distress syndrome, Crit Care Med., 34(8), 2090-2098

Levenberg K. (1944) A Method for the Solution of Certain Problems in Least Squares., Quart. Appl. Math., 2(164-168

Marquardt D.W. (1963) An Algorithm for Least-Squares Estimation of Nonlinear Parameters., SIAM J. Optim. Appl. Math., 11(2), 431-441

Dormand J.R., Prince P.J. (1980) A family of embedded Runge-Kutta formulae, J. Comp. Appl. Math., 6(19-26

Schranz C., Knöbel C., Kretschmer J., et al. (2011) Hierarchical Parameter Identification in Models of Respiratory Mechanics, IEEE Trans Biomed Eng., 58(11), 3234-3241