Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Using the Adapted Levenberg-Marquardt method to determine the validity of ignoring insulin and glucose data that is affected by mixing

Nicholas Lam*. Paul D. Docherty*. Rua Murray. J. Geoffrey Chase*, Lisa Te Morenga*****

**Department of Mechanical Engineering, University of Canterbury, Christchurch,Private Bag 4800, New Zealand (ph:+6433692230 nicholas.lam@pg.canterbury.ac.nz, paul.docherty@canterbury.ac.nz, geoff.chase@canterbury.ac.nz)*
***School of Mathematics and Statistics, University of Canterbury, Christchurch, (rua.murray@canterbury.ac.nz)*
****Faculty of Health, Victoria University of Wellington, New Zealand (lisa.temorenga@vuw.ac.nz)*

Abstract - Most parameter ID methods use least squares criterion to fit parameter values to observed behavior. However, the least squares criterion can be heavily influenced by outlying data or un-modelled effects. In such cases, least squares estimation can yield poor results. Outlying data is often manually removed to avoid inaccurate outcomes, but this process is complex, tedious and operator dependent.

This research presents an adaptation of the Levenberg-Marquardt (L-M) parameter identification method that effectively ignores least-square contributions from outlying data. The adapted method (aL-M) is capable of ignoring outlier data in accordance with the coefficient of variation of the residuals and was thus, capable of operator independent omission of outlier data using the 3 standard deviation rule. The aL-M was compared to the original Levenberg-Marquardt (L-M) method in C-peptide, insulin and glucose data. In total three cases were tested: L-M in the full dataset, L-M in the same data where the points that were suspected to be affected by incomplete mixing at the depot site were removed, and the aL-M in the full data set.

There were strong correlations between the aL-M and the reduced dataset from [0.85, 0.71] for the clinically valuable glucose parameters. In contrast, the unreduced data yielded poor residuals and poor correlations with the aL-M [0.44, 0.33]. The aL-M approach provided strong justification for consistent removal of data that was deemed to be affected by mixing.

Keywords: Glycemic modelling, least squares estimation, outlier data, numerical optimisation

1. INTRODUCTION

Parameter identification methods are used to determine optimal parameter values such that models can accurately capture some observed behaviour (Carson and Cobelli, 2001). Most parameter identification algorithms identify these parameter values by minimising a least squares objective/penalty function (Bard, 1970, Davidon, 1991, Docherty et al., 2012, Levenberg, 1944, Marquardt, 1963, Steihaug, 1983). This means doubling the distance of a data point from the modelled behaviour will lead to four times the influence from the objective function.

This approach works well with most datasets, but is a cause of inaccurate parameter identification when outlying data is present (Sheiner and Beal, 1985). Outliers can cause least-squares optimal parameter sets to diverge from an optimal parameter set defined by 'inlying' data points. This issue is overcome by performing inverse problems over a number of observations, determining the variance of the residuals, then defining points outside 3 standard deviations from simulated behaviour to be outliers and omitting them from subsequent iterations (Pukelsheim, 1994, Bakar et al., 2006). This process is time-consuming and can lead to ambiguous outcomes and diminished operator independence.

We previously presented an adaption of the Gauss-Newton gradient-descent parameter identification method that reduces the contribution of outliers to the inverse problem (Gray et al., 2016). Subsequently, this adaptation was compared with a typical approach through modelling a cohort of C-peptide and insulin data and showed that the adapted method can capture model parameters obscured by outlier points (Docherty et al., 2014). This analysis compares very similar methods with the addition of glucose modelling, and also compares the adapted method to the typical method where the main unmodelled outlier data has been manually removed in all datasets.

2. METHODS

2.1 Clinical Protocol

This analysis used data from a dietary intervention study that measured the effect of dietary fibre in females at risk of developing type 2 diabetes. The outcomes of the trial were presented by TeMorenga et al. (2010). Eighty-three individuals underwent the DISST (Lotz et al., 2010) at weeks 0, 12, and 24. Some participants did not attend followup appointments and a total of 218 DISST procedures were undertaken.

Participants fasted from 10 p.m. the night before the test and attended the clinic in the morning. During the test, participants

sat in a relaxed position. They had a cannula placed in their antecubital-fossa to administer glucose and insulin boluses, and draw blood samples. This ultimately led to high local depot concentrations of insulin and glucose after administration. A 10 g glucose bolus (50% dextrose) was administered at 6 minutes, and a bolus of actrapid insulin was administered at 6 minutes. Blood samples were taken at t = 0, 5, 10, 15, 20, 25, 30, 35, 40, and 50 minutes. Glucose was measured at the bedside (Enzymatic glucose hexokinase assay, Abbot Labs, Illinois, USA), and samples were then spun and frozen for batch assays of insulin and C-peptide (ELISA Immunoassay, Roche, Germany).

2.2 DISST Model

The DISST model defines glucose, insulin, and C-peptide kinetics (Lotz et al., 2010). The models are defined:

$$\dot{C} = k_2 Y - (k_1 + k_3)C + U_N \tag{1}$$

$$\dot{Y} = k_1 C - k_2 Y \tag{2}$$

$$U_N = U_B + U_1(t) + U_2(t) + U_3(t) \tag{1a}$$

$$U_{\mathrm{B}} = k_3 C_0 \tag{1b}$$

$$U_1(t) = \begin{cases} \theta_1, & t = 6 \\ 0, & otherwise \end{cases} \tag{1c}$$

$$U_2(t) = \begin{cases} \theta_2(60 - t)/54, & 6 \le t \le 60 \\ 0, & t < 6 \end{cases} \tag{1d}$$

$$U_3(t) = \begin{cases} \theta_3(t - 6)/54, & 6 \le t \le 60 \\ 0, & t < 6 \end{cases} \tag{1e}$$

$$\dot{I} = \frac{n_I}{V_p} Q - \left(\theta_4 + \frac{n_I}{V_p}\right) I + \theta_5 U_N + \frac{U_X}{V_P} \tag{3}$$

$$\dot{Q} = \frac{n_I}{V_Q} I - (n_C + \frac{n_I}{V_Q}) Q \tag{4}$$

$$\dot{G} = p_G(G - G_0) - \theta_6(GQ - G_0 Q_0) + \theta_7 P_X \tag{5}$$

where: U_N is the endogenous insulin production comprised of the basal rate (U_B), and first and second phases of insulin release (U_{1-3}) (pmol·L^{-1}·min^{-1}); U_X is the exogenous insulin dose (mU·min^{-1}); C is the plasma C-peptide concentration (pmol·L^{-1}); Y is the interstitial C-peptide concentration (pmol·L^{-1}); I is the plasma insulin concentration and Q is the interstitial insulin concentration (mU·L^{-1}); G is the blood glucose concentration (mmol·L^{-1}); P_X is the exogenous glucose dose (g·min^{-1}) V_P is the plasma insulin distribution volume (L); V_Q is the interstitial insulin distribution volume (L); k_{1-3} are the C-peptide kinetic parameters (min^{-1}); n_I is the plasma-interstitial diffusion rate (L·min^{-1}); n_C is the interstitial insulin degradation rate (min^{-1}); p_G is the non-insulin mediated glucose disposal rate (min^{-1}).

θ_{1-7} are the lumped parameters identified in this analysis. θ_1 is the first phase insulin release, θ_2 and θ_3 are the start and finish of the second phase insulin release, respectively. θ_4 is a combined metric for renal and hepatic insulin clearance. θ_5 is equal to 1 minus the first pass hepatic extraction of insulin. θ_6 is the insulin sensitivity of the subject and is the key metric of clinical interest. θ_7 is the inverse of the glucose distribution volume. The remaining parameters from (1) and (2) were determined *a-priori* via the methods of Van Cauter *et al.* (Lotz et al., 2010, Van Cauter et al., 1992).

2.3 Parameter identification methods

This analysis compares the outcomes of the adapted Levenberg-Marquardt method (aL-M) with the original approach (L-M). The original Levenberg-Marquardt parameter identification approach iterates towards the optimal parameter set ($\boldsymbol{\theta_{opt}}$) with the iterative process:

$$\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i - (\mathbf{J^T J} + \lambda \cdot diag(\mathbf{J^T J}))^{-1} \mathbf{J^T \psi} \tag{6}$$

where:

$$\mathbf{J} = \begin{bmatrix} \frac{\partial \psi_1}{\partial \theta_1} & \frac{\partial \psi_1}{\partial \theta_2} & \cdots & \frac{\partial \psi_1}{\partial \theta_n} \\ \frac{\partial \psi_2}{\partial \theta_1} & \frac{\partial \psi_2}{\partial \theta_2} & \cdots & \frac{\partial \psi_2}{\partial \theta_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \psi_m}{\partial \theta_1} & \frac{\partial \psi_m}{\partial \theta_2} & \cdots & \frac{\partial \psi_m}{\partial \theta_n} \end{bmatrix} \tag{6a}$$

$$\boldsymbol{\psi} = \left[X(\boldsymbol{\theta_i}, t_j) - X_{M,j}\right] = \begin{bmatrix} X(\boldsymbol{\theta_i}, t_1) - X_{M,1} \\ X(\boldsymbol{\theta_i}, t_2) - X_{M,2} \\ \vdots \\ X(\boldsymbol{\theta_i}, t_m) - X_{M,m} \end{bmatrix} \tag{6b}$$

and $\mathbf{J}$ is a Jacobian, $\boldsymbol{\psi}$ is the residual vector, X is the measured property; j is the sample index from 1 up to the number of samples (m, $j = 1..m$); $X(\boldsymbol{\theta_i}, t_j)$ is the simulated value of X at $t = t_j$; and $X_{M,j}$ is the measured value of X at $t = t_j$. The damping term λ scales based on the Jacobian value. In contrast to typical implementation of L-M, λ is set as a constant in this analysis. This sacrificed some convergence speed but enabled more stable and consistent iteration.

The aL-M was designed to dissipate the contribution of outlying data on the identification of $\boldsymbol{\theta_{opt}}$. The Gauss-Newton method defines the optimal direction, given by the combined Jacobian terms ($(\mathbf{J^T J})^{-1}\mathbf{J^T}$), for reducing residuals of each data point, for each parameter. These direction vectors are multiplied by the residual matrix to determine the ideal direction for convergence. The adapted method modulates the effect of the residual vector by the residuals, but still uses the Jacobian in (6*a*). This contrasts with other robust estimation methods which use the long-established class of M-estimators (Farcomeni and Ventura, 2012, Banaś and Ligas, 2014). Hence, the adapted method reduces the effect of outliers on $\boldsymbol{\psi}$ in (6), while inheriting the robustness properties of Gauss-Newton, substituting $\boldsymbol{\psi}$ for $\widehat{\boldsymbol{\psi}}$:

$$\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i - (\mathbf{J}^{\mathbf{T}}\mathbf{J} + \lambda \cdot diag(\mathbf{J}^{\mathbf{T}}\mathbf{J}))^{-1}\mathbf{J}^{\mathbf{T}}\hat{\boldsymbol{\psi}} \quad (7)$$

where:

$$\hat{\boldsymbol{\psi}} = \boldsymbol{\psi} \odot \exp\left(\frac{-|\boldsymbol{\psi}|}{\beta|\psi_M|}\right) \quad (7a)$$

$|\psi_M|$ is the median of the absolute values of the residuals and β is a scaling factor that determines the width of the peak as a function of $|\psi_M|$. In this analysis, $\beta = 3$. This value provides maximal objective function contributions at $\psi = \pm\beta|\psi_M|$, shown in Fig. 1. In contrast, the objective surface of typical Gauss-Newton minimises least-squares residuals and thus follows ψ^2 as ψ increases. The choice of β significantly down-weights outlier data over three standard deviations, in accordance with accepted statistics for rejection (Pukelsheim, 1994, Bakar et al., 2006).

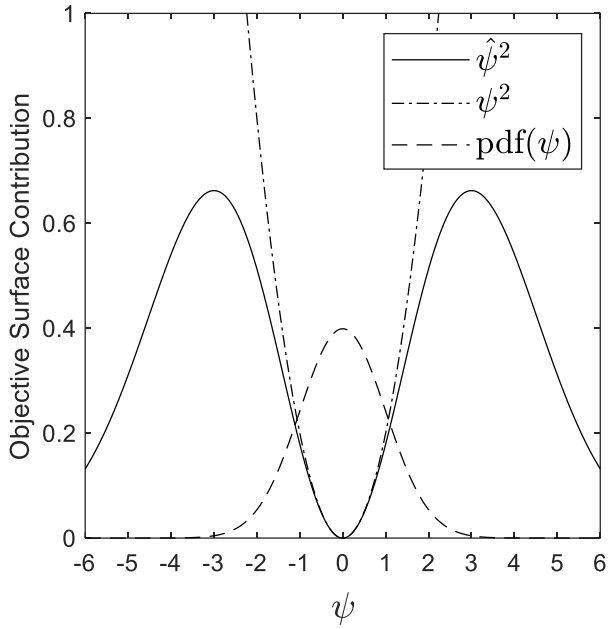

Fig. 1. Objective contributions from (6) and (7).

Three identification approaches were used to identify $\boldsymbol{\theta}$: L-M with the full dataset (L-Mf), L-M with a down-sampled dataset (L-Mds), and the aL-M with the full dataset. First, the L-M and aL-M approaches were used to identify the parameter set $\boldsymbol{\theta} = [\theta_1, \theta_2, \theta_3]^T$, to determine the contributions to U_N using the C-peptide data. This generated two sets of U_N profiles and two sets of residuals for each DISST trial. The U_N profiles were then used to identify the insulinaemic pharmaco-kinetic parameters $\boldsymbol{\theta} = [\theta_4, \theta_5]^T$. The U_N profile determined via the aL-M methodology was used for the aL-M estimation in the full insulin dataset. The U_N profile determined using L-M was used to identify the insulinaemic parameters using the L-M method, and the full insulin data set, then in a dataset that had the point 5 minutes after insulin administration manually removed. Finally, the modelled interstitial insulin concentration ($Q(t)$) was used to identify $\boldsymbol{\theta} = [\theta_6, \theta_7]^T$ using the glucose data. The aL-M derived $Q(t)$ profile was used to identify glycaemic parameters with aL-M. The $Q(t)$ profile from L-Mf was used to identify glucose parameters in the full data set. Finally, the $Q(t)$ profile from the L-Mds was used to aid identification of glycaemic parameters in a dataset that had the glucose samples 5 and 10 minutes after glucose administration removed. Visual inspection showed that glucose had slower mixing behaviour than insulin.

2.4 Evaluation

The 3 approaches were assessed qualitatively based on the model residuals. Since the typical approach minimises $\|\boldsymbol{\psi}\|_2$ and the adapted approach minimises $\|\hat{\boldsymbol{\psi}}\|_2$, the methods cannot be quantitatively compared. To highlight the differences in the approach outcomes, both summary statistics of absolute ψ values and residuals as a function of time ($\psi(t)$) will be presented.

3. RESULTS

Correlations between the methods are presented in Table 1. Samples at $t = 20$ minutes for insulin, and $t = 10$, 15 minutes for glucose were removed from the datasets to form the 'downsampled' parameter identification set. The adapted results correlated well to the downsampled glucose parameters, but were not so well correlated for the insulin parameters.

Table 1. Summary statistics of parameter correlations

Set 1	Set 2	Parameter Correlations
L-Mf	aL-M	[0.96, 0.92, 0.91, 0.44, 0.08, 0.44, 0.33]
L-Mds	aL-M	[0.96, 0.92, 0.91, 0.27, 0.26, 0.85, 0.71]
L-Mds	L-Mf	[1.00, 1.00, 1.00, 0.21, 0.11, 0.68, 0.15]

Summary statistics of the absolute residual data are presented in Table 2. This residual data was not moderated by equation 7a in any case, indicating a discordance across the objective function and residuals recorded for the aL-M.

Fig. 2 shows a set of responses in which outliers appear at the locations where the residuals are biased. Fig. 3 shows the distribution of the residuals about the measured points ($\boldsymbol{\psi}$). C-peptide samples are relatively well centred about zero and follow a seemingly normal distribution. In contrast, both the insulin and glucose residuals were sporadic, showing biases during the mixing phases at t=20 minutes for insulin, and t=10 minutes for glucose.

4. DISCUSSION

The adaptation to the Levenberg-Marquardt parameter identification method yielded different results to the typical approach. By minimising a residual that represents a majority of the datapoints and mitigating the contribution from outliers, the adapted method produced residuals that were qualitatively different from the original method (Figs. 2 and 3). The residual curves of the aL-M approach resembled those of the L-Mds, which identified parameters in data wherein the most common outliers had been removed.

The adapted method performs similarly to the original approach in noisy data that does not contain outliers, providing no benefit nor deleterious outcomes. However, when data contains known outliers, the adapted method considerably improves identification. The correlations in C-peptide parameters θ_{1-3} were between 0.91 and 0.96 (Table 1). These high correlations occurred as the C-peptide data contained measurement noise but no significant outliers, and the adapted method varied minimally from the original approach.

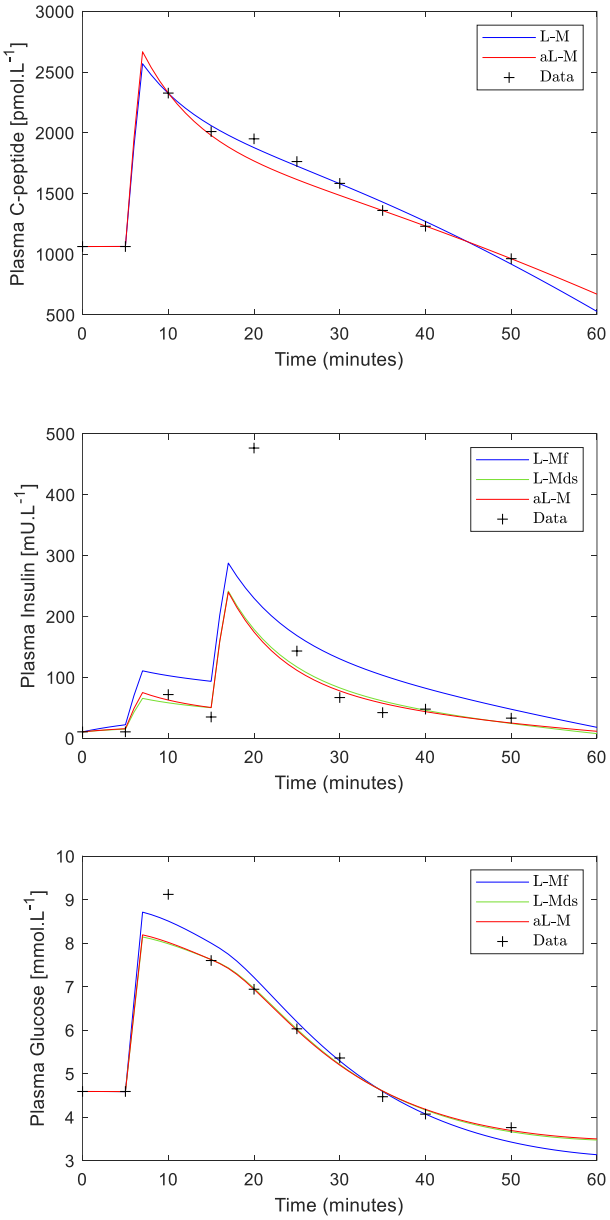

Fig. 2. Plasma C-peptide, insulin, and glucose responses of a patient response to the DISST test with noticeable outliers in the insulin and glucose data.

However, the insulin and glucose data both had un-modelled mixing behaviour that led to outlying data. This caused significant divergence in parameter values obtained (R=0.08 – 0.44 for the insulin parameters, and R=0.33 – 0.44 in glucose parameters). The aL-M method recognised the outlier points and minimised their contribution to convergence in the parameter space. Fig. 3 shows where biases occur in the models for the insulin and glucose data. In the insulin data, the original approach models the t = 20 minutes data point more accurately than the adapted approach. However, this data point is affected by incomplete mixing of insulin at the depot site, and is thus, an unmodelled phenomenon. Hence, this point was a consistent outlier in measured data, and an ideal fit for this data would not be influenced by this point. The adapted method increased the apparent residuals at the outlier data point in order to improve fit for the points that were well approximated by the model. Fig. 2 and Fig. 3 indicate performance of the aL-M was comparable to the manual removal of outliers undertaken in the L-Mds approach.

Table 2. Summary statistics of model residuals

Model	Approach	Percentiles of Relative Residuals (as percentage of observed data) [$\psi 25$, $\psi 50$, $\psi 75$, $\psi 95$, $\psi 99$]
C-peptide [pmol.L^{-1}]	L-M	[0.55, **1.69**, 3.62, 8.24, 13.6]
	aL-M	[0.08, **0.96**, 3.58, 9.89, 20.6]
Insulin [mU.L^{-1}]	L-Mf	[13.2, **34.8**, 78.0, 304, 1219]
	L-Mds	[6.94, **17.4**, 35.6, 100, 283]
	aL-M	[4.82, **14.8**, 36.4, 75.9, 138]
Glucose [mmol.L^{-1}]	L-Mf	[0.86, **2.82**, 5.54, 15.5, 31.3]
	L-Mds	[0.40, **1.52**, 3.35, 12.4, 24.4]
	aL-M	[0.28, **1.54**, 4.52, 14.0, 28.7]

The glucose data also benefits from the use of the adapted method, though not as much as the insulin. The first two datapoints have the most bias due to mixing at t = 10, 15 minutes, and adapted method again decreases their contribution to the objective surface and follows the remaining datapoints closely. Table 1 shows that the aL-M method displays good correlation with the L-Mds method (R=0.85, R=0.71), compared to the original approach (R=0.44, R=0.33). The higher correlation was for insulin sensitivity, which is the primary metric of interests in glycaemic modelling (Ferrannini et al., 1997, Haffner et al., 1999, Bergman et al., 1979).

This study considered applying a simple Gauss-Newton algorithm similar to the methods from Gray et al. (2016). However, this method was susceptible to instability when outliers were particularly severe in data. While adapted Gauss-Newton was stable, the un-adapted method became unstable and led to extreme residuals for several trials, preventing proper comparisons. Hence, a simplified version of the Levenberg-Marquart algorithm was used. This required more iterations to converge, but allowed consistent and stable results.

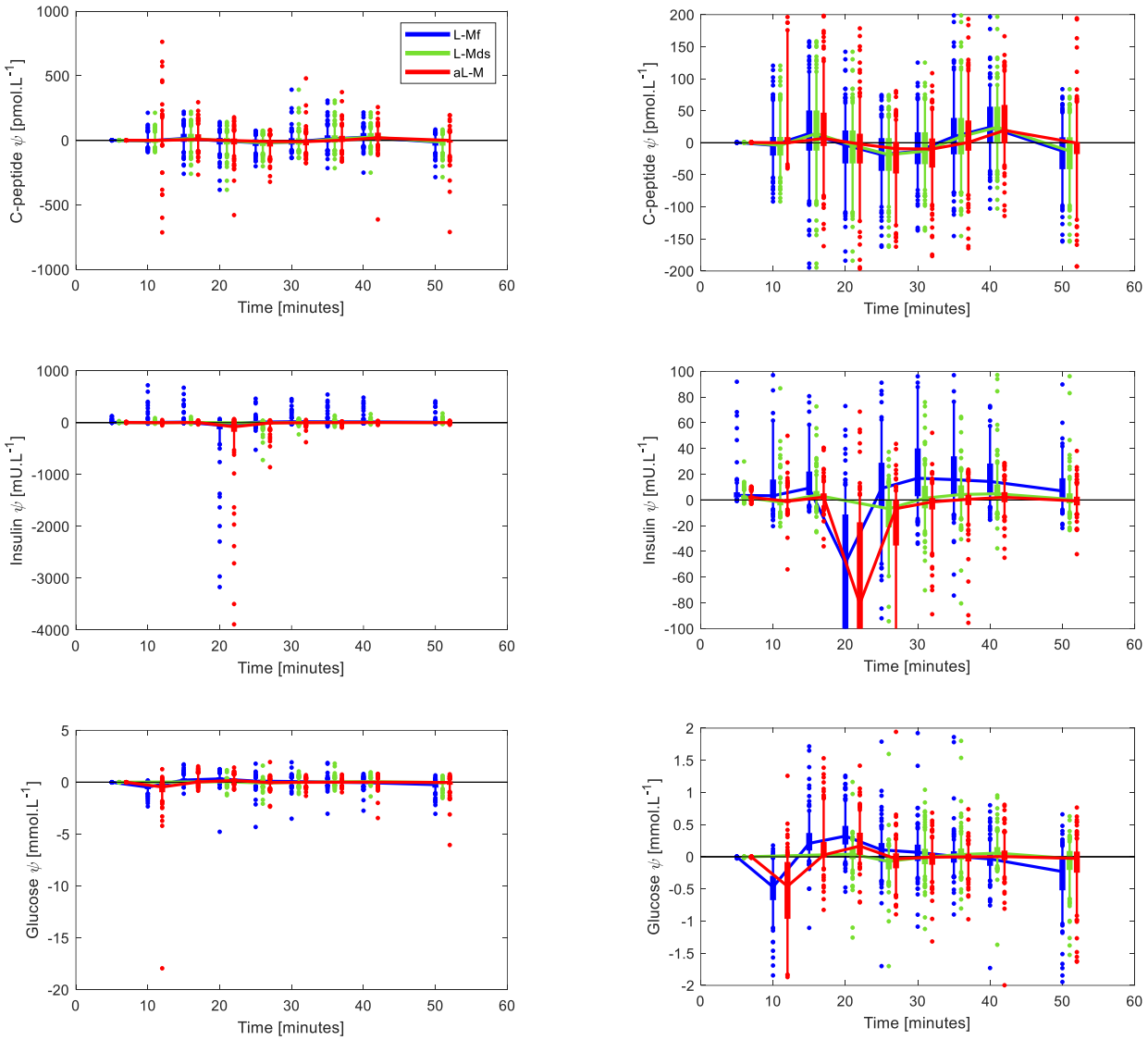

Fig. 3. Residual plots for C-peptide, insulin, and glucose. The plots on the right are cropped to show the general behaviour. The thick error bars show the interquartile range, thin error bars show the 5th to 95th percentile range, and the dots show the outlying points. The time points are offset for L-Mds and aL-M to enable clearer observation.

C-peptide data did not benefit from the adaptive method overall. While the interquartile ranges remained similar, the extreme outliers out spread further, creating outlier residuals. This suggests the adapted method is not beneficial in every case, and that the nature of noise in the data should be considered before choosing to use the method. In particular, the method should be used when there are significant outliers or unmodelled effects. This may be due to the method relying on inlier datapoints meeting or exceeding the number of identified parameters. With three parameters to identify and random noise sometimes creating too many points that classify as outliers, the method may return a poor model, this could potentially be ameliorated by shifting to a 75$^{\text{th}}$ percentile residual in Eq. 7a rather than the median, 50$^{\text{th}}$ percentile.

The adapted method is relatively simple to add to the parameter identification methods, and computationally inexpensive. The adapted method removes need for the manual removal of outlier data. Removing outlier data using other applications often requires two runs of the inverse problem: once to simulate a model, then again after datapoints that fall at more than three standard deviations from the model have been removed. This is costly in operator time, and risks reducing operator independence.

The presence of outliers is often suspected *a priori*. In particular, outliers can be observed when recording data, or through plotting the data before initiating parameter identification. However, it can be difficult to determine which data points should be declared outliers with statistically justified scientific integrity. This analysis shows that removing

points directly after mixing is justified according to the operator independent aL-M algorithm that automatically applies well-known statistical justification (Pukelsheim, 1994, Bakar et al., 2006). Further research could be performed to directly compare the adapted method with a downsampled method that does two runs of the inverse problem to properly remove outlier points.

5. CONCLUSIONS

In this analysis we tested the aL-M method against a more typical Levenberg-Marquardt method where outliers were both kept in, and manually removed. The methods were tested for both noisy data that contained outliers and un-modelled effects, and data that were only noisy. The aL-M method captured observed behaviour better in data that contained un-modelled effects or outlying data, performing similarly to the typical method following a manual removal of outliers. It provided minimal changes in identification to data that was just noisy.

Overall, this analysis showed that the data immediately after bolus administration are consistently affected by unmodelled mixing. This paper provides a statistical justification for consistent removal of these points in accordance with the 3-sigma rule outlier detection rule.

REFERENCES

BAKAR, Z. A., MOHEMAD, R., AHMAD, A. & DERIS, M. M. A Comparative Study for Outlier Detection Techniques in Data Mining. Cybernetics and Intelligent Systems, 2006 IEEE Conference on, 7-9 June 2006 2006. 1-6.

BANAŚ, M. & LIGAS, M. 2014. Empirical tests of performance of some M – estimators. *Geodesy and Cartography,* 63**,** 127-146.

BARD, Y. 1970. Comparison of Gradient Methods for the Solution of Nonlinear Parameter Estimation Problems. *SIAM Journal on Numerical Analysis,* 7**,** 157-186.

BERGMAN, R., IDER, Y., BOWDEN, C. & COBELLI, C. 1979. Quantitative estimation of insulin sensitivity. *Am J Physiol,* 236**,** E667 - 677.

CARSON, E. R. & COBELLI, C. 2001. *Modelling methodology for physiology and medicine,* San Diego, Academic Press.

DAVIDON, W. 1991. Variable Metric Method for Minimization. *SIAM Journal on Optimization,* 1**,** 1-17.

DOCHERTY, P. D., CHASE, J. G. & DAVID, T. 2012. Characterisation of the iterative integral parameter identification method. *Medical and Biological Engineering and Computing,* 50**,** 127-134.

DOCHERTY, P. D., GRAY, R. A. L. & MANSELL, E. 2014. Reducing the Effect of Outlying Data on the Identification of Insulinaemic Pharmacokinetic Parameters with an Adapted Gauss-Newton Approach. *In:* BOJE, E. (ed.) *IFAC 19th World Congress.* Cape Town, South Africa.

FARCOMENI, A. & VENTURA, L. 2012. An overview of robust methods in medical research. *Statistical Methods in Medical Research,* 21**,** 111-133.

FERRANNINI, E., NATALI, A., BELL, P., CAVALLO-PERIN, P., LALIC, N. & MINGRONE, G. 1997. Insulin resistance and hypersecretion in obesity. European Group for the Study of Insulin Resistance (EGIR). *Journal of Clinical Investigation,* 100**,** 1166-1173.

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.

HAFFNER, S. M., D'AGOSTINO, R., JR., MYKKANEN, L., TRACY, R., HOWARD, B., REWERS, M., SELBY, J., SAVAGE, P. J. & SAAD, M. F. 1999. Insulin sensitivity in subjects with type 2 diabetes. Relationship to cardiovascular risk factors: the Insulin Resistance Atherosclerosis Study. *Diabetes Care,* 22**,** 562-568.

LEVENBERG, K. 1944. A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied mathematics,* 2**,** 164-8.

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). *J Diabetes Sci Technol,* 4**,** 1408-1423.

MARQUARDT, D. W. 1963. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *SIAM Journal on Applied Mathematics,* 11**,** 431-441.

PUKELSHEIM, F. 1994. The Three Sigma Rule. *The American Statistician,* 48**,** 88-91.

SHEINER, L. B. & BEAL, S. L. 1985. Pharmacokinetic parameter estimates from several least squares procedures: Superiority of extended least squares. *Journal of Pharmacokinetics and Biopharmaceutics,* 13**,** 185-201.

STEIHAUG, T. 1983. The Conjugate Gradient Method and Trust Regions in Large Scale Optimization. *SIAM Journal on Numerical Analysis,* 20**,** 626-637.

TEMORENGA, L., WILLIAMS, S. M., BROWN, R. & MANN, J. I. 2010. Effect of a relatively high protein, high fiber diet on body composition and metabolic risk factors in overweight women. *Eur J Clin Nutr,* 64**,** 1323-31.

VAN CAUTER, E., MESTREZ, F., STURIS, J. & POLONSKY, K. 1992. Estimation of insulin secretion rates from C- peptide levels. Comparison of individual and standard kinetic parameters for C-peptide clearance. *Diabetes,* 41**,** 368 - 377.