engineering proceedings

MDPI

Article

Comparing two fitting algorithms for determining the Cole-Cole parameters in blood glucose problems

Roberto Dima [1], Giovanni Buonanno [1] and Raffaele Solimene [1]

[1] Department of Engineering, University of Campania, Aversa, 81031 Italy; roberto.dima@unicampania.it, giovanni.buonanno@unicampania.it, raffaele.solimene@unicampania.it

* Correspondence: roberto.dima@unicampania.it

Abstract: The paper addresses the non-linear inverse problem of estimating the parameters of the Cole-Cole model used to describe the behaviour of the complex permittivity of blood samples. Such a model provides an efficient and accurate representation of biological tissues in the entire frequency band considered and reduces the complexity of the experimental data to a few parameters. In this way, it is possible to extract a "synthetic view" of the dielectric properties of tissues in such a way that more information on the glucose concentration can be derived, in addition to the resonance peak or phase shift. In order to perform the fitting of the Cole-Cole model, two different algorithms are used and compared: the Levenberg-Marquardt and the Variable Projection algorithms. The synthetic data present in the literature are used to evaluate the performances obtainable with these methods. In particular, Monte Carlo analysis is used in order to evaluate the accuracy and the precision that these two methods provide in the process of estimating the parameters involved, with respect to the starting points of the parameters. The results obtained show that the variable projection algorithm always outperforms the Levenberg-Marquardt one, although the former has a greater computational burden than the latter.

Keywords: glucose measurement; Cole-Cole model; Levenberg-Marquardt algorithm; Variable Projection algorithm; blood dielectric properties; non-linear fitting problem

Citation: Dima, R.; Buonanno, G.; Solimene, R., "Comparing two fitting algorithms for determining the Cole-Cole parameters in blood glucose problems". *Appl. Sci.* **2021**, *1*, 0. https://doi.org/

Received:
Accepted:
Published:

1. Introduction

Diabetes is a metabolic disorder that afflicts millions of people in the world. It degrades the cell's ability to absorb glucose from the bloodstream because of the improper regulation of insulin hormone. For this reason, great efforts have been dedicated to the development of non-invasive glucose monitoring devices, which may considerably improve the quality of life for diabetics [1].

The present work, in particular, relates to microwave sensor technology that relies on the change in the dielectric and conductivity properties of blood plasma as a function of the glucose concentration in order to track such a change.

In this framework, developing accurate and precise fitting methods for blood models, at different glucose concentrations, is essential for the development of robust electromagnetic (EM) based techniques that could be employed for non-invasive, continuous glucose monitoring. Indeed, accurate electromagnetic tissue modeling is of paramount importance since it affects the simulation stage required for sensor design [2]. Moreover, extracting a "synthetic view" (in terms of a few parameters) of the sensor response data is essential for analyzing patterns and possibly extract more information, besides resonance peak or phase shift, about glucose concentration.

In this paper, the aim is just to address the fitting problem. More in details, starting from the dielectric spectrum, which is assumed known over a certain number of frequencies, we aim at estimating the parameters of a single-pole Cole-Cole model. As is well known, this entails solving a non-linear inverse problem which here is addressed

by two different methods: the classical Levenberg-Marquardt method [3] [4] and the Variable Projection algorithm [5]. We evaluate how sensitive the two methods are with respect to the starting points of the parameters and with what accuracy and precision these parameters can be estimated. In order to check the two methods, we first generate synthetic relative permittivities by employing a single-pole Cole-Cole model, using data from the literature [6] as *true values* for its parameters, and then solve an inverse problem in order to trace these values by resorting to the two aforesaid methods.

Although higher order models can be more performing, we consider a first order model to perform the comparison in the simplest possible case.

2. Methods

2.1. Cole-Cole model

The Cole-Cole model [7] is widely used to describe the complex relative permittivity of biological tissues, $\varepsilon_{\mathrm{r}}(\omega) = \varepsilon(\omega)/\varepsilon_0$, and its equation is

$$\varepsilon_{\mathrm{r}}(\omega) = \varepsilon_\infty + \sum_{n=1}^{N} \frac{\varepsilon_{sn} - \varepsilon_\infty}{1 + (j\omega\tau_n)^{1-\alpha_n}} - \frac{\sigma}{j\omega\varepsilon_0} \tag{1}$$

in which N is the number of poles and thence the order of the model, $\varepsilon_\infty = \lim_{\omega\to\infty} \varepsilon_{\mathrm{r}}(\omega)$ is the permittivity at high frequencies, σ is the static ionic conductivity and $\varepsilon_{sn} = \lim_{\omega\to 0} \varepsilon_{\mathrm{r}}(\omega)$, τ_n and α_n are the static permittivity, the relaxation time constant and the so-called distribution parameter of the n-th addend of the summation, respectively. Such a model incorporates the Debye model [8]. Indeed, the main difference between the Debye and the Cole-Cole models is that the latter includes the exponent $1 - \alpha$, with $0 \leq \alpha \leq 1$. When the exponent becomes smaller, the relaxation time distribution becomes broader, *i.e.*, the transition between low- and high-frequency values becomes wider and the peak on imaginary part of the spectrum also becomes wider.

The complexity of both the structure and composition of biological material is such that dispersion region of each pole may be broadened by multiple contributions to it. The broadening of the dispersion could be empirically accounted for by using the Cole-Cole model [9]. It is for that reason that the Cole-Cole model is expected to give more accurate dielectric spectrum curve-fitting.

2.2. Curve fitting Algorithms

Let be $\boldsymbol{x}$ the vector of model parameters and P its length, M the number of frequency points the measures are taken. We define the data vector (^{T} stands for transposition)

$$\boldsymbol{y} = \left[y(\omega_1) \dots y(\omega_m) \dots y(\omega_M)\right]^\top \tag{2}$$

in which the mth component of the vector $\boldsymbol{y}$ is the observed value $y(\omega_m)$. Let also be

$$\boldsymbol{\varepsilon}_{\mathrm{r}} = \left[\varepsilon_{\mathrm{r}}(\omega_1;\boldsymbol{x}) \dots \varepsilon_{\mathrm{r}}(\omega_m;\boldsymbol{x}) \dots \varepsilon_{\mathrm{r}}(\omega_M;\boldsymbol{x})\right]^\top \tag{3}$$

the model vector, here given by eq. (1), with $\varepsilon_{\mathrm{r}}(\omega_m;\boldsymbol{x})$ being is the estimation at ω_m.

Solving the least squares problem means finding $\hat{\boldsymbol{x}}$ such that

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}\in\mathbb{R}^P} \left\{ \frac{1}{2} \|\boldsymbol{\varepsilon}_{\mathrm{r}}(\boldsymbol{x}) - \boldsymbol{y}\|_2^2 \right\} \tag{4}$$

in which the function to minimize, $\Psi = \frac{1}{2}\|\boldsymbol{\varepsilon}_{\mathrm{r}}(\boldsymbol{x}) - \boldsymbol{y}\|_2^2$, is the ℓ_2 quadratic norm of the misfit $\boldsymbol{r} = \boldsymbol{\varepsilon}_{\mathrm{r}}(\boldsymbol{x}) - \boldsymbol{y}$, which is a non-linear function such that $\boldsymbol{r} : \mathbb{R}^P \mapsto \mathbb{R}^M$ with $P \ll M$.

We address the non-linear fitting problem with two methods: the Levenberg-Marquardt Algorithm (LMA) and the Variable Projection Algorithm (VPA).

The Levenberg-Marquardt Algorithm [3] [4] acts more like a gradient-descent method when the parameters are far from their optimal value, and acts more like the

Gauss-Newton method when the parameters are close to their optimal value [10]. The equation for the step $\boldsymbol{h}$ at the kth iteration is

$$\Big(J(\boldsymbol{x}_k)^\top J(\boldsymbol{x}_k) + \lambda_k I\Big)\boldsymbol{h} = -J(\boldsymbol{x}_k)^\top \boldsymbol{f}(\boldsymbol{x}_k) \tag{5}$$

where J is the Jacobian of $\boldsymbol{f}$ and λ_k is the damping parameter. It controls both the magnitude and direction of $\boldsymbol{h}$ and it is chosen at each iteration. It can be shown [4] that, at each iteration, eq. (5) solves the minimization problem over a reduced set of admissible solutions, *i.e.*, those that satisfy $\|\boldsymbol{h}\| \leq R(\lambda)$, limiting the correction step to within a region near $\boldsymbol{x}_k$. The radius of the trust region $R = R(\lambda)$ is a strictly decreasing function with $\lim_{\lambda\to\infty} R(\lambda) = 0$. When $\lambda_k = 0$, the step $\boldsymbol{h}$ is identical to that of Gauss-Newton method, *i.e.*, the same direction and maximum magnitude. As $\lambda \to \infty$, $\boldsymbol{h}$ tends towards the steepest descent direction, with magnitude tending towards 0.

Based on the above, we infer the qualitative update rule for λ_{k+1}: if $\Psi(\boldsymbol{x}_k + \boldsymbol{h}) < \Psi(\boldsymbol{x}_k)$ then the quadratic approximation works well and we can extend the trust region, *i.e.* it will be $\lambda_{k+1} < \lambda_k$. Otherwise, the step is unsuccessful and we reduce the trust region, i.e. it will be $\lambda_{k+1} > \lambda_k$; in this way the next step tends towards the negative gradient method and a lower value of Ψ is more likely to be found.

The MATLAB implementation has been used, in particular the `lscurvefit` function with the Levenberg-Marquardt option [11].

The Variable Projection Algorithm [5] is a method used to solve separable nonlinear least squares problems. The least squares problem is said to be separable when the model parameters can be separated into two sets of parameters, one that enter linearly into the model, $\boldsymbol{c} = [c_1, \ldots, c_k]$, and another set of parameters that enter the model non linearly, $\boldsymbol{a} = [a_1, \ldots, a_l]$, so that $\boldsymbol{x} = [\boldsymbol{c}, \boldsymbol{a}]$. For each observation y_m of a separable nonlinear least squares problems, the model is a linear combination of nonlinear functions that depend on non linear parameters, and the model function can be written as

$$\varepsilon_{\mathrm{r}}(\omega) = \sum_{j=1}^{k} c_j\, \phi_j(\omega; \boldsymbol{a})$$

The functional Ψ is written in terms of residual vector $\boldsymbol{r}$ as

$$\Psi(\boldsymbol{a}, \boldsymbol{c}) = \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{\Phi}(\boldsymbol{a})\boldsymbol{c}\|^2 \tag{6}$$

in which the columns of the matrix $\boldsymbol{\Phi}$ are the non linear functions $\phi_j(\omega; \boldsymbol{a})$. The linear parameters $\boldsymbol{c}$ could be obtained from the knowledge of $\boldsymbol{a}$, by solving the linear least squares problem:

$$\boldsymbol{c} = \boldsymbol{\Phi}(\boldsymbol{a})^\dagger \boldsymbol{y} \tag{7}$$

which stands for the minimum-norm solution of the linear least squares problem for fixed $\boldsymbol{a}$, where $\boldsymbol{\Phi}(\boldsymbol{a})^\dagger$ is the Moore-Penrose generalized inverse of $\boldsymbol{\Phi}(\boldsymbol{a})$. By replacing this in eq. (6), we obtain the Variable Projection functional

$$\Psi_{\mathrm{VP}}(\boldsymbol{a}) = \frac{1}{2}\Big\|\boldsymbol{y} - \boldsymbol{\Phi}(\boldsymbol{a})\boldsymbol{\Phi}(\boldsymbol{a})^\dagger \boldsymbol{y}\Big\|^2 \tag{8}$$

The Variable Projection algorithm consists of two steps: first minimizing eq. (8) with an iterative non linear method and then using the optimal value found for $\boldsymbol{a}$ to solve for $\boldsymbol{c}$ in eq. (7) [12]. The principal advantage is that the iterative nonlinear algorithm used to solve the first minimization problem works in a reduced space and less initial guesses are necessary. A robust implementation in MATLAB, called `VARPRO` [13], has been adapted and used to deal with complex-value problems, choosing the Levenberg-Marquardt option for the solution of eq. (8).

2.3. Numerical Simulations

The generation of the synthetic complex relative permittivity of blood plasma relies on *in-vitro* data reported in [6], and precisely on data relating to concentrations that are more realistic from the point of view of human physiology (*i.e.*, 250 mg/dl and 500 mg/dl). The data vector consists of $M = 1000$ points in the frequency range 500MHz $-$ 20GHz.

In gradient-like algorithms (such as those used in this paper), the choice of the initial point is a crucial factor for the convergence of the procedure. For the single-pole model case, it is fairly easy to exploit the physical meaning of the parameters to infer an initial estimate. However, since the noise can invalidate the initial estimate, we propose to study the robustness of the two algorithms with respect to random initial points. To this end, we consider N=1000 uniformly distributed random initial points arranged in a 5D hypercube of the parameter space. Each side of the hypercube represents an interval containing the range of variation of each parameter for the glucose concentrations considered.

The intervals for generating the random initial value for each parameter (of the Cole-Cole model) are chosen from the data tabulated in [6]. In particular, the widths of these intervals are the same for each glucose concentration and are: $[1,5]$ for ε_∞, $[1,150]$ for ε_s, $[1\times10^{-14}, 1\times10^{-11}]$ for τ, $[0.1-1\times10^{-9}, 0.1+1\times10^{-9}]$ for α and $[0,5]$ for σ. These intervals are relatively large compared to the values taken from [6] in order to test the two algorithm in sufficiently stressful situations. Only the range of variation of α is extremely small because the model used in [6] practically fixes it a priori to 0.1. Obviously, it must be taken into account that VPA requires only to generate the values for τ and α.

For an initial bland qualitative assessment, we established evaluation intervals (the same for generations) for the estimated parameters so that we could assert that a reconstruction is "good" if it falls within these ranges, "wrong" otherwise.

Now, let $\hat{\boldsymbol{x}}^{(i)} = [\widehat{\varepsilon}_\infty^{(i)}, \widehat{\varepsilon}_s^{(i)}, \widehat{\tau}^{(i)}, \widehat{\alpha}^{(i)}, \widehat{\sigma}^{(i)}]$ be the vector of the parameter estimates returned by the two algorithms at the i-th simulation and let $\hat{x}^{(i)}$ denote one of its five elements. Moreover, let $\langle x\rangle = (1/N_{\text{sim}})\sum_{i=1}^{N_{\text{sim}}}\hat{x}_i$ and $\sigma_x = \sqrt{[1/(N_{\text{sim}}-1)]\sum_{i=1}^{N_{\text{sim}}}|\hat{x}_i - \langle x\rangle|^2}$ be the sample mean and standard deviation, respectively, calculated for each parameter.

For a quantitative evaluation of the performance of the two algorithms, we then define multiple figures of merit for characterising the results. For each parameter, eqs. (9a) and (9b) define measures of accuracy and precision, respectively, defined over the entire set of reconstructions. However, such measures can be greatly affected by estimates that are very far from the true value, x_{true}, which the latter represents one of the five elements of the vector of reference values $\hat{\boldsymbol{x}}_{true} = [\widehat{\varepsilon}_{\infty_{true}}, \widehat{\varepsilon}_{s_{true}}, \widehat{\tau}_{true}, \widehat{\alpha}_{true}, \widehat{\sigma}_{true}]$. For this reason, we also introduce eqs. (10a) and (10b) in order to define accuracy and precision measures, respectively, that instead dampen the effect of the above isolated events. They are calculated on a subset obtained by eliminating the ζ% of reconstructions with lower values and ζ% of reconstructions with higher values, in which $0 < \zeta < 50$.

$$\mathcal{A} = \left|\frac{x_{\text{true}} - \langle x\rangle}{x_{\text{true}}}\right| \times 100\% \tag{9a}$$

$$\mathcal{P} = \frac{\sigma_x}{\langle x\rangle} \times 100\% \tag{9b}$$

$$\mathcal{A}_{cut} = \left|\frac{x_{\text{true}} - \langle x\rangle_{cut}}{x_{\text{true}}}\right| \times 100\% \tag{10a}$$

$$\mathcal{P}_{cut} = \frac{\sigma_{x_{cut}}}{\langle x\rangle_{cut}} \times 100\% \tag{10b}$$

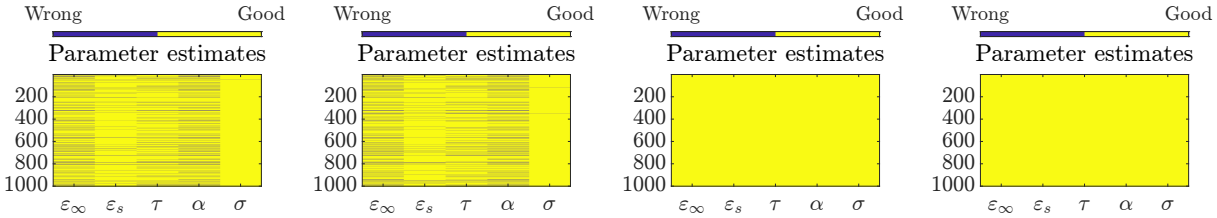

Figure 1. Graphical representations of convergence of LMA and VPA for 1000 simulations and 2 different glucose concentrations: (**a**) LMA for 250 mg/dl, (**b**) LMA for 500 mg/dl, (**c**) VPA for 250 mg/dl, (**d**) VPA for 500 mg/dl;

Table 1. Figures of merit (in %) of the two algorithm in which the glucose concentration is 250 mg/dl.

		ε_∞	ε_s	τ [s]	α	$\sigma\,[S/m]$
	$\boldsymbol{x}_{\text{true}}$	2.04	7.21×10^{1}	8.62×10^{-12}	0.1	1.96
LMA	$\mathcal{A}$	2.32×10^{7}	5.91×10^{7}	3.76×10^{8}	1.51×10^{8}	7.32
LMA	$\mathcal{A}_{cut,25\%}$	3.21×10^{-7}	3.65×10^{-10}	2.60×10^{-9}	5.25×10^{-9}	8.61×10^{-10}
LMA	$\mathcal{P}$	2.79×10^{3}	3.16×10^{3}	3.29×10^{3}	3.79×10^{3}	1.62×10^{2}
LMA	$\mathcal{P}_{cut,25\%}$	8.37×10^{-7}	9.60×10^{-10}	6.75×10^{-9}	1.24×10^{-8}	2.27×10^{-9}
VPA	$\mathcal{A}$	1.64×10^{-7}	4.48×10^{-10}	7.03×10^{-9}	2.39×10^{-8}	5.42×10^{-10}
VPA	$\mathcal{P}$	4.03×10^{-7}	1.11×10^{-9}	1.72×10^{-8}	5.86×10^{-8}	1.34×10^{-9}

Table 2. Figures of merit (in %) of the two algorithm in which the glucose concentration is 500 mg/dl.

		ε_∞	ε_s	τ [s]	α	$\sigma\,[S/m]$
	$\boldsymbol{x}_{\text{true}}$	2.67	73.1	8.88×10^{-12}	0.1	1.93
LMA	$\mathcal{A}$	2.33×10^{4}	3.35×10^{4}	2.43×10^{6}	2.07×10^{6}	2.09
LMA	$\mathcal{A}_{cut,25\%}$	3.11×10^{-7}	3.84×10^{-10}	1.65×10^{-9}	5.85×10^{-9}	1.08×10^{-9}
LMA	$\mathcal{P}$	1.52×10^{4}	1.87×10^{3}	3.16×10^{3}	2.66×10^{3}	7.19
LMA	$\mathcal{P}_{cut,25\%}$	7.98×10^{-7}	1.15×10^{-9}	4.47×10^{-9}	1.59×10^{-8}	2.90×10^{-9}
VPA	$\mathcal{A}$	1.09×10^{-7}	4.02×10^{-10}	6.04×10^{-9}	2.11×10^{-8}	4.98×10^{-10}
VPA	$\mathcal{P}$	3.32×10^{-7}	1.28×10^{-9}	1.83×10^{-8}	6.53×10^{-8}	1.57×10^{-9}

3. Results

We have conducted many numerical simulations by widening more and more the generation intervals. In this paper we report the case where the generation intervals are very large except for the α interval, due to the above explanation. In all these experiments, following the qualitative criterion mentioned above, VPA has always provided good estimations while the same is not true for LMA. In particular, in the case considered, LMA provided wrong estimates in about 270 simulations out of 1000, for each of the two glucose concentrations, while no wrong estimate was returned by the VPA. Here, by "wrong" estimate we mean that at least one component of the parameters vector $\boldsymbol{x}$ has a value that is outside its generation range. Graphical representations of those qualitative results are provided in Figure 1.

Consistent with the qualitative results, considering the whole set of 1000 estimates, VPA exhibits excellent accuracy and precision, while this is not the case for LMA, as can be observed in in Table 1 and Table 2. For this reason, for the VPA only the results calculated by means of the eqs. (9a) and (9b) are reported, whilst for the LMA the results deriving from eqs. (10a) and (10b), obtained by cutting 25% of the lowest values and 25% of the highest values, are also considered.

4. Discussion

In this paper we faced the problem of fitting the dielectric spectrum of blood sample in order to estimate the parameters of the single-pole Cole-Cole model. In particular, we compared the performance of two different algorithms, LMA and VPA, in terms of accuracy and precision with respect to the starting points of the parameters.

For the parameter range considered, VPA outperforms LMA in robustness with respect to the initial point of the algorithm. However, analyzing the figures of merit related to LMA, it becomes clear that there are erroneous reconstructions so far from the true value such that they heavily deteriorate the (standard) accuracy and precision while the cut versions do not suffer from this problem. In fact, once the erroneous ones are removed, in all other simulations the algorithm converges to the true values.

On the other hand, VPA gives these good results because less initial guesses are necessary and because the iterative nonlinear algorithm used to solve the first minimization problem works in a reduced space. The big disadvantage of VPA, or at least of the implementation used in this paper, is the execution time. In our test we took advantage of the Parallel Computing Toolbox, using `parfor` loop for running the 1000 simulations. LMA took 7 seconds to finish them while VPA took around 170 seconds on a machine with Intel i9-10850K (10 physical cores), 32GB RAM and Ubuntu 21.04. This is certainly due to the numerous SVDs that the algorithm calculates in its runtime.

The results are promising and the research will continue by evaluating the algorithms in increasingly realistic scenarios, including adding noise on the synthetic data.

Author Contributions: Conceptualization, R.S.; methodology, R.D. and G.B.; software, R.D.; validation, R.D. and G.B.; formal analysis, R.S., R.D. and G.B.; investigation, R.D. and G.B.; resources, R.S.; data curation, R.D.; writing—original draft preparation, R.D.; writing—review and editing, G.B. and R.D.; visualization, R.D. and G.B.; supervision, R.S. and G.B.; project administration, R.S.; funding acquisition, R.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by MIUR, project grant PRIN 2017 "Microwave Biosensors: Enhanced Non-Invasive Methodology for Blood Glucose Monitoring". The APC was funded by the above project grant PRIN 2017.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lin, T. Non-Invasive Glucose Monitoring: A Review of Challenges and Recent Advances. *6*. doi:10.19080/CTBEB.2017.06.555696.
2. Costanzo, S.; Cioffi, V. Dielectric Models for the Accurate Design of Wearable Diabetes Sensors. 2019 23rd International Conference on Applied Electromagnetics and Communications (ICECOM). IEEE, pp. 1–3. doi:10.1109/ICECOM48045.2019.9163626.
3. Levenberg, K. A Method for the Solution of Certain Non-Linear Problems in Least Squares. *2*, 164–168.
4. Marquardt, D.W. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *11*, 431–441. doi:10.1137/0111030.
5. Golub, G.H.; Pereyra, V. The Differentiation of Pseudo-Inverses and Nonlinear Least Squares Problems Whose Variables Separate. *10*, 413–432. doi:10.1137/0710036.
6. Topsakal, E.; Karacolak, T.; Moreland, E.C. Glucose-Dependent Dielectric Properties of Blood Plasma. 2011 XXXth URSI General Assembly and Scientific Symposium. IEEE, pp. 1–4. doi:10.1109/URSIGASS.2011.6051324.
7. Cole, K.S.; Cole, R.H. Dispersion and Absorption in Dielectrics I. Alternating Current Characteristics. *9*, 341–351. doi:10.1063/1.1750906.
8. Debye, P. *Polar Molecules*; The Chemical Catalog Company, Inc.
9. Gabriel, S.; Lau, R.W.; Gabriel, C. The Dielectric Properties of Biological Tissues: III. Parametric Models for the Dielectric Spectrum of Tissues. *41*, 2271–2293. doi:10.1088/0031-9155/41/11/003.
10. Gavin, H.P. The Levenberg-Marquardt Algorithm for Nonlinear Least Squares Curve-Fitting Problems. p. 19.
11. Matlab. *Optimization Toolbox*.
12. Golub, G.; Pereyra, V. Separable Nonlinear Least Squares: The Variable Projection Method and Its Applications. *19*, R1–R26. doi:10.1088/0266-5611/19/2/201.
13. O'Leary, D.P.; Rust, B.W. Variable Projection for Nonlinear Least Squares Problems. *54*, 579–593. doi:10.1007/s10589-012-9492-9.