DETERMINATION OF PARAMETERS IN CHEMICAL AND BIOCHEMICAL NON-LINEAR MODELS USING SIMULATED DATA WITH GAUSSIAN NOISE PERTURBATION

Victor MARTINEZ-LUACES[1]
Electrochemistry Engineering Multidisciplinary Research Group. UdelaR. Uruguay

E-mail: victorml@fing.edu.uy

Leticia SILVA[2]
Antel. Uruguay

Abstract

The Michaelis-Menten kinetics is a well-known model in biochemistry, widely used in enzyme-substrate interaction (Nelson and Cox, 2008). The same mathematical formula is called Langmuir equation (Masel, 1996) when is used to model generic adsorption of chemical species, and finally, an empirical equation of this form is applied to microbial growth and it is called J. Monod kinetics (Martinez-Luaces, 2008).

A typical problem in chemistry and/or biochemistry consists in determining the parameters of these equations from experimental data. In order to solve this problem, several methods were proposed, Lineweaver-Burk, Hanes-Woolf, Hofstee, Scatchard and Cornish-Bowden-Eisenthal are the most important ones (Nelson and Cox, 2008).

In this paper, all these methods are analysed and compared in terms of exactitude and precision. For this purpose, simulated data were generated and perturbed using Gaussian noise with different amplitudes. The same methodology was used in a previous work (Martinez-Luaces, 2009).

Absolute and relative errors of the different methods are compared, and taking into account the results, general conclusions about their robustness are obtained. This is particularly important in order to choose the best method when the relation between trend and noise tends to increase.

Keywords: *Chemical and biochemical models, Data simulation, Gaussian noise.*

Introduction

The non-linear mathematical formula $y = \frac{a\,x}{x+b}$ (Eq. 1) is widely used in chemistry and biochemistry for different purposes. For instance, the Michaelis-Menten kinetics is a well-known model in biochemistry of the form $v_0 = \frac{v_{\max}[S]}{K_m + [S]}$ (Eq. 2) where v_0 and $v_{\max}$ are the initial and the maximum velocity of the enzymatic reaction, $[S]$ is the substrate concentration and k_m is a constant (called Michaelis constant), which depends on the enzymatic reaction considered (Nelson and Cox, 2008).

Irving Langmuir, a Nobel Prize winner in chemistry, developed an equation that relates the coverage or adsorption of molecules on a solid surface to gas pressure or concentration of a medium above the solid surface at fixed temperature (Masel, 1996). The equation is $\theta = \frac{\alpha P}{1+\alpha P}$ (Eq. 3), where θ is the fractional coverage of the surface, P is the gas pressure (or concentration in the case of liquids) and α is a constant. A very simple algebraic manipulation gives $\theta = \frac{P}{\frac{1}{\alpha} + P}$ (Eq. 4) which is just a particular case of (Eq. 1).

The last example is a mathematical model for the growth of microorganisms proposed by Jacques Monod in 1949. The mathematical formula is $\mu = \frac{\mu_{\max} S}{K_S + S}$ (Eq. 5) where μ is the specific growth rate of microorganisms and $\mu_{\max}$ represents its maximum value, S is the concentration of the limiting substrate for growth and, K_S is called the "half-velocity constant" (Martinez-Luaces, 2008 and Martinez-Luaces, 2009) since it corresponds to the value of S when $\frac{\mu}{\mu_{\max}} = \frac{1}{2}$ as well as the constant K_m in (Eq. 2).

The Monod equation has the same form as the Michaelis-Menten equation, but it was developed empirically whereas the Michaelis-Menten model is based on theoretical considerations.

A typical problem that arises in the treatment of data corresponding to these equations is the parameters determination since all of them are non-linear models. In order to solve this problem, several methods were proposed to linearize these equations, Lineweaver-Burk, Hanes-Woolf, Eadie-Hofstee, Scatchard, and Eisenthal and Cornish-Bowden are the most important ones.

In this paper, all these methods will be compared in terms of exactitude and precision, using simulated data perturbed with Gaussian noise with different amplitudes. The details of this procedure will be described in the following section.

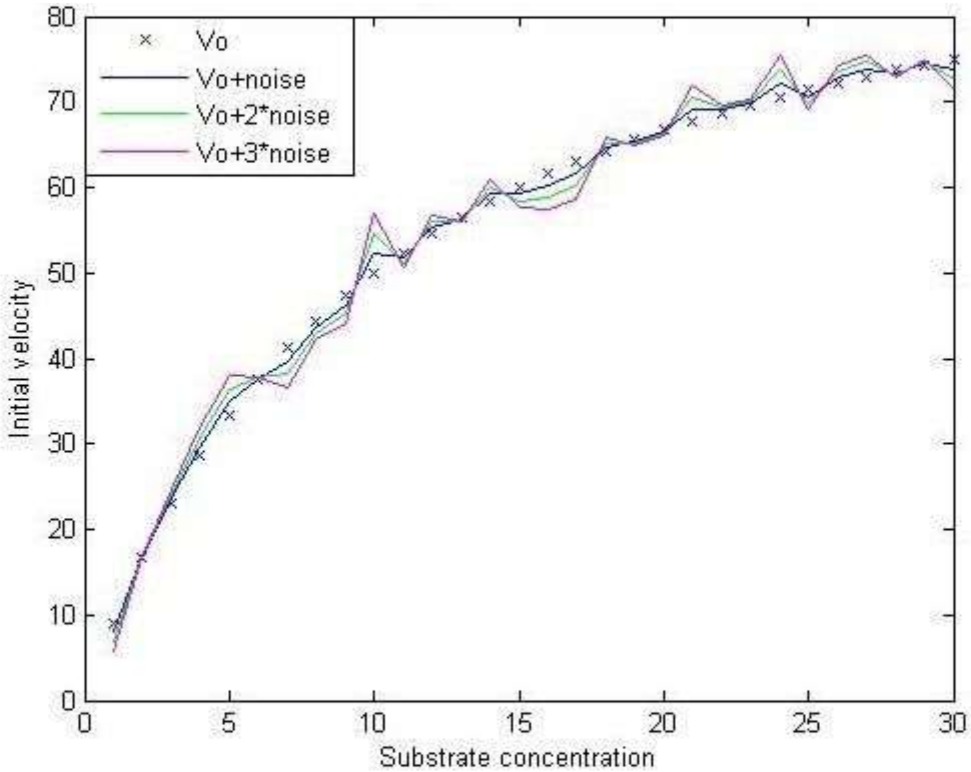
Data Simulation

Since equations (2), (3) and (4) represent the same mathematical model (Eq 1), we choose one of them (the Michaelis-Menten equation) to show the methodology to be followed, so the parameters will be K_M and $V_{\max}$ and variables will be [S] and v_0. For simplicity the notation in this paper will be K_M, V_M, S and V_0, respectively.

Different enzymes have different K_M values. They typically range from 10^{-1} to $10^{-7} M$, so $K_M = 10$ can be considered as a typical value depending on the units employed (for example $10\ mM = 10^{-2} M$ or even $10\ \mu M = 10^{-5} M$ are both possible values, although they are very different). On the other hand V_M , the maximum velocity depends on a constant named K_{Cat}, and $[E_t]$ the total enzyme concentration , like in $V_M = K_{Cat}[E_t]$ (Eq. 6). The constant $K_{C\,a}$ can vary between $0.5\left(s^{-1}\right)$ and $40000000\ \left(s^{-1}\right)$ so, once again, it is difficult to propose a "typical value" for V_M . Thus, we decided to consider $V_M = 100$ which may be taken as a typical value if units are $\mu M /_{\min}$.

Then, a Gaussian noise with different amplitudes was superimposed to the theoretical data obtained from (Eq 2) with $V_M = 100$, $K_M = 10$ and S varying from 0 to 30. The graphics in Figure 1 show the simulated curves with the Gaussian noise multiplied by 2, 3 and 4.

Figure 1. Initial velocity vs Substrate concentration. Simulated curves.



These simulated data will take the place of the real experimental data and they will be used to determine the parameters K_M and V_M ,which real values are known, so the different methods could be compared in terms of exactitude and precision.

A similar methodology was followed in a previous paper (Martinez-Luaces et al., 2006) for Electrochemical Noise studies. More recently, in a mathematical modelling paper (Martinez-Luaces, 2015) this methodology was utilized for educational purposes.

Methods for obtaining the parameters in Michaelis-Menten equation

Several methods were proposed for linearizing the Michaelis-Menten equation. Perhaps the simplest one is the Lineweaver-Burk (or double reciprocal plot), which is a graphical representation of $\frac{1}{V_0}$ vs. $\frac{1}{S}$ (Nelson and Cox, 2008). It is easy to observe that the reciprocal of (Eq 2) gives $\frac{1}{V_0} = \frac{K_M}{V_M}\frac{1}{S} + \frac{1}{V_M}$ (Eq.7) so, the $x-$ intercept of the graph represents $-\frac{1}{K_M}$ and the $y-$ intercept is equivalent to the inverse of V_M. An alternative way is to obtain the coefficients of a linear regression (i.e. $\frac{K_M}{V_M}$ and $\frac{1}{V_M}$) and finally get the K_M and V_M.

The obtained results are summarized in Table 1, with the corresponding absolute and relative errors.

Table 1. Absolute and relative errors for the Lineweaver-Burk method

Noise	**Lineweaver-Burk Method**					
	Km	**Vm**	**Absolute error Km***	**Absolute error Vm***	**Relative error Km***	**Relative error Vm***
Amplitud Noise 1	12.5936237	112.8575325	2.5936	12.8575	0.2594	0.1286
Amplitud Noise 2	17.1811787	135.3073339	7.1812	35.3073	0.7181	0.3531
Amplitud Noise 3	27.4897736	185.3756073	17.4898	85.3756	1.7490	0.8538

A second methodology was posed by Hanes and Woolf. These researchers proposed to plot $\frac{S}{V_0}$ against S, since a rearrangement of (Eq. 1) gives $\frac{S}{V_0} = \frac{S}{V_M} + \frac{K_M}{V_M}$ (Eq. 8). Once again, a linear regression gives the coefficients $\frac{1}{V_M}$ and $\frac{K_M}{V_M}$ and lastly they can be used straightforward to obtain V_M and K_M.

The results of this method can be observed in Table 2

Table 2. Results corresponding to Hanes-Woolf method

Noise	**Hanes-Woolf Method**					
	Km	**Vm**	**Absolute error Km***	**Absolute error Vm***	**Relative error Km***	**Relative error Vm***
Amplitud Noise 1	10.1840993	100.685067	0.1841	0.6851	0.0184	0.0069
Amplitud Noise 2	10.4974793	101.751327	0.4975	1.7513	0.0497	0.0175
Amplitud Noise 3	11.0019191	103.400162	1.0019	3.4002	0.1002	0.0340

A third method is due to Eadie and Hofstee. They inverted (Eq. 2) and multiplied by V_M obtaining $\frac{V_M}{V_0}=\frac{K_M+S}{S}$ (Eq. 9) and a rearrange gives $V_M=\frac{K_M V_0}{S}+V_0$ (Eq. 10) or $V_0=-K_M\frac{V_0}{S}+V_M$ (Eq. 11).

A plot of V_0 against $\frac{V_0}{S}$ will yield V_M as the $y-$intercept and $-K_M$ as the slope of the straight line. Alternatively, a linear regression will give coefficients $-K_M$ and V_M from where the parameters are easily obtained as it is showed in Table 3 with the corresponding absolute and relative errors.

Table 3. Parameters and errors corresponding to Eadie-Hofstee method

Noise	Eadie and Hofstee Method					
	Km	Vm	Absolute error Km*	Absolute error Vm*	Relative error Km*	Relative error Vm*
Amplitud Noise 1	10.16691318	100.637843	0.1669	0.6378	0.0167	0.0064
Amplitud Noise 2	9.911745044	99.38537855	0.0883	0.6146	0.0088	0.0061
Amplitud Noise 3	9.231519446	96.2503578	0.7685	3.7496	0.0768	0.0375

The Scatchard plot can be obtained from (Eq 11), that can be multiplied by $-\frac{1}{K_M}$ to give: $-\frac{1}{K_M}V_0=\frac{V_0}{S}-\frac{V_M}{K_M}$ (Eq. 12) or $\frac{V_0}{S}=-\frac{1}{K_M}V_0+\frac{V_M}{K_M}$ (Eq. 13). Once again, a linear regression will give a slope $-\frac{1}{K_M}$ and a $y-$intercept $\frac{V_M}{K_M}$.

As in other methods K_M and V_M can be reached from these coefficients and compared with the theoretical values $K_M=10$ and $V_M=100$. Table 4 shows the results.

Table 4. Results corresponding to Scatchard method

Noise	Scatchard Method					
	Km	Vm	Absolute error Km*	Absolute error Vm*	Relative error Km*	Relative error Vm*
Amplitud Noise 1	10.4434129	101.877116	0.4434	1.8771	0.0443	0.0188
Amplitud Noise 2	11.0831493	104.616199	1.0831	4.6162	0.1083	0.0462
Amplitud Noise 3	11.9679929	108.424515	1.9680	8.4245	0.1968	0.0842

The final methodology to be discussed here is the Eisenthal and Cornish-Bowden direct linear plot. In this original approach (Eq. 4) is rearranged to give $\frac{V_M}{V_0}-\frac{K_M}{S}=1$ (Eq. 14) so if V_M is plotted against K_M a straight line is obtained and the $x-$intercept is $-S$ while the $y-$intercept is V_0 .

Then, for each observation (S,V_0) a straight line is obtained. Theoretically all these lines will intersect at a common point, whose co-ordinates (K_M,V_M) provide the values of the parameters. When the observations are subject to error there will be $\binom{n}{2}=\frac{1}{2}n(n-1)$ (Eq. 15) intersections.

Each intersection provides an estimate of K_M and an estimate of V_M, then the corresponding medians will give the best estimate for K_M and V_M, respectively. The results of this procedure are given in Table 5.

Table 5. Results and errors corresponding to Cornish-Bowden and Eisenthal method

Noise	Cornish-Bowden & Eisenthal Method					
	Km	Vm	Absolute error Km*	Absolute error Vm*	Relative error Km*	Relative error Vm*
Amplitud Noise 1	10.0167	100.3243	0.0167	0.3243	0.0017	0.0032
Amplitud Noise 2	9.8727	99.7177	0.1273	0.2823	0.0127	0.0028
Amplitud Noise 3	9.6059	98.5973	0.3941	1.4027	0.0394	0.0140

Results

The results of five different methods were showed in Tables 1-5 in the previous section. In order to compare all these results, Table 6 shows the minimum absolute and relative errors in K_M and V_M and which methodology was the best in each case, depending on the noise amplitude.

Table 6. Comparison of the results for the different methods

Noise	Min. abs. error Km	Min. abs. error Vm	Min. rel. error Km	Min. rel. error Vm	Min. abs. Error Method Km	Min. abs. Error Method Vm	Min. rel. Error Method Km	Min. rel. Error Method Vm
Amplitud Noise 1	0.0167	0.3243	0.0017	0.0032	CBE	CBE	CBE	CBE
Amplitud Noise 2	0.0883	0.2823	0.0088	0.0028	H	CBE	H	CBE
Amplitud Noise 3	0.3941	1.4027	0.0394	0.0140	CBE	CBE	CBE	CBE

The method due to Eisenthal and Cornish-Bowden was the best one in all cases, except when the Gaussian noise had double amplitude and K_M is the considered parameter. In this last case, Eadie and Hofstee's method obtained the minimum absolute and relative error in the parameter K_M, but not in V_M where once again Eisenthal and Cornish-Bowden gave the best estimate.

This last method does not seem to be as simple as the others and the number of intersections grows quadratically – since $\binom{n}{2}=\frac{1}{2}n(n-1)=O(n^2)$ – demanding more computation time than other simpler methodologies. On the other hand, it gives the best results in all the cases except one, for both parameters V_M and K_M.

Conclusions

In the previous sections several data for initial velocity and substrate concentration were simulated. For this purpose, typical values of the parameters K_M and V_M were proposed and the theoretical values obtained were perturbed with Gaussian noise of different amplitudes. These simulated values were used to check five different linearization methods for the Michaelis-Menten equation. The best results were obtained by the methodology proposed by Eisenthal and Cornish-Bowden.

This result can be explained because the original data (S,V_0) are not transformed like in other methods where reciprocal quantities, product, etc., are performed before plotting the data, diminishing the possibility of errors propagation. Moreover, taking medians of the intersection points co-ordinates may give more robustness to this method, since medians are not sensible to extreme values (Janke and Tinsley, 2005). The same situation happens with outliers, because the mean can be completely upset by a single outlier, while the sample median is little affected for these values that are absolutely different in relation to the majority of the sample.

Another aspect that must be considered is that several methods with a poor performance like Lineweaver-Burk may be useful for other purposes. For instance, Lineweaver-Burk's plots allow the researcher to know if there is an inhibitor and if it is competitive or uncompetitive (Nelson and Cox, 2008). This possibility was analyzed in detail in a paper written by Dixon (1953).

A different approach was proposed 20 years later by A. Cornish-Bowden who provided a simple way of determining the inhibition constant of an uncompetitive, mixed or non-competitive inhibitor (Cornish-Bowden, 1974).

Taking into account the previous comments, the errors (relative and absolute) of the different methods are an important aspect to be considered, but not the unique one, particularly if inhibition is taking place. As a consequence, the final election of the methodology to be used will depend on the objectives of the research project.

References

1. Cornish-Bowden, A., **A simple graphical method for determining the inhibition constants of mixed, uncompetitive and non-competitive inhibitors**, Biochemical Journal, Volume 137, No. 1, 1974, pp. 143-144.
2. Dixon, M., **The determination of enzyme inhibitor constants**, Biochemical Journal, Vol. 55, No. 1, 1953, pp. 170-171.
3. Janke, S. J. and Tinsley, F. **Introduction to Linear Models and Statistical Inference**. New York: John Wiley and Sons. 2005
4. Martinez-Luaces, M., Martinez-Luaces, V. & Ohanian, M., **Trend-removal with Neural Networks: data simulation, preprocessing and different training algorithms applied to Electrochemical Noise studies**, WSEAS Transactions on information science and applications. Vol. 3, Issue 4, 2006, pp. 810-817.
5. Martinez-Luaces, V., **Sistemas E.D.O. lineales y no lineales: aplicación a un problema de Bioingeniería**, in Cuadrado, G. (ed.) "Educación Matemática en Carreras de Ingeniería", Mendoza, Argentina: FRM-UTN, 2008.

6. Martinez-Luaces, V., **Aplicaciones y Modelado: Ecuaciones Diferenciales Ordinarias, Transformación de Laplace, Ecuaciones en Derivadas Parciales**, Ed. ISC-Delta, Montevideo, Uruguay. 2009.
7. Martinez-Luaces, V., **Mathematical models for Chemistry and Biochemistry service courses**. Proceedings of the ninth Congress of the European Society for Research in Mathematics Education, Prague: CERME, 2015, pp. 904-909
8. Masel, R., **Principles of adsorption and reaction on solid surfaces**, Wiley Interscience, New York. 1996.
9. Nelson, D.L. and Cox, M.M., **Lehninger principles of biochemistry**, 5th Edition, W.H. Freeman & Company, New York. 2008

[1] **Victor Martinez-Luaces** has obtained three degrees from the State University of Uruguay (UdelaR), in Chemistry, Chemical Engineering and Mathematics and he has worked in that university for more than 30 years. He published more than ten books and book chapters in Croatia, Germany, U.S.A. and Uruguay and his papers have been published worldwide in the field of Mathematics, Chemistry, Engineering and Education. He organised Delta '07 in Argentina and he was co-chair of the Topic Study Group about Teaching and Learning of Calculus in Korea (ICME-12) and Germany (ICME-13).

[2] **Leticia Silva** has obtained a degree in Electrical Engineering, specialised in Telecommunications, in the State University of Uruguay (UdelaR). She also completed a course at Universidade Federal de Paraíba, Brazil. She is attending a Master in Technology Management, in UdelaR. At present, she works for ANTEL (Uruguayan Government-owned Telecommunication Company) in Mobile Network Planning Area.