Accepted version of Electrochim. Acta, 317 (2019) 648-653. doi: 10.1016/j.electacta.2019.06.030

The Akaike information criterion in weighted regression of immittance data

Mats Ingdal[a], Roy Johnsen[a], David A. Harrington[*,b]

[a] *Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, Trondheim, 7491, Norway.*
[b] *Department of Chemistry, University of Victoria, Victoria, British Columbia, V8W 2Y2, Canada.*

Abstract

The Akaike Information Criterion (AIC) is a powerful way to distinguish between models. It considers both the goodness-of-fit and the number of parameters in the model, but has been little used for immittance data. Application in the case of weighted complex nonlinear least squares regression, as typically used in fitting impedance or admittance data, is considered here. AIC can be used to compare different weighting schemes as well as different models. These ideas are tested for simulated and real transadmittance data for hydrogen diffusion through an iron foil in a Devanathan-Stachurski cell.

Key words: Akaike Information Criterion, weighted regression, error structure, impedance, maximum likelihood

Introduction

The Akaike Information Criterion (AIC) was developed by Akaike in 1973 [1] and is widely used in modeling of biological systems to decide between different models [2]. For models with the same number of adjustable parameters, it selects the model with the best least-squares fit. Models with more parameters typically fit better, but addition of more parameters may not be statistically justifiable. The common way to prevent this overfitting is the F-test, in which one tests whether the better goodness of fit as measured by decrease in χ^2 of the fit is significant (caused by the added parameters) or could arise by chance at a particular confidence level. This has been used for impedance data for a long time, e.g., [3], but the choice of a confidence level is somewhat arbitrary, and it is only applicable for nested models, in which a parameter or two are added while retaining the original model parameters. The AIC value, Eq. (1), disfavours

*Corresponding author. Tel.: +1-250-7217166
Email addresses: mats.ingdal@ntnu.no (Mats Ingdal), roy.johnsen@ntnu.no (Roy Johnsen), dharr@uvic.ca (David A. Harrington)

additional parameters in a simple way that does not require a confidence level and statistical test, and may be applied to unrelated (non-nested) models.

$$AIC = -2 \ln \widehat{L} + 2k \tag{1}$$

Specifically, the first term involves the likelihood estimator $\widehat{L}$ which for normally distributed errors is maximized when the residual sum of squares is minimized, so a better fitting model makes the first term more negative. The second term is twice the number of parameters and a simpler model (smaller number of parameters) makes AIC more negative. Therefore the best model in terms of simplicity and goodness-of-fit is the one with the minimum AIC. Overfitting makes the first term more negative but the second more positive, and leads to a compromise between these two effects.

The biological literature typically uses unweighted data, and consideration of weights in the calculation of AIC appears to only include the simple case of weighting to account for the number of replicated data points. Weighting to account for heteroscedastic data, where different points have different variances in their stochastic noise, requires either replicate data to estimate the variances or a postulated functional form for the variance. There are several methods that iteratively reweight during the minimization process, such as iteratively reweighted least squares [4] or the inside-variance estimation method [5], but these require replicate data, which is often not available. In the area of electrochemical impedance spectroscopy (EIS), immittance (impedance or admittance) data and their variances often vary over several orders of magnitude, and weighting is esential. There are several weighting schemes typically used that will be discussed here, though a more detailed determination of the error structure through replicate measurements should really be considered [6].

In the EIS literature, Ciucci has recently advocated for AIC [7] and Orazam and Tribollet mention it in their book [8], but no examples of its use in EIS are given. However, Ritzberger et al. [9] used it with an online time series method of estimating electrochemical impedance of PEM fuel cells; Mu et al. [10] used it in a wavelet transform method for Li-ion batteries, and Qui and Orazem [11] used it with (non-impedance) cathodic protection data. In all three cases, the complication of heteroscedasticity did not arise, though the latter two papers mentioned weighted least squares for fitting (weighting unspecified). To our knowledge, AIC has not been used with weighted complex nonlinear least squares (CNLS) fitting of immittance data.

1. Theory

The derivation of least squares from the maximum likelihood principle is well known for the unweighted case, but is briefly given here as a preamble for the weighted case. The notation is general and will be specialized to the case of immittance data later where necessary. A single measurement of y_i, e.g., the real part of admittance, at a particular value of the independent variable x_i, e.g., a particular frequency, is considered a sample from a normally distributed random

variable with (true) mean μ_i and variance σ_i^2. For the perfect measurement $y_i = \mu_i$ and the probability density, Eq. (2), is at a maximum.

$$\rho_i = \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left(-\frac{(y_i - \mu_i)^2}{2\sigma_i^2}\right) \tag{2}$$

$$L(\boldsymbol{\theta}|\boldsymbol{x}, \boldsymbol{y}, g) = \prod_{i=1}^{n} \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left(-\frac{(y_i - g(x_i;\boldsymbol{\theta}))^2}{2\sigma_i^2}\right) \tag{3}$$

In practice, we have a model $g(x_i;\boldsymbol{\theta})$ for the true value μ_i, and for given (non-ideal) data we adjust the vector $\boldsymbol{\theta}$ containing p model parameters to maximuize the probability density, which for given data and model as a function of the parameters is known as the likelihood. For many measurements at different values of the independent variables x_i, the joint likelihood is the product of the probability densities ρ_i under the assumption that the measurements at different x_i are uncorrelated. This gives the overall likelihood of Eq. (3), where n is the number of data points (twice the number of frequencies). For impedance data, there is empirical evidence that measurements at different frequencies have uncorrelated errors, as do measurements of real and imaginary parts at the same frequency [6].

The parameters that maximize the likelihood need to be found. It is easier to maximize the log likelihood, or minimize its negative; the factor of 2 is then introduced in Eq. (4) in anticipation of the first term of AIC. If we knew the true variances σ_i^2 for a given data set, then the first two terms would be known and fixed, and to maximize the likelihood, we would minimize the last term. That is, this equation shows that least squares minimization of a weighted sum of squares with weights equal to the reciprocal of the true variances maximizes the likelihood.

$$-2\ln L = n\ln(2\pi) + \sum_{i=1}^{n} \ln\sigma_i^2 + \sum_{i=1}^{n} \frac{(y_i - g(x_i;\boldsymbol{\theta}))^2}{\sigma_i^2} \tag{4}$$

In practice, the variances are unknown and need to be estimated. In weighted least squares, the objective function that is minimized is not the last term above, but a weighted residual sum of squares, Eq. (5). The choice of weights w_i implies a model for the variances that needs to be considered in estimating the likelihood in AIC. There are several common weighting schemes that imply models for the error stracture, i.e., models for how σ_i^2 varies across the data.

$$S = \sum_{i=1}^{n} w_i \left(y_i - g(x_i;\boldsymbol{\theta})\right)^2 \tag{5}$$

1.1. Unit weighting

The most common error structure model is to assume that the variances are the same for all data points, i.e., all the σ_i take a common but unknown value σ. This leads to the unweighted (U) case in which all points are weighted

equally, $w_i = 1$. The last term of Eq. (4) is therefore estimated as S/σ^2 leading to Eq. (6). This is minimized by finding when all the derivatives with respect to the parameters are zero. For the model parameters $\boldsymbol{\theta}$ this leads to the so-called normal equations with solutions that give least-squares estimates of the parameters $\widehat{\boldsymbol{\theta}}$ in the usual way. However, σ is also a parameter for which the likelihood needs to be maximized. The same differentiation procedure yields an estimate $\widehat{\sigma}^2$ for the variance (Eqs. (7)-(8)). This maximum likelihood estimate $\widehat{\sigma}^2$ is substituted into the log likelihood expression and then the well-known result Eq. (10) is found for AIC. The number of parameters k is the number of model parameters p plus one for the σ parameter.

$$-2\ln L = n\ln(2\pi) + n\ln\sigma^2 + S/\sigma^2 \tag{6}$$

$$\partial(-2\ln \mathfrak{L})/\partial(\sigma^2) = n/\sigma^2 - S/\left(\sigma^2\right)^2 = 0 \tag{7}$$

$$\widehat{\sigma}^2 = S/n \tag{8}$$

$$-2\ln \widehat{L} = n\ln(2\pi) + n\ln(S/n) + n \tag{9}$$

$$AIC = n\ln(2\pi) + n\ln S - n\ln n + n + 2(p+1) \tag{10}$$

1.2. Proportional weighting

In proportional experimental (PE) weighting, the weights are chosen by $w_i = 1/y_i^2$. The implied model of the error structure is that the relative errors are constant: the σ_i are a constant fraction α of the true measured quantity and an estimate of σ_i^2 is given by $\widehat{\sigma}_i^2 = \alpha^2 y_i^2 = \alpha^2/w_i$. Therefore Eq. (4) becomes Eq. (11) An estimate of the error structure parameter α is again found by differentiation and used to find AIC, Eqs. (13)-(15).

$$-2\ln L = n\ln(2\pi) + \sum_{i=1}^{n} \ln \alpha^2/w_i + \sum_{i=1}^{n} \frac{(y_i - g(x_i;\boldsymbol{\theta}))^2}{\alpha^2/w_i} \tag{11}$$

$$= n\ln(2\pi) + n\ln\alpha^2 - \sum_{i=1}^{n} \ln w_i + S/\alpha^2 \tag{12}$$

$$\partial\left(-2\ln L\right)/\partial(\alpha^2) = n/\alpha^2 - S/\left(\alpha^2\right)^2 = 0 \tag{13}$$

$$\widehat{\alpha}^2 = S/n \tag{14}$$

$$AIC = n\ln(2\pi) + n\ln S - n\ln n + n - \sum_{i=1}^{n} \ln w_i + 2(p+1) \tag{15}$$

A variation of proportional weighting, proportional theory (PT) weighting, is to use $w_i = 1/g(x_i;\boldsymbol{\theta})^2$ instead $w_i = 1/y_i^2$, i.e., to estimate the true values of y by the fitted model values rather than by the experimental values y_i. Since the model parameters are not initially known but are refined during the iterative minimizaton procedure, the weights are similarly refined as the least squares optimization proceeds, a procedure known as iterated generalized least squares (GLS) [12]. The computational effort is increased, but the weights vary more smoothly across the data set, and this method has been advocated as superior

to using the experimental values [12–14]. The same derivation applies, giving the AIC in Eq. (15). To strictly give the maximum likelihood estimator, the optimum weights must be obtained not from a least-squares minimization of S, but by a minimization of an objective function that includes both S and the term $\sum_i \ln g(x_i;\boldsymbol{\theta})^2$, because this term for the weights in Eq. (12) depends on the model parameters. Such a minimization is the basis of the extended least squares (ELS) method [14]. However, ELS has been criticized and needs to be used with caution, as it can lead to inconsistent parameter estimates if the wrong variance model is used [12]. For simplicity, we use Eq. (15) with the theory weights determined from the weighted least squares minimization rather than ELS, and assume the AIC value calcuated in this way is only slightly higher than the true value.

1.3. Magnitude weighting

Immittance data is complex, and real and imaginary parts are considered as separate data points. In experimental magnitude (ME) weighting, the real and imaginary parts are given the same weighting based on the magnitude: $w_{j,\mathrm{Re}} = w_{j,\mathrm{Im}} = 1/(y_{j,\mathrm{Re}}^2 + y_{j,\mathrm{Im}}^2)$, implying a model for variance $\sigma_{j,\mathrm{Re}}^2 = \alpha^2\left(y_{j,\mathrm{Re}}^2 + y_{j,\mathrm{Im}}^2\right) = \alpha^2/w_{j,\mathrm{Re}}$ and the same for $\sigma_{j,\mathrm{Im}}^2$. (The index j here labels the $n/2$ frequencies and not the n data points.) Like proportional weighting, it also has a theory weighting equivalent (MT). Empirically, magnitude weighting has been found to facilitate fitting and it has been advocated by Zoltowski [15], though it is less clear how good a model for variance it is. Nonetheless, AIC can be used to compare it with other weighting schemes. A similar derivation to the above shows that AIC is again given by Eq. (15).

2. Results and discussion

2.1. Simulated data

We use throughout the formula for AIC for weighted data, Eq. (15); it also applies for unweighted (unit weighted) data because the additional term with the weights is zero in that case. We note that although the values of the weights and weighted sums of squares are very different for the different weighting schemes, AIC compares the combination $n \ln S - \sum_i \ln w_i$ which is scale independent, e.g., doubling the weights leaves this combination unchanged.

We first illustrate the use of AIC for simulated data for the immittance function

$$y = \frac{y_0}{\cosh(\sqrt{\mathrm{i}\omega\tau_d})\,(1+\mathrm{i}\omega\tau_\theta)} \tag{16}$$

This is derived as a special case of the general theory for electrochemical production of hydrogen on one side of an iron foil and its diffusion through the foil with detection at the exit side in a Devanathan-Stachurski cell [16, 17]. (Details of the experiments and theory presented here will be given elsewhere.) The diffusivity may be determined from τ_d and the time constant τ_θ is related to the kinetics of adsorbed hydrogen. The immittance here is a "transadmittance",

a transfer function for oscillating current at the exit side relative to oscillating potential on the entrance side. The complexity of the measurement is such that we have no simple expectation of how the variances will vary with frequency.

Synthetic data were generated with random normally-distributed errors with 10% standard deviation, i.e., $\alpha = 0.1$ in the proportional model. This is larger than the errors in the real data, but enables comparisons of fitting methods under challenging conditions. Other details of the data are shown in Fig. 1, where fits using five weighting schemes to the simpler (incorrect) function in Eq. (17) are shown. The fitting was done using Maple's NLPSolve routine [18], with the nonlinear simplex method option, and with a custom calling program to generate the objective function and produce the statistics. Standard errors and confidence intervals were calculated from partial derivatives at the minimum using the usual first-order analysis [19].

$$y = \frac{y_0}{\cosh(\sqrt{\mathrm{i}\omega\tau_d})} \tag{17}$$

Visually the fits seem poor (Fig. 1), and unit (red) and magnitude (green) weighting schemes seem superior to the proportional weighting (blue) by which the samples were actually generated. The AIC values for fitting to Eq. (17) with different weighting schemes are shown in the top half of Table 1. The AIC values agree with the visual observations, indicating increasingly better fits in the order: proportional, magnitude, unit. Within this ranking, theory weighting is slightly worse than data weighting. Examination of the plots for the real and imaginay parts, and the residual plots, e.g., Fig. 1 shows the reason for the poor performance of the proportional: good fits of the imaginary parts at low frequencies and moderately-good fits of the real parts at higher frequencies lead to a minimum that has the fitted function far from the true one. This feature of proportional weighting in which tradeoffs between real and imaginary parts give additional flexibility is also seen in the large difference between the theory and data fits for this weighting (solid and dashed blue lines in Fig. 1). The conclusion from AIC is to use unit weighting in this case, and it is interesting that this is the only weighting scheme for which the 95% confidence intervals ($\approx \pm 2\sigma$) enclose the true values for both parameters. The τ_d parameter is overestimated in most cases, owing to the systematic error of using the wrong model.

Perhaps not surprisingly, refitting to the correct model, Eq. (16), gives fits that are visually much better (not shown), and for which all AIC values are more negative than those for the wrong model, clearly justifying the addition of the parameter τ_θ. More interesting is that the proportional weighting is now superior to the magnitude weighting, i.e., AIC now finds the correct error model. Theory weighting is now superior to data weighting in both the proportional and magnitude cases, though the AIC difference is small. The 95% confidence intervals for all parameters now enclose the true values, with the exception of the proportional theory case, where the interval for y_0 ($303 - 318$ μS cm^{-2}) is slightly high.

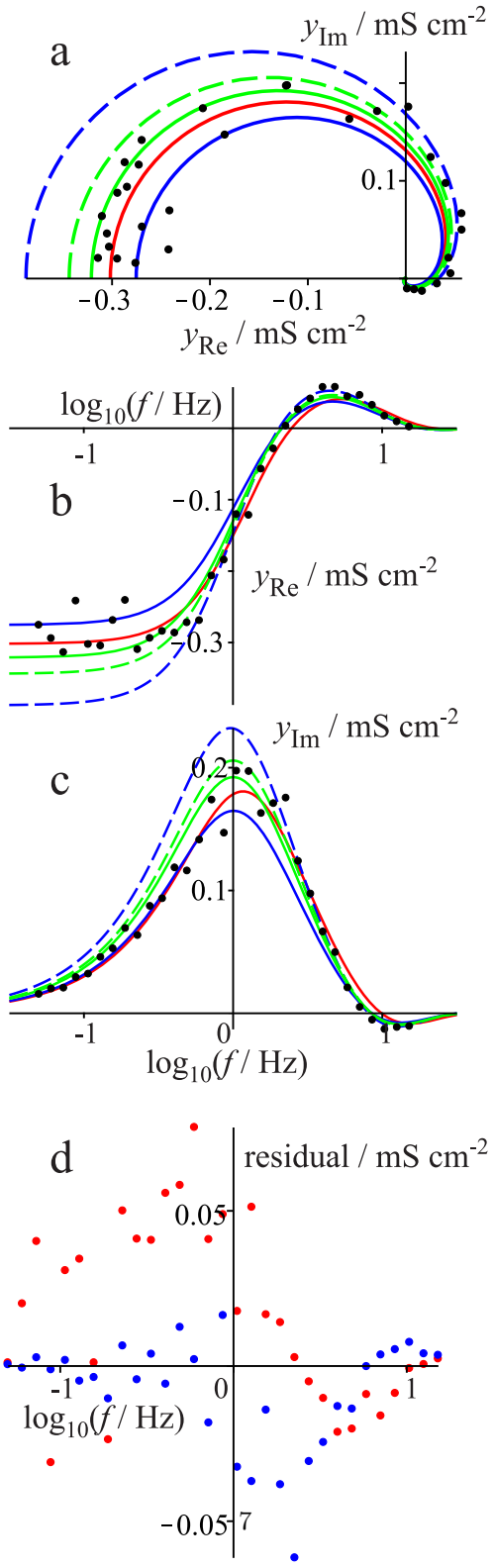

Figure 1: Fits to simulated data. (a)-(c) data (circles) and fits with different weighting: unit (red), magnitude (green), proportional (blue); solid lines use experimental weighting, dashed lines use theory weighting. (d) real (red) and imaginary (blue) residuals for PE weighting.

model Eq.	weighting	$-y_0$ / μS cm^{-2}	τ_d / ms	τ_θ / ms	AIC	Δ	W
16		300	300	10			
17	U	302 ± 4	318 ± 9		-323	116	0.00
17	PE	275 ± 12	368 ± 3		-290	149	0.00
17	PT	388 ± 17	386 ± 5		-266	173	0.00
17	ME	321 ± 8	370 ± 8		-314	125	0.00
17	MT	344 ± 9	372 ± 8		-309	130	0.00
16	U	296 ± 4	272 ± 16	15 ± 5	-337	102	0.00
16	PE	302 ± 4	303 ± 3	9.6 ± 0.5	-437	2	0.25
16	PT	311 ± 4	303 ± 3	9.6 ± 0.5	-439	0	0.75
16	ME	295 ± 5	280 ± 10	13 ± 2	-388	52	0.00
16	MT	301 ± 5	280 ± 10	13 ± 2	-392	47	0.00

Table 1: Fitting results for simulated data. First line shows the true parameters. Errors are the standard errors (1σ) from the regression. AIC is calculated according to Eq. (15), Δ is the difference between AIC and the minimum AIC, W is the Akaike weight, Eq. (18).

The absolute values of AIC are dependent on the scaling and units, and so it is common to tabulate the difference from the best fitting model, $\Delta_i = AIC_i - \min_i(AIC_i)$ as in Table 1. Now smaller values indicate better models, with zero for the best model. Another measure widely used is Akaike weights [2], Eq. (18).

$$W_i = \exp\left(-\Delta_i/2\right) / \sum_i \exp\left(-\Delta_i/2\right) \tag{18}$$

These are given in the last column of Table 1 and may roughly be interpreted as the probability that the model is correct. These probabilities are the quantitative statement of confidence in different models, and replace acceptance/rejection statistical tests in classical statistics, such as the F-test. Although they do not give a definitive decision on acceptance/rejection of a model as does a statistical test, they do not have the requirement for a confidence level, whose choice can be somewhat arbitrary.

While it is not entirely clear that Akaike weights for different weighting schemes should be compared in this way, it is clear that the proportional weighting for the model with adsorption is much superior to other options. Examination of the standard errors for the parameters might not have been enough to come to this conclusion. While other tools such as examination of the residuals must complement any regression analysis, the AIC analysis proves to be simple and effective for both model and weighting discrimination.

2.2. Real data

With the above analysis of simulated data as a preamble, we turn now to analysis of real transadmittance data, where there are two objectives: (i) extraction of the diffusivity of hydrogen through an Fe foil from accurate determination of the τ_d parameter, and (ii) extraction of kinetic information from τ_θ or other parameters in more complicated models. The principal difficulty in (i), as shown above, is that τ_d may be biased away from its true value because of the presence of unknown kinetic or other phenomena. While the presence of adsorbed

H on the entry side leads to an expectation that Eq. (16) is probably correct, the adsorption semicircle is masked in the regular entry-side impedance, which suggests that τ_θ is small and perhaps not detectable. If, in contrast, another parameter can be deduced, the next most complex variation of the theory leads to the expectation:

$$y = \frac{a + b\left(1 + \mathrm{i}\omega\tau_\theta\right)^{-1}}{\cosh(\sqrt{\mathrm{i}\omega\tau_d})} \tag{19}$$

Another possibility is that there is some time-constant dispersion leading to CPE-like behavior, which can be accommodated by raising one or other of $\mathrm{i}\omega\tau_d$ or $\mathrm{i}\omega\tau_\theta$ to an exponent ϕ that is typically slightly less than one. All these possibilities are covered by Eq. (20), where $y_0 = a + b$ is the zero-frequency intercept as before.

$$y = \frac{a\left(\mathrm{i}\omega\tau_\theta\right)^{\phi_\theta} + y_0}{\cosh(\sqrt{\left(\mathrm{i}\omega\tau_d\right)^{\phi_d}})\left(1 + \left(\mathrm{i}\omega\tau_\theta\right)^{\phi_\theta}\right)} \tag{20}$$

The simpler models are attained by setting $a = 0$, $\tau_\theta = 0$, $\phi_d = 1$ and/or $\phi_\theta = 1$ as appropriate. This leads to the following nine models, in which different parameters were varied: 1. y_0, τ_d; 2. y_0, τ_d, ϕ_d; 3. y_0, τ_d, τ_θ; 4. y_0, τ_d, τ_θ, ϕ_θ; 5. y_0, τ_d, τ_θ, ϕ_d, ϕ_θ; 6. a, y_0, τ_d, τ_θ; 7. a, y_0, τ_d, τ_θ, ϕ_d; 8. a, y_0, τ_d, τ_θ, ϕ_θ; 9. a, y_0, τ_d, τ_θ, ϕ_d, ϕ_θ, and in each case the three weighting schemes U, ME, PE were used. (Theory weighting was omitted since the simulated data study showed the same AIC ordering for theory or experimental weighting, and convergence was more difficult for theory weighting.) Data at three different potentials (vs $\mathrm{Hg|HgSO_4}$ reference electrode) were investigated. As usual with nonlinear least squares, different results (false local minima) were sometimes attained for different starting values (especially for the PE weighting), and so several combinations of starting values were investigated. The following strategy was adopted: Runs giving physically unreasonable parameters were rejected (τ_θ or τ_d negative, ϕ_d or ϕ_θ outside the range $0.5 - 1.0$). Then if there were different results for the same model/weighting combination for different initial estimates, only the result for the most negative AIC was retained. The details of these retained fits are given in the supplementary information.

The simplest way to compare the remaining valid models is to rank them by AIC for each potential. (AIC is data dependent, so AIC values for different potentials cannot be directly compared.) This leads to the conclusion that model 7 with unit weighting was best for the -1.5 V data, model 6 with ME weighting was best for the -1.6 V data, and model 8 with ME weighting was best for the -1.7 V data, and these fits are shown in Fig. 2. The fact that the best -1.5 V fit has a different weighting scheme from the other two potentials is simply a reflection that it is a better model for the error structure in that case; in the information theoretic sense the comparison between different error structure models and physical models are on the same footing.

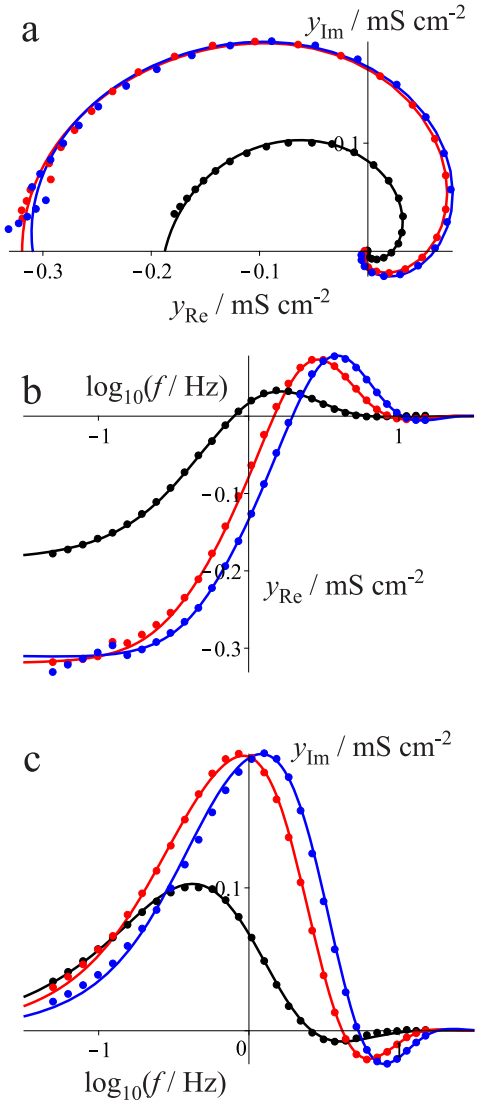

Figure 2: Fits to real data. Best AIC fits for -1.5 V (black), -1.6 V (red), -1.7 V (blue).

Since all these models contain the kinetic parameters τ_θ and a, these parameters can reliably be determined, in addition to the parameters y_0 and τ_d of the basic model. However, these models differ in the significance of the dispersion parameters ϕ_d and ϕ_θ so a closer look is in order.

Of the retained runs, for -1.6 V and -1.7 V most models had values of ϕ_d and/or ϕ_θ fixed at 1 or close to it (> 0.87). For -1.6 V, the second ranked model has $\phi_\theta = 0.995 \pm 0.015$, but ϕ_θ is fixed at 1 in the top ranked model. These are the only models with significant Akaike weights (0.72 and 0.28), but since the fitted value of ϕ_θ is not significantly different from 1, both models indicate no dispersion. For the -1.7 V data, the two only models of significance have ϕ_d fixed at 1 and $\phi_\theta = 0.933 \pm 0.012$ or 0.944 ± 0.006, though these models differ only in the weighting scheme. The third ranked model has ϕ_θ fixed at 1, and all three models yield similar values of the other parameters, so a small amount of dispersion of no consequence for extracting kinetic quantities is indicated. The most significant dispersion was in the -1.5 V data, where the top model had $\phi_d = 0.837 \pm 0.012$. Compared to the same model without ϕ_d, the time constant τ_d is hardly changed, and although the values of τ_θ and a change, their estimated errors also significantly decreased. For the second and fourth models at -1.5 V, low ϕ_θ values correlated with higher τ_θ and lower τ_d values than other models, making the addition of dispersion seem anomalous, though in any case their Akaike weights were very small ($< 10^{-7}$).

With regard to the other parameters, the values of y_0, the zero-frequency real axis intercept, were the most consistent across models. The time constant τ_d showed more variability, up to a factor of two, but the model ranking showed a trend to more consistency for higher ranked models. This illustrates how using AIC to rank a large number of competing models and weighting schemes can give subjective confidence in the chosen parameter values. The time constant τ_θ showed more variability but again consistency for the higher ranked models, with anomalous values usually associated with standard errors exceeding 50% for some parameter(s) of that fit (corresponding approximately to acceptance at the 95% confidence level in a t-test). The parameter a was least consistent, and its value also changed sign: from small and positive at -1.5 V to negative for -1.6 V and -1.7 V. However, this change in value could be a real effect of potential, and suggests further data collection is warranted.

There is a bias correction available for AIC for when the number of degrees of freedom is small. The corrected value, $AICc$, Eq. (21), is recommended when the ratio of the number of data points to number of parameters is less than 40 [2], as is the case here. It was also calculated but made no difference in the rankings.

$$AIC_c = -2 \ln \widehat{L} + 2(p+1) \left(\frac{n}{n-p-2} \right) \tag{21}$$

The usual method of deciding whether to add a parameter is the F-test, and this test showed that adding τ_θ and then a to the base model 1 was justified at the 99% confidence level, except for adding a in the case of the -1.5 V data. Unit weighting was used, with the residual sum of squares S used as a scaled version

of χ^2. Although different weighting schemes cannot be directly compared, the conclusion for the unit weighting case is the same as found by ranking with AIC.

Finally, we note that the data are Kramers-Kronig compliant insofar as they fit well to a function which satisfies the Kramers-Kronig stability and finiteness conditions (provided τ_θ is positive). This was confirmed using Boukamp's KK-test program (in admittance mode) based on the algorithm described in Ref. [20].

3. Conclusions

An expression for AIC was derived for the common weighting schemes used in complex nonlinear least-squares fitting of immittance data. AIC provides a single number to conveniently compare many models with different weighting schemes, and was demonstrated for simulated and real data for a transadmittance function. Other information theoretic criteria such as the Bayesian infomation criterion (BIC) or the weighted information criterion [21] (WIC, a weighted average of AIC and BIC) may also be useful. However, AIC supplemented by rejection of physically inappropriate models is a useful tool for fitting immittance data.

4. Acknowledgements

Financial support for this work is acknowledged from the Natural Sciences and Engineering Council of Canada (Discovery Grant RGPIN-2017-04045) and the Research Council of Norway (INTPART CANOPENER project 261620 and HyF-Lex project 244068/E30). D.H. thanks Peter Harrington for useful discussions.

A. Supplementary Information

Supplementary material consists of a Excel spreadsheet with parameters, standard errors, AIC and AICc values for all the retained models, and data for the F-tests. This material can be found, in the online version,
at http://dx.doi.org/10.1016/j.electacta.XXXX.

References

[1] H. Akaike, Information theory as an extension of the maximum likelihood principle, in B.N. Petrov, F. Csaki, eds, Second International Symposium on Information Theory, Akademiai Kiado, Budapest, 1973.

[2] K.P. Burnham, D.R. Anderson, Model Selection and Multimodel Inference, 2nd ed., Springer, NY, 2002.

[3] M.E. van der Geest, N.J. Dangerfield, D.A. Harrington, An a.c. Voltammetry Study of Pt Oxide Growth, J. Electroanal. Chem. 420 (1997) 89-100.

[4] Z. Gao, J.W. Green, J. vanderborght, W. Schimdt, Improving uncertainty analysis in kinetic evaluations using iteratively reweighted least squares, Environmental Toxicology and Chemistry, 30 (2011) 2363–2371.

[5] V.R. Vasquez, W.B. Whiting, Regression of binary interaction parameters for thermodynamic models using an inside-variance estimation method (IVEM), Fluid Phase Equilibria 170 (2000) 235–253.

[6] M.E. Orazem, A systematic approach toward error structure identification for impedance spectroscopy, J. Electroanal. Chem. 572 (2004) 317–327.

[7] F. Ciucci, Modeling electrochemical impedance spectroscopy, Curr. Opinion Electrochem., doi: 10.1016/j.coelec.2018.12.003.

[8] M.E. Orazem, B. Tribollet, Electrochemical Impedance Spectroscopy, 2nd ed, Wiley, Hoboken, NJ, 2017.

[9] D. Ritzberger, M. Striednig, C. Simon, C. Hametner, S. Jakubek, Online estimation of the electrochemical impedance of polymer electrolyte membrane fuel cells using broad-band current excitation, J. Power Sources 405 (2018) 150–161.

[10] D. Mu, Z. Dong, M. Wang, A separated-frequency parameters identification method of Li-ion battery model for electric vehicles, Applied Mechanics and Materials, 672-674 (2014) 727-730.

[11] C. Qui, M.E. Orazem, A weighted nonlinear regression-based inverse model for interpretation of pipeline survey data, Electrochim. Acta, 49 (2004) 3965–3975.

[12] J.C. van Houwelingen, Use and abuse of variance models in regression, Biometrics 44 (1988) 1073-1081.

[13] J.R. Macdonald, J.R. Potter, Jr, A flexible procedure for analyzing impedance spectroscopy results: description and illustrations, Solid State Ionics, 23 (1987) 61-79.

[14] L.B. Sheiner, S.L. Beal, Pharmacokinetic parameter estimates from several least squares procedures: superiority of extended least squares. J. Pharmacokinetics Biopharmaceutics 13 (1985) 185-201.

[15] P. Zoltowski, The error function for fitting of models to immittance data, J. Electroanal. Chem. 178 (1984) 11-19.

[16] A. Lasia, Applications of the Electrochemical Impedance Spectroscopy to Hydrogen Adsorption, Evolution and Absorption into Metals, in: B.E. Conway, R.E. White, eds, Modern Aspects of Electrochemistry, Kluwer/Plenum, New York, v. 35 (2002) pp. 1-49.

[17] C. Gabrielli, G. Maurin, L. Mirkova, H. Perrot, B. Tribollet, Transfer function analysis of hydrogen permeation through a metallic membrane in a Devanathan cell. I. Theory, J. Electroanal. Chem. 590 (2006) 1–14.

[18] Maple version 2017, Maplesoft, a division of Waterloo Maple Inc., Waterloo, Ontario, <http://www.maplesoft.com>.

[19] D.M. Bates, D.G. Watts, Nonlinear Regression Analysis and its Applications, Wiley, 1988, p 53.

[20] B.A. Boukamp, A Linear Kronig-Kramers Transform Test for Immittance Data Validation, J. Electrochem. Soc., 142 (1995) 1885-1894.

[21] T.-J. Wu, A. Sepulveda, The weighted average information criterion for order selection in time series and regression models, Statistics Probability Lett. 39 (1998) 1-10.

E / V	model	weight	y_0 / mS cm^{-2}	std err	τ_d / s	std err	ϕ_d	std err	τ_θ / s	std err	ϕ_θ	std err	a / mS cm-2	std err	AIC	AICc	Δ	exp(-Δ/2)	W
-1.5	7	U	-0.187452	0.001022	0.893163	0.016149	0.836553	0.011972	0.073069	0.0067			0.0609907	0.00521	-641.121	-639.593	0	1.00E+00	1.000
-1.5	8	U	-0.193227	0.002658	0.348304	0.040872			0.213515	0.03935	0.721446	0.044506	0.0412816	0.008168	-606.49	-604.963	34.631	3.02E-08	0.000
-1.5	5	U	-0.183263	0.001724	0.912232	0.027168	0.899128	0.017696	0.06412	0.012541					-558.017	-556.945	83.104	9.00E-19	0.000
-1.5	4	U	-0.189141	0.002939	0.55266	0.128211			0.1472	**0.087945**	0.686878	0.097944			-519.499	-518.427	121.622	3.89E-27	0.000
-1.5	3	U	-0.173401	0.001491	0.919334	0.020117			0.019401	0.004913					-503.816	-503.114	137.305	1.53E-30	0.000
-1.5	6	U	-0.173206	0.001477	0.919643	0.019111			0.010526	**0.011156**			0.156093	**0.320362**	-503.144	-502.072	137.977	1.09E-30	0.000
-1.5	1	U	-0.176115	0.001589	0.989031	0.01405									-487.748	-487.334	153.373	4.96E-34	0.000
-1.5	6	PE	-0.164543	0.012837	1.37804	0.026592			0.158915	0.044295			-0.299207	0.027975	-451.424	-450.352	189.697	6.42E-42	0.000
-1.5	3	PE	-0.160009	0.010989	0.913307	0.060994			0.029143	0.007864					-445.502	-444.8	195.619	3.33E-43	0.000
-1.5	3	ME	-0.169829	0.010032	0.83542	0.058083			0.043806	0.01111					-445.084	-444.382	196.037	2.70E-43	0.000
-1.5	6	ME	-0.172769	0.014272	0.863814	0.130651			0.041204	0.01611			-0.00605002	**0.021737**	-443.366	-442.294	197.755	1.14E-43	0.000
-1.5	5	ME	-0.171102	0.016043	0.846764	0.135293	0.997361	0.025231	0.04335	0.012973					-443.095	-442.024	198.026	9.98E-44	0.000
-1.5	8	PE	-0.138964	0.019782	1.74102	**0.787251**			1.13E-06	**0.000963**	0.568586	0.089431	-146.355	**70761.7**	-441.752	-440.225	199.369	5.10E-44	0.000
-1.5	7	ME	-0.205761	0.031238	2.86523	1.02628	0.886396	0.045951	0.184184	0.053295			-0.742259	0.360196	-440.439	-438.911	200.682	2.65E-44	0.000
-1.5	1	ME	-0.199343	0.013559	1.25886	0.036824									-429.523	-429.109	211.598	1.13E-46	0.000
-1.5	1	PE	-0.160153	0.014876	1.25822	0.026841									-418.148	-417.734	222.973	3.82E-49	0.000
																	Total:	**1.00E+00**	

(bold std errors are those > 50%)

Supplementary material for M. Ingdal, R. Johnsen, D.A. Harrington, The Akaike information criterion in weighted regression of immittance data, Electrochim. Acta, 2019.

E / V	model	weight	y_0 / mS cm^{-2}	std err	τ_d / s	std err	ϕ_d	std err	τ_θ / s	std err	ϕ_θ	std err	a / mS cm-2	std err	AIC	AICc	Δ	exp(-Δ/2)	W
-1.6	6	ME	-0.319192	0.002205	0.864819	0.003196			0.096073	0.002263			-0.854287	0.010405	-579.059	-577.988	0	1.00E+00	0.719
-1.6	8	ME	-0.318655	0.002815	0.865605	0.004067			0.095356	0.003205	0.99515	0.015382	-0.858083	0.016021	-577.175	-575.648	1.884	3.90E-01	0.280
-1.6	6	U	-0.317486	0.00089	0.879974	0.010777			0.098031	0.002349			-0.855326	0.023814	-557.407	-556.336	21.652	1.99E-05	0.000
-1.6	6	PE	-0.323789	0.003252	0.857057	0.000565			0.105553	0.001966			-0.807786	0.008041	-530.671	-529.599	48.388	3.11E-11	0.000
-1.6	7	U	-0.321016	0.001656	0.389153	0.012864	0.892765	0.014	0.041087	0.004946			0.135346	0.013843	-518.439	-516.912	60.62	6.86E-14	0.000
-1.6	5	U	-0.31523	0.003059	0.415039	0.020497	0.966716	0.0204	0.035672	0.009269					-432.609	-431.538	146.45	1.58E-32	0.000
-1.6	3	U	-0.310659	0.002281	0.422949	0.012143			0.026931	0.003529					-428.116	-427.415	150.943	1.67E-33	0.000
-1.6	3	ME	-0.348308	0.009583	0.503122	0.026645			0.020718	0.004064					-367.741	-367.039	211.318	1.30E-46	0.000
-1.6	1	U	-0.31933	0.003877	0.516489	0.011014									-357.726	-357.312	221.333	8.67E-49	0.000
-1.6	3	PE	-0.315052	0.019484	0.603846	0.026101			0.014861	0.003444					-300.2	-299.498	278.859	2.80E-61	0.000
-1.6	1	ME	-0.383788	0.018869	0.701042	0.022004									-293.792	-293.378	285.267	1.14E-62	0.000
-1.6	1	PE	-0.272241	0.028574	0.834876	0.010397									-254.412	-253.998	324.647	3.19E-71	0.000
																	Total:	**1.39E+00**	

E / V	model	weight	y_0 / mS cm^{-2}	std err	τ_d / s	std err	ϕ_d	std err	τ_θ / s	std err	ϕ_θ	std err	a / mS cm-2	std err	AIC	AICc	Δ	exp(-Δ/2)	W
-1.7	8	ME	-0.309831	0.002631	0.729371	0.00473			0.052039	0.002276	0.933338	0.012083	-1.2079	0.034144	-554.366	-552.838	0	1.00E+00	0.998
-1.7	8	PE	-0.316717	0.002113	0.727546	0.003624			0.054025	0.001601	0.944122	0.005849	-1.17181	0.024355	-541.634	-540.107	12.732	1.72E-03	0.002
-1.7	6	ME	-0.318264	0.002396	0.712291	0.003871			0.062307	0.001578			-1.07562	0.018693	-529.954	-528.883	24.412	5.00E-06	0.000
-1.7	6	U	-0.315835	0.001197	0.680961	0.011535			0.072711	0.002383			-0.930158	0.036311	-511.836	-510.765	42.53	5.82E-10	0.000
-1.7	7	U	-0.316551	0.001689	0.69157	0.020995	0.995836	0.006783	0.074074	0.003281			-0.936741	0.038034	-510.213	-508.686	44.153	2.58E-10	0.000
-1.7	9	U	-0.316936	0.002348	1.31532	**0.96747**	0.904519	0.085317	0.0407	**0.03871**	0.873994	0.077514	-2.63233	**3.57785**	-509.319	-507.245	45.047	1.65E-10	0.000
-1.7	6	PE	-0.317728	0.003227	0.701429	0.002746			0.06721	0.001419			-1.0035	0.014035	-490.012	-488.94	64.354	1.06E-14	0.000
-1.7	3	U	-0.310104	0.002376	0.297344	0.010604			0.024223	0.00327					-414.356	-413.654	140.01	3.96E-31	0.000
-1.7	5	U	-0.312336	0.003123	0.293344	0.017674	0.982179	0.021736	0.027859	0.007941					-414.008	-412.936	140.358	3.32E-31	0.000
-1.7	1	U	-0.319137	0.004245	0.379627	0.009256									-338.733	-338.32	215.633	1.50E-47	0.000
-1.7	3	ME	-0.344899	0.01158	0.364332	0.028549			0.01909	0.004968					-311.143	-310.442	243.223	1.53E-53	0.000
-1.7	3	PE	-0.286633	0.020386	0.330671	0.039045			0.025569	0.007706					-256.61	-255.908	297.756	2.20E-65	0.000
-1.7	1	ME	-0.374574	0.020013	0.522401	0.021874									-250.497	-250.083	303.869	1.04E-66	0.000
-1.7	1	PE	-0.241813	0.027304	0.533224	0.016829									-218.504	-218.09	335.862	1.17E-73	0.000
																	Total:	**1.00E+00**	

(bold std errors are those > 50%)

E / V	model	weight	y_0 / mS cm^{-2}	std err	τ_d / s	std err	ϕ_d	std err	τ_θ / s	std err	ϕ_θ	std err	ɪ / mS cm-:	std err	AIC	AICc	RSS	DOF	F	p
-1.5	1	U	-0.176115	0.001589	0.989032	0.01405	0	0	0	0	0	0	0	0	-487.748	-487.334	0.001263	60		
-1.5	3	U	-0.173401	0.001491	0.919337	0.020117	0	0	0.0194	0.004913	0	0	0	0	-503.816	-503.114	0.000944	59	19.962	0.000
-1.5	6	U	-0.173205	0.001477	0.919644	0.01911	0	0	0.010524	0.011155	0	0	0.156158	0.320499	-503.144	-502.072	0.000924	58	1.2554	0.267
-1.6	1	U	-0.31933	0.003877	0.516489	0.011014	0	0	0	0	0	0	0	0	-357.726	-357.312	0.010283	60		
-1.6	3	U	-0.31066	0.002281	0.422952	0.012143	0	0	0.02693	0.003528	0	0	0	0	-428.116	-427.415	0.003199	59	130.64	0.000
-1.6	6	U	-0.317485	0.00089	0.879971	0.010777	0	0	0.098032	0.002349	0	0	-0.85532	0.023814	-557.407	-556.336	0.000385	58	424.05	0.000
-1.7	1	U	-0.319137	0.004245	0.379627	0.009256	0	0	0	0	0	0	0	0	-338.733	-338.32	0.013968	60		
-1.7	3	U	-0.310104	0.002376	0.297343	0.010604	0	0	0.024223	0.00327	0	0	0	0	-414.356	-413.654	0.003994	59	147.34	0.000
-1.7	6	U	-0.315835	0.001197	0.680964	0.011535	0	0	0.072711	0.002383	0	0	-0.93017	0.036312	-511.836	-510.765	0.000803	58	230.57	0.000

RSS = residual sum of squares
DOF = degrees of freedom