Efficient prediction of x-axis intercepts of discrete impedance spectra

Thomas Schmid¹, Dorothee Günzel² and Martin Bogdan¹

1- Universität Leipzig - Department of Computer Engineering Augustusplatz 10, 04109 Leipzig - Germany
2- Charité - Institute of Clinical Physiology Campus Benjamin Franklin, 12200 Berlin - Germany

Abstract. In impedance spectroscopy of epithelial cell layers, it is a common task to extrapolate discrete two-dimensional plots in order to determine electrical properties associated with axis intercepts. Here, we investigate how implicit properties of such curves can be used to predict the x-axis intercept where explicitly determined properties fail to do so. We perform feature extraction, algorithmic feature ranking and dimension reduction on model impedance spectra derived from a tissue-equivalent electric circuit. Selected feature subsets are assessed by training artificial neural networks to predict the intercept. Results show that subsets of three or less implicit features provide a reasonable basis for predictions.

1 Introduction

Impedance spectroscopy is the measurement of current-voltage relationships under alternate current (AC). Spectra are typically obtained by measuring complex impedances at 40 to 50 frequencies between 1 Hz and 100 kHz [6]. They are often presented as Nyquist diagrams (Fig. 1a), in which real part Z^{re} and imaginary part Z^{im} are plotted against each other. The theoretical impedance of an electric circuit can be calculated if the values of all circuit components are known.

Electric circuits of different degrees of complexity are used to describe epithelial cell layers [4]. The simplest circuit that explicitly reflects asymmetry of such polar cell layers consists of two resistor-capacitor (RC) subcircuits in series and a resistor in parallel to them (Fig. 1b). In addition, electric behavior of the so-called subepithelium may be considered by a further resistor in series.

In measurements, epithelial cell layers often yield semicircular impedance spectra (Fig. 1a) and the real part of the closest, complex impedance (e.g. at 1.3 Hz) is a reliable predictor for the x-axis intercept or the epithelial resistance R respectively [3]. It becomes increasingly unreliable, however, the more a spectrum deviates from a semicircular shape (Fig 1a,c).

To achieve efficient predictions in this context, we suggest to train artificial neural networks (ANNs) with theoretical spectra of tissue-equivalent circuits and corresponding target values. ANNs have proven a reliable pattern recognition method [2], even though reliablity is limited when imbalanced training data is used [5]. Due to the infamous curse of dimensionality [1], generation of training data balanced regarding all circuit parameters plus the target domain is challenging in practice. To allow efficient predictions, we search for a minimal subset of features extracted from the up to 100 features measured per spectrum.



Fig. 1: (a) Overlap of two Nyquist diagrams reflecting impedance measurements on low-resistent epithelial cell layers at 42 frequencies between 1.3 and 16,000 Hz. The semicircular spectrum reflects physiological conditions ($\tau_a \approx \tau_b$), where Z^{re} or r at 1.3 Hz can be used as predictor for the epithelial resistance $R = R_p(R_a + R_b)/(R_p + R_a + R_b)$. If R_a is decreased by drug application, τ_a decreases and a non-semicircular shape is obtained. (b) Equivalent electric circuit discriminating between apical ($\tau_a = R_a C_a$) and basolateral ($\tau_b = R_b C_b$) properties of an epithelial cell layer. (c) Alternative representation of the non-semicircular spectrum from Fig. 1a. Complex impedances (Z^{re}, Z^{im}) are transformated into polar coordinates (magnitude r, phase ϕ).

2 Methods

2.1 Modeling impedance spectra

The tissue-equivalent circuit considered here (Fig. 1b) consists of two RC subcircuits a (R_a , C_a) and b (R_b , C_b) located in series and a resistor in parallel (R_p). Using Kirchhoff's laws, the complex impedance Z of an electric circuit at an angular frequency ω can be derived from the impedances of its components:

$$Z(\omega) = \frac{R_p(R_a + R_b) + i\omega[R_p(R_a\tau_b + R_b\tau_a)]}{R_a + R_b + R_p(1 - \omega^2\tau_a\tau_b) + i\omega[R_p(\tau_a + \tau_b) + R_a\tau_b + R_b\tau_a]}$$
(1)

where $i = \sqrt{-1}$, and $\tau_a = R_a C_a$ and $\tau_b = R_b C_b$. Results are transformed into polar coordinates, i.e. into phase ϕ and magnitude r.

Using n=42 frequency-dependent impedances $Z(\omega_0), ..., Z(\omega_{n-1})$ with $\omega_0 = 2\pi (1.3 \text{ Hz})$ and $\omega_j = \omega_0 \cdot 10^{j \cdot 0.1}$ as one spectrum, then establishes the relation

$$(R_a, C_a, R_b, C_b, R_p) \sim (\phi(\omega_0), r(\omega_0), \dots, \phi(\omega_{n-1}), r(\omega_{n-1}))$$

$$(2)$$

By varying all five circuit parameters, measurements on distinct epithelial cell lines under a variety of experiment conditions are mimicked: ESANN 2013 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 24-26 April 2013, i6doc.com publ., ISBN 978-2-87419-081-0. Available from http://www.i6doc.com/en/livre/?GCOI=28001100131010.

- a primarily transcellularily (T) conducting type, where $R_p > R_a + R_b$;
- a primarily paracellularily (P) conducting type, where $R_p < R_a + R_b$.

To identify predictors reliable in the given context, only spectra are used where Z^{re} and r at the lowest frequency (1.3 Hz) differ by at least five percent from the target value. Dataset T contains 17,954 spectra with 410 $\Omega \text{cm}^2 < R <$ 1252 Ωcm^2 , dataset P contains 16,213 spectra with 3324 $\Omega \text{cm}^2 < R <$ 9203 Ωcm^2 . A third dataset is created by merging these two (Mix).

2.2 Extracting and ranking implicit features

Phases and magnitudes are handled as separate feature subsets S_{ϕ} and S_r :

$$S_{\phi} = \{\phi(\omega_0), ..., \phi(\omega_{n-1})\}$$
(3)

$$S_r = \{r(\omega_0), ..., r(\omega_{n-1})\}$$
(4)

Two further sets of inherent features are extracted that represent the n-1 distances between two consecutive features of the original subsets:

$$S_{\Delta\phi} = \{\Delta\phi | \Delta\phi_i = \phi(\omega_{i+1}) - \phi(\omega_i), 0 \le i < n-1\}$$
(5)

$$S_{\Delta r} = \{ \Delta r | \Delta r_i = r(\omega_{i+1}) - r(\omega_i), 0 \le i < n-1 \}$$
(6)

From all four sets, primary statistical parameters (minimum, maximum, average, median, variance, standard deviation) were calculated individually per impedance spectrum. This resulted in a total of 24 implicit curve features that were considered for feature selection. Each feature was normalized individually before selecting sample sets and applying feature selection.

Due to the continuous nature of the target domain, an algorithm for nonlinear regression problems is required to rank features. A method that has been tested successfully on non-trivial data is preferable. Both applies to the "Regression, Gradient-guided feature selection" (RGS) by Navot et al., which is based on a feature-weighted version of the k-nearest-neighbor algorithm [7].

To apply RGS, we use MATLAB code provided by the authors¹. Used parameters are: number of neighbors k = 50; number of iterations T = 1; Gaussian decay factor β is set to half the mean distance between points and their k neighbors; number of starts was 2. Three sets of 5,000 samples each are chosen randomly from the datasets T, P and Mix after pre-selection and normalization.

2.3 Evaluating feature subsets with artificial neural networks

The most interesting feature sets (subsets a and b, cf. Sections 3.1 and 4.1) are tested as predictors for resistance R or the x-axis intercept respectively. In addition, the feature best ranked among a and b (maximum r) was assessed as one-dimensional predictor. Feed-forward networks are employed with standard

 $^{^1\}mathrm{Available}$ online at www.cs.huji.ac.il/labs/learning/code/fsr/

backpropagation as learning algorithm. As smaller networks tend to generalize better, the number of hidden units are kept at a minimum by employing 1-2-1, 2-2-1 and 3-2-1 architectures respectively, where the input units reflect the size of the assessed feature subset. Input and output units use linear activation functions, hidden units sigmoidal ones. Features of the assessed subset are extracted for all samples and squashed to match input ranges of the activation functions; squashing is applied to the target, too. Each obtained dataset is then split into training data (50 %) and test data (50 %). After training at a learning rate of 0.00001 for 5,000 epochs, predictions for the test data are evaluated.

3 Results

3.1 Feature ranking

Application of RGS yielded consistent Top 5 feature rankings for datasets T and Mix (Table 1). Ranks four and five were non-consistent for dataset P, where the respective features were interchanged in one out of three runs. Within the Top 3, rankings were consistent among all three datasets. For datasets P and Mix, not only the actual Top 3 were consistent, but also RGS evaluation function values differed only slightly over all runs with varying sample sets (Table 2).

We used a threshold of 2.0 to select subsets as predictor candidates (cf. section 4.1), i.e. only features with evaluation function values consistently higher than 2.0 were selected. By this means, no predictor is obtained from dataset T (Table 2). P and Mix reveal a two- and a three-dimensional predictor:

- a) maximum r, average Δr ;
- b) maximum r, average Δr , maximum ϕ .

		Т			Р		Mix			
rank	run 1	run 2	run 3	run 1	run 2	run 3	run 1	run 2	run 3	
1	$\max \phi$	$\max \phi$	$\max \phi$	$\max r$	$\max r$	$\max r$	$\max r$	$\max r$	$\max r$	
2	$\max r$	$\max r$	$\max r$	avg Δr	avg Δr	avg Δr	avg Δr	avg Δr	avg Δr	
3	avg Δr	avg Δr	avg Δr	avg $\Delta \phi$	avg $\Delta \phi$	avg $\Delta \phi$	$\max \phi$	$\max \phi$	$\max \phi$	
4	var Δr	var Δr	var Δr	var r	$\max \phi$	var r	var Δr	var Δr	var Δr	
5	min Δr	min Δr	$\min \Delta r$	$\max\phi$	var \boldsymbol{r}	$\max\phi$	std Δr	st d Δr	st d Δr	

Table 1: Top 5 rated features based on three independent RGS runs with 5,000 randomly chosen impedance spectra from the datasets T, P and Mix.

	Т				Р		Mix			
rank	run 1	run 2	run 3	run 1	run 2	run 3	run 1	run 2	run 3	
1	1.9661	2.0305	1.9386	2.1471	2.1340	2.1470	2.5517	2.3867	2.4980	
2	1.7651	1.7379	1.7802	2.1432	2.1300	2.1430	2.5445	2.3807	2.4913	
3	1.7614	1.7344	1.7762	1.5759	1.6055	1.5906	2.1069	2.0984	2.1146	
4	1.6620	1.6366	1.6927	1.5634	1.5919	1.5839	1.7816	1.6898	1.7896	
5	1.5485	1.5423	1.5502	1.5510	1.5835	1.5773	1.5918	1.5349	1.5898	

Table 2: Values of the RGS evaluation function for the Top 5 rated features of the datasets T, P and Mix (cf. Table 1).

3.2 Evaluating feature subsets with artificial neural networks

Within a given dataset (T, P, Mix), average errors as well as minimum and maximum errors were relatively consistent in absolute values for all three predictor candidates (Table 3). In relative errors, the three candidates yielded average errors of about 3.3% (dataset T), 2% (P) and 4% (Mix) where reference predictors differed by at least five percent from the target value (Table 4).

Applying the value of r(1.3Hz), which is obtained from simply representing Z(1.3Hz) in polar coordinates, directly as predictor yielded smaller average and maximum errors than applying the value $Z^{re}(1.3\text{Hz})$ directly as predictor but notably larger errors than all assessed feature subsets.

	Т				Р			Mix		
Predictor	avg.	min.	max.	avg.	min.	max.	avg.	min.	max.	
max r	31.75	0.00	146.8	143.3	0.01	871.0	98.59	0.00	973.3	
a	31.94	0.01	147.7	140.6	0.08	891.0	96.49	0.00	956.3	
ь	29.50	0.00	141.3	136.3	0.00	817.6	96.96	0.00	940.8	
r(1.3 Hz)	86.85	21.29	248.7	543.2	169.6	1507	303.4	21.29	1507	
$Z^{re}(1.3 \text{Hz})$	138.6	24.76	447.5	769.8	203.2	2002	438.2	24.76	2002	

Table 3: Absolute errors (Ωcm^2) of ANN predictions with maximum r or feature subsets a, b vs. static reference predictors (on datasets T, P and Mix).

	Т			Р			Mix		
Predictor	avg.	min.	max.	avg.	min.	max.	avg.	min.	max.
max r	3.37	0.00	11.73	2.00	0.00	9.47	4.06	0.00	32.56
a	3.38	0.00	11.80	1.95	0.00	9.80	3.96	0.00	33.95
b	3.24	0.00	13.91	1.93	0.00	8.88	4.25	0.00	54.09
r(1.3 Hz)	8.52	5.00	19.88	7.20	5.00	16.58	7.89	5.00	19.88
$Z^{re}(1.3 \text{Hz})$	13.60	5.58	35.77	10.17	5.89	21.96	11.97	5.58	35.77

Table 4: Relative errors (%) of ANN predictions with maximum r or feature subsets a, b vs. static reference predictors (on datasets T, P and Mix).

4 Discussion

4.1 Identifying feature subsets

It is common practice to perform feature selection on the basis of a threshold [8]. An intuitive way to do so, is to use the most notable consistent difference between the evaluation function values of the ranked features as criterion.

In dataset P, this is observed between ranks 3 and 2, suggesting a threshold between about 1.6 and about 2.1 (Table 2). A smaller but still notable jump can be identified in dataset Mix between ranks 4 and 3, suggesting a threshold between about 1.8 and about 2.1 (Table 2). By this, we justify a general threshold value of 2.0 and thereby to discard ranks with lower RGS evaluation function values. In particular, the one-dimensional predictor maximum ϕ is ignored that would be suggested from the jump between rank 2 and 1 in dataset T (Table 2).

4.2 Evaluating feature subsets

All networks trained with the predictor candidates yielded significantly better accuracy than the reference method of using $Z^{re}(1.3\text{Hz})$ directly as predictor. While using r(1.3Hz) gave slightly better predictions than $Z^{re}(1.3\text{Hz})$, average relative errors for T, P and Mix were about two to three times larger than from ANN predictions (less than five percent, Table 4). Maximum relative errors were significantly lower than those of the references methods, too (Table 4).

On average, subset b yielded lower errors than a for datasets T and P, but larger errors for Mix. At the same time, average and maximum errors of a and b increased with Mix. This could be due to the larger size of the Mix dataset (containing 17,954+16,213 spectra). On the other hand, a gap in the target domain is induced when combining T ($R < 1252 \ \Omega \text{cm}^2$) and P ($R > 3324 \ \Omega \text{cm}^2$), rendering the training data imbalanced and possibly affecting generalization [5].

Training with maximum r did not yield best results for T and P, but performed better than subset a on T and better than a and b on Mix, matching its dominance in the RGS ranking. While maximum r and r(1.3Hz) are identical for most curves, the ANN trained gave better predictions than the directly used value r(1.3Hz). Therefore, the influence of r(1.3Hz) is likely of non-linear nature.

5 Conclusions

We have demonstrated that the x-axis intercept of an impedance spectrum can be predicted with reasonable accuracy from a set of three or less implicit curve features; such features describe a spectrum as a whole and can be calculated easily from measured or modeled data. It has been shown that this approach improves estimating the biomedically relevant resistance R of epithelial cell layers where explicit parameters like the real part of a complex impedance at 1.3 Hz are unreliable. Use of balanced training data is likely to further improve accuracy.

References

- R. E. Bellman. Adaptive Control Processes: A Guided Tour. Princeton University Press, 1961.
- [2] C. M. Bishop. Neural Networks for Pattern Recognition. Clarendon Press, 1996.
- [3] A. H. Gitter et al. Impedance analysis for the determination of epithelial and subepithelial resistance in intestinal tissues. J Biochem Biophys Methods, 37(1-2):35–46, 1998.
- [4] D. Günzel et al. From TER to trans- and paracellular resistance: lessons from impedance spectroscopy. Ann N Y Acad Sci, 1257:142–151, 2012.
- [5] H. He and E. Garcia. Learning from Imbalanced Data. Knowledge and Data Engineering, IEEE Transactions on, 21(9):1263 –1284, 2009.
- [6] J. R. Macdonald and W. B. Johnson. Fundamentals of Impedance Spectroscopy, pp. 1–26. John Wiley & Sons, Inc., 2005.
- [7] A. Navot et al. Nearest neighbor based feature selection for regression and its application to neural activity. In: Advances in Neural Information Processing Systems, vol. 18, pp. 995–1002. MIT Press, 2006.
- [8] Y. Saeys et al. A review of feature selection techniques in bioinformatics. *Bioinformatics*, 23(19):2507–2517, 2007.