Automated Quantification of the Capacitance of Epithelial Cell Layers from an Impedance Spectrum

Thomas Schmid

Department of Computer Engineering
Universität Leipzig
Leipzig, Germany
schmid@informatik.uni-leipzig.de

Dorothee Günzel

Institute of Clinical Physiology
Charité
Berlin, Germany
dorothee.guenzel@charite.de

Martin Bogdan

Department of Computer Engineering
Universität Leipzig
Leipzig, Germany
bogdan@informatik.uni-leipzig.de

Abstract—**Quantifying the intestinal surface area of epithelia is crucial to assess changes in protein expression during disease. A convenient alternative to microscopic evaluation of serial sections is capacitance measurement by impedance spectroscopy. While the underlying theoretical relations are well-known, in practice data scatter considerably decreases precision of estimations. Estimations are even less precise if obtained impedance spectra cannot be approximated by a semicircle. Here, we demonstrate that using machine learning techniques together with detailed modeling of cell layers allows reliable predictions of epithelial capacitance. Our results show that estimates for modeled impedance spectra can be obtained with less than 20 percent relative deviation from the target value. In particular, this is shown for spectra that deviate from a semicircular shape.**

Keywords–Physiology, Epithelia, Impedance Spectroscopy, Artificial Neural Networks, Clustering.

I. INTRODUCTION

The intestinal epithelium is the inner-most cell layer lining the gut wall and forms the primary barrier between the gut contents and the body. To maintain a tight barrier against toxins and pathogens, neighbouring epithelial cells are connected by tight junctions, arrays of transmembrane proteins that seal the space between two neighbouring cells. Acute (e.g., norovirus infection, giardiasis) and chronic intestinal diseases (e.g., Crohns disease, celiac disease) cause a restructuring of the gut mucosa due to loss of damaged surface cells and compensatory cell division within the crypts. Depending on the different rates, mucosal area may be enlarged [1] or reduced [2][3]. When investigating molecular processes underlying these diseases, an exact knowledge of changes in mucosal area is indispensable.

Commonly, mucosal area is determined morphometrically, i.e., by microscopic evaluation of serial sections, a process that is both tedious and time consuming. A much more elegant way is to determine the epithelial capacitance as a surrogate marker for the epithelial surface area. Capacitative properties of a cell are due to the lipid bilayer of the cell membrane. Capacitance of the unit cell membrane is in the order of 1 μF/cm^2 [4] and considered to be widely constant. In epithelial cells, the tight junction divides the plasma membrane into two compartments and as a consequence the total epithelial capacitance (C^{epi}) is composed of two capacitances in series. This subdivision is asymmetrical, as the tight junction is located close the apical side facing the outer environment. Intestinal epithelial cells are columnar (height $\gg$ diameter), therefore the apical membrane area is considerably smaller than the opposing basolateral membrane and C^{epi} is, as a first approximation, proportional to the apical membrane area.

A fast, convenient and noninvasive method to determine electrical properties of tissues is impedance spectroscopy. By measuring current-voltage relationships under alternate current (AC) at frequencies between 1 Hz and 100 kHz, typically 40 to 50 complex impedance values Z are obtained [5]. These spectra are often displayed in so-called Nyquist diagrams (Figure 1a), where the real part $\Re$ of each impedance value is plotted against its imaginary part $\Im$. To explain properties of the measured samples, it is common to derive an equivalent electric circuit. To describe epithelial cell layers, circuits of different degrees of complexity are used [6]. The simplest circuit that incorporates C^{epi} is a resistor-capacitor (RC) circuit (Figure 1b). To reflect physiological polarity of epithelial cells explicitly, two RC subcircuits in series and a resistor in parallel are used (Figure 1c). Electric behavior of the subepithelium may be considered by a further resistor in series.

In previous work, we have demonstrated that conventional analysis of impedance spectra like visual extrapolation of plots can lead to non-neglectable errors in parameter estimations. At the same time, we have shown that the precision in estimating epithelial and subepithelial resistance can be improved substantially for the epithelial cell lines HT-29/B6 and IPEC-J2 by using machine learning techniques [7][8]. Rationale behind this approach is that for a given electric circuit, the theoretical impedance at a given frequency can be calculated if the values of all circuit components are known. As exact target values are known for such synthetic data, too, this data can be used to train artificial neural networks or other machine learning techniques. In order to draw conclusions for data from laboratory measurements, however, optimal modeling of such training data according to the respective cell line and cell conditions is required.

Here, we adapt this approach to efficiently predict C^{epi} from an impedance spectrum. The tasks to be carried out include cell line modeling, feature selection, training and evaluation of the predictions. As model cell line, we investigate Madin-Darby canine kidney cells type I (MDCK-I). These cells have been studied since the 1960s [9] and are typically described as possessing a high transepithelial resistance [10].
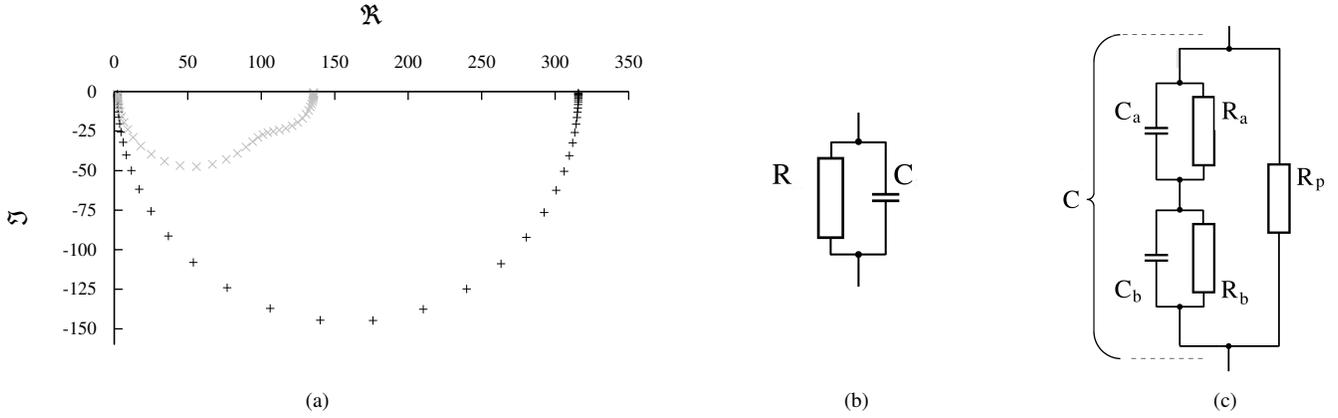
Figure 1. (a) Overlay of a semi- and a nonsemicircular Nyquist plot where real ($\Re$) and imaginary ($\Im$) part of each complex-valued impedance Z are plotted against each other. The displayed impedance spectra reflect AC application at $n=42$ frequencies between 1.3 and 16,000 Hz on an epithelial cell layer with capacitance $1/C = 1/C_a + 1/C_b$. (b) A simple resistor-capacitor (RC) circuit that can be used as equivalent circuit. This circuit yields semicircular impedance spectra. (c) Equivalent electric circuit discriminating between apical and basolateral properties of an epithelial cell layer. This circuit can yield semicircular or nonsemicircular impedance spectra.

II. METHODS

A. Modeling Impedance Spectra

To model realistic impedance spectra for a given cell line, three prerequisites are required: a) a matching equivalent electric circuit, b) appropriate ranges for the parameters of the circuit, and c) an error model that reflects the data scatter intrinsic to the electrophysiological measurement set-up.

The equivalent circuit considered here (Figure 1c) 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 impedance Z of an electric circuit at an angular frequency ω can be derived from the impedances of its components:

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

where $i = \sqrt{-1}$, and $\tau_a = R_aC_a$ and $\tau_b = R_bC_b$.

In practice, however, the electrophysiological set-up used for measurements induces data scatter and thus systematic deviation from the theoretical impedance value. In order to mimic realistic data, such systematic deviations can be modeled as function of the transepithelial resistance R^T [7] and added to $Z(\omega)$. For simplicity, such data scatter was not considered here.

Using $n = 42$ frequencies (1.3 to 16,350 Hz), n tupels of real and imaginary parts $((\Re(\omega_0), \Im(\omega_0)), \ldots, (\Re(\omega_{n-1}), \Im(\omega_{n-1})))$, are obtained from measurements or calculations, respectively. Alternatively, complex impedance values can be transformed into polar coordinates, i.e., into phase angle ϕ and magnitude r $(((\phi(\omega_0), r(\omega_0)), \ldots, (\phi(\omega_{n-1}), r(\omega_{n-1}))))$.

Real and imaginary parts of a spectrum can be regarded as separate feature sets $S_\Re$ and $S_\Im$:

$$S_\Re = \{\Re(\omega_0), \ldots, \Re(\omega_{n-1})\} \quad (2)$$

$$S_\Im = \{\Im(\omega_0), \ldots, \Im(\omega_{n-1})\} \quad (3)$$

Analogously for phase angles and magnitudes:

$$S_\phi = \{\phi(\omega_0), \ldots, \phi(\omega_{n-1})\} \quad (4)$$

$$S_r = \{r(\omega_0), \ldots, r(\omega_{n-1})\} \quad (5)$$

B. Sampling the MDCK-I Cell Line

For MDCK-I cells published values for the transepithelial resistance R^T range from 1500 to 14000 Ωcm^2 [11]. For the parameters R_p, R_a, R_b, C_a and C_b to the best of our knowledge published estimates exist neither for physiological conditions nor for drug applications. Therefore these parameter ranges were initially estimated from laboratory experiences and evaluated and optimized analogously to [7]. As two distinct cell conditions, physiological conditions ("Control") and conditions after the application of EGTA ("EGTA") were modeled. A table of the final parameters can be found in the appendix.

Estimating C^{epi} for semi-circular spectra is often considered a simple task with little error potential. At the same time, reliable estimation of C^{epi} for spectra deviating from this shape is thought to be considerably more difficult. Reassessing this assumption, we considered both cases individually and separated spectra reflecting control and EGTA conditions accordingly. As separation criterion, we assumed that spectra possessing greatly asymmetrical time constants express a non-semicircular shape. To this end, we defined the τ ratio of the used electric circuit (Figure 1c) as the larger time constant divided by the smaller time constant, and a nonsemicircular shape was assumed for data with a τ ratio greater than five. This parameter can not only be calculated directly for modeled impedance spectra, but also be predicted with good precision for measured spectra [12].

C. Reference Methods to Determine Epithelial Capacity

Analogously to our previous work [7], we used two different conventional approaches as reference methods for estimating the parameter C^{epi}. Additionally, we employed a theoretical relation of the underlying circuit.

1) Nearest Data Point (Method M1): Assuming that a semicircular shape results from a single RC circuit (Figure 1b), $C^{epi} = 1/(\omega_c R)$ holds true. ω_c is the characteristic frequency, at which the spectrum reaches its minimal turning point. The frequency related the minimum of $S_\Im$ was used to approximate ω_c and the maximum value of $S_\Re$ to approximate $R = R^{epi}$.

2) Frequency-blind Circle Fit (Method M2): Analogously to M1, C^{epi} was calculated from the substitute parameters ω_c and R^{epi}. A Cole-Cole fit [13] was carried out on a Nyquist diagram, i.e., a circle was fitted as described by Kasa [14]. The frequency of the data point nearest to the circle center was used to approximate ω_c; the intercept with the x-axis at the low frequency end was taken as R^{epi}.

3) High-Frequency Limit Approximation (Method M3): Given the electric circuit in Figure 1b or 1c, respectively, the theoretical high-frequency limit for the imaginary part of the impedance is the reciprocal of the overall capacitance [15]:

$$-\lim_{\omega\to\infty}\omega\mathfrak{I}(\omega)=\frac{1}{C} \tag{6}$$

Of all 42 impedances obtained here, the data point with the highest frequency was $Z_{42}=\mathfrak{R}(\omega_{42})-i\mathfrak{I}(\omega_{42})$. Thus, we used the value given by $-1/(\omega_{42}\mathfrak{I}(\omega_{42}))$ to approximate C^{epi}.

D. Machine Learning Approach

The given reference methods represent two distinct estimation approaches: solving either a primarily geometric fitting problem (M2) or an idealized physics formula with error-prone data (M3). Therefore, we investigated these two representations of the same problem by two individual machine learning approaches. Also, we considered semicircular and nonsemicircular spectra as separate problem domains. For each representation and domain, we assessed the prediction quality of the respective problem representation by decision trees (using *R* and the package *rpart*), artificial neural networks with backpropagation (using the FORWISS Artificial Neural Network Toolbox [16]) and random forests (using *R* and the package *randomForest* [17]).

1) Training data: For the semicircular, as well as for the nonsemicircular domain, a sample of 30,000 random spectra was selected each. Data for each domain was split into a training dataset of 20,000 spectra (circa 66 percent) and a test dataset of 10,000 spectra (circa 33 percent), respectively.

As geometric data representations, either cartesian ($\perp$) or polar ($\angle$) coordinates were used where

$$S_\perp=\{\mathfrak{R}_0,...,\mathfrak{R}_{n-1},\mathfrak{I}_0,...,\mathfrak{I}_{n-1}\} \tag{7}$$

$$S_\angle=\{r_0,...,r_{n-1},\phi_0,...,\phi_{n-1}\} \tag{8}$$

As representation related to the high-frequency limit of the imaginary part, we transformed the spectra into the preciprocal products of frequencies and imaginary parts (cf. (6)):

$$S_{\omega\mathfrak{I}}=\left\{-\frac{1}{\omega_0\mathfrak{I}(\omega_0)},...,-\frac{1}{\omega_{n-1}\mathfrak{I}(\omega_{n-1})}\right\} \tag{9}$$

Note that while polar and cartesian representations possessed a total of $2\cdot n$ features, $S_{\omega\mathfrak{I}}$ possesses only n features.

2) Algorithm settings: Given a multivariate regression task, decision trees were created using analysis of variance ("anova" method). As ANNs, multilayer perceptrons (MLP) with one hidden layer were used. Depending on the number of input features, a $2n$-20-1 or n-10-1 architecture was used where hidden units employed sigmoid and input and output units linear activation functions; as learning algorithm Quickprop [18] was used. For random forests, 50 trees were used and variable importance was assessed by 25 consecutive runs of

the *Boruta* algorithm (using the *R* package *Boruta* which searches all relevant variables by iterative removal of features that are statistically less relevant than random probes [19]). For evaluation, only test data was used. As exact target values were known, predictions were evaluated by relative deviation from the target, i.e., in percent.

3) Clustering: Using the best performing data representation, the sample was clustered by *k-means* where $k=\{1,..,10\}$; semicircular und nonsemicircular spectra were clustered separately. For each clustering, individual clusters were evaluated by decision trees. The best clustering for semi- and nonsemicircular data was determined by the least predictive cluster of each clustering and re-evaluated with ANNs.

III. Results

A. Evaluation of Reference Methods

Estimations of C^{epi} for nonsemicircular spectra showed in general greater deviations from the target value than those for semicircular spectra. For semicircular spectra, M2 showed least maximum deviations, while M3 showed least interquartile distance (Figure 2). For nonsemicircular spectra, M3 showed both least maximum deviations and least interquartile distance (Figure 3). Numerically, the maximum deviations of M3 was 90% for semicircular and 222% for nonsemicircular spectra.
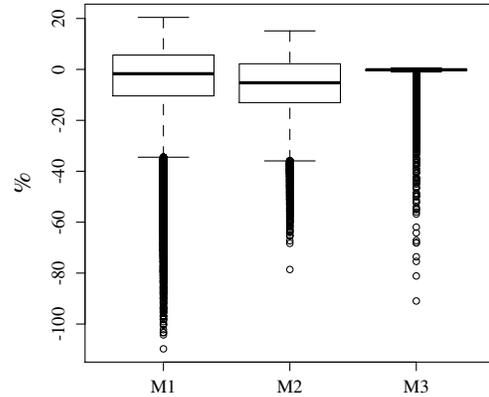
Figure 2. Relative deviation from the target value using reference methods M1, M2 and M3 for semicircular test data (n=30,000).
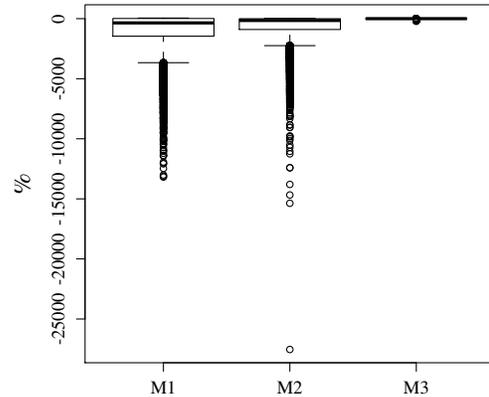
Figure 3. Relative deviation from the target using reference methods M1, M2 and M3 value for nonsemicircular test data (n=30,000).

B. Predictions of C^{epi}

Application of decision trees yielded maximum relative deviations from the target between 42 and 80 percent for semicircular and between 80 and 229 percent for nonsemicircular spectra. Application of ANNs yielded maximum relative deviations between 23 and 40 for semicircular and between 37 and 167 percent for nonsemicircular spectra. Application of random forests yielded constantly maximum relative deviations larger than the largest maximum relative deviation observed for decision trees; these results are therefore omitted here.

TABLE I. RELATIVE DEVIATION OF PREDICTIONS FROM THE TARGET VALUE C^{epi} FOR SEMICIRCULAR SPECTRA (IN PERCENT).

	cartesian		polar		high-frequency limit	
	tree	ANN	tree	ANN	tree	ANN
Minimum	-15.4	-39.1	-14.7	-40.3	-15.1	-23.5
1. Quartile	-4.6	-1.0	-4.1	-0.8	-4.2	-0.2
Median	0.1	0.1	0.4	0.1	0.3	-0.1
Mean	0.5	0.0	0.6	0.0	0.7	-0.1
3. Quartile	4.4	1.0	4.3	0.8	4.4	0.1
Maximum	80.7	29.9	42.1	30.7	80.6	11.4

TABLE II. RELATIVE DEVIATION OF PREDICTIONS FROM THE TARGET VALUE C^{epi} FOR NONSEMICIRCULAR SPECTRA (IN PERCENT).

	cartesian		polar		high-frequency limit	
	tree	ANN	tree	ANN	tree	ANN
Minimum	-22.5	-68.6	-20.0	-70.5	-22.2	-37.3
1. Quartile	-5.3	-1.7	-4.1	-1.6	-5.4	-0.1
Median	0.1	-0.2	0.1	-0.2	0.0	0.0
Mean	1.2	-0.1	0.6	-0.1	1.4	0.0
3. Quartile	4.8	1.7	4.4	1.4	5.0	- 0.2
Maximum	229.4	28.1	80.8	36.6	167.1	21.6

C. Variable Importance

Assessing the three data representations $S_\perp$, $S_\angle$ and $S_{\omega\mathfrak{I}}$ using the *Boruta* algorithm, neither for the semicircular nor for nonsemicircular domain a relevant number of features was removed. For the high-frequency limit representation $S_{\omega\mathfrak{I}}$ of the semicircular domain, e.g., only one feature was removed while 41 features were kept as relevant. Consequently, these findings were not used for explicit feature selection.

In all three data representations, however, analysis of variable importance showed that features reflecting the highest five frequencies yielded higher variable importance than the remaining features. This was observed for the semicircular, as well as for the nonsemicircular domain and was exploited in the next step of the analysis.

D. Cluster evaluation

K-means clustering was applied either to the full (84 or 42 features, respectively) or partial data representation (five features). When evaluating with decision trees, most clusterings did not yield less maximum deviations from the target value C^{epi} than seen in previous evaluations (Table I and II). An exception was the $S_{\omega\mathfrak{I}}$ representation clustered by the five features related to the five highest frequencies. For a number of five clusters, the highest maximum deviation observed for all clusters was 30.1 percent for semicircular spectra (Figure 4) and 58.8 percent for nonsemicircular spectra (Figure 5).
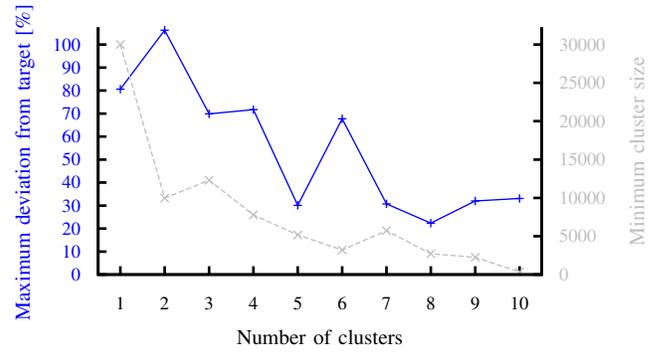
Figure 4. Cluster analysis for semicircular test data split into a variable number of clusters by k-means.
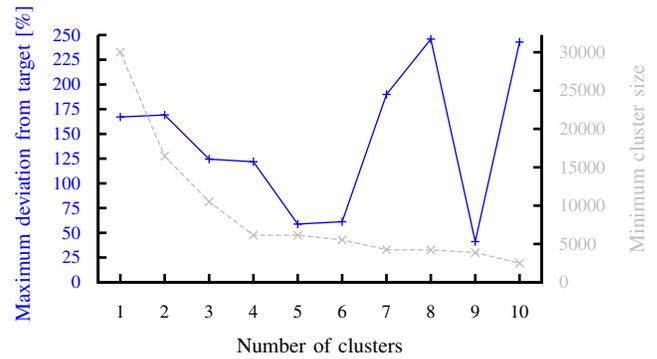
Figure 5. Cluster analysis for nonsemicircular test data split into a variable number of clusters by k-means.

E. Cluster-based estimations

Employing the findings from the cluster analysis, we split the $S_{\omega\mathfrak{I}}$ representation of the sample data into five clusters (cf. section III.D). For each cluster, an individual ANN was trained (analogously to section II.D). Relative deviations of the predictions did not exceed ±15.9 percent for any cluster of semicircular spectra (Table III), and did not exceed ±18.1 percent for nonsemicircular spectra (Table IV).

TABLE III. RELATIVE DEVIATION OF PREDICTIONS FROM C^{epi} FOR CLUSTERED SEMICIRCULAR SPECTRA (IN PERCENT).

Cluster	1	2	3	4	5
Minimum	-5.1	-15.9	-16.4	-3.3	-2.7
1. Quartile	-0.1	-0.1	-0.2	-0.1	-0.2
Median	0.0	0.0	0.0	0.0	0.0
Mean	0.0	0.0	0.0	0.0	0.0
3. Quartile	0.1	0.1	0.2	0.1	0.2
Maximum	8.9	6.3	6.2	8.3	2.3
$n_{cluster}$	7413	4526	7659	5240	5162

TABLE IV. RELATIVE DEVIATION OF PREDICTIONS FROM C^{epi} FOR CLUSTERED NONSEMICIRCULAR SPECTRA (IN PERCENT).

Cluster	1	2	3	4	5
Minimum	-9.3	-18.1	-7.4	-8.2	-11.3
1. Quartile	-0.1	-0.1	-0.2	-0.2	-0.1
Median	0.0	0.0	0.0	0.0	0.0
Mean	0.0	0.0	0.0	0.0	0.0
3. Quartile	0.1	0.2	0.2	0.2	0.1
Maximum	9.6	15.0	7.9	10.8	5.3
$n_{cluster}$	7191	7190	2969	6518	6132

(a)

(b)
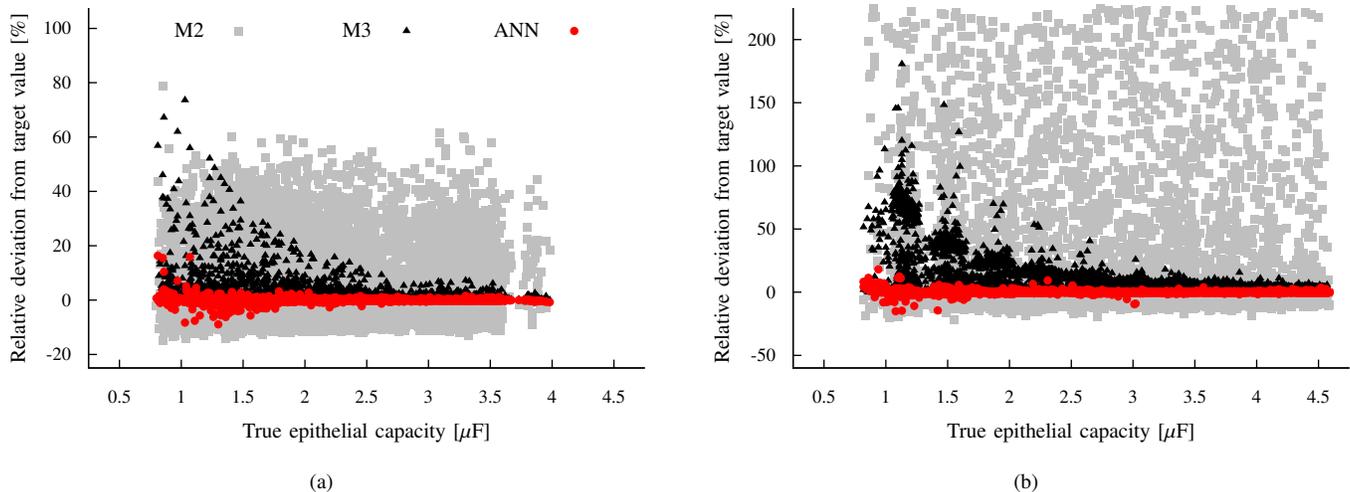
Figure 6. Relative deviations of estimated from true values for C^{epi} using reference methods M2 (grey ■) and M3 (black ▲) and a cluster-based ANN approch (red ●) are plotted against the respective true target value; method M1 is omitted as estimations are considerably less precise than those by M2 and M3 (cf. Figures 2 and 3). Values shown refer to (*a*) semicircular and (*b*) nonsemicircular spectra obtained either under physiological conditions (control) or after application of EGTA; to discriminate semi- from nonsemicircular spectra the relation between apical and basolateral time constants was used, i.e., a value greater than five was taken as indicator of nonsemicircular shape.

IV. DISCUSSION

A. Evaluation of Reference Methods

As expected, the naive reference methods M1 and M2 failed to provide precise estimations for C^{epi} on nonsemi-circular spectra. Interestingly, however, estimations can also exhibit large relative deviations of more than 50 percent from the target value when applied to semicircular spectra. And while method M3 is also a rather rough estimation method with up to 80 percent relative deviation in the same task, less than 5 percent for quartiles of relative deviations indicate a remarkable specificity compared to M1 and M2. This becomes even more obvious with differences in prediction quality for the nonsemicircular spectra. In both domains, however, maximum relative deviations of M3 are by far too great to allow reliable predictions.

B. Comparison of Problem Representations

In the first step of this study, we compared learning from geometric data representations $S_\perp$ and $S_\angle$ to learning from the physical data representation $S_{\omega\Im}$. ANNs yielded better predictions than decision trees in both representations, as well as in both problem domains. In the semicircular, as well as in the nonsemicircular domain, median and average relative deviations of all three representations did not exceed one percent of the target value (in one case: 1.2 percent). On the maximum relative deviation, however, ANNs using $S_{\omega\Im}$ performed better than ANNs using $S_\perp$ and $S_\angle$. Usefulness of this representation is not surprising considering that estimations with reference method M3 already exhibited little inter-quartile distance. It is likely that this excellent performance is due to the immediate physical relation between the features of $S_{\omega\Im}$ and C^{epi}.

C. Clustering

In the second step of this study, we aimed to optimize ANN predictions based on the physical representation $S_{\omega\Im}$.

As optimization criterion, we considered the predictiveness of the least predictive cluster, respectively; this was measured by decision trees and the relative deviation of predicted values form the target value. At the same time, the number of spectra per cluster was intended to be as large as possible; naturally, the number of spectra decreases with increase of the number of clusters. As can be seen in Figures 4 and 5, both goals are achieved by choosing *k-means* clustering with $k = 5$; in particular, this holds true for both semicircular and nonsemicircular spectra.

D. Cluster-based Estimations

While clustering analysis was carried out with decision trees, the optimal clustering ($k = 5$) was evaluated by ANNs afterwards to further improve predictions. As in the first step, evaluations were performed for semicircular and nonsemicircular spectra separately. ANN estimations yielded maximum relative deviations of less than 20 percent within all clusters. Moreover, in all cases median and average relative deviations were 0 percent and inter-quartile distance less than 0.5 percent points. Compared to ANN predictions on unclustered data (Tab. I and II), this is a notable improvement. Even more remarkable is the improvement compared to the two best performing reference methods M2 and M3 (Figure 6).

V. CONCLUSIONS

Impedance spectroscopy is a convenient method to determine the capacitance of an epithelial tissue. In practice, however, this clinically important parameter can only be roughly approximated from impedance data, as common estimation methods fail to provide reliable estimations. Here, we have shown that our approach of modeling cell properties and applying machine learning techniques is a fruitful approach for this task. For impedance spectra modeled after the epithelial cell line MDCK-I, we developed a cluster-based neural network approach that shows a maximum relative deviation from the

theoretical target of less than 20 percent. In future work, we will apply this approch to other epithelial cell lines, as well as native tissue and further optimize estimations.

APPENDIX

Impedance spectra for the cell line MDCK-I were calculated for two distinct cell states. Physiological conditions (control) and conditions after application of EGTA (EGTA) were modeled separately according to (1) using the parameter ranges in Table I. Note that for the given electric circuit $R_t = R_a + R_b$ and $C^{epi} = \frac{C_a \cdot C_b}{C_a + C_b}$.

The parameter interval for R^{epi} was 10 to 2,000 Ωcm^2 and 10 to 200 Ωcm^2 for control and EGTA, respectively; all other intervals were chosen dynamically to yield ten values per range. By this, a total of 1,865,823 and 1,684,784 spectra were produced for control and EGTA, respectively.

TABLE V. PARAMETER RANGES FOR A MDCK-I-EQUIVALENT CIRCUIT.

	R^{epi}	R_p [Ωcm^2]	R_t	C^{epi}	C_a [$\mu F/cm^2$]	C_b
Control	10–2000	10–10000	10–5000	0.5–5.0	1–5	1–75
EGTA	10–200	10–250	10–5000	0.5–5.0	1–5	1–75

From all of these spectra, 137,162 possessed a τ quotient less than five ($\sim$ semicircular) and 3,413,445 possessed a τ quotient larger than five ($\sim$ nonsemicircular). From these, a sample of 45,000 semicircular and a sample of 45,000 nonsemicircular spectra were randomly selected and analyzed.

To confirm correctness of our model, congruency with impedance measurements from laboratory experiments on MDCK-I cells was evaluated graphically as previously described [7]. During these experiments, R^{para} had been manipulated by the application of EGTA; 56 spectra were recorded before EGTA application, 49 after application. As reference data 25,000 modeled spectra from each condition were used.

REFERENCES

[1] S. Zeissig et al., "Changes in expression and distribution of claudin 2, 5 and 8 lead to discontinuous tight junctions and barrier dysfunction in active Crohn's disease," Gut, vol. 56, no. 1, 2007, pp. 61–72.

[2] H. Troeger et al., "Structural and functional changes of the duodenum in human norovirus infection." Gut, vol. 58, no. 8, 2009, pp. 1070–1077.

[3] H. Troeger et al., "Effect of chronic Giardia lamblia infection on epithelial transport and barrier function in human duodenum," Gut, vol. 56, no. 3, 2007, pp. 328–335.

[4] H. Fricke, "The electric capacity of suspensions of red corpuscles of a dog," Physical Review, vol. 26, no. 5, 1925, pp. 682–687.

[5] J. R. Macdonald and W. B. Johnson, Fundamentals of Impedance Spectroscopy. John Wiley & Sons, Inc., 2005, pp. 1–26.

[6] D. Günzel et al., "From TER to trans-and paracellular resistance: lessons from impedance spectroscopy," Annals of the New York Academy of Sciences, vol. 1257, no. 1, 2012, pp. 142–151.

[7] T. Schmid, M. Bogdan, and D. Günzel, "Discerning apical and basolateral properties of HT-29/B6 and IPEC-j2 cell layers by impedance spectroscopy, mathematical modeling and machine learning," PLOS ONE, vol. 8, no. 7, 2013, p. e62913.

[8] T. Schmid, D. Günzel, and M. Bogdan, "Efficient prediction of x-axis intercepts of discrete impedance spectra," in Proceedings of the 21st European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN), 2013, pp. 185–190.

[9] C. R. Gaush, W. L. Hard, and T. F. Smith, "Characterization of an established line of canine kidney cells (MDCK)." Proc Soc Exp Biol Med, vol. 122, no. 3, 1966, pp. 931–935.

[10] B. R. Stevenson, J. M. Anderson, D. A. Goodenough, and M. S. Mooseker, "Tight junction structure and ZO-1 content are identical in two strains of Madin-Darby canine kidney cells which differ in transepithelial resistance." J Cell Biol, vol. 107, no. 6 Pt 1, 1988, pp. 2401–2408.

[11] M. Furuse, K. Furuse, H. Sasaki, and S. Tsukita, "Conversion of zonulae occludentes from tight to leaky strand type by introducing claudin-2 into Madin-Darby canine kidney I cells," Journal of Cell Biology, vol. 153, no. 2, 2001, pp. 263–272.

[12] T. Schmid, D. Günzel, and M. Bogdan, "Automated quantification of the relation between resistor-capacitor subcircuits from an impedance spectrum," in Proceedings of the 7th International Conference on Bio-Inspired Systems and Signal Processing, 2014, pp. 141–148.

[13] K. S. Cole and R. H. Cole, "Dispersion and absorption in dielectrics I. alternating current characteristics," The Journal of Chemical Physics, vol. 9, no. 4, 1941, pp. 341–351.

[14] I. Kasa, "A circle fitting procedure and its error analysis," Instrumentation and Measurement, IEEE Transactions on, vol. 1001, no. 1, 1976, pp. 8–14.

[15] M. E. Orazem and B. Tribollet, Eds., Electrochemical Impedance Spectroscopy. Wiley, 2008, ch. Methods for Representing Impedance, pp. 309–331.

[16] M. K. Arras and K. Mohraz, FORWISS Artificial Neural Network Simulation Toolbox v.2.2, Bayerisches Forschungszentrum für wissensbasierte Systeme, Erlangen, Germany, 1996.

[17] A. Liaw and M. Wiener, "Classification and regression by randomforest," R news, vol. 2, no. 3, 2002, pp. 18–22.

[18] S. E. Fahlman, "Faster-learning variations on Back-propagation: An empirical study," in Proceedings of the 1988 Connectionist Summer School. San Mateo, CA: Morgan Kaufmann, 1988, pp. 38–51.

[19] M. B. Kursa and W. R. Rudnicki, "Feature selection with the Boruta package," Journal of Statistical Software, vol. 36, no. 11, 2010, pp. 1–13.