Time Series Features for Classification of Contaminated Cell Cultures

Laura L. Tupper
Mount Holyoke College
South Hadley, Massachusetts
ltupper@mtholyoke.edu

Charles R. Keese
Applied BioPhysics, Inc.
Troy, New York
keese@biophysics.com

David S. Matteson
Cornell University
Ithaca, New York
matteson@cornell.edu

ABSTRACT

We examine the use of time series data, derived from Electric Cell-substrate Impedance Sensing (ECIS), to differentiate between standard mammalian cell cultures and those infected with a mycoplasma organism. We perform feature-based classification, extracting interpretable features from the ECIS time courses, and achieve high classification accuracy using only two features at a time. Initial results also show the existence of experimental variation between plates and suggest types of features that may prove more robust to such variation. Our paper is the first to perform a broad examination of ECIS time course features in the context of detecting contamination, and to describe and suggest possibilities for ameliorating plate-to-plate variation.

CCS CONCEPTS

• **Mathematics of computing** → **Cluster analysis**; *Exploratory data analysis*; • **Applied computing** → *Computational biology*.

KEYWORDS

Time series classification, Feature-based classification, Contamination of cell cultures, Biophysics, Electric Cell-substrate Impedance Sensing

ACM Reference Format:
Laura L. Tupper, Charles R. Keese, and David S. Matteson. 2021. Time Series Features for Classification of Contaminated Cell Cultures. In *MileTS '21: 7th KDD Workshop on Mining and Learning from Time Series, August 14th,2021, Singapore.* ACM, New York, NY, USA, 7 pages.
https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

1.1 Motivation: cell contamination

The study of cells in culture is a vital component of biological research, allowing the examination of cells' physical morphology, their patterns of growth and process through the life cycle, and their responses to stimuli and the environment. Yet there is a reproducibility crisis in cell culture research, in large part due to misidentification and contamination of cell samples (see [13]). One common issue is the contamination of mammalian cells with other microorganisms such as *Mycoplasma*, which can flourish in the medium used to grow the cells and create misleading results.

In this paper, we address the issue of contaminated cell cultures as a classification problem. We draw on an automated, non-invasive data collection method, electric cell-substrate impedance sensing, to generate time series corresponding to individual cell cultures. We then investigate the performance of features derived from these time series in classifying cell cultures as infected or uninfected.

1.2 Electric Cell-substrate Impedance Sensing

Electric cell-substrate impedance sensing, or ECIS®, is an established tool for measuring the behavior and characteristics of cells over time. An introduction can be found in [10], and surveys of some of the many applications can be found in [18] and [12]. ECIS is far less labor intensive than optical assessment of cell cultures, requires no chemical labels or markers to be applied to the cells, and does not damage the cells or change their morphology, allowing for sustained observation of the same culture.

In the ECIS technique, an experimenter cultivates a cell culture using a tray of several wells, each of which is fitted with an electrode. The tray is placed in a machine that passes a weak AC current through the electrodes, at a chosen frequency, and repeatedly measures the impedance, resistance and capacitance over time. Over time, the cells multiply, move, and eventually reach *confluence*, complete coverage of cells across the entire well. If desired, the machine can *wound* the cultures by briefly applying high voltage. This process kills most of the cells located on top of the electrodes, causing further changes in impedance as they are replaced by new cells.

Many studies have used ECIS data to examine specified cell types or processes, but quantitative discussion has been sparse. [8] appear to be the first to examine quantitative features of ECIS data across multiple cell lines, with the goal of identifying unknown or potentially-mislabeled cultures. [28] examine contamination data, comparing long-memory behavior and the timing of the confluence stage in infected and uninfected cells. We build on this work by bringing in many other types of features, and by comparing their efficacy in identifying contamination in multiple cell types.

1.3 The current dataset

In this paper, we use data provided by Applied BioPhysics, Inc., generated using the ECIS Zθ machine. Cells are grown on arrays, or plates, each containing 96 separate wells; measurements are obtained at several AC frequencies between 500 and 32000 Hz.

We primarily examine MDCK II cells, some of which are contaminated with *Mycoplasma hominis*. Cells are grown on one of two substrates: either with a gelatin coating, or with an adsorbed layer of BSA (bovine serum albumin). Some wells of each type are left

MileTS '21, August 14th, 2021, Singapore

ACM ISBN 978-1-4503-9999-9/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

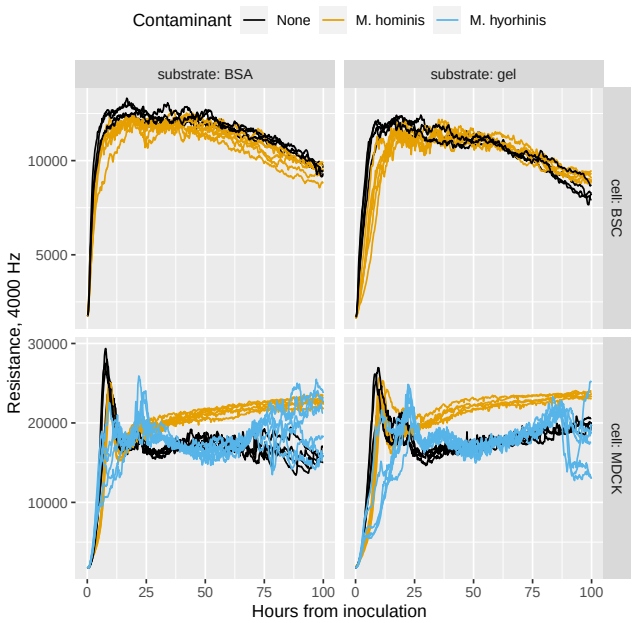

Figure 1: Sample ECIS data: time courses of resistance. These time courses differ based on the cell type, the substrate on which it is grown, and the presence and type of contaminating organisms.

empty, containing substrate but no cells, to provide a baseline. As an extension to our original investigation, we also examine cultures from a different cell line, BSC-1, some of which are infected with *M. hominis*, and MDCK II cells infected with a different species of mycoplasma, *M. hyorhinis*.

2 FEATURE-BASED CLASSIFICATION

2.1 The feature-based approach

Instead of calculating similarity or distance "pointwise" between individual time points in the time series, we reduce the effective dimension by generating a limited set of features from the original data. We may then perform classification or clustering using these features, or a subset of them. There is an extensive body of literature on feature generation and selection; an overview and typography of many approaches appears in [4]. [20] provide an extensive discussion of feature-based approaches specifically for the classification of time series data.

One strength of the feature-based approach is the reduction of dimensionality, as argued in [27], which can avoid many of the problems involved in high-dimensional clustering (for examples, see [1]). We also avoid the problematic requirement that the time series under consideration be exactly the same length, as mentioned in [15]. Moreover, the feature-based approach means that the user need only consider a few features to understand why an observation is classified in a particular way. We can easily visualize the distribution of scores across observations, and if we perform classification using pairs of features, we can also visualize the classification regions in the feature space.

2.2 Types of features

For this project, we generate a variety of features from the ECIS time courses. Some reflect characteristics of ECIS data mentioned as useful in previous studies, while others are based on time series analysis or on capturing particular stages of the cells' behavior. The basic types of features are described here; a more complete and detailed list can be found in the supplemental materials.

Several of our literature-inspired features look at the level of resistance at a certain time and frequency, or simple differences in resistance over time. The rate of increase in resistance early on, shortly after inoculation, is mentioned in studies such as [21], [16], and [22]. Post-confluence levels, meanwhile, are used in [16], [11], and [2], while the overall peak value of resistance appears in [21]. Several studies (including [18], [11], and [24]) address cells' responses to wounding, both its immediate effects and the time to recovery; we extend this idea by comparing the post-wounding recovery process to the initial growth stage.

To characterize post-confluence behavior more generally, we turn to time series methods. Using ARIMA model coefficients and errors for clustering has appeared in a variety of time series application areas (see [19] and [5]), while [7], [27], [3], and [6] use the ACF or PACF. The estimates of error variance from ARIMA models can also be useful in reflecting short-term erratic behavior.

In addition to attachment and spreading behavior, cells also exhibit *micromotion*: small-scale shifts in their attachments to the substrate, even after confluence has been reached and large-scale movement is impossible. Several papers (such as [10], [18], and [25]) examine this behavior, which can be observed in ECIS data as small fluctuations in impedance after a stable level has been reached. [17] and [23] examine long-memory behavior, while [28] looks at the regime change between early growth and a post-confluence long-memory system.

2.3 Variations of feature types

For the broad feature search, we calculate each feature on the time course of resistance. Basic feature types are calculated both for the raw time courses and using a rolling windowed average over 5 consecutive time points, intended to make the feature scores robust to momentary noise in the data. We also introduce a normalization procedure to help account for variations in conditions over time, by taking the ratio of each time course to measurements from "empty wells" on the same tray. All features are calculated for all frequencies, since measurements at different frequencies may reflect different components of cell behavior, as discussed in [18] and [26].

3 CLASSIFICATION AND RESULTS

3.1 Classification process

We attempted three methods for performing classification on the feature data: classification trees, linear discriminant analysis (LDA), and quadratic discriminant analysis (QDA). A thorough discussion of these methods appears in [8], including an exploration of different parameter choices for discriminant analysis. From the 476 features and feature variations, we obtain over 113,000 possible feature pairs on which to base classifiers. For each feature pair, we split the observations evenly into training and testing sets, and

record the success rate. Performance values given here are averages over 10 different training/test splits, using LDA.

3.2 General classification results

An example of the LDA classification results is shown in Figure 2, using two simple features: resistance at 2 hours after inoculation (measured at a frequency of 4000 Hz) and the maximum resistance during the first 24 hours after inoculation (at 32000 Hz). The classifier uses the training observations (square points) to generate a posterior probability of contamination for each point in the feature space, and a linear division of the feature space into infected and uninfected regions. The test observations (round points) are all correctly classified based on their coordinates in this feature space, and we can see that there is very clear separation between contaminated and uncontaminated cultures in terms of these features.

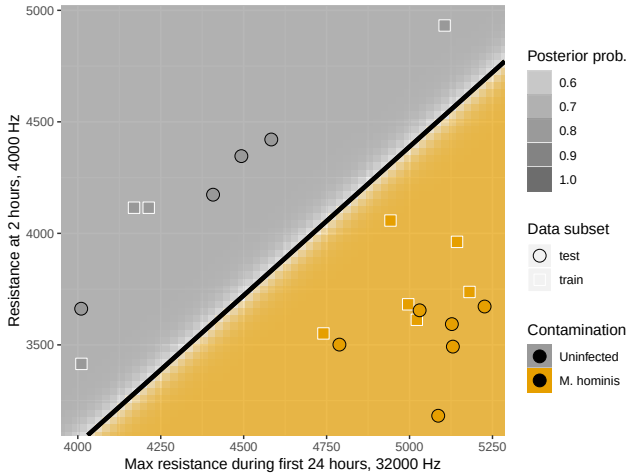

Figure 2: LDA classification regions based on one pair of features. This pair yields perfect classification of the test observations.

In general, classification accuracy is high. There is no "best pair" of features: many different feature pairs offer strong, and in many cases equal, performance. For example, on our sample of BSA-treated wells, about 27% of all feature pairs yield perfect classification accuracy. This is a promising indication of the effectiveness of ECIS measurements for distinguishing contaminated cultures, though the problem becomes more difficult once we consider experimental variation across plates, as discussed below.

The protein coat on which the cells are grown does affect their behavior: some features are more effective for classification of cells grown with BSA, while others are more effective on gel-treated wells. Since it is a simple matter in practice to grow cells using one or the other substrate, high performance on either group of cultures is equally useful.

3.3 Feature combinations

Figure 3 shows a heatmap of performance values for a subset of feature pairs, with the brighter tiles representing feature pairs with higher accuracy. This subset illustrates the high performance that can be achieved with many different features, but also suggests rules for combining features, based on the "sub-diagonals" of poorer performance that can be seen throughout.

Some of these low-performance areas correspond to pairs consisting of a feature and a modified version of the same feature: for example, *at2hR32000* (resistance at 2 hours, 32000 Hz) paired with *w5c2hR32000* (resistance averaged over a window at 2 hours, 32000 Hz). Other low-performance cases appear when the pairing consists of the same feature calculated on two different versions of the data: the original time series and the version normalized to empty wells (features beginning with "EN"). Evidently, these feature modifications do not greatly change the information in the feature, so that pairing a feature and its modified version is essentially equivalent to using a single feature for classification.

To a lesser extent, we also see relatively weak performance when a feature is paired with itself at a similar frequency, or with a modification of the same feature at a similar frequency. For example, *at2hR32000* and *at2hR16000* are not an effective pair, but each yields perfect classification accuracy if paired with *at2hR500*. This provides further evidence that more divergent frequencies provide more useful information, by reflecting different characteristics of the cells' behavior.

3.4 Individual feature performance

We can obtain a simple measure of the effectiveness of individual features by averaging the classification accuracy over all pairs involving a given feature. Several different types of features prove effective on this problem, with accuracy averages of 0.97 or above (given the optimal frequency and substrate). These include the level of resistance at two hours after inoculation; the initial growth rate; the value of the first peak; the post-wounding recovery rate; and the ratio of post-wounding recover rate to initial growth rate. Notably, fitted ARIMA coefficients perform quite poorly: though the cell cultures' ARIMA models are substantially different from white noise and from models fitted to empty wells, there are no reliable differences between infected and uninfected model coefficients. The estimated error variance from these models, however, *is* effective, reflecting differences in post-confluence micromotion.

4 VARIATIONS

4.1 Cell and contaminant types

When we extend our analysis to include a comparator cell type (BSC-1) and contaminant (*M. hyorhinis*), it is immediately evident that both factors change the contamination problem. For example, with MDCK II cells and *M. hyorhinis*, the strongest-performing individual features are those focusing on the early peak behavior and recovery from wounding. In contrast, all of the features with high average performance for BSC-1 cells are either coefficients from a quadratic fit to the first two hours of growth, or measurements of local variation/micromotion after confluence. Indeed, for BSC-1 cells the problem is overall more difficult: far fewer feature pairs are able to give good classification performance and, due to the different growth pattern, the early levels and first-peak features that worked well for MDCK II cells are not distinctive. These results underline the importance of establishing the most effective and

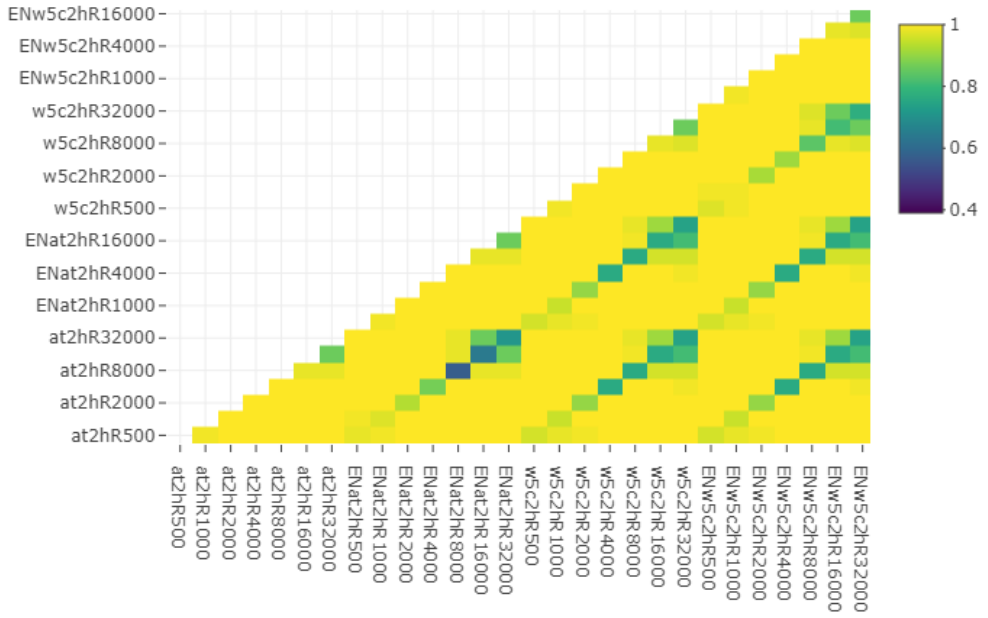

Figure 3: Heatmap matrix of classification accuracy (on gel) for a few feature pairs. Many pairs have high accuracy, but there are "sub-diagonals" of feature pairs with poorer performance.

characteristic features for individual cell lines, and of exploring a variety of possible feature types.

4.2 Experimental variation

In laboratory use, the cell cultures that require classification may often be grown on a separate plate than the training set of cultures known to be infected or uninfected. The test set may even be measured at a substantially different time, or under different conditions. It is therefore desirable to examine the robustness of classification to variation across plates, and experimental variation in general.

To this end, we examine four separate plates of MDCK II cells/*M. hominis*, each plate using cultures taken from the same frozen source, but grown and measured at a separate time. We can immediately see that plate-to-plate variation exists, and makes the classification problem substantially more difficult. Figure 4 shows the same set of features and classification procedure as Figure 2, but this time, all observations on one plate are used as the training set, while the test observations come from a different plate. The accuracy is far worse than the within-plate case.

Most features show a large drop in accuracy for cross-plate classification. A noticeable standout is the "7 hours minus 2 hours" family of features, which reflect the rate of initial cell growth (for example, average accuracy of 0.93 on gel-treated cultures at 4000 Hz). Though simple, this feature type performs well when paired with several other types of features, or even with itself at two distinct frequencies. Such a feature draws on the information provided by the early growth stage while accounting for the level reached during the first two hours, making the values less sensitive to the initial conditions of each cell culture.

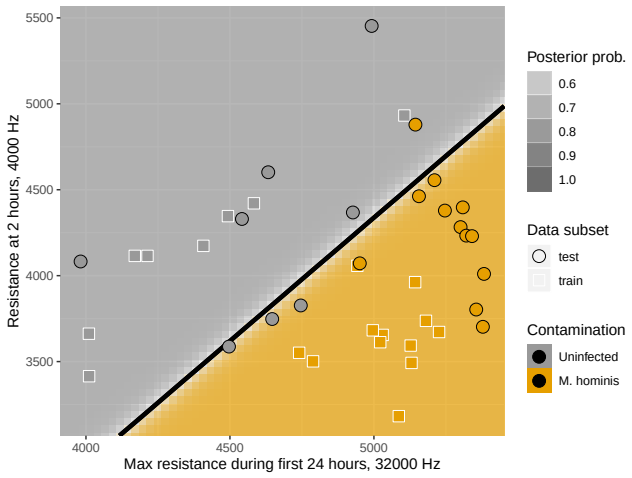

Figure 4: LDA classification regions, with training set from one plate and test set from a second plate. Many test observations are incorrectly classified.

In general, features focusing on later-stage behavior also appear to be more robust to plate-to-plate variation. The process of each culture reaching confluence, or being wounded with the same high-voltage current, appears to equalize the effects of initial conditions. Indeed, features measuring local variability after confluence are among the top overall performers (such as $\hat{\sigma}^2$ from an ARIMA(0,1,1) model, with an accuracy of 0.92 on BSA at 500 Hz). While the current dataset is small, these are promising indicators of useful features.

5 CONCLUSION

This paper demonstrates a methodology for differentiating between standard cell cultures and those infected with a mycoplasma organism, by comparing features of their ECIS® time courses. Detecting contamination in this way is particularly desirable since the ECIS data gathered during the process can also be used to investigate other scientific questions about cell morphology, without requiring a separate tool, or labeling or destroying cells.

We have determined that high classification accuracy can be achieved using a straightforward classification algorithm such as LDA, applied to a small number of interpretable features. The best individual features depend on the cell type and contaminant; typically, combining features at different frequencies is most effective.

We can also see that there is substantial variation across plates. Since this experimental variation is to be expected in practice, where cultures will be grown at different times and in multiple laboratories, it is an important topic for further investigation. We have already shown that certain features appear to be more robust to this "plate effect" than others, especially those that reduce the effect of initial conditions; it may also be possible to normalize the measurements in some way, though a naive method is not effective. A larger dataset will be needed to study this plate effect more fully.

ACKNOWLEDGMENTS

The authors thank Applied BioPhysics, Inc. for providing the data used in this work, and gratefully acknowledge financial support from the Cornell University Institute of Biotechnology, the New York State Foundation of Science, Technology and Innovation (NYSTAR), a Xerox PARC Faculty Research Award, National Science Foundation Awards 1455172, 1934985, 1940124, and 1940276, USAID, and Cornell University Atkinson Center for a Sustainable Future.

REFERENCES

[1] Charu C. Aggarwal, Alexander Hinneburg, and Daniel A. Keim. 2001. On the Surprising Behavior of Distance Metrics in High Dimensional Space. *International Conference on Database Theory* (2001).

[2] Pierre O. Bagnaninchi and Nicola Drummond. 2011. Real-time label-free monitoring of adipose-derived stem cell differentiation with electric cell-substrate impedance sensing. *Proceedings of the National Academy of Sciences* 108, 16 (2011), 6462–6467.

[3] Jorge Caiado, Nuno Crato, and Daniel Peña. 2006. A periodogram-based metric for time series classification. *Computational Statistics and Data Analysis* 50 (2006), 2668–2684.

[4] M. Dash and H. Liu. 1997. Feature Selection for Classification. *Intelligent Data Analysis 1* (1997), 131–156.

[5] Pierpaolo D'Urso, Livia De Giovanni, and Riccardo Massari. 2015. Time series clustering by a robust autoregressive metric with application to air pollution. *Chemometrics and Intelligent Laboratory Systems* 141 (2015), 107–124.

[6] Pierpaolo D'Urso and Elizabeth Ann Maharaj. 2009. Autocorrelation-based fuzzy clustering of time series. *Fuzzy Sets and Systems* 160 (2009), 3565–3589.

[7] Pedro Galeano and Daniel Peña. 2000. Multivariate analysis in vector time series. *Resenhas* 4 (2000), 383–404.

[8] Megan L. Gelsinger, Laura L. Tupper, and David S. Matteson. 2019. Cell Line Classification using Electric Cell-Substrate Impedance Sensing (ECIS). *The International Journal of Biostatistics* 16, 1 (2019).

[9] Ivar Giaever and Charles R. Keese. 1984. Monitoring Fibroblast Behavior in Tissue Culture with an Applied Electric Field. *Proceedings of the National Academy of Sciences, USA* 81 (1984), 3761–3784.

[10] Ivar Giaever and Charles R. Keese. 1991. Micromotion of mammalian cells measured electrically. *Proceedings of the National Academy of Sciences USA* 88 (1991), 7896–7900.

[11] I.H. Heijink, S.M. Brandenburg, D.S. Postma, and A.J.M. van Oosterhout. 2012. Cigarette smoke impairs airway epithelial barrier function and cell–cell contact recovery. *European Respiratory Journal* 39, 2 (2012), 419–428.

[12] Jongin Hong, Karthikeyan Kandasamy, Mohana Marimuthu, Cheol Soo Choi, and Sanghyo Kim. 2011. Electrical cell-substrate impedance sensing as a non-invasive tool for cancer cell study. *Analyst* 136 (2011), 237–245.

[13] Peyton Hughes, Damian Marshall, Yvonne Reid, Helen Parkes, and Cohava Gelber. 2007. The costs of using unauthenticated, over-passaged cell lines: hos much more data do we need? *Biotechniques* 43, 5 (2007), 575–582.

[14] Charles R. Keese, Joachim Wegener, Sarah R. Walker, and Ivar Giaever. 2004. Electrical wound-healing assay for cells *in vitro*. *Proceedings of the National Academy of Sciences, USA* 101 (2004), 1554–1559.

[15] Eamonn Keogh, Stefano Lonardi, Chotirat Ann Ratanamahatana, Li Wei, Sang-Hee Lee, and John Handley. 2007. Compression-based data mining of sequential data. *Data Mining and Knowledge Discovery* 14, 1 (2007), 99–129.

[16] M. Kowolenko, C.R. Keese, D.A. Lawrence, and I. Giaever. 1990. Measurement of macrophage adherence and spreading with weak electric fields. *Journal of Immunological Methods* 127 (1990), 71–77.

[17] Douglas C. Lovelady, Jennifer Friedman, Sonali Patel, David A. Rabson, and Chun-Min Lo. 2009. Detecting effects of low levels of cytochalasin B in 3T3 fibroblast cultures by analysis of electrical noise obtained from cellular micromotion. *Biosensors and Bioelectronics* 24 (2009), 2250–2254.

[18] Sonja Lukic and Joachim Wegener. 2015. Impedimetric Monitoring of Cell-Based Assays. *eLS* (2015).

[19] Elizabeth Ann Maharaj. 2000. Clusters of Time Series. *Journal of Classification* 17 (2000), 297–314.

[20] Elizabeth Ann Maharaj, Pierpaolo D'Urso, and Jorge Caiado. 2019. *Time Series Clustering and Classification.* CRC Press.

[21] Giljun Park, Chang K. Choi, Anthony E. English, and Tim E. Sparer. 2009. Electrical impedance measurements predict cellular transformation. *Cell Biology International* 33 (2009), 429–433.

[22] Michael J. Rutten, Bryan Laraway, Cynthia R. Gregory, Hua Xie, Christian Renken, Charles Keese, and Kenton W. Gregory. 2015. Rapid assay of stem cell functionality and potency using electric cell-substrate impedance sensing. *Stem Cell Research and Therapy* 6, 192 (2015).

[23] David Schneider, Marco Tarantola, and Andreas Janshoff. 2011. Dynamics of TGF-β induced epithelial-to-mesenchymal transition monitored by Electric Cell-Substrate Impedance Sensing. *Biochimica et Biophysica Acta* 1813 (2011), 2099–2107.

[24] Judith A. Stolwijk, Christoph Hartmann, Poonam Balani, Silke Albermann, Charles R. Keese, Ivar Giaever, and Joachim Wegener. 2011. Impedance analysis of adherent cells after in situ electroporation: Non-invasive monitoring during intracellular manipulations. *Biosensors and Bioelectronics* 26 (2011), 4720–4727.

[25] Howell Tong and P Dabas. 1990. Clusters of Time Series Models: An Example. *Journal of Applied Statistics* 17 (1990), 187–198.

[26] Lei Wang, Lei Wang, Hongying Yin, Wanli Xing, Zhongyao Yue, Min Guoe, and Jing Cheng. 2010. Real-time, label-free monitoring of the cell cycle with a cellular impedance sensing chip. *Biosensors and Bioelectronics* 25 (2010), 990–995.

[27] Xiaozhe Wang, Kate Smith, and Rob Hyndman. 2006. Characteristic-Based Clustering for Time Series Data. *Data Mining and Knowledge Discovery* 13 (2006), 335–364.

[28] Wenyu Zhang, Maryclare Griffin, and David S. Matteson. 2021. Modeling Nonlinear Growth Followed by Long-Memory Equilibrium with Unknown Change Point. *In submission* (2021).

A SUPPLEMENTARY MATERIALS

A.1 Methodology of ECIS data collection

A.1.1 Preparation of the wells. Proteins were adsorbed to the surface of wells in 96W1E+ (Applied BioPhysics, Inc.) arrays using sterile solutions of bovine serum albumin (BSA) (Pentex® brand Miles Laboratories, Inc.) and of gelatin (Fisher purified grade gelatin; 275 bloom) in 0.15 M NaCl at a concentration of 200 micrograms/ml. Half of the wells in each plate received 200 microliters of either BSA or gelatin solutions at room temperature. Following 15 minutes to allow protein adsorption, the solutions were aspirated from the wells, and all wells received 200 microliters of a sterile solution of 10 mM cysteine in distilled water (Electrode Stabilizing solution, Applied BioPhysics) to stabilize the gold electrodes. The cysteine solution remained in the wells for at least 30 minutes and was aspirated just prior to cell inoculation.

A.1.2 Cell culture. BSC-1 cells and MDCK II cells were obtained from the American Type Culture Collection (Manassas, VA) and grown in DMEM (low glucose, SIGMA D6046) with 10% fetal bovine serum. Antibiotics were not used to eliminate any possible effects upon the mycoplasma infections being studied. Cells were grown to near confluency, and cell suspensions were prepared. BSCI-1 cell layers were disrupted using trypsin/EDTA, and MDCK II were exposed first to EDTA for 10 min before the introduction of trypsin/EDTA. Detached cells were harvested using complete medium and centrifuged (250 x g, 5 mins). The cell pellets were resuspended in complete medium at 37° C, counted, and plated into the ECIS wells. The arrays were mounted in a 96 well ECIS Station that was maintained in a tissue culture incubator at 37° C with 5% carbon dioxide and 95% relative humidity.

A.1.3 Infection of cell layers with mycoplasma. Freeze-dried samples of *M. hominis* and *M. hyorhinis* were obtained from the American Type Culture Collection (Manassas, VA) in sealed vials and opened under aseptic conditions. The dried samples were divided into two approximately equal parts. One was placed in a sealed glass vial and stored under liquid nitrogen vapor for future inoculation. The other was split and used to inoculate cultures of either BSC-1 cells or MDCK II cells. The lyophilized material was dispersed into the medium of 50% confluent culture to initiate the infection. All infected cultures were passaged several times using fastidious care to prevent contamination of any uninfected cultures.

Tests were routinely run to verify infection of the inoculated cultures and to assure that uninfected cultures were mycoplasma free. We used one of two different commercially available kits to detect mycoplasma: the MycoAlert® Detection Kit by Lonza and the PlasmoTest™ by InvivoGen.

A.1.4 Measurements. All impedance measurements were obtained using the ECIS® Zθ instrumentation using ninety-six well 96W1E+ arrays, provided by Applied BioPhysics, Inc., Troy, NY. ([9], [10], [24])

Following the addition of cell suspensions to the wells, time course impedance data were gathered at low non-invasive current (2.5 microamperes) at seven different AC frequencies, doubling each frequency from 500 Hz to 32 kHz. The impedance (both the resistive and reactive components) recorded the attachment and spreading of the cells and the formation of a confluent cell monolayer. At 48 hours post-inoculation, half of the wells received a brief invasive high current pulse (3000 microamperes at 60 kHz for 20 sec), killing the cells attached to the electrodes. The system then reverted to the low current mode, and cellular migration was then monitored as the normal cells on the periphery of the electrode migrated inward to reestablish a monolayer upon the electrodes ([14]).

A.2 List of features

The following features were extracted from the time series of resistance values for each cell culture. Each feature was calculated using the time series at each available AC frequency. In addition, we calculated most features using both the original measured values, and values normalized to the empty wells, though typically there were only small differences in performance (see 3).

(1) Value at 2 hours after inoculation
   (a) *Windowed* value at 2 hours: average value over 5 consecutive time points, centered at 2 hours after inoculation
(2) Value at 24 hours after inoculation
   (a) Windowed value at 24 hours
(3) Second-order coefficient of a quadratic fit to the first 2 hours after inoculation
(4) First-order coefficient of a quadratic fit to the first 2 hours after inoculation
(5) Maximum value up to 24 hours after inoculation
   (a) Maximum value up to 24 hours after inoculation of a *smoothed* time series (rolling average over 5 time points)
(6) Maximum value up to 48 hours after inoculation
   (a) Maximum value up to 48 hours after inoculation of a smoothed time series
(7) Features of the first peak in a smoothed time series after inoculation (these features use a rolling window of 15 time points, for reduced sensitivity to noise in the early series):
   (a) Value at this peak
   (b) Time when this peak occurs
(8) Value at 7 hours minus value at 2 hours
   (a) Windowed value at 7 hours minus windowed value at 2 hours
(9) Value at 57 hours minus value at 52 hours
   (a) Windowed value at 57 hours minus windowed value at 52 hours
(10) Value at 57 hours over value at 52 hours
   (a) Windowed value at 57 hours over windowed value at 52 hours
(11) Ratio of "value at 57 hours minus value at 52 hours" to "value at 7 hours minus value at 2 hours"
   (a) Ratio of "windowed value at 57 hours minus windowed value at 52 hours" to "windowed value at 7 hours minus windowed value at 2 hours"
(12) Coefficients from ARIMA models fit to post-confluence/pre-wounding data, from 24 to 36 hours:
   (a) $ARIMA(1,1,0)$ ($AR(1)$ model with first differencing), $\hat{\phi}_1$
   (b) $ARIMA(1,1,0)$, constant term
   (c) $ARIMA(0,1,1)$ ($MA(1)$ model with first differencing), $\hat{\theta}_1$
   (d) $ARIMA(0,1,1)$, constant term
(13) Estimated error variance from these fitted ARIMA models:
   (a) $ARIMA(1,1,0)$, $\hat{\sigma}^2$
   (b) $ARIMA(0,1,1)$, $\hat{\sigma}^2$
(14) Features measuring autocorrelation time series from 24 to 36 hours, after first differencing:
   (a) ACF at lag 1
   (b) ACF at lag 2
   (c) PACF at lag 1
   (d) PACF at lag 2
(15) Sample variance of once-differenced time series from 24 to 36 hours
(16) Sample standard deviation of once-differenced time series from 24 to 36 hours
(17) Features of the first peak in a smoothed time series after wounding:
   (a) Value at this peak
   (b) Time when this peak occurs

(18) Minimum value after wounding
   (a) Minimum value after wounding of a smoothed time series
(19) Features from the Growth-to-Confluence Detector of Zhang et al., applied to post-wounding data:
   (a) τ, the detected time for the beginning of the confluent phase
   (b) d, the estimated long-memory parameter for the series following confluence

A.3 Additional results

A.3.1 Results for classification algorithms. The three classification algorithms used in this study are classification trees, linear discriminant analysis (LDA), and quadratic discriminant analysis (QDA). Following the findings of [8], we do not investigate further possibilities for the parameters of the discriminant analysis beyond LDA and QDA, since these were not found to have a great effect on the classification of cells.

Overall, LDA was the most effective classification algorithm on our main dataset. The average performance of the LDA classifiers over all feature pairs was 0.882, as compared to 0.855 for QDA classifiers and 0.857 for trees. In a small number of cases, the discriminant analysis algorithms are unable to define classification regions, thanks to a small sample size and small numerical differences between individual wells' feature scores. The tree method can be used in these cases.

The difference in performance between the methods is, indeed, not large. We find LDA to be desirable, however, not only for its slight advantage in performance but for the nature of its classification regions, which are easy to describe and visualize.

A.3.2 Results for different cell/contaminant types. We summarize additional results from the within-plate classification task for the three different combinations of cell type and contaminant.

For MDCK-II cells with *M. hominis*, 27% of all feature pairs yielded perfect classification accuracy on BSA-treated cultures, and 25% of all pairs did so on gel-treated cultures. Some examples of pairs with perfect accuracy include:

- 57 hours - 52 hours, 32000 Hz and Value of initial peak, 1000 Hz; BSA
- Sample variance of differenced series, 8000 Hz and Value of post-wound peak, 2000 Hz; BSA
- Value at 2 hours, 500 Hz and PACF of differenced series, lag 2, 32000 Hz; gel

There are many individual features with average accuracy rates extremely close to 1 (above 0.99). These include:

- (57 hours - 52 hours)/(7 hours - 2 hours),500 Hz, gel
- 57 hours - 52 hours, 500 Hz, BSA
- Quadratic coefficient from fit to first 2 hours, 100 Hz, gel
- Value of post-wound peak, 1000 Hz, gel

For MDCK-II cells with *M. hyorhinis*, 13% of all feature pairs yielded perfect classification accuracy on BSA-treated cultures, and 38% of all pairs did so on gel-treated cultures. Some examples of pairs with perfect accuracy include:

- Quadratic coefficient from fit to first 2 hours, 16000 Hz and 57 hours/52 hours, 32000 Hz, gel
- Max over first 24 hours, 4000 Hz and Quadratic coefficient from fit to first 2 hours, 1000 Hz, gel
- Minimum value after wounding, 500 Hz and Post-wounding long-memory parameter d, 1000 Hz, BSA

Again, many individual features have average accuracy rates above 0.99. These include:

- Max over first 24 hours, 8000 Hz, BSA
- 57 hours - 52 hours, 32000 Hz, gel
- Quadratic coefficient from fit to first 2 hours, 400 Hz, gel
- Minimum value after wounding, 500 Hz, gel

For BSC-1 cells with *M. hominis*, 0.08% (888/113469) of all feature pairs yielded perfect classification accuracy on BSA-treated cultures, and only 4 pairs did so on gel-treated cultures. Some examples of pairs with perfect accuracy include:

- Max over first 24 hours, 32000 Hz and Max over first 24 hours, 8000 Hz, gel
- Sample variance of differenced series, 500 Hz and Value at 24 hours, 500 Hz, BSA
- Quadratic coefficient from fit to first 2 hours, 1000 Hz and 7 hours - 2 hours, 16000 Hz, BSA

Overall, the individual features with the highest average accuracy rate are:

- Quadratic coefficient from fit to first 2 hours, 500 Hz, BSA: 0.964
- Sample standard deviation of differenced series, 32000 Hz, BSA: 0.957
- Sample standard deviation of differenced series, 1600 Hz, BSA: 0.950
- Linear coefficient from quadratic fit to first 2 hours, 1000 Hz, BSA: 0.943

A.3.3 Results for cross-plate classification. For the cross-plate classification task, few feature pairs had perfect average classification accuracy across train-test combinations: 1 for BSA-treated cells and 97 for gel-treated cells. Some examples with accuracy above 0.99 include:

- $\hat{\sigma}^2$ from ARIMA(1,1,0), 500 Hz and Value of post-wound peak, 500 Hz, BSA
- 7 hours - 2 hours, 4000 Hz and Constant from ARIMA(0,1,1), 16000 Hz, gel
- Sample standard deviation of differenced series, 500 Hz and Value at 2 hours, 1000 Hz, gel

Overall, the individual features with the highest average accuracy rate are:

- 7 hours - 2 hours, 4000 Hz, gel: 0.933
- Sample standard deviation of differenced series, 500 Hz, BSA: 0.926
- 7 hours - 2 hours, 2000 Hz, gel: 0.925
- $\hat{\sigma}^2$ from ARIMA(0,1,1) model, 5000 Hz, BSA: 0.924