IMPEDANCE SPECTROSCOPY FOR THE CHARACTERIZATION OF SWINE REPRODUCTIVE STATES

BY

DANIEL J. GLASS

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Agricucltural and Biological Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2019

Urbana, Illinois

Master's Committee:

Associate Professor Kaustubh Bhalerao, Chair
Professor Robert V Knox
Professor Tony Grift

Abstract

This study uses machine learning to examine data acquired from a frequency sweeping electrical impedance spectrometer taking readings from sows in order to potentially determine if the resultant data can be used to predict a sow's date of estrus, date of farrowing, and if the inseminated sow has become pregnant. The frequencies analyzed were between 1000 Hz and 29000 Hz. The models were logistic regressions which were reduced by a stepwise Akaike Information Criterion. The data set for the fertility model included 250 frequency sweeps across 85 animals, for farrowing there were 921 observations across 421 animals, and for the pregnancy data set there were 395 observations across 227 animals. The fertility and farrowing detection models displayed substantial strength, with the latter holding particular promise in the face of conventional methods. The fertility model had a high Area Under Curve (AUC) of 0.946, while the two tested farrowing models had an AUC of 0.839 and 0.834. Pregnancy detection models were skewed by the retrieved dataset containing few non-pregnant animals, with only 50 observations across 29 non-pregnant animals as compared to 345 observations across 198 pregnant animals, and had an AUC of 0.839. Separating between late and early readings in the model is considered to be a potential improvement to the pregnancy model, with an AUC of 0.939, although additional observations, particularly of non-pregnant animals, are necessary. The groundwork was laid for further work to be done, particularly regarding whether an animal that has failed to become pregnant is going to return to estrus. Future models that incorporate more information about individual animals such as the number of litters previously birthed, as well as expanding this concept into other animals, are recommended.

To Hallie Krex and Karen, Donald, and Andrew Glass.

Acknowledgments

I would like to thank my advisor, Kaustubh Bhalerao, who gave me this opportunity and helped me make this a reality, as well as my other committe members, Robert Knox and Tony Grift, for their work and consideration. Finally, I owe a great debt of gratitude to my family and Hallie Krex, for unwaveringly supporting me in this endeavor.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AC	Alternating Current.
AIC	Akaike Information Criterion.
AUC	Area Under Curve
DC	Direct Current.
FPR	False Positive Rate.
FSH	Follicle-Stimulating Hormone.
EIS	Electrical Impedance Spectroscopy.
EP	Electrode Polarization.
LAB	Lactic Acid Bacteria.
LH	Luteinizing Hormone.
ML	Machine Learning.
ROC	Receiver Operating Characteristic.
TPR	True Positive Rate.

Chapter 1

Introduction

The idea of measuring the reproductive state of mammals has held significant promise in the possibilities it could elicit. By being able to predict fertility, for example, costs could be saved on artificial insemination, a method which has come to dominate the livestock industry. One tool that has seen particular interest in this field is Electrical Impedance Spectroscopy (EIS), which indirectly measures the reproductive state of the animal by detecting changes in impedance which correspond to changes in reproductive state [4]. EIS is a well-established method of data acquisition and material characterization [5, 6], having been used in the fields of biology [7, 8, 9], materials [10, 11], and electrochemistry [12, 13].

Distinct patterns in vaginal impedance over the estrous cycle have been seen throughout economically important mammalian livestock, e.g. cattle [14, 15, 16] and sheep [8, 17], though the nature of the pattern varies between species. As such, EIS holds great promise as a method by which to detect reproductive states for the meat industry [18].

Another animal that displays such patterns is swine [9, 7]. The pork industry is both massive, with over 100 million tons in Carcass Weight Equivalent produced and consumed per year from over a billion pigs produced. It also continues to grow due to shifting consumer preferences increasing demand, and Chinese importation demand is additionally expected to rise at a fast pace due to massive losses of domestic stock to African Swine Fever [19].

Early models met with results not better than already-known techniques and often were bulky and expensive [20]. Since, advances in computational technology have enabled the rise of Machine Learning (ML) to quickly and effectively analyze data with smaller, more user-friendly equipment. This, combined with falling costs for the parts that would make up a spectroscope, have greatly increased the viability of such a device on a commercial scale. The device in question collects data that has applications not only in potentially detecting the state of the estrous cycle, but also whether or not the animal in question is pregnant, and if she is nearing her farrowing date.

1.1 Preceding Work

In 1984, Zink, M.F. and Diehl, J.R. published an article to the Journal of Animal Science [20], testing an electrical conductivity meter for predicting the best time to inseminate a pig. To do this, they performed two experiments. In the first, they took ten healthy gilts and checked them using a boar for the onset of behavioral estrus, i.e. when they accepted mounting from the boar, and measured the conductivity of their vaginal mucus using their Ovascan device. In the second experiment, they used 46 gilts and 17 sows. Half of the gilts were brought into synchronization of their estrous cycle by chemical means. The females were also either mated manually or with artificial insemination, or bred twice based on the detection of estrus via the aforementioned boar method, or were bred once based on the readings from their Ovascan device. They found that the variables of insemination method or estrous synchronization had no effect, but also found that their device yielded consistent, repeatable results, and was not significantly better or worse than the boar detection method. They considered that in the future, refinements would be needed, as the scanner was no better at providing a time for insemination than the current common practices and thumb rules.

The promise of being able to detect multiple reproductive states from such a device has continued to endure as a possibility for reducing costs in the pork industry. The device that collected the data analyzed is from the work of Seokchan Yoo in his 2017 thesis: A Low-Cost, Portable, Web-Based Impedance Spectroscope for Agricultural Applications. The goal was the development of a novel system design for a low-cost, portable, wireless electrical impedance spectroscope for measuring swine vaginal impedance, and interpreting the results. This feat would bring a cumbersome $30,000 piece of equipment for impedance analysis to a portable, $200 design[1]. A four-electrode design was used, as this design has lower variance than a two-electrode design [18], and solves the error caused by electrode polarization [21].

The probe took readings deeper within the vagina, past the urethra and near the cervix, and was shown to be able to measure impedance and phase to a high degree of accuracy. With the information provided by the probe, there were clear patterns emerging in the impedance of spectrum over the days prior to, upon, and after the onset of estrus over tested frequency values, with a sharp decline following a usual maximum at four days until estrus. Clear patterns emerged with days before farrowing, with impedance increasing until two days before farrowing. However, there was a failure in the device, but the problem was not reproducible, and measures to add strain relief to the wire and make the electrode wire more robust to ensure the contact of the circuitry was maintained were taken. These did not solve the problem, so subsequently the battery type and the instrumental amplifier chip were changed, and this solved the problem [1].

1.2 Theory of Device Operation

1.2.1 Reproductive Environment

The swine estrous cycle is normally 21 days in length, though it can vary from 19 to 23 days [3]. The profile of a normal estrous cycle is displayed in Figure 1.1.

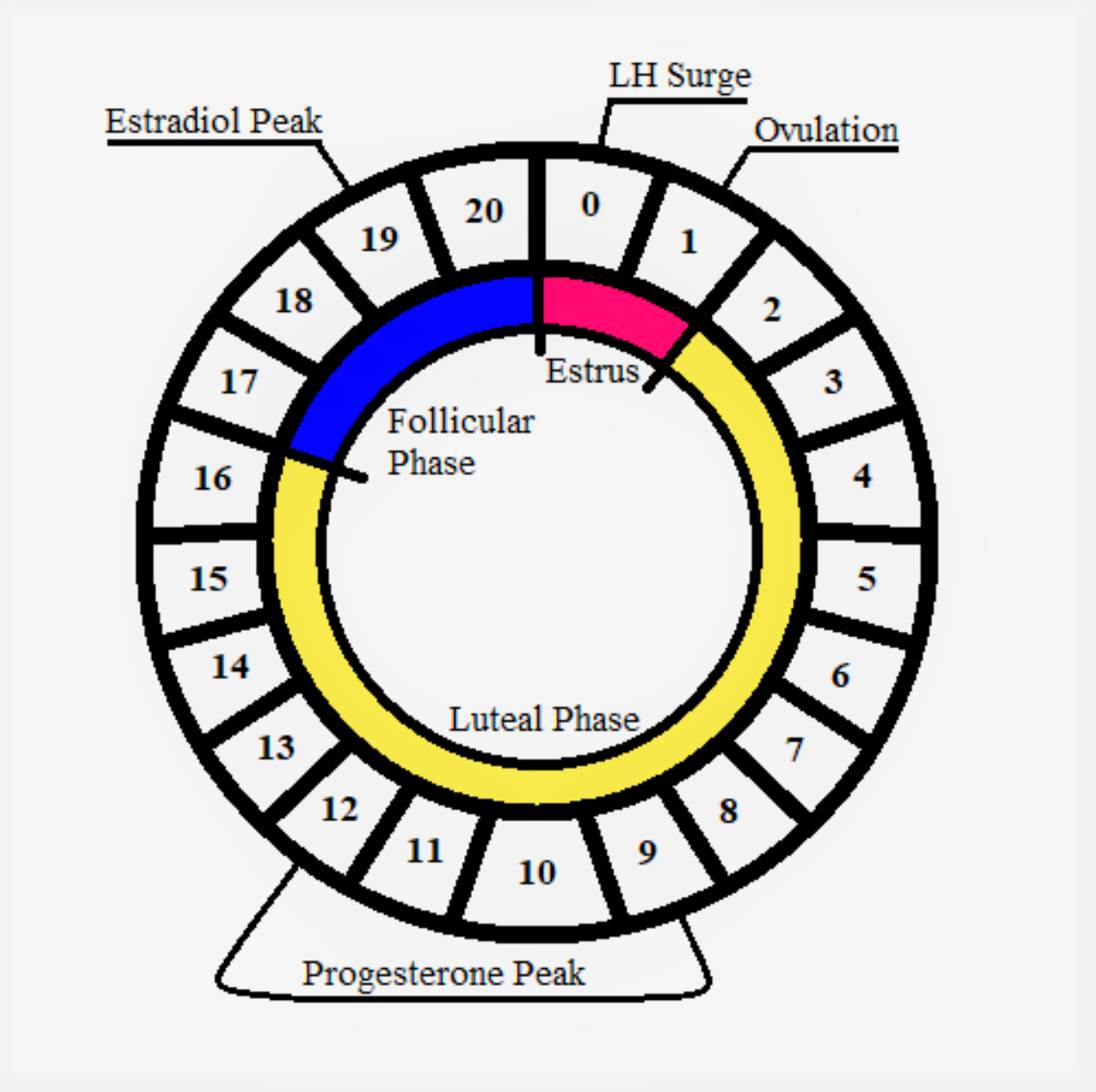


Figure 1.1: The estrous cycle of swine, with hormonal labelling. The cycle is variously described as day 0-20 and day 1-21. [1] [2] [3]

It has been observed that, among a host of eutherian mammals, electrical impedance reaches a minimum

during the follicular phase, and is highest during the luteal phase of the estrous cycle, and this pattern is present in swine [9, 7]. Four signaling compounds are of primary interest to this situation: the steroids estradiol and progesterone, and the protein hormones Follicle-Stimulating Hormone (FSH) and Luteinizing Hormone (LH) [18, 7, 3]. The levels of these are described in Figure 1.2. FSH stimulates the growth of ovarian follicles and estradiol production, while LH triggers the release of the eggs [3]. Levels of progesterone, estradiol, and LH have, in particular, been investigated in terms of their correlation to vaginal impedance. The precipitous drop appears to correlate tightly with the sharp rise in the concentration of estradiol in the work of Řezáč [18, 9], as shown in Figure 1.3, and LH spikes at day 0, and it is around this time that the impedance quickly rises again [18, 7].

Of course, the question arises as to why it is that the vaginal impedance decreases during this point in the cycle. The vaginal mucosal environment consists primarily of water, proteins, other biomolecules, and ions, and adjacent to it is the wall of tissue. The composition of the mucus, or the makeup of the tissue, must change in order for the impedance to fall, and then rise. Likely, this is due to changes in protein and ion presence.

Unfortunately, investigating tissue changes is invasive. Additionally, there is a relative lack of prior work done in analyzing the changes in the composition of vaginal mucus during the estrous cycle in sows, though there are some studies for other animals. One example is in detecting the changes of ionic concentrations in oviduct fluid, uterine fluid, and blood serum in bovines. However, this study only looks at the estrus and luteal phases for the oviduct and uterine fluid respectively, whereas the primary period of interest is the follicular phase [22]. Another bovine-based study broke down changes in cervical mucus biochemical composition, evaluating it for predictive capability. It indicated that, in bovines, cholesterols and proteins decline as estrus nears, but sorbitol spikes in the days leading up to estrus [23].

So, without much background to draw from, the conjectural hypothesis decided upon for this is that the two stages - the rise of estradiol, and the surge of LH - each have a respective effect on vaginal health and fertility. Normally, the vagina is rather acidic, which prevents it from being infected [24]. With the rise of estradiol, the ionic profile and changes in biomolecule presence may have an effect on the local microbiota, which could then change the acidity. Changes in the profile of vaginal fluids in terms of ionic composition during the cycle have been noted in humans [25], but if there is a significant difference in the readings based upon the frequency, then larger biomolecules could be more to blame. During this first phase, changes in the mucosal composition may allow for the proliferation of microflora. Among the common flora in the sow are species of *Streptococcus* and *Enterococcus*, Lactic Acid Bacteria (LAB), which have been observed to make up around 23.2 percent of the microflora [26]. The proliferation and metabolism of LAB would produce

excess lactic acid, acidifying the vagina.

The estradiol peak phase has been shown to correlate with lowest vaginal pH in the cycle of many mammals [27]. As such, the estradiol peak is hypothesized to change the ionic and biomolecular composition of the mucus to suit the native LAB in acidification. This acidification prior to estrus may serve as a preventative disinfection prior to estrus, as during this subsequent LH surge stage, the impedance rises dramatically, at which time the vagina will be more vulnerable to the proliferation of unwanted foreign microbial life. The deacidification during the LH surge is necessary, as sperm cells prefer an environment that is neutral to alkaline. The rise in pH of the reproductive tract immediately prior to, and during, estrus would make it a more habitable environment for sperm cells, increasing their motility [28], and likely, their survival time. As such, this process increases the likelihood that the animal will successfully reproduce following a mating. It is likely, then, that the LH surge induces changes in the composition once more, diminishing the population of its LAB, deacidifying it and preparing for fertilization.

The swine gestation period is 114-115 days [1]. Estrogens and progesterone increase in early pregnancy, while animals which fail to become pregnant will experience a fall in progesterone, and estradiol will not rise generally, but willspike shortly before reentering estrus. The days leading up to farrowing are characterized by a steep rise in relaxin and prolactin, and a similar fall in progesterone [3].

1.2.2 Electrical Physics of Impedance Measurement

Impedance (Z) itself is the complex extension of resistance, the latter's function being described in Ohm's Law in Equation 1.1, which states that between two points in a circuit, the current (I) running through them is directly proportional to the voltage (V), and inversely proportional to the resistance (R), including the ability of capacitors and inductors to affect the availability of electrical energy.

$$I = \frac{V}{R} \tag{1.1}$$

The impedance measurement's primary area of action is between a cathode and an anode. The medium that is thus run through is a local tissue and its mucus, a substance consisting of water, ions, and biomolecules. These ions and biomolecules have charges, which attracts them to the cathode or anode based on their charge, at which point an electron transfer occurs. When this happens, their new charge leads them to be attracted to the pole of the opposite charge. Loss may, however, occur if particles diffuse away from the area. A model architecture useful for examining the electrical physics of EIS is the Randles Circuit, incorporating mixed charge transfer and diffusion [29], and is shown in Figure 1.4. The solution's electrolytic resistance, R_s, is characterized by the solution's concentration of charge carriers and the geometry of the environment

between the electrodes. The electrode-electrolytic interface's double layer capacitance, C_{dl}, is a result of charge carriers 'sticking' to the surface of the electrode. The charge transfer resistance, R_{ct}, is a result of there being a particular speed to the transfer of electrons. Finally, the Warburg impedance, Z_W, is a result of diffusion within the system, and is inversely related to frequency [30].

One element not discussed above is the effect of Electron Polarization (EP). EP occurs when the current must flow through electrochemical reactions occurring at the surface of the electrode, whether the result is surface corrosion or not [30]. EP is a process which will confound the data acquired from two-electrode probe structures, particularly in mediums with low impedance, and at frequencies below 100 kHz. These lower frequencies are often the ones of interest for biosensing applications, and as such, EP is a major problem that will confound readings in biological mediums [31, 32]. To combat this problem, Schwan proposed and created a four-electrode impedance measurement device [21].This model was followed up on and used by Chang, Pop, and Meijer, who confirmed the mitigation of the effect of electrode polarization on a biological sample [33]. This four-electrode model, which uses an outer pair of current electrodes and an inner pair of voltage electrodes, makes up the basis of the probe used for the collection of the impedance data [1]. It is shown in Figure 1.5, with current running from the top electrode to the bottom. The impedances before the middle electrodes of the feedback probe are the result of the electrochemical reactions of EP on the outer elecrodes, while the impedance between the two is the result of the medium. These middle two electrodes have a voltage drop due to the medium impedance, but not the EP-based impedance, as current does not run into them, and thus the electrochemical reactions due to current do not occur. The voltage difference between the two electrodes can then be used to find the medium impedance.

In a situation of direct current, the ions and biomolecules thus move from one end to another, completing the circuit with an impedance attached. DC is not, however, used in this probe. The primary reason that AC is used is because it also affects the impedance based on its frequency. At low frequencies, AC functions similarly to DC, in that the different charge carriers, from massive proteins to light ions, will have time to be able to move between the poles to receive and deposit electrons. However, at higher frequencies, a new phenomenon arises, in that the proteins may end up being trapped between the cathode and anode without reaching either as their polarities shift. The proteins in this mucosal environment are massive, and while they can carry a far greater electrochemical potential than the ions, they have far lower potential per mass. In addition, their volume is far larger than that of an ion, reulting in slower diffusion. As such, they will migrate through the medium towards the charged pole far more slowly than the ions do. In a DC situation, the proteins are able to go between the poles in their slow manner, migrating large quantities of charge when doing so. This is the same in the low frequency AC situation, as the reversing of the polarity of the

electrodes. In the high frequency AC situation, however, the electrodes' charges change too quickly for the proteins to be able to migrate all the way, which increases the impedance due to an increase in capacitance. However, the Warburg impedance will decrease as the frequency increases. Because of this, higher frequency readings exhibit both upward and downward pressures on impedance. This opens up a new range of data to acquire, as the differing frequencies will each have their own respective curves. Certain frequencies may hold more predictive data than others, and this could even vary from animal to animal. It is for this reason that a frequency sweep is employed - by conducting readings on a wide range of frequencies, differences that may be missed on any one read frequency may show up in one of the many others.

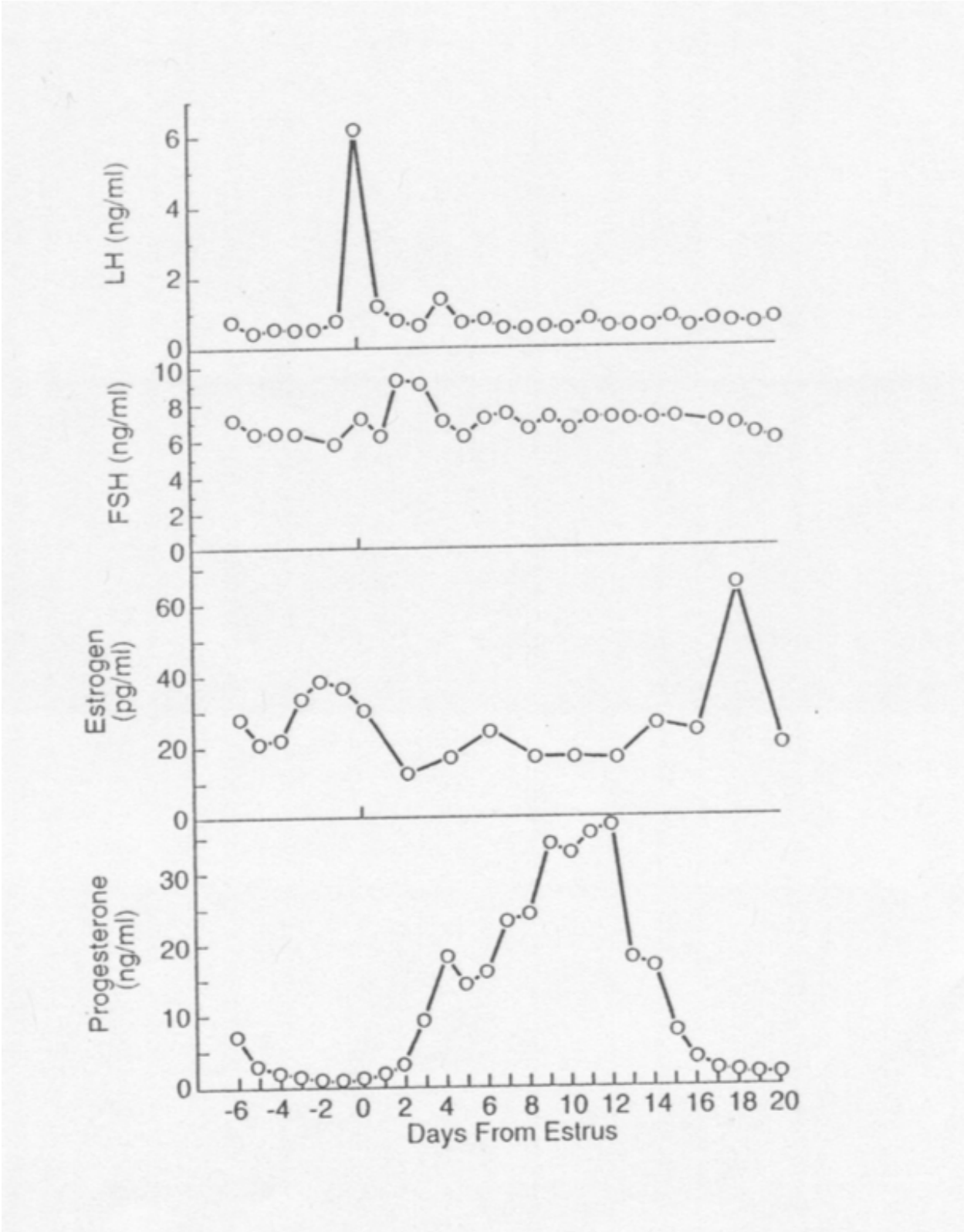


Figure 1.2: Variation of LH, FSH, Estradiol, and Progesterone during the estrous cycle of the sow. Reprinted from Animal Industry Report, Volume 665, Anderson, Lloyd L., Reproductive Biology of Pigs, Page No. 66, Copyright 2009, with permission from Iowa State University.

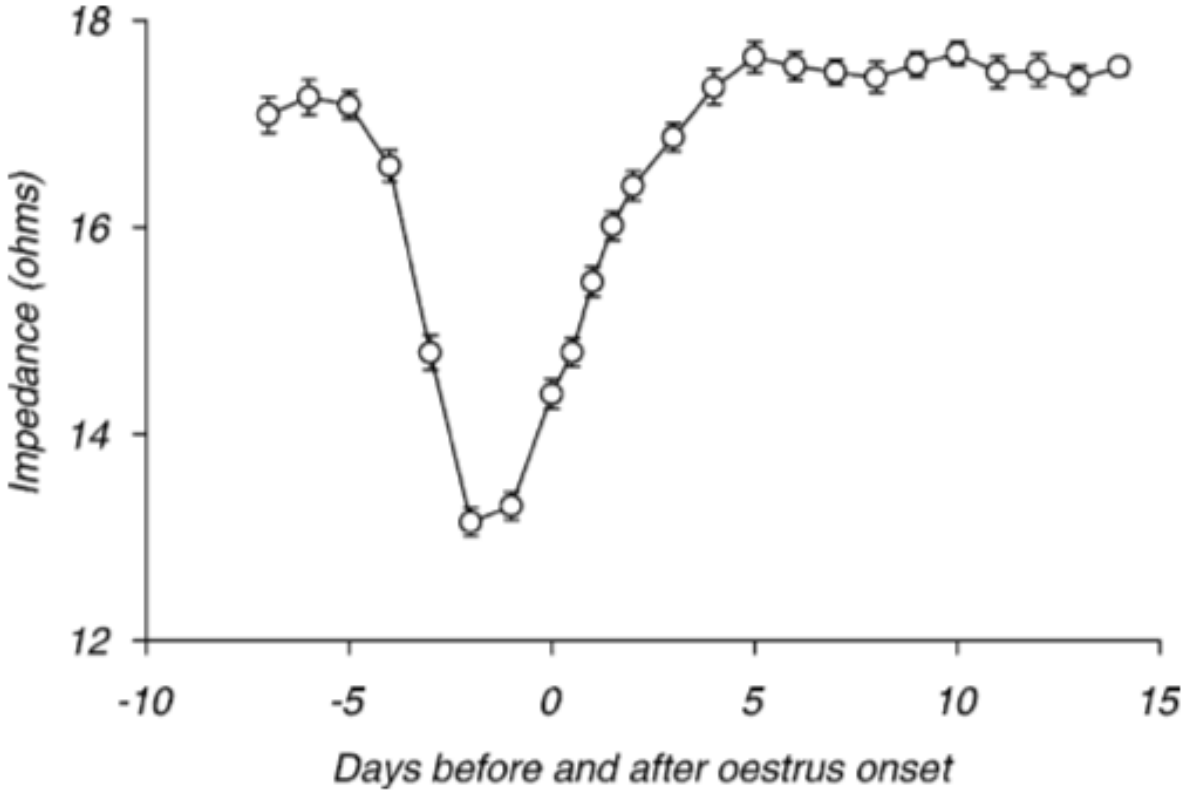


Figure 1.3: Variation in the impedance of the swine vaginal environment during the estrous cycle, mean values retrieved via EIS at 5700 Hz across 177 animals. Reprinted from Animal Reproduction Science, Volume 72, Řezáč, P. et al., Effect of sow parity on vaginal electrical impedance, Pages No. 223-234, Copyright 2002, with permission from Elsevier.

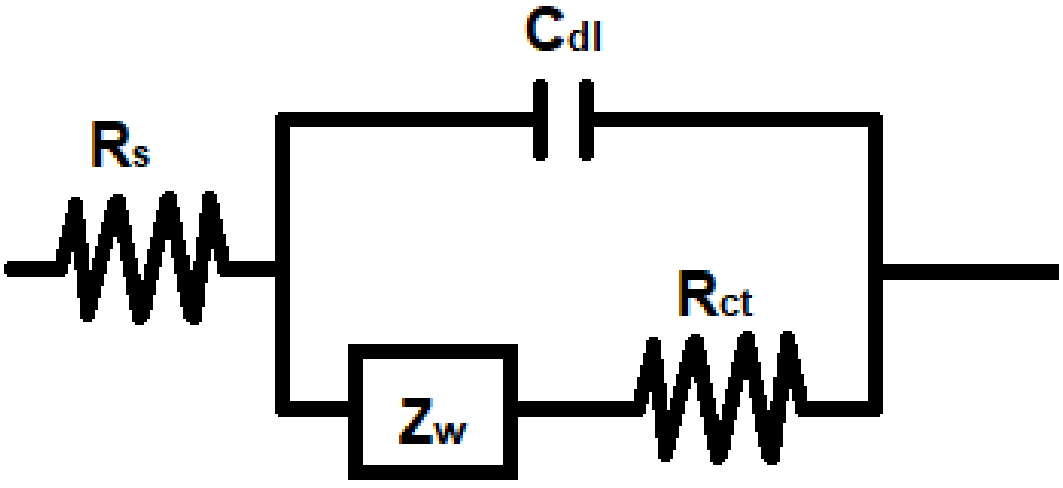


Figure 1.4: The Randles Circuit model.

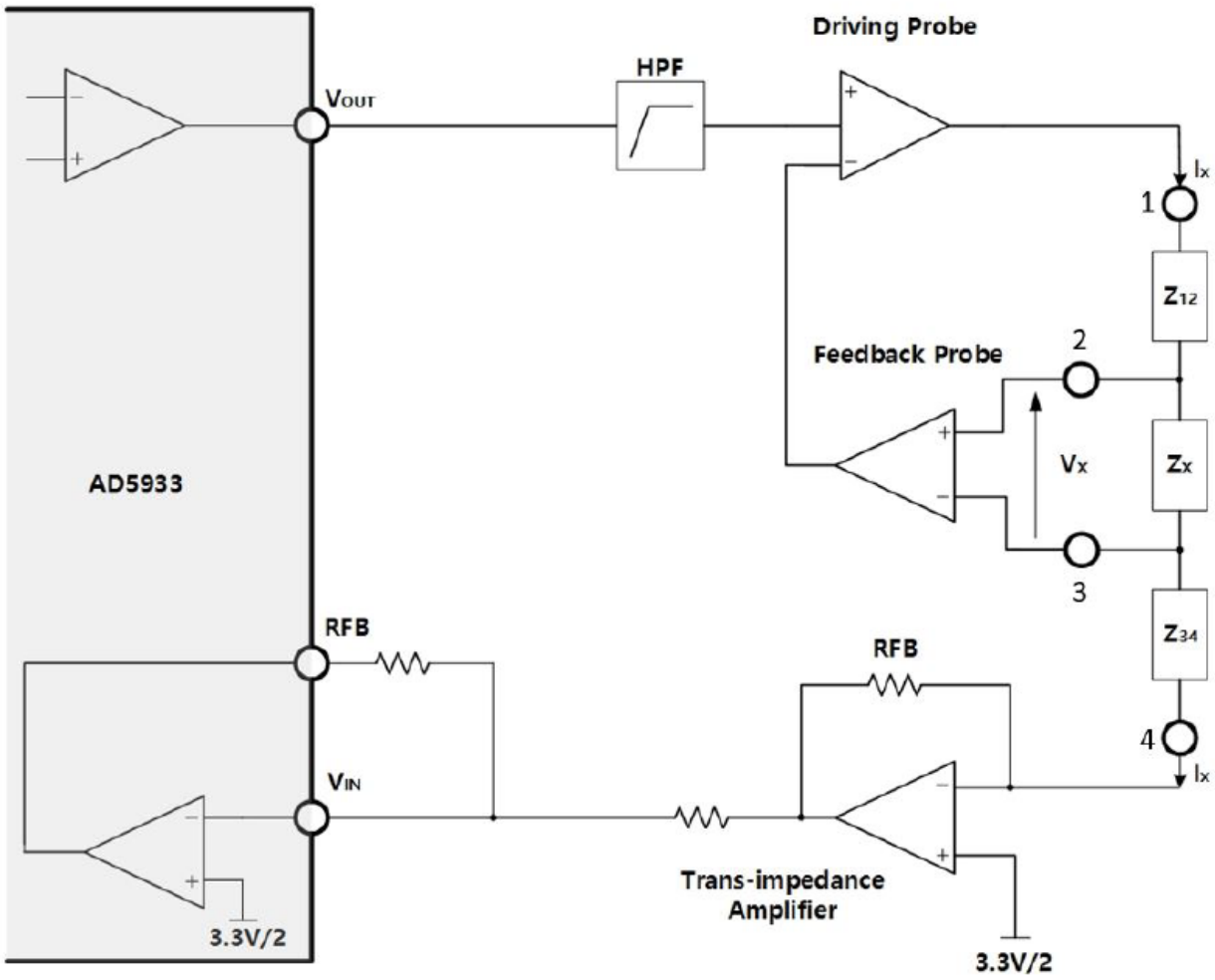


Figure 1.5: Reprinted from Seokchan Yoo et al. A low-cost, portable, web-based impedance spectroscope for agricultural applications. M.S. Thesis., 2017.

Chapter 2

Data Acquisition, Processing, and Visualization

The electrical impedance spectrometer has an AD5933 high-precision impedance measurement system-on-chip for the measurements. It uses a Bluetooth Low Energy capable microcontroller (Adafruit Bluefruit LE) for its fast data transfer speed, high processing speed, and its security, for the purpose of simplifying data logging and visualization. The device was built with low cost and ease of use in mind, factors with which many earlier devices struggled. [1]. Housed within the device's tip is a 100 ohm resistor, used for calibration. The reading is taken in the deep region of the vagina, past the urethra and near the cervix.

One setback that occurred during testing of devices between that of Yoo and the current model is that the animals began to experience a painful electrical shock due to an increase in the power used. This problem was rectified by decreasing the surface area of the electrode rings, after which adverse reactions from the sows ceased.

The capabilities of the device allowed for a frequency sweep to be performed in gathering data. On each measurement, 74 different frequencies were used: Every 100 Hz between 1k Hz and 6 kHz, then every 1 kHz between 6 kHz and 29 kHz. The frequency sweep is a step forward in EIS methods for reproductive predictions that this device explores [1], as some frequencies may be more useful than others.

The R language was used, not only due to its strong statistical capabilities [34], but also its wide use [35], which will make it easier for future work to take place. For its feedback provision and ease of use, the integrated development environment RStudio [36] was used.

Each model was created using a logistic regression between binary values. The function used was glm(), which is able to create models of various kinds depending on the family argument chosen. In this case, binomial was used, which links to a logistic model [37]. The result is that each frequency, an intercept, and potentially any other data inputs would each generate their own coefficient to be added into the logistic model. Each model then was subjected to stepwise Akaike Information Criterion (AIC) [38] reduction via the function stepAIC [39]. AIC is a method by which the quality of an individual model can be assessed, and is shown in Equation 2.1. In it, k is the number of parameters or coefficients, M is the maximum likelihood determined by the model, and AIC is the output variable, which is desired to be minimized.

$$AIC = -2 * \log(M) + 2k \tag{2.1}$$

The method employed was one of backwards steps. During each step, the model with all coefficients is tested against reduced models, each with one of the coefficients dropped, which reduces k by 1. That model which has the greatest drop to the AIC then replaces the previous one, and the next step repeats this process. This continues until the potential decline in M is so great that no coefficient can be dropped without raising the value of AIC. The reduction of k is considered important to the quality of the model because large numbers of parameters are likely to cause overfitting, a phenomenon in which a model produces excellent results for the data, but very poor results elsewhere. This is due excessive number of variables accommodating the noise of the data set, and thus creating a tortured fit which suits only the data set.

Two main types of graphical representation were used to analyze the models' performance: violin plots, and receiver operating characteristic (ROC) curves, through the ggplot function [40] in the ggplot2 package for R. Violin plots are an expansion on the box plot, displaying the probability density function graphically and reflecting it into a 'Violin' for ease of understanding [41]. ROC curves are a mechanism for the display of classifier systems' characteristics [42]. ROC is frequently used in a wide variety of fields, including medicine [43], fire and landslide hazard prediction [44, 45], meteorological forecasting [46]. In this case, ROC curves are primarily used to visualize the vital properties of the models - their True Positive Rate (TPR) and False Positive Rate (FPR), from which other properties such as sensitivity, selectivity, and precision can be derived at different points. Among these additional properties is the Area Under the ROC Curve (AUC), which varies from 0 to 1, with a value of 0.5 indicating model that is no better than randomly guessing, and a value of 1 indicating a model that is perfect (although this, in reality, an indicator that there is too little data in the set). AUC is similar to the Wilcoxon signed-ranks test in that higher values indicate a greater propensity to correctly rate a randomly chosen actual positive higher than a randomly chosen actual negative [47].

2.1 Frequency Data Processing

Prior to the specific analysis of any one stage, the relationship between frequency and impedance in the different stages was given scrutiny to see if any pattern would emerge, as shown in Figure 2.1. The result was a distinct trend in the post-weaning (or, estrus) data of decreasing impedance with higher frequencies used, while the early pregnancy data seemed to slightly increase, stabilize, and then finally fall starting at around 10 kHz. The pre-farrowing data displayed a trend between the two, largely stable with a small

downward trend up until above 10 kHz, at which point it dropped precipitously. This indicates that, at high frequencies, the drop in the diffusion-based Warburg impedance may dominate the increased capacitance due to large biomolecules being trapped between the electrodes.

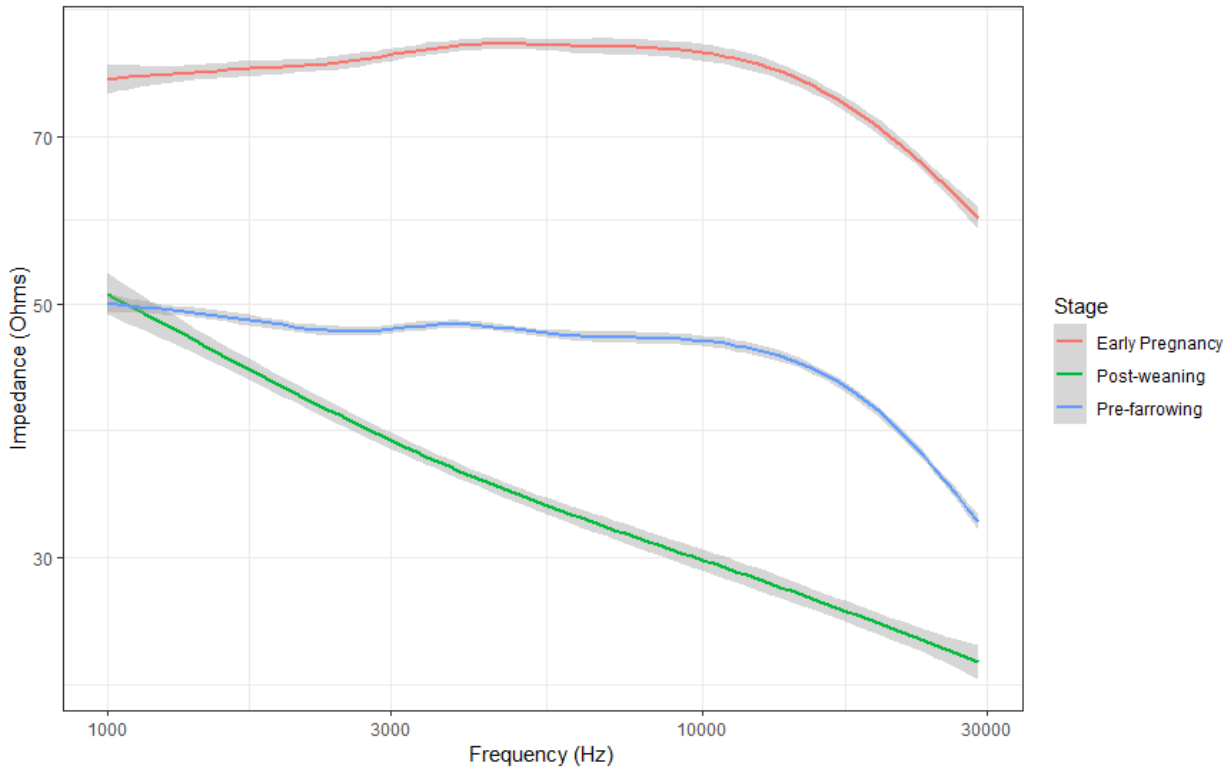


Figure 2.1: The association between impedance read and frequency used in each stage.

However, the data were not so simple. The sets of data gathered that make up the whole, from three different operators with months between, show that the readings varied between them in Figure 2.2. Here, the three different stages show a similar trend per each operator, but the most recent battery of data acquisition, the set gathered by Andre, shows a distinctly different pattern than the prior ones. This final set shows each as being constantly decreasing, and additionally, shows the post-weaning and pre-farrowing stages in a reversed order of impedance compared to the other two, which show the same general trend of similar impedance readings for most of the frequencies until a strong decrease at the highest levels.

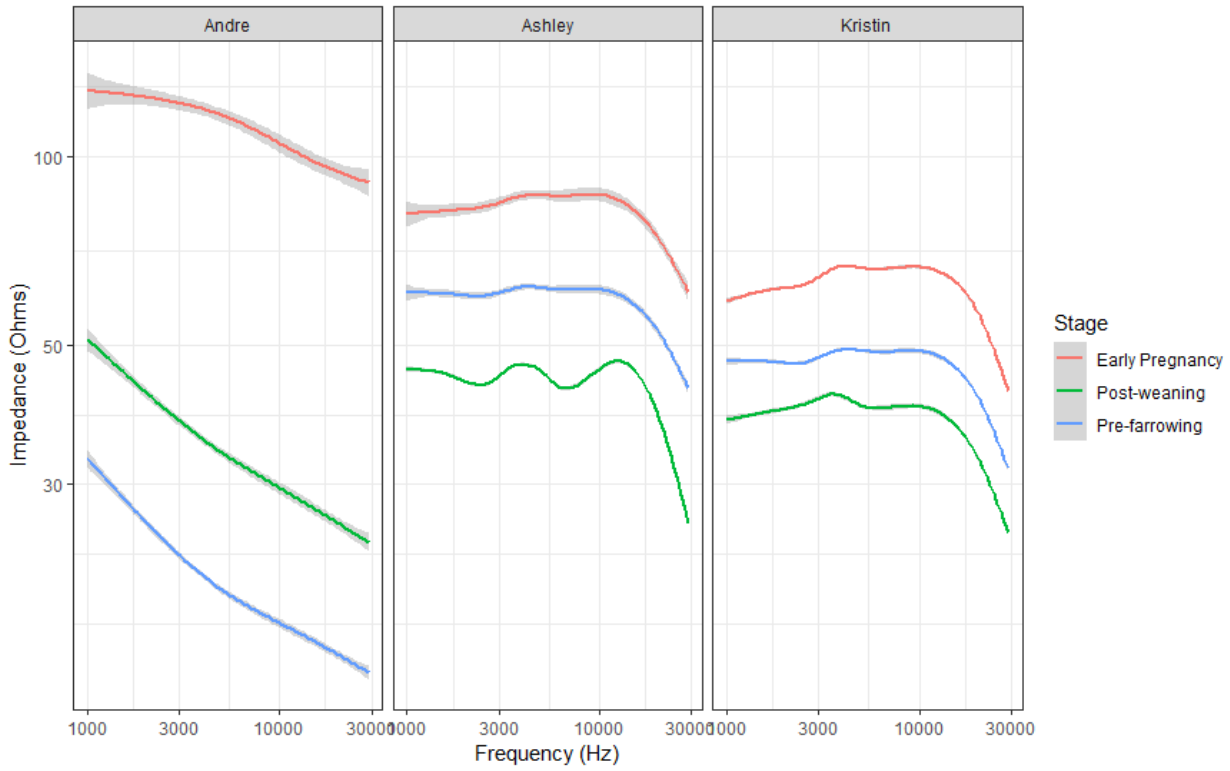


Figure 2.2: The association between impedance read and frequency used in each stage, by operator.

2.2 Fertility Data Processing

For data acquisition, the EIS device took 250 frequency-sweeping readings across 85 different animals. Of these readings, only 2 observations were taken six days before the onset of estrus, only 5 observations five days before, 45 observations four days, 82 observations three days, 80 observations two days, and 38 were taken one day prior to estrus. As such, data from the earliest days is sorely lacking, and thus the margin of error on such readings is enormous. It is, however, the later days that are more pressing to take readings on, so the comparative abundance of readings in the latter days is useful.

Contrary to the expected outcome of a strong dip in impedance prior to onset, as shown in Figure 1.3, the dip in impedance observed was, though present on each frequency, so small as to be insignificant in the face of the error, as shown in Figure 2.3. Starting at day -3 (labelled in relation to day 0 being the actual day of estrus), impedance rose across all frequencies, though not as strongly in the higher frequencies as in the lower frequencies. Additionally, higher frequencies were consistently observed to create lower impedances than higher frequencies throughout the period leading up to estrus, though the margin of error between particular frequencies meant that this trend was only somewhat applicable.

The model was based upon a classification of requiring imminent service, which was defined as the animal

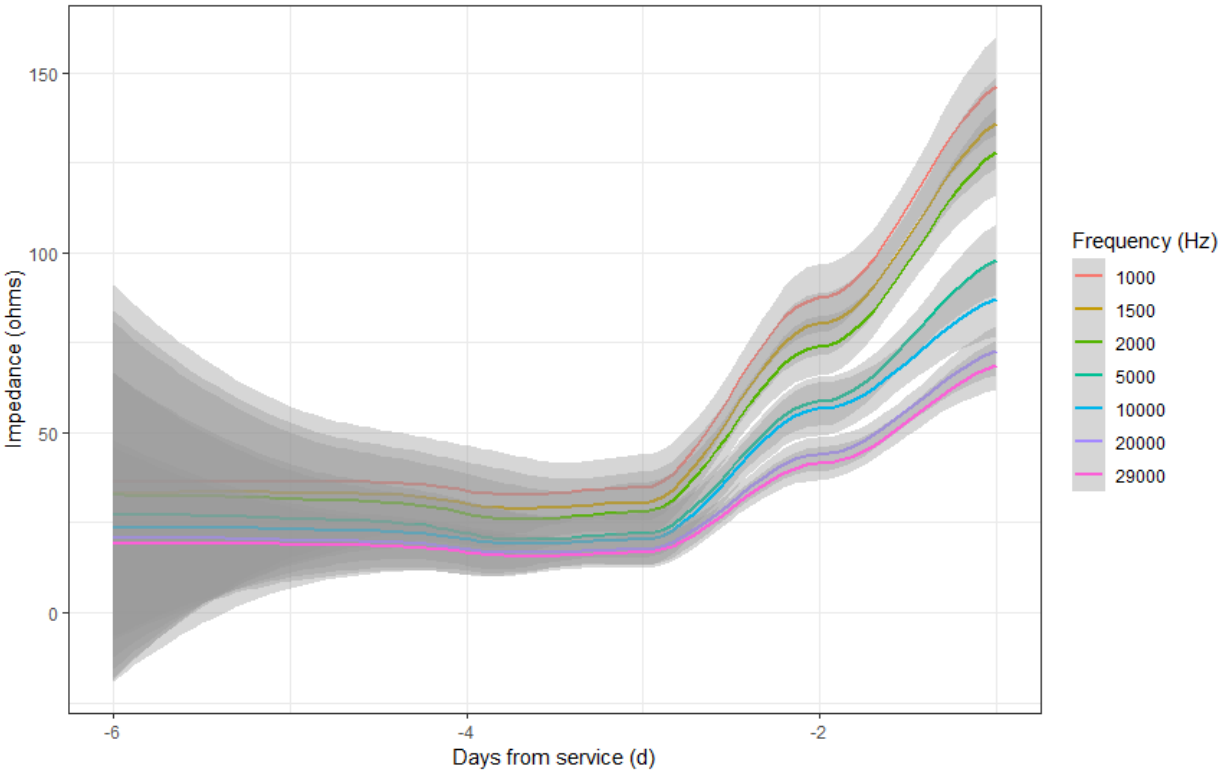


Figure 2.3: Graph tracking the impedance given at different frequencies leading up to estrus.

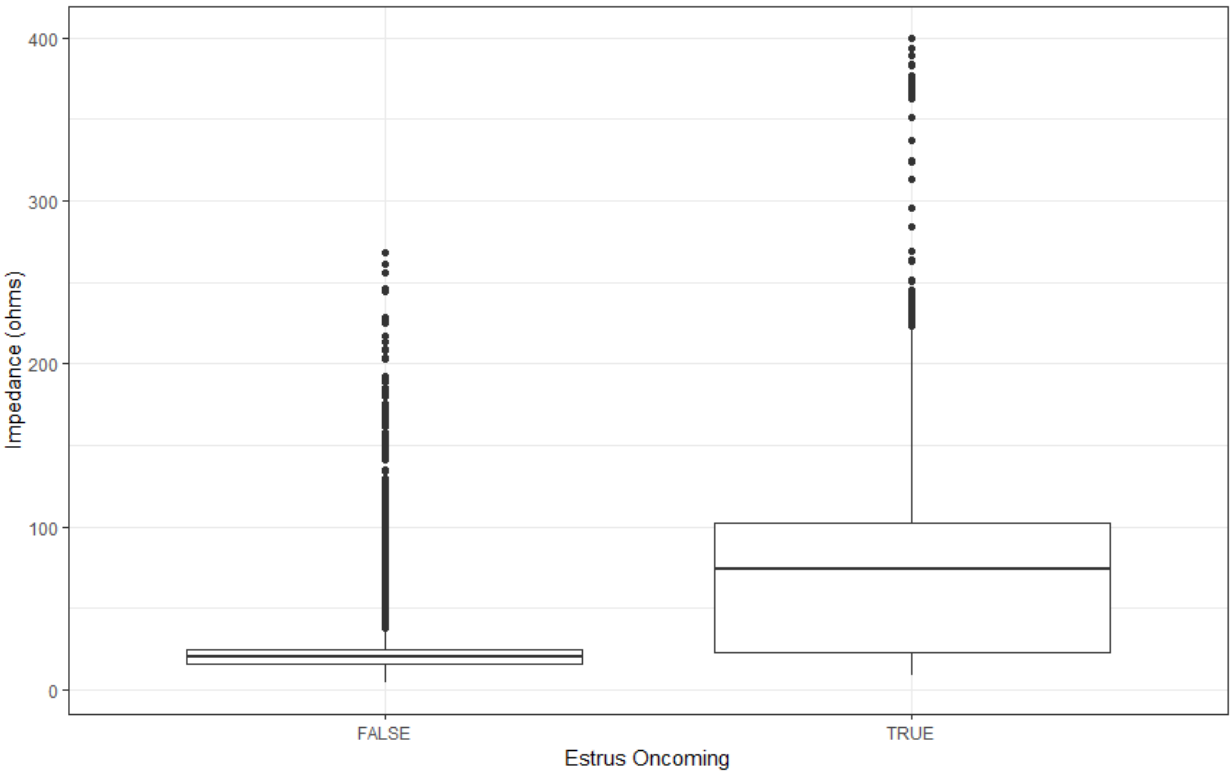


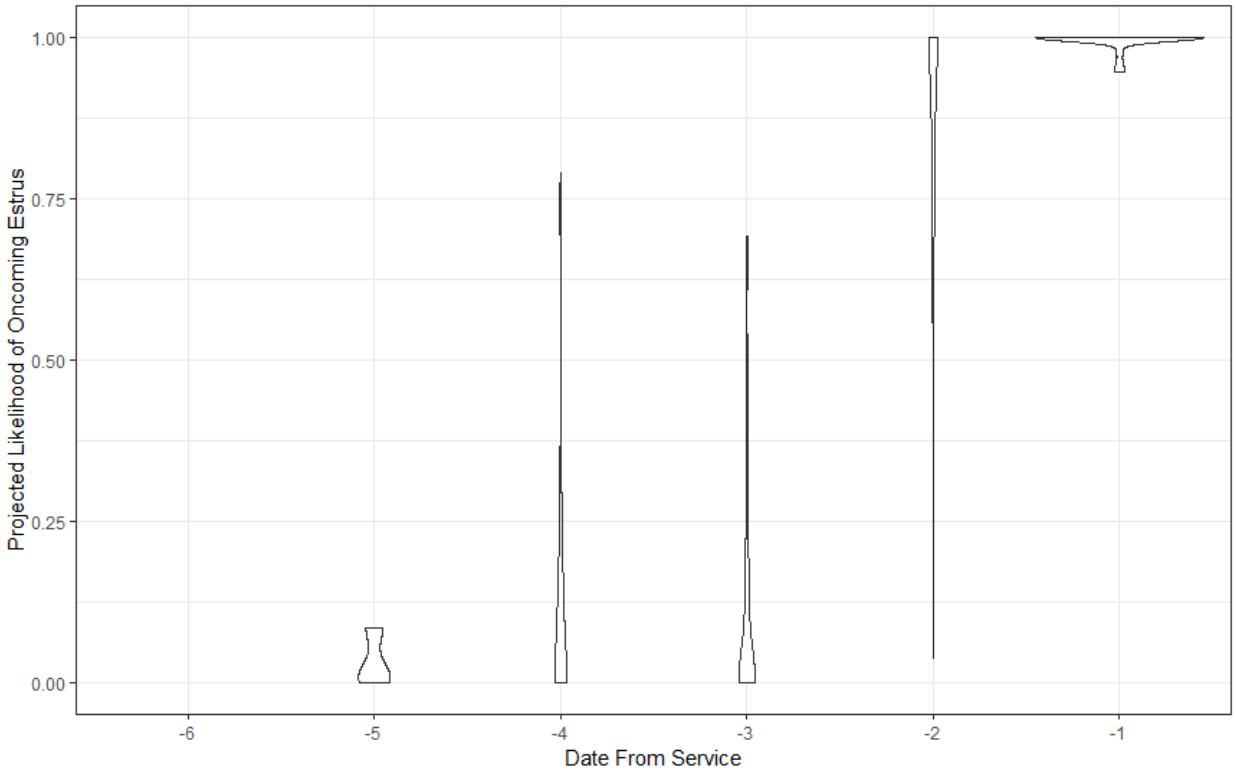
Figure 2.4: Box plot for all frequencies together.

being within two days of beginning estrus. Following logistic regression and reduction via the stepAIC function, the coefficients, shown in figure 2.1, are numerous. Within the table, the $Estimate$ is the value of the coefficient in the model. Positive values indicate that the effect of an increase in the impedance on that frequency indicates a greater likelihood of a positive result of coming back into estrus, negative values indicating the opposite relationship. The $StandardError$ is the error on this value, reflecting the level of uncertainty that the value of this coefficient is correct, with larger values indicating a larger confidence interval. The $z - value$ is the standard score, indicating how many standard deviations away from the non-predictive zero the value of the coefficient is, with higher magnitudes indicating a greater likelihood of the coefficient's significance to the model. Finally, the $Pr(=0)$ is the probability that the coefficient's effect on the model is no greater than zero, with lower values indicating more strongly indicitave results. Those with the strongest significance values included high frequency readings (particularly 20000 and 23000 Hz), as well as low frequency readings (such as 1800 and 4800 Hz).

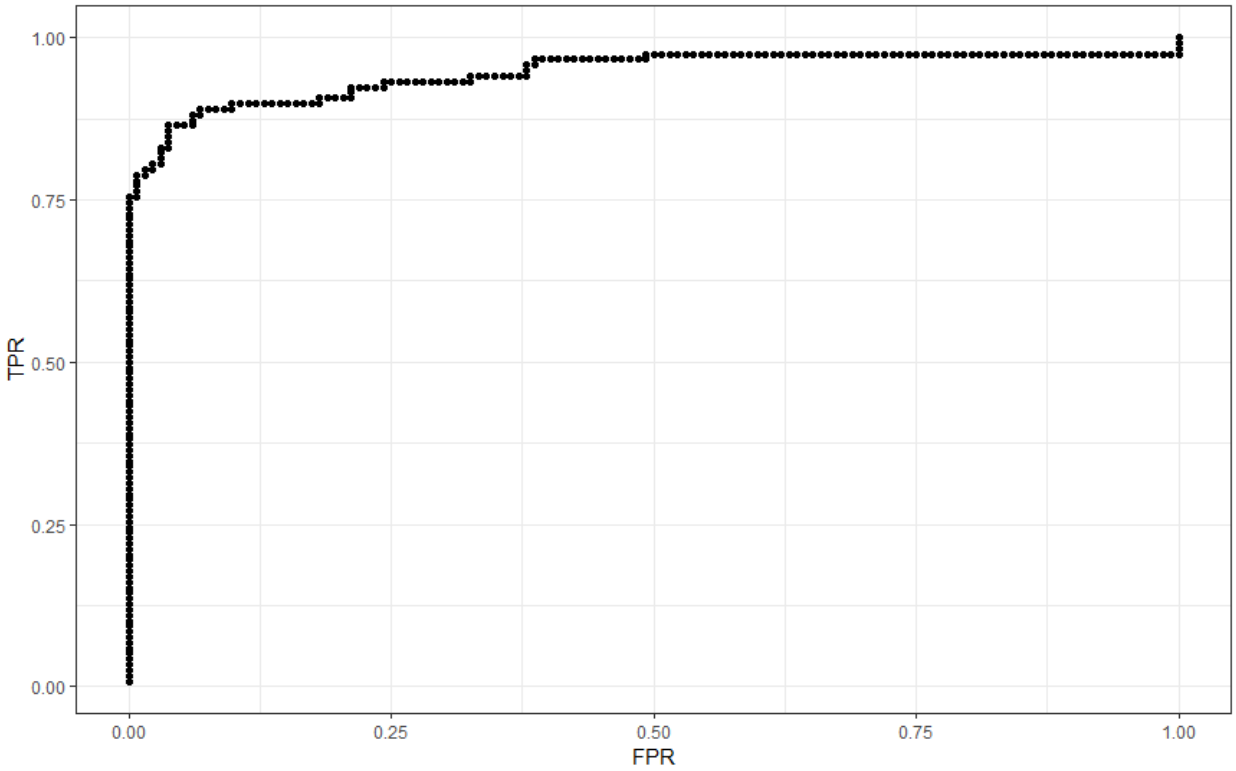
The resultant model displayed a strong ability to discern which measurements were within two days, as demonstrated in the violin diagram, Figure 2.5a. It shows a high TPR floor of 0.75 at a 0 FPR in the ROC curve, Figure 2.5b. Increases in TPR given a minute increase in FPR are substantial at first, but quickly level into very slow rises, with an AUC of 0.946.

Table 2.1: Table of coefficients for the estrus detection logistic model.

Frequency (Hz)	Estimate	Standard Error	z-value	Pr(=0)
(Intercept)	5·52	2·80	1·97	4.84E-02
1100	−2·95	0·929	−3·18	1.49E-03
1200	7·42	2·25	3·29	1.00E-03
1300	−5·81	1·93	−3·01	2.57E-03
1600	−2·64	1·03	−2·57	1.02E-02
1700	14·7	3·87	3·80	1.43E-04
1800	−12·1	3·08	−3·93	8.56E-05
2400	3·32	1·57	2·11	3.49E-02
2500	−1·25	0·816	−1·53	1.27E-01
2600	1·15	1·03	1·11	2.67E-01
2900	−3·27	1·54	−2·12	3.44E-02
3000	7·81	2·20	3·55	3.82E-04
3100	−5·87	2·04	−2·88	4.03E-03
3300	−5·90	2·78	−2·12	3.38E-02
3400	6·04	2·42	2·49	1.26E-02
3500	6·23	2·28	2·73	6.26E-03
3600	−18·3	5·86	−3·13	1.76E-03
3700	10·3	4·91	2·10	3.60E-02
3800	8·21	4·82	1·70	8.87E-02
3900	−9·77	3·67	−2·66	7.84E-03
4000	3·26	1·34	2·43	1.52E-02
4400	6·18	2·29	2·69	7.04E-03
4600	−9·08	3·90	−2·33	1.99E-02
4800	−19·1	5·51	−3·46	5.48E-04
4900	18·2	4·83	3·76	1.67E-04
5000	4·27	1·42	2·99	2.75E-03
5300	5·31	2·47	2·15	3.14E-02
5600	−11·3	4·33	−2·61	9.05E-03
5700	23·1	7·12	3·25	1.17E-03
5800	−25·0	7·70	−3·25	1.16E-03
5900	10·0	4·20	2·39	1.70E-02
6000	−4·91	1·98	−2·48	1.30E-02
9000	6·96	3·03	2·30	2.15E-02
10000	−7·85	3·09	−2·54	1.10E-02
11000	3·03	1·94	1·56	1.19E-01
12000	−4·73	2·36	−2·00	4.52E-02
13000	−5·45	3·85	−1·42	1.57E-01
14000	12·8	6·60	1·94	5.20E-02
15000	−10·3	5·58	−1·85	6.43E-02
16000	17·4	6·99	2·48	1.31E-02
19000	−27·6	8·65	−3·19	1.44E-03
20000	35·0	8·57	4·09	4.30E-05
23000	−21·2	6·30	−3·36	7.66E-04
24000	−13·3	6·06	−2·19	2.87E-02
25000	14·1	5·95	2·38	1.74E-02
26000	11·5	7·50	1·53	1.26E-01
28000	−24·7	9·02	−2·74	6.13E-03
29000	15·4	5·69	2·71	6.74E-03



(a) Violin diagram, showing the probability density of estimates for true positives and negatives.



(b) ROC curve, visualising the diagnostic capacity of the model.

Figure 2.5: Violin diagram and ROC curve for Estrus, using 250 observations across 85 animals.

2.3 Pregnancy Data Processing

Following the removal of outlying and erroneous reads, there were 395 useful observations gathered across 227 animals. Of these, 50 observations were of 29 animals that failed to become pregnant, and correspondingly, 345 observations taken of pregnant states across 198 animals. Of the readings, 26 were taken on animals 16 days after insemination, 60 observations on day 17, 7 observations on day 18, 108 observations on day 19, 113 observations on day 20, and 81 observations were taken 21 days after insemination.

The coefficients, following the AIC reduction, covered a wide range of frequencies, as shown in Table 2.2, with low frequencies (such as 2500, 3000, and 3300 Hz) as well as high frequencies (such as 20000, 22000, and 24000 Hz) being highly significant.

Table 2.2: Table of coefficients for the pregnancy detection logistic model.

Frequency (Hz)	Estimate	Standard Error	z-value	Pr(=0)
(Intercept)	−3·58	1·19	−3·01	2.58E-03
1100	−0·166	0·104	−1·60	1.10E-01
1300	0·617	0·281	2·19	2.82E-02
1400	−0·461	0·226	−2·04	4.12E-02
2500	0·377	0·104	3·62	3.00E-04
3000	−0·614	0·164	−3·74	1.84E-04
3200	0·770	0·401	1·92	5.45E-02
3300	−1·72	0·510	−3·38	7.20E-04
3400	0·909	0·297	3·06	2.24E-03
3700	0·748	0·264	2·83	4.63E-03
4200	−0·523	0·200	−2·61	9.03E-03
4400	0·480	0·286	1·68	9.33E-02
4500	−0·295	0·173	−1·70	8.90E-02
4700	−0·885	0·304	−2·91	3.60E-03
4800	0·870	0·344	2·53	1.14E-02
5000	−0·193	0·105	−1·85	6.47E-02
5100	1·04	0·390	2·67	7.69E-03
5200	−1·09	0·387	−2·81	4.91E-03
5600	1·207	0·426	2·84	4.58E-03
5700	−1·96	0·509	−3·84	1.24E-04
5800	0·957	0·425	2·25	2.43E-02
10000	−0·460	0·179	−2·57	1.02E-02
12000	0·737	0·311	2·37	1.79E-02
14000	−0·457	0·312	−1·46	1.43E-01
16000	0·452	0·263	1·72	8.53E-02
20000	−1·18	0·299	−3·96	7.59E-05
22000	2·05	0·466	4·40	1.06E-05
24000	−1·16	0·291	−3·99	6.66E-05

Taken across the entire range, the model shows a strong TPR, however, its FPR varies wildly by day. This is likely due to two factors - variance in the number of observations per day, and a separation in the patterns between early and late days in the observed period, likely due to the onset of estrus in non-pregnant

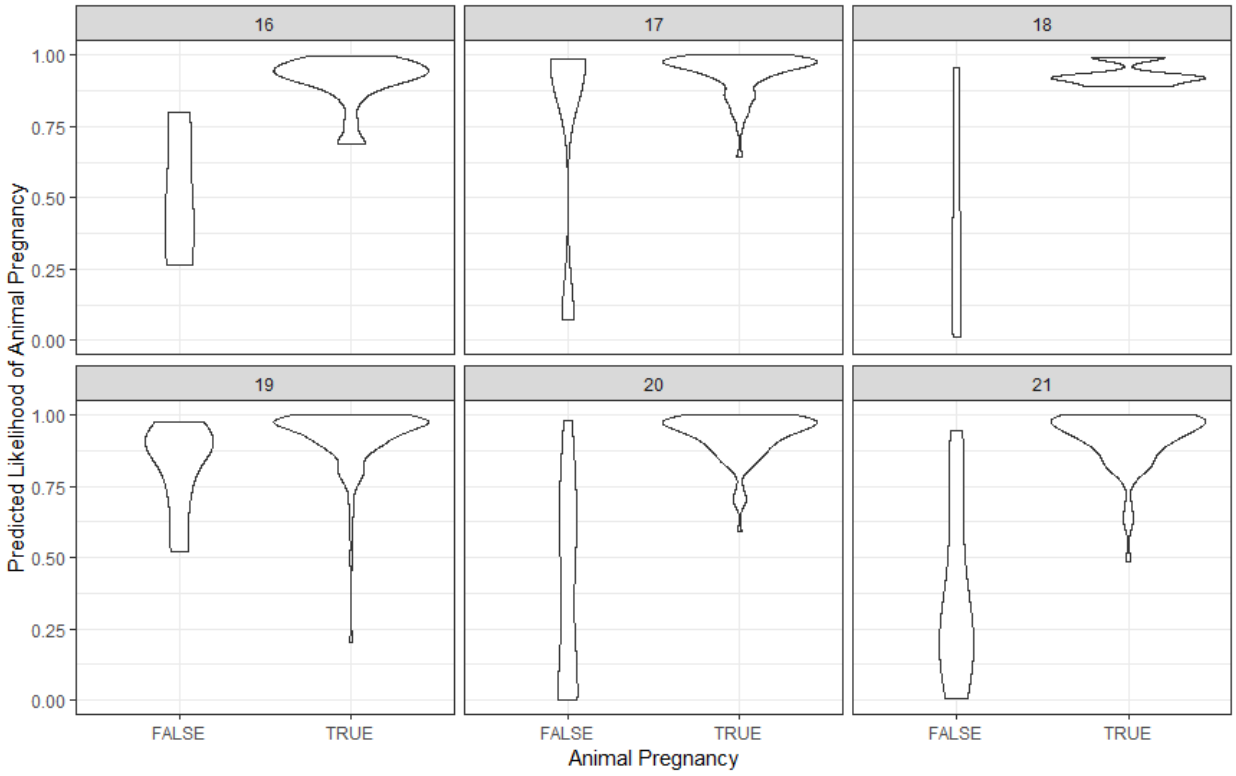
animals. As shown in Figure 2.6a, the model has a strong tendency to give false positives on days 17 and 19, reflected by the wide violins at high p-values on the false columns. Because of this, the ROC curve of Figure 2.6b does not give an accurate picture of what is happening beneath the surface with its AUC of 0.839. The particular nature of the days with regards to the accuracy of the model indicated a possible separation in the predictive power of different frequencies and impedances on different days. The next steps, then, were to divide the data set and its models into smaller sets of days. Many of these came back as models which appeared to predict with perfect accuracy, but these were discarded as the results of overfitting.

The model on the range of days 17-20 produced similarly poor results for day 17, but greatly reduced the FPR of predictions on day 19, as shown in Figure 2.7a. This improvement with regards to day 19 is shown in the ROC curve of Figure 2.7b, which increases in TPR with regards to FPR far more quickly than the entire set's model in Figure 2.6b. For the range of days 17-21, however, the model was no better for days 17 or 19 than for the full model, as shown when one compares Figures 2.6a and 2.8a. Matching this, it produced a similar ROC curve, Figure 2.8b. Days 18-21 also showed little improvement in the model's tendency towards false positives on day 19, as shown in Figure 2.9a, but its lack of confounding from day 17's poor showing yielded a modestly improved ROC curve in Figure 2.9b.

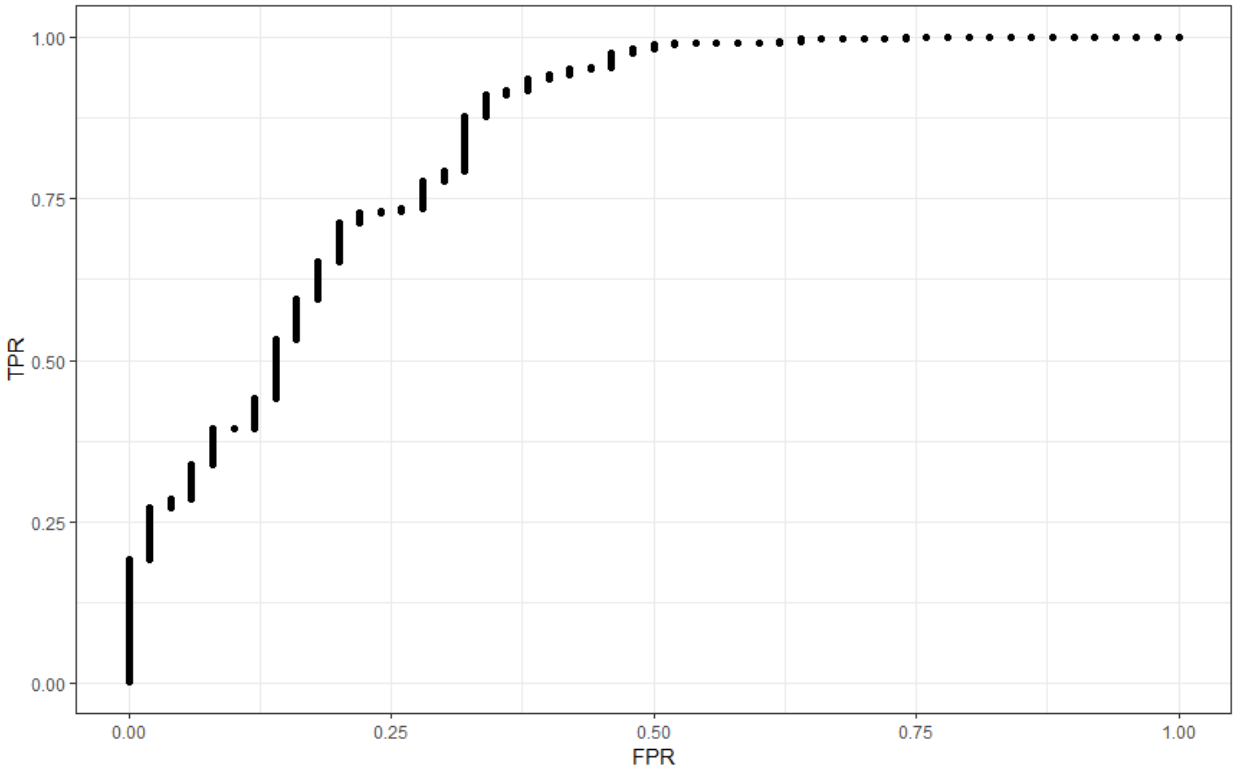
Perhaps the greatest indicator of a separation between early and late data, however, was that of the model built on days 19-21.Notably, this model includes most of the data, with 302 observations across 162 animals, reducing the likelihood that the improvements in the model are due simply to overfitting. The violin plots of Figure 2.10a show a far lower FPR in day 19, and this model in particular yields the strongest ROC curve in Figure 2.10b, with an AUC of 0.939. The reason for this is believed to be due to the onset of estrus in non-pregnant animals creating significantly different data for the purposes of prediction than those animals in the earlier data sets. As such, the different levels of importance ascribed to different pieces of data result in some days being represented poorly. However, when only late data is shown, the model is much better, as it is able to ignore the different trends present in the early data.

With this separation between early and late data on non-pregnant animals, the possibility that the onset of estrus is a key factor in this separation is a highly likely one. This supplied an even greater possibility to explore if the estrus detection model would be useful for the pregnancy data. After all, for those animals considered to not be pregnant, it is of considerable importance to detect whether or not the animal is coming back into heat. Animals which are not coming back into heat doubly display reproductive failure: failure to become pregnant, and failure to reenter estrus. It is important to identify such animals, as they are unproductive and costly to maintain. The model for estrus detection was thus applied to the non-pregnant animals, with the results shown in Figure 2.11, five observations, across four animals, were determined to

not be going into estrus by the model, while there were 45 observations determined to be going into estrus across 25 animals. This is not conclusive, though, as data with regards to whether or not these animals did or did not enter estrus is not available. However, it does indicate that the estrus model could potentially be applicable to animals determined not to be pregnant, or to reinforce the evaluation when they are pregnant.

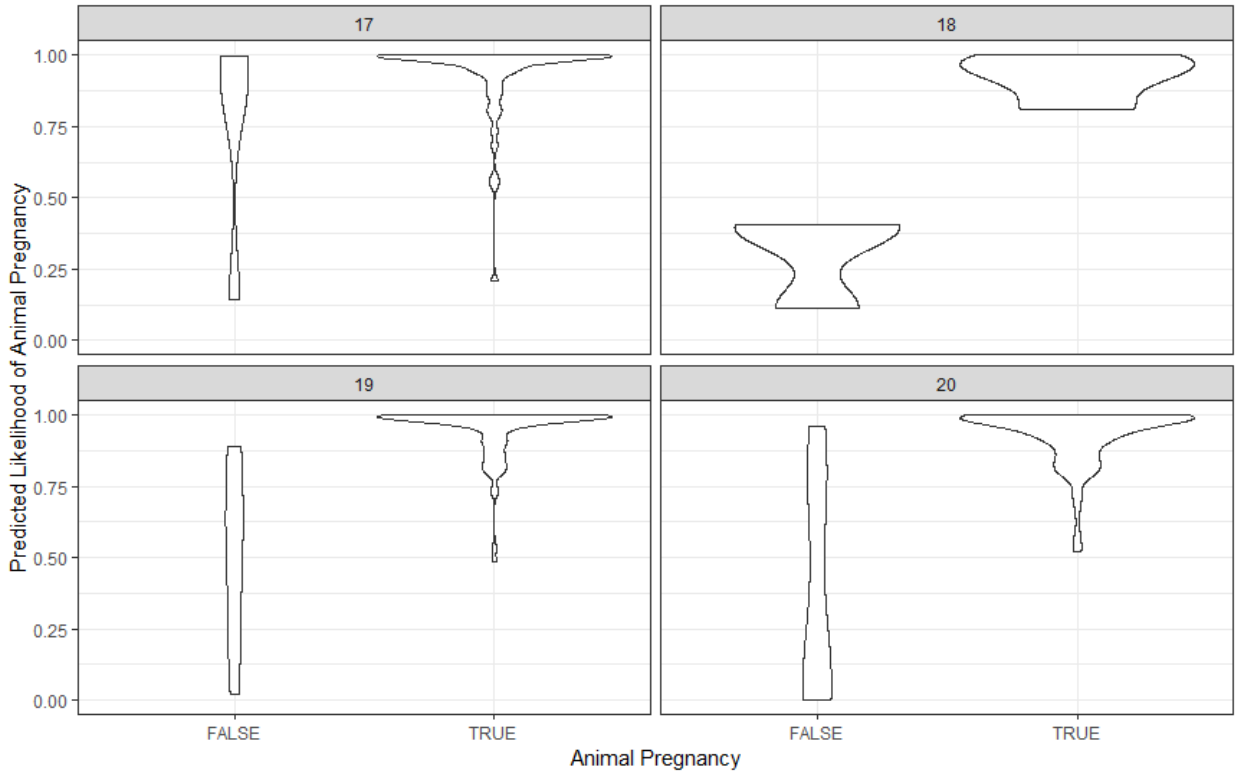


(a) Violin diagram, showing the probability density of estimates for true positives and negatives.

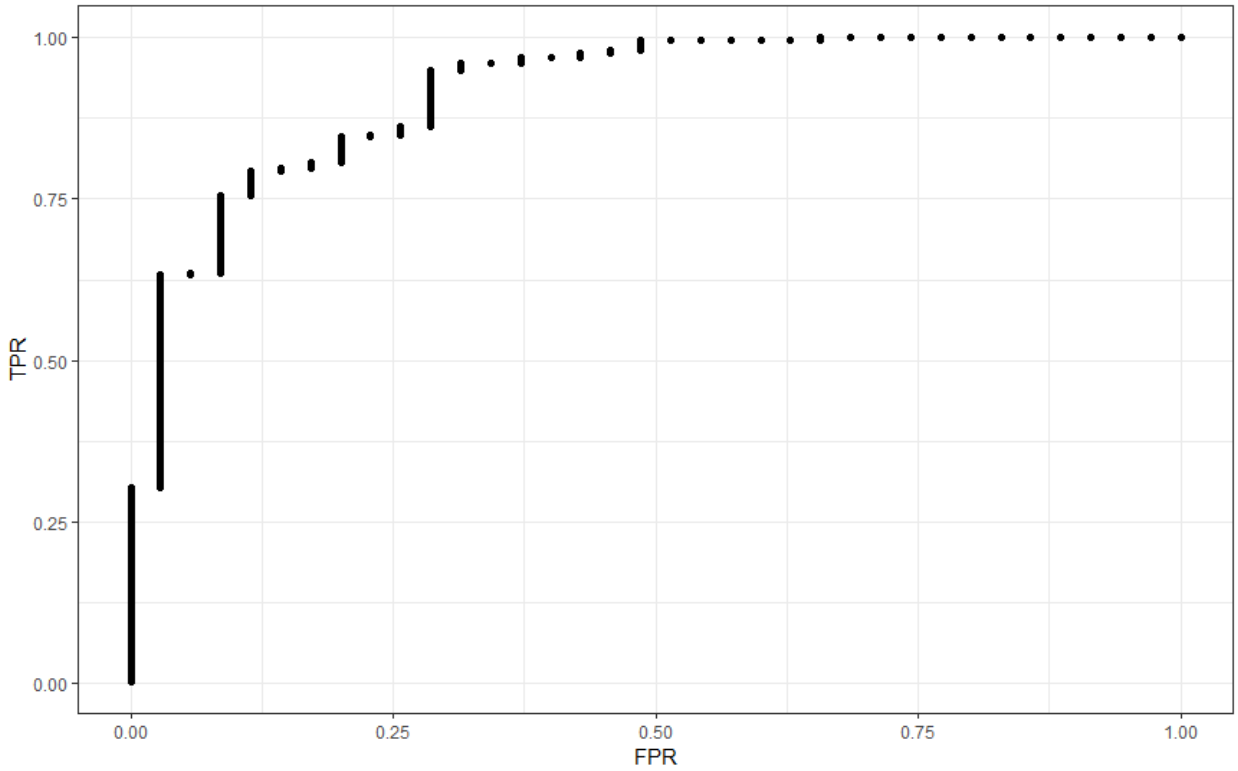


(b) ROC curve, visualising the diagnostic capacity of the model.

Figure 2.6: Violin diagram and ROC curve for whether pregnant (True) or non-pregnant (False) in the 16-21 day range, using 395 observations across 227 animals.

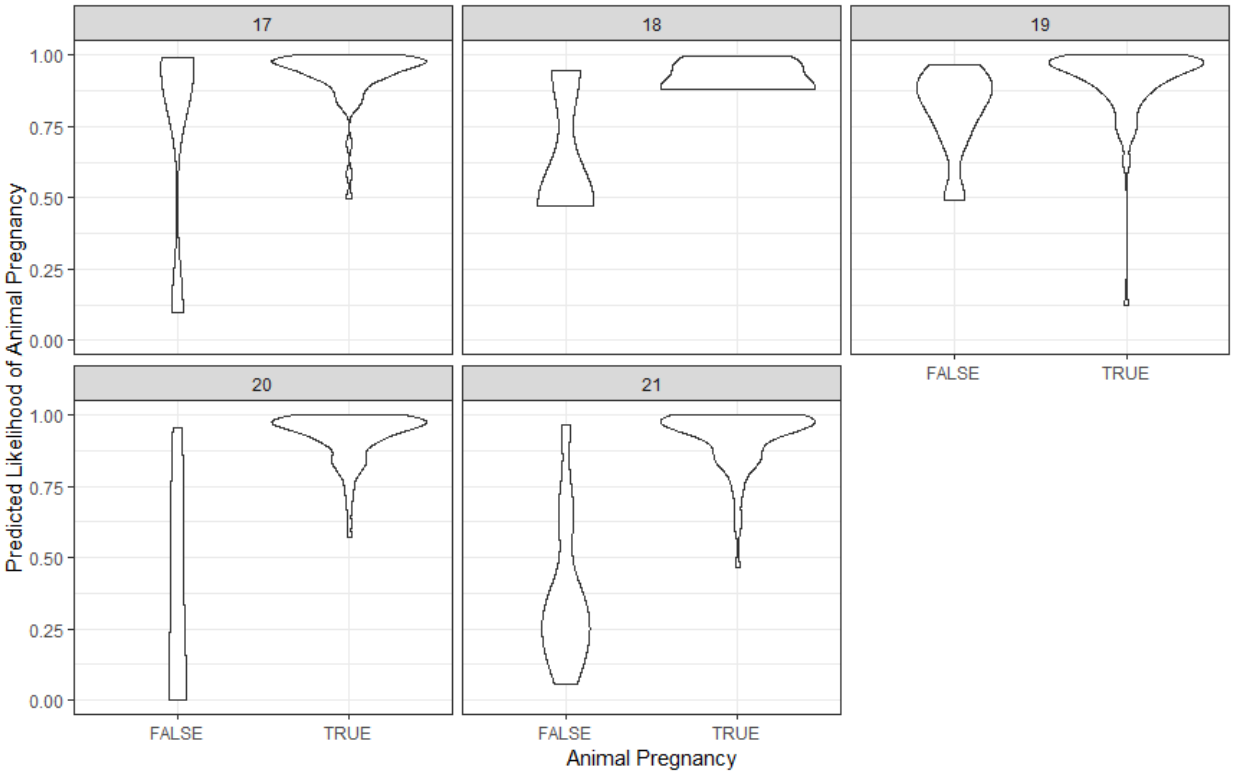


(a) Violin diagram, showing the probability density of estimates for true positives and negatives.

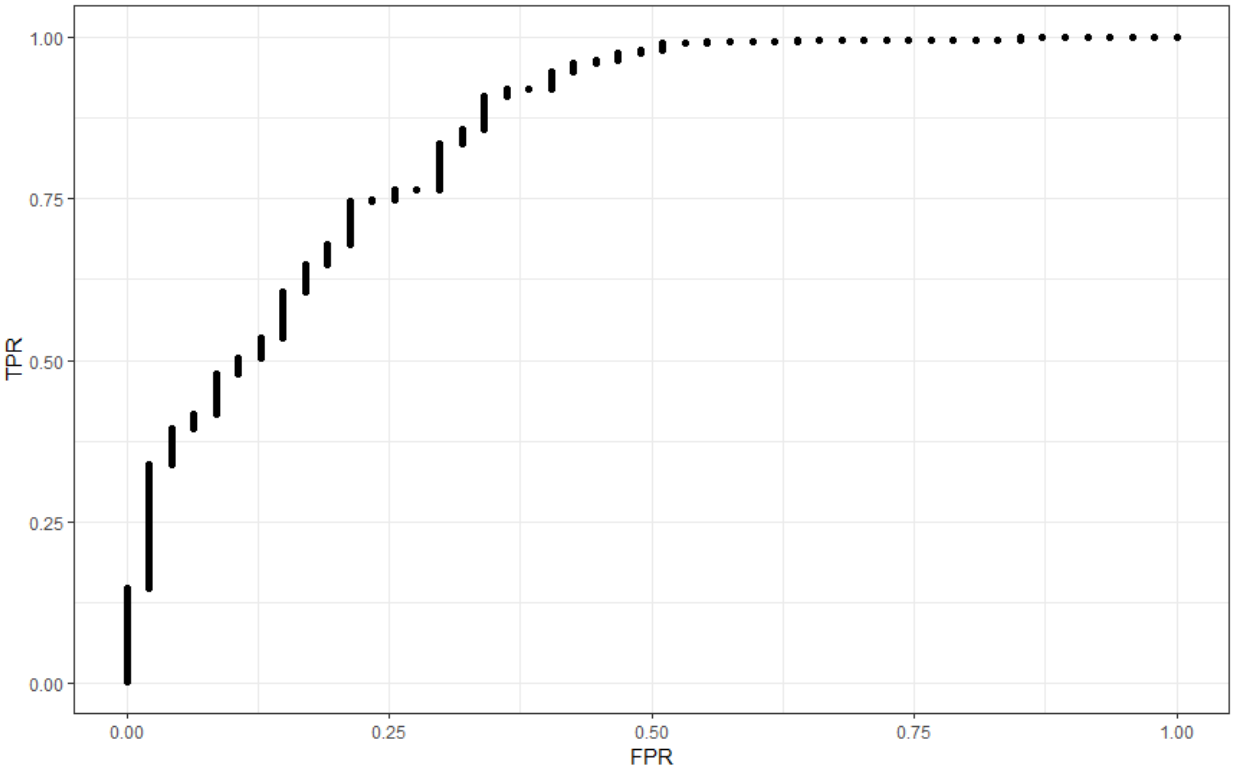


(b) ROC curve, visualising the diagnostic capacity of the model.

Figure 2.7: Violin diagram and ROC curve for whether pregnant (True) or non-pregnant (False) in the 17-20 day range, using 288 observations acros 221 animals.

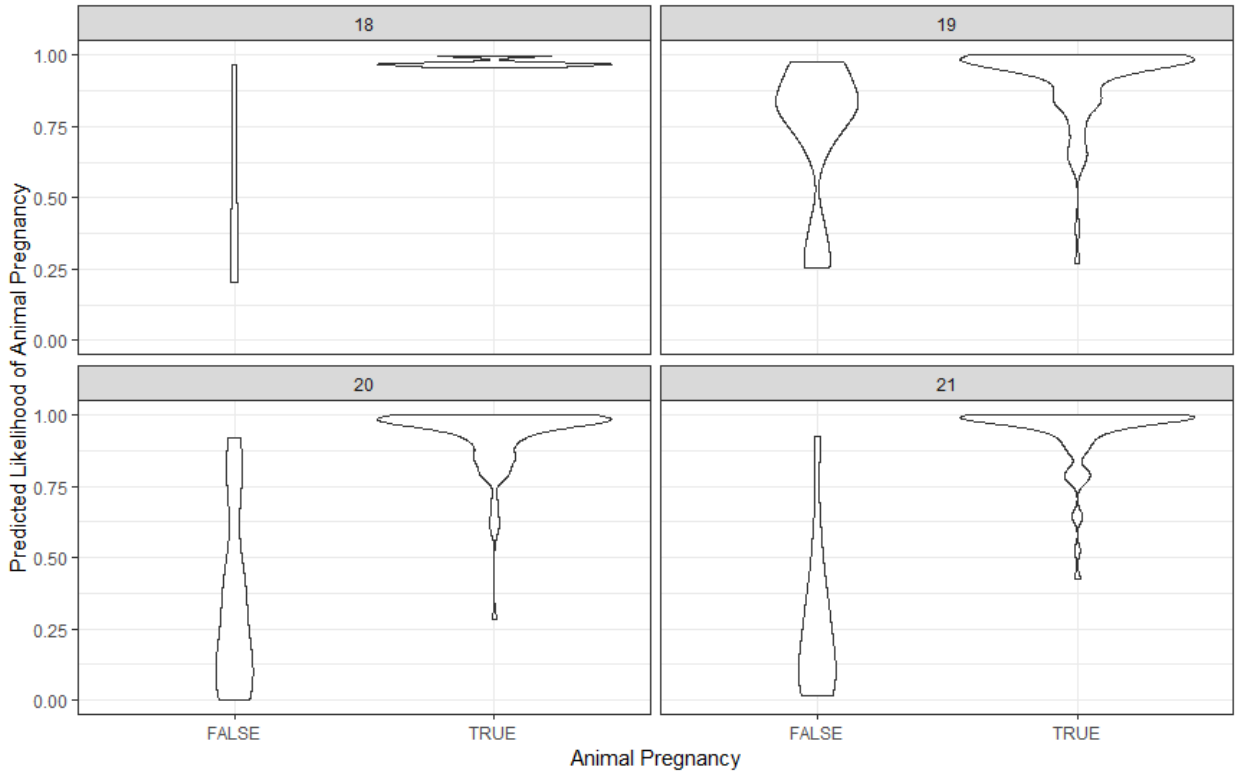


(a) Violin diagram, showing the probability density of estimates for true positives and negatives.

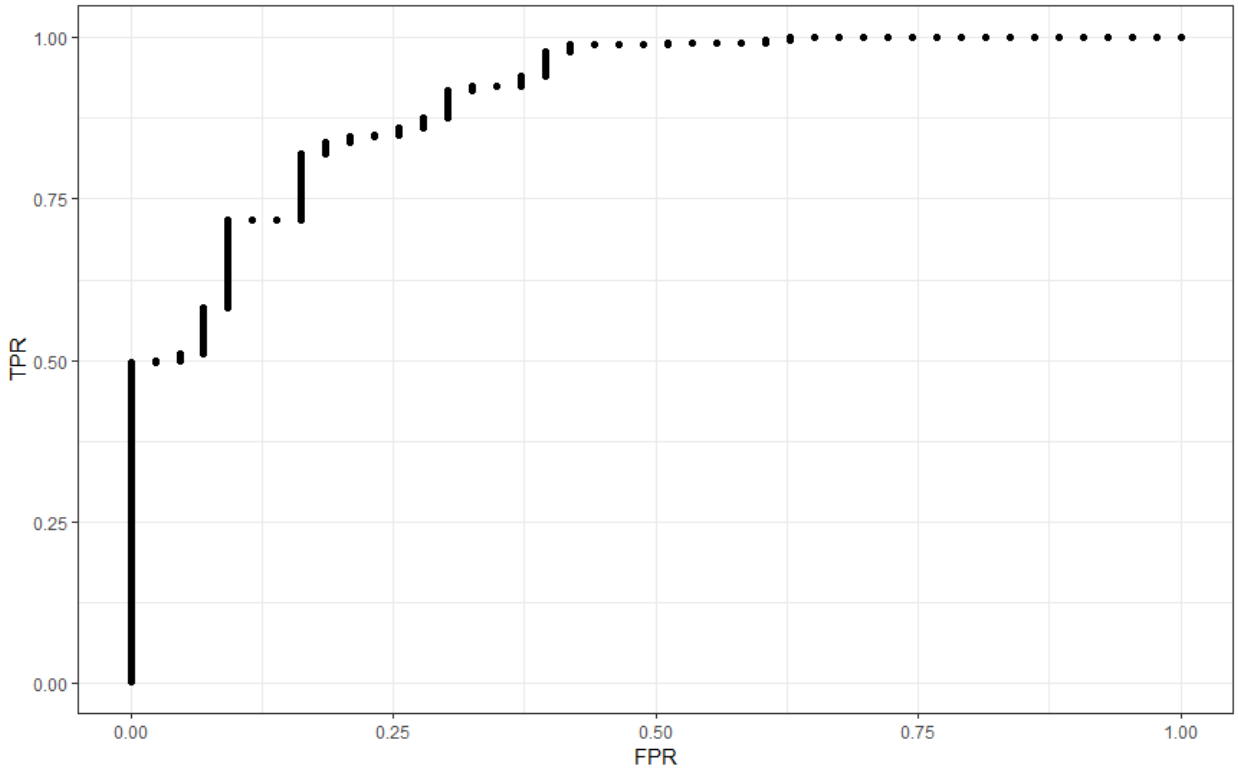


(b) ROC curve, visualising the diagnostic capacity of the model.

Figure 2.8: Violin diagram and ROC curve for whether pregnant (True) or non-pregnant (False) in the 17-21 day range, using 369 observations across 224 animals.

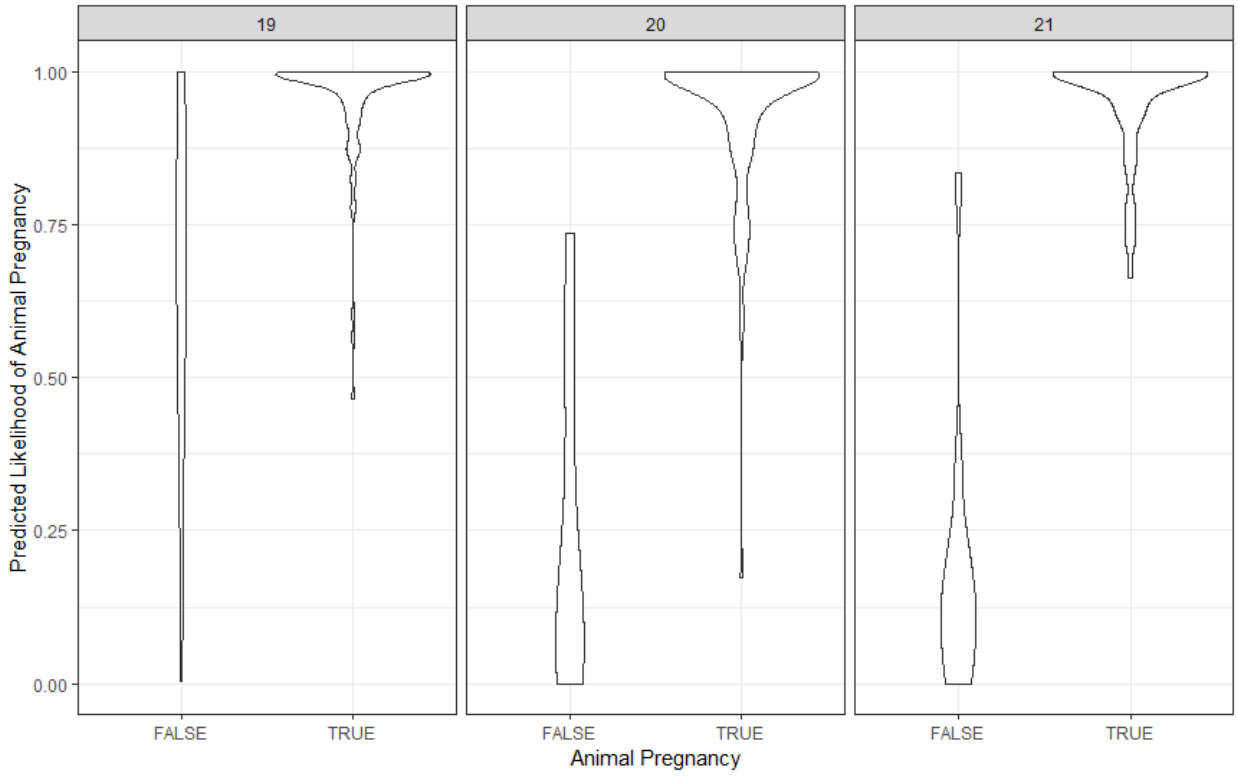


(a) Violin diagram, showing the probability density of estimates for true positives and negatives.

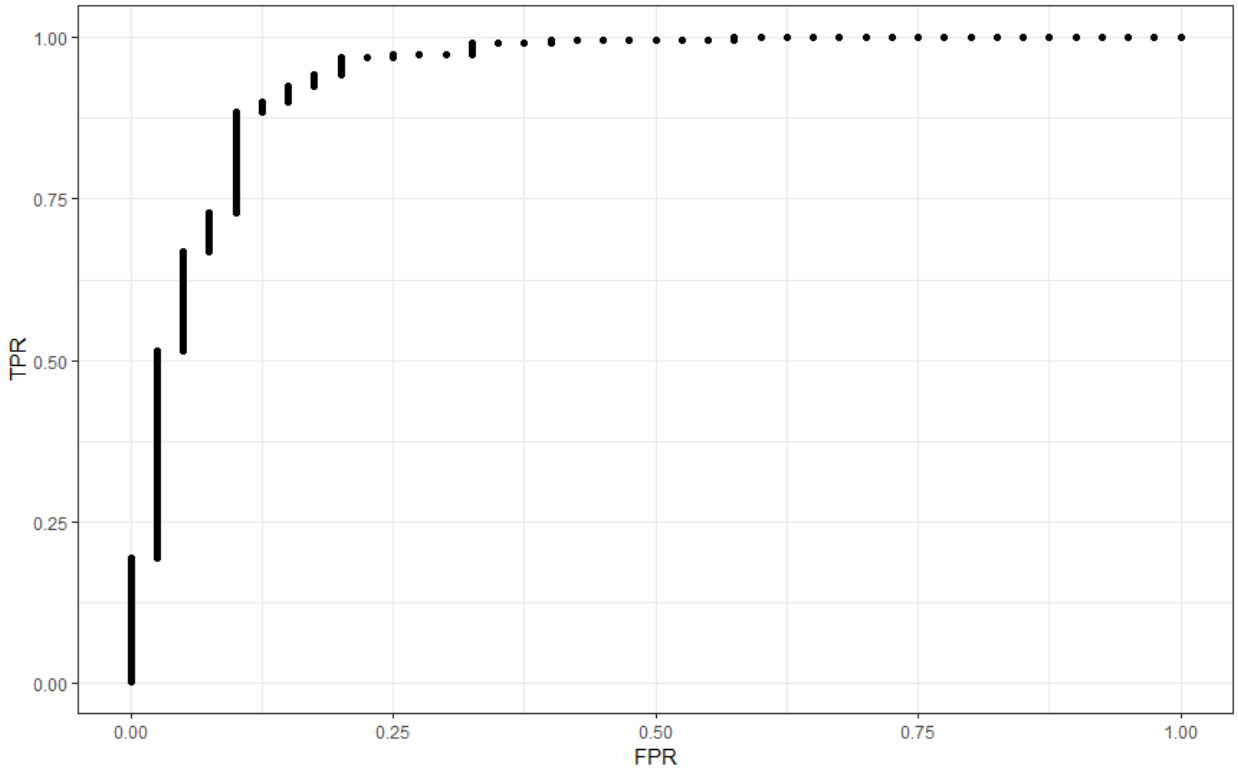


(b) ROC curve, visualising the diagnostic capacity of the model.

Figure 2.9: Violin diagram and ROC curve for whether pregnant (True) or non-pregnant (False) in the 18-21 day range, using 309 observations across 164 animals.



(a) Violin diagram, showing the probability density of estimates for true positives and negatives.



(b) ROC curve, visualising the diagnostic capacity of the model.

Figure 2.10: Violin diagram and ROC curve for whether pregnant (True) or non-pregnant (False) in the 19-21 day range, using 302 observations across 162 animals.

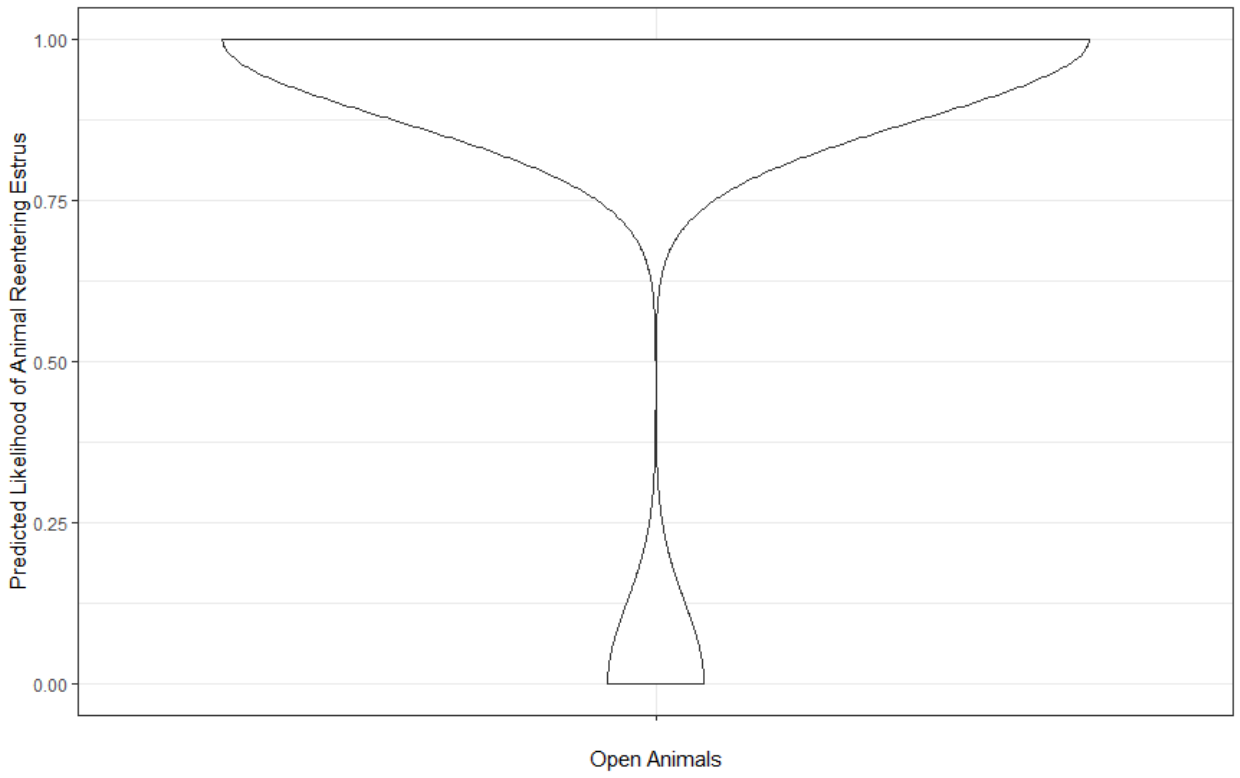


Figure 2.11: Violin diagram showing the results of applying the model for estrus to the non-pregnant animals from the pregnancy detection data.

2.4 Farrowing Data Processing

427 different animals nearing their estimated date of farrowing had readings taken with the device, with 999 readings taken between them. The retrieved data was then pruned down to animals that were 1 to 5 days prior to their actual date of farrowing for analysis and model creation, leaving 921 observations across 421 animals. Of the readings, 117 were taken five days prior to the animal farrowing, 169 were four days, 236 were three days, 248 were two days, and 151 were one day from farrowing.

Figure 2.12 shows the monotonically increasing impedance over the days prior to farrowing. Notably, the lower frequency data had a consistently higher impedance, with example low frequencies of 1000, 1500, 2000, and 10000 Hz all being very similar, but the higher frequencies of 20000 and 29000 Hz were significantly different from both the low frequencies and each other. Additionally, trends of interest were found in the phase angle, with variations indicating changes in capacitance,with the high frequencies having negative phases in a 'w' shape, the lowest frequencies rising and positive except for a dip on the third day prior to farrowing, and similar, more muted patterns in the significantly different frequencies of 5000 and 10000.

The model itself was based on a notion of imminent farrowing, that is, animals that are predicted to be within 2 days of farrowing by the model are said to be imminently farrowing, and should thus be moved to the farrowing crates. As for the frequencies that became relevant in the coefficients created for the model, lower frequencies made the bulk of the pack, though the most significant coefficients included not only the magnitudes of 2700 and 3600, but also the more intermediate phase of 9000, and the high frequency magnitudes of 19000 and 20000. As expected, the time of gestation was by far the strongest and most predictive variable. The total AUC of the model was 0.839.

One question of particular interest was the question of whether having a detection based on two days prior to farrowing, or three days prior to farrowing, would produce better results. As such, a model testing with a cutoff three days prior to farrowing was also created, and its coefficients (Table 2.4), violin diagram (Figure 2.15a), and ROC curve (Figure 2.15b) were retrieved. The three-day model appears to have higher TPR and FPR than the two-day model, making it a more sensitive one, and its AUC was only slightly lower at 0.834.

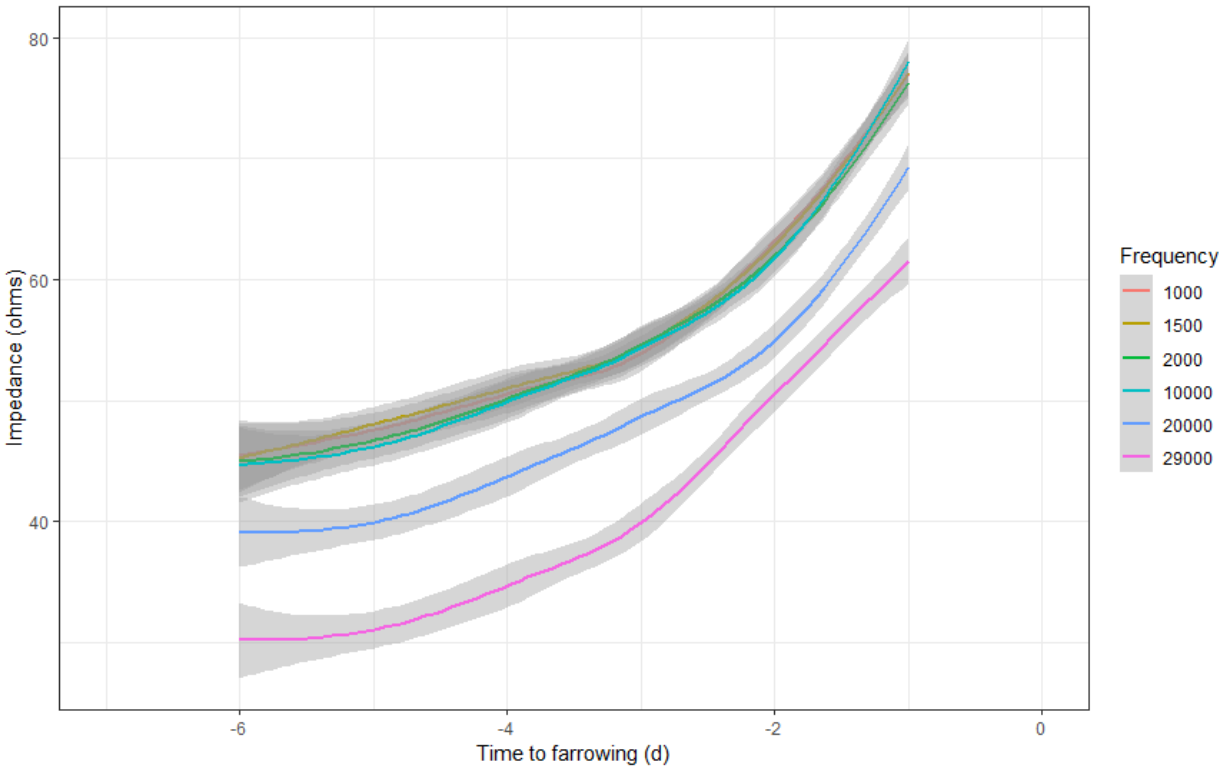


Figure 2.12: Plot of impedance values read by day.

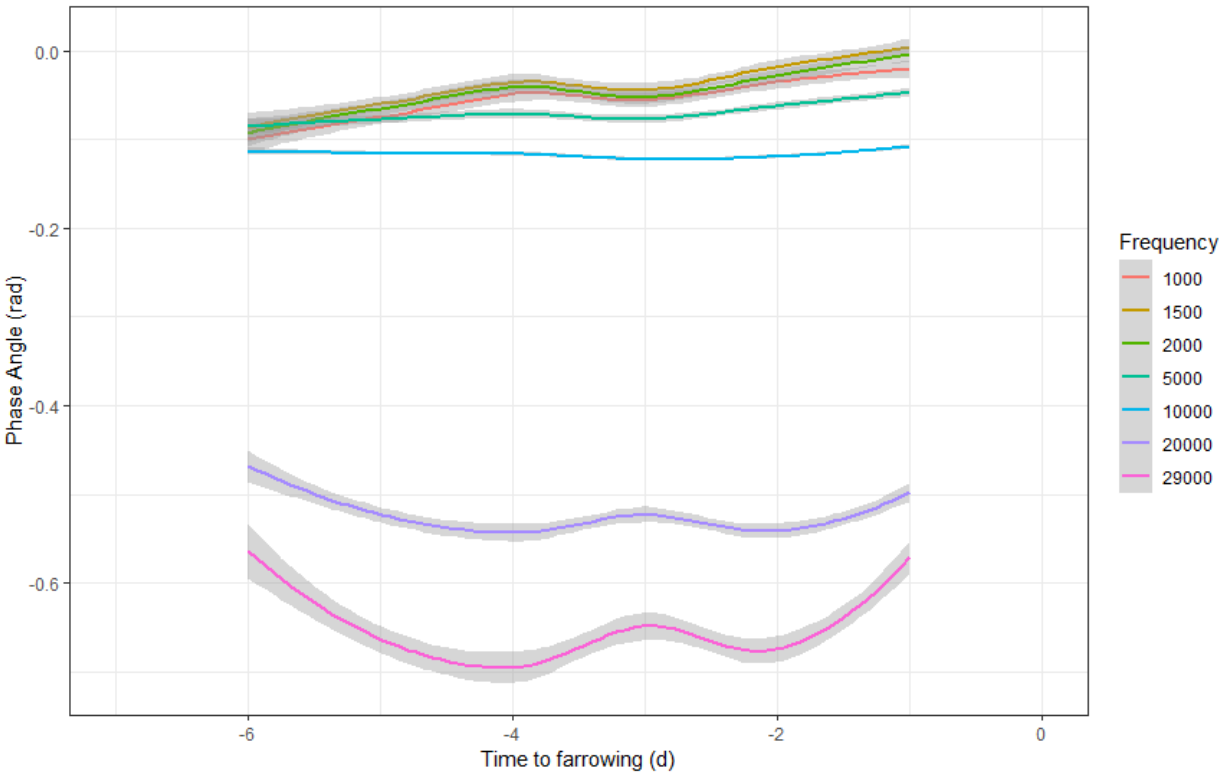
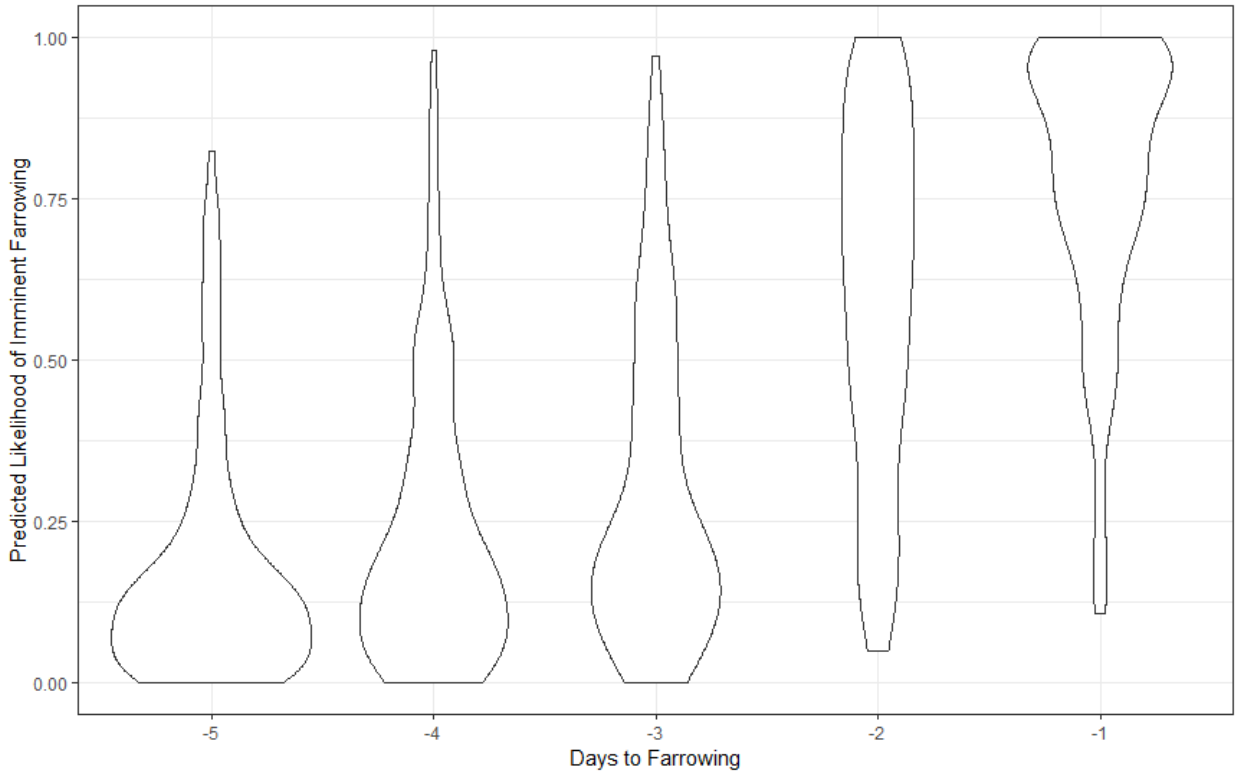


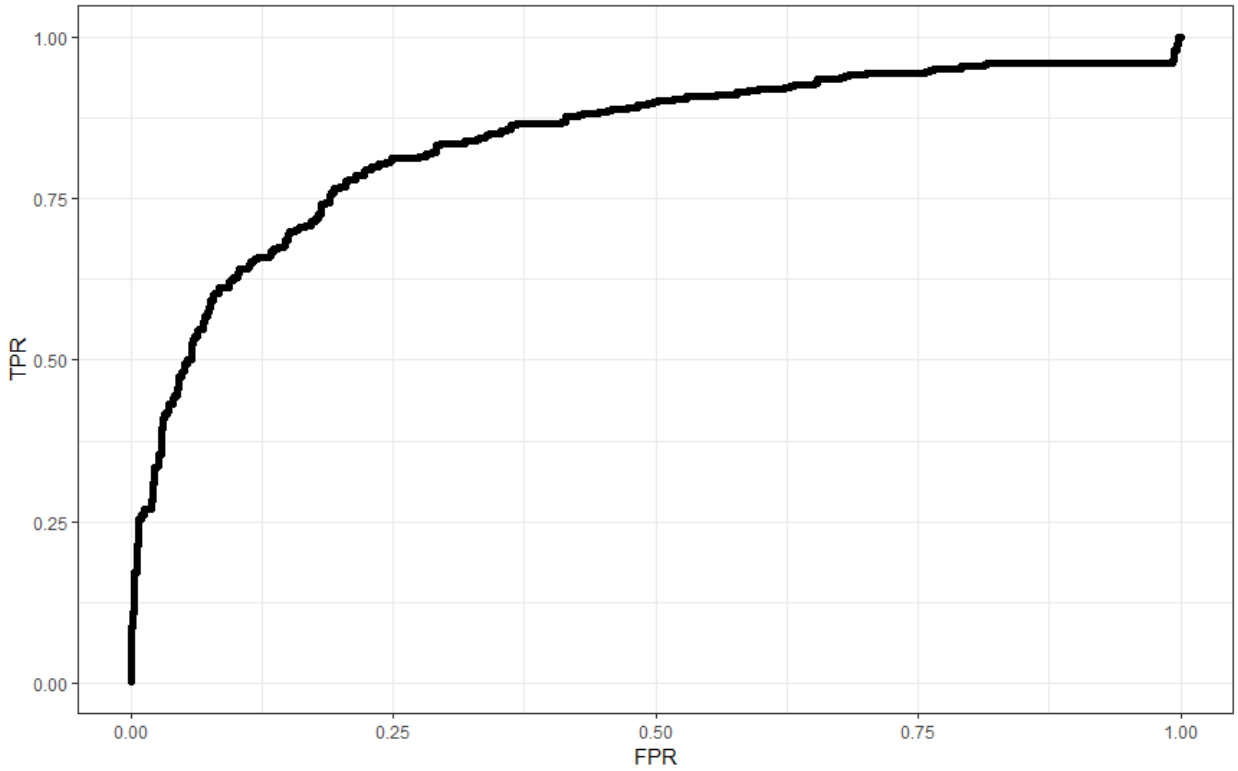
Figure 2.13: Plot of phase values read by day.

Table 2.3: Table of coefficients for the -2 days farrowing logistic model.

Frequency (Hz)	Estimate	Standard Error	z-value	Pr(=0)
(Intercept)	−166	12·9	−12·9	4.32E-38
gestation	−1·45	0·113	−12·8	2.12E-37
1000	0·0375	0·0155	2·42	1.54E-02
1900	−0·0676	0·0378	−1·79	7.38E-02
2300	−0·230	0·114	−2·02	4.32E-02
2400	0·352	0·119	2·95	3.16E-03
2500	−0·0688	0·0418	−1·64	1.00E-01
2700	−0·561	0·129	−4·34	1.42E-05
2800	0·434	0·144	3·01	2.60E-03
2900	0·165	0·110	1·50	1.33E-01
3300	−0·285	0·0931	−3·06	2.17E-03
3600	0·351	0·0980	3·58	3.43E-04
3900	−0·157	0·0945	−1·66	9.65E-02
4600	−0·368	0·114	−3·24	1.21E-03
4700	0·327	0·113	2·89	3.91E-03
5000	0·0904	0·0375	2·41	1.59E-02
5300	−0·307	0·138	−2·22	2.65E-02
5400	0·342	0·127	2·69	7.19E-03
7000	−0·135	0·0679	−1·98	4.73E-02
8000	0·131	0·0540	2·43	1.50E-02
15000	−0·112	0·0539	−2·08	3.74E-02
19000	0·491	0·137	3·59	3.35E-04
20000	−0·435	0·117	−3·73	1.94E-04
Frequency (Phase)				
1000	−4·80	1·87	−2·57	1.01E-02
2100	29·3	11·5	2·55	1.07E-02
2200	−28·2	11·7	−2·41	1.59E-02
4000	−14·1	7·62	−1·85	6.38E-02
4300	32·7	11·6	2·83	4.67E-03
4600	−32·2	12·0	−2·69	7.22E-03
4900	20·4	9·59	2·13	3.30E-02
5000	9·80	5·17	1·89	5.83E-02
5300	−22·9	11·2	−2·04	4.18E-02
5900	24·7	9·58	2·58	1.00E-02
9000	−28·0	6·42	−4·37	1.27E-05
14000	15·6	7·40	2·10	3.55E-02
16000	−7·89	4·56	−1·73	8.35E-02



(a) Violin diagram, showing the probability density of estimates for true positives and negatives.

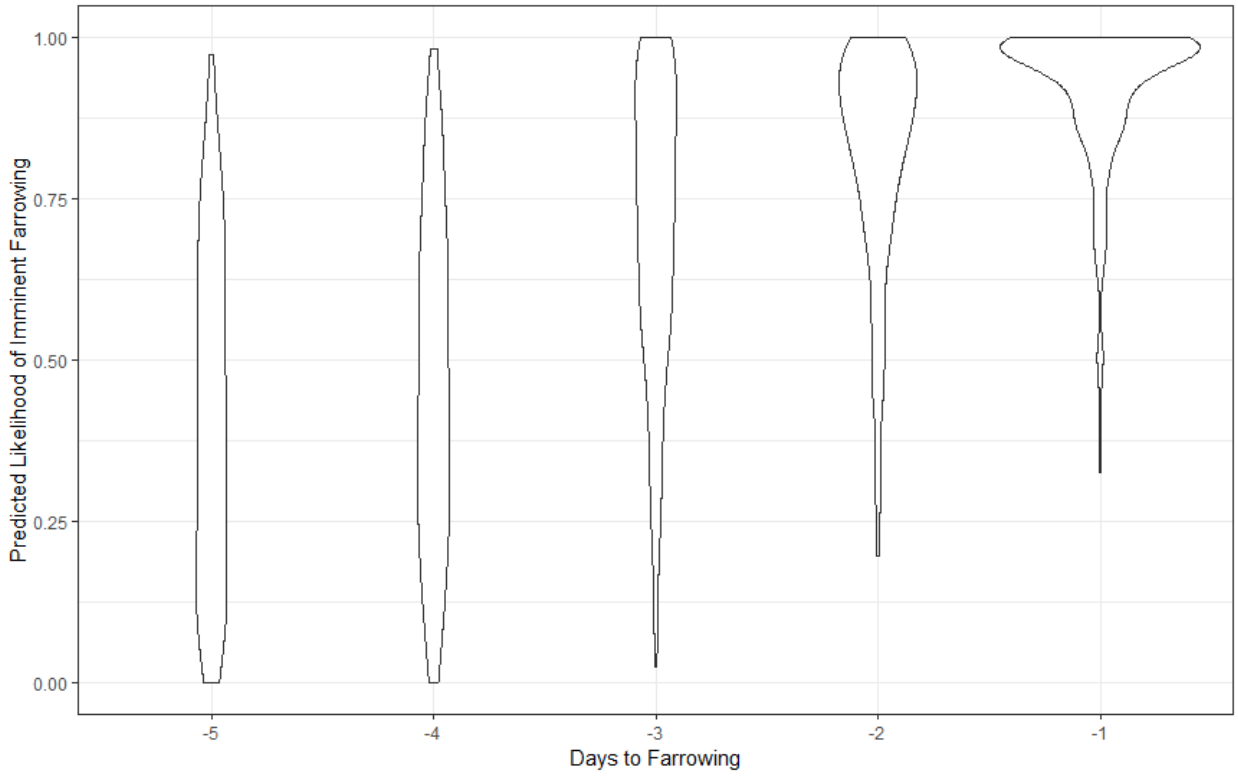


(b) ROC curve, visualising the diagnostic capacity of the model.

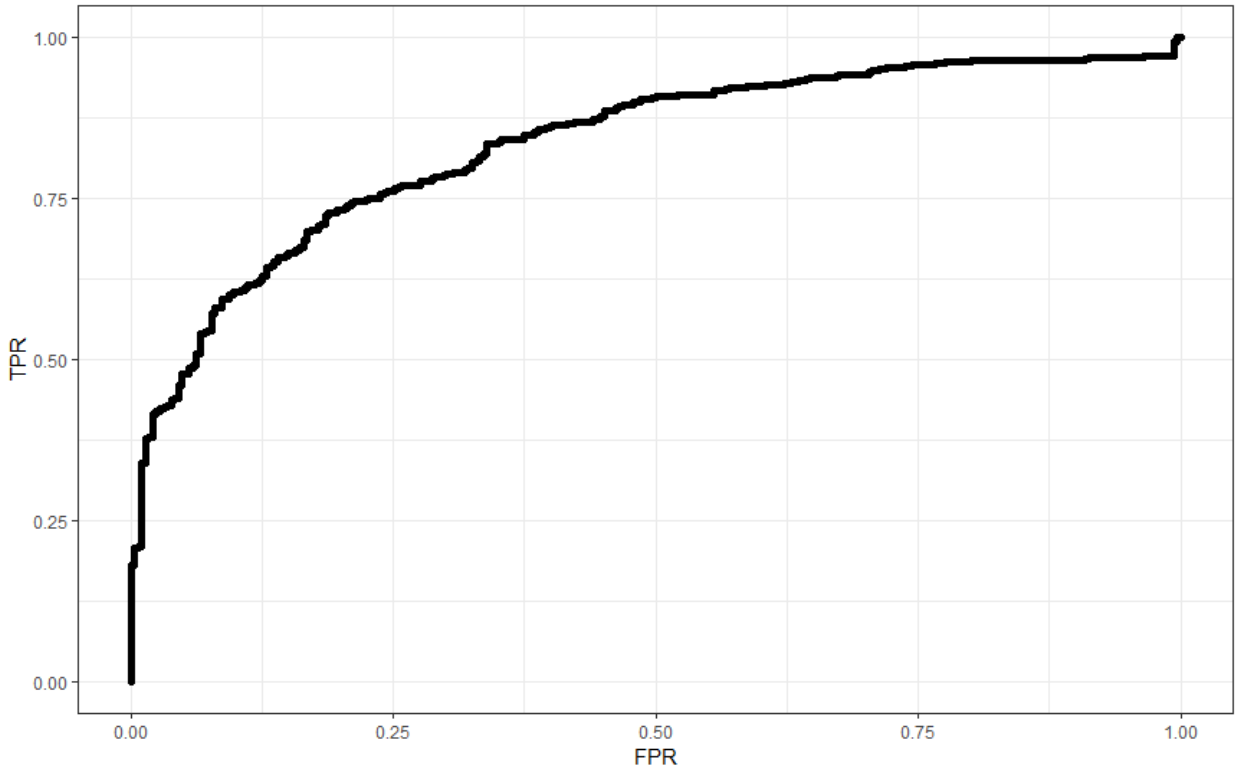
Figure 2.14: Violin diagram and ROC curve for Imminent Farrowing at -2 days, using 921 observations across 421 animals.

Table 2.4: Table of coefficients for the -3 days farrowing logistic model.

Frequency (Hz)	Estimate	Standard Error	z-value	Pr(=0)
(Intercept)	−136	12·8	−10·6	2.44E-26
gestation	−1·20	0·113	−10·6	2.01E-26
1000	0·0318	0·0214	1·48	1.38E-01
1300	−0·0643	0·0368	−1·75	8.03E-02
1600	0·0896	0·0513	1·74	8.10E-02
2000	0·0859	0·0931	0·923	3.56E-01
2100	−0·372	0·146	−2·55	1.07E-02
2200	0·314	0·150	2·08	3.71E-02
2300	−0·361	0·178	−2·02	4.33E-02
2400	0·287	0·127	2·25	2.43E-02
2500	−0·117	0·0512	−2·28	2.27E-02
3700	0·441	0·169	2·62	8.92E-03
3800	−0·373	0·178	−2·09	3.63E-02
4400	−0·141	0·0963	−1·46	1.44E-01
4700	0·285	0·113	2·52	1.17E-02
5000	−0·142	0·0633	−2·24	2.53E-02
5300	−0·367	0·197	−1·87	6.21E-02
5400	0·520	0·228	2·28	2.25E-02
5500	0·250	0·118	2·12	3.39E-02
5600	−0·453	0·239	−1·90	5.77E-02
5700	0·447	0·188	2·39	1.71E-02
5800	−0·412	0·159	−2·59	9.70E-03
8000	0·133	0·0575	2·31	2.08E-02
9000	0·0952	0·0492	1·93	5.31E-02
10000	−0·150	0·0839	−1·79	7.40E-02
12000	−0·419	0·140	−3·00	2.73E-03
14000	0·293	0·127	2·30	2.13E-02
17000	0·407	0·152	2·67	7.63E-03
18000	−0·464	0·216	−2·15	3.14E-02
19000	0·293	0·184	1·59	1.11E-01
21000	−0·166	0·102	−1·63	1.04E-01
Frequency (Phase)				
1000	−9·98	6·86	−1·46	1.46E-01
1100	11·2	7·62	1·47	1.43E-01
1800	−25·8	9·53	−2·71	6.76E-03
1900	32·09	10·2	3·14	1.69E-03



(a) Violin diagram, showing the probability density of estimates for true positives and negatives.



(b) ROC curve, visualising the diagnostic capacity of the model.

Figure 2.15: Violin diagram and ROC curve for Imminent Farrowing at -3days, using 921 observations across 421 animals.

Chapter 3

Discussion and Conclusions

3.1 Fertility Detection

The data retrieved about the changes in impedance over the course of the days before estrus did not appear to produce the significant expected dip in impedance as shown by sources such as Řezáč [9, 18] and Yoo [1]. Although a small decrease around day -4 does appear, it is so small that error precludes it from being relevant. However, the small number of animals tested prior to day -4 resulted in very high error margins, as is visible in Figure 2.3. Because of this, it cannot be said with any certitude whether or not a dip occurred prior to day -4. What was certainly observed, however, was the rapid rise in impedance as estrus nears. As figure 2.3 shows, after day -3, the impedance rises dramatically. This finding was expected from, and corroborates, earlier impedance measurements [9, 7].

The model displayed a strong AUC of 0.946, as well as highly distinct separation in the violin diagram of Figure 2.5a. One problem with this, however, is the high number of coefficients used in the model, at 48. This is higher than any of the other models, which leads to there being a potential problem of overfitting, particularly considering that this model had the least data, only 250 observations across 85 animals. However, model does itself did show a separation in animals that had failed to become pregnant, giving promise for the prediction of whether or not non-pregnant animals are reentering estrus, as per Figure 2.11. That this model produced this separation is an encouraging sign that it may not be overfitted, however, the lack of labels makes it impossible to draw certain conclusions. This potential predictive capability could substantially improve the efficiency of handling reproductively failing animals.

3.2 Pregnancy Detection

One obstacle to the pregnancy detection model is the relatively low proportion of inseminated animals that truly failed to become pregnant, which makes gathering enough observations to put this to a rigorous test will be challenging and require a substantial amount of time. The predictive power of the baseline model

from the entire 16-21 day dataset was worthwhile, but obviously flawed. Subsequent models built from data gathered from the later days of 19-21 only were significantly better than those encompassing all or most of the days. This is considered to be because of the changes that occur in non-pregnant animals when they are preparing to reenter estrus, as they would in the latter days. All points to this aspect of the device having substantial potential, but further data is needed, including a more substantial population of truly non-pregnant animals.

The groundwork for detecting estrus in animals determined to not be pregnant was laid, as can be seen in Figure 2.11. This would be a very useful capability of the device, as being able to eliminate animals that are failing to go into estrus early would be a substantial improvement to efficiency. However, without labels to apply to these animals to get a sense of what the TPR and FPR look like, this important matter is best trusted to a future data gathering and work.

3.3 Farrowing Prediction

The continually-rising impedance results from pre-farrowing animals over time, in Figure 2.12, was completely different from that of the previous data presented by Yoo, which indicated a rise until three days prior to farrowing, followed by a fall [1]. The lack of data from the date of farrowing and later does not explain this phenomenon, as the decrease is present in the Yoo data set starting between two and one day prior. The most likely reason for this is the massive difference in sample size. The data set analyzed here consisted of 427 different animals, while the previous work examined only 10.

The question of whether two days prior or three days prior to farrowing made a better model ended inconclusively. The violin diagrams of Figures 2.14a and 2.15a imply that, in comparison to a two-day model, a three-day model has a stronger tendency towards correctly identifying animals truly going into farrowing, as well as misidentifying those that were not as doing so. As such, the resulting ROC curves of Figures 2.14b and 2.15b are not significantly different, and either is useful for the task of predicting farrowing, albeit in different capacities.

These different capacities are outlined by the consequences of false negatives and false positives, which differ drastically in the case of farrowing, far more than the other reproductive state predictions examined. This is because a false negative on the final day could result in the loss of an entire litter, as the sow that farrows outside of the farrowing crate may end up crushing the piglets. False positives are also expensive, however, as farrowing crates are limited and more expensive to use and maintain. Those who wish to begin testing their animals on earlier dates, or are willing to do so every day after their date of choice, may be

better served by the two-day model, as the chances of a false positive on any day is slim due to its higher selectivity, while a false negative is itself very unlikely to occur two days in a row, though this depends upon one's threshold. The preferences of the barn manager are the ultimate arbiter, and the models shown, or potential future ones, can accommodate their individual preferences.

3.4 Future Work

The conjectural hypothesis about the nature of the change in impedance being due to changes in biochemical and ionic composition in order to control the population and metabolism of LAB, and thus acidify and deacidify the mucus, to cleanse and then accommodate sperm, requires testing. A future experiment should be designed in which mucus is collected by day, and then the ionic and biochemical concentrations measured, along with the presence of bacterial flora.

With the use of the estrus detection model on the non-pregnant animals of the pregnancy data, as shown in Figure 2.11, the possibility of detecting estrus in animals that failed to become pregnant has been laid. Future data on inseminated animals should be obtained, with animals which fail to become pregnant being observed for estrus in the future. This would grant a labelled data set for the production of a model of whether animals predicted to not be pregnant are reentering estrus or not. Such a model could be useful for those in the industry to make a quick and informed decision as to whether or not a particular animal is failing reproductively.

Further development of the farrowing model is recommended. Reframing the model from an instance of 'Yes' and 'No' to a multi-tier system, such as a 'High', 'Medium', and 'Low' likelihood of farrowing within the next given period may be of greater use due to the particularly high costs in this problem, both in the extended stays of false positives and the lost litters of those that run false negatives. For example, in such a three-answer model, those with the 'High' likelihood are to be moved to the farrowing crate immediately, those with the 'Low' likelihood are to be left and it may even be acceptable to extend the time between tests. 'Medium' results would then be up to the particular judgment of the user, potentially moving early, or simply keeping a closer eye on that animal and cutting time in between tests on that particular animal.

Additionally, it is suggested that models be tested that incorporate further information about the individual animals, such as their parity, average litter size, and even their weight. This additional information is likely to be readily available to barn managers, and may improve the predictive power of each of the constructed models. Further, the phenomenon of impedance changes is common throughout domestic mammals, and as such this technology could be expanded to other livestock such as cattle, sheep, and goats.

References

[1] S. Yoo et al., "A low-cost, portable, web-based impedance spectroscope for agricultural applications.," *M.S. Thesis.*, 2017.

[2] R. N. Kirkwood, "Pharmacological intervention in swine reproduction," *Swine health and production*, vol. 7, pp. 29–36, 1999.

[3] L. L. Anderson, "Reproductive biology of pigs," *Animal Industry Report*, vol. 655, no. 1, p. 66, 2009.

[4] G. Lamb and N. Dilorenzo, *Current and future reproductive technologies and world food production*. NY: Springer Science and Business Media, 2014.

[5] A. Lasia, "Electrochemical impedance spectroscopy and its applications," in *Modern aspects of electrochemistry*, pp. 143–248, Springer, 2002.

[6] B.-Y. Chang and S.-M. Park, "Electrochemical impedance spectroscopy," *Annual Review of Analytical Chemistry*, vol. 3, pp. 207–229, 2010.

[7] L. Dusza et al., "The relationship between electrical resistance of vaginal mucus and plasma hormonal parameters during periestrus in sows," *Theriogenology*, vol. 45, no. 8, pp. 1491–1503, 1996.

[8] L. Adam et al., "Impedometric properties of the vulvar and vaginal tissues of ewes during the oestrous cycle," *Reproduction*, vol. 61, no. 1, pp. 11–17, 1981.

[9] P. Řezáč et al., "Effect of sow parity on vaginal electrical impedance," *Animal reproduction science*, vol. 72, no. 3-4, pp. 223–234, 2002.

[10] J. T. Irvine et al., "Electroceramics: characterization by impedance spectroscopy," *Advanced Materials*, vol. 2, no. 3, pp. 132–138, 1990.

[11] F. Fabregat-Santiago et al., "Influence of electrolyte in transport and recombination in dye-sensitized solar cells studied by impedance spectroscopy," *Solar Energy Materials and Solar Cells*, vol. 87, no. 1-4, pp. 117–131, 2005.

[12] P. Taberna et al., "Electrochemical characteristics and impedance spectroscopy studies of carbon-carbon supercapacitors," *Journal of The Electrochemical Society*, vol. 150, no. 3, pp. A292–A300, 2003.

[13] T. Springer et al., "Characterization of polymer electrolyte fuel cells using ac impedance spectroscopy," *Journal of the Electrochemical Society*, vol. 143, no. 2, pp. 587–599, 1996.

[14] G. Lewis et al., "Changes in electrical resistance of vulvar tissue in holstein cows during ovarian cycles and after treatment with prostaglandin f2α," *Animal Reproduction Science*, vol. 18, no. 1-3, pp. 183–197, 1989.

[15] R. Marshall et al., "Electrical conductivity probes for detection of estrus in cattle," *Transactions of the ASAE*, vol. 22, no. 5, pp. 1145–1151, 1979.

[16] P. Carter and J. Dufty, "Assessment of vaginal impedance measurements as an indicator of oestrus in cattle," *Australian veterinary journal*, vol. 56, no. 7, pp. 321–323, 1980.

[17] P. Bartlewski et al., "The relationship between vaginal mucous impedance and serum concentrations of estradiol and progesterone throughout the sheep estrous cycle," *Theriogenology*, vol. 51, no. 4, pp. 813–827, 1999.

[18] P. Řezáč, "Potential applications of electrical impedance techniques in female mammalian reproduction," *Theriogenology*, vol. 70, no. 1, pp. 1–14, 2008.

[19] USDA, "Livestock and poultry: World markets and trade," *United States Department of Agriculture Foreign Agricultural Service*, 2019.

[20] M. F. Zink and J. R. Diehl, "Efficacy of using vaginal conductivity as an indicator of the optimum time to breed in swine," *Journal of Animal Science*, vol. 59, no. 4, pp. 869–874, 1984.

[21] H. P. Schwan and C. D. Ferris, "Four-electrode null techniques for impedance measurement with high resolution," *Review of scientific instruments*, vol. 39, no. 4, pp. 481–485, 1968.

[22] S. Hugentobler et al., "Ion concentrations in oviduct and uterine fluid and blood serum during the estrous cycle in the bovine," *Theriogenology*, vol. 68, no. 4, pp. 538–548, 2007.

[23] D. Zaaijer et al., "Changes in the composition of cervical mucus of the cow during the estrous cycle as parameters for predicting potential fertility," *Theriogenology*, vol. 39, no. 3, pp. 569–580, 1993.

[24] S. Boris and C. Barbés, "Role played by lactobacilli in controlling the population of vaginal pathogens," *Microbes and infection*, vol. 2, no. 5, pp. 543–546, 2000.

[25] G. Wagner and R. Levin, "Electrolytes in vaginal fluid during the menstrual cycle of coitally active and inactive women," *Reproduction*, vol. 60, no. 1, pp. 17–27, 1980.

[26] M. Bara et al., "A study of the microbial flora of the anterior vagina of normal sows during different stages of the reproductive cycle," *Australian veterinary journal*, vol. 70, no. 7, pp. 256–259, 1993.

[27] E. A. Miller, D. E. Beasley, R. R. Dunn, and E. A. Archie, "Lactobacilli dominance and vaginal ph: why is the human vaginal microbiome unique?," *Frontiers in microbiology*, vol. 7, p. 1936, 2016.

[28] J. Zhou et al., "The semen ph affects sperm motility and capacitation," *PloS one*, vol. 10, no. 7, p. e0132974, 2015.

[29] V. F. Lvovich, *Impedance spectroscopy: applications to electrochemical and dielectric phenomena*. John Wiley & Sons, 2012.

[30] G. Instruments, "Basics of electrochemical impedance spectroscopy," *G. Instruments, Complex impedance in Corrosion*, pp. 1–30, 2007.

[31] H. P. Schwan, "Alternating current electrode polarization," *Biophysik*, vol. 3, no. 2, pp. 181–201, 1966.

[32] H. P. Schwan, "Electrode polarization impedance and measurements in biological materials,," *Annals of the New York Academy of Sciences*, vol. 148, no. 4, pp. 191–209, 1968.

[33] Z. Chang et al., "A comparison of two- and four-electrode techniques to characterize blood impedance for the frequency range of 100 hz to 100 mhz," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 3, pp. 1247–1249, 2008.

[34] R-Core-Team, "R: A language and environment for statistical computing." https://www.r-project.org/, 2016. Date last accessed 1-July-2019.

[35] D. Smith, "R: A language and environment for statistical computing." http://java.sys-con.com/node/2288420, 2012. Date last accessed 1-July-2019.

[36] RStudio, "Rstudio open source and enterprise-ready professional software for r." https://www.rstudio.com/, 2018. Date last accessed 1-July-2019.

[37] RDocumentation, "Fitting generalized linear models." https://www.rdocumentation.org/packages/stats/versions/3.6.0/topics/glm. Date last accessed 30-June-2019.

[38] H. Akaike, "A new look at the statistical model identification.," *Selected Papers of Hirotugu Akaike*, vol. 55, no. 3, pp. 215–22, 1974.

[39] RDocumentation, "Choose a model by aic in a stepwise algorithm." https://www.rdocumentation.org/packages/MASS/ versions/7.3-51.4/topics/stepAIC. Date last accessed 1-July-2019.

[40] RDocumentation, "ggplot." https://www.rdocumentation.org/packages/ggplot2/versions/3.2.0/ topics/ggplot Date last accessed 22-July-2019.

[41] J. L. Hintze and R. D. Nelson, "Violin plots: a box plot-density trace synergism," *The American Statistician*, vol. 52, no. 2, pp. 181–184, 1998.

[42] B. J. McNeil and J. A. Hanley, "Statistical approaches to the analysis of receiver operating characteristic (roc) curves," *Medical decision making*, vol. 4, no. 2, pp. 137–150, 1984.

[43] M. H. Zweig and G. Campbell, "Receiver-operating characteristic (roc) plots: a fundamental evaluation tool in clinical medicine.," *Clinical chemistry*, vol. 39, no. 4, pp. 561–577, 1993.

[44] L. Gudmundsson et al., "Predicting above normal wildfire activity in southern europe as a function of meteorological drought," *Environmental Research Letters*, vol. 9, no. 8, p. 084008, 2014.

[45] D. Peres and A. Cancelliere, "Derivation and evaluation of landslide-triggering thresholds by a monte carlo approach," *Hydrology and Earth System Sciences*, vol. 18, no. 12, pp. 4913–4931, 2014.

[46] V. V. Kharin and F. W. Zwiers, "On the roc score of probability forecasts," *Journal of Climate*, vol. 16, no. 24, pp. 4145–4150, 2003.

[47] T. Fawcett, "An introduction to roc analysis," *Pattern recognition letters*, vol. 27, no. 8, pp. 861–874, 2006.