Respiratory Disease Diagnosis for Dolphin using Breath Data

Chinh Lai [*]   Lei Jin [*]   Andreas Fahlman [†]

Abstract

Respiratory disease is common in cetaceans both in the wild and under human care. Diagnosing lung disease is complicated, and recent development of spirometry in dolphins may provide an alternative minimally invasive, cheap and logistically feasible method to assess lung disease. Data from dolphins under managed care are used to measure baseline respiratory lung function under stress-free conditions. Because of new features in the data, new statistical methods are required for the breath data analysis. In this paper, we investigate one potential method for analyzing breath data. We consider an entire breath cycle to be one unit of observation. Starting and ending points of breath cycles can be difficult to determine, and cause a large amount of variation in size and shape of breath curves. To reduce cycle to cycle variability, we apply curve registration to synchronize a set of breath cycles. Breath cycles are described using magnitude information and geometric shape information. We propose three shape models, namely, simple oval model, quadratic spline model, and piecewise linear model. Furthermore, principal component analysis is applied to the magnitude/shape descriptors to obtain main features of breath cycles. Criteria for disease diagnosis are developed by identifying key differences among these main features between healthy and unhealthy animals. The proposed methods were applied to check whether two testing animals are diseased or not. The results were consistent with the status of both animals.

Key Words: Respiratory disease, disease diagnosis, breath data, functional data, functional data analysis, dolphin

1. Introduction

Respiratory disease in marine mammals evokes strong public attention as well as worthwhile scientific interest. The growing number of marine mammals suffering morbidity and mortality from respiratory disease implies a growing need for diagnosing the disease (Smith, Solano, Lutmerding, Johnson, Meegan, Le-Bert, Emory-Gomez, Cassle, Carlin, and Jensen, 2012; Sweeney, and Ridgway, 1976; Waltzek, Cortes-Hinojosa, Wellehan Jr, and Gray, 2012). Traditional methods for animal disease diagnosis include blood test, ultrasound, and computed tomography scan (Smith, Solano, Lutmerding, Johnson, Meegan, Le-Bert, Emory-Gomez, Cassle, Carlin, and Jensen, 2012). These methods require invasive equipment to perform, and cannot be applied to free-swimming animals. Compared to the traditional methods, the analysis of breath data is a non-invasive way for diagnosing respiratory disease in dolphin (Gans, 2013; Van Elk, Epping, and Gans, 2001). Respiration cycles of dolphin are measured by a combined flow-meter (pneumotachometer), allowing collection of breath data from trained dolphins participating voluntarily. Then, functional data analysis with computational statistics is introduced to provide tools to analyze breath data. For disease diagnosis of dolphin, differences between distributions of scores for healthy and diseased dolphins are revealed. The findings and results of the study will help researchers to diagnose health status of marine mammals, which can be used in studying health status of different populations. Thus, the diagnostic tools developed here may provide an important resource to assess environmental health.

The rest of this paper is organized as follows. Sections 2 introduces the methods to decompose time series into individual breath cycles, and to analyze the breath data. We

---

[*]Department of Mathematics and Statistics, Texas A&M University-Corpus Christi
[†]Fundacin Oceanogrc de la Comunitat Valenciana

apply curve registration, where multiple adjustments are applied to reduce variability in the breath cycles. We propose three geometric models to describe magnitude/shape features of the breath cycles. Next, principal component analysis (PCA) is used to obtain main features from magnitude/shape descriptors. Then, distribution of scores from the breath data of healthy animals is constructed. Finally, criteria for disease diagnosis are defined to detect health condition of dolphin. Section 3 provides numerical results, which will be used to detect the health condition of two testing animals. Finally, concluding remarks are provided in Section 4.

2. Methods

2.1 Breath Data

Breath data were measured on both healthy and diseased dolphins. Our analysis will focus on two specific measurements, volume and flow-rate. Volume, measured in liters, is the amount of air in the lung at different times as a dolphin exhales and inhales. Flow-rate, measured in liter per second, is the change in volume per unit time. Because measurements are taken at the constant frequency rate, for convenience, we just use the temporal order to denote the time information.

In this paper, we define a breath cycle to begin with exhalation and end with inhalation. When the dolphin begins to exhale, the flow-rate of a breath cycle starts decreasing from zero. The flow-rate declines to a minimum value before increasing to a maximum value, and then returns to zero again. Consequently, the sign of flow-rate changes from negative to positive as the respiration cycle of the dolphin changes from exhalation to inhalation.

Original breath data contains a sequence of breath cycles. Figure 1 shows different views of a sequence of breath cycles. In particular, Figure 1a shows breath cycles with respect to flow-rate versus time, which form fluctuating curve moving in forwarding direction. The breath cycles are attached by flat segments. However, since respiration cycles occur when flow-rate is changing over time, these steady segments are not connected with the breath cycles. Eliminating the segments from the dataset helps to decompose the time series into individual breath cycles. Besides that, Figure 1b shows the flow-rate of breath cycles plotted against volume, which results in closed and asymmetrical curves. In addition, we define half-up cycles are where flow-rate is greater than or equal to zero, and half-down cycles are where flow-rate is less than zero.
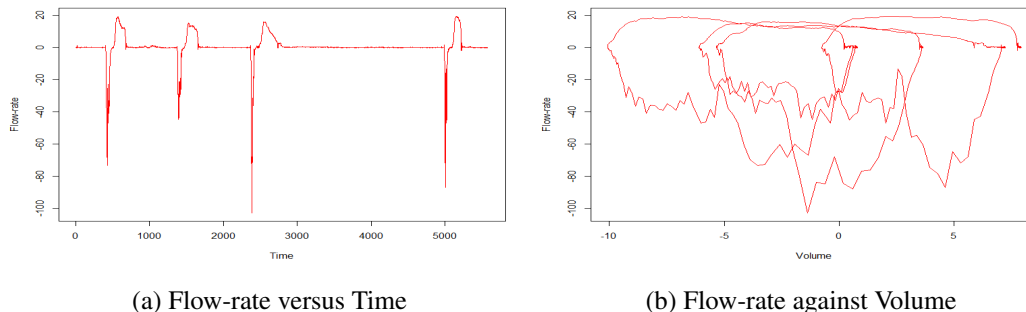

(a) Flow-rate versus Time   (b) Flow-rate against Volume

Figure 1: Breath Curves

2.2 Curve Registration

Consulting Figure 1b, we can that the raw breath curves contain substantial imperfections that would interfere with any synthesis of a common pattern. These imperfections include the differences in starting/ending points and sizes of the breath cycles, and the local intra-subject variability of the breath curves. The purpose of curve registration is to remedy these imperfections without losing any important information. To begin the registration process, we apply moving averages to moderate the intrasubject variability of breath cycles. Moving averages replace a given dataset with sets of averages of consecutive values in that dataset (Booth, Mount, and Viers, 2006). Since the half-up of the breath curve typically has less variation than the half-down cycle, two different equations are used to describe how moving averages are applied in each half cycle. To illustrate, let $\{X_1, X_2, \ldots X_n\}$ be a set of breath data measurements before moving averages. Then, $\{Y_1, Y_2, \ldots, Y_{n-1}\}$ and $\{Z_1, Z_2, \ldots, Z_{n-2}\}$ describe two new sets of breath data measurements after moving average, which are computed by

$$Y_{i-1} = \frac{X_{i-1} + X_i}{2}, 2 \le \text{i} \le \text{n}, \qquad Z_{i-2} = \frac{X_{i-2} + X_{i-1} + X_i}{3}, 3 \le \text{i} \le \text{n}$$

where Y_{i-1} is used to compute the new set of breath data for the top half, and Z_{i-2} is used to compute the new set of breath data for the down half. Figure 2 shows that after moving average, the adjusted curve has less local intrasubject variation than the original.
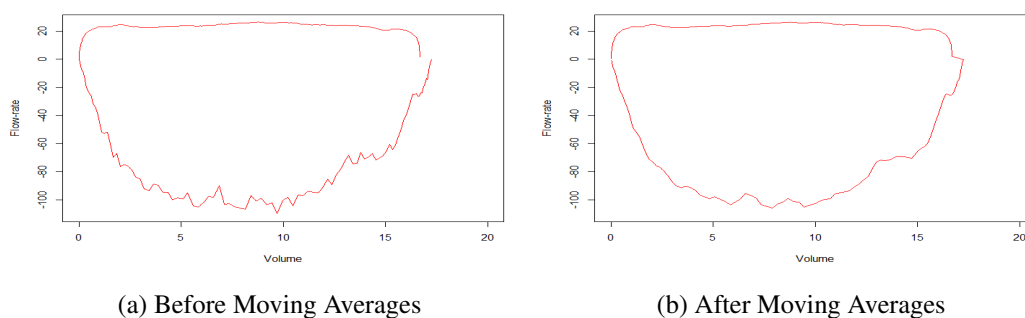

(a) Before Moving Averages   (b) After Moving Averages

Figure 2: A Breath Cycle before and after Moving Averages

Figure 1b shows that the magnitude of the breath cycles are different. Our next step is to resize inner area of the breath cycles. Each half of the breath cycle is rescaled separately. We use ratios r_1 and r_3 to rescale the half-up cycle to have maximum flow-rate at 100, and ending point at (200,0). We use ratios r_2 and r_4 to rescale the half-down cycle to have minimum flow-rate at -100, and ending point at (200,0). The equations for computing the ratios are described as

$$r_1 = \frac{100}{F_{mu}}, \qquad r_2 = -\frac{100}{F_{md}}, \qquad r_3 = \frac{200}{V_{mu}}, \qquad r_4 = \frac{200}{V_{md}}$$

where F_{mu} is maximum flow-rate of the half-up cycle, and F_{md} is minimum flow-rate of the half-down cycle. The quantity volume is defined in a similar fashion. After being resized, maximum flow-rate is equal to 100, minimum flow-rate is equal to -100, and maximum volume is equal to 200. Figure 3 shows the breath cycles after curve registration.

2.3 Modeling Size and Shape Information

In this section, we describe three models that provide simplified descriptions of the breath cycles. The first, a simple oval model, uses only the magnitude information of the breath cycles. The remaining two models, a quadratic spline model and a piecewise linear model, augment magnitude information with mark registration of the half-up breath cycles.

2.3.1 Simple Oval Model

The first geometric model constructed in this study is a simple oval model. To represent the model, a vector of five components is used, which includes area a_e, and the ratios (r_1, r_2, r_3, and r_4). In detail, the ratios r_1, r_2, r_3, and r_4 are used to rescale the magnitude of the breath cycles stay. After being rescaling, the shape of the breath cycles look relatively similar to oval shape, which names the method. Besides that, area a_e, represents inner area bounded by the curve of flow-rate of the breath cycles plotted against volume, is used to describe the magnitude of the entire breath cycles before moving averages and rescaling. The area is calculated by using the trapezoid sum, which is described as

$$area = \frac{US + LS}{2}$$

where the upper sum $US = \sum_{i=2}^{i=n} |F_i| \delta V$, the lower sum $LS = \sum_{i=1}^{i=n-1} |F_i| \delta V$, and $\delta V = V_{i+1} - V_i$.

2.3.2 Quadratic Spline Model

The second geometric model constructed in this study is a quadratic spline model. To represent the model, a vector of ten components is used. Firstly, to summarize the magnitude of the breath cycles, we use the inner area a_u of the half-up cycles before curve registration, and the ratios (r_1 and r_3). The fourth component of the vector is x_m, the volume after curve registration at while the maximum re-scaled flow rate of 100 occurs. Secondly, we use a_1, b_1, a_2, and b_2 to summarize the geometric shape of half-up cycles after curve registration. The name "quadratic spline", itself, represents that quadratic equation is used to construct the model. The system of equations below describes how quadratic equation is applied to construct the model.

$$\begin{cases} f_1(x) &= a_1 x^2 + b_1 x + c_1, \text{when x} \leq x_m \\ f_2(x) &= a_2 x^2 + b_2 x + c_2, \text{when x} \geq x_m \\ f_1'(x_m) &= f_2'(x_m) \end{cases}$$
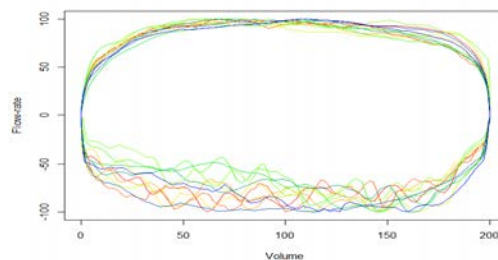

Figure 3: Breath Cycles after Curve Registration

where a_1 and b_1 are the coefficients of the first quadratic equation, which passes through the origin and maxima at $(x_m, 100)$ after curve registration; and a_2 and b_2 are the coefficients of the second quadratic equation, which passes through the maxima and ending point at $(200, 0)$ after curve registration. Additionally, at the maximum rescaled flow-rate, the first derivative of the two quadratic equations are equal to each other.

The last two components of the vector representing the quadratic spline model are R_1 and R_2, which summarize deviation from the quadratic spline model. The equations for computing R_1 and R_2 are described as

$$R_1 = y_2 - y_1, \qquad\qquad R_2 = y_4 - y_3$$

where x_1 is the midpoint of the interval $[0, x_m]$, y_2 is the positive flow-rate at x_1, and y_1 is the predicted flow-rate for the quadratic spline at x_1. The quantity R_2 is computed in a similar fashion on the interval $[x_m, 200]$. Figure 4a shows an example of the quadratic spline model.
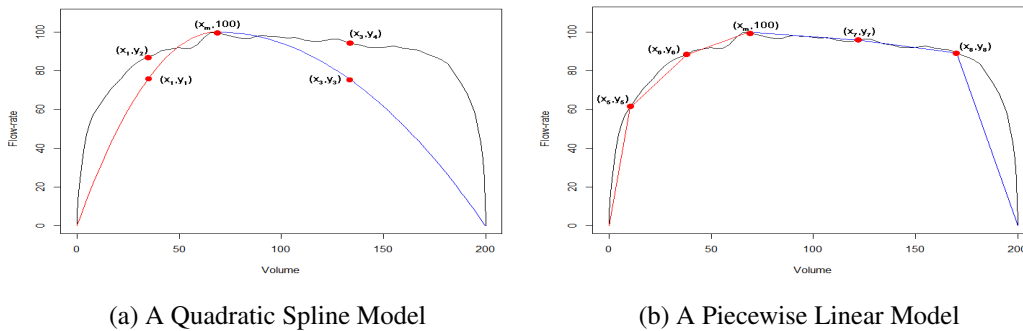

(a) A Quadratic Spline Model
(b) A Piecewise Linear Model

Figure 4: Geometric Models

2.3.3 Piecewise Linear Model

The third geometric model constructed in this study is a piecswise linear model. To represent the model, a vector of twelve components is used. There are similarities and differences between the quadratic spline model and the piecewise linear model. Regarding the similarities, three identical marks appear in both models, which are the origin, maximum, and ending point. Thus, they share four identical components including the area a_u, ratios r_1 and r_3, and variable x_m, which are used to summarize the magnitude of the half-up cycles in both models.

The remaining eight elements of the vector representing the piecewise linear model are x_5, y_5, x_6, y_6, x_7, y_7, x_8, and y_8, defined as follows. x_5 and x_6 are the trisectors of the interval $[0, x_m]$, and y_5 and y_6 are the positive rescaled flow-rates at those two points. x_7, y_7, x_8, and y_8 are defined similarly on the interval $[x_m, 200]$. Figure 4b shows an example of the piecewise linear model.

2.4 Principal Component Analysis

Let X_0 be a $d \times n$ data matrix, to begin dimension reduction process, X_0 is normalized, which is described as

$$X = \frac{X_0 - \bar{X}}{sd(X_0)}$$

where $\hat{X}$ is the mean of the column of the matrix X_0, and $sd(X_0)$ is the standard deviation of the column of the matrix X_0.

Then, correlation matrix $\Sigma_{XX} = X^T X$ is constructed to describe linear dependence between variables in the matrix X. By diagonalizing the correlation matrix Σ_{XX}, we obtain loading W as eigenvectors, and variances Λ of new variables as eigenvalues, which is described as

$$\Sigma_{XX} = W\Lambda W^T$$

where eigenvalues and eigenvectors are in the descending order of eigenvalues.

For each of the models described above, a model matrix X_0 is formed by using the vectors from all breath cycles as row vectors. The R function `prcomp()` is used to compute the singular value decomposition for these three model matrices.

By dropping some of the least important eigenvectors, PCA reduces the number of dimensions of the breath data. The number of dimensions can be determined by comparing the proportion of variation remaining in the models with the total variation of the dataset. Figure 5 shows three scree plots, which displays PCA results. A scree plot is a simple line segment plot, which shows the eigenvalues/variances of individual components in descending order (Cattell, 1966).
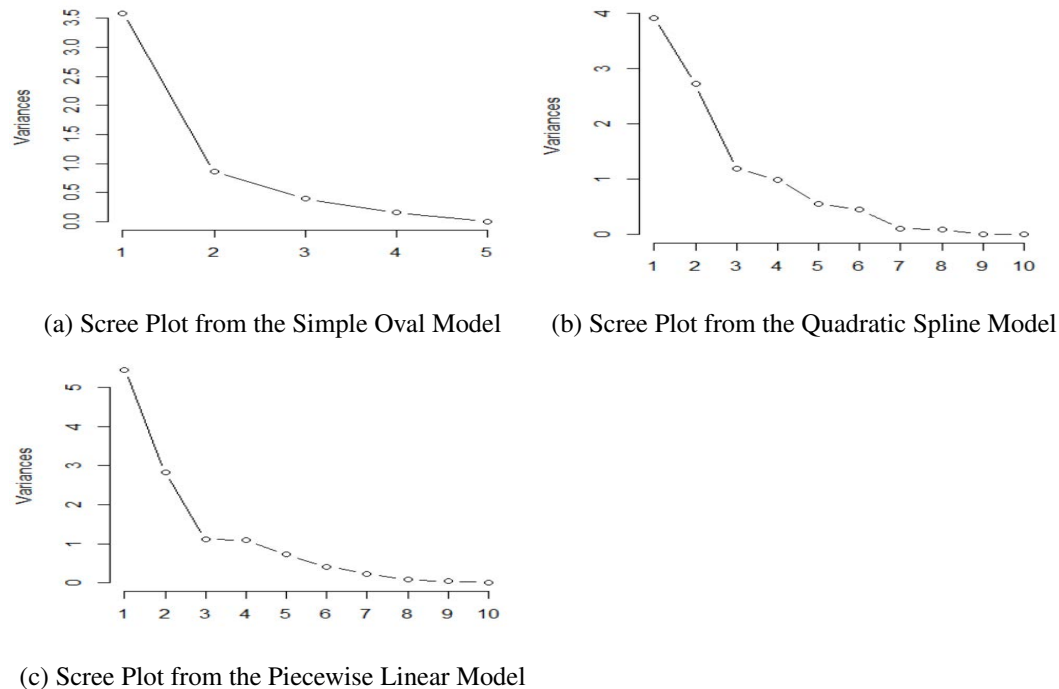

(a) Scree Plot from the Simple Oval Model   (b) Scree Plot from the Quadratic Spline Model

(c) Scree Plot from the Piecewise Linear Model

Figure 5: PCA Results from Different Models

According to the screen plots in Figure 5, the variances of PC1 and PC2 are significantly higher than the other PCs in individual models. Therefore, there are respectively two equations for computing $score_1$ and $score_2$ from each model, where Equation 1 and 2 are from the simple oval model, Equation 3 and 4 are from the quadratic spline model, and Equation 5 and 6 are from the piecewise linear model.

$$score_1 = 0.495a_u - 0.379r_1 - 0.400r_2 - 0.478r_3 - 0.472r_4 \qquad (1)$$

$$score_2 = -0.033a_u + 0.581r_1 + 0.493r_2 - 0.445r_3 - 0.469r_4 \qquad (2)$$

$$score_1 = 0.149a_u - 0.059r_1 - 0.157r_3 - 0.473x_m - 0.433a_1 + 0.461b_1 + 0.352a_2$$
$$- 0.342b_2 + 0.088R_1 - 0.277R_2 \tag{3}$$

$$score_2 = -0.538a_u + 0.429r_1 + 0.491r_3 - 0.126x_m + 0.024a_1 + 0.020b_1 + 0.296a_2$$
$$- 0.302b_2 - 0.012R_1 + 0.298R_2 \tag{4}$$

$$score_1 = 0.085a_u + 0.004r_1 - 0.098r_3 - 0.419x_m - 0.404x_5 - 0.233y_5 - 0.417x_6$$
$$+ 0.031y_6 - 0.423x_7 - 0.173y_7 - 0.413x_8 + 0.188y_8 \tag{5}$$

$$score_2 = -0.550a_u + 0.438r_1 + 0.510r_3 - 0.057x_m - 0.153x_5 - 0.333y_5 - 0.095x_6$$
$$- 0.085y_6 - 0.013x_7 + 0.142y_7 + 0.093x_8 - 0.252y_8 \tag{6}$$

2.5 Distributions of Scores for Healthy Animals

In this section, to determine scores characterizing the breath cycle of healthy animals, a set of 92 breath cycles taken from five healthy animals are combined into a training set referred to below as **animal H**. We construct bivarate boxplots to describe two-dimensional distributions of scores from the breath data of animal H (Rousseeuw, Ruts, and Tukey, 1999). From the bivariate boxplot, we can visualize the spread and outlier of the breath data. Figure 6 show two-dimensional distributions of scores from breath data of healthy animals from the three models. We also use the Shapiro-Wilk test and Anderson-Darling test to test for normality (Shapiro and Wilk, 1965; Anderson and Darling, 1952). Table 1 shows the normality test results.
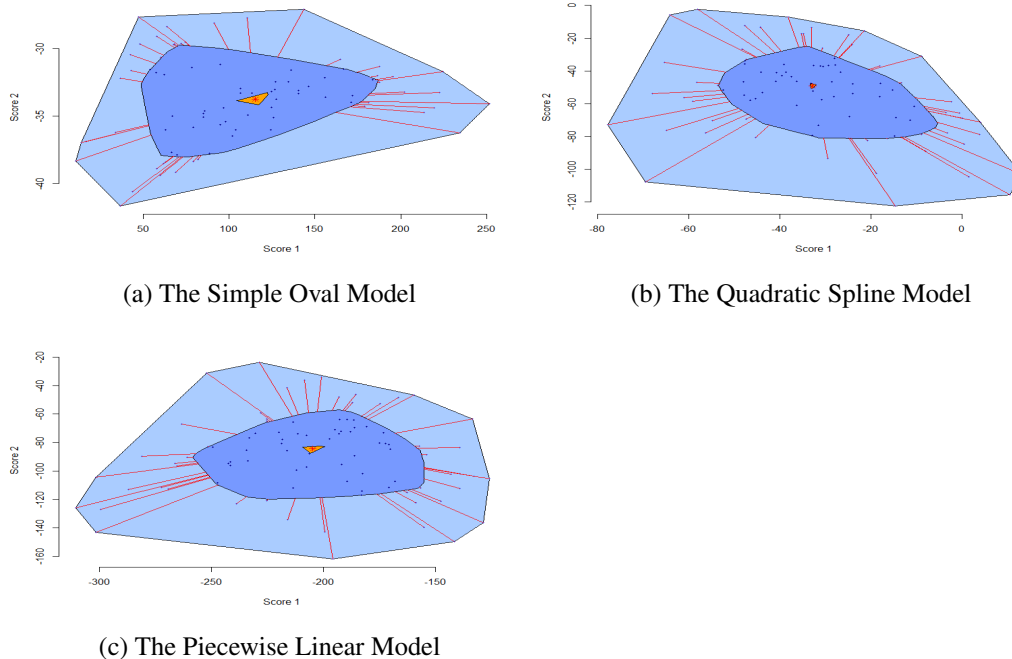

(a) The Simple Oval Model

(b) The Quadratic Spline Model

(c) The Piecewise Linear Model

Figure 6: Bivariate Boxplots of the first Two Scores from Different Models for Healthy Animals

According to Figure 6, neither plots reveal unusual qualities, such as outliers or gaps. The bivariate boxplots show that the distributions of scores are slightly skewed; however, their means are located approximately at the centers of the boxplots, and the spread of the

data is small as the majority lie relatively close to the means. In addition, according to the normality test results from Shapiro-Wilk and Anderson-Darling tests, all p-values are greater than the 1% significant level. These results/findings indicate that the distributions of scores from the breath data of healthy animals are approximately normal or not too far away from normal distribution.

Table 1: Normality Test Results from Different Models

Model	Simple oval		Quadratic spline		Piecewise linear	
Score	$score_1$	$score_2$	$score_1$	$score_2$	$score_1$	$score_2$
Shapiro-Wilk Test	0.082	0.616	0.836	0.286	0.100	0.970
Anderson-Darling Test	0.029	0.603	0.934	0.314	0.264	0.950

2.6 Criteria for Disease Diagnosis

In this section, based on the PCs and scores collected from PCA, we define criteria for disease diagnosis. Because the breath data of different animals have different numbers of breath cycles measured, we define criteria for disease diagnosis according to the number of breath cycles. When there are only one or a very few cycle available for a testing animal, the criterion is that the animal is diseased if the score/or the average score is out of the normal range shown in Figure 6 (or one-dimensional boxplot). The normal ranges of the healthy population are defined by 3 standard deviations from the mean criterion or 1.5 interquartile range criterion in a modified boxplot. Additionally, when a dataset from a testing animal contains many cycles, two-sample t-test and Hotelling's t-squared test are used to test the difference between healthy and testing animals (Snedecor and Cochran, 1989; Hotelling, 1931). While the two-sample t-test is used to compare the two population means of one score from healthy and testing animals, the Hotelling's t-squared test is used to test the mean vector of the two sets of score values from healthy and testing animals. The equation for computing diseased index for the two-sample t-test is defined as

$$\text{Diseased index T} = \frac{\overline{\mathbf{x}} - \overline{\mathbf{y}}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where $\overline{\mathbf{x}}$ is the average score of a testing animal, $\overline{\mathbf{y}}$ is the average score of healthy animals, n_1 is the number of breath cycles from a testing dataset, n_2 is the number of breath cycles from healthy dataset, s_1 is the standard deviation of scores of a testing animal, and s_2 is the standard deviation of scores of healthy animals.

The equation for computing diseased index for the Hotelling's t-squared test is defined as

$$\text{Diseased index } T^2 = (\overline{\mathbf{x}} - \overline{\mathbf{y}})^T (\hat{\Sigma}(\frac{1}{n_1} + \frac{1}{n_2}))^{-1}(\overline{\mathbf{x}} - \overline{\mathbf{y}})$$

where $\overline{\mathbf{x}} = \frac{1}{n_x} \sum_{i=1}^{n_x} \mathbf{x}_i$ is the average score vector of a testing animal, $\overline{\mathbf{y}} = \frac{1}{n_y} \sum_{i=1}^{n_y} \mathbf{y}_i$ is the average score vector of healthy animals, and $\hat{\Sigma} = \frac{n_x \hat{\Sigma}_{\mathbf{x}} + n_y \hat{\Sigma}_{\mathbf{y}}}{n_x + n_y - 2}$ is the covariance matric

where $\hat{\Sigma}_{\mathbf{x}} = \frac{1}{n_x-1}\sum_{i=1}^{n}(\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})'$ is the covariance matrix of a testing animal and $\hat{\Sigma}_{\mathbf{y}} = \frac{1}{n_y-1}\sum_{i=1}^{n}(\mathbf{y}_i - \overline{\mathbf{y}})(\mathbf{y}_i - \overline{\mathbf{y}})'$ is the covariance matrix of healthy animals. Under the null hypothesis, the statistics T^2 is related to the F-distribution with p and $n - p$ degrees of freedom, where $p = 2$ for comparing bivariate scores.

3. Disease Diagnosis for Animals

In this section, based on the scores defined in equations 1 to 6, we compute scores for testing animals. Then we test the claim that the mean for a testing animal is the same as, or different from, the mean for the healthy population. Results from the test will be used to detect the health condition of two testing animals which we will refer to as **animal T1** and **animal T2**. T1 has one breath cycle measured, while T2 has 59 breath cycles measured. Figure 7 shows breath cycles of T1 and T2 after curve registration. In addition, scatter plots are used to compare the two-dimensional measurements from healthy animals with testing animals, as shown in Figure 8.
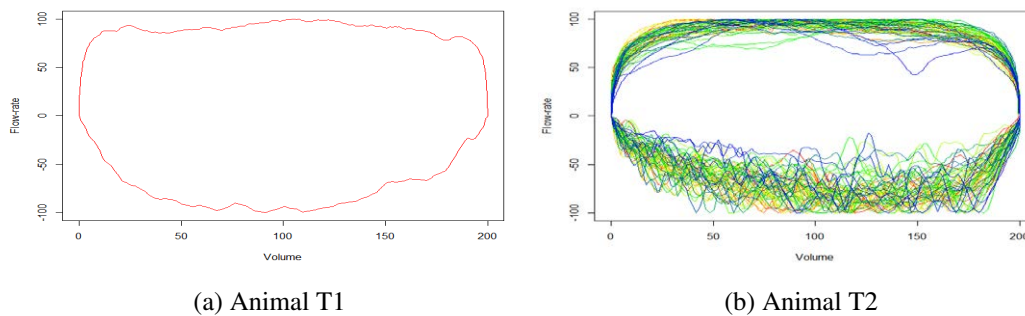

(a) Animal T1   (b) Animal T2

Figure 7: Breath Cycles of Testing Animals after Curve Registration

For the disease diagnosis for animal T1, as can be seen in Figures 8, the scores of animal T1 appear to be outliers in the plots. That indicates some differences between healthy animals and animal T1. For the disease diagnosis for animal T2, since it has 59 breath cycles measured, we use two-sample t-test and Hotelling's t-squared test to test the differences between healthy animals and testing animal T2. Two-sample t-test is used to compare the two population means of one score from healthy animals and animal T2. The null hypothesis H_0 states that the mean for animal T2 is the same as the mean for the healthy population. The alternative hypothesis H_1 states that the mean for animal T2 is different from the mean for the healthy population. Table 2 shows results from the two-sample t-test.

Table 2: Two-sample T-test Results from Different Models

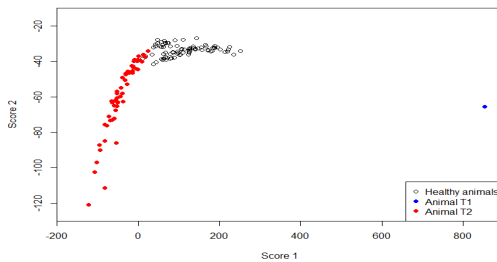| Model	Simple oval		Quadratic spline		Piecewise linear	
Score	$score_1$	$score_2$	$score_1$	$score_2$	$score_1$	$score_2$
P-value	< 2.2e-16	1.262e-13	2.449e-11	< 2.2e-16	0.004	< 2.2e-16

Hotelling's t-squared test is used to test the mean vector of the two sets of score values from healthy animals and animal T2. The null hypothesis H_0 states that the mean vector for testing animal is the same as the mean vector for the healthy population. The alternative

hypothesis H_1 states that the mean vector for animal T2 is different from the mean vector for the healthy population. Table 3 shows results from the Hotelling's t-squared test.

Table 3: Hotelling's T-squared Test Results from Different Models

| Model	Simple oval	Quadratic spline	Piecewise linear
P-value	< 2.2e-16	< 2.2e-16	< 2.2e-16

As can be seen from Tables 2 and 3, all the p-values are less than the 1% significant level, so we reject the null hypothesis H_0 (Snedecor and Cochran, 1989; Hotelling, 1931). In addition, according to Figure 8, the distributions of scores for animal T2 are separated from the distributions of scores for healthy animals. These findings/results indicate some differences between healthy animals and animal T2.

(a) The Simple Oval Model
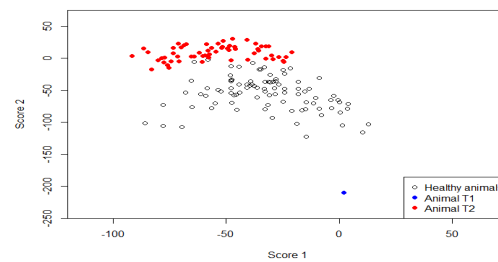
(b) The Quadratic Spline Model
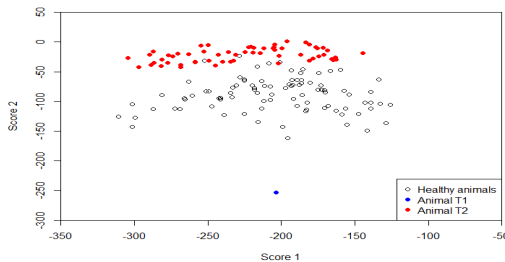
(c) The Piecewise Linear Model

Figure 8: Scatter Plots of the first Two Scores from Different Models for Healthy and Testing Animals

4. Conclusion and Future Work

We propose some statistical methods for disease diagnosis on dolphin by analyzing the breath data. In the analysis, a total of 92 breath cycles of five healthy dolphins have been analyzed to build the baseline of healthy animals. The ratios/area are used to describe the magnitude of the breath curves. Three models are constructed to describe the geometric shape of the breath curves. Then PCA is applied to extract important features from the magnitude/shape descriptors. Criteria for disease diagnosis are defined using boxplots and bivariate boxplots to display the distributions of scores for healthy and testing animals, the two-sample t-test to compare the two population means of one score from healthy and

testing animals, and the Hotelling's t-squared test to test the mean vector of the two sets of score values from healthy and testing animals. The proposed methods were applied to detect the health condition of two testing animals. The results/findings indicate that there are some differences between healthy and testing animals.

REFERENCES

Anderson, T. W., and Darling, D. A. (1952), "Asymptotic Theory of Certain Goodness-of-fit," *Annals Mathematical Statistics*, 23(2), 193–212.

Booth, E. G., Mount, J. F., and Viers, J. H. (2006), "Hydrologic Variability of the Cosumnes River Floodplain," *San Francisco Estuary and Watershed Science*, 4(2), 1–19.

Cattell, R. B. (1966), "The Scree Test for the Number of Factors," *Multivariate Behavioral Research*, 1(2), 245–276.

Gans, S. (2013), "Lung Function Measurements Using Spirometry in Small Cetaceans: Tursiops Truncatus and Phocoena Phocoena," University of Utrecht.

Hotelling, H. (1931), "The generalization of student's ratio," *Annals Mathematical Statistics*, 2(3), 360–378.

Rousseeuw, P. J., Ruts, I., and Tukey, J. W. (1999), "The Bagplot: A Bivariate Boxplot," *The American Statistician*, 53(4), 382–387.

Shapiro, S. S., and Wilk, M. B. (1965), "An Analysis of Variance Test for Normality," *Biometrika*, 52(3/4), 591–611.

Smith, C. R., Solano, M., Lutmerding, B. A., Johnson, S. P., Meegan, J. M., Le-Bert, C. R., Emory-Gomez, F., Cassle, S., Carlin, K., and Jensen, E. D. (2012), "Pulmonary Ultrasound Findings in a Bottlenose Dolphin Tursiops Truncatus Population," *Dis Aquat Organ*, 101(3), 243–255.

Snedecor, G. W., and Cochran, W. G. (1989), *Statistical Methods*, Ames, Iowa: Iowa State University Press.

Sweeney, J. C., and Ridgway, S. H. (1975), "Common Diseases of Small Cetaceans," *Journal of the American Veterinary Medical Association*, 167, 533–540.

Van Elk, C. E., Epping, N. and Gans, S. J. (2001), "Pulmonary Function Measurements in Dolphins Using Capnography," *Veterinary Records*, 149(10), 308-309.

Waltzek, T. B., Cortés-Hinojosa, G., Wellehan Jr, J. F., and Gray, G. C. (2012), "Marine Mammal Zoonoses: A Review of Disease Manifestations," *Zoonoses and Public Health*, 59(8), 521–535.