AIRFLOW ESTIMATION FROM RESPIRATORY SOUNDS

by

İlhan Yıldırım

B.S., Electrical & Electronics Engineering, Boğaziçi University, 2013

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Electrical & Electronics Engineering
Boğaziçi University
2018

AIRFLOW ESTIMATION FROM RESPIRATORY SOUNDS

APPROVED BY:

Prof. Yasemin Z. Kahya   . . . .
(Thesis Supervisor)

Assist. Prof. İpek Şen   . . . .
(Thesis Co-supervisor)

Prof. Murat Saraçlar   . . . .

Assoc. Prof. Rifat Koray Çiftçi   . . . .

DATE OF APPROVAL: 24.5.2018

ACKNOWLEDGEMENTS

Acknowledgements come here...

ABSTRACT

AIRFLOW ESTIMATION FROM RESPIRATORY SOUNDS

One page abstract will come here.

ÖZET

SOLUNUM SESLERİNDEN SOLUK AKIŞI KESTİRİMİ

Bir sayfa uzunluğunda özet gelecektir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

a_{ij}	Description of a_{ij}
A	State transition matrix of a hidden Markov model
α	Blending parameter *or* scale
$\beta_t(i)$	Backward variable
Θ	Parameter set

LIST OF ACRONYMS/ABBREVIATIONS

2D	Two Dimensional
3D	Three Dimensional
AAM	Active Appearance Model
ASM	Active Shape Model

1. INTRODUCTION

Listening to respiratory sounds to gain information about respiratory diseases is a method which has been applied for at least 2400 years. Since then, different auscultation techniques have been used. The most widely used tool developed for this purpose is called "stethoscope" and it is invented in 1816 by Rene Laënnec [2]. Medicine doctors still use this device to diagnose several diseases such as asthma, bronchitis etc.

In recent decades, engineers started to work on respiratory sounds and they produced electronic stethoscopes. These electronic stethoscopes have some advantages over the traditional ones. First of all, the frequency band is not limited [3] by the mechanic structure of stethoscope and the sounds can be recorded for reuse. Another advantage is that the respiratory sound recordings made analyzing the sound with the help of computers possible.

In recent years, together with developing classification techniques, there has been a great effort in automatic classification of respiratory sounds as healthy or sick [4]. Some of these methods are dependent not only on sounds but also on the airflow information [5]. Airflow is usually recorded by a tool which is called pneumotachograph. This tool is not useful in case of subject who have some disabilities.

Main motivation of this thesis is to estimate the airflow and respiratory phases from the respiratory sounds recorded at chest wall without using pneumotachograph to enable airflow measurements to be easier and a development of single handheld electronic stethoscope whose recording will be enough for automatic diagnose.

Before giving the preview of the work has been done over this thesis, we may give a brief review of the literature about the relation between flow, phase and respiratory sounds.

Lessard and Wong [6] reported that the relation between the spectral parameters

which are mean frequency, frequency of maximum power and the highest frequency at which the power in the spectrum is at least 10 percent of maximum power, and flow rate is not linear and spectral parameters saturates as the flow rate goes beyond 0.75 l/s.

Yadollahi and Moussavi [7] reported that using entropy of bandpass filtered tracheal sounds by using overlapping windows whose durations are 100 milliseconds. They achieved average error of 7.3 % and 7.4 % for inspiratory and expiratory phases after calibrating the model which uses the entropy information.

Huq and Moussavi [8] proposed using log of variance (LV) first to detect onsets. For phase identification, they used 4 parameters calculated over LV curve and the duration of phase. They developed and tested the method for tracheal sounds and it is reported that 95.6% accuracy was achieved for phase identification after onsets are verified by visual inspection.

Moussavi et al. [9] suggested using tracheal sounds for onset detection and the power difference in 150-450 Hz at the "best recording position", where the difference in power for inspiration and expiration is greatest, for phase identification. It is reported that the success rate in phase identification is 100% after the onsets are found.

Golabbakhsh et al. [10] suggested using the average power (P_{ave}) of tracheal sounds to estimate the respiratory flow. They have two approaches, first one is expressing the flow as a linear function of log of P_{ave} as in 1.1 and the other one is training an adaptive filter with three taps and whose input is P_{ave} as in 1.2.

$$\hat{F} = c_0 + c_1 \log(P_{ave}) \tag{1.1}$$

$$\hat{F} = \sum_{i=1}^{M} w_i P_{ave} \tag{1.2}$$

Çiftçi and Kahya [11] modeled the sounds recorded at both trachea and chest

as time varying autoregressive (TVAR) processes by using Fourier basis functions and used the first autoregressive (AR) coefficient vector as the estimate of absolute airflow curve. Reported correlations for sounds recorded at trachea and chest wall were 0.9 and about 0.6 (extracted from figure in paper) respectively.

Organization of thesis will be explained in following paragraphs.

In chapter 2, the experimental setup and data will be explained.

In chapter 3, we will describe AR and TVAR processes and the methods to find TVAR coefficients of a signal, the methods we will describe and use are windowing based AR modeling, TVAR modeling with basis functions and Kalman filter. We will also present Short Time Fourier Transform (STFT) and Wiener filter which will be used to unify different estimations. Finally experiments and results will be presented.

In chapter 4, the method for period estimation will be given first. Then neural networks approach and features will be explained. After neural networks, the method which estimates the transition points and then identifies the phases between estimated transition points will be described. Lastly, experiments and results will be documented.

In the conclusion chapter, the last words about this work will be said.

2. EXPERIMENTAL SETUP AND DATA

The data used throuhout this thesis is taken from the database belonging Boğaziçi University Lung Acoustics Laboratory. All sound and airflow samples were recorded with the device developed in this laboratory. (İpek Sen Masters) The recording device has 14 channels. Each channel has the same analog interface, which includes a Sony ECM-44 BPT electret microphone followed by an analog amplifier with gain 100 which has a instrumentation amplifier based design. The amplification unit is followed by a 6th order Butterworth low pass filter which have 4000 Hz cut-off and a 8th order Bessel high pass filter whose cutoff is 80 Hz. The airflow recordings was done with a Validyne CD379 pneumatachograph. The digitization unit is the National Instrument's DAQCard-6024E which is operating at 9600 Hz and has 12 bit resolution. Each recording has a duration of 15 seconds [12].

. There are 23 healthy subjects in this study. The sounds were recorded at 14 different locations on chest wall. Locations of channels are shown in 2.1.
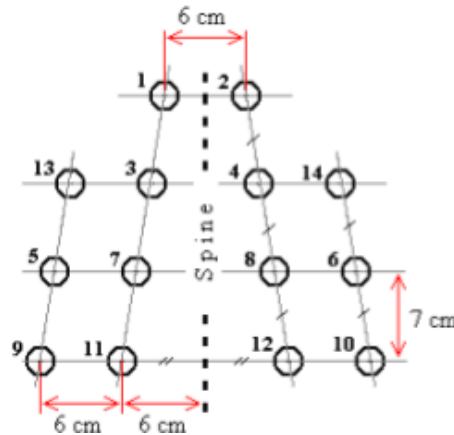
Figure 2.1. Microphone Locations on the Chest Wall [1]

3. AIRFLOW CURVE ESTIMATION

3.1. Autoregressive Models

Autoregressive modeling is a widely used method in spectral signal analysis. It is also used in respiratory signal analysis and since the respiratory sounds have spectral characteristics changing over time, either we divide the sound into small pieces and assume the small pieces are stationary or we model the respiratory sounds as time varying autoregressive processes. It is reported that when a respiratory sound signal is modeled as TVAR process, the first AR coefficient is highly correlated with the airflow for the sounds recorded at trachea and have less but still correlation for the sounds recorded at chest wall [11]. We will test this for the sounds recorded at chest wall and try to find best settings for most correlation in this subsection. We will try to extract the AR coefficients with three methods:

- Windowing Based Autoregressive Modeling: Divide the signal into overlapping windows assume each window is independent from others.
- Time Varying Autoregressive Modeling with Basis Functions: Assume the AR parameters are combination of sinusoidal signals.
- Time Varying Autoregressive Modeling with Kalman Filter: Assume the AR parameters are changing slowly.

3.1.1. Univariate Autoregressive Model

Univariate AR model is a difference equation where the current measurement of the signal is a linear combination of past measured values and a white Gaussian noise. The model is equivalent to a filter with infinite impulse response where the input is white Gaussian noise. The equation of an AR model is given in (3.1). In this equation a_i is the ith order AR coefficient, N is the order of model and e is the noise term where

$\mu_e = 0$ and σ_e^2 is constant [13].

$$x(n) = \sum_{i=1}^{N} a_i x(n-i) + e(n) \tag{3.1}$$

Before looking at mean, variance, autocorrelation and spectral content of a general AR model, we must set some conditions on a_i's to make the model stationary. The stationarity requires the mean, variance and autocorrelation function is constant. If the mean is constant then (3.2) must hold. We know that μ_e is 0, then either $\sum_{i=1}^{N} a_i = 1$ or $\mu_x = 0$ is true.

$$\mu_x = \sum_{i=1}^{N} a_i \mu_x + \mu_e \tag{3.2}$$

We require $E[X(0)] = 0$ for initial conditions for AR processes and it ensures $\mu_x = 0$. Then for variance we can write the equations in (3.3), (3.4) and (3.5).

$$\sigma_X^2 = E[X(n)X(n)] - \mu_X^2 \tag{3.3}$$

$$\sigma_X^2 = \sum_{i=1}^{N} a_i E[X_n X_{n-i}] + E[X_n e_n] - \mu_X^2 \tag{3.4}$$

$$\sigma_X^2 = \sum_{i=1}^{N} a_i R_{XX}(i) + \sigma_e^2 \tag{3.5}$$

We know that autocorrelation coefficients are very important in analysis of an AR model. The AR coefficients can be directly calculated in presence of autocorrelation

coefficients.

$$r_{xx}(i) = R_{xx}(i)/R_{xx}(0) \tag{3.6}$$

$$r_{xx}(i) = E[x(n)x(n-i)]/R_{xx}(0) \tag{3.7}$$

$$r_{xx}(i) = \frac{\sum_{j=1}^{N} E[a_j x(n-j)x(n-i)]}{R_{xx}(0)} \tag{3.8}$$

$$r_{xx}(i) = \sum_{j=1}^{N} a_j r_{xx}(i-j) \tag{3.9}$$

The equation in (3.9) is very useful since we can rewrite it as in (3.10), this equation is called Yule-Walker equation. So, given a signal generated with an AR model, we can find the coefficients of that model by solving the equation in (3.10). Let's call this equation as $r = Ra$, the the coefficients can easily be found with $a = R^{-1}r$.

$$
\begin{pmatrix} r_{XX}(1) \\ r_{XX}(2) \\ \vdots \\ r_{XX}(N-1) \\ r_{XX}(N) \end{pmatrix} = \begin{pmatrix} 1 & r_{XX}(1) & \dots & r_{XX}(N-2) & r_{XX}(N-1) \\ r_{XX}(1) & 1 & \dots & \dots & r_{XX}(N-2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{XX}(N-2) & \dots & \dots & 1 & r_{XX}(1) \\ r_{XX}(N-1) & r_{XX}(N-2) & \dots & r_{XX}(1) & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{N-1} \\ a_N \end{pmatrix} \tag{3.10}
$$

Although solving the equation (3.10) seems to be very straightforward, it requires exact knowledge on correlation coefficients. However, when we are given a signal with finite length we can just estimate the correlation coefficients, and the error in calculated AR coefficients will depend on the condition number of R matrix. To solve this equation, there are several methods in literature, throughout this thesis we will use Burg's Method to find the AR coefficients, since it is more stable than the others [14].

Algorithm 1 AR Coefficients Estimation With Overlapping Windows
---

1: **procedure** WINDOWEDAR(signal, order, winLen, overlap)

2:     $windowStart \leftarrow 1; L \leftarrow Length(signal);$

3:     **while** $windowStart \leq L - winLen$ **do**

4:         $temp \leftarrow signal(windowStart : (windowStart + windowLength));$

5:         $AR(:, i) \leftarrow EstimateAR(temp, N);$

6:         $windowStart \leftarrow windowStart + windowLength;$

7:     **end while**

8:     **return** $AR;$

9: **end procedure**

3.1.2. Time Varying Autoregressive Model

While univariate autoregressive model is very useful for many signals, the method doesn't use the continuity of the signal and treats each segment independently. Assuming the coefficients at different time instants are correlated with each other and building some structured models is another method which is widely applied for nonstationary signals. TVAR modeling were applied to respiratory signals and satisfactory results have been obtained in many papers. The equation describing a TVAR process is given in (3.11).

$$x(n) = \sum_{i=1}^{N} a_i(n)x(n-i) + e(n) \tag{3.11}$$

We will use two methods for solving this equation for a_is. One of them is modeling a_is as combination of sinusoids, the other one is modeling a_is as slowly changing parameters.

3.1.2.1. TVAR Coefficients Estimation With Fourier Basis Functions. In this method, the AR coefficients are assumed to be combination of sinusoidal functions. Then equa-

tions in (3.12) and (3.13) can be written.

$$x(n) = \sum_{i=1}^{N} x(n-i) \sum_{j=1}^{M} c_{ij} u_j(n-i) + e(n) \qquad (3.12)$$

$$a_i(n) = \sum_{j=1}^{M} c_{ij} u_j(n-i) \qquad (3.13)$$

Equation (3.12) can be rewritten in vector form as in (3.14) where X_i is the diagonal matrix with the diagonal elements are adjacent elements of x starting from i, U is the matrix whose columns are basis vectors, and the c_i's are the unknown parameters in this equation.

$$x = \sum_{i=1}^{N} X_i U c_i + e \qquad (3.14)$$

Let $Y_i = X_i U$, the equation can be written in matrix form as in

$$x = \left(Y_1 | Y_2 \ldots Y_N \right) \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_M \end{pmatrix} + e \qquad (3.15)$$

The equation in (3.15) describes an overdetermined set of equations, we will use least squares approach to solve this. After finding $\mathbf{c}$ vector, which has NxM elements, calculating $a_i(n)$, TVAR coefficients, will be straightforward.

$$\mathbf{c} = (Y^T Y)^{-1} Y^T x \qquad (3.16)$$

We will try to find the optimum values for AR order, number of basis vectors,

and the range of frequencies spanned by basis vectors to get the best correlation.

Algorithm 2 TVAR Coefficients Estimation With Fourier Basis Functions

1: **procedure** TVARFOURIERBASIS(signal, order, freqRange, numBasis)

2:     $\delta_f = freqRange/numBasis$;

3:     **for** i=1:numBasis **do**

4:         $U(:, 2 * i - 1) = sin(i * \delta_f)$;

5:         $U(:, 2 * i) = cos(i * \delta_f)$;

6:     **end for**

7:     $Generate\ Y\!:\ Project\ signal\ onto\ U$;

8:     $c = (Y^T Y)^{-1} Y^T x$

9:     **for** i=1:N **do**

10:         $AR(:, i) = U c_i$;

11:     **end for**

12:     **return** AR;

13: **end procedure**

3.1.2.2. TVAR Coefficients Estimation With Kalman Filter. Kalman filter is a very old but still popular algorithm in field of information processing. It is the minimum mean square estimator for the state of linear dynamical systems [15]. It is used everywhere from tracking applications to computer games.

Before explaining how Kalman filter works, we must first give the required equations to describe the model where it is applied. Kalman filter is invented to work on dynamic systems where we can record the noisy observations of transformations of the process in interest. So we may give two equations, for both measurement (3.17) and process (3.18) model [16].

$$y_n = H_n x_n + v_n \tag{3.17}$$

$$x_n = F_n x_{n-1} + B_n u_n + w_n \tag{3.18}$$

In (3.18) x is the state vector, u is control input to the system, B is the control matrix, F is the state transition matrix and w is process noise. In (3.17) the H is the transform matrix and v is the measurement noise. When a model desribed in (3.17) and (3.18) is present, we can observe y and we are after estimatin x, we can apply Kalman filter. When the process and measurement noises are Gaussian then Kalman is the optimal estimator.

In our problem, there is not any control input to system and equations convert to (3.19) and (3.20). In these equations y is the recorded sound amplitude, x is the AR coefficients, H is equal to raw vector containing previous N values of y where N is the AR order.

$$y_n = H_n x_n + v_n \tag{3.19}$$

$$x_n = F_n x_{n-1} + w_n \tag{3.20}$$

Kalman filter includes two stages, prediction and measurement update. The prediction equations are given in (3.21) and (3.22). The measurement update equations are given in (3.23), (3.24) and (3.25). In these equations $\hat{x}_{n|m} = E[x_n|y_{1:m}]$ and $P_{n|m} = E[(x_n - \hat{x}_{n|m})(x_n - \hat{x}_{n|m})^T|y_{1:m}]$.

Prediction:

$$\hat{x}_{n|n-1} = F\hat{x}_{n-1|n-1} \tag{3.21}$$

$$P_{n|n-1} = FP_{n-1|n-1}F^T + \sigma_w^2 \tag{3.22}$$

Measurement Update:

$$K_n = P_{n|n-1}H(H_n^T P_{n|n-1} H_n + \sigma_v^2)^{-1} \tag{3.23}$$

$$\hat{x}_{n|n} = \hat{x}_{n|n-1} + K_n(y_n - H_n^T \hat{x}_{n|n-1}) \tag{3.24}$$

$$P_{n|n} = (I - K_n H_n^T) P_{n|n-1} \tag{3.25}$$

In the equations above, the filter uses only measurement values before N to estimate a value of x_N. In order to add the information from complete signal, the estimations can be smoothed with Rauch-Tung-Striebel backward recursions given in (3.26), (3.27) and (3.28).

$$J_n = P_{n|n}F^T P_{n+1|n}^{-1} \tag{3.26}$$

$$\hat{x}_{n|N} = \hat{x}_{n|n} + J_n(x_{n+1|N} - x_{n+1|n}) \tag{3.27}$$

$$P_{n|N} = P_{n|n} + J_n(P_{n+1|N} - P_{n+1|n})J_n^T \tag{3.28}$$

There process noise, state transition matrix and measurement noise are not being updated, and we must tune them. The tuning process and results will be given in Experiments & Results section.

3.2. Time Frequency Analysis

Short Time Fourier Trasnform is a technique which is widely used and very useful for analysis of signals with a time varying spectral characteristics. STFT is used for analysis of respiratory signals since they have a nonstationary nature.

$$X_m(f) = \sum_{n=-\infty}^{\infty} x(n)w(n - mR)e^{-j2\pi fn} \tag{3.29}$$

STFT is defined in (3.29). It can be seen as a sliding Fourier Transform (FT) [17].
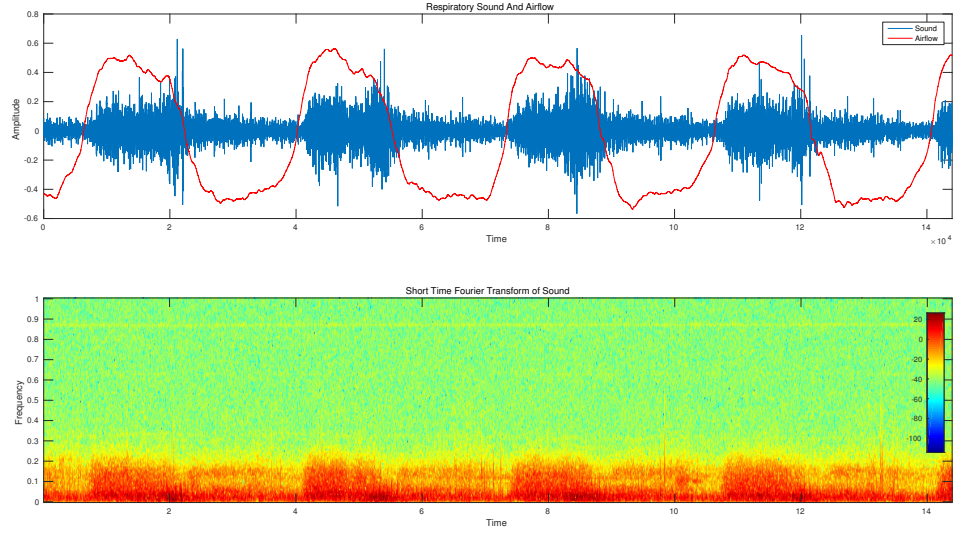
Figure 3.1. A respiratory sound, corresponding airflow and STFT magnitude plot

In this equation, R is hop size and w is the window function which zero outside of a predefined range and is used to pick the part of signal which will be input of FT and smooth the signal to make it stationary. The window is an important parameter for STFT, it is the effective parameter to adjust time-frequency resolution. First of all, the window length must be so small that it must ensure that the selected portion is stationary. There is also a trade-off between resolution in time and resolution in frequency, this trade-off is also controlled with window length. As the window length increases the frequency resolution increases and the time resolution decreases. So, for wideband signals one can use smaller window lengths whereas for narrowband signals greater window lengths may be used.

The output of STFT operation is a two dimensional complex matrix, which describes the magnitude and phase of frequency band component. For most of the time and for respiratory sounds too, we need the magnitude information and we convert the complex matrix to a real matrix which will give information about energy directly. The resulting matrix is usually visualized by a heatmap as in 3.1.

When we look at figure 3.1, respiratory sound, magnitude plot of its STFT and corresponding airflow we can see a trend seems related to airflow in evolution of some

horizontal lines. They correspond to the energy at a frequency band over the time. We will run experiments to test the correlation between the evolution of energy at each frequency band with the airflow. We will tune STFT by tweaking the window type, window length and number of FFT bins in Experiment & Results.

3.3. Unifying Estimations

When we have an estimation problem with desired output d and observations in x vector if each observation in x and d are jointly wide sense stationary (wss) the optimum estimator for mean square error (MSE) is the Wiener filter [18]. We can write down the equations and derive the optimum filter. The estimation equation is given in (3.30), $\hat{d}$ is the estimation in this equation. Estimation error is given in (3.31) and the error measure we are interested in is given in (3.33)

$$\hat{d} = \sum_{i=1}^{N} w_i x_i = \mathbf{w^T x} \tag{3.30}$$

$$e = d - \hat{d} = d - \mathbf{w^T x} \tag{3.31}$$

$$E[e^2] = E[|d - \hat{d}|^2] \tag{3.32}$$

$$E[e^2] = E[d^2 - 2d\mathbf{x^T w} + \mathbf{w^T x x^T w}] \tag{3.33}$$

In order to minimize the error we can differentiate with respect to $\mathbf{w}$ and find the value where derivative is zero. For error to be at its minimum the expectation in (3.34) must be zero, this requires $\mathbf{x}$ and e to be uncorrelated.

$$\frac{\partial E[e^2]}{\partial \boldsymbol{w}} = 2E[e \frac{\partial e}{\partial \boldsymbol{w}}] = -2E[e x_i] \tag{3.34}$$

If we rewrite error as difference between d_n and $\mathbf{w_{opt}^T x}$, then we can state the

equations in (3.35) and (3.36).

$$E[(d_n - \mathbf{w_{opt}^T}\mathbf{x})\mathbf{x}] = 0 \tag{3.35}$$

$$E[\mathbf{xx^T}]\mathbf{w_{opt}} = E[\mathbf{x}d_n] \tag{3.36}$$

The first expectation in (3.36) is autocovariance matrix of $\mathbf{x}$ and the second expectation is crosscovariance between x and d. The filter w_{opt} is the called as Wiener filter.

In previous sections we end up with vectors which are estimations of airflow. We will use a weighted sum of these estimations as our observations and weights will be decided by using this Wiener filter approach.

3.4. Experiments & Results

In this section, the experiments to reach the best tuning parameters and the results will be given for the methods explained in this chapter.

3.4.1. Univariate Autoregressive Model

There are 4 parameters to be optimized in univariate autoregressive modeling method. These parameters are the coefficient degree, AR order, window length and overlap.

Firstly, we will try to answer the question about which AR order and which coefficient order is best. For this purpose, experiments will be done for AR orders from 1 to 15 with three different window length 500, and 50% overlap.

The inference from figures is that the correlation is decreasing with increasing coefficient order and the correlation with absolute value of flow is significantly greater.

Average Correlation of Evolution of Coefficient with Absolute Airflow, $WindowLength = 500\ Overlap = \%50$
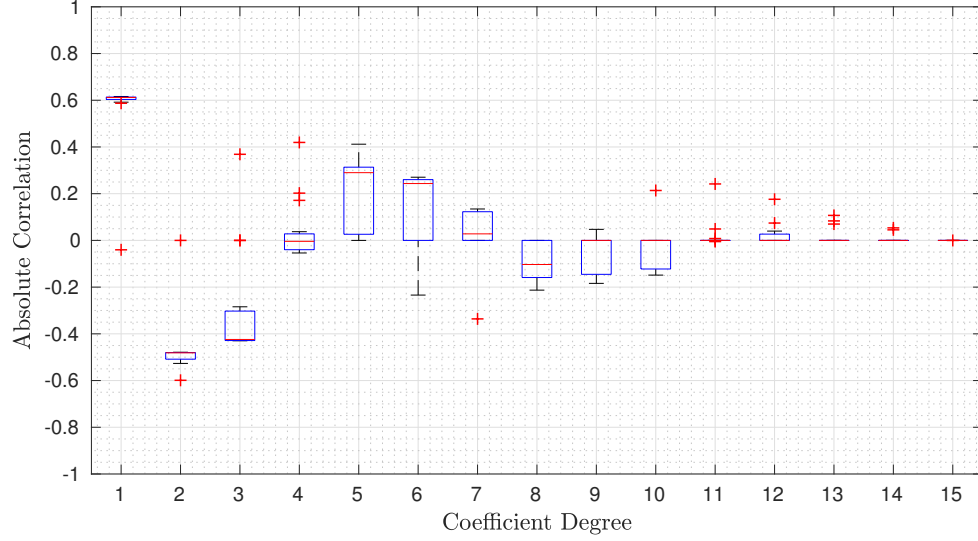

Figure 3.2. Boxplot for correlation coefficient of AR coefficient evolution with absolute value of airflow for different coefficient orders

Average Correlation of Evolution of Coefficient with Airflow, $WindowLength = 500\ Overlap = \%50$
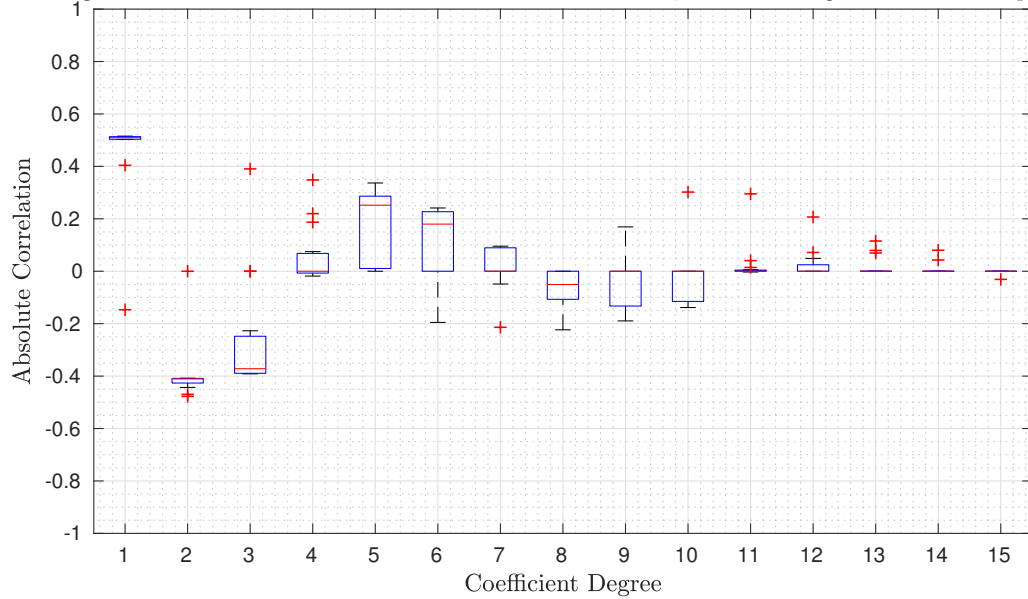
Figure 3.3. Boxplot for correlation coefficient of AR coefficient evolution with airflow for different coefficient orders

Correlation of Evolution of First AR Coefficient with Absolute Airflow, $WindowLength = 500$ $Overlap = \%50$
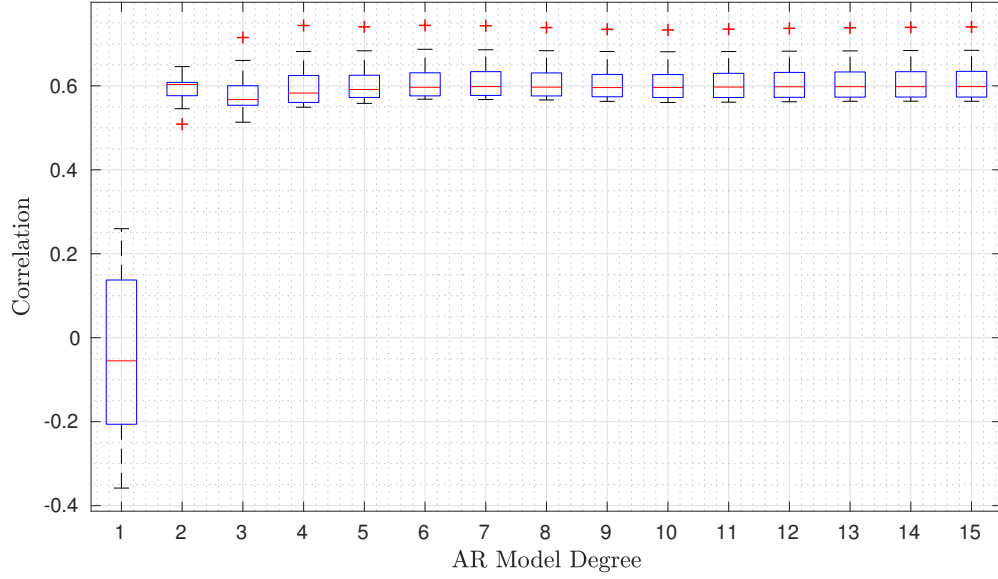

Figure 3.4. Boxplot for correlation coefficient of AR coefficient evolution with airflow for different AR model orders
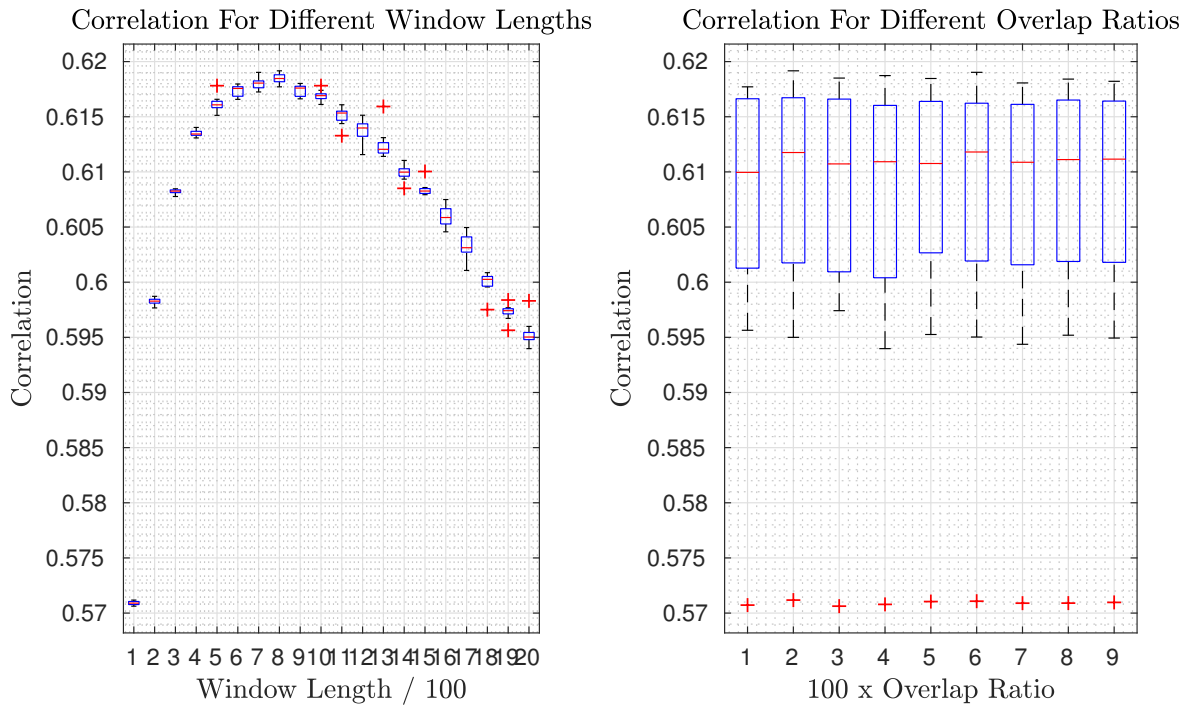

Figure 3.5. Boxplot for correlation coefficient of first AR coefficient evolution with absolute airflow for different window lengths and overlap ratios

We will continue our analysis with absolute value of airflow for AR estimators. Now we will continue the analysis to find the best model order. For this, we run experiments with window length 500 and an overlap of %50 with model orders from 1 to 15. The results are shown in figure 3.4. As can be seen from figure, there isn't any significant difference after sixth order AR model, so we chose 6 as our model order. After selecting model order, the window length and overlap ratio are left to be decided on. In order to select them we run experiments with window lengths from 100 to 2000 with a seperation of 100 and overlap ratios from 10% to 90% with a separation of 10%. The results for different window lengths and overlap ratios are summarized in figure 3.5, according to the results best window length is 800 and difference in overlap ratio doesn't generate any difference in the correlation. We decided to use 90% overlap to increase the resolution.

The resulting decision for univariate autoregressive solution with overlapping windows is 6, 800, 90% for model order, window length and overlap ratio respectively.

3.4.2. Time Varying Autoregressive Model with Basis Functions

The parameters for this method are model order, number of basis functions and the frequency difference in adjacent basis functions.

First, we run experiments to determine the best model order and for this purpose run experiments with model orders from 1 to 15 where the number of basis functions are 201 (100 sines, 100 cosines and a constant) and the separation in frequency is 0.04 Hz. Experiment results are given in figure 3.6. Most correlation is achieved with the model order of 6 again.

After deciding on model order, we need to decide on number of basis functions and frequency separation of basis functions. Before doing this, we run experiments to determine the frequency coverage for best correlation and run experiments with separation of 0.025 Hz and different number of basis frequencies from 50 to 300 with steps of 50. The result is given in figure 3.7. According to test results 5 Hz is enough to
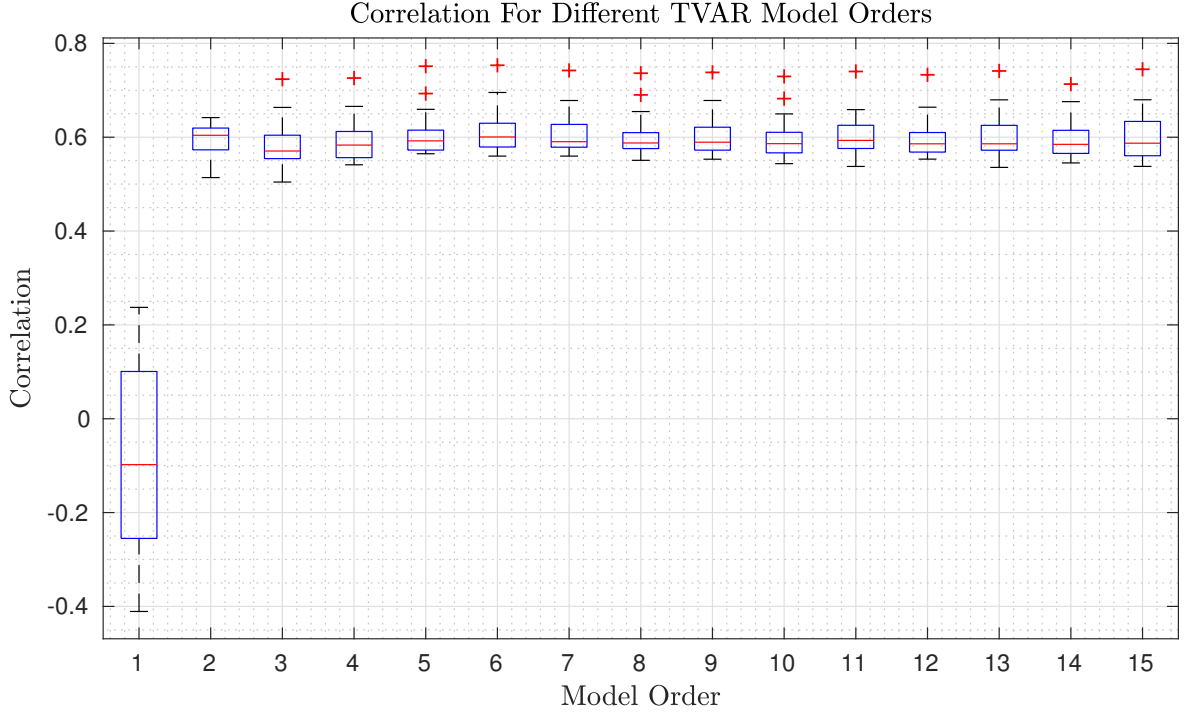
Figure 3.6. Boxplot for correlation coefficient of first AR coefficient evolution with absolute airflow for different TVAR model orders

cover for best results. Finally we run simulations to decide both the number of basis frequencies and frequency separation and for each tuning we covered the frequency range from 0 to 5 Hz. The results are given in figure 3.8, it shows that it doesn't make significant change to choose number of basis functions as 200 or 400. We chose it to be 250 for minimum standard deviation.

To summarize, 6, 250, 0.02 is chosen for model order, number of basis frequencies and frequency separation respectively.

3.4.3. Time Varying Autoregressive Model with Kalman Filter

For this method, we used the noise estimated by windowing based AR modeling as the measurement noise variance, and look for the best noise variance for process noise. We run simulations where the state uncertainity goes from 0.0002 to 0.004 with steps of 0.0002. The results are given in 3.9. We chose 0.004 as our process variance.

Correlation For Different Coverages

Figure 3.7. Boxplot for correlation coefficient of first AR coefficient evolution with absolute airflow for different frequency coverage

Correlation For Different Basis Functions and Same Frequency Coverage

Figure 3.8. Boxplot for correlation coefficient of first AR coefficient evolution with absolute airflow for number of basis functions
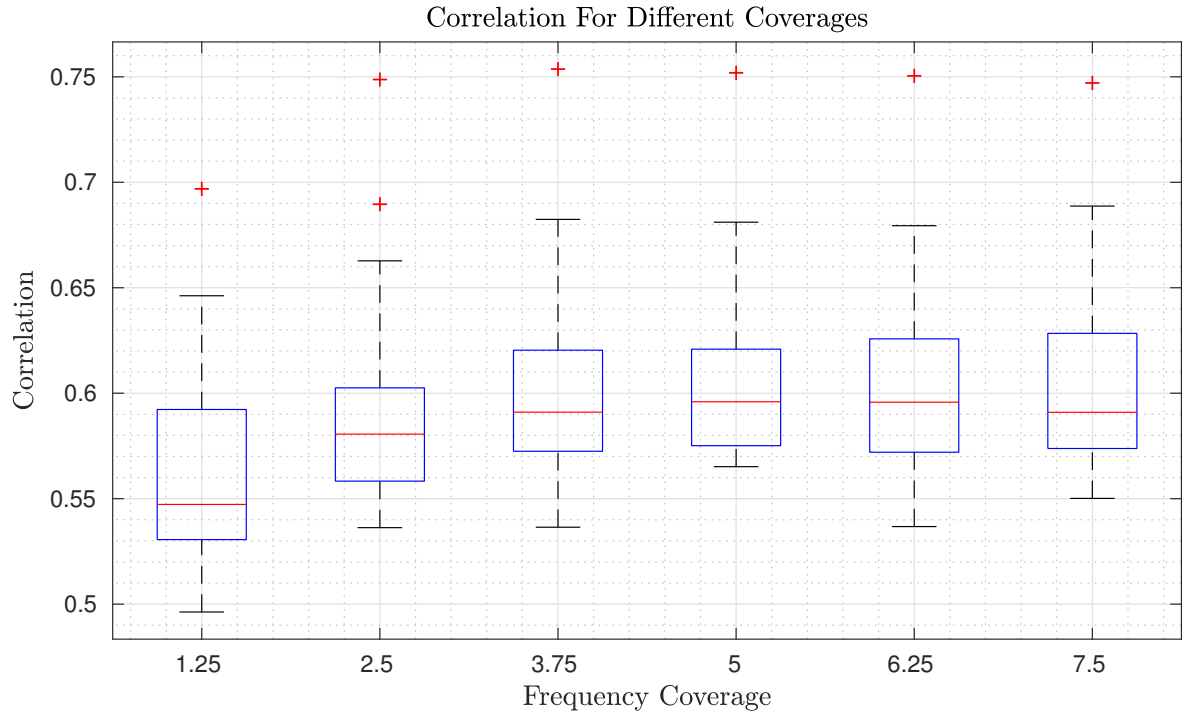
Figure 3.9. Boxplot for correlation coefficient of first AR coefficient evolution with absolute airflow for different process noise variances

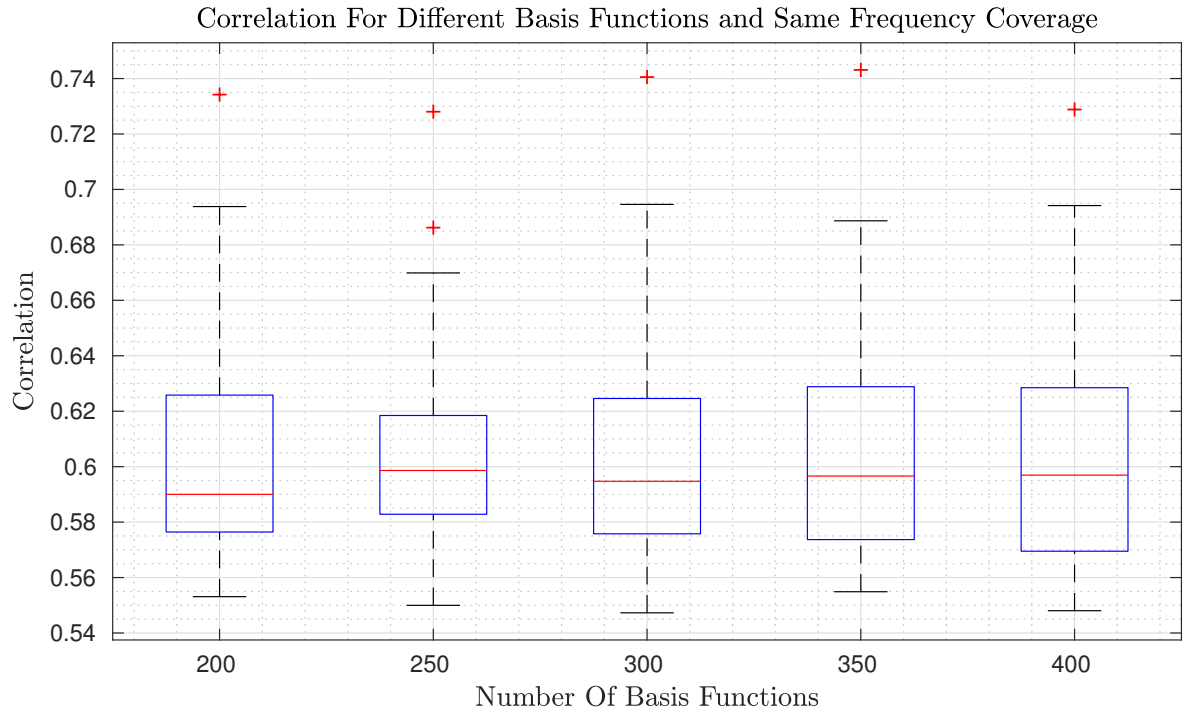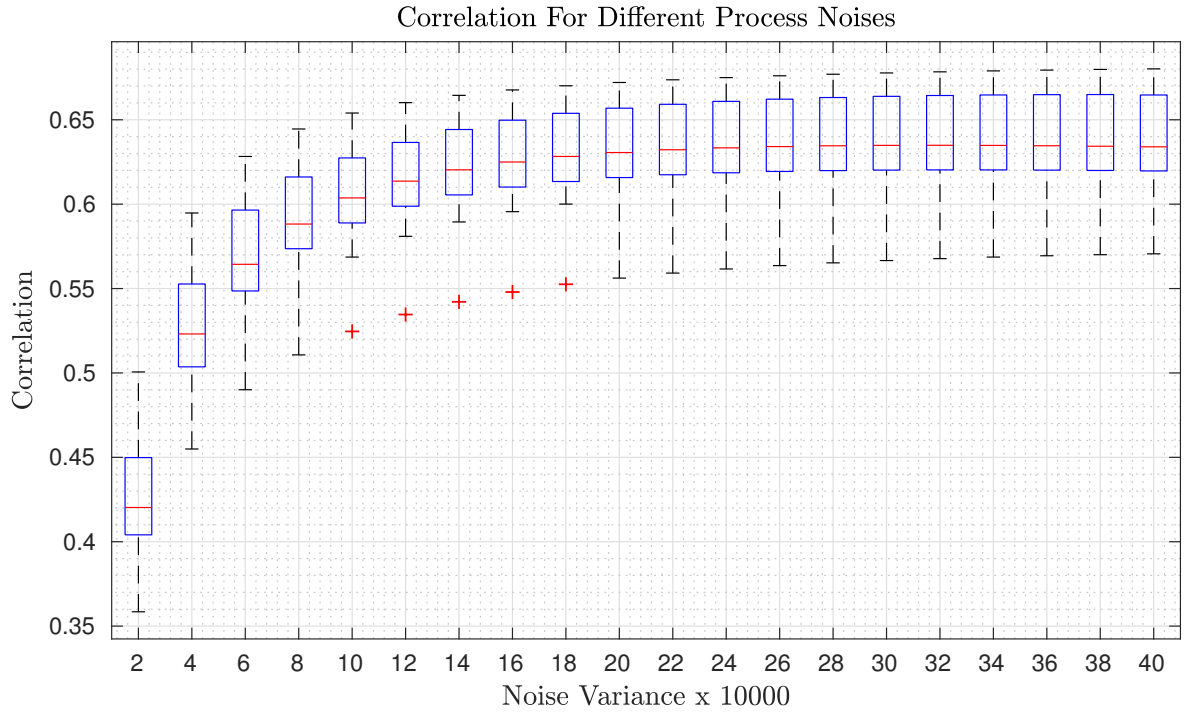### 3.4.4. Short Time Fourier Transform

Figure 3.10. Mean of correlations for each band for STFT method with 64 fft bins, window lengths and overlap ratios

Correlation Of Different Bands For 128 Point STFT With Different Window Lengths and Overlap Ratios

Figure 3.11. Mean of correlations for each band for STFT method with 128 fft bins, window lengths and overlap ratios

Correlation Of Different Bands For 256 Point STFT With Different Window Lengths and Overlap Ratios
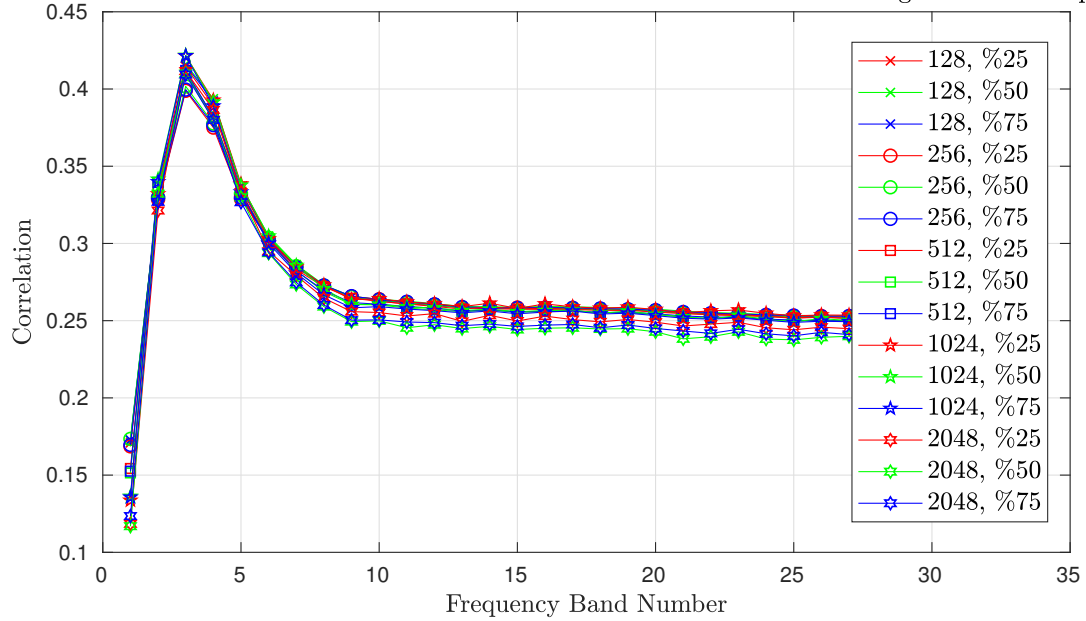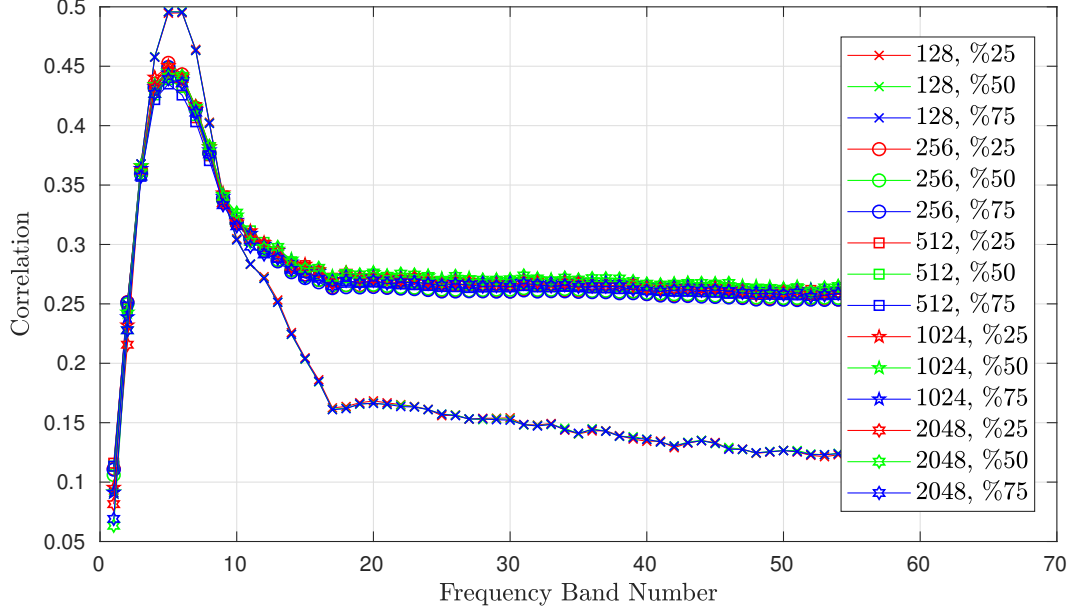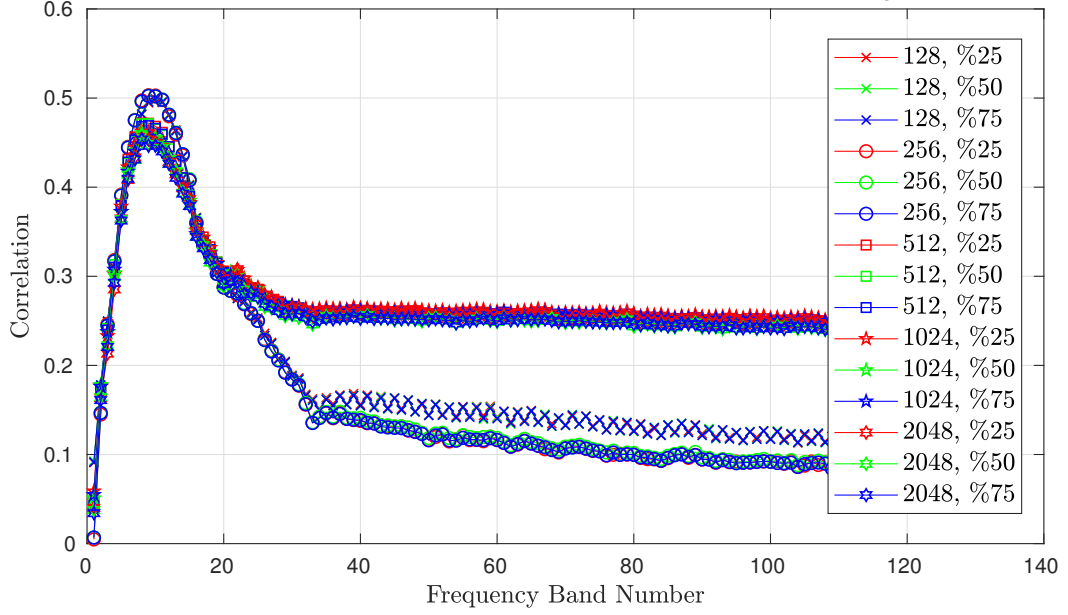

Figure 3.12. Mean of correlations for each band for STFT method with 256 fft bins, window lengths and overlap ratios

3.4.5. Unifying Estimations

We run simulations with different number of vectors to be unified, where the vectors are outputs of univariate AR and STFT methods.
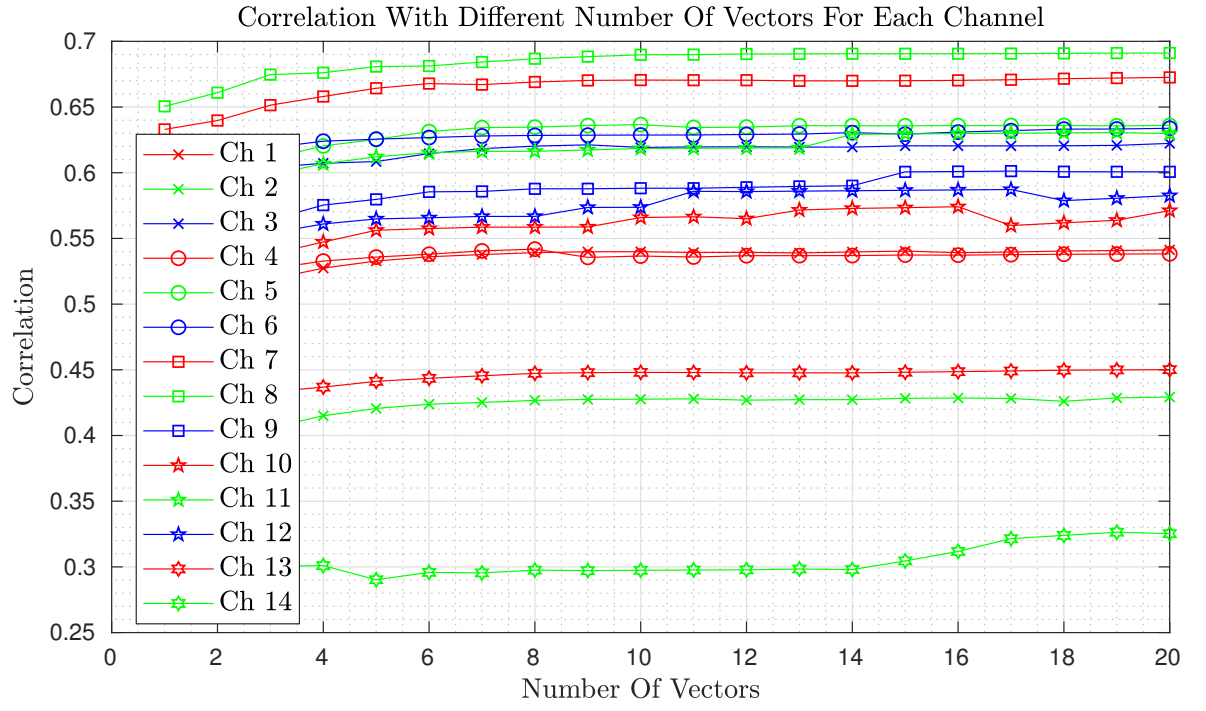
Figure 3.13. Mean of correlations for each channel with different number of vectors unified

3.4.6. Result

Channel	Univariate AR	Basis Functions	Kalman
1	0.6871	0.6862	0.6290
2	0.7440	0.7280	0.6802
3	0.6800	0.6699	0.6647
4	0.6310	0.6185	0.6536
5	0.6153	0.6047	0.6649
6	0.5935	0.5963	0.6134
7	0.6031	0.6031	0.6328
8	0.5980	0.5885	0.6672
9	0.5761	0.5828	0.6351
10	0.5900	0.6009	0.6328
11	0.5681	0.5500	0.6197
12	0.5688	0.5719	0.6362
13	0.5733	0.5652	0.6127
14	0.5952	0.5916	0.5706

Table 3.1. Correlation For Different Methods for All Channels

4.   AIRFLOW PHASE ESTIMATION

We will present more than one methods to estimate the respiratory phases in this chapter. In one approach we will use the estimated airflow and in the other approach we will not use that, we directly use neural networks. In both methods we will use the estimation of breathing period.

4.1. Period Estimation

We assume that the period of breathing doesn't change abruptly and stay almost constant for each recording. In order to estimate the period we look at the autocorrelation function. For a periodic wave, the autocorrelation function is also periodic with the same period and there are peaks at the integer products of period. So, the period can be estimated by calculating the distances between the peaks. However, since the signal is noisy there are many local maxima, so instead of estimating the distance between peaks, we looked at the autocorrelation function of autocorrelation function of the signal that is being analyzed, and use the location of the maximum of local peaks in the part that is determined by minimum and maximum period. The equation for estimated period is given in (4.1)

$$T' = period_{min} + argmax(R_{R_x}[period_{min} : period_{max}]) \qquad (4.1)$$

4.2. Phase Estimation With Neural Networks

We try to use a neural network to decide the phase of a given signal segment. We decided to use the following features: AR coefficients, Shannon entropy estimate, percentile frequencies (f_{25}, f_{50}, f_{75} and f_{90}) and the ratio between each of them, variance, spectral magnitude and kurtosis. These features change for each recording, bu the difference between inspiration and expiration doesn't change, so we used z-score
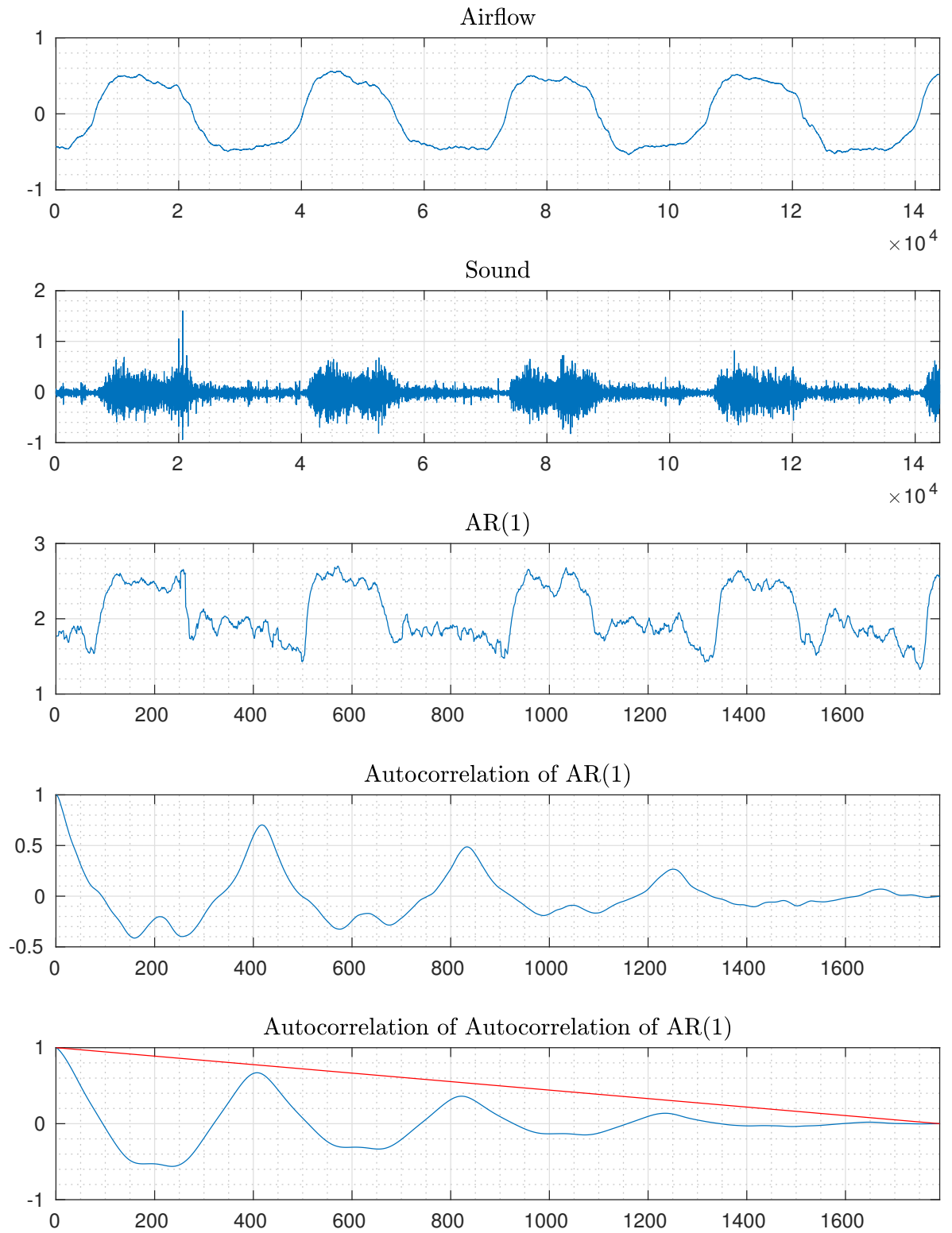
Figure 4.1. Visual Description Of Period Estimation

before giving these features to neural network. A visual description of estimating the period is given in Figure 4.1. We are looking at the ratio of blue function to the red line at the end.

4.2.1. Description of Neural Networks

A neural network is a function with many inputs and a single output. Formulation and development of neural networks are inspired from biological nervous system. Input values of a neural network is processed by dozens of interconnected functions, and these functions are called "neuron"s. A neuron can be thought as a function with multiple outputs and a single output. A neuron's output is determined by the activation function [19].

4.2.1.1. Neuron.   Mathematical expression of a neuron's activation function is given in (4.2). Here β is the activation function. Activation function may be in several different forms. It is selected based on the application where the neural network is used and the distribution of input values. Activation functions are generally selected or constructed in a way that they are easily differentiable. Some of popular activation functions are given in figure 4.3.

$$y = \beta(\sum_{i=1}^{N}(w_i x_i))\tag{4.2}$$
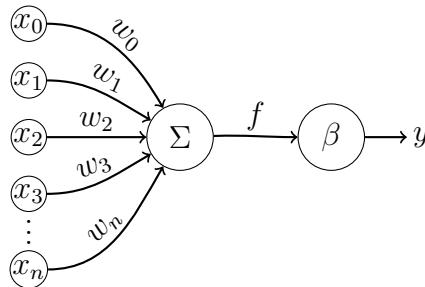
Figure 4.2. Diagram of A Neuron
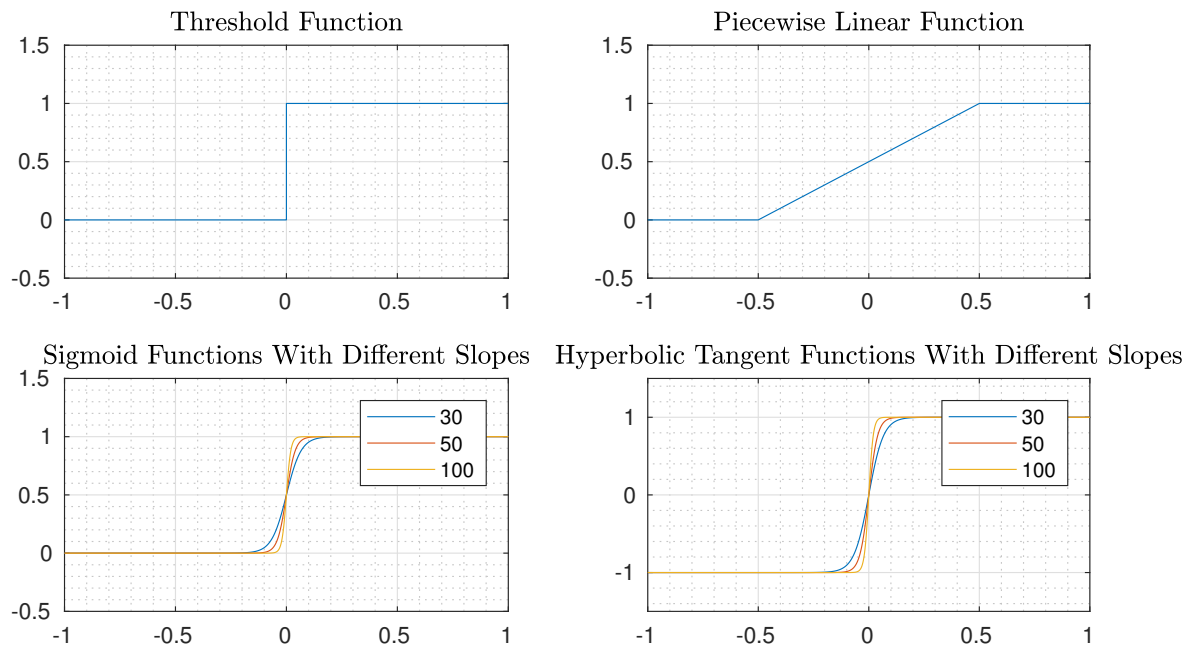
Figure 4.3. Some Activation Functions

4.2.1.2. Layers In Neural Networks. There are three different layers in neural networks, input layer, output layer and hidden layers. Input and output layers are where the input values are received and output values are given out respectively. Hidden layers are where the information from input layer or previous hidden layer is processed. Hidden layers may have arbitrary number of neurons and there may be any number of hidden layers in a neural network.

Figure 4.4. A Neural Network With One Hidden Layer

4.2.1.3. Learning Process In Neural Networks. Learning is the primary property of neural networks; it enables a neural network to improve its performance.

The learning processes can be divided into three categories with respect to the learning paradigms. These categories are supervised, reinforcement and unsupervised learning. In supervised learning, the desired output is available so the error is known in the learning process. In reinforcement learning there is a performance measure which is being tried to be maximized. In unsupe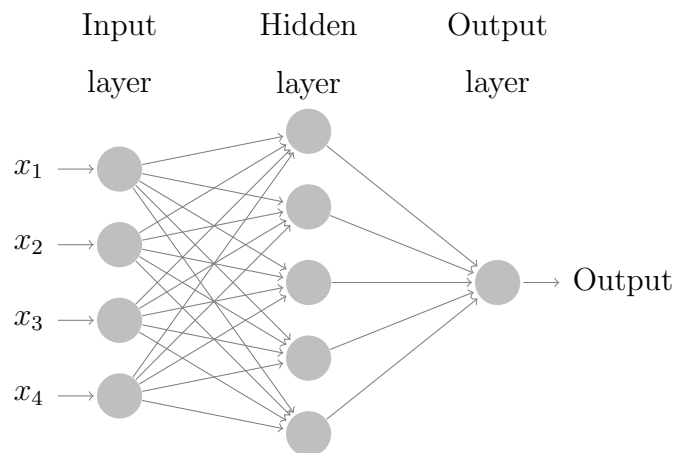rvised learning, neither desired output nor performance measure is available. We will use supervised learning since the correct phases are known to us.

Learning methods can be divided into four sections, error-correction learning, memory based learning, Hebbian learning, competitive learning and Boltzmann learning. We will use error-correction learning and try to minimize the total error measure by adjusting the synaptic weights.

4.2.2. Features

In this section, we will look at the features that will be used in neural network approach. Each feature is found for all the sounds. However while looking at features, one can see that the patterns don't change however the range of values change. To overcome this issue each feature series from each sound gets normalized by using z-score. Z-score is the measure which gives how much distance a random sample has to its source mean. The formula for z-score is given in (4.3)

$$z_i = \frac{x_i - \hat{\mu}_X}{\hat{\sigma}_X} \tag{4.3}$$

We also used the Kullback-Leibler (KL) divergence to quantify how much the distributions for inspiration and expiration differ from each other for each feature. KL-divergence is a definition from information theory and it gives a distance between

probability distributions. The formula for KL-divergence is given in (4.4) [20].

$$D_{KL}(p(x)||q(x)) = \sum_{x \in X} p(x) \log \frac{q(x)}{p(x)} \tag{4.4}$$

4.2.2.1. AR Coefficients. When we modeled the respiratory sounds as AR processes and solve for the AR coefficients, we saw that estimated AR coefficients are behaving differently for inspiration and expiration. To give an example, for the case of first AR coefficient it seems as the coefficients are attenuated for expiration phases. So we decided to use AR coefficients as a feature to the neural network. A detailed explanation about AR models and methods to estimate AR coefficients is given in Section 3. Distributions of AR coefficients during inspiration and expiration for third channel is given in Figure 4.5

4.2.2.2. Shannon Entropy Estimate. Entropy is a measure which gives information about the complexity of the samples collected. As the randomness of the samples increases entropy increases too. The formula to calculate entropy is given in (4.5) [20]. In general we have a wider range for signal values in inspiration, so we expect entropy to be greater for the segments belonging to inspiration.

Distributions of Shannon's entropy estimates during inspiration and expiration for third channel is given in Figure 4.6

$$H(X) = - \sum_{x \in X} p(x) \log(x) \tag{4.5}$$

4.2.2.3. Percentile Frequencies. A percentile frequency is defined together with the percent. It is the frequency until which the cumulative sum of power spectral density is equal to the defined percent of total power. In order to calculate them, we first estimate the power spectral density according to the formula in Percentile frequencies and their relations are measures which are widely used in respiratory signal analysis [21,22].
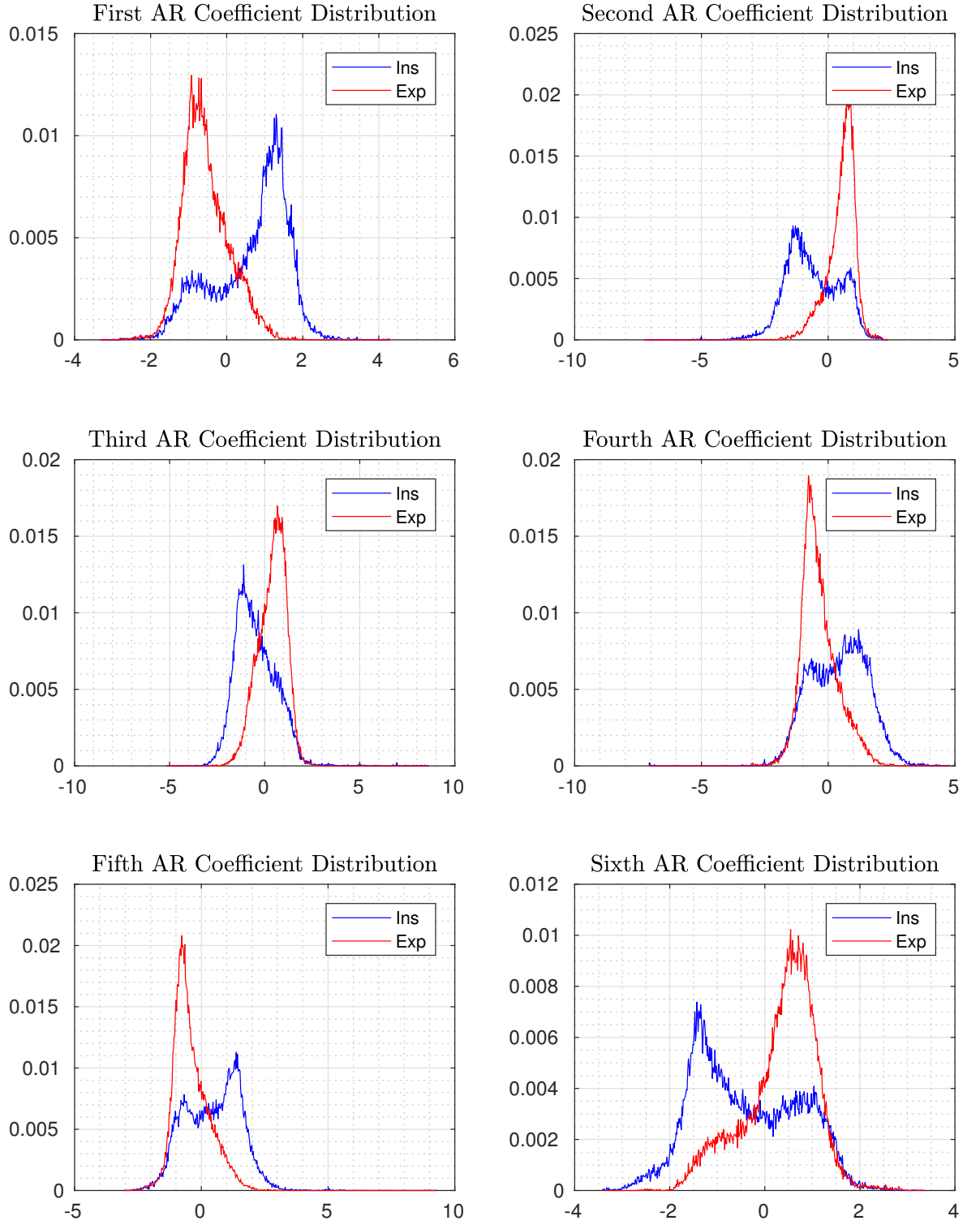
Figure 4.5. Distribution of AR Coefficients For Inspiration and Expiration

Since the expiration and inspiration phases have different time-frequency content, percentile frequencies may be used as input to classifier neural network. Definition of a percentile frequency is given in (4.8)
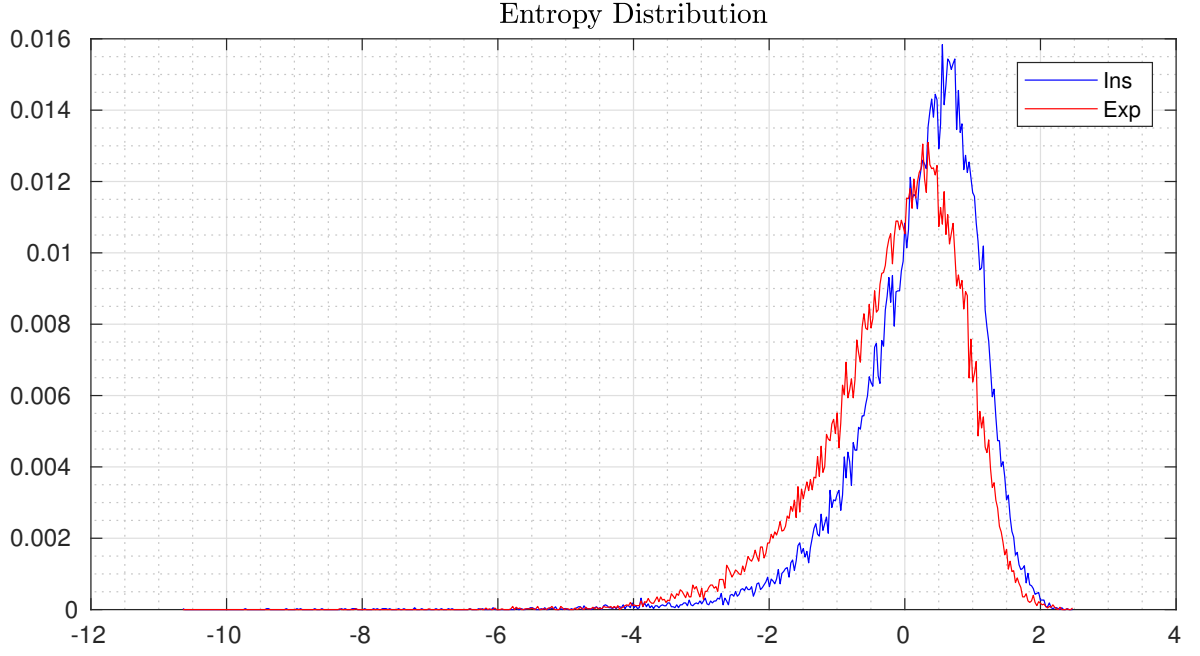
Figure 4.6. Distribution of Shannon Entropy Estimate For Inspiration and Expiration

Distributions of percentile frequencies and their ratios between each other during inspiration and expiration for third channel is given in Figure 4.7.

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-2\imath f t\pi} dt \tag{4.6}$$

$$P_{XX}(f) = X(f)^2 \tag{4.7}$$

$$f_y = \arg\min_{x} \frac{\int_0^x P_{XX}(f)df}{\int_0^\infty P_{XX}(f)df} > \frac{y}{100} \tag{4.8}$$

4.2.2.4. Variance. Variance corresponds to power and observations show that power for different respiratory phases differs. So we decided to use variance as a supportive feature. Sample variance is calculated as in (4.10).

Distributions of variance of windows during inspiration and expiration for third channel

Figure 4.7. Distribution of Percentile Frequency Measures For Inspiration and Expiration

Variance Distribution
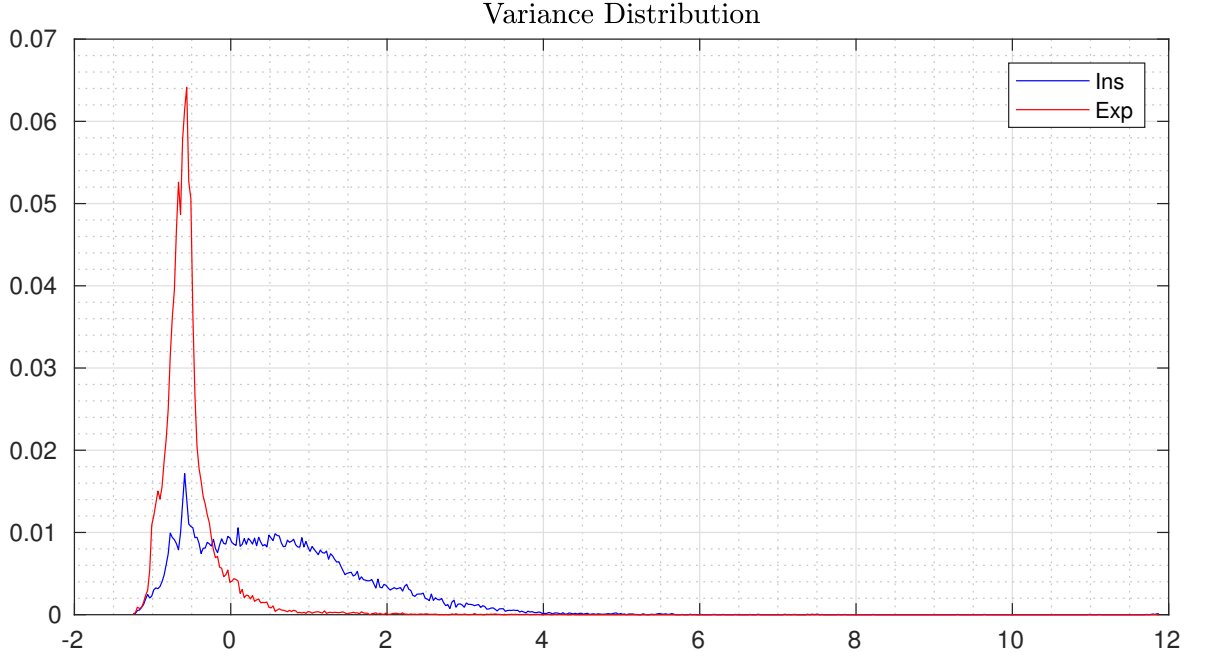
Figure 4.8. Distribution of Variance For Inspiration and Expiration

is given in Figure 4.7.

$$\hat{\mu}_X = \frac{\sum_{x \in X} x}{|X|} \tag{4.9}$$

$$\hat{\sigma}_X^2 = \frac{\sum_{x \in X} (x - \hat{\mu}_X)^2}{|X| - 1} \tag{4.10}$$

4.2.2.5. Spectral Magnitude. The magnitudes of several frequency bands are calculated for each segment using FFT. We calculated the spectral band's magnitude by taking FFT of windows after multiplying them a Hamming window for smoothing. We used 128 points FFT and each band represents a band of 75 Hz. We used the following bands' magnitudes: {150 Hz-225 Hz, 225 Hz-300 Hz, 300 Hz-375 Hz, 375 Hz-450 Hz}. Distributions of spectral magnitudes of windows during inspiration and expiration for third channel is given in Figure 4.9.
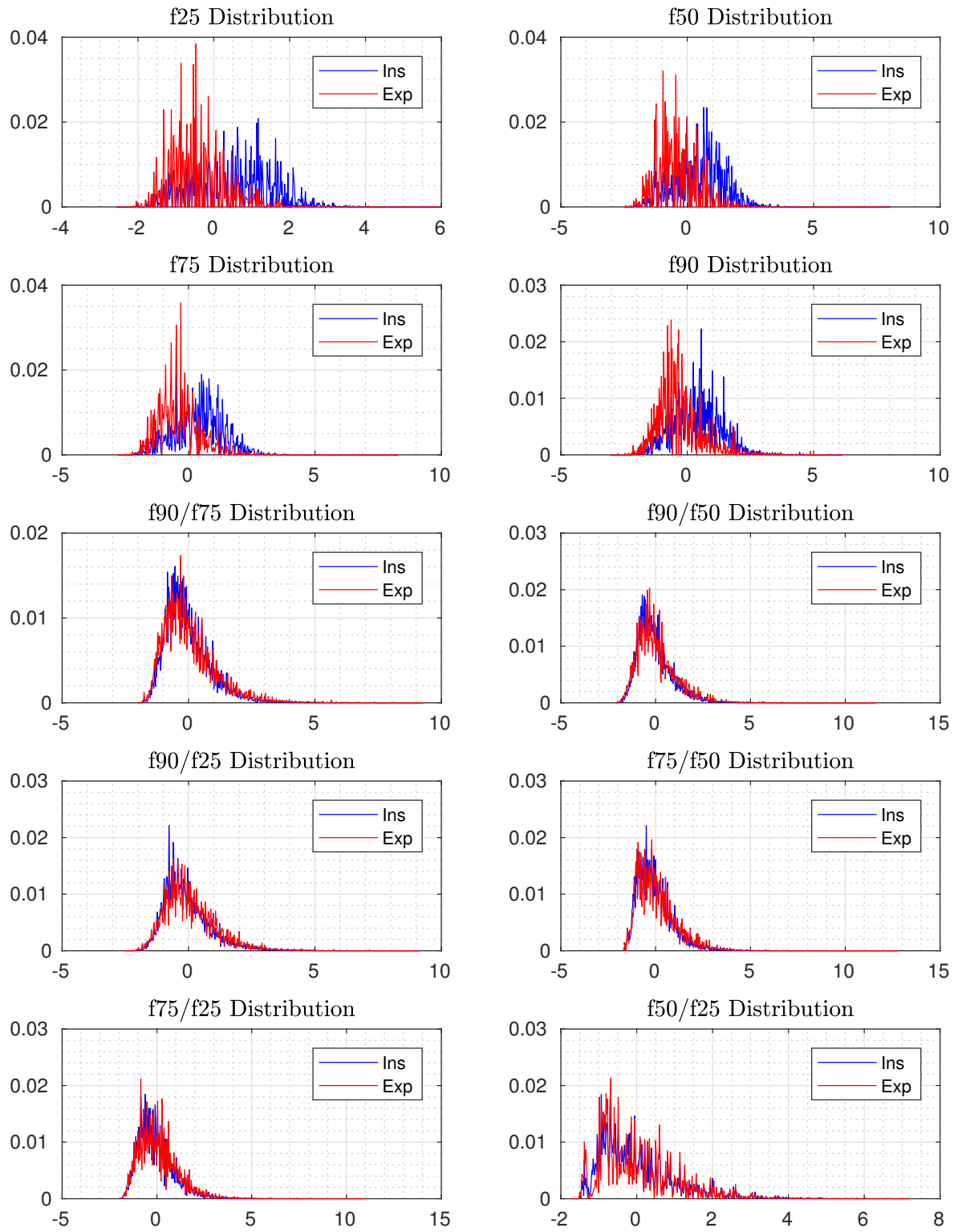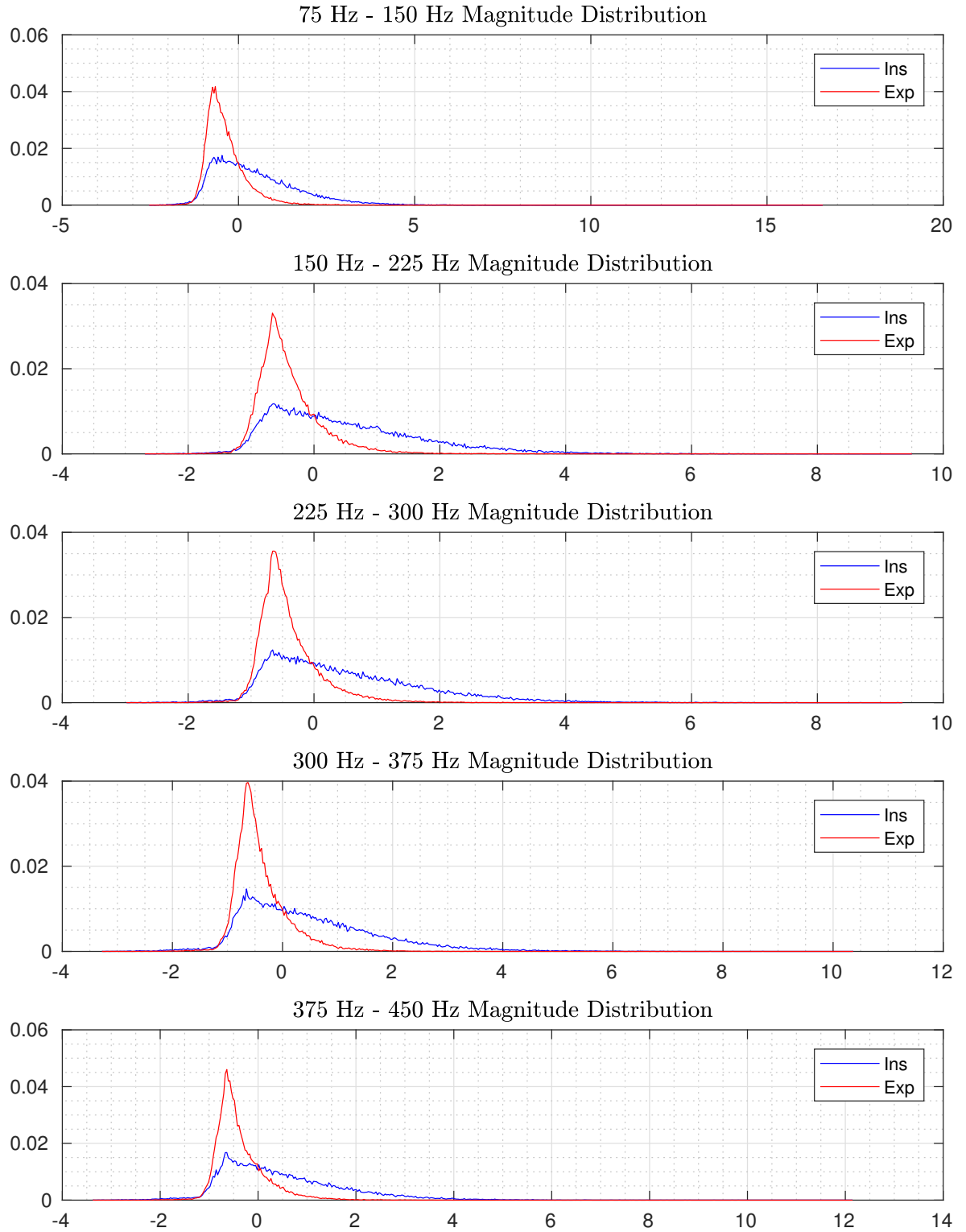
Figure 4.9. Distribution of Spectral Magnitudes For Inspiration and Expiration
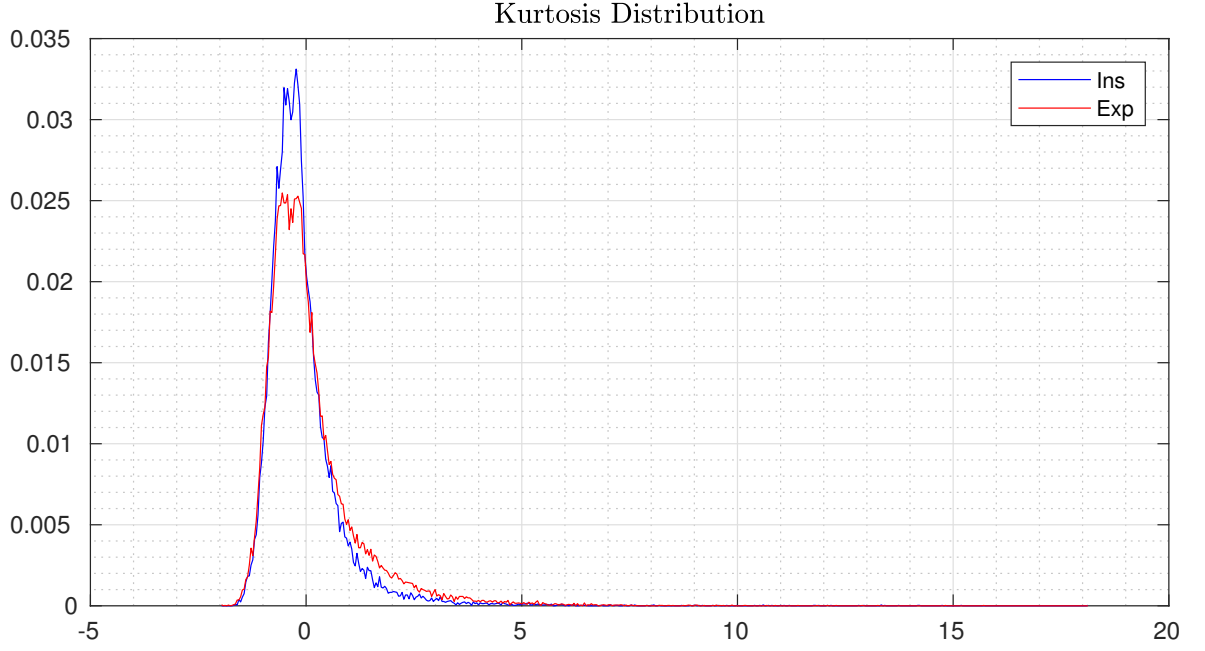
Figure 4.10. Distribution of Kurtosis For Inspiration and Expiration

4.2.2.6. Kurtosis. Kurtosis is the ratio of the fourth moment to square of second moment. It is a statistical feature which have been used in respiratory signal analysis. Formula for kurtosis is given in (4.11)

Distributions of kurtosis of windows during inspiration and expiration for third channel is given in Figure 4.10.

$$K = |X| \frac{\sum_{x \in X} (x - \hat{\mu}_X)^4}{(\sum_{x \in X} (x - \hat{\mu}_X)^2)^2} \tag{4.11}$$

The average KL-divergence for features is given in figure 4.11. The feature indices are consistent with the order features are explained above.

4.2.3. Denoisining The Output of Neural Network

The output of neural network is not very good in terms of estimating the phase correctly and it is alternating too much even in short durations. In order to overcome this issue, a filtering was needed. Linear filters don't perform well since they show many

KL-Divergence Between Distributions for Features Calculated in Inspiration and Expiration

Figure 4.11. KL-Divergence Between Distributions for Features Calculated in Inspiration and Expiration

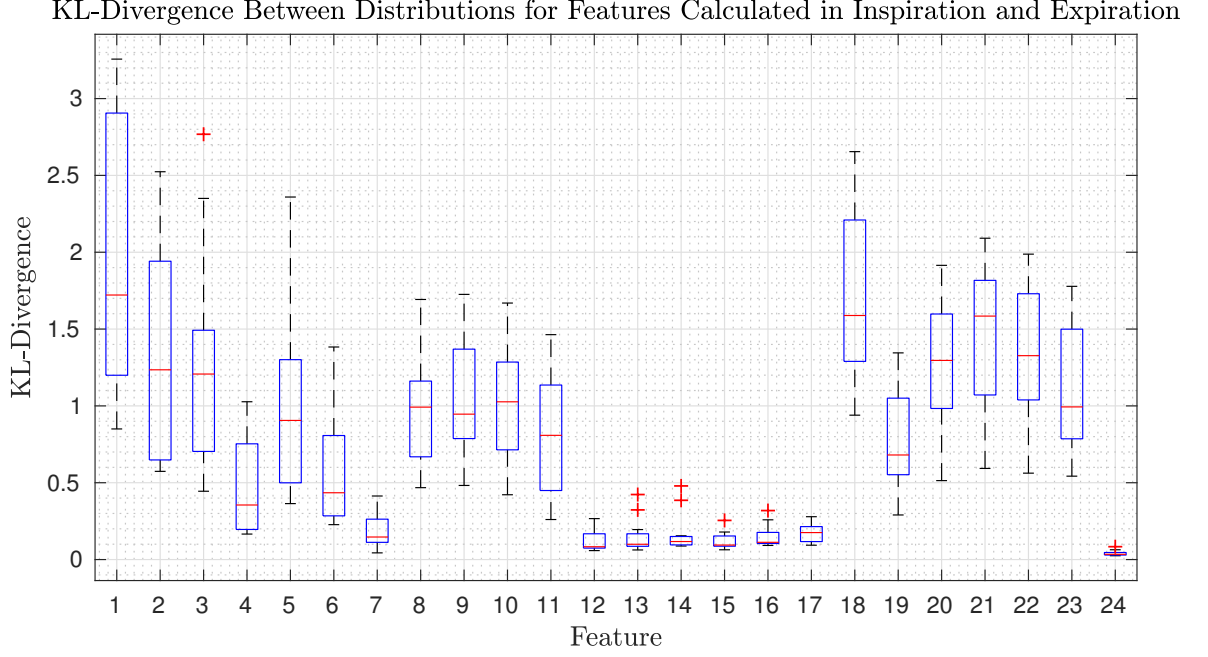discontiunity, so we decided to use a filter that counts the contiunity into account and since we had estimated the period earlier, we tried to use that for filtering purposes. Then we created a periodic square wave with the period equal to the the estimated period and 50% duty cycle. We calculated the cross correlation between this square wave and the output of neural network. This shows us the required amount of shift that is needed for the square wave to fit on to the phase of airflow. We must note that, here we assume that the airflow is periodic and its duty cycle is 50%.

4.3. Phase Estimation With Estimated Airflow Curve

We try to estimate the airflow phase curve using the estimated airflow curves. In order to do that, we used the estimation which has the most correlation with the absolute airflow curve, so we used the first AR coefficient of consecutive and overlapping windows. It is reported that the onsets can be found by looking at local minima of estimated absolute airflow curve. In our case, local minima analysis only didn't solve the problem and we developed some heuristic algorithms based on our observations.

4.3.1. Prefiltering

We try to locate the local minima, but because of the noisy nature of estimated airflow, there are many false minima. Then we applied a median and moving average filter. We chose to use a median filter to filter out the spikes and a moving average filter to smooth the signal. Describing formulae for median and moving average filters with $2K + 1$ taps are given in (4.12) and (4.13).

$$y_i = med(x_{\max(0,i-K)}, x_{\max(0,i-K+1)}, ...x_i, ..., x_{\min(L,i+K-1)}, x_{\min(L,i+K)}) \qquad (4.12)$$

$$y_i = \frac{\sum_{j=\max(0,i-K)}^{\min(L,i+K)} x_j}{\min(L, i + K) - \max(0, i - K) + 1} \qquad (4.13)$$

. A figure showing unfiltered and filtered signal is given in Figure 4.12

4.3.2. Local Minima Extraction

The number of local minima is decreased by filtering, however, it is still too much to process. So, we decided to add a heuristic algorithm to selection algorithm of local minima. We decided to use two types of local minimum, one type, let's call it *type*-1, is mostly seen in transition from expiration to inspiration and the other, *type*-2, is mostly seen in transition from inspiration to expiration. We don't expect any minima in the right neighborhood of *type*-1 minimum and in the left neighborhood of *type*-2 minimum for a certain amount of distance. The examples to *type*-1 and *type*-2 minima is shown in figure 4.13.

4.3.3. Selection of Transition Points

We use the period estimation and local minima together to decide on the location of transition points. While all the local minima are candidate transition points, we
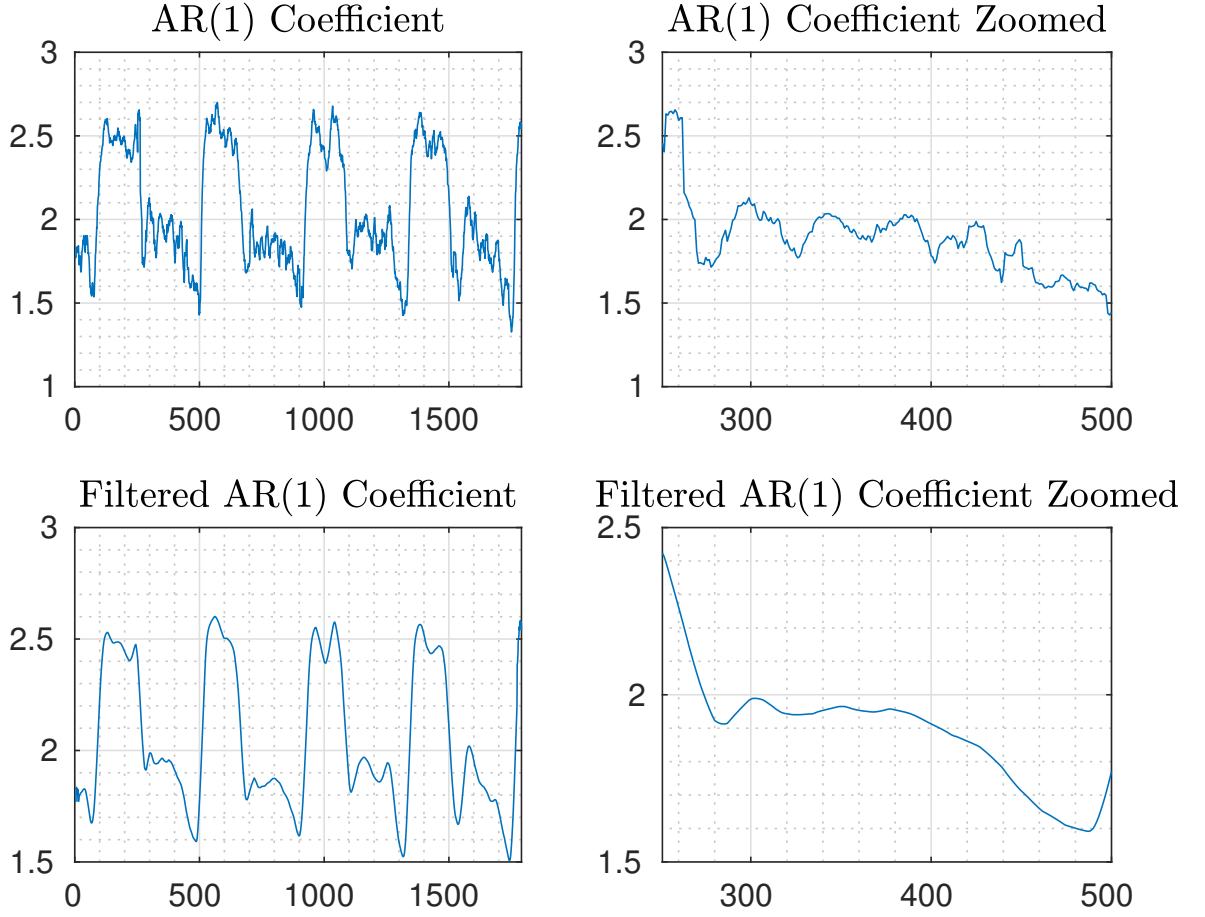
Figure 4.12. Filtering The Flow Estimate

assume that the distribution of the distances between consecutive transitions from inspiration to expiration and expiration to inspiration have a distribution with a small variance and a mean which is equal to the period estimation. So, firstly, we decide on the number of transitions in each direction. Then we list all the subsets of local minima which has the number of transition elements. Finally we calculate the likelihoods of each subset according to (4.14) and select the one which has most. After selecting the first transition points set, in order to decide on the set belonging to the other transition type, we also look at the distance to the set of first transition type. In this method we don't assume any value for the duty cycle. An example of change point estimation is given in 4.14.

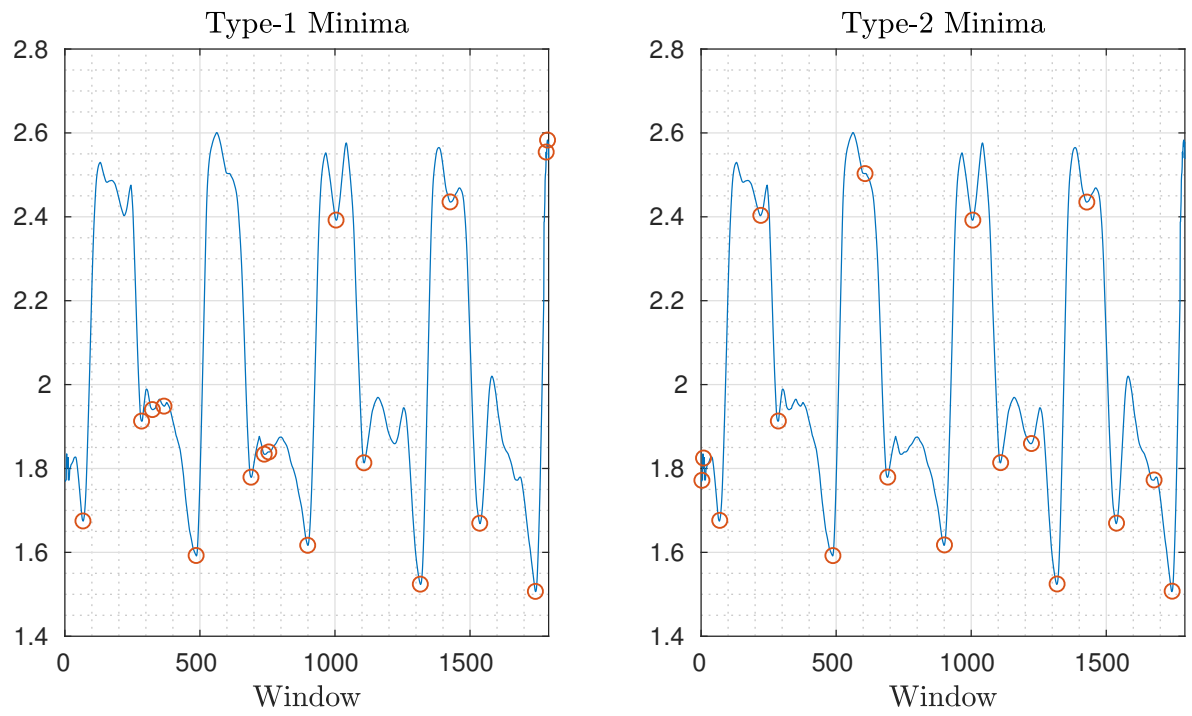$$L(\vec{d}) = \frac{\exp((\vec{d} - T')^T(\vec{d} - T')/\sigma^2)}{\sigma} \tag{4.14}$$
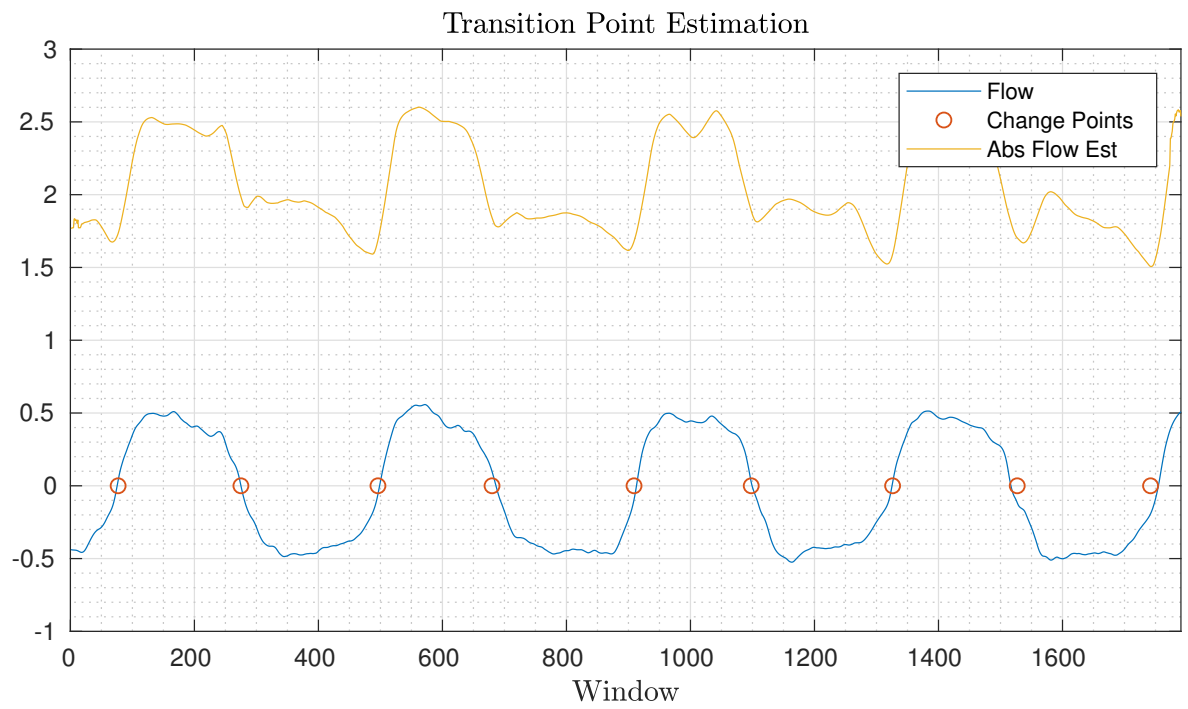
Figure 4.13. Type-1 and Type-2 Minima

Figure 4.14. Transition Point Estimation

4.3.4. Estimating The Phases Given Transition Points

We estimate the phases of segments after deciding on transition points by using the first AR coefficient of segments divided by the transitions points. We create two group and each group includes nonconsecutive segments. We then calculate the AR coefficients for each segment, and we label the group with a greater average AR coefficient as inspiration. This method works with a success rate of 96%.

4.4. Experiments & Results

4.4.1. Neural Networks

Matlab's neural network toolbox is used in training process. We selected the performance function to be cross-entropy since it is a standard [**?**]nd training function to be scaled conjugate gradient. We used neural networks with different number of hidden neurons and the correlation of neural network's output and phase is not changing significantly with the number of hidden neurons. This experiment is done on third channel of 10 subjects. The boxplot showing results for different number of hidden neurons is given in figure 4.15. So we decided to use 20 neurons. We measured the performance of neural networks by using correlation measure. The boxplot of correlation between neural network's output and the phase plot is given in 4.16. We then corrected the output of neural networks by using the period information, and the result is shown in 4.17.

. In addition to correlation measure we also measured the deviation of detected transition points from real transition points. We introduced a deadband which is equal to 5% of peak-to-peak voltage around zero before calculating the true transition points.
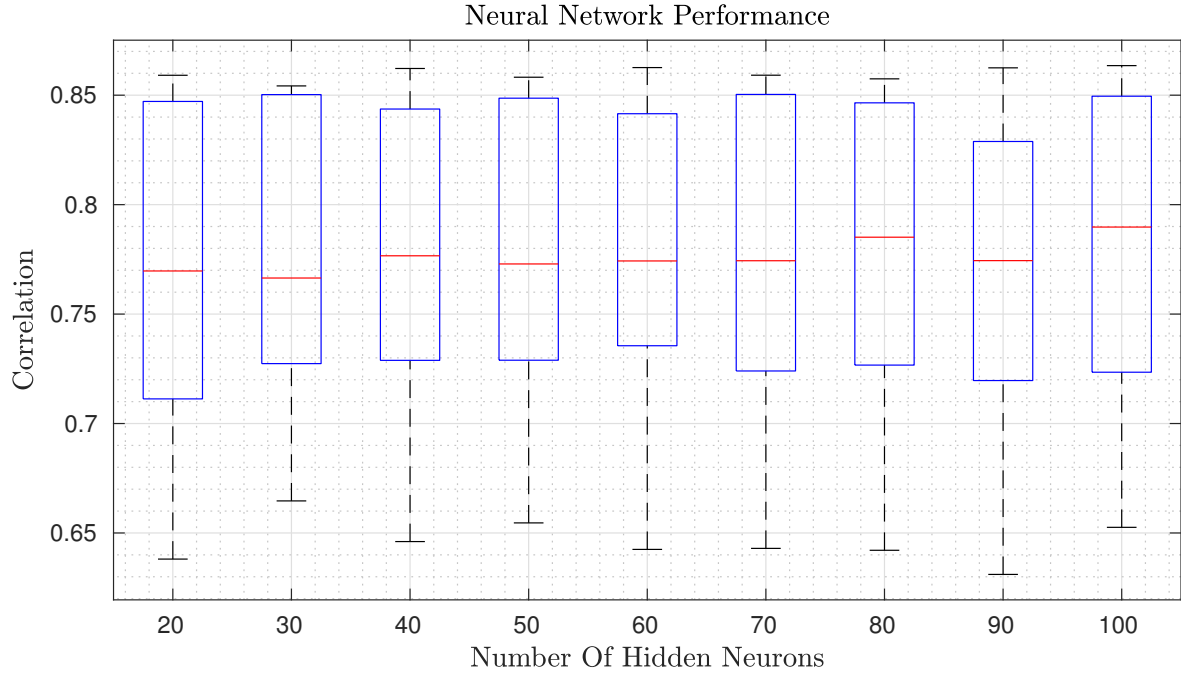
Figure 4.15. Performance of Neural Network vs Number of Hidden Neurons

Channel	1	2	3	4	5	6	7	8	9	10	11	12	Average
Ins to Exp	74	183	76	87	77	74	80	71	146	115	86	82	96.8
Exp to Ins	64	150	66	77	63	67	62	65	126	123	68	65	83.2

Table 4.1. Deviation From True Transitions in Milliseconds

4.4.2. Transition Points Detection Based On Local Minima

We used first AR coefficient and period information to estimate transition points and calculated the first AR coefficient of the segments between transition points to decide on phase. The correlation between esimated and true flow phase is given in figure 4.18 and the deviation from true transitions is given in table 4.2.

Channel	1	2	3	4	5	6	7	8	9	10	11	12	Average
Ins to Exp	112	121	130	110	96	117	159	102	88	176	110	118	119.76
Exp to Ins	106	186	94	198	113	117	157	149	104	143	131	131	130.73

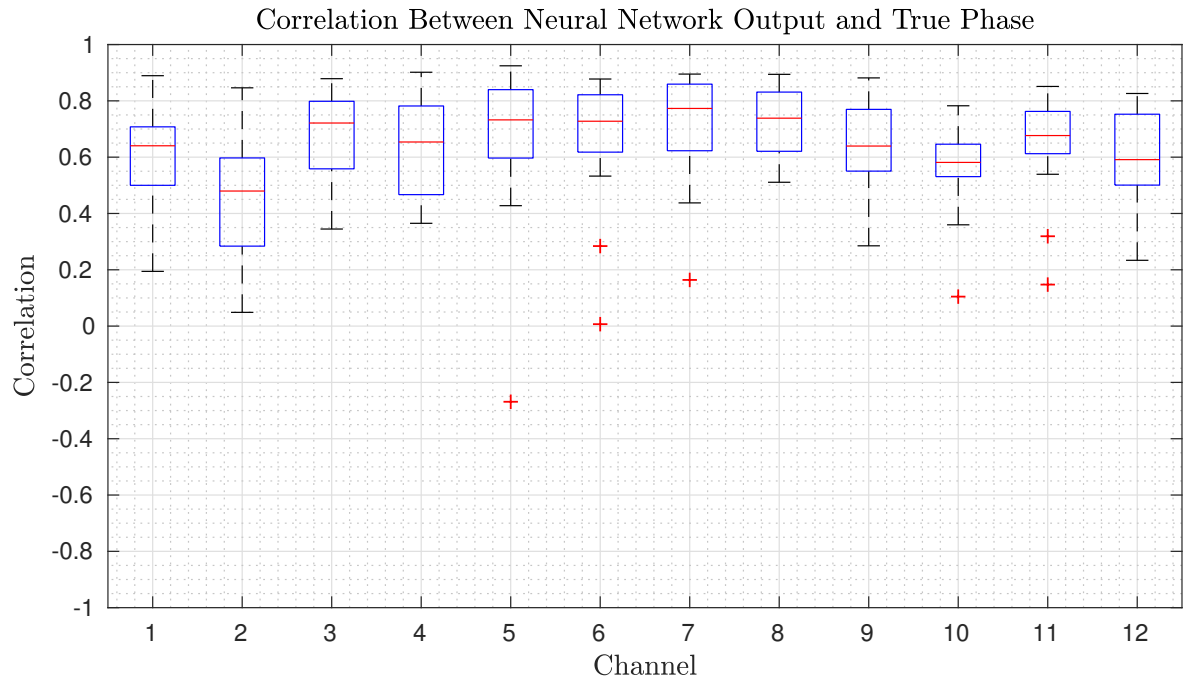Table 4.2. Deviation From True Transitions in Milliseconds

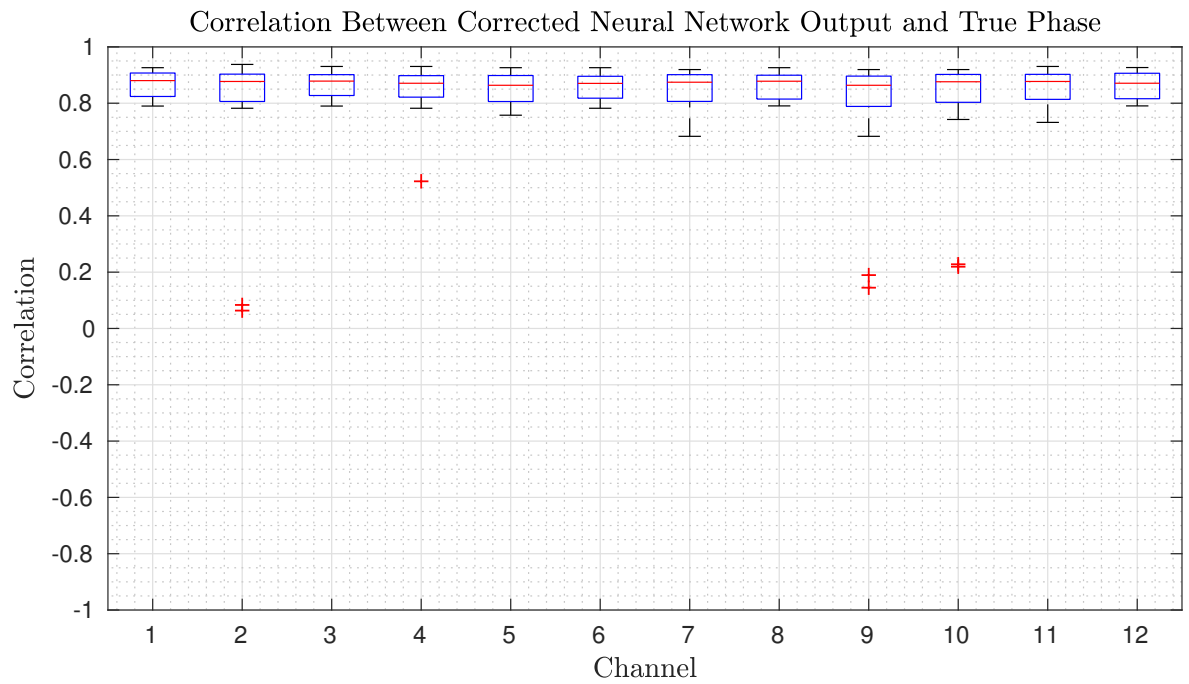Figure 4.16. Performance of Neural Network vs Channels

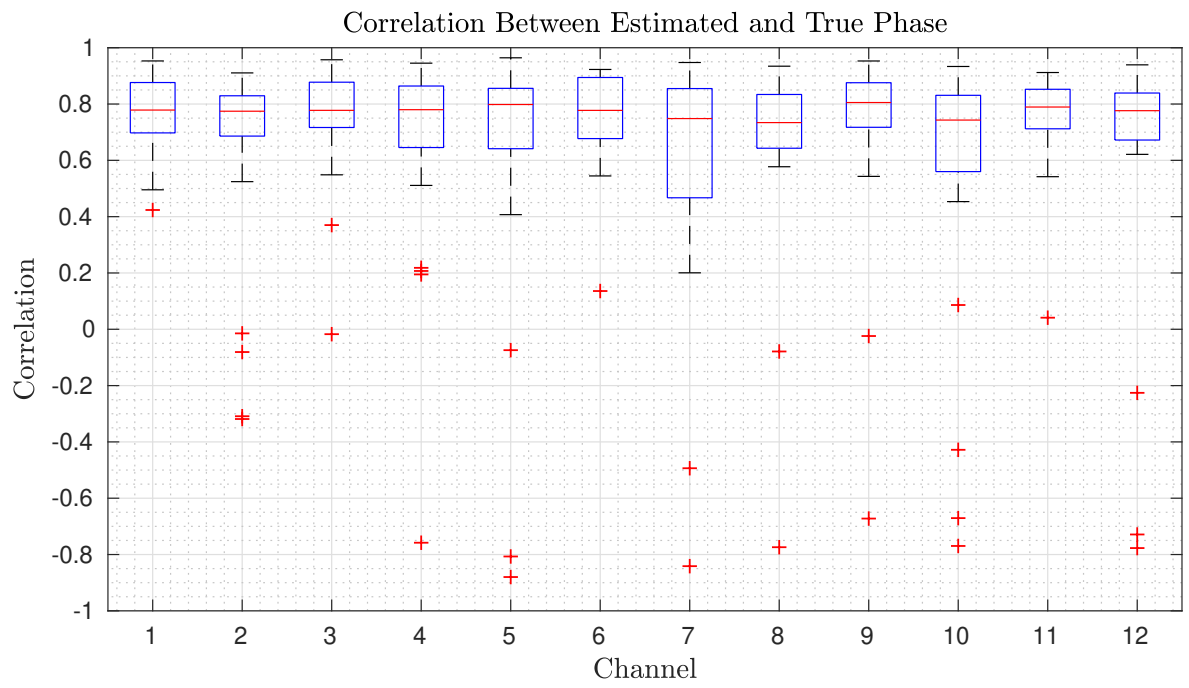Figure 4.17. Performance of Neural Network vs Channels After Correction

Figure 4.18. Performance of Phase Estimation Based On Local Minima

5. CONCLUSION

In this thesis, we suggested and tested methods, some of which were already suggested, to estimate the curve and phase of the airflow over the mouth by using the sounds recorded on chest wall. The airflow and its phase have diagnostic value, so the accurate estimation of them may increase pulmonary disease detector's performance in case there is no airflow measurement.

In chapter 3, we tested the TVAR modeling of respiratory sounds approach after we gave description of AR and TVAR processes. The method in (Koray) uses the basis functions to estimate the TVAR coefficients, we also implemented the windowing based and Kalman filter approaches and measured the performance by looking at the correlation coefficient between estimation and absolute value of airflow. From the results, it can be said that, all three approaches have similar performance. Later, we tested the correlation of magnitudes of different frequency bands with the airflow itself. Lastly, we introduced the Wiener filter approach to unify different estimations of airflow.

In chapter 4, we first gave a method to estimate the period of breathing from the estimation of airflow, which was the first AR coefficient. Then, we worked with a neural network for classification of inspiration and expiration. We presented the histograms of features (AR, Time-Frequency, Percentile Frequencies, Variance, Entropy and Kurtosis) from inspiration and expiration parts. We then presented a method to denoise the output of neural networks by using the period information and assuming a 50% duty cycle. We also try to estimate the transition points by using the absolute airflow estimation and the period information.

To sum up, the relation between respiratory sounds, their features and the airflow is explored throughout this thesis. It can be concluded that, the phase information can be extracted from respiratory sounds with a good performance by using the techniques in chapter 4 and the airflow curve can be estimated with the techniques in chapter 3

but with a less performance compared to phase estimation.

REFERENCES

1. Sen, I., *Multivariate Modeling and Diagnostic Classification of Pulmonary Sounds*, Ph.D. Thesis, Bogazici University, 2013.

2. Gavriely, N., *Breath Sounds Methodology*, CRC Press, 1995.

3. H Pasterkamp, S. K. and G. Wodicka, "Respiratory sounds. Advances beyond the stethoscope.", *American Journal of Respiratory and Critical Care Medicine*, Vol. 157, pp. 974–987, 1997.

4. Sen, I., M. Saraclar and Y. P. Kahya, "A Comparison of SVM and GMM-Based Classifier Configurations for Diagnostic Classification of Pulmonary Sounds", *IEEE Transactions on Biomedical Engineering*, Vol. 62, No. 7, pp. 1768–1776, July 2015.

5. Yeginer, M., K. Ciftci, U. Cini, I. Sen, G. Kilinc and Y. P. Kahya, "Using lung sounds in classification of pulmonary diseases according to respiratory subphases", *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 1, pp. 482–485, Sept 2004.

6. Lessard, C. S. and W. C. Wong, "Correlation of Constant Flow Rate with Frequency Spectrum of Respiratory Sounds When Measured at the Trachea", *IEEE Transactions on Biomedical Engineering*, Vol. BME-33, No. 4, pp. 461–463, April 1986.

7. Yadollahi, A. and Z. Moussavi, "Robust Respiratory Flow Estimation Using Statistical Properties of Tracheal Sounds", *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, pp. 1106–1109, Jan 2005.

8. Huq, S. and Z. Moussavi, "Acoustic breath-phase detection using tracheal breath sounds", *Medical & Biological Engineering & Computing*, Vol. 50, No. 3, pp. 297–308, Mar 2012, https://doi.org/10.1007/s11517-012-0869-9.

9. Moussavi, Z. K., M. T. Leopando, H. Pasterkamp and G. Rempel, "Computerised acoustical respiratory phase detection without airflow measurement", *Medical and Biological Engineering and Computing*, Vol. 38, No. 2, pp. 198–203, Mar 2000, `https://doi.org/10.1007/BF02344776`.

10. Golabbakhsh, M., Z. Moussavi and M. Aboofazeli, "Respiratory Flow Estimation from Tracheal Sound by Adaptive Filters", *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, pp. 4216–4219, Jan 2005.

11. Ciftci, K. and Y. P. Kahya, "Respiratory airflow estimation by time varying autoregressive modeling", *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 347–350, Aug 2008.

12. Sen, I. and Y. P. Kahya, "A Multi-Channel Device for Respiratory Sound Data Acquisition and Transient Detection", *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, pp. 6658–6661, Jan 2005.

13. Shanmugan, K. and A. M. Breipohl, *Random Signals: Detection, Estimation, and Data Analysis*, Wiley, 1988.

14. OON, T. H. J. J. V. D. H. A. M. J. L. D. H., H. S. CHOONEWELLE and H. V. D. A. Interfaculty, "Should Not Be Used for Autoregressive Modelling", .

15. Haykin, S., *Kalman Filtering and Neural Networks*, Wiley, 1988.

16. Tary, J. B., R. H. Herrera and M. van der Baan, "Time-varying autoregressive model for spectral analysis of microseismic experiments and long-period volcanic events", *Geophysical Journal International*, Vol. 196, No. 1, pp. 600–611, Jan 2014.

17. Allen, J., "Short term spectral analysis, synthesis, and modification by discrete Fourier transform", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. 25, No. 3, pp. 235–238, Jun 1977.

18. Haykin, S., *Adaptive Filter Theory*, Prentice-Hall Inc., 1996.

19. Haykin, S., *Neural Networks*, Prentice-Hall Inc., 2005.

20. Cover, T. M. and J. A. Thomas, *Elements Of Information Theory*, Wiley, 2006.

21. Aydore, S., *Wheeze Detection in Respiratory Sounds via Statistical Signal Modeling*, Master's Thesis, Bogazici University, 2009.

22. Sankur, B., Y. P. Kahya, E. Çağatay Güler and T. Engin, "Comparison of AR-based algorithms for respiratory sounds classification", *Computers in Biology and Medicine*, Vol. 24, No. 1, pp. 67 – 76, 1994, http://www.sciencedirect.com/science/article/pii/0010482594900388.