

2019-01-01

Computer-Aided Classification Of Impulse Oscillometric Measures Of Respiratory Small Airways Function In Children

Nancy Selene Avila

University of Texas at El Paso, nsavila74@hotmail.com

Follow this and additional works at: https://digitalcommons.utep.edu/open_etd



Part of the [Biomedical Commons](#), and the [Computer Sciences Commons](#)

Recommended Citation

Avila, Nancy Selene, "Computer-Aided Classification Of Impulse Oscillometric Measures Of Respiratory Small Airways Function In Children" (2019). *Open Access Theses & Dissertations*. 37.

https://digitalcommons.utep.edu/open_etd/37

This is brought to you for free and open access by DigitalCommons@UTEP. It has been accepted for inclusion in Open Access Theses & Dissertations by an authorized administrator of DigitalCommons@UTEP. For more information, please contact lweber@utep.edu.

COMPUTER-AIDED CLASSIFICATION OF IMPULSE OSCILLOMETRIC
MEASURES OF RESPIRATORY SMALL AIRWAYS FUNCTION
IN CHILDREN

NANCY SELENE AVILA RODRIGUEZ
Doctoral Program in Biomedical Engineering

APPROVED:

Vladik Kreinovich, Chair, Ph.D.

Heidi A. Taboada, Ph.D.

Ricardo von Borries, Ph.D.

Erika Meraz, Ph.D.

Nelly Gordillo, Ph.D.

Charles Ambler, Ph.D.
Dean of the Graduate School

Copyright ©

by

Nancy Selene Ávila Rodríguez

2019

to my beloved sons

JOSHUA and DEREK

COMPUTER-AIDED CLASSIFICATION OF IMPULSE OSCILLOMETRIC
MEASURES OF RESPIRATORY SMALL AIRWAYS FUNCTION
IN CHILDREN

by

NANCY SELENE AVILA RODRIGUEZ

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

Doctoral Program in Biomedical Engineering

THE UNIVERSITY OF TEXAS AT EL PASO

May 2019

Acknowledgements

Quiero expresar toda mi gratitud a mi madre Guadalupe Rodríguez que ha sido un pilar en mi vida y ha estado en los momentos más difíciles, apoyandome cada momento para poder superarme. Mamá has tenido siempre las palabras perfectas para alentarme a seguir adelante y motivarme a conseguir mis sueños, gracias por estar siempre para mí y para mis hijos, este logro es tambien tuyo.

I would also like to express my gratitude to you Liborio Nuñez, my life partner. Thank you for taking care of our sons when I had to work on my research and dissertation. Thank you for being my right hand and being there for us all the time. For your understanding, your support, and for your positive words to achieve this goal, I am infinitely grateful to you.

To my former advisor Dr. Homer Nazeran from the Department of Metallurgical, Materials and Biomedical Engineering and also the Department of Electrical and Computer Engineering, thank you for giving me the opportunity to work in this research that I have found very passionate. Thank you for sharing your knowledge about Impulse Oscillometry. This research has been a great opportunity for me to make a contribution that would potentially improve asthmatic's children lives. I really appreciate the opportunity you gave me. Thank you for your continuous support and understanding, I would have liked to finish this dissertation work together. However, I know that sometimes life gives us difficult times and it is necessary to prioritize and define what is best for our lives. I really hope that your retirement gives you the time to take care of yourself and your health. I wish you the best.

To my advisor, Dr. Vladik Kreinovich from the Department of Computer Science Department, thank you for taking over in such short notice after Dr. Nazeran's retirement. I had the opportunity to work with you for a short period of time, but this time was enough to finalize my dissertation with your advice. Thank you for being a great example of getting things done. Many thanks Dr. Kreinovich and also thanks to Dr. Julio Urenda from the

Department of Mathematics, your help was invaluable to translate all my research into a mathematical language.

To Dr. Edgar Martínez from the University of Ciudad Juárez (UACJ) for being my co-advisor for a year. Thanks to you I have learned the basics about machine learning and programming. Thank you for helping me to nail down my ideas through your advice.

Also thanks to Dr. Nelly Gordillo at the UACJ, for being my co-advisor during these years, your input and guidance have been very important to me. Your vision about research and soft computing techniques have been essential to my research. Thank you for all your time, availability and understanding. I would not have made it without you. I will always be grateful to you.

Thanks to Dr. Erika Meraz at the UACJ, for your great work in Impulse Oscillometry, without your work my research would have not been possible. Thank you for your advice, your kindness, your knowledge, and your disposition to answer all my questions.

I also wish to thank the other members of my committee, Dr. Heidi Taboada, Associate Dean for Research & Graduate Studies and Dr. Ricardo von Borries from the Department of Electrical and Computer Engineering for your open disposition and contribution to my research. A special note to Dr. Taboada who without hesitations volunteered to be part of my dissertation committee when Dr. Nazeran retired, and for all the additional guidance, support, and expertise she provided to complete this work.

Finally, I would like to acknowledge the National Council of Science and Technology of México (CONACYT) for my doctoral fellowship (No. 310901) and the Research Program on Migration and Health (PIMSA), a program funded by the University of California, Berkeley, Health Initiative of the Americas (HIA) that partially supported my research.

NOTE: This dissertation was submitted to my Dissertation Committee on the March 06, 2019.

Abstract

Computer-aided classification of respiratory small airways dysfunction is not an easy task. There is a need to develop more robust classifiers, specifically for children as the classification studies performed to date have the following limitations: 1) they include features derived from tests that are not suitable for children and 2) they cannot distinguish between mild and severe small airway dysfunction.

This dissertation describes the classification algorithms with high discriminative capacity to distinguish different levels of respiratory small airways function in children (Asthma, Small Airways Impairment, Possible Small Airways Impairment, and Normal lung function). This ability came from innovative feature selection, where features were derived from the child-friendly and reliable Impulse Oscillometry (IOS) technique. The feature selection process included deep statistical analyses and a proposed novel invariance-based pre-processing approach in the study of IOS features. The results are 100% accurate, sensitive and specific to classify normal lung function vs. small airways dysfunction; and 92%-95% accurate, 73%-100% sensitive, and 80%-100% specific for classifying a specific type of small airways dysfunction. These results are better than any of the previous computer-aided classification of small airways dysfunction results.

Table of Contents

	Page
Acknowledgements	v
Abstract	vii
Table of Contents	viii
List of Tables	xi
List of Figures	xiii
Chapter	
1 Introduction	1
1.1 Background and Significance	1
1.2 Relevance	3
1.3 Potential Social Impact	4
1.3.1 Immigration and Policy Component	4
1.3.2 Implications for Mexican Immigrants	6
2 State of the Art	7
2.1 Pulmonary Function Tests	7
2.1.1 Spirometry and Its Parameters	7
2.1.2 IOS and Its Derived Parameters	8
2.2 Computer-aided Classification Studies using IOS Features: A Review . . .	13
2.2.1 Methods	13
2.2.2 Results	14
2.2.3 Discussion	20
2.2.4 Conclusion	21
3 Formulation of the Problem, Objective and Research Questions	22
3.1 Peripheral Lung Dysfunctions	22
3.2 Diagnosis of Lung Dysfunctions	22

3.2.1	How Different Dysfunctions are Diagnosed	23
3.2.2	Children Diagnostics: A Serious Problem	23
3.2.3	Difficulty to Diagnose based on IOS Data	23
3.2.4	How IOS-Based Diagnosis could be Performed?	24
3.3	Resulting Problem	25
3.4	Research Objective	25
3.5	Research Questions	26
3.5.1	Research Questions	26
3.5.2	Hypothesis Testing	26
4	Data and Methodology	28
4.1	Data	28
4.1.1	Subjects	28
4.1.2	IOS Equipment and Data Acquisition	29
4.2	Methodology	31
4.2.1	Feature Selection	32
4.2.2	Supervised Classification	33
4.2.3	ANN Performance Evaluation and Model Selection	33
5	Feature Selection: Conventional Approach	34
5.1	IOS Data Characterization for Dimensionality Reduction	34
5.2	Results	36
5.3	Discussion	39
5.4	Conclusions	40
6	Feature Selection: Data Pre-Processing Approach	43
6.1	IOS Data Behavior	44
6.2	Typical Class Functions	47
6.2.1	Which Order Polynomials shall be Used?	48
6.2.2	Evaluating Goodness of Fit	49
6.2.3	Graphical Assessment of Cubic Functions	52

6.2.4	Typical Functions for Area and Slope	56
6.3	Similarity/Dissimilarity Measures	58
6.3.1	Similarity Measure Estimation	60
6.4	Preliminary Classification using Similarity Measures	63
6.5	Discriminative Similarity Measures	65
6.6	Conclusions	66
7	Results and Discussion	68
7.1	Baseline Classification Work	68
7.2	ANN Classification	68
7.2.1	ANN Classification - First Stage	68
7.2.2	ANN Classification – Second Stage	70
7.2.3	ANN Classification - Final Stage	72
8	Conclusions	77
8.1	Research Conclusions	77
8.1.1	Hypothesis Testing	77
8.1.2	Research Questions	79
8.2	Novelty	80
8.3	Significance of the Result- Contribution to Society	81
8.4	Future Work	81
	References	83
Appendix		
A	Polynomial Approximation Graphs of Resistance per Condition	92
B	Polynomial Approximation Graphs of Reactance per Condition	93
C	Similarity Graphs of Resistance per Condition	94
D	Similarity Graphs of Reactance per Condition	95
E	C++ Code Flow Diagram	96
	Curriculum Vitae	97

List of Tables

2.1	Studied Populations in Current Research Works of IOS Lung Function Classification	17
2.2	Results of Current Research Work of IOS Lung Function Classification	19
4.1	Study's Population Demographics.	28
4.2	Demographics for Clinician's Classification.	30
4	Mean and Coefficient of Variation of each IOS Parameter.	31
3	IOS Resistance Parameters Comparison-Matrix.	36
5.2	IOS Reactance Parameters Comparison-Matrix.	37
5.3	IOS Derived Parameters Comparison-Matrix.	37
5.4	eRIC Parameters Comparison-Matrix.	38
5.5	aRIC Parameters Comparison-Matrix.	38
5.6	Ranking of Potentially Discriminative IOS Parameters.	39
5.7	Findings of Previous Studies Related to IOS Discriminative Parameters.	41
6.1	R^2 and R^2_{adj} Results per Class.	50
6.2	R^2 and R^2_{adj} Results for the mean of the data points for each Class.	52
6.3	Similarity Measures of Typical Class Functions	62
6.4	Preliminary ANN Classification Results using Similarity Measures.	64
6.5	Resistance Similarity Measures Comparison-Matrix.	65
6.6	Reactance Similarity Measures Comparison-Matrix.	66
7.1	First Classification Results	68
7.2	ANN Classification Results for Asthma, SAI, PSAI, and Normal.	70
7.3	ANN Classification Results for Asthma vs. Normal.	71
7.4	ANN Classification Results Normal vs. Peripheral Airways Obstruction.	72

7.5	ANN Classification Results PSAI vs. Peripheral Airways Obstruction.	73
7.6	ANN Classification Results SAI vs. Asthma.	74
7.7	TP, FN, TN, and FP Table	76
7.8	Sensitivity and Specificity of ANN1 and ANN2	76
7.9	Sensitivity and Specificity of ANN3	76

List of Figures

2.1	Resistance IOS curves of Children with Peripheral Obstruction or No Obstruction (healthy).	10
2.2	Reactance IOS curves of Children with Peripheral Obstruction or No Obstruction (healthy).	11
2.3	AX and <i>Fres</i> of Children with Peripheral Obstruction (a) and Healthy Condition (b)	12
2.4	eRIC Model of the Human Respiratory System.	13
2.5	aRIC Model of the Human Respiratory System.	13
3.1	Curve Behavior of Resistance (a) and Reactance (b) per Class	24
4.1	Study Demographics by Ethnicity and Gender.	29
4.2	Study Demographics by Age.	29
4.3	Clinician's Classification Breakdown by Ethnicity.	30
4.4	Clinician's Classification Breakdown by Gender.	30
5.1	3D Plot of Resistance and Reactance vs. Frequency.	35
6.1	Resistance (a) and Reactance (b) Maximum, Middle and Minimum Patients' Curves per Class	46
6.2	Curve Behavior of Resistance (a) and Reactance (b) Average per Class	46
6.3	Cubic vs Quartic vs a Patient's Dataset	49
6.4	Cubic Resistance Functions per Class	53
6.5	Cubic Reactance Functions per Class	54
6.6	Actual Patients' Curves Maximum, Middle and Minimum vs Typical Class Function for Asthma (a), SAI (b), PSAI (c), and Normal (d)	55

6.7	Potential Geometrical Features of Resistance Curves for Asthma, SAI, PSAI, and Normal.	56
6.8	Asthmatic Child vs Typical Resistance (a) and Reactance (b) Functions	59
6.9	Similarity Measures of Resistance for Normal Class.	63
6.10	Similarity Measures of Reactance for Normal Class.	63
7.1	Final ANN Structure	74
A.1	Polynomial Regression Curves of Resistance data for (a) Asthma, (b) SAI, (c) PSAI, and (d) Normal.	92
B.1	Polynomial Regression Curves of Reactance data for (a) Asthma, (b) SAI, (c) PSAI, and (d) Normal.	93
C.1	Similarity Measures of Resistance for (a) Asthma, (b) SAI, (c) PSAI, and (d) Normal.	94
D.1	Similarity Measures of Reactance for (a) Asthma, (b) SAI, (c) PSAI, and (d) Normal.	95
E.1	C++ Code Flow Diagram	96

Chapter 1

Introduction

1.1 Background and Significance

Asthma causes airway hyper-responsiveness, inflammation, and airway obstruction, with a predominant airflow resistance at the small airways, also called peripheral or distal airways with an inner diameter of less than 2 mm [1, 2, 3, 4]. The early manifestation of these conditions prior to an asthma diagnosis could be early Small Airway Disease (SAD), also known as Small Airway Impairment (SAI), which is a chronic obstructive bronchitis with narrowing of the bronchioles and small bronchi [2, 5]. If inflammation persists during SAI, asthma will appear. Therefore, the early evaluation and treatment of the small airways might be even more effective when initiated earlier in the course of asthmatic disease [5, 6].

The timely diagnosis of asthma is a challenging task since its symptoms are similar to other respiratory conditions. Additionally, the diseases affecting the small airways are difficult to detect by traditional diagnostic tests [7]. Early childhood is a critical period to assess pulmonary function since those suffering from asthma usually face the onset of their symptoms during this time [8, 9].

Spirometry is a Pulmonary Function Test (PFT) that quantifies the volume and flow of air inhaled and exhaled as a function of time, and it is the most common PFT used by primary medical practitioners to diagnose asthma. Spirometry requires significant cooperation from patients to perform extreme inhalation and exhalation maneuvers and is often unsuitable for young children who cannot follow exact instructions. It has been demonstrated that when performing spirometry, both pre-school and school-age children have difficulty meeting some of the quality-control criteria to achieve reliable and repeatable

results [8, 9]. These challenges may have led this method to be significantly under-utilized when treating children. It is estimated that only 21% of primary care practitioners consistently use spirometry in the diagnosis of asthma in children and only 8.3% consistently use it for routine monitoring [10]. Frequently, children of this age are under-diagnosed and poorly controlled [11], creating a substantial burden that includes a decreased quality of life [11, 12, 13]. Therefore, robust quantitative techniques to determine airway obstruction to reliably diagnose and monitor asthma at an early stage are needed.

The IOS could be an alternative solution for the asthma diagnosis and control in children since it offers considerable advantages compared to spirometry. The IOS, which measures the respiratory impedance (pressure/flow as a function of frequency), is a child-friendly, noninvasive, and well validated technique requiring only the subject's passive cooperation with fast, easy, and reproducible measurements [14, 15]. The value added of the IOS over spirometry, the gold standard for pulmonary function testing, is that the IOS provides objective and reliable information about the small airways. The IOS has been found to be more sensitive and specific to detect small airway obstruction in children, since unnoticeable changes in a patient's airway function may be detected earlier than spirometry [15, 16, 17, 18, 19]. Additionally, the IOS provides objective information in cases in which spirometry has been either normal or could not have been performed [15].

The disadvantage of the IOS technique is in the interpretation of its results, as these are difficult to understand by ordinary clinicians, since the IOS is highly technical and relies on mechanical and electrical models of the human respiratory system with multiple output variables. Additionally, the dispersion of IOS testing data makes the analysis of the IOS test results difficult. These considerations hinder the broad acceptance of the IOS in the clinical field despite its objectivity and patient-friendliness. To this end, computer-aided decision systems could improve the utility of the IOS and help clinicians to strengthen the way that peripheral airway obstruction diseases (such as asthma) in children are diagnosed, monitored and controlled.

Recognizing the need to improve the clinical utility of the IOS, this dissertation focuses

on the development of intelligent computational classification algorithms to distinguish between four different degrees of small airways obstruction which are the Asthma (A), Small Airway Impairment (SAI), Possible Small Airway Impairment (PSAI), and Normal (N) respiratory conditions. This research intends to assist in the reduction of the burden that asthma causes in children by facilitating the interpretation of the impulse oscillometric test results, and providing clinicians with a reliable and proven method for accurate classification of children’s lung function for an early asthma detection, diagnosis, and control.

1.2 Relevance

Asthma is a major chronic obstructive pulmonary disease affecting around 235 million people on a global scale and it is considered the most common chronic non-communicable disease among children [11]. Children suffering from this condition are often under-diagnosed and poorly controlled, creating a substantial social and economic burden in terms of a decreased quality of life, increased school absences with work loss for parents, elevated health care expenditures, high rates of hospitalizations and emergency room visits, among others [3, 11, 12, 13, 20].

In 2016, the asthmatic population in the United States (US) was estimated to be 26 million, with children comprising about 6.1 million and a prevalence of 8.3% [21] (meaning that 8.3% of US children have asthma). Additionally, asthma is considered the 2nd most costly condition among children aged 0-17 with a total expenditure in 2012 of approximately 8.3 billion dollars [22].

In México, the prevalence of asthma in the general population of the country or in defined groups of age and sex has not been estimated [23]. However, estimates of the prevalence of asthma have been published for different areas of the country, where the estimated prevalence of asthma in children based on these studies ranges from 4.5% to 12.5% [23, 24].

In reference to our region, according to the “Texas Asthma Burden Report” published in

2014 by the Texas Department of State Health Services, the general prevalence of asthma in children in Texas was 9.1%, estimating Hispanics prevalence at 7.5%. Specifically, for the El Paso, TX region the prevalence in children was estimated at 12.3%, which was higher than the general prevalence found in Texas. In Juárez, MX, a study to assess the prevalence of asthma in schoolchildren residing in this city, determined an estimated prevalence at 6.8% [25].

Regardless of its impact, asthma remains a poorly controlled disease [12].

1.3 Potential Social Impact

Regarding immigration, several studies [26, 27, 28, 29, 30, 31, 32] have been performed to analyze the associations between asthma diagnosis and immigration variables. The findings of these studies consistently show that those who were born in the United States or had immigrated as children had a higher asthma prevalence than those who had immigrated as adults. Specifically, for Mexican-Americans living in the US, the prevalence rates of asthma results were higher in those born in the United States than in their peers born in Mexico, and the prevalence of asthma in Mexico-born immigrants increased with longer residence in the United States. [26, 27, 28, 29, 30, 31, 32]. Mexican American children who lived in the United States in their first year of life were more likely to have physician-diagnosed asthma than their peers who lived in Mexico in their first year of life [29]. Also, Mexico-born individuals who moved to the US before 2 years of age were almost twice as likely to experience asthma compared with Mexico-born children who moved to the US at the age of 2 or above [28].

1.3.1 Immigration and Policy Component

The uncertain political climate that currently exists in the United States in relation to immigration policies is discouraging for Mexican immigrants. An example of this, is the recent repeal and potential elimination of the DACA program, which has provided a tem-

porary protection from deportation for immigrant youth in the United States. This change in political climate, indicates that immigration policies are creating an environment of instability rather than one of wellness among immigrant populations. The panorama regarding health care is no different. Contrary to popular beliefs, unauthorized immigrants have very limited access to health care program due to their legal status in this country [33, 34]. Based on the current political scenario, and the immigration health care policy history, it is unlikely that a new policy will mitigate this problem anytime soon. The barriers for immigrants to obtain access to formal health care, include, but are not limited to, the risk and the fear of being arrested and deported while approaching official institutions [34]. The fear is not particular to unauthorized immigrants, but also to legal immigrants some of whom are afraid of using benefits for fear that this could hinder their ability to obtain citizenship or keep permanent residence in the future [34]. All of these barriers to health care access contribute to a significant disparity between non-citizen immigrants vs. native-born citizens.

According to U.S. Census data for 2011, 44 percent of noncitizen immigrants in the United States are uninsured, compared to 13 percent of native-born citizens. In particular, the disparity of immigrant children vs. native-born children is seen with the estimated measure of the share of the population who has made at least one office-based medical visit in a year. Ku et al indicate that 47 percent of immigrant children made a medical visit in 2010, compared to 69 percent of native-born children [33]. Fortunately, one particularly important federal initiative is the Community Health Center Program, which provides affordable non-profit primary care in clinics designed to help vulnerable populations. These clinics provide care to those in need of medical care including unauthorized immigrants since these clinics do not ask people for proof of immigration status [33, 34]. Unauthorized immigrants often use these community health centers to access health care because they have few alternatives.

1.3.2 Implications for Mexican Immigrants

Given the analysis of contemporary anti-immigrant policies discussed earlier, this research could be especially beneficial to the low-income and unauthorized Mexican immigrant populations whose access to health care is limited. Even though this research is purely technical in relation to the diagnosis and timely control of asthma, it is foreseen that the development of our software could facilitate the implementation of IOS tests in community clinics since the test would be less dependent on highly specialized doctors. Furthermore, our software could be deployed in mobile applications if future research work allows the IOS technique to be adapted and performed by mobile devices. In this manner, it would be possible to self-monitor the children's lung function for a timely treatment, contributing to the well-being of those immigrant children suffering from asthma, especially those children who are low-income and unauthorized immigrants. This research could potentially reduce the health disparities that could be caused by the limited health care access, by allowing immigrants and their children better access to monitoring asthma in a timely fashion, before complications that merit emergency health care arise.

The present research could be very relevant in terms of immigration, especially for those individuals immigrating at an early age and whose asthma risk increases when immigrating to the United States. This research could help in the diagnosis and control of asthma not only among immigrant children but also for the second generation of immigrants that are expected to assimilate the effects of asthma prevalence at the host country. Since this research intends to provide an early diagnosis to improve the quality of life of those suffering from asthma, it can help to reduce the disparity and the burden that asthma currently causes not only for immigrants, but also for the asthma children population in general.

Chapter 2

State of the Art

This chapter presents a methodological review of research works related to the computer-aided classification of peripheral airway obstruction using the IOS technique, which is focused on but not limited to asthmatic children. The main objective here is to understand the current computer-aided classification research work on this topic. In addition, we intend to provide an overview of the most common PFT features used by different classifiers used in the selected articles for this review, which include: spirometry, IOS, and respiratory model parameters derived from the IOS.

2.1 Pulmonary Function Tests

2.1.1 Spirometry and Its Parameters

Spirometry is a pulmonary function test (PFT) that measures lung function by quantifying the volume and flow of air inhaled and exhaled as a function of time. In other words, it determines how much air can be inhaled and exhaled by an individual, and how fast [35, 36]. This test is highly dependent on patient cooperation since extreme maneuvers such as a maximal forced exhalation after a maximum deep inspiration are required. Spirometry is the PFT most commonly used for the diagnosis and management of chronic obstructive respiratory diseases such as asthma as well as chronic obstructive pulmonary disease (COPD) [1, 36]. Forced vital capacity (FVC), forced expiratory volume in 1 second (FEV_1), and the FEV_1/FVC ratio are the most important spirometric indices used to assess lung function. These parameters are compared to predicted values based upon demographic data (age,

height, sex, and ethnicity) to determine the severity of airway obstruction [35].

Forced Vital Capacity (FVC)

The FVC is the total volume of air that the patient can forcefully exhale in one breath. To obtain this measurement, the patient has to perform a spirometric maneuver, namely, to deeply inhale air. Then the FVC is obtained when the patient exhales as long and as forcefully as possible. Values of FVC are measured in liters and are also expressed as a percentage of the predicted values for that individual [35, 36].

Forced Expiratory Volume in 1 Second (FEV_1)

The FEV_1 is the amount of air exhaled during the first second of the FVC maneuver. In obstructive airway diseases (such as asthma), the FEV_1 value tends to be reduced. Values of FEV_1 are measured in liters and are also expressed as a percentage of the predicted values for an individual [35, 36].

FEV_1 /FVC ratio

The FEV_1 /FVC is the ratio of FEV_1 to FVC expressed as a fraction and is used to determine if the pattern is obstructive, restrictive, or normal. A reduced FEV_1 /FVC ratio indicates airflow limitation. In children, an FEV_1 /FVC ratio below 0.90 represents airway obstruction. In adults, the FEV_1 /FVC ratio is normally greater than 0.75; smaller values suggest airflow limitation [35, 36, 37, 38].

2.1.2 IOS and Its Derived Parameters

The IOS is based on the forced oscillation technique (FOT), which measures the mechanical properties of the lungs by using the noninvasive superimposition of air pressure fluctuations applied to the airways over the subject's normal breathing. An essential aspect to be understood is that the IOS graphically displays frequency-dependent curves that are of

the utmost importance in the diagnosis of distal obstruction, because changes in lung function are evaluated based on the visual interpretation of their shape and magnitude [15]. Consequently, the interpretation of IOS results must be performed by trained physicians who understand the specifics of the test.

Impedance

The IOS uses sound waves to calculate the respiratory impedance (Z) by injecting short impulses of air pressure into the mouth and measuring the consequent air flow at the mouth. Periodic brief pulses (rectangular waves) of pressure oscillations are applied with a fixed period of 200 msec from which a range of frequencies of interest is derived (at 5, 10, 15, 20, 25, and 35 Hz). The respiratory (pulmonary) impedance is the sum of all the *resistive* and *reactive* forces that oppose the pressure impulses (oscillations) and are calculated from the ratio of pressure and flow at each frequency [17]. The components of impedance (Z) are: resistance (R) and reactance (X).

Resistance

Resistance is the in-phase component of respiratory impedance and reflects information about the forward pressure of the conducting airways [17]. The resistance at 5 Hz (R_5) represents the total airway resistance, and the resistance at 20 Hz (R_{20}) represents the resistance of the large airways. The small airways resistance could be calculated by subtracting R_{20} from R_5 ($R_5 - R_{20}$). The larger the difference between $R_5 - R_{20}$ is, the greater the peripheral airways resistance is. $R_5 - R_{20}$ is also considered an IOS-derived parameter known as the frequency-dependence of resistance (fdR) [17, 39]. For instance, resistance is independent of the frequency in healthy subjects, while for small airway obstruction, the resistance at lower frequencies increases, but it is unchanged at higher frequencies that do not reach the small airways [40]. In other words, distal obstructive diseases such as asthma result in a frequency-dependent increase in resistance at lower frequencies, because the pressure signal wave propagating out to the lung periphery (R_5) encounters greater re-

sistance than the more proximal higher-frequency (R20) impulse [15]. Figure 2.1 shows the resistance IOS curves of children with peripheral obstruction and no obstruction (healthy).

Reactance

Reactance is the out-of-phase component of respiratory impedance and reflects the capacitive (C) and inertive (I) properties of the airways. Reactance can be viewed as the rebound resistance, or an echo, giving information about the distensible airways [17]. Unlike resistance, reactance is always frequency-dependent. In healthy individuals, the absolute value (magnitude) of the reactance is smaller, while in individuals with peripheral obstruction the absolute value of the reactance increases, making the reactance more negative. Specifically, the difference in the reactance magnitude between healthy and individuals with peripheral obstruction is more pronounced at the lower frequencies, where the capacitive properties of the lungs dominate and in which the small airways play an important role as storing capacity energy primarily relies on the periphery of the lungs [15, 17, 40]. Figure 2.2 shows the reactance IOS curves of children with peripheral obstruction and no obstruction (healthy).

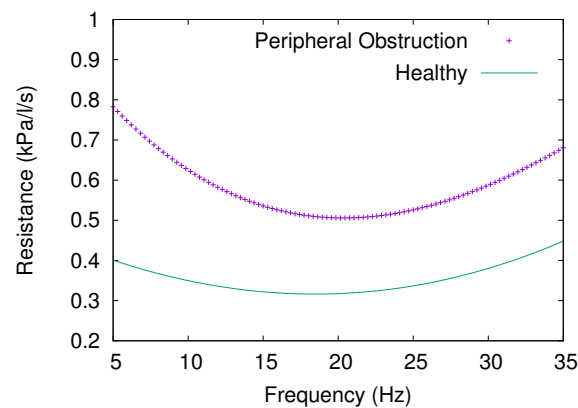


Figure 2.1: Resistance IOS curves of Children with Peripheral Obstruction or No Obstruction (healthy).

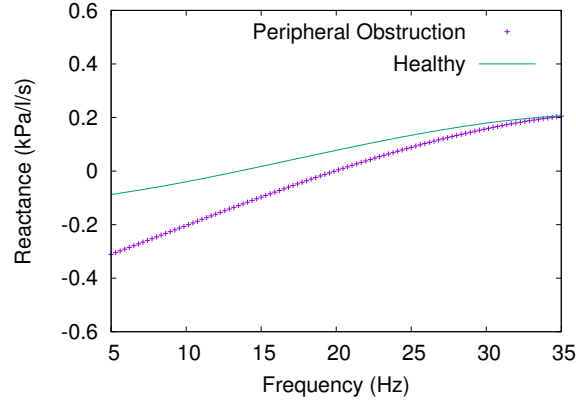


Figure 2.2: Reactance IOS curves of Children with Peripheral Obstruction or No Obstruction (healthy).

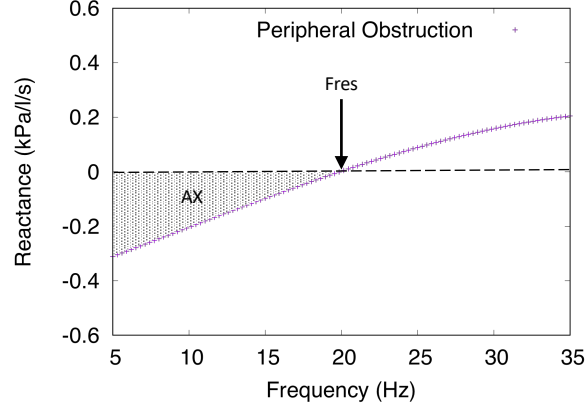
Resonant Frequency and Reactance Area

The resonant frequency (F_{res}) is the frequency at which the reactance's components: C and I have the same magnitude, since the I component represents a positive value and capacitance C represents a negative value, with the value of reactance at F_{res} equal to zero. F_{res} separates the low-frequency from high-frequency impedance. Respiratory restrictive and obstructive conditions cause the F_{res} value to be increased; moreover, F_{res} is usually higher in children and decreases with age [17, 39].

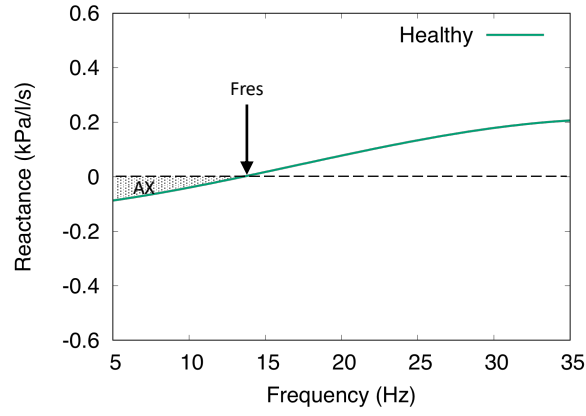
The reactance area (AX), also known as the “Goldman triangle”, is the area under the curve between 5 Hz and F_{res} (please refer to Figure 2.3). This is a single-magnitude parameter (an index), which integrates the reactance at all frequencies where the capacitance (elasticity of the airways) dominates the inertance. As with reactance at low frequencies, the AX parameter provides important information about small airways obstruction [17, 39].

eRIC and aRIC Respiratory Model Parameters

The extended Resistance Inductance Capacitance (eRIC) and the augmented RIC (aRIC) are equivalent electrical circuit models for the human respiratory system impedance which are derived from IOS parameters. are physiologically realistic and in line with the expected



(a) Child with Peripheral Obstruction



(b) Child with Healthy Condition

Figure 2.3: AX and F_{res} of Children with Peripheral Obstruction (a) and Healthy Condition (b)

values in normal subjects and those suffering from pulmonary dysfunction [39, 41, 42].

The eRIC model is an improvement of the RIC model and its components include the representation of large airway resistance R , peripheral resistance R_p , large airway inertance I , and peripheral airway compliance C_p . Specifically, the added peripheral resistance R_p represents the small airways in the respiratory system which could not be captured by the previously developed RIC models.

The aRIC model was developed and validated as an augmentation of the eRIC model. The additional element C_e in the aRIC model represents the extrathoracic compliance mainly due to the upper airways shunt effects [39, 41, 42]. Figures 2.4 and 2.5 show the enhanced and validated IOS-based equivalent electric circuit models of the human

respiratory system.

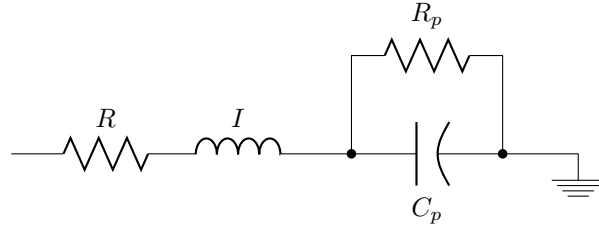


Figure 2.4: eRIC Model of the Human Respiratory System.

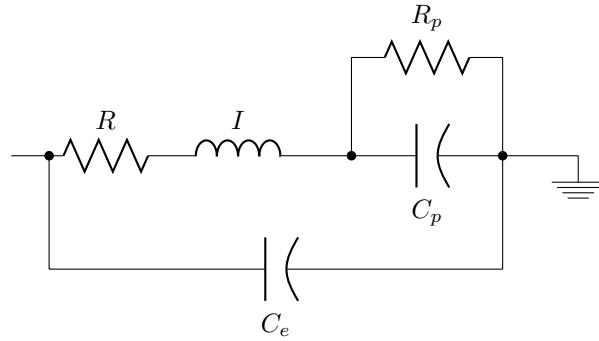


Figure 2.5: aRIC Model of the Human Respiratory System.

2.2 Computer-aided Classification Studies using IOS Features: A Review

2.2.1 Methods

The literature review was performed using the following parameters and scientific databases: “All fields” in PubMed, “Full-Text & Metadata” in IEEE Xplore, and “All Databases” in Web of Knowledge. The search included all languages, however, articles written in a different language than English and Spanish were discarded. The words and logic operators used for the search in each of the databases were “asthma” AND “classification” AND “oscillometry”. Additional searches were performed using the same methodology, but substituting in first place the word “asthma” by the phrase “small airways”, the next search was done

by replacing “small airways” by “peripheral airways”, and in the last search “peripheral airways” was replaced by “distal airways”. The last search performed for this review was completed on June 29, 2018. A total of 34 articles were found by the search. The title and abstract of these articles were screened and selected based on the following eligibility criteria: 1) Publications that focused on the computer-aided classification of asthma or small airway impairment, 2) Computer-aided classification that included impulse oscillometric features. The bibliography of the selected articles was also screened to find other relevant articles.

2.2.2 Results

Out of the 34 articles identified using scientific web databases, only 7 met the eligibility criteria, and an additional article was found through the screening of selected articles’ bibliography for a total of 8 articles.

Based on the selected articles, the most recent research efforts regarding computer-aided classification of lung function using IOS features have been conducted by Badnjevic et al. [43, 44, 45, 46, 47], Hafezi et al. [48], and Barua et al. [49, 50].

Studied conditions, population, classifiers’ input features and soft-computing techniques used in the different classification studies are summarized in Tables 2.1 and 2.2 and further described in the upcoming sections.

Studied Conditions

In the selected articles, the conditions studied are asthma, COPD, small airway impairment (SAI), possible small airway impairment (PSAI), central obstruction, peripheral and healthy conditions. Even though peripheral obstruction related to asthma and SAI is of our main interest, the results of classification of COPD which is a respiratory condition characterized by chronic airflow obstruction mainly due to long-term smoking [51, 52], are presented because in the selected study [46] this condition was investigated along with

asthma. COPD is not part of the scope of this review; therefore, the works presented in here do not represent all computer-aided classification works using IOS features for this condition.

Most of the Badnjevic et al. studies focused on a bi-class classification which included the asthma and healthy conditions [43, 44, 45, 47]. There was only one article from Badnjevic et al. that studied three classes, in this article COPD was studied along with the asthma and healthy conditions[46].

The classification efforts conducted by the University of Texas at El Paso (UTEP) research group included the work performed by Barua et al. and Hafezi et al. The first two efforts were conducted by Barua et al. in 2004 and 2005 [49, 50]. In both cases, the studies differentiated between 2 classes. In the 2004 classifier, the classes studied were central and peripheral diseases, while the classification in 2005 included asthmatic and non-asthmatics. The last classification effort by UTEP’s researchers was conducted in 2009 by N. Hafezi et al. [48], the classes studied were the asthma, SAI, PSAI, and healthy conditions.

Studied Populations

Regarding age, Badnjevic et al. in [45] and [46] studied asthmatic subjects with an average age of 19.85 years \pm standard deviation (SD) 8.18, while normal subjects were in average 30.03 years \pm SD 11.83. In addition, these researchers in [46] studied COPD patients with an average age of 52.25 years \pm SD 7.636. The other studies performed by Badnjevic et al. [43, 44, 47] did not provide information regarding age. In [48] Hafezi et al. studied children from 5 to 17 years old; while the studies performed by Barua et al. included children from 2 to 8 years old in [49] and in [50] the subjects ages ranged from 13 to 85 years old. In summary, from the demographics provided by the different studies, only two papers [48] and [49] addressed early childhood, which is considered a critical period to assess pulmonary function since those suffering from asthma usually face the onset of their symptoms during this time [8, 9]. Three papers [45, 46, 50] included subjects in a later childhood stage and adulthood, while the three remaining articles [43, 44, 47] did not

provide information about age.

The articles that reported information about gender were the studies performed by Badnjevic et al. [43, 44, 45, 46] and Barua et al. [49, 50]. These papers included a balanced female and male population, except for [49], where females constituted 66.8% of the study population while males represented 33.2%. The remaining two studies [47] and [48] did not include gender information.

The papers published by Barua et al. [49, 50] are the ones that provided more specifics about the demographics of the studied population. These papers also included subject's information about height and weight. In [49], the subject's height and weight ranged from 0.88 to 1.4 meters and 12 to 32.7 kilograms respectively, and in [50] from 1.4 to 1.85 meters and 35 to 176 kilograms, respectively.

A summary of the studied population in each of the selected articles is presented in Table 2.1.

Input Features and Computer-aided Classification Techniques

In the selected articles, different spirometric and IOS parameters were used as input features for the classifiers presented by the authors. These input features were derived from what are called static and/or dynamic assessments. A static assessment is based solely on the first PFT measurements (IOS and/or spirometry). On the other hand, a dynamic assessment of the patient takes into consideration the application of bronchial dilation (BDT) and bronchial provocation (BPT) tests. After BDT and/or BPT treatment and after the second and/or third measurements of pulmonary function, potential changes of lung parameter values are evaluated, from which physicians get accurate information on the specifics of the disease [43, 44, 45, 46].

In the studies reviewed, Badnjevic et al. were the only authors who performed a dynamic assessment in [43, 44, 45, 46, 47] when inconclusive results were obtained with the static assessment. The common input features used in the studies by Bandjevic et al. included $R5$, $R20$, $R5-R20$, $X5$ and $Fres$ from IOS; FVC , FEV_1 , FEV_1/FVC from spirometry, along

Table 2.1: Studied Populations in Current Research Works of IOS Lung Function Classification

Author	Year	Reference	Conditions Studied	Number of Subjects (N)	N per Gender	Age (Years)	Height (m)	Weight (Kg)
A. Badnjević et al	2016	[24]	Asthma & Healthy	N=1250. Asthma: 728 Healthy: 522	Male: 601 Female: 649	Not reported	Not reported	Not reported
A. Badnjević et al	2016	[25]	Asthma & Healthy	N=1250. Asthma: 728 Healthy: 522	Male: 601 Female: 649	Not reported	Not reported	Not reported
A. Badnjević et al	2015	[26]	Asthma & Healthy	N=289 Asthma: 72 Healthy: 217	Male: 142 Female: 147	Asthma : 19.85 +/- SD 8.18 Healthy: 30.03 +/- SD 11.83.	Not reported	Not reported
A. Badnjević et al	2015	[27]	Asthma, COPD & Healthy	N= 455 Asthma: 170 COPD: 248 Healthy: 37	Male: 244 Female: 211	Asthma : 19.85 +/- SD 8.18 COPD: 52.25 +/- SD 7.636 Healthy: 30.03 +/- SD 11.83.	Not reported	Not reported
A. Badnjević et al	2013	[28]	Asthma & Healthy	N=156 Asthma: 72 Healthy: 84	Not reported	Not reported	Not reported	Not reported
Nazila Hafezi et al	2009	[29]	Asthma, SAI, Mild SAI & Healthy	N=112	Not reported	5-17	Not reported	Not reported
Barúa, Miroslava et al	2005	[30]	Asthmatic Constricted & Asthmatic Non-Constricted	N= 361 IOS patterns from 41 subjects. Constricted: 168 Non-constricted: 193	Male: 120 Female: 241	2-8	0.88-1.4	12-32.7
Barúa, Miroslava et al	2004	[31]	Central & Peripheral Diseases	N=131	Male : 64 Female: 67	13-85	1.4 - 1.85	35 - 176

with symptoms and allergy history features. In addition to these features, in [43] Bandjevic et al. outlined the use of the PEF spirometric parameter as an input feature and in [45], they included the utilization of body plethysmographic features, while in [44] no specifics about the input parameters used were provided. In these studies, the classification was performed using algorithms based on Artificial Neural Networks (ANN) [43], Fuzzy logic [44] or Neuro-Fuzzy techniques [45, 46, 47], and, as previously mentioned, most of the classifiers were bi-class.

Studies conducted by N. Hafezi et al. and Barua et al. were based on the static assessment of IOS derived features. In [48], N. Hafezi et al. developed two Neuro-Fuzzy

multi-class classifiers using the following IOS parameters: R5-R15, AX, and extended parameters from the eRIC model (R , R_p , I , and C_p), or augmented parameters from the aRIC model (R , R_p , I , C_p , and C_e). In this publication, four different degrees of peripheral obstruction were assessed and the reported results only showed the performance of the classifier on the training data and no validation with unseen data was performed.

On the other hand, Barua et al. studies [49, 50] were based on ANN algorithms where the input features used for the bi-class classification were the IOS parameters R5, R10, R15, R20, R25, R35, and X5, X10, X15, X20, X25, X35 along with demographic features of age, gender, height and weight; and for the particular case of study [50] smoking status features were considered.

The selected studies were based on a supervised training approach for the learning algorithms.

Classification Performance

The studies performed by Badnjevic et al. achieved the best classification results in the analysis presented in Table 2.2. In their articles, the best accuracy results were achieved when both static and dynamic assessment features were used, producing overall classifier's accuracy results ranging from 92.3% to 99.34% in the different studies performed. The best result achieved was in [46], where a neuro-fuzzy algorithm was used to classify asthma, COPD and healthy conditions. It is important to mention that Badnjevic et al. did not present the overall classifier's accuracy results in their articles, instead, they presented the classification performance per class for both static and dynamic assessment approaches. As the authors provided all the information required to calculate the overall classifier's accuracy, we calculated it to be able to compare the different studies, Table 2.2 outlines the condition-specific accuracy results reported by Badnjevic et al. as well as the overall classifier's accuracy calculated by us for both static and dynamic assessment approaches.

As mentioned previously, Badnjevic et al. achieved the best results after performing static and then dynamic testing. However, in 3 out of 5 papers, these authors also reported

Table 2.2: Results of Current Research Work of IOS Lung Function Classification

Author	Year	Ref.	Conditions Studied	Diagnostic Techniques Used	Assessment Type	Input Parameters Used	Classification Technique	Accuracy by Condition after Static Assessment			Accuracy by Condition after Static & Dynamic Assessments			Overall Classifier's Accuracy after:		Sensitivity	Specificity
								Asthma	COPD	Healthy	Asthma	COPD	Healthy	Static Assessment	Static & Dynamic Assessments		
A. Badnjević et al	2016	[24]	Asthma & Healthy	IOS, SPIR, BDT, BPT	Static & Dynamic	Symptoms IOS: not specified SPIR: not specified	ANN	Not reported	Not reported	Not reported	97.11%	N/A	98.85%	Not reported	97.84%	97.11%	98.85%
A. Badnjević et al	2016	[25]	Asthma & Healthy	IOS, SPIR, BDT, BPT	Static & Dynamic	Symptoms IOS: R5, R20, X5, R5-R20, Fes SPIR: FVC, FEV1, FEV1/FVC, PEF	Fuzzy Logic	8.65 % (63/728)	N/A	89.08% (465/522)	91.89%	N/A	95.01%	42.24%	93.20%	N/A	N/A
A. Badnjević et al	2015	[26]	Asthma & Healthy	IOS, SPIR, BDT, BPT	Static & Dynamic	Symptoms and allergy history IOS: R5, R20, R5-R20, X5, Fes SPIR: FEV1, FEV1/FVC Body Plethysmography	Neuro-fuzzy	11.43% (8/72)	N/A	Not reported	97.22%	N/A	98.61%	Could not be estimated with the information reported	98.20%	98.15%	98.62%
A. Badnjević et al	2015	[27]	Asthma, COPD & Healthy	IOS, SPIR, BDT, BPT	Static & Dynamic	Symptoms IOS: R5, R20, R5-R20, X5, Fes SPIR: FEV1, FVC, FEV1/FVC	Neuro-fuzzy	87.65% (149/170)	85.5% (212/248)	Not reported	99.41%	99.19%	100%	**86.3%	99.34%	99.28 % (For Asthma & COPD only)	100% (For Asthma & COPD)
A. Badnjević et al	2013	[28]	Asthma & Healthy	IOS, SPIR, BDT, BPT	Static & Dynamic	Symptoms, allergies and risk factors IOS: R5, R20, R5-R20, X5, Fes SPIR: FEV1, FVC, FEV1/FVC	Neuro-fuzzy	10.70%	N/A	93.67%	90.25%	N/A	94.04%	51.92%	92.30%	N/A	N/A
Nazila Hafezi et al	2009	[29]	Asthma, SAI, Mild SAI & Healthy	IOS	Static	IOS: R5-R15, AX eRIC (R, Rp, I, Cp).	Co-Active Neuro-fuzzy Inference System (CANFIS)	Not reported	N/A	Not reported	N/A	N/A	N/A	*95.54%	N/A	N/A	N/A
						IOS: R5-R15, AX, aRIC (R, Rp, I, Cp, Ce)		Not reported	N/A	Not reported	N/A	N/A	N/A	*97.32%	N/A	N/A	N/A
Barúa, Miroslava et al	2005	[30]	Asthmatic Constricted & Asthmatic Non-Constricted	IOS	Static	IOS: R5, R10, R15, R20, R25, R35, X5, X10, X15, X20, X25, X35 General: Age, gender, height, weight.	ANN	Not reported	N/A	Not reported	N/A	N/A	N/A	98.61%	N/A	N/A	N/A
Barúa, Miroslava et al	2004	[31]	Central & Peripheral Diseases	IOS	Static	IOS: R5, R10, R15, R20, R25, R35, X5, X10, X15, X20, X25, X35 General: Smoking status, age, gender, height, weight.	ANN	Not reported	N/A	Not reported	N/A	N/A	N/A	61.53%	N/A	N/A	N/A

* Accuracy of training data. No validation results available.

** Only taking into consideration the Asthmatic and COPD populations as Healthy results were not reported by the authors

the classification results after performing only static testing. For comparative purposes, this is very useful since the rest of the studies involved in this review solely used input variables derived from static data. To this end, the overall classifier's accuracy results for Badnjevic et al. studies after the static assessment ranged between 42.24% to 86.3%; while Nafezi et al. obtained overall accuracy results on training data (no validation testing was reported) were 95.54% and 97.32% for the two classifiers described in Nafezi's article [48]. Lastly, the overall classification accuracy obtained by Barua et al. from unseen data in 2004 [50] and 2005 [49] were 61.53% and 98.61%, respectively.

Badnjevic et al. were the only authors that reported results in three of their articles

in terms of sensitivity and specificity [43, 45, 46]. In these papers, the results obtained ranged from 97.11% to 99.28% for the classifier’s sensitivity and from 98.62% to 100% for specificity. The best classifier results were reported in article [46], this article studied three conditions: asthma, COPD and healthy subjects. However, in the calculations for both sensitivity and specificity, only the results from asthma and COPD were considered. Therefore, the overall performance in terms of these two metrics is unknown as healthy subjects results were not reported.

2.2.3 Discussion

Different computer-aided classification efforts have been conducted to improve the diagnosis of peripheral obstruction conditions using IOS features. However, these studies are limited regarding the number of publications. According to our research criteria, only eight papers were found published from three different main authors.

Most of the papers reviewed provided limited information about the demographics of the studied population. An important aspect is to understand the demographics behind each of the classification studies, as they could play an important role in the performance of the IOS-acquired values. Particularly, height has been found to strongly correlate with different IOS variables [53]. Therefore, a comprehensive assessment of discriminative anthropometric variables could potentially help to obtain better computer-aided classification results. Moreover, early childhood is considered a critical period to assess pulmonary function since those suffering from asthma usually face the onset of their symptoms during this time [8, 9]. Additionally, the diagnostic utility of the IOS has a greater impact on children population, as the use of spirometry for this population could provide subjective and unreliable results. To this end, only two articles [48] and [49] addressed the asthmatic children’s population. Even though Badnjevic et al. studies were not specifically focused in children, we need to emphasize that the dynamic assessment used by these authors to obtain the classifier’s input features, especially the one based on the BPT is unsuitable for an important asthma population which is young children. The BPT testing is phys-

ically demanding and depends on the patient’s ability to perform acceptable spirometric maneuvers, and the patient’s fatigue to perform repetitive spirometry testing [54, 55].

The areas of opportunities for improvement in the reviewed papers relate to the selection of the relevant features. As supervised classification was used by all authors, the use of discriminative parameters is an important aspect that impacts the classifier’s computational and quality performance. To this end, none of the studies presented in this review provided a rationale or a method for feature selection that explained why the input features used were chosen or considered relevant.

Additionally, to address both computational and clinical points of view, it is highly desirable to use the minimum number of PFT sources for input feature acquisition to perform classification. To this end, only three papers used solely IOS-derived features as the classifier’s inputs while five used a combination of input variables from different sources: IOS and spirometry derived from static and dynamic testing. In view of the difficulty of using spirometry for children, the use of one source of input variables is more desirable.

2.2.4 Conclusion

In conclusion, there is still a great opportunity to improve the utility of IOS by developing more robust classifiers, specifically for the children asthmatic population as the classification studies performed to date have the following limitations: 1) they are very few in number, 2) they include features derived from tests that are not optimally suitable for children; 3) they are solely bi-class (mostly asthma and non-asthma) and do not have the ability to distinguish between mild and severe small airway dysfunction; and 4) they produce modest results or lack the validation testing using unseen data for the multi-class classification of the different degrees of peripheral airway obstruction.

Chapter 3

Formulation of the Problem, Objective and Research Questions

3.1 Peripheral Lung Dysfunctions

One of the major lung dysfunctions is asthma, a long-term inflammatory disease of the airways of the lungs [11].

Asthma may be preceded by Small Airway Impairment (SAI), a chronic obstructive bronchitis. If inflammation persists during SAI, it could cause asthma.

SAI, in its turn, may be preceded by a less severe condition called Possible Small Airways Impairment (PSAI).

3.2 Diagnosis of Lung Dysfunctions

Diagnosis of different lung dysfunctions is difficult but important. All lung dysfunctions lead to similar symptoms like wheezing, coughing, etc. As a result, it is difficult to distinguish between these dysfunctions – and it is also difficult to distinguish these chronic dysfunctions from a common short-term respiratory disease.

However, the diagnosing of these diseases is very important, because in general, for different diseases, different treatments are efficient.

3.2.1 How Different Dysfunctions are Diagnosed

Since it is difficult to diagnose different dysfunctions solely based on the symptoms, the corresponding diagnostics involves measuring airflow in different situations. The most effective diagnostic comes from active measurements – spirometry. A patient is asked to deeply inhale, to hold their breath, and to exhale as fully as possible – and the corresponding instrument is measuring the airflow following all these instructions. Based on these measurements, symptoms, and clinical history, medical doctors come up with a diagnosis of different dysfunctions.

3.2.2 Children Diagnostics: A Serious Problem

Unfortunately, the spirometry technique described above does not work with little children, especially children of pre-school age, since it not easy to make them follow the corresponding instructions; see, e.g., [9]. The same problem occurs with elderly patients and patients with certain limitations.

An additional problem is that even when children follow instructions during the spirometry, spirometry results are not sensitive enough to detect obstruction of the small airways (2 mm or less in diameter); see, e.g., [15, 16, 17, 18, 19].

Since we cannot use active measuring techniques, techniques that require children’s active participation, we have to rely on passive techniques, i.e., techniques that do not require such participation and are reliable in the detection of peripheral obstruction such as Impulse Oscillometry.

3.2.3 Difficulty to Diagnose based on IOS Data

If we plot the IOS data corresponding to patients with the different peripheral lung dysfunction and normal subjects studied in here, we see that the corresponding ranges of the real $R(f)$ and imaginary $X(f)$ parts of the impedance have a huge intersection, all the lung

function classes overlap. This shows that it is not easy to diagnose a patient based on IOS data.

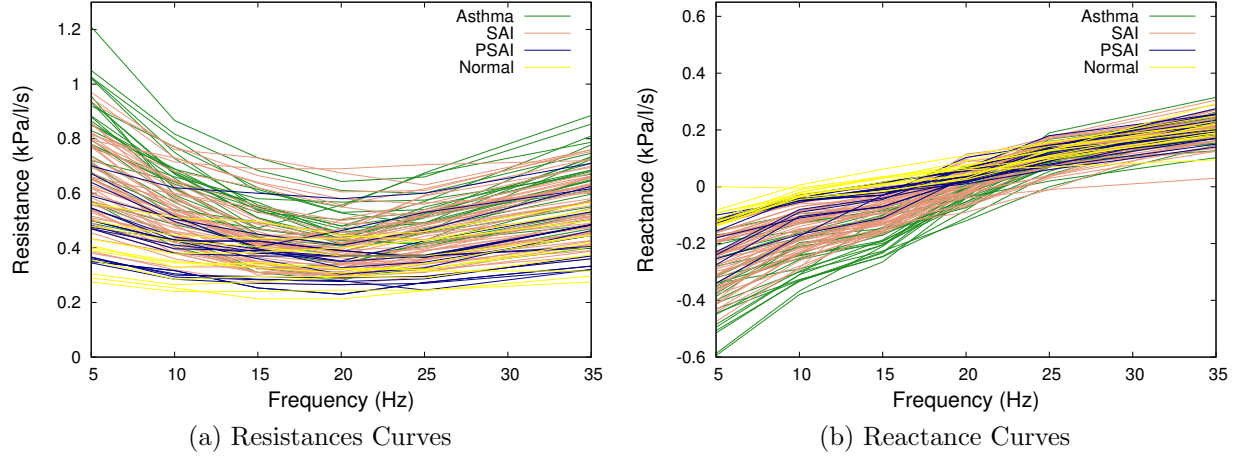


Figure 3.1: Curve Behavior of Resistance (a) and Reactance (b) per Class

3.2.4 How IOS-Based Diagnosis could be Performed?

There are no exact formulas that describe the diagnosis based on the resistance and reactance values $R(f)$ and $X(f)$; the research about the clinical applications of the IOS parameters is still ongoing. However, we do have 288 data sets from 112 patients for which, on the one hand, we know the corresponding values $R(f)$ and $X(f)$, and, on the other hand, we have a diagnosis provided by a skilled medical doctor. It is therefore reasonable to use machine learning and train the system to be able to classify the lung function of a patient.

This has indeed been done: researchers have used either all or some of the 12 numbers ($R(f)$ and $X(f)$) and other IOS derived parameters such F_{res} and $R5 - R20$ as input features and tried to train the classifier to learn the diagnosis in children patients.

The resulting diagnostic systems are, however, not yet perfect. For adult patients, if spirometry as well as IOS static features (solely based on the first pulmonary function testing) are used, the classification accuracy of asthma from healthy is 42-52%, this clas-

sification improved to 92-98% when dynamic features (derived from the application BDT and BPT tests as well as second and/or third measurements of pulmonary function) were added as the classifier inputs; see Chapter 2. It is important to mention that from both computational and clinical points of view, the use of dynamic features is not desirable, since these features are not appropriate for small children.

On the other hand, when using only IOS static features, the current system's testing-data accuracy in distinguishing peripheral lung dysfunctions such as asthma, SAI, and PSAI from patients who do not have any of these diseases is only close to 60% [50].

3.3 Resulting Problem

It is therefore desirable to come up with better computer-aided diagnostic systems based on IOS-derived features and improve the utility of the IOS. Our approach is to help a neural network by selecting discriminative IOS features. The selection of the features was done following two main approaches, the Conventional approach and the Data Pre-Processing approach that are further described Section 4.1.

3.4 Research Objective

The goal is to develop computational classification algorithms with high discriminative capacity (sensitivity, specificity, and accuracy) to distinguish between Asthma, Small Airway Impairment, Possible Small Airway Impairment, and Normal lung function. The expected outcome is an intelligent system that facilitates the difficult task of interpreting the IOS data and provides clinicians with a reliable and proven method for accurate classification of children's lung function.

3.5 Research Questions

3.5.1 Research Questions

The research questions that will be answered in this dissertation are as follows:

- a) What are the IOS derived features that best classify respiratory small airway function in children?
- b) Does computer classification of IOS data improve the diagnostic utility of this child-friendly lung function testing?

3.5.2 Hypothesis Testing

The hypotheses to be tested are as follows:

For the Conventional Approach:

- 1) Conventional-derived IOS features reliably classify Asthmatic, SAI, PSAI, and Normal children based on Artificial Neural Networks (ANN) derived algorithms.

For Data Pre-Processing:

- 2) Data pre-processing based on IOS resistance values reliably classify Asthmatic, SAI, PSAI, and Normal children based on Artificial Neural Networks (ANN) derived algorithms.
- 3) Data pre-processing based on IOS reactance values reliably classify Asthmatic, SAI, PSAI, and Normal children based on ANN derived algorithms.
- 4) Performance of data pre-processing-based classification of Asthmatic, SAI, PSAI, and Normal improves when using features based on both IOS resistance and reactance values.

General:

- 5) The best classification performance is achieved when using IOS discriminative features derived from both the conventional and data pre-processing approaches.

Chapter 4

Data and Methodology

4.1 Data

4.1.1 Subjects

The IOS dataset acquired as part of an NIH-funded study (“Asthma on the Border”) was used for this study. The data were collected in El Paso, Texas by Erika Meraz with the approval of the University of Texas at El Paso Institutional Review Board (IRB). An informed consent form was given to every parent and their child, providing them with a detailed description of the study. This unique IOS database consists of 112 records of male and female Caucasian and Hispanic children from 5 to 17 years of age. The demographics of the studied population are further described in Table 4.1 and Figure 4.2.

Table 4.1: Study’s Population Demographics.

Age (Years)	Height (cm)	Weight (kg)	Body Mass Index (Kg/m ²)	Ethnicity	Gender	Age by Gender (Years)	Children Tested	Datasets
Range: 5 - 17	Range: 101.6 - 183.4	Range: 14.5 - 93.8	Range: 12.7 - 37.6	Caucasian	Male	5-17	38	78
					Female	5-17	26	137
Mean ± SD: 9.88 ± 3.62	Mean ± SD: 139.8 ± 21.31	Mean ± SD: 41.02 ± 19.94	Mean ± SD: 19.66 ± 4.73	Hispanic	Male	5-17	22	42
					Female	5-16	26	31
						Total	112	288

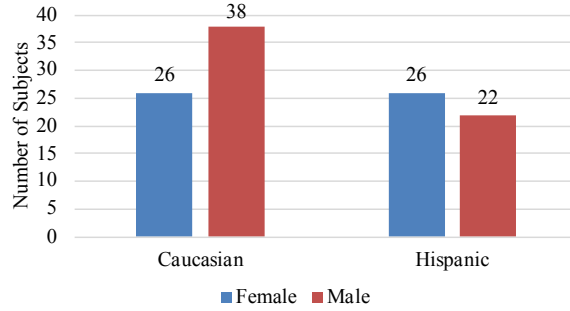


Figure 4.1: Study Demographics by Ethnicity and Gender.

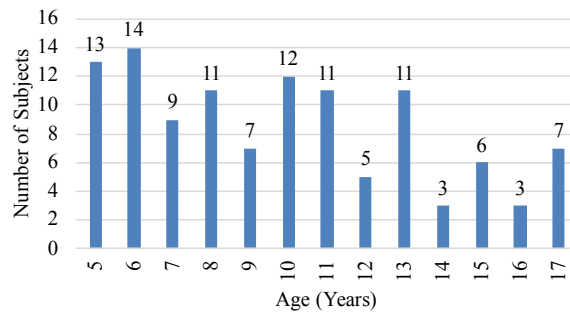


Figure 4.2: Study Demographics by Age.

4.1.2 IOS Equipment and Data Acquisition

The equipment used for the study was a Jaeger MasterScreen IOS (manufactured by Viasys Healthcare, Höchberg, Germany). The system was calibrated every day before data collection using a 3-L syringe for volume calibrations and a reference resistance (0.2 kPa/L/s) for pressure calibrations. Children were asked to wear a nose clip, while breathing normally through a mouthpiece, and were instructed to close their lips tightly around it to avoid air leakage. During data collection, three to five IOS test repetitions were performed on each subject to ensure reproducible tests without artifacts. In each IOS test, impulses were applied for a period of 30 to 45 seconds. The data were then carefully reviewed (quality-assured) offline by our expert clinician to ensure that they were artifact free (no air leaks, no swallowing effects, no breath holding, no vocalization), segments containing

artifacts were rejected and were not used for the present study. After the review of every test, each child’s data were classified into one of four conditions based on the clinician’s expertise and experience: Normal (N), Possible Small Airway Impairment (PSAI), Small Airway Impairment (SAI), or Asthmatic (A). Table 4.2 and Figures 4.3 and 4.4 detail the demographics by class.

Table 4.2: Demographics for Clinician’s Classification.

Classification	Age (Years)			Body Mass Index (Kg/m ²)		
	Range	Mean	SD	Range	Mean	SD
Asthma (n=30)	5 - 13	8.1	2.5	14.35 - 37.57	18.89	5.35
SAI (n=54)	5 - 17	9.3	3.4	12.73 - 31.21	19.31	4.49
PSAI (n=17)	5 - 17	12.6	3.8	14.64 - 30.52	20.16	3.93
Normal (n=11)	11 - 17	13.4	2.5	16.14 - 30.15	22.79	4.55

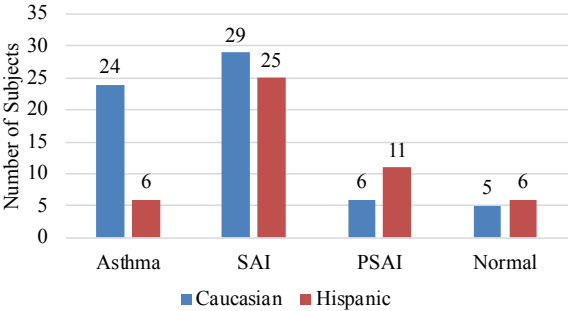


Figure 4.3: Clinician’s Classification Breakdown by Ethnicity.

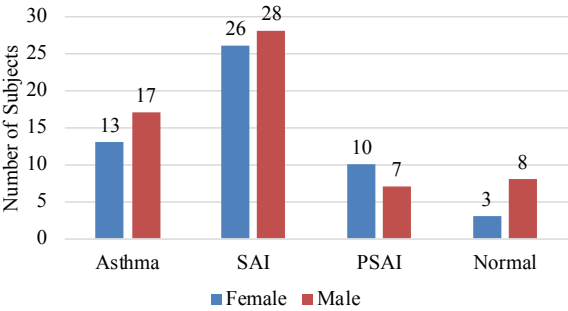


Figure 4.4: Clinician’s Classification Breakdown by Gender.

The collected data from IOS testing included resistance and reactance measurements at frequencies 5, 10, 15, 20, 25, and 35 Hz, resonant frequency (Fres), and the reactance area (AX). Additionally, R5-R20 (fdR), the eRIC, and the aRIC parameters were calculated. The methodology for parameters estimation of the eRIC and aRIC respiratory models are further described by Diong et al. in [41] and [42], respectively. In total, 24 IOS derived features for each child were considered. Table 4 lists all parameters obtained for the present study and shows the mean and coefficient of variation of each IOS parameter per each of the studied conditions.

Table 4: Mean and Coefficient of Variation of each IOS Parameter.

IOS Parameter	Units	Mean				Coefficient of Variation (%)			
		Asthma	SAI	PSAI	Normal	Asthma	SAI	PSAI	Normal
R5	kPa/l/s	0.823	0.654	0.493	0.400	20.2	22.2	23.0	22.8
R10	kPa/l/s	0.629	0.519	0.410	0.348	18.2	21.5	22.5	24.8
R15	kPa/l/s	0.504	0.430	0.374	0.342	19.7	24.7	22.4	24.9
R20	kPa/l/s	0.450	0.402	0.354	0.332	19.1	24.2	23.6	22.3
R25	kPa/l/s	0.498	0.447	0.369	0.342	17.2	21.5	25.4	21.2
R35	kPa/l/s	0.650	0.551	0.461	0.417	18.0	19.7	24.4	23.1
X5	kPa/l/s	-0.362	-0.263	-0.168	-0.106	32.2	34.3	42.7	37.9
X10	kPa/l/s	-0.240	-0.156	-0.072	-0.023	32.2	42.3	61.5	65.9
X15	kPa/l/s	-0.163	-0.091	-0.026	0.018	35.6	49.7	130.8	128.6
X20	kPa/l/s	-0.025	0.021	0.040	0.062	172.4	214.7	62.1	41.9
X25	kPa/l/s	0.100	0.111	0.114	0.120	42.1	41.8	29.2	22.4
X35	kPa/l/s	0.199	0.206	0.210	0.206	22.5	22.4	18.7	23.9
R5-R20	kPa/l/s	0.373	0.253	0.139	0.068	31.3	35.5	48.5	49.8
Fres	Hz	20.914	19.334	16.697	13.145	8.0	12.6	11.6	18.9
AX	kPa/l	3.015	1.887	0.866	0.366	32.5	39.5	58.0	32.9
eRIC R	kPa/l/s	0.423	0.389	0.347	0.329	18.6	23.5	24.5	22.8
eRIC Rp	kPa/l/s	0.836	0.637	0.444	0.497	31.5	37.6	48.3	103.4
eRIC I	kPa/l/s ²	0.001	0.001	0.001	0.001	28.6	30.8	31.6	16.0
eRIC Cp	l/kPa	0.040	0.063	0.112	0.188	32.9	44.2	34.2	34.8
aRIC R	kPa/l/s	0.371	0.340	0.324	0.310	27.1	33.1	28.2	26.4
aRIC Rp	kPa/l/s	0.795	0.597	0.408	0.309	28.9	31.7	41.3	34.6
aRIC I	kPa/l/s ²	0.002	0.002	0.001	0.001	39.3	45.1	45.2	24.2
aRIC Cp	l/kPa	0.032	0.051	0.100	0.152	40.7	59.0	41.5	20.4
aRIC Ce	l/kPa	0.003	0.003	0.002	0.002	68.1	74.1	95.9	96.3

4.2 Methodology

The methodology used in the deployment of the IOS-based classification algorithms include the deep analysis of previously collected IOS data, which is described in the Section 4.1

of this dissertation and are further explained in the following sections. A Conventional approach, a Data Pre-Processing approach, and a combination of both was used for feature selection and classification. The conventional approach was based on the individual analysis of IOS parameters, including those derived from the Human Respiratory Systems Models previously developed by UTEP’s research group. deterministic approach was one based on reference models or functions; and the last approach included the combination of both. The intention was to methodologically compare the ANN classifiers’ effectiveness and performance. The classification work included two main phases: the Training Phase and the Prediction Phase.

4.2.1 Feature Selection

The biomedical raw data was statistically assessed and transformed by geometric, algebraic, and other mathematical techniques for dimensionality reduction, features selection and classification. During this stage, IOS features were deeply studied using Conventional and Pre-Processing approaches further explained in the following sections.

Conventional Approach

The conventional approach was based on the individual analysis of the 24 IOS parameters listed in Section 4.1, these include the parameters derived from the Human Respiratory Systems Models previously developed by UTEP’s research group. The objective was to identify the IOS direct features that demonstrated statistical significance in the differentiation between Asthma (A), Small Airway Impairment (SAI), Possible Small Airway Impairment (PSAI), and Normal (N) respiratory conditions.

Data Pre-Processing Approach

The Pre-Processing approach was based on the analysis of the actual shape of the IOS curves. Discriminative parameters were obtained based on the development of reference

models or typical functions. In addition, statistical analysis was performed to select the most discriminative parameters.

4.2.2 Supervised Classification

The classification work performed included two main phases: the training phase and the prediction phase. During training phase, 75% of the labeled IOS data sets were used, while in the prediction Phase 25% of the IOS data sets were used as validation data sets. Artificial Neural Networks (ANN) classifiers were developed using input features derived from the 1) Conventional approach, 2) Pre-Processing Approach, and 3) a combination of both.

4.2.3 ANN Performance Evaluation and Model Selection

Accuracy of each of the ANN classifiers developed were assessed to determine the IOS derived parameters and the ANN that best classified the lung conditions studied. After selection of the ANN(s) the overall system was assessed in terms of sensitivity and specificity.

Chapter 5

Feature Selection: Conventional Approach

This chapter focuses on the characterization of the IOS features using conventional approach which was based on the individual analysis of IOS parameters. This was a preliminary step for our computer-aided classification of lung function in children. The aim of this work was to identify those IOS features that demonstrate statistical significance pertaining to the differentiation between four different degrees of small airways obstruction which are the Asthma (A), Small Airway Impairment (SAI), Possible Small Airway Impairment (PSAI), and Normal (N) respiratory conditions.

5.1 IOS Data Characterization for Dimensionality Reduction

The high dimensionality of the IOS parameters, as well as the dispersion of the data generated for the different classes produce an overlapping effect. This makes computer-aided classification of multiple classes with different degrees of severity in peripheral obstruction difficult. The complexity of the data used for this investigation plotted in terms of the respiratory impedance components R and X , as a function of the frequency f is shown in Figure 5.1.

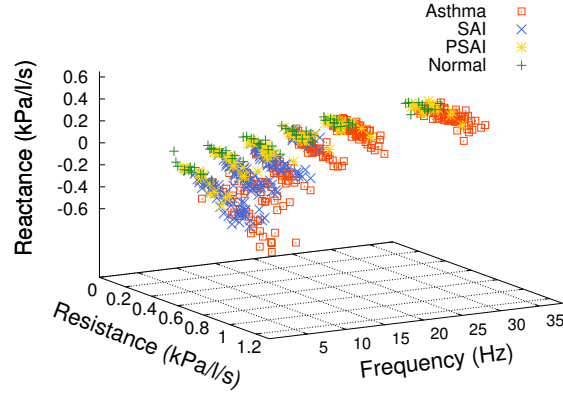


Figure 5.1: 3D Plot of Resistance and Reactance vs. Frequency.

It is observed that the data from the different classes overlap, thus making the class differentiation a challenging task. Therefore, a statistical analysis of the IOS parameters is required to find discriminating features that could help in the accurate classification of classes with different degrees of distal obstruction. To this end, statistical analysis was performed using the MINITAB 18 Statistical Software (Minitab, Inc., State College, USA). ANOVA One-way (Analysis of Variance) was the test used to determine statistical differences between data sets. In other words, ANOVA evaluates the importance of one or more factors by comparing the means to determine if they are different. With this test, if the obtained p-value was less than 0.05, the Null Hypothesis H_0 was rejected, and it could be firmly concluded that there was statistical evidence that the means of the compared groups were different. However, if the p-value was greater than 0.05, then the test failed to reject H_0 , and it could then be concluded that there was no significant statistical evidence to infer that the means between the compared groups were different. The statistical analysis was performed using a confidence level of 95%.

The following hypotheses were tested:

$$\text{Null Hypothesis} \quad H_0 : \mu_1 = \mu_2 \quad (5.1)$$

$$\text{Alternative Hypothesis} \quad H_i : \mu_1 \neq \mu_2 \quad (5.2)$$

In this analysis, each IOS parameter for each class was compared against the same IOS parameter for a different class, until the parameter was compared for all classes. A comparison-matrix for each of the main IOS parameters was completed, where the p-values obtained for each of the comparisons were listed in the corresponding matrix. Please refer to Table 3 for the resistance parameters comparison, Table 5.2 for the reactance parameters, Table 5.3 for other IOS derived parameters, Table 5.4 for the eRIC, and Table 5.5 for the aRIC human respiratory model parameters. For an IOS parameter to be considered potentially discriminative all of its p-values must be less than 0.05. The corresponding tables have all cells where the p-values are greater than 0.05 grayed out, which render that specific IOS feature non-discriminative. Any IOS-parameter that had at least one cell in gray was discarded.

5.2 Results

For resistance, it is observed in Table 3 that the resistance parameters R5, R10, R15, R20, R25, and R35 are potentially discriminative to differentiate between all classes except SAI vs. PSAI and PSAI vs. Normal; whereas R15 and R20 cannot differentiate between SAI vs. PSAI, and R10, R15, R20, R25, and R35 cannot differentiate between PSAI vs. Normal. Therefore, the only resistance parameter found to be potentially discriminative for all conditions was R5.

Table 3: IOS Resistance Parameters Comparison-Matrix.

IOS Parameter	Asthma vs. SAI	Asthma vs. PSAI	Asthma vs. Normal	SAI vs. PSAI	SAI vs. Normal	PSAI vs. Normal
R5	0.000	0.000	0.000	0.000	0.000	0.031
R10	0.000	0.000	0.000	0.001	0.000	0.086
R15	0.002	0.000	0.000	0.055	0.012	0.324
R20	0.027	0.001	0.000	0.073	0.029	0.490
R25	0.018	0.000	0.000	0.005	0.001	0.422
R35	0.000	0.000	0.000	0.004	0.000	0.289

For reactance, it is observed in Table 5.2 that the reactance parameters X5, X10, X15 are potentially discriminative to differentiate all classes, while X20 differentiates all classes except SAI vs. PSAI. X25 and X35 do not differentiate any of the classes. Therefore, the reactance parameters found to be potentially discriminative for all conditions were X5, X10, and X15.

Table 5.2: IOS Reactance Parameters Comparison-Matrix.

IOS Parameter	Asthma vs. SAI	Asthma vs. PSAI	Asthma vs. Normal	SAI vs. PSAI	SAI vs. Normal	PSAI vs. Normal
X5	0.000	0.000	0.000	0.000	0.000	0.015
X10	0.000	0.000	0.000	0.000	0.000	0.002
X15	0.000	0.000	0.000	0.000	0.000	0.001
X20	0.000	0.000	0.000	0.105	0.005	0.003
X25	0.280	0.239	0.149	0.802	0.540	0.627
X35	0.497	0.428	0.687	0.790	0.974	0.821

Table 5.3: IOS Derived Parameters Comparison-Matrix.

IOS Parameter	Asthma vs. SAI	Asthma vs. PSAI	Asthma vs. Normal	SAI vs. PSAI	SAI vs. Normal	PSAI vs. Normal
R5-R20	0.000	0.000	0.000	0.000	0.000	0.003
Fres	0.002	0.000	0.000	0.000	0.000	0.000
AX	0.000	0.000	0.000	0.000	0.000	0.003

For the other IOS derived parameters category, which included R5-R20 (fdR), Fres, and AX, it is observed in Table 5.3 that for all instances these parameters are potentially discriminative since all p-values were nearly equal to zero in all cases.

For the eRIC parameters, it is observed in Table 5.4 that the C_p parameter is potentially discriminative to differentiate all classes, while R_p and I differentiate all classes except PSAI vs. Normal, and the R_c parameter does not differentiate any of the classes. Therefore, the only eRIC parameter found to be potentially discriminative for all conditions was C_p .

Table 5.4: eRIC Parameters Comparison-Matrix.

eRIC Parameter	Asthma vs. SAI	Asthma vs. PSAI	Asthma vs. Normal	SAI vs. PSAI	SAI vs. Normal	PSAI vs. Normal
R	0.091	0.003	0.003	0.097	0.073	0.669
Rp	0.001	0.000	0.000	0.004	0.000	0.205
I	0.026	0.000	0.000	0.019	0.011	0.463
Cp	0.000	0.000	0.000	0.000	0.000	0.001

For the aRIC parameters, it is observed in Table 5.5 that the C_p parameter is potentially discriminative to differentiate all classes. R_p differentiates all classes except PSAI vs. Normal, while I differentiates between all classes except Asthma vs. SAI and PSAI vs. Normal. The R_c and C_e parameters do not differentiate any of the classes. Therefore, the only aRIC parameter found to be potentially discriminative for all conditions was C_p .

Table 5.5: aRIC Parameters Comparison-Matrix.

aRIC Parameter	Asthma vs. SAI	Asthma vs. PSAI	Asthma vs. Normal	SAI vs. PSAI	SAI vs. Normal	PSAI vs. Normal
R	0.208	0.121	0.082	0.611	0.418	0.689
Rp	0.000	0.000	0.000	0.000	0.000	0.094
I	0.053	0.000	0.000	0.008	0.004	0.334
Cp	0.001	0.000	0.000	0.000	0.000	0.002
Ce	0.738	0.112	0.455	0.070	0.347	0.610

Table 5.6 summarizes the IOS derived-parameters that have discriminative capacity to statistically differentiate all four classes based on their means. It is also important to understand the dispersion of the data for each parameter, since less variation will imply less data overlapping between groups. Therefore, these IOS parameters were ranked based on the dispersion of their data measured in terms of the average of the coefficient of variation for the different classes. Table 5.6 shows the ranking based on the coefficient of variation.

Table 5.6: Ranking of Potentially Discriminative IOS Parameters.

IOS Parameter	Coefficient of Variation (%)				CV Avg	Ranking
	Asthma	SAI	PSAI	Normal		
Fres	8.0	12.6	11.6	18.9	12.8	1
R5	20.2	22.2	23.0	22.8	22.0	2
eRIC Cp	32.9	44.2	34.2	34.8	36.5	3
X5	32.2	34.3	42.7	37.9	36.8	4
aRIC Cp	40.7	59.0	41.5	20.4	40.4	5
AX	32.5	39.5	58.0	32.9	40.7	6
R5-R20	31.3	35.5	48.5	49.8	41.3	7
X10	32.2	42.3	61.5	65.9	50.5	8
X15	35.6	49.7	130.8	128.6	86.2	9

5.3 Discussion

Based on the analysis performed, nine parameters were found sensitive to discriminate between the Asthma, SAI, PSAI and Normal pulmonary conditions. We demonstrated that these parameters statistically differentiate between four levels of peripheral lung function instead of just two (asthmatics and non-asthmatics) as previously presented by other studies [5, 16, 18, 56, 57].

From the analysis of the resistance parameters, R5 was found to be discriminative. This finding is supported by previous studies that have found R5 to be statistically significant to differentiate IOS bronchodilator responses for asthmatic and non-asthmatic children. In addition to R5, R10 has also been found to be an important parameter to differentiate between these two groups. However, in our study it was found that R10 has no discriminative capacity to distinguish between PSAI and Normal conditions [16, 18, 19, 56, 58].

Regarding reactance, we found X5, X10, and X15 to be discriminative parameters, in previous studies X5 and X10 have been observed to be sensitive to differentiate IOS bronchodilator responses for asthmatic and non-asthmatic children [16, 18, 19, 58]. In addition, X5 has been shown to correlate with improvement in the peripheral lung function due to the use of systemic drugs [57].

R5-R20 (fdR), Fres, and AX have also been found to be sensitive measures to detect lung function changes in children in previous studies [5, 6, 19, 39, 58, 59, 60].

The findings related to the equivalent electrical circuit model parameters for the human respiratory system are supported by previous studies performed by our research group. The baseline IOS measures and estimated model parameters have previously been analyzed to evaluate their discriminative capacity to track changes in lung function in children [5, 39, 59]. In these studies, IOS measures and estimated model parameters of Asthma, SAI, PSAI, and Normal children were statistically analyzed to evaluate their discriminative capacity by comparing pre and post-bronchodilator responses for each group. According to these studies, the data for Asthma and SAI seemed to fall into one category (Small Airway Impaired), while PSAI and Normal could be grouped into a different one (Healthy). A statistical assessment was then performed to identify the IOS and estimated model parameters that were discriminative to distinguish between the two groups, it was concluded that AX and the eRIC C_p estimated model parameter were the most sensible and reliable measures to statistically distinguish between the two groups [5]. With the present study, we were able to confirm that the eRIC C_p estimated model parameter not only differentiates between two conditions but also differentiates between the four conditions presented in here. In addition to the eRIC C_p , the aRIC C_p estimated model parameter was also found discriminative to distinguish between the four groups.

Table 5.7 provides a summary of previous studies where IOS parameters have been found to be discriminative to differentiate mostly between two groups (asthmatic and non-asthmatic children) or evaluate significant differences between pre and post-bronchodilation data within an specific group. These studies in one way or another support the IOS parameters found to be potentially discriminative in this study.

5.4 Conclusions

In this work, we were able to identify the IOS parameters that statistically differentiate between four levels of peripheral lung function in children (Asthma, SAI, PSAI and Normal). Out of the 24 IOS parameters studied, only 9 were found to be sensitive to differentiate

Table 5.7: Findings of Previous Studies Related to IOS Discriminative Parameters.

Author	Year	Conditions Studied	**Type of Assessment	Type of Comparisson	Statistical Technique Used	IOS Parameters Studied	Discriminative IOS Parameters Found
Schulze et al	2016	Exacerbated and Non-Exacerbated Asthmatic Children	Static	Bi-Class	Fisher's exact test	R5, X5, R5-R20	R5, X5, R5-R20
Komarow et al	2012	Asthmatic & Non-asthmatic Children	Dynamic	Bi-Class	Wilcoxon sign rank tests	R5, R10, R20, R5-R20, X5, AX	* Δ R5 (%), * Δ R10 (%),* Δ X5 (%), * Δ AX (%)
Meraz et al	2012	Asthma, SAI, PSAI, Normal (Children)	Dynamic	***Bi-Class	T-test	R3, R5,R10,R15,R20, R35, R3-R20, R5- R20, X3, X5, X10, X15, X20, X35 AX, Fres, eRIC R, eRIC I, eRIC Cp, eRIC Rp, aRIC R, aRIC I, aRIC Cp, a RIC Rp, and aRIC Ce	Asthma: All parameters except aRIC Ce and eRIC I SAI: R3,R5,X3,X5,X10,X15,R3-R20,R5-R20,AX,aRIC Rp, eRIC Rp, eRIC I and eRIC Cp PSAI: R5, R35, eRIC Rc Normal: None
Shi et al	2011	Healthy, Asthmatic Controlled, and Asthmatic Uncontrolled Children	Static & Dynamic	Tri-Class	Mann-Whitney U test	R5, R20 R5-R20, Fres, X5, AX	R5, R5- 20, Fres, X5, AX * Δ R5(%), * Δ R20(%), * Δ R5-20(%),* Δ Fres (%), * Δ X5 (%), * Δ AX (%)
Meraz et al	2011	SAI and Healthy (Children)	Static & Dynamic	Bi-Class	T- Test	R5, R5-R20, AX , aRIC Rp, aRIC Cp, eRIC Rp, eRIC Cp	R5, R5-R20, AX , aRIC Cp, and eRIC Cp
Meraz et al	2009	Normal-PSAI and Asthma-SAI (Children)	Dynamic	***Bi-Class	T-test	R3, R5,R10,R3-R20, R5- R20, X3, X5, X10, AX, eRIC R, eRIC I, eRIC Cp, eRIC Rp, aRIC R, aRIC I, aRIC Cp, aRIC Rp, and aRIC Ce	Normal-PSAI group: None Asthma-SAI group: All parameters track changes except aRIC R, aRIC Ce and eRIC I.
Song et al	2008	Asthmatic & Non-asthmatic Children	Static & Dynamic	Bi-Class	T-tests	R5, R10, R20, R35, Fres, X5, Z	R10 * Δ R5(%), * Δ R10(%), * Δ R20(%), * Δ R35(%)
Nieto et al	2006	Asthmatic Chidren threatened with Montelukast .	Static	Bi-Class	T-test, Wilcoxon rank test	R5, R20, R5-R20, Fres, X5, Z	R5, R20, R5-R20, Fres, X5, Z
Marotta et al	2003	Asthmatic & Non-asthmatic Children	Dynamic	Bi-Class	Wilcoxon rank sum test	R5, R10, X5, X10, Fres	* Δ R5 (%), * Δ R10 (%)
Marotta et al	2003	Atopic Children with and without Asthma	Dynamic	Bi-Class	Wilcoxon rank sum test	R5, R10, X5, X10, Fres	* Δ R5 (%), * Δ R10 (%)

*Results based on percentage change between pre and post-bronchodilator responses.

** Static assessment is based only on first IOS testing. Dynamic assessment uses Pre and Post Bronchodilatation tests .

***Pre and Post Bronchodilator response comparisson within each group.

between the four respiratory conditions studied in here. The discriminative IOS parameters identified and listed from higher to lower sensitivity (based on the coefficient of variation – a measure of dispersion of their data) were: F_{res} , R_5 , C_p from the eRIC model, X_5 , C_p from the aRIC model, AX , R_5 - R_{20} , X_{10} , and X_{15} . Even though the parameters C_p from the eRIC model and C_p from the aRIC model were found discriminative, these parameters were not used in further classification work, since individual values for each IOS test were not available in the database used: these parameters were calculated based on the average of the patient’s datasets. Therefore only 7 parameters were used as input features in further computer-aided classification work to best distinguish normal lung function and different degrees of small airways obstruction in children and they were: F_{res} , R_5 , X_5 , AX , R_5 - R_{20} , X_{10} , and X_{15} .

Chapter 6

Feature Selection: Data Pre-Processing Approach

Several clinical studies have been done to understand the resulting IOS testing values and to be able to relate them to the lung function condition studied. In these studies, specific IOS parameters have been evaluated/assessed individually or independently. For example, most of these studies have assessed R5 and X5 individually and independently, and reference equations have been developed to determine the belonging to certain condition for each of the parameters. By doing an assessment based on specific IOS features, other patient's demographic/anthropometric parameters need to be taken into consideration. For example, the height of a patient has been found to be an important aspect to determine the relationship of the IOS parameters values and the lung function for a specific patient. For instance, healthy small children present higher resistance values than adults, the explanation is because the anatomy of the airways is smaller in children than adults, and naturally present a higher resistance to the airflow. Therefore, the resulting IOS parameter values depend, for this particular case, on the height of the subject. The dependency on anthropometric variables also adds complexity to the IOS data and contributes to the lack of an agreement about the ability of the standardized IOS values to accurately diagnose airway obstruction. Therefore, an improvement of the clinical utility of the IOS could be achieved by developing robust computer-aided classifiers based on input features derived from the actual IOS patient's curve.

This chapter is a preliminary step for the computer-aided classification of lung function in children. Here, a new approach in the use of the IOS resistance (R) and reactance

(X) parameters is provided. Potential discriminative features based on the shape of the curves can be obtained by using a data pre-processing method for feature extraction and selection, this approach implies a new vision of how to assess the IOS parameters. More than taking into consideration individual values (i.e. $R5$), it takes into consideration the actual behavior of the patient's curves of resistance or reactance, which already have the individual's anthropometric information intrinsic to the IOS curve. This is the first research work where the shape of the IOS frequency-dependent curves has been analyzed as a whole towards the classification of lung function in children using the IOS parameters.

The aim of this work is to present the development of reference models or typical functions for each of the classes to extract discriminative features that demonstrate statistical significance pertaining to the differentiation between four different degrees of small airways obstruction which are the Asthma (A), Small Airway Impairment (SAI), Possible Small Airway Impairment (PSAI), and Normal (N) respiratory conditions.

6.1 IOS Data Behavior

An essential aspect to be understood is that the IOS graphically displays frequency-dependent curves that are of the utmost importance in the diagnosis of distal obstruction; changes in lung function are evaluated based on the visual interpretation of their shape and magnitude [15]. Recapitulating from Chapter 2, resistance is independent of the frequency in healthy subjects, while for small airway obstruction, the resistance at lower frequencies increases, but it is unchanged at higher frequencies that do not reach the small airways [40]. In other words, distal obstructive diseases such as Asthma result in a frequency-dependent increase in resistance at lower frequencies, because the pressure signal wave propagating out to the lung periphery ($R5$) encounters greater resistance than the more proximal higher-frequency ($R20$) impulse [15]. Unlike resistance, reactance is always frequency-dependent. In healthy individuals, the magnitude of reactance is smaller, while in individuals with peripheral obstruction the magnitude (absolute value) of the reactance

increases and the reactance itself becomes more negative. Specifically, the difference in the reactance magnitude between healthy and individuals with peripheral obstruction is more pronounced at the lower frequencies where the capacitive properties of the lungs dominate and in which the small airways play an important role as storing capacity energy primarily relies on the periphery of the lungs [15, 17, 40].

In an effort to understand the overlapping of the classes, three data sets of resistance values were chosen per class; the minimum, the middle and maximum R5 value of each class data set was selected. In other words, for each class, the data of the child whose R5 was the largest in this class was selected to be plotted. The same approach was taken for the minimum value, and for the value R5 which is closest to the midpoint between the smallest and the largest values of R5 from this class. The actual curves for all classes were plotted into Figure 6.1. It is observed that all classes overlap at some point. But it can clearly be observed that the pattern of the curves of each class seems to be a potential feature for discrimination. By plotting the IOS data corresponding to patients with the different conditions studied here, the corresponding ranges of values for Resistance $R(f)$ and Reactance $X(f)$ have a huge intersection. This overlapping is attributed to the dispersion of the data generated and shows that it is not easy to diagnose a patient based on IOS data.

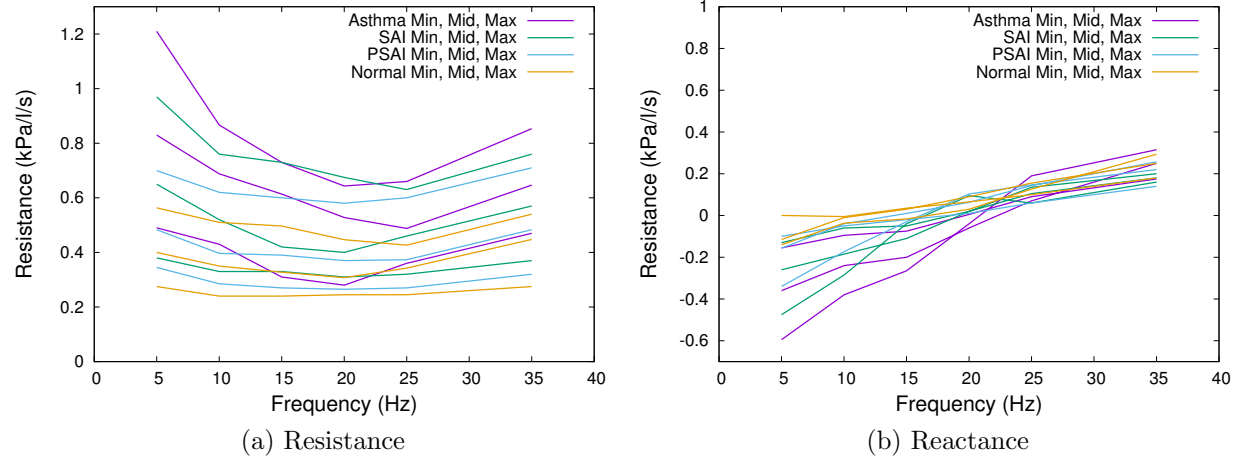


Figure 6.1: Resistance (a) and Reactance (b) Maximum, Middle and Minimum Patients' Curves per Class

To better understand the resistance and reactance behavior of the collected data for each class (Asthma, SAI, PSAI, and Normal), the class average for each of R and X parameters at the different frequencies was calculated (R5, R10, R15, R20, R25, R5, R10, R15, R20, R25, R35, and X5, X10, X15, X20, X25, X35). The resistance and reactance curves were plotted separately using the obtained average results per class per parameter and are shown in fig. 6.2.

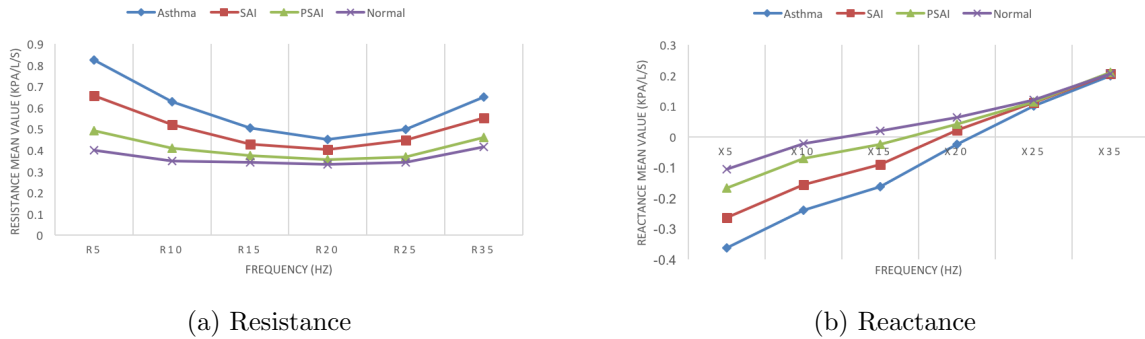


Figure 6.2: Curve Behavior of Resistance (a) and Reactance (b) Average per Class

The mean curves of R and X at different frequencies for the different classes are in agreement with what is expected. For R, the curves present a different degree of frequency-

dependency at low frequencies (R5, R10, R15), that varies depending on the severity of the small airways obstruction. Meanwhile, X decreases when the severity of the condition evaluated increases.

Derived from the previous analysis it is confirmed that the shape of the curves could provide important information for the classification of the 4 different levels of severity of distal obstruction that are studied here.

6.2 Typical Class Functions

From the previous average analysis, it makes sense to find a function that could represent each of the classes studied. Polynomial regressions were performed in order to obtain reference models or typical class functions. Quadratic, cubic and quartic order regressions were done for each class data set (Asthma, SAI, PSAI, and Normal), to find the function that best fits the IOS parameters of resistance (R) and reactance (X) using frequency as independent variable ($R(f)$, $X(f)$).

Polynomial regressions were computed given the general polynomial form:

$$y_i = a_0 + a_1 f_i + a_2 f_i^2 + .. + a_n f_i^n \quad (6.1)$$

where, in this particular scenario:

f_i = Frequency in Hz,

y_i = Resistance $R(f)$ or reactance $X(f)$ in terms of frequency (f) in kPa/l/s.

By using the Least Squares method, we get the following linear system:

$$\mathbf{y} = \mathbf{A} \cdot \mathbf{a}, \quad (6.2)$$

$$\begin{pmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n f_i y_i \\ \sum_{i=1}^n f_i^2 y_i \\ \vdots \\ \sum_{i=1}^n f_i^n y_i \end{pmatrix} = \begin{pmatrix} n & \sum_{i=1}^n f_i & \cdots & \sum_{i=1}^n f_i^n \\ \sum_{i=1}^n f_i & \sum_{i=1}^n f_i^2 & \cdots & \sum_{i=1}^n f_i^{n+1} \\ \sum_{i=1}^n f_i^2 & \sum_{i=1}^n f_i^3 & \cdots & \sum_{i=1}^n f_i^{n+2} \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{i=1}^n f_i^n & \sum_{i=1}^n f_i^{n+1} & \cdots & \sum_{i=1}^n f_i^{n+n} \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \quad (6.3)$$

so that:

$$\mathbf{A}^{-1} \cdot \mathbf{y} = \mathbf{a} \quad (6.4)$$

6.2.1 Which Order Polynomials shall be Used?

Considerations to select the order of the polynomial use the fact the larger the order, the better the polynomial will fit all data. However, having polynomials of too large degree will over-fit the data. The whole purpose of the polynomial is to de-noise the signal. If all the values are kept intact, all the noise will be retained, and it will be difficult to use the polynomial to make future generalizations. Thus, polynomials of too high order should not be used. On the other hand, if polynomials of too low order are used, information contained in the original signal may be lost. For this particular case, a natural idea is to take into account the general monotonicity of IOS curves that is described in the IOS literature [15]. Therefore, the largest polynomial order shall be selected for which the curve's monotonicity pattern is conserved.

The quadratic, cubic and quartic regression functions obtained for R and X are were plotted against the actual data in Figures A.1 and B.1 in Appendices A and B. We can observe that the cubic and quartic functions behave similarly, the main difference by visually assessing the curves, is that the cubic function presents a monotonic behavior, while the quartic function does not. For example, for the resistance $R(f)$, the corresponding 4th order polynomial first decreases, then increases, but then decreases again; see Figure 6.3. The monotonic behavior of the cubic function is more in accordance with what is described

in the IOS literature about distal. Therefore, we selected the highest polynomial order that conserved the IOS curve's monotonicity, which is the cubic polynomial. Because of this, in this work, we approximate the functions $R(f)$ and $X(f)$ by *cubic polynomials*. Let us denote the corresponding approximating polynomials by $R^d(f)$ and $X^d(f)$, where d stands for disease: a = asthma, s = SAI, p = PSAI and n = Normal.

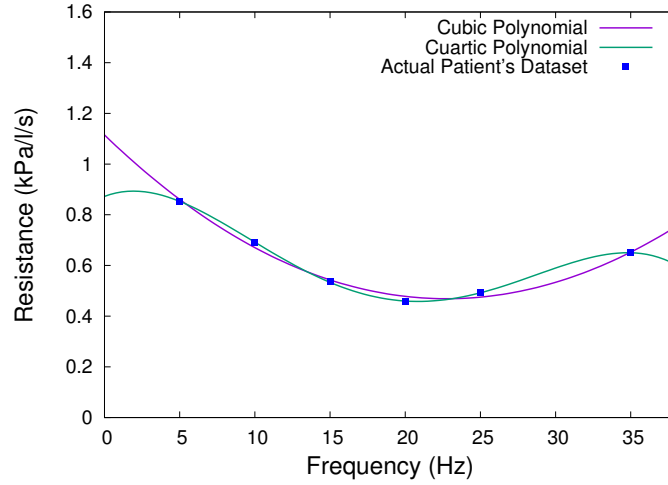


Figure 6.3: Cubic vs Quartic vs a Patient's Dataset

6.2.2 Evaluating Goodness of Fit

R-Squared and R-Squared adjusted are parameters commonly used in polynomial regressions to evaluate the goodness of fit. Therefore, it makes sense to also evaluate the goodness of fit of our polynomial regressions based on these two parameters. R-squared (Coefficient of Determination) R^2 and R-squared adjusted R^2_{adj} were computed for each class data set. The computation for R^2 was performed based on the following equations:

$$R^2 = 1 - \frac{SSE}{SST} \quad (6.5)$$

given:

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (6.6)$$

$$SSE = \sum_{i=1}^n (y_i - \bar{y})^2 \quad (6.7)$$

where:

SSE = the error sum of squares,

SST = the total sum of squares,

y_i = the actual observed value,

\hat{y}_i = estimated value using the regression model,

\bar{y} = mean of all actual data observations.

The computation for R_{adj}^2 was performed based on the following equations:

$$R_{adj}^2 = 1 - \frac{(1 - R^2)(n - 1)}{n - k - 1} \quad (6.8)$$

where:

n = number of observations in the data sample,

k = number of regression model predictors.

Table 6.1 shows the computation results of R^2 and R_{adj}^2 per class and polynomial regression order.

Table 6.1: R^2 and R_{adj}^2 Results per Class.

Condition	# Data sets (subjects)	Total data points	R^2			R^2 adjusted		
			Cuadratic	Cubic	Cuartic	Cuadratic	Cubic	Cuartic
Asthma	30	180	0.539	0.548	0.548	0.536	0.545	0.546
SAI	54	324	0.363	0.372	0.374	0.361	0.370	0.372
PSAI	17	102	0.224	0.225	0.225	0.216	0.217	0.217
Normal	11	66	0.135	0.135	0.137	0.121	0.122	0.123

It is observed in Table 6.1 that the quadratic, cubic, and quartic order R^2 and R_{adj}^2 values are very similar. The cubic and quartic approximations present the highest R^2 and R_{adj}^2 values. The difference between them is negligible.

Now, let us understand these results, R^2 and R_{adj}^2 values that are closer to 1 imply that

the model explains all the variability of the response data around its mean, and theoretically, if a model could explain 100% of the variance, the fitted values would always equal the observed values and, therefore, all the data points would fall on the fitted regression line. However, R^2 and R_{adj}^2 do not precisely indicate whether a regression model is adequate, other considerations need to be taken in place. It is important to understand that polynomial regression gives you the line that is closer to all data points in the y axis, in this particular case. Therefore, a regression model will be the function that represents the mean in the y axis in relation to the independent variable. For example, the smaller values obtained in our regression models were for the Normal class, where R^2 and R_{adj}^2 values were 0.13 and 0.12 respectively; do these values mean that we have an inadequate model? Not exactly; by analyzing further the data and looking at Equations (6.5), (6.7) and (6.8), a low R^2 value here indicates that:

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2 \approx \sum_{i=1}^n (y_i - \bar{y})^2 \quad (6.9)$$

Then,

$$\hat{y}_i \approx \bar{y} \quad (6.10)$$

For this particular case, we are looking for the function that reflects the behavior (form of the curve) of each of the classes and approximates the mean of all the data points; therefore the fact that $\hat{y}_i \approx \bar{y}$ is appropriate, and the low values obtained for the R^2 and R_{adj}^2 come from the dispersion of the actual data points versus the mean, which, as previously described, is one of the main reasons why different classes overlap. The adequacy of the cubic regression models for each of the classes is further assessed by obtaining the R^2 and R_{adj}^2 values for the mean of the classes that are reflected on Table 6.2:

Table 6.2: R^2 and R_{adj}^2 Results for the mean of the data points for each Class.

Condition	Total data points	R^2	R^2 adjusted
Asthma	6	0.994	0.992
SAI	6	0.992	0.990
PSAI	6	0.997	0.997
Normal	6	0.978	0.973

In summary, R^2 and R_{adj}^2 results along with the graphical assessment based on Figures A.1 and B.1 provided a base for the selection of the cubic model regression that was the more optimal polynomial order for the data presented in this work. The adequacy of the cubic order for its intended use was confirmed by the results shown in Table 6.2.

6.2.3 Graphical Assessment of Cubic Functions

The cubic equations of resistance obtained for Asthma Equation (6.11), SAI Equation (6.12), PSAI Equation (6.13) and Normal Equation (6.14) were plotted in Figure 6.4; observe that no overlap occurred between classes. Figure 6.4 shows the resistance behavior curves for each class.

$$R_a(f) = 1.152 - 7.842 \times 10^{-2}f + 2.686 \times 10^{-3}f^2 - 2.443 \times 10^{-5}f^3 \quad (6.11)$$

$$R_s(f) = 8.960 \times 10^{-1} - 5.738 \times 10^{-2}f + 2.067 \times 10^{-3}f^2 - 2.024 \times 10^{-5}f^3 \quad (6.12)$$

$$R_p(f) = 6.076 \times 10^{-1} - 2.717 \times 10^{-2}f + 8.278 \times 10^{-4}f^2 - 4.888 \times 10^{-6}f^3 \quad (6.13)$$

$$R_n(f) = 4.612 \times 10^{-1} - 1.508 \times 10^{-2}f + 4.789 \times 10^{-4}f^2 - 2.424 \times 10^{-6}f^3 \quad (6.14)$$

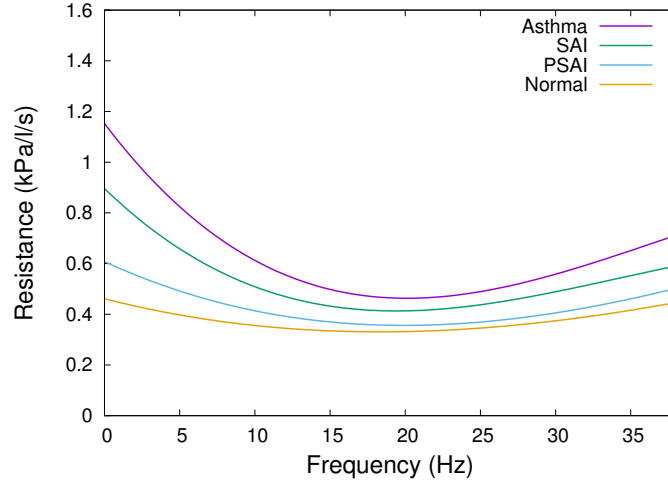


Figure 6.4: Cubic Resistance Functions per Class

Similarly, the cubic equations of reactance obtained for Asthma (6.15), SAI (6.16), PSAI (6.17), and Normal (6.18) were plotted in Figure 6.4, which also shows the reactance performance curve for each class.

$$X_a(f) = -4.0500 \times 10^{-1} + 4.4909 \times 10^{-3}f + 1.2294 \times 10^{-3}f^2 - 2.4657 \times 10^{-5}f^3 \quad (6.15)$$

$$X_s(f) = -3.2507 \times 10^{-1} + 1.0693 \times 10^{-2}f + 5.9596 \times 10^{-4}f^2 - 1.3338 \times 10^{-5}f^3 \quad (6.16)$$

$$X_p(f) = -2.4920 \times 10^{-1} + 1.8110 \times 10^{-2}f - 1.9477 \times 10^{-4}f^2 + 1.5170 \times 10^{-6}f^3 \quad (6.17)$$

$$X_n(f) = -1.9135 \times 10^{-1} + 2.0045 \times 10^{-2}f - 4.8540 \times 10^{-4}f^2 + 6.7960 \times 10^{-6}f^3 \quad (6.18)$$

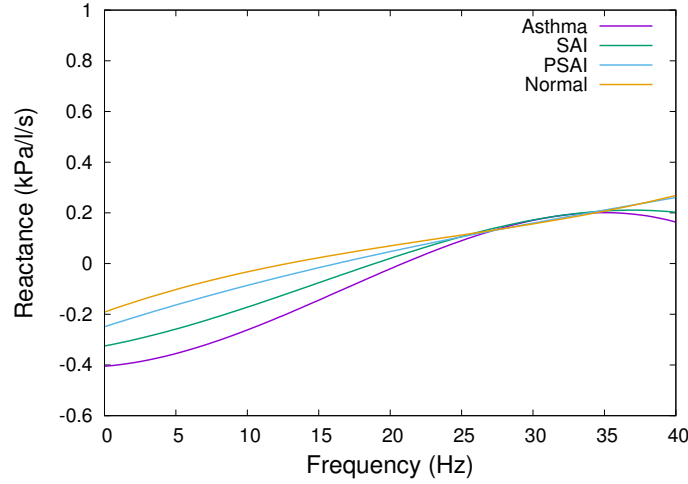


Figure 6.5: Cubic Reactance Functions per Class

From the typical functions obtained for each class, it is observed that while the Asthma category, which is the worst pathological condition evaluated, presented the the most pronounced curve with the largest airways resistance in all frequencies recorded, the Normal class presented the most flattened curve with the least airways resistance in all frequencies recorded. Additionally, it was confirmed that the resistance at low frequencies (R5, R10, and R15) in all pathological classes is frequency-dependent. Comparing these curves to the typical resistance curves in the IOS literature [6, 15, 40], we can see that the pathological curves (Asthma, SAI, and PSAI) match the pattern of distal obstruction, which represents the obstruction of the small airways with different degrees of severity. The resistance pattern for the Normal category also matches the typical linear resistance pattern for a normal condition, whereas the Asthma class presented the smallest airways reactance at all frequencies recorded, the Normal curve exhibited the highest airways reactance. Comparing these curves with typical reactance curves, we can see that the pathological patterns (Asthma, SAI, and PSAI) also match the reactance pattern of distal obstruction. The reactance curve for the Normal category matches the typical reactance curve for a normal condition.

Once the typical functions for each class were obtained, the actual behavior of the

patients' resistance curves was compared to the function obtained for each class. For this graphical comparison, the minimum, middle and maximum curve data sets described in Section 6.1 were plotted along with the corresponding typical function of each class. Figure 6.6 shows the graphical comparison.

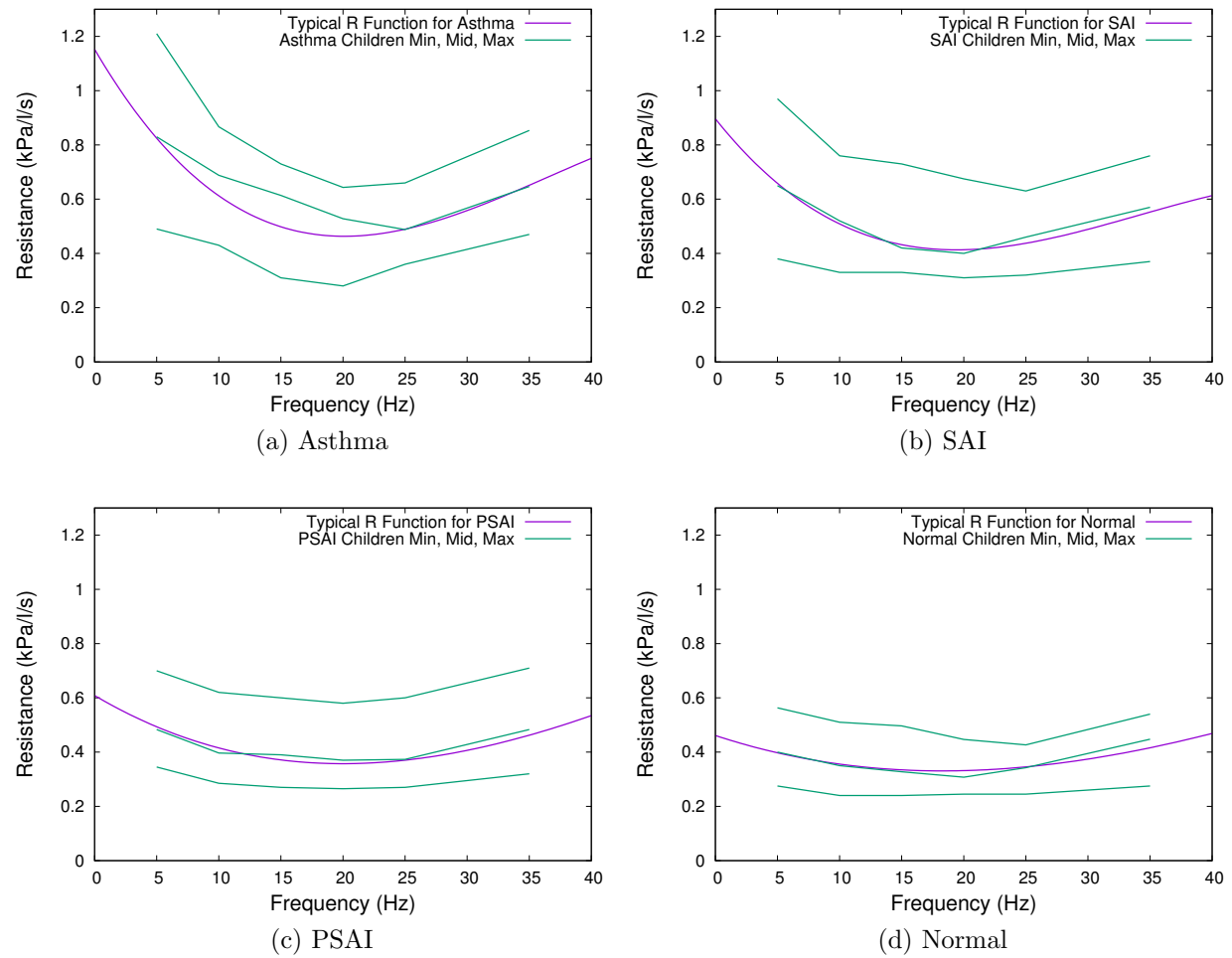


Figure 6.6: Actual Patients' Curves Maximum, Middle and Minimum vs Typical Class Function for Asthma (a), SAI (b), PSAI (c), and Normal (d)

It is observed that the actual maximum, middle and minimum curves of each class present a similar behavior (pattern) than the typical class function. Therefore, these cubic typical functions could be used as reference models for future classification work because no intersection occurs between them and the actual patients' curves seem to match the

typical function pattern.

6.2.4 Typical Functions for Area and Slope

Further observing Figure 6.4, we can see that slope and area under the curve can potentially help with the differentiation between classes; see Figure 6.7.

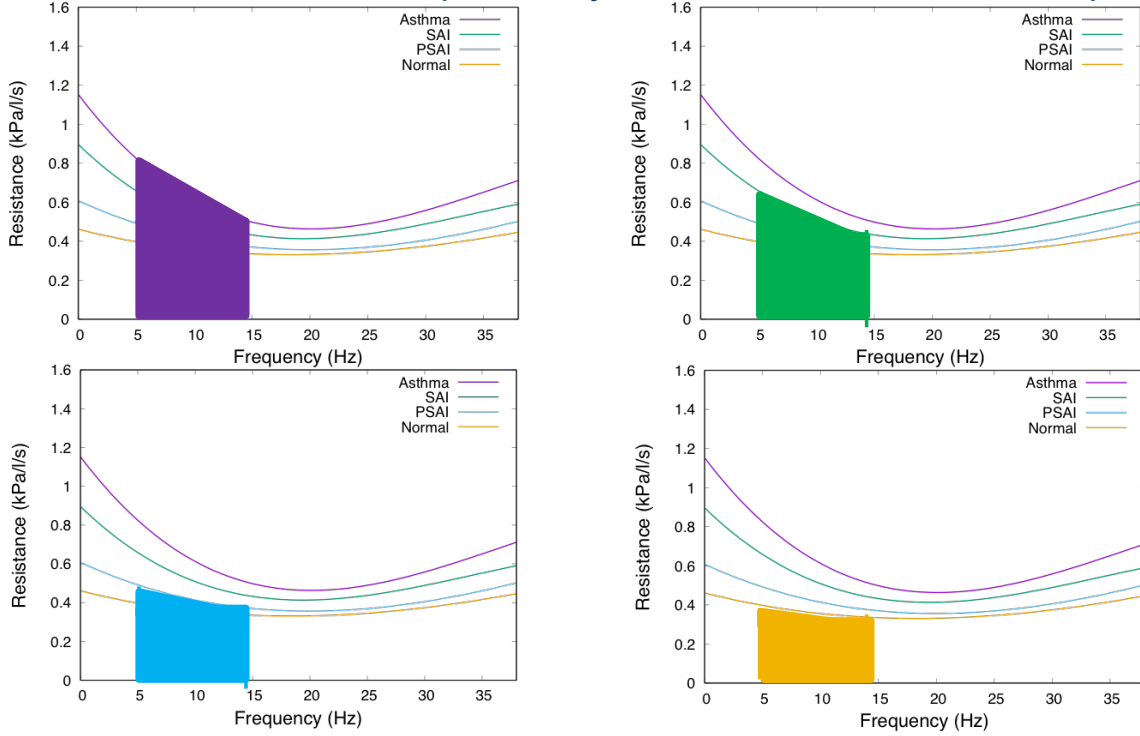


Figure 6.7: Potential Geometrical Features of Resistance Curves for Asthma, SAI, PSAI, and Normal.

Typical class functions for the area under the Resistance $R_d(f)$ curve were obtained based on the following formulas:

$$I_{R,a}(f) = \int R_a(f) df \quad (6.19)$$

$$I_{R,s}(f) = \int R_s(f) df \quad (6.20)$$

$$I_{R,p}(f) = \int R_p(f) df \quad (6.21)$$

$$I_{R,n}(f) = \int R_n(f) df \quad (6.22)$$

Typical class functions for the area under the Reactance $X_d(f)$ curve were obtained based on the following formulas:

$$I_{X,a}(f) = \int X_a(f) df \quad (6.23)$$

$$I_{X,s}(f) = \int X_s(f) df \quad (6.24)$$

$$I_{X,p}(f) = \int X_p(f) df \quad (6.25)$$

$$I_{X,n}(f) = \int X_n(f) df \quad (6.26)$$

Typical class functions for the slope of the Resistance $R_d(f)$ curve were obtained based on the following formulas:

$$D_{R,a}(f) = \frac{dR_a(f)}{df} \quad (6.27)$$

$$D_{R,s}(f) = \frac{dR_s(f)}{df} \quad (6.28)$$

$$D_{R,p}(f) = \frac{dR_p(f)}{df} \quad (6.29)$$

$$D_{R,n}(f) = \frac{dR_n(f)}{df} \quad (6.30)$$

Typical class functions for the slope of the Reactance $R_d(f)$ curve were obtained based on the following formulas:

$$D_{X,a}(f) = \frac{dX_a(f)}{df} \quad (6.31)$$

$$D_{X,s}(f) = \frac{dX_s(f)}{df} \quad (6.32)$$

$$D_{X,p}(f) = \frac{dX_p(f)}{df} \quad (6.33)$$

$$D_{X,n}(f) = \frac{dX_n(f)}{df} \quad (6.34)$$

In summary, we have 24 functions that describe the geometrical patterns of each of the classes studied:

- 4 functions (one per class) describing the Typical Resistance Function $R_d(f)$.
- 4 functions (one per class) describing the Typical Resistance Area Function $I_{R,d}(f)$.
- 4 functions (one per class) describing the Typical Resistance Slope Function $D_{Rd}(f)$.
- 4 functions (one per class) describing the Typical Reactance Function $X_d(f)$.
- 4 functions (one per class) describing the Typical Reactance Area Function $I_{X,d}(f)$.
- 4 functions (one per class) describing the Typical Reactance Slope Function $D_{Xd}(f)$.

6.3 Similarity/Dissimilarity Measures

How can we use the 24 functions that describe geometrical features? We could compare the typical function of each class against the actual patient's curve. To do so, we need to be consistent with the pre-processing approach previously used for typical functions. Therefore, the same methodology as the typical functions was followed. Cubic polynomial

regressions were performed for each patient's data set (112 typical functions were obtained, 112 slope functions and 112 area functions). Each of the subject's functions obtained were graphically compared with each of the typical functions. This comparison was done to visually examine the behavior of each child against the four different models, and make conclusions about the similarities between individual and typical functions. Figure 6.8 illustrates the graphical comparisons performed for the Resistance and Reactance curves for an asthmatic patient.

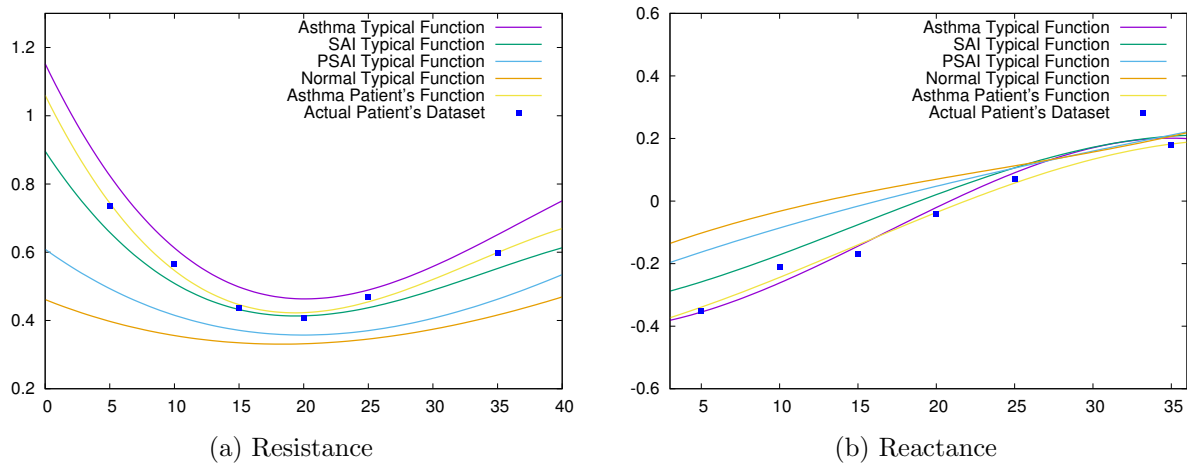


Figure 6.8: Asthmatic Child vs Typical Resistance (a) and Reactance (b) Functions

When each of the 112 individual functions was compared to each of the typical functions, it was observed that there is a considerable similarity between the subject's behavior and the typical functions where the patient belongs. Additionally, it was observed that the similarity within the range from R5 to R15 was consistent when the individual model was compared to the reference model. This finding is congruent with previous research work, where R5 to R15 has been identified as potential discriminant features.

6.3.1 Similarity Measure Estimation

In order to estimate similarity/dissimilarity measures based on the functions developed, let's denote each of the patient's functions as: $R_i(f)$, $I_{R,i}(f)$, and $D_{R,i}(f)$ for resistance; and $X_i(f)$, $I_{X,i}(f)$, and $D_{X,i}(f)$ for reactance.

As was mentioned earlier, the most informative part of the IOS results correspond to the 5-15 Hz range. The central point of this range is the value 10 Hz. It is reasonable to take the difference $\Delta R_{id} \stackrel{\text{def}}{=} R_d(10) - R_i(10)$ of the values corresponding to this central frequency as an estimate for a constant difference. Thus, we should compare the typical function $R_d(f)$ not to the actual patient's function $R_i(f)$, but with the "shifted" function $R_i(f) + \Delta R_{id}$, shifted to best match the typical function. Same analogy applies for all other patient's functions $I_{R,i}(f)$, $D_{R,i}(f)$, $X_i(f)$, $I_{X,i}(f)$, and $D_{X,i}(f)$.

Therefore, the similarity measures for each of the classes corresponding to the comparison of the patient's functions vs typical functions are:

Resistance:

Similarity Index for Typical Functions:

$$S_{R_d} = \int_5^{15} |R_d(f) - (R_i(f) + \Delta R_{id})| df, \quad \text{where} \quad \Delta R_{id} \stackrel{\text{def}}{=} R_d(10) - R_i(10) \quad (6.35)$$

Similarity Index for Area Functions:

$$S_{I_{R,d}} = \int_5^{15} |I_{R,d}(f) - (I_{R,i}(f) + \Delta I_{Rid})| df, \quad \text{where} \quad \Delta I_{Rid} \stackrel{\text{def}}{=} I_{R,d}(10) - I_{R,i}(10) \quad (6.36)$$

Similarity Index for Slope Functions:

$$S_{D_{R,d}} = \int_5^{15} |D_{R,d}(f) - (D_{R,i}(f) + \Delta D_{Rid})| df, \quad \text{where} \quad \Delta D_{Rid} \stackrel{\text{def}}{=} D_{R,d}(10) - D_{R,i}(10) \quad (6.37)$$

Reactance:

Similarity Index for Typical Functions:

$$S_{X_d} = \int_5^{15} |X_d(f) - (X_i(f) + \Delta X_{id})| df, \quad \text{where} \quad \Delta X_{id} \stackrel{\text{def}}{=} X_d(10) - X_i(10) \quad (6.38)$$

Similarity Index for Area Functions:

$$S_{I_{X,d}} = \int_5^{15} |I_{X,d}(f) - (I_{X,i}(f) + \Delta I_{X_{id}})| df, \quad \text{where} \quad \Delta I_{X_{id}} \stackrel{\text{def}}{=} I_{X,d}(10) - I_{X,i}(10) \quad (6.39)$$

Similarity Index for Slope Functions:

$$S_{D_{X,d}} = \int_5^{15} |D_{X,d}(f) - (D_{X,i}(f) + \Delta D_{X_{id}})| df, \quad \text{where} \quad \Delta D_{X_{id}} \stackrel{\text{def}}{=} D_{X,d}(10) - D_{X,i}(10) \quad (6.40)$$

From each of the above formulas, there are 4 resulting values, one for each of the classes studied here (Asthma, SAI, PSAI and Normal). Therefore, in total, there are 24 similarity features derived from the data Pre-Processing approach so far. In summary, these similarity measures compare the form of the functions by summing the difference of different data points between the patient's function and typical function. The closer the value is to 0, the higher the similarity between curves. The larger the result, the higher the dissimilarity. The previously described translation of the curves provides a more robust similarity measure.

To prove that the similarity measures behave as expected, the typical functions for each condition d were compared against each other. It is observed that when comparing the same class typical functions, the similarity index is zero, and it increases according to the dissimilarity against the condition compared; see Table 6.3.

Table 6.3: Similarity Measures of Typical Class Functions

Condition	Asthma	SAI	PSAI	Normal
Asthma	0	1.02624	2.49132	3.07514
SAI	1.02624	0	1.46508	2.0489
PSAI	2.49132	1.46508	0	0.583818
Normal	3.07514	2.0489	0.583818	0

Data Generation

Now that similarity measures have been defined, an extensive amount of data can be used. In order to calculate the previously explained similarity measures, data points were generated based on the different typical functions and patient's functions. As previously noted, the range selected for the data generation was between R5 and R15, which was the range that presented the major consistency in terms of similarity. The resolution chosen were frequency increments of 0.05; through experimentation, it was observed that smaller resolutions did not make a difference. As discussed before, 24 similarity resulting values were obtained.

Graphical Assessment of Similarity Measures

Once similarity measures were obtained, these were plotted per condition and are shown in Figures C.1 and D.1 in Appendices C and D. It is observed that the similarity measures could provide more separability between classes, specially for the Normal class, since a clear separation between classes is observed in Figures 6.9 and 6.10.

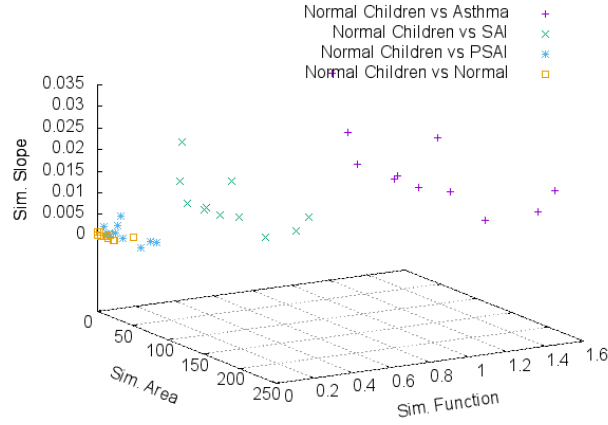


Figure 6.9: Similarity Measures of Resistance for Normal Class.

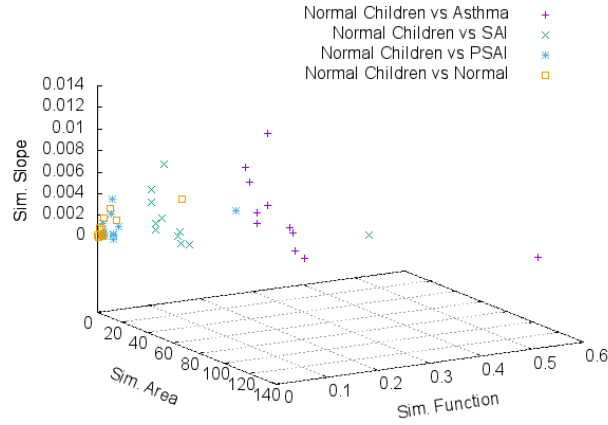


Figure 6.10: Similarity Measures of Reactance for Normal Class.

6.4 Preliminary Classification using Similarity Measures

The 24 similarity features derived from the pre-processing approach seem to be potentially discriminative. Therefore, to evaluate their discriminative capacity, these 24 features were

used as input features in preliminary ANN classification. The development of the ANN classifiers included the use of the similarity measures derived from 1) Resistance (R) (12 features), 2) Reactance (X) (12 features), and 3) combination of both R and X (24 features).

In the development of each of the above described ANN classifiers, several experiments were performed, where different values of the ANN parameters (training error and number of hidden neurons) were tested in order to determine the ANN parameters that provided the best possible classification. A design of experiments approach was taken to evaluate the parameters, and about 27 different combinations were tested to find the optimal ANN in each case. As described in Section 4.1, 75% of the data sets were used to train each of the ANN and 25% of the data was used for validation of the ANN. Table 6.4 shows the best results found for each of the ANN in the classification of the 4 conditions: Asthma, SAI, PSAI and Normal.

Table 6.4: Preliminary ANN Classification Results using Similarity Measures.

Type of feature	# Input Features	# Classes	Classes	Training Samples	Validation Samples	Validation Accuracy (%)
Pre-Processed (R)	12	4	Asthma, SAI, PSAI , Normal	214	74	63.7
Pre-Processed (X)	12	4	Asthma, SAI, PSAI , Normal	214	74	68.91
Pre-Processed (R and X combined)	24	4	Asthma, SAD, PSAI , Normal	214	74	71.62

The results obtained show that the best classification accuracy was achieved when using both R and X similarity measures. However, these results are not optimal, it may be because too many input features (up to 24) were used. Therefore, a reduction in the number of classifier's input features might improve the performance of the ANN.

6.5 Discriminative Similarity Measures

To find the most discriminative similarity measures, a statistical analysis was performed using the same methodology described in Chapter 5. In this analysis, the values of each similarity measure for each class were compared against the same similarity measure values for a different class, until the parameter was compared for all classes. A comparison-matrix for each of the similarity measures was completed, where the p-values obtained for each of the comparisons for Resistance similarity measures were listed Table 6.5 and for Reactance similarity measures were listed in Table 6.5. For a similarity parameter to be considered potentially discriminative of its p-values must be less than 0.05. The corresponding tables have all cells where the p-values are greater than 0.05 grayed out, which would render that specific similarity measure undiscriminative.

Table 6.5: Resistance Similarity Measures Comparison-Matrix.

Similarity Indices:		Asthma vs. SAI	Asthma vs. PSAI	Asthma vs. Normal	SAI vs. PSAI	SAI vs. Normal	PSAI vs. Normal
Typical Function (R)	S_{Ra}	0.042	0.000	0.000	0.000	0.000	0.000
	S_{Rs}	0.000	0.237	0.014	0.000	0.000	0.000
	S_{Rp}	0.000	0.000	0.000	0.000	0.000	0.801
	S_{Rn}	0.000	0.000	0.000	0.000	0.000	0.000
Area Function (R)	S_{IRa}	0.000	0.000	0.000	0.000	0.000	0.008
	S_{IRs}	0.000	0.000	0.240	0.137	0.000	0.005
	S_{IRp}	0.000	0.000	0.000	0.002	0.006	0.204
	S_{IRn}	0.000	0.000	0.000	0.000	0.000	0.033
Slope Function (R)	S_{DRa}	0.419	0.030	0.000	0.067	0.000	0.010
	S_{DRs}	0.403	0.419	0.818	0.850	0.332	0.091
	S_{DRp}	0.032	0.000	0.000	0.000	0.000	0.101
	S_{DRn}	0.024	0.000	0.000	0.000	0.000	0.019

Table 6.6: Reactance Similarity Measures Comparison-Matrix.

Similarity Indices:		Asthma vs. SAI	Asthma vs. PSAI	Asthma vs. Normal	SAI vs. PSAI	SAI vs. Normal	PSAI vs. Normal
Typical Function (X)	S_{Xa}	0.220	0.773	0.507	0.148	0.071	0.611
	S_{Xs}	0.019	0.133	0.027	0.810	0.680	0.456
	S_{Xp}	0.001	0.001	0.000	0.189	0.000	0.002
	S_{Xn}	0.001	0.000	0.000	0.036	0.000	0.002
Area Function (X)	S_{IXa}	0.006	0.000	0.000	0.000	0.000	0.000
	S_{IXs}	0.000	0.071	0.007	0.000	0.000	0.000
	S_{IXp}	0.000	0.000	0.000	0.000	0.001	0.254
	S_{IXn}	0.000	0.000	0.000	0.000	0.000	0.000
Slope Function (X)	S_{DXa}	0.629	0.118	0.002	0.051	0.001	0.135
	S_{DXs}	0.064	0.359	0.581	0.546	0.435	0.786
	S_{DXp}	0.006	0.000	0.000	0.073	0.009	0.209
	S_{DXn}	0.006	0.000	0.000	0.017	0.000	0.072

Another consideration is the fact that the similarity measures for a specific characteristic are related. This was previously explained and can further be observed in Table 6.3. In other words, similarity measures for typical functions are related to each other, the same applies to the area functions, and slope functions. Therefore, it makes sense to select all the similarity measures when most of its p-values are less than 0.05 for a specific type of function, in order to avoid losing information in the relationship between similarity measures for that specific type of function. Based on this approach, the parameters that were found to be potentially discriminative are the similarity measures related to the Typical Function and the Area Function of Resistance, and the Area Function for Reactance.

6.6 Conclusions

Based on the analysis performed in this chapter, 12 features derived from the IOS data were obtained using a data pre-processing approach and were considered potentially discriminative to classify Asthma, SAI, PSAI and Normal lung functions. These features shall

be used in further classification work: $S_{R_a}, S_{R_s}, S_{R_p}, S_{R_n}, S_{I_{R,a}}, S_{I_{R,s}}, S_{I_{R,p}}, S_{I_{R,n}}, S_{I_{X,a}}, S_{I_{X,s}}, S_{I_{X,p}},$ and $S_{I_{X,n}}.$

Chapter 7

Results and Discussion

7.1 Baseline Classification Work

Our first attempt to classify Asthma, SAI, PSAI, and Normal lung function, was done using machine learning techniques. In this preliminary work, all the Resistance and Reactance IOS were directly used (without pre-processing) as input features of the classifier. For this particular work, only the training phase of the classification was completed, as the results obtained in terms of accuracy were very low. These results are described in Table 7.1

Table 7.1: First Classification Results

Clinician Classification		Algorithm Classification
Condition	Number of Subjects	RX Classifier
Asthma	30	86.6%
SAD	54	44.4%
Mild SAD	17	35.3%
Normal	11	0.0%

After doing this first classification attempt, it was concluded that further work needed to be done to best understand the IOS parameters and to find the most discriminative ones.

7.2 ANN Classification

7.2.1 ANN Classification - First Stage

Let's summarize the results obtained in Chapters 5 and 6 for feature selection. The different types of IOS derived features that were identified as discriminative to distinguish between

the four different degrees of respiratory small airways function in children (Asthma, SAI, PSAI, and Normal lung function) were:

- 1) From the Conventional approach (Direct IOS features): F_{res} , R_5 , X_5 , AX , R_5-R_{20} , X_{10} , and X_{15} .
- 2) From Resistance Pre-Processing: S_{R_a} , S_{R_s} , S_{R_p} , S_{R_n} , $S_{I_{R,a}}$, $S_{I_{R,s}}$, $S_{I_{R,p}}$, and $S_{I_{R,n}}$.
- 3) From Reactance Pre-Processing: $S_{I_{X,a}}$, $S_{I_{X,s}}$, $S_{I_{X,p}}$, and $S_{I_{X,n}}$.

These 19 features were used as input features in ANN classification. The development of the ANN classifiers included the use of the similarity measures as input features as described above, and also included the combinations of the different types of features.

In the development of each of the above described ANN classifiers, several experiments were performed, where different values of the ANN parameters (training error and number of hidden neurons) were tested in order to determine the ANN parameters that provided the best possible classification. A design of experiments approach was taken to evaluate the parameters, and about 27 different combinations were tested to find the optimal ANN in each case. As described in Section 4.1, 75% of the data sets were used to train each of the ANN and 25% of the data was used for validation of the ANN. Table 7.2 shows the best results found for each of the ANNs in the classification of the 4 conditions: Asthma, SAI, PSAI, and Normal.

Table 7.2: ANN Classification Results for Asthma, SAI, PSAI, and Normal.

Type of Feature	# Features	# Classes	Classes	Training Samples	Validation Samples	Hidden Neurons	Training Error	Validation Accuracy (%)
Conventional (IOS)	7	4	Asthma, SAI, PSAI , Normal	214	74	28	0.035	75.67
Pre-Processed R	12	4	Asthma, SAI, PSAI , Normal	214	74	25	0.035	63.7
Pre-Processed X	12	4	Asthma, SAI, PSAI , Normal	214	74	30	0.045	68.91
Pre-Processed R & X	24	4	Asthma, SAI, PSAI , Normal	214	74	23	0.01	71.62
Conventional (IOS) & Pre-Processed R	19	4	Asthma, SAI, PSAI , Normal	214	74	25	0.02	64.86
Conventional (IOS) & Pre-Processed X	19	4	Asthma, SAI, PSAI , Normal	214	74	28	0.035	63.51
Conventional (IOS) & Pre-Processed R & X	31	4	Asthma, SAI, PSAI , Normal	214	74	25	0.035	71.62

The obtained results show that the best classification accuracy (75.67%) was achieved when using only IOS direct features derived from the Conventional approach.

It is important to emphasize that no previous validated work has been done to distinguish between 4 different levels of small airways lung function. Therefore, it is not possible to compare these results with a previous work. However, from the literature review presented in Chapter 2, we could compare it to the closest work in classifying peripheral diseases such the ones studied in here, the work performed by Barua et. al [50]. In this study, the authors obtained a classification accuracy of 61.53% in separating peripheral diseases from central diseases. Even though our results have a better classification accuracy, these results are still not optimal.

7.2.2 ANN Classification – Second Stage

Based on the previous classification results, another classification attempt was done. This time to classify only the groups of Asthma versus Normal. This new approach was taken to evaluate the classification performance using the IOS derived features previously selected and compare if there was any improvement against what was previously done by other authors. This could be a good reference since most of the classification studies performed

so far distinguished Asthma from Healthy. Table 7.3 shows the best results found for each of the ANNs in the classification of the 2 conditions: Asthma and Normal.

Table 7.3: ANN Classification Results for Asthma vs. Normal.

Type of Feature	# Features	# Classes	Classes	Training Samples	Validation Samples	Hidden Neurons	Training Error	Validation Accuracy (%)
Conventional (IOS)	7	2	Asthma, Normal	81	28	10	0.01	100
Pre-Processed R	8	2	Asthma, Normal	81	28	15	0.001	100

By classifying Asthma vs. Normal, perfect results were obtained for both scenarios, the Conventional and Resistance Pre-Processed approaches. In both cases, an accuracy of 100% was achieved. These results were compared to previous work. Since our classification is based on static lung function assessment, the first comparison was done against those works that studied these two conditions (Asthma and Healthy) and had results using only static features. To this end, the overall classifier's accuracy results for Badnjevic et al. studies [44, 47] after the static assessment were between 42-52%. Therefore, our classifier results are better than Badnjevic's previous work that used a static assessment of lung function. Now, let us compare against the best classification performance obtained in the differentiation of Asthma vs Healthy, independently to the type of lung function assessment used (static and/or dynamic). To this end, the studies performed by Badnjevic et al. obtained accuracy results ranging from 92-98.20% when both static and dynamic assessments were used. With these results, we can conclude that this work is better than any results from the previous classification work that attempted to distinguish Asthma from Normal lung function. However, this research is not focused only on the differentiation of these two classes, the objective is to differentiate 4 different levels of small airways lung function. Therefore further analysis is required to improve the performance of the classification of multiple classes.

Note that classification using Reactance Pre-Processed features or the combination of different types of features was not done, as the only objective was to prove the performance by classifying Asthma vs. Healthy for an initial assessment.

7.2.3 ANN Classification - Final Stage

Per previous assessment, it was observed that the performance of the ANN classifier greatly improved when analyzing two classes instead of 4. Now, the new attempt to classify the different classes studied in here is to have multiple ANN bi-class classifiers. In other words, an ANN will be used to classify Normal vs. Peripheral Lung Dysfunction (Asthma, SAI, and PSAI), another ANN to classify PSAI vs. Severe Peripheral Lung Dysfunction (Asthma and SAI), and the last ANN to classify Asthma vs. SAI.

ANN1: Normal vs. Peripheral Lung Dysfunction

Given the results in Table 7.4, it can be observed that the best classification result was obtained when Resistance Pre-Processed derived features were used. The purpose of the neural network was to separate patients with Normal lung function from patients with peripheral lung dysfunction. During the training phase of this ANN, the best result was obtained when 50 neurons in the hidden layer and a training error of 0.001 were used. On the validation data set, this neural network achieved 100% accuracy on the validation set: all 74 cases were classified correctly. These results are better in performance than any of the classification results described in Chapter 2. This neural network was denoted by ANN1; it was used for the final ANN classifier.

Table 7.4: ANN Classification Results Normal vs. Peripheral Airways Obstruction.

Type of Feature	# Features	# Classes	Classes	Training Samples	Validation Samples	Hidden Neurons	Training Error	Validation Accuracy (%)
Conventional (IOS)	7	2	1: Asthma, SAI, PSAI 2: Normal	214	74	15	0.001	95.94
Pre-Processed R	8	2	1: Asthma, SAI, PSAI 2: Normal	214	74	50	0.001	100
Pre-Processed X	4	2	1: Asthma, SAI, PSAI 2: Normal	214	74	10	0.005	93.24
Pre-Processed R & X	12	2	1: Asthma, SAI, PSAI 2: Normal	214	74	15	0.001	95.94

ANN2: PSAI vs. Severe Peripheral Lung Dysfunction

Given the results in Table 7.5, it is observed that the best classification result was obtained when direct IOS features derived from the Conventional approach were used. The purpose of this neural network was to separate patients with PSAI from patients with severe peripheral lung dysfunction. During the training phase of this ANN, the best result was obtained when 10 neurons in the hidden layer and a training error of 0.02 were used. On the validation data set, this neural network achieved 95.45% accuracy on the validation set: 63 cases out of 66 were classified correctly. This neural network was denoted as ANN2 and was used for the final ANN classifier.

Table 7.5: ANN Classification Results PSAI vs. Peripheral Airways Obstruction.

Type of Feature	# Features	# Classes	Classes	Training Samples	Validation Samples	Hidden Neurons	Training Error	Validation Accuracy (%)
Conventional (IOS)	7	2	1: Asthma, SAI 2: PSAI	191	66	10	0.02	95.45
Pre-Processed R	8	2	1: Asthma, SAI 2: PSAI	191	66	15	0.005	87.87
Pre-Processed R & X	12	2	1: Asthma, SAI 2: PSAI	191	66	20	0.001	89.39
Conventional (IOS) & Pre-Processed R	15	2	1: Asthma, SAI 2: PSAI	191	66	15	0.001	89.39
Conventional (IOS) & Pre-Processed R & X	19	2	1: Asthma, SAI 2: PSAI	191	66	15	0.01	92.42

ANN3: SAI vs. Asthma

Given the results in Table 7.6, it is observed that the best classification result was obtained when direct IOS features derived from the Conventional approach were used. The purpose of this neural network was to separate patients with SAI from patients with Asthma. During the training phase of this ANN, the best result was obtained when 35 neurons in the hidden layer and a training error of 0.1 were used. On the validation data set, this neural network achieved 92.73% accuracy on the validation set: 51 cases out of 55 were classified correctly. This neural network was denoted as ANN3 and was used for the final ANN classifier.

Table 7.6: ANN Classification Results SAI vs. Asthma.

Type of Feature	# Features	# Classes	Classes	Training Samples	Validation Samples	Hidden Neurons	Training Error	Validation Accuracy (%)
Conventional (IOS)	7	2	Asthma, SAD	160	55	35	0.1	92.73
Pre-Processed R & X	12	2	Asthma, SAD	160	55	15	0.05	81.81

Final ANN

Figure 7.1 shows the final classification system, it includes the 3 ANN previously selected: ANN1, ANN2 and ANN3. In all the ANN the training algorithm used was the iRPROP and the activation function used for the hidden and output neurons was the Symmetric Sigmoid Function, which is the most commonly used activation function in ANNs.

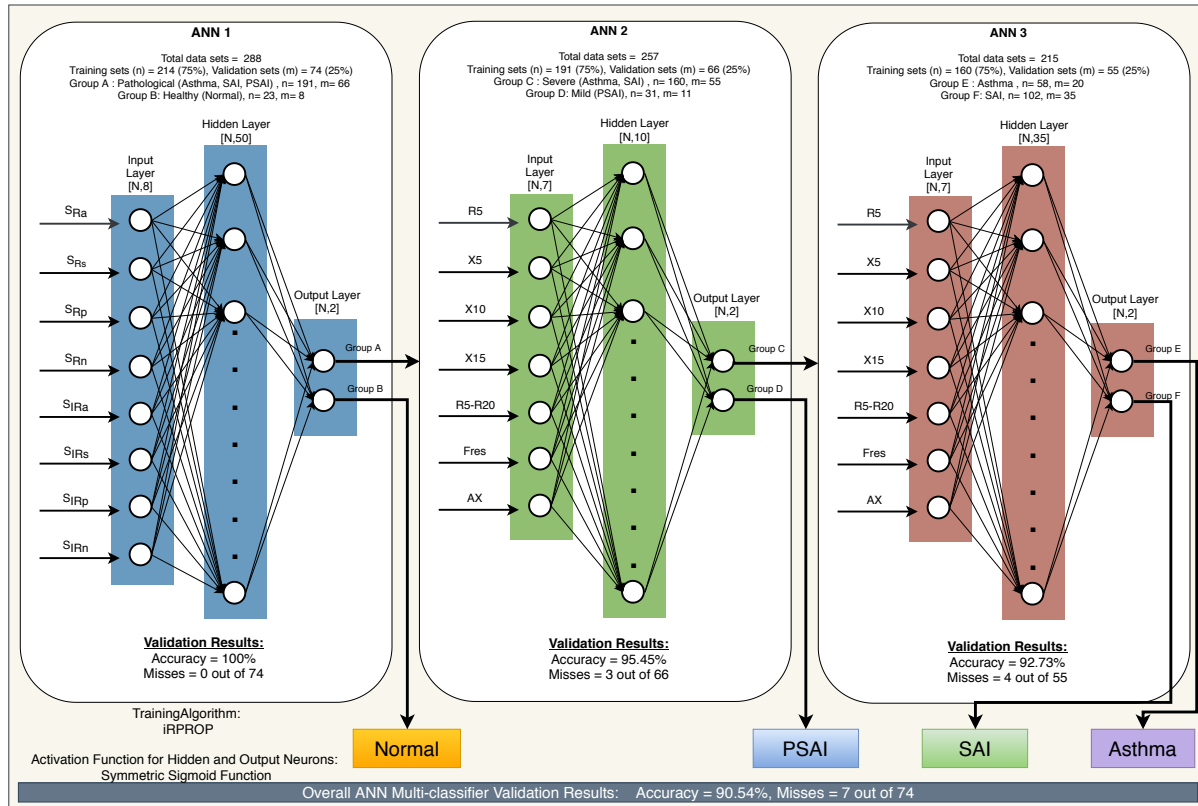


Figure 7.1: Final ANN Structure

Overall the system is 100% accurate to detect peripheral lung dysfunctions and 92% to

95% detect the specific type of peripheral lung dysfunction.

The ANN1 uses the following 8 input features derived from pre-processing approach for Resistance: S_{Ra} , S_{Rs} , S_{Rp} , S_{Rn} , $S_{I_{R,a}}$, $S_{I_{R,s}}$, $S_{I_{R,p}}$, and $S_{I_{R,n}}$.

ANN2 y ANN2 use the 7 features derived from the Conventional approach: Fres, R5, X5, AX, R5-R20, X10, and X15.

The ANN classifier was written in C++, the algorithm for the programming code can be found in Figure E.1, Appendix E.

ANN Sensitivity and Specificity

The diagnosis system developed was evaluated in terms of sensitivity and specificity to determine how well the system identifies true positives and true negatives.

Sensitivity and Specificity are defined as:

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

where

TP = True Possitives

FN = False Negatives

TN = True Negatives

FP = False Positives;

see Tables 7.7 to 7.9 for calculations and results.

Table 7.7: TP, FN, TN, and FP Table

		Actual Condition	
		+	-
Diagnosis	+	TP	FP
	-	FN	TN

Table 7.8: Sensitivity and Specificity of ANN1 and ANN2

ANN 1				ANN 2			
		Actual Condition				Actual Condition	
		Healthy	Peripheral Dysfunction			PSAI	Severe Dysfunction
Diagnosis	Healthy	66	0	Diagnosis	PSAI	8	0
	Peripheral Dysfunction	0	8		Severe Dysfunction	3	55
Sensitivity =		100%		Sensitivity =		73%	
Specificity=		100%		Specificity=		100%	

Table 7.9: Sensitivity and Specificity of ANN3

ANN 3							
		Actual Condition				Actual Condition	
		SAI	Asthma			Asthma	SAI
Diagnosis	SAI	35	4	Diagnosis	Asthma	16	0
	Asthma	0	16		SAI	4	35
Sensitivity =		100%		Sensitivity =		80%	
Specificity=		80%		Specificity=		100%	

It is concluded that the presented ANN classifier is accurate, sensitive and specific to identify Asthma, SAI, PSAI, and Normal lung functions.

Chapter 8

Conclusions

8.1 Research Conclusions

Based on the research questions and hypothesis testing defined in Chapter 3, the following is concluded:

8.1.1 Hypothesis Testing

Hypothesis 1

Conventional-derived IOS features reliably classify Asthmatic, SAI, PSAI, and Normal children based on Artificial Neural Networks (ANN) derived algorithms.

In this work, the IOS features selected through the conventional approach, which is the one derived from the study of the Resistance and Reactance direct IOS measurements, were found the best to reliably classify the specific type of small airways obstruction. The 7 features that accurately classified PSAI, SAI and Asthma dysfunctions were: Fres, R5, X5, AX, R5-R20, X10, and X15. On the other hand, these features were not found as sensitive as the pre-processed derived features to distinguish Normal lung function (75% vs 100%).

Hypothesis 2

Data pre-processing based on IOS resistance values reliably classify Asthmatic, SAI, PSAI, and Normal children based on Artificial Neural Networks (ANN) derived algorithms.

The Resistance features derived from the pre-processing approach were found to be the ones with the highest discriminative capacity to distinguish Normal lung function from

small airways obstruction, the accuracy, sensitivity and specificity obtained for this classification was 100%. However, for classifying the specific type of small airways obstruction, the pre-processed derived features presented a reduced performance when compared to the IOS conventionally-derived features. In total, 8 pre-processed features were discriminative. Specifically, these were the similarity measures derived when the Resistance typical functions and area typical functions were compared to the actual patient's curve and area functions (S_{R_a} , S_{R_s} , S_{R_p} , S_{R_n} , $S_{I_{R,a}}$, $S_{I_{R,s}}$, $S_{I_{R,p}}$, and $S_{I_{R,n}}$).

Hypothesis 3

Data pre-processing based on IOS reactance values reliably classify Asthmatic, SAI, PSAI, and Normal children based on ANN derived algorithms.

The Reactance parameters derived from the pre-processing approach were found less discriminative than conventionally-derived features or Resistance pre-processed features, therefore they were not used in the neural networks as their use reduced the performance of the ANN.

Hypothesis 4

Performance of data pre-processing-based classification of Asthmatic, SAI, PSAI, and Normal improves when using features based on both IOS resistance and reactance values.

The data pre-processing-based classification did not improve when using IOS resistance pre-processed and reactance pre-processed values. Adversely, the performance of the ANNs usually worsened when both types of features were used.

Hypothesis 5

The best classification performance is achieved when using IOS discriminative features derived from both the conventional and data pre-processing approaches.

Indeed, the best classification performance was achieved when IOS discriminative features derived from both the conventional and data pre-processing approaches were used. To

be more specific, the classification improved when using specifically conventionally-derived and Resistance data pre-processed features. In addition, the best classification performance was achieved when multiple bi-class ANNs classifiers were used instead of one multi-class ANN classifier.

8.1.2 Research Questions

Research Question 1

What are the IOS derived features that best classify respiratory small airway function in children?

In total, 15 features were selected based on the Conventional and Pre-processing approach presented and the performance of different ANN classifiers. These features included 7 features derived from the study of the Resistance and Reactance direct IOS measurements (Conventional approach) and 8 derived from the pre-processing approach that included pre-processed features (similarity measures) based on the Resistance typical functions for the different classes, as well as the Resistance typical area functions.

The IOS derived features that were identified as discriminative to distinguish between the four different degrees of respiratory small airways function in children (Asthma, SAI, PSAI, and Normal lung function) were:

From the Conventional approach: F_{res} , R_5 , X_5 , AX , R_5-R_{20} , X_{10} , and X_{15} .

From the Pre-processing approach: S_{R_a} , S_{R_s} , S_{R_p} , S_{R_n} , $S_{I_{R,a}}$, $S_{I_{R,s}}$, $S_{I_{R,p}}$, and $S_{I_{R,n}}$.

Research Question 2

Does computer classification of IOS data improve the diagnostic utility of this child-friendly lung function testing?

Indeed, computer-aided classification improves the diagnostic utility of the IOS technique. The computational algorithms developed will help clinicians to interpret IOS data and potentially will improve the IOS acceptance in the medical field.

In summary, a diagnostic support system with high discriminative capacity in terms of accuracy, sensitivity, and specificity was developed. The system is 100% accurate, sensitive and specific to classify Normal function vs. small airways dysfunction, and 92%- 95% accurate, 73%-100% sensitive, and 80%-100% specific for classifying a specific type of small airways dysfunction.

The resulting classification algorithm performs better than any of the previously proposed classification algorithms that use IOS features. Another advantage of this system is relies only on IOS testing, while other classifiers are based on multiple pulmonary tests (IOS, Spirometry, BPT, and BDT). This advantage is important because only IOS tests are reliable for small children.

8.2 Novelty

Biomedical Novelty

1. This work presented the first successful algorithm for enhancing diagnostics of different degrees of small airways obstruction: Asthma, SAI, PSAI, and Normal lung function.

Computational Novelty

1. This work presented the use of innovative pre-processing techniques in machine learning: statistical and scale-invariance-based.
2. This is the first research work to assess lung function using IOS curve-shape-derived features.

8.3 Significance of the Result- Contribution to Society

1. This research will assist clinicians with a reliable and proven method for accurate classification of children's lung function.
2. This research improves the clinical utility of the IOS since the IOS test would be less dependent on highly specialized doctors.
3. This research allows the diagnosis and also the timely control of asthma. This could potentially result in the improvement of the quality of life of asthmatic children and their parents.
4. This research is relevant for those Mexican individuals who immigrated at an early age and whose asthma risk has increased over time they lived in the United States. It is foreseen that this research could facilitate the implementation of IOS tests in community clinics, since the test would be less dependent on highly specialized doctors. This research could potentially reduce health disparities that could be caused by the limited health care access by allowing vulnerable populations, such e.g., immigrants and their children, better access to monitoring asthma in a timely fashion, before complications that merit emergency health care arise.
5. This research could potentially reduce the health care expenditures (8 billion dollars per year) by providing a reliable screening tool for asthma control.

8.4 Future Work

The computational algorithms developed will need to be tested at a greater scale to confirm its utility and enhance it as required. This could be done by collaborating with the National Institute of Respiratory Diseases (INER) in Mexico and the National Jewish Health Institute in Denver, Colorado. Both institutions are dedicated to medical research and

treatment of patients with respiratory, cardiac, immune and related disorders and have shown interest in continuing this research.

Additionally, the scope of this research work can be increased by studying other populations and other pulmonary conditions such as Chronic Obstructive Pulmonary Disease (COPD) and pulmonary hypertension. IOS also provides important information about small airways obstruction in these two conditions.

References

- [1] Bonini M, Usmani OS. The role of the small airways in the pathophysiology of asthma and chronic obstructive pulmonary disease. *Therapeutic Advances in Respiratory Disease*. 2015;9(6):281-293. DOI: 10.1177/ 1753465815588064.
- [2] McNulty W, Usmany OS. Techniques of assessing small airways dysfunction. *European Clinical Respiratory Journal*. 2014; 1(1). DOI: 10.3402/ecrj.v1.25898.
- [3] Noutsios GT, Floros J. Childhood asthma: causes, risks, and protective factors; a role of innate immunity. *Swiss Medical Weekly*. 2014;144(w14036). DOI: 10.4414/smw.2014.14036.
- [4] Yanai M, Sekizawa K, Ohnui T, Sasaki H, Takishima T. Site of airway obstruction in pulmonary disease: direct measurement of intrabronchial pressure. *Journal of Applied Physiology*. 1992;72(3):1016-1023. DOI: 10.1152/jappl.1992.72.3.1016.
- [5] Meraz E, Nazeran H, Ramos C, Nava P, Diong B, Goldman M. Analysis of impulse oscillometric measures of lung function and respiratory system model parameters in small airway-impaired and healthy children over a 2-year period. *BioMedical Engineering OnLine*. 2011;10(1):1-21. DOI: 10.1186/1475-925X-10-21.
- [6] Meraz E, Nazeran H, Goldman M, Nava P, Diong B. Impulse oscillometric features of lung function: Towards computer-aided classification of respiratory diseases in children. In: *30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Vancouver: IEEE; 2008. p.2443-2446. DOI: 10.1109/IEMBS.2008.4649693
- [7] Hansell DM. Small airways diseases: detection and insights with computed tomography. *European Respiratory Journal*. 2001;17(6):1294-1313. ISSN 0903-1936. DOI: 10.1183/09031936.01.00206101.

- [8] Meraz E, Nazeran H, Edalatpour R, Rodriguez C, Montano K, Aguilar C, et al. Reference equations for impulse oscillometric and respiratory system model parameters in anglo and hispanic children. *Revista Mexicana de Ingeniera Biomedica*. 2016;37:49-61. DOI: 10.17488/rmib.37.1.6.
- [9] Mochizuku H, Hirai K, Tabata H. Forced Oscillation Technique and Childhood Asthma. *Allergology International*. 2012;61(3):373-383. DOI: 10.2332/allergolint.12-rai-0440.
- [10] Finkelstein JA, Lozano P, Shulruff R, Inui TS, Soumerai SB, Ng M, et al. Self-reported physician practices for children with asthma: are national guidelines followed?. *Pediatrics*. 2000;106(4):886-896.
- [11] World Health Organization. 10 facts on asthma. [cited 2018 June 29]. Available from: <http://www.who.int/features/factfiles/asthma/en/>.
- [12] Luo G, Stone BL, Fassl B, Maloney CG, Gesteland PH, Yerram SR, et al. Predicting asthma control deterioration in children. *BMC Medical Informatics and Decision Making*. 2015;15(1):1-84. DOI: 10.1186/s12911-015-0208-9.
- [13] Partridge ME, Dolen WK. Best estimates of asthma control in children. In: Pawankar R, Holgate ST, Rosenwasser LJ, (eds.). *Allergy Frontiers: Therapy and Prevention*. Tokio: Springer; 2009: p.565-576. DOI: 10.1007/978-4-431-99362-9_34.
- [14] Youssef HA, Shalaby A, Kaddah S, Selim S, Saad M. Impulse oscillometry as an alternative modality to the conventional pulmonary function tests in chronic obstructive pulmonary disease. *Egyptian Journal of Chest Diseases and Tuberculosis*. 2016;66(3):395-399. DOI: 10.1016/j.ejcdt.2016.08.004.
- [15] Komarow HD, Myles IA, Uzzaman A, Metcalfe DD. Impulse oscillometry in the evaluation of diseases of the airways in children. *Annals of Allergy, Asthma & Immunology*. 2011;106(3):191-199. DOI: 10.1016/j.anai.2010.11.011.

- [16] Marotta A, Klinnert MD, Price MR, Larsen GL, Liu AH. Impulse oscillometry provides an effective measure of lung dysfunction in 4-year-old children at risk for persistent asthma. *Journal of Allergy and Clinical Immunology*. 2003;112(2):317-322. DOI: 10.1067/mai.2003.1627.
- [17] Bickel S, Popler J, Lesnick B, Eid N. Impulse oscillometry: Interpretation and practical applications. *Chest*. 2014;146(3):841-847. DOI: 10.1378/chest.13-1875.
- [18] Komarow HD, Skinner J, Young M, Gaskins D, Nelson C, Gergen PJ et al. A study of the use of impulse oscillometry in the evaluation of children with asthma: Analysis of lung parameters, order effect, and utility compared with spirometry. *Pediatric Pulmonology*. 2012;47(1):18-26. DOI: 10.1002/ppul.21507.
- [19] Shi Y, Aledia AS, Tatavoosian AV, Vijayalakshmi S, Galant SP, George SC. Relating small airways to asthma control by using impulse oscillometry in children. *Journal of Allergy and Clinical Immunology*. 2012;129(3):671-678. DOI: 10.1016/j.jaci.2011.11.002.
- [20] Loftus PA, Wise SK. Epidemiology and economic burden of asthma. *International Forum of Allergy and Rhinology*. 2015;5(S1):S7-S10. DOI: 10.1002/alr.21547.
- [21] Centers for Disease Control and Prevention. Most Recent Asthma Data. [cited 2018 June 29]. Available from: https://www.cdc.gov/asthma/most_recent_data.htm.
- [22] Soni, A. *Top Five Most Costly Conditions among Children, Ages 0-17, 2012: Estimates for the U.S. Civilian Noninstitutionalized Population*. Agency for Healthcare Research and Quality, Rockville, MD. Statistical Brief number: 472, 2015. Available from: http://www.meps.ahrq.gov/mepsweb/data_files/publications/st472/stat472.pdf
- [23] Del Ro-Navarro E, Hidalgo-Castro EM, Sienra-Monge JJJ. Astma. *Boletín Médico del*

Hospital Infantil de México. 2009;66(1). Available from http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1665-11462009000100002.

- [24] Ortiz MI, Cano-Fragoso CV, Lazcano-Ortiz M, Romo-Hernandez G, Escamilla-Acosta MA. Factores de riesgo en niños asmáticos. Conocimiento que tienen sus padres sobre asma. *Acta Pediátrica de México*. 2012;33(3). Available from: <https://www.redalyc.org/pdf/4236/423640337005.pdf>
- [25] Barraza-Villarreal A, Sanin-Aguirre LH, Tellez-Rojo MM, Lacasanna-Navarro M, Romieu I. Prevalencia de asma y otras enfermedades alérgicas en niños escolares de Ciudad Juárez, Chihuahua. *Salud Pública Mex*. 2001;43:433-443. Available from: https://www.scielosp.org/article/ssm/content/raw/?resource_ssm_path=/media/assets/spm/v43n5/6722.pdf.
- [26] Graham Barr R, Avilés-Santa L, Davis SM, Aldrich TK, Gonzalez F, Henderson AG et al. Pulmonary Disease and Age at Immigration among Hispanics Results from the Hispanic Community Health Study/Study of Latinos. *American Journal of Respiratory and Critical Care Medicine*. 2016;193(4):386-395. DOI: 10.1164/rccm.201506-1211OC.
- [27] Cabieses B, Uphoff E, Pinart M, Anto JM, Wright J. A Systematic Review on the Development of Asthma and Allergic Diseases in Relation to International Immigration: The Leading Role of the Environment Confirmed. *PLOS One*. 2014;9(8):e105347. DOI: 10.1371/journal.pone.0105347.
- [28] Eldeirawi K, McConnell R, Furner S, Freels S, Stayner L, Hernandez E, et al. Associations of doctor-diagnosed asthma with immigration status, age at immigration, and length of residence in the United States in a sample of Mexican American School Children in Chicago. *Journal of Asthma*. 2009;46(8):796-802.
- [29] Eldeirawi K, Persky VW. Associations of physician-diagnosed asthma with country of residence in the first year of life and other immigration-related factors: Chicago

Asthma School Study. *Annals of Allergy, Asthma & Immunology*. 2007;99(3):236-243. DOI: 10.1016/S1081-1206(10)60659-X.

- [30] Lara M, Akinbami L, Flores G, Morgenstern H. Heterogeneity of childhood asthma among Hispanic children: Puerto Rican children bear a disproportionate burden. *Pediatrics*. 2006;117(1):43-53. DOI: 10.1542/peds.2004-1714.
- [31] Holguin F, Mannino DM, Antó J, Mott J, Ford ES, Teague WG, et al. Country of birth as a risk factor for asthma among Mexican Americans. *American Journal of Respiratory and Critical Care Medicine*. 2005;171(2):103-108. DOI: 10.1164/rccm.200402-143OC.
- [32] Eldeirawi K, McConnell R, Freels S, Persky VW. Associations of place of birth with asthma and wheezing in Mexican American children. *Journal of Allergy and Clinical Immunology*. 2005;116(1):42-48. DOI: 10.1016/j.jaci.2005.03.041.
- [33] Ku L, Jewers M. *Health Care for Immigrant Families: Current Policies and Issues*. Migration Policy Institute. 2013. Available from: <https://www.migrationpolicy.org/research/health-care-immigrant-families-current-policies-and-issues>.
- [34] Heyman JM, Nunez GG, Talavera V. Healthcare Access and Barriers for Unauthorized Immigrants in El Paso County, Texas. *Family & Community Health*. 2009;32(1):4-21. DOI: 10.1097/01.fch.0000342813.42025.a3.
- [35] Al-Ashkar F, Mehra R, Mazzone PJ. Interpreting pulmonary function tests: Recognize the pattern, and the diagnosis will follow. *Cleveland Clinic Journal of Medicine*. 2003;70(10):866-881. Available from: <https://pdfs.semanticscholar.org/2fbc/0355ee1f9bd6a11188b640132904da5e5431.pdf>
- [36] Global Initiative for Chronic Obstructive Lung Disease (GOLD). Spirometry for healthcare providers. [cited 2018 June 29]. Available from: https://goldcopd.org/wp-content/uploads/2016/04/GOLD_Spirometry_2010.pdf.

- [37] Global Initiative for Asthma (GINA). Global strategy for asthma management and prevention (2018 update). [cited 2018 June 29]. Available from: https://ginasthma.org/wp-content/uploads/2018/04/wms-GINA-2018-report-tracked_v1.3.pdf.
- [38] Brigham EP, West NE. Diagnosis of asthma: diagnostic testing. *International Forum of Allergy & Rhinology*. 2015;5(S1):S27-S30. DOI: 10.1002/alr.21597.
- [39] Meraz E, Nazeran H, Ramos C, Rodriguez L, Madrigal L, Gordillo N. Impulse Oscillometric Features and Respiratory System Models Track Small Airway Function in Children. In: Andrade AO, Soares A (eds.). *Practical Applications in Biomedical Engineering*. Rijeka: INTECH; 2012. DOI: 10.5772/52579.
- [40] Desiraju K, Agrawal A. Impulse oscillometry: The state-of-art for lung function testing. *Lung India*. 2016;33(4):410-416. DOI: 10.4103/0970-2113.184875.
- [41] Diong B, Nazeran H, Nava P, Goldman M. Modeling Human Respiratory Impedance: Comparing the Best Method with the Least Estimation Errors. *IEEE Engineering in Medicine and Biology Magazine*. 2007;26:48-55.
- [42] Diong B, Rajagiri A, Goldman M, Nazeran H. The augmented RIC model of the human respiratory system. *Medical & Biological Engineering & Computing*. 2009;47(4):395-404. DOI: 10.1007/s11517-009-0443-2.
- [43] Badnjevic A, Gurbeta L, Cifrek M, Marjanovic D. Classification of Asthma Using Artificial Neural Network. In: *39th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)-Proceedings*. Opatija: IEEE. 2016;p.387-390. DOI: 10.1109/MIPRO.2016.7522173.
- [44] Badnjevic A, Gurbeta L, Cifrek M, Marjanovic D. Diagnostic of asthma using fuzzy rules implemented in accordance with international guidelines and physicians experience. In: *39th International Convention on Information and Communication*

Technology, Electronics and Microelectronics (MIPRO)-Proceedings. Opatija: IEEE. 2016;p.375-380. DOI: 10.1109/MIPRO.2016.7522171.

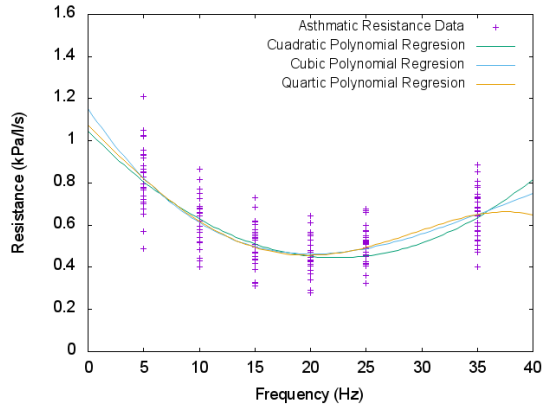
- [45] Badnjevic A, Cifrek M. Classification of Asthma Utilizing Integrated Software Suite. In: *6th European Conference of the International Federation for Medical and Biological Engineering, IFMBE Proceedings*. Dubrovnik: Springer. 2015;45:p.415-418. DOI: 10.1007/978-3-319-11128-5_104.
- [46] Badnjevic A, Cifrek M, Koruga D, Osmankovic D. Neuro-fuzzy classification of asthma and chronic obstructive pulmonary disease. *BMC Medical Informatics and Decision Making*. 2015;15(3). DOI: 10.1186/1472-6947-15-S3-S1
- [47] Badnjevic A, Koruga D, Cifrek M, Smith HJ, Bego T. Interpretation of pulmonary function test results in relation to asthma classification using integrated software suite. In: *36th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*. Opatija: IEEE. 2013;p.140-144.
- [48] N. Hafezi, H. Nazeran, E. Meraz, P. Nava, M. Goldman. An Integrated Software Package to Classify Human Respiratory Diseases. In: *25th Southern Biomedical Engineering Conference, IFMBE Proceedings*. Berlin: Springer. 2009;24. DOI: 10.1007/978-3-642-01697-4_112.
- [49] Barua M, Nazeran H, Nava P, Diong B, Goldman M. Classification of Impulse Oscillometric Patterns of Lung Function in Asthmatic Children using Artificial Neural Networks. In: *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*. Shanghai: IEEE. 2005;p.327-331. DOI: 10.1109/IEMBS.2005.1616411.
- [50] Barua M, Nazeran H, Nava P, Granda V, Diong B. Classification of pulmonary diseases based on impulse oscillometric measurements of lung function using neural networks. In: *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. San Francisco: IEEE. 2004;p.3848-3851. DOI: 10.1109/IEMBS.2004.1404077.

- [51] World Health Organization. COPD Definition. [cited 2018 June 29]. Available from <http://www.who.int/respiratory/copd/definition/en/>.
- [52] Liu Y, Pleasants R, Croft J, Wheaton A, Heidari K, Malarcher A, et al. Smoking duration, respiratory symptoms, and COPD in adults aged ≥ 45 years with a smoking history. *International Journal of Chronic Obstructive Pulmonary Disease*. 2015;10:1409-1416. DOI: 10.2147/copd.s82259.
- [53] De Assumpcao MS , da Silva Goncalves E, Oliveira MS, Ribeiro JD, Dalbo Contrera Toro AA, Azevedo Barros-Filho A, et al. Impulse Oscillometry System And Anthropometric Variables Of Preschoolers, Children And Adolescents: Systematic Review. *Current Pediatric Reviews*. 2017;13(2):126-135. DOI: 10.2174/1573396313666170622075940.
- [54] Birnbaum S, Barreiro, TJ. Methacholine Challenge Testing: Identifying Its Diagnostic Role, Testing, Coding, and Reimbursement. *Chest*. 2007;131(6):1932-1935. DOI: 10.1378/chest.06-1385.
- [55] Coates AL, Wanger J, Cockcroft DW, Culver BH, Diamant Z, Gauvreau G, et al. ERS technical standard on bronchial challenge testing: general considerations and performance of methacholine challenge tests. *European Respiratory Journal*. 2017;49(5). DOI: 10.1183/13993003.01526-2016.
- [56] Song TW, Kim KW, Kim ES, Park JW, Sohn MH, Kim KE. Utility of impulse oscillometry in young children with asthma. *Pediatric Allergy and Immunology*. 2008;19(8):763-768. DOI: 10.1111/j.1399-3038.2008.00734.x.
- [57] Nieto A, Pamies R, Oliver F, Medina A, Caballero L, Mazon A. Montelukast improves pulmonary function measured by impulse oscillometry in children with asthma (Mio study). *Respiratory Medicine*. 2006;10(7):1180-1185. DOI: 10.1016/j.rmed.2005.10.025.

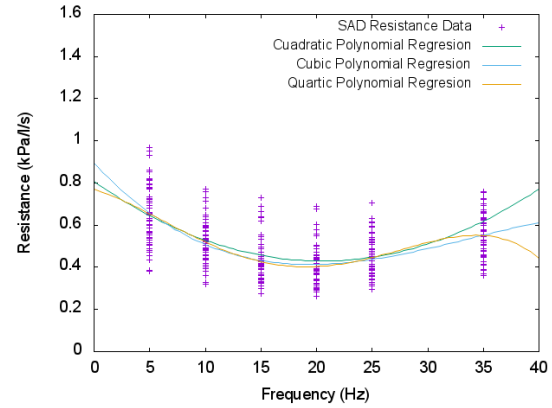
- [58] Schulze J, Biedebach S, Christmann M, Herrmann E, Voss S, Zielen S. Impulse Oscillometry as a Predictor of Asthma Exacerbations in Young Children. *Respiration*. 2016; 91(2):107-114. ISBN: 0025-7931. DOI: 10.1159/000442448.
- [59] Meraz E, Nazeran H, Goldman M, Diong B. Respiratory System Model Parameters Track Changes in Lung Function after Bronchodilation. In: *IFMBE Proceedings 25th Southern Biomedical Engineering Conference*. Miami: Springer. 2009;24:p.319-322. ISBN: 978-3-642-01697-4. DOI: 10.1007/978-3-642-01697-4_113.
- [60] Goldman M, Carter R, Klein R, Fritz G, Carter B, Pachuki P. Within- and Between-Day Variability of Respiratory Impedance, Using Impulse Oscillometry in Adolescent Asthmatics. *Pediatric Pulmonology*. 2002;34(4):312-319. DOI: 10.1002/ppul.10168.

Appendix A

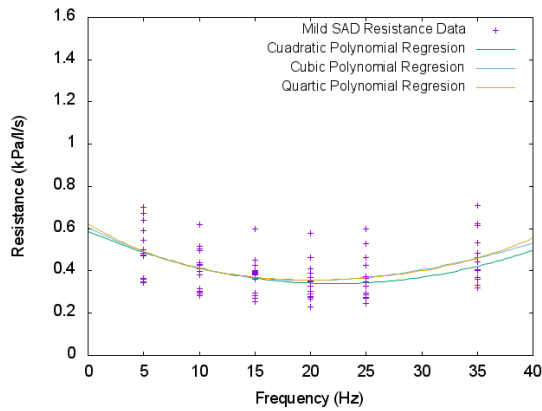
Polynomial Approximation Graphs of Resistance per Condition



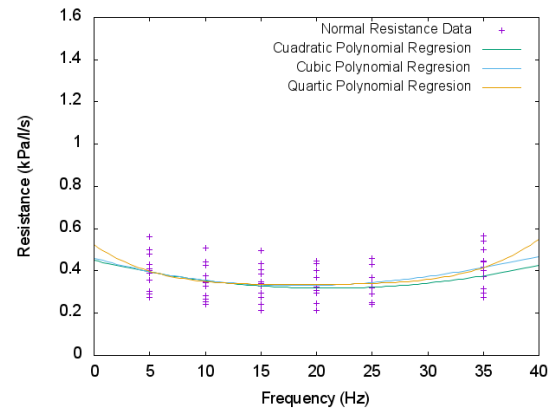
(a) Asthma



(b) SAI



(c) PSAI



(d) Normal

Figure A.1: Polynomial Regression Curves of Resistance data for (a) Asthma, (b) SAI, (c) PSAI, and (d) Normal.

Appendix B

Polynomial Approximation Graphs of Reactance per Condition

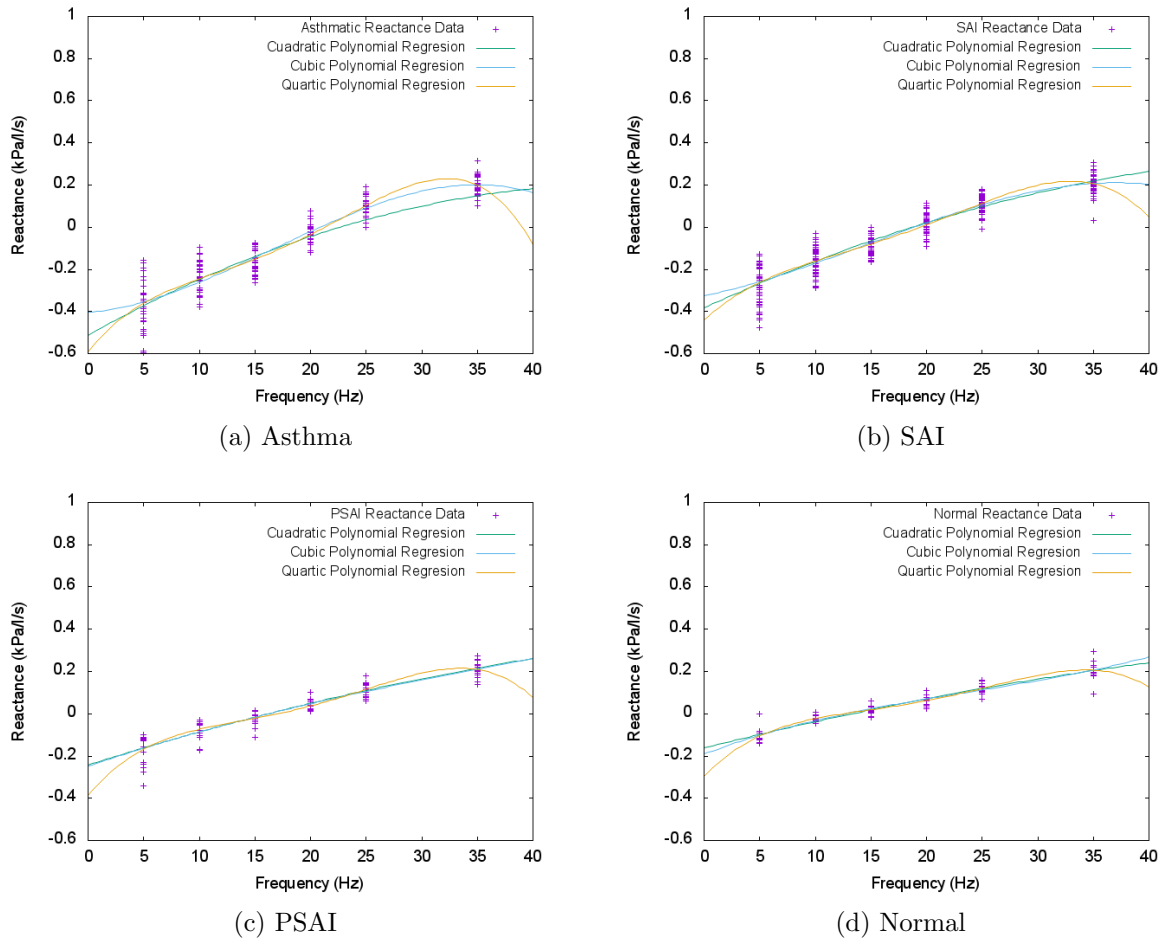
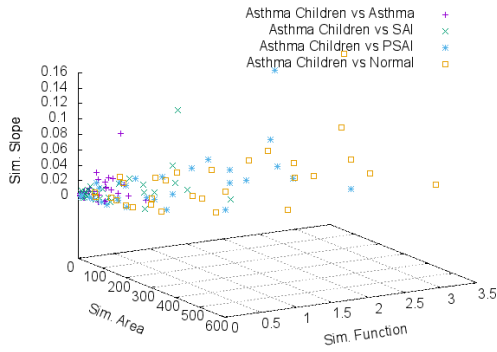


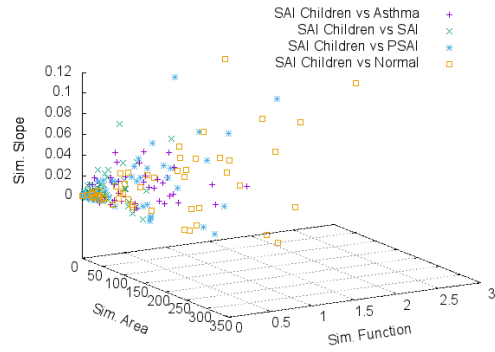
Figure B.1: Polynomial Regression Curves of Reactance data for (a) Asthma, (b) SAI, (c) PSAI, and (d) Normal.

Appendix C

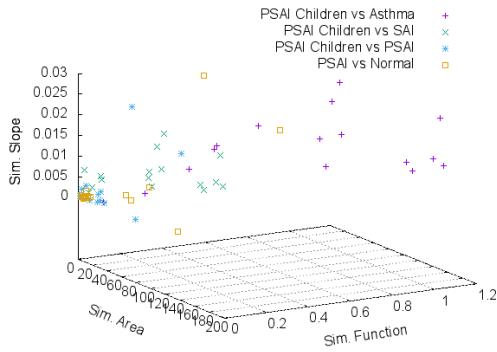
Similarity Graphs of Resistance per Condition



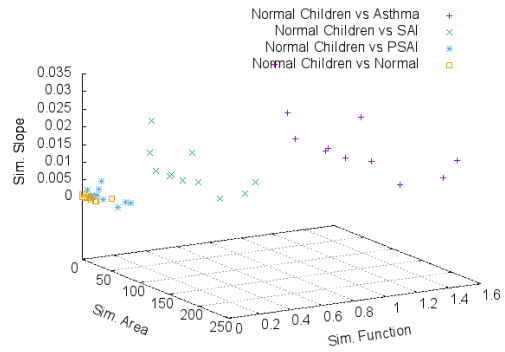
(a) Asthma



(b) SAI



(c) PSAI

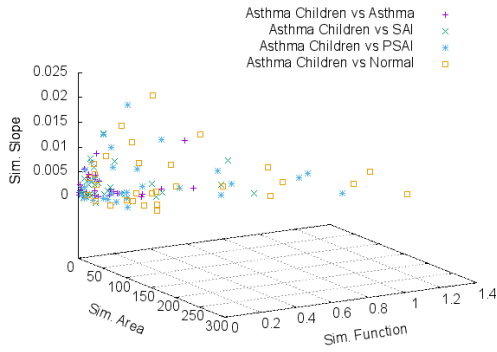


(d) Normal

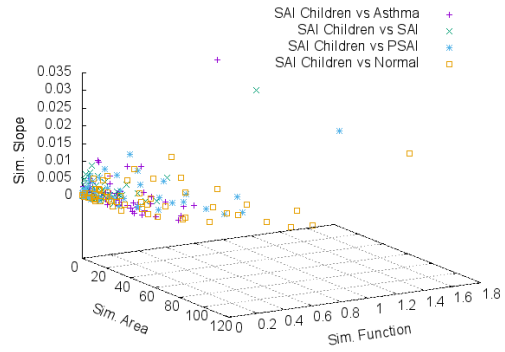
Figure C.1: Similarity Measures of Resistance for (a) Asthma, (b) SAI, (c) PSAI, and (d) Normal.

Appendix D

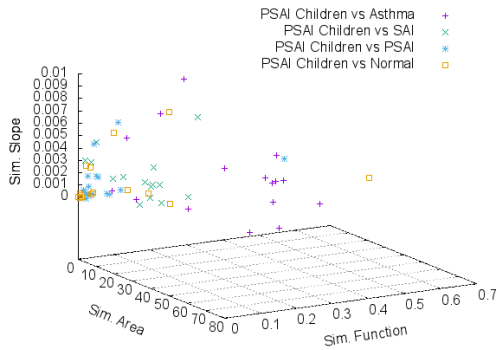
Similarity Graphs of Reactance per Condition



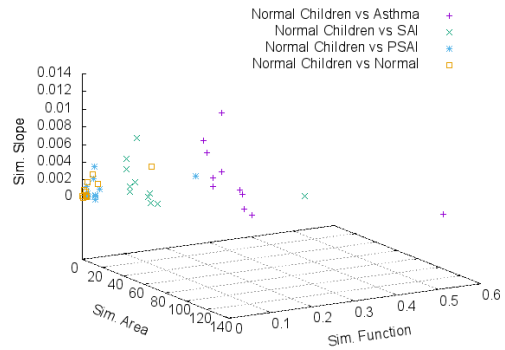
(a) Asthma



(b) SAI



(c) PSAI



(d) Normal

Figure D.1: Similarity Measures of Reactance for (a) Asthma, (b) SAI, (c) PSAI, and (d) Normal.

C++ Code Flow Diagram

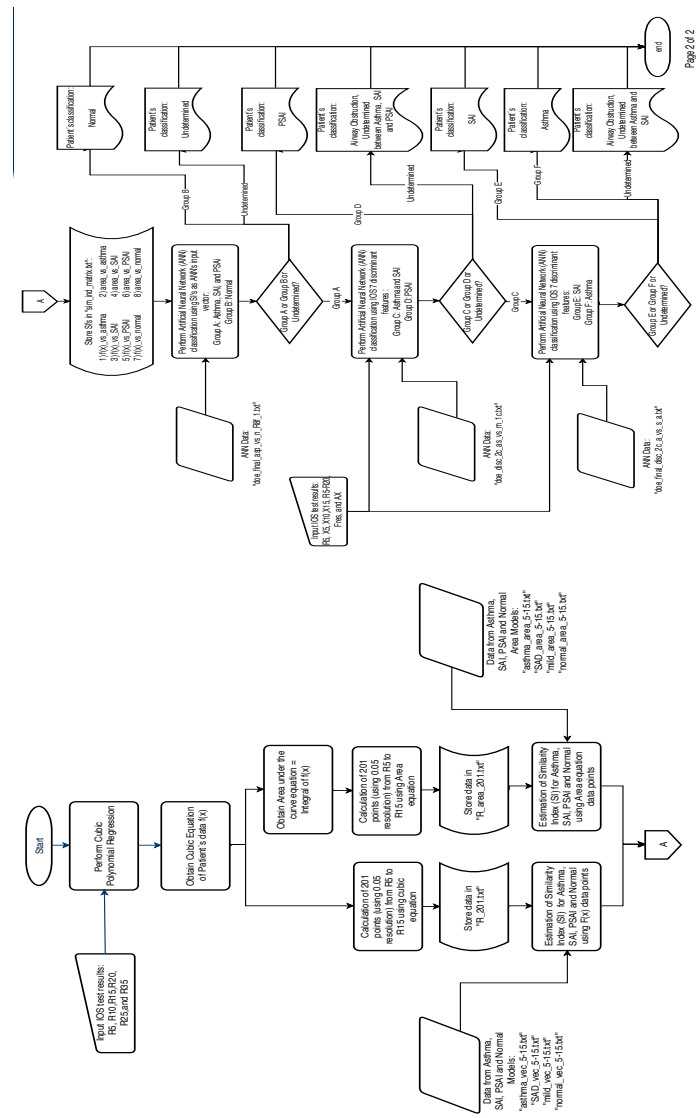


Figure E.1: C++ Code Flow Diagram

Curriculum Vitae

Nancy Avila was born on April 8, 1974. She holds B.Sc. degree in Industrial Engineering (with Honors), from the University of Ciudad Juárez, México and a M.Sc. degree in Industrial Engineering from the University of Texas at El Paso. She has over 22 years of experience in the manufacturing industry, being the last 13 years in the Medical Industry field, mainly in the Quality area. She has held several positions, each with increasing responsibility, including management positions at a senior level.

In the fall of 2014, she entered the Doctoral Program in Biomedical Engineering at The University of Texas at El Paso. She was awarded a fellowship by the Council of Science and Technology of México (CONACYT) to complete and achieve doctoral degree studies. She is the recipient of the Graduate Student Award administered by the Health Initiative of the Americas – UC Berkeley and Health and Migration Research Program (PIMSA, for its Spanish Acronym). During her doctoral studies she submitted 6 publications to high quality peer reviews journals; in 3 of them, she has been the first author. To date, 3 of her publications have been accepted and 3 of them are in peer review status. Nancy is an active member of the Latin American Society of Respiratory Physiology and is the founder of the DATAA company that stands for Development and Application of Machine Learning Techniques for its acronym in Spanish.

Permanent address: Rancho Santa Clara 2908

Ciudad Juárez, Chihuahua 32612