

**“Evaluación Acústica de las  
Neumopatías Intersticiales Difusas  
Utilizando Análisis Multicanal”**

Tesis que para obtener el título de  
**Maestra en Ciencias (Ingeniería Biomédica )**  
presenta:

**Graciela E. Martínez Hernández**

Directores de tesis:

**Dr. Tomás Aljama Corrales**

**Dr. Ramón González Camarena**

Sinodales:

**Dr. Ramón González Camarena (Presidente)**

**Dra. María Teresa García González (Secretaria)**

**Dra. Georgina Chi Iem (Vocal)**

**M. en I. Ricardo Rodríguez Vera (Vocal)**



UNIVERSIDAD AUTÓNOMA METROPOLITANA  
Iztapalapa

DIVISIÓN DE CIENCIAS BÁSICAS E INGENIERÍA

## **“Evaluación Acústica de las Neumopatías Intersticiales Difusas Utilizando Análisis Multicanal”**

---

Tesis que para obtener el título de  
**Maestra en Ciencias (Ingeniería Biomédica )**  
presenta:

**Graciela E. Martínez Hernández**

Directores de tesis:

Dr. Tomás Aljama Corrales  
Dr. Ramón González Camarena

México, D. F. A 23 de Marzo del 2006

## AGRADECIMIENTOS

- ❖ Mi enorme gratitud al Dr. Tomás Aljama, Al Dr. Ramón González y a la Dra. Sonia Charleston, por su dirección, apoyo y paciencia durante la realización de este trabajo. Gracias a sus consejos he adquirido más que conocimientos a lo largo de mi estancia en esta institución. Un reconocimiento a su labor como investigadores y dedicación con sus alumnos.
- ❖ Especiales agradecimientos a la Dra. Georgina Chi Lem por su valiosa colaboración durante la adquisición de los registros en el INER y por sus sugerencias durante la realización del presente trabajo.
- ❖ Gracias a la Dra. María Teresa García y al M. en I. Ricardo Rodríguez por sus sugerencias y críticas constructivas sobre el trabajo de tesis.
- ❖ Agradezco el apoyo generoso y desinteresado que me brindó en todo momento la Dra. Raquel Valdés.
- ❖ A todos mis profesores del posgrado en Ingeniería Biomédica por heredarme parte de sus conocimientos.
- ❖ Gracias a Esperanza Ibáñez por su ayuda y calidez en su trato.
- ❖ A todos mis amigos y compañeros del posgrado, especialmente a Carla Mántaras, Paola Ponce, Yudibeth Sánchez, Paola Coello, Luis Domínguez, Oswaldo Sánchez, Gabriel Gentiletti, Marcos López, Sergio Solís, Oscar Marrufo, Joaquín Cruz y Teodoro Solís, gracias por su amistad, afecto y motivación.
- ❖ Gracias a mis compañeros y amigos del Laboratorio de Procesamiento Digital de Señales e Imágenes Biomédicas por su apoyo.
- ❖ Agradezco a Iris Núñez y Francis Plaza por su afecto y compañía.
- ❖ A mis amigas Eréndira García, Yesenia González, Liliana Sánchez y Karen Gordillo por su amistad, apoyo constante y cariño. A mis amigos Álvaro Anzueto y Pedro Rojas por alentarme siempre.

- ❖ Gracias a la Sra. Emilia Frías y al Sr. Ángel Salazar por brindarme su amistad y hacer de su casa una cálida estancia para mí durante los últimos años.
- ❖ A Riurik Domínguez y a Juana Rojas quienes ahora son parte de mi familia, gracias por su apoyo.
- ❖ De todo corazón agradezco las enseñanzas sobre la vida, el amor y apoyo incondicional de mis padres Antonia Hernández y René Martínez, a quienes debo todo lo que he logrado.
- ❖ A mis hermanas Maricela e Yvonne y a mi hermano René por ser mis eternos amigos y mi sostén en cualquier situación. A mis sobrinas Ana Isabel y Abigail por sus enseñanzas de amor.

Dedico con mucho cariño esta tesis a mi familia, ... lo más valioso de mi vida.

## **RECONOCIMIENTOS**

Este trabajo fue realizado en el Laboratorio de Procesamiento Digital de Señales e Imágenes Biomédicas (LPDSIB) de la Universidad Autónoma Metropolitana y fue financiado por una beca de posgrado otorgada por el CONACyT.

Los registros de pacientes y sujetos sanos fueron adquiridos en el Instituto Nacional de Enfermedades Respiratorias (INER).

## CONTENIDO

<b>AGRADECIMIENTOS</b> .....	<b>2</b>
<b>RECONOCIMIENTOS</b> .....	<b>3</b>
<b>RESUMEN</b> .....	<b>9</b>
<b>CAPÍTULO I</b>	
<b>INTRODUCCIÓN</b> .....	<b>13</b>
<b>1.1 NEUMOPATÍAS INTERSTICIALES DIFUSAS</b> .....	<b>13</b>
<i>1.1.1 Definición de Neumopatía Intersticial Difusa</i> .....	13
<i>1.1.2 Causas de Neumopatía Intersticial Difusa</i> .....	14
<i>1.1.3 Proceso Fisiopatológico de las NID</i> .....	15
<i>1.1.4 Cuadro clínico de las NID</i> .....	15
<i>1.1.5 Métodos Diagnósticos de las NID</i> .....	16
1.1.5.1 Pruebas de función pulmonar .....	17
1.1.5.2 Radiología .....	17
1.1.5.3 Lavado broncoalveolar .....	18
1.1.5.4 Biopsia Pulmonar.....	18
<i>1.1.6 Epidemiología</i> .....	19
<b>1.2 AUSCULTACIÓN Y SONIDOS RESPIRATORIOS</b> .....	<b>20</b>
<i>1.2.1 Introducción</i> .....	20
<i>1.2.2 Características de los sonidos respiratorios normales y adventicios</i> .....	21
1.2.2.1 Sonidos Respiratorios Normales .....	22
1.2.2.2 Sonidos Adventicios .....	24
1.2.2.2.1 Sibilancias.....	24
1.2.2.2.2 Crepitancias .....	25
<b>1.3 REFERENCIAS</b> .....	<b>28</b>

## CAPÍTULO II

<b>2.1</b>	<b>PLANTEAMIENTO DEL PROBLEMA.....</b>	<b>30</b>
<b>2.2</b>	<b>HIPÓTESIS .....</b>	<b>31</b>
<b>2.3</b>	<b>OBJETIVO.....</b>	<b>31</b>
2.3.1	<i>Objetivo General.....</i>	<i>31</i>
2.3.2	<i>Objetivos Específicos.....</i>	<i>32</i>

## CAPÍTULO III

<b>TÉCNICAS DE PARAMETRIZACIÓN .....</b>	<b>33</b>	
<b>3.1</b>	<b>INTRODUCCIÓN .....</b>	<b>33</b>
<b>3.2</b>	<b>DESCRIPCIÓN DE MÉTODOS .....</b>	<b>35</b>
3.2.1	<i>Medidas clásicas en el dominio de la frecuencia .....</i>	<i>35</i>
3.2.1.1	<i>Método de Burg para estimación espectral.....</i>	<i>36</i>
3.2.2	<i>Modelo Autorregresivo Monocanal (AR) .....</i>	<i>36</i>
3.2.2.1	<i>Orden del modelo AR.....</i>	<i>40</i>
	<i>Criterio de Información de Akaike.....</i>	<i>40</i>
3.2.3	<i>Modelo Autorregresivo Multicanal.....</i>	<i>41</i>
3.2.3.1	<i>Reducción De La Dimensionalidad Del Espacio De Características.....</i>	<i>44</i>
3.2.3.1.1	<i>Eigenvalores Y Eigenvectores .....</i>	<i>44</i>
3.2.3.1.2	<i>Descomposición en Valores Singulares.....</i>	<i>44</i>
3.2.3.1.3	<i>Análisis de componentes principales (PCA) .....</i>	<i>46</i>
<b>3.3</b>	<b>REDES NEURONALES ARTIFICIALES .....</b>	<b>48</b>
3.3.1	<i>Algoritmo de Retropropagación.....</i>	<i>51</i>
3.3.2	<i>Modos de entrenamiento.....</i>	<i>54</i>
3.3.3	<i>Criterio para detener el proceso de entrenamiento.....</i>	<i>54</i>
<b>3.4</b>	<b>REFERENCIAS .....</b>	<b>56</b>

## CAPÍTULO IV

<b>ESTADO DEL ARTE .....</b>	<b>58</b>
<b>4.1 PROBLEMA DE CLASIFICACIÓN DE SONIDOS RESPIRATORIOS.....</b>	<b>58</b>
<b>4.2 REFERENCIAS .....</b>	<b>67</b>

## CAPÍTULO V

<b>DESCRIPCIÓN DE LA METODOLOGÍA .....</b>	<b>70</b>
<b>5.1 METODOLOGÍA GENERAL.....</b>	<b>70</b>
<b>5.1 ADQUISICIÓN MULTICANAL .....</b>	<b>71</b>
5.1.1 <i>Equipo .....</i>	<i>71</i>
5.1.2 <i>Sujetos .....</i>	<i>71</i>
5.1.3 <i>Adquisición.....</i>	<i>72</i>
5.1.4 <i>Programa de adquisición .....</i>	<i>74</i>
<b>5.2 PREPROCESAMIENTO.....</b>	<b>76</b>
5.2.1 <i>Calibración .....</i>	<i>76</i>
5.2.2 <i>Eliminación automática de periodos.....</i>	<i>77</i>
<i>de apnea y segmentación de fases inspiratorias y espiratorias. ....</i>	<i>77</i>
5.2.3 <i>Filtrado de las señales para la reducción de ruido cardíaco y muscular .....</i>	<i>78</i>
5.2.4 <i>Eliminación de artefactos .....</i>	<i>79</i>
<b>5.3 FORMACIÓN DEL BANCO DE DATOS.....</b>	<b>79</b>
5.3.1 <i>Criterios de Inclusión: .....</i>	<i>79</i>
5.3.2 <i>Criterios de Exclusión:.....</i>	<i>79</i>
<b>5.4 PARAMETRIZACIÓN DE LOS DATOS .....</b>	<b>80</b>
5.4.1 <i>Eigenvalores de la matriz de covarianza .....</i>	<i>80</i>
5.4.2 <i>Medidas clásicas en el dominio de la frecuencia .....</i>	<i>82</i>
5.4.3 <i>Modelo AR monocanal .....</i>	<i>84</i>
5.4.4 <i>Modelo AR multicanal .....</i>	<i>86</i>

5.4.4.1	Reducción del espacio de características .....	87
5.4.4.1.1	Matriz individual de coeficientes del modelo AR multicanal.....	89
5.4.4.1.2	Matriz total de coeficientes del modelo AR multicanal.....	90
<b>5.5</b>	<b>RECONOCIMIENTO DE PATRONES .....</b>	<b>91</b>
<b>5.6</b>	<b>VALIDACIÓN GENERAL .....</b>	<b>94</b>
<b>5.7</b>	<b>REFERENCIAS.....</b>	<b>96</b>
 <b>CAPÍTULO VI</b>		
<b>RESULTADOS.....</b>		<b>98</b>
<b>6.1</b>	<b>INTRODUCCIÓN .....</b>	<b>98</b>
<b>6.2</b>	<b>BASE DE DATOS .....</b>	<b>98</b>
<b>6.3</b>	<b>PARAMETRIZACIÓN Y CLASIFICACIÓN DE LAS SEÑALES DE SONIDO</b>	
<b>RESPIRATORIO .....</b>		<b>103</b>
6.3.1	<i>Resultados utilizando los Parámetros de Eigenvalores de la matriz de covarianza .....</i>	<i>109</i>
6.3.2	<i>Resultados Utilizando Parámetros en el Dominio de la Frecuencia.....</i>	<i>113</i>
6.3.3	<i>Resultados aplicando el Modelo AR Monocanal (ARU).....</i>	<i>117</i>
6.3.4	<i>Resultados aplicando el modelo AR Multicanal (ARM).....</i>	<i>120</i>
6.3.4.1	Matriz individual de coeficientes del modelo AR multicanal.....	121
6.3.4.1.1	Reducción de Características por Descomposición en Valores Singulares (SVD) de la matriz de coeficientes AR .....	121
6.3.4.1.2	Reducción de Características mediante la obtención de los eigenvalores de la matriz de covarianza de la matriz de coeficientes AR .....	123
6.3.4.1.3	Reducción de Características por medio de la proyección de eigenvectores de la matriz de covarianza del vector V.....	127
6.3.4.2	Matriz Total de coeficientes del modelo AR multicanal.....	129
6.3.4.2.1	Reducción de características aplicando Análisis en Componentes Principales (PCA) .....	129

<i>TABLA 35 RESUMEN DEL DESEMPEÑO DE LA RED NEURONAL</i> .....	132
<i>TABLA 36 PROMEDIO DEL DESEMPEÑO DE LA RED NEURONAL</i> .....	133
<i>TABLA 37 RESUMEN DE VALIDACIÓN GENERAL</i> .....	134

## **CAPÍTULO VII**

<b>DISCUSIÓN</b> .....	<b>135</b>
<b>7.1 MEDIDAS CLÁSICAS EN EL DOMINIO DE LA FRECUENCIA</b> .....	<b>139</b>
<b>7.2 EIGENVALORES DE LA MATRIZ DE COVARIANZA DE LA SEÑAL DE SONIDO RESPIRATORIO</b> .....	<b>140</b>
<b>7.3 MODELO AR MONOCANAL</b> .....	<b>140</b>
<b>7.4 MODELO AR MULTICANAL</b> .....	<b>141</b>
<b>7.5 RED NEURONAL</b> .....	<b>142</b>
7.5.1 <i>Fase de entrenamiento</i> .....	143
7.5.2 <i>Validación General</i> .....	144
<b>7.6 REFERENCIAS</b> .....	<b>145</b>

## **CAPÍTULO VIII**

<b>CONCLUSIONES Y PERSPECTIVAS</b> .....	<b>147</b>
<b>ANEXO 1</b>	
<b>PARTICIPACIONES EN CONGRESOS</b> .....	<b>149</b>
A.1.1 <i>Clasificación de Información Acústica en Neumopatías Intersticiales Difusas Mediante un Análisis Autorregresivo Multicanal</i> .....	150
A.1.2 <i>A Multichannel Autoregressive Model Applied to Acoustic Information In Difuse Interstitial Pneumonia</i> .....	154
A.1.3 <i>Computerized Classification of Normal and Abnormal Lung Sounds by Multivariate Linear Autoregressive Model</i> .....	155

## RESUMEN

En la presente tesis se estudia el caso específico de un tipo de enfermedades respiratorias que impactan considerablemente a la salud en nuestro país, las Neumopatías Intersticiales Difusas (NID). Estas neumopatías constituyen un grupo de enfermedades que comparten rasgos fisiopatológicos, clínicos y radiológicos. Por tratarse de un grupo heterogéneo, presentan un difícil diagnóstico en etapas tempranas del padecimiento, lo que conduce a la falta de tratamiento adecuado y el desarrollo progresivo de fibrosis pulmonar difusa, destruyendo las unidades alveolo-capilares y alterando el intercambio gaseoso, provocando que la enfermedad sea irreversible y mortal para el paciente en un plazo de tiempo relativamente corto.

Actualmente se utilizan una serie de rigurosos procedimientos y técnicas, cuyos resultados soportan el diagnóstico de NID. Uno de los procedimientos básicos durante la exploración física del paciente, es la auscultación clínica en la que el médico escucha los sonidos respiratorios con ayuda del estetoscopio en un esfuerzo por detectar la presencia de anomalías. Debido a los cambios histológicos y fisiológicos que suceden en los pulmones con el desarrollo de NID, además del sonido respiratorio de base se producen sonidos anormales conocidos como estertores crepitantes, que pueden ser escuchados sobre la pared torácica del sujeto enfermo. La distribución en el tórax de los estertores crepitantes y el número de estertores que se escuchan, puede relacionarse con la severidad de la enfermedad. Sin embargo, a pesar de su utilidad, el estetoscopio posee ciertas desventajas técnicas que llevan al médico a un diagnóstico subjetivo de la enfermedad, basado en su habilidad y experiencia. Otros de los exámenes de rutina que se realizan para el diagnóstico de las NID son las pruebas de función respiratoria, los rayos X y la tomografía axial computarizada. No obstante, en ocasiones es necesario recurrir a la biopsia pulmonar como procedimiento definitivo en el diagnóstico y especificidad del tipo de NID.

Las técnicas anteriores son útiles para evaluar la extensión y la severidad de la enfermedad para proporcionar tratamiento y seguimiento al paciente, sin embargo, el diagnóstico de las NID generalmente se realiza en etapas medias o avanzadas de la enfermedad, y el diagnóstico puede

ser confuso o en el peor de los casos, erróneo.

Considerando los puntos anteriores, el objetivo de la presente tesis es evaluar la información acústica multicanal a través de técnicas de procesamiento digital de señales y comprobar si es posible utilizarla como herramienta para apoyar el diagnóstico de las Neumopatías Intersticiales Difusas, con una técnica simple y no invasiva. Específicamente el trabajo se enfoca al análisis y comparación de diferentes técnicas de parametrización para caracterizar información acústica respiratoria y que en combinación con un clasificador ayude a diferenciar sonidos normales y anormales.

En la literatura se reportan diversos trabajos relacionados con la clasificación de sonidos respiratorios normales y aquellos que son consecuencia de alguna anomalía respiratoria. Se han utilizado diversas técnicas de procesamiento extrayendo características de las señales en el dominio del tiempo y de la frecuencia, en uno o varios canales de forma independiente y se han evaluado diferentes algoritmos de clasificación, tales como clasificadores bayesianos y redes neuronales, sin embargo, los resultados obtenidos no son completamente satisfactorios.

La presente investigación comienza con la revisión del estado del arte en cuestión de clasificación de sonidos normales y anormales, la adquisición u obtención de las señales de sonido respiratorio en sujetos sanos y pacientes con NID del Instituto Nacional de Enfermedades Respiratorias (INER). Una parte destacada de esta investigación es el uso de una técnica *multicanal* utilizando 25 sensores sobre la pared posterior del tórax del sujeto, lo que proporciona información acústica espacio-temporal. La formación de una base de datos con los expedientes clínicos proporcionados por el INER.

Se realizó la comparación de cuatro diferentes técnicas de parametrización. Para llevar a cabo la parametrización de segmentos aproximadamente estacionarios de la señal respiratoria multicanal, se evaluaron medidas clásicas en el dominio de la frecuencia (frecuencias percentiles), así como eigenvalores de la matriz de covarianza, modelos autorregresivos monocanales y multicanales. La etapa de reconocimiento de patrones se realizó a través de una red neuronal de retropropagación utilizando el algoritmo de aprendizaje de Levenberg-Marquadt. La clasificación de las características de las señales de sonido respiratorio en

normales y anormales se obtuvo sobre la base de ciertos patrones definidos por el historial clínico, los exámenes clínicos y decisiones de un médico experto. Finalmente se realizó la validación general de la red neuronal simulando a la entrada vectores pertenecientes a un solo sujeto sano o enfermo y que nunca habían sido “vistos” por la red neuronal.

Las técnicas de parametrización mostraron diferencias en su desempeño, sin embargo, al comparar los resultados obtenidos durante la fase de entrenamiento, prueba y validación de la red, se concluye que todas las técnicas de parametrización presentaron un buen desempeño. Los porcentajes promedio obtenidos, van desde el 78.52 en el peor de los casos, hasta el 100% de sensibilidad y especificidad de la red neuronal. Sin embargo, al realizar la validación general por sujeto, utilizando vectores pertenecientes a un solo sujeto sano o enfermo, los porcentajes obtenidos varían de 53.47% hasta el 97.01% al clasificar los vectores de sujetos sanos y porcentajes de 16.69% a 98.75% para la clasificación de vectores de sujetos enfermos con NID.

La red neuronal utilizada es de tipo supervisada y cada vector de características se etiquetó como normal si provenía de un registro de sujeto sano o como anormal si pertenecía a un registro de paciente con NID. Es recomendable modificar el método para etiquetar a los vectores de características ya que presenta limitaciones al establecer los patrones, principalmente para los pacientes con NID, debido a que existen segmentos de la señal de sonido respiratorio normales y anormales dentro de un mismo registro.

En general, se concluye que la información acústica multicanal efectivamente proporciona información espacio-temporal que ayuda a mejorar la clasificación de sonidos normales y anormales. Según los resultados, la técnica de parametrización mediante los eigenvalores es un método útil, ya que mostró buen desempeño durante la fase de entrenamiento, sin embargo el número de vectores no clasificados es bastante grande lo que se refleja en la validación general al haber obtenido los porcentajes más bajos. Una técnica interesante es la de medidas en la frecuencia, debido a que evidentemente existen componentes en frecuencia que ayudan a discriminar sonidos normales y anormales, pero es necesario mayor cantidad de datos o sujetos para entrenar a la red y que conduzca a una mejor generalización del clasificador. La utilización

del modelo AR monocanal resultó consistente, con excelente desempeño durante la fase de entrenamiento, prueba y validación de la red neuronal y también durante la validación general con datos nunca antes vistos por la red. La utilización del modelo autorregresivo (AR) multicanal requirió de un gran costo computacional e hizo necesario la utilización de técnicas de reducción de la dimensionalidad, dentro de las cuales el método de análisis de componentes principales (PCA) obtuvo mejor desempeño. Los resultados obtenidos con el modelo AR multicanal en combinación con PCA son ligeramente más bajos en porcentaje que los resultados obtenidos con el modelo AR monocanal, como se verá más adelante, en la sección de resultados. Esto provoca cierta confusión, ya que se esperaba que el modelo AR multicanal proporcionará un mejor comportamiento que el AR monocanal, ya que el modelo multicanal incluye al monocanal.

# **CAPÍTULO I**

## **INTRODUCCIÓN**

### **1.1 NEUMOPATÍAS INTERSTICIALES DIFUSAS**

Los problemas respiratorios son en conjunto una de las principales causas de enfermedad en todo el mundo, y en México forman parte de la lista de las diez enfermedades más comunes, según el reporte del Sistema Nacional de Vigilancia Epidemiológica de 2003 [1]. En la presente tesis se investiga la clasificación de sonidos respiratorios provenientes de sujetos sanos y enfermos bajo el esquema de un grupo de enfermedades denominadas Neumopatías Intersticiales Difusas (NID), que afectan a un gran número de personas en nuestro país y tienen un difícil diagnóstico en etapas tempranas del padecimiento lo que provoca la falta de tratamiento adecuado y el desarrollo potencial de la enfermedad hasta hacerla irreversible y mortal para el paciente, en un plazo de tiempo relativamente corto.

#### **1.1.1 DEFINICIÓN DE NEUMOPATÍA INTERSTICIAL DIFUSA**

Las Neumopatías Intersticiales Difusas (NID) también denominadas Enfermedades Pulmonares Intersticiales Difusas, Enfermedades Fibrosantes del Pulmón o Neumonías Difusas Intersticio Alveolares (NDIA) [2], [3], constituyen un grupo heterogéneo de padecimientos respiratorios agudos, subagudos y crónicos, que se caracterizan por afectar difusamente al parénquima pulmonar, primordialmente al intersticio (espacio tisular que se encuentra entre las paredes alveolares) y a los espacios alveolares, aunque muchas de ellas también producen lesión de los bronquiolos terminales, el tejido conjuntivo peribronquial, los vasos pulmonares y en ocasiones, el espacio pleural. Por consiguiente, los términos “enfermedad pulmonar intersticial” o “neumopatía intersticial”, ampliamente utilizados por numerosos autores y tratados médicos, no describen anatómicamente a la enfermedad y

“neumonías difusas intersticio-alveolares” es el término apropiado, sin embargo, a lo largo de la presente tesis se seguirá empleando, por simplicidad, el término “neumopatías intersticiales difusas” [3].

Estas enfermedades presentan en común rasgos fisiopatológicos, clínicos y radiológicos, y morfológicamente, después de un daño inicial, se presenta un proceso inflamatorio intersticial e intra-alveolar difuso (cuyas características varían según el tipo de NID) que si no es tratado oportuna y adecuadamente, suele conducir a una fibrosis pulmonar difusa, engrosando las unidades alveolo-capilares y alterando gravemente el intercambio gaseoso [2], [3].

Por tratarse de un grupo heterogéneo, el estudio de las NID se dificulta y no existen reportes estadísticos precisos. Algunas NID progresan y ocasionan la muerte en un periodo de tiempo relativamente corto: 2 a 5 años, como ocurre con la fibrosis pulmonar idiopática. Otras, como la sarcoidosis tienen mejor pronóstico y un porcentaje variable de pacientes se curan o mejoran [2]. Por citar un caso, se ha considerado a la alveolitis alérgica extrínseca como una enfermedad benigna, sin embargo, en México el 40% de los pacientes evoluciona hacia la fibrosis y el 30% fallece en un promedio de 5 años. Alrededor del 40% sufre alteraciones clínicas, radiológicas o funcionales de diferentes grados y solamente un 20% se cura [5].

### **1.1.2 CAUSAS DE NEUMOPATÍA INTERSTICIAL DIFUSA**

Existen más de 150 padecimientos y/o agentes agresores que son capaces de producir NID. La Asociación Americana del Pulmón sugiere que las causas más conocidas son: Ocupacionales, de exposición al ambiente, infecciones, enfermedades del tejido conectivo y del colágeno, drogas y genético familiar. En México, una clasificación basada en las causas que la provocan, ha sido propuesta por Selman y colaboradores, agrupando las NID de la forma siguiente: Fibrosis pulmonar idiopática, enfermedades hereditarias, enfermedades colágeno-vasculares y síndromes nefro-pulmonares, causas inhalatorias y enfermedades causadas por entidades específicas, por ejemplo proteinosis alveolar [2].

### **1.1.3 PROCESO FISIOPATOLÓGICO DE LAS NID**

Los eventos iniciales en la patogénesis de la fibrosis intersticial son difíciles de analizar, debido a que son enfermedades de comienzo aparentemente benigno pero diagnosticadas en etapas avanzadas. Independientemente de su etiología, el proceso de una NID está integrado primero por un componente inflamatorio (alveolitis), y otro dinámico, de reparación y fibrosis. En la figura 1 se muestra el proceso típico del desarrollo de fibrosis: el tejido pulmonar, al igual que todos los tejidos del organismo está constantemente siendo agredido por diversas partículas externas y microorganismos. Afortunadamente, los diferentes mecanismos de defensa controlan gran parte de estas agresiones. Por otro lado existen algunas partículas que logran hacer efecto pero con una adecuada atención médica, el paciente logra restablecerse. Sin embargo, si estas partículas persisten en el tejido pulmonar o la lesión por alguna enfermedad no se restablece correctamente, la agresión pulmonar inicial puede provocar lesión en el endotelio capilar y en las células epiteliales y activar los mecanismos de la inflamación; es decir, la alveolitis. Estos mecanismos llevan a una proliferación de fibroblastos con alteraciones en la síntesis y degradación del colágeno, produciendo una matriz extracelular, sobretodo de fibras colágenas que se extienden por una gran zona del pulmón formando una cicatriz. El proceso de fibrosis se manifiesta tanto de forma intra-alveolar como intersticial y posteriormente conduce al pulmón al estadio final de panalización [5].

### **1.1.4 CUADRO CLÍNICO DE LAS NID**

En cualquier tipo de evolución, el síntoma principal es la disnea de esfuerzo progresiva. Otros síntomas asociados son la tos seca, hiporexia, astenia y pérdida de peso, el hipocratismo digital y en menor medida la osteoartropatía hipertrófica. Durante la auscultación se encuentran estertores crepitantes bibasales (tipo velcro) y conforme progresa la enfermedad, los estertores se extienden hacia los ápices pulmonares.

Varios trabajos demuestran que el pronóstico depende en muchos aspectos del grado

morfológico de fibrosis al momento del diagnóstico, estableciendo la importancia de hacer un reconocimiento temprano en los casos de NID.

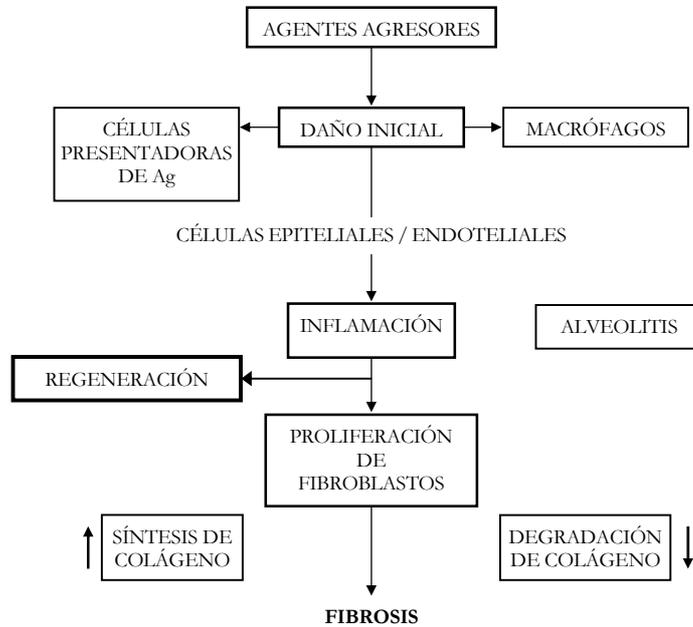


Fig. 1 Proceso de la Fibrosis Intersticial Difusa. Extraído de [2]

### 1.1.5 MÉTODOS DIAGNÓSTICOS DE LAS NID

Si el paciente presenta disnea de esfuerzo progresiva, la base fundamental para el diagnóstico de las NID es el grado de sospecha clínica la cual depende de la experiencia del médico [2]. Usualmente se lleva a cabo un diagnóstico por etapas: Primero, se realiza la evaluación de la sintomatología actual, así como del historial ocupacional, ambiental, del uso de medicamentos, y antecedentes familiares de enfermedades similares y hereditarios, construyendo el expediente clínico del paciente. Posteriormente, se emplean técnicas no invasivas, es decir, se lleva a cabo la exploración física por parte del personal médico especializado, que consiste en aplicar la técnica de la *auscultación*. El diagnóstico se conjunta con los resultados de las pruebas de función pulmonar, los exámenes de sangre, las técnicas de imagenología tales como *Rx*, *TACAR* (Tomografía Axial Computarizada de Alta Resolución) e *IRM* (Imágenes de

Resonancia Magnética). En segunda instancia, se procede a utilizar métodos semi-invasivos como la broncoscopia por fibra óptica y el lavado broncoalveolar. Por último, los pacientes son expuestos a métodos invasivos tales como la biopsia transbronquiolo-alveolar (BTB), la biopsia a cielo abierto (BPCA) o cirugía torácica video-asistida [2], [4], [6]. A continuación se presenta un resumen de las técnicas más importantes utilizadas para diagnosticar una NID.

#### **1.1.5.1 PRUEBAS DE FUNCIÓN PULMONAR**

Desde el punto de vista funcional respiratorio estas entidades presentan un trastorno predominantemente restrictivo caracterizado por una progresiva disminución de los volúmenes y capacidades a consecuencia de la fibrosis y la disminución de la distensibilidad. En algunos casos puede coexistir disminución del flujo aéreo. En etapas tempranas de la enfermedad lo más significativo es la hipoxemia durante el ejercicio y posteriormente con el avance de la fibrosis se presenta incluso la hipoxemia durante el reposo.

El estudio del intercambio gaseoso también revela desaturación de la Hb en ejercicio a través de la prueba de caminata de 6 minutos, recordando que el ejercicio tiene un profundo efecto sobre la función respiratoria. En general se produce una respiración rápida y superficial con una caída significativa de la  $P_aO_2$  con aumento de la diferencia alvéolo-arterial de oxígeno (A-a  $O_2$ ).

Es importante señalar que las alteraciones observadas en la mecánica respiratoria y en el intercambio de gases son compartidas por la mayoría de las NID y por lo tanto no aportan mayor información sobre el diagnóstico específico, pero sirven para evaluar la gravedad y controlar la evolución de la enfermedad y la respuesta al tratamiento [2].

#### **1.1.5.2 RADIOLOGÍA**

El estudio de los aspectos radiológicos es fundamental en la evaluación de estas enfermedades. El examen radiológico demuestra reducción del volumen pulmonar. Sin embargo, la radiografía de tórax es normal en el 15 % de los casos y en caso de existir anomalías visibles, la Rx de

tórax tiene limitada especificidad diagnóstica en pacientes con NID.

La TACAR ocupa un sitio muy importante en el estudio de NID correlacionando significativamente la severidad de los síntomas y el deterioro funcional y es particularmente útil para detectar una enfermedad pulmonar en pacientes asintomáticos o con Rx de tórax normal, también nos determina la distribución de las lesiones y nos puede sugerir el grado de actividad, la presencia o ausencia de alteraciones reversibles (agudas o activas) o irreversibles (fibróticas) y por consiguiente los mejores sitios para la toma de biopsias [3].

En las primeras etapas las imágenes que predominan son las de vidrio despolido, compatible con el estadio de alveolitis. El hallazgo de este patrón posiblemente indica una mejor respuesta al tratamiento esteroide, un mejor pronóstico y altas posibilidades de afección reversible. Cuando la enfermedad evoluciona al estadio terminal es visible el patrón de "pulmón en panal" [2].

#### **1.1.5.3 LAVADO BRONCOALVEOLAR**

En el lavado broncoalveolar (BAL), se observa un incremento del número de neutrófilos, eosinófilos y linfocitos con diferentes patrones para cada etiología. La especificidad es relativa ya que requiere un procesamiento técnico riguroso y laboratorios bien equipados con personal entrenado [2].

#### **1.1.5.4 BIOPSIA PULMONAR**

El método diagnóstico definitivo es la biopsia pulmonar, ya sea por BTB, videotoracoscopia o a cielo abierto. Es la herramienta más precisa, porque no sólo sirve para el diagnóstico, sino que también desempeña un papel fundamental en la planificación del tratamiento, evaluando la extensión y gravedad de la enfermedad y orientando el pronóstico. La biopsia pulmonar a cielo abierto frecuentemente es el elemento que se requiere para establecer el diagnóstico de enfermedad intersticial y en los casos de duda diagnóstica es el método de referencia debido a que el material obtenido por otros medios es insuficiente en cantidad [2] - [4].

### 1.1.6 EPIDEMIOLOGÍA

Existen diferencias entre estudios epidemiológicos debido tanto a diferencias reales en la incidencia, como a cambios en la definición y clasificación de la enfermedad y diferencias en el diseño del estudio.

La incidencia y predominio de las NID no se conoce con precisión, esto es debido a varios factores:

Se requiere de un buen grado de sospecha clínica para realizar el diagnóstico, obteniéndose después de numerosos estudios y en clínicas especializadas; no obstante, existe la probabilidad de un diagnóstico erróneo.

Las estadísticas son poco confiables en la mayoría de los países en vías de desarrollo y en un cierto número de países desarrollados [2].

Los registros de las diferentes enfermedades han sido realizados por grupos médicos que desconocen o subestiman la incidencia real de las NID [6].

En la tabla 1, se muestran datos estadísticos que arrojó el estudio en Bernalillo County, región de Nuevo México, acerca de la incidencia y la prevalencia de las NID. El registro se estableció sobre una región cercana a los 500,000 habitantes. El diagnóstico más frecuente fue el de fibrosis pulmonar y fibrosis pulmonar idiopática (45% del total de casos) [6].

Tabla 1. Predominio e incidencia de Enfermedades Intersticiales Pulmonares de octubre 1988 hasta septiembre 1990

AUTORES	ESTUDIO	INCIDENCIA		PREVALENCIA		PREVALENCIA DE CASOS NO DIAGNOSTICADOS
		sexo / población / año	sexo / población / año	sexo / población	sexo / población	
Coultas y asociados	Bernalillo County. USA	MUJ/ 10 <sup>5</sup> /AÑO 26.1	HOM/10 <sup>5</sup> /AÑO 31.5	MUJ/10 <sup>5</sup> 67.2	HOM/10 <sup>5</sup> 80.9	1.80%

En México, en el Instituto Nacional de Enfermedades Respiratorias, la NID que se observa con mayor frecuencia es la alveolitis alérgica extrínseca, una enfermedad producida por la respuesta inmune exagerada a la inhalación repetida de una gran variedad de partículas

orgánicas: proteínas de aves y mamíferos, hongos y bacterias termofílicas. El promedio de casos nuevos es de 60-70 casos por año y su predominio es mayor en el género femenino con una relación mujer/hombre de 8/1 [5].

## **1.2 AUSCULTACIÓN Y SONIDOS RESPIRATORIOS**

### **1.2.1 INTRODUCCIÓN**

El estetoscopio es una de las formas más antiguas de instrumentación biomédica, desde su invención en 1816 por el médico francés René Théophile Hyacinthe Laënnec (1781-1826) [7]. La auscultación del pulmón es uno de los métodos más importantes, simples y no invasivos para la detección de enfermedades [8]. Proporciona información directa sobre la estructura y la función del pulmón que no se puede obtener con ningún otro método simple y no invasivo [9]. Varios cambios patológicos en el pulmón, que producen sonidos característicos, se pueden detectar más fácilmente por la auscultación que por la radiografía debido a la conveniencia y la disponibilidad del estetoscopio [8]. Como es obvio resulta muy útil estudiar, analizar y comprender los sonidos provenientes del pulmón.

Durante el procedimiento de auscultación pulmonar el médico escucha los sonidos producidos por la actividad respiratoria en la superficie torácica del paciente, en un esfuerzo por detectar la presencia de anormalidades. El diagnóstico correspondiente se realiza en términos de su habilidad y experiencia [10]. Sin embargo, el estetoscopio posee la desventaja de no proporcionar fidelidad acústica, amplificar o atenuar los sonidos dentro del espectro clínico de interés (amplifica el sonido por debajo de los 112 Hz y lo atenúa en altas frecuencias, aproximadamente por encima de 120 Hz) [11], [12].

Los sonidos respiratorios contienen información relevante acerca de la fisiología y patología de los pulmones y vías aéreas. Estos sonidos son producidos por el movimiento de aire a través de las vías aéreas y transmitidos a lo largo de las estructuras pulmonares. Las características

acústicas de los sonidos respiratorios varían significativamente entre sonidos normales y anormales (adventicios). Los sonidos anormales pueden poseer características acústicas que indican la presencia, severidad y localización de la enfermedad en el sistema respiratorio [13].

El propósito del procesamiento digital de las señales de sonidos respiratorios ha sido disminuir el carácter subjetivo de la técnica de auscultación pulmonar por medio de la extracción y cuantificación de características importantes que no pueden ser apreciadas por el médico con el estetoscopio.

### **1.2.2 CARACTERÍSTICAS DE LOS SONIDOS RESPIRATORIOS NORMALES Y ADVENTICIOS**

El análisis de los sonidos respiratorios es un campo multidisciplinario, por esto los términos y conceptos involucrados deben estar claramente definidos [14]. Sin embargo la terminología ha variado dependiendo de la fecha del estudio y de la nacionalidad del autor [10]. A continuación se describen los términos que se manejan durante el desarrollo de la presente tesis basándose en los términos y técnicas estándares recomendados por la Sociedad Europea Respiratoria (ERS Task Force Report. Computerized Respiratory Sounds Analysis, CORSA), [15].

Revisiones recientes definen a los sonidos respiratorios como todos los sonidos relacionados con la respiración, incluyendo sonidos ventilatorios (breath sounds), sonidos adventicios, sonidos producidos por la tos, los ronquidos, los estornudos y los músculos respiratorios. Los sonidos ventilatorios son aquellos producidos por la respiración, excluyendo a los sonidos adventicios, escuchados sobre la pared torácica, la tráquea y la boca. La generación de los sonidos ventilatorios está relacionada con el flujo de aire en el tracto respiratorio. Acústicamente se caracterizan por presentar un amplio espectro de frecuencia cuyo rango depende de la posición donde se escuche o grabe el sonido [14].

Por otro lado, cuando existe un desorden o patología pulmonar, usualmente se producen otro tipo de sonidos adicionales a los sonidos ventilatorios, que son llamados sonidos adventicios y pueden ser sonidos continuos (sibilancias), discontinuos (estertores crepitantes) o una combinación de ambos (squawks) escuchados también sobre la pared torácica, la tráquea y la boca. Finalmente, todos aquellos sonidos escuchados sobre la pared torácica o dentro del tórax

son conocidos como sonidos pulmonares [14]. En la figura 2 se indican los fundamentos para definir adecuadamente la terminología utilizada en el estudio de los sonidos respiratorios.

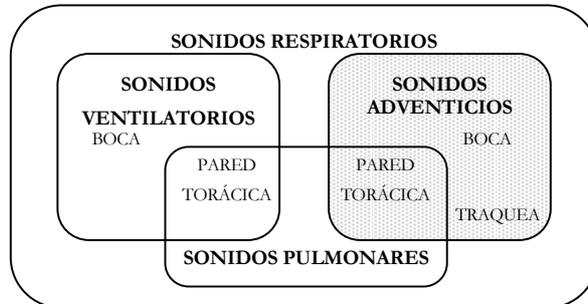


Fig. 2 Relación entre sonidos ventilatorios, sonidos adventicios, sonidos pulmonares y sonidos respiratorios. Extraído de referencia [14].

### 1.2.2.1 SONIDOS RESPIRATORIOS NORMALES

Los sonidos producidos por la actividad respiratoria en sujetos sanos son llamados Sonidos Respiratorios Normales y se diferencian según su localización en la superficie torácica y varían de acuerdo al ciclo respiratorio y primariamente con el cuadrado del flujo de aire [14]. En el origen de los sonidos generados por ventilación están involucrados múltiples mecanismos por lo que no está completamente claro [13].

Los sonidos ventilatorios son producidos probablemente por la turbulencia del aire a nivel de los bronquios lobulares o segmentarios. En bronquios más pequeños, la velocidad del aire es menor que la velocidad mínima requerida para inducir turbulencias. Por lo tanto, el flujo de aire en vías aéreas pequeñas se considera que es de tipo laminar y silencioso [13].

Se consideran dos tipos básicos de sonidos respiratorios, los *sonidos traqueo-bronquiales*, escuchados sobre la pared torácica en la proyección de dichas estructuras y los *sonidos vesiculares* que se perciben sobre la pared torácica en sitios periféricos a las vías aéreas primarias, principalmente sobre las bases pulmonares [18]. El rango en frecuencias de los sonidos respiratorios se encuentra entre 100 a 2000 Hz y el mayor contenido espectral se encuentra por debajo de los 1000 Hz en el caso de los sonidos traqueo-bronquiales y menor a 500 Hz en el de

los sonidos vesiculares. Sin embargo, esta banda de frecuencias también contiene componentes provenientes de la actividad de los músculos respiratorios y del corazón [13], [14], [18].

El sonido proveniente de las vías aéreas más grandes posee un espectro amplio en frecuencia, el cual es transmitido a la superficie torácica después de ser filtrado, en forma pasa-bajas por los pulmones y la pared del tórax. El espectro en frecuencia de los sonidos pulmonares se encuentra normalmente entre 100 y 200 Hz, pero pueden ser detectados en 1000 Hz o más con micrófonos sensibles en una sala libre de ruidos ambientales [13].

El sonido respiratorio registrado en la superficie torácica presenta una fase inspiratoria más prolongada en tiempo que la fase espiratoria, con un cociente inspiración / espiración alrededor de 2/1, durante la respiración corriente. Los sonidos respiratorios no presentan características uniformes a lo largo de la superficie pulmonar. En el ápice pulmonar, el sonido es menos intenso durante una inspiración realizada a partir del volumen residual. En contraste, en la base pulmonar el sonido es menos intenso al inicio de la fase inspiratoria y, gradualmente se incrementa hasta alcanzar su máximo alrededor del 50% de la capacidad vital [13].

Cuando el sonido es grabado sobre la tráquea, el sonido no es filtrado, por lo tanto, debido al fenómeno de resonancia, el espectro en frecuencia contiene un rango de ~100 Hz a 1500 Hz [12], la banda de energía principal se extiende de 850 a 1000 Hz [13]. En la figura 3 se muestra un ejemplo de sonido pulmonar en un sujeto sano.

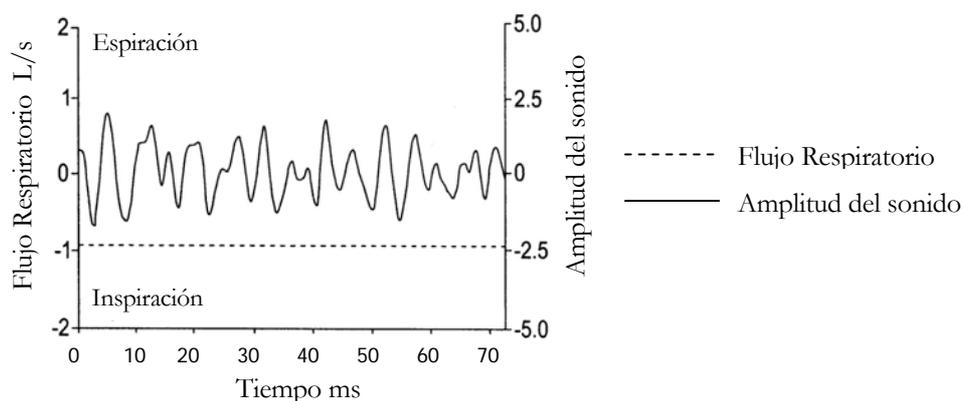


Fig. 3 Ejemplo de sonido pulmonar de un sujeto sano. Extraído de referencia [13].

### 1.2.2.2 SONIDOS ADVENTICIOS

Los cambios en la estructura pulmonar que ocurren cuando existe alguna enfermedad afectan la amplitud y el comportamiento de la transmisión del sonido de las vías aéreas a la superficie torácica. Los sonidos adventicios pueden clasificarse principalmente en sonidos continuos como las sibilancias y discontinuos como los estertores crepitantes [12].

A continuación se indican las características de los tipos más importantes de sonidos adventicios.

#### 1.2.2.2.1 Sibilancias

Las sibilancias son sonidos adventicios del tipo continuo, que se superponen a los sonidos respiratorios normales. La morfología de un sonido silbante se asemeja a la morfología de un sonido sinusoidal (Figura 4). De acuerdo a la definición de la Sociedad Torácica Americana (STA), la palabra “continuo” significa que la duración de una sibilancia es mayor a los 250 mseg, con la característica de ser sonidos de tono alto (sonidos agudos) y contiene una frecuencia dominante mayor o igual a los 400 Hz [12].

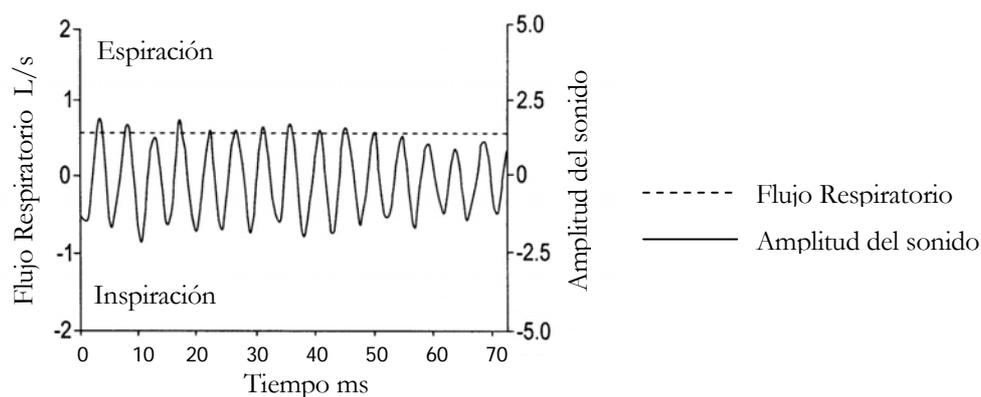


Fig. 4 Ejemplo de sonido pulmonar de un paciente asmático presentando una sibilancia. Extraído de referencia [13].

La frecuencia dominante de una sibilancia es usualmente mayor que 100 Hz con una duración mayor a 100 ms, de acuerdo con las definiciones de CORSA. Pueden ser monofónicas (una frecuencia) o polifónicas (varias frecuencias) [15].

Recientemente, Grotberg y Gavriely propusieron un modelo en el cual las sibilancias son producidas por la agitación de las vías aéreas. Las oscilaciones comienzan cuando la velocidad del flujo de aire alcanza un valor crítico. Las sibilancias van siempre acompañadas de una limitación del flujo, sin embargo, la limitación de flujo no implica la producción de sibilancias.

Las sibilancias se pueden presentar en diversas enfermedades, son signos clínicos comunes en pacientes con enfermedades obstructivas, y se ha demostrado en varios estudios una asociación entre el grado de obstrucción bronquial y la presencia de sibilancias (tiempo sibilancias/tiempo ciclo respiratorio) [13].

#### **1.2.2.2.2 Crepitancias**

Los estertores crepitantes son sonidos adventicios discontinuos de carácter explosivo y transitorio. Su duración es menor a los 20 mseg y su contenido en frecuencia se ha considerado de 100 hasta los 2000 Hz [13], [16]. Se pueden distinguir dos tipos de crepitancias: gruesas y el finas. Se asume que ambos tipos de crepitancias se originan por la energía acústica generada por la eualización de presiones, o por un cambio en la tensión elástica, después de una apertura repentina de vías aéreas que se encuentran anormalmente cerradas. En la figura 5 se muestra la señal de sonido respiratorio de un paciente con Alveolitis Alérgica Extrínseca con 2 crepitancias finas típicas [13].

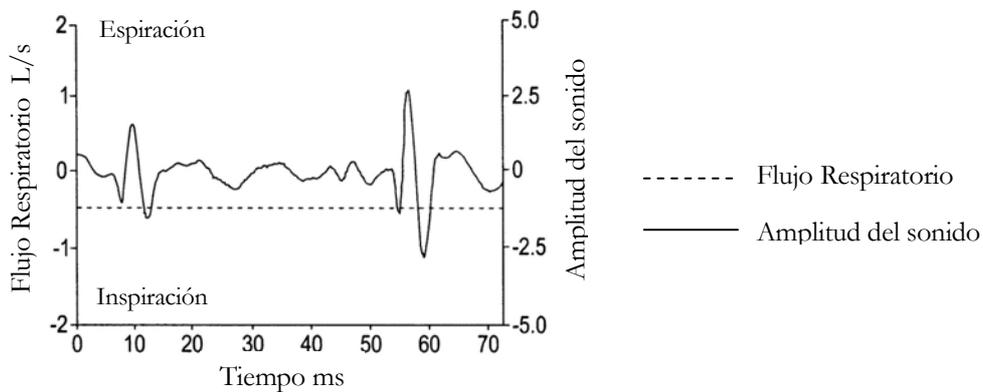


Fig. 5 Ejemplo de sonido pulmonar de un paciente con Alveolitis Alérgica presentando crepitancias. Extraído de referencia [13].

En pacientes con fibrosis pulmonar las crepitancias son típicamente finos y repetitivos, que ocurren al final de la fase inspiratoria [13], [17]. La presencia de crepitancias puede ser un signo temprano de enfermedad, como en el caso de la asbestosis. Las crepitancias tienden a ocurrir primero en las bases de los pulmones y posteriormente, cuando progresa la enfermedad, aparecen también en zonas apicales. Además, el número de crepitancias por respiración se asocia con la severidad de la enfermedad en pacientes con enfermedades pulmonares intersticiales. La forma de onda y localización de las crepitancias puede tener influencia en el diagnóstico de enfermedades cardio-respiratorias [13], [16].

En la siguiente tabla se presentan en forma resumida las principales características que ayudan a diferenciar los sonidos normales y adventicios durante la auscultación del tórax [12], [15], [19] y [20].

Tabla 1.- Categorías de los sonidos respiratorios frecuentemente escuchados durante la auscultación del tórax [12, 15, 19, 20].

	<b>Sonidos Respiratorios</b>	<b>Origen</b>	<b>Mecanismo de producción</b>	<b>Características Acústicas</b>	<b>Relevancia Clínica</b>
<b>Sonidos Normales</b>	Traqueal Normal	Faringe, laringe tráquea, bronquios principales	Flujo turbulento, impacto del flujo sobre la pared de vías aéreas	<100 a >3000 Hz Sonido con resonancias	Configuración de vías aéreas superiores
	Pulmonar Normal	Vías aéreas centrales (expiración) Vías aéreas lobares a segmentarias (inspiración), tensión alveolar? Otras	Flujo turbulento, vórtices, mecanismos desconocidos	<100 a >1000 Hz Sonido sometido a filtrado pasa-bajas	Ventilación regional, calibre de vías aéreas
<b>Sonidos Adventicios</b>	Sibilantes	Vías aéreas centrales y periféricas	Oscilaciones de la pared de vías aéreas, dispersión de un vórtice	Sinusoide, continuo, ~100 a 1000 Hz, duración >80 ms	Aumento de la resistencia al flujo por obstrucción de vías aéreas
	Roncantes	Vías aéreas superiores y centrales	Ruptura súbita de capas de líquido, vibraciones de vías aéreas superiores o inferiores	Serie de sinusoides amortiguados, <300 Hz, duración >100 ms	Aumento de la resistencia por colapsabilidad de vías aéreas, secreciones
	Crepitantes	Vías aéreas centrales y periféricas	Choque de presiones (tensión-relajación con oscilaciones bruscas de vías aéreas)	Ondas rápidamente amortiguadas, 200 a >1000 Hz, duración <20 ms	Cierre y posterior apertura súbita de vías aéreas, secreciones

### 1.3 REFERENCIAS

- [1] Sistema Nacional de Epidemiología, “Morbilidad 2003.” *Epidemiología, Sistema Único de Información*, 21(15), pp. 1-42, 2003.
- [2] L. M. Selman, “Neumopatías Intersticiales Difusas”, 1ª Edición, Editorial Médica Panamericana, 1996.
- [3] J. Giannattasio, A. Miranda, L. Zabala, “Neumonías Difusas Intersticio Alveolares.” *Rev. Cubana Med*, 40(2), pp. 135-44, 2001.
- [4] L. M. Selman, R. Barrios, “Intertitial pulmonary diseases: Selected Topics.” CRC Press, Inc., 1991.
- [5] R. Chapela, H. J. Salas, R. Sansores, R. Pérez-Padilla, G. Carrillo, M. Sánchez, R. Pérez-Chavira, “Alveolitis Alérgica Extrínseca.” *Neumología y Cirugía de Tórax*, 50(2), pp. 9-27, 1991.
- [6] M. Demedts, A. U. Well, J. M. Antó, U. Costabel, R. Hubbard, P. Cullinan, H. Slabbynck, G. Rizzato, V. Poletti, E. K. Verbeken, M. J. Thomeer, J. Kokkarinen, J. C. Dalphin, A. Newman Taylor, “Interstitial lung diseases: an epidemiological overview.” *Eur. Respir. J.*, 8(32), pp. 2-6, 2001.
- [7] R. T. H. Laënnec, “A treatise on the diseases of the chest.” New York: Hafner, pp. 282-284, 1962.
- [8] R. G. Loudon, “The lung speaks out.” *Am. Rev. Respir. Dis.* 27, pp. 411-412, 1982.
- [9] G. Charbonneau, J. L. Racineux, M. Sudraud, and E. Tuchais, “An accurate recording system and its use in breath sounds spectral analysis.” *J. Appl. Physiol.* 55, pp. 1120-1127, 1982.
- [10] L. R. Waitman, “Representation and classification of breath sounds recorded in an intensive care setting using neural networks.” Master thesis. Vanderbilt University, 1998.
- [11] A. R. A. Sovijärvi, J. Vanderschoot, J. E. Earis, “Standardization of computerized respiratory sound analysis.” *European Respiratory Review*, 10(77), pp. 585, 2000.

- [12] H. Pasterkamp, S. Kraman, G. Wodicka, "Respiratory sounds. Advances beyond the stethoscope." *American Journal of Respiratory and Critical Care Medicine*, 156(3), pp. 974-987, 1997.
- [13] A. R. A. Sovijärvi, L. P. Malmberg, G. Charbonneau, J. Vanderschoot, F. Dalmaso, C. Sacco, M. Rossi, J. E. Earis, "Characteristics of breath sounds and adventitious respiratory sounds." *European Respiratory Review*, 10(77), pp. 591-596, 2000.
- [14] A. R. A. Sovijärvi, F. Dalmaso, J. Vanderschoot, L. P. Malmberg, G. Righini, S. A. T. Stoneman, "Definition of terms for applications of respiratory sounds." *European Respiratory Review*, 10(77), pp. 597-610, 2000.
- [15] (ERS Task Force Report. Computerized Respiratory Sounds Analysis CORSA): recommended standards for terms and techniques. *Eur. Respir. Rev.* 77, 2000.
- [16] P. Forgacs, "Crackles and wheezes." *The Lancet*, 1967.
- [17] G. R. Epler, C. B. Carrington, E. A. Gaensler, "Crackles (rales) in the interstitial pulmonary diseases." *Chest*, 73, pp. 333-339, 1978.
- [18] R. González-Camarena, S. Carrasco, S. Charleston, J. Gaitán, M. G. Ramos, G. Chillem, J. Pérez-Padilla, "Sistema para el análisis cuantitativo de los sonidos respiratorios." *Revista Mexicana de Ingeniería Biomédica*, 18: 2, pp. 7-23, 1997.
- [19] R. L. H. Murphy, S. K. Holford, "Lung sounds." *Basics of RD ATS News*; 8, pp. 24-29, 1980.
- [20] R. Mikami, et al International symposium on lung sounds: Synopsis of proceedings. *Chest*, 92, pp. 342-345, 1987.

## CAPÍTULO II

### 2.1 PLANTEAMIENTO DEL PROBLEMA

El diagnóstico de NID es un proceso largo y complicado que se determina a través de varios estudios médicos, con personal entrenado y en clínicas especializadas. Esta enfermedad se caracteriza por ser de comienzo insidioso y normalmente se detecta en etapas avanzadas, lo cual provoca la falta de tratamiento adecuado y el desarrollo progresivo de la enfermedad hasta hacerla irreversible y mortal para el paciente en un lapso de tiempo relativamente corto.

En la Figura 1 se presenta, a manera de ejemplo, la estadística para un tipo de NID, la Alveolitis Alérgica Extrínseca (AAE), una enfermedad de las más comunes de la población que asiste para atención médica al Instituto Nacional de Enfermedades Respiratorias (INER). El 40% de los pacientes con AAE evoluciona hacia la fibrosis, de los cuales un 30% fallece en un promedio de 3 a 5 años. Alrededor de un 40% queda con alteraciones clínicas, radiológicas o funcionales de diferentes grados y, solamente un 20% se recupera [1].

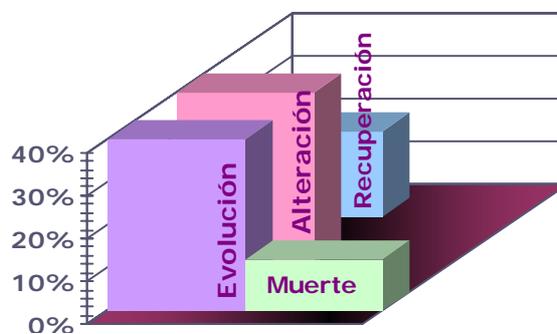


Fig. 1 Estadísticas de la NID Alveolitis Alérgica Extrínseca en el INER.

En la actualidad no existe una técnica precisa para el diagnóstico específico del tipo de NID. Sin embargo, la biopsia pulmonar a cielo abierto (BPCA) proporciona información directamente del tejido y es el método de referencia para el diagnóstico definitivo del padecimiento. La BPCA es utilizada como último recurso porque posee la desventaja de ser puntual e invasiva, razón por la cual no puede efectuarse de forma regular para realizar un

seguimiento de la evolución de la enfermedad. En estas condiciones se dificulta el diagnóstico, tratamiento y seguimiento del paciente y es factible cometer equivocaciones que conduzcan al empeoramiento de la enfermedad [2], [3], [4]. Una de las técnicas más simples y no invasivas, es la auscultación clínica, pero tiene ciertas limitaciones para integrar de manera cuantitativa la información acústica espacial del tórax debido principalmente a desventajas técnicas que presenta el estetoscopio, que conducen al médico a un diagnóstico subjetivo de la enfermedad. Por lo anterior, es importante el desarrollo de nuevas técnicas no invasivas que permitan el análisis cuantitativo de la información espacial del tórax de un sujeto para contribuir a una mejor evaluación del estado clínico del paciente.

## **2.2 HIPÓTESIS**

El análisis multicanal de los sonidos respiratorios proporciona información importante que por sí misma, o en combinación con otro tipo de descripciones (datos clínicos), permite diferenciar sonidos respiratorios normales de los anormales, en presencia de NID durante diferentes etapas de lesión.

## **2.3 OBJETIVO**

### **2.3.1 OBJETIVO GENERAL**

Evaluar la información acústica multicanal a través de técnicas de procesamiento digital de señales y comprobar si es posible utilizarla como herramienta para apoyar el diagnóstico de las Neumopatías Intersticiales Difusas por medio de la caracterización cuantitativa y la clasificación de los sonidos respiratorios adquiridos que incluyen al sonido respiratorio básico y a los sonidos adventicios conocidos como estertores crepitantes.

### **2.3.2 OBJETIVOS ESPECÍFICOS**

1. Adquirir en forma multicanal y simultánea información acústica en diferentes posiciones del tórax en sujetos sanos y en pacientes con NID.
2. Programar una interfaz gráfica amigable al usuario médico que permita manipular, analizar y procesar las señales acústicas multicanales.
3. Procesar las señales de sonido respiratorio utilizando diferentes técnicas de parametrización univariadas y multivariadas para representarlas mediante un número reducido de datos.
4. Encontrar el número adecuado de características que representen a la señal de sonido respiratorio sin pérdida importante de información para la discriminación del sonido normal y anormal.
5. Evaluar cada una de las técnicas utilizando los parámetros obtenidos como entradas de un clasificador y comparar los resultados obtenidos, analizando los porcentajes que indican el desempeño del clasificador para la correcta clasificación del sonido normal y anormal.

## CAPÍTULO III

### TÉCNICAS DE PARAMETRIZACIÓN

#### 3.1 INTRODUCCIÓN

Desde el punto de vista de procesamiento de señales, los sonidos respiratorios se consideran señales no estacionarias, cuasiperiódicas y con distribución gaussiana [1].

Una señal aleatoria, estocástica o no determinística es aquella que varía de tal forma, que no puede ser descrita por una ecuación matemática, tampoco es posible predecir su comportamiento mediante una regla, sino que es descrita en términos de sus propiedades estadísticas formando un modelo matemático conocido como *proceso aleatorio* [2], [3].

Los sonidos respiratorios poseen una naturaleza *aleatoria* debido a la generación y propagación del sonido dentro del tórax y no es posible predecir su comportamiento por una ecuación, sin embargo presentan características similares de un ciclo respiratorio a otro. El ciclo de la respiración consiste de una inspiración seguida de una espiración. Sin embargo, la respiración es variable, por lo que cada ciclo es diferente en duración, amplitud y morfología. Por este motivo, la señal de sonido respiratorio es una señal *cuasiperiódica*.

La estacionaridad implica que los parámetros estadísticos de la señal no varían con el tiempo. Es decir, la función de densidad o los momentos, son independientes de donde sean evaluados a lo largo del proceso aleatorio [4]. La estadística debe ser dependiente solamente de los espacios entre las muestras y no de las posiciones absolutas. Las señales respiratorias son inherentemente *no estacionarias* debido al cambio en el flujo de aire y volumen del pulmón durante un ciclo respiratorio [5].

Un proceso aleatorio discreto está formado por un conjunto de secuencias representativas de un experimento aleatorio. Dicho conjunto es denominado el *ensamble* del proceso aleatorio. Se dice que un proceso aleatorio es *ergódico* si es estacionario y si el promedio del *ensamble* es igual a los promedios temporales, obtenidos a partir de una realización. La ergodicidad implica

estacionaridad, sin embargo, lo recíproco no siempre se cumple [4], [6].

En la figura 1 se muestra el sonido respiratorio de un sujeto sano registrado sobre la pared torácica.

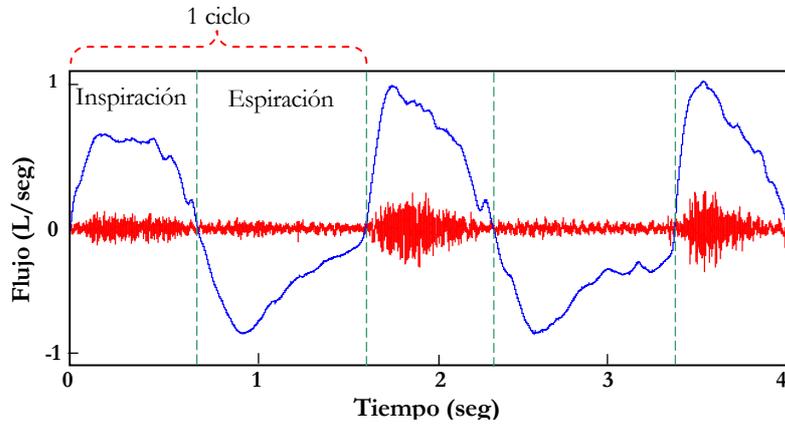


Fig.1 Señal de sonido respiratorio. Se muestra un ciclo formado por una inspiración seguida de una espiración.

La distribución de probabilidad de los sonidos respiratorios se considera aproximadamente *normal o gaussiana* [1], en la figura 2 se ilustra un segmento de señal de sonido respiratorio con su correspondiente distribución de probabilidad.

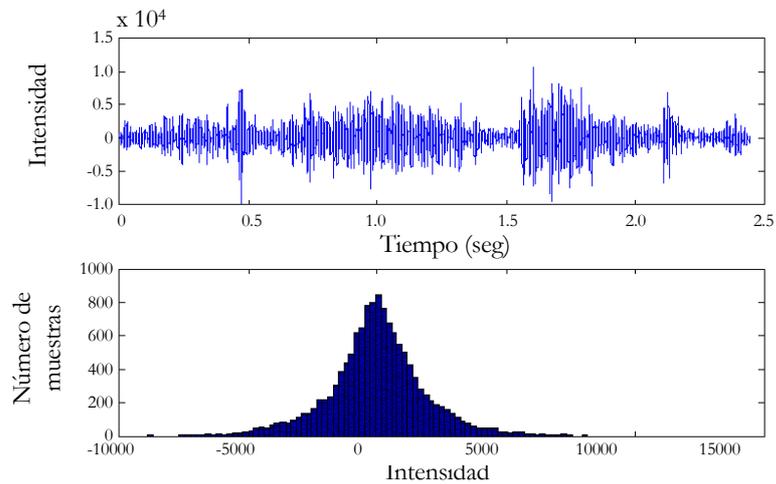


Fig. 2 Distribución de probabilidad de la señal de sonido respiratorio. Extraído de referencia [1]

## 3.2 DESCRIPCIÓN DE MÉTODOS

### 3.2.1 MEDIDAS CLÁSICAS EN EL DOMINIO DE LA FRECUENCIA

Este método ha sido utilizado para parametrizar señales en el dominio de la frecuencia a través de valores de frecuencia denominadas *frecuencias percentiles*. Para el cálculo de los parámetros en el dominio de la frecuencia se realiza la estimación del espectro de potencia (EP) del vector de datos. Los valores de frecuencia del espectro estimado, se integran para formar el área total bajo la curva del espectro. Una frecuencia percentil [12], representa un determinado porcentaje del área total bajo la curva del espectro de potencia (EP), como se ilustra en la figura 3, en donde  $f_{25}$  es el valor de frecuencia que representa el 25% del área total bajo la curva del EP y así sucesivamente.

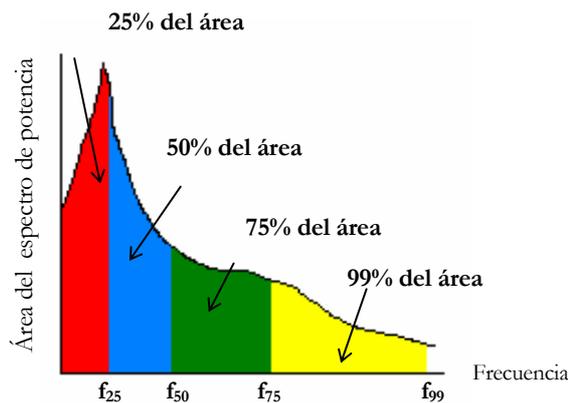


Fig. 3 Frecuencias percentiles

Existen diferentes métodos para estimar el espectro de potencia. Para el cálculo de las frecuencias percentiles se utilizó un espectro de potencia estimado mediante el método de Burg que se describe a continuación.

### 3.2.1.1 MÉTODO DE BURG PARA ESTIMACIÓN ESPECTRAL

John Parker Burg desarrolló un método para estimación espectral basado en el modelo AR al que denominó “método de máxima entropía”. Este método estima los parámetros AR directamente de los datos sin pasos intermedios de cálculo de la matriz de correlación y resolución de las ecuaciones de Yule-Walker. El método de Burg está basado en la recursión de Levinson-Durbin y en la estructura de Lattice [3], [11].

En contraste con otros métodos de estimación espectral basados en modelos AR, el método de Burg asegura un modelo AR estable y es computacionalmente eficiente. La precisión del método es menor para modelos de alto orden, datos de gran longitud y relación señal a ruido alto, en estos casos el método genera picos en el espectro estimado [11].

Cuando el número de datos es pequeño, el método de Burg proporciona mayor resolución en frecuencia. Sin embargo, es importante considerar que el resultado en el dominio de la frecuencia, puede ser muy diferente dependiendo de la elección del orden del modelo AR.

La estimación espectral  $\hat{S}_x(e^{j\omega})$  mediante modelado AR [4], se expresa de la siguiente forma:

$$\hat{S}_x(e^{j\omega}) = \frac{\sigma_\omega^2}{|A(e^{j\omega})|^2} \quad (1)$$

Donde  $\sigma_\omega^2$  es la varianza del error de predicción y  $A(e^{j\omega})$  es igual a  $A(z)$ , descrita por el polinomio  $1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_p z^{-p}$ , que se explicará con más detalle en la siguiente sección.

### 3.2.2 MODELO AUTORREGRESIVO MONOCANAL (AR)

Para formular el modelo AR, primero definamos el filtro de error de predicción que produce la salida  $\varepsilon[n]$  de la entrada  $x[n]$ , este filtro es de tipo FIR cuya salida es [4]:

$$\varepsilon[n] = x[n] + a_1^* x[n-1] + a_2^* x[n-2] + \dots + a_p^* x[n-p] \quad (2)$$

En la figura 4 se representa el filtro cuya función de transferencia está dada por:

$$A(z) = 1 + a_1^* z^{-1} + a_2^* z^{-2} + \dots + a_p^* z^{-p} \quad (3)$$

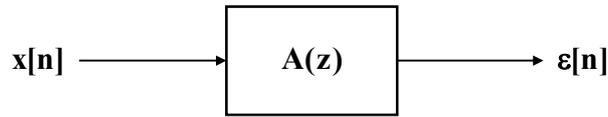


Fig. 4 Modelo de predicción lineal. Filtro de error de predicción

Si el orden de predicción  $p$  del filtro es adecuado, entonces la salida es aproximadamente un proceso de ruido blanco con varianza  $\sigma_{\epsilon}^2$ .

Si el filtro es invertido y excitado con una secuencia de ruido blanco  $\omega[n]$  con varianza  $\sigma_{\omega}^2 = \sigma_{\epsilon}^2$ , como se muestra en la figura 5:

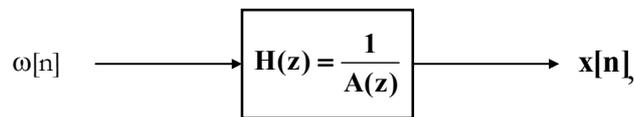


Fig. 5 Modelo Autorregresivo

este sistema producirá una secuencia aleatoria con las mismas características estadísticas que la secuencia original  $x[n]$ . Entonces, cualquier proceso aleatorio estacionario puede ser representado como la salida de un filtro lineal autorregresivo (AR), invariante al corrimiento que ha sido excitado por ruido blanco.

El filtro lineal AR es un filtro IIR con función de transferencia:

$$\mathbf{H}(z) = \frac{1}{\mathbf{A}(z)} \quad (4)$$

donde:

$$\mathbf{A}(z) = 1 + \mathbf{a}_1 z^{-1} + \mathbf{a}_2 z^{-2} + \dots + \mathbf{a}_p z^{-p} \quad (5)$$

y como puede observarse,  $\mathbf{A}(z)$  involucra únicamente potencias negativas de  $z$ , y por lo tanto, el filtro tiene  $p$  ceros en el origen y  $p$  polos distribuidos en el plano  $z$ . Por este motivo, el modelo AR es un sistema de únicamente polos.

Si los coeficientes  $\mathbf{a}_i$  son encontrados resolviendo las ecuaciones normales con una matriz de correlación definida positiva, todos los polos caen dentro del círculo unitario.

Por otro lado, si se considera al sistema escrito en forma de regresión estadística:

$$\mathbf{x}[n] = -\mathbf{a}_1\mathbf{x}[n-1] - \mathbf{a}_2\mathbf{x}[n-2] - \dots - \mathbf{a}_p\mathbf{x}[n-p] + \boldsymbol{\omega}[n], \quad (6)$$

la variable dependiente  $\mathbf{x}[n]$  está representada como una combinación lineal de las variables independientes  $\mathbf{x}[n-1]$  hasta  $\mathbf{x}[n-p]$  ponderadas por un parámetro escalar  $\mathbf{a}_1$  hasta  $\mathbf{a}_p$ , más una señal de entrada aleatoria, no correlacionada  $\boldsymbol{\omega}[n]$ . Debido a que las variables pertenecen a un mismo proceso aleatorio,  $\mathbf{x}[n]$  el proceso se denomina AUTORREGRESIVO (AR) [4].

La forma simplificada del modelo AR está especificada por

$$\mathbf{x}[n] = \sum_{i=1}^p \mathbf{a}_i \mathbf{x}[n-i] + \boldsymbol{\omega}[n] \quad (7)$$

donde:  $\mathbf{a}_i$  es el término escalar que representa a los coeficientes del modelo AR,  $p$  es el orden del modelo AR y  $\boldsymbol{\omega}[n]$  es la secuencia de ruido blanco.

Existen varias técnicas para estimar los coeficientes AR. Las principales categorías son mínimos cuadrados (Least Squares LS) y el método de Burg, cada técnica con sus respectivas variantes, el método más común, LS está basado en las ecuaciones de Yule Walker.

Los coeficientes del modelo AR,  $\mathbf{a}_1 \dots \mathbf{a}_p$  y  $\sigma_{\boldsymbol{\omega}}^2$  también pueden ser estimados mediante la resolución de las ecuaciones normales. Si se escribe el modelo para el proceso aleatorio de la ecuación (6) de la forma:

$$\mathbf{x}[n] + \mathbf{a}_1\mathbf{x}[n-1] + \mathbf{a}_2\mathbf{x}[n-2] + \dots + \mathbf{a}_p\mathbf{x}[n-p] = \boldsymbol{\omega}[n], \quad (8)$$

entonces, la función de autocorrelación del proceso aleatorio satisface la ecuación de diferencias:

$$\mathbf{R}_x[\mathbf{t}] + \mathbf{a}_1\mathbf{R}_x[\mathbf{t}-1] + \dots + \mathbf{a}_p\mathbf{R}_x[\mathbf{t}-p] = \mathbf{R}_{\boldsymbol{\omega}x}[\mathbf{t}], \quad (9)$$

en donde  $\mathbf{t}$  es el espacio entre muestras.

Si  $\mathbf{h}[n]$  es la respuesta a impulso del modelo para un proceso aleatorio, entonces  $\mathbf{R}_{x\boldsymbol{\omega}}[\mathbf{t}]$  está dado por:

$$\mathbf{R}_{x\boldsymbol{\omega}}[\mathbf{t}] = \mathbf{h}[\mathbf{t}] * \mathbf{R}_{\boldsymbol{\omega}}[\mathbf{t}] = \mathbf{h}[\mathbf{t}] * \sigma_{\boldsymbol{\omega}}^2 \boldsymbol{\delta}[\mathbf{t}] = \sigma_{\boldsymbol{\omega}}^2 \mathbf{h}[\mathbf{t}] \quad (10)$$

Recordando que por tratarse de un proceso aleatorio estacionario y real:  $\mathbf{R}_{x\boldsymbol{\omega}}[\mathbf{t}] = \mathbf{R}_{\boldsymbol{\omega}x}[-\mathbf{t}]$ :

$$\mathbf{R}_{\omega x}[\iota] = \sigma_{\omega}^2 \mathbf{h}^*[-\iota] \quad (11)$$

sustituyendo (10) en (8):

$$\mathbf{R}_x[\iota] + \mathbf{a}_1 \mathbf{R}_x[\iota-1] + \dots + \mathbf{a}_p \mathbf{R}_x[\iota-p] = \sigma_{\omega}^2 \mathbf{h}^*[-\iota] \quad (12)$$

y considerando que  $\mathbf{h}[\mathbf{n}]$  es la respuesta a impulso del filtro causal,  $\mathbf{h}[\mathbf{n}]$  es igual a cero para  $\mathbf{n} < \mathbf{0}$  y del teorema del valor inicial:

$$\mathbf{h}[0] = \lim_{z \rightarrow \infty} \mathbf{H}(z) = \lim_{z \rightarrow \infty} \frac{1}{1 + \mathbf{a}_1 z^{-1} + \dots + \mathbf{a}_p z^{-p}} = 1 \quad (13)$$

donde  $\sigma_{\omega}^2 \mathbf{h}^*[-\iota]$  toma el valor de:

$$\begin{array}{ll} \sigma_{\omega}^2 & \text{para } \iota = 0 \\ 0 & \text{para } \iota > 0 \end{array} \quad \text{evaluada para } \iota = 0, 1, 2, \dots, p \quad (14)$$

Si se representa en forma de matriz a (11), se obtienen las ecuaciones de Yule-Walker para el modelo AR:

$$\begin{bmatrix} \mathbf{R}_x[0] & \mathbf{R}_x[-1] & \dots & \mathbf{R}_x[-p] \\ \mathbf{R}_x[1] & \mathbf{R}_x[0] & \dots & \mathbf{R}_x[-p+1] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_x[p] & \mathbf{R}_x[p-1] & \dots & \mathbf{R}_x[0] \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_p \end{bmatrix} = \begin{bmatrix} \sigma_{\omega}^2 \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix} \quad (15)$$

Conjugando ambos sitios y tomando en cuenta que  $\mathbf{R}_x^*[\iota] = \mathbf{R}_x[-\iota]$  se obtienen las ecuaciones normales de predicción lineal:

$$\begin{bmatrix} \mathbf{R}_x[0] & \mathbf{R}_x[1] & \dots & \mathbf{R}_x[p] \\ \mathbf{R}_x[-1] & \mathbf{R}_x[0] & \dots & \mathbf{R}_x[p-1] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_x[-p] & \mathbf{R}_x[-p+1] & \dots & \mathbf{R}_x[0] \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{a}_1^* \\ \vdots \\ \mathbf{a}_p^* \end{bmatrix} = \begin{bmatrix} \sigma_{\omega}^2 \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix} \quad (16)$$

Las ecuaciones de Yule Walker involucran parámetros del modelo AR y la matriz de correlación, mientras que las ecuaciones normales involucran parámetros de predicción lineal y la matriz de correlación con las columnas y renglones en orden invertido. En casos reales, la matriz es simétrica en vez de Hermitian simétrica y ambos conjuntos de ecuaciones son idénticos. Una matriz es Hermitian cuando es igual a su conjugada compleja.

Los coeficientes AR son estimados con el método de Burg el cual estima directamente los coeficientes de reflexión y utiliza la recursión de Levinson para obtener los parámetros AR [2].

### 3.2.2.1 ORDEN DEL MODELO AR

#### CRITERIO DE INFORMACIÓN DE AKAIKE

Un modelo es una aproximación de la realidad. Uno de los aspectos más importantes del uso de modelos AR es la selección del orden  $\mathbf{p}$ , que permite el mejor ajuste de los datos minimizando el error entre los datos estimados por los modelos propuestos y los datos reales. La característica de este error, es que generalmente decrece conforme el orden del modelo se incrementa. Cuando la velocidad de decrecimiento del error es relativamente lenta, se estabiliza o vuelve a incrementarse, el proceso de selección del orden del modelo se detiene.

Dos de los criterios más conocidos para la selección del orden del modelo han sido propuestos por Akaike:

El primero es el Criterio del Error de Predicción Final (FPE) [9], que selecciona el orden del modelo AR de tal forma que la varianza del error de predicción es minimizado sin conducir a un alto orden innecesario. El FPE puede ser calculado de la forma:

$$\mathbf{FPE}[\mathbf{p}] = \sigma_{\omega}^2 \left( \frac{N + (\mathbf{p} + 1)}{N - (\mathbf{p} + 1)} \right) \quad (17)$$

Utilizando el segundo Criterio de Información de Akaike (AIC) [10], el orden óptimo del modelo AR es aquel que minimiza la expresión:

$$\mathbf{AIC}[\mathbf{p}] = N \ln \sigma_{\omega}^2 + 2\mathbf{p} \quad (18)$$

donde  $N$  es el número de muestras,  $\sigma_{\omega}^2$  es la varianza del error residual y  $\mathbf{p}$  es el orden del modelo.

Si el número de muestras es grande, estos dos criterios arrojan resultados similares. En la figura 6 se grafican el AIC y FPE para un mismo modelo, en este caso el orden seleccionado es seis.

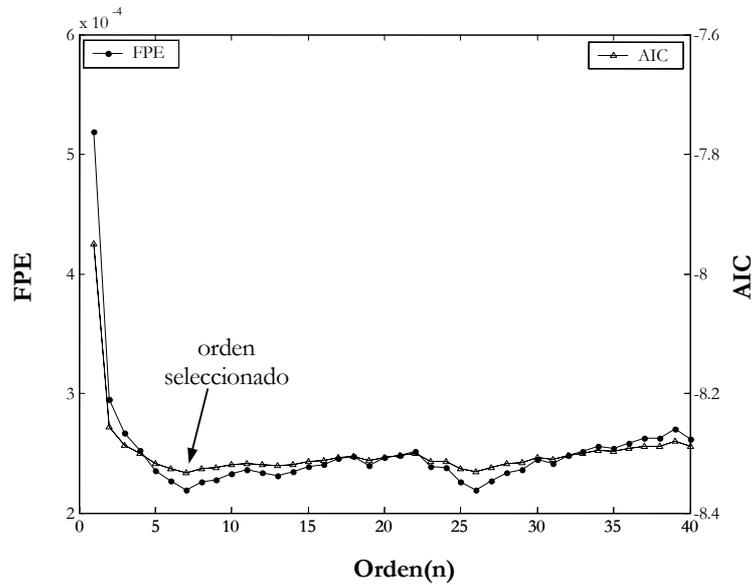


Fig. 6 Gráficas de AIC y FPE para selección del orden del modelo AR

Evaluar la bondad del ajuste del modelo AR, consiste en calcular el error residual dado el orden  $p$  y graficar la función de autocorrelación, y comprobar el comportamiento del error como ruido blanco.

### 3.2.3 MODELO AUTORREGRESIVO MULTICANAL

Un proceso aleatorio multicanal  $\mathbf{U}$  está compuesto de varios procesos aleatorios que suceden simultáneamente y que pueden estar o no correlacionados. La siguiente expresión representa a la muestra  $\mathbf{n}$  del proceso aleatorio multicanal:

$$\mathbf{u}[\mathbf{n}] = \begin{bmatrix} \mathbf{u}_1(\mathbf{n}) \\ \mathbf{u}_2(\mathbf{n}) \\ \vdots \\ \mathbf{u}_d(\mathbf{n}) \end{bmatrix} \quad (19)$$

donde,  $\mathbf{u}[\mathbf{n}]$  es un vector compuesto por un conjunto de muestras que se definen como  $\mathbf{u}_1(\mathbf{n})$  hasta  $\mathbf{u}_d(\mathbf{n})$ , que representan una muestra de  $d$  diferentes canales de información pertenecientes a un mismo proceso aleatorio.

El modelo AR multicanal para el vector  $\mathbf{u}[\mathbf{n}]$  está descrito por la ecuación de diferencias:

$$\mathbf{u}[\mathbf{n}] = -\mathbf{A}_1 \mathbf{u}[\mathbf{n}-1] - \mathbf{A}_2 \mathbf{u}[\mathbf{n}-2] - \dots - \mathbf{A}_p \mathbf{u}[\mathbf{n}-p] + \mathbf{w}[\mathbf{n}] \quad (20)$$

en donde,  $\mathbf{A}_i$  son las matrices de coeficientes y  $\mathbf{w}[\mathbf{n}]$  es un proceso multicanal de ruido blanco con función de correlación:

$$\mathbf{E}\{\mathbf{w}[\mathbf{n}]\mathbf{w}^{*T}[\mathbf{n}-\mathbf{t}]\} = \sum_w \delta[\mathbf{t}] \quad (21)$$

La matriz de correlación para un proceso aleatorio multicanal es una matriz de matrices. Aunque el proceso aleatorio multicanal de ruido blanco es no correlacionado con respecto a  $\mathbf{t}$  (espacio entre muestras), puede estar correlacionado entre canales.

Los parámetros del modelo son encontrados de las ecuaciones normales de predicción lineal, las cuales tienen la forma:

$$\begin{bmatrix} \mathbf{R}_u[0] & \mathbf{R}_u[1] & \dots & \mathbf{R}_u[p] \\ \mathbf{R}_u[-1] & \mathbf{R}_u[0] & \dots & \mathbf{R}_u[p-1] \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{R}_u[-p] & \mathbf{R}_u[-p+1] & \dots & \mathbf{R}_u[0] \end{bmatrix} \begin{bmatrix} \mathbf{1} \\ \mathbf{A}_1^{*T} \\ \vdots \\ \mathbf{A}_p^{*T} \end{bmatrix} = \begin{bmatrix} \sum_w \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix} \quad (22)$$

La matriz de correlación que aparece en esta ecuación es Toeplitz, donde las matrices  $\mathbf{R}[\cdot]$  representan la función de autocorrelación multicanal del proceso aleatorio  $\mathbf{u}[\mathbf{n}]$ .

$$\mathbf{R}_u[\mathbf{t}] = \mathbf{E}\{\mathbf{u}[\mathbf{n}]\mathbf{u}^{*T}[\mathbf{n}-\mathbf{t}]\} \quad (23)$$

La función de autocorrelación puede ser estimada de la siguiente forma:

$$\hat{\mathbf{R}}_u[\mathbf{t}] = \frac{1}{N} \sum_{n=0}^{N-1-\mathbf{t}} \mathbf{u}[\mathbf{n}+\mathbf{t}]\mathbf{u}^{*T}[\mathbf{n}]; \quad \mathbf{0} \leq \mathbf{t} < N \quad (24)$$

Para encontrar la solución de las ecuaciones normales puede utilizarse la recursión multicanal de Levinson-Wiggins-Robinson (LWR) o una forma multicanal del algoritmo de Burg conocido como Nuttall-Strand.

En la siguiente ecuación se observa que la muestra  $\mathbf{n}$  del proceso AR multicanal  $\mathbf{u}[\mathbf{n}]$ , no es un valor escalar, sino es un vector. El número de matrices de coeficientes  $\mathbf{A}$  es igual al orden del modelo AR multicanal. Por lo tanto existen  $p$  matrices  $\mathbf{A}$  de tamaño  $\mathbf{d} \times \mathbf{d}$ .

$$\begin{aligned}
\mathbf{u}[n] = \begin{bmatrix} \mathbf{u}_1(n) \\ \mathbf{u}_2(n) \\ \vdots \\ \mathbf{u}_d(n) \end{bmatrix} = & - \begin{bmatrix} \mathbf{a}_{11} & \mathbf{a}_{21} & \cdots & \mathbf{a}_{d1} \\ \mathbf{a}_{12} & \mathbf{a}_{22} & \cdots & \mathbf{a}_{d2} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{a}_{1d} & \mathbf{a}_{2d} & \cdots & \mathbf{a}_{dd} \end{bmatrix}_1 \begin{bmatrix} \mathbf{u}_1(n-1) \\ \mathbf{u}_2(n-1) \\ \vdots \\ \mathbf{u}_d(n-1) \end{bmatrix} - \begin{bmatrix} \mathbf{a}_{11} & \mathbf{a}_{21} & \cdots & \mathbf{a}_{d1} \\ \mathbf{a}_{12} & \mathbf{a}_{22} & \cdots & \mathbf{a}_{d2} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{a}_{1d} & \mathbf{a}_{2d} & \cdots & \mathbf{a}_{dd} \end{bmatrix}_2 \begin{bmatrix} \mathbf{u}_1(n-2) \\ \mathbf{u}_2(n-2) \\ \vdots \\ \mathbf{u}_d(n-2) \end{bmatrix} - \\
\cdots - & \begin{bmatrix} \mathbf{a}_{11} & \mathbf{a}_{21} & \cdots & \mathbf{a}_{d1} \\ \mathbf{a}_{12} & \mathbf{a}_{22} & \cdots & \mathbf{a}_{d2} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{a}_{1d} & \mathbf{a}_{2d} & \cdots & \mathbf{a}_{dd} \end{bmatrix}_p \begin{bmatrix} \mathbf{u}_1(n-p) \\ \mathbf{u}_2(n-p) \\ \vdots \\ \mathbf{u}_d(n-p) \end{bmatrix} + \begin{bmatrix} \mathbf{w}_1(n) \\ \mathbf{w}_2(n) \\ \vdots \\ \mathbf{w}_d(n) \end{bmatrix}
\end{aligned} \tag{25}$$

en forma simplificada, el modelo AR multicanal se representa por la expresión:

$$\hat{\mathbf{u}}[n] = -\sum_{i=1}^p \mathbf{A}[i] \mathbf{u}[n-i] \tag{26}$$

La matriz de correlación  $\tilde{\mathbf{R}}$  está compuesta de matrices de tamaño  $\mathbf{p} \times \mathbf{d}$

$$\tilde{\mathbf{R}} = \begin{bmatrix} \mathbf{R}(0) & \mathbf{R}(1) & \cdots & \mathbf{R}(p-1) \\ \mathbf{R}(1) & \mathbf{R}(0) & \cdots & \mathbf{R}(p-2) \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{R}(p-1) & \mathbf{R}(p-2) & \cdots & \mathbf{R}(0) \end{bmatrix}_{(\mathbf{p} \times \mathbf{d}) \times (\mathbf{p} \times \mathbf{d})} \tag{27}$$

La muestra al tiempo  $\mathbf{n}$  del modelo AR multicanal  $\mathbf{u}_1(\mathbf{n})$  correspondiente al canal  $\mathbf{1}$  se describe de la forma:

$$\begin{aligned}
\mathbf{u}_1(\mathbf{n}) = & \overset{*1}{\mathbf{a}_{11}} \mathbf{u}_1(\mathbf{n}-1) + \overset{*1}{\mathbf{a}_{12}} \mathbf{u}_2(\mathbf{n}-1) + \cdots + \overset{*1}{\mathbf{a}_{1d}} \mathbf{u}_d(\mathbf{n}-1) + \\
& \overset{*2}{\mathbf{a}_{11}} \mathbf{u}_1(\mathbf{n}-2) + \overset{*2}{\mathbf{a}_{12}} \mathbf{u}_2(\mathbf{n}-2) + \cdots + \overset{*2}{\mathbf{a}_{1d}} \mathbf{u}_d(\mathbf{n}-2) + \\
& \vdots \\
& \overset{*p}{\mathbf{a}_{11}} \mathbf{u}_1(\mathbf{n}-p) + \overset{*p}{\mathbf{a}_{12}} \mathbf{u}_2(\mathbf{n}-p) + \cdots + \overset{*p}{\mathbf{a}_{1d}} \mathbf{u}_d(\mathbf{n}-p)
\end{aligned} \tag{28}$$

### 3.2.3.1 REDUCCIÓN DE LA DIMENSIONALIDAD DEL ESPACIO DE CARACTERÍSTICAS

#### 3.2.3.1.1 Eigenvalores Y Eigenvectores

En ciertas ocasiones es deseable transformar un vector aleatorio  $\mathbf{x}$  en otro vector aleatorio  $\mathbf{x}'$ , cuyos componentes tengan la propiedad de ortogonalidad, es decir que:

$$\mathbf{E}\{\mathbf{x}'_k \mathbf{x}'_l\} = \mathbf{0} \quad \mathbf{k} \neq \mathbf{l} \quad \mathbf{E}\{\ } = \text{Valor esperado} \quad (29)$$

Dada una matriz  $\mathbf{M}$  de tamaño  $\mathbf{n} \times \mathbf{n}$ , por ejemplo la matriz de correlación  $\mathbf{R}_x$ , o la matriz de covarianza  $\mathbf{C}_x$  que cumplen con la propiedad de ser una matriz diagonal y si además la matriz  $\mathbf{M}$  es hermitiana simétrica, es decir que sea igual a su conjugada compleja, es posible encontrar  $\mathbf{n}$  eigenvectores ortonormales  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_N$  y el correspondiente conjunto de eigenvalores  $\lambda_1, \lambda_2, \dots, \lambda_N$  los cuales son todos escalares y reales [13], [14] que satisfacen la ecuación:

$$\mathbf{M}\mathbf{e} = \lambda\mathbf{e} \quad (30)$$

donde,  $\mathbf{M}$  es considerada una transformación lineal, que mapea el eigenvector  $\mathbf{e}$  en una versión escalada de sí mismo. Los eigenvectores cambian en magnitud pero no de dirección. Si la matriz es diagonal, entonces los eigenvectores son paralelos a los ejes coordenados.

Los eigenvectores de la matriz de correlación  $\mathbf{R}_x$  del vector aleatorio  $\mathbf{x}$  pueden ser utilizados para construir una transformación unitaria que se interpreta como una rotación del sistema de coordenadas en el cual el vector  $\mathbf{x}$  está representado.

#### 3.2.3.1.2 Descomposición en Valores Singulares

La descomposición en valores singulares (SVD) es una herramienta común para diagonalizar matrices cuando el sistema es singular, o numéricamente muy próximo a ser singular. En estos

casos, se puede hacer una resolución por mínimos cuadrados del sistema, o una resolución del sistema por SVD.

También cuando se requieren calcular los eigenvalores y eigenvectores con mayor precisión numérica, se desarrolla una SVD de la matriz de datos que establece que cualquier matriz  $\mathbf{X}$  de tamaño  $\mathbf{k} \times \mathbf{n}$  sea singular o no, puede ser descompuesta y escrita como el producto de matrices:

$${}^1 \mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{*T} \quad (31)$$

donde  $\mathbf{U}$  es la matriz unitaria de tamaño  $\mathbf{k} \times \mathbf{k}$  de los vectores singulares izquierdos.  $\mathbf{V}$  es la matriz unitaria de tamaño  $\mathbf{n} \times \mathbf{n}$  de los vectores singulares derechos y  $\mathbf{\Sigma}$  es la matriz de tamaño  $\mathbf{k} \times \mathbf{n}$  de valores singulares reales no negativos.  $\mathbf{U}$  y  $\mathbf{V}$  son matrices ortogonales, por lo que sus inversas son iguales a sus transpuestas y  $\mathbf{\Sigma}$  es una matriz diagonal.

En la ecuación anterior se asume que  $\mathbf{k} \geq \mathbf{n}$ , que  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_k$  y que algunos valores  $\sigma_k$  pueden ser cero. En general, existen  $r$  valores singulares, donde  $r$  es el rango de  $\mathbf{X}$  (el número de columnas independientes).

La descomposición de  $\mathbf{X}^{*T}\mathbf{X}$  corresponde a una descomposición única:

$$\mathbf{X}^{*T}\mathbf{X} = \mathbf{V} \begin{bmatrix} \sigma_1^2 & \mathbf{0} & \vdots & \mathbf{0} \\ \mathbf{0} & \sigma_2^2 & \vdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \vdots & \sigma_n^2 \end{bmatrix} \mathbf{V}^{*T}, \quad (32)$$

en donde la matriz  $\mathbf{V}$ , es la matriz de eigenvectores, y los  $\sigma_k^2$  son los eigenvalores. Cuando la matriz de correlación es estimada mediante  $\hat{\mathbf{R}}_x = \frac{1}{\mathbf{k}} \mathbf{X}^{*T}\mathbf{X}$  los eigenvalores y eigenvectores están dados por:

---

<sup>1</sup> \*T Inversa transpuesta de una matriz

$$\begin{aligned} \mathbf{e}_k &= \mathbf{V}_k & \mathbf{k} &= 1, 2, \dots, \mathbf{n} \\ \lambda_k &= \frac{1}{\mathbf{k}} \sigma_k^2 & \mathbf{k} &= 1, 2, \dots, \mathbf{n} \end{aligned} \quad (33)$$

Los vectores singulares derechos  $\mathbf{V}_k$  son los eigenvectores  $\mathbf{e}_k$  y los  $\sigma_k^2$  son los eigenvalores  $\lambda_k$  de la matriz  $\mathbf{X}\mathbf{X}^{*T}$  de tamaño  $\mathbf{k} \times \mathbf{k}$ , entonces existen  $\mathbf{k}$  eigenvalores. Si  $\mathbf{k} > \mathbf{n}$ , entonces existen  $\mathbf{n}$  valores singulares. Por el contrario,  $\mathbf{k} - \mathbf{n}$  eigenvalores son cero [4].

### 3.2.3.1.3 Análisis de componentes principales (PCA)

Hotelling desarrolló una transformación lineal que remueve la correlación entre los elementos de un vector aleatorio y lo denominó “el método de componentes principales”. Posteriormente Karhunen y Loéve desarrollaron una transformación análoga para señales continuas [14].

El análisis de componentes principales (PCA) es un método no supervisado para encontrar las características importantes de un conjunto de datos. El PCA consiste en representar los datos n-dimensionales, en un espacio de características más pequeño. El método de componentes principales o transformada Karhunen-Loéve conceptualmente consiste en:

- a) Primero, se determina el valor medio del vector n-dimensional y su matriz de covarianza de tamaño  $\mathbf{n} \times \mathbf{n}$ .
- b) Posteriormente, los eigenvectores y eigenvalores son calculados y ordenados en forma decreciente de acuerdo a los eigenvalores. El eigenvector  $\mathbf{e}_1$  con su respectivo eigenvalor  $\lambda_1$ , el eigenvector  $\mathbf{e}_2$  con su respectivo eigenvalor  $\lambda_2$  y así sucesivamente. Después, se seleccionan los  $\mathbf{m}$  eigenvectores más grandes observando un espectro de los eigenvectores y se forma una matriz  $\mathbf{A}$  de tamaño  $\mathbf{m} \times \mathbf{m}$  cuyas columnas corresponden a los eigenvectores.

El vector transformado que minimiza el criterio del error cuadrático  $e = |\mathbf{x} - \mathbf{x}'|^2$  es:

$$\mathbf{x}' = \mathbf{A}^t(\mathbf{x} - \boldsymbol{\mu}) \quad (34)$$

Desde la perspectiva del reconocimiento de patrones, el valor práctico del PCA es que proporciona una técnica efectiva de reducción de la dimensionalidad. Podemos reducir el número de características necesarias para una representación de datos efectiva descartando aquellas combinaciones lineales que tienen varianzas pequeñas y reteniendo solo aquellos términos que tienen varianzas grandes [13].

Sean  $\lambda_0, \lambda_1, \dots, \lambda_{m-1}$  los eigenvalores más grandes de la matriz de correlación  $\mathbf{R}$ , podemos aproximar el vector de datos  $\mathbf{x}$  truncando su expansión después de  $m$  términos, convirtiéndose en  $\mathbf{x}'$  como sigue:

$$\mathbf{x} = \mathbf{U}\mathbf{a} = \sum_{j=0}^{p-1} \mathbf{a}_j \mathbf{U}_j \quad \text{se convierte en} \quad \mathbf{x}' = \sum_{j=0}^{m-1} \mathbf{a}_j \mathbf{U}_j, \quad m < p \quad (35)$$

que representa una transformación de coordenadas, en donde  $\mathbf{U}_j$  representa una base del espacio de datos. El punto  $\mathbf{x}$  en el espacio de datos es transformado en el correspondiente punto  $\mathbf{a}$  en el espacio de características.

La varianza total de los componentes del vector aleatorio  $\mathbf{x}$  es:

$$\sum_{j=0}^{p-1} \sigma_j^2 = \sum_{j=0}^{p-1} \lambda_j \quad (36)$$

donde  $\sigma_j^2$  es la varianza del  $j$ -ésimo componente principal  $\mathbf{a}_j$ .

La varianza total de los  $m$  elementos del vector de aproximación  $\mathbf{x}'$ :

$$\sum_{j=0}^{m-1} \sigma_j^2 = \sum_{j=0}^{m-1} \lambda_j \quad (37)$$

La varianza total de los  $\mathbf{p} - \mathbf{m}$  elementos en el vector de error de aproximación  $\mathbf{e} = (\mathbf{x} - \mathbf{x}')$  es:

$$\sum_{j=m}^{p-1} \sigma_j^2 = \sum_{j=m}^{p-1} \lambda_j \quad (38)$$

Los eigenvalores  $\lambda_m, \dots, \lambda_{p-1}$  son los eigenvalores  $(\mathbf{p} - \mathbf{m})$  más pequeños de la matriz de correlación  $\mathbf{R}$ , que corresponden a los términos descartados.

En resumen, el procedimiento de la descomposición en componentes principales consiste en calcular los eigenvalores y eigenvectores de la matriz de correlación del vector de datos de entrada. Posteriormente, en proyectar los datos de entrada ortogonalmente sobre el subespacio abarcado por los eigenvectores pertenecientes a los eigenvalores más grandes. Este método de representación de datos es comúnmente referido como descomposición del subespacio [13], [15].

### 3.3 REDES NEURONALES ARTIFICIALES

Una red neuronal es un procesador de información distribuido paralelamente, que posee la capacidad natural para almacenar conocimientos basados en la experiencia, haciéndolos disponibles para su uso.

En una red neuronal artificial, los elementos de procesamiento de información, llamados perceptrones, nodos o neuronas, se utilizan para simular una neurona biológica y esas unidades son interconectadas para formar redes, imitando la interacción neuronal en un sistema nervioso biológico. El conocimiento es adquirido por la red a través de un proceso de aprendizaje. Las conexiones interneuronales conocidas como pesos sinápticos, son utilizadas para almacenar el conocimiento.

Cada neurona recibe estimulación de las entradas del sistema o de otras neuronas. En la figura 7 se muestra el modelo de una neurona. Cada entrada es multiplicada por un factor de ponderación (pesos)  $\mathbf{w}$  y el producto entrada-peso son sumados. Este valor pasa a través de una función de activación  $\phi$  no lineal tal como una sigmoidea o tangente hiperbólica [16].

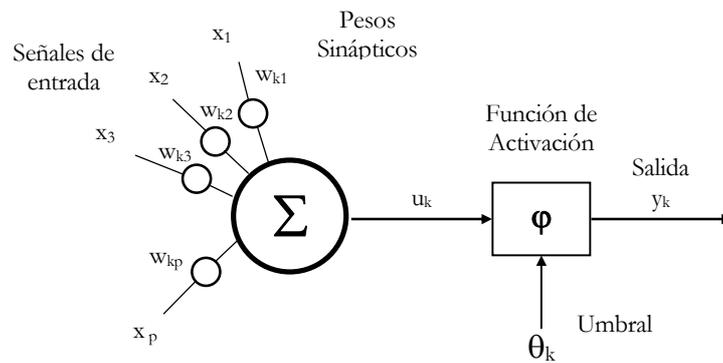


Fig. 7 Modelo no lineal de una neurona

Matemáticamente se puede describir a una neurona mediante las siguientes ecuaciones:

$$\mathbf{u}_k = \sum_{j=1}^p \mathbf{w}_{kj} \mathbf{x}_j \quad (39)$$

$$\mathbf{y}_k = \phi(\mathbf{u}_k - \theta_k), \quad (40)$$

en donde,  $\mathbf{x}_p$  son las entradas y  $\mathbf{w}_{kp}$  son los pesos asociados con las entradas para la neurona  $\mathbf{k}$ . Además,  $\mathbf{u}_k$  es la suma del producto entrada-peso y  $\mathbf{y}_k$  es la salida de la neurona  $\mathbf{k}$ . Frecuentemente una de las entradas es un umbral  $\theta_k$  que se encarga de disminuir el valor de la entrada de la función de activación o un término de sesgo que es lo contrario del umbral y tiende a incrementar el valor de entrada de la función de activación, generalmente tiene el valor de uno, y no corresponde a una entrada actual.

Las neuronas son combinadas entre sí para formar capas, que posteriormente llegan a formar redes. Típicamente la red consiste de un conjunto de unidades sensoriales que constituyen la capa de entrada, una o más capas ocultas y una capa de salida. Este tipo de redes son comúnmente llamados perceptrones multicapa. En la figura 8 se representa gráficamente la arquitectura de un perceptrón multicapa [16].

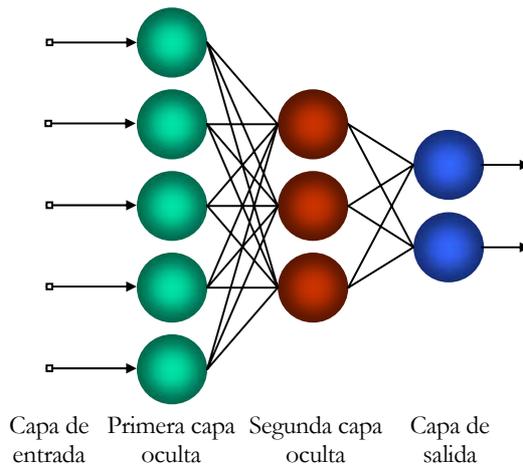
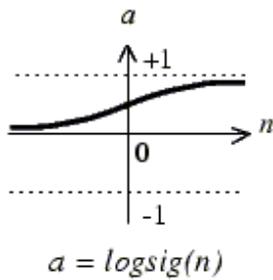
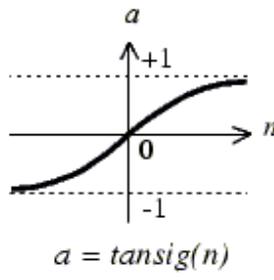


Fig. 8 Arquitectura de un perceptrón multicapa con dos capas ocultas

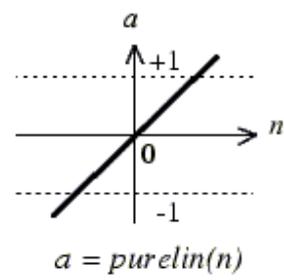
Las funciones de activación comúnmente usadas son:



$$\text{logsig}(n) = \frac{1}{1 + e^{-n}}$$



$$\text{tansig}(n) = \frac{2}{(1 + e^{-2n}) - 1}$$



$$\text{purelin}(n) = n$$

En la figura 9 se muestra la arquitectura de una red utilizada como clasificador y el valor de la salida de la capa oculta y la capa de salida.

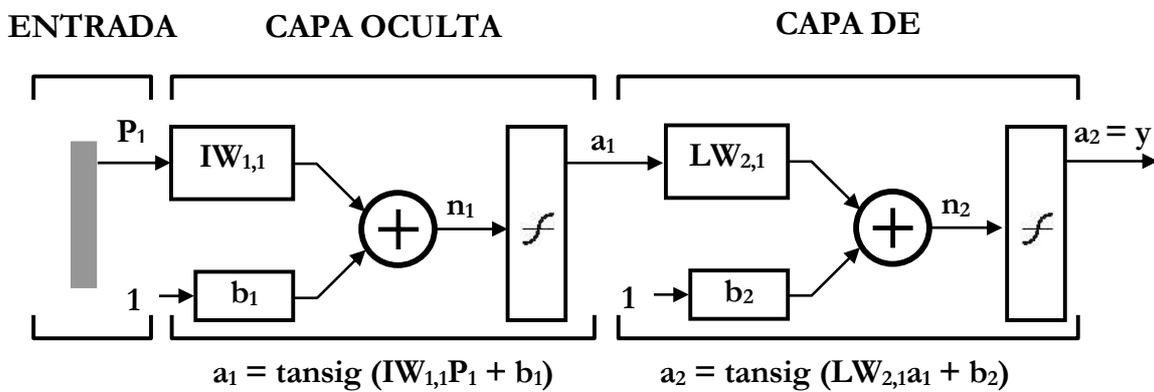


Fig. 9 Arquitectura de la RNA utilizada como clasificador.

Donde  $\mathbf{P}$  representa al vector de entrada con un número determinado de parámetros,  $\mathbf{IW}$  y  $\mathbf{LW}$  son las matrices de pesos que ponderan a un valor de entrada,  $\mathbf{b}$  es el sesgo,  $\mathbf{a1}$  y  $\mathbf{a2}$  son las salidas de la capa oculta y capa de salida respectivamente. Se ha demostrado que una capa puede clasificar patrones que son linealmente separables, requiriéndose más de una capa para resolver problemas de clasificación no lineales. La red neuronal más utilizada es la red multicapa de retropropagación.

Los perceptrones multicapa han sido satisfactoriamente aplicados para resolver diversos y difíciles problemas de clasificación, entrenándolos de forma supervisada con el *algoritmo del error de retropropagación* que se describe en la siguiente sección y está basado en la regla de aprendizaje de corrección del error [15].

### 3.3.1 ALGORITMO DE RETROPROPAGACIÓN

Básicamente el algoritmo de retropropagación consiste de dos pasos a través de diferentes capas de la red [15], [16]:

- a) Un paso hacia delante, en donde un patrón o vector de entrada es aplicado a los nodos sensoriales de la red y su efecto se propaga a través de la red capa por capa, mientras los pesos sinápticos son actualizados. Finalmente, un conjunto de salida es producido como la respuesta actual de la red.
- b) Un paso hacia atrás, donde los pesos sinápticos son ajustados de acuerdo con la regla de corrección del error. La respuesta actual de la red es sustraída de una respuesta deseada para producir una señal de error. Este error se propaga hacia atrás de la red.

El algoritmo de retropropagación está basado en mínimos cuadrados medios (Least Mean Square, LMS) también conocido como algoritmo delta (delta rule). Si  $\mathbf{d}$  es la respuesta deseada y  $\mathbf{y}$  es la salida actual, entonces el error  $\mathbf{e}$  en la salida de la neurona  $\mathbf{j}$  en la iteración  $\mathbf{n}$ , está definido por:

$$\mathbf{e}_j(\mathbf{n}) = \mathbf{d}_j(\mathbf{n}) - \mathbf{y}_j(\mathbf{n}) \quad (41)$$

y el error cuadrático se expresa como:

$$\mathbf{J} = \frac{1}{2} \mathbf{E}[e^2] \quad (42)$$

donde  $\mathbf{E}$  indica el valor esperado o esperanza matemática. EL objetivo del LMS y el algoritmo de retropropagación es encontrar el valor mínimo para  $\mathbf{J}$  en un sistema. Para una neurona, minimizar el error corresponde a descender sobre el borde de una parábola. Para una red de neuronas significa encontrar el valle más profundo en un mapa topográfico o superficie. En problemas complejos, el error encontrado puede corresponder a un mínimo local. Muchas estrategias de redes neuronales han sido desarrolladas para evitar los mínimos locales y continuar buscando por el mínimo global en el espacio total del problema.

La suma instantánea de errores cuadráticos es el error calculado después de aplicar un patrón de entrenamiento a la red y es obtenido sumando el error cuadrático medio sobre todas las neuronas en la capa de salida, es decir:

$$\boldsymbol{\varepsilon}(\mathbf{n}) = \frac{1}{2} \sum_{j \in \mathbf{C}} e_j^2(\mathbf{n}) \quad (43)$$

donde la iteración  $\mathbf{n}$  corresponde al  $\mathbf{n}$ -ésimo patrón de entrenamiento presentado a la red y el subíndice  $\mathbf{j}$  indica una neurona específica de salida. En consecuencia, el error es sumado a través del conjunto  $\mathbf{C}$  de todas las neuronas en la capa de salida de la red neuronal. Considerando  $\mathbf{n}$  iteraciones, el promedio del error cuadrático  $\boldsymbol{\varepsilon}_{\text{prom}}$  se calcula mediante la suma instantánea de errores cuadráticos sobre todas las  $\mathbf{n}$  y normalizando con respecto al conjunto de entrenamiento de tamaño  $\mathbf{N}$ .

$$\boldsymbol{\varepsilon}_{\text{prom}}(\mathbf{n}) = \frac{1}{\mathbf{N}} \sum_{n=1}^{\mathbf{N}} \boldsymbol{\varepsilon}(\mathbf{n}) \quad (44)$$

El proceso de aprendizaje tiene por objetivo minimizar la suma promedio del error cuadrático, ajustando los pesos de cada neurona. La entrada  $\mathbf{v}$  a la función no lineal  $\boldsymbol{\varphi}$ , y la salida  $\mathbf{y}_j$  de la neurona  $\mathbf{j}$  se puede expresar:

$$\mathbf{v}_j(\mathbf{n}) = \sum_{i=0}^{\mathbf{P}} \mathbf{w}_{ij}(\mathbf{n}) \mathbf{y}_i(\mathbf{n}) \quad (45)$$

$$\mathbf{y}_j(\mathbf{n}) = \boldsymbol{\varphi}_j(\mathbf{v}_j(\mathbf{n})) \quad (46)$$

donde  $\mathbf{P}$  es el número total de entradas (excluyendo el umbral) aplicado a la neurona  $\mathbf{j}$ . El peso sináptico  $w_{j_0}$  corresponde a la entrada  $y_0 = -1$ , equivalente al umbral  $\theta_j$ . Para minimizar el error, cada peso es ajustado de una forma proporcional al gradiente  $\partial \epsilon(\mathbf{n}) / \partial w_{ji}(\mathbf{n})$ . El gradiente puede escribirse:

$$\frac{\partial \epsilon(\mathbf{n})}{\partial w_{ji}(\mathbf{n})} = \frac{\partial \epsilon(\mathbf{n})}{\partial e_j(\mathbf{n})} \frac{\partial e_j(\mathbf{n})}{\partial y_j(\mathbf{n})} \frac{\partial y_j(\mathbf{n})}{\partial v_j(\mathbf{n})} \frac{\partial v_j(\mathbf{n})}{\partial w_{ji}(\mathbf{n})} \quad (47)$$

Aplicando la regla de la cadena, sustituyendo términos y multiplicando el gradiente por una constante  $\eta$ , llamada razón de aprendizaje, la cual controla la velocidad de descenso del gradiente, la corrección aplicada a los pesos se expresa:

$$\Delta w_{ji}(\mathbf{n}) = \eta \delta_j(\mathbf{n}) y_i(\mathbf{n}) \quad (48)$$

$$\text{donde } \delta_j = - \frac{\partial \epsilon(\mathbf{n})}{\partial e_j(\mathbf{n})} \frac{\partial e_j(\mathbf{n})}{\partial y_j(\mathbf{n})} \frac{\partial y_j(\mathbf{n})}{\partial v_j(\mathbf{n})} = e_j(\mathbf{n}) \phi_j'(v_j(\mathbf{n})) \quad (49)$$

Los pesos son ajustados basados en el parámetro de razón de aprendizaje, el gradiente local y la señal de entrada a cada neurona. Existen tres opciones para decidir cuando detener el entrenamiento.

1. El error predeterminado es alcanzado
2. El ajuste sucesivo de los pesos es mínimo
3. La red clasifica apropiadamente todos los vectores de entrenamiento

Después del entrenamiento, la red se valida con vectores que nunca ha visto, pero que son similares a los vectores de entrenamiento por pertenecer a una misma población de datos.

Teóricamente, una red con una capa oculta puede resolver cualquier problema de clasificación no lineal, si posee las neuronas necesarias. Sin embargo, agregando capas y neuronas ocultas puede hacer que el tiempo de entrenamiento disminuya y realzar el desempeño de la red neuronal en la práctica. La teoría sugiere que solamente se requieren un par de neuronas en la capa oculta para una clasificación no lineal, sin embargo, en la práctica se requieren más neuronas en la capa oculta de la red neuronal. Por lo tanto, la configuración y arquitectura de

una red es de naturaleza experimental. Primero se propone una configuración inicial y posteriormente se realizan cambios hasta obtener el desempeño deseado para una apropiada clasificación de los datos [15].

### **3.3.2 MODOS DE ENTRENAMIENTO**

El aprendizaje es el resultado de presentar muchas veces a la red, un conjunto de vectores de entrenamiento. Una presentación completa del conjunto de entrenamiento durante el proceso de aprendizaje se conoce como una época. El proceso de aprendizaje se mantiene hasta que los pesos sinápticos y los niveles del umbral de la red se estabilizan y el promedio del error cuadrático sobre el conjunto total de entrenamiento converge hacia un valor mínimo.

En la práctica, es recomendable que el orden de presentación de vectores de entrenamiento sea de modo aleatorio de una época a la siguiente.

Existen dos formas de presentar a la red los vectores de entrenamiento.

1. Modo patrón: La actualización de los pesos se desarrolla después de la presentación de cada vector de entrenamiento.
2. Modo grupal: El proceso de actualización de los pesos se lleva a cabo después de la presentación de todos los vectores de entrada, que constituye una época.

Desde un punto de vista operacional, el modo patrón es preferido porque requiere menor almacenamiento local para cada conexión sináptica. Además, dado que los patrones son presentados a la red de manera aleatoria, el uso del entrenamiento patrón por patrón hace menos probable que el algoritmo de retropropagación quede atrapado en un mínimo local durante la búsqueda en el espacio de los pesos. Por otro lado, el uso del modo grupal de entrenamiento proporciona un estimado preciso del vector gradiente, lo que acelera el proceso de entrenamiento [15].

### **3.3.3 CRITERIO PARA DETENER EL PROCESO DE ENTRENAMIENTO**

No existe un criterio bien definido que indique el momento en que debe detenerse la fase de

entrenamiento y el ajuste de los pesos en una RNA. Sin embargo, existen criterios prácticos que sugieren cuando puede detenerse el proceso de entrenamiento [15]. Un método para mejorar la generalización consiste en dividir los datos disponibles en tres subconjuntos. El primer subconjunto es el conjunto de entrenamiento, el cual es usado para calcular el gradiente y actualizar los pesos y términos de sesgo de la red neuronal. El segundo subconjunto es el conjunto de validación. El error alcanzado con el conjunto de validación es monitoreado durante el proceso de entrenamiento, y normalmente decrece durante la fase inicial de entrenamiento al igual que el error con el conjunto de entrenamiento. Sin embargo, cuando la red comienza a sobre-ajustar los datos, el error con el conjunto de validación comenzará a incrementarse. Si esto sucede para un número específico de iteraciones, se sugiere detener el entrenamiento. El conjunto de prueba no es usado durante el entrenamiento, pero es utilizado para comparar diferentes modelos o arquitecturas de red.

En la figura 10 se observa una gráfica de la curva de aprendizaje de la red, con el conjunto de entrenamiento la red se desempeña muy bien, sin embargo al evaluar la red con los conjuntos de prueba y validación el error se incrementa.

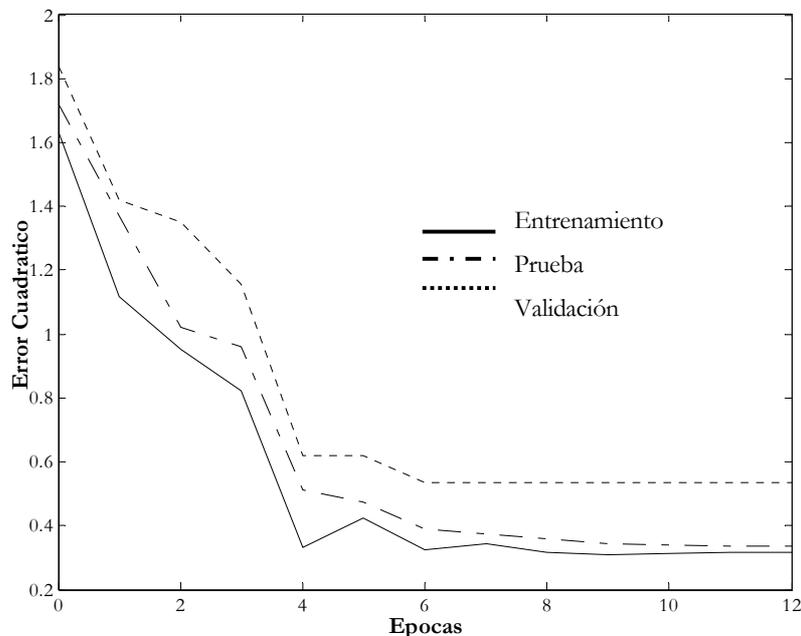


Fig. 10 Curva de aprendizaje de una red con conjuntos de entrenamiento, prueba y validación

### 3.4 REFERENCIAS

- [1] L. R. Waitman, "Representation and classification of breath sounds recorded in an intensive care setting using neural networks." MS thesis. Vanderbilt University, 1998.
- [2] Z. Peyton, J. R. Peebles, "Probability, random variables and random signal principles." McGraw-Hill Series in Electrical and Computer Engineering 3a. edición, 1993.
- [3] J. G. Proakis, "Digital signal Processing." Editorial Prentice Hall, 3a. edición, pp. 11-14, 1996.
- [4] C. W. Therrien, "Discrete random signals and statistical signal processing." Editorial Prentice Hall, 1992.
- [5] B. Sankur, Y. P. Kahya, E. C. Gulert, T. Engels, "Feature extraction and classification of nonstationary signals based on the multiresolution signal decomposition." International Conference on Pattern Recognition, Jerusalem, pp. 592-595, 1994.
- [6] E. N. Bruce, "Biomedical signal processing and signal modeling." Edit. John Wiley & Sons, Inc., pp. 356-360, 2001.
- [7] "Introduction to maximum likelihood estimators and information theoretic methods." wild 7960, Analysis of wildlife populations.
- [8] L. Ljung, "System Identification." Editorial Prentice Hall, 1987.
- [9] H. Akaike, "Power Spectrum Estimation through Autoregression Model Fitting." Annals of the Institute of Statistical Mathematics, vol. 21, pp. 407-419, 1969.
- [10] H. Akaike, "A New Look at the Statistical Model Identification." IEEE Trans. Autom. Control, vol. 19, no. 6, pp. 716-723, 1974.
- [11] S. L. Marple, "Digital Spectral Analysis." Englewood Cliffs, NJ: Prentice Hall, 1987.
- [12] Y. P. Kahya, U. Çini, Ö. Cerid, "Real-time Regional Respiratory Sound Diagnosis Instrument." 25th Annual IEEE-EMBS Conference, Cancun Mex., pp. 3098-3101, 2003.
- [13] R. O. Duda, P. E. Hart, D. G. Stork, "Pattern Classification." John Wiley & sons Inc, New York, 2ª edición, pp. 609-610, 2001.

- [14] K. R. Castleman, "Digital Image Processing." Edit. Prentice Hall, New Jersey, pp. 295-649, 1996.
- [15] S. Theodoridis, K. Konstantinos, "Pattern Recognition." Edit. Academic Press, San Diego. Pp. 215-218, 2003.
- [16] S. Haykin, "Neural networks. A comprehensive foundation." Edit. Mac Millan, pp. 367-369, 1994.

## CAPÍTULO IV

### ESTADO DEL ARTE

#### 4.1 PROBLEMA DE CLASIFICACIÓN DE SONIDOS RESPIRATORIOS

Los sonidos respiratorios han sido ampliamente estudiados como indicadores del estado fundamental de los pulmones [1]. Muchos estudios consideran importantes las características de estertores debido a asma, enfermedades crónicas obstructivas pulmonares (COPD), fallas cardíacas, fibrosis pulmonar y neumonía [2], [3], [4] y [5]. Otros han estudiado patologías como enfisema y sonidos respiratorios disminuidos [6]. La mayoría de las investigaciones realizadas en clasificación de sonidos respiratorios se ha centrado en comparar el sonido de una patología específica versus sonidos respiratorios normales.

Un equipo de investigadores de la universidad de Chicago, N. Asada y colaboradores [7], realizaron un diagnóstico diferencial de enfermedades pulmonares intersticiales. Estudiaron nueve diferentes patologías y caracterizaron un segmento de señal proveniente de los pacientes considerando datos clínicos y extrayendo información de imágenes radiográficas. Los vectores de características formados con estos parámetros fueron clasificados usando una red neuronal y los resultados fueron analizados mediante curvas ROC (Receiver Operating Characteristic). Los hallazgos revelan que los resultados de la red neuronal son comparables con los diagnósticos propuestos por los expertos radiólogos, pero son superiores al diagnóstico de médicos residentes.

Sankur y Kahya han realizado varias investigaciones acerca de sonidos respiratorios, en particular sobre clasificación. En 1994 compararon dos clasificadores, el de vecino más cercano (Nearest Neighbour K-NN) con el método de distancia de Itakura y clasificadores cuadráticos de mínima distancia de Mahalanobis, para clasificar pacientes con enfermedades pulmonares

crónicas obstructivas o enfermedades restrictivas y sujetos normales. Las señales de sonido fueron caracterizadas por un modelo AR de orden 6, obteniendo un porcentaje de clasificaciones correctas de 93.75% con el clasificador K-NN y de 87.5% con el clasificador de mínima distancia de Mahalanobis [8].

Oud y colaboradores [9], presentaron en la década pasada, trabajos sobre clasificación de sonidos respiratorios en pacientes con asma. En una de las investigaciones reportadas, se intentó separar dos o tres clases de etapas de asma. Los sonidos fueron caracterizados mediante información espectral utilizándose redes neuronales supervisadas y no supervisadas. Los resultados de las redes neuronales fueron comparados con aquellos obtenidos con un clasificador KNN. Se obtuvo un desempeño similar entre el clasificador KNN y la red neuronal al separar dos clases. Sin embargo, en la clasificación de tres clases, los resultados con el KNN superan a los de la red neuronal.

En la tesis realizada por Waitman 0, de la Universidad de Vanderbilt, se clasificaron sonidos respiratorios normales y anormales en una unidad de cuidados intensivos (UCI) para monitoreo continuo de los pacientes en un intento por evitar la aparición de infecciones nosocomiales respiratorias que podrían provocarles la muerte. Los sonidos respiratorios fueron adquiridos por medio de dos micrófonos colocados en la parte anterior de la superficie torácica. En este trabajo se realizó un análisis monocanal, caracterizando a las señales de sonido respiratorio por sus componentes en frecuencia a través de la FFT en un rango de 0 a 800 Hz. Waitman encontró que el número óptimo de características fue de 40. También concluyó que segmentar la señal de sonidos en fase inspiratoria y espiratoria, el uso de ventanas (Blackman) contribuyó a mejorar el desempeño del clasificador. Waitman probó diferentes arquitecturas de redes neuronales para la clasificación de las señales de sonidos respiratorio, encontrando que la de mejor desempeño fue la red neuronal de 1 capa oculta compuesta de 14 neuronas, utilizando el algoritmo de entrenamiento Delta-Bar-Delta Extendido. Los resultados muestran un porcentaje promedio de clasificaciones correctas de 73% con las señales de sonido

respiratorio adquiridas en la UCI, y un porcentaje de 91% con archivos de sonidos respiratorio disponibles comercialmente. Ambas poblaciones incluyeron una gran variedad de sonidos anormales (Estertores, sibilancias, etc).

En otros estudios, Kahya y colaboradores [11], desarrollaron una clasificación jerárquica de sonidos respiratorios. El método consistió en dividir los ciclos respiratorios de pacientes en fases, y clasificar cada fase con un perceptrón multicapa de acuerdo a una “fase experta”. La información correspondiente a una fase consistió de la representación paramétrica de varios segmentos de tiempo (60 segmentos), que se obtuvo calculando los parámetros del modelo AR de orden 6 y los coeficientes cepstrales. Se estudiaron pacientes con enfermedades obstructivas (18 sujetos), enfermedades restrictivas (19 sujetos) y normales (20 sujetos). Se aplicó una decisión experta por cada fase y fue combinada con un esquema de fusión de decisión simulando una sesión de consulta. Con este método se encontró que el porcentaje de clasificaciones correctas fue de 60% y utilizando esquemas de consulta se obtuvo 80% de clasificaciones correctas.

Asimismo en otra investigación de Oud et. al [12], se comparó la capacidad de expertos humanos y redes neuronales para clasificar tres diferentes clases de sonidos respiratorios pertenecientes a niños con asma. Se formaron dos conjuntos de vectores de características calculando el espectro de potencia de Fourier para entrenamiento y prueba de una red neuronal no supervisada auto-organizada (self organization map). Se encontraron tres clases de sonidos con un porcentaje de clasificaciones correctas de 95% con el conjunto de entrenamiento y de 43% con el conjunto de prueba. En este estudio se concluyó que las redes neuronales superaron a los expertos humanos en la clasificación de sonidos.

Sezgin et. al. [13], utilizaron una red neuronal para la clasificación de sonidos respiratorios de pacientes con asma y sujetos normales. El estudio consistió en adquirir el sonido a través de un estetoscopio electrónico con una frecuencia de muestreo de 22,050 Hz de 10 pacientes

asmáticos y 10 personas sanas. Posteriormente se aplicó la transformada wavelet a cada ventana de señal mediante el método de Daubechies-2 y se calcularon los valores RMS de los seis coeficientes wavelet para formar el vector de características. El número de vectores de características en el conjunto total fue de 300. En este trabajo se propuso la implementación de una red neuronal tipo GAL (Grow and Learn) debido a que presenta ciertas ventajas sobre las redes neuronales convencionales. Con este esquema de clasificación se obtuvo un porcentaje promedio de 98% de clasificaciones correctas utilizando una red neuronal GAL con 2 nodos y se estimó que el tiempo de entrenamiento de la red fue menor a un segundo.

Un estudio realizado en la Universidad de Québec, encabezado por M. Bahoura, y C. Pelletier [14], propuso un nuevo enfoque de clasificación de sonidos respiratorios basado en análisis cepstral. El sonido fue dividido en segmentos y posteriormente caracterizados mediante un número reducido de coeficientes cepstrales. Estos segmentos son clasificados como sonidos anormales (sibilancias) o como sonidos normales utilizando el método de vector de cuantización (VQ). Este método es probado y comparado con diferentes métodos de extracción de características como transformada Wavelet, FFT y otros. La frecuencia de muestreo que se manejó durante la adquisición de los sonidos fue de 6 Khz. Las secuencias estuvieron formadas por 512 muestras traslapadas 50% utilizando una ventana de Hamming. Cada segmento estuvo caracterizado por 20 coeficientes cepstrales. El porcentaje promedio de clasificaciones correctas para los sonidos normales fue de 75.8% y para sibilancias de 77.5%. Con el método de Transformada Wavelet se obtuvo un porcentaje de 60.3% para sonidos normales y de 58.0% para sibilancias. Los resultados obtenidos con la FFT alcanzaron 90.6% con sonidos normales y 44.8% con sibilancias.

Baydar [15] analizó y clasificó sonidos respiratorios normales y anormales de diferentes patologías. La adquisición de las señales fue a través de dos canales con frecuencia de muestreo de 6 KHz, la duración del registro fue de 13 segundos. La parametrización de las señales de sonido se efectuó mediante el método de coherencia de la señal y se redujo la dimensión del

espacio de características aplicando análisis de componentes principales. Además, se utilizó un clasificador K-NN de distancia euclidiana para separar las clases. Baydar obtuvo los mejores resultados con un segmento de espiración, aplicando el método de Yule-Walker y Burg para estimar el periodo fundamental de las señales, el porcentaje de clasificaciones correctas fue de 71%. Con la fase inspiratoria encontró un porcentaje de 69% con los mismos métodos para estimar el periodo fundamental de la señal.

En trabajos más recientes, Y.P. Kahya y colaboradores [16], grabaron simultáneamente cuatro posiciones del tórax posterior de 18 pacientes y 18 personas sanas. La frecuencia de muestreo fue de 4.8 KHz. El objetivo consistió en clasificar ciclos respiratorios pertenecientes a cada sujeto, cada ciclo dividido en 60 segmentos traslapados 25% y cada segmento estuvo representado por tres diferentes conjuntos de vectores de características formados a través de: a) parametrización del sonido por medio de un modelo autorregresivo de orden 6, b) frecuencias percentiles correspondientes al 25, 50, 75, 90 por ciento y por último, c) componentes principales (PCA) proyectando los eigenvectores en un espacio de tres dimensiones, sobre el vector original de parámetros AR. Para separar los vectores de características entrenaron un clasificador KNN de distancia euclidiana con  $K=5$ , utilizando el método de dejar uno fuera. Los mejores resultados revelan que la sensibilidad del clasificador fue de 72% con los parámetros AR en la posición abajo-izquierda y PCA en la posición arriba-izquierda. Con los tres métodos el clasificador mostró una especificidad del 100% en la posición abajo-derecha y el desempeño del clasificador se considera moderado con 86% utilizando parámetros de PCA con el sensor ubicado en la parte arriba-izquierda.

Es importante mencionar que la metodología de los trabajos anteriores se basa en la parametrización de las señales extrayendo características de canales individuales, y en caso de grabar varios canales, estos se concatenan juntos para entrenar al clasificador.

Hasta ahora, no se habían considerado modelos multicanales aplicados al problema de clasificación de sonidos respiratorios, sin embargo, en otros campos, como el de clasificación

de señales electroencefalográficas, se han hecho estudios basados en el modelo AR monocal y multicanal. Anderson y colaboradores [17], compararon los resultados de entrenar una red neuronal con vectores obtenidos por modelo AR multicanal y monocal, eigenvalores de la matriz de correlación, transformada Karhunen-Loève de los coeficientes AR multicanales. Los vectores de características definidos por estas representaciones fueron clasificados con una red neuronal estándar utilizando el algoritmo de retropropagación. En este trabajo los cuatro métodos arrojaron resultados similares, pero el desempeño de la red con los parámetros obtenidos con el modelo AR multicanal fue ligeramente mejor que los demás, con un porcentaje promedio de clasificación de 91.4%

Recientemente, nuestro equipo de investigación reportó el trabajo que discute un esquema de clasificación de los sonidos respiratorios en sujetos con Neumopatía Intersticial Difusa (NID) sobre la base de vectores de características obtenidos por modelado AR multicanal. Un arreglo de 12 sensores sobre la superficie torácica de sujetos sanos y enfermos permitió adquirir fononeumogramas entre 0.6-1.0 L/seg. Los sonidos fueron adquiridos, mediante un sistema Biopac® MP100 de 16 bits, durante 19 segundos, a una frecuencia de muestreo de 7000 Hz. La discriminación entre patrones normales y anormales se realizó con una red neuronal supervisada utilizando el algoritmo de aprendizaje de retropropagación, basado en la regla de Levenberg-Marquardt. Aunque el porcentaje promedio de clasificaciones correctas fue del 81%, se consideró que era posible realizar depuraciones metodológicas que mejoraran el desempeño del modelo propuesto. La presente tesis aborda el problema de manera similar, con una adquisición multicanal, pero utilizando un arreglo de 25 micrófonos y las señales respiratorias son adquiridas a una frecuencia de muestreo mayor [18], 0.

El problema de clasificación de sonidos respiratorios abarca una extensa área de investigación, tomando en cuenta que existen numerosas patologías respiratorias, cada una con sus variantes y detalles específicos. Por otro lado, dado que las técnicas actuales de diagnóstico no proporcionan suficiente información o se trata de un método invasivo para el paciente, se

advierte la importancia de proporcionar un diagnóstico objetivo basado en características cuantitativas, lo que ha llevado a la búsqueda de técnicas de extracción de características por las técnicas mencionadas, de las cuales, el método más utilizado es la parametrización en la frecuencia por medio del espectro de potencia.

El uso de clasificadores de distancia y redes neuronales artificiales es muy común, pero aún no se han reportado resultados completamente satisfactorios.

En la tabla 1 se resumen cronológicamente los trabajos más importantes revisados durante la presente investigación.

TABLA 1  
Resumen de los trabajos más importantes revisados durante la presente investigación.

Titulo	Autor	Año	Lugar	Objetivo	Método	Clasificador	Resultado
Neural network approach for differential diagnosis of interstitial lung diseases [7].	Asada N, Doi K, MacMahon H, Montner S, Giger ML, Abe C, Wu Y	1990	Radiology Clinical Faculty, The University of Chicago	Realizar un diagnóstico diferencial de enf. Pulm. Intersticiales (9 enfermedades)	Información clínica y radiográfica (imágenes)	RN. Curvas ROC	RN comparable a los radiólogos RN superior a residentes
Comparison of AR based algorithms for respiratory sound classification [8].	Sankur, B. Kahya, Y.P. Güler, E.C. Engin, T.	1994	Universidad Bogazici Bebek, Istanbul, Turkey	Encontrar el orden óptimo del modelo AR	Modelo AR	1.- KNN 2.- Cuadrático	Modelo AR óptimo: orden 6
Classification and recognition of asthmatic breathing sounds [9].	M. Oud	1996	Malsch TechnoValuation, Utrecht, The Netherlands	Separar 2 ó 3 etapas de Asma	Información espectral	RN supervisada y no supervisada	KNN similar a RN en separar 2 clases. Para 3 clases mejores resultados con KNN. Óptimos resultados con Power Spectra on a near-decibel scale
Representation and classification of breath sounds recorded in an intensive care setting using neural networks (thesis) 0.	L. R. Waitman	1998	Vanderbilt University, Nashville, Tennessee	Clasificar sonidos respiratorios normales y anormales grabadas en una U.C.I. 2 canales. (crackles, wheezes, snores, diminishes breath sounds, ventilator turbulence, patients on ventilators, diaphragm breathing, normal breath sounds)	Power spectrum	RN backpropagation connect prior Extended Delta-Bar-Delta	Número óptimo de características igual a 40. Arquitectura de red una capa oculta con 14 neuronas. 73% de clasificaciones correctas
Hierarchical classification of respiratory sounds [11].	Kahya, Y.P. Güler, E.C Sankur, B. Raudys, S.	1998	Bogazici University Bebek, Istanbul, Turkey	Clasificar cada fase de sonido utilizando una "fase experta" Pacientes obstructivos (18), restrictivos (19) y normales (20)	Parámetros AR Cepstral coefficients (obtenidos de los coef. de predicción lineal)	Perceptrón Multicapa	Modelo AR orden 6, 6 coeficientes Cepstrales, 60 segmentos representan una fase. 60% de clasificaciones correctas. 80% utilizando esquemas de consulta.

TABLA 1  
Resumen de los trabajos más importantes revisados durante la presente investigación.

Titulo	Autor	Año	Lugar	Objetivo	Método	Clasificador	Resultado
Classification of asthmatic breath sound by human examiners versus artificial neural networks [12].	S. Rietveld, M. Oud	1999	Malsch TechnoValuation, Utrecht, The Netherlands	Comparar la capacidad de expertos humanos y redes neuronales para clasificar 3 clases de sonidos respiratorios. 50 niños con asma, 10 de control. Grabación 20 seg.	Power Spectrum	RN no supervisada: Self organization Map (SOM)	Se encontraron 3 clases, RN: 95% (train) clasificación correcta, 43% (test) clasificación correcta Las RNs superaron la clasificación humana
Classification of respiratory sounds by using an artificial neural network [13]	M.C. Sezgin, Z. Dokur, T. Ölmez, M. Korürek	2001	Istanbul Technical University, Istanbul, Turkey	Clasificación de sonidos respiratorios en pacientes con asma (10) y sujetos normales (10) Fm= 22050 Hz, un canal	Transformada Wavelet a una ventana de 256 muestras. Método de Daubechies-2	Grow and Learn neural network (GAL)	98% de clasificaciones correctas RN con 2 nodos 300 vectores de características
New Parameters for respiratory sounds classification [14]	M. Bahoura, C. Pelletier	2003	Québec University, Canada	Clasificar sonidos resp. como sibilancias (12) y sonidos normales (12). Comparar con método AR y wavelet . Fm=6KHz. Secuencias de 512 muestras, tralapadas 50%, utilizando ventana de hamming	Frequency Cepstrum Coefficients (MFCC). Cada segmento caracterizado por 20 coeficientes	Vector Quantification	Normales: 75.8% Sibilancias: 77.5%
Analysis and classification of respiratory sounds by signal coherence method [15]	K.S Baydar, A.Ertuzun, Y.P. Kahya	2003	Bogazici University Bebek, Istanbul, Turkey	Clasificación de sujetos normales (9) y pacientes con dif. enfermedades (11). Fm=6KHz, 2 canales, 13 seg. de grabación,	Coherencia de la señal. Reducción de la dimensionalidad mediante PCA	Euclidian Distance Classifier (nearest mean classifier)	Mejores resultados: 71% espiración, método de Yule Walker y Burg para estimar el periodo fundamental 69% inspiración, método de Yule Walker y Burg para estimar el periodo fundamental
Comparison of different feature sets for respiratory sounds classifiers [16]	Y.P. Kahya, E. Bayatli, M. Yeginer, K. Ciftci, G. Kilinc	2003	Bogazici University Bebek, Istanbul, Turkey	4 canales, Fm=4.8KHz, clasificar ciclos pertenecientes a c/sujeto, dividido en 6 segmentos y se calculan 3 dif. vectores de características. 18 normales y 18 pacientes	Modelo AR(6), frecuencias percentiles (25,50,75,90) y componentes principales(proyección del vector AR en 3 dimensiones)	K-NN, Euclidian Distance, k=5 Leave-one-out	Sensitividad: 72%, AR (abajo-izq) y PCA (arriba-izq.) Especificidad: 100% todos los métodos, posición abajo-der. Desempeño del clasificador: 86% PCA, posición arriba-izq.
Clasificación de Información Acústica en Neumopatías Intersticiales Difusas Mediante un Análisis Autorregresivo Multicanal[18], 0	Martínez, H.G., Aljama, C.T., González-Camarena, R., Chi-Lem, G. y Charleston V. S.	2004	UAM Iztapalapa; INER, D.F., México	12 canales, Fm= 7KHz, clasificar vectores de características mediante modelado AR multicanal. Flujo=1.0 L/seg	Modelo AR multicanal, orden 4. Reducción de la dimensión mediante valores singulares. Vectores de tamaño 48. Conjunto total de 960 vectores	RN de retropropagación, algoritmo de aprendizaje basado en Levenberg-Marquardt	81% de clasificaciones correctas Sensitividad de 85% y Especificidad de 88%

## 4.2 REFERENCIAS

- [1] H. Pasterkamp, S. Kraman, G. Wodicka, "Respiratory sounds. Advances beyond the stethoscope." *American Journal of Respiratory and Critical Care Medicine*, 156(3), pp. 974-987, 1997.
- [2] P. Piirila, A. Sovijarvi, T. Kaisla, H. M. Rajala, T. Katila, "Crackles in Patients with Fibrosing Alveolitis, Bronchiectasis, COPD, and Heart Failure." *Chest*, 99(5), pp. 1076-1083, 1991.
- [3] P. Piirila, "Changes in Crackle Characteristics during the Clinical Course of Pneumonia." *Chest*, 102(1), pp. 176-183, 1992.
- [4] T. Kaisla, A. R. A. Sovijarvi, P. Piirila, H. M. Rajala, S. Haltsonen, T. Rosqvist, "Validated Methods for Automatic Detection of Lung Sound Crackles." *Medical and Biological Engineering and Computing*, 29, pp. 517-521, 1991.
- [5] M. Munakata, H. Ukita, I. Doi, Y. Ohtsuka, Y. Masaki, Y. Homma, Y. Kawakami, "Spectral and Waveform Characteristics of Fine and Coarse Crackles." *Thorax*, 46, pp. 651-657, 1991.
- [6] H. J. W. Schreur, P. J. Sterk, J. Vanderschoot, H. C. van Klink, E. van Vollenhoven, J. H. Dijkman, "Lung Sound Intensity in Patients with Emphysema and in Normal Subjects at Standardized Airflows." *Thorax*, 47, pp. 674-679, 1992.
- [7] N. Asada, K. Doi, H. MacMahon, S. Montner, M. L. Giger, C. Abe, Y. Wu, "Neural network approach for differential diagnosis of interstitial lung diseases." *Proc. SPIE* 1233, pp. 45-50, 1990.
- [8] B. Sankur, Y. P. Kahya, E. C. Guler, T. Engin, "Comparison of AR-Based Algorithms for Respiratory Sounds Classification." *Computers in Biology and Medicine*, 24(1), pp. 67-76, 1994.
- [9] M. Oud, "Classification and recognition of asthmatic breathing sounds." *EANN Conf. Proc.*, pp. 414-418, 1996.

- [10] L. R. Waitman, "Representation and classification of breath sounds recorded in an intensive care setting using neural networks." MS thesis. Vanderbilt University, 1998.
- [11] Y. P. Kahya, E. C. Güler, B. Sankur, S. Raudys, "Hierarchical classification of respiratory sounds." Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and biology society, 20:3, pp. 1598-1601, 1998.
- [12] S. Rietveld and M. Oud, "Classification of asthmatic breath sounds spectra by human examiners versus artificial neural networks." *Comp. Biomed. Research* 32, pp. 440-448, 1999.
- [13] M. C. Sezgin, Z. Dokur, T. Ölmez, M. Korürek, "Classification of respiratory sounds by using an artificial neural network." Proceedings of the Annual EMBS International Conference, Istanbul, Turkey, pp 697-699, 2001.
- [14] M. Bahoura, C. Pelletier, "New Parameters for respiratory sounds classification." Canadian Conference on Electrical and Computer Engineering, Montreal, Canada, vol. 3, pp 1457-1460, 2003.
- [15] K. S. Baydar, A. Ertuzun, Y. P. Kahya, "Analysis and classification of respiratory sounds by signal coherence method." Proceedings of EMBS, IEEE, Cancún, Mexico. pp. 2950-2953, 2003.
- [16] Y. P. Kahya, E. Bayatli, M. Yeginer, K. Ciftci, G. Kilinc, "Comparison of different feature sets for respiratory sound classifier." Proceedings of EMBS, IEEE, Cancún, Mexico. pp. 2853-2856, 2003.
- [17] C. W. Anderson, E. A. Stolz, S. Shamsunder, "Multivariate Autoregressive models for classification of spontaneous electroencephalographic signals during mental tasks." *IEEE Transactions on Biomedical Engineering*, 45 (3), pp. 277–286, 1998.
- [18] G. Martinez, T. Aljama, R. González, G. Chi-Lem, S. Charleston, "A multichannel autoregressive model applied to acoustic information in diffuse interstitial pneumonia." Proc. 29th Annual Conf. Int. Lung Sounds Assoc., Glasgow, Scotland, September 2004.

- [19] G. Martínez, T. Aljama, R. González, G. Chi-Lem, S. Charleston, “Clasificación de información acústica en Neumopatías intersticiales difusas mediante un análisis autorregresivo multicanal.” Jornadas del Posgrado Divisional en Ciencias Básicas e Ingeniería, Universidad Autónoma Metropolitana-Iztapalapa, México, pp. 135-138, 2004.

# CAPÍTULO V

## DESCRIPCIÓN DE LA METODOLOGÍA

### 5.1 METODOLOGÍA GENERAL

La metodología general del presente trabajo incluye el diseño del estudio, los criterios para la selección y entrenamiento de sujetos, el protocolo de adquisición y almacenamiento de datos, procesamiento de las señales adquiridas, clasificación y validación. En el diagrama a bloques de la figura 1, se resume la metodología propuesta.

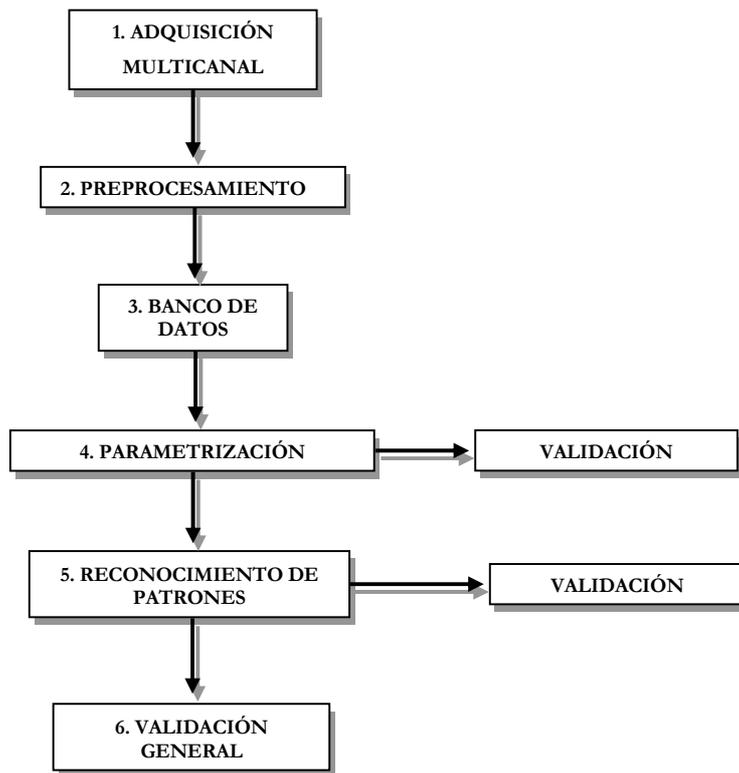


Fig. 1 Metodología general para la clasificación de información acústica normal y anormal con NID.

## **5.1 ADQUISICIÓN MULTICANAL**

### **5.1.1 EQUIPO**

- Se utilizaron 25 micrófonos subminiatura tipo electret BT-1834 Knowles® con respuesta en frecuencia de 100 a 10 KHz, acoplados cada uno a una campana de acrílico con un disco de doble adhesivo en su superficie que permitió la colocación sobre el tórax de los sujetos.
- Para la adquisición de las señales se utilizó una computadora personal estándar con procesador AMD Duron de 1GHz de velocidad y 250 MB de memoria RAM.
- La digitalización se realizó a través de una tarjeta de conversión A/D National Instruments PCI-6071E de 12 bits de resolución. Este sistema tiene una capacidad máxima de 64 canales analógicos con opción de ganancias independientes seleccionadas mediante software. La máxima frecuencia es de 625 KHz y captura hasta  $1.25 \times 10^6$  muestras/segundo.
- Para los registros se utilizaron 28 canales de entrada a una frecuencia de muestreo de 10 KHz. La tarjeta se configuró en modo REFERENCED SINGLE ENDED y el rango de los canales de entrada analógicos es de  $-10V$  a  $+10V$ .
- El equipo de flujo fue calibrado con un rotámetro y se adquirió mediante un neumotacógrafo tipo Fleish, acoplado al transductor de presión Validyne®.
- La maniobra respiratoria se controló mediante la visualización de la señal de flujo en la pantalla de un osciloscopio.

### **5.1.2 SUJETOS**

Se obtuvieron registros de sonidos respiratorios de 19 pacientes del Instituto Nacional de Enfermedades Respiratorias (INER) con un diagnóstico de NID, realizado por neumólogos del Instituto sobre la base de las características clínicas y funcionales, a datos obtenidos por

rayos x, tomografía y biopsia a cielo abierto. También se adquirieron registros de 10 sujetos sanos diagnosticados sobre la base de la historia clínica, pruebas funcionales y rayos X.

Todos los sujetos fueron informados sobre los requerimientos y objetivos del experimento obteniéndose así el consentimiento informado de acuerdo a la declaración de Helsinki .

### 5.1.3 ADQUISICIÓN

Durante la adquisición de los registros, estuvieron presentes médicos encargados de la preparación, interacción y cuidado de los pacientes y sujetos sanos.

Se colocaron 25 micrófonos sobre la parte posterior del tórax del paciente con una distancia entre micrófonos de 5 cm [5]. Este procedimiento tuvo una duración de 15 minutos aproximadamente.

Las posiciones de los micrófonos fueron propuestas considerando abarcar sonidos bronquiales, bronquio-vesiculares y vesiculares. La nomenclatura que define cada una de las posiciones está integrada por las siglas: P (Posterior), L (Left), R (Right), X (Axilar), C (Clavicular), M (Media) y un número del 1 al 5 según la figura 2.

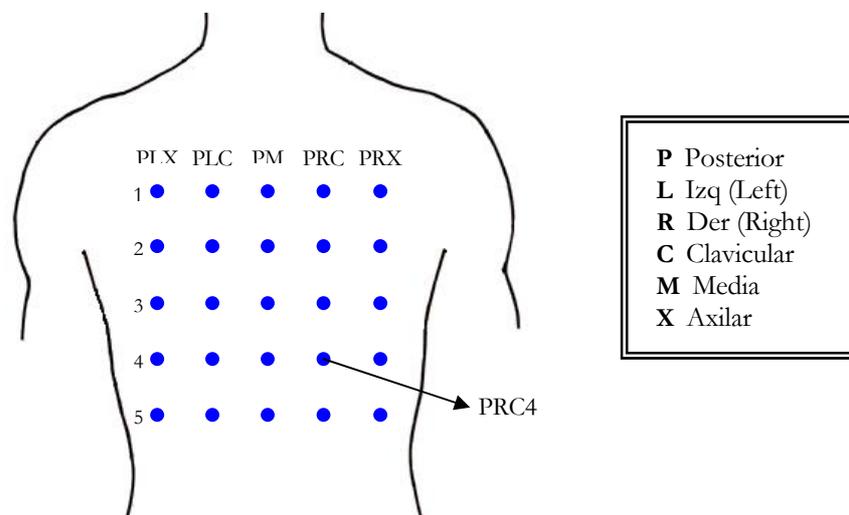


Fig. 2 Nomenclatura y posición de los micrófonos sobre el tórax

Los registros fueron obtenidos mientras los sujetos se encontraban sentados, el médico les indicó a los pacientes la maniobra a realizar. La maniobra tuvo una duración de 15 segundos y consistió de 5 segundos de apnea inspiratoria, seguida de 5 segundos de respiración normal aproximadamente a un flujo de ~1.5 Litros/seg y finalmente el paciente realizó 5 segundos de apnea espiratoria.

Se adquirieron simultáneamente 28 canales que incluyen a las señales de sonido respiratorio provenientes de los micrófonos (25 canales), la señal de flujo, la señal de electrocardiograma y una señal de línea de base. La frecuencia de muestreo fue de 10 KHz por canal.

Es importante mencionar que el paciente con los micrófonos y el equipo de cómputo con la tarjeta de adquisición se encontraron en salas aisladas e independientes para evitar que los micrófonos captaran ruidos de interferencia debido a la operación del equipo.

En la figura 4 se observa un paciente con los micrófonos colocados sobre su espalda y en la figura 5 un paciente respirando a través del equipo de flujo visualizando la señal en la pantalla del osciloscopio.

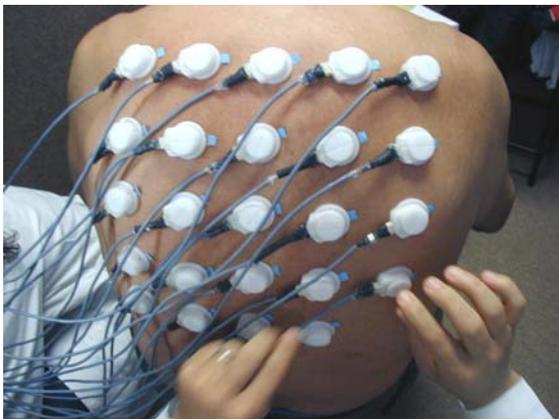


Fig. 3 Micrófonos sobre la espalda de un paciente con NID.



Fig. 4 Paciente respirando a través del equipo de flujo visualizando la señal en la pantalla del osciloscopio

#### 5.1.4 PROGRAMA DE ADQUISICIÓN

El programa de adquisición y visualización en pantalla de las señales de sonido respiratorio se desarrolló como un proyecto independiente en el Laboratorio de Procesamiento de Señales e Imágenes Biomédicas de la UAM-I [6]. El instrumento virtual ACQANARS4 es un programa diseñado en ambiente LabView® 6i para llevar a cabo la adquisición de ruidos respiratorios a través de un total de 64 canales independientes, con los cuales se pueden registrar datos en forma continua o finita.

Pueden seleccionarse diferentes ganancias para cada uno de los canales, las cuales se especifican en la tabla 1.

Tabla 1

Ganancias correspondientes a los canales utilizados en la adquisición

CANAL	SALIDA BIPOLAR	GANANCIA
0 (Mic. línea de base)	$\pm 1$ V	5
25 (Micrófonos)	$\pm 1$ V	5
60 (ECG)	$\pm 1$ V	5
61 (Flujo)	$\pm 10$ V	0.5

La frecuencia de muestreo se fijó en 10KHz y el número de muestras a adquirir fue de 150 000 muestras. La información adquirida puede ser almacenada en archivos de dos tipos: archivos binarios y archivos de texto. Estos últimos tienen la ventaja de poder ser procesados mediante programas de hojas de cálculo, Matlab®, etc. Sin embargo poseen la desventaja de ocupar más espacio de memoria, en comparación con los archivos de tipo binario.

En la figura 6 se presenta la estructura de la carátula de edición del ACQANARS4 donde pueden configurarse parámetros como número, posición y nomenclatura de los canales, modo de adquisición (continua o finita), formato de grabación del archivo (Binario o texto), frecuencia de muestreo, número de muestras a adquirir, tamaño del buffer, ganancias independientes por canal y datos personales y clínicos del sujeto.

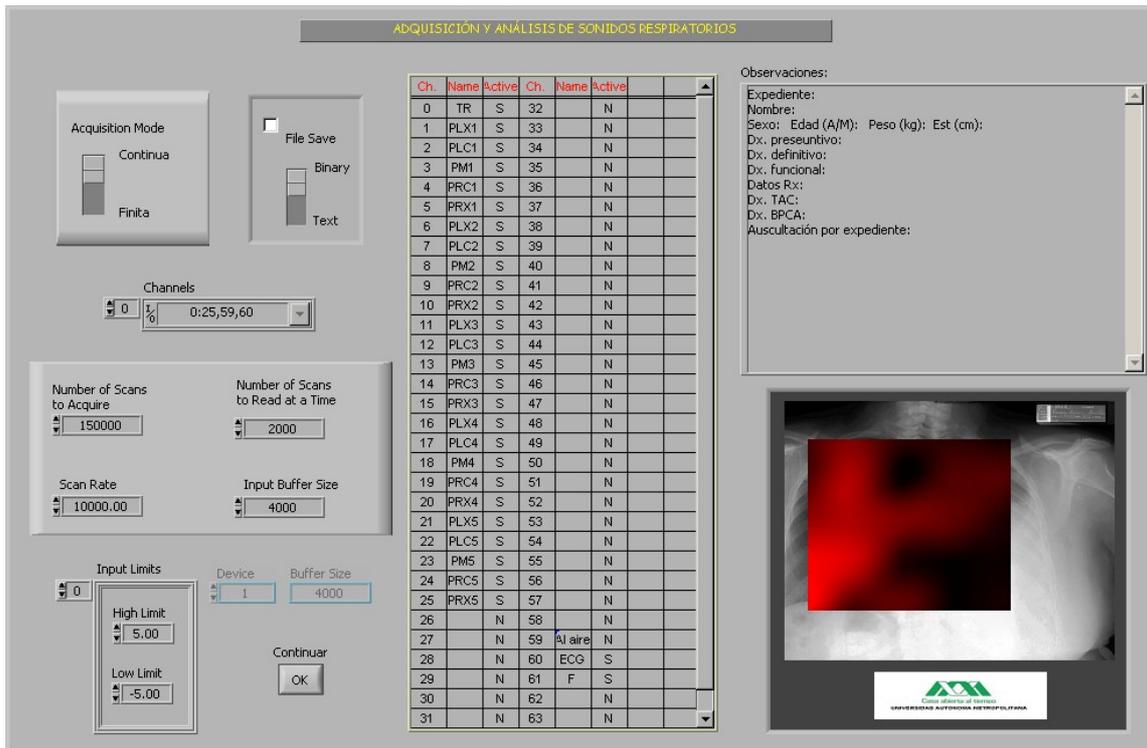


Fig. 6 Estructura de la carátula de edición del ACQANARS4

En la figura 7 se muestra la estructura de la carátula principal del programa en LabView llamado Lec\_Bin que tiene como finalidad convertir el archivo tipo binario a archivo tipo texto, para posteriormente ser procesado por otra herramienta computacional. En la parte izquierda se observan todos los datos configurados en el ACQANARS4 y del lado derecho las gráficas de todas las señales adquiridas con el número de canal en el extremo derecho.

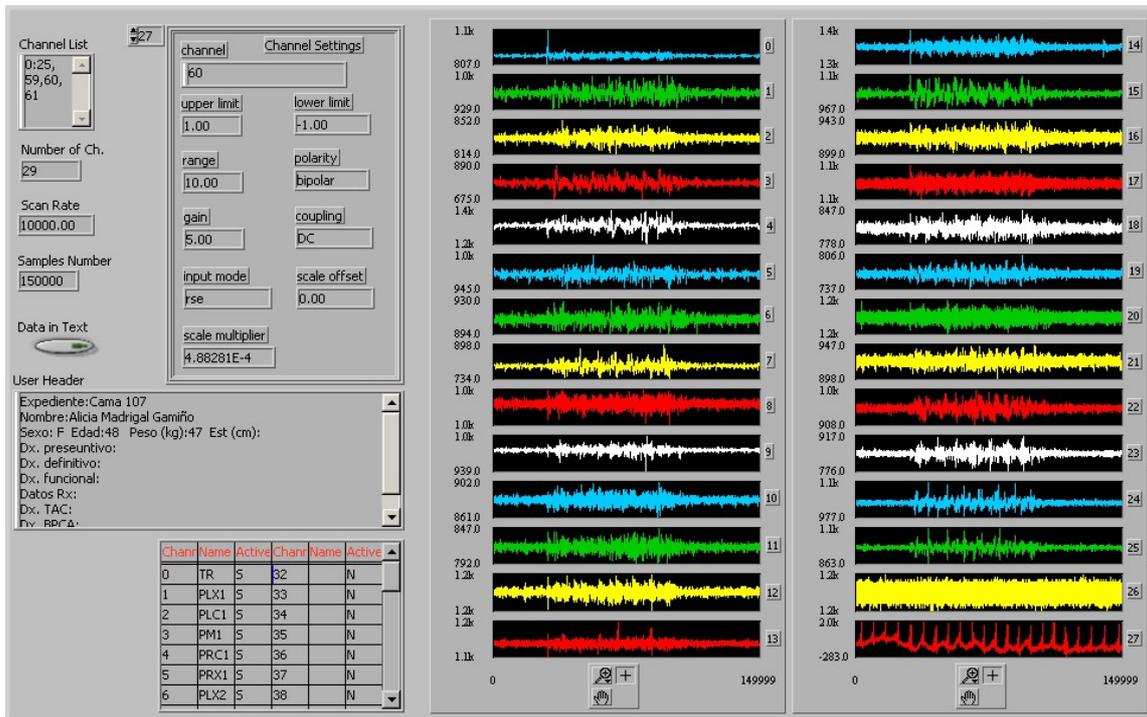


Fig. 7 Estructura de la carátula principal del Lec\_Bin

## 5.2 PREPROCESAMIENTO

Se programó una interfaz gráfica en ambiente MATLAB® para Windows® como herramienta computacional para manipular, analizar y procesar las señales digitalizadas. Este programa realiza diversas tareas como calibración, eliminación de los periodos de apnea, eliminación de artefactos de movimiento, filtrado de las señales, segmentación de las señales en fase inspiratoria y espiratoria y parametrización de las señales por diferentes métodos.

### 5.2.1 CALIBRACIÓN

Los micrófonos que se utilizaron fueron elegidos de modo que todos tuvieran ganancia y respuesta en frecuencia similares, sin embargo para evitar diferencias en los niveles de amplitud se adquirieron señales de calibración grabando la señal de un tono puro simultáneamente en todos los canales. Con base en estas señales se calculó un factor de escalamiento que multiplicó a la señal del canal correspondiente.

### **5.2.2 ELIMINACIÓN AUTOMÁTICA DE PERIODOS DE APNEA Y SEGMENTACIÓN DE FASES INSPIRATORIAS Y ESPIRATORIAS.**

El periodo de apnea no es útil para el presente estudio, por tal motivo es eliminado. Las señales de sonido respiratorio son segmentadas automáticamente en fases inspiratoria y espiratoria de acuerdo a la señal de flujo, considerando los cruces por cero. Las fases espiratorias son descartadas debido a que no aportan información relevante para la clasificación de los sonidos en el caso de enfermedades como las NID [7], [8]. En la figura 8, se visualizan los sonidos respiratorios multicanales y la señal de flujo de un sujeto con NID con su respectiva nomenclatura en el extremo derecho, por medio de la interfaz gráfica se realiza la segmentación automática de las fases inspiratorias y espiratorias

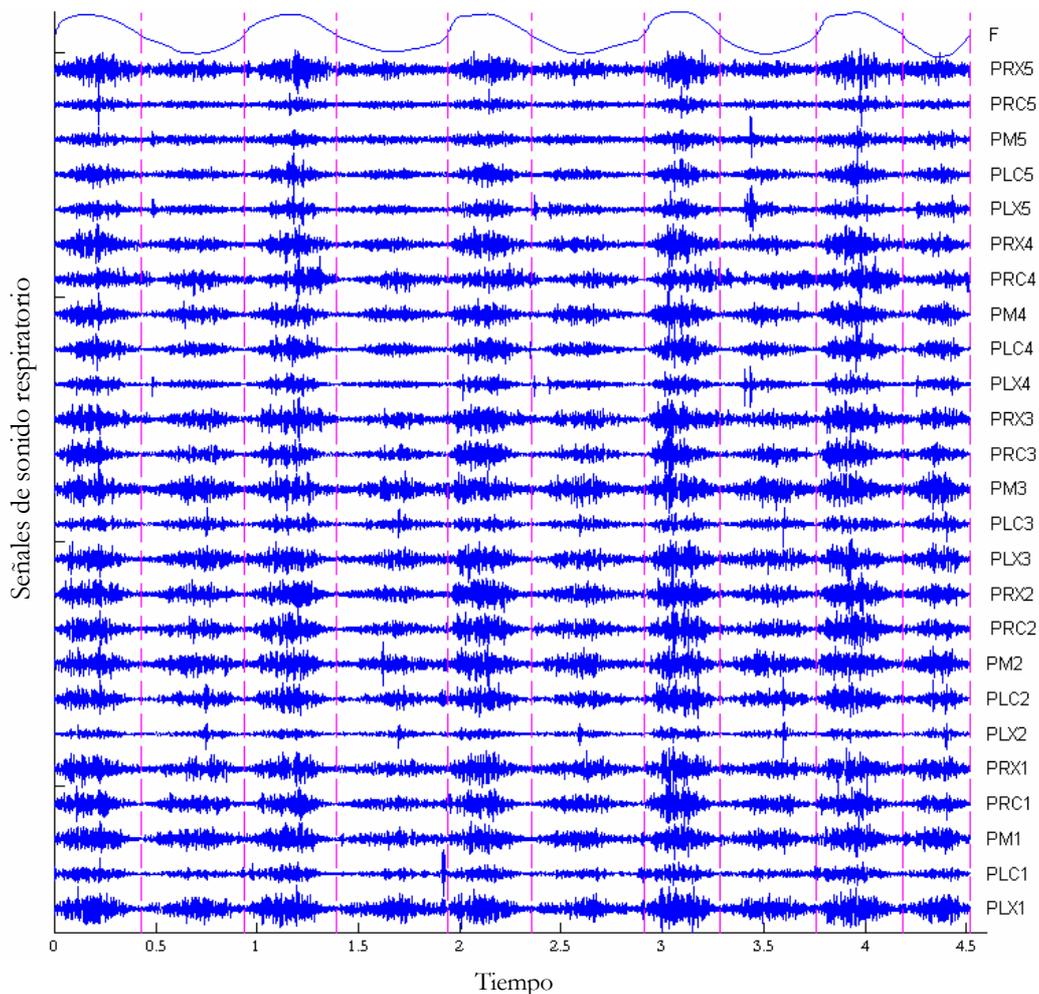


Fig. 8 Señales de sonido respiratorio multicanales, segmentadas en fases inspiratorias y espiratorias con base en la señal de flujo. Las señales se muestran filtradas con un filtro pasabandas de 75 a 2000 Hz y cada una se encuentra escalada con respecto a la máxima amplitud de los sonidos respiratorios.

### 5.2.3 FILTRADO DE LAS SEÑALES PARA LA REDUCCIÓN DE RUIDO CARDÍACO Y MUSCULAR

Para reducir el ruido cardíaco y muscular de las señales de sonido adquiridas se filtraron mediante un filtro FIR pasabanda de fase lineal con frecuencias de corte  $W_{C1}=75$  Hz y  $W_{C2}=2000$  Hz. El número de coeficientes del filtro fue 500 y se utilizó una ventana de Hamming. La Ganancia normalizada del filtro en  $W_C$  es de -6 dB.

Se utilizó un filtro pasabajas para la señal de flujo, con las mismas características pero con una frecuencia de corte  $F_c = 1$  Hz.

#### **5.2.4 ELIMINACIÓN DE ARTEFACTOS**

Durante la adquisición de las señales de sonido existen movimientos propios de la fricción de cables, movimientos del paciente, etc. Los cuales pueden ser removidos automática y manualmente utilizando el programa de interfaz gráfica. En este trabajo de investigación se evitó eliminar artefactos, únicamente fueron descartadas las fases en donde se detectaron por inspección visual artefactos con una amplitud considerable, con relación a la señal de información. Las fases que estuviesen contaminadas con uno o más artefactos, fueron excluidas del estudio junto con las fases correspondientes en los demás canales.

### **5.3 FORMACIÓN DEL BANCO DE DATOS**

Se obtuvieron 19 registros de pacientes y 10 registros de sujetos sanos. A continuación se muestran los criterios de inclusión y exclusión de pacientes que fueron considerados y aprobados por el comité de ética del INER.

#### **5.3.1 CRITERIOS DE INCLUSIÓN:**

**Pacientes:** Consentimiento informado, historia clínica, pruebas de función respiratoria, Rx, ECG, TAC, broncoscopía, lavado broncoalveolar, biopsia a cielo abierto.

**Sanos:** Consentimiento informado, historia clínica, espirometría normal y electrocardiograma normal.

#### **5.3.2 CRITERIOS DE EXCLUSIÓN:**

**Pacientes:** Incapacidad para realizar las maniobras respiratorias, padecimiento infeccioso agregado, registros con pobre relación señal a ruido (RSR).

**Sanos:** Incapacidad para realizar las maniobras respiratorias, registros con baja RSR.

## 5.4 PARAMETRIZACIÓN DE LOS DATOS

La parametrización de las señales de sonido respiratorio es parte fundamental en el tratamiento de los datos y consiste en representar de forma reducida a una señal o parte de una señal, mediante un conjunto de parámetros o características sin pérdida importante de información. La definición de los parámetros de las señales acústicas se realizó mediante vectores de características que fueron obtenidos a través de diferentes técnicas:

### 5.4.1 EIGENVALORES DE LA MATRIZ DE COVARIANZA

Primero fueron segmentadas las fases inspiratorias y espiratorias de todos los canales de cada uno de los sujetos sanos y pacientes. Posteriormente sólo se seleccionaron las fases inspiratorias y fueron divididas en 30 ventanas traslapadas 25%. Se calculó la función de autocovarianza XCOV del vector compuesto por las muestras de una ventana monocanal y se formó una matriz Toeplitz COV con la secuencia XCOV. Los eigenvalores de la matriz de covarianza COV fueron calculados, de los cuales se eligieron los seis eigenvalores de mayor importancia para formar el vector de características correspondiente a esa ventana de tiempo monocanal. Finalmente los vectores de cada ventana se concatenaron junto con los vectores que correspondieran a la misma ventana de tiempo pero de los demás canales. Por simplicidad, más adelante se proporciona un diagrama que explica la forma en que se concatenaron los vectores. El conjunto de vectores de características se dividió para el entrenamiento, prueba y validación del clasificador.

En la figura 9 se esquematiza la forma en que se obtienen los eigenvalores que constituyen al vector de características. La señal de sonido respiratorio se divide en  $M$  ventanas de tiempo traslapadas 25% formando un vector. Se calcula la matriz de covarianza a partir de este vector y se obtienen  $N$  eigenvalores por cada ventana monocanal. Se seleccionan los eigenvalores más importantes y se concatenan con los eigenvalores de los demás canales, pero correspondientes a la misma ventana de tiempo.

A la derecha se muestra la concatenación de los eigenvalores de los  $d$  canales que forman la matriz de datos, que es introducida al clasificador, en donde cada columna representa a una ventana de tiempo multicanal.

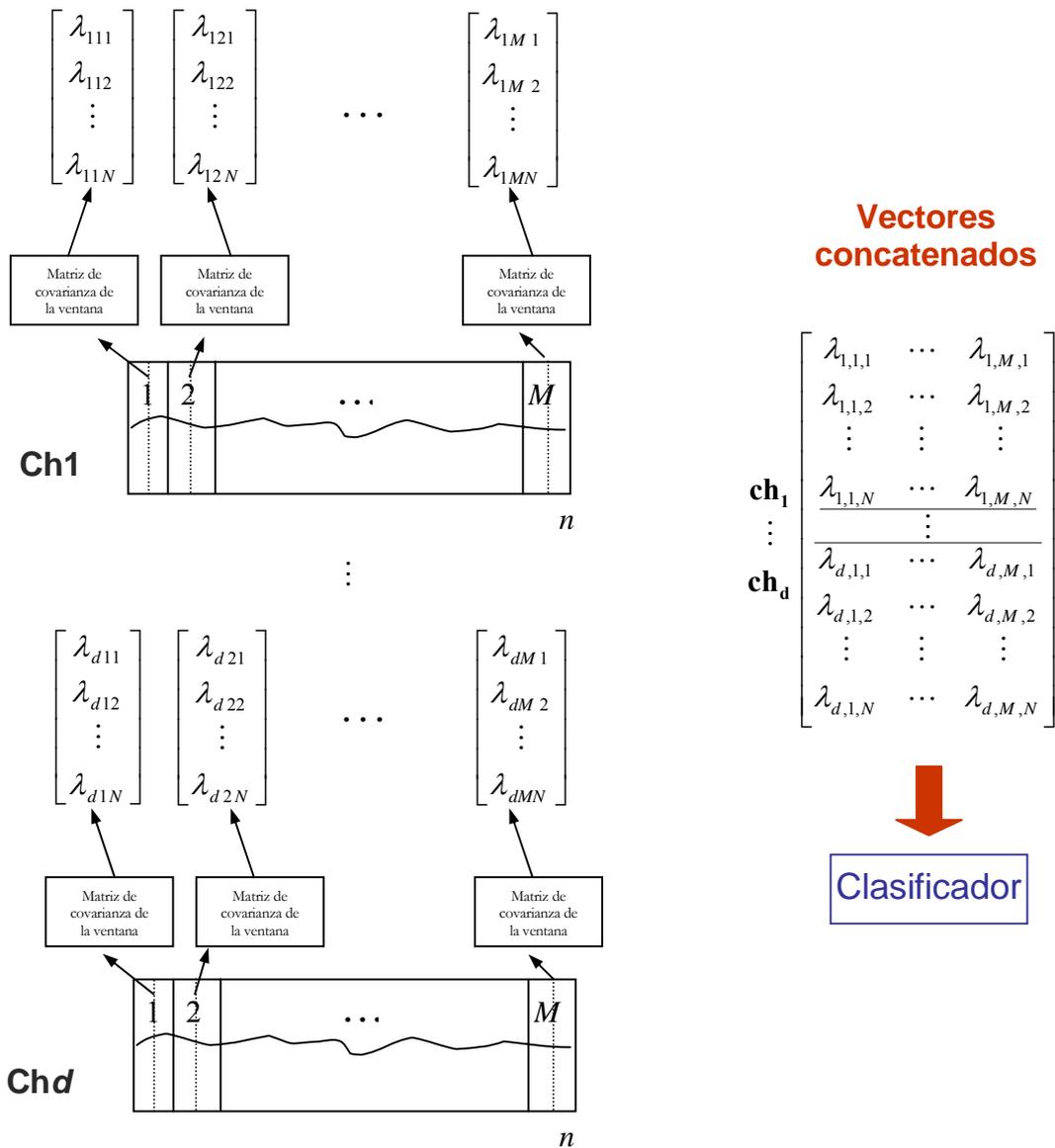


Fig. 9 Concatenación de los vectores de características utilizando los eigenvalores de la matriz de covarianza de la señal de sonido respiratoria. En la gráfica  $\lambda$  representa un eigenvalor de la matriz de covarianza,  $N$  número de eigenvalores,  $M$  número de ventana,  $d$  número de canal y  $n$  es en número de muestras.

### 5.4.2 MEDIDAS CLÁSICAS EN EL DOMINIO DE LA FRECUENCIA

Se calculó el espectro de potencia (EP) de cada fase inspiratoria monocanal a partir de los coeficientes del modelo AR de orden 4, mediante el método de Burg con la función PBURG. El número de puntos de frecuencia calculados fue de 10000 puntos, los cuales fueron integrados para formar el área total bajo la curva del EP. Una frecuencia percentil representa un determinado porcentaje del área total bajo la curva del EP [9]. El vector de parámetros que caracteriza a cada fase inspiratoria se formó con las frecuencias percentiles que corresponden al 25%, 50%, 75%, 90%, 95% y 99% del área total bajo la curva del espectro de potencia. Por lo tanto el vector de características consiste de seis valores de frecuencias percentiles, como se ejemplifica en la figura 10.

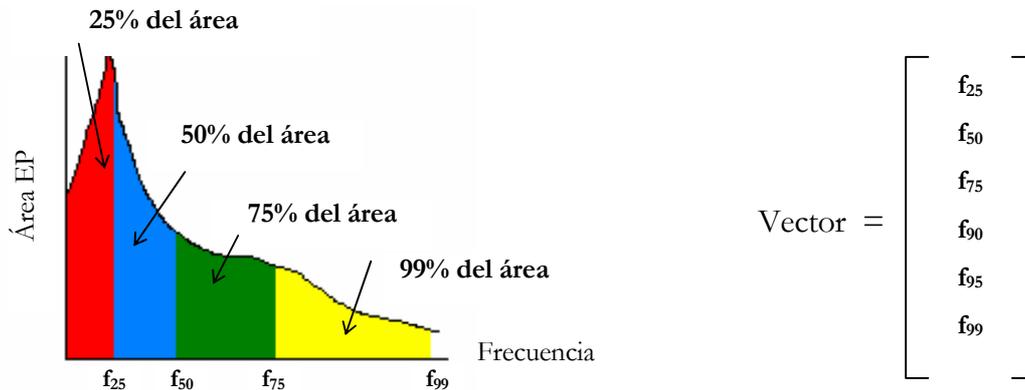


Fig. 10 Frecuencias percentiles

Los parámetros obtenidos en cada fase inspiratoria de 1 hasta N inspiraciones, fueron concatenados verticalmente junto con los parámetros correspondientes a la misma fase pero de diferentes canales, del canal 1 hasta el canal 25 (Figura 11). El conjunto de vectores de características fue dividido en tres subconjuntos para entrenamiento, prueba y validación del clasificador.

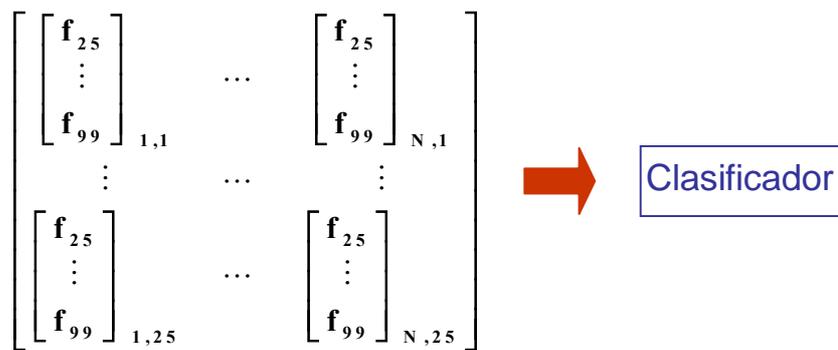
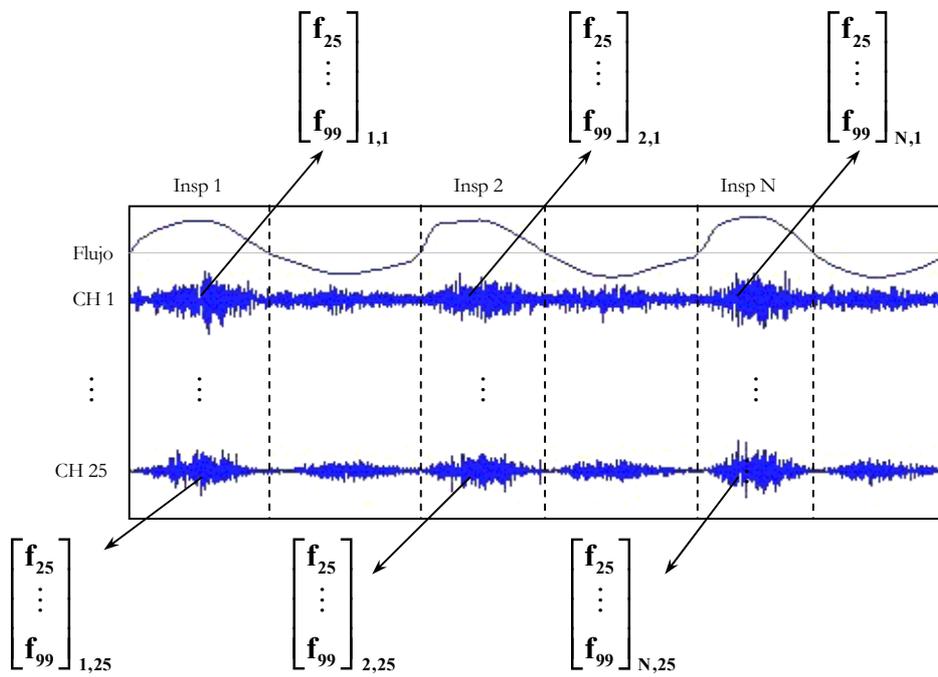


Fig. 11 Concatenación de los vectores de características en la técnica de medidas clásicas en el dominio de la frecuencia.

### 5.4.3 MODELO AR MONOCANAL

Las señales de sonido de cada uno de los sujetos sanos y pacientes fueron segmentadas en fases inspiratorias y espiratorias. Posteriormente se seleccionaron únicamente las fases inspiratorias y fueron divididas en 30 ventanas traslapadas 25%, cada ventana formó un vector  $V$ . Previamente se determinó el orden del modelo AR monocanal de acuerdo al criterio de información de Akaike y se evaluó la bondad del ajuste mediante la función de autocorrelación de los residuos y la correlación cruzada entre la serie y los residuos, asegurando que los valores se encontraran dentro del intervalo de confianza (95%). El orden del modelo seleccionado fue de 4.

Se calculó el modelo AR monocanal de cada ventana monocanal mediante la función ARX que utiliza el método de mínimos cuadrados (Least Squares) para estimar los coeficientes del modelo. Para una serie de tiempo, ARX calcula el modelo AR de orden  $p = na$  el cual tiene la forma:

$$\mathbf{A}(\mathbf{q})\mathbf{x}(\mathbf{n}) = \mathbf{e}(\mathbf{n}) \quad (1)$$

$$\text{donde: } \mathbf{A}(\mathbf{q}) = 1 + \mathbf{a}_1\mathbf{q}^{-1} + \dots + \mathbf{a}_{na}\mathbf{q}^{-na} \quad (2)$$

$$\text{si } p = na; \quad \mathbf{x}[\mathbf{n}] = -\mathbf{a}_1\mathbf{x}[\mathbf{n} - 1] - \mathbf{a}_2\mathbf{x}[\mathbf{n} - 2] - \dots - \mathbf{a}_p\mathbf{x}[\mathbf{n} - p] + \mathbf{e}[\mathbf{n}] \quad (3)$$

En donde  $\mathbf{x}[\mathbf{n}]$  es el valor de  $\mathbf{x}$  dada la muestra  $\mathbf{n}$ . En la ecuación (3),  $\mathbf{a}_1$  hasta  $\mathbf{a}_p$  son los coeficientes escalares del modelo AR monocanal y  $\mathbf{e}[\mathbf{n}]$  es una variable aleatoria no correlacionada con media cero.

En la figura 12 se esquematiza una señal dividida en ventanas de tiempo monocanales.  $\mathbf{M}$  indica el número de ventanas,  $\mathbf{p}$  el orden del modelo AR monocanal y  $\mathbf{d}$  el número de canales. Se calculan  $\mathbf{p}$  coeficientes del modelo AR monocanal por ventana. Del lado derecho se observa la forma en que son concatenados los vectores de características de cada ventana monocanal para formar el conjunto total de vectores utilizados para entrenamiento del clasificador.

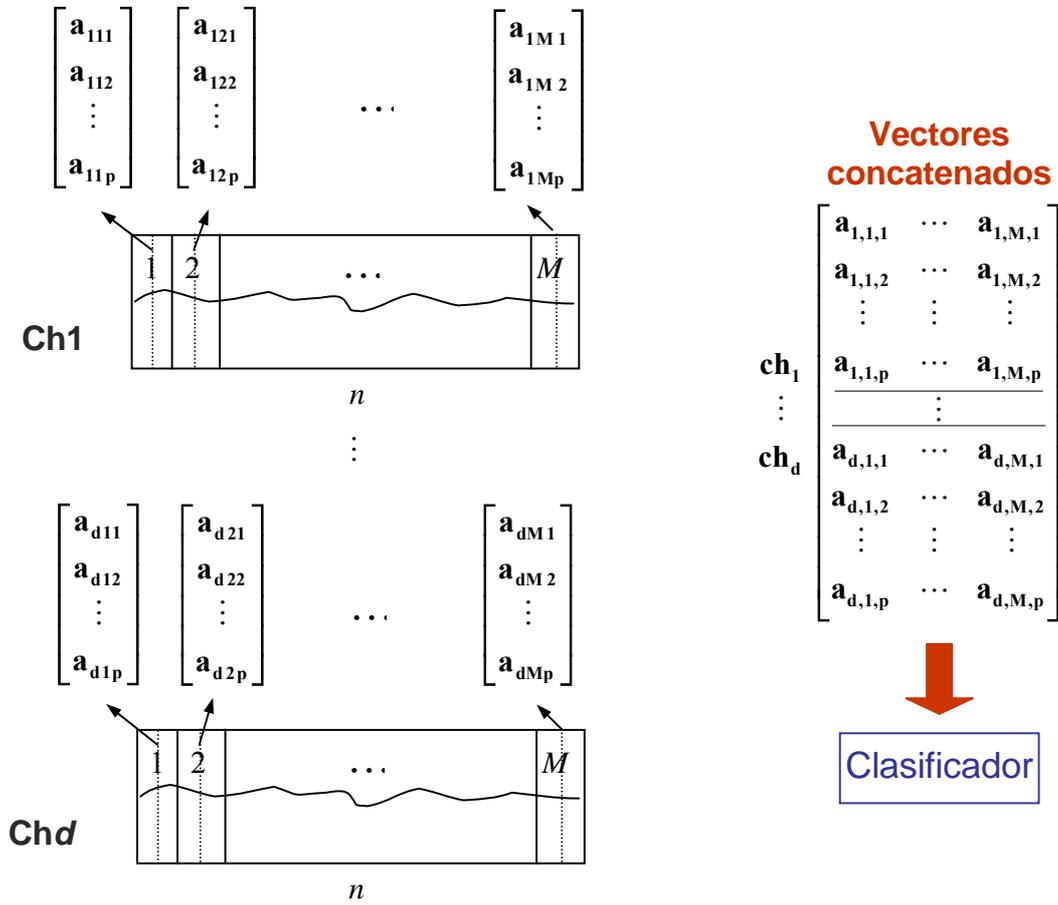


Fig. 12 Concatenación de los vectores de características con la técnica de modelo AR

#### 5.4.4 MODELO AR MULTICANAL

Del mismo modo que el modelo monocanal las señales fueron segmentadas en fases inspiratorias y espiratorias de cada uno de los sujetos sanos y pacientes. Posteriormente se seleccionaron las fases inspiratorias y fueron divididas en 30 ventanas multicanales traslapadas 25%, cada ventana formó una matriz en donde cada columna pertenece a un canal y cada renglón representa observaciones.

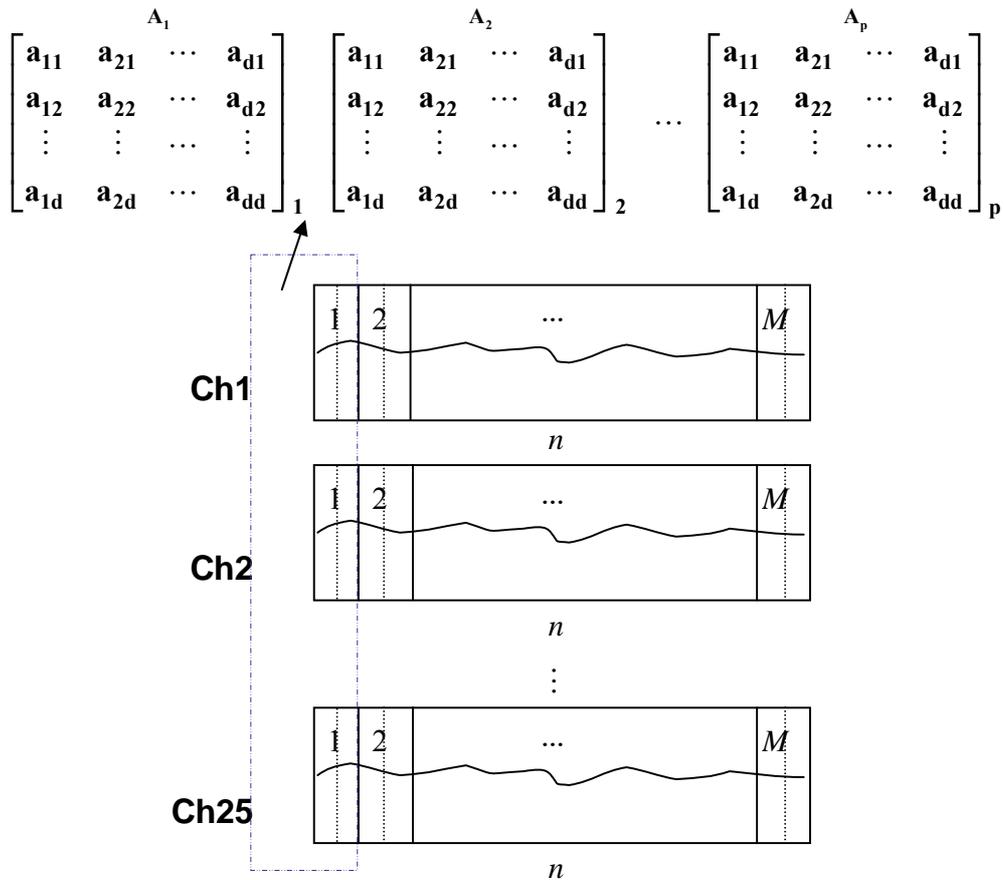
El orden óptimo del modelo AR multicanal se determinó sobre la base del criterio de información de Akaike evaluando aleatoriamente vectores de diferentes sujetos y canales. Se implementó un modelo AR multicanal de orden 4.

El modelo AR multicanal se representa mediante la siguiente ecuación, en donde  $\bar{\mathbf{x}}(\mathbf{k})$  y  $\bar{\mathbf{e}}(\mathbf{k})$  son vectores de dimensión  $\mathbf{d}$ , y  $\mathbf{A}(\mathbf{1})$  hasta  $\mathbf{A}(\mathbf{p})$  son las matrices de coeficientes de dimensión  $\mathbf{d} \times \mathbf{d}$ .  $\mathbf{p}$  es el orden del modelo AR multicanal.

$$\bar{\mathbf{x}}(\mathbf{k}) = -\mathbf{A}(\mathbf{1})\bar{\mathbf{x}}(\mathbf{k} - \mathbf{1}) - \mathbf{A}(\mathbf{2})\bar{\mathbf{x}}(\mathbf{k} - \mathbf{2}) + \dots - \mathbf{A}(\mathbf{p})\bar{\mathbf{x}}(\mathbf{k} - \mathbf{p}) + \bar{\mathbf{e}}(\mathbf{k}) \quad (4)$$

De este modo se obtuvieron 4 matrices de tamaño 25X25 por cada ventana multicanal. La forma en que se organizan las ventanas de tiempo multicanales se ejemplifica en la figura 13.

Para la estimación de los coeficientes del modelo AR multicanal se utilizó la colección de módulos ARFIT para ambiente MATLAB® [10], que utiliza el algoritmo de mínimos cuadrados (LS) para la estimación de las matrices de coeficientes del modelo AR multicanal y la matriz de covarianza del ruido.



5.4. Fig. 13 Cálculo de las matrices de coeficientes en la técnica de modelo AR multicanal

La dimensión de los vectores de características obtenidos mediante el modelo AR multicanal resulta poco práctico introducirlas al clasificador debido a limitaciones computacionales y fue necesario recurrir a técnicas de reducción del espacio de características para facilitar el manejo de los datos. La reducción se realizó de dos formas:

- a) Reduciendo el número de coeficientes de cada matriz individual  $\mathbf{A}$  por separado y concatenando verticalmente los vectores reducidos.
- b) Concatenando verticalmente las columnas de  $\mathbf{A}(1)$  hasta  $\mathbf{A}(4)$  y formando un vector total por cada ventana multicanal. Ver figura 14.

Matrices de coeficientes del modelo AR multicanal

$$A_1 = \begin{bmatrix} \mathbf{a}_{11} & \mathbf{a}_{21} & \cdots & \mathbf{a}_{d1} \\ \mathbf{a}_{12} & \mathbf{a}_{22} & \cdots & \mathbf{a}_{d2} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{a}_{1d} & \mathbf{a}_{2d} & \cdots & \mathbf{a}_{dd} \end{bmatrix}_1$$

$$A_2 = \begin{bmatrix} \mathbf{a}_{11} & \mathbf{a}_{21} & \cdots & \mathbf{a}_{d1} \\ \mathbf{a}_{12} & \mathbf{a}_{22} & \cdots & \mathbf{a}_{d2} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{a}_{1d} & \mathbf{a}_{2d} & \cdots & \mathbf{a}_{dd} \end{bmatrix}_2$$

$$A_3 = \begin{bmatrix} \mathbf{a}_{11} & \mathbf{a}_{21} & \cdots & \mathbf{a}_{d1} \\ \mathbf{a}_{12} & \mathbf{a}_{22} & \cdots & \mathbf{a}_{d2} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{a}_{1d} & \mathbf{a}_{2d} & \cdots & \mathbf{a}_{dd} \end{bmatrix}_3$$

$$A_4 = \begin{bmatrix} \mathbf{a}_{11} & \mathbf{a}_{21} & \cdots & \mathbf{a}_{d1} \\ \mathbf{a}_{12} & \mathbf{a}_{22} & \cdots & \mathbf{a}_{d2} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{a}_{1d} & \mathbf{a}_{2d} & \cdots & \mathbf{a}_{dd} \end{bmatrix}_4$$

a)  $V =$  Concatenación vertical de la matriz individual de  $A_1$

$$V = \begin{bmatrix} \mathbf{a}_{11} \\ \mathbf{a}_{12} \\ \vdots \\ \mathbf{a}_{1d} \\ \mathbf{a}_{21} \\ \mathbf{a}_{22} \\ \vdots \\ \mathbf{a}_{2d} \\ \vdots \\ \mathbf{a}_{d1} \\ \mathbf{a}_{d2} \\ \vdots \\ \mathbf{a}_{dd} \end{bmatrix}_1$$

b)  $V_T =$  Concatenación vertical de todas las matrices individuales de  $A_1$  hasta  $A_4$

$$V_T = \begin{bmatrix} \mathbf{a}_{111} \\ \mathbf{a}_{121} \\ \vdots \\ \mathbf{a}_{1d1} \\ \mathbf{a}_{211} \\ \mathbf{a}_{221} \\ \vdots \\ \mathbf{a}_{2d1} \\ \vdots \\ \mathbf{a}_{d11} \\ \mathbf{a}_{d21} \\ \vdots \\ \mathbf{a}_{dd1} \\ \vdots \\ \mathbf{a}_{114} \\ \mathbf{a}_{124} \\ \vdots \\ \mathbf{a}_{1d4} \\ \mathbf{a}_{214} \\ \mathbf{a}_{224} \\ \vdots \\ \mathbf{a}_{2d4} \\ \vdots \\ \mathbf{a}_{d14} \\ \mathbf{a}_{d24} \\ \vdots \\ \mathbf{a}_{dd4} \end{bmatrix}$$

Fig. 14 Concatenación de columnas de las matrices  $A$ . Una línea indica cambio de columna dentro de una misma matriz y una doble línea indica cambio de matriz.

#### 5.4.4.1.1 Matriz individual de coeficientes del modelo AR multicanal

##### 5.4.4.1.1.1 REDUCCIÓN POR DESCOMPOSICIÓN EN VALORES SINGULARES (SVD) DE LA MATRIZ $A$

1. - Se calculan  $K$  valores singulares de cada una de las matrices de coeficientes del modelo AR multicanal.
2. - Se concatenan los vectores de cada matriz individual, obteniendo una vector columna de  $4K$  parámetros por cada ventana multicanal

##### 5.4.4.1.1.2 REDUCCIÓN POR EIGENVALORES DE LA MATRIZ DE COVARIANZA DE $A$

1. - Obtención de la matriz de covarianza  $covA$  de cada matriz de coeficientes  $A$  del modelo AR multicanal
2. - Se calculan los eigenvalores de la matriz de covarianza y se ordenan en forma descendente.
3. -Se concatenan los vectores de eigenvalores de cada matriz individual.

##### 5.4.4.1.1.3 REDUCCIÓN POR PROYECCIÓN DE EIGENVECTORES DE LA MATRIZ DE COVARIANZA DE LOS COEFICIENTES AR CONCATENADOS VERTICALMENTE

1. - Se formó un vector  $V$  por cada matriz individual con la concatenación vertical de las columnas de la matriz  $A$ .
2. - Cálculo de la función de autocovarianza  $xcovV$  del vector resultante  $V$
3. - Se obtuvo la matriz Toeplitz  $covV$  de la función  $xcovV$  que equivale a calcular la matriz de covarianza de  $V$
4. - Posteriormente se calcularon los eigenvalores y eigenvectores de la matriz  $covV$  ordenándolos en forma descendente
5. - Se seleccionaron los  $N$  eigenvectores principales de acuerdo a sus eigenvalores, y se

realizaron las proyecciones de los eigenvectores sobre los vectores originales para reducir la dimensión, de este modo se obtiene un vector de tamaño reducido por cada matriz individual.

6. - Se concatenan verticalmente los 4 vectores provenientes de las 4 matrices individuales, entonces se obtiene un vector columna de  $4N$  parámetros por cada ventana multicanal.

#### **5.4.4.1.2 Matriz total de coeficientes del modelo AR multicanal**

##### **5.4.4.1.2.1 REDUCCIÓN POR COMPONENTES PRINCIPALES (PCA)**

1. - Se formó un vector  $V$  ( $625 \times 1$ ) por cada matriz con la concatenación vertical de las columnas de cada matriz individual, posteriormente se concatenan los cuatro vectores provenientes de las matrices individuales, obteniendo de esta manera un vector total  $V_T$  de tamaño  $2500 \times$  cada ventana multicanal.

2. - Se forma el conjunto total de vectores de características de todas las ventanas multicanales de sujetos sanos y enfermos.

3. - Los datos del conjunto de entrada son normalizados con media cero y desviación estándar igual a la unidad, utilizando la función PRESTD del toolbox de redes neuronales.

4. - Se calculan los componentes principales con la función PREPCA cuya salida corresponde a los vectores transformados y la matriz de transformación, que será utilizada para transformar cualquier vector que vaya a ser introducido a la red neuronal. El componente de mínima fracción de varianza de los datos que se conservan se consideró del 2%.

5. - Finalmente se obtiene el conjunto total de entrada de 136 características

## 5.5 RECONOCIMIENTO DE PATRONES

Para “reconocer” si un vector que representa a un segmento de la señal, es sonido respiratorio normal o anormal y decidir si un sujeto está sano o enfermo, se implementó una red neuronal artificial (RNA) del tipo supervisado (Figura 15). La etapa de reconocimiento de patrones tiene como finalidad la clasificación de los vectores de características utilizando patrones o valores deseados (targets) asignados a cada vector de entrada, tomando en cuenta la información a priori basada en el diagnóstico médico.

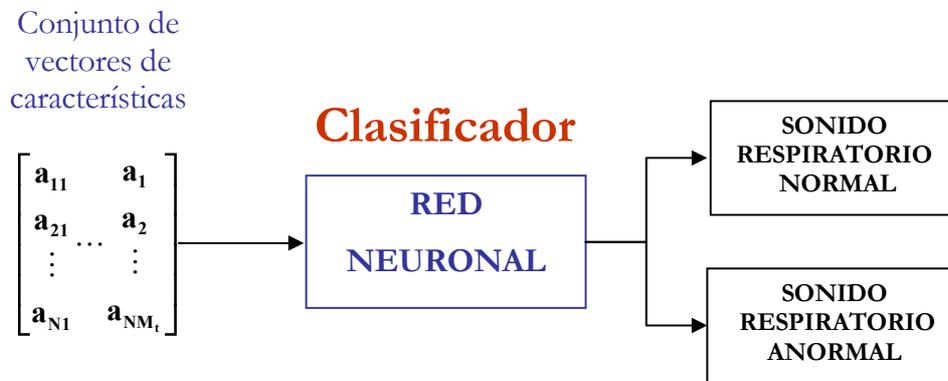


Fig. 15 Etapa de Reconocimiento de Patrones que utiliza una RNA para clasificar sonidos respiratorios normales y anormales representados por vectores de características, en donde N es el número de características y  $M_i$  es el número de ventanas multicanales del conjunto total de vectores.

Las redes neuronales fueron desarrolladas y entrenadas con 100 épocas aproximadamente, mediante la interfaz gráfica NNTOOL de MATLAB® que permite al usuario:

- Crear una red neuronal, importar datos (entrada o valores deseados) y RNAs del espacio de trabajo o de un archivo, exportar datos (salida) y RNAs al espacio de trabajo o a un archivo, inicializar los datos de la RNA, entrenar y simular la red seleccionada.

Se entrenó a la RNA con el algoritmo de aprendizaje de retropropagación (Feed Forward Backpropagation, FFB) utilizando la función LEARNGDM (Gradient Descent with Momentum weight/bias Learning) y función de entrenamiento sobre la base de la regla de Levenberg-Marquardt [11].

En la arquitectura de la RNA se consideró una capa oculta y una capa de salida con un solo nodo o neurona. La función de transferencia no lineal de las neuronas ocultas y de salida fue una tangente sigmoidea.

La nomenclatura utilizada en el toolbox de redes neuronales se muestra en la figura 16, el número de capas ocultas y de neuronas ocultas se varió de acuerdo al número de características del vector de entrada.

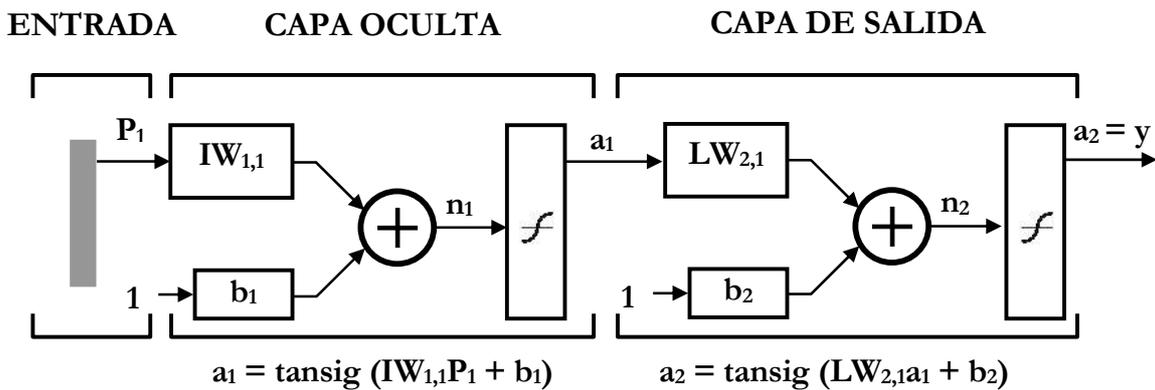


Fig. 16 Arquitectura de la RNA utilizada como clasificador.

En la figura 16  $\mathbf{P}_1$  es el vector de entrada con un número determinado de parámetros o características,  $\mathbf{IW}$  y  $\mathbf{LW}$  son las matrices de pesos que ponderan a un valor de entrada,  $\mathbf{a}_1$  y  $\mathbf{a}_2$  son las salidas de la capa oculta y capa de salida respectivamente [12].

Los valores máximo y mínimo para los valores deseados o patrones se consideraron de +1 para los sujetos sanos y -1 para los sujetos enfermos. Los valores de la salida de la red neuronal fueron clasificados mediante un umbral. Los vectores que se encontraron fuera del intervalo del umbral fueron descartados.

Vector	Valor deseado	Umbral de clasificación
Sano	+ 1	$\geq + 0.8$
Enfermo	- 1	$\leq - 0.8$

El desempeño de la red fue evaluado según la función del Error Cuadrático Medio (MSE), observando la curva de aprendizaje de la RNA.

Por otra parte, no existe un criterio bien definido que indique el momento en que debe detenerse la fase de entrenamiento y el ajuste de los pesos en una RNA. Sin embargo, existen criterios prácticos que sugieren cuando puede detenerse el proceso de entrenamiento [13]. Un método para mejorar la generalización consiste en dividir los datos disponibles en tres subconjuntos. El primer subconjunto es el conjunto de entrenamiento, el cual es usado para calcular el gradiente y actualizar los pesos (**W**) y términos de sesgo (**b**) de la red. El segundo subconjunto es el conjunto de validación. El error alcanzado con el conjunto de validación es monitoreado durante el proceso de entrenamiento, y normalmente decrece durante la fase inicial de entrenamiento al igual que el error con el conjunto de entrenamiento. Sin embargo, cuando la red comienza a sobre-ajustar los datos, el error con el conjunto de validación comenzará a incrementarse. Si esto sucede para un número específico de iteraciones, se sugiere detener el entrenamiento. El conjunto de prueba no es usado durante el entrenamiento, pero es utilizado para comparar diferentes modelos o arquitecturas de red.

El conjunto total de vectores de características fue dividido aleatoriamente en tres subconjuntos: entrenamiento (80% del total), prueba (10%) y validación (10%). El procedimiento de partición del conjunto de vectores es similar a la herramienta estadística de validación cruzada (Cross Validation) [11], en donde el modelo de RNA se valida con un conjunto de datos diferente del que se utilizó para la estimación de los parámetros de la RNA.

En la figura 17 se observa una gráfica de la curva de aprendizaje de la red, con el conjunto de entrenamiento la red se desempeña muy bien, sin embargo al evaluar la red con los conjuntos de prueba y validación el error se incrementa.

Finalmente, para explorar la generalización que realiza la red se utilizó el procedimiento de validación cruzada y el desempeño de la red fue evaluado en base a las siguientes ecuaciones:

$$SENSITIVIDAD = \frac{norm\_norm}{norm\_norm + enf\_norm} \times 100$$

$$ESPECIFICIDAD = \frac{enf\_enf}{enf\_enf + norm\_enf} \times 100$$

$$DIAGNÓSTICO DE DESEMPEÑO = \frac{Sensitividad + Especificidad}{2}$$

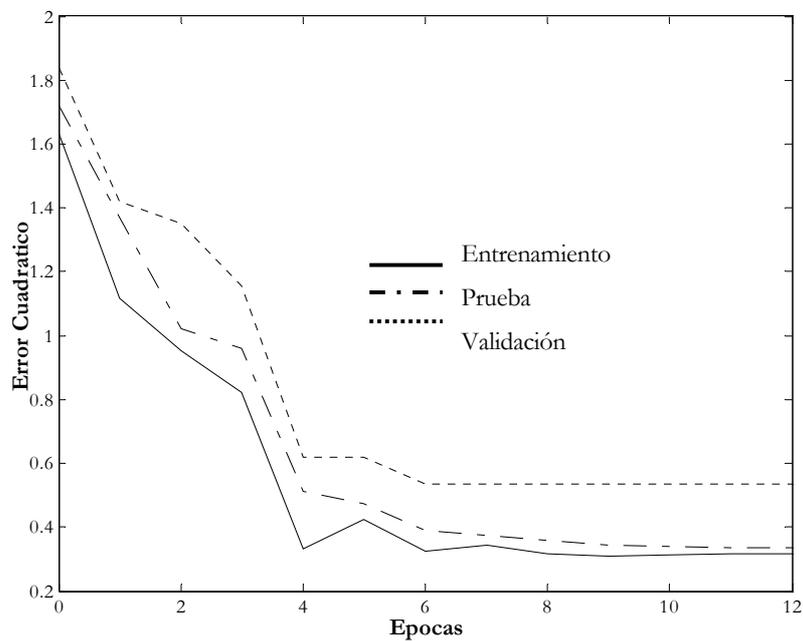


Fig. 17 Curva de aprendizaje de una red con conjuntos de entrenamiento, prueba y validación

## 5.6 VALIDACIÓN GENERAL

El objetivo de esta etapa es validar cada técnica de parametrización presentando al clasificador datos de sujetos no considerados en procesos previos. Para entrenamiento, prueba y validación de la red neuronal se utilizaron 4 registros de sujetos sanos y 4 registros de pacientes con NID lo que representa a una determinada población, compuesta de una enorme cantidad de datos,

necesaria para una buena generalización del clasificador. La validación general consistió en introducir datos de registros de sujetos sanos y pacientes que no se utilizaron en etapas previas, para analizar el comportamiento el clasificador con datos de diferentes. Es importante resaltar que en los trabajos anteriores, no se ha considerado analizar los resultados del clasificador con nuevos datos.

## 5.7 REFERENCIAS

- [1] R. Chapela, J. H. Salas, R. Sansores, R. Pérez-Padilla, G. Carrillo, M. Sánchez y R. Pérez-Chavira, Rosario. “Alveolitis Alérgica Extrínseca.” *Neumología y Cirugía de Tórax*.1991; 50:2, 9-27.
- [2] L. M. Selman, “Neumopatías Intersticiales Difusas.” 1ª Edición, Editorial Médica Panamericana, 1996.
- [3] J. Giannattasio, A. Miranda, L. Zabala, “Neumonías Difusas Intersticio Alveolares.” *Rev. Cubana Med*, 40(2), pp. 135-44, 2001.
- [4] L. M. Selman, R. Barrios, “Intertitial pulmonary diseases: Selected Topics” CRC Press, Inc. 1991
- [5] M. Kompis, G. R. Wodicka, “Coherence of inspiratory and expiratory Breath Sounds as a Function of Inter-Microphone Distance.” *Proceeding of EMBC*, 1995.
- [6] G. T. Flores, “Sistema de adquisición multicanal para sonidos respiratorios en LabView®.” Proyecto terminal de licenciatura en Ingeniería Biomédica. Universidad Autónoma Metropolitana, Unidad Iztapalapa, 2004.
- [7] A. R. A. Sovijärvi, L. P. Malmberg, G. Charbonneau, J. Vanderschoot, F. Dalmaso, C. Sacco, M. Rossi, J. E. Earis, “Characteristics of breath sounds and adventitious respiratory sounds.” *European Respiratory Review*, 10(77), pp. 591-596, 2000.
- [8] G. R. Epler, C. B. Carrington, E. A. Gaensler, “ Crackles (rales) in the interstitial pulmonary diseases.” *Chest*, 73, pp. 333-339, 1978;
- [9] Y. P. Kahya, U. Çini, Ö. Cerid, “Real-time Regional Respiratory Sound Diagnosis Instrument.” 25th Annual IEEE-EMBS Conference, Cancun Mex., pp. 3098-3101, 2003.
- [10] T. Schneider, A. Neumaier, “ARFIT – A Matlab package for the estimation of parameters and eigenmodes of multivariate autoregressive models.” Courant Institute of Mathematical Sciences, New York University, 1995.

- [11] S. Haykin, "Neural Networks: A comprehensive foundation." New York, NY, Macmillan, 1994.
- [12] Neural Network Toolbox: Backpropagation: Fundamentals
- [13] Neural Network Toolbox Documentation: Backpropagation: Sample Training Session. Referencia ( PuLu92 ).

# CAPÍTULO VI

## RESULTADOS

### 6.1 INTRODUCCIÓN

En este capítulo se presentan los resultados de la clasificación de los sonidos respiratorios normales y anormales bajo un esquema de NID representadas por parámetros extraídos a partir de las cuatro técnicas estudiadas en este trabajo: Eigenvalores de la matriz de covarianza, medidas clásicas en el dominio de la frecuencia, modelo autorregresivo (AR) monocanal y modelo AR multicanal.

### 6.2 BASE DE DATOS

Se adquirieron 19 registros multicanales de sujetos enfermos con edad promedio de  $53.25 \pm 11.72$  diagnosticados con NID, siendo la alveolitis alérgica extrínseca (AAE) el tipo encontrado con mayor frecuencia y 8 registros multicanales de sujetos sanos con edad promedio de  $63.4 \pm 6.69$ . En la tabla 1 se resumen los principales datos acerca de los sujetos enfermos con NID cuyos registros de señales respiratorias multicanales fueron utilizados en la presente investigación. Un sujeto enfermo puede ser diagnosticado con una o varias enfermedades por el personal médico del INER con base en el análisis del historial clínico, las pruebas funcionales respiratorias, rayos X, tomografía axial computarizada y lavado broncoalveolar. En ninguno de los pacientes se realizó biopsia quirúrgica antes de la adquisición de las señales de sonido respiratorio.

TABLA 1  
Características de Sujetos enfermos con NID

	EDAD	PESO	TALLA
<b>MEDIA (<math>\mu</math>)</b>	53.35	61.93	1.56
<b>DESV. EST (<math>\sigma</math>)</b>	$\pm 11.72$	$\pm 10.94$	$\pm 0.10$
<b>DIAGNÓSTICO</b>	NID en estudio, AAE, fibrosis pulmonar secundaria, neumonía cavitada, FPI, insuficiencia respiratoria, tromboembolia pulmonar, silicosis, proteinosis alveolar		

En la tabla 2 se muestran los datos de edad, peso y talla de los sujetos sanos cuyos registros se utilizaron en este trabajo. El diagnóstico se realizó basándose en pruebas de función respiratoria, electrocardiograma y auscultación clínica.

TABLA 2  
Características de Sujetos sanos

	<b>EDAD</b>	<b>PESO</b>	<b>TALLA</b>
<b>MEDIA (<math>\mu</math>)</b>	63.40	74.25	1.62
<b>DESV. EST (<math>\sigma</math>)</b>	$\pm 6.69$	$\pm 14.41$	$\pm 0.08$
<b>DIAGNÓSTICO</b>	Sanos, no fumadores		

No se descartó ningún registro, ya que todos cumplieron con los criterios de inclusión y exclusión. Sin embargo se utilizaron únicamente las fases inspiratorias y se excluyeron las fases que presentaron artefactos de ruido sobresalientes. Si una fase estaba contaminada con uno o más artefactos, se eliminan también las fases correspondientes a los demás canales.

En la figura 1 se muestran 25 canales obtenidos de un sujeto enfermo segmentados en fases inspiratoria y espiratoria en base a la señal de flujo, en el extremo derecho se observa la nomenclatura convenida en la sección 5.1.3.

En la figura 2 se presentan los fononeumogramas correspondientes a un ciclo respiratorio de un sujeto sano (figura 2.a) y un sujeto con NID (figura 2.b) adquiridos en la pared torácica posterior. En la figura se distinguen, a partir de la señal de flujo, las fases inspiratoria y espiratoria de la información acústica. Así también se señala la posición de al menos un estertor crepitante, localizado por inspección visual. En la fase espiratoria no se distinguieron estertores crepitantes que reunieran los criterios de Murphy et al, en ninguno de los dos sujetos.

En la figura 2.c, se presenta la expansión en tiempo de un segmento de la figura 2.b para observar con mayor detalle la morfología clásica de un estertor crepitante el cual se presentó al final de la fase inspiratoria. Las oscilaciones que se observan al inicio de la figura 2.c podrían corresponder a un estertor crepitante.

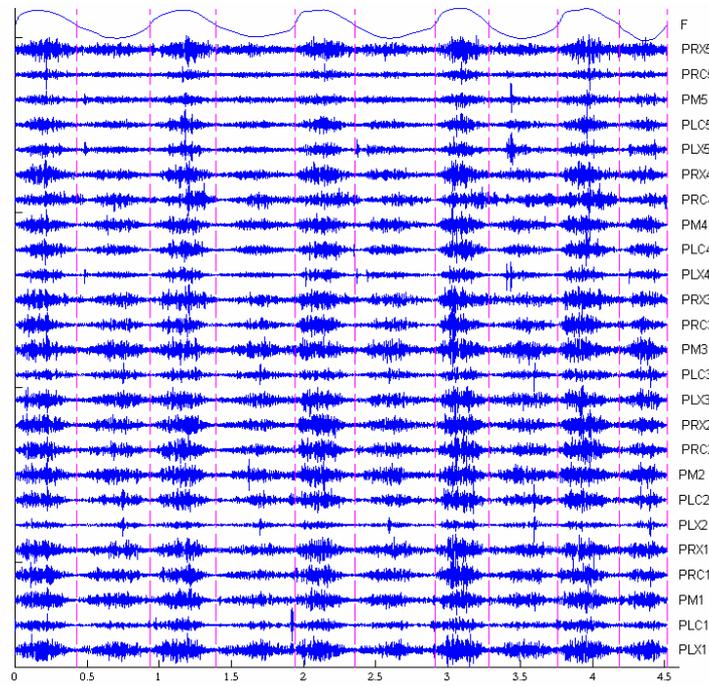


Fig.1 Señales respiratorias multicanales de un sujeto enfermo con NID

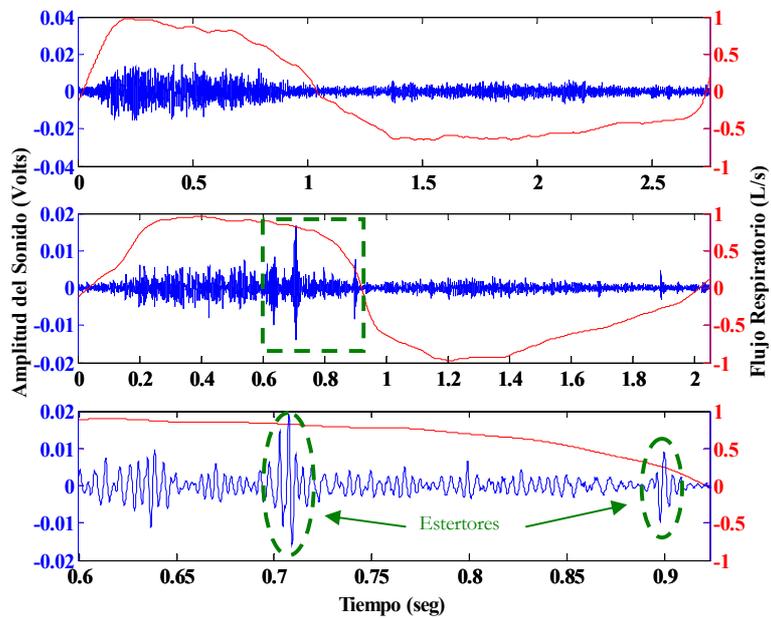


Fig. 2 Fononeumogramas de sonidos respiratorios: a) Sujeto sano, b) Sujeto con NID y c) Segmento de b con estertores crepitantes al final de la inspiración



En la figura 3, se muestra un segmento multicanal de sonidos respiratorios durante la fase de inspiración, con la correspondiente nomenclatura para cada canal. a) En algunos canales se observan ondas que no corresponden a la morfología típica de estertores crepitantes, este tipo de señal fue encontrada también en otros registros de pacientes. b) Se observan sibilancias en varios canales de mayor o menor intensidad. Con esto podemos deducir que la enfermedad de NID no solamente posee estertores crepitantes como sonidos anormales característicos.

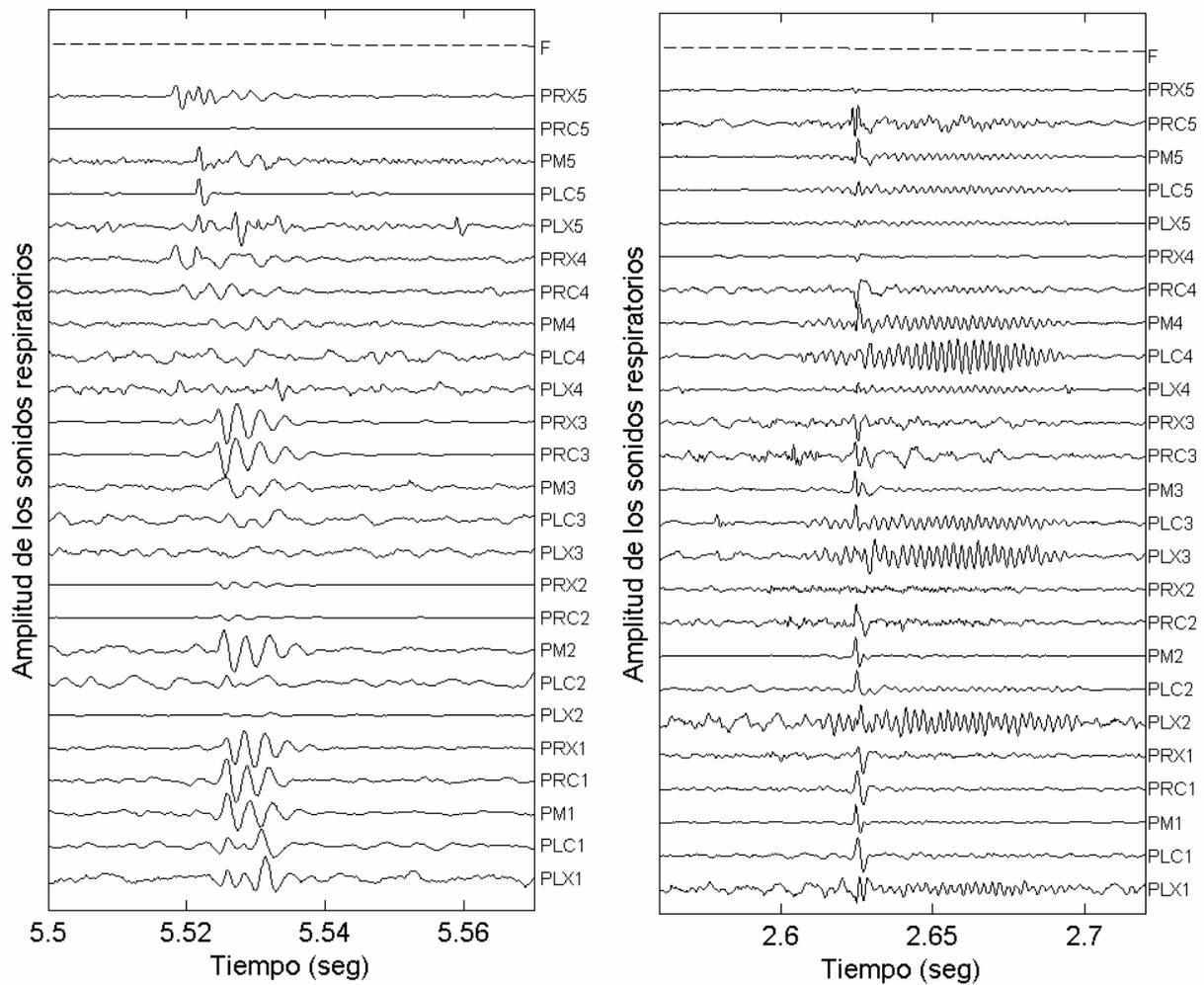


Fig. 3 Segmentos de la señal de sonido respiratorio en pacientes con NID. Se presentan 25 canales de sonido y el canal de Flujo. a) Se observan ondas que no cumplen con la morfología típica, pero que también son sonidos anormales. b) Se encontraron ondas cuya morfología es similar a las sibilancias, pero de mayor duración.

Abajo en el figura 4, se ilustran otros tipos de sonidos anormales encontrados en los registros de los pacientes con NID.

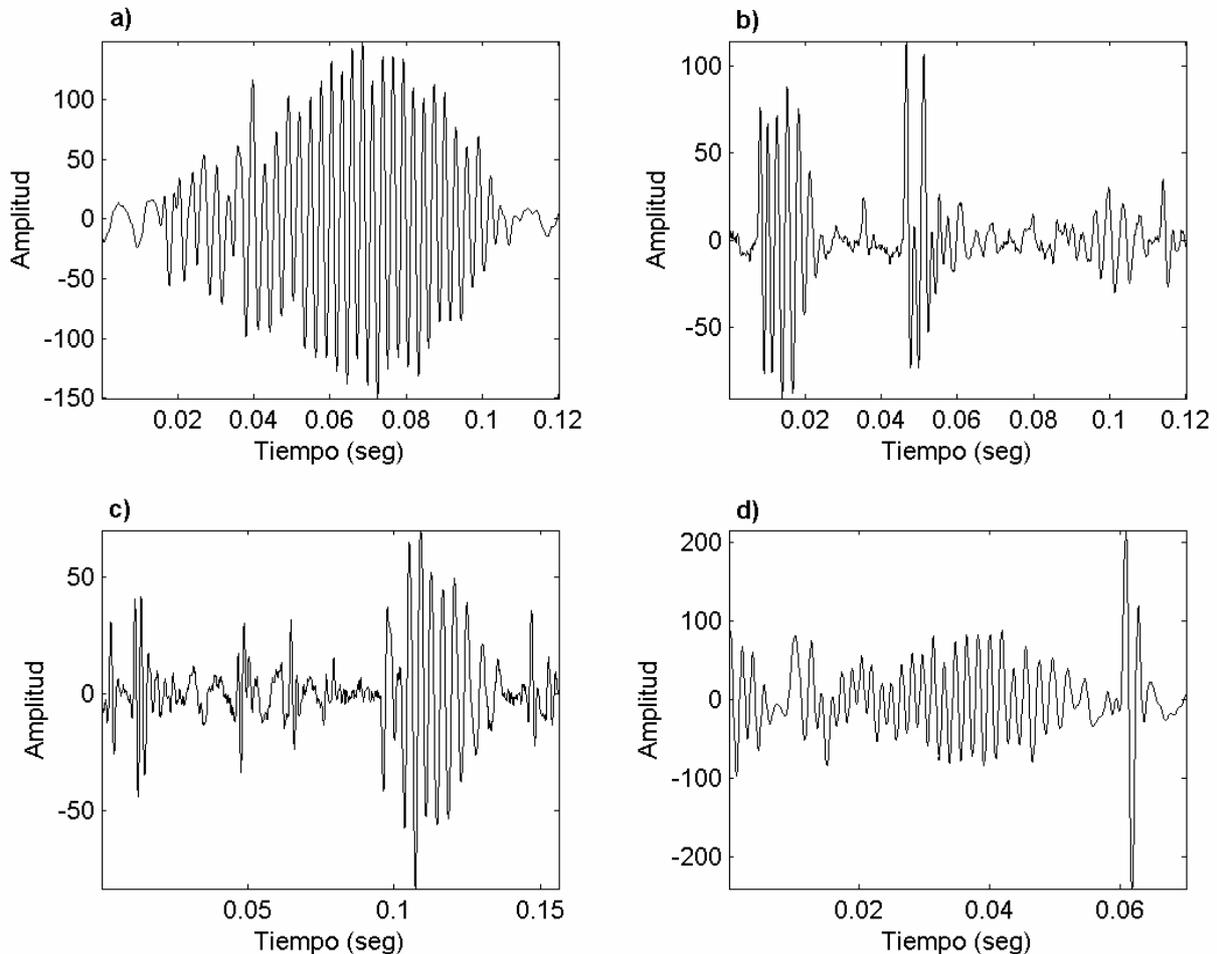


Fig. 4 Sonidos anormales encontrados en pacientes con NID. a) Aparente crepitancia seguido de una sibilancia de menor duración que 40 mseg b) Sonidos que no cumplen con ningún criterio c) Crepitancias y una sibilancia de muy corta duración. d) sibilancia seguida de una crepitancia.

### 6.3 PARAMETRIZACIÓN Y CLASIFICACIÓN DE LAS SEÑALES DE SONIDO RESPIRATORIO

A continuación se presentan los resultados de cada técnica de parametrización y se comparan para determinar la utilidad de cada una de ellas en la extracción de características, analizando tablas de desempeño de la red neuronal que clasifica los vectores de características (VC).

El conjunto de VC fue dividido en tres diferentes subconjuntos, de acuerdo al procedimiento de validación cruzada, para entrenamiento, prueba y validación de la red neuronal. Se introdujo un número igual de VC normales y anormales, etiquetados en base al diagnóstico médico proporcionado. El número de nodos ocultos se varió de 10 a 30. Las redes neuronales fueron entrenadas con 100 épocas en promedio. Sin embargo, el número de épocas fue dependiente de la duración del entrenamiento, que manual o automáticamente se detuvo debido a las razones mencionadas en las secciones 3.3.1 y 3.3.3. Los vectores de características se presentaron aleatoriamente y de modo grupal a la red neuronal. La función de entrenamiento está basada en la regla de Levenberg-Marquadt (TRAINLM) y la función de aprendizaje utilizada fue de descenso por gradiente y momento (LEARNGDM). Se utilizó la herramienta computacional de redes neuronales (NNTOOL) de MATLAB® con la siguiente configuración:

TABLA 3  
Configuración, arquitectura y funciones de entrenamiento de la red neuronal

PARÁMETRO NNTOOL®	CONFIGURACION
Network Type	Feed-forward backprop
Input ranges	Get from input (ENTRENAMIENTO)
Training function	TRAINLM
Adaption learning function	LEARNGDM
Performance function	MSE
Number of layers	3
Layer1	10 - 30
Number of neurons Layer2	5 - 15
Layer3	1
Transfer function	TANSIG

Los parámetros de entrenamiento fueron los mismos para todas las redes neuronales

TABLA 4  
Parámetros de entrenamiento de la red neuronal

PARÁMETRO	VALOR
Épocas	100
Goal	0
Max_fail	5
Mem_reduc	1
Min_grad	1e-010
mu	0.001
mu_dec	0.1
mu_inc	10
mu_max	10000000000
Show	1
Time	inf

Se ha explicado en secciones anteriores, la metodología empleada para el proceso de clasificación. La población total de la base de datos se conformó con 19 pacientes con NID y 8 sujetos sanos, la cual se dividió en dos grandes grupos, el primero para llevar a cabo el entrenamiento, prueba y validación de la red, y el segundo para la evaluación por sujeto.

Para analizar el comportamiento y desempeño de la red neuronal se utilizaron las GRÁFICAS DE ENTRENAMIENTO DE LA RED NEURONAL, los resultados en esta sección se muestran a través de tablas denominadas en este trabajo como TABLAS DE DESEMPEÑO, que resumen los porcentajes obtenidos durante las fases de entrenamiento, prueba y validación de la red neuronal para cada una de las técnicas de parametrización. Los resultados obtenidos en la etapa de evaluación por sujeto, considerando datos no “vistos” por la red neuronal, se resumen en las TABLAS DE VALIDACIÓN GENERAL, igualmente para cada una de las técnicas de parametrización. A continuación se describe la nomenclatura que se utiliza para la mejor comprensión de los resultados.

## GRÁFICAS DE ENTRENAMIENTO DE LA RED NEURONAL

En las siguientes figuras se muestra la gráfica de desempeño de una red neuronal durante el entrenamiento. Se utilizaron al mismo tiempo el subconjunto de prueba y de validación para probar la generalización de la red y para detener el proceso de entrenamiento (early stopping) respectivamente. En la presente tesis se entrenaron y probaron diferentes configuraciones de redes neuronales hasta encontrar la arquitectura óptima que mejor se desempeñe de acuerdo con los índices de sensibilidad, especificidad y desempeño.

En la figura 5 se grafican a manera de ejemplo, las tres curvas del error cuadrático medio (MSE) para los subconjuntos de entrenamiento, prueba y validación, obsérvese que el error decrece aproximadamente de igual forma en los tres subconjuntos.

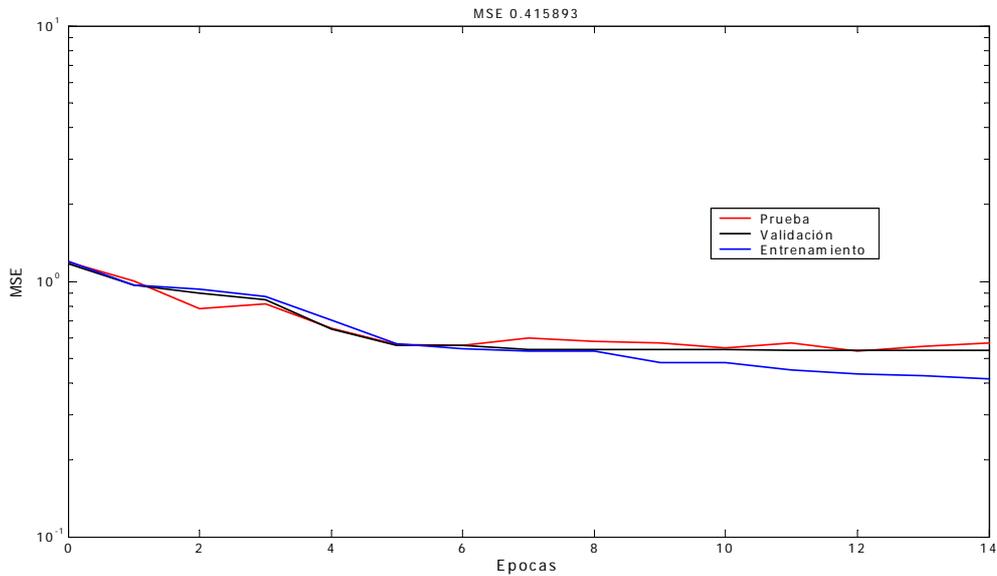


Fig 5 Gráfica del error cuadrático medio (MSE) en las curvas de entrenamiento, prueba y validación.

En la figura 6, durante las primeras ocho épocas el error decrece en la misma magnitud, después la curva del subconjunto de entrenamiento sigue descendiendo mientras que en las dos restantes, el error no disminuye y por el contrario, tiende a subir y el entrenamiento se detiene.

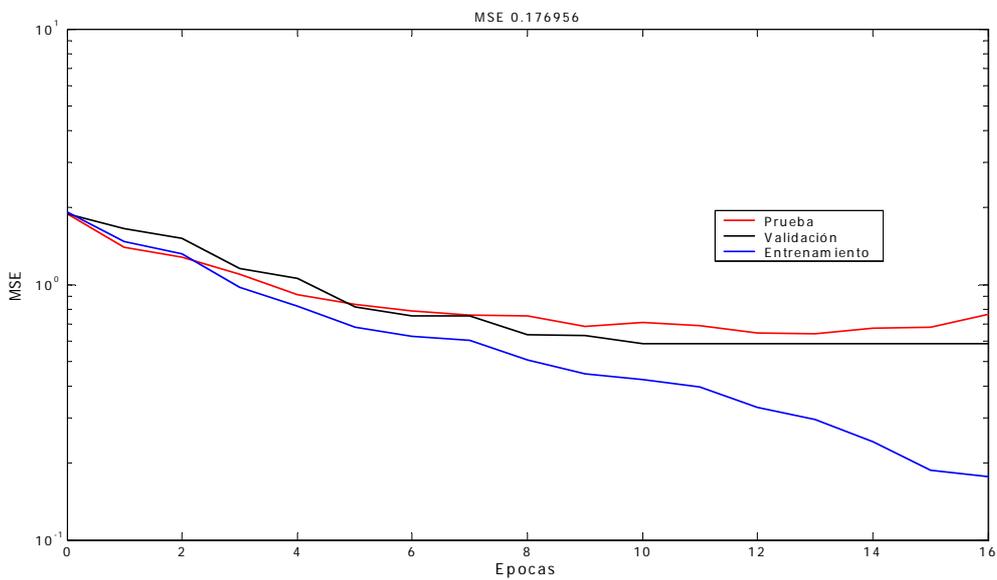


Fig. 6 Gráfica del error cuadrático medio (MSE) en las curvas de entrenamiento, prueba y validación mostrando early stopping.

## TABLAS DE DESEMPEÑO

En las Tablas de Desempeño se muestra el número de vectores de características (VC) de entrenamiento, prueba y validación que fueron aplicados como datos de entrada a la red neuronal. **N-N** es el número de VC normales correctamente clasificados como normales. **N-A** es el número de VC normales incorrectamente clasificados como anormales o enfermos (Falsos negativos). **A-A** corresponde al número de VC anormales correctamente clasificados como anormales. **A-N** indica el número de VC anormales incorrectamente clasificados como normales (Falsos positivos) y **NC** se refiere al número de VC que no fueron clasificados como normales y tampoco como anormales.

En la figura 7 se muestra la salida de la red neuronal ya entrenada que se simula con nuevos datos para valorar la generalización de la red. Debido a que la red fue entrenada con una función de transferencia tangente sigmoidea con valores máximos de +1 y -1, los valores esperados a la salida deben ser +1 salida -1, y se consideró un umbral de +0.8 y -0.8, de esta manera todos los vectores cuya salida corresponda a  $\geq +0.8$  se consideraron normales y aquellos vectores cuya salida sea  $\leq -0.8$  se consideraron anormales. Estos valores de umbral han sido propuestos en la literatura y son considerados como un índice de confiabilidad. En muchas ocasiones la salida se encuentra entre los valores de los umbrales, es decir no se consideran normales ni anormales, y son descartados para el porcentaje de clasificación (%NC).

También se muestran los índices de sensibilidad (SEN%), especificidad (ESP%) y desempeño.

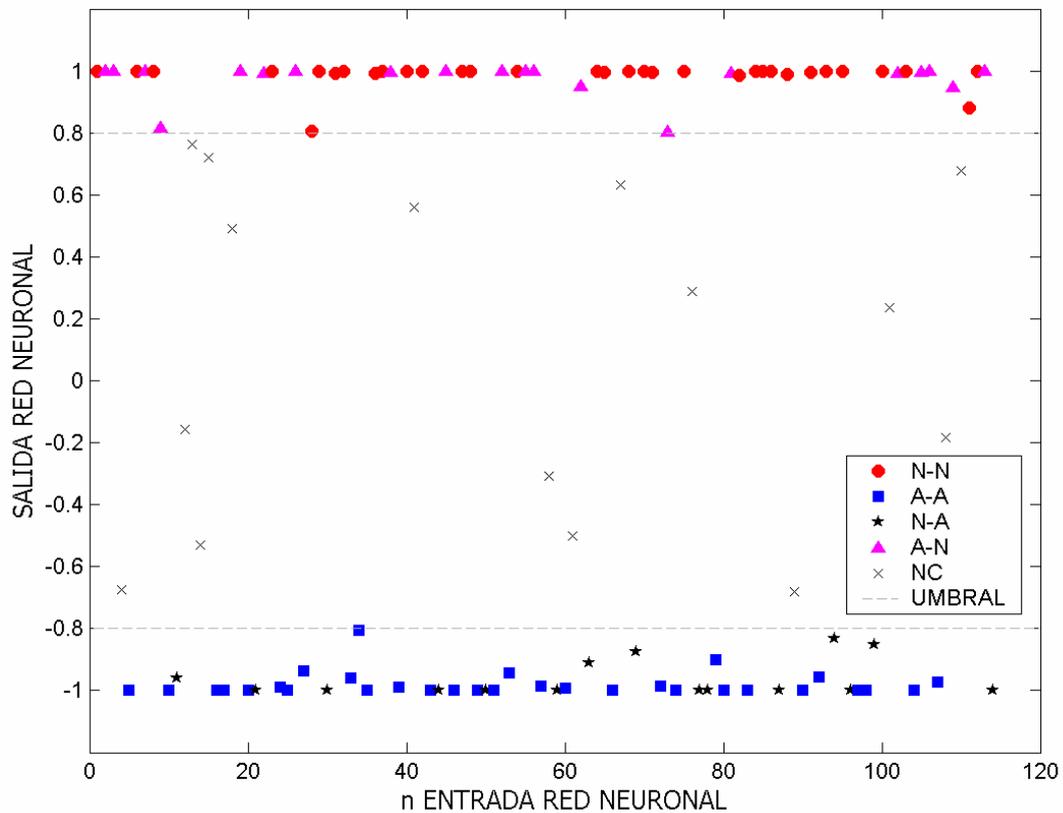


Fig. 7 Gráfica de la salida de una red neuronal tras la simulación con un conjunto de vectores de características.

**VC:** Vectores de características; **N-N:** Número de clasificaciones correctas para VC normales; **A-A:** Número de clasificaciones correctas para VC anormales; **N-A:** Número de VC normales clasificados como anormales; **A-N:** Número de VC anormales clasificados como normales; **NC:** Número de VC que no fueron clasificados como normales y tampoco como anormales; **UMBRAL:** Valores de umbral de  $\pm 8$  para considerar a los VC como normales o anormales.

## TABLAS DE VALIDACIÓN GENERAL

Al presentar los resultados por cada técnica de parametrización, se muestran dos Tablas de Validación General, una correspondiente a los sujetos sanos y otra para los sujetos enfermos y se resumen los porcentajes de clasificación para cada uno de los sujetos que fueron utilizados en la primera fase de entrenamiento, prueba y validación y también se incluyen sujetos que nunca han sido “vistos” por la red, la cual es simulada con el conjunto de vectores de

características obtenidos del registro del sujeto en cuestión. Calificando con +1 ó -1 (umbral +0.8 ó -0.8) si son clasificados como normal o anormal respectivamente. El porcentaje es calculado aplicando la regla de la mayoría al vector de datos de salida de la red y se toma la decisión de clasificar al sujeto de acuerdo al número de vectores considerados como normales (%NORMAL) o anormales (%ANORMAL). Cada tabla está acompañada por una gráfica de barras (**Gráfica de Validación General**) para visualizar los porcentajes de vectores normales (%NORMAL), anormales (%ANORMAL) y vectores no clasificados (%NC) por la red neuronal.

### 6.3.1 RESULTADOS UTILIZANDO LOS PARÁMETROS DE EIGENVALORES DE LA MATRIZ DE COVARIANZA

A continuación se presentan los resultados obtenidos con la técnica de eigenvalores de la matriz de covarianza de la señal respiratoria.

a) El vector de características consistió de 150 eigenvalores. La red neuronal que mejor desempeño mostró se construyó con 20 neuronas de entrada, 10 neuronas ocultas y 1 neurona de salida. El error alcanzado por la red entrenada durante 18 épocas fue de 0.2402, evaluado mediante la función del error cuadrático medio (MSE). En la Tabla 5 se resumen los datos de clasificación con esta técnica.

Tabla 5 Desempeño de la Red Neuronal  
Desempeño de la Red neuronal para la clasificación de vectores parametrizados mediante eigenvalores de la matriz de covarianza

VC(150X1)	TAMAÑO	N-N	A-A	N-A	A-N	N.C	%SENS	%ESP	%DES
Entrenamiento	864	354	146	0	10	354	97.25	100	98.63
Prueba	108	48	19	1	4	36	92.31	95	93.65
Validación	108	41	20	0	1	46	97.62	100	98.81

Tabla 6 Validación General

Validación general de cada sujeto sano por la técnica de eigenvalores de la matriz de covarianza de la señal respiratoria multicanal

SUJETO SANO	%NORMAL	%ANORMAL
AAC	46.67	0.28
BTC	96.56	0.00
CMR	71.11	1.94
GMR	82.22	0.89
LMDL	53.42	0.50
LO	24.80	6.13
MVCC	64.33	0.44
PRC	5.09	4.09
<b>TOTAL</b>	<b>55.52</b>	<b>1.78</b>

Validación por sujeto

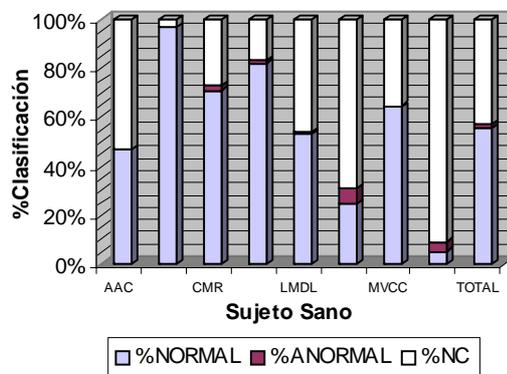


Fig 8 Gráfica de Validación general de sujetos sanos por la técnica de eigenvalores de la matriz de covarianza de la señal respiratoria.

Tabla 7 Validación General

Desempeño de la Red neuronal para la clasificación mediante eigenvalores de la matriz de covarianza

SUJETO	%NORMAL	%ANORMAL
AMA	0.00	35.33
ASR	9.33	30.67
CC	3.89	26.02
CVG	10.50	24.33
EFC	8.19	20.97
EM	0.28	28.33
EMG	12.40	9.33
EV	10.93	11.78
FVM	2.33	22.56
GAG	7.64	14.72
ICG	7.89	23.44
JGM	18.80	6.30
JHL	11.28	6.06
JVG	13.33	7.78
LHM	16.52	4.82
MFC	24.92	6.94
MLB	28.50	14.72
MPM	60.89	5.33
TEA	16.94	17.67
<b>TOTAL</b>	<b>13.92</b>	<b>16.69</b>

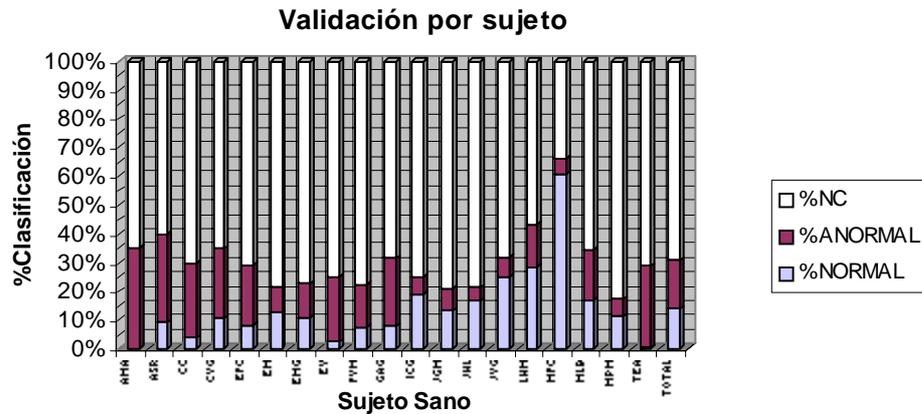


Fig 9 Gráfica de Validación general de sujetos Enfermos por la técnica de eigenvalores de la matriz de covarianza de la señal respiratoria.

b) Se consideraron 10 eigenvalores por ventana, cada fase respiratoria se dividió en 30 ventanas. Concatenando los parámetros de los 25 canales, el vector de características que representa a una ventana multicanal se formó de 250 eigenvalores.

La red neuronal que mejor desempeño mostró fue con 20 neuronas de entrada, 10 neuronas ocultas y 1 neurona de salida. El error alcanzado de la red entrenada durante 100 épocas fue de 0.0679.

Tabla 8 Desempeño de la Red Neuronal para la clasificación de vectores parametrizados mediante eigenvalores de la matriz de covarianza

VC(250X1)	TAMAÑO	N-N	A-A	N-A	A-N	N.C	%SEN	%ESP	%DES
Entrenamiento	864	402	410	3	1	48	99.75	99.27	99.51
Prueba	108	37	43	13	9	6	80.43	76.79	78.61
Validación	108	38	45	10	8	7	82.61	81.82	82.21

Tabla 9 Validación General  
Técnica de eigenvalores de la matriz de covarianza de la señal respiratoria multicanal

SUJETO	%NORMAL	%ANORMAL
AAC	71.25	21.81
BTC	80.56	12.28
CMR	66.11	24.44
GMR	75.56	21.33
LMDL	54.96	41.08
LO	49.80	39.50
MVCC	59.00	36.22
PRC	38.20	54.69
<b>TOTAL</b>	<b>61.93</b>	<b>31.42</b>

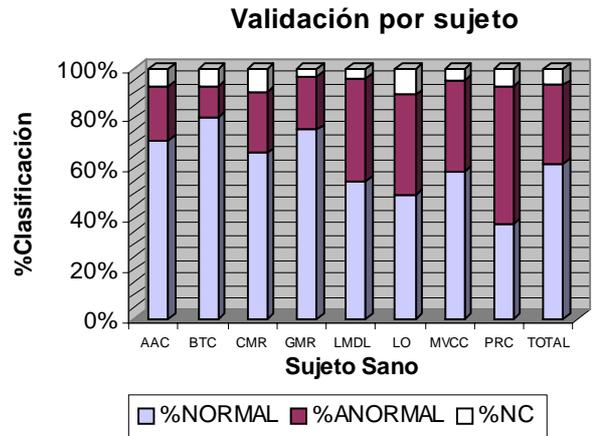


Fig. 10 Gráfica de Validación general de sujetos sanos por la técnica de eigenvalores de la matriz de covarianza de la señal respiratoria.

Tabla 10 Validación General  
Clasificación mediante eigenvalores de la matriz de covarianza

SUJETO	%NORMAL	%ANORMAL
AMA	13.56	81.11
ASR	13.58	76.67
CC	14.91	79.35
CVG	18.00	78.83
EFC	16.94	76.39
EMG1	26.53	66.67
EM	7.17	89.39
EV1	21.26	73.33
FVM1	13.39	80.39
GAG1	24.44	68.89
ICG1	22.65	68.94
JGM1	17.69	76.67
JHL	22.42	69.86
JVG2	41.67	50.00
LHM1	29.29	63.75
MFC	36.41	57.28
MLB	41.00	50.56
MPM	54.89	38.67
TEA	24.67	71.72
<b>TOTAL</b>	<b>24.23</b>	<b>69.39</b>

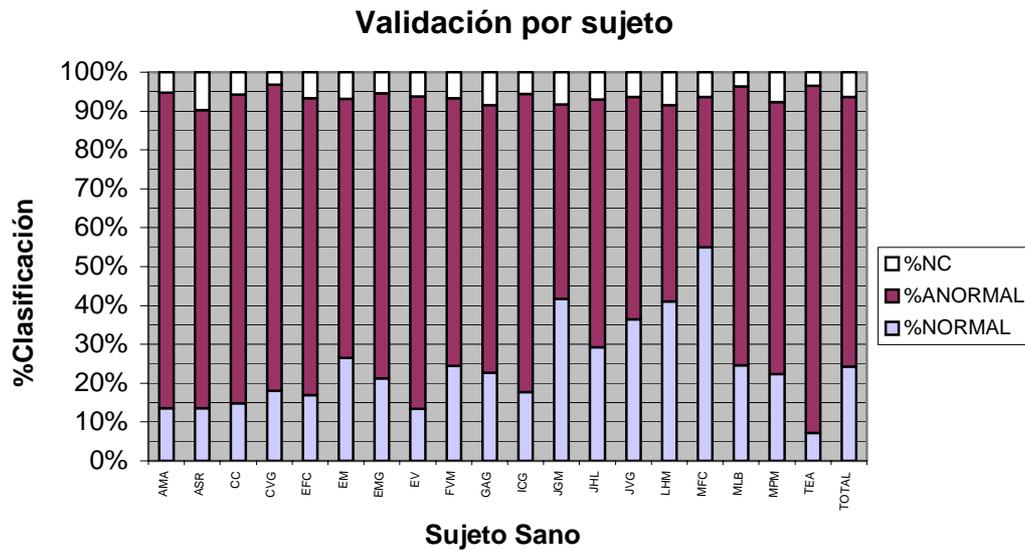


Fig. 11 Gráfica de Validación general de sujetos sanos por la técnica de eigenvalores de la matriz de covarianza de la señal respiratoria.

### 6.3.2 RESULTADOS UTILIZANDO PARÁMETROS EN EL DOMINIO DE LA FRECUENCIA

Se realizaron dos pruebas con los parámetros en frecuencia. Se calcularon 6 parámetros de frecuencia correspondientes al 25, 50, 75, 90, 95 y 99 por ciento del área total bajo la curva de potencia espectral calculada mediante el modelo AR y el método de Burg.

a) Se utilizaron los parámetros obtenidos de 4 sujetos sanos y 4 sujetos enfermos para entrenamiento, prueba y validación de la red de igual manera que se realizó en las demás técnicas. El conjunto total consistió de 38 vectores de características de 150 parámetros de frecuencia. Se entrenaron varias redes neuronales, y la arquitectura de la red óptima se eligió mediante prueba y error. La red neuronal que alcanzó el mínimo error de  $1.36 \times 10^{-5}$  consistió de 20 neuronas de entrada, 10 neuronas ocultas y una neurona de salida entrenada durante 100 épocas. Los resultados se muestran a continuación:

Tabla 11 Desempeño de la Red Neuronal  
Clasificación de mediante parámetros clásicos de frecuencia

VC (150X1)	TAMAÑO	N-N	A-A	N-A	A-N	N.C	SENS	% ESP	% DES%
Entrenamiento	30	15	15	0	0	0	100	100	100
Prueba	4	2	2	0	0	0	100	100	100
Validación	4	2	2	0	0	0	100	100	100

Tabla 12 Validación General  
Técnica de parámetros clásicos en la frecuencia

SUJETO	%NORMAL	%ANORMAL
AAC	100	0
BTC	100	0
CMR	100	0
GMR	100	0
LMDL	100	0
LO	40	40
MVCC	100	0
PRC	0	100
<b>TOTAL</b>	<b>60</b>	<b>35</b>

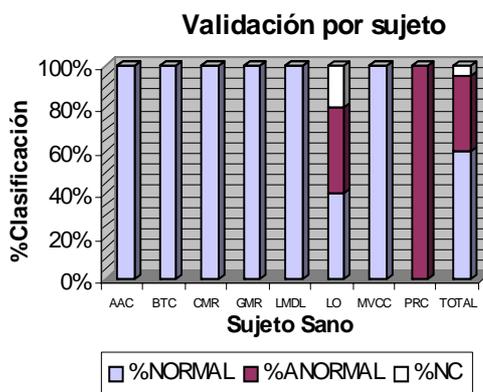


Fig. 12 Gráfica de Validación general de sujetos sanos por la técnica de parámetros clásicos en la frecuencia

Tabla 13 Validación General  
Clasificación mediante parámetros clásicos de frecuencia

SUJETO	%NORMAL	%ANORMAL
AMA	0	100
ASR	0	100
CC	0	100
CVG	0	100
EFC	0	100
EM	0	100
EMG	0	100
EV	0	100
FVM	0	100
GAG	0	75
ICG	0	75
JGM	0	66.67
JHL	0	100
JVG	80	0
LHM	0	100
MFC	0	28.57
MLB	0	50
MPM	20	0
TEA	0	100
<b>TOTAL</b>	<b>5.26</b>	<b>78.70</b>

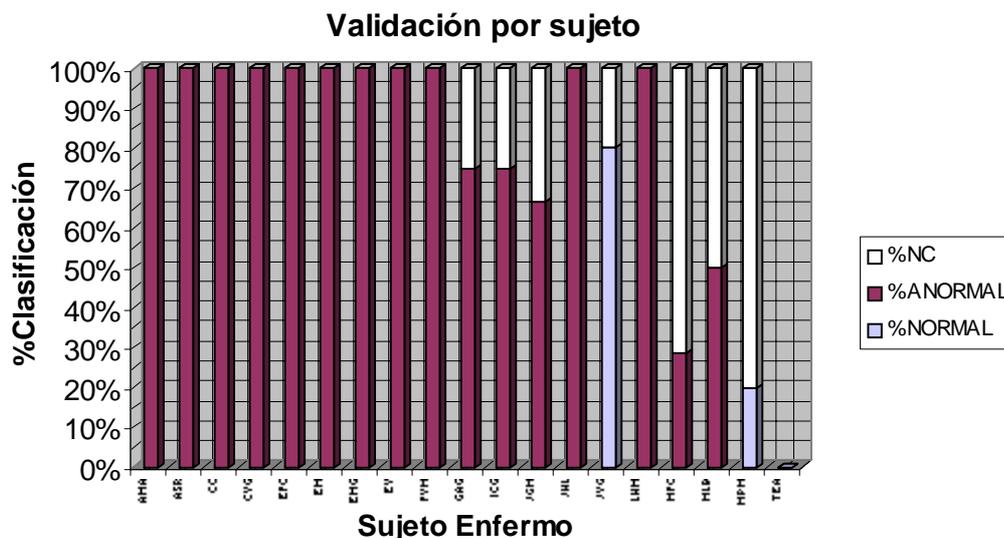


Fig. 13 Gráfica de Validación general de sujetos sanos por la técnica de parámetros clásicas de frecuencia

b) Se obtuvieron los parámetros de los 19 sujetos enfermos y 8 sujetos sanos. Se obtuvieron 208 vectores de características de 150 parámetros de frecuencia.

La red neuronal que mejor desempeño demostró consistió de 20 neuronas de entrada, 10 neuronas ocultas y 1 neurona de salida.

El error alcanzado por la red neuronal entrenada durante 100 épocas, según la función del error cuadrático medio fue de 0.00023.

Tabla 14 Desempeño de la Red Neuronal  
Desempeño de la Red neuronal para la clasificación mediante parámetros clásicos de frecuencia

V.C (150X1)	TAMAÑO	N-N	A-A	N-A	A-N	N.C	SEN%	ESP%	DES%
Entrenamiento	166	83	83	0	0	0	100	100	100
Prueba	20	9	9	1	1	0	90	90	90
Validación	22	11	11	0	0	0	100	100	100

Tabla 15 Validación General  
Clasificación mediante  
parámetros clásicos de frecuencia

SUJETO	%NORMAL	%ANORMAL
AAC	100.00	0.00
BTC	100.00	0.00
CMR	100.00	0.00
GMR	100.00	0.00
LMDL	100.00	0.00
LO	94.44	5.56
MVCC	100.00	0.00
PRC	81.67	0.00
<b>TOTAL</b>	<b>97.01</b>	<b>0.69</b>

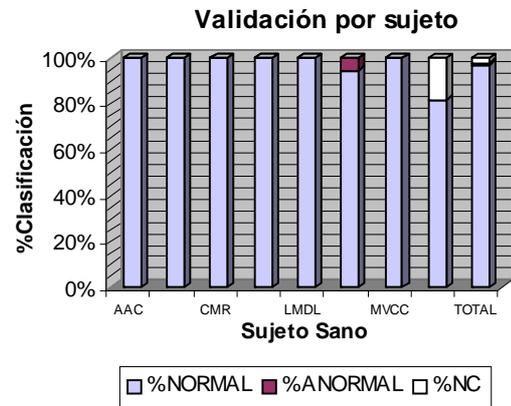


Fig. 14 Gráfica de Validación general de sujetos sanos por la técnica de parámetros clásicos de frecuencia

Tabla 16 Validación General  
Clasificación mediante parámetros clásicos de frecuencia

SUJETO	%NORMAL	%ANORMAL
AMA	0	100
ASR	0	100
ASR	0	100
CC	0	100
CVG	0	100
CVG	0	100
EFC	0	100
EM	0	100
EMG	0	100
EV	0	100
GAG	0	100
ICG	25	75
JGM	0	100
JHL	0	100
JVG	0	100
LHM	0	100
MFC	0	100
MLB	0	100
MPM	0	100
TEA	0	100
<b>TOTAL</b>	<b>1.25</b>	<b>98.75</b>

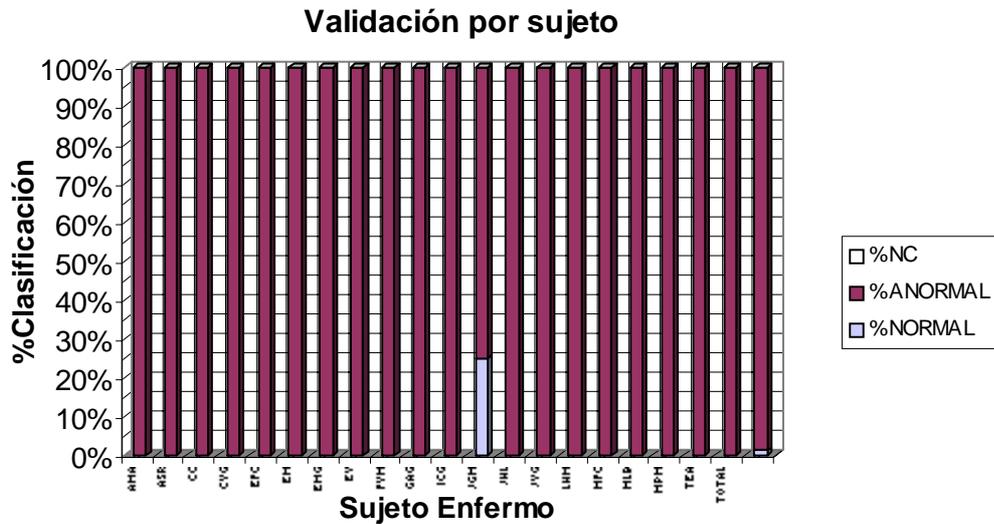


Fig. 15 Gráfica de Validación general de sujetos sanos por la técnica parámetros clásicos de frecuencia

### 6.3.3 RESULTADOS APLICANDO EL MODELO AR MONOCANAL (ARU)

Se seleccionaron aleatoriamente 4 registros de sujetos sanos y 4 registros de sujetos enfermos, el número de canales es igual a 25 y las fases inspiratorias de cada registro y por cada canal, fueron divididas en 30 ventanas de tiempo traslapadas 25%. Se obtuvieron en total 1140 ventanas de tiempo, de longitud variable, aproximadamente de 200 muestras.

Se estimaron los coeficientes AR mediante la función ARX de Matlab®, se seleccionó un orden 4 para el modelo AR monocanal y se evaluó la bondad del ajuste asegurando que los valores se encontraran dentro del 95% del intervalo de confianza.

El conjunto de vectores de características resultó de 1140 vectores con 100 coeficientes AR, el cual fue dividido en tres subconjuntos para entrenamiento, prueba y validación de la red neuronal.

La arquitectura de red neuronal que mostró mejor desempeño durante la clasificación

utilizando la parametrización AR monocal fue una red con 20 neuronas en la capa de entrada, 10 neuronas en la capa oculta y una neurona de salida. La red se entrenó con 25 épocas y se evaluó mediante la función del MSE con error de entrenamiento de  $9.97e-17$ .

Tabla 17 Desempeño de la Red Neuronal  
Desempeño de la Red neuronal para la clasificación de vectores parametrizados mediante el modelo AR monocal

VC (100X1)	TAMAÑO	N-N	A-A	N-A	A-N	N.C	SEN %	ESP%	DES%
Entrenamiento	912	456	456	0	0	0	100	100	100
Prueba	114	53	55	3	1	2	98.15	94.83	96.49
Validación	114	53	56	3	1	1	98.15	94.92	96.53

Tabla 18 Validación General  
Validación general de cada sujeto sano por la técnica de modelo AR monocal de la señal respiratoria

SUJETO	%NORMAL	%ANORMAL
AAC	97.5	1.67
BTC	100.00	0.00
CMR	99.44	0.56
GMR	94.67	4.67
LMDL	47.33	46.67
LO	60.83	35.83
MVCC	99.17	0.00
PRC	86.11	10.55
<b>TOTAL</b>	<b>73.66</b>	<b>23.26</b>

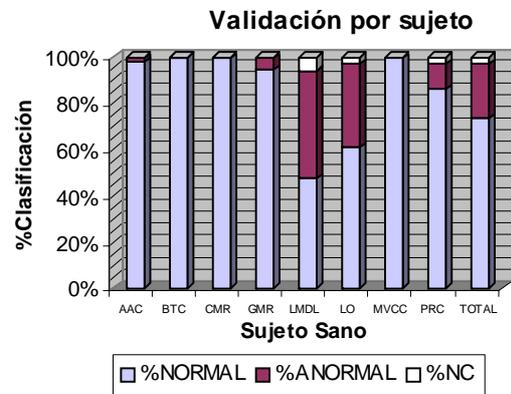


Fig. 16 Gráfica de Validación general de sujetos sanos por la técnica AR monocal de la matriz de covarianza de la señal respiratoria.

Tabla 19 Validación General  
Validación general de cada sujeto enfermo por la técnica de modelo AR monocanal de la señal respiratoria

SUJETO	%NORMAL	%ANORMAL
AMA	0.00	100.00
ASR	1.33	98.67
CC	0.00	100.00
CVG	0.00	99.33
EFC	4.44	90.56
EM	1.33	98.00
EMG	14.00	78.67
EV	1.11	98.89
FVM	12.78	83.89
GAG	0.83	97.5
ICG	2.50	97.5
JGM	2.22	97.78
JHL	26.67	67.33
JVG	14.00	82.00
LHM	8.57	90.00
MFC	18.10	79.05
MLB	51.11	45.00
MPM	12.67	85.33
TEA	8.33	90.00
<b>TOTAL</b>	<b>11.91</b>	<b>85.43</b>

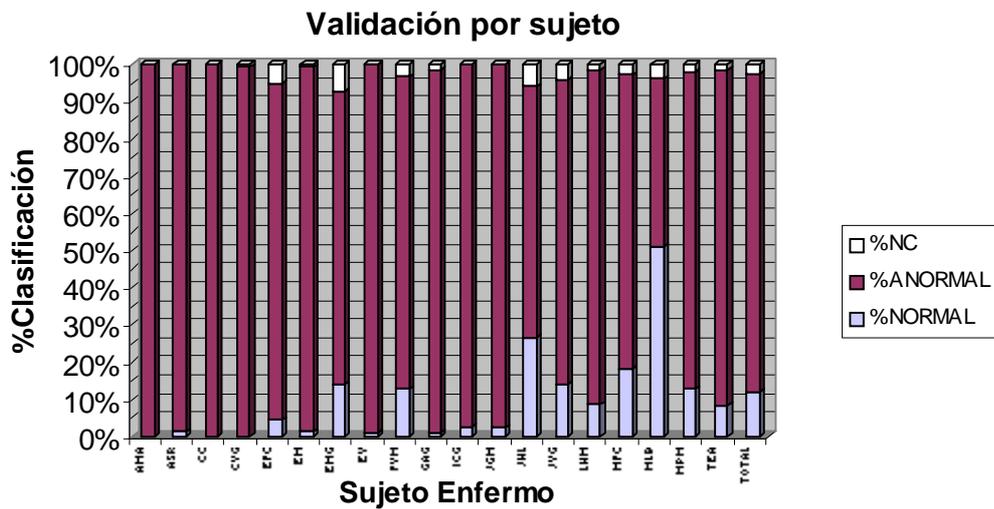


Fig. 17 Gráfica de Validación general de sujetos sanos por la técnica de modelo AR monocanal de la señal respiratoria.

### 6.3.4 RESULTADOS APLICANDO EL MODELO AR MULTICANAL (ARM)

Cada ventana multicanal de sonido respiratorio fue representada por las matrices  $A(1)$  hasta  $A(4)$ . El tamaño de la matriz  $A$  resultó ser de  $25 \times 25$ .

En la figura 18 se presentan las magnitudes de las matrices  $A(1)$  hasta  $A(4)$ , en forma de imágenes, para observar los cambios que contienen. En la figura 18(a) se muestran las magnitudes de matrices de coeficientes AR para un segmento de información acústica proveniente de la fase inspiratoria de un sujeto sano mientras que las gráficas en la figura 18(b) corresponden a un sujeto con NID. La magnitud de las superficies en ambos casos tiende a disminuir conforme se va de  $A(1)$  hasta  $A(4)$ , por lo que el orden del modelo se considera adecuado.

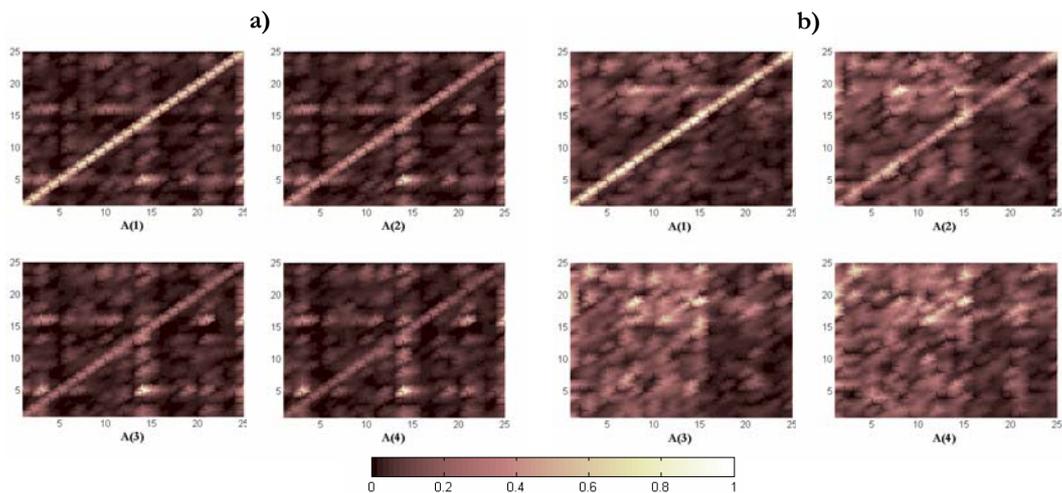


Fig. 18 Magnitud de los coeficientes de la matriz  $A(1)$  hasta  $A(4)$ : a) Sonido respiratorio normal, b) Sonido respiratorio anormal.

La reducción de la dimensión del vector de características se realizó de dos formas diferentes:

#### 6.3.4.1 MATRIZ INDIVIDUAL DE COEFICIENTES DEL MODELO AR MULTICANAL

##### 6.3.4.1.1 Reducción de Características por Descomposición en Valores Singulares (SVD) de la matriz de coeficientes AR

Se calcularon 25 valores singulares por cada matriz A, que fueron concatenados para obtener un vector de características de 100 coeficientes por cada ventana multicanal.

Diferentes configuraciones de redes neuronales fueron entrenadas con los vectores de características de tamaño 100x1 haciendo un conjunto total de 1140 vectores de características. La arquitectura de la red neuronal que mejor desempeño mostró para los datos reducidos por SVD durante la etapa de validación fue con 20 nodos de entrada, 10 nodos ocultos y un nodo de salida, evaluada según la función del error cuadrático medio que indicó un error de 1.79e-02.

Tabla 20 Desempeño de la Red Neuronal ARM en conjunto con SVD

VC(100X1)	Tamaño	N-N	A-A	N-A	A-N	N.C	SEN%	ESP%	DES%
Entrenamiento	912	454	452	1	3	2	99.34	99.78	99.56
Prueba	114	43	49	12	7	3	86	80.33	83.16
Validación	114	41	44	16	10	3	80.39	73.33	76.86

Tabla 21 Validación General  
Clasificación de vectores parametrizados mediante modelo ARM y SVD

SUJETO	%NORMAL	%ANORMAL
AAC	73.33	22.78
BTC	56.83	42.33
CMR	87.50	11.67
GMR	73.78	22.44
LMDL	56.17	39.33
LO	46.48	48.06
MVCC	19.00	76.89
PRC	32.94	63.92
<b>TOTAL</b>	<b>55.75</b>	<b>40.93</b>

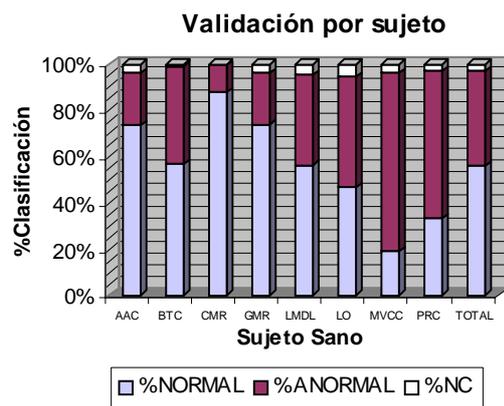


Fig. 19 Gráfica de Validación general de sujetos sanos por la técnica ARM

Tabla 22 Validación General  
para la clasificación de vectores parametrizados mediante ARM en conjunción con SVD

SUJETO	%NORMAL	%ANORMAL
AMA	2.67	97.33
ASR	10.94	87.44
CC	2.50	96.67
CVG	4.50	93.83
EFC	85.00	13.33
EM	32.67	62.67
EMG	68.67	28.00
EV	5.56	90.00
FVM	48.89	48.89
GAG	5.00	90.83
ICG	59.17	37.50
JGM	0.00	100.00
JHL	97.33	1.33
JVG	12.67	84.00
LHM	99.05	0.95
MFC	37.62	59.52
MLB	65.56	30.56
MPM	100.00	0.00
TEA	95.00	5.00
<b>TOTAL</b>	<b>43.83</b>	<b>54.10</b>

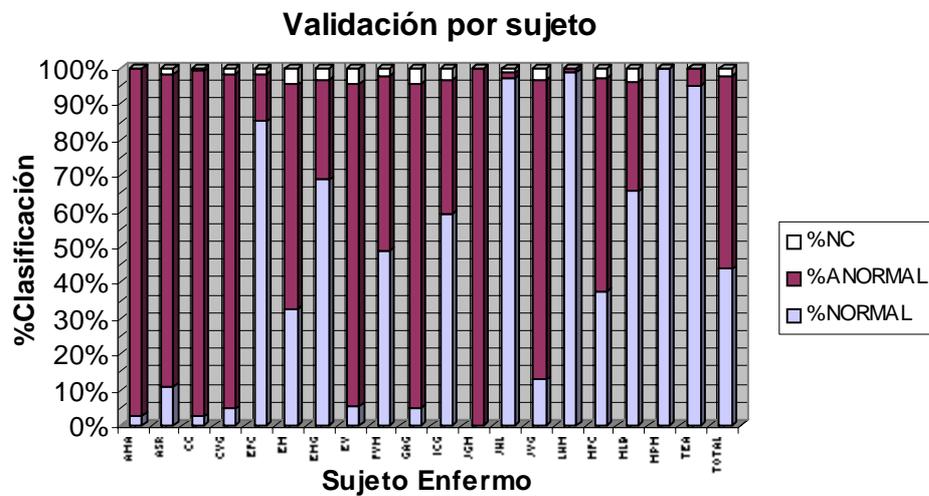


Fig. 20 Gráfica de Validación general de sujetos enfermos por la técnica ARM

### 6.3.4.1.2 Reducción de Características mediante la obtención de los eigenvalores de la matriz de covarianza de la matriz de coeficientes AR

a) Se tomaron los 25 eigenvalores más importantes de cada matriz de covarianza de A y se concatenaron en un vector de 100 eigenvalores. El conjunto total de vectores de características se formó de 1140 vectores.

Se utilizó el método de validación cruzada dividiendo el conjunto total en entrenamiento, prueba y validación de la red neuronal.

La red neuronal que mejor desempeño mostró fue con 25 neuronas de entrada, 10 neuronas ocultas y una neurona de salida. El error de entrenamiento es de 0.0254 evaluado con la función del error cuadrático medio. Cada vector de características estuvo conformado por 100 parámetros.

Tabla 23 Desempeño de la Red Neuronal  
ARM en conjunción con eigenvalores de la matriz de covarianza de la matriz A

VC (100X1)	TAMAÑO	N-N	A-A	N-A	A-N	N.C	SEN %	ESP%	DES%
Entrenamiento	912	444	446	4	1	17	99.78	99.11	99.44
Prueba	114	42	40	10	12	10	77.78	80	78.89
Validación	114	41	42	7	8	16	83.67	85.71	84.69

Tabla 24 Validación General  
Técnica ARM en conjunción con eigenvalores de la matriz de cov. de la matriz A

SUJETO	%NORMAL	%ANORMAL
AAC	63.19	23.89
BTC	62.72	27.83
CMR	87.50	10.28
GMR	67.78	16.22
LMDL	51.54	32.00
LO	50.56	36.24
MVCC	18.06	66.67
PRC	26.44	62.67
<b>TOTAL</b>	<b>53.47</b>	<b>34.47</b>

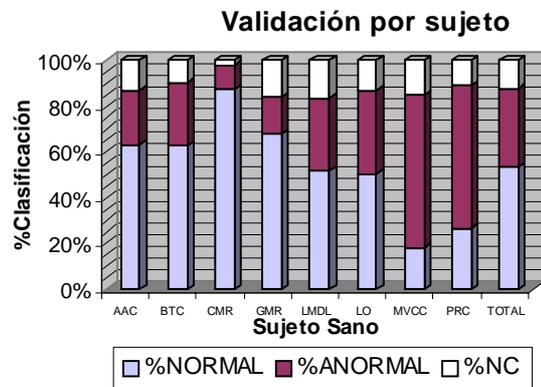


Fig. 21 Gráfica de Validación general de sujetos sanos por la técnica ARM en conjunción con eigenvalores de la matriz de cov. de la matriz A.

Tabla 25 Validación General  
Técnica ARM en conjunción con eigenvalores de la matriz de cov. de la matriz A

SUJETO	%NORMAL	%ANORMAL
AMA	2.00	92.67
ASR	11.17	81.89
CC	5.00	93.33
CVG	3.00	94.56
EFC	83.89	8.89
EM	20.67	62.00
EMG	60.67	26.67
EV	3.33	88.89
FVM	52.22	38.33
GAG	5.00	90.83
ICG	44.17	43.33
JGM	0.00	100.00
JHL	100.00	0.00
JVG	13.33	76.67
LHM	100.00	0.00
MFC	32.86	54.76
MLB	63.33	22.22
MPM	100.00	0.00
TEA	95.56	0.56
<b>TOTAL</b>	<b>41.90</b>	<b>51.35</b>

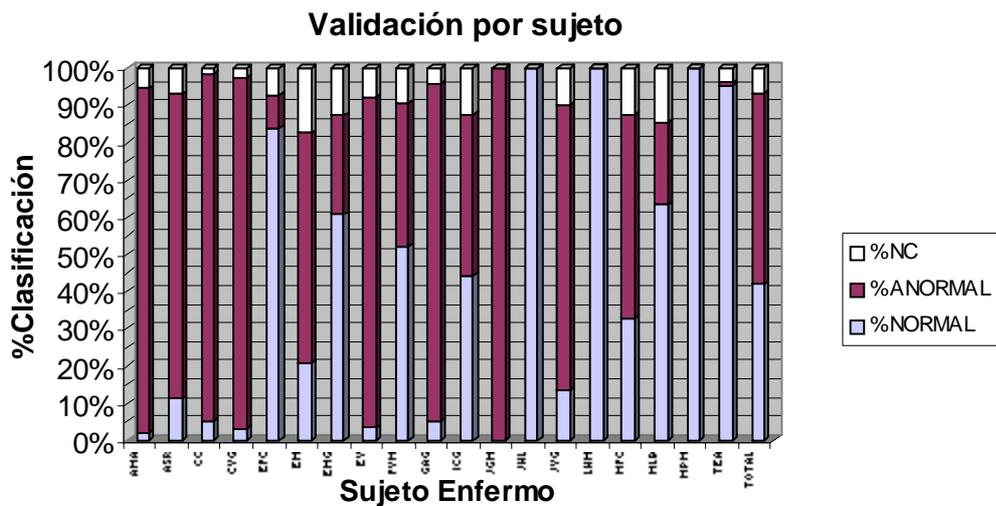


Fig. 22 Gráfica de Validación general de sujetos enfermos por la técnica ARM en conjunción con eigenvalores de la matriz de cov. de la matriz A

En la figura 23 se observan los eigenvalores de la matriz de covarianza de las matrices de coeficientes A1 hasta A4. Se tomaron los 25 eigenvalores más importantes de cada matriz de covarianza de A y se concatenaron en un vector de 100X1.

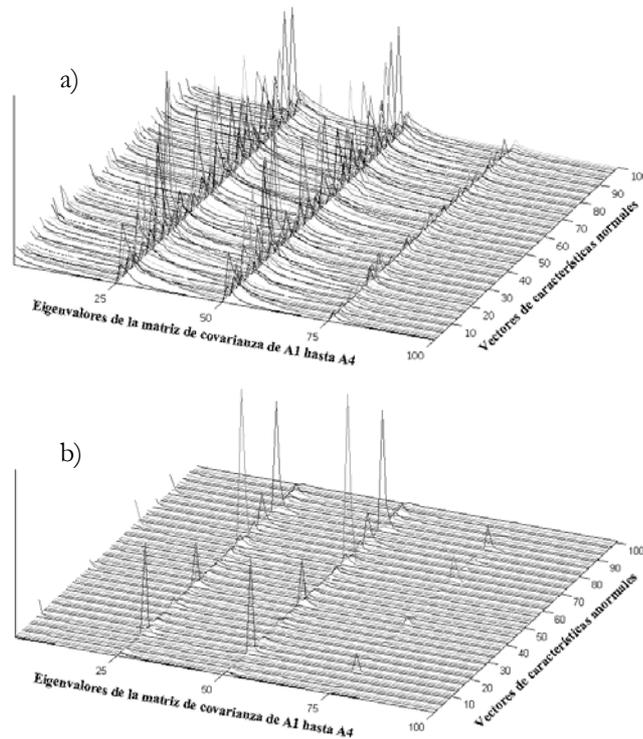


Fig. 23 Eigenvalores de la matriz de covarianza de la matriz de coeficientes A de: a) Sujeto sano y de b) Sujeto con NID

b) Se tomaron los 25 eigenvalores más importantes de cada matriz de covarianza de A y se concatenaron en un vector de  $100 \times 1$ . Posteriormente se reordenaron en orden descendente tomando los 50 eigenvalores más grandes para formar el vector de características que representa a una ventana multicanal.

El vector de características se conformó de 50 eigenvalores y el conjunto total estuvo integrado por 1140 vectores. La red neuronal que mejor desempeño mostró fue con 20 neuronas de entrada, 10 neuronas ocultas y una neurona de salida. El error de entrenamiento es de 0.0746 evaluado con la función del error cuadrático medio.

Tabla 26 Desempeño de la Red Neuronal  
MAR en conjunción con eigenvalores de la matriz de covarianza de A 50 eigenvalores

VC (50X1)	TAMAÑO	N-N	A-A	N-A	A-N	N.C	SEN %	ESP%	DES%
Entrenamiento	912	447	9	448	8	0	98.24	98.03	98.14
Prueba	114	41	15	51	5	2	89.13	77.27	83.20
Validación	114	42	15	46	10	1	80.77	75.41	78.09

Tabla 27 Validación General  
Técnica ARM en conjunción con eigenvalores de la matriz de cov. de la matriz A

SUJETO	%NORMAL	%ANORMAL
AAC	75.14	22.92
BTC	58.06	40.78
CMR	87.50	12.50
GMR	73.56	22.00
LMDL	61.42	35.42
LO	56.19	42.94
MVCC	21.28	76.28
PRC	35.14	62.25
<b>TOTAL</b>	<b>58.53</b>	<b>39.39</b>

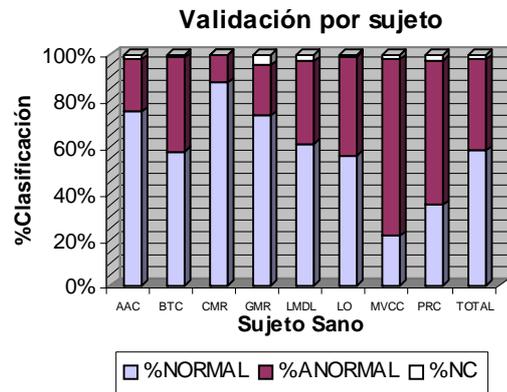


Fig. 24 Gráfica de Validación general de sujetos sanos por la técnica ARM en conjunción con eigenvalores de la matriz de cov. de la matriz A.

Tabla 28 Validación General  
Técnica ARM en conjunción con eigenvalores de la matriz de covarianza de la matriz A

SUJETO	%NORMAL	%ANORMAL
AMA	3.33	96.67
ASR	15.50	83.39
CC	2.50	95.83
CVG	3.83	95.06
EFC	88.33	9.44
EM	30.00	68.67
EMG	70.67	27.33
EV	14.44	81.11
FVM	51.67	46.11
GAG	10.83	89.17
ICG	56.67	42.50
JGM	2.22	97.78
JHL	100.00	0.00
JVG	14.67	84.67
LHM	99.52	0.00
MFC	37.14	61.43
MLB	71.11	27.78
MPM	100.00	0.00
TEA	95.00	5.00
<b>TOTAL</b>	<b>45.65</b>	<b>53.26</b>

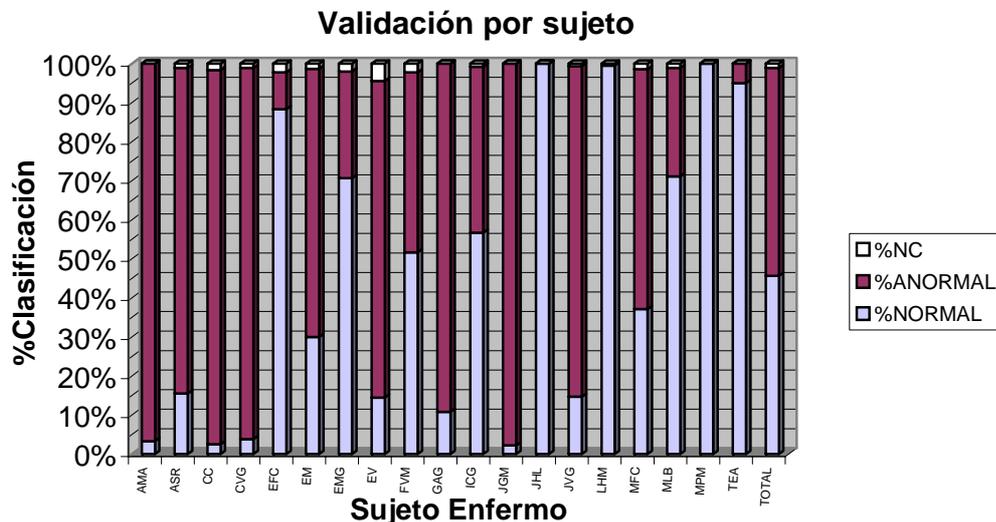


Fig. 25 Gráfica de Validación general de sujetos enfermos por la técnica ARM en conjunción con eigenvalores de la matriz de covarianza de la matriz A

### 6.3.4.1.3 Reducción de Características por medio de la proyección de eigenvectores de la matriz de covarianza del vector V

Cada matriz de coeficientes A se descompone en un vector V concatenando sus columnas, el vector resultante es de tamaño 625, se calcula la matriz de covarianza a partir de este vector y los eigenvalores y eigenvectores de la matriz de covarianza, se seleccionan los eigenvalores más importantes y se realizan las proyecciones de los eigenvectores sobre los vectores originales para reducir la dimensión, de este modo se obtiene un vector de 100 parámetros

La red neuronal que se desempeñó mejor fue con una arquitectura de 20 neuronas de entrada, 10 neuronas ocultas y una neurona de salida, entrenada durante 31 épocas, el error obtenido fue  $4.047e-7$  evaluado con la función del error cuadrático medio.

Tabla 29 Desempeño de la Red Neuronal.  
ARM en conjunción con proyección de eigenvectores de la matriz de covarianza de V

VC (100X1)	TAMAÑO	N-N	A-A	N-A	A-N	N.C	SEN %	ESP%	DES%
Entrenamiento	912	456	0	456	0	0	100	100	100
Prueba	114	28	17	39	11	19	71.79	69.64	70.72
Validación	114	33	15	31	20	15	62.26	67.39	64.83

Tabla 30 Validación General  
ARM en conjunción con proyección de eigenvectores de la matriz de covarianza de V

SUJETO	%NORMAL	%ANORMAL
AAC	67.22	18.33
BTC	59.67	33.72
CMR	82.78	9.17
GMR	64.44	21.11
LMDL	51.00	30.88
LO	40.56	42.59
MVCC	35.78	48.72
PRC	37.75	46.36
<b>TOTAL</b>	<b>54.90</b>	<b>31.36</b>

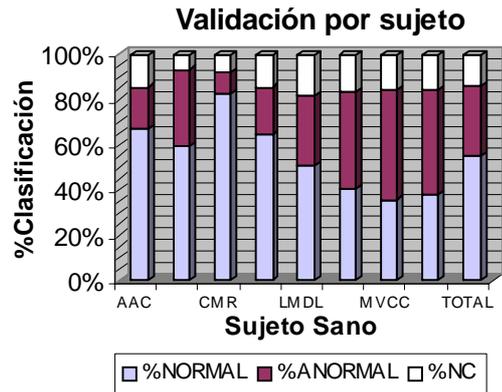


Fig. 26 Gráfica de Validación general de sujetos sanos por la técnica ARM en conjunción con proyección de eigenvectores de la matriz de covarianza de V.

Tabla 25 Validación General  
Técnica ARM en conjunción con eigenvalores de la matriz de cov. de la matriz A

SUJETO	%NORMAL	%ANORMAL
AMA	2.00	92.67
ASR	11.17	81.89
CC	5.00	93.33
CVG	3.00	94.56
EFC	83.89	8.89
EM	20.67	62.00
EMG	60.67	26.67
EV	3.33	88.89
FVM	52.22	38.33
GAG	5.00	90.83
ICG	44.17	43.33
JGM	0.00	100.00
JHL	100.00	0.00
JVG	13.33	76.67
LHM	100.00	0.00
MFC	32.86	54.76
MLB	63.33	22.22
MPM	100.00	0.00
TEA	95.56	0.56
<b>TOTAL</b>	<b>41.90</b>	<b>51.35</b>

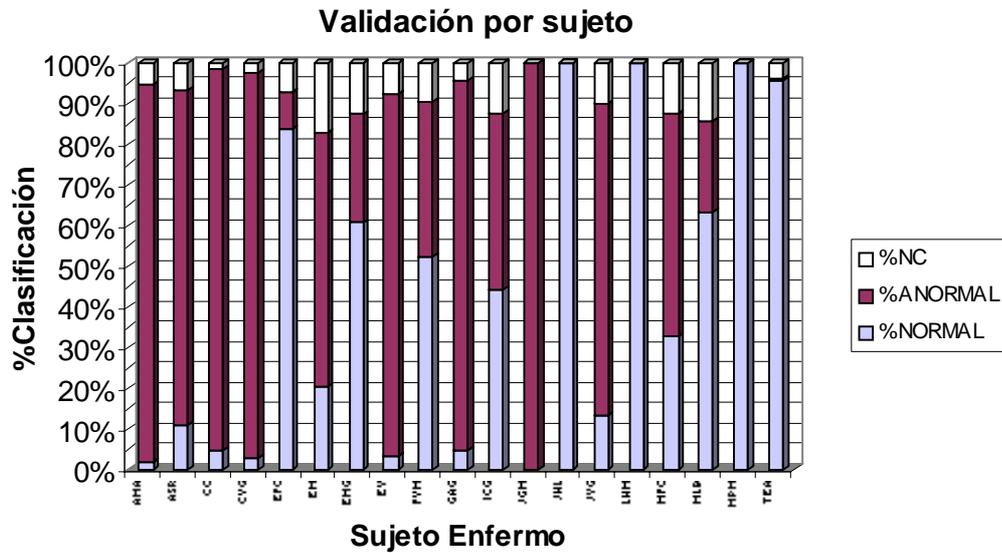


Fig. 27 Gráfica de Validación general de sujetos enfermos por la técnica ARM en conjunción con eigenvalores de la matriz de cov. de la matriz A

### 6.3.4.2 MATRIZ TOTAL DE COEFICIENTES DEL MODELO AR MULTICANAL

#### 6.3.4.2.1 Reducción de características aplicando Análisis en Componentes Principales (PCA)

Cada matriz A1... A4 fue desdoblada en  $25 \times 25 = 625$  y fueron concatenadas y se obtuvieron vectores de 2500 características. Se reunieron en total 1140 vectores los cuales redujeron su dimensión de 2500 a 136 características por medio del Análisis de Componentes Principales PCA utilizando un factor de 0.02% de varianza.

Para los datos reducidos por PCA la arquitectura de red consistió de 25 nodos de entrada, 10 nodos ocultos y un nodo de salida, evaluada también por FPE indicó un error de entrenamiento de  $6.39e-16$  durante 28 épocas.

Tabla 32 Desempeño de la Red Neuronal ARM en conjunción con PCA

VC(100X1)	Tamaño	N-N	A-A	N-A	A-N	N.C	SEN%	ESP%	DES%
Entrenamiento	912	456	456	0	0	0	100	100	100
Prueba	114	53	50	3	5	3	91.38	94.34	92.86
Validación	114	48	46	9	11	0	81.36	83.64	82.5

Tabla 33 Validación General para la clasificación de vectores parametrizados mediante modelo ARM en conjunción con PCA

SUJETO	%NORMAL	%ANORMAL
AAC	86.11	12.50
BTC	83.94	14.67
CMR	94.44	4.17
GMR	86.67	11.56
LMDL	77.46	19.33
LO	71.69	25.02
MVCC	20.44	76.83
PRC	51.78	44.50
<b>TOTAL</b>	<b>71.57</b>	<b>26.07</b>

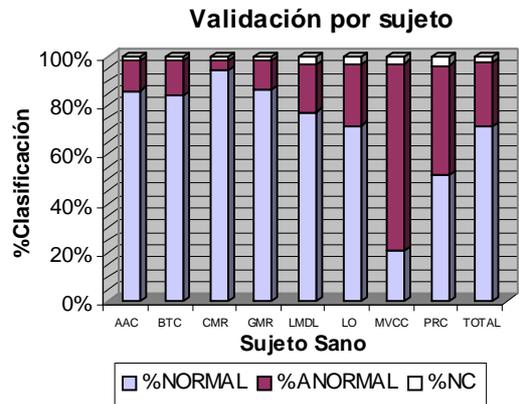


Fig. 28 Gráfica de Validación general de sujetos sanos por la técnica ARM en conjunción con PCA.

Tabla 34 Validación General para la clasificación de vectores parametrizados mediante ARM en conjunción con PCA

SUJETO	%NORMAL	%ANORMAL
AMA	4.67	95.33
ASR	5.89	92.67
CC	4.17	95.00
CVG	7.67	88.17
EFC	87.78	8.33
EM	12.00	84.67
EMG	82.00	16.67
EV	5.56	91.11
FVM	21.67	72.78
GAG	1.67	98.33
ICG	46.67	46.67
JGM	4.44	93.33
JHL	94.67	2.67
JVG	21.33	75.33
LHM	100.00	0.00
MFC	35.71	61.90
MLB	67.78	28.89
MPM	100.00	0.00
TEA	98.89	0.56
<b>TOTAL</b>	<b>42.24</b>	<b>55.39</b>

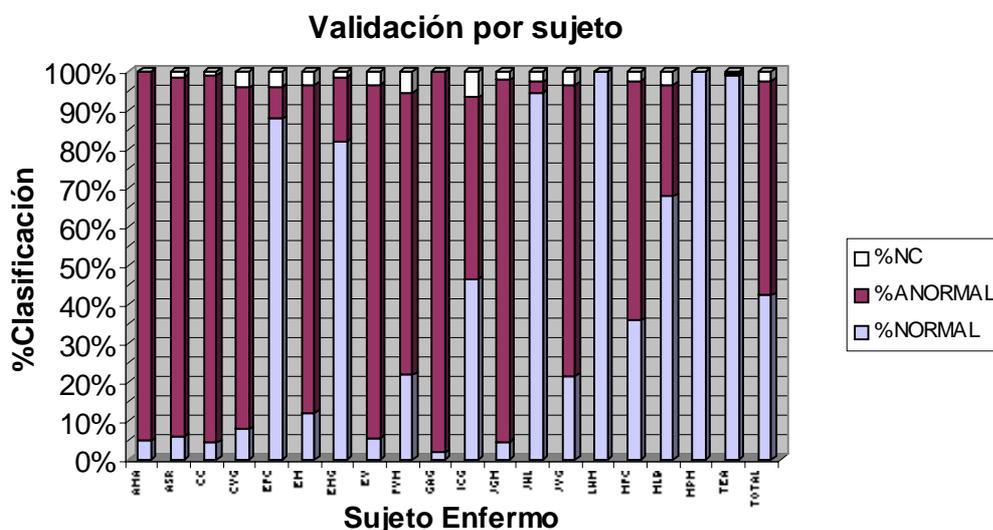


Fig. 29 Gráfica de Validación general de sujetos enfermos por la técnica ARM en conjunción con PCA

A continuación se presenta un resumen del desempeño de la red neuronal con cada una de las técnicas estudiadas en este trabajo. En la tabla 35 se muestra el porcentaje de sensibilidad, especificidad y desempeño de la red neuronal para las fases de entrenamiento, prueba y validación.

En la tabla 36 se muestra el promedio del desempeño de la red neuronal y el promedio total para evaluar y comparar los resultados que arroja cada técnica. En la tabla 36 aparecen solo los promedios del desempeño debido a que este índice es en sí un promedio de la sensibilidad y especificidad del clasificador.

En la tabla 37 se resumen los porcentajes para la validación general y únicamente se muestran los promedios totales obtenidos en cada técnica.

**TABLA 35 RESUMEN DEL DESEMPEÑO DE LA RED NEURONAL**

MÉTODO DE PARAMETRIZACIÓN	ENTRENAMIENTO			PRUEBA			VALIDACIÓN		
	%S	%E	%D	%S	%E	%D	%S	%E	%D
1. Eigenvalores de la matriz de covarianza (6 eigenvalores)	97.25	100	98.63	92.31	95.00	93.65	97.62	100	98.81
2. Eigenvalores de la matriz de covarianza (10 eigenvalores)	99.75	99.27	99.51	80.43	76.79	78.61	82.61	81.82	82.21
3. Medidas clásicas en el dominio de la frecuencia (4 registros)	100	100	100	100	100	100	100	100	100
4. Medidas clásicas en el dominio de la frecuencia (todos registros)	100	100	100	90.00	90.00	90.00	100	100	100
5. Medidas clásicas en el dominio de la frecuencia utilizando segmentación por ventanas	100	99.10	99.55	100	97.67	98.84	100	97.56	98.78
6. Modelo AR Monocanal	100	100	100	98.15	94.83	96.49	98.15	94.92	96.53
7. ARM con Reducción por Descomposición en Valores Singulares (SVD) de la matriz A	99.34	99.78	99.56	86.00	80.33	83.16	80.39	73.33	76.86
8. ARM con Reducción mediante la obtención de los eigenvalores de la matriz de covarianza de la matriz de coeficientes (100 eigenvalores)	99.78	99.11	99.44	77.78	80.00	78.89	83.67	85.71	84.69
9. ARM con Reducción mediante la obtención de los eigenvalores de la matriz de covarianza de la matriz de coeficientes (50 eigenvalores)	98.24	98.03	98.14	89.13	77.27	83.2	80.77	75.41	78.09
10. ARM con Reducción por Proyección de eigenvectores de la matriz de covarianza del vector V	100	100	100	71.79	69.64	70.72	62.26	67.39	64.83
11. ARM con Reducción Análisis en Componentes Principales (PCA)	100	100	100	91.38	94.34	92.86	81.36	83.64	82.5

<p><b>S%</b> Sensitividad  <b>E%</b> Especificidad  <b>D%</b> Desempeño</p>
---

**TABLA 36 PROMEDIO DEL DESEMPEÑO DE LA RED NEURONAL**

MÉTODO DE PARAMETRIZACIÓN	ENTRENAMIENTO	PRUEBA	VALIDACIÓN	%PROMEDIO
	%DESEMPEÑO	%DESEMPEÑO	%DESEMPEÑO	
1. Eigenvalores de la matriz de covarianza (6 eigenvalores)	98.63	93.65	98.81	<b>97.03</b>
2. Eigenvalores de la matriz de covarianza (10 eigenvalores)	99.51	78.61	82.21	<b>86.78</b>
3. Medidas clásicas en el dominio de la frecuencia (4 registros)	100	100	100	<b>100.00</b>
4. Medidas clásicas en el dominio de la frecuencia (todos registros)	100	90	100	<b>96.67</b>
5. Medidas clásicas en el dominio de la frecuencia utilizando segmentación por ventanas	99.55	98.84	98.78	<b>99.06</b>
6. Modelo AR Monocanal	100	96.49	96.53	<b>97.67</b>
7. ARM con Reducción por Descomposición en Valores Singulares (SVD) de la matriz A	99.56	83.16	76.86	<b>86.53</b>
8. ARM con Reducción mediante la obtención de los eigenvalores de la matriz de covarianza de la matriz de coeficientes (100 eigenvalores)	99.44	78.89	84.69	<b>87.67</b>
9. ARM con Reducción mediante la obtención de los eigenvalores de la matriz de covarianza de la matriz de coeficientes (50 eigenvalores)	98.14	83.2	78.09	<b>86.48</b>
10. ARM con Reducción por Proyección de eigenvectores de la matriz de covarianza del vector V	100	70.72	64.83	<b>78.52</b>
11. ARM con Reducción Análisis en Componentes Principales (PCA)	100	92.86	82.5	<b>91.79</b>

**TABLA 37 RESUMEN DE VALIDACIÓN GENERAL**

MÉTODO DE PARAMETRIZACIÓN	SANOS		ENFERMOS	
	%Normal	%Anormal	%Normal	%Anormal
1. Eigenvalores de la matriz de covarianza (6 eigenvalores)	<b>55.52</b>	1.78	13.92	<b>16.69</b>
2. Eigenvalores de la matriz de covarianza (10 eigenvalores)	<b>61.93</b>	31.42	24.23	<b>69.39</b>
3. Medidas clásicas en el dominio de la frecuencia (4 registros)	<b>60.00</b>	35.00	5.26	<b>78.70</b>
4. Medidas clásicas en el dominio de la frecuencia (todos registros)	<b>97.01</b>	0.69	1.25	<b>98.75</b>
5. Medidas clásicas en el dominio de la frecuencia utilizando segmentación por ventanas	<b>59.29</b>	21.44	7.42	<b>74.91</b>
6. Modelo AR Monocanal	<b>73.66</b>	23.26	11.91	<b>85.43</b>
7. ARM con Reducción por Descomposición en Valores Singulares (SVD) de la matriz A	<b>55.75</b>	40.93	43.83	<b>54.10</b>
8. ARM con Reducción mediante la obtención de los eigenvalores de la matriz de covarianza de la matriz de coeficientes (100 eigenvalores)	<b>53.47</b>	34.47	41.90	<b>51.35</b>
9. ARM con Reducción mediante la obtención de los eigenvalores de la matriz de covarianza de la matriz de coeficientes (50 eigenvalores)	<b>58.53</b>	39.39	45.65	<b>53.26</b>
10. ARM con Reducción por Proyección de eigenvectores de la matriz de covarianza del vector V	<b>54.90</b>	31.36	41.90	<b>51.35</b>
11. ARM con Reducción Análisis en Componentes Principales (PCA)	<b>71.57</b>	26.07	42.24	<b>55.39</b>

## CAPÍTULO VII

### DISCUSIÓN

Las NID constituyen un grupo heterogéneo de enfermedades que comparten aspectos clínicos, radiológicos y funcionales, y presentan además como rasgo característico, el desarrollo progresivo de fibrosis pulmonar. El proceso fisiopatológico de la fibrosis se caracteriza por modificar gravemente el tejido pulmonar y el intercambio gaseoso, provocando que durante la respiración se presenten sonidos anormales conocidos como estertores crepitantes.

Las NID afectan a un número considerable de personas en nuestro país, que tienen un difícil diagnóstico en etapas tempranas y comúnmente se detecta en etapas avanzadas de la enfermedad, cuando es ya irreversible y mortal para el paciente en un lapso de tiempo relativamente corto. Es por ello que resulta interesante el diseño de una metodología simple y no invasiva que apoye el diagnóstico de las NID en cualquier etapa de la enfermedad y que sirva como complemento para el diagnóstico que se lleva a cabo por las otras técnicas mencionadas en el Capítulo I, sección 1.1.5.

En la literatura se han reportado numerosos trabajos sobre clasificación de sonidos respiratorios, pero bajo diferentes condiciones de enfermedad de los sujetos, los estudios más comunes abarcan sonidos provenientes de:

- a) Una patología en específico, como en el caso del asma [1], [2], [3].
- b) Un grupo de enfermedades del tipo obstructivas y restrictivas que incluyen una amplia gama de sonidos anormales como estertores, sibilancias, sonidos roncales, sonidos disminuidos, entre otros [4], [5], [6], [7], [8], [9].

El trabajo que aquí se presenta, es la base para el desarrollo de una nueva herramienta que apoye el diagnóstico de las NID. Se comenzó por obtener muestras de sonidos respiratorios de pacientes con NID, se exploraron diversas técnicas de procesamiento para caracterizar o parametrizar a los sonidos respiratorios de forma reducida y finalmente se llevó a cabo la tarea de reconocer entre sonidos normales o anormales mediante un algoritmo de clasificación.

Esta sección tratará diversos temas que no son presentados en orden de importancia, pero siguen la progresión de la tesis. En primer lugar se discutirán los hallazgos encontrados en las señales de sonido respiratorio, es decir, el tipo de sonidos encontrados en las señales grabadas de los pacientes con NID y aspectos relacionados con la adquisición multicanal de los sonidos respiratorios en el Instituto Nacional de Enfermedades Respiratorias.

Se han publicado en la literatura, trabajos que describen que para la identificación visual de los estertores crepitantes en la información acústica respiratoria y su correspondiente clasificación, el médico cuenta con una caracterización temporal del estertor en términos de amplitudes y duración de las ondas presentes en el estertor [10], [11]. Con la ayuda de esta caracterización, y después de un procedimiento manual laborioso que incluye también la correlación con la etapa del ciclo respiratorio, se intenta determinar el número de estertores presentes en la señal respiratoria anormal para estimar la gravedad de la enfermedad. Sin embargo, este procedimiento de evaluación de los estertores, visual y manual, implica posibles errores de clasificación pues varias morfologías podrían no reunir los requisitos de una crepitancia típica. Lo anterior nos da la idea de que el análisis de los sonidos respiratorios puede utilizarse como una herramienta para determinar el estado de un paciente. Sin embargo, sugiere la necesidad de atenuar la subjetividad del método y desarrollar una herramienta computacional que realice las tareas en forma automática, disminuya el tiempo y la tasa de error.

Durante la estancia en el INER del grupo de investigación, se observó que en la práctica clínica el médico ausculta en diferentes puntos sobre la superficie pulmonar para obtener información de la extensión de la enfermedad pulmonar. La revisión de los expedientes clínicos revela que los médicos adscritos del INER, hacen referencia a la técnica de auscultación, anotando el hallazgo de estertores crepitantes escapulares o subescapulares bilaterales. El diagnóstico se lleva a cabo mediante el análisis del expediente clínico y evaluando los resultados de las técnicas como pruebas funcionales, rayos X, tomografía axial computarizada y lavado bronco-alveolar. Como un estudio complementario para el historial clínico se realizó la auscultación en los mismos puntos donde se colocaron los micrófonos para la adquisición de los sonidos respiratorios, esto también se utilizará como una guía para intentar correlacionar los ruidos

respiratorios escuchados con los resultados obtenidos en este trabajo.

A través de la técnica de auscultación en la parte posterior del tórax de los pacientes con NID en estudio, se encontraron en su mayoría sonidos respiratorios de intensidad disminuida. Estertores crepitantes inspiratorios en la base y puntos superiores de los pulmones, en pacientes graves se encontraron en casi todos los puntos de auscultación. Además, se encontraron estertores de tipo seco, graves o tipo velcro. También en algunos pacientes se encontraron sibilancias cortas en la fase inspiratoria

Utilizando la interfaz gráfica desarrollada en el presente trabajo se pudo confirmar la descripción anterior, y además se proporcionó la observación de varios canales simultáneos. Además es posible aplicar diferentes técnicas de procesamiento como filtros digitales, expansión en el tiempo, etc. que hacen más fácil el estudio de las señales acústicas respiratorias para observar las características morfológicas de la señal, duración, frecuencia, etc.

Los registros de los pacientes muestran sonidos anormales de acuerdo con la literatura, se encontraron estertores crepitantes y sibilancias. Sin embargo la frecuencia más alta de los estertores no sobrepasa los 1000 Hz. Tampoco se encontraron estertores que cumplen estrictamente con el criterio de Murphy para los estertores clásicos [10], [11]. Se localizaron trenes de estertores con diferente morfología a los estudiados en la literatura, sin embargo al escucharlos fue posible encontrar el patrón de estertores tipo velcro.

Es posible que el procedimiento computacional para la obtención de los vectores de características incluya a los estertores típicos y además a los de difícil reconocimiento, remarcando así mayores diferencias entre la información acústica normal y anormal.

Para todas las técnicas, excepto en la aplicación de la técnica de medidas clásicas en el dominio de la frecuencia, se obtuvieron suficientes ventanas de tiempo de 200 muestras aproximadamente. Dado que se caracterizan inspiraciones completas para la obtención de los parámetros de frecuencias, la población resultó insuficiente para una adecuada generalización del clasificador y se sugiere realizar mayor cantidad de registros para estudios futuros. El número de pacientes o sujetos sanos, debe incrementarse gradualmente hasta obtener resultados satisfactorios o bien, calcular la N estadística.

Por otro lado, en lo que se refiere a las técnicas de procesamiento, otros autores han empleado esquemas de clasificación basados en información clínica y radiográfica [4]; modelos autorregresivos (AR) monocanales de orden 6 [12], [6], [9]; características espectrales [1], espectro de potencia [5], [2]; transformada wavelet [3]; coherencia de la señal [8] y frecuencias percentiles [9], entre las más importantes y recientes, y en estos trabajos se han encontrado porcentajes de aciertos que van desde el 43% hasta el 98% en la etapa de entrenamiento y prueba del clasificador. En un intento previo por nuestro grupo de investigación, empleando modelo AR multicanal en combinación con descomposición en valores singulares (SVD) para reducir el espacio de características, se encontró un porcentaje promedio de clasificaciones correctas de 81% utilizando un arreglo de micrófonos de 12 sensores sobre la superficie torácica posterior [13], [14].

De acuerdo al modelo propuesto en el presente trabajo, comparando las cuatro técnicas de parametrización y una red neuronal como clasificador, los porcentajes de desempeño varían del 97.67% al 100%. La técnica de parametrización utilizando el modelo AR monocanal resultó ser una de las mejores técnicas entre las estudiadas en la presente tesis. Si se compara con los resultados obtenidos con 12 micrófonos en el estudio previo, es importante mencionar que la incorporación de mayor información multicanal probablemente influyó para incrementar el porcentaje de aciertos en la clasificación.

En la literatura se reporta el uso de clasificadores de distancia y redes neuronales artificiales, los porcentajes de clasificación que se presentan son resultados obtenidos durante la etapa de entrenamiento, prueba y validación del clasificador.

La adquisición multicanal de los sonidos respiratorios permite la exploración de técnicas de extracción de características cuantitativas que incorporen información espacial y temporal de forma simultánea. La incorporación de información espacial y temporal es relevante en el apoyo al diagnóstico de enfermedades como las NID dado que en el curso del desarrollo de la enfermedad, ésta presenta un cierto comportamiento espacio-temporal.

A continuación se resume las ventajas y desventajas de cada una de las técnicas empleadas en la presente investigación:

## 7.1 MEDIDAS CLÁSICAS EN EL DOMINIO DE LA FRECUENCIA

Este método ha sido empleado por otros investigadores utilizando cuatro sensores y reportan que el desempeño promedio durante el entrenamiento del clasificador fue de 73.5%, en el cual la clasificación se realizó para cada sensor y no utilizaron la información multicanal [9]. Aunque no reportan los resultados de la fase de prueba y validación. En el presente trabajo, se encontró: a) un desempeño promedio de 100% en las tres fases utilizando los registros de 4 sujetos obteniendo 38 vectores, y b) un porcentaje promedio de 100%, 90% y 100% para entrenamiento, prueba y validación, respectivamente, utilizando los registros de todos los sujetos disponibles con un total de 208 vectores de características.

Al realizar la validación con registros que no había visto la red y con vectores de características pertenecientes a una población diferente, se encontró que en el primer caso los sujetos sanos se clasificaron como sanos de acuerdo a la *regla de la mayoría*, en promedio 60% sanos y 35% enfermos, y los sujetos enfermos se clasificaron en promedio como enfermos 78.70% y 5.26% como sanos.

En el segundo caso, al analizar la tabla de validación general, se observa que los resultados son muy altos, (97.01% en sujetos sanos y 98.75% en pacientes con NID) sin embargo, no son muy confiables, debido a que se obtuvieron con vectores de características pertenecientes a la misma población utilizada para entrenamiento, prueba y validación de la red neuronal.

Se considera que este método es bastante prometedor en cuanto a que existen parámetros percentiles en el dominio de la frecuencia que son útiles para discriminar entre sonidos normales y anormales. En este estudio no fue posible obtener mayor número de registros y dado que tenemos un número limitado de vectores de características para obtener una adecuada generalización de la red neuronal con datos ajenos al conjunto de entrenamiento, será necesario realizar pruebas en sujetos que no han sido “vistos” por la red, en el caso de la validación general. Otra alternativa consideraría dividir cada fase en ventanas de tiempo para obtener una mayor cantidad de vectores de características, tal y como se realizó en las demás técnicas. Al aplicar la división en ventanas de tiempo, se obtuvo un porcentaje promedio de

59.29% para sujetos sanos y 74.91 para pacientes con NID en la validación general.

## **7.2 EIGENVALORES DE LA MATRIZ DE COVARIANZA DE LA SEÑAL DE SONIDO RESPIRATORIO**

La obtención de los eigenvalores de la matriz de covarianza de una ventana de la señal de sonido respiratorio es un método simple, que demostró buenos resultados en general. Tiene la ventaja de evitar la aplicación y validación de un modelo a la señal, obteniendo los parámetros en forma más directa. Los resultados obtenidos con 10 eigenvalores representando a una ventana de tiempo son superiores a los encontrados con 6 eigenvalores, en la tabla de desempeño de la red neuronal puede observarse que aunque los porcentajes de sensibilidad, especificidad y desempeño de la red son altos, el número de vectores descartados o no clasificados como normales o anormales es considerable (40.97%, 33.3% y 42.59% de vectores descartados durante la fase de entrenamiento, prueba y validación de la red neuronal, respectivamente), lo cual implica una desventaja.

En la tabla de validación general el resultado para clasificación de sujetos sanos es bajo (55.52%), mientras que para los sujetos enfermos es prácticamente inaceptable (16.69%). Esto significa que con 6 eigenvalores representando a una ventana, el clasificador no es capaz de generalizar adecuadamente. Si se entrena al clasificador con 10 eigenvalores el costo computacional se incrementa y los resultados no mejoran significativamente (61.93% para sujetos sanos y 69.39% para pacientes con NID).

## **7.3 MODELO AR MONOCANAL**

La técnica de parametrización utilizando un modelo AR de orden 4 resultó ser la mejor técnica según la tabla de desempeño de la red con porcentaje global de 97.67% y validación general de 73.66% para clasificación de sujetos sanos y de 85.43% para sujetos enfermos. Otros autores, quienes han empleado esquemas de clasificación basados en modelos AR monocanales y

componentes principales, obtuvieron resultados de 81.9% en el desempeño promedio realizando la clasificación de manera individual en cada uno de los cuatro micrófonos [9], 85.7% de desempeño promedio utilizando un solo micrófono [12] y un porcentaje promedio de 79.85% utilizando modelo AR y PCA [9]. En todos los trabajos utilizaron un modelo AR de orden 6.

La aplicación de un modelo AR de orden 4 a las señales de sonido respiratorio se considera adecuado según el Criterio de Akaike, el orden del modelo parece ajustarse bien a los datos reales y no se requiere demasiado esfuerzo computacional para el cálculo de los coeficientes AR, ya que la cantidad de datos es manejable (los coeficientes AR monocanales son escalares) y al concatenar los coeficientes AR obtenidos en un canal con los demás canales de la misma ventana de tiempo, se adquiere la información espacial que se requiere para la clasificación. Este método presenta la ventaja sobre el modelo AR multicanal de que no es necesario la implementación de métodos adicionales para la reducción de características.

Por otro lado, se considera que dentro de la etapa de reconocimiento de patrones, el método para etiquetar los vectores de características como normales o anormales debe ser mejorado, debido a que existe ambigüedad al establecer los patrones si se considera que existen ventanas normales dentro de registros anormales o viceversa, y que están siendo etiquetados erróneamente.

#### **7.4 MODELO AR MULTICANAL**

No es posible realizar comparaciones de los resultados obtenidos en el presente trabajo con estudios de otros autores, ya que no existen reportes de la utilización de esta técnica aplicada a clasificación de sonidos respiratorios.

El uso del modelo AR multicanal incluye información espacial y temporal de la información acústica la cual es modelada simultáneamente para todos los canales englobando características correlacionadas y no correlacionadas. Sin embargo, la aplicación del modelo AR multicanal exige un esfuerzo computacional considerable, debido al manejo de grandes volúmenes de

datos, como sabemos, una ventana multicanal está representada por 4 matrices de tamaño 25 X 25, y resulta difícil introducir estos datos al clasificador. Entonces, es necesario recurrir a técnicas adicionales para la reducción de características.

De las técnicas de reducción del espacio de características, la que mejor resultados mostró fue la técnica de *análisis de componentes principales (PCA)* con 91.79% de porcentaje promedio en el desempeño de la red neuronal y la que presentó el menor desempeño, fue la de *reducción por proyección de eigenvectores de la matriz de covarianza de  $V$*  con 78.52% de porcentaje promedio en el desempeño de la red neuronal.

También en la tabla de validación general se observa que la técnica de PCA presentó resultados superiores (71.57% para Sanos, 55.39% para Enfermos) a los presentados por las demás técnicas de reducción de la dimensionalidad, las cuales mostraron porcentajes muy bajos y similares entre sí.

Se considera que el comportamiento del modelo AR multicanal durante la validación por sujeto, se debe a la ocurrencia espacio temporal de los eventos que caracterizan a los sonidos normales y anormales, lo anterior justifica también la razón de que los sujetos sanos son clasificados con mayor porcentaje (71.57% ARM con PCA), en donde la información acústica espacio temporal es más homogénea que en el caso de los pacientes con NID (55.39% ARM con PCA), en donde la información acústica anormal se presenta en forma localizada. Ver tabla 37.

## **7.5 RED NEURONAL**

Se ha mencionado la utilización de una red neuronal como clasificador en la presente tesis. Varias redes neuronales fueron entrenadas con el algoritmo de retropropagación, que posee la ventaja de ser simple y eficiente para la discriminación de sólo dos clases: normales y anormales. La función de entrenamiento basada en la regla de Levenberg-Marquardt se caracteriza por acelerar la convergencia hacia un mínimo global de la superficie de error, reduciendo el tiempo de entrenamiento. También se probaron diferentes arquitecturas de redes

neuronales, utilizando una capa de entrada, una capa de salida y una capa oculta, variando el número de nodos de entrada (NE) de 10 a 30, y el de nodos ocultos (NO) de 5 a 15. La configuración que contribuyó a la obtención de mejores resultados fue la de 20 NE, 10 NO y 1 neurona de salida (NS); o bien, 25 NE, 10 NO y 1 NS.

La red neuronal utilizada es de tipo supervisada y cada vector de características se etiquetó como normal si provenía de un registro de sujeto sano o como anormal si pertenecía a un registro de paciente con NID. Es recomendable modificar el método para etiquetar a los vectores de características ya que presenta limitaciones al establecer los patrones generalizados, principalmente para los pacientes con NID, debido a que existen segmentos de la señal de sonido respiratorio normales y anormales dentro de un mismo registro.

### **7.5.1 FASE DE ENTRENAMIENTO**

Durante la fase de entrenamiento las mejores técnicas que clasificaron correctamente todos los vectores presentados fueron: Medidas clásicas en el dominio de la frecuencia, modelo AR monocanal, modelo AR multicanal con reducción por PCA y reducción por proyección de eigenvectores de la matriz de covarianza del vector de coeficientes del modelo AR multicanal, todas obtuvieron un desempeño promedio del 100%. Seguidas por el modelo AR multicanal con reducción de la dimensión con descomposición en valores singulares 99.56% en el promedio de su desempeño. Después, la técnica de *eigenvalores de la matriz de covarianza*, utilizando 10 eigenvalores obtuvo 99.51%

Es importante aclarar que el método con las medidas clásicas de frecuencia no es comparable con las demás técnicas debido a la metodología empleada al calcular los parámetros. La técnica de medidas en la frecuencia es aplicada a toda la fase respiratoria (inspiración), mientras que las demás se aplicaron utilizando ventanas cortas de tiempo. Por lo tanto, el número de vectores de características en ambos conjuntos es diferente. Con la técnica de medidas en la frecuencia se pueden obtener pocos vectores, y fue necesario reunir todos los registros disponibles para utilizarlos en la fase de entrenamiento, prueba y validación (208 vectores), aún así, el número de vectores no se compara en cantidad con los utilizados en las demás técnicas (1140 vectores,

modelo AR monocal y multicanal) y con la técnica de eigenvalores se utilizaron 1080 vectores en el conjunto total.

En la fase de prueba, después de la técnica de medidas en la frecuencia, la técnica del modelo AR monocal fue la que mejor resultado obtuvo con porcentaje de 96.49%, seguida de la técnica de eigenvalores de la matriz de covarianza utilizando 6 eigenvalores y la de modelo AR multicanal con reducción de la dimensión utilizando PCA.

Al comparar los resultados de la fase de validación, de nuevo se observa que la técnica de medidas en la frecuencia tiene excelentes resultados (100%). En segundo sitio se encuentra la de eigenvalores de la matriz de covarianza utilizando 6 eigenvalores para caracterizar a la señal respiratoria (98.81%) y en tercer lugar la técnica del modelo AR monocal (96.53%)

### **7.5.2 VALIDACIÓN GENERAL**

De acuerdo con las tablas de validación general, se observa que la técnica de medidas en la frecuencia se comporta relativamente bien, con porcentajes de 97.01% en el promedio de clasificación para sujetos sanos y 98.75% para clasificar sujetos enfermos. Existen características en la frecuencia que logran diferenciar entre sonidos normales y anormales, sin embargo, estos resultados no son objetivos, ya que los datos ya habían sido “vistos” por la red en la primera fase (entrenamiento, prueba y validación). Aunque esta técnica ha demostrado resultados muy favorables para la clasificación, no fue posible en este trabajo realizar una validación con mayor cantidad de registros.

Con la técnica del modelo AR monocal se encontró que un sujeto sano puede clasificarse como tal en un 73.66% y un paciente puede clasificarse como enfermo con un porcentaje de 85.43%. También el modelo AR en combinación con PCA reveló que en el caso de los sanos 71.57% y la técnica de eigenvalores de la matriz de covarianza contribuyó a que los enfermos clasificaran con 69.39%.

Los resultados obtenidos con las técnicas de AR monocal y AR multicanal en combinación con PCA son más consistentes que las demás técnicas porque presentan buen desempeño en la etapa de entrenamiento, prueba y validación de la red neuronal y también durante la etapa de

validación general, en cuanto al número de vectores descartados los porcentajes son muy pequeños, del orden del 0% al 0.3%, y en el caso de la técnica de eigenvalores de la matriz de covarianza el número de vectores descartados, que no fueron clasificados como normales y tampoco como anormales es amplio con porcentajes del 33.3% al 40.97%.

## 7.6 REFERENCIAS

- [1] M. Oud, "Classification and recognition of asthmatic breathing sounds." EANN Conf. Proc. pp. 414-418, 1996.
- [2] S. Rietveld and M. Oud, "Classification of asthmatic breath sounds spectra by human examiners versus artificial neural networks." *Comp. Biomed. Research* 32, pp. 440-448, 1999.
- [3] M. C. Sezgin, Z. Dokur, T. Ölmez, M. Korürek, "Classification of respiratory sounds by using an artificial neural network." *Proceedings of the Annual EMBS International Conference, Istanbul, Turkey*. Pp. 697-699, 2001.
- [4] N. Asada, K. Doi, H. MacMahon, S. Montner, M. L. Giger, C. Abe, Y. Wu, "Neural network approach for differential diagnosis of interstitial lung diseases." *Proc. SPIE* 1233, pp. 45-50, 1990.
- [5] L. R. Waitman. L. "Representation and classification of breath sounds recorded in an intensive care setting using neural networks." MS thesis. Vanderbilt University, 1998.
- [6] Y. P. Kahya, E. C. Güler, B. Sankur, S. Raudys, "Hierarchical classification of respiratory sounds." *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and biology society*, 20:3, pp. 1598-1601, 1998.
- [7] M. Bahoura, C. Pelletier, "New Parameters for respiratory sounds classification." *Canadian Conference on Electrical and Computer Engineering, Montreal, Canada*, Vol. 3, pp 1457-1460, 2003.

- [8] K. S. Baydar, A. Ertuzun, Y. P. Kahya, "Analysis and classification of respiratory sounds by signal coherence method." Proceedings of EMBS, IEEE, Cancún, Mexico. Pp. 2950-2953, 2003.
- [9] Y. P. Kahya, E. Bayatli, M. Yeginer, K. Ciftci, G. Kilinc, "Comparison of different feature sets for respiratory sound classifier." Proceedings of EMBS, IEEE, Cancún, Mexico. Pp. 2853-2856, 2003.
- [10] R. Murphy, S. Holford, W. Knowle, "Visual lung-sound characterisation by time-expanded wave-form analysis", *N. Eng. J. Med.* 296:2968-71, 1977.
- [11] R. Murphy, E. Del-Bono, F. Davidson, "Validation of an automatic crackle (rale) counter", *Am. Rev. Respir. Dis.* 140:1017-20, 1989.
- [12] B. Sankur, Y.P. Kahya, E. C. Guler, T. Engin, "Comparison of AR-Based Algorithms for Respiratory Sounds Classification." *Computers in Biology and Medicine*, 24(1), pp. 67-76. , 1994.
- [13] G. Martínez, T. Aljama, R. González, G. Chi-Lem, S. Charleston, "A multichannel autoregressive model applied to acoustic information in diffuse interstitial pneumonia." Proc. 29th Annual Conf. Int. Lung Sounds Assoc., Glasgow, Scotland, September 2004.
- [14] G. Martínez, T. Aljama, R. González, G. Chi-Lem, S. Charleston, "Clasificación de información acústica en neumopatías intersticiales difusas mediante un análisis autorregresivo multicanal." *Jornadas del Posgrado Divisional en Ciencias Básicas e Ingeniería*, Universidad Autónoma Metropolitana-Iztapalapa, México, pp. 135-138, 2004.

## CAPÍTULO VIII

### CONCLUSIONES Y PERSPECTIVAS

La adquisición de información acústica respiratoria multicanal permitió la exploración de técnicas de extracción de características cuantitativas que incorporan información espacial y temporal de forma simultánea. La incorporación de información espacial y temporal es relevante en el apoyo al diagnóstico de enfermedades como las NID dado que en el curso del desarrollo de la enfermedad, ésta presenta un cierto comportamiento espacio-temporal.

El desempeño de las técnicas de parametrización de la señal de sonido respiratorio multicanal presentó diferencias y según los porcentajes de clasificación la técnica que aporta mayor utilidad para discriminar sonidos normales y anormales es la técnica de medidas clásicas en el dominio de la frecuencia, sin embargo, se ha discutido que se requiere una cantidad mayor de vectores de características para entrenamiento y validación de la red que la cantidad utilizada.

La técnica que empleó el modelo AR monocanal de orden 4 presentó muy buen desempeño global y es un método simple de aplicar a las señales respiratorias sin máximos requerimientos computacionales.

La aplicación del modelo AR multicanal, en conjunto con el análisis de componentes principales para la reducción de la dimensionalidad del conjunto de vectores de características, para la clasificación de sonidos respiratorios en sujetos con NID mostró un buen desempeño. Pero en la validación general presentó un porcentaje medio. La utilización del modelo AR multicanal requirió de un gran costo computacional e hizo necesario la utilización de técnicas de reducción de la dimensionalidad.

Los eigenvalores de la matriz de covarianza de la señal de sonido respiratoria parece ser una buena técnica para parametrizar y clasificar sonidos normales y anormales, sin embargo, es necesario realizar mejoras en la metodología para reducir el número de vectores no clasificados por la red neuronal.

Consideramos que es necesario aplicar otro método para definir o etiquetar las ventanas de tiempo y establecer patrones deseados utilizando clasificadores supervisados o bien utilizar una red no supervisada u otro tipo de clasificadores (máquinas de soporte vectorial).

Debido a las maniobras realizadas, varios registros mostraron artefactos de ruido de difícil eliminación, por ello, consideramos que se debe tener un mejor control de las condiciones experimentales, asimismo una mayor depuración de los registros para la eliminación de ruidos y el controlar de manera más estricta las maniobras a realizar, lo cual podría conducir a registros de mayor calidad y, consecuentemente, a mejorar el desempeño del esquema de clasificación.

En forma general, los resultados indican que es necesario trabajar en diferentes aspectos de la metodología para que pueda aplicarse satisfactoriamente en el ambiente clínico como un complemento para el diagnóstico de las NID, que actualmente se realiza por otras técnicas. Este trabajo proporcionó la base para seguir explorando otras metodologías, puesto que es un logro importante el reconocimiento entre sonidos normales y anormales con un alto porcentaje de aciertos, en cualquier etapa de la enfermedad.

Más adelante se podría correlacionar el grado de enfermedad con los porcentajes obtenidos y así poder clasificar diferentes etapas de lesión durante la evolución de la enfermedad, y esto impactará sobretodo, en el reconocimiento de pacientes en etapas tempranas de NID, para establecer un tratamiento oportuno que evite el desarrollo irreversible de fibrosis pulmonar.

Con el fin de facilitar el manejo de datos evitando la reducción de características se recomienda la división espacial del tórax y después utilizar un conjunto de clasificadores por cada región torácica. Con esto es posible lograr un manejo localizado de la información acústica espacio-temporal que proporciona la técnica multicanal.

## ANEXO 1

### PARTICIPACIONES EN CONGRESOS

Jornadas del Posgrado Divisional con el trabajo titulado: “*Clasificación de Información Acústica en Neumopatías Intersticiales Difusas Mediante un Análisis Autorregresivo Multicanal.*”

◆ Colegio Nacional, México, D. F., 8 de Septiembre de 2004.

29<sup>th</sup> Annual Conference of the International Lung Sounds Association con el trabajo titulado: “*A Multichannel Autoregressive Model Applied to Acoustic Information In Difuse Interstitial Pneumonia.*”

◆ Glasgow, Escocia., 9 y 10 de Septiembre 2004.

27<sup>th</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society con el trabajo titulado: “*Computerized Classification of Normal and Abnormal Lung Sounds by Multivariate Linear Autoregressive Model.*”

◆ Shanghai, China., 1 al 4 de Septiembre de 2004.

